



Search for SUSY in Same-Sign Di-Lepton Events with the CMS Detector

Dissertation
zur Erlangung des Doktorgrades
des Department Physik
der Universität Hamburg

vorgelegt von
Matthias Stein
aus Göttingen

Hamburg
2012

Gutachter der Dissertation:	Dr. Isabell Melzer-Pellmann Prof. Dr. Peter Schleper
Gutachter der Disputation:	Dr. Isabell Melzer-Pellmann JProf. Dr. Christian Sander
Datum der Disputation:	21. Juni 2012
Vorsitzender des Prüfungsausschusses:	Dr. Georg Steinbrück
Vorsitzender des Promotionsausschusses:	Prof. Dr. Peter Hauschildt
Leiterin des Fachbereichs Physik:	Prof. Dr. Daniela Pfannkuche
Dekan der MIN-Fakultät:	Prof. Dr. Heinrich Graener

Abstract

In this thesis a search for new physics is presented, based on events with two isolated same-sign leptons (muons and electrons), jets, and missing transverse energy. The analyzed data corresponds to an integrated luminosity of 1.1 fb^{-1} and is collected with the CMS detector in proton-proton collisions at the LHC with a center-of-mass energy of 7 TeV. A counting experiment is established by developing an event selection in order to select signal events in three different channels ($\mu\mu$, $e\mu$, and ee), while rejecting the Standard Model background as efficiently as possible. Standard Model background contributions are classified into several categories and estimated using different techniques. A total of 7 events is observed in data, being in agreement with the Standard Model background prediction of $11 \pm 1(\text{stat.}) \pm 2.7(\text{syst.})$ events. Hence, no evidence for new physics is found. The observations are utilized to constrain the cMSSM by setting upper exclusion limits on possible signal contributions.

The technical part of this theses contributes to the planned upgrade of the CMS experiment, by giving a motivation for a certain readout design of the hadronic calorimeter, in order to improve the performance of the detector in future.

Kurzfassung

In der vorliegenden Arbeit wird eine Suche nach neuer Physik vorgestellt, die auf der Selektion von Ereignissen mit zwei isolierten, gleich geladenen Leptonen (Myonen und Elektronen), Jets und fehlender transversaler Energie basiert. Die analysierten Daten wurden mit dem CMS Detektor in Proton-Proton Kollisionen am LHC bei einer Schwerpunktsenergie von 7 TeV aufgenommen und entsprechen einer integrierten Luminosität von 1.1 fb^{-1} . Ein Zählexperiment wird durchgeführt, indem eine Ereignisselektion mit dem Ziel entwickelt wird, Signalereignisse in drei unterschiedlichen Kanälen zu selektieren ($\mu\mu$, $e\mu$ und ee), während der Standardmodell-Untergrund so effektiv wie möglich unterdrückt wird. Standardmodell-Untergrundsbeiträge werden kategorisiert und mit Hilfe überwiegend datengetriebener Techniken abgeschätzt. Insgesamt 7 Ereignisse werden in den Daten beobachtet, was in Übereinstimmung mit der Standardmodell-Untergrundsvorhersage von $11 \pm 1(\text{stat.}) \pm 2.7(\text{syst.})$ Ereignissen steht. Da keine Evidenz für neue Physik gefunden wurde, wird die Beobachtung dazu verwendet, um das cMSSM zu beschränken und obere Ausschlussgrenzen für ein mögliches Signal zu setzen.

Der technische Teil dieser Arbeit trägt zum geplanten Upgrade des CMS Experiments bei, indem ein bestimmtes Auslesedesign des hadronischen Kalorimeters motiviert wird, um die Leistungsfähigkeit des Detektors in Zukunft zu verbessern.

Contents

1	Introduction	1
2	Particle Physics	3
2.1	Standard Model of Particle Physics	3
2.1.1	Particles	3
2.1.2	Interactions	4
2.2	Open Questions	7
2.3	Supersymmetry	9
2.4	Supersymmetric Breaking and the Constrained Supersymmetric Model . .	12
2.5	Searches for Supersymmetry at the LHC	14
3	Experimental Setup	17
3.1	The Large Hadron Collider	17
3.2	The Compact Muon Solenoid	20
3.2.1	Silicon Tracking System	22
3.2.2	Electromagnetic Calorimeter	24
3.2.3	Hadronic Calorimeter	24
3.2.4	Superconducting Magnet	26
3.2.5	Muon System	28
3.2.6	Trigger System	29
4	Energy Weighting for the Upgrade of the Hadronic Calorimeter	33
4.1	Methodology	34
4.2	Implementation	35
4.2.1	The Geant3 Standalone Simulation	36
4.2.2	Investigated Readout Schemes	37
4.2.3	Calibration	37
4.2.4	Determination of the Weighting Factors	39
4.3	Application of the Energy Weighting Method	42
4.3.1	Average Energy Density	42
4.3.2	Energy Resolution and Linearity	42
4.4	Next Steps for Future Investigations	46
5	Event Simulation, Reconstruction, and Selection	47
5.1	Event Simulation and Simulated Processes	47
5.2	Pile-up Weighting	50
5.3	Used Data and Trigger Strategy	52
5.4	Physics Objects Reconstruction in CMS	54

5.4.1	Muon Reconstruction	55
5.4.2	Electron Reconstruction	57
5.4.3	Jet Reconstruction	61
5.4.4	Vertex Reconstruction	64
5.4.5	Missing Transverse Energy	64
5.4.6	H_T	65
5.5	Preselection	65
5.6	Event Selection	66
5.7	Event Yield	67
6	Background	71
6.1	Background Contributions	71
6.2	Background Estimation Methods	74
6.2.1	Extrapolation Method	74
6.2.2	t-Tag-and-Probe Method	75
6.2.3	b-Tag-and-Probe Method	76
6.3	Tight-to-Loose Method	78
6.3.1	Fake-ratio Estimation	80
6.3.2	Prompt-ratio Estimation	81
6.3.3	Background Prediction	84
6.4	Charge Flip	87
6.5	Irreducible Background	90
7	Uncertainties	91
7.1	Jet Energy Scale	91
7.2	Lepton Reconstruction	93
7.3	Parton Distribution Functions	94
7.4	Luminosity	94
7.5	Trigger	95
7.6	Pile-up	95
7.7	Systematic Uncertainties of the Background Estimation	96
7.7.1	Irreducible Background	96
7.7.2	Tight-to-Loose Method	97
7.7.3	Charge Flip	99
8	Result	101
8.1	Summary of the Event Yield	101
8.2	Exclusion Limit for the cMSSM	102
9	Summary and Outlook	105
A	Additional Investigations for the Energy Weighting of the HCAL	107
A.1	Shower profiles	107
A.2	Impact of a readout channel defect	108
A.3	Weighting without the first layer	112
A.4	Weighting without ECAL	112

A.5	Weighting with readout design 17x1	113
B	Details of the Event Selection Flow	115
C	Details of Uncertainty Studies	121
C.1	Trigger Efficiencies	121
C.2	Jet Energy Scale	122
C.3	Parton Distribution Functions	124
C.4	Pile-up	125
C.5	TL Method	126
D	Examples for Rare Standard Model Processes	129
E	Consistency Check of the Background Prediction	131
F	Limit Setting Procedure	133
	List of Figures	135
	List of Tables	139
	Bibliography	141
	Acknowledgements	151

1 Introduction

*“Discovering particles
takes a certain amount of concentration.”
— Bill Bryson*

One of the most challenging scientific endeavors is the discovery and investigation of the building blocks of nature. The Standard Model (SM) of particle physics summarizes today’s grasp for the elementary particles and the interactions between them. However, there are still unanswered questions which puzzle physicists since several decades. Hence, huge efforts are made to build larger and larger machines to solve some of the riddles, like e.g. the nature of the dark matter, the existence of supersymmetric particles, or the behavior of a quark-gluon plasma, similar to the state of the early universe.

Pushing technological limits onward, with the Large Hadron Collider (LHC), higher energies than ever before are reached at the realm of the Terascale. It is a tool which opens a new era for investigating the fundamental particles of our world. The LHC machine performs very well and recent studies of the CMS and ATLAS collaboration show first hints for the existence of the Higgs boson [1–3].

In this analysis, a search for physics beyond the Standard Model (SM) is performed, leading to a same-sign di-lepton signature ($\mu\mu, e\mu, ee$), associated with jets and missing transverse energy. This search is strongly motivated by the fact, that this signature is very rare in the SM, but appears naturally in many new physics (NP) scenarios. Hence, this search has a good discovery potential for NP.

As an example, here, the search is interpreted in the context of the constrained Minimal Supersymmetric Model (cMSSM). However, it is more general: Every model, leading to a same-sign di-lepton signature can be probed by the presented analysis. Alternative NP models which also lead to this signature are e.g. universal extra dimensions [4], same-sign top quark production from flavor changing neutral currents in the top sector [5, 6], pair production of $T_{5/3}$ (a fermionic partner of the top quark) [7], heavy Majorana neutrinos [8], and same-sign top-pair resonances as predicted in theories with warped extra dimensions [9].

The principle of the presented analysis is straight forward: It is a counting experiment, in which the number of events in the signal region is observed. By estimating the SM background contribution to this region, it is possible to judge, if an excess is found in data, or if a model can be excluded for a certain region of its parameter space.

This thesis is organized as follows: In Chapter 2 the SM of particle physics is introduced and open questions are discussed. The extension of the SM to the supersymmetric model is pointed out and searches for Supersymmetry at the LHC are depicted with an emphasis on the same-sign di-lepton topology.

The experimental setup, the LHC along with its experiments, are introduced in Chapter 3. Especially the CMS detector and its components that are relevant for this thesis are described in detail.

In Chapter 4 the technical part of this thesis is presented: A simulation study is performed for the planned upgrade of the CMS experiment, with the aim to optimize the readout of the hadronic calorimeter in order to improve its performance in future. The applied energy weighting method is documented and its results are shown.

The search for NP is described in detail in Chapter 5 by documenting the trigger strategy and the object and event reconstruction, as well as a description of the event simulation, used data, and the simulated processes.

Chapter 6 contains detailed information about background estimation methods with an emphasis on the Tight-to-Loose method, used in the presented analysis.

The investigation of the systematic uncertainties can be found in Chapter 7.

In Chapter 8 the event yields for the different decay channels and their combination are displayed, as well as the determination of an exclusion limit for the cMSSM.

Finally, the summary of the results can be found in Chapter 9. An outlook for possible extensions of the analysis closes this Chapter.

2 Particle Physics

The Standard Model of particle physics is introduced in this Chapter and open questions are addressed. They can mostly be answered by the constrained Minimal Supersymmetric Model (cMSSM), which is probed in this analysis. Finally, the search for new particles at the Large Hadron Collider, predicted by the cMSSM, is discussed with a focus on the same-sign di-leptonic topology, since that is the subject of this analysis.

2.1 Standard Model of Particle Physics

Today’s knowledge about particle physics is represented by the Standard Model (SM) of particle physics [10, 11]. It describes the observed building blocks of matter of the universe and the interactions between them, namely the electromagnetic, the weak, and the strong interaction. The fourth interaction, gravity, is not included in the SM.¹ However, gravity is many orders of magnitude weaker than the other three interactions and can be neglected in particle physics processes. More detailed descriptions of the SM than the one presented in the following Sections can be found in various textbooks, e.g. [12].

2.1.1 Particles

Figure 2.1 shows an illustration of the particles in the SM, their names, and some of their properties. In the first three columns, 12 different fermions are displayed. They are classified into 6 charged quarks (purple) and 6 leptons (green), of which three are charged. Quarks interact through the strong, the electromagnetic, and the weak force. Charged leptons interact electromagnetically and weakly, uncharged leptons only weakly. The three columns represent three generations, each consisting of two quarks (one “up-type” quark with a positive electric charge and one “down-type” quark with a negative electric charge), one charged lepton, and one uncharged lepton, the neutrino. The particles of each generation differ only by their masses — particles of the first generation are lighter than the corresponding particles of the second generation, and so on². The first-generation particles are stable and do not decay into other particles. Hence, all ordinary matter consists of such particles.

Each fermion has a corresponding antifermion, which differs only by its opposite charge. The spin S , or intrinsic angular momentum, of a particle is a fundamental property. All fermions have a half integer spin of $S = 1/2$, which implies that they satisfy the Pauli exclusion principle [13]: Two fermions cannot occupy the same state. As a consequence,

¹There have been efforts to develop a quantum field theory of gravity to include it into the SM, which have not yet been successful.

²Note, that the mass hierarchy of the neutrinos has not been measured yet.

$$\psi(\vec{r}_1, S_1; \vec{r}_2, S_2) = -\psi(\vec{r}_2, S_2; \vec{r}_1, S_1) \Rightarrow \psi(\vec{r}_1, S_1; \vec{r}_1, S_1) = 0. \quad (2.1)$$

The right most column of Fig. 2.1 represents the gauge bosons, the carriers of the forces. The massless photon (γ) mediates the electromagnetic force, the gluon (g) the strong force, and the massive Z^0 - and $W^\pm$ bosons the weak force, respectively. They have a spin of $S = 1$ and do not satisfy the Pauli exclusion principle.

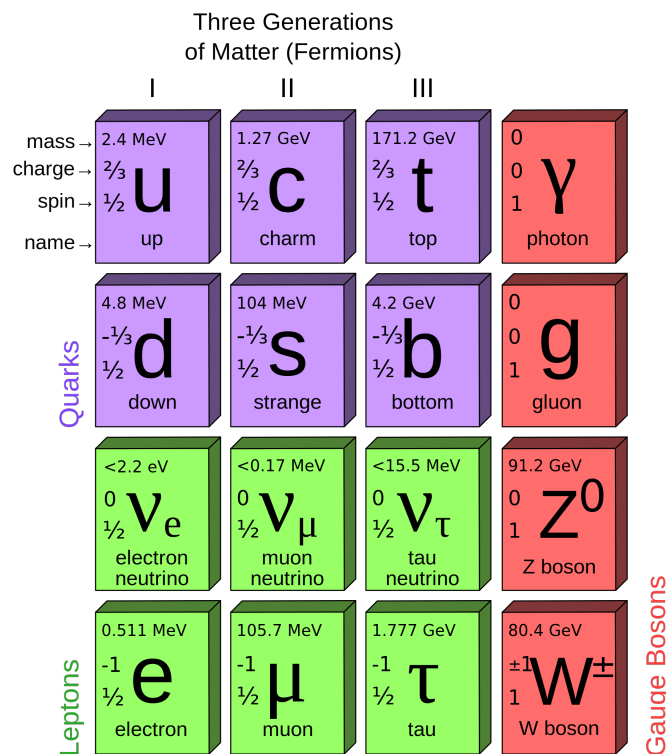


Figure 2.1: Representation of the SM particles. Its particles and some of their properties are illustrated. *Purple*: quarks, *green*: leptons, *red*: gauge bosons. The three generations are indicated in the first three columns. The last column displays the gauge bosons. From [14].

2.1.2 Interactions

Particles influence other particles via interactions. Table 2.1 gives a short summary of the four fundamental forces and some of their properties.

The first three forces are described by quantum field theories. The theory describing the electromagnetic interaction is called Quantum Electrodynamics (QED). It is unified

Table 2.1: The fundamental forces, their mediators, and some of their properties. Gravitation is not included in the SM.

Force	Mediator	Rel. strength	Range	Mass
Strong	8 gluons g	1	10^{-15} m	$0 \text{ GeV}/c^2$
Electromagnetic	photon γ	10^{-2}	∞	$0 \text{ GeV}/c^2$
Weak	$W^\pm, Z^0$	10^{-6}	10^{-18} m	$\approx 80, 91 \text{ GeV}/c^2$
Gravitation	graviton G (?)	$< 10^{-40}$	∞	$0 \text{ GeV}/c^2$

with the weak force resulting in the electroweak force. The strong interaction is described by Quantum Chromo Dynamics (QCD).

The field theories rely on certain gauge symmetries. According to Noether's theorem, symmetries imply conserved quantities. Therefore, the gauge symmetries of the SM lead to invariant quantum numbers, identified as charges of the interacting particles.

Electromagnetic Interaction

The electromagnetic interaction is the consequence of gauge invariance under $U(1)$ transformations. The conserved quantity is the electric charge. The coupling constant α is given by

$$\alpha = \frac{e^2}{4\pi\hbar c} \approx \frac{1}{137.036\dots} \ll 1, \quad (2.2)$$

where e is the electric charge in units of electron charge, $\hbar$ the reduced Planck constant, and c the speed of light in vacuum.

QED is one of the most precise tested theories in physics. The small value of α allows to perform perturbative calculations up to high orders. This can be illustrated via a quantity called “ $g - 2$ ”, which refers to an anomaly of the magnetic moment $\mu = g\mu_B S$ of a Dirac lepton, where μ_B is the Bohr magneton, $S = 1/2$ the spin of the lepton, and g equals two in the Dirac theory. However, there are higher order corrections to the magnetic moment.

To illustrate what these higher order corrections are, in Fig. 2.2 a leading order (LO) and a next-to-leading order (NLO) Feynman diagram for electron-positron scattering in the s -channel are shown. This also demonstrates, how the coupling constant α enters physics processes.

At each vertex, the coupling constant enters the calculation of the cross-section for the process. Hence, for NLO processes α enters quadratically, for next-to-next-to-leading order (NNLO) processes as a power of three, etc. Since α is small, higher order corrections have, in general, a small contribution to the total cross-section.

For the calculation of the magnetic moment of Dirac electrons, $(g - 2)/2$ represents the anomaly due to higher order corrections. From theory

$$\left(\frac{g - 2}{2}\right)^{\text{theory}} = (115965218279 \pm 771) \cdot 10^{-14} \quad (2.3)$$

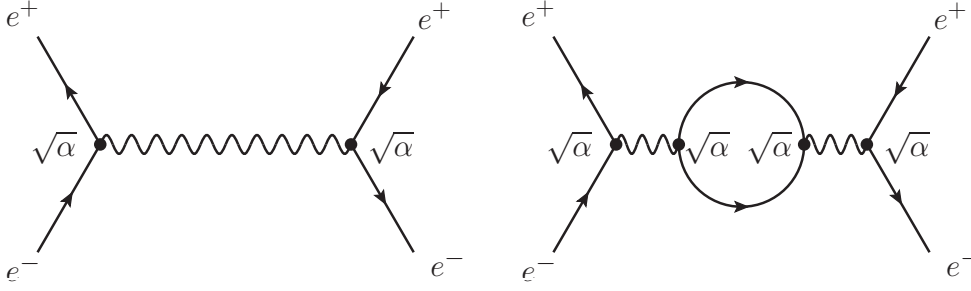


Figure 2.2: Feynman diagram for electron-positron scattering in the s -channel. *Left:* Leading Order (LO), *right:* Next-to-Leading Order (NLO). At the vertices the coupling constant α enters.

is obtained [15]. Comparing the experimental measured value [16]

$$\left(\frac{g-2}{2}\right)^{\text{expt}} = (115965218073 \pm 28) \cdot 10^{-14} \quad (2.4)$$

gives a remarkable agreement to the theoretical prediction.³ This illustrates the predictive power of quantum field theory.

Strong Interaction

The strong interaction is the consequence of gauge invariance under $SU(3)$ transformations. The conserved quantity is the *color charge*. Three different color eigenstates exist, called *red* (r), *green* (g), and *blue* (b). The wave functions of gluons are $3^2 - 1 = 8$ linear combinations of color-anticolor pairs. Gluons couple to colored particles — this also implies that they couple to themselves.

In analogy to the electromagnetic interaction, the strong coupling constant α_s is given by

$$\alpha_s = \frac{g_s}{4\pi} \approx 1 \text{ (for } Q \propto \Lambda_{\text{QCD}}), \quad (2.5)$$

where g_s is defined as the strong charge of the constituent quarks. The strong coupling constant depends on the momentum transfer Q^2 . At low Q^2 (corresponding to large distances of $> \mathcal{O}(10^{-15} \text{ m})$) α_s becomes large. Consequently, gluons and quarks are not observed in nature as free particles. Colorless compositions of quarks are built, namely mesons, containing one color-anticolor pair of quarks, and baryons, made of three colored quarks (or three anticolored antiquarks), which combine to a colorless state. This phenomenon is called *confinement*.

Similarly, at large Q^2 values (corresponding to small distances of $< \mathcal{O}(10^{-15} \text{ m})$) α_s becomes small. This phenomenon is called *asymptotic freedom*.

Due to confinement a *hadronization* of quarks and gluons occurs in particle collisions. When a quark is separated at a large enough distance from its original hadron, it becomes energetically more favorable to create a new quark pair, which combines to a colorless state.

³The larger error in the predicted number originates from the experimental uncertainty in the value of α .

Gluons also share this property of being confined within hadrons. One consequence is that gluons are not directly involved in the nuclear forces between hadrons. Instead mesons are built as force mediators.

Weak Interaction

The weak interaction is the consequence of gauge invariance under $SU(2)$ transformations. The conserved quantity is the *weak isospin* T . The weak interaction acts on all twelve SM particles, which are arranged into pairs, or isospin *doublets*:

$$\begin{pmatrix} u_L \\ d'_L \end{pmatrix}; \quad \begin{pmatrix} c_L \\ s'_L \end{pmatrix}; \quad \begin{pmatrix} t_L \\ b'_L \end{pmatrix}; \quad \begin{pmatrix} \nu_e \\ e_L \end{pmatrix}; \quad \begin{pmatrix} \nu_\mu \\ e_\mu \end{pmatrix}; \quad \begin{pmatrix} \nu_\tau \\ e_\tau \end{pmatrix}, \quad (2.6)$$

where the index L refers to the left-handedness of the particles. Note, that down-type quarks of the isospin doublets (d' , s' , and b') are linear combinations of the corresponding eigenstates (d , s , and b). Hence, the mass eigenstates of quarks differ from their weak eigenstates, allowing that quarks from one generation can transform into those of another. This is realized by charged currents, converting up-type quarks into down-type quarks (and vice versa) by coupling to a $W^\pm$. The mixing of these eigenstates is given by the Cabibbo-Kobayashi-Maskawa (CKM) matrix [17]

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}. \quad (2.7)$$

It quantifies the transition probability between different quark flavors. In the SM the CKM matrix is assumed to be unitary.

Further characteristics of the weak interaction are: It violates parity [18] since it couples only to left-handed fermions and right-handed antifermions. Neutral currents, explicable by the existence of Z bosons, have been detected with the Gargamelle experiment [19] in reactions between neutrinos and nuclei. Neutrino oscillations have been observed in various experiments [20, 21]. They imply that neutrinos are not massless and different neutrino flavors mix, according to the corresponding mixing matrix, the Pontecorvo-Maki-Nakagawa-Sakata (PMNS) matrix [22].

2.2 Open Questions

Although the SM describes well the realm of particle physics, there are still open questions.

Higgs Boson

The missing piece of the SM is the Higgs boson (H), which has not yet been discovered. It is a hypothetical massive scalar elementary particle with a spin of $S = 0$. The Higgs mechanism is currently the favored theory for the spontaneous breaking of the electroweak gauge symmetry. The corresponding Higgs field has a non-zero vacuum expectation value. It gives mass to every particle that couples to the Higgs field, including the Higgs boson itself. Therefore, the mass of the gauge bosons of the weak interaction can be explained.

Limits have been set on the Higgs boson mass m_H : A lower limit of $114 \text{ GeV}/c^2 < m_H$ is obtained from direct experimental exclusion by experiments of the LEP accelerator. An upper limit was estimated by theoretical calculations to be $m_H < 1 \text{ TeV}/c^2$ [12].⁴ An additional exclusion in the range of $147 < m_H < 179 \text{ GeV}/c^2$ has been performed by the experiments of the Tevatron accelerator [23]. The Higgs boson mass is connected to the mass of the W boson and the top-quark. If the mass of any two of the three particles is known, then the mass of the third particle can be calculated. Figure 2.3 depicts the mass of the Higgs boson as a function of the top-quark and W boson mass.

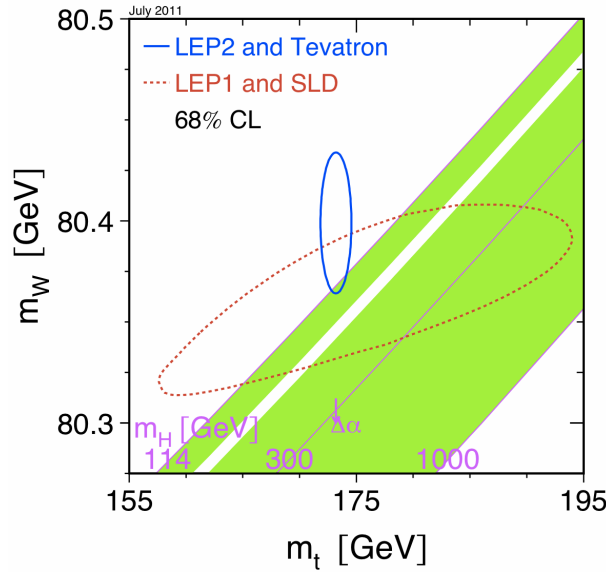


Figure 2.3: Higgs mass as a function of the top-quark and W boson mass. *Diagonal green lines* represent a single Higgs boson mass. *Blue solid ellipse*: most likely values for the top-quark and W boson masses, based on all available experimental measurements from the Tevatron and LEP, *red dotted ellipse*: top-quark and W boson masses, based on Lep data (run 1) and data from, both at 68 % C. L. From [24].

Each diagonal line represents a single possible Higgs boson mass which is not yet excluded. The blue solid ellipse indicates the most likely values for the top-quark and W boson masses at a 68 % percent confidence level, based on all available experimental measurements from the Tevatron and LEP. From this plot it can be seen that a light Higgs mass (at about $120 \text{ GeV}/c^2$) is favored.

⁴For this estimation, the WW scattering is considered. The cross-section of $WW \rightarrow Z, \gamma \rightarrow WW$ diverges for collisions with large center-of-mass energies. This can be compensated by the Higgs boson via the analogous process, $WW \rightarrow H \rightarrow WW$ as long as it fulfills

$$m_H < \left(\frac{8\pi\sqrt{2}}{2G_F} \right)^{1/2} \approx 1 \text{ TeV}/c^2, \quad (2.8)$$

where G_F is the Fermi coupling constant.

Recent combined searches for the SM Higgs Boson of the CMS and ATLAS collaboration [1–3] show hints for an excess of events for a hypothesized Higgs boson mass of about $m_H \approx 125 \text{ GeV}/c^2$. The SM Higgs boson can also be excluded at a 99 % confidence level in the mass range of $128 - 525 \text{ GeV}/c^2$ [3]. Similar results have been observed at the experiments of the Tevatron [23].

Hierarchy Problem

The electroweak scale $\mathcal{O}(10^2 \text{ GeV})$ and the GUT scale⁵ $\mathcal{O}(10^{16} \text{ GeV})$ are separated by 14 orders of magnitude. On the one hand, the mass scale of the Higgs boson and the gauge bosons, responsible for the weak interaction, is of the order of $\mathcal{O}(10^2 \text{ GeV})$. On the other hand, the Higgs boson is a scalar particle and therefore, there are quadratic loop corrections to the Higgs boson mass of the order of the GUT scale (see also Section 2.3). This discrepancy of the large difference of scales is known as the so-called hierarchy problem [25].

Fine-tuning Problem

To solve the hierarchy problem within the SM, quadratically divergent loop corrections to the Higgs boson mass should be controlled at the Planck scale. An enormous degree of *fine-tuning* of the SM parameters is necessary in order to get a Higgs boson mass which is close to the electroweak scale. This is known as the fine-tuning problem.

Dark Matter

Astrophysical observations have been performed indicating that the baryonic matter only contributes with about 4.6 % to the total mass of the visible universe [26]. About 23 % is made of “dark matter” and the remaining 72 % out of “dark energy”. Neutrinos and black holes can only explain a small fraction of the dark matter. Dark matter interacts only by its gravitation and has been observed indirectly from e.g. rotation curves of galaxies or the behavior of two colliding galaxies.

Unification of the Interactions

An extrapolation of the coupling constants to high energies can be performed due to their running behavior. The unification of the electromagnetic, weak, and strong interaction is one major goal of particle physics. However, it can be seen, that in the SM the three interactions do not meet in one point at high energies (see Figure 2.4), whereas in a Supersymmetric model they converge.

2.3 Supersymmetry

At present, Supersymmetry (SUSY) [28] is the most popular theory of physics beyond the SM. It predicts a fundamental symmetry between matter and interactions by introducing

⁵The GUT (Grand Unification Theory) scale $\Lambda_{\text{GUT}} \approx 10^{16} \text{ GeV}$ is the scale where the electroweak and the strong interaction become equal in strength.

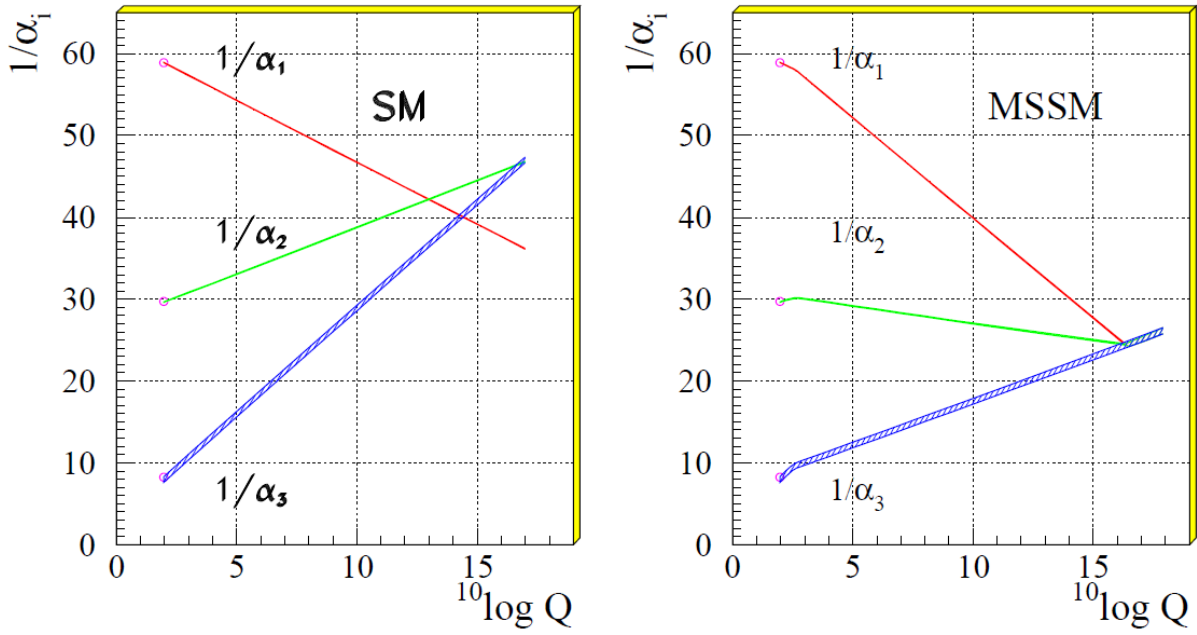


Figure 2.4: Running behavior of the electromagnetic (*blue*), weak (*red*), and strong (*blue*) interaction as a function of energy. *Left:* SM, *right:* supersymmetric model. From [27].

a mechanism to convert bosons into fermions and vice versa

$$Q|\text{Fermion}\rangle = |\text{Boson}\rangle; \quad Q|\text{Boson}\rangle = |\text{Fermion}\rangle, \quad (2.9)$$

where the operators Q are the generators of SUSY. The generators Q are related (via anti-commutation) to the Poincaré-group which opens the possibility to include gravity into the theory. SUSY also solves the hierarchy and fine-tuning problem by the cancellation of loop corrections due to the introduction of new particles. Figure 2.5 (*left*) shows a correction to the Higgs mass m_H from a loop containing a Dirac fermion f with mass m_f .

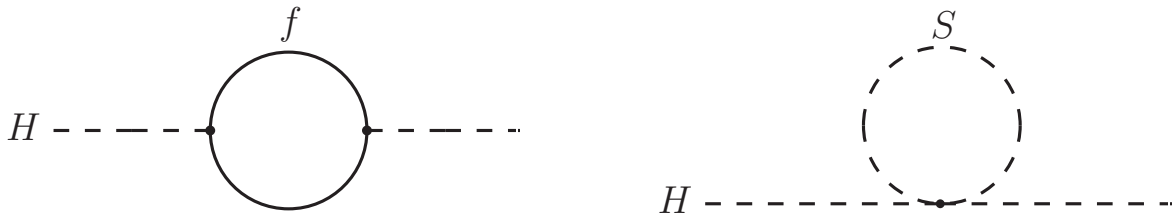


Figure 2.5: One-loop quantum corrections to the Higgs squared mass parameter m_H^2 , due to a Dirac fermion (*left*) and a scalar S (*right*).

The correction yields

$$\Delta m_H^2 = -\frac{|\lambda_f|^2}{8\pi^2} \Lambda_{\text{UV}}^2 + \dots, \quad (2.10)$$

where Λ_{UV} is ultraviolet momentum cutoff used to regulate the loop integral and λ_f represents the Yukawa coupling for each fermion. As can be seen from equation (2.10), the quantum loop correction to m_H^2 has a quadratic sensitivity to the energy scale cutoff Λ_{UV} . This is a problem, because if Λ_{UV} is at the order of the Planck scale ($\mathcal{O}(10^{19} \text{ GeV})$), the quantum correction to m_H^2 is some 30 orders of magnitude larger than the required value of about $-(100 \text{ GeV})^2$ (see [28]). Fermions also suffer indirectly from this problem, since they couple directly to the Higgs boson via the Yukawa coupling. However, if SUSY is a property of the universe, there are additional loop corrections to m_H^2 (Fig. 2.5, right)

$$\Delta m_H^2 = \frac{\lambda_S}{16\pi^2} [\lambda_{\text{UV}}^2 - 2m_S^2 \ln(\lambda_{\text{UV}}/m_S) + \dots], \quad (2.11)$$

where the index S refers to a scalar particle with a mass of m_S and a Yukawa coupling of λ_{UV} . In SUSY each of the quarks and leptons of the SM is accompanied by a pair of complex scalars with $\lambda_S = 2|\lambda_f|^2$, which neatly cancel out the loop corrections (compare the relative minus sign between equations (2.10) and (2.11) of the leading term).

Another feature of SUSY is its impact on the running behavior of the coupling constants, which permits unification of the interactions at high energies. This is illustrated in Fig. 2.4. SUSY also provides a natural candidate for dark matter (see Section 2.4). Hence, there is a strong motivation to search for SUSY particles.

The SUSY extension of the SM introduces a large set of free parameters whose values are not known, resulting in a large number of possible SUSY models. In the Minimal Supersymmetric Standard Model (MSSM) each SM fermion (boson) has a bosonic (fermionic) superpartner. The following naming convention is used:

- Supersymmetric bosons are denoted with an “s” in front (e.g. sparticles, squarks, sleptons, selectron, etc.). They have a spin of $S = 0$.
- Supersymmetric fermions are denoted with an “ino” at the end (e.g. gaugino, gluino, wino, etc.). They have a spin of $S = 1/2$.
- A tilde represents supersymmetric particles in the abbreviation with letters.

In the MSSM the SM Higgs sector must be extended, too, since otherwise new divergences would occur.⁶ Hence, two complex Higgs doublets are introduced

$$\begin{pmatrix} H_2^+ \\ H_2^0 \end{pmatrix}; \quad \begin{pmatrix} H_1^0 \\ H_1^- \end{pmatrix}, \quad (2.12)$$

where the first (second) doublet has a hypercharge of $Y = 1$ ($Y = -1$). The Higgs fields mix to the five observable Higgs bosons

$$h^0, H^0, A, H^+, H^-, \quad (2.13)$$

where the first two neutral Higgs bosons have even CP-Parity (h^0 being the lighter one).

⁶In e.g. Compton scattering, a contribution of higher order are triangle diagrams. The Higgsino, having a hypercharge of $Y = 1$, contributes to these diagrams. However, in order to avoid divergences, $\sum_{\text{fermions}} Y = 0$ must be fulfilled. This is only possible if a new Higgs doublet is introduced, which has a hypercharge of $Y = -1$.

The analogous superpartners of the Higgs fields are the Higgsinos $\tilde{H}_2^+$, $\tilde{H}_2^0$, $\tilde{H}_1^0$, and $\tilde{H}_1^-$, having a spin of $S = 1/2$. The charged Higgsinos $\tilde{H}_2^+$ and $\tilde{H}_1^-$ mix with the Winos $\tilde{W}^\pm$ to four *charginos* $\tilde{\chi}_1^\pm$ and $\tilde{\chi}_2^\pm$. Equally, the uncharged Higgsinos $\tilde{H}_1^0$ and $\tilde{H}_2^0$ mix with the neutral gauginos $\tilde{W}^0$ and $\tilde{B}^0$ to four uncharged *neutralinos* $\tilde{\chi}_1^0$, $\tilde{\chi}_2^0$, $\tilde{\chi}_3^0$, and $\tilde{\chi}_4^0$. Table 2.2 summarizes the particles of the MSSM.

Table 2.2: Summary of SUSY particles in the MSSM and the extended Higgs sector.

Names	Gauge eigenstates	Mass eigenstates	Spin
squarks	$\tilde{u}_L \tilde{u}_R \tilde{d}_L \tilde{d}_R$	(same)	0
	$\tilde{c}_L \tilde{c}_R \tilde{s}_L \tilde{s}_R$	(same)	
	$\tilde{t}_L \tilde{t}_R \tilde{b}_L \tilde{b}_R$	$\tilde{t}_1 \tilde{t}_2 \tilde{b}_1 \tilde{b}_2$	
sleptons	$\tilde{e}_L \tilde{e}_R \tilde{\nu}_e$	(same)	0
	$\tilde{\mu}_L \tilde{\mu}_R \tilde{\nu}_\mu$	(same)	
	$\tilde{\tau}_L \tilde{\tau}_R \tilde{\nu}_\tau$	$\tilde{\tau}_1 \tilde{\tau}_2 \tilde{\nu}_\tau$	
neutralinos	$\tilde{B}^0 \tilde{W}^0 \tilde{H}_u^0 \tilde{H}_d^0$	$\tilde{\chi}_1^0 \tilde{\chi}_2^0 \tilde{\chi}_3^0 \tilde{\chi}_4^0$	1/2
charginos	$\tilde{W}^\pm \tilde{H}_u^\pm \tilde{H}_d^\pm$	$\tilde{\chi}_1^\pm \tilde{\chi}_2^\pm$	1/2
gluinos	$\tilde{g}$	(same)	1/2
Higgs bosons	$H_u^0 H_d^0 H_u^+ H_d^-$	$h^0 H^0 A^0 H^\pm$	0

Taking into account the number of new particles, fields, mixing-angles, and phases introduced by the MSSM, the SM is extended by a total of 105 parameters. It is not feasible to cope with such a large number of new parameters. Therefore, a more constrained MSSM is necessary in order to make concrete predictions which can be probed in an experiment.

2.4 Supersymmetric Breaking and the Constrained Supersymmetric Model

So far, no SUSY particles have been observed. Hence, the mass of SUSY particles must be larger than the mass of SM particles — otherwise SUSY particles would have been discovered already. Therefore, SUSY must be broken resulting in sparticle masses above the current observation limits.

Various breaking scenarios are possible. In this analysis, the prominent scenario of the constrained Minimal Supersymmetric Model (cMSSM) is considered which is split up into a visible sector (including SM and SUSY particles) and a hidden sector, where the breaking takes place, mediated by gravity. In the cMSSM soft SUSY breaking terms are naturally generated, where “soft” refers to the aspect that no quadratic divergences are introduced, stabilizing the Higgs sector. As a consequence, the renormalization group equations result in radiative breaking of electroweak symmetry.

Due to its predictive power, the cMSSM is one of the most widely investigated models of particle physics. The 105 parameters of the MSSM are reduced to 5 remaining parameters:

- m_0 : Unified mass of sfermions at the GUT scale.
- $m_{1/2}$: Unified mass of gauginos at the GUT scale.
- A_0 : Trilinear coupling of the scalar superfields at the GUT scale.
- $\tan \beta$: Ratio of the vacuum expectation values of the Higgs doublets.
- μ : Sign of the Higgs mass parameter.

Several “Low Mass” (LM) benchmark points have been defined in the CMS collaboration [29]. In Table 2.3 some points are listed.

Table 2.3: “Low Mass” (LM) benchmark points of the cMSSM of the CMS collaboration. For all points, it is $\text{sign}(\mu) = +1$.

Benchmark	m_0 [GeV/c ²]	$m_{1/2}$ [GeV/c ²]	A_0 [GeV]	$\tan \beta$
LM0	200	160	−400	10
LM1	60	250	0	10
LM2	185	350	0	35
LM3	330	240	0	20
LM4	210	285	0	10
LM5	230	360	0	10
LM6	85	400	0	10
LM7	3000	230	0	10
LM8	500	300	−300	10
LM9	1450	175	0	50
LM10	3000	500	0	10
LM11	250	325	0	35
LM12	2544.58	246.564	−865.752	47.5897
LM13	270	218	−553	40

In SUSY models, baryon number and lepton number are not conserved by all of the renormalizable couplings in the theory, which allows the proton to decay.⁷ R-parity R is a symmetry acting on the MSSM fields that forbids these couplings and is defined as

$$R = (-1)^{3B-L-2S}, \quad (2.14)$$

where B is the baryon number (1/3 for quarks, −1/3 for antiquarks), L is the lepton number (1 for leptons, −1 for antileptons), and S is the spin quantum number. For SM particles the R-parity is $R = +1$ and for SUSY particles $R = -1$, respectively. If R-parity is conserved, SUSY particles are always produced in pairs (there must be an even number of SUSY particles at a vertex), and the Lightest Supersymmetric Particle (LSP) is stable, since it cannot decay into a (lighter) SM particle. If the LSP is electromagnetically neutral, it only interacts weakly and cannot be observed with the CMS detector. In this analysis, R-parity is assumed to be conserved.

⁷However, a lower limit of the proton half-life of $> 6.6 \cdot 10^{33}$ years has been set experimentally [30].

2.5 Searches for Supersymmetry at the LHC

In many SUSY scenarios sparticles have a mass at the TeV energy scale. Hence, they can be produced and searched for at the Large Hadron Collider (LHC).

Production of Sparticles

The initial partons are gluons or quarks from the colliding protons, where for the energies at the LHC the gluon production dominates. In Figure 2.6 the most dominant production mechanisms of SUSY particles via the strong interaction are shown. Some diagrams will dominate over others, depending on the mass hierarchy of the gluino with respect to the squarks. Three different cases can be considered:

- $m_{\tilde{g}} \gg m_{\tilde{q}}$: The production of two squarks is dominant (Fig. 2.6 (a) and (c)).
- $m_{\tilde{g}} \ll m_{\tilde{q}}$: The production of two gluinos is dominant (Fig. 2.6 (b) and (d)).
- $m_{\tilde{g}} \approx m_{\tilde{q}}$: All diagrams in Fig. 2.6 have non-trivial contributions to the sparticle production.

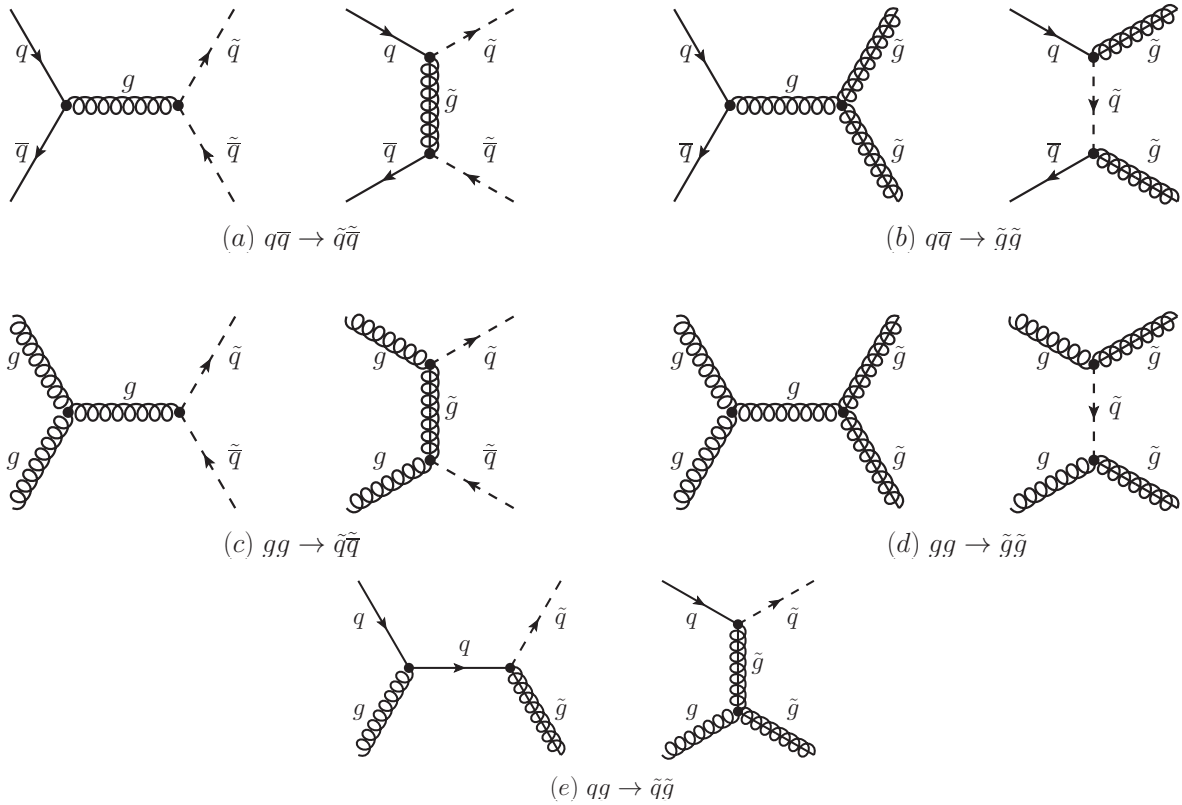


Figure 2.6: Some LO Feynman diagrams of SUSY particle production. The s - and t -channel are shown.

In the case that the squarks and gluinos are too heavy to be produced at the LHC, lighter superpartners (neutralinos, charginos) can be produced directly via the electroweak

interaction (see Fig. 2.7). Depending on the mass spectrum of the colored superpartners, these diagrams may be negligible or significant.

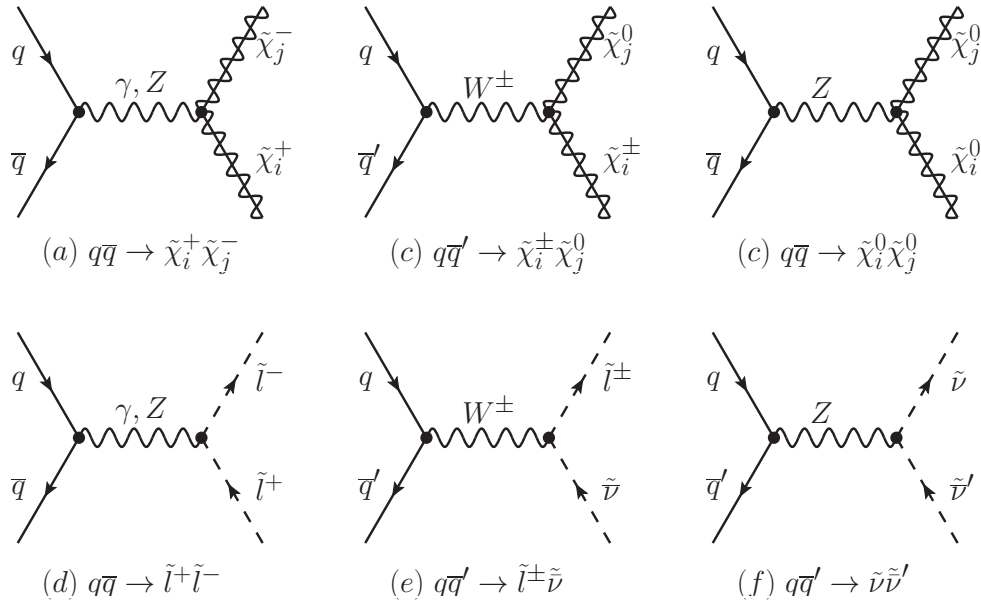


Figure 2.7: Dominant weak production mechanisms of SUSY particles. *Top*: production of charginos and neutralinos, *bottom*: production of sleptons and sneutrinos.

Decay of Sparticles

Figure 2.8 illustrates a possible decay of a gluino and a squark, respectively.

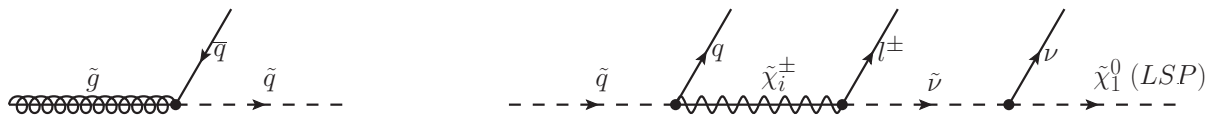


Figure 2.8: Possible decay of a gluino (*left*) and a squark (*right*).

If R-parity is conserved, at one vertex always two SUSY particles must be involved (see Section 2.4). The gluino decays into a quark-squark pair which then undergoes a whole decay cascade, leading to SUSY decay topologies.

Supersymmetric Topologies

According to the decay cascades of SUSY particles, searches for supersymmetry involve a broad range of signatures with jets, leptons, photons, and missing transverse energy. A variety of analyses have been developed, based on certain signatures, which are not discussed here. However, general characteristics of SUSY signatures involve:

- High jet multiplicity, due to the long decay cascades involving multiple quarks. This also results in a large scalar sum of the transverse momentum of jets.

- Large missing transverse momentum, due to the undetectable LSP and possible neutrinos.
- The hadronic activity is in general more central, since the SUSY particles have a larger mass than SM particles and therefore, less energy is available for a boost of the sparticles.
- A natural possibility to produce leptons with the same charge.

The latter feature is the basis for the presented analysis of the SUSY same-sign dileptonic topology.

Same-Sign Di-leptonic Topology

A decay of a squark or a gluino can involve the production of a lepton with a certain charge (see Fig. 2.8). Since sparticles are produced in pairs, each decaying branch can lead to a lepton with the *same charge*. Hence, the possibility to naturally produce same-sign di-lepton pairs in SUSY exists [31–33]. Figure 2.9 illustrates such a decay.

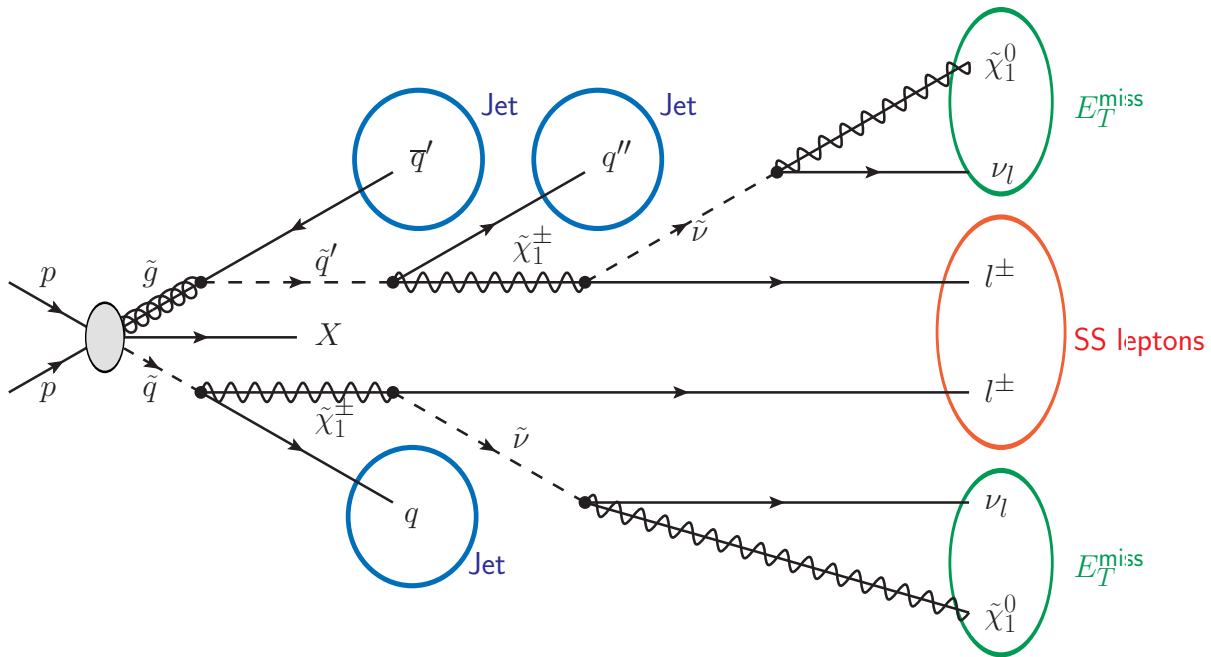


Figure 2.9: SUSY decay cascade leading to the signature of a same-sign di-lepton pair. The lepton pair is accompanied by hadronic activity and missing transverse energy.

In this thesis, muons and electrons are considered (not taus), leading to the three decay channels $\mu\mu$, $e\mu$, and ee . Hence, “leptons” always refers to muons and electrons, if not stated differently.

The signature also includes hadronic activity from released quarks and missing transverse energy from the LSPs. Note that there are only very rare SM process in which a same-sign di-lepton pair is produced. Therefore, this SUSY decay chain yields a very characteristic signature in a detector, which has a large potential to be discovered at a particle experiment.

3 Experimental Setup

In this Chapter the experimental setup is described. Since particle physics pushes technical requirements to their limits, large facilities and collaborations are needed in order to realize a particle physics experiment at the energy frontier. In the following, the Large Hadron Collider and its experiments are introduced. In particular, the Compact Muon Solenoid detector is described in more detail, since this thesis is based on this experiment.

3.1 The Large Hadron Collider

The Large Hadron Collider (LHC) [34–37] is a ring-shaped proton-proton accelerator with a circumference of 27 km. It is hosted in a tunnel of 3.8 m diameter, 45 to 170 m below ground at the CERN Laboratory (European Organization for Nuclear Research) in Geneva, Switzerland (see Fig. 3.1). The LHC is designed to accelerate two proton beams up to an energy of 7 TeV each, at an instantaneous luminosity of $\mathcal{L} = 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$.



Figure 3.1: Areal view of the LHC ring in the region of Geneva. The red line marks where the tunnel is hosted underground, small circles indicate where the experiments are located. From [38].

The luminosity can be related to the rate dN/dt of a process with a cross-section σ

$$\frac{dN}{dt} = \mathcal{L} \cdot \sigma \quad \Rightarrow \quad L := \int \mathcal{L} dt = \frac{N}{\sigma}, \quad (3.1)$$

where L is the integrated luminosity. $\mathcal{L}$ can also be written as

$$\mathcal{L} = \frac{\gamma f n_B N_{p1} N_{p2}}{4\pi \sigma_x^* \sigma_y^*} F, \quad (3.2)$$

where γ is the Lorentz factor, f is the revolution frequency, n_B are the numbers of bunches, N_{p1} and N_{p2} is the number of protons of the colliding bunches (see below), $\sigma_{x,y}^* = \sqrt{\epsilon_{x,y} \cdot \beta_{x,y}^*}$ is the beam cross-section at the interaction point in x - and y -direction, where $\epsilon_{x,y}$ and $\beta_{x,y}^*$ are the emittance and the betatron function, respectively, and F is the reduction factor due to the crossing angle. The betatron function characterizes the influence of external forces (originating from the focusing magnets) on the orbit of the accelerated particles.

Protons pass a chain of pre-accelerators before they are injected into the LHC:

- Linear accelerator LINAC 2 ($\rightarrow 50$ MeV),
- Proton Synchrotron Booster (PSB) ($\rightarrow 1.4$ GeV),
- Proton Synchrotron (PS) ($\rightarrow 26$ GeV),
- Super Proton Synchrotron (SPS) ($\rightarrow 450$ GeV).

In the LHC ring the two counter-rotating beams travel inside an ultrahigh vacuum of about 10^{-10} mbar. Different types of magnets are located intermittently along the beam pipe. Radio Frequency (RF) cavities provide the proton acceleration resulting in an increased energy of 0.5 MeV/turn. A large magnetic field is needed in order to keep the proton beams in their orbits. A total of 1,232 superconducting dipole magnets of about 15 m length are operating at currents of 11,850 A, resulting in a magnetic field up to 8.3 T in order to achieve sufficient bending power. Super-fluid Helium II at a temperature of 1.9 K is used for the cooling. For these magnets a special twin-bore technique is used, exploiting efficiently the limited space of the preset tunnel. They embed both beam pipes and combine the required magnetic field using the same magnetic yoke. A total of 392 main quadrupole magnets with a length varying between 5 and 7 m focus the beam, increasing the probability of an interaction when they collide.

The beams consist of “bunches” of protons. At nominal luminosity 2,808 bunches of about 10^{11} protons with a time spacing of 25 ns circulate in the LHC, leading to a bunch crossing frequency of 40 MHz. The bunch pattern depends on the injection of the pre-accelerators and contains various gaps, which are used for synchronization and acquiring of calibration data.

In a second operation mode the LHC accelerates heavy ions, reaching an energy of 2.76 TeV per nucleon. This mode is dedicated to produce a quark-gluon plasma, a high energetic state which resembles the condition of the universe shortly after the Big Bang, where gluons and quarks were not confined.

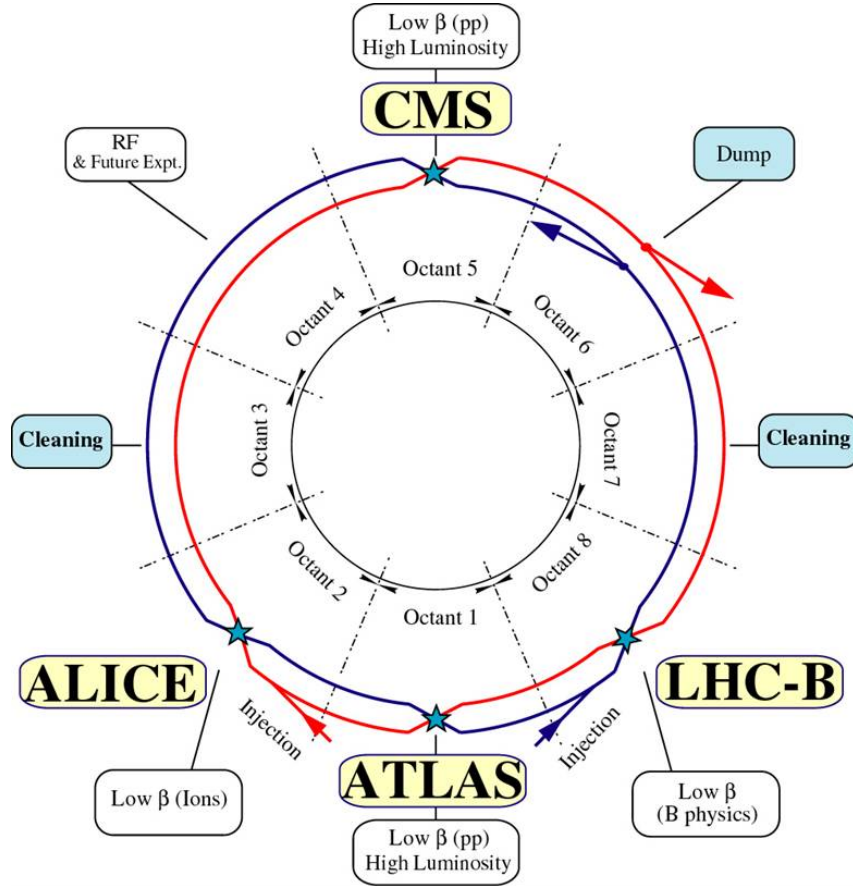


Figure 3.2: Schematic view of the LHC. The two proton beams are indicated via the red line and the blue line, respectively. The four experiments are labeled and the eight octants are marked. From [39].

The LHC consists of eight separate sectors (“octants”) which are cooled independently (see Fig. 3.2). In one particular octant the beams can be dumped while the cleaning is realized in two opposing octants.

At four different interaction points the beams are brought to collision where the experiments are built, situated in four underground caverns: CMS with TOTEM, ATLAS with LHCf, ALICE, and LHCb.

ATLAS – A Toroidal LHC ApparatuS

ATLAS [40–42] is a general-purpose detector. With its 44 m length and 25 m in diameter it is the experiment with the largest extension in space at the LHC. It is complementary to the CMS experiment. One main difference is its magnet system — it consists of one huge torodial magnet and one relatively small solenoid. The hadronic liquid argon calorimeter has a better energy resolution than the hadronic calorimeter of CMS, while the momentum resolution of the tracking detectors is worse.

ALICE – A Large Ion Collider Experiment

ALICE [43, 44] is a special-purpose experiment, built for studying quark-gluon plasma which is produced in heavy-ion collisions. It is optimized to resolve high track multiplicities. A specialized time projection chamber is used to measure many tracks in one event. However, it is not designed for the high event rate achieved in proton collisions.

LHCb – Large Hadron Collider beauty experiment

The focus of LHCb [45, 46] is the measurement of the properties of b -hadrons. Studies to determine the parameters of Charge Parity (CP) violation and rare charm and beauty meson decays are performed. In contrast to the above experiments, the LHCb detector is an asymmetric one-arm spectrometer of 20 m length. Since at the LHC charm and beauty mesons are predominantly produced in the forward direction, LHCb covers only a small angle close to the beam line.

LHCf – Large Hadron Collider forward experiment

LHCf [47, 48] is a special-purpose experiment in which hadron interactions at high energies in the very forward region are measured. These processes resemble particle showers of the earth’s atmosphere and help to interpret and calibrate large-scale cosmic-ray experiments. It consists of two calorimeters installed 140 m downstream on both sides of the ATLAS interaction point.

TOTEM – TOTAal cross-section and Elastic scattering Measurement experiment

The TOTEM [49, 50] experiment is affiliated to CMS. Its purpose is to measure the effective total proton-proton cross-section with an absolute error of 1 mb, using a luminosity independent method. Furthermore, elastic scattering and hard diffractive dissociation processes will be studied. “Roman Pots” [51] are used to measure particles very close to the LHC beam line.

3.2 The Compact Muon Solenoid

The Compact Muon Solenoid (CMS) [52–55] detector is one of the two general-purpose experiments at the LHC. Its objective is to investigate proton-proton collisions at the TeV energy scale. The design of the detector differs considerably from that of the second general-purpose experiment, ATLAS, which allows the results to be complementary. Three distinct features of the CMS detector are already covered in its name:

- It is relatively compact (especially compared to the ATLAS detector).
- An emphasis has been set to the reconstruction of muons, in particular with high transverse momenta.
- The large superconducting solenoid magnet encloses a large part of the detector, including the calorimeters, providing a magnetic field of up to 3.8 T.

The CMS detector is 21.6 m long, has a diameter of 14.6 m, and weights about 12,500 tons. Figure 3.3 shows an overview of the apparatus and most of its components. The

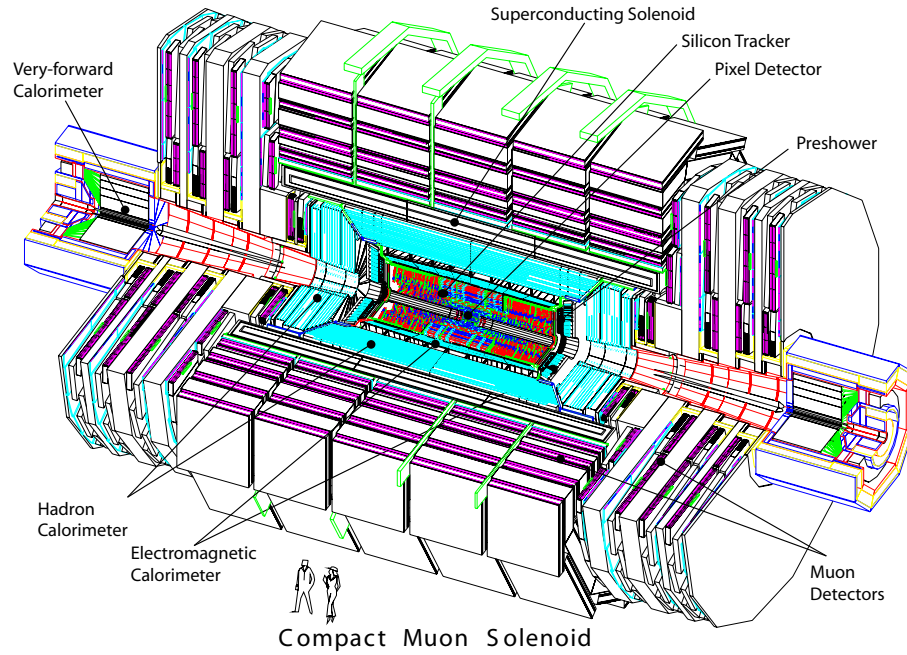


Figure 3.3: Schematic view of the CMS detector. Single components are labeled. To illustrate the size of the detector two persons are sketched. From [54].

CMS detector has the typical onion-like structure of a hermetic particle detector – different components for different purposes surround the interaction point in a series of cylindrical layers. The main principle of all the detector components is to measure the charge released by the interaction of particles from the collision with the detector material. The whole detector can be separated into five distinct segments, which allows the inner parts of the detector to be accessed during technical stops, e.g. for upgrades.

The origin of the CMS coordinate system is the nominal interaction point in the middle of the detector. The x -axis points horizontally towards the center of the LHC ring, the y -axis upwards, and the z -axis in the direction of the beam line with the positive direction counter-clockwise the LHC ring. The azimuth angle ϕ is defined as the angle in the x - y -plane starting from the x -axis, while θ is the angle to the z -axis. The transverse distance to the z -axis is denoted with r . Other helpful quantities are the rapidity¹ y

$$y := \frac{1}{2} \ln \left(\frac{E + p_z}{E - p_z} \right), \quad (3.3)$$

where E and p_z are the energy and the momentum component of a particle in z -direction, the pseudo-rapidity η , defined as

$$\eta := -\ln \left[\tan \left(\frac{\theta}{2} \right) \right], \quad (3.4)$$

and the distance ΔR given by

$$\Delta R := \sqrt{\Delta\phi^2 + \Delta\eta^2}. \quad (3.5)$$

¹Rapidity differences are invariant under Lorentz transformations.

The CMS detector comprises two additional projects: CASTOR and ZDC.

CASTOR – Centauro And Strange Object Research

CASTOR [56, 57] is a very forward Cherenkov sampling calorimeter for the CMS experiment. It is located behind the HF and the T2, a tracking station of the TOTEM experiment, covering a range of $-6.6 < \eta < -5.2$. Its purpose is to search for exotic objects with unusual longitudinal shower profile properties. It measures low- x parton dynamics, minimum bias event structure, diffraction, cosmic ray related physics in low-luminosity proton-proton, and heavy ion collisions.

ZDC – Zero Degree Calorimeters

Two Zero Degree Calorimeters (ZDCs) [58] are designed to complement the CMS very forward region for pseudo-rapidities of $|\eta| > 8.3$. Each of them comprises of two parts: an electromagnetic and a hadronic sampling calorimeter using tungsten and quartz fibers. It provides heavy ion and proton-proton diffractive studies. Measurements of real-time luminosity, beam tuning and accelerator monitoring are performed at the ZDCs.

3.2.1 Silicon Tracking System

The innermost part of the CMS detector is the Silicon Tracking System (STS) [59]. It fulfills several requirements:

- Fine granularity, since it is the part closest to the beam axis and about $\mathcal{O}(1,000)$ particles will occur in one bunch crossing.
- Fast readout, since the bunch crossing occurs with a frequency of 40 MHz.
- Radiation hardness, since a considerable amount of radiation is released in the collisions, especially close to the beam axis.
- Effective cooling, down to -10°C to reduce radiation damages and noise.

The STS takes all these requirements into account and is a good compromise between granularity, space, and material budget. It consists of two sub-detectors: The Pixel Detector (PD) and the Silicon Strip Detector (SSD). Figure 3.4 shows the layout of the STS. Its overall size is 5.5 m in length and at a radius of 4 cm to 1.1 m from the beam axis, covering a pseudo-rapidity range of $|\eta| < 2.5$, whereby tracks with $|\eta| > 2.4$ only cross few layers. The momentum resolution of the STS reaches about 1 % for tracks with a momentum of 100 GeV [60, 61]. In order to exploit the resolution of the single modules of the STS, the knowledge of their geometrical position (“alignment”) is needed. A large effort is done to perform a precise alignment [62].

Pixel Detector

The PD consists of three cylindric layers in the barrel region (covering $|\eta| < 2.2$) and two endcap disks (covering $2.2 < |\eta| < 2.5$) at each side of the detector, perpendicular to the beam axis. The barrel layers are installed at radii of 4.3 cm, 7.2 cm, and 10 cm from the

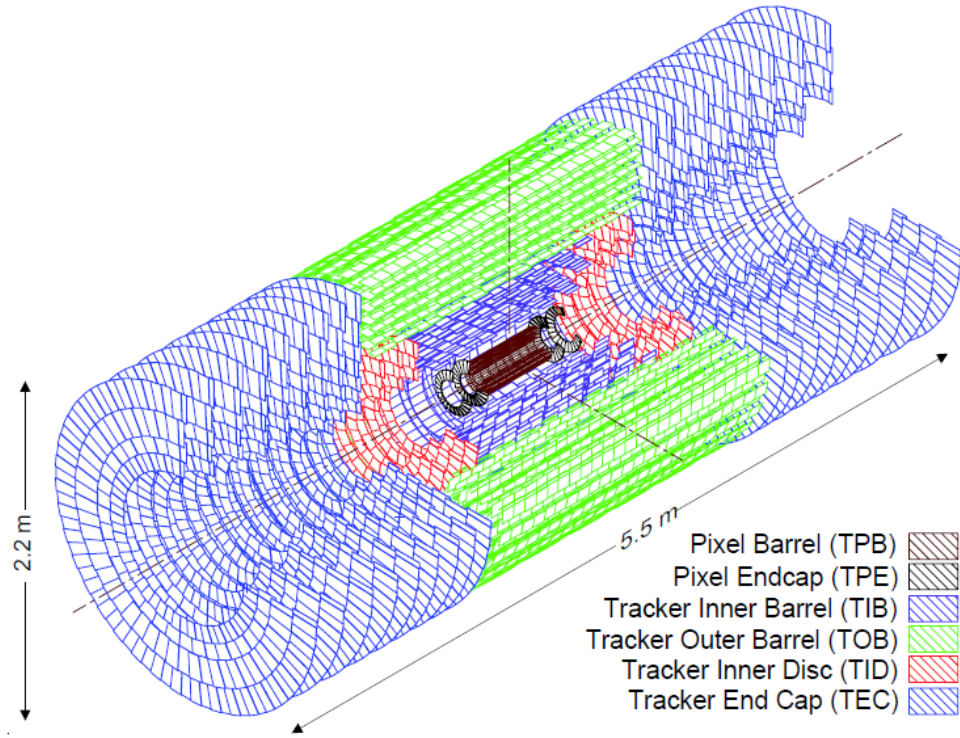


Figure 3.4: Layout of the STS of the CMS detector. Different components are illustrated in different colors. From [54].

beam axis, and the endcaps at a z -position of 32.5 cm and 46.5 cm. The PD is composed of 1,440 pixel modules having 66 million readout channels in total.

The spacial resolution of a PD module achieves $15\text{--}20\,\mu\text{m}$. Its total area corresponds to 1 m^2 . One main task of this component is to resolve primary and secondary vertices, which is an important contribution to e.g. the identification of b -hadrons.

Silicon Strip Detector

The SSD envelops the PD. It consists of multiple layers in the central barrel region and of endcap discs in the r - ϕ -plane. It is divided into several components:

- Tracker Inner Barrel (TIB), consisting of four layers with the two inner ones being double-sided.
- Tracker Inner Detector (TID), consisting of three disks on each side.
- Tracker Outer Barrel (TOB), consisting of six layers with the two inner ones being double-sided.
- Tracker End Cap (TEC), consisting of nine disks on each side.

The resolution of the SSD is of the same order of magnitude as the resolution of the PD.

3.2.2 Electromagnetic Calorimeter

A focus of the CMS detector is set on the high quality Electromagnetic Calorimeter (ECAL) [63], in order to be sensitive to signatures like $H \rightarrow \gamma\gamma$. Hence, it consists entirely of lead tungstate crystals (PbWO_4) which provide an excellent energy resolution. Lead tungstate crystals (see Fig. 3.5) have the advantage to be at the same time absorber material, stopping electromagnetic interacting particles, and active material which can be read out directly (such that no sampling fluctuations occur due to passive material, which would decrease the energy resolution). Furthermore, 80% of the emitted light can be collected within 25 ns, which significantly reduces the interplay of multiple bunch crossings. In addition, this material is very radiation hard and has a very short radiation length of $\chi_0 = 0.89\text{ cm}$ and a Molière radius of only 2.2 cm. A simulation of an electron shower in lead tungstate is shown in Fig. 3.6. A benefit of the small radiation length is the moderate size of the crystals with $25.8\chi_0$.

The ECAL envelops the STS, covering the range of $|\eta| < 1.479$ in the barrel (EB) region and $1.479 < |\eta| < 3.0$ in the endcap (EE) region (see Fig. 3.7). In total 75,848 lead tungstate crystals are installed, each covering a solid angle of $\Delta\eta \times \Delta\phi = 0.0174 \times 0.0174$ to ensure a fine granularity. The emitted light is detected by avalanche photodiodes (in EB) and vacuum phototriodes (in EE). Preshower detectors (ES) are mounted in front of EE, extending over a range of $1.653 < |\eta| < 2.6$. These are sampling calorimeters equipped with silicon strip sensors. Their main purpose is the identification of neutral pions and the improvement of the position measurement of electromagnetic showers from electrons and photons.

The energy resolution of the ECAL [54] can be parameterized by

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{2.8\% \sqrt{\text{GeV}}}{\sqrt{E}}\right)^2 + \left(\frac{0.12\text{ GeV}}{E}\right)^2 + (0.26\%)^2, \quad (3.6)$$

where the energy is given in GeV. The first term represents the stochastic term, the second one arises from electronic noise, digitalization, and pile-up, and the constant term refers to non-uniformity of the ECAL and mis-calibration.

3.2.3 Hadronic Calorimeter

The purpose of the Hadronic Calorimeter (HCAL) [66] is the energy measurement of hadronic interacting particles. Its design is strongly constrained by the choice of magnet parameters, since most of the CMS calorimetry is located inside the volume of the solenoid magnet. Hence, it is a challenge to design a calorimeter small enough to fit into the solenoid magnet and at the same time to ensure enough stopping power to obtain a sufficient shower containment. The HCAL consists of several components:

- HCAL Barrel (HB) in a region of $|\eta| < 1.4$ at a radius range of 1.8-2.9 m.
- HCAL Endcap (HE), extending the range to $1.3 < |\eta| < 3.0$, having a small overlap with EB to ensure hermeticity.

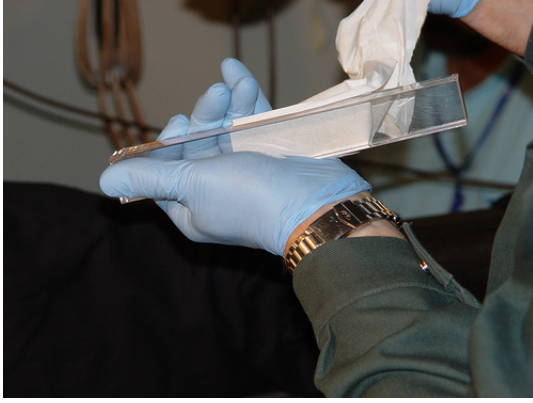


Figure 3.5: Lead tungstate crystal of the ECAL barrel. It has a size of $2.2 \times 2.2 \times 23 \text{ cm}^3$. From [64].

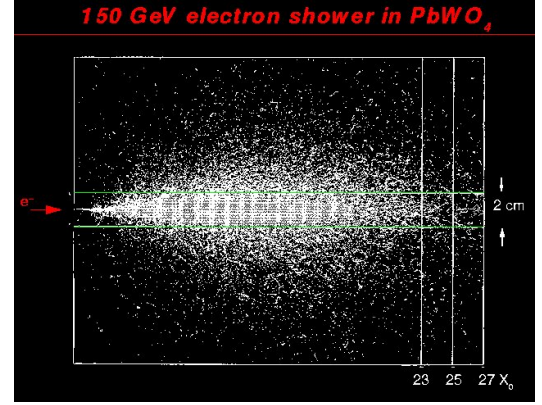


Figure 3.6: Simulation of an electromagnetic shower in lead tungstate, caused by an electron with an energy of 150 GeV. Green lines illustrate the size of a crystal in the ECAL of CMS. From [65].

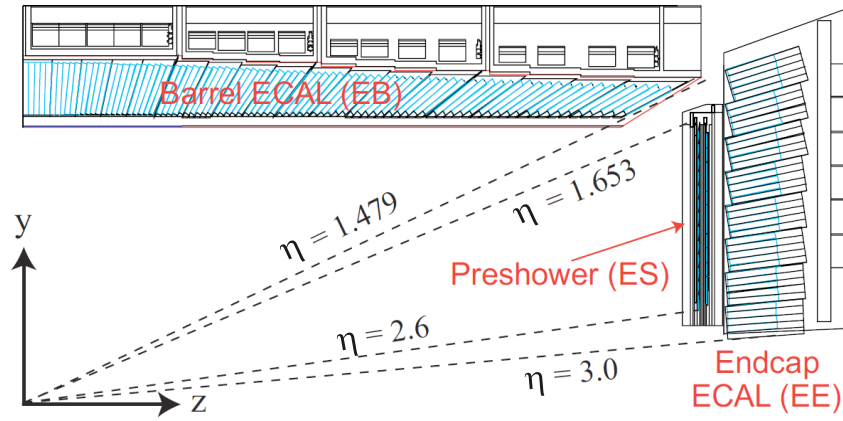


Figure 3.7: y - z -view of a quadrant of the CMS ECAL. EB, EE, and ES are labeled, the η -region is indicated. From [54].

- HCAL Outer (HO), a “tail-catcher”, installed outside the solenoid to improve the shower containment in the region of $|\eta| < 1.26$ and to measure late showering particles which even punch through the magnet.
- HCAL Forward (HF) for measuring particles in a region of $3.0 < |\eta| < 5.0$ and increasing the hermeticity of the missing transverse energy measurement, being installed 11.2 m from the interaction point.

HB and HE surround the ECAL completely and are enclosed by the solenoid magnet. It is a non-compensating sampling calorimeter. Consequently, the response for electromagnetic and hadronic showers differ. This is quantified by the e/π -ratio which is about

1.2 at 50 GeV [67]. A software based compensation method to improve the energy measurement after the HCAL upgrade is subject of the technical work of this thesis and will be discussed in detail later in Chapter 4.

The CMS sampling calorimeter consists of about 5 cm thick passive layers made of copper, interlayed with plastic scintillator tiles, which are connected to wavelength shifting fibers. One segment (“tower”) covers a solid angle of $\Delta\eta \times \Delta\phi = 0.087 \times 0.087$, corresponding to a 5×5 block of crystals in the ECAL. They are read out by hybrid photodiodes (HPD). In the central region, the thickness of HB is only $5.82 \lambda_I$ in interaction lengths, increasing with $1/\sin\theta$. The inclusion of HO extends the total depth of the HCAL to $11 \lambda_I$ for $|\eta| < 1.26$.

HB is divided into 32 segments in η -direction and 72 segments in ϕ -direction, respectively (see Fig. 3.8). In the longitudinal direction it consists of 17 layers which are – so far – read out as a whole. A segmented readout for HB is under consideration for the planned detector upgrade [68], improving its longitudinal granularity up to a factor of four and replacing the HPDs with Silicon Photo Multipliers (SiPM), which have much better properties. A possible readout scheme is shown in Fig. 3.9 for illustration.

An estimation of the HCAL energy resolution has been performed [54] using pions of 30-300 GeV and yields

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{120\% \sqrt{\text{GeV}}}{\sqrt{E}}\right)^2 + (6.9\%)^2, \quad (3.7)$$

where the first term refers to the stochastic term and the constant term accounts for the non-uniformity of the HCAL and mis-calibration.

3.2.4 Superconducting Magnet

The central piece of CMS is the superconducting solenoid magnet [69] (see Fig. 3.10). It provides a magnetic field of $B = 3.8 \text{ T}$, which bends the trajectories of charged particles perpendicular to the beam line due to the Lorentz force. This opens the possibility to apply the fundamental principle to determine the particles’ momenta p by exploiting the relation

$$R = \frac{p}{eB}, \quad (3.8)$$

where R is the track radius and e the electric charge.

The solenoid magnet consists of a cylinder of roughly 13 m length and a diameter of more than 6 m and weights about 220 tons. Nonetheless, the coil is thin in the sense that the ratio $\Delta R/R$ is only about 0.1. A total of 19,500 A electrical current flows through superconducting cables, cooled by liquid Helium down to 4.6 K. The iron return yoke is provided by five barrel rings around the coil and three endcap discs on each side. With 10,000 tons its mass has by far the largest contribution to CMS’s total mass. The stored energy of the magnetic field is about 2.7 GJ.

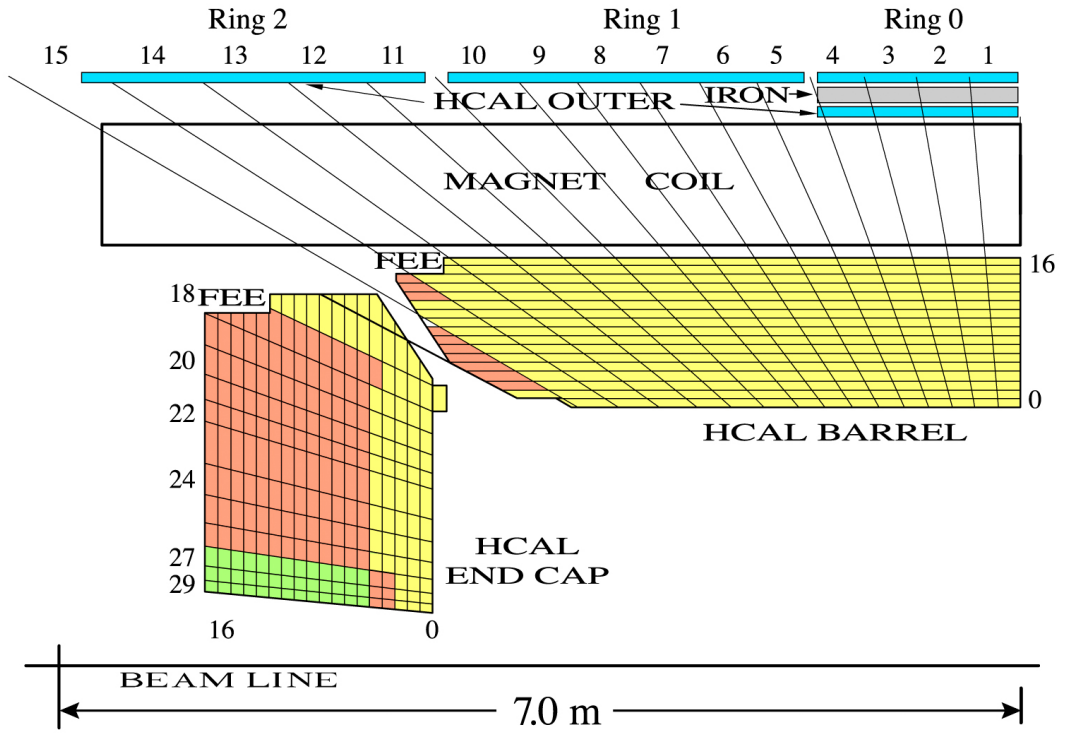


Figure 3.8: y - z -view of a quadrant of the CMS HCAL. HB, HE, and HO are labeled. Each color of a tower represents a readout channel (“depth”). From [53].

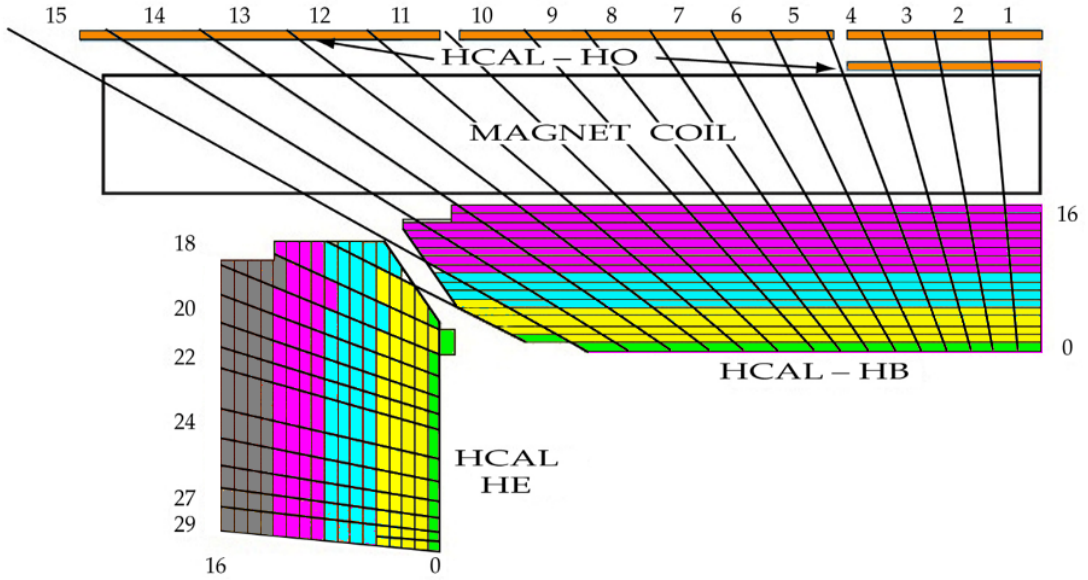


Figure 3.9: y - z -view of a quadrant of the CMS HCAL after a possible upgrade. In the barrel region, the “1-4-4-8” readout scheme is applied. Each color of a tower represents a readout channel (“depth”). From [68].

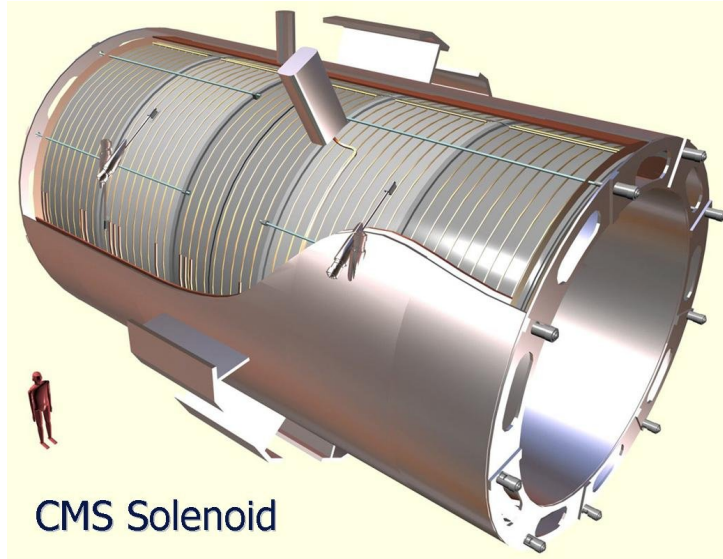


Figure 3.10: Drawing of the CMS solenoid coil. A man is sketched to illustrate its dimensions. From [70].

3.2.5 Muon System

The Muon System (MS) is installed outside the magnet coil and inside the iron return yoke. Its performance requirements are in particular driven by the $H \rightarrow 4l$ channel for which a good resolution for muons with a high transverse momentum is needed, as well as a highly reliable muon identification. The MS consists of a barrel detector (MB) in the region of $|\eta| < 1.2$ and an endcap detector (ME) covering $1.2 < |\eta| < 2.4$.

Figure 3.11 shows the cross-section of a quarter of the MS. It consists of three different types of gaseous detectors:

- Drift Tubes (DTs),
- Cathode Strip Chambers (CSCs),
- Resistive Plate Chambers (RPCs).

The choice of the detector technologies has been driven by the large surface to be covered and the different radiation environments.

Drift Tubes

DTs are used in the central barrel region, since there the neutron induced background is small and both, the muon rate and the residual magnetic field in the chambers are low. A total of 250 chambers of DTs are organized in the gaps of the solenoid return yoke in four layers at radii of about 4.0, 4.9, 5.9 and 7.0 m from the beam axis. Chambers in different layers are staggered such that at least three out of four chambers must be crossed by traversing muons. One chamber consists of three superlayers each containing four layers of DTs, where the outer superlayers are aligned along the beam line and the middle one is rotated by 90 degrees in order to measure the muons' z -coordinates.

Cathode Strip Chambers

CSCs are deployed in the endcaps, since the muon rate, the neutron induced background rate, and the magnetic field are high in this region. The ME system contains 468 CSCs in the two endcaps. A CSC comprises six gas gaps, each having a plane of radial cathode strips and a plane of anode wires, being almost perpendicularly to the strips. Most of the CSCs have an overlap in ϕ to avoid gaps in the muon acceptance. A charged particle traversing a CSC initiates an ionization of the gas, causing an electron avalanche. The charge of the released electrons is collected at the anode wire and an image charge on a group of cathode strips is measured. Determining the center-of-gravity of the charge distribution induced on the cathode strips allows to perform a precise position measurement. The spatial resolution of a single CSC is typically about $200\ \mu\text{m}$ and the angular resolution in ϕ is $\mathcal{O}(10)$ mrad.

Resistive Plate Chambers

RPCs are used in both the barrel region and the endcap region. The main feature of RPCs is their fast response with a good time resolution. Therefore, they can identify unambiguously the correct bunch crossing for a measured muon. However, they have a coarser position resolution than the DTs or CSCs. They are installed in the pseudo-rapidity region up to $|\eta| < 2.1$ and consist of double-gap chambers with common readout strips. The width of the gaps is only 2 mm leading to the fast timing performance. RPCs play an important role for the Trigger System.

The combination of the silicon tracking system with the MS improves the precision of the momentum measurement compared to the tracker-only reconstruction, in particular for muons with a high transverse momentum (see Fig. 3.12).

3.2.6 Trigger System

The purpose of the Trigger System (TS) [71, 72] is to decide during data taking (“online”), which events to record. This is necessary because of the high bunch crossing rate of 40 MHz — it is not possible to store every event since otherwise the needed disk space would be by far too large.²

The TS reduces the event rate from initially 40 MHz down to $\mathcal{O}(100)$ Hz. It consists of two stages:

- Level-1 (L1) trigger, using custom-designed flexible hardware. Reduction: from 40 MHz to 50 kHz.
- High Level Trigger (HLT), fully software based. Reduction: from 50 kHz to $\mathcal{O}(100)$ Hz.

L1 Trigger

The hardware-based L1 trigger consists of Field Programmable Gate Arrays (FPGAs) which ensure a high flexibility of the configurable hardware. The decisions of the L1 trigger must be made very fast. Hence, it is implemented close to the detector and uses only coarsely reconstructed information from the calorimeters and the MS, and not from

²One event (of pp interactions) requires about 1.5 MByte of disk space in CMS.

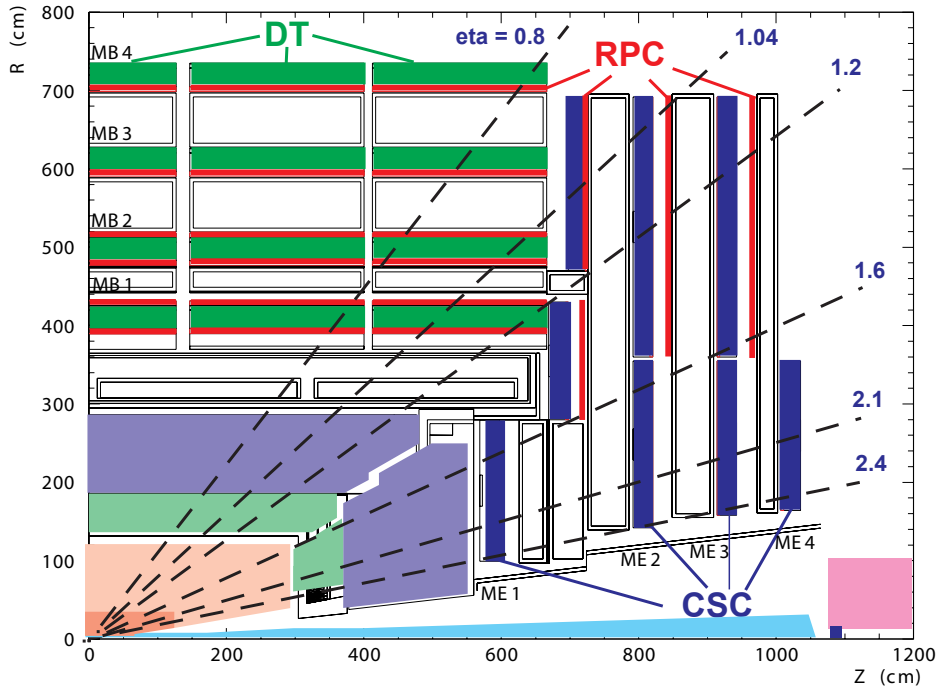


Figure 3.11: y - z -view of a quadrant of the CMS MS. The different muon detectors are labeled and some η -regions are marked. From [54].

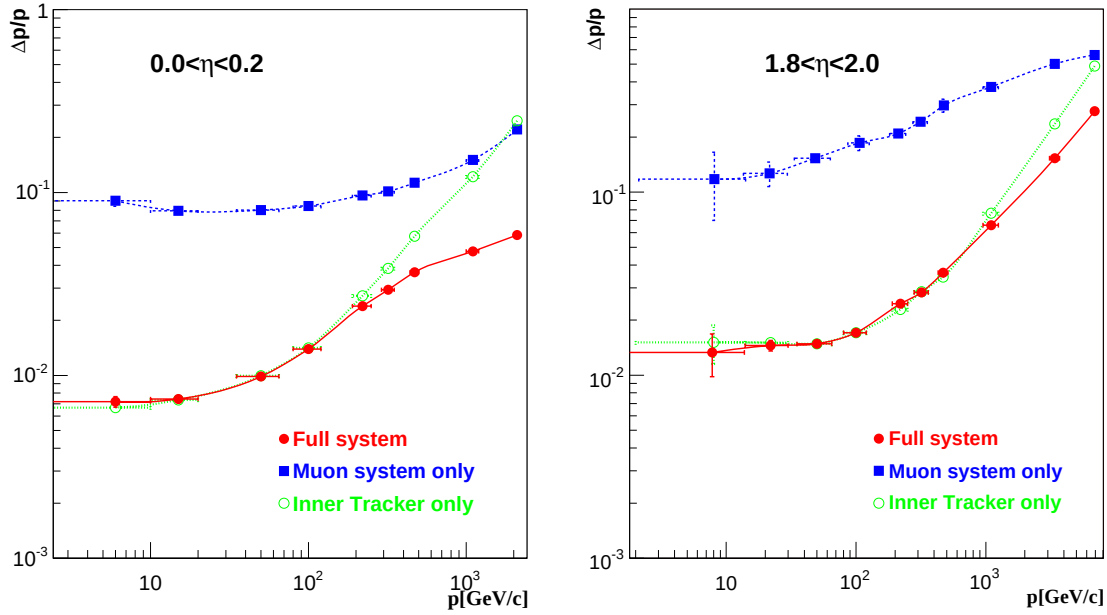


Figure 3.12: Muon momentum resolution as a function of p_T in two different η -regions: *Left:* $|\eta| < 0.2$, *right:* $1.8 < |\eta| < 2.0$. *Blue squares with dashed line:* muon system only, *green circles with dotted line:* inner tracker only, *red full circles with solid line:* combination of muon system and tracker. From [54]

the tracker since the track reconstruction takes too long. For a latency of only $3.2\mu\text{m}$ the detector output of the front-end electronics is stored in pipelined memories. Within this time-window the decision of the L1 trigger must be made. This decision is based on the presence of Trigger Primitive Objects (TPOs) like photons, electrons, muons, and jets above certain thresholds of their transverse momentum, as well as on global sums of the transverse and missing transverse energy. TPOs are collected for the global L1 which is located about 90 m from the experiment. There the decision is made if the event is further analyzed by the HLT.

HLT Trigger

The HLT is entirely software-based using the same algorithms as for the final reconstruction. Data from the pipelines of the L1 trigger is transferred to front-end readout buffers, to be investigated by the HLT. An event building switch distributes data from a given event to a processor of a computing farm with about 1,000 cores. The main strategy is to discard events as soon as possible. First of all, only calorimeter and muon chamber information is considered. Subsequently, tracker information is added and eventually, the full detector information is used.

This whole procedure has maximum flexibility since it has complete freedom in the selection of the data to access, as well as in the realization of the algorithms. If the event is accepted by the HLT it is passed to the global Data Acquisition (DAQ) and written to disc.

4 Energy Weighting for the Upgrade of the Hadronic Calorimeter

Detector upgrades must be planned well in advance due to the complexity of today's particle physics experiments. Two mayor upgrades [68] are planned which will take place during long shutdowns of the LHC:

- *Phase I*: 2013/2014: Technical stop after data taking until end of 2012. The LHC will be shut down for 1.5 to 2 years to make revisions, necessary to run at higher energies.
- *Phase II*: 2017/2018: Preparation of the LHC to operate at and possibly above the design luminosity of $L = 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$.

The detector work for the upgrades involves the hadron calorimeters, the muon detectors, the pixel detector, the beam radiation monitoring and luminosity measurement system, the trigger, the data acquisition system, and the CMS infrastructure and facilities. Such upgrades must be prepared carefully in advance and are motivated by physics.

The HCAL upgrade, relevant for the presented studies, will take place in Phase II. The Hybrid Photo Diodes (HPDs) of the HCAL readout electronics will be replaced by Silicon Photo Multipliers (SiPM), where a segmented readout is under consideration, improving its longitudinal granularity up to a factor of four. A possible readout scheme is shown in Fig. 4.1 for illustration (see also Fig. 3.9). The CMS HCAL is a non-compensating sampling calorimeter with an e/π -ratio of about 1.2 at 50 GeV [67] (see Section 3.2.3). Consequently, the response for electromagnetic energy depositions is larger than for hadronic ones, which influences the energy measurement. A software based compensation of the e/π -ratio is possible if one can identify the electromagnetic or hadronic origin of the energy deposition within a hadronic shower. Therefore, parts of a particle shower need to be resolved. This will be possible after the detector upgrade, providing an additional segmentation in longitudinal direction.

The motivation for these studies (published in [73, 74]) is to give a *proof of principle* that a tabulated weighting method can successfully be applied to the signals of the CMS HCAL after the planned upgrade. The corrected energy will lead to an improved energy measurement which is from importance for most physics analyses at CMS. Furthermore, various readout designs are investigated in order to find the one that yields the best results with respect to energy resolution. The impact of a defect of one readout channel is tested and documented for two different readout schemes in the Appendix A.2. Note that here it is not intended to present a closed study with full error propagation and discussion of systematic errors. The proof of principle is sufficient to motivate a certain readout scheme for the Phase II upgrade.

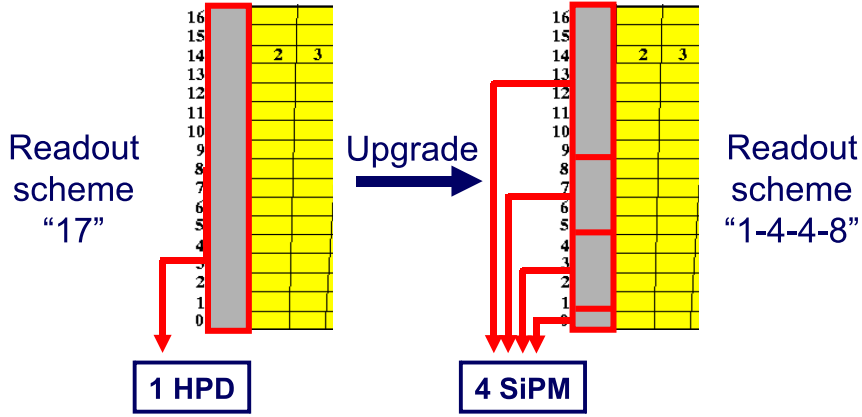


Figure 4.1: Schematic illustration of the HCAL readout before and after the Phase II upgrade. Here, the possible readout scheme “1-4-4-8” is shown. The currently used hybrid photodiodes (HPDs) will be replaced by silicon photo multipliers (SiPM).

In the following, the methodology of the tabulated weighting factors is introduced. A detailed description of the implementation of the method is given afterwards. Subsequently, results of the applied weighting method are presented. Finally, next steps for future investigations are discussed. The Appendix A provides some additional detailed investigations.

4.1 Methodology

The method of the tabulated Weighting Factors (WFs) [75, 76] is a software based method, aiming to compensate for the e/π -ratio of a calorimeter. The basic idea is to distinguish between electromagnetic and hadronic energy depositions and to find appropriate WF for the compensation. The discrimination criterion is the measured *energy density*

$$\rho_{\text{meas}}^i = E_{\text{meas}}^i / V^i, \quad (4.1)$$

where E_{meas}^i is the measured energy and V^i a measure for the volume in arbitrary units, both for a readout channel i .

The weighting method is based on the fact that the average energy density of electromagnetic depositions is larger than for hadronic ones. In a Monte Carlo simulation it is possible to obtain WFs w^i as a function of the energy density ρ_{meas}^i (see Fig. 4.2) and the fixed energy of simulated incident pions E_π

$$w^i(\rho_{\text{meas}}^i, E_\pi) = \left\langle \frac{E_{\text{dep}}^i}{E_{\text{meas}}^i} \right\rangle \quad (\text{obtained WFs}), \quad (4.2)$$

where E_{dep}^i is the deposited energy given by

$$E_{\text{dep}}^i = E_{\text{abs}}^i + E_{\text{sci}}^i + E_{\text{inv}}^i, \quad (4.3)$$

where E_{abs}^i is the energy deposited in the absorber, E_{sci}^i the energy deposited in the scintillator, and E_{inv}^i the invisible energy (originating from neutrinos, nuclear excitation, etc.).

Once the WFs are obtained from a Monte Carlo simulation, they can be applied to data (here: simulated data) to yield a weighted energy

$$E_{\text{weight}}^i = E_{\text{meas}}^i \cdot w^i(\rho_{\text{meas}}^i, E_{\text{meas}}). \quad (4.4)$$

Note that by *applying* the WFs, E_π is replaced by the total measured shower energy $E_{\text{meas}} = \sum_{\text{cluster}} (\sum_i E_{\text{meas}}^i)$, received from a 3×3 -cluster around the tower of the maximal energy deposition.

$$w^i(\rho_{\text{meas}}^i, E_{\text{meas}}) = \left\langle \frac{E_{\text{dep}}^i}{E_{\text{meas}}^i} \right\rangle \quad (\text{applied WFs}). \quad (4.5)$$

The WFs are illustrated in Fig. 4.2 and discussed in more detail in Section 4.2.4.

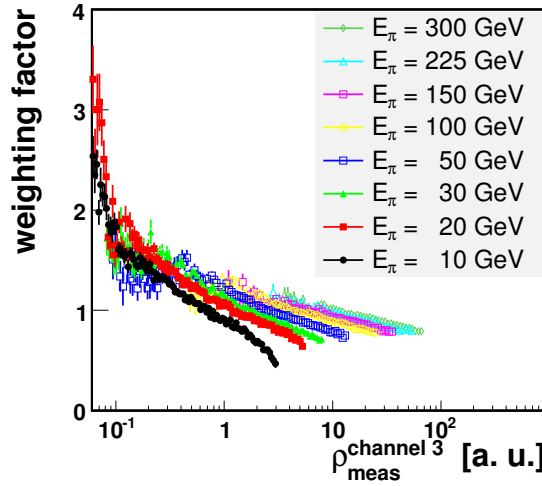


Figure 4.2: WFs for the third channel of the readout scheme “1-4-4-8” as a function of the energy density ρ_{meas}^i . The color code refers to the energy E_π of the simulated incident pions.

4.2 Implementation

In order to implement the tabulated weighting method, a simulation of the CMS HCAL is necessary, because the deposited energy E_{dep}^i can only be extracted from simulation (see equation (4.2)). Since the energy deposition in the absorber E_{abs}^i , being essential for the calculation of E_{dep}^i (see equation (4.3)), is presently not available in the CMS software [54], it was necessary to develop a standalone simulation of the CMS calorimeter.

In the next Section the standalone simulation will be described in more detail, followed by a Section in which the calibration of the simulated HCAL is illustrated. The determination of the WFs can be found in Section 4.2.4. Finally, the investigated readout schemes are shortly introduced.

4.2.1 The Geant3 Standalone Simulation

The standalone simulation of the CMS calorimeter was provided in GEANT3 [77]. Four different hadronic shower models are available and were used: GCALE, GHEISHA, FLUKA and MICAP. However, GCALE is chosen as shower model for all studies presented here, as it has been shown to be the most realistic one [78].

The simulation has been implemented on the basis of the CMS detector design described in [53]. As this study concentrates on single pion showers, it is not necessary to simulate the whole HCAL barrel region of the CMS detector. Thus, only a piece of the CMS HCAL in the central region is implemented — ten towers in η -direction and four towers in ϕ -direction.

For a more realistic simulation, the CMS ECAL has been simulated, too. However, it is very roughly represented by a simple block of matter made of a 23 cm thick lead tungstate (PbWO_4) crystal without any readout.¹ This corresponds to about 1.1λ interaction lengths of material (see [53]).

Between the ECAL and HCAL a piece of 10 cm copper is included in order to take the passive material into account, corresponding to 0.65λ interaction lengths for $\eta = 0$ [79]. Finally, the HCAL follows containing 17 layers of scintillators and 16 layers of absorber. Figure 4.3 shows a schematic section view of the simulated geometry in the η - ϕ -plane. Details can be extracted from Table 4.1.

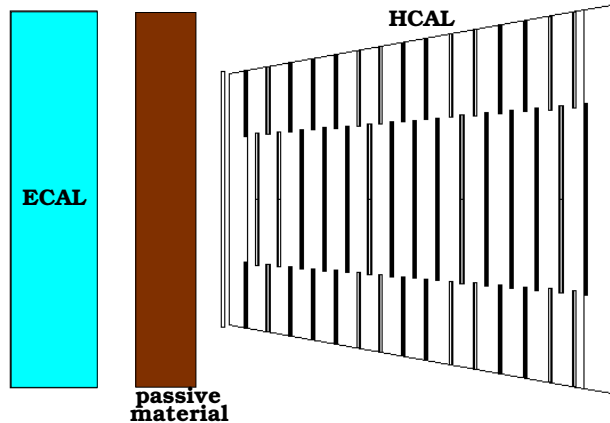


Figure 4.3: Schematic section view of the geometry of the GEANT3 standalone simulation in the η - ϕ -plane.

Note that the specifications of the simulation differ in two points from the detector design described in [53], where the first scintillator of the CMS HCAL is specified to be made of Bicron BC408 (which yields 1.4 times more light than Kuraray SCSN81 for the same thickness) and the thickness of the first stainless steel absorber is 40 mm (compare with Table 4.1).

Birk's law [80] is also implemented in the simulation, which is an important feature.

¹Consequently, only the information of the true deposited energy exists for the ECAL and no simulated measured energy.

Table 4.1: Single layers of the GEANT3 standalone simulation of the HCAL. The brass of the HCAL absorber is a chemical composition of 70 % Cu and 30 % Zn with a density of 8.53 g/cm, a radiation length of 1.49 cm, and an interaction length of 16.42 cm.

layer	material	thickness [mm]
ECAL	lead tungstine (PbWO_4)	230
passive material	copper	100
1st scintillator	Kuraray SCSN81	9
scintillator 1-16	Kuraray SCSN81	3.7
scintillator 17	Kuraray SCSN81	9
1st absorber	stainless steel	49.5
absorber 1-8	brass	50.5
absorber 9-15	brass	56.5
absorber 16	stainless steel	75

Since the HCAL standalone simulation is not as detailed as the CMS software simulation, no electronic noise is included and hence no threshold for noise reduction is applied. Furthermore, the magnetic field is not simulated.

4.2.2 Investigated Readout Schemes

An important goal of this study is to find the optimal readout scheme for the realization of the method of tabulated WFs, being constrained to a certain number of readout channels. On the one hand, a high granularity and therefore a large number of readout channels is desired. On the other hand, it is not possible to read out each of the 17 layers separately, because the space for the readout electronics is limited, as well as its power supply, the cooling, and the material budget. Figure 4.4 shows a schematic view of the different readout schemes of the CMS HCAL which are investigated here.

Currently, a segmentation with three or four readout channels is under discussion. Various readout schemes with four channels are investigated here. The design with 17 readout channels represents the design with the highest possible segmentation and is investigated for comparative reasons. The design with six readout channels is a compromise between both, high granularity and moderate number of readout channels.

4.2.3 Calibration

The calibration of the simulated CMS HCAL is performed with 50 GeV pions, entering into the middle of one tower. Figure 4.5 shows the simulated scintillator signal E_{sci} (left) and illustrates the simulated measured energy E_{meas} after the calibration (right).

The simulated measured energy E_{meas} consists of the scintillator signal times a calibration factor “C” plus a small contribution of energy depositions from the ECAL (for details see Section 4.2.4)

$$E_{\text{meas}} = E_{\text{sci}} \cdot C + E_{\text{ECAL}}. \quad (4.6)$$

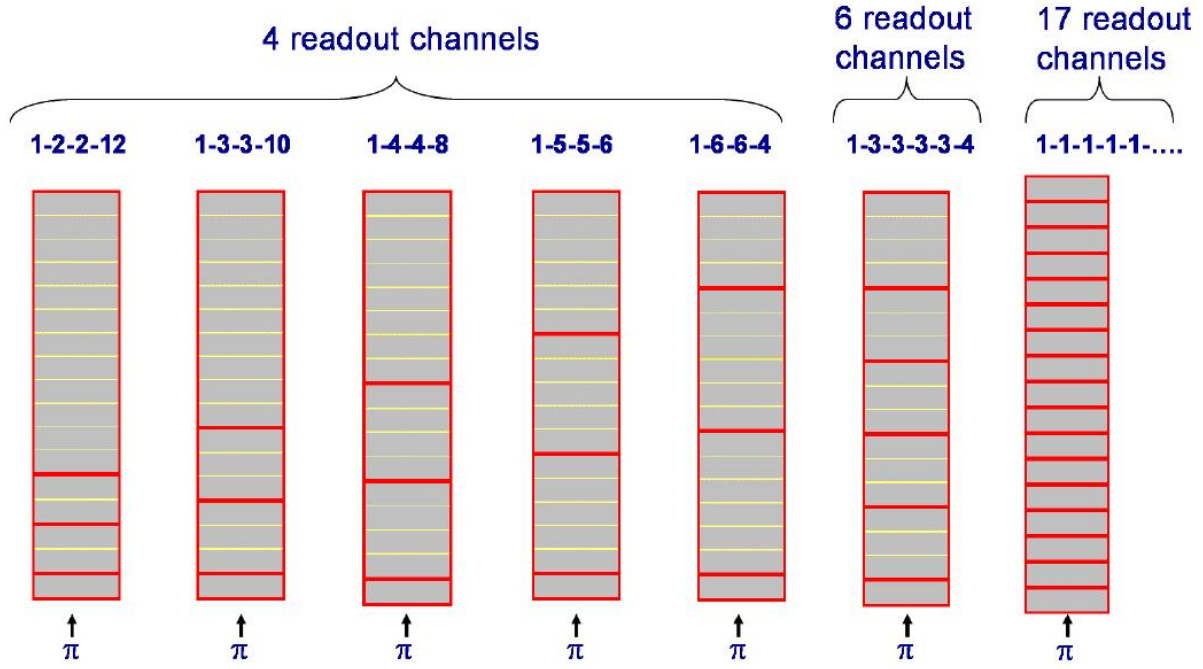


Figure 4.4: Schematic view of the different readout schemes of the CMS HCAL which are investigated in this study. Here, exemplarily one tower is shown. The red border marks the readout cells which are combined to build one readout channel. The pions enter the tower from below.

Furthermore, for the first and last layer, the scintillator signal is multiplied by a factor of 3.7/9. The calibration factor is

$$C = 50/0.398 \approx 126. \quad (4.7)$$

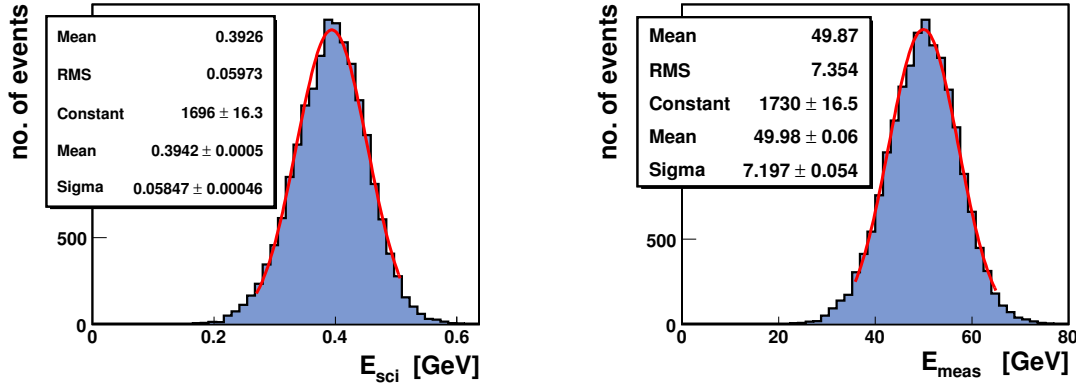


Figure 4.5: Illustration of the calibration of the simulated CMS HCAL with 50 GeV pions. *Left:* scintillator signal before calibration. *Right:* simulated measured energy, consisting of the scintillator signal times a calibration factor “C” plus a small contribution of energy depositions from the ECAL. The red line represents a Gaussian fit to the distribution.

4.2.4 Determination of the Weighting Factors

Detailed studies have been performed in order to develop the weighting configuration presented in the following (some of them can be found in the Appendix A).

Simulated pions enter the middle of one HCAL tower in the central region. The ECAL and the passive material are included in the setup described in Section 4.2.1. Two cuts are applied:

1. $E_{\text{ECAL}} < 2 \text{ GeV}$,
2. $E_{\text{leakage}} < 20 \% \cdot E_{\pi}$.

The first cut is applied to avoid a large energy deposition in the ECAL. The second cut ensures that most of the energy is deposited within the HCAL, which is necessary to develop the WFs for a particular incident pion energy E_{π} . Table 4.2 gives the number of generated events and remaining events after applied cuts.

Table 4.2: Number of simulated pions for different energies and remaining pions after applied cuts.

$E_{\pi} [\text{GeV}]$	10	20	30	50	100	150	225	300
no. of events before cuts	50,000	50,000	50,000	50,000	40,000	35,000	30,000	30,000
no. of events after cuts	22,193	20,792	20,082	19,662	15,106	12,592	10,179	10,267

Simulation studies have shown that the energy weighting yields better results when the first layer is not weighted (see Appendix A.3). Thus, the first layer is never weighted.

For the tabulation of the WFs a *binning* of the energy density is needed (see Section 4.2.4). Since the distribution of the entries of the corresponding histograms is not uniform and the value of the WFs depends more strongly on lower energy densities, it is appropriate to choose a *logarithmic* scale for the WFs (see also [75]).

The WFs depend on the shower energy of the incident pions (see equation (4.2)). Hence, it is necessary to determine a set of them for several energies E_{π} . A *set of WFs* refers to the WFs binned in the energy density for each channel i of a particular readout design and a certain pion energy E_{π} . For each of the following energies a set of WFs is obtained: 10, 20, 30, 50, 100, 150, 225, 300 GeV.

Though, in an experiment it is a priori not clear which set of WFs should be applied for a particular energy deposition. The best hint for the choice of a set of WFs is the measured total shower energy E_{meas} of the particle (see equation (4.5)). However, the WFs are obtained only for several *fixed* pion energies E_{π} . Therefore, an *interpolation* of the WFs between these fixed pion energies is needed. This is done depending on the binned energy density ρ^i and a linear interpolation in E_{π} .²

²Some other interpolations, like a quadratic one or square-root-like interpolation has been briefly investigated. However, the performance of a linear interpolation yields the best results.

A measure for the volume of a readout channel is needed in order to calculate the energy density. Here, the *number of layers* within a readout channel is used. The different thickness of the first and last scintillator is taken into account by a factor of 3.7/9 (see Table 4.1).

According to equation (4.2), the ratio $E_{\text{dep}}^i/E_{\text{meas}}^i$ is needed in order to establish the WFs; the scatter plots in Fig. 4.6 show this ratio as a function of the energy density for each channel of the “1-4-4-8” design, originating from simulated 50 GeV pions.

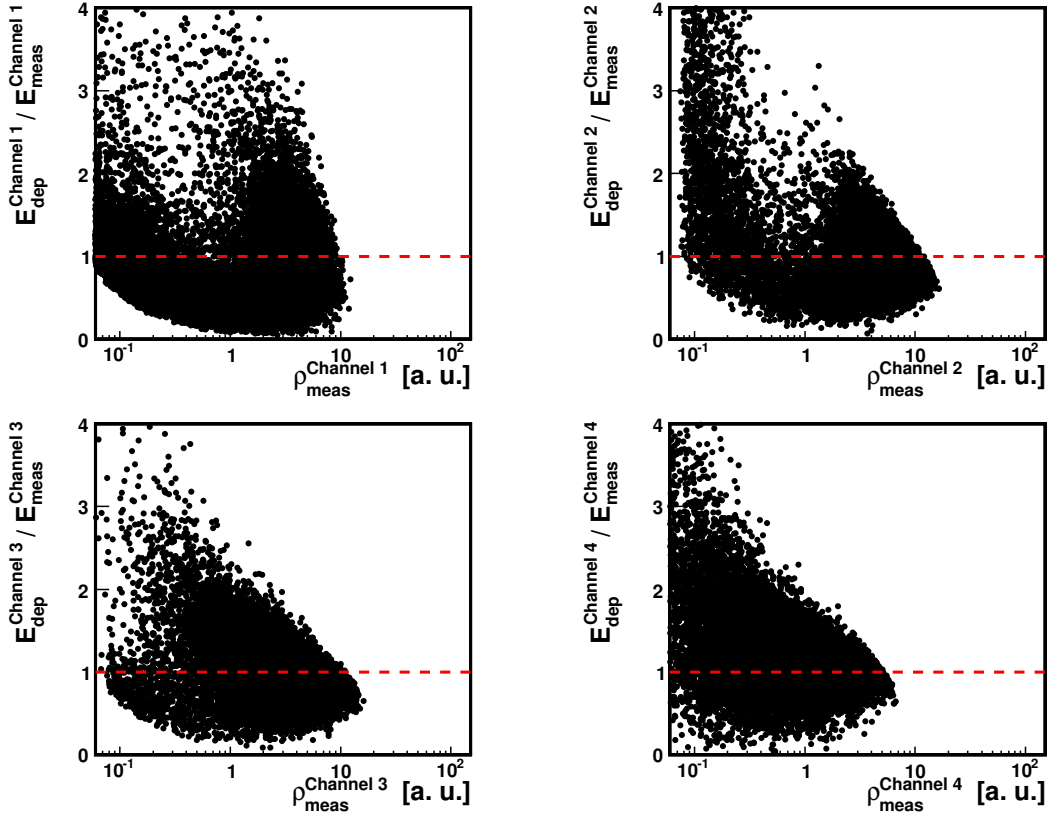


Figure 4.6: Scatter plots of the ratio $E_{\text{dep}}^i/E_{\text{meas}}^i$ for each channel of the “1-4-4-8” design of 50 GeV pions as a function of the energy density. These plots are essential to establish the WFs. The red dashed line marks the ‘1’ where a perfectly measured energy value would be located.

Two populations can be identified: one for small energy densities, caused by late showering particles (especially for channel 1 and channel 2) and one for large energy densities, caused by early showering particles.³ To obtain the WFs from these plots, the x -axis is divided into 150 intervals in a logarithmic scale and plotted as a profile plot (see Fig. 4.7).

If an energy density interval contains only a few entries the WF might be considerably shifted due to statistical fluctuations. To avoid using unrealistic WFs because of these fluctuations, two quality criteria are applied:

³For further investigations a discrimination between these populations could be fruitful. However, in these studies no differentiation is made.

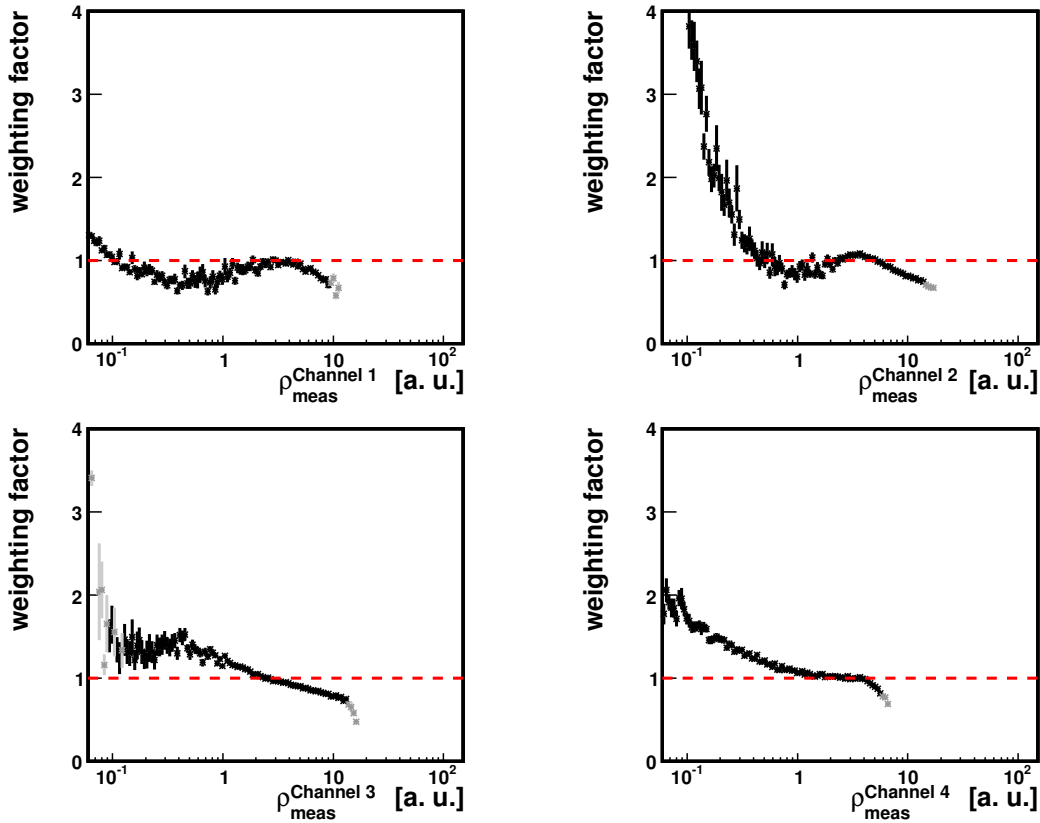


Figure 4.7: Profile plots of the WFs according to Fig. 4.6. The red line marks the '1'. The gray entries are rejected WFs by the quality cuts.

1. Number of entries in an energy density interval > 20 ,
2. $(\text{Statistical error of WF}) / \sqrt{\text{no. of entries}} > 3$.

If these criteria are not fulfilled, the WF '1' is used. The first criterion is chosen to provide a minimum number of entries in a single energy density interval in order to avoid too large statistic fluctuations. The second one refers to the spread of the entries of an energy interval and is chosen according to experience. Entries which appear gray in the plots of Fig. 4.7 did not survive these quality criteria.

A WF can only be applied if there are *two* sets of WFs available for the corresponding energy density. Otherwise, no interpolation of the WFs can be performed. However, the WFs do not cover the complete energy density range (see Fig. 4.2). This is the reason why for the energies 10 and 300 GeV no weighting procedure can be realized — as a boundary effect, for most energy densities of 10 and 300 GeV only *one* set of WFs exists, not two, and thus an interpolation is not possible for these energies (see Fig. 4.2). Note that it would be a conceptual mistake, if the WFs are used without an interpolation in the case where only one set for a certain energy density exists — the WFs would be chosen based on the wrong total shower energy, namely on a certain fixed pion energy.

4.3 Application of the Energy Weighting Method

To demonstrate the effect of the weighting method, the WFs, determined in Section 4.2.4, are applied to statistical independent simulated data samples yielding the following results.

4.3.1 Average Energy Density

The average energy density of 50 GeV pions before and after the energy weighting in comparison with the true energy density (known from simulation) are shown in Fig. 4.8 for the “1-4-4-8” readout design as a function of the layer. Since the thickness of the first and the last scintillator differs from the others, the volume of these channels is multiplied by a factor of 3.7/9 (see Table 4.1). Consequently, the calculated energy density in these layers is smaller, which can be seen in the plot for the first layer.

From Fig. 4.8 it can be seen that the average energy density after the weighting matches much better the true energy density than before the weighting. Only for the first readout channel the energy density before and after the weighting are identical because the first layer is not weighted (see Section 4.2.4).

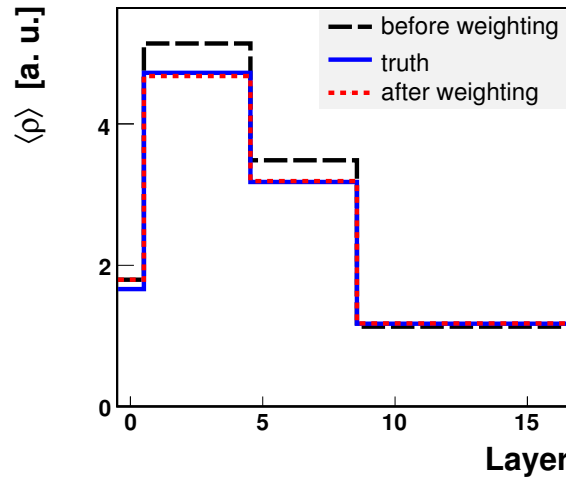


Figure 4.8: Average energy density for 50 GeV pions of each readout channel of the “1-4-4-8” readout design before weighting (black dashed), after weighting (red dotted), and from the simulated true information (blue).

4.3.2 Energy Resolution and Linearity

The energy resolution and linearity before and after applying the weighting method is obtained from the energy distributions of simulated pions with energies E_π . Figure 4.9 illustrates the energy distribution for $E_\pi = 50$ GeV with respect to the “1-4-4-8” readout design.

In this example it can be seen that the width of the distribution is improved after the weighting (right plot), compared to the energy distribution before the weighting (left

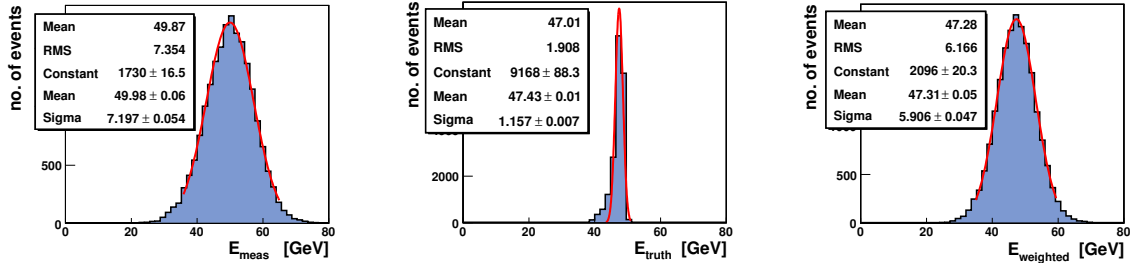


Figure 4.9: Energy distributions of simulated pions of 50 GeV with respect to the “1-4-4-8” readout design before (left) and after (right) the weighting. For comparison the true energy distribution (middle) is shown, too. A Gaussian fit is applied to the distributions in the region indicated by a red line.

plot). Note that the mean of the weighted energy distribution is slightly shifted. This effect originates from the interpolation of the WFs, which can still be optimized and is easily removed by a recalibration of the HCAL. Due to leakage and energy depositions in front of the HCAL the true energy deposition (middle plot) is not a delta peak but has a certain width.

From histograms similar to those of Fig. 4.9, both, the energy resolution and the linearity are derived as a function of the pion energy. This is done via two methods: On the one hand they are obtained from a Gaussian fit to energy distributions. On the other hand they are calculated using the RMS value and the mean of the histogram. Figure 4.10 shows the energy resolution and linearity before and after the weighting for the “1-4-4-8” readout design for both methods. For the linearity, after applying the weighting method, the HCAL was recalibrated to 50 GeV pions. As already mentioned, there are no entries for 10 and 300 GeV due to conceptional considerations (see Section 4.2.4).

Here, the energy resolution is parameterized with:

$$\left(\frac{\sigma_E}{E}\right)^2 = \underbrace{\left(\frac{a [\%]}{\sqrt{E}}\right)^2}_{\text{sampling term}} + \underbrace{(c [\%])^2}_{\text{constant term}}. \quad (4.8)$$

In Table 4.3 and Table 4.4 the energy resolution before and after weighting is shown for a Gaussian fit and for the RMS value and the mean of the energy distributions for different readout designs.

The largest improvement of the energy resolution with four readout channels is obtained using the “1-4-4-8” readout design. The linearity, obtained from the RMS values and the mean of the histograms in Fig. 4.10 (bottom right) also improves. The deviation of the linearity decreases from $\Delta l = 6\%$ ($96\% - 102\%$) to 4% ($97\% - 101\%$).

Using six readout channels, the “1-3-3-3-3-4” shows a larger improvement than any four channel design. However, the improvement of the energy resolution for the readout design “17x1” in which every layer is read out is smaller than for the “1-3-3-3-3-4” readout design. This can be explained by the sampling fluctuations which have a larger influence for the “17x1” readout design than for a four channel readout design (see Appendix A.5).

As a consistency check a statistically independent test sample of 80 GeV has been produced for which no WFs exist. In Fig. 4.11 the energy resolution and linearity including

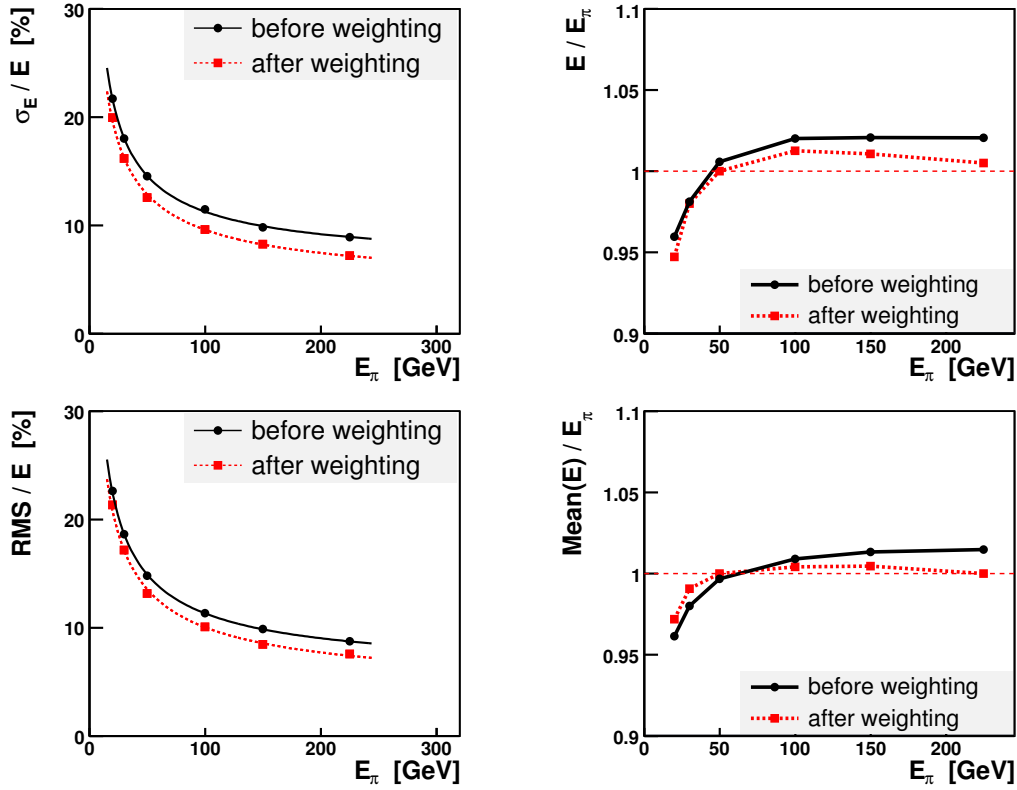


Figure 4.10: Energy resolution (left column) and linearity (right column) before (black) and after the energy weighting (red dashed) for the “1-4-4-8” readout design. The first row shows the results obtained from a Gaussian fit to the energy distributions, while the second row is obtained from the RMS value and mean of the corresponding histograms. The statistical errors are so small that they are lay within the markers.

Table 4.3: Energy resolution before and after weighting obtained from a *Gaussian fit* for different readout designs. The last column shows the relative improvement of the sampling term in percent.

Design	sampling term [%]	constant term [%]	rel. impr. [%]
before weighting	92.2 ± 0.6	6.5 ± 0.1	—
1-2-2-12	86.0 ± 0.5	5.4 ± 0.1	6.7 ± 0.6
1-3-3-10	87.4 ± 0.5	4.8 ± 0.1	5.2 ± 0.6
1-4-4-8	85.4 ± 0.5	4.4 ± 0.1	7.4 ± 0.6
1-5-5-6	86.9 ± 0.5	4.6 ± 0.1	5.8 ± 0.6
1-6-6-4	88.4 ± 0.5	4.4 ± 0.1	4.0 ± 0.6
1-3-3-3-3-4	84.3 ± 0.5	4.5 ± 0.1	8.5 ± 0.6
17x1	86.4 ± 0.5	4.1 ± 0.1	6.3 ± 0.6

Table 4.4: Energy resolution before and after weighting obtained from the *RMS value and mean* of the energy distributions for different readout designs. The last column shows the relative improvement of the sampling term in percent.

Design	sampling term [%]	constant term [%]	rel. impr. [%]
before weighting	96.7 ± 0.4	5.9 ± 0.1	—
1-2-2-12	91.5 ± 0.4	5.0 ± 0.1	5.4 ± 0.4
1-3-3-10	93.8 ± 0.4	4.3 ± 0.1	3.1 ± 0.4
1-4-4-8	90.8 ± 0.4	4.3 ± 0.1	6.1 ± 0.4
1-5-5-6	92.4 ± 0.4	4.1 ± 0.1	4.5 ± 0.4
1-6-6-4	92.8 ± 0.4	4.4 ± 0.1	4.1 ± 0.4
1-3-3-3-3-4	90.1 ± 0.4	4.2 ± 0.1	6.9 ± 0.4
17x1	90.4 ± 0.4	4.4 ± 0.1	6.6 ± 0.4

the entry at 80 GeV are shown. The weighting procedure also works for 80 GeV. For the linearity a small kink is observed which is a hint that the interpolation of the WFs can still be optimized.

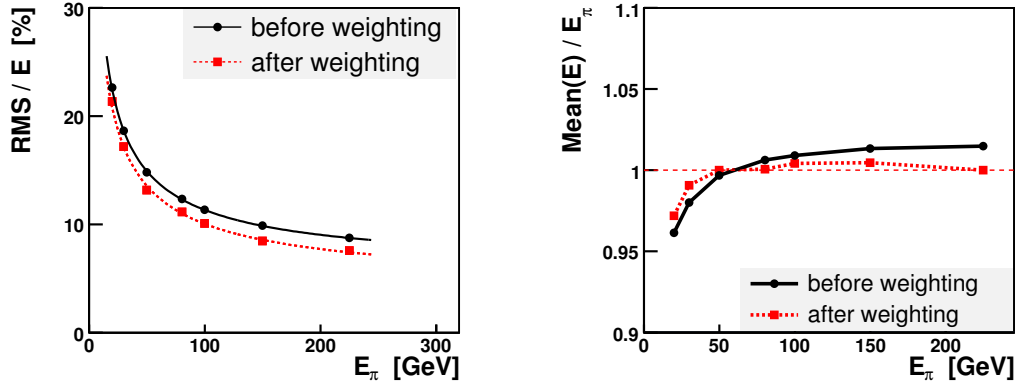


Figure 4.11: Energy resolution (left) and linearity (right) before (solid black) and after weighting (red dashed) for the “1-4-4-8” readout design. As a consistency check the entry at 80 GeV has been added. For this energy no WFs exist. The errors are so small that they lay within the markers.

In summary, the weighting improves the energy resolution and linearity of the simulated CMS HCAL. The “1-4-4-8” readout design yields the best results for the four channel option. Using the six channel “1-3-3-3-3-4” readout design, the improvement of the weighting method is even larger.

4.4 Next Steps for Future Investigations

There are still several aspects which can be investigated in order to improve the energy weighting method. The interpolation of the WFs can be optimized, either by investigating the interpolation in detail (applying an interpolation other than a linear one), or by establishing a two-dimensional fit of the WFs (a function of the energy density and total shower energy).

Another important improvement is a more detailed simulation of the detector. Changes to the standalone simulation code can be performed in order to make the simulation as realistic as possible. However, the best possibility is a modification of the official CMS software such that the simulated energy deposited in the absorber material and the invisible energy can be written out for each HCAL cell. In addition, this would allow to test the method on particle jets. Analogue investigations [76] have shown that the positive effect of the weighting is even more pronounced for jets than for single pions because the inhomogeneities within jets are larger than for single pion showers.

It might be worthwhile to distinguish between early-showering and late-showering particles, since they represent two different populations for the WFs (see Fig 4.6). An evaluation of different WFs for these populations could optimize the weighting method. In this way the energy weighting might also work for the first layer.

Further investigations with different shower models (GHEISHA, FLUKA, MICAP, etc.) can serve to estimate the systematic error.

To test the weighting method with non-simulated data, the WFs can be applied to test beam data of a CMS HCAL test setup. First investigations look very promising and showed an even larger improvement of the energy measurement, using a two dimensional fit of the WFs instead of a one dimensional linear interpolation [81].

5 Event Simulation, Reconstruction, and Selection

The basis of the presented analysis is the event selection with well reconstructed objects. The used objects are muons, electrons, jets, and vertices. Furthermore, the transverse energy and the scalar sum of the jet momenta are relevant for this analysis. These entities are described in the following Sections after introducing the event simulation, the pile-up weighting procedure, the used data, and the trigger strategy, respectively. The preselection and event selection are documented thereafter. Eventually, the event yield of the applied event selection is shown.

5.1 Event Simulation and Simulated Processes

In order to test our understanding of the underlying physics, it is essential to improve or to falsify theoretical models. These models can be simulated with Monte-Carlo (MC) generators. Proton-proton collisions up to the final state as measured in the detector are modeled. For that purpose, the simulation is factorized into several steps:

- Hard scattering process,
- parton showering,
- hadronization,
- reconstruction level.

For the simulation of the hard scattering process, the generators MADGRAPH [82] and PYTHIA6 [83] are used in this analysis. The hard scattering processes are calculated with up to n particles in the final state, depending on the MC generator and the process. MADGRAPH calculates up to $2 \rightarrow 9$ processes, while PYTHIA6 calculates the hard scattering process without any additional partons.

The parton showering step takes initial and final state radiation (ISR/FSR) into account. A matching between the hard partons of the matrix element calculation and the partons of the parton showering step is performed using the MLM algorithm [84] which introduces a phase space cutoff for energetic partons.

In the hadronization step the partons are matched to colorless hadrons. Furthermore, the simulation of the underlying event, considering softer interactions of secondary partons inside the same pp collision and proton remnants, is implemented at this point.

Eventually, all final state particles are passed through a full simulation of the CMS detector (including a trigger chain), based on GEANT4 [85] (except for the rare SM processes explained below and the SUSY scan in Chapter 8).

In Table 5.1 the main SM processes, which are considered in this analysis, are listed. The matrix element generator, the number of simulated events N_{events} , the cross-section σ up to a certain order, the integrated luminosity L of the sample is stated, as well as the corresponding weight, which has to be applied to the simulated events in order to normalize them to the integrated luminosity of 1.1 fb^{-1} . The single top sample is split into six separate samples (s-, t-, and tW-channel for the top and antitop quark.).

For the cross-sections, the most precise calculations available at the time of writing are used. For the $t\bar{t}$, single-top, and WW , the parton-level event integrator MCFM [86] was used. For $W \rightarrow l\nu$ and $Z/\gamma^* \rightarrow l^+l^-$ FEWZ [87] was used, providing full NNLO precision.

The simulated QCD samples are listed in Table 5.2, separately. They are split into “ $\hat{p}_T$ ” bins, which refer to the sum of the transverse momentum of the partons of the hard interaction. This splitting allows to simulate more rare events at high transverse momenta with a relatively small cross-section. Furthermore, these QCD samples are enriched with leptonic processes (one set with muons, the other with electrons). Therefore, a filter efficiency ϵ_{filter} is applied to the cross-section in order to take the branching ratio of the enriched processes into account.

In this analysis, rare SM processes play an important role. In Table 5.3 the rare processes are summarized, comprising also di-boson processes. All rare processes (except for WW) are simulated using MADGRAPH and a simplified detector simulation (“fast simulation”). Some Feynman diagrams of rare processes can be found in the Appendix D.

In Table 5.4 the SUSY signal processes of the CMS benchmark points of the cMSSM (see Section 2.4), are listed. The cross-section is given at LO. A k -factor is applied to the cross-section in order to get a rough estimate of the NLO cross-section. All these samples are generated with PYTHIA6.

For simplicity, some processes are combined in this analysis: “QCD” summarizes all QCD-samples (μ -enriched and e -enriched for all “ $\hat{p}_T$ ” bins); “Single Top” summarizes all channels of the single top samples; “Rare” comprises all rare processes given in Table 5.3.

Table 5.1: Simulated SM processes used in this analysis. The weight corresponds to a normalization of the simulated events to the integrated luminosity of $L = 1.1 \text{ fb}^{-1}$.

Process	Generator	N_{events}	σ [pb]	order	L [pb $^{-1}$]	weight
$t\bar{t}$	MADGRAPH	3,701,947	157.5	NLO	23,504	$48 \cdot 10^{-3}$
single top (s)	MADGRAPH	259,971	2.72	NLO	113,031	$9.7 \cdot 10^{-3}$
single top (t)	MADGRAPH	390,017	42.6	NLO	120,748	$9.1 \cdot 10^{-3}$
single top (tW)	MADGRAPH	814,390	5.3	NLO	153,658	$7.1 \cdot 10^{-3}$
single antitop (s)	MADGRAPH	137,980	1.49	NLO	59,991	$18 \cdot 10^{-3}$
single antitop (t)	MADGRAPH	1,944,826	22	NLO	60,211	$18 \cdot 10^{-3}$
single antitop (tW)	MADGRAPH	809,984	5.3	NLO	152,827	$7.2 \cdot 10^{-3}$
$W \rightarrow l\nu$	MADGRAPH	81,302,581	31,314	NNLO	2,598	$422 \cdot 10^{-3}$
$Z/\gamma^* \rightarrow l^+l^-$	MADGRAPH	35,427,961	3,048	NNLO	11,623	$94 \cdot 10^{-3}$

Table 5.2: QCD samples split into $\hat{p}_T$ bins and enriched with leptons, which is taken into account by a filter efficiency ϵ_{filter} . For all these samples, the PYTHIA6 generator was used and the cross-section is calculated to LO.

QCD sample	N_{events}	σ [pb]	ϵ_{filter}	L [pb ⁻¹]	weight
μ -enriched (15-20)	2, 536, 447	579, 200, 000	0.00254	1.7	636
μ -enriched (20-30)	10, 076, 800	236, 300, 000	0.00518	8.2	133
μ -enriched (30-50)	10, 898, 867	53, 070, 000	0.01090	18	58
μ -enriched (50-80)	10, 830, 209	6, 351, 000	0.02274	75	15
μ -enriched (80-120)	8, 313, 422	785, 100	0.03700	286	3.8
μ -enriched (120-150)	8, 013, 763	92, 950	0.04777	1, 805	0.6
μ -enriched (150- ∞)	4, 045, 565	47, 580	0.05964	1, 426	0.8
e -enriched (20-30)	35, 729, 669	236, 000, 000	0.0104	15	75
e -enriched (30-80)	70, 142, 060	59, 480, 000	0.065	18	60
e -enriched (80-170)	8, 150, 672	900, 000	0.155	58	19

Table 5.3: Simulated rare SM processes used in this analysis. “ WW ” refers to W^+W^- (which does *not* contribute to the irreducible background, see Chapter 6) and “ $WW(DPS)$ ” to $W^\pm W^\pm$ from double parton scattering. The weight corresponds to a normalization of the simulated events to the integrated luminosity of $L = 1.1 \text{ fb}^{-1}$.

Process	Generator	N_{events}	σ [pb]	order	L [pb ⁻¹]	weight
WW	PYTHIA6	4, 187, 885	43	NLO	97, 393	$11 \cdot 10^{-3}$
WW (DPS)	MADGRAPH	400, 000	0.3778	NLO	1, 058, 761	$422 \cdot 10^{-3}$
$GV\text{Jets}$	MADGRAPH	1, 067, 879	56.64	NLO	18, 854	$58 \cdot 10^{-3}$
$t\bar{t}\gamma$	MADGRAPH	51, 691	0.6545	NLO	78, 978	$14 \cdot 10^{-3}$
$t\bar{t}W$	MADGRAPH	1, 089, 608	0.1633	NLO	6, 672, 431	$164 \cdot 10^{-6}$
$t\bar{t}Z$	MADGRAPH	1, 467, 136	0.139	NLO	10, 554, 935	$104 \cdot 10^{-6}$
$W\gamma^* \rightarrow l\nu\mu\mu$	MADGRAPH	99, 990	1.604	NLO	62, 338	$1.8 \cdot 10^{-3}$
$W\gamma^* \rightarrow l\nu\tau\tau$	MADGRAPH	48, 188	0.287	NLO	167, 902	$6.5 \cdot 10^{-3}$
$W\gamma^* \rightarrow l\nu ee$	MADGRAPH	266, 882	5.546	NLO	48, 122	$23 \cdot 10^{-3}$
$WW\gamma$	MADGRAPH	55, 430	0.177	NLO	313, 164	$3.5 \cdot 10^{-3}$
WWW	MADGRAPH	1, 391, 777	0.038	NLO	36, 625, 711	$30 \cdot 10^{-6}$
WWZ	MADGRAPH	380, 423	0.0268	NLO	14, 194, 888	$77 \cdot 10^{-6}$
$WZ + \text{Jets} \rightarrow 3l\nu$	MADGRAPH	231, 134	0.856	NLO	270, 016	$4.1 \cdot 10^{-3}$
WZZ	MADGRAPH	2, 020, 469	0.0088	NLO	229, 598, 750	$4.8 \cdot 10^{-6}$
$ZZ + \text{Jets} \rightarrow 4l$	MADGRAPH	345, 188	0.076	NLO	4, 541, 947	$241 \cdot 10^{-6}$
ZZZ	MADGRAPH	2, 005, 143	0.00288	NLO	696, 230, 208	$1.6 \cdot 10^{-6}$

Table 5.4: SUSY signal processes of the CMS benchmark points (LM points) of the cMSSM. If the k -factor is multiplied with the cross-section, given at LO, a rough estimate of the NLO cross-section is obtained.

Process	N_{events}	σ [pb]	k-factor	L [pb^{-1}]	weight
LM0	425,825	38.93	1.41	7,758	0.14
LM1	55,000	4.888	1.34	8,397	0.13
LM2	11,000	0.6027	1.33	13,723	0.08
LM3	36,475	3.438	1.40	7,578	0.14
LM4	21,190	1.879	1.35	8,354	0.13
LM5	10,595	0.4734	1.34	16,702	0.07
LM6	8,570	0.3104	1.30	21,238	0.05
LM7	16,500	1.209	1.11	12,295	0.09
LM8	10,595	0.7300	1.41	10,293	0.11
LM9	79,665	7.134	1.48	7,545	0.15
LM11	11,000	0.8236	1.35	9,893	0.11
LM12	48,690	4.414	1.34	8,232	0.13
LM13	77,000	6.899	1.42	7,860	0.14

5.2 Pile-up Weighting

With increasing luminosity of the recorded data, the pile-up distribution changes and becomes more important. The simulated processes must be adapted to the pile-up distribution in data in order to avoid a systematic bias of the simulation. This is done by performing a weighting procedure of the simulated processes on an event-by-event basis, before any selection requirement is applied.

In Fig. 5.1 the normalized pile-up distribution for data is shown in comparison with the normalized *unweighted* pile-up distribution of the simulated SM processes (for a better illustration the distributions are displayed in two different views). Thereby, “number of pile-up” is the number of pile-up interactions within one bunch crossing in addition to the primary vertex. The data distribution and the distribution from the simulated processes show a considerable difference, pointing out the necessity of a pile-up weighting procedure.

In Fig. 5.2 the normalized pile-up distribution for data is displayed in comparison with the normalized *weighted* pile-up distribution of the simulated SM processes. Consequently, these distributions agree, showing that the pile-up weighting procedure is successfully performed.

Note, that the pile-up distribution in data is not perfectly known and therefore, the uncertainty due to the pile-up weighting must be considered (see Section 7.6). The pile-up distribution in simulation is roughly flat up to a number of pile-up of 10, followed by a Poisson distribution with an average of 10 interactions per crossing for larger pile-up numbers.¹

¹The recommendation for the simulated pile-up distribution might change with new MC productions.

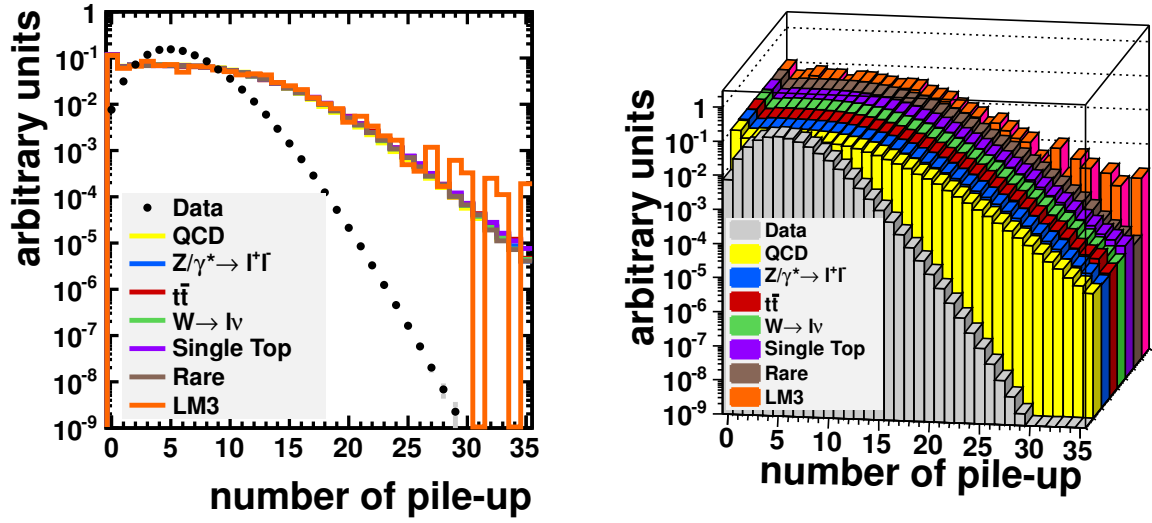


Figure 5.1: Pile-up distributions for data and simulated SM processes without a weighting procedure. All these distributions are normalized to unity. “number of pile-up” is the number of pile-up interactions within one bunch crossing in addition to the primary vertex. For a better illustration both plots show the same in a different view, *left*: 2D-view, *right*: 3D-view.

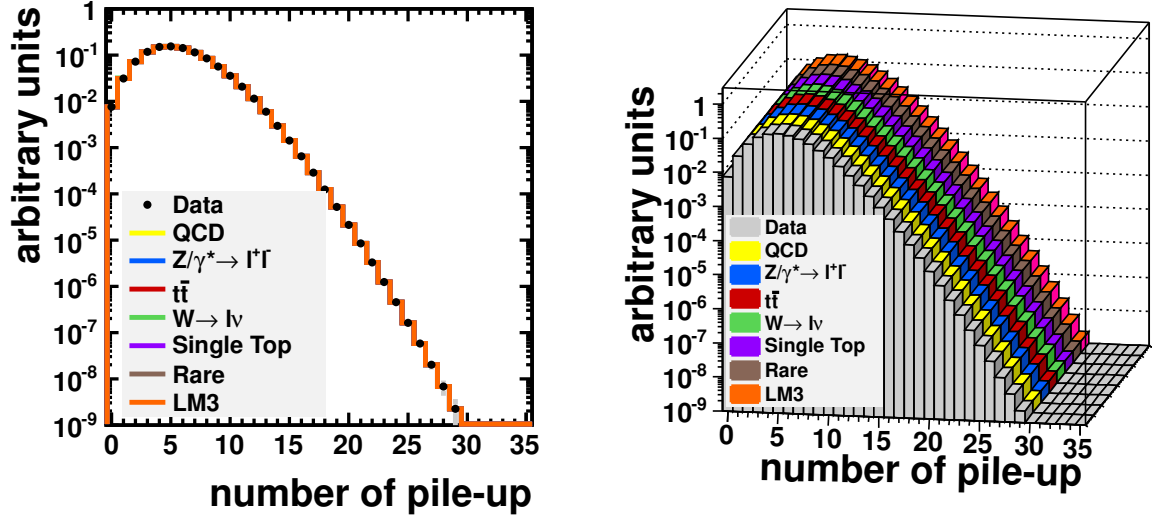


Figure 5.2: Pile-up distributions for data and simulated SM processes including a weighting procedure. All these distributions are normalized to unity. “number of pile-up” is the number of pile-up interactions within one bunch crossing in addition to the primary vertex. For a better illustration both plots show the same in a different view, *left*: 2D-view, *right*: 3D-view.

5.3 Used Data and Trigger Strategy

This analysis is based on a data sample corresponding to an integrated luminosity of about $L = 1.1 \text{ fb}^{-1}$ collected in pp collisions by the CMS experiment at a center-of-mass energy of $\sqrt{s} = 7 \text{ TeV}$. The data is divided into “runs” in which the beam conditions and detector status are stable. The runs, in turn, are divided into “lumi sections”, being defined as 2^{18} beam orbits which corresponds to about 23.3 seconds of CMS data taking.² Each run is certified in order to decide, whether it can be used in a physics analyses, based on the status of the individual detector components.

The run range used in this analysis is 160,431 – 167,913. From these runs, the events of interest have to be selected. This is done via triggers (see Section 3.2.6). For distinct triggers certain requirements must be fulfilled. They have e.g. a certain threshold for the transverse momentum of a lepton or a requirement on the hadronic activity (H_T) in the detector. In order to ensure a full efficiency of the triggers, the object and event selection must be chosen well above the trigger thresholds. Hence, the trigger strategy is closely related to the object definition and event selection.

In principle, the trigger strategy together with the object and event selection are chosen in order to probe a kinematic phase space as large as possible, for being most sensitive to a potential SUSY signal. Therefore, two different selection strategies [88, 89] are defined: *High- p_T selection*, using leptonic triggers with a relatively high lepton p_T threshold (but without a H_T requirement), *low- p_T selection* (not applied in this analysis), using cross triggers, allowing to choose a lower lepton p_T threshold, but implying a H_T requirement, which reduces the kinematical phase space of the hadronic activity of an event.

Another aspect of the kinematic phase space of leptonic SUSY analysis is illustrated in Fig. 5.3. The masses of different SUSY particles are denoted as m_A , m_B , and m_C and their mass differences as $\Delta m_{AB} := m_A - m_B$, etc. The high- p_T and low- p_T selection can be considered as follows:

- *high- p_T selection*: $\Delta m_{AB} < \Delta m_{BC}$: The mass difference Δm_{AB} is small, while Δm_{BC} is large. Hence, most energy is released by the decay of particle B . Therefore, the signature is expected to have relatively low hadronic activity and high E_T^{miss} (originating from the released neutrino and LSP).
- *low- p_T selection*: $\Delta m_{AB} > \Delta m_{BC}$: The mass difference Δm_{AB} is large, while Δm_{BC} is small. Hence, most energy is released by the decay of particle A . Therefore, the signature is expected to have relatively large hadronic activity and low E_T^{miss} .

There is also a third search region aiming to select τ -enhanced final states, leading to the decay channels $\tau\tau$, $\mu\tau$, and $e\tau$. Since the decay of τ s results in additional hadronic activity, the search region implies a large H_T requirement and hadronic cross triggers are used, allowing to lower the p_T requirement for leptons. τ -enhanced final states are favored in SUSY scenarios with a relatively high value for the $\tan\beta$ parameter.

This analysis follows the *high- p_T selection* strategy, using di-lepton triggers. Table 5.5 shows these triggers for the different decay channels of this analysis.

²The lumi sections consist of “lumi nibbles”. One nibble contains the information of four bits. Hence, the definition of a lumi section is a power of 2.

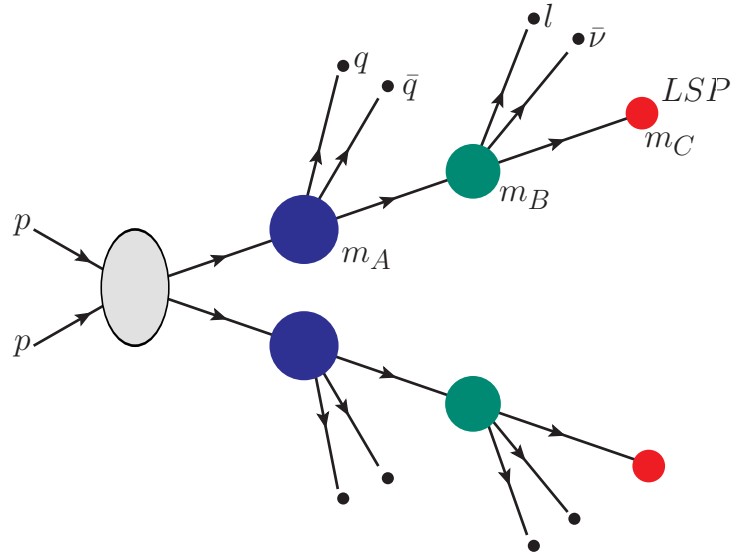


Figure 5.3: Sketch illustrating the kinematics of a leptonic SUSY decay cascade. The variables m_A , m_B , and m_C stand for the masses of different SUSY particles.

Table 5.5: Used high-level triggers for the different decay channels. The abbreviations Mu and Ele stand for muon and electron, respectively. The subsequent number gives the threshold for the transverse momentum of the lepton. To simplify the electron trigger names, the following abbreviations are used: XX = CaloIdL_CaloIsoVL, YY = CaloIdT_TrkIdVL_CaloIsoVL_TrkIsoVL.

Channel	Trigger
$\mu\mu$	HLT_DoubleMu7 HLT_Mu13_Mu8
$e\mu$	HLT_Mu8_Ele17_CaloIdL HLT_Mu17_Ele8_CaloIdL
ee	HLT_Ele17_XX_Ele8_XX HLT_Ele17_YY_Ele8_YY

Trigger efficiencies

The triggers used in this analysis are applied in data and simulation. However, the simulated efficiencies do not necessarily correspond to the efficiency in data. Hence, a data-to-MC scale-factor might be used, to account for the differences. For data the trigger efficiencies of the presented triggers have been investigated in considerable detail [90, 91]. A simple test in simulation (see Appendix C.1) yields trigger efficiencies which are in good agreement with the efficiencies in data. Hence, no trigger efficiency scale-factor needs to be applied.

5.4 Physics Objects Reconstruction in CMS

Objects like muons, electrons, or jets leave a characteristic signature in the detector. They must be *reconstructed* from the raw data of an event, which merely contains information like e.g. single hits of a module or the energy deposition in a certain calorimeter cell. Hence, reconstruction methods must be developed in order to obtain physics objects which can be investigated in an analysis.

Different kinds of particles leave signatures in different parts of the detector, corresponding to their underlying interaction. In Fig. 5.4 an illustration is given on how the particles are detected in the CMS experiment.

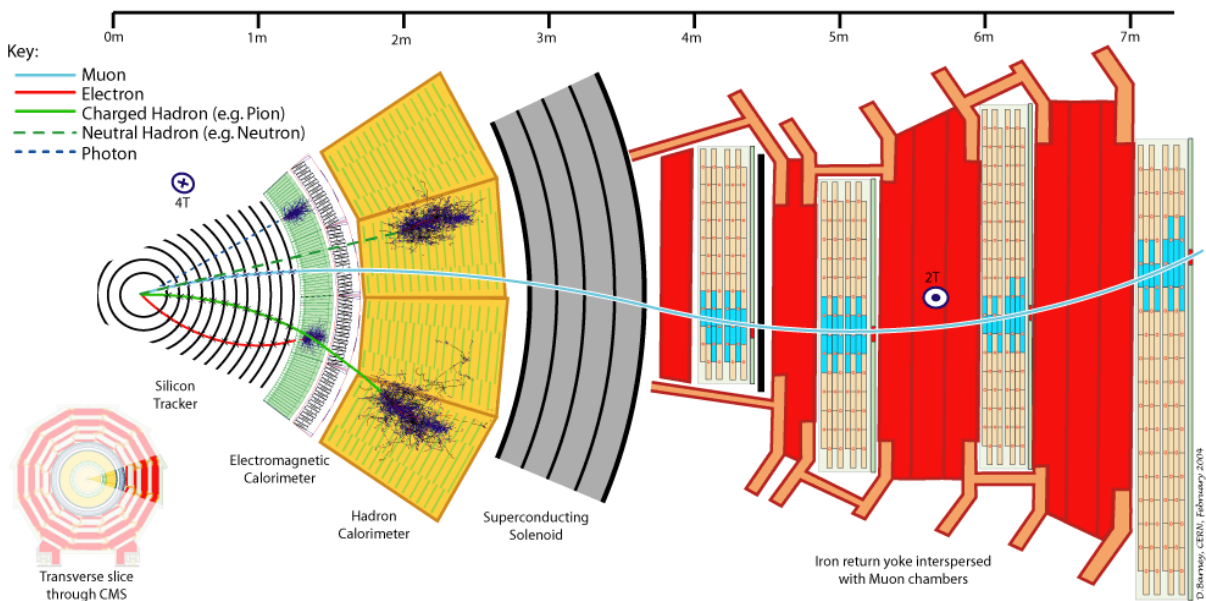


Figure 5.4: Transverse slice through the CMS detector. The different signatures of different particles are illustrated. From [92].

Information from different subsystems can be combined in order to reconstruct physics objects. There is a wide range of possibilities to realize the reconstruction. However, in the following, only the object reconstruction relevant for the presented analysis is described.

There are also ambiguities in the object reconstruction, e.g. it is not always clear, which hit in the pixel detector, or which energy deposition in the calorimeter system belongs to which object, especially when they overlap. Therefore, it is necessary to apply quality requirements (“object selections”), in order to obtain a reliable object reconstruction. These requirements are also documented in the following. The particular values for the requirements (recommended by the CMS collaboration) are the result of detailed studies. They are optimized to yield the best compromise between purity and efficiency of the reconstruction. For this analysis, the default recommendation by the CMS collaboration for the object reconstruction is used and documented in detail in the following.

5.4.1 Muon Reconstruction

Muons cannot be stopped in the CMS detector and pass all its components. They are expected to leave only a small amount of energy in the calorimeters since they are Minimum Ionizing Particles (MIP) in a large energy range (see the Bethe-Bloch formula [93, 94]).

For the muon reconstruction [95], the information of the MS and the STS is exploited. Two reconstruction algorithms are used: *global muon* (outside-in) and *tracker muon* (inside-out).

The global muon reconstruction starts from the MS. A track reconstruction is initiated from a seed in a segment of a DT or of a CSC. In an iterative way, the track is extrapolated layer by layer using a helix parametrization. Subsequently, this track is extrapolated to the STS. If there is a matching track in the STS, the information of both, the MS and STS is combined to perform a global fit of the track. This procedure improves the muon momentum resolution considerably for transverse momenta of $p_T > 200$ GeV/c compared to the momentum resolution which is achieved using only one sub-detector (see Fig. 3.12).

The tracker muon reconstruction starts from a reconstructed track in the STS, which is extrapolated to the MS. Energy loss due to interactions with the material between the STS and MS is taken into account. An uncertainty is applied based on possible multiple scattering processes. If a matching track is found in the MS within the applied uncertainty, a tracker muon is reconstructed. In case of ambiguities concerning track reconstruction or matching, the candidate with the smallest χ^2 -value³ is chosen.

In rare cases (about 1%) a muon can only be reconstructed in the MS. These muons are referred to as *standalone muons*. However, muons originating from cosmic radiation are mostly reconstructed as standalone muons. This is due to the fact, that the volume of the MS is much larger than the volume of the STS. Thus, the geometric acceptance of the MS exceeds the acceptance of the STS by far.

A key quantity of reconstructed muon candidates is the isolation (see Fig. 5.5 for illustration). It characterizes the amount of activity in the vicinity of a muon. The isolation can be calculated by summing up energy depositions in a cone of a certain size around the muon candidate, except the muon itself. This is done for the STS, ECAL, and HCAL separately, where for the STS the transverse momentum p_T^{STS} of the neighboring tracks is used and for the ECAL and HCAL the transverse energy E_T^{ECAL} and E_T^{HCAL} is applied, respectively.

The *relative* isolation I_{rel} is obtained by dividing the isolation by the transverse momentum p_T^μ of the muon

$$I_{\text{rel}} := \frac{1}{p_T^\mu} \sum_{\Delta R < 0.3} (p_T^{\text{STS}} + E_T^{\text{ECAL}} + E_T^{\text{HCAL}}), \quad (5.1)$$

where all energy deposits with a maximal distance of $\Delta R = 0.3$ to the muon direction at the vertex are summed up. The smaller the value of I_{rel} the more likely it is that the muon is a real muon and not e.g. a mis-identified jet.

To ensure a high quality muon reconstruction, additional muon selection criteria are required.

³The χ^2 -value is the quadratic sum of the deviations of single hits from the reconstructed track (“residuals”). Thus, it yields the possibility to evaluate how good a track fits to the measured hits.

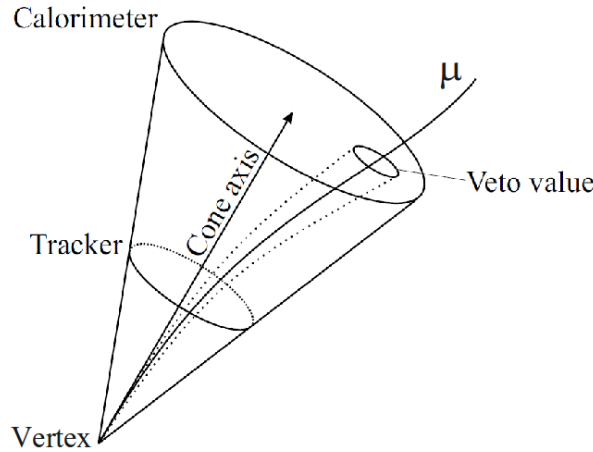


Figure 5.5: Illustration of muon isolation. Based on [54].

Muon Selection

Muons are required to be successfully reconstructed [96] as a global muon and a tracker muon. The transverse momentum requirement is chosen with respect to the trigger thresholds. To account for the detector acceptance, muons are only considered in a range of $|\eta| < 2.4$. They must have a minimum number of hits in the silicon tracker $N_{\text{hits, STS}}$ and minimum number of hits in the muon detectors $N_{\text{hits, muon}}$. The requirement of a maximum χ^2/N_{dof} value ensures a high quality of the muon track fit, based on the single hits in the STS. Furthermore, calorimeter energy deposits of muons are required to be consistent with originating from a minimum ionizing particle and therefore, the energy deposits have an upper threshold (ECAL/HCAL non-MIP veto). In order to ensure that muons originate from the primary interaction, the distance of the muon impact parameter $|d_{0, \text{pv}}|$ to the primary vertex must not be too large. Finally, muons need to fulfill an isolation requirement.⁴

Details of the muon requirements are given in Table 5.6.

Muon Reconstruction Efficiency

The muon reconstruction efficiency has been studied in considerable detail using a T&P method [90, 97]. It has been shown, that the muon reconstruction efficiency in data and simulation (identification and isolation as a function of muon p_T) agrees well within the uncertainty and no scale-factor needs to be applied. However, an *uncertainty* due to the reconstruction efficiency can be applied (see Section 7.2). Note, that a correction due to inefficiencies would be mandatory (since they lower the measured event yields) if the *absolute* events yield are of importance, as e.g. in the case of a cross-section measurement.

⁴Note, that for illustrative purposes, the isolation requirement might be applied in the last step of the event selection (see Section 5.6).

Table 5.6: Muon selection requirements. ID refers to the identification of a muon candidate as a global or tracker muon.

Observable	Value or Range
p_T	$> 10 \text{ GeV}/c$
$ \eta $	< 2.4
χ^2/N_{dof}	< 10
ID	Tracker and Global
$N_{\text{hits, STS}}$	≥ 11
$N_{\text{hits, muon}}$	≥ 1
$ d_{0, \text{pv}} $	< 0.02
ECAL/HCAL non-MIP veto	$< 4/6 \text{ GeV}$
I_{rel}	< 0.15

5.4.2 Electron Reconstruction

The main component to detect electrons is the ECAL. Furthermore, they leave a signature in the STS since they are charged. Electrons entering the ECAL interact with the dense detector material (PbWO_4) and produce electromagnetic showers of electron-positron pairs and bremsstrahlung photons before they are stopped. Thus, they deposit energy in ECAL cells. Neighboring cells containing energy deposits form *clusters*, multiple clusters can be combined to Super Clusters (SC). They are spread in ϕ direction because of bremsstrahlung, which is radiated from electrons as they are bent in the magnetic field (see Fig. 5.6).

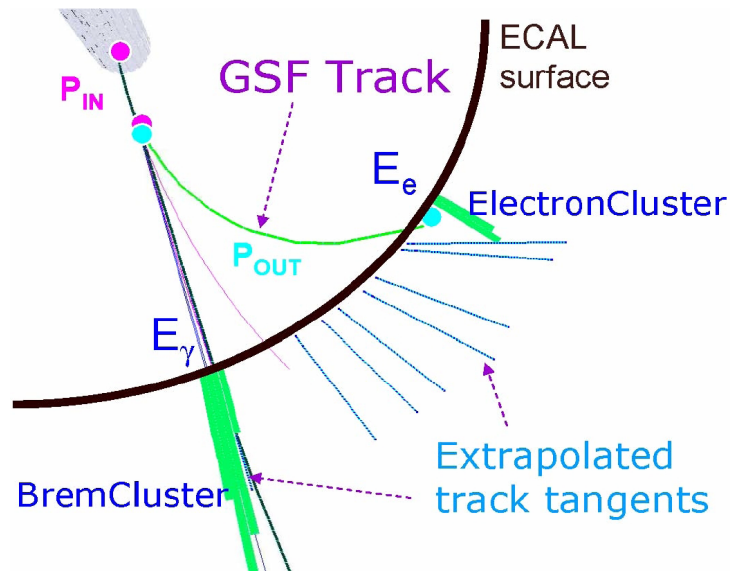


Figure 5.6: Illustration of an electron radiating a photon while traveling in the magnetic field of the CMS detector, leaving a signature in the ECAL. From [98].

The electron reconstruction [99] starts with SCs in the ECAL, which are matched to tracks reconstructed in the STS. These tracks are obtained by performing a fit of the observed hits in single STS modules with a Gaussian Sum Filter (GSF) [100]. The GSF can also handle tracks of electrons which radiate photons. Electrons that are seeded by a SC in the ECAL and which can successfully be matched to a GSF-track are called GSF-electrons.

In analogy to the muon reconstruction, a key quantity of reconstructed electron candidates is the relative isolation I_{rel} ⁵ defined as

$$I_{\text{rel}} := \frac{1}{p_{\text{T}}^e} \sum_{\Delta R < 0.3} (p_{\text{T}}^{\text{STS}} + E_{\text{T}}^{\text{ECAL}} + E_{\text{T}}^{\text{HCAL}}), \quad (5.2)$$

where all energy deposits with a maximal distance of $\Delta R = 0.3$ to the electron direction at the vertex are summed up. The smaller the value of I_{rel} the more likely it is that the electron is a real electron and not e.g. a mis-identified jet.

For the electron charge reconstruction there are three different methods (requiring that all three charge measurements agree strongly reduces the probability of charge a mis-measurement):

- GSF-track: The curvature of the reconstructed GSF-track defines the sign of the reconstructed charge.
- CTF-track (“Combinatorial Track Finder”): Track algorithm which uses both, Pixel Detector and Silicon Strip Chambers for seed finding and an iterative process going from layer to layer of the STS, taking into account multiple scattering using Kalman Filter (KF) [101] techniques. The curvature of the reconstructed CTF-track defines the sign of the reconstructed charge.
- SC: A charge estimation can also be performed from the SC. It is assumed that the largest energy deposition of an electron originates from the electron itself rather than from its bremsstrahlung. Hence, from the distribution of the energy depositions a charge estimation is performed.

To ensure a high quality electron reconstruction, additional electron selection criteria are required.

Electron Selection

The electron reconstruction [102] must be seeded by a cluster in the ECAL. Further electron identification variables, based on the shower shape, are applied to distinguish real electrons from jets. There are several levels of requirements, corresponding to different electron identification (ID) efficiencies. Here, the “working point 80” (WP80) criteria are used, designed to be 80 % efficient at selecting real electrons. To be more explicit, these variables are:

⁵Note that in the presented analysis, GSF-electrons are used, and not particle flow electrons (for an explanation of particle flow objects, see jet reconstruction). This is important because particle flow electrons also imply certain isolation requirements. However, in this analysis, isolation is a key quantity and a delicate subject to manipulate. Thus, using particle flow electrons could bias the background estimation.

- $\sigma_{i\eta i\eta}$: Measures the width of the electromagnetic cluster in the ECAL in η -direction ($i\eta$ refers to the i th detector element in η -direction). A single electron is expected to leave a deposit with a small spread in η -direction.⁶
- $\Delta\phi_{\text{In}}$: Is the difference between the electron *inner* track, extrapolated from the vertex, and the ϕ value of the SC in the calorimeter.
- $\Delta\eta_{\text{In}}$: Is the same as $\Delta\phi_{\text{In}}$, but in η -direction.
- H/E : Is the ratio of energy deposit in the HCAL and ECAL. This quantity is also characteristic to distinguish electromagnetic from hadronic showers.

In Table 5.7 the values for the shower shape variables are given.

Table 5.7: Electron shower shape requirements for the working point with 80 % identification efficiency. B/E stands for Barrel/Endcap.

Observable	Value or Range
$\sigma_{i\eta i\eta} (B/E)$	$< 0.01/0.03$
$\Delta\phi_{\text{In}} (B/E)$	$< 0.06/0.03$
$\Delta\eta_{\text{In}} (B/E)$	$< 0.004/0.007$
$H/E (B/E)$	$< 0.04/0.025$

For the background prediction, looser shower shape requirements are used with a working point of 90 %. The corresponding values are given in Table 5.8 for completeness.

Table 5.8: Loose electron identification requirements for the working point with 90 % electron identification efficiency. B/E stands for Barrel/Endcap.

Observable	Value or Range
$\sigma_{i\eta i\eta} (B/E)$	$< 0.01/0.03$
$\Delta\phi_{\text{In}} (B/E)$	$< 0.08/0.7$
$\Delta\eta_{\text{In}} (B/E)$	$< 0.007/0.009$
$H/E (B/E)$	$< 0.12/0.05$
$I_{\text{rel}} (B/E)$	$< 1.0/0.6$

A conversion rejection is applied to reject electrons originating from converted photons, rather than from the primary interaction. The conversion rejection comprises three criteria:

⁶Due to the magnetic field of the solenoid, bremsstrahlung is released in ϕ -direction. Therefore, the spread in ϕ -direction can be very large and no requirement is set.

- No missing hits in the STS: Since photons are neutral and do not leave a track in the STS, electrons produced in photon conversions typically miss hits in the tracker.
- d_0 : This is a geometrical distance variable in the r - ϕ -plane. If an electron (positron) originates from photon conversion, it is expected to have a partner track from a positron (electron). If a nearby track is found with opposite charge (within a cone of $\Delta R < 0.3$ to the GSF-track), the electron is rejected. The distance characterizes the distance of the tracks and is defined to be negative in case the two tracks overlap.
- $\Delta \cot \theta$: This is another geometrical variable in the r - ϕ -plane which also characterizes a possible partner track of the electron.

In Table 5.9 the criteria for the conversion rejection are given.

Table 5.9: Electron conversion rejection requirements.

Observable	Value or Range
Missing pixel hits	= 0
$ \Delta \cot \theta $	< 0.02
$ d_0 $	< 0.02

In order to ensure that electrons originate from the primary interaction, the distance of the electron impact parameter $|d_{0, \text{pv}}|$ to the primary vertex must not be too large. A cross cleaning with muons is performed to remove electron candidates coming from muon bremsstrahlung — if a muon is found within a cone of $\Delta R < 0.1$ around the electron axis, the electron is removed. To account for the detector acceptance, electrons are only considered in a range of $|\eta| < 2.4$, excluding the range of $1.4442 < |\eta| < 1.566$ to account for a small ECAL gap. The transverse momentum requirement is chosen with respect to the trigger thresholds. To avoid a mis-reconstruction of the electron charge, the charge measurement of the three different reconstruction algorithms must be consistent. Finally, electrons need to fulfill an isolation requirement.⁷

Details of the electron requirements are given in Table 5.10.

Electron Reconstruction Efficiency

The electron reconstruction efficiency has been studied together with the muon reconstruction efficiency [90, 97] (see previous Section) and shows analogous results (agreement in data and simulation within the uncertainties; no scale-factor needs to be applied).

⁷Note, that for illustrative purposes, the isolation requirement might be applied in the last step of the event selection (see Section 5.6).

Table 5.10: Electron selection requirements. WP80 refers to the electron identification working point with an efficiency of 80 %.

Observable	Value or Range
p_T	$> 10 \text{ GeV}/c$
$ \eta $	$< 2.4, \notin [1.4442, 1.566]$
Electron ID	WP80 (shower shape)
$ d_{0, \text{pv}} $	< 0.02
$\Delta R(\mu)$	> 0.1
conversion rejection	true
Seed	ECAL-driven
I_{rel}	< 0.15
charge consistency among CTF, GSF and SC	

5.4.3 Jet Reconstruction

A jet is a bundle of (mostly) collimated hadrons, originating from a single quark or gluon of a hard process. Due to confinement single quarks or gluons from the hard interaction hadronize and many hadrons are created. In general, these hadrons travel into a similar direction as the original particle (because of momentum conservation). However, it is possible that some hadrons have a larger momentum relative to the original particle than usual. Hence, it can be difficult to assign single hadrons to a certain jet. Similarly, an overlap of two or more jets can lead to ambiguities in the attribution of hadrons to jets. Furthermore, Final State Radiation (FSR) possibly leads to a loss of particles of a jet, whereas Initial State Radiation (ISR), pile-up, and underlying events might contribute to a reconstructed jet.

Therefore, it is a challenge to develop appropriate algorithms in order to reconstruct jets [103] in a reliable way. Other important requirements for jet algorithms are *infrared safety* and *collinear safety*. Infrared safety refers to the robustness against the addition of soft (low-energetic) particles originating from long distance interactions. Collinear safety implies the stability of the algorithm in cases where a hard (energetic) particle splits into two or more softer collinear ones.⁸

Here, in particular the anti- k_T algorithm and the Particle Flow (PF) jet type are discussed, since they are used in this thesis. In principle, there are two types of jet algorithms:

- Cone algorithms (e.g. SIS-Cone [104], Iterative Cone [105]), which maximize the energy flow within a cone of a certain radius R .
- Clustering algorithms (e.g. k_T [106], anti- k_T [107]), which successively combine quantities based on their distance to each other. These algorithms are infrared and collinear safe.

⁸If an algorithm has an energy threshold it is not collinear safe. In cases where a particle produces two collinear particles from which one does not pass the threshold, the algorithm yields a different result from what would have been measured if the original particle did not split.

Anti- k_T algorithm

The anti- k_T algorithm [107] clusters particles into jets based on their distance in space and transverse momentum. Therefore, two distance quantities are defined

$$\begin{aligned} d_i &:= k_{Ti}^{2p} \\ d_{ij} &:= \min(k_{Ti}^{2p}, k_{Tj}^{2p}) \frac{\Delta_{ij}^2}{R^2}, \end{aligned} \quad (5.3)$$

where $\Delta_{ij}^2 := (y_i - y_j)^2 + (\phi_i - \phi_j)^2$ and k_{Ti} , η_i , and ϕ_i are the transverse momentum, rapidity, and azimuthal angle of particle i . R corresponds to a cone radius and is set to $R = 0.5$ in the presented study. The relative power of the energy compared to the geometrical scale Δ_{ij}^2 is given by p and has a value of $p = -1$ in the case of the anti- k_T algorithm.

The algorithm compares d_i with d_{ij} :

- If $\exists j \mid d_{ij} < d_i$, i and j are merged.
- If $d_{ij} > d_i \forall j$, i is called a jet and removed from the list of quantities.

By this procedure collinear and infrared safe jets are built.

Particle Flow

The concept of Particle Flow (PF) [108] is an approach to reconstruct all stable particles of an event in a consistent way. Thereby, the information of sub-detectors is combined in a sophisticated way, resulting in an improvement of the measurement. Fundamental elements like tracks and clusters are reconstructed in the first step of the PF algorithm. Based on their position in ϕ and η they are linked to each other. Subsequently, particles are reconstructed from these blocks. These particles can be taken as input, e.g. for the anti- k_T algorithm to form jets.

Jet Energy Corrections

Jet Energy Corrections (JECs) [109] need to be applied to jets since there are various effects which distort the measurement and make it more difficult to translate the measured jet energy to the true particle or parton energy. The goal is, that the corrected jet energy is on average closer to the energy of the original particle than the uncorrected one. The correction is applied as a multiplicative factor C to each component of the raw jet four-momentum vector p^{raw} . The factor C can depend on various jet related quantities (p_T , η , flavor, etc.). A factorization approach is performed in order to realize the JECs. Each correction level takes care of a different effect. Multiple corrections are applied sequentially, having a fixed order. Although there are data-driven techniques, studies based on simulation are used to derive the correction factors, as for the latter the uncertainties of the correction factors are smaller, since the amount of available data is limited and the detector response is well simulated.

The following correction steps are performed in this analysis:

- *Level 1*: Offset correction (C_{offset}), to subtract energy not associated with the high- p_T scattering. This energy originates from both, pile-up (multiple interactions in the same bunch crossing) and electronic noise.

- *Level 2*: Relative correction (C_{rel}), to account for a non-uniform calorimeter response in η . The correction factor depends on η and can be obtained utilizing a p_{T} -balance technique. Due to momentum conservation, the two jets of di-jet events are expected to be back-to-back in their center-of-mass frame. Therefore, a correction factor can be derived by comparing a central jet with its corresponding partner jet in an arbitrary η -region.
- *Level 3*: Absolute correction (C_{abs}), to account for a non-linear calorimeter response. The goal of this correction is to achieve a flat jet response in p_{T} . Correction factors can be derived from γ + jet-balance methods. The jet is balanced by a photon. The photon four-momentum vector can be measured very accurately in the ECAL. Using momentum conservation, the jet parameters can be estimated from the photon four-momentum and compared to the measured jet value.
- *Simulation calibration*: Residual Correction (C_{MC}), applied to data only, to correct for the difference in simulation and data as a function of p_{T} and η . It corrects the energy of the reconstructed jets such that it is equal on average to the energy of the simulated particle jets.

Thus, the corrected four-momentum p^{cor} is obtained by

$$p^{\text{cor}} = p^{\text{raw}} \cdot C_{\text{offset}}(p_{\text{T}}^{\text{raw}}) \cdot C_{\text{rel}}(\eta) \cdot C_{\text{abs}}(p'_{\text{T}}) \cdot C_{\text{MC}}(p''_{\text{T}}, \eta), \quad (5.4)$$

where p'_{T} is the transverse momentum of the jet after applying the offset correction and p''_{T} the transverse momentum of the jet after all previous corrections. JECs are of the order of 5 % (10 %) in the barrel (endcaps).

To ensure a high quality jet reconstruction, additional jet selection criteria are required.

Jet Selection

The jet reconstruction is performed using the particle flow algorithm. A hard transverse momentum requirement of $p_{\text{T}} > 40 \text{ GeV}/c$ is chosen due to large uncertainties for low jet energy values to obtain a reliable jet energy reconstruction. To account for the detector acceptance, jets are only considered in a range of $|\eta| < 2.5$. Furthermore, loose Jet ID criteria must be fulfilled, following the recommendation of the CMS collaboration based on detailed investigations [110, 111]. To be more explicit, these recommendations refer to:

- EMF: Electromagnetic fraction of jet energy contributed by ECAL energy deposits.
- n_{hits}^{90} : Minimum number of hits which contribute with 90 % to the jet energy.
- f_{HPD} : Maximum fraction of energy contributed by the highest energy HPD readout.

In Table 5.11 the values for the Jet ID requirements are given.

A cross cleaning with leptons is performed — if a lepton is found within a cone of $\Delta R < 0.4$ around the jet axis, the jet is removed. The Jet requirements are summarized in Table 5.12.

Table 5.11: Loose Jet ID requirements.

Observable	Value or Range
EMF	> 0.01
n_{hits}^{90}	≥ 2
f_{HPD}	< 0.98

Table 5.12: Jet Selection requirements.

Observable	Value or Range
p_{T}	$> 40 \text{ GeV}/c$
$ \eta $	< 2.5
loose Jet ID	true
$\Delta R(\mu)$	> 0.4
$\Delta R(e)$	> 0.4

5.4.4 Vertex Reconstruction

Primary vertices are reconstructed [61] from a set of tracks. These tracks are chosen based on the transverse impact parameter significance with respect to the beam line, number of strip and pixel hits, and the normalized track χ^2 . Vertex candidates are obtained by grouping the tracks, separated in their z -coordinate by $\Delta z < 1 \text{ cm}$ from their nearest neighbor. Subsequently, the vertices are fit with an adaptive vertex fit [112].

In this analysis, a “good” vertex fulfills three more requirements, listed in Table 5.13, where $|d_z|$ and d_ρ are the longitudinal and transverse distance with respect to the nominal interaction point and N_{dof} is the number of degrees of freedom of the reconstructed vertex.

Table 5.13: Vertex Selection requirements.

Observable	Value or Range
$ d_z $	$< 15(24) \text{ cm}$ in MC (data)
d_ρ	$< 2 \text{ cm}$
N_{dof}	> 4

5.4.5 Missing Transverse Energy

Missing transverse energy [113–115], denoted as $E_{\text{T}}^{\text{miss}}$, is a measure of the imbalance of an event in the transverse plane. The incoming protons do not have a (noteworthy) transverse momentum contribution. Hence, according to momentum conservation, the

sum of the momenta of all particles released in the collision is zero. E_T^{miss} is defined as

$$E_T^{\text{miss}} := \left| - \sum_i^n \vec{p}_{Ti} \right|, \quad (5.5)$$

where (in this analysis) the sum runs over all n PF jets, unclustered energy depositions, and leptons of an event. Sources of E_T^{miss} in the detector are only weakly interacting particles like neutrinos, which cannot be measured with the CMS detector, carrying away momentum. Furthermore, a limited acceptance of the detector, momentum mis-measurements, and defect calorimeter cells can contribute to E_T^{miss} .

JECs influence the reconstruction of E_T^{miss} , too. Hence, there are possible corrections for E_T^{miss} :

- *Type-1*: E_T^{miss} is adjusted for JECs (due to the non-linear response of the calorimeter) above given thresholds of the jets. These thresholds are: $p_T'' > 20 \text{ GeV}$ and $\text{EMF} < 0.9$, where p_T'' is defined as in the previous Section and EMF is the electromagnetic fraction obtained by dividing the jet energy measured in the ECAL by the jet energy in the HCAL. Jets not fulfilling these criteria are considered as uncorrected.
- *Type-2*: E_T^{miss} corrections with respect to unclustered energy in the calorimeter, or jets below the thresholds of the *Type-1* correction.

However, these corrections are not applied in this analysis, since the *Type-2* correction is controversial and a better performance of the physics results has been obtained without the *Type-1* correction.

5.4.6 H_T

H_T is defined as the scalar sum of the transverse momentum of all n reconstructed jets in an event, fulfilling a p_T requirement

$$H_T := \sum_i^n |\vec{p}_{Ti}|, \quad \forall |\vec{p}_{Ti}| > 40 \text{ GeV}/c. \quad (5.6)$$

It is a measure of the hadronic activity of an event and is often used in analyses to discriminate signal from background.

5.5 Preselection

A preselection is defined, since a standard event cleaning is applied for every investigation of this analysis (for the signal selection, as well as for the selection of control samples for the background estimation, or for the estimation of systematic uncertainties).

For the preselection, a trigger, appropriate for the decay channel, is applied (see Section 5.3). An event is also required to have at least one good vertex. Furthermore, two recommended cleaning requirements are applied to remove non-collision events and events

which can not be reconstructed correctly: A noise filter rejects events with significant noise in the HB and HE (for details, see [111]) and 25 % of all tracks in the events are required to be of high quality by a “pure tracks” requirement, removing beam scraping events. In Table 5.14 the preselection requirements are summarized.

Table 5.14: Preselection requirements applied to every selection used in this analysis. This preselection contains a standard event cleaning.

Observable	Value or Range
Trigger	true
N_{vertex}	≥ 1
Noise filter (HB, HE)	true
Pure Tracks	true

5.6 Event Selection

The aim of the event selection is to suppress the SM background as efficiently as possible, while selecting SUSY same-sign di-lepton events.

After passing the preselection, an event must contain at least two leptons. The lepton with the larger transverse momentum has to fulfill $p_T > 20 \text{ GeV}/c$, while the lepton with the second largest transverse momentum is required to have $p_T > 10 \text{ GeV}/c$ (this is already ensured by the lepton object definition). This p_T requirement is chosen with respect to the trigger thresholds (see Section 5.3). There is no veto imposed on additional leptons in the event. However, for the leptons, all possible combinations of lepton pairs are built, and a cleaning is performed, rejecting leptons originating from Z bosons or low mass resonances:

- Z-veto for opposite-sign, same-flavor isolated leptons: Remove leptons pairs with an invariant mass M_{inv} close to the Z -mass: $|M_{\text{inv}} - M_Z| \leq 15 \text{ GeV}/c^2$.
- Low mass resonance veto for same-sign isolated leptons: Remove leptons pairs with a small invariant mass: $M_{\text{inv}} < 5 \text{ GeV}/c^2$

After this cleaning, all di-lepton same-sign candidates are built from the remaining leptons. If multiple candidates are found, a priority is given in the order $\mu\mu$, $e\mu$, and ee . Hence, the three decay channels are considered exclusively. If there are several candidates for the same channel, the one with the highest transverse momentum is chosen. The event passes the selection, if such a di-lepton same-sign candidate is found.

Furthermore, an event is required to have at least two jets, which implies a requirement of $H_T > 80 \text{ GeV}/c$, since jets are reconstructed with $p_T > 40 \text{ GeV}/c$. Finally, events are required to contain missing transverse energy of above 100 GeV . For illustration and for the background estimation it is useful to apply the relative isolation requirement for the

leptons in the last step.⁹ Therefore, the isolation requirement for leptons is not applied in the object selection, but performed in the event selection as the final requirement, for the better isolated lepton of the same-sign candidate ($I_{\text{Rel}}(\text{1st iso lep}) < 0.15$) and eventually for the less isolated lepton ($I_{\text{Rel}}(\text{2nd iso lep}) < 0.15$). Table 5.15 summarizes the event selection.

Table 5.15: Event selection requirements, where N_l is the number of leptons in an event, $p_{\text{T}}(l)$ the lepton transverse momentum, $N_{\text{SS candidates}}$ the number of di-lepton same-sign candidates in an event, N_{Jets} the Jet multiplicity, and $I_{\text{Rel}}(\text{1st iso lep})$ and $I_{\text{Rel}}(\text{2nd iso lep})$ the relative isolation of the better and less isolated lepton of the di-lepton same-sign candidate.

Observable	Value or Range
Preselection	
N_l	≥ 2
$p_{\text{T}}(l)$	$\geq 20(10) \text{ GeV}/c$
$N_{\text{SS candidates}}$	$= 1$
N_{Jets}	≥ 2
$E_{\text{T}}^{\text{miss}}(PF)$	$> 100 \text{ GeV}$
$I_{\text{Rel}}(\text{1st iso lep})$	< 0.15
$I_{\text{Rel}}(\text{2nd iso lep})$	< 0.15

5.7 Event Yield

In Fig. 5.7 the event yield of the event selection flow for the different MC samples, a SUSY signal sample, and data is shown exemplarily for the $\mu\mu$ -channel (the other channels and the single numbers for the event selection flow can be found in the Appendix B).

Note, that here it is not intended to perform a validation of the simulation with the data, since this analysis is essentially data-driven (see also next Chapter). Differences in simulation and data underline the necessity of *data-driven* analyses.

However, some aspects can be seen from Fig. 5.7: The selection requirements of two same-sign leptons in an event, $N_{\text{Jets}} \geq 2$, and $E_{\text{T}}^{\text{miss}}(PF) > 100 \text{ GeV}$ strongly suppress SM processes but affect the SUSY signal sample relatively slightly. From special importance is the last selection step, $I_{\text{Rel}}(\text{2nd iso lep}) < 0.15$, the requirement that the less isolated lepton is also well isolated — this requirement directly refers to the different nature of processes: SM processes tend to have a worse isolated second lepton than SUSY processes (see next Chapter). Hence, the last requirement of the event selection has a strong separation power between SUSY signal and SM background.

⁹Note, that the main background originates from semi-leptonic $t\bar{t}$ -decays associated with a second lepton from a heavy flavor decay (see Chapter 6). Hence, the isolation requirement for the less isolated lepton has a large separation power between signal and background.

In Table 5.16 a summary of the event yield after the last selection step is given for the three channels and their combination. Although this result already indicates, that the event yield found in data agrees with the SM prediction from simulation and no SUSY signal is found, a data-driven background estimation is necessary to perform a reliable background prediction (see next Chapter). Furthermore, a careful investigation of systematic uncertainties is needed, documented in Chapter 7.

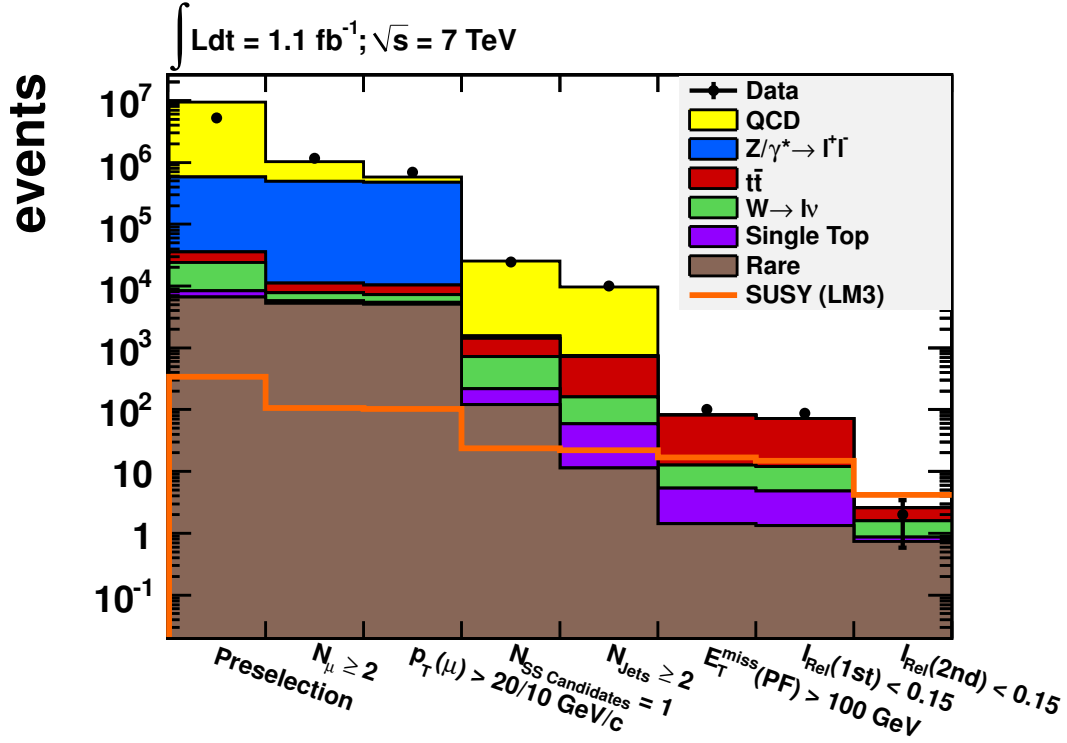


Figure 5.7: Event selection flow for the $\mu\mu$ -channel. Different MC samples are stacked in different colors, a SUSY signal sample is shown as a line, and the data yield is displayed with black points. Everything is normalized to the integrated data luminosity of $L = 1.1 \text{ fb}^{-1}$.

Table 5.16: Summary of the event yield after the last selection step for the three different decay channels. The event yield is shown for single simulated SM processes, some SUSY benchmark points, and 1.1 fb^{-1} of data. In the last column the channels are summed up. Numbers are summed up before they are rounded. The given uncertainty is statistical only.

sample	$\mu\mu$	$e\mu$	ee	sum
QCD	0	0	0	0
$Z/\gamma^* \rightarrow l^+l^-$	0	0.2 ± 0.2	0	0.2 ± 0.2
$W \rightarrow l\nu$	0.7 ± 0.7	0	0	0.7 ± 0.7
$t\bar{t}$	1.0 ± 0.3	2.2 ± 0.4	0.6 ± 0.2	3.8 ± 0.5
Single top	0.1 ± 0.0	0.2 ± 0.1	0.1 ± 0.0	0.4 ± 0.1
Rare	0.7 ± 0.1	1.5 ± 0.1	0.7 ± 0.1	2.9 ± 0.2
sum (MC)	2.6 ± 0.8	4.1 ± 0.5	1.4 ± 0.2	8.1 ± 1.0
Signal (LM1)	13.3 ± 1.7	20.6 ± 2.1	7.9 ± 1.3	41.8 ± 3.0
Signal (LM3)	4.2 ± 0.9	9.8 ± 1.6	5.9 ± 1.3	19.9 ± 2.3
Signal (LM8)	1.8 ± 0.6	5.5 ± 1.1	1.6 ± 0.6	8.9 ± 1.4
Data	2	3	2	7

6 Background

Backgrounds are processes which mimic the signature of the signal process (but have a different origin), such that they also appear in the signal region. Only by estimating the contribution of backgrounds to the signal region, it is possible to judge if there is an excess in the measurement, or if a new physics scenario can be excluded.

Simulation studies might be performed in order to test background estimation methods. However, it is a priori not clear if the simulation reflects reality. While the hard interaction of partons is well described, their hadronization is very difficult to model. Furthermore, the processes are merely simulated for the first leading orders. There is also the chance that the simulations' underlying model (see Chapter 2) is wrong and represents only an approximation to reality. These difficulties can partly be taken into account by systematic studies (see chapter 7). However, the simulation of a complex detector like CMS is non-trivial and cannot be performed to perfection. In particular, modeling key quantities of an analysis, like e.g. the isolation, E_T^{miss} , or the hadronic activity is in general not reliable. Hence, in principle, it is not sufficient to perform a background estimation based on simulation only. Consequently, so-called *data-driven* techniques have to be developed in order to yield a reliable result. These techniques rely on data only. Usually, the background estimation is performed in a control region, and not in the signal region. This prevents the background estimation to be influenced by the signal.

In this analysis, a Tight-to-Loose method is applied in order to perform the estimation of the main background, which originates from semi-leptonic $t\bar{t}$ -decays associated with a second lepton from a heavy flavor decay. The key quantity for the background prediction is the relative isolation of leptons. The estimation is performed in a region similar to the signal region (see Section 6.3). This is possible due to the particular estimation method: Every possible contribution (from signal and background) is considered and estimated at the same time in a consistent way.

In the following, the main background contributions to the SUSY same-sign di-lepton signal are named and discussed. Several background estimation methods are introduced thereafter. Subsequently, the Tight-to-Loose method is presented in detail. The Chapter is closed by describing the method to measure the electron charge flip rate and how it contributes to the signal region, as well as an estimation of the contribution from the irreducible SM background (explained in the next Section).

6.1 Background Contributions

In the context of the presented SUSY same-sign di-lepton search, background contributions originate from SM processes. In order to characterize these backgrounds, two items need to be introduced: *Prompt leptons* and *fake leptons*.

Prompt Leptons

Prompt leptons are leptons which originate from a hard process, e.g. from $W^\pm \rightarrow l^\pm \nu_l$, or $Z \rightarrow l^+ l^-$. In principle, they are isolated since no further particle activity is expected in their vicinity.

Fake Leptons

In this analysis, fake leptons comprise all non-prompt leptons, which also includes real leptons not stemming from hard processes. They can originate from

- heavy flavor decays,
- kaon/ pion decays (meson in-flight decays),
- electrons from unidentified photon conversions,
- mis-reconstructed hadrons (jets).

Fake leptons tend to be not isolated. In principle, they are surrounded by additional activity from e.g. the b -jet in case of a heavy flavor decay.

In this analysis, *two* isolated same-sign leptons are required. Thus, the potential background sources can be classified into three categories:

- *fake-fake*: QCD,
- *prompt-fake*: $t\bar{t}$ (semi-leptonic decay), W +Jets, Single-top+Jets,
- *prompt-prompt*: Drell-Yan, di-leptonic $t\bar{t}$ -decays, W^+W^- (charge flip) and irreducible, rare SM processes.

The key quantity to distinguish prompt leptons from fake leptons is the relative isolation. It is used for both separation of signal and background, and background estimation. Although requiring two *isolated* same-sign leptons strongly reduces the SM background, in rare cases, events from SM processes can pass the signal selection.

Fake-fake Background

The fake-fake background from QCD is expected to be negligible. Lepton isolation requirements, as well as H_T and E_T^{miss} selections, very strongly suppress the QCD background. Nonetheless, it is also the most challenging background to model and is estimated by the Tight-to-Loose method (see Section 6.3).

Prompt-fake Background

Sources of prompt-fake lepton pairs yield the largest background contribution in this analysis. It originates mainly from semi-leptonic $t\bar{t}$ -decays associated with a second lepton, originating from a heavy flavor decay as illustrated in Fig. 6.1. A top-quark almost always decays into a b -quark and a W boson (see [17]). The W boson can decay hadronically into a $q\bar{q}$ -pair or leptonically into a lepton and the corresponding neutrino. In the semi-leptonic $t\bar{t}$ -decay one W decays leptonically and the other hadronically, leading to the signature of two b -jets, two light-quark jets, one lepton, and missing transverse energy due to the

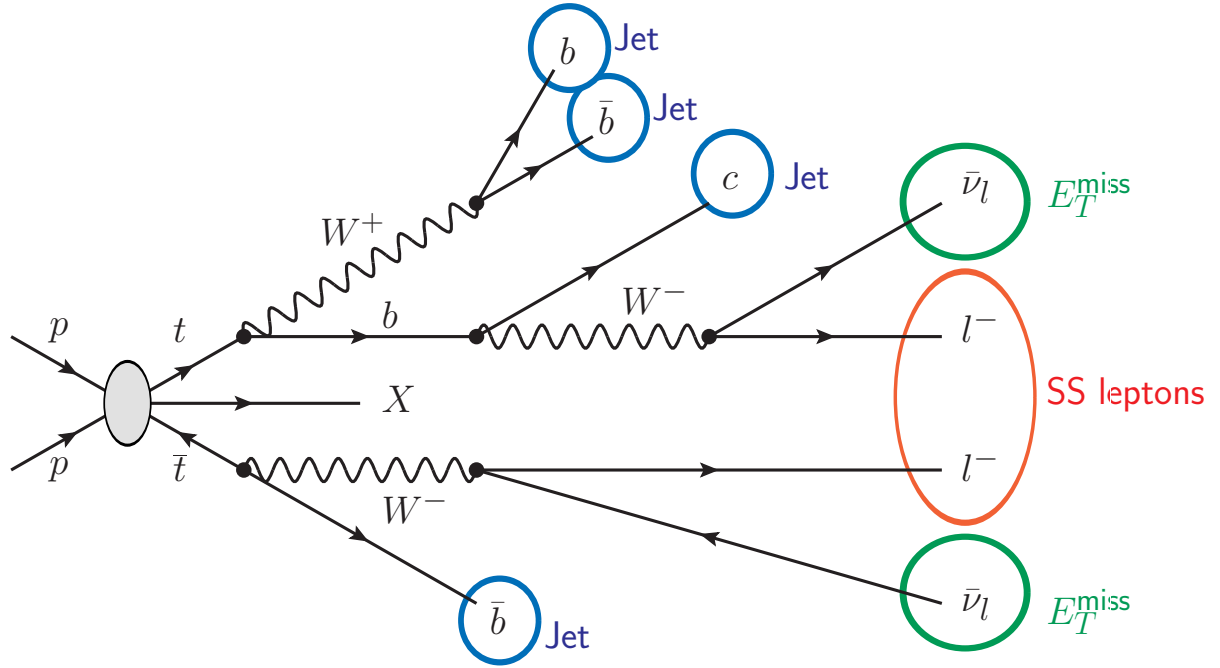


Figure 6.1: Sketch illustrating the background contribution from the semi-leptonic $t\bar{t}$ -decay. The b -quark of the branch with the hadronically-decaying W boson undergoes a heavy flavor decay, leading to a signature of a prompt-fake same-sign di-lepton pair. The fake lepton is in general not well isolated.

neutrino ($t\bar{t} \rightarrow b\bar{b}q\bar{q}l\bar{\nu}_l$). When the heavy flavor decay of the b -quark of the branch with the hadronically-decaying W yields an associated lepton, a *same-sign* di-lepton pair is produced.¹

The $t\bar{t}$ -decay is the main background for many SUSY searches, since it is associated with high jet activity and missing transverse energy, which is characteristic for many SUSY scenarios as well. Another potential prompt-fake background source is the decay of W bosons associated with jets: The prompt lepton originates from the W boson-decay ($W^\pm \rightarrow l^\pm \nu_l$), yielding also missing transverse energy, the fake lepton from a jet (either mis-reconstructed or from a heavy flavor decay). However, this process is stronger suppressed by the H_T and E_T^{miss} requirement of the signal selection than the $t\bar{t}$ -decay and thus plays a sub-dominant role. Similarly, single-top processes associated with jets can lead to prompt-fake same-sign lepton pairs, the prompt lepton originating from the hard process while the fake lepton stems from a jet. Due to the small cross-section of this process, it is also a minor source of background.

Prompt-prompt Background

Charge mis-reconstruction of leptons in di-leptonic $t\bar{t}$ -decays, Drell-Yan processes, or di-boson decays can lead to prompt-prompt same-sign lepton pairs. Although it is not very likely that such events pass the signal selection (including Z -veto, E_T^{miss} requirement, etc.), the rate of the charge mis-reconstruction is a priori not negligible and must be estimated.

¹The branching ratio of a b -quark decaying into a lepton is about 11% [17].

Furthermore, there are a number of irreducible, rare SM sources (having very small cross-sections) that can contribute to the same-sign signature via the production of prompt-prompt same-sign lepton pairs, e.g. WZ , ZZ , $2 \times W$ (double parton scattering), $t\bar{t}W$, $t\bar{t}Z$, $t\bar{t}\gamma$, $WW\gamma$, WWW , WWZ , WZZ , ZZZ , etc. Most of these processes have never been directly observed and their cross-section is known only from theory. Their contribution is estimated from simulation, since the yields do not depend on the difficult simulation of the production mechanisms for fake leptons in hadronic showers.² All processes listed in Table 5.3 (except for the WW -sample) contribute to this irreducible background.

6.2 Background Estimation Methods

In the following, data-driven background estimation methods for the SUSY same-sign di-lepton analysis are described. These methods utilize the lepton isolation, since the main background in this analysis originates from the prompt-fake source of semi-leptonic $t\bar{t}$ -decays with an additional lepton from a heavy flavor decay (see previous Section).

6.2.1 Extrapolation Method

Background contributions from fake leptons can be estimated from relative isolation distributions in background-dominated control regions by an extrapolation to the signal region [116]. The signal region, containing predominantly prompt leptons, is narrow (see Fig. 6.2, the small red area around $I_{\text{rel}} < 0.15$) and therefore, an extrapolation of the relative isolation distribution can be performed.

In Fig. 6.2 the extrapolation method is illustrated. It is assumed that the relative isolation distribution of non-prompt leptons from the background can be described by a function (here: Landau function³). Furthermore, it is assumed that in the region of large relative isolation values, the number of signal events are negligible. The idea is to fit the relative isolation distribution in the background region — the extrapolation of the fit-function into the signal region provides the background prediction. This procedure is performed for both, electrons and muons, separately.

The extrapolation method is applicable if the amount of data is sufficient to yield a smooth shape of the relative isolation distribution in order to perform a reliable fit. However, in the case of the same-sign di-leptonic analysis, the number of remaining events is too low to obtain an adequate relative isolation distribution which can be fitted. Therefore, other methods must be developed to provide enhanced statistical sensitivity for this analysis.

²An uncertainty of 50% is assigned to the yields obtained for this background contribution.

³Note, that there is no physical motivation to choose in particular a Landau-function. Any function which describes the shape of the isolation distribution reasonably well is in principle suitable.

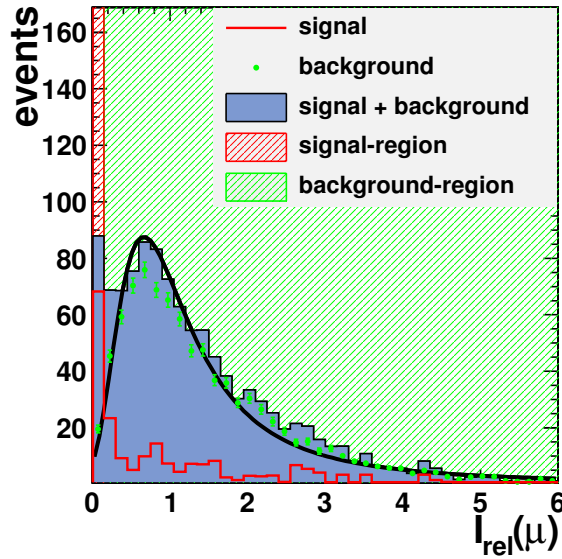


Figure 6.2: Illustration of the extrapolation method: Relative isolation distribution of muons from a simulated $t\bar{t}$ -sample. Red region: signal region ($I_{\text{rel-iso}} < 0.15$), green region: background region ($I_{\text{rel-iso}} > 0.15$). A Landau function is fitted to the background region. The prediction of the background contribution is the integral of the fit function in the signal region.

6.2.2 t-Tag-and-Probe Method

The principle of the t-Tag-and-Probe (t-T&P)⁴ method is to measure the number of *non-isolated* ($I_{\text{rel}} > 0.15$) leptons $N_{I_{\text{rel}} > 0.15}^{\text{signal}}$ in the SUSY same-sign di-lepton signal region to estimate the number of *isolated* leptons $N_{I_{\text{rel}} < 0.15}^{\text{signal}}$ (originating from a heavy flavor decay in a semi-leptonic $t\bar{t}$ -event) by

$$N_{I_{\text{rel}} < 0.15}^{\text{signal}} = p \cdot N_{I_{\text{rel}} > 0.15}^{\text{signal}}, \quad (6.1)$$

where p is the probability that a fake lepton is isolated ($I_{\text{rel}} < 0.15$). The estimation must be performed for electrons and muons separately.

The key challenge of the method is to obtain the probability p . In principle, it can be estimated from a $t\bar{t}$ -enriched sample. Since the process of interest for the background contribution are semi-leptonic $t\bar{t}$ -decays with an associated heavy flavor decay, *two* leptons per event must be required. Therefore, this sample also contains events of di-leptonic $t\bar{t}$ -decays. Simulation studies indicate that these events are effectively removed by requiring the two leptons to have the same charge. The less isolated lepton is assumed to originate from a heavy flavor decay. By counting the number of events, in which the less isolated lepton in the $t\bar{t}$ -enriched sample is isolated $N_{I_{\text{rel}} < 0.15}^{t\bar{t}}$ and non-isolated $N_{I_{\text{rel}} > 0.15}^{t\bar{t}}$, the probability p can be derived from

$$p = \frac{N_{I_{\text{rel}} < 0.15}^{t\bar{t}}}{N_{I_{\text{rel}} > 0.15}^{t\bar{t}}}. \quad (6.2)$$

⁴The naming “T&P” is adopted from the b-T&P method in which a T&P method is applied.

However, simulation studies indicate that at least 10fb^{-1} of data is needed for the t-T&P method to be applied. Therefore, another approach has been developed in order to be less limited by the amount of data: The b-T&P method opens the possibility to derive p from a $b\bar{b}$ -enriched data sample (which contains more events than a $t\bar{t}$ -enriched data sample) at the cost of relying on some information from simulation (see next Section).

6.2.3 b-Tag-and-Probe Method

The procedure of the b-Tag-and-Probe (b-T&P) method [88, 116] works in analogy to the t-T&P method, apart from the determination of p , the probability that a lepton from a heavy flavor decay in semi-leptonic $t\bar{t}$ -events is isolated. This alternative method is motivated by the fact that the t-T&P method requires a large amount of integrated luminosity, while for the b-T&P method, less data is sufficient. Here, p is estimated from a $b\bar{b}$ -enriched sample, which is obtained from data by applying appropriate selection requirements (details can be found in the references [88, 116]). The idea is to provide one very pure b-tagged jet (“Tag”) and a lepton (“Probe”) with a minimum distance in the η - ϕ -plane of $\Delta R > 1$ to the b-tagged jet. The lepton stems mainly from the $b\bar{b}$ -pair and hence, originates from a heavy flavor decay.

However, the kinematics of the $b\bar{b}$ -enriched sample is different from the kinematics of the main background, $t\bar{t}$. They differ in the lepton transverse momentum p_T -spectrum and the average number of Jets N_{Jets} of an event. Figure 6.3 shows an illustration of the muon p_T -spectrum and N_{Jets} for a simulated $b\bar{b}$ -enriched and a $t\bar{t}$ sample.⁵

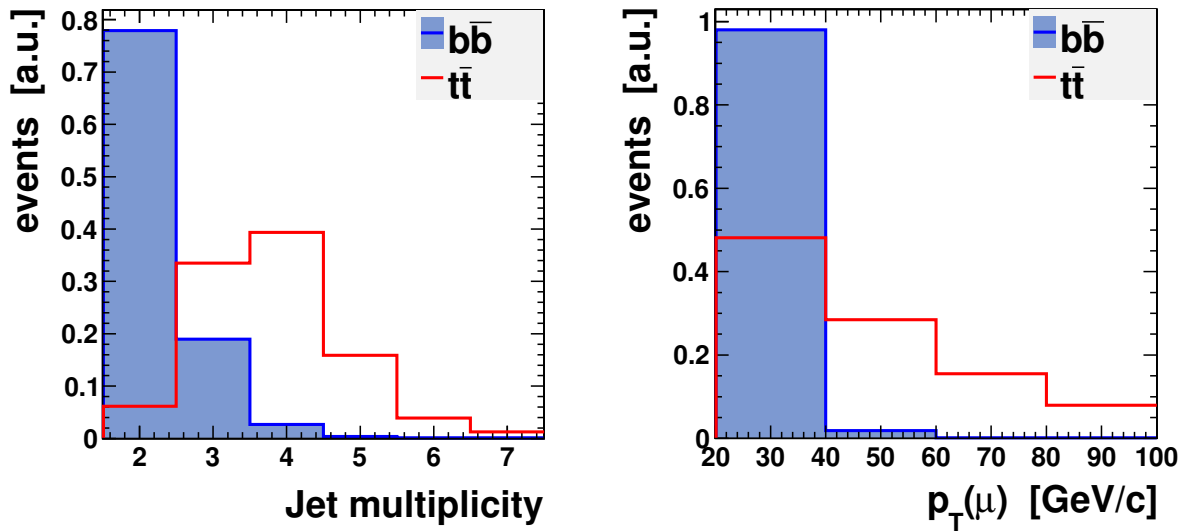


Figure 6.3: Illustration of the jet multiplicity (*left*) and lepton p_T (*right*) for muons in a simulated $t\bar{t}$ (*red line*) and QCD sample (*filled blue*). The distributions are normalized to unity. The distributions of the $t\bar{t}$ and QCD samples differ considerably.

⁵For the $b\bar{b}$ -enriched sample, the selection documented in [88, 116] has been applied while for the $t\bar{t}$ sample the signal selection was performed (see Chapter 5).

Due to the different kinematics of the samples, the relative lepton isolation is also different. Therefore, a re-weighting procedure must be performed. Figure 6.4 illustrates the workflow of the b-T&P method.

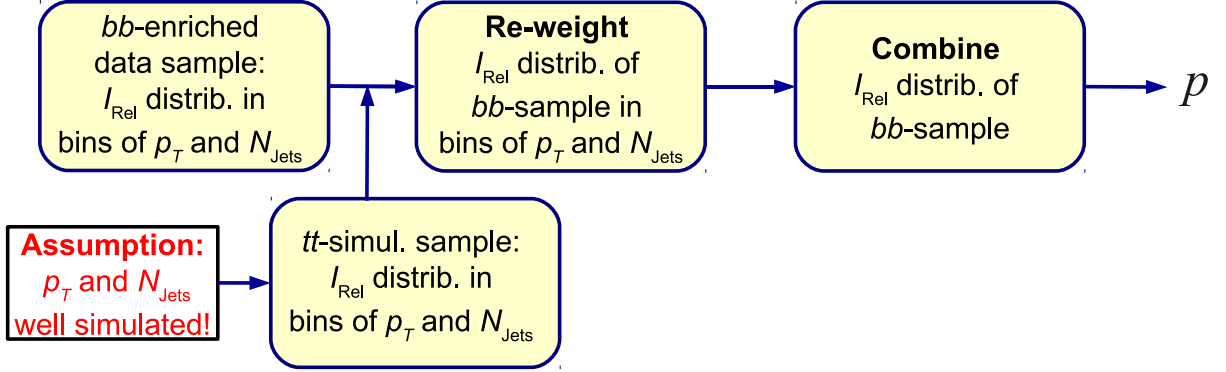


Figure 6.4: Sketch to illustrate the workflow of the b-T&P method to estimate p , the probability that a lepton of a heavy flavor decay is isolated. The different lepton p_T and N_{Jets} in the $b\bar{b}$ -enriched sample and a simulated $t\bar{t}$ -enriched sample is taken into account by a re-weighting procedure.

The relative isolation distribution of the leptons is split up into bins of lepton transverse momentum p_T , to account for the lepton kinematics, and into the number of jets N_{Jets} , to account for the hadronic activity (which also affects the relative isolation of leptons). This is done for both, the $b\bar{b}$ -enriched sample from data and a $t\bar{t}$ -enriched sample from simulation.⁶ As an important consistency check, the *shape* of the relative isolation distributions in the two samples must agree in the single kinematic bins — if they do, it is legitimate to re-weight the $b\bar{b}$ -enriched sample according to the $t\bar{t}$ -sample (see Fig. 6.5).

The split relative isolation distributions of the $b\bar{b}$ -enriched sample are normalized to unity and weighted by

$$w(p_T; N_{\text{Jets}}) = \frac{N^{t\bar{t}}(p_T; N_{\text{Jets}})}{N^{t\bar{t}}}, \quad (6.3)$$

where $N^{t\bar{t}} = \sum_{p_T; N_{\text{Jets}}} N^{t\bar{t}}(p_T; N_{\text{Jets}})$ is the number of events of the whole $t\bar{t}$ -enriched sample and $N^{t\bar{t}}(p_T; N_{\text{Jets}})$ the number of events in one kinematic bin of this sample. Note that at this point, it is assumed that the jet multiplicity, as well as the lepton p_T spectrum are well modeled in simulation, since it enters the re-weighting approach. Finally, the re-weighted distributions of the $b\bar{b}$ -enriched sample are combined into one distribution. From this distribution $N_{I_{\text{rel}} < 0.15}^{b\bar{b}}$ and $N_{I_{\text{rel}} > 0.15}^{b\bar{b}}$ are obtained and p can be estimated by

$$p = \frac{N_{I_{\text{rel}} < 0.15}^{b\bar{b}}}{N_{I_{\text{rel}} > 0.15}^{b\bar{b}}}. \quad (6.4)$$

Eventually, the estimation of the background contribution is performed by measuring $N_{I_{\text{rel}} > 0.15}^{\text{signal}}$ in the same-sign di-lepton signal region and multiplying this number by the probability p (see equation 6.1).

⁶Note that the integrated luminosity of the simulated $t\bar{t}$ -sample exceeds the integrated luminosity of the $t\bar{t}$ -enriched data sample.

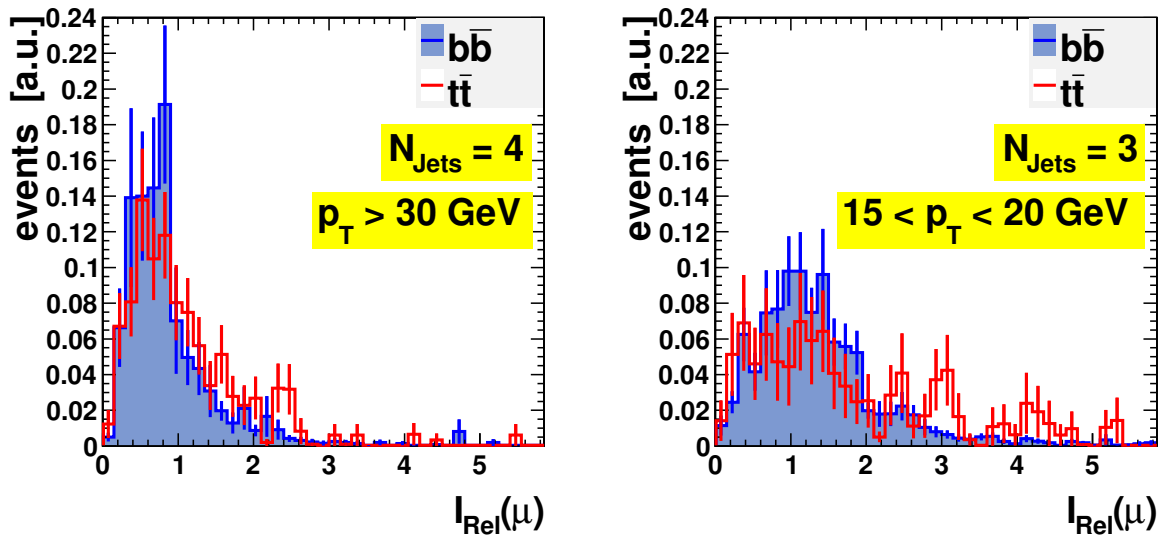


Figure 6.5: Illustration of the relative muon isolation in single kinematic bins. *Left:* $N_{\text{Jets}} = 4$; $p_T > 30$ GeV, *right:* $N_{\text{Jets}} = 3$; $15 < p_T < 20$ GeV. *Filled blue:* simulated $b\bar{b}$ -enriched sample, *red line:* simulated $t\bar{t}$ sample in the signal region. The distributions are normalized to unity. The samples agree within single kinematic bins. The shape of the distributions are different in different kinematic bins.

Although the b-T&P method has been shown to work [88, 116], it still has the drawback that it relies on some simulation information, namely the jet multiplicity and the lepton p_T spectrum of $t\bar{t}$ -decays. Therefore, it should be replaced by the t-T&P method as soon as sufficient data is available.

6.3 Tight-to-Loose Method

The purpose of the Tight-to-Loose (TL) method [116, 117] in this analysis is to estimate the prompt-fake and fake-fake SM background contribution to the same-sign di-leptonic signal region. However, this technique is more general. It can also be applied to final states with an arbitrary number of leptons, or used for different purposes like b -tagging studies or the identification of hadronically-decaying taus.

For leptons, a loose and a tight object selection are defined. The tight selection is identical to the object definition of the same-sign di-lepton signal selection (see Section 5.4). This is essential, because the probability of background contributions (fake lepton sources), leading to tight leptons — and therefore mimicking a signal-like signature — is estimated. It is measured, how likely it is, that loose leptons from prompt and fake sources also pass the tight selection. The loose selection differs from the tight selection for muons only by its relative isolation

- $I_{\text{rel}}(\mu) < 1$,

while for electrons also the identification criteria is relaxed

- $I_{rel}(e) < 1(0.6)$ in Barrel (Endcap),
- Electron identification (ID) with 90 % efficiency including conversion rejection (for details see Section 5.4.2).

The relative isolation criterion addresses to usually non-isolated fake leptons, originating from background processes like heavy flavor decays. The electron ID accounts for the shower profile: Jets having energy depositions also in the electromagnetic calorimeter can be matched to a track with a large transverse momentum and therefore, have a chance to be reconstructed as an electron. However, the shower profile is a criterion to distinguish such jets from electrons.

The probability that a prompt (fake) lepton is reconstructed as a tight lepton is given in terms of a “prompt-ratio” p (“fake-ratio” f), which is obtained from a Z -enriched control sample (QCD-enriched control sample). p and f depend on lepton kinematics (p_T and η): Fake leptons, originating from a heavy flavor decay, tend to have a small p_T . Hence, leptons with a high p_T mainly stem from prompt leptons and therefore, are more isolated and are likely to pass the tight selection. Furthermore, the activity in the forward region of the detector is larger than in the central region, which also affects the lepton isolation. Therefore, the prompt- and fake-ratio are also measured as a function of η .

Using the probabilities p and f obtained in control regions allows to estimate the SM background contribution in the signal region. Thereby, it is assumed that the probability is the same in both, the control and the signal region. In the signal region, the number of events containing tight-tight (N_{TT}), tight-loose (N_{TL}), and loose-loose (N_{LL}) lepton pairs are measured. Using these numbers together with the prompt-ratio, fake-ratio, and some algebra, the background prediction is performed. Figure 6.6 illustrates the workflow of the TL method.

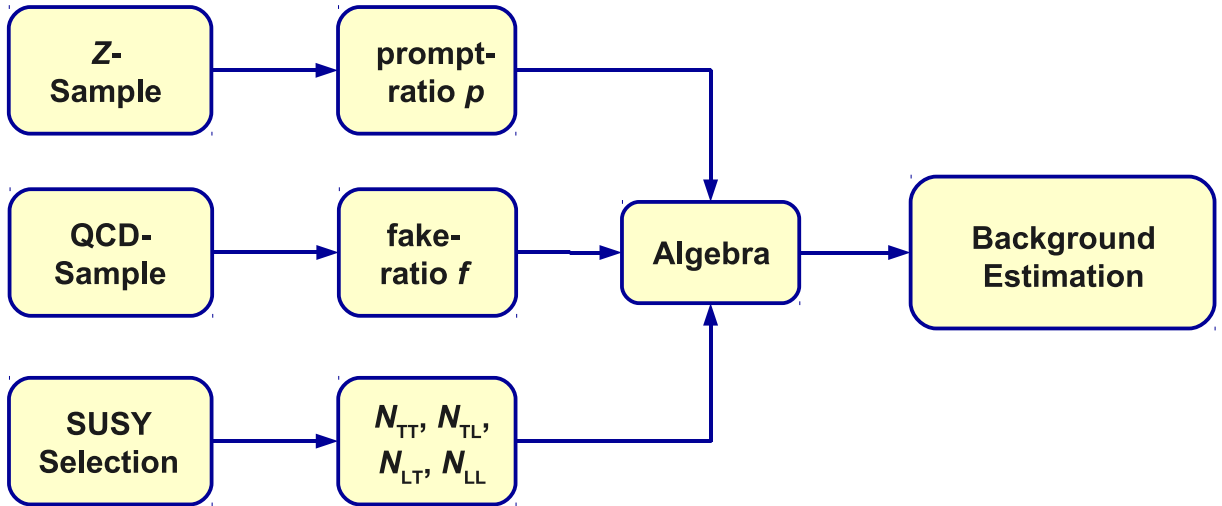


Figure 6.6: Sketch to illustrate the workflow of the TL method. The prompt-ratio and fake-ratio are obtained from a Z - and a QCD-enriched control sample, respectively. In the signal selection the number of events containing tight-tight, tight-loose, and loose-loose lepton pairs are measured. Combining these quantities and using some algebra results in the background prediction.

6.3.1 Fake-ratio Estimation

The fake-ratio f is defined as the probability that a loose *fake* lepton also passes the tight criteria. This ratio is measured in a QCD-enriched control sample. It is assumed that all leptons of this control sample are fake leptons. Thus, f is expected to be close to zero, because it is not likely that a fake lepton is isolated. In Table 6.1 the event selection, applied on data in order to obtain the QCD-enriched control sample, is listed.

Table 6.1: Event selection requirements in order to select a QCD-enriched control sample from which the fake-ratio is estimated.

Observable	Value or Range
Preselection (see Section 5.5) using lepton triggers	
$N_l(\text{loose})$	$= 1$
$M_T(\text{lepton}, E_T^{\text{miss}})$	$< 20 \text{ GeV}/c^2$
E_T^{miss}	$< 20 \text{ GeV}$
N_{Jets}	≥ 2

The requirement to have exactly one lepton reduces the contribution from Z -decays. The M_T and E_T^{miss} selection reduces the events containing W bosons and imposing at least two jets ensures, that the hadronic activity in the control sample is similar to the hadronic activity in the signal region. In Fig. 6.7 the event selection flow of the QCD-enriched control sample for muons and electrons is shown. Simulation indicates that the dominant contribution of the control sample originates from QCD processes. A SUSY signal sample is plotted here as well, indicating that it does not contribute to the estimation of the fake-ratio.

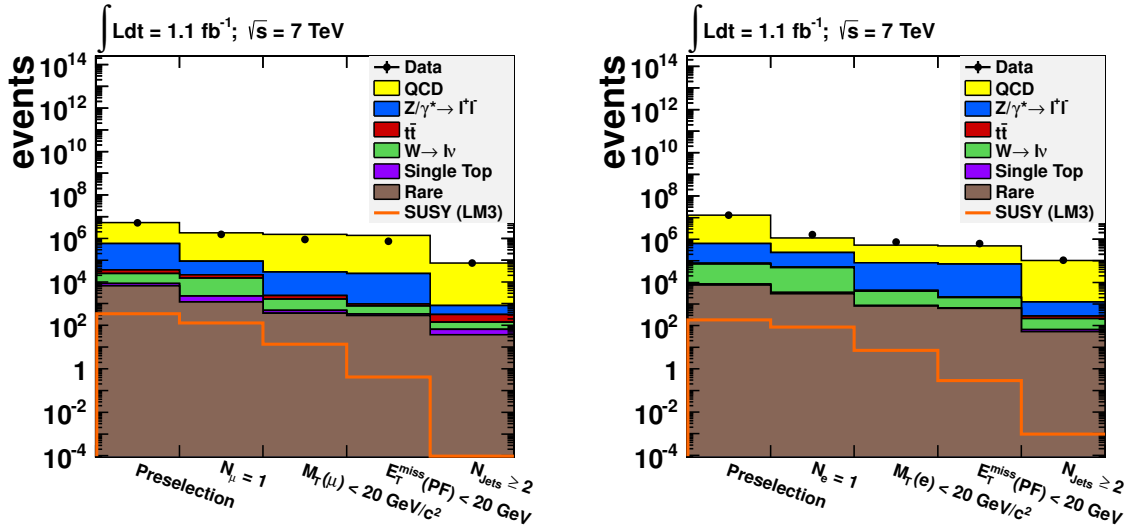


Figure 6.7: Event selection flow of the QCD-enriched control sample for muons (*left*) and electrons (*right*) from which the fake-ratio is estimated.

The fake-ratio f is defined as

$$f = \frac{N_T}{N_T + N_L}, \quad (6.5)$$

where N_L is the number of events, which contain exactly one loose, and not tight, lepton and N_T is the number of events in which the lepton also passes the tight selection. Note that by definition, this ratio is always between 0 and 1, since the numerator is a subset of the denominator.

As the fake-ratio also depends on the lepton flavor as well as on the lepton kinematics, it is measured for electrons and muons separately in bins of lepton p_T and η (see Fig. 6.8). The upper value of the last p_T bin is chosen to be 60 GeV/c. This value is motivated from simulation studies, which indicate that the contamination of the QCD-enriched control sample with prompt leptons is not negligible for higher p_T values (see Section 7.7.2). This choice is in agreement with other studies [97, 118]. In the case, where the ratios are applied to leptons with $p_T > 60$ GeV/c, the ratio of the last bin is utilized.

In Table 6.2 the mean of the fake-ratio is given.

Table 6.2: Mean of the fake-ratio for muons and electrons. The given uncertainty is statistical only.

Flavor	Mean fake-ratio [%]
μ	7.6 ± 0.1
e	7.3 ± 0.1

6.3.2 Prompt-ratio Estimation

The prompt-ratio p is defined as the probability that a loose *prompt* lepton also passes the tight criteria. The principle to measure p is performed in close analogy to the measurement of the fake-ratio. However, the prompt-ratio is derived from a Z -enriched control sample, which is obtained by using a Tag-and-Probe (T&P) method. This implies that there are *two* leptons per event, making the procedure more complicated, but the idea is still straight forward. It is assumed that all leptons of the Z -enriched control sample are prompt leptons. Thus, p is expected to be close to one, because it is likely that a prompt lepton is isolated. Note that $p = 1$ is used in other analyses [97, 119] as an approximation. Nonetheless, here the TL method is applied in detail and this approximation is not used.

In Table 6.3 the event selection, applied on data in order to obtain the Z -enriched control sample, is listed. The three cuts after the preselection are imposed to enrich the sample with Z events and to reject other SM contributions. The M_T requirement reduces events containing W bosons and imposing at least two jets ensures, that the hadronic activity of the events is similar to the hadronic activity in the signal region.

Note that requiring at least one tight lepton is part of the T&P method. Having a tight lepton (“Tag”) in the event considerably increases the probability that a Z -decay took place. Thus, the loose lepton (“Probe”) is very likely to originate from the Z , too. In general, this allows to investigate properties of the Probe lepton, which would not be

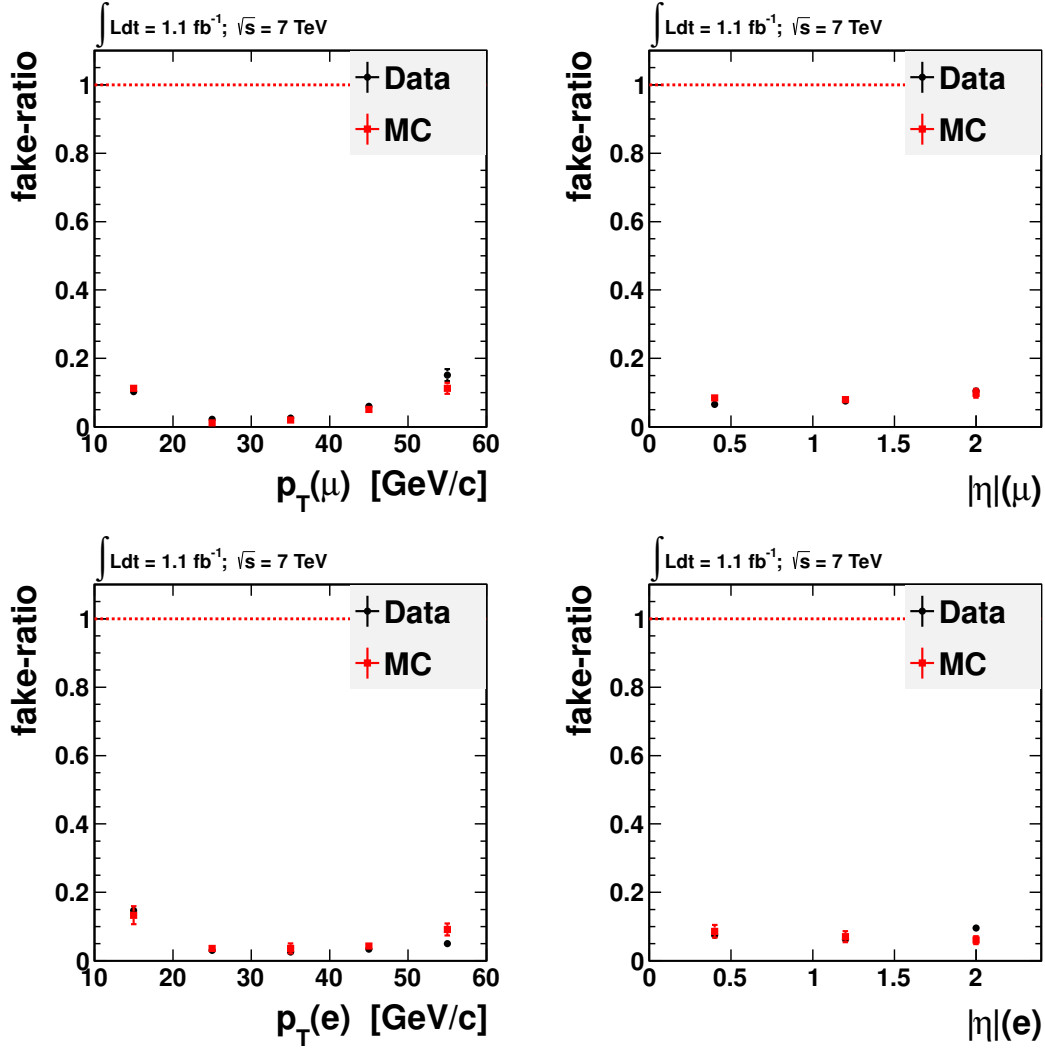


Figure 6.8: Fake-ratio for muons (*top*) and electrons (*bottom*) as a function of lepton p_T (*left*) and η (*right*). *Black circles*: ratios estimated from data, *red squares*: ratios estimated from simulation. *Red dashed line*: marks the fake-ratio value of one.

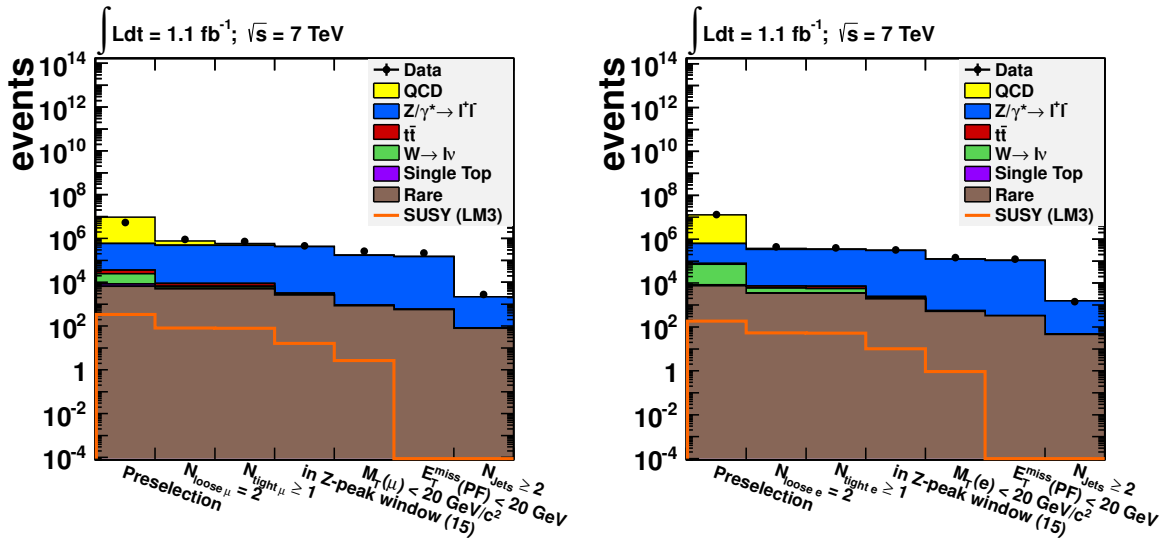
accessible if its requirements were as tight as for the Tag lepton. Here, the purpose of imposing at least one tight lepton is to obtain a pure Z -enriched sample and to measure the prompt-ratio from the Probe lepton.

In Fig. 6.9 the event selection flow of the Z -enriched control sample for muons and electrons is shown. Simulation indicates that the dominant contribution of the control sample originates from Drell-Yan processes. A SUSY signal sample is plotted here as well, indicating that it does not contribute to the estimation of the prompt-ratio.

In analogy to the fake-ratio, the prompt-ratio p is measured by counting how often the loose (Probe) lepton passes or fails the tight selection. In events with both leptons being tight, it is not obvious which lepton to identify as Tag and which as Probe. The solution is simple: both possibilities are used. To clarify this point, some notation is needed. In the following definitions, the first index always refers to the lepton with the higher transverse

Table 6.3: Event selection requirements in order to select a Z -enriched control sample from which the prompt-ratio is estimated.

Observable	Value or Range
Preselection (see Section 5.5) using lepton triggers	
$N_l(\text{loose, same flavor, opposite charge})$	$= 2$
$N_l(\text{tight})$	≥ 1
$ M_{\text{inv}} - M_Z $	$\leq 15 \text{ GeV}/c^2$
$M_T(\text{lepton}, E_T^{\text{miss}})$	$< 20 \text{ GeV}/c^2$
E_T^{miss}	$< 20 \text{ GeV}$
N_{Jets}	≥ 2

**Figure 6.9:** Event selection flow of the Z -enriched control sample for muons (*left*) and electrons (*right*) from which the prompt-ratio is estimated.

momentum:⁷

N_{TT} := number of events with both leptons being tight

N_{TL} := number of events with only the first lepton being tight

N_{LT} := number of events with only the second lepton being tight

The prompt-ratio p can be measured by

$$p = \frac{2 \cdot N_{TT}}{2 \cdot N_{TT} + N_{LT} + N_{TL}}, \quad (6.6)$$

where the factor of 2 originates from the fact that in events with two tight leptons both possibilities must be considered: The higher- p_T lepton is the Tag with the lower- p_T lepton

⁷This is an arbitrary definition. The purpose is to sort the leptons according to a well-defined rule. Sorting by the larger angle φ e.g. leads to the same result.

being the Probe and vice versa. Note that by definition, this ratio is always between 0 and 1, since the numerator is a subset of the denominator.

As the prompt-ratio also depends on the lepton flavor as well as on the lepton kinematics, it is measured for electrons and muons separately in bins of lepton p_T and η (see Fig. 6.10). The binning of the prompt-ratio is the same as the binning of the fake-ratio. The mean of the prompt-ratio is given in Table 6.2.

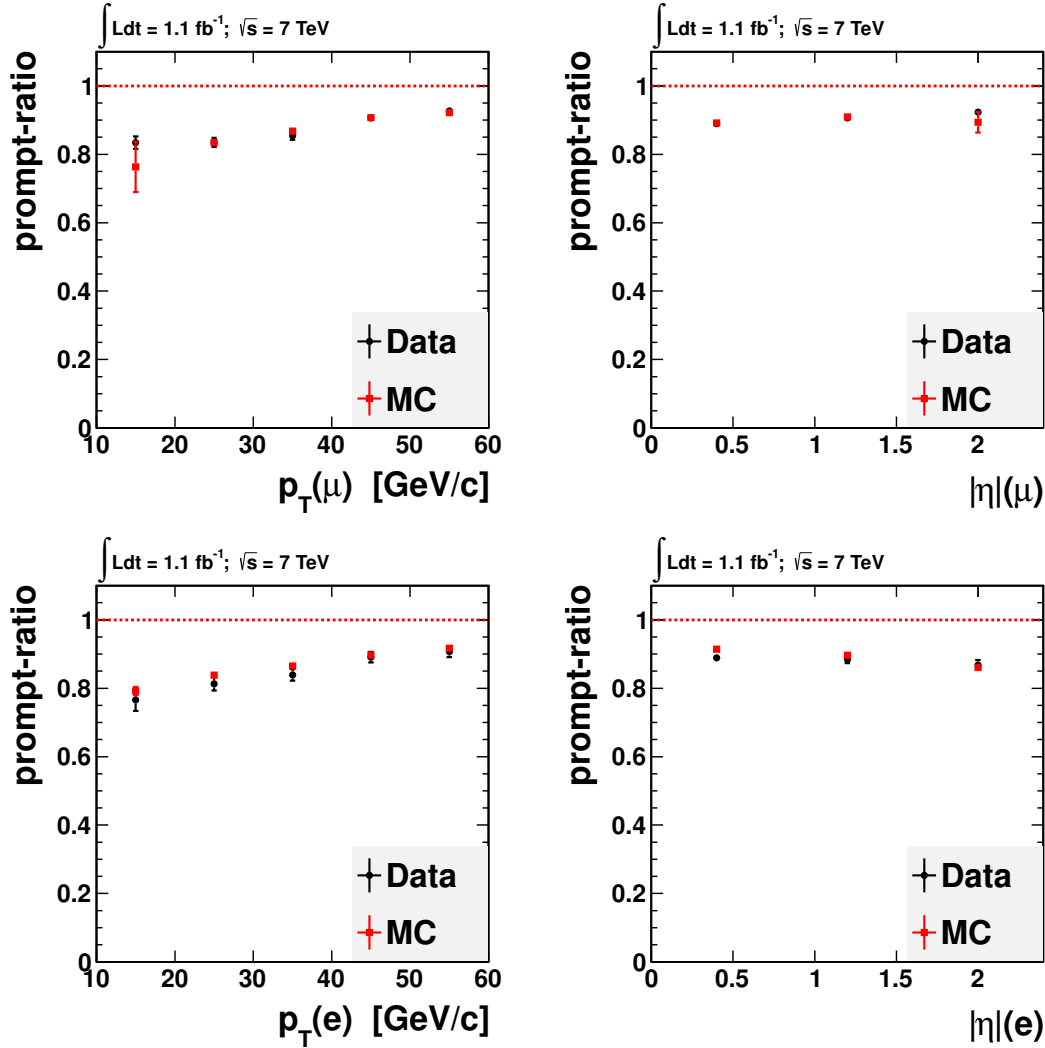


Figure 6.10: Prompt-ratio for muons (*top*) and electrons (*bottom*) as a function of lepton p_T (*left*) and η (*right*). *Black circles:* ratios estimated from data, *red squares:* ratios estimated from simulation. *Red dashed line:* marks the prompt-ratio value of one.

6.3.3 Background Prediction

Having estimated the prompt- and fake-ratio, some algebra must be introduced in order to obtain a background prediction from the TL method. As explained in Section 6.1 the background originates from prompt-fake ($t\bar{t}$ and W +Jets) and fake-fake (QCD) lepton

Table 6.4: Mean of the prompt-ratio for muons and electrons. The given uncertainty is statistical only.

Flavor	Mean prompt-ratio [%]
μ	87.2 ± 0.5
e	85.1 ± 0.8

pairs. It is not *measurable* whether a di-lepton pair is prompt-prompt, prompt-fake, or fake-fake. However, they can be related to the observable numbers of tight-tight, tight-loose and loose-loose lepton pairs in the signal region.

To elaborate the background prediction, the definition of some additional quantities is needed:

- N := number of events containing inclusively two loose leptons (possibly tight)
- N_{LL} := number of events containing exclusively two loose leptons (no tight leptons)
- N_{pp} := number of events containing two prompt leptons
- N_{pf} := number of events with only the first lepton being prompt
- N_{fp} := number of events with only the second lepton being prompt
- N_{ff} := number of events with both leptons being fake

It is obvious that the number of events N containing two loose leptons (which can also be tight), is the sum of the events in which the leptons are exclusively tight-tight, tight-loose, loose-tight, and loose-loose:

$$N = N_{TT} + N_{TL} + N_{LT} + N_{LL} = N_{pp} + N_{pf} + N_{fp} + N_{ff}. \quad (6.7)$$

These events consist of prompt-prompt, prompt-fake, fake-prompt, and fake-fake lepton pairs. The numbers N_{TT} , N_{TL} , N_{LT} , and N_{LL} can be directly measured, N_{pp} , N_{pf} , N_{fp} , and N_{ff} not. It is possible to relate these quantities. In the following, this is discussed in detail for N_{TT} . The other quantities are obtained in analogy.

Note that the two leptons can be distinguished in terms of flavor (in the case of the $e\mu$ -channel) or by their kinematic bins (p_T and η). Thus, a different prompt-/ fake- ratio must be used for each of the leptons. In accordance to the previous Section, the first index always refers to the lepton with the higher transverse momentum.

The number of events with two tight leptons consist of processes with prompt-prompt (signal), prompt-fake ($t\bar{t}$, W +Jets), and fake-fake leptons (QCD):

$$N_{TT} = p_1 p_2 N_{pp} + p_1 f_2 N_{pf} + f_1 p_2 N_{fp} + f_1 f_2 N_{ff}, \quad (6.8)$$

where the indices of p and f refer to the first or second lepton, respectively. The idea of equation (6.8) is straight forward: the probability that a prompt-prompt lepton pair is measured as a tight-tight lepton pair, is $p_1 \cdot p_2$. Similarly, the probability that a fake-fake lepton pair is measured as a tight-tight lepton pair, is $f_1 \cdot f_2$. The same principle is applied to the mixed terms.

Continuing with this procedure also for the other quantities, a system of equations is obtained. It is convenient to use matrix notation:

$$\begin{pmatrix} N_{TT} \\ N_{TL} \\ N_{LT} \\ N_{LL} \end{pmatrix} = \mathbf{M} \cdot \begin{pmatrix} N_{pp} \\ N_{pf} \\ N_{fp} \\ N_{ff} \end{pmatrix}, \quad (6.9)$$

with

$$\mathbf{M} := \begin{pmatrix} p_1 p_2 & p_1 f_2 & f_1 p_2 & f_1 f_2 \\ p_1(1-p_2) & p_1(1-f_2) & f_1(1-p_2) & f_1(1-f_2) \\ (1-p_1)p_2 & (1-p_1)f_2 & (1-f_1)p_2 & (1-f_1)f_2 \\ (1-p_1)(1-p_2) & (1-p_1)(1-f_2) & (1-f_1)(1-p_2) & (1-f_1)(1-f_2) \end{pmatrix}. \quad (6.10)$$

This system of equation consists of four equations and contains four unknowns (N_{pp} , N_{pf} , N_{fp} , and N_{ff}). Thus, it is solvable. Terms like $(1-p)$ or $(1-f)$ reflect the probability of a lepton *not* to pass the tight selection, e.g. the probability that a prompt-prompt lepton pair is measured as a loose-loose lepton pair is $(1-p_1) \cdot (1-p_2)$. Inverting the matrix $\mathbf{M}$ yields the solution of the system of equations

$$\begin{pmatrix} N_{pp} \\ N_{pf} \\ N_{fp} \\ N_{ff} \end{pmatrix} = \frac{1}{(p_1 - f_1)} \frac{1}{(p_2 - f_2)} \cdot \mathbf{A} \cdot \begin{pmatrix} N_{TT} \\ N_{TL} \\ N_{LT} \\ N_{LL} \end{pmatrix}, \quad (6.11)$$

with

$$\mathbf{A} := \begin{pmatrix} (1-f_1)(1-f_2) & -(1-f_1)f_2 & -f_1(1-f_2) & f_1 f_2 \\ -(1-f_1)(1-p_2) & (1-f_1)p_2 & f_1(1-p_2) & -f_1 p_2 \\ -(1-p_1)(1-f_2) & (1-p_1)f_2 & p_1(1-f_2) & -p_1 f_2 \\ (1-p_1)(1-p_2) & -(1-p_1)p_2 & -p_1(1-p_2) & p_1 p_2 \end{pmatrix}. \quad (6.12)$$

Equation (6.11) gives the recipe of how to estimate the number of prompt-prompt, prompt-fake, fake-prompt, and fake-fake lepton pairs from measured quantities of the loose selection. To predict how many of these pairs are reconstructed as tight-tight pairs (i.e. signal-like, which contribute to the signal region), these numbers must simply be multiplied with the prompt- and fake-ratios:

$$\begin{aligned} N_{fake-fake} &:= f_1 f_2 N_{ff} && \text{(fake-fake background),} \\ N_{prompt-fake} &:= (p_1 f_2 N_{pf} + f_1 p_2 N_{fp}) && \text{(prompt-fake background).} \end{aligned} \quad (6.13)$$

To account for the binning of the TL method, it is summed over all “ i ” p_T -bins and “ j ” η -bins of the two leptons:

$$N_{fake-fake} = \sum_{i_1, j_1, i_2, j_2} [f(i_1, j_1) f(i_2, j_2) N_{ff}(i_1, j_1, i_2, j_2)] \quad (6.14)$$

$$N_{prompt-fake} = \sum_{i_1, j_1, i_2, j_2} [p(i_1, j_1) f(i_2, j_2) N_{pf}(i_1, j_1, i_2, j_2) + f(i_1, j_1) p(i_2, j_2) N_{fp}(i_1, j_1, i_2, j_2)]. \quad (6.15)$$

To conclude: The prompt- and fake-ratio are obtained in a Z - and QCD-enriched control region, respectively. The observable quantities N_{TT} , N_{TL} , N_{LT} , and N_{LL} are measured in the signal region. Using these numbers together with the fake- and prompt-ratio, the TL method can be established developing some algebra and the background contribution in the signal region can be estimated. The result of the background prediction for 1.1 fb^{-1} of collected data is listed in Table 6.5.

Table 6.5: Background prediction of the TL method for the *prompt-fake* and *fake-fake* contribution to the signal region for the three channels and their combination. The given uncertainty is statistical only.

Contribution	$\mu\mu$	$e\mu$	ee	Sum
$N_{\text{prompt-fake}}$	2.91 ± 0.56	4.03 ± 0.73	0.65 ± 0.21	7.58 ± 0.94
$N_{\text{fake-fake}}$	-0.014 ± 0.009	-0.019 ± 0.014	-0.002 ± 0.006	-0.035 ± 0.018

The prediction of $N_{\text{prompt-fake}}$ agrees well with previous investigations (compare with e.g. Table 5.16). The prediction of $N_{\text{fake-fake}}$ shows *negative* numbers. However, they are very small and consistent with zero within the uncertainty (an additional systematic uncertainty of $\pm 50\%$ must be applied to the prediction, see Section 7.7.2). Note, that the analytical formulas of the TL method allow for negative results (similar results have also been observed in other analyses, e.g. in [118]). The fact, that the contribution of fake-fake processes is consistent with zero agrees with the considerations in Section 6.1: Fake-fake di-lepton events — essentially from QCD — are very unlikely to be found in the signal region. Furthermore, this also agrees with the result of the background contribution from simulation (compare Table 5.16, where no QCD contribution is found in the signal region). Hence, the background contribution due to fake-fake processes is found to be negligible.

6.4 Charge Flip

As described in Section 6.1, the charge mis-measurement (“charge flip”) of a lepton can lead to a same-sign di-lepton signature in the detector, which ultimately propagates to the signal region. Therefore, the charge flip is also a potential source of background.

The requirement for electrons to agree in all three charge measurements (see Section 5.4.2), reduces the rate of charge mis-measurement considerably. However, due to hard bremsstrahlung a charge flip might occur. Previous studies [88] have shown that the electron charge flip rate contributes less than 10 % to the total background. Although this background is small, it is not necessarily well-described in simulation. Thus, a data-driven method to estimate the background due to the electron charge flip rate is needed. Note, that the muon charge flip rate is not considered, since it is so small that it can be neglected.

The background prediction of the electron charge flip is performed in three steps:

- Obtaining the electron charge flip rate from a Z -enriched control sample.

- Counting the number of *opposite-sign* di-lepton pairs in the signal region.
- Calculating the background contribution from the two previous steps.

In order to obtain a Z -enriched control sample, the same selection is applied as in Section 6.3.2, except that

- both electrons must be tight,
- there is no charge requirement for the two leptons,
- no cut is set on the number of jets.

The charge flip rate r is obtained from the measured number of same-sign di-lepton pairs N_{SS} and the number of opposite-sign di-lepton pairs N_{OS} in the control region. For that purpose, some notation is needed:

$$\begin{aligned}
 N &:= N_{SS} + N_{OS} = \text{number of all events} \\
 N_{00} &:= \text{number of events with no charge flip} \\
 N_{\times 0} &:= \text{number of events with one charge flip} \\
 N_{\times \times} &:= \text{number of events with two charge flips}
 \end{aligned}$$

It is obvious that the total number of events N is the sum of the events with no, one, and two charge flips:

$$N = \underbrace{N_{00} + N_{\times \times}}_{=N_{OS}} + \underbrace{N_{\times 0}}_{=N_{SS}}. \quad (6.16)$$

The number of events with no and two charge flips can be estimated using the charge flip rate r :

$$\begin{aligned}
 N_{00} &:= (1 - r)^2 \cdot N, \\
 N_{\times \times} &:= r^2 \cdot N.
 \end{aligned} \quad (6.17)$$

Combining equation (6.16) and (6.17) yields a quadratic equation

$$N = (1 - r)^2 \cdot N + r^2 \cdot N + N_{SS}, \quad (6.18)$$

with the solution

$$r = \frac{1}{2} - \sqrt{\frac{1}{4} - \frac{N_{SS}}{2 \cdot (N_{SS} + N_{OS})}}. \quad (6.19)$$

Here, the solution with the positive sign has been discarded, since it is not physically motivated (the charge flip rate is small and cannot be larger than 50%). Thus, the estimation of the electron charge flip rate r can be achieved by counting N_{SS} and N_{OS} in the control sample. This has been done using 1.1 fb^{-1} of data, summarized in Table 6.6.

For the background prediction N_{OS} must be obtained in the signal region. For that purpose, the complete signal selection is applied, except for requiring two *opposite-sign* leptons in the final state instead of two same-sign leptons. Note, that for the counting procedure, also the Z -veto must be removed. This is important, because in the signal region, only *opposite-sign* di-leptons are required to be rejected if their invariant mass is

Table 6.6: Estimation of the electron charge flip rate r from a Z -enriched data control sample by counting the number of opposite-sign di-electrons N_{OS} and the number of same-sign di-leptons N_{SS} in the Z -mass window. The given uncertainty is statistical only.

N_{OS}	N_{SS}	Charge flip rate [%]
$94,500 \pm 307$	167 ± 13	0.088 ± 0.007

in the Z -mass window of ± 15 GeV. However, this does not affect *same-sign* di-leptons with a charge flip. Hence, N_{OS} must also be counted in the Z -peak region.

The counting procedure is performed for the ee - and $e\mu$ -channel, since the electron charge flip rate does not contribute in the $\mu\mu$ -channel. A relation is needed in order to estimate N_{SS} from N_{OS} and the charge flip rate r for the two channels ee and $e\mu$. The true number of opposite-sign events N_{OS}^{true} contributes to the measured number of same-sign N_{SS} and opposite-sign N_{OS} lepton pairs. They are related by the charge flip rate. In analogy to equation (6.16) for the ee -channel it is

$$\begin{aligned}
 N_{OS}^{\text{true}} &= \overbrace{N_{00} + N_{\times\times}}^{=N_{OS}^{ee}} + \overbrace{N_{\times 0}}^{=N_{SS}^{ee}} \\
 N_{OS}^{ee} &= [(1-r)^2 + r^2] \cdot N_{OS}^{\text{true}} \\
 N_{SS}^{ee} &= [1 - [(1-r)^2 + r^2]] \cdot N_{OS}^{\text{true}} = 2r(1-r)N_{OS}^{\text{true}}
 \end{aligned} \tag{6.20}$$

and thus

$$\Rightarrow N_{SS}^{ee} = \frac{2r(1-r)}{(1-r)^2 + r^2} \cdot N_{OS}^{ee}. \tag{6.21}$$

For the $e\mu$ -channel the similar calculation yields

$$\begin{aligned}
 N_{OS}^{\text{true}} &= \overbrace{N_{00}}^{=N_{OS}^{e\mu}} + \overbrace{N_{\times 0}}^{=N_{SS}^{e\mu}} \\
 N_{OS}^{e\mu} &= (1-r) \cdot N_{OS}^{\text{true}} \\
 N_{SS}^{e\mu} &= r \cdot N_{OS}^{\text{true}}
 \end{aligned} \tag{6.22}$$

with the solution

$$\Rightarrow N_{SS}^{e\mu} = \frac{r}{(1-r)} \cdot N_{OS}^{e\mu}. \tag{6.23}$$

Therefore, by counting N_{OS} in the signal region, the contribution of N_{SS} due to electron charge flip can be estimated. This has been done using 1.1 fb^{-1} of data, yielding:

$$N_{OS}(e\mu) = 314 \pm 18; \quad N_{OS}(ee) = 141 \pm 12. \tag{6.24}$$

Using these numbers together with equations (6.23) and (6.21) the estimation of the background contribution (N_{SS}) of an electron charge flip is performed. The result is listed in Table 6.7.

Table 6.7: Background prediction due to an electron charge flip in the signal region for all channels. The number of opposite-sign di-leptons N_{OS} is measured in data and the prediction is performed using equations (6.23) and (6.21). Numbers are summed up before they are rounded. The given uncertainty is statistical only.

Contribution	$\mu\mu$	$e\mu$	ee	Sum
Charge flip	–	0.28 ± 0.03	0.25 ± 0.03	0.52 ± 0.05

6.5 Irreducible Background

As already explained in Section 6.1, there are rare SM processes which contribute to the signal region (all processes listed in Table 5.3 except for the WW -sample). Since they have a *prompt-prompt* same-sign di-lepton signature, they are irreducible and have to be estimated from simulation. In Table 6.8 the event yield after the whole event selection is listed for the three channels and their combination.

Table 6.8: Summary of the event yield of simulated irreducible SM background processes for the three decay channels and their combination. The given uncertainty is statistical only. Numbers are summed up before they are rounded.

Sample	$\mu\mu$	$e\mu$	ee	Sum
Irreducible	0.7 ± 0.1	1.4 ± 0.1	0.7 ± 0.1	2.9 ± 0.2

7 Uncertainties

The significance of an experimental result depends on the accuracy of the measurement. The statistical uncertainty due to a limited amount of data is taken into account by the standard deviation σ given by

$$\sigma = \sqrt{N}, \quad (7.1)$$

where N is the number of entries of a counting experiment and Poisson statistics is used.

However, in an experiment, there are also systematic sources of inaccuracy that need to be considered in order to estimate its total uncertainty. They originate from often unavoidable biases of the measurement, but also from theoretical models. In this Chapter, an estimation of the relevant systematic uncertainties for this analysis is performed in a conservative way to prevent an underestimation of the uncertainty. A summary of these uncertainties is given in Table 7.1 and Table 7.2 and will be discussed in the following Sections.

The uncertainties stated in Table 7.1 only affect a comparison between data and simulation (which is from minor importance for the presented analysis, since it is essentially data-driven), and for the setting of SUSY exclusion limits (see Section 8.2). The uncertainties stated in Table 7.2 are of major importance, since they reflect the accuracy of the estimation of all background contributions.

A general way of estimating the effect of an uncertainty on an analysis is to vary a quantity within its uncertainty (e.g. the jet energy scale) and then to observe its influence on the final result (here, the event yield after the whole event selection). This is done for the jet energy scale uncertainty, pile-up uncertainty, and the parton density distribution uncertainty on the $t\bar{t}$ sample, since it has the largest contribution in the signal region and hence the largest number of entries after the event selection and therefore the largest statistical significance.¹

The uncertainties have been studied for all three channels, $\mu\mu$, $e\mu$, and ee , respectively, and show similar results. For simplicity, only the $\mu\mu$ -channel is shown in this Chapter and the other channels can be found in the Appendix.

7.1 Jet Energy Scale

In order to perform an energy measurement in a particle physics experiment, the amount of collected charge from calorimeter cells must be translated into an energy deposit. This is done by the Jet Energy Scale (JES). The estimation of the JES is not trivial and requires a careful calibration, and its uncertainty needs to be considered in an analysis. In principle, the JES uncertainty depends on the flavor, p_T , and η of the jet. However, it

¹With the QCD sample e.g. this investigation would not be possible, since none of its simulated events survives the whole event selection flow.

Table 7.1: Summary of the relevant uncertainties for the simulated SM processes (excluding rare processes) for all channels ($\mu\mu$, $e\mu$, ee). The propagation of the asymmetric PDF uncertainty yields a total uncertainty of $^{+20}_{-18}$ %. For simplicity and as a conservative approach a total uncertainty of ± 20 % is chosen.

Source of uncertainty	Value
JES	± 10 %
Lepton isolation modeling	± 10 %
Lepton reconstruction	± 8 %
PDF	$^{+10}_{-5}$ %
Luminosity	± 4.5 %
Trigger efficiency	± 3 %
Pile-up	± 2 %
Total	20 %

Table 7.2: Uncertainties relevant for the background prediction and the setting of exclusion limits for all decay channels ($\mu\mu$, $e\mu$, ee).

Source of uncertainty	Value
TL method	± 50 %
Rare SM processes (MC)	± 50 %
Charge flip estimation	± 20 %

has been shown [120] that the absolute JES uncertainty for all jet types is smaller than 2 % for jets with a $p_T > 40$ GeV. Hence, the JES uncertainty is studied by varying the JES by 2 % up and down. The event yield of the $t\bar{t}$ sample with and without the variation is compared and shown in Table 7.3 for the $\mu\mu$ -channel (for the $e\mu$ - and ee - channel see Appendix C.2). Note, that the variation of the JES is propagated to the E_T^{miss} calculation.

To indicate the statistical significance of the numbers, the number of entries after each selection step is also given. Note, that in particular for the last selection step, the number of entries is very low and thus, large statistical fluctuations occur. Therefore, a simple consistency test is performed: The order of the selection steps is changed. The cuts which are sensitive to a variation of the JES ($N_{\text{jets}} \geq 2$, $E_T^{\text{miss}}(PF) > 100$ GeV, and possibly the lepton isolation requirement) are applied as soon as possible in order to obtain higher statistical power in these selection steps. The result of this consistency test, shown in Table 7.4, is in agreement with the previous one (see Table 7.3).

With respect to both Tables (and the $e\mu$ - and ee -channel from Appendix C.2) the systematic uncertainty due to the JES variation is conservatively estimated to be ± 10 % for all three channels.

Table 7.3: JES variation for the $\mu\mu$ -channel and its impact on the event yield of the $t\bar{t}$ sample. “Entries” gives the number of entries in each selection step, “Norm” is the normalized event yield without a variation of the JES (normalized to $L = 1.1 \text{ fb}^{-1}$), and “Up” (“Down”) is the relative change of the event yield of the up(down)-variation with respect to Norm.

Cut	Entries	Norm	Up [%]	Down [%]
Preselection	233, 137	11, 066	0	0
$N_\mu \geq 2$	69, 880	3, 321	0	0
$p_T(\mu) > 20/10 \text{ GeV}/c$	64, 926	3, 081	0	0
$N_{\text{SS candidates}} = 1$	14, 773	704	0	0
$N_{\text{Jets}} \geq 2$	11, 747	558	1.1	−1.3
$E_T^{\text{miss}}(PF) > 100 \text{ GeV}$	1, 446	64.4	9.9	−6.6
$I_{\text{Rel}}(\text{1st iso lep})$	1, 228	55.2	9.9	−6.7
$I_{\text{Rel}}(\text{2nd iso lep})$	23	1	7.5	0

Table 7.4: Results of the $\mu\mu$ -channel in analogy to Table 7.3. The order of the selection steps has been changed. The cuts which are sensitive to a variation of the JES are applied as soon as possible in order to obtain higher statistical power in these selection steps.

Cut	Entries	Norm	Up [%]	Down [%]
Preselection	233, 137	10, 902	0	0
$N_{\text{Jets}} \geq 2$	191, 820	8, 987	0.9	−1
$E_T^{\text{miss}}(PF) > 100 \text{ GeV}/c$	27, 191	1, 267	6.6	−6
$I_{\text{Rel}}(\text{1st iso lep})$	23, 421	1, 094	6.6	−6.1
$I_{\text{Rel}}(\text{2nd iso lep})$	585	27.5	6.1	−8
$N_\mu \geq 2$	159	7.3	6.5	−7.2
$p_T(\mu) > 20/10 \text{ GeV}$	133	6.1	6.8	−5.2
$N_{\text{SS candidates}} = 1$	23	1	8.4	0

7.2 Lepton Reconstruction

The muon and electron reconstruction uncertainties have been investigated in considerable detail [90, 91]. A Tag-and-Probe method based on Z -events was used to measure the reconstruction efficiency and uncertainty as a function of lepton p_T and η for the lepton identification and isolation separately. It has been shown that an overall uncertainty of 4 % is sufficient to cover the lepton reconstruction uncertainty for both, muons and electrons. Since in this analysis *two* leptons are required in the signal region, an uncertainty of 8 % is applied for all channels due to lepton reconstruction.

Furthermore, an additional systematic uncertainty of 5 % per lepton is applied to account for a potential mis-modeling of the lepton isolation efficiency between data and simulation, since the hadronic activity is smaller in the control region (Z -events) than in

the signal region (dominated by $t\bar{t}$ -events).

7.3 Parton Distribution Functions

The theoretical calculation of cross-sections has an additional uncertainty: the uncertainty of the parton distribution function (PDF), which is used for the simulation of particle physics processes. The PDF is flavor dependent and defined as the probability density for finding a parton within a proton with a certain momentum fraction x (the so-called “Björken- x ”) at momentum transfer Q^2 .

PDFs cannot be obtained by perturbative QCD, because of the inherent non-perturbative effect in a QCD binding state. Instead they are obtained by using experimental data, due to the limitations in present lattice QCD calculations and therefore, they comprise uncertainties.

There are different sets of PDFs provided by different collaborations, e.g. CTEQ6.6 [121], MRST [122], or NNPDF2.0 [123]. The CTEQ6.6 PDF set was applied for the production of the physics processes used in this analysis. However, there is no reason to favor this particular PDF set. Hence, the effect of the PDF uncertainty on the analysis is studied by varying three different PDF sets (CTEQ6.6, MRST, and NNPDF2.0). For the $t\bar{t}$ sample the PDFs are varied for the set of simulated colliding partons which survive a selection step and the largest deviation from the original event yield is determined.

In Table 7.5 the event yield of the $t\bar{t}$ sample with and without PDF variation is listed for the $\mu\mu$ -channel (for the $e\mu$ - and ee -channel see Appendix C.3). To indicate the statistical significance of the numbers, the number of entries after each selection step is also given. The systematic uncertainty due to the PDF variation is estimated to be $+10\% - 5\%$ for all three channels.

Note, that the central value of the strong coupling constant α_s is different for the different PDF sets (which effectively affects the cross-section of a process). The CTEQ6.6 PDF set has the smallest central α_s value. Therefore, the variation with the MRST and NNPDF2.0 gives a shift to upper values of the event yield and dominate the upper limit of the variation. Hence, the PDF variation gives an asymmetric result.

7.4 Luminosity

The luminosity of the simulated processes is normalized to a recorded data luminosity of $L = 1.1 \text{ fb}^{-1}$. However, the measurement of the luminosity underlies uncertainties. Two methods, based on signals from the HF, are used to obtain a real-time instantaneous luminosity:²

- “Zero counting method”: The average fraction of empty HF towers is used to infer the mean number of interactions per bunch crossing.

²The measured number of reconstructed vertices per bunch crossing gives an offline cross check with independent systematic uncertainty. It is known that the vertex reconstruction efficiency has a value of about 70 %. However, a good estimation of the number of additional interactions per bunch crossing is obtained by dividing the average number of reconstructed vertices by 0.7.

Table 7.5: PDF variation for the $\mu\mu$ -channel and its impact on the event yield of the $t\bar{t}$ sample. “Entries” gives the number of entries in each selection step, “Norm” is the normalized event yield without a variation of the JES (normalized to $L = 1.1 \text{ fb}^{-1}$), and “Up” (“Down”) is the relative change of the event yield of the up(down)-variation with respect to Norm.

	Cut	Entries	Norm	Up [%]	Down [%]
	Preselection	233, 137	11, 066	9.5	−3.6
	$N_\mu \geq 2$	69, 880	3, 321	9.7	−3.5
	$p_T(\mu) > 20/10 \text{ GeV}/c$	64, 926	3, 081	9.7	−3.5
	$N_{\text{SS candidates}} = 1$	14, 773	704	9.8	−3.6
	$N_{\text{Jets}} \geq 2$	11, 747	558	9.7	−3.5
	$E_T^{\text{miss}}(PF) > 100 \text{ GeV}$	1, 446	64.4	9.2	−4.4
	$I_{\text{Rel}}(\text{1st iso lep})$	1, 228	55.2	9.4	−4.4
	$I_{\text{Rel}}(\text{2nd iso lep})$	23	1	9.4	−3.3

- A linear relationship between the luminosity and the average transverse energy per tower is used.

An absolute calibration of the luminosity is achieved performing a Van Der Meer scan [124].³

The total uncertainty of the luminosity measurement is estimated to be 4.5 % [125].

7.5 Trigger

The trigger uncertainty for the used triggers in this analysis has been investigated in considerable detail [90, 91]. It has been shown, that the trigger uncertainty for all three channels is within $\pm 3 \%$.

7.6 Pile-up

The pile-up distribution in data is estimated by using the measured instantaneous luminosity $\mathcal{L}$ for each bunch crossing (taking the revolution frequency into account). The number of interactions for a bunch-bunch collision is obtained multiplying $\mathcal{L}$ by the total inelastic cross-section. This method relies on the knowledge of $\mathcal{L}$ for the colliding bunches and of the total inelastic cross-section. Therefore, the pile-up distribution of data is not perfectly known and contains uncertainties, and the uncertainty due to the pile-up weighting must be considered (see Section 5.2).

The pile-up uncertainty is studied by varying the mean of its distribution by ± 0.6 . Figure 7.1 shows the pile-up distribution from data as well as its variation by ± 0.6 .

³During these scans, the beams are moved transversely to each other and the interaction rate is measured. This allows to measure the shape and the size of the interaction region. Eventually, the reduction factor F from equation (3.2) is determined and hence the instantaneous luminosity can be calculated.

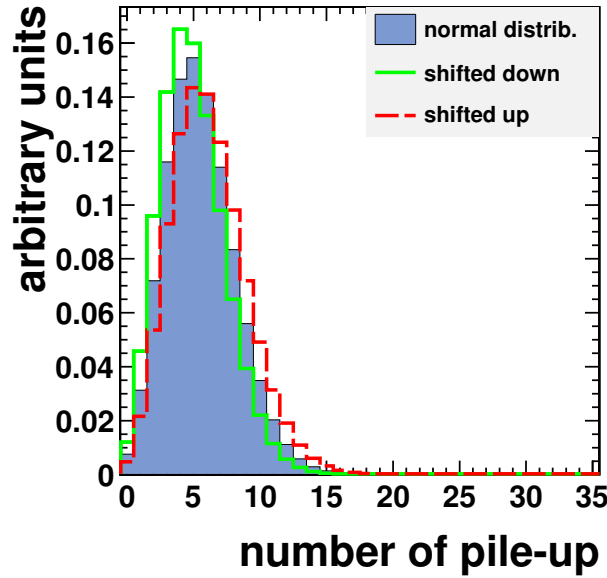


Figure 7.1: Pile-up distribution of the used data (*filled blue*) and its variation of the mean by $+0.6$ (*red dashed line*) and by -0.6 (*green solid line*).

The event yield of the $t\bar{t}$ sample using the shifted distributions compared to the original distribution for the $\mu\mu$ -channel is shown in Table 7.6 (for the $e\mu$ - and ee -channel see Appendix C.4).

To indicate the statistical significance of the numbers, the number of entries after each selection step is also given. Note, that in particular for the last selection step, the number of entries is very low and thus, large statistical fluctuations occur. With respect to the previous selection steps the systematic uncertainty due to the pile-up variation is estimated to be $\pm 2\%$ for all three channels.

7.7 Systematic Uncertainties of the Background Estimation

The total background estimation comprises of the simulated irreducible background in rare SM processes, the estimation of the prompt-fake background via the TL method, and the estimation of the opposite-sign prompt-prompt background due to a charge flip, respectively. For these three background contributions the systematic uncertainties are studied in the following, where a *conservative* estimation is performed in order to ensure that the uncertainty is not underestimated.

7.7.1 Irreducible Background

As already mentioned in Section 6.1, the contribution of rare SM processes, resulting in an irreducible prompt-prompt background, is estimated from simulation, being in agreement

Table 7.6: Pile-up variation for the $\mu\mu$ -channel and its impact on the event yield of the $t\bar{t}$ sample. “Entries” gives the number of entries in each selection step, “Norm” is the normalized event yield without a variation of the JES (normalized to $L = 1.1 \text{ fb}^{-1}$), and “Up” (“Down”) is the relative change of the event yield of the up(down)-variation with respect to Norm.

Cut	Entries	Norm	Up [%]	Down [%]
Preselection	233, 137	11, 066	−0.15	0.17
$N_\mu \geq 2$	69, 880	3, 321	−0.29	0.32
$p_T(\mu) > 20/10 \text{ GeV}/c$	64, 926	3, 081	−0.25	0.29
$N_{\text{SS candidates}} = 1$	14, 773	704	−0.22	0.24
$N_{\text{Jets}} \geq 2$	11, 747	558	−0.35	0.32
$E_T^{\text{miss}}(PF) > 100 \text{ GeV}$	1, 446	64.4	−0.46	0.47
$I_{\text{Rel}}(\text{1st iso lep})$	1, 228	55.2	−0.56	0.46
$I_{\text{Rel}}(\text{2nd iso lep})$	23	1	2.65	−2.5

with other studies [89, 119]. Since most of these processes have never been directly observed, their cross-section is known only from theory and a systematic uncertainty of 50 % is assigned on the yields obtained directly from simulation. Simulation uncertainties due to JES, PDF, etc. are assumed to be covered by this uncertainty.

7.7.2 Tight-to-Loose Method

The background prediction of the TL method is determined utilizing the mathematical concept explained in Section 6.3 by simply counting the tight and loose leptons in the signal region, once the prompt- and fake-ratios (p and f) have been obtained from control samples. Hence, the uncertainty of the TL method arises from the estimation of p and f . There are two main sources which potentially influence the values of p and f :

- Different kinematics/ hadronic activity in the control- and the signal region.
- Contamination of the control samples by undesired processes.

The first influence refers to the fact, that p (f) is obtained from a Z -enriched (QCD-enriched) sample, while it is applied to the signal region, where the dominant contribution originates from $t\bar{t}$ -decays. Hence, the environment in the signal region has an increased hadronic activity and is particularly enriched with b -jets, which also influences lepton isolation. In order to estimate the uncertainty due to this effect, the control samples are required to have at least one b -tagged⁴ jet.

The second influence refers to the composition of the control samples. The fake-ratio is estimated from a QCD-enriched control sample, which is expected to be dominated by fake leptons (see Section 6.3.1). However, with increasing lepton transverse momentum

⁴For the b -tagging a track counting high efficiency algorithm is used.

the contamination with prompt leptons becomes larger (mainly originating from Drell-Yan processes). This can be seen from simulation studies, shown in Fig. 7.2, where the p_T for muons and electrons of the control sample for the fake-ratio is plotted. An increased contamination with prompt leptons will artificially increase the fake-ratio, since prompt leptons are likely to pass the tight selection. Therefore, the choice of the binning range of the fake-ratio also influences the background prediction. However, it is not obvious up to which lepton p_T this effect is negligible. For that purpose, the upper p_T range for f is varied by ± 20 GeV/c to estimate its systematic uncertainty.

To be consistent, the same change in the binning is performed for the prompt-ratio. This is not expected to have a large impact, since the contamination of the Z -enriched control sample with fake leptons only weakly depends on the lepton p_T and therefore, the prompt-ratio is more robust. This can be seen from simulation studies, shown in Fig. 7.3, where the p_T for muons and electrons of the control sample for the prompt-ratio is plotted.

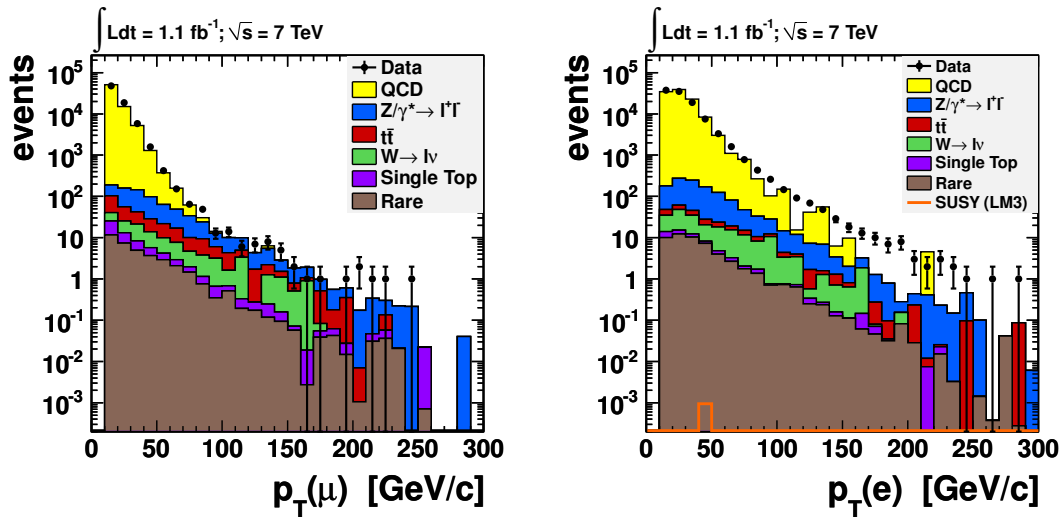


Figure 7.2: Simulation study for the lepton p_T in the QCD-enriched control sample from which the fake-ratio is estimated. *Left:* muons, *right:* electrons. For the electron case, there are too few simulated events in the QCD sample in the higher p_T region, but the trend is similar to what is observed in data.

The varied prompt- and fake-ratios due to the studies explained above can be found in the Appendix C.5. They are applied to the TL method and the change in the prompt-fake background prediction for all three channels is listed in Table 7.7. From this investigation a rough estimate of the systematic uncertainty of $\pm 50\%$ is performed, being in agreement with other studies [89, 119].

From Table 7.7 it can be seen, that the additional b-Tag requirement has a contrary effect on muons compared to electrons: While in the muon case ($\mu\mu$ -channel) the fake-ratio is increased by an additional b-Tag requirement (leading to a larger prompt-fake background prediction), it decreases for electrons (ee -channel). A possible explanation attributes this effect to the different reconstruction of muons and electrons. The worst isolated fake electrons stem from a heavy flavor decay of a b -quark. They tend to be

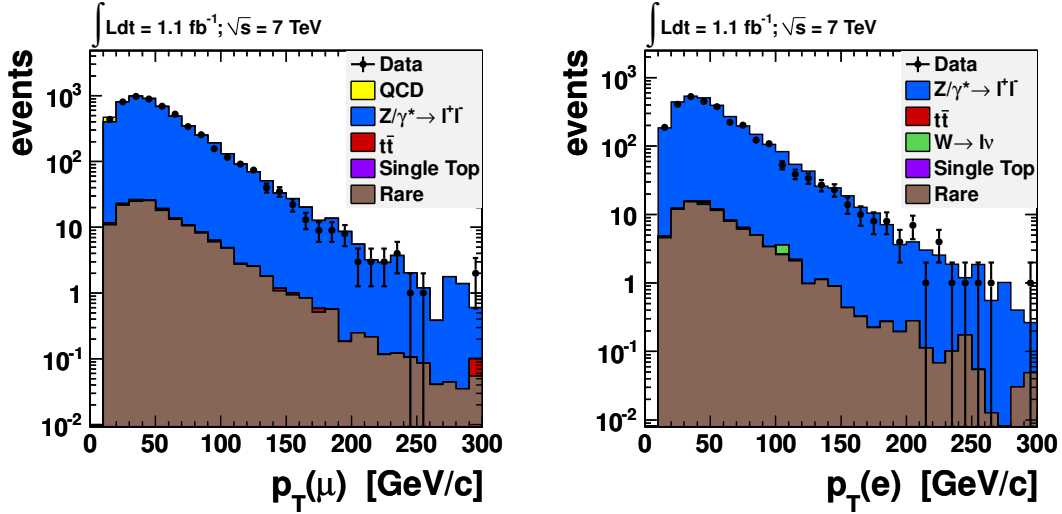


Figure 7.3: Simulation study for the lepton p_T in the Z -enriched control sample from which the prompt-ratio is estimated. *Left:* muons, *right:* electrons.

Table 7.7: Variation of the prompt-fake background prediction due to systematic studies of the TL method. The varied prediction is compared to the prediction with the default parameters: $p_T \leq 60$ GeV and without a b-Tag. The variation is given in percent.

Variation	$\mu\mu$	$e\mu$	ee
$p_T \leq 40$ GeV/c	-7.2	-7.5	-17.2
$p_T \leq 80$ GeV/c	10.9	18.5	38.9
with b-Tag	27.9	0.5	-17.9

surrounded by a large hadronic activity (compared to fake electrons from e.g. photon conversion). Hence, requiring an additional b-Tag enriches the sample with less isolated electrons and thus the fake-ratio decreases (since less electrons also pass the tight selection).

Muons leave a characteristic signature in the muon chambers. Fake muons which do *not* originate from a heavy flavor decay of a b -quark, are muons where the reconstruction failed due to a major problem (wrong assignment of the track, punch-through of other particles into the muon system, etc.). Hence, requiring an additional b-Tag reduces these completely wrong reconstructed muons, enriches the sample with better isolated muons, and thus, the fake-ratio increases (since more muons also pass the tight selection).

In the $e\mu$ -channel both objects, muons and electrons, are present, and the effect due to an additional b-Tag cancels out.

7.7.3 Charge Flip

The background prediction due to a charge mis-measurement is based on the estimation of the charge flip rate r (see Section 6.4). This rate has been extracted from a Z -enriched

control sample, containing di-leptons with an invariant mass in the Z -mass window of $\pm 15 \text{ GeV}/c^2$. To study the effect of a possible contamination of the sample, the size of the window is varied to $m_Z \pm 10 \text{ GeV}/c^2$ and $m_Z \pm 20 \text{ GeV}/c^2$. In Table 7.8 the result of this test is given.

Table 7.8: Uncertainty due to a change of the charge flip rate estimation, based on the variation of the Z -mass window in the event selection (see Section 6.4). The last column gives the relative change of the charge flip prediction using the varied charge flip rate.

Window	N _{OS}	N _{SS}	Flip Rate r	Rel.Variation of pred. [%]
$10 \text{ GeV}/c^2$	89,534	145	$8.09094 \cdot 10^{-4}$	-8.9
$15 \text{ GeV}/c^2$	94,506	167	$8.82763 \cdot 10^{-4}$	-
$20 \text{ GeV}/c^2$	96,957	176	$9.06797 \cdot 10^{-4}$	2.7

With respect to this study and in order to cover possible additional effects, an uncertainty of $\pm 20 \%$ is assigned to the charge flip estimation (which is in agreement with other studies, e.g. [116]).

8 Result

8.1 Summary of the Event Yield

The event yield of the three decay channels ($\mu\mu$, $e\mu$, and ee), as well as their combination, obtained after the whole signal selection is summarized in Fig. 8.1. The different SM background predictions (prompt-fake, charge flip and irreducible background) are summed up and the total uncertainty is shown (statistical and systematic uncertainties are added in quadrature). The comparison with the event yield from data is also displayed.

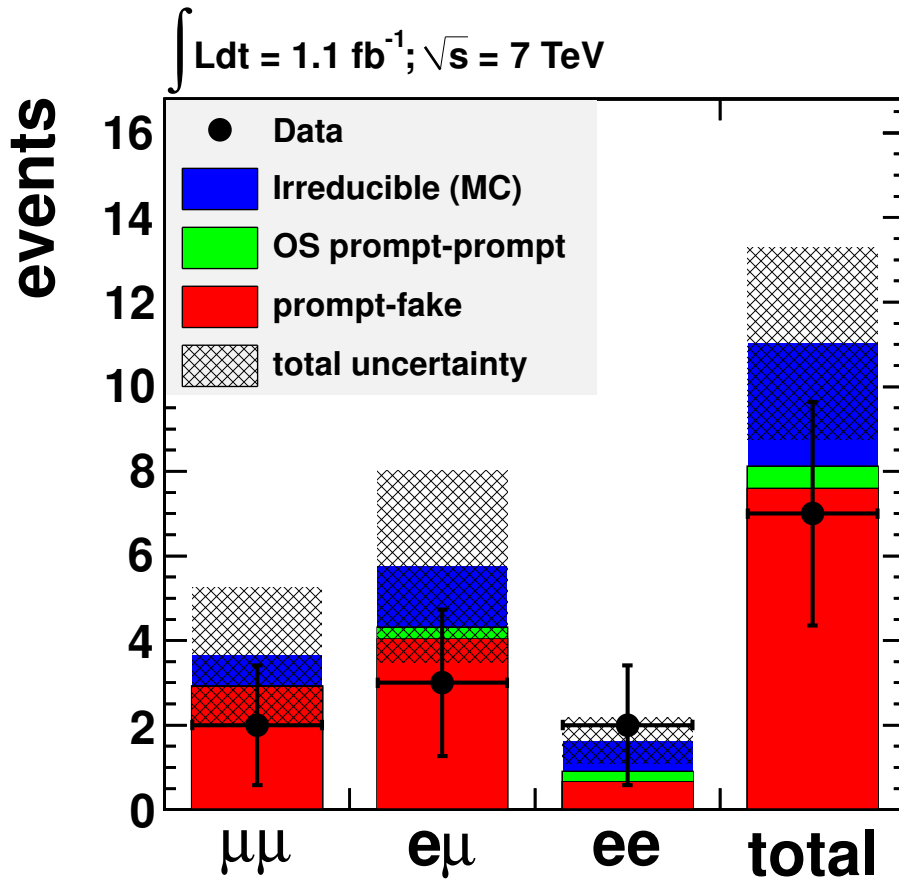


Figure 8.1: Event Yield after the whole signal selection. The different SM background predictions are stacked and displayed in different colors. Its total uncertainty is shown as a shaded area. The measured event yield from data is depicted in black circles.

The SM background prediction is in agreement with the measured result from data within the uncertainties. The background prediction for the ee -channel is smaller than for the $\mu\mu$ -channel. From the physics point of view, there is no reason why the ee -channel should be favored. A possible explanation for this difference is the worse reconstruction efficiency [90, 97] for electrons (which enters the ee -channel quadratically, since it comprises of *two* electrons). Furthermore, the fake-ratio for muons is larger than for electrons for high p_T values (see Fig. 6.8). This might be caused by a contamination of the fake-ratio control sample with prompt leptons, especially in the muon case (see Fig. 7.2).

The largest event yield is obtained in the $e\mu$ -channel. This is expected because the branching ratio for this channel with a mixed flavor is larger than for the same flavor.

For completeness, the event yield for the combined channels is also given in Table 8.1 (more detailed numbers can be found in the Appendix B.4).

Table 8.1: Summary of the event yield after the last selection step for the three combined channels. The event yield is shown for the sum of the simulated samples, some SUSY benchmark points, all SM background predictions (prompt-fake, charge flip and irreducible background) and 1.1 fb^{-1} of data. In the last column the channels are summed up. The first error is the statistical uncertainty and the second one the systematic uncertainty, respectively.

sample	Combined channels
Sum all (MC)	$8.1 \pm 1.0 \pm 1.0$
Signal (LM1)	$41.8 \pm 3.0 \pm 5.2$
Signal (LM3)	$19.9 \pm 2.3 \pm 2.4$
Signal (LM8)	$8.9 \pm 1.4 \pm 1.2$
Prompt-fake pred.	$7.6 \pm 0.9 \pm 2.5$
Charge Flip pred.	$0.5 \pm 0.0 \pm 0.1$
Irreducible pred.	$2.9 \pm 0.2 \pm 0.9$
Sum pred.	$11.0 \pm 1.0 \pm 2.7$
Data	7

8.2 Exclusion Limit for the cMSSM

Since in this analysis no evidence for an excess over the SM expectation is observed in data and thus no new physics is found, the result is interpreted in the context of the cMSSM and an exclusion limit is set. In order to exclude parts of the parameter space of the cMSSM, three parameters are fixed ($\tan \beta = 10$, $A_0 = 0$, and $\text{sign}(\mu) = +1$) and m_0 and $m_{1/2}$ are varied. The variable m_0 is varied between $100 \text{ GeV}/c^2 - 3,000 \text{ GeV}/c^2$ in steps of $20 \text{ GeV}/c^2$ and $m_{1/2}$ between $100 \text{ GeV}/c^2 - 1,000 \text{ GeV}/c^2$ in steps of $20 \text{ GeV}/c^2$. A total of 10,000 events are simulated for each point of this grid. They are enriched with

di-lepton events and the corresponding filter-efficiency is considered for each point in the plane, as well as the NLO cross-section.¹

A scan in the m_0 - $m_{1/2}$ -plane is performed (combining all three decay channels), based on the following input:

- *Signal efficiency* for each point: It is obtained by dividing the event yield before the signal selection (10,000 events) by the event yield after the signal selection (all three decay channels are added up). The uncertainty for all points is taken to be $\pm 20\%$ (see Chapter 7).
- *Background prediction*: It is assumed, that the background prediction is not affected by a SUSY signal and taken as constant for all points in the m_0 - $m_{1/2}$ -plane ($11.0 \pm 1.0 \pm 2.7$, compare Table 8.1). This approximation is discussed in the Appendix E.
- *Measurement* from 1.1 fb^{-1} of data (7 events, compare Table 5.16).
- *Luminosity* and its uncertainty of 4.5% (see Section 7.4).

Based on this input, for each point of the grid the maximum cross-section σ_{max} for a signal is estimated (some details can be found in Appendix F).² This is the cross-section for which the signal added to the background prediction is still in agreement with the measurement. The estimation is performed utilizing Bayesian statistics and the Upper Limit (UL) for the cross-section is set at a 95 % Credibility Level (CL). A point of the grid in the m_0 - $m_{1/2}$ -plane is excluded if the NLO cross-section of the SUSY signal is larger than σ_{max} . In Fig. 8.2 the exclusion line (blue) in the m_0 - $m_{1/2}$ -plane is shown. The exclusion line has been set in the middle of the bin with the largest $m_{1/2}$ -value for a given m_0 , which was excluded by the described procedure.

Note, that the uncertainty of $\pm 20\%$ for the signal efficiency (taken from Chapter 7) is a conservative estimate, since e.g. the effect of the JES uncertainty is expected to be smaller in SUSY scenarios than for SM processes from which the uncertainty was estimated.³ Furthermore, the signal efficiency uncertainty already contains the uncertainty of the luminosity. Moreover, the theoretical uncertainty of the PDF might be taken into account at a later step, when an error is set to the exclusion line. Based on variations of the PDFs, a varied cross-section of a SUSY signal can be utilized to set a theoretical uncertainty for the exclusion line.

To estimate the uncertainty of the obtained exclusion line, more detailed investigations have to be performed. However, Fig. 8.2 gives an idea of the exclusion power of the presented analysis in the cMSSM.

¹In principle, it is possible to distinguish between the different production subprocesses. However, here, an overall NLO cross-section is utilized.

²Since the background prediction, the measurement, and the luminosity are taken to be constant for each point in the m_0 - $m_{1/2}$ -plane, σ_{max} only depends on their errors and the signal efficiency.

³In general, in SUSY models, the hadronic activity is larger than in SM processes. Hence, the efficiency to pass a jet multiplicity requirement is not so strongly affected by a variation of the JES in SUSY models than for SM processes than in SM processes.

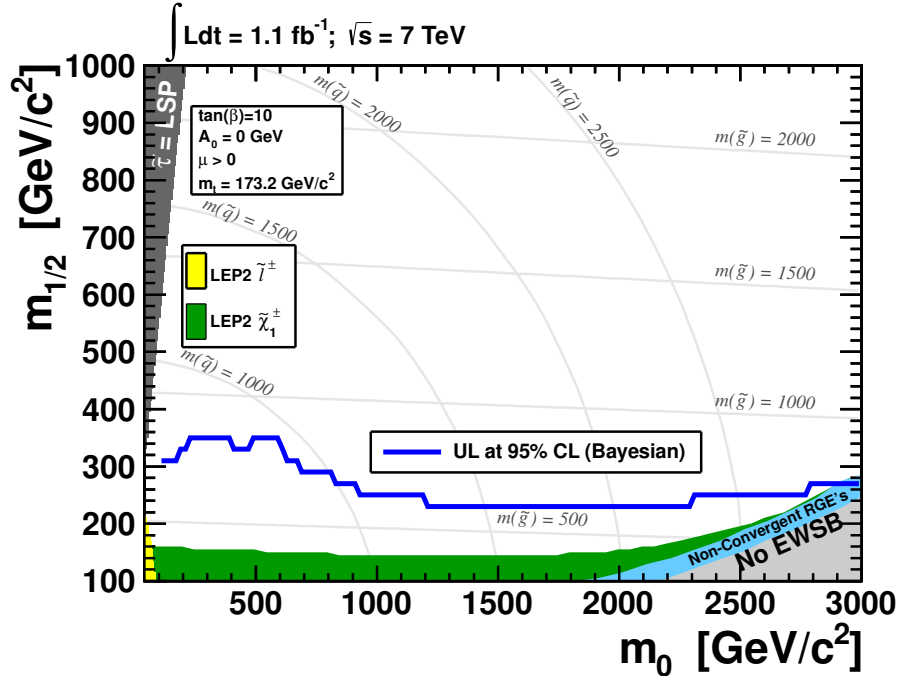


Figure 8.2: Exclusion line (blue) in the m_0 - $m_{1/2}$ -plane of the cMSSM. The line corresponds to an Upper Limit (UL) for the exclusion at a 95% Credibility Level (CL) in Bayesian statistics. The area below the blue line is excluded. The yellow and green bands indicate regions excluded by LEP 2. For the right lower corner no reasonable solution exists within the cMSSM. The gray lines mark constant gluino ($\tilde{g}$) and squark ($\tilde{q}$) masses.

9 Summary and Outlook

This thesis was carried out during the first two years of data-taking of the LHC. In this phase, the search for new physics is interesting, since the collected data allows to search for new particles, or to exclude new physics scenarios, respectively.

The technical part of this theses contributes to the planned upgrade of the CMS experiment, by giving a motivation for a certain readout design (“1-4-4-8”) of the HCAL, in order to improve the performance of the detector in future. A simulation study of an energy weighting method of the HCAL is performed, investigating various readout designs and their enhanced resolution and linearity. An additional study illustrates the impact of a channel breakdown to the energy measurement. Detailed suggestions for improvements of the energy weighting method are given in Section 4.4.

The analysis presented in this thesis is a search for SUSY in same-sign di-lepton events, utilizing 1.1 fb^{-1} of data collected by the CMS experiment at a center-of-mass energy of 7 TeV. A counting experiment is established by developing an event selection in order to select signal events in three different channels ($\mu\mu$, $e\mu$, and ee), while rejecting the SM background as efficiently as possible. The trigger strategy is chosen to investigate *high- p_T* leptons ($p_T > 20(10) \text{ GeV}$), allowing to probe kinematic regions with a low H_T requirement, and the object selection is defined accordingly. The remaining SM background contributions are classified into several categories and estimated using different techniques:

- *Prompt-fake Background*: Mostly from semi-leptonic $t\bar{t}$ -decays associated with a leptonic heavy flavor decay. It is estimated via the data-driven TL method and contributes with about 70 % to the total background.
- *Prompt-Prompt Background (irreducible)*: Originating from rare SM sources, like e.g. $W^\pm W^\pm$ (double parton scattering), WWW , WWZ , etc. It is estimated from simulation¹ and contributes with about 25 % to the total background.
- *Prompt-Prompt Background (charge Flip)*: Originating from SM sources, like e.g. Z or W^+W^- , in which the charge of one lepton is mis-measured. It is estimated via a data-driven Tag-and-Probe method and contributes with about 5 % to the total background.
- *Fake-fake Background*: Mostly from QCD sources. It is estimated via the data-driven TL method and found to be negligible.

¹It is assumed that in particular these processes can be reasonably well modeled, since they comprise *prompt* processes. It is more difficult to simulate processes associated with hadronic activity, as in the case of a leptonic heavy flavor decay, leading to *fake* leptons (especially the lepton isolation properties are difficult to simulate). However, a conservative uncertainty of 50 % is assigned to this background contribution.

The observed yields agree with the SM prediction. Hence, no evidence for new physics is found. However, the observations are used to set upper limits on the cMSSM in order to exclude parts of its parameter space.

In principle, the analysis is in a good shape to be applied to more data. However, for future investigations additional studies can be performed:

- Instead of the binned prompt and fake-ratio a two dimensional function from a fit in the p_T - η -plane of electrons and muons can be developed. This prevents the method to be biased by binning effects.
- Additional constraints can be introduced to the TL method in order to ensure that it always yields physically plausible results. One constraint e.g. could set $N_{ff} = 0$ if the analytical solution yields a negative result for this number (see Section 6.3.3). A subsequent fit of the other variables should yield an improvement of the background prediction.
- Consistency of the TL method with the irreducible background prediction from simulation: In principle, the TL method also yields a prediction for the irreducible prompt-prompt background (ppN_{pp}). However, the statistical significance for the data used in this analysis is not sufficient to perform the prompt-prompt background prediction from the TL method. To be consistent with the prediction from simulation (irreducible background prediction), N_{pp} could be set to the value which is obtained from simulation. A subsequent fit of the other variables should yield an improved of the background prediction.²
- The setting of exclusion limits of the cMSSM can be improved by evaluating the background prediction of signal plus background for each point of the m_0 - $m_{1/2}$ -plane, as well as estimating its uncertainty.
- The analysis can be extended by a study of simplified models. These are more general phenomenological models which cover a wider range of new physics scenarios.
- Analyzing more data allows to define additional signal regions (e.g. with stronger H_T and E_T^{miss} requirements) which are potentially more sensitive to a NP signal.

The exploration of more data involves the utilization of new triggers and therefore an adaption of the object selection. However, future analyses with an increased integrated luminosity will enlighten the realm of new physics scenarios.

²Note, that the total background prediction tends to be overestimated. This might be prevented by the suggested procedure.

A Additional Investigations for the Energy Weighting of the HCAL

Additional investigations for the energy weighting of the HCAL are presented in the following. First of all, simulated shower profiles within the HCAL are shown. It is important to get an idea of the showering of particles in the simulation in order to verify its reliability. The impact of a readout channel defect for the “1-4-4-8” and “1-3-3-3-4” readout design is investigated thereafter. The motivation for this study is, that there are other competing readout designs, which would barely be affected by such a readout channel defect.¹ Subsequently, it is shown why the first layer is not weighted in the energy weighting method and it is explained, why the ECAL could not be included into the weighting procedure in this simulation study. Finally, the most detailed readout design “17x1”, in which every layer is read out separately, is investigated.

A.1 Shower profiles

In order to characterize the showering within the HCAL, simulated pions enter the HCAL without ECAL and without passive material in front. No cuts are applied here (compare with Section 4.2.4). Figure A.1 shows the average energy deposition per layer $\langle E_{\text{layer}} \rangle$ originating from simulated pion showers in the CMS HCAL for different energies.

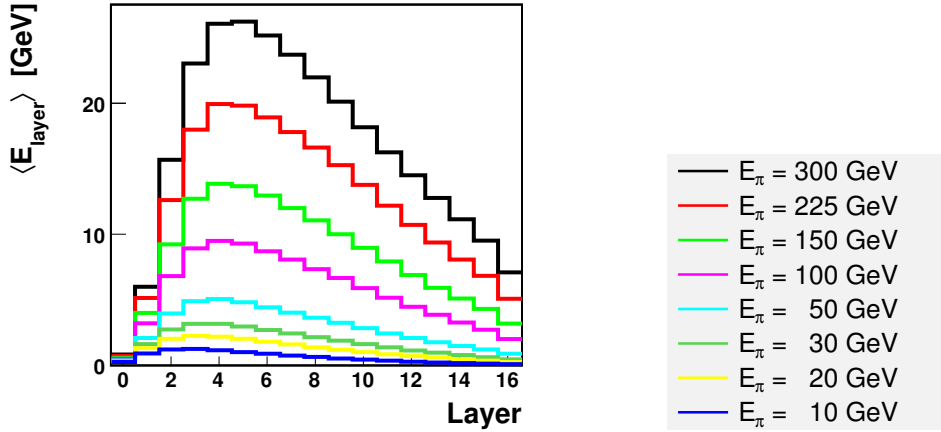
In this simulation, the pions are shot *directly* into the HCAL (without ECAL and without passive material in front of the HCAL) in order to characterize solely the showering within the HCAL. No cuts are applied. Note that the average energy deposition in the first layer is quite small. This can be explained by the fact that there is no absorber in front of the first layer, but only a scintillator which is hit by the pions first. Therefore, most of the pions pass this layer without any interaction.

To yield the average energy deposition per layer, thousands of pions have been generated (see Table A.1). Due to technical reasons, the number of events is smaller for higher energies.

¹In these designs, the readout channels combine alternating layers, e.g. one channel combines every second layer, another channel every second plus one layer. Hence, one readout channel defect could mostly be compensated by using such a design. However, depth information of a particle shower would be lost and therefore, applying an energy weighting method would be impossible for such a readout design. Additionally, the concern of a readout channel defect mainly arises from experiences with the delicate HPDs. These, however, will be replaced by much more robust SiPMs and thus, a readout channel defect is expected to be no issue anymore.

Table A.1: Number of simulated pions which have been generated for different energies.

E_π [GeV]	10	20	30	50	100	150	225	300
no. of events	50,000	50,000	50,000	50,000	40,000	35,000	30,000	30,000

**Figure A.1:** Pion shower profiles of the CMS HCAL for different simulated pion energies E_π (color coded), shot directly onto the HCAL. The average energy deposition as a function of the HCAL layer is shown. One layer corresponds to more than three radiation lengths and to less than a third interaction length.

A.2 Impact of a readout channel defect

In order to investigate the impact of a readout channel defect, in the following scenario a total failure of one readout channel within a tower is simulated for the readout designs “1-4-4-8” and “1-3-3-3-4”. A total failure means that a readout channel gives no signal at all. Note that this scenario is the worst case and the most conservative scenario because no software compensation for the simulated defect is applied here. In reality, one would correct it with an estimation of the energy of the defect readout channel which is not addressed here.²

Figure A.3 shows systematically the effect of a dead channel of the “1-4-4-8” readout design and Fig. A.4 for the “1-3-3-3-4” readout design, respectively. The peak position of the energy distribution with no dead channel is marked as a red line. A dead channel is marked with an ‘X’. Considering the “1-4-4-8” readout design, the largest deviation from the energy distribution without a readout channel defect is observed when the second channel is giving no signal. This is expected as the shower maximum is observed at this position on average (see Fig A.1). Figure A.2 shows the impact of a readout channel defect for the energy resolution of the “1-4-4-8” readout design. For comparison the energy resolution without a defect is plotted. Since it is often not possible to perform an

²A possible software compensation could be: If an energy deposition is measured in front and behind a dead channel, one could think of a method to estimate the energy deposition within the defect channel.

appropriate Gaussian fit to the energy distributions, here, the energy resolution is derived from the RMS- and mean-value of the histograms.

In analogy to the “1-4-4-8” readout design, the largest impact of a channel breakdown for the “1-3-3-3-4” readout design is also found in the second channel (see Fig. A.4). However, the effect is larger for the “1-4-4-8” readout design, since the second channel contains on average more energy.

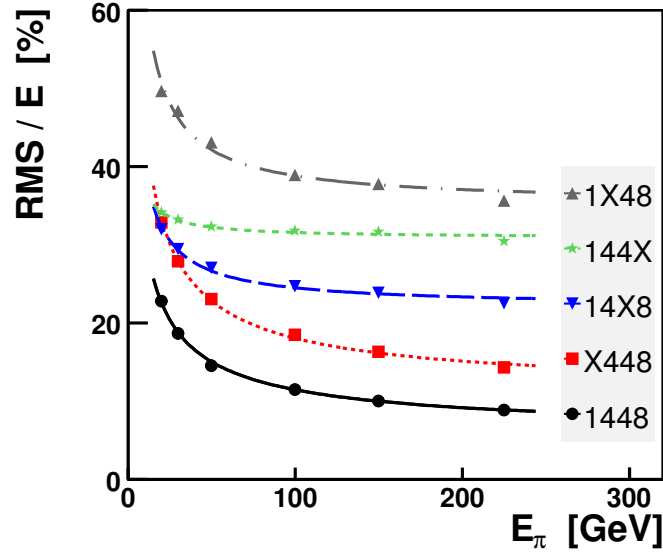


Figure A.2: Illustration of the impact of a readout channel defect for the “1-4-4-8” readout design, regarding the energy resolution. For comparison the energy resolution without a defect is plotted.

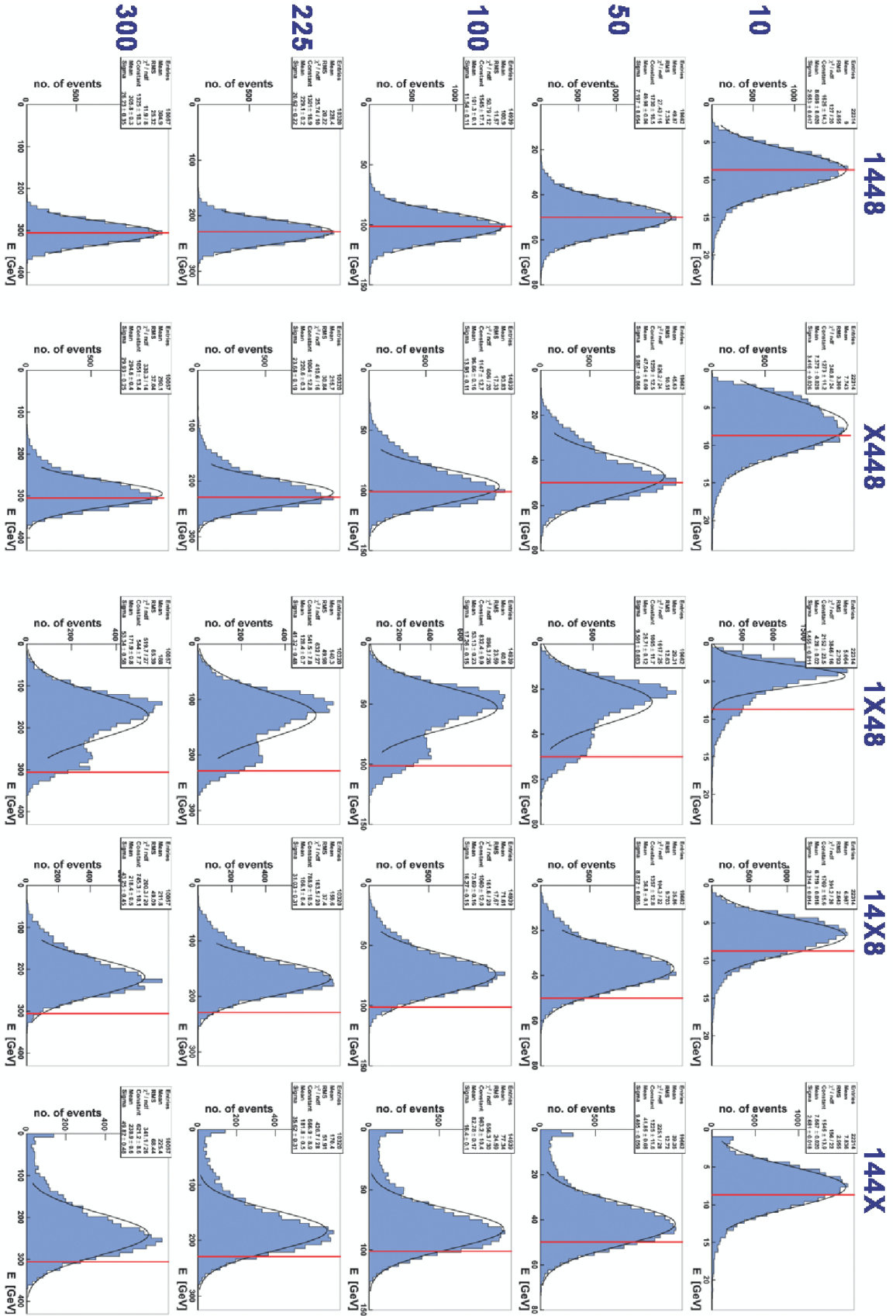


Figure A.3: Illustration of the impact of a readout channel defect for the “1-4-4-8” readout design for the energies 10, 50 100, 225, and 300 GeV. ‘X’ marks the channel which is broken. The red line refers to the peak position of the energy distribution in which every readout channel works.

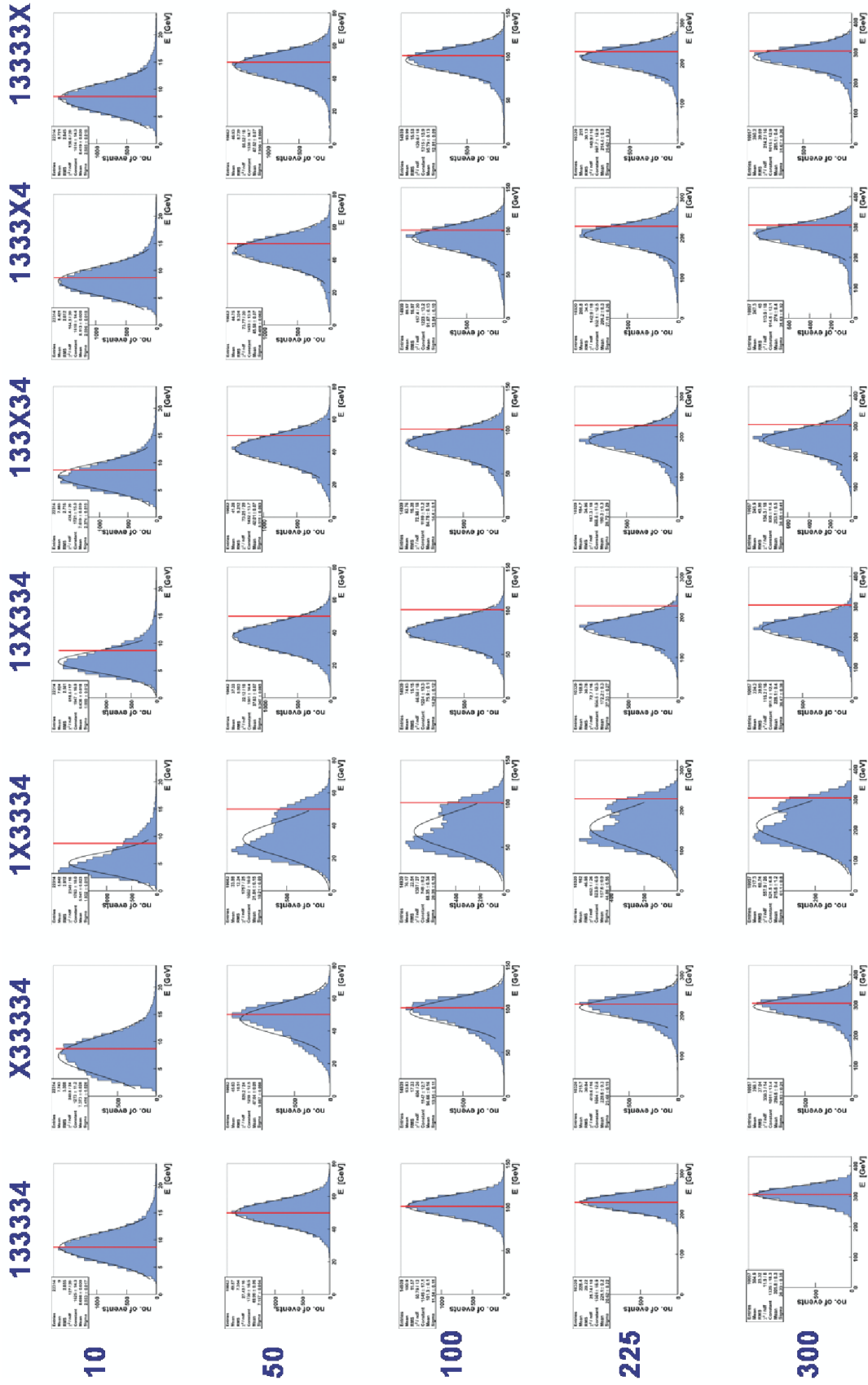


Figure A.4: Illustration of the impact of a readout channel defect for the “1-3-3-3-4” readout design for the energies 10, 50, 100, 225, and 300 GeV. ‘X’ marks the channel which is broken. The red line refers to the peak position of the energy distribution in which every readout channel works.

A.3 Weighting without the first layer

As mentioned in Section 4.2.4, the weighting method is implemented without weighting the first layer. This is due to the fact that the first layer differs from the other layers (see Section 4.2.1) and thus no improvement of the energy measurement is made by including it in the weighting procedure. This can be seen from Fig. A.5, in which the energy distributions for 50 GeV pions with respect to the "1-4-4-8" readout design is shown. The left plot shows the energy distribution before weighting, the middle plot after weighting without the first layer, and the right plot after weighting with the first layer. The relative energy resolution for the distribution without weighting the first layer $\sigma_E/E = (12.6 \pm 0.6) \%$ is smaller than the relative energy resolution of the distribution with weighting the first layer $\sigma_E/E = (13.1 \pm 0.7) \%$. Although this effect is not large, it can be observed systematically for every used pion energy and every investigated readout scheme of these studies. Thus, the weighting is more successful without weighting the first layer.

The upper left plot of Fig. 4.6 is also instructive to elucidate this aspect. Evaluating weighting factors from this plot is a bad compromise between the two populations. A profile plot does not take the discrepancy of these populations into account.

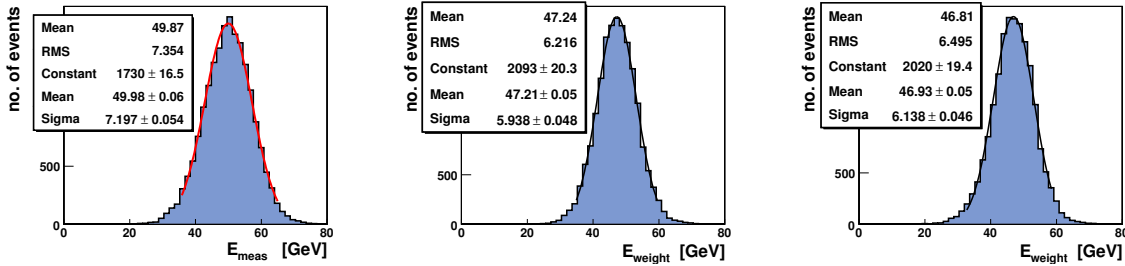


Figure A.5: Energy distributions for 50 GeV pions with the "1-4-4-8" readout design before weighting (left), after weighting without the first layer (middle), and after weighting with the first layer (right). A Gaussian fit is applied to the distributions in the region indicated by a red line. The corresponding relative energy resolutions are: $\sigma_E/E = (14.4 \pm 0.9) \%$ (left), $\sigma_E/E = (12.6 \pm 0.6) \%$ (middle), and $\sigma_E/E = (13.1 \pm 0.7) \%$ (right).

A.4 Weighting without ECAL

As mentioned in Section 4.2.4, the weighting method is implemented demanding that the energy fraction left in the ECAL is smaller than 2 GeV. This is due to the fact that for the ECAL no readout is implemented in the simulation and thus only the Monte Carlo truth information exists for this component. As a consequence every energy deposition in the ECAL contributes as a *perfectly measured* fraction to the total shower energy. If the weighting scenario would include the ECAL (no ECAL cut), a fraction of about 20 %–30 % of the total shower energy will come from the energy depositions in the ECAL. This introduces a bias to the weighting factors: As a result smaller weighting factors would

always be favored, because then the relative fraction of the ECAL measurement would be larger and the total energy measurement appears to be more accurate. Note that the weighting method, as it is applied in these studies, only affects the energy measurement of the HCAL and not of the ECAL.

This feature has been investigated using a *constant* (fake-)weighting factor of 0.8 for *all* energy depositions of the HCAL. In this way — by simply multiplying the HCAL energy depositions with the factor of 0.8 — one would a priori not expect an improvement of the relative energy resolution, but only a shift of the whole energy distribution to smaller values. However, the (*perfectly known*) energy depositions of the ECAL remain unchanged and thus the relative fraction of the ECAL measurement would be larger and therefore the total energy measurement appears to be more accurate and the relative energy resolution seemingly improves.

This is the reason, why the effect of the tabulated weighting method can only be studied in the scenario including the ECAL cut in this simulation.

A.5 Weighting with readout design 17x1

In principle, having more readout channels results in a larger improvement of the energy measurement. However, in these studies it turned out that the most detailed readout design “17x1” yields a smaller improvement of the energy resolution than the “1-4-4-8” readout design. An explanation can be the sampling fluctuations which are more pronounced for the “17x1” readout design. Note that one HCAL layer has a thickness of less than a third interaction length. Figure A.6 shows scatter plots from which the weighting factors are developed. The left plot illustrates the fourth channel of the “1-4-4-8” readout design while the right one shows the 15th channel of the “17x1” readout design. The “cloud” of points of the left plot is more situated around the value ‘1’ while the spread of the entries of the right plot is considerably larger. These two plots illustrate that the influence of sampling fluctuations is large for the “17x1” readout design. Thus the smaller improvement of the weighting method with this readout design can be explained.

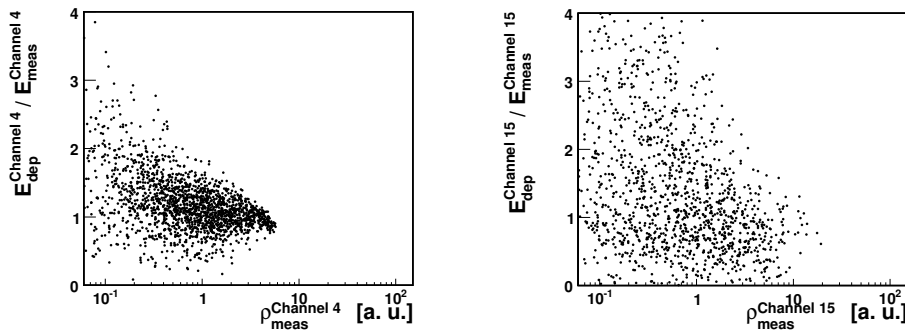


Figure A.6: Scatter plots from which the weighting factors are developed; here, for illustration reasons only for 5,000 pions. *Left:* last Channel of the “1-4-4-8” readout design. *Right:* Channel 15 for the “17x1” readout design.

B Details of the Event Selection Flow

Figures B.1 and B.2 show the event yield of the event selection flow for the $e\mu$ - and the ee -channel, respectively (the $\mu\mu$ -channel can be found in Section 5.7). The detailed numbers for all channels are listed in Tables B.1, B.2, and B.3 thereafter. A summary of the event yield after the whole selection flow, including the systematic uncertainties and a combination of the three channels is given in Table B.4.

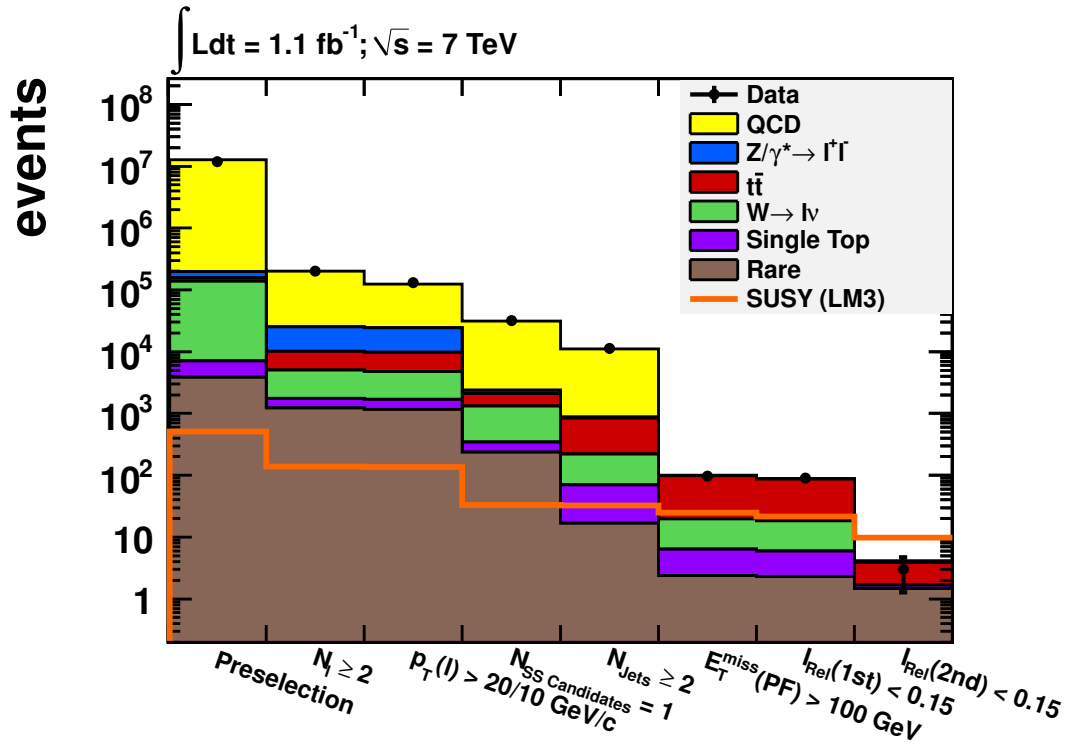


Figure B.1: Event selection flow for the $e\mu$ -channel. Different MC samples are stacked in different colors, a SUSY signal sample is shown as a line, and the data yield is displayed with black points. Everything is normalized to the integrated data luminosity of $L = 1.1 \text{ fb}^{-1}$.

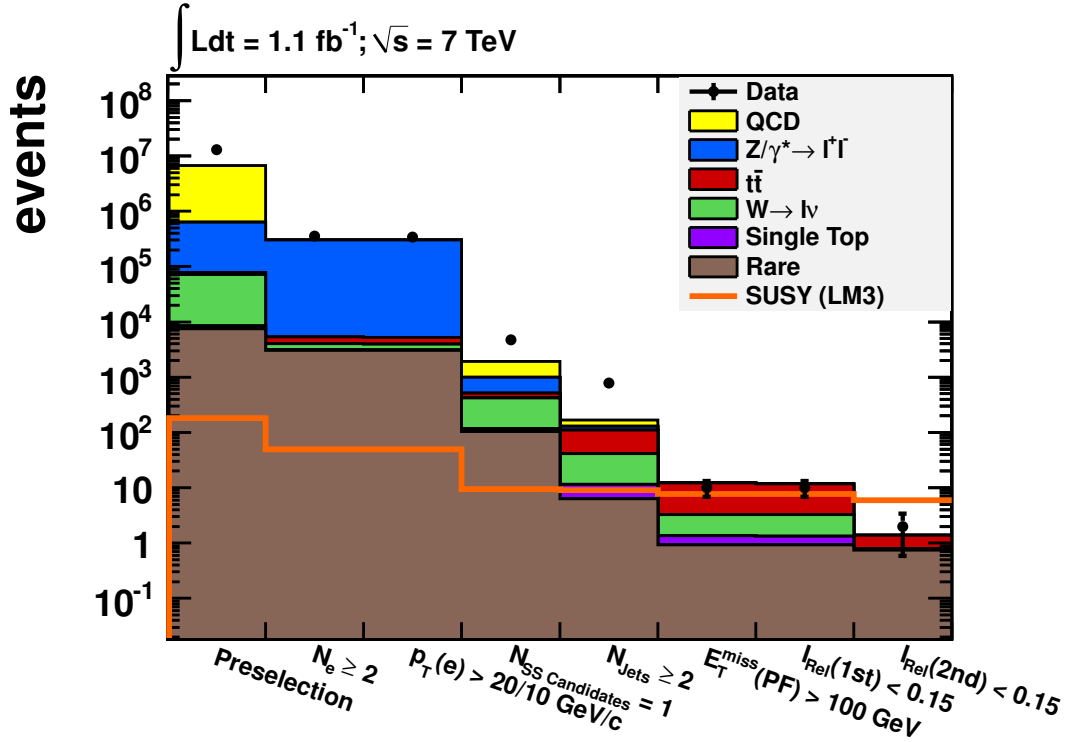


Figure B.2: Event selection flow for the ee -channel. Different MC samples are stacked in different colors, a SUSY signal sample is shown as a line, and the data yield is displayed with black points. Everything is normalized to the integrated data luminosity of $L = 1.1 \text{ fb}^{-1}$. The discrepancy between data and simulation does not affect the result, since the background prediction is essentially data-driven. It probably arises from the QCD electron enriched sample. Using a newer production of the simulated sample (with new trigger menus etc.) might solve the discrepancy. However, a discrepancy due to QCD samples does not affect the final selection, since they do not contribute to the signal region.

Table B.1: Event Yield of the event selection flow for the $\mu\mu$ -channel. The event yield is shown for single simulated SM processes, its sum, some SUSY benchmark points, and 1.1 fb^{-1} of data. $N_\mu(20/10)$ refers to the requirement of at least two muons in an event, where at least one muon fulfills $p_T > 20\text{ GeV}/c$, and both fulfill $p_T > 10\text{ GeV}/c$, and where the muons were cleaned by the Z-veto, low mass resonance veto and a priority of the channels is given in the order $\mu\mu$, $e\mu$, and ee . The given uncertainty is statistical only. Numbers are summed up before they are rounded.

Sample	Preselection	$N_\mu(20/10) \geq 2$	$N_{\text{SS cand}} = 1$	$N_{\text{Jets}} \geq 2$	$E_T^{\text{miss}} > 100$	$I_{\text{Rel},1} < 0.15$	$I_{\text{Rel},2} < 0.15$
QCD	8900798 ± 43230	104724 ± 2514	24019 ± 1295	8906 ± 558	1.2 ± 1	0	0
$Z/\gamma^* \rightarrow l^+l^-$	546872 ± 292	467170 ± 269	126 ± 4.4	22 ± 1.8	0.3 ± 0.2	0.3 ± 0.2	0
$t\bar{t}$	11066 ± 29	3081 ± 15	704 ± 7	565 ± 7	69 ± 2.3	59.2 ± 2.1	1.0 ± 0.3
$W \rightarrow l\nu$	15832 ± 104	1829 ± 35	507 ± 19	100 ± 8	7.2 ± 2.3	7.2 ± 2.3	0.7 ± 0.7
Single top	1800 ± 6	375 ± 3	98 ± 1	48 ± 0.9	4.0 ± 0.2	3.5 ± 0.2	0.1 ± 0
Rare	6608 ± 18	5077 ± 16	120 ± 1.4	11.4 ± 0.4	1.4 ± 0.1	1.3 ± 0.1	0.7 ± 0.1
sum (MC)	9482976 ± 43678	582255 ± 2852	25574 ± 1328	9653 ± 576	83.1 ± 6.1	71.5 ± 4.9	2.6 ± 1.1
Signal (LM1)	444 ± 10	187 ± 6.4	29.4 ± 2.5	25.1 ± 2.3	22.5 ± 2.2	21.2 ± 2.1	13.3 ± 1.7
Signal (LM3)	339 ± 9.1	102 ± 5.0	23.5 ± 2.4	22.0 ± 2.3	16.8 ± 2.0	14.9 ± 1.9	4.2 ± 0.9
Signal (LM8)	116 ± 4.8	37.4 ± 2.7	6.3 ± 1.2	5.7 ± 1.1	4.4 ± 1.0	3.7 ± 0.9	1.8 ± 0.6
Data	5278946 ± 2298	692300 ± 832	24613 ± 157	9968 ± 100	101 ± 10	88 ± 9	2 ± 1

Table B.2: Event Yield of the event selection flow for the $e\mu$ -channel. The event yield is shown for single simulated SM processes, its sum, some SUSY benchmark points, and 1.1 fb^{-1} of data. $N_{\ell}(20/10)$ refers to the requirement of at least two leptons in an event, where at least one lepton fulfills $p_T > 20\text{ GeV}/c$, and both fulfill $p_T > 10\text{ GeV}/c$, and where the leptons were cleaned by the Z-veto, low mass resonance veto and a priority of the channels is given in the order $\mu\mu$, $e\mu$, and ee . The given uncertainty is statistical only. Numbers are summed up before they are rounded.

Sample	Preselection	$N_{\ell}(20/10) \geq 2$	$N_{SS\text{ cand}} = 1$	$N_{\text{Jets}} \geq 2$	$E_{\text{T}}^{\text{miss}} > 100$	$I_{\text{Rel},1} < 0.15$	$I_{\text{Rel},2} < 0.15$
QCD	12462172 ± 30731	100622 ± 2841	28858 ± 1443	10126 ± 660	0.5 ± 0.5	0	0
$Z/\gamma^* \rightarrow l^+l^-$	38171 ± 77	144467 ± 47	261 ± 6.4	28.8 ± 2.1	0.5 ± 0.3	0.5 ± 0.3	0.2 ± 0.2
$t\bar{t}$	19480 ± 39	4924 ± 19	795 ± 7.8	631 ± 6.9	78.9 ± 2.5	68.8 ± 2.3	2.2 ± 0.4
$W \rightarrow l\nu$	130977 ± 300	3093 ± 46	961 ± 25.7	148 ± 9.9	13.4 ± 3.0	12.4 ± 2.9	0
Single top	3259 ± 7.6	516 ± 2.9	110 ± 1.4	53.8 ± 1.0	4.0 ± 0.2	3.6 ± 0.2	0.2 ± 0.1
Rare	3851 ± 14.3	1168 ± 6.5	238 ± 2.6	16.8 ± 0.5	2.4 ± 0.2	2.3 ± 0.2	1.5 ± 0.1
sum (MC)	12657911 ± 31167	124770 ± 2963	31223 ± 1487	11005 ± 681	99.6 ± 6.6	87.6 ± 5.9	4.1 ± 0.8
Signal (LM1)	470 ± 10.1	144 ± 5.6	46.9 ± 3.2	39.2 ± 3.0	34.0 ± 2.8	32.1 ± 2.7	20.6 ± 2.1
Signal (LM3)	511 ± 11.2	136 ± 5.8	33.3 ± 2.9	32.3 ± 2.8	25.1 ± 2.5	21.5 ± 2.3	9.8 ± 1.6
Signal (LM8)	185 ± 6.1	55.3 ± 3.3	16.1 ± 1.9	14.7 ± 1.8	12.2 ± 1.6	9.5 ± 1.4	5.5 ± 1.1
Data	11806375 ± 3436	131152 ± 362	31822 ± 178	11219 ± 106	97 ± 10	90 ± 9	3 ± 2

Table B.3: Event Yield of the event selection flow for the ee -channel. The event yield is shown for single simulated SM processes, its sum, some SUSY benchmark points, and 1.1 fb^{-1} of data. $N_e(20/10)$ refers to the requirement of at least two electrons in an event, where at least one electron fulfills $p_T > 20\text{ GeV}/c$, and both fulfill $p_T > 10\text{ GeV}/c$, and where the electrons were cleaned by the Z-veto, low mass resonance veto and a priority of the channels is given in the order $\mu\mu$, $e\mu$, and ee . The given uncertainty is statistical only. Numbers are summed up before they are rounded.

Sample	Preselection	$N_e(20/10) \geq 2$	$N_{\text{SS cand}} = 1$	$N_{\text{Jets}} \geq 2$	$E_T^{\text{miss}} > 100$	$I_{\text{Rel},1} < 0.15$	$I_{\text{Rel},2} < 0.15$
QCD	6048772 ± 24409	2465 ± 488	922 ± 276	40.0 ± 24.5	0	0	0
$Z/\gamma^* \rightarrow l^+l^-$	557205 ± 294	2947656 ± 214	487 ± 8.7	17.6 ± 1.7	0	0	0
$t\bar{t}$	5704 ± 20.9	1281 ± 9.9	93.8 ± 2.7	70.2 ± 2.3	9.0 ± 0.8	8.6 ± 0.8	0.6 ± 0.2
$W \rightarrow l\nu$	62977 ± 208	793.1 ± 23.5	301 ± 14.5	29.8 ± 4.7	1.9 ± 1.3	1.9 ± 1.3	0
Single top	855.8 ± 3.8	103.2 ± 1.2	12.4 ± 0.5	5.1 ± 0.3	0.4 ± 0.1	0.4 ± 0.1	0.1 ± 0.0
Rare	7546 ± 20.7	3023 ± 11.9	104.6 ± 1.9	6.4 ± 0.4	0.9 ± 0.1	0.9 ± 0.1	0.7 ± 0.1
sum (MC)	6683061 ± 24956	302432 ± 749	1920 ± 304.5	169.0 ± 33.9	12.3 ± 2.3	11.9 ± 2.3	1.4 ± 0.3
Signal (LM1)	306.2 ± 8.2	114.0 ± 5.0	13.2 ± 1.7	10.1 ± 1.5	8.9 ± 1.4	8.9 ± 1.4	7.9 ± 1.3
Signal (LM3)	182.6 ± 6.7	49.2 ± 3.6	9.4 ± 1.6	9.0 ± 1.6	7.8 ± 1.5	7.7 ± 1.5	5.9 ± 1.3
Signal (LM8)	63.1 ± 3.5	21.1 ± 2.0	4.7 ± 1.0	4.2 ± 0.9	3.1 ± 0.8	3.1 ± 0.8	1.6 ± 0.6
Data	13071565 ± 3615	342535 ± 585	4703 ± 69	782 ± 28	10 ± 3	10 ± 3	2 ± 1

Table B.4: Summary of the event yield after the last selection step for the three channels and its combination. The event yield is shown for the sum of the simulated samples, some SUSY benchmark points, all SM background predictions (prompt-fake, charge flip and irreducible background) and 1.1 fb^{-1} of data. In the last column the channels are summed up. The first error is statistical uncertainty and the second one the systematic uncertainty, respectively. Numbers are summed up before they are rounded.

sample	$\mu\mu$	$e\mu$	ee	Combined
Sum all (MC)	$2.6 \pm 0.8 \pm 0.4$	$4.1 \pm 0.5 \pm 0.9$	$1.4 \pm 0.2 \pm 0.4$	$8.1 \pm 1.0 \pm 1.0$
Signal (LM1)	$13.3 \pm 1.7 \pm 2.7$	$20.6 \pm 2.1 \pm 4.1$	$7.9 \pm 1.3 \pm 1.6$	$41.8 \pm 3.0 \pm 5.2$
Signal (LM3)	$4.2 \pm 0.9 \pm 0.8$	$9.8 \pm 1.6 \pm 2.0$	$5.9 \pm 1.3 \pm 1.2$	$19.9 \pm 2.3 \pm 2.4$
Signal (LM8)	$1.8 \pm 0.6 \pm 0.4$	$5.5 \pm 1.1 \pm 1.1$	$1.6 \pm 0.6 \pm 0.3$	$8.9 \pm 1.4 \pm 1.2$
Prompt-fake pred.	$2.9 \pm 0.6 \pm 1.5$	$4.0 \pm 0.7 \pm 2.0$	$0.6 \pm 0.2 \pm 0.3$	$7.6 \pm 0.9 \pm 2.5$
Charge Flip pred.	—	$0.3 \pm 0.0 \pm 0.1$	$0.2 \pm 0.0 \pm 0.0$	$0.5 \pm 0.0 \pm 0.1$
Irreducible pred.	$0.7 \pm 0.1 \pm 0.4$	$1.4 \pm 0.1 \pm 0.7$	$0.7 \pm 0.1 \pm 0.4$	$2.9 \pm 0.2 \pm 0.9$
Sum pred.	$3.6 \pm 0.6 \pm 1.5$	$5.7 \pm 0.7 \pm 2.1$	$1.6 \pm 0.2 \pm 0.5$	$11.0 \pm 1.0 \pm 2.7$
Data	2	3	2	7

C Details of Uncertainty Studies

In the following, details for the uncertainty studies are shown. In particular, the results for the $e\mu$ - and ee -channel are given, which are similar to the results for the $\mu\mu$ -channel shown in Chapter 7. In addition, a trigger efficiency test is documented in the following Section (referring to Section 5.3).

C.1 Trigger Efficiencies

A simple simulation test has been performed in order to quantify, if the simulated trigger efficiencies correspond to the trigger efficiencies estimated from data [90, 91], or if a scale-factor has to be applied to the simulation. The number of triggered and untriggered events containing two electrons (defined as documented in Section 5.4.2 and 5.4.1) have been counted in the simulated $t\bar{t}$ sample. In Table C.1 the result of this test is summarized, where the trigger efficiency from data is cited from detailed investigations in [90].

Table C.1: Simple test of the trigger efficiency ϵ in data and simulation. The efficiency from data is taken from [90] while the trigger efficiency from MC is calculated from the number of triggered events and the event where no trigger was required. The ratio of the efficiencies is consistent with one within the uncertainty.

Channel	Untriggered	Triggered	ϵ_{MC} [%]	ϵ_{Data} [%]	$\epsilon_{\text{Data}}/\epsilon_{\text{MC}}$
$\mu\mu$	$24,570 \pm 157$	$23,463 \pm 153$	95.5 ± 0.01	$91.7 \pm 0.4 \pm 2.6$	$0.96 \pm 0.01 \pm 0.03$
$e\mu$	$41,984 \pm 205$	$39,257 \pm 198$	93.5 ± 0.01	$92.9 \pm 0.7 \pm 2.9$	$0.99 \pm 0.01 \pm 0.02$
ee	$17,779 \pm 133$	$17,746 \pm 133$	99.8 ± 0.01	$99.4 \pm 0.2 \pm 1.5$	$1 \pm 0.01 \pm 0.03$

The ratios of the trigger efficiencies in data and simulation agree with “1” within their uncertainty. However, possible differences of the efficiencies are accounted for by the applied trigger uncertainty of 3% (see Section 7.5). Hence, no scale-factor needs to be applied to account for a difference of the trigger efficiency.

C.2 Jet Energy Scale

In Tables C.2, C.3, C.4, and C.5 the impact on the event yield due to a variation of the JES by $\pm 2\%$ is shown for the $e\mu$ - and ee -channel (in analogy to the $\mu\mu$ -channel discussed in Section 7.1).

Table C.2: JES variation for the $e\mu$ -channel and its impact on the event yield of the $t\bar{t}$ sample. “Entries” gives the number of entries in each selection step, “Norm” is the normalized event yield without a variation of the JES (normalized to $L = 1.1 \text{ fb}^{-1}$), and “Up” (“Down”) is the relative change of the event yield of the up(down)-variation with respect to Norm.

Cut	Entries	Norm	Up [%]	Down [%]
Preselection	413,935	19,480	0	0
$N_l \geq 2$	106,578	5,064	0	0
$p_T(l) > 20/10 \text{ GeV}/c$	103,644	4,924	0	0
$N_{\text{SS candidates}} = 1$	16,577	795	0	0
$N_{\text{Jets}} \geq 2$	13,109	631	1.3	−1.6
$E_T^{\text{miss}}(PF) > 100 \text{ GeV}$	1,668	78.9	6.5	−6.6
$I_{\text{Rel}}(\text{1st iso lep})$	1,461	68.8	6.7	−7
$I_{\text{Rel}}(\text{2nd iso lep})$	38	2.2	8.9	−8.7

Table C.3: Results of the $e\mu$ -channel in analogy to Table C.2. The order of the selection steps has been changed. The cuts which are sensitive to a variation of the JES are applied as soon as possible in order to obtain higher statistical power in these selection steps.

Cut	Entries	Norm	Up [%]	Down [%]
Preselection	413,935	19,183	0	0
$N_{\text{Jets}} \geq 2$	344,245	15,957	0.9	−1
$E_T^{\text{miss}}(PF) > 100 \text{ GeV}/c$	45,512	2,101	6.6	−6.5
$I_{\text{Rel}}(\text{1st iso lep})$	40,202	1,856	6.5	−6.6
$I_{\text{Rel}}(\text{2nd iso lep})$	1,374	63.7	6.3	−7.3
$N_l \geq 2$	358	17	8.3	−8
$p_T(l) > 20/10 \text{ GeV}$	344	16.4	7.4	−8.3
$N_{\text{SS candidates}} = 1$	38	2.2	9.6	−9.4

Table C.4: JES variation for the ee -channel and its impact on the event yield of the $t\bar{t}$ sample. “Entries” gives the number of entries in each selection step, “Norm” is the normalized event yield without a variation of the JES (normalized to $L = 1.1 \text{ fb}^{-1}$), and “Up” (“Down”) is the relative change of the event yield of the up(down)-variation with respect to Norm.

Cut	Entries	Norm	Up [%]	Down [%]
Preselection	119,427	5,704	0	0
$N_e \geq 2$	26,998	1,293	0	0
$p_T(e) > 20/10 \text{ GeV}/c$	26,743	1,281	0	0
$N_{\text{SS candidates}} = 1$	1,927	93.8	0	0
$N_{\text{Jets}} \geq 2$	1,420	70.2	1.3	-2.2
$E_T^{\text{miss}}(PF) > 100 \text{ GeV}$	187	9	4.9	-7.1
$I_{\text{Rel}}(\text{1st iso lep})$	179	8.6	5.1	-6.4
$I_{\text{Rel}}(\text{2nd iso lep})$	13	0.6	0	0

Table C.5: Results of the ee -channel in analogy to Table C.4. The order of the selection steps has been changed. The cuts which are sensitive to a variation of the JES are applied as soon as possible in order to obtain higher statistical power in these selection steps.

Cut	Entries	Norm	Up [%]	Down [%]
Preselection	119,427	5,616	1.2	-1.3
$N_{\text{Jets}} \geq 2$	91,478	4,303	6.7	-6.3
$E_T^{\text{miss}}(PF) > 100 \text{ GeV}/c$	14,050	659	6.8	-6.2
$I_{\text{Rel}}(\text{1st iso lep})$	13,155	616	5.9	-7.3
$I_{\text{Rel}}(\text{2nd iso lep})$	790	39.1	7.5	-8.6
$N_e \geq 2$	179	9.3	7.5	-8.6
$p_T(e) > 20/10 \text{ GeV}$	177	9.3	0	0
$N_{\text{SS candidates}} = 1$	13	0.5	0	0

C.3 Parton Distribution Functions

In Tables C.6 and C.7 the impact on the event yield due to a variation of the PDF is shown for the $e\mu$ - and ee -channel (in analogy to the $\mu\mu$ -channel discussed in Section 7.3).

Table C.6: PDF variation for the $e\mu$ -channel and its impact on the event yield of the $t\bar{t}$ sample. “Entries” gives the number of entries in each selection step, “Norm” is the normalized event yield without a variation of the JES (normalized to $L = 1.1 \text{ fb}^{-1}$), and “Up” (“Down”) is the relative change of the event yield of the up(down)-variation with respect to Norm.

Cut	Entries	Norm	Up [%]	Down [%]
Preselection	413,935	19,480	9.6	−3.6
$N_l \geq 2$	106,578	5,064	9.9	−3.5
$p_T(l) > 20/10 \text{ GeV}/c$	103,644	4,924	9.9	−3.5
$N_{\text{SS candidates}} = 1$	16,577	795	10.3	−3.4
$N_{\text{Jets}} \geq 2$	13,109	631	10.3	−3.5
$E_T^{\text{miss}}(PF) > 100 \text{ GeV}$	1,668	78.9	9.4	−4.2
$I_{\text{Rel}}(\text{1st iso lep})$	1,461	68.8	9.4	−4.3
$I_{\text{Rel}}(\text{2nd iso lep})$	38	2.2	7.1	−4.9

Table C.7: PDF variation for the ee -channel and its impact on the event yield of the $t\bar{t}$ sample. “Entries” gives the number of entries in each selection step, “Norm” is the normalized event yield without a variation of the JES (normalized to $L = 1.1 \text{ fb}^{-1}$), and “Up” (“Down”) is the relative change of the event yield of the up(down)-variation with respect to Norm.

Cut	Entries	Norm	Up [%]	Down [%]
Preselection	119,427	5,704	9.2	−3.6
$N_e \geq 2$	91,478	1,293	10.1	−3.5
$p_T(e) > 20/10 \text{ GeV}/c$	14,050	1,281	10.1	−3.5
$N_{\text{SS candidates}} = 1$	13,155	93.8	10.1	−3.4
$N_{\text{Jets}} \geq 2$	790	70.2	10.3	−3.5
$E_T^{\text{miss}}(PF) > 100 \text{ GeV}$	179	9	10.4	−4.2
$I_{\text{Rel}}(\text{1st iso lep})$	177	8.6	10.3	−4.1
$I_{\text{Rel}}(\text{2nd iso lep})$	13	0.6	11.5	−4.5

C.4 Pile-up

In Tables C.8 and C.9 the impact on the event yield due to a variation of the mean of the pile-up distribution by ± 0.6 is shown for the $e\mu$ - and ee -channel (in analogy to the $\mu\mu$ -channel discussed in Section 7.3).

Table C.8: Pile-up variation for the $e\mu$ -channel and its impact on the event yield of the $t\bar{t}$ sample. “Entries” gives the number of entries in each selection step, “Norm” is the normalized event yield without a variation of the JES (normalized to $L = 1.1 \text{ fb}^{-1}$), and “Up” (“Down”) is the relative change of the event yield of the up(down)-variation with respect to Norm.

Cut	Entries	Norm	Up [%]	Down [%]
Preselection	413,935	19,480	0.27	−0.26
$N_l \geq 2$	106,578	5,064	−0.11	0.11
$p_T(l) > 20/10 \text{ GeV}/c$	103,644	4,924	−0.12	0.12
$N_{\text{SS candidates}} = 1$	16,577	795	−0.33	0.31
$N_{\text{Jets}} \geq 2$	13,109	631	−0.55	0.55
$E_T^{\text{miss}}(PF) > 100 \text{ GeV}$	1,668	78.9	−0.69	0.78
$I_{\text{Rel}}(\text{1st iso lep})$	1,461	68.8	−0.81	0.96
$I_{\text{Rel}}(\text{2nd iso lep})$	38	2.2	−2.66	2.45

Table C.9: Pile-up variation for the ee -channel and its impact on the event yield of the $t\bar{t}$ sample. “Entries” gives the number of entries in each selection step, “Norm” is the normalized event yield without a variation of the JES (normalized to $L = 1.1 \text{ fb}^{-1}$), and “Up” (“Down”) is the relative change of the event yield of the up(down)-variation with respect to Norm.

Cut	Entries	Norm	Up [%]	Down [%]
Preselection	119,427	5,704	−0.42	0.42
$N_e \geq 2$	26,998	1,293	−0.71	0.77
$p_T(e) > 20/10 \text{ GeV}/c$	26,743	1,281	−0.71	0.77
$N_{\text{SS candidates}} = 1$	1,927	93.8	−1.78	1.86
$N_{\text{Jets}} \geq 2$	1,420	70.2	−2.2	2.22
$E_T^{\text{miss}}(PF) > 100 \text{ GeV}$	187	9	−1.03	0.70
$I_{\text{Rel}}(\text{1st iso lep})$	179	8.6	−0.96	0.65
$I_{\text{Rel}}(\text{2nd iso lep})$	13	0.6	6.24	−6.70

C.5 TL Method

In Fig. C.1 and C.2 the effect of an additional b-Tag requirement in the selection of the control samples on the prompt- and fake-ratios is shown, while in Fig. C.3 and C.4 the effect of a changed p_T -range in the fake- and prompt-ratio is depicted.

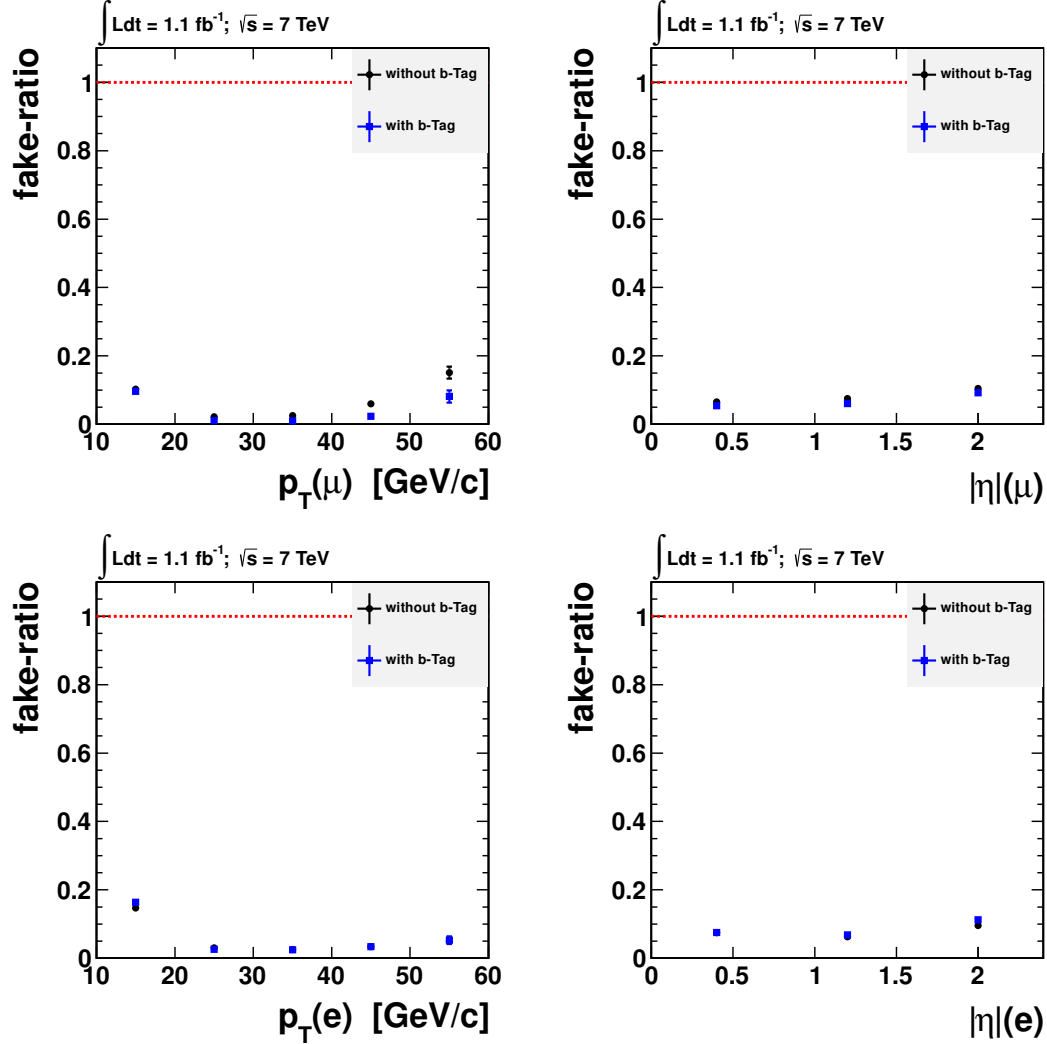


Figure C.1: Fake-ratio for muons (*top*) and electrons (*bottom*) with and without an additional b-Tag requirement as a function of lepton p_T (*left*) and η (*right*). *Black circles:* ratios estimated from data with the default selection (see Section 6.3.1), *blue squares:* with additional b-Tag.

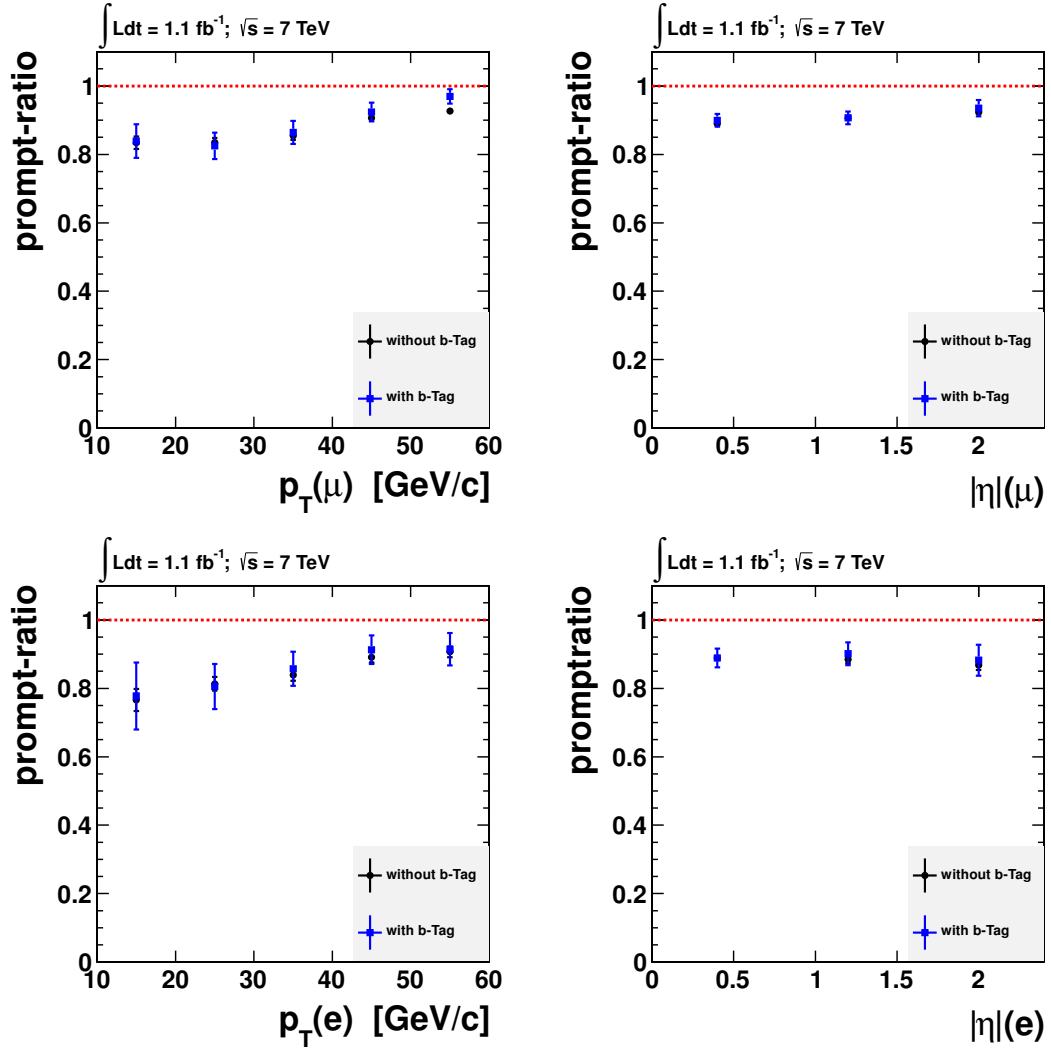


Figure C.2: Prompt-ratio for muons (*top*) and electrons (*bottom*) with and without an additional b-Tag requirement as a function of lepton p_T (*left*) and η (*right*). *Black circles:* ratios estimated from data with the default selection (see Section 6.3.1), *blue squares:* with additional b-Tag.

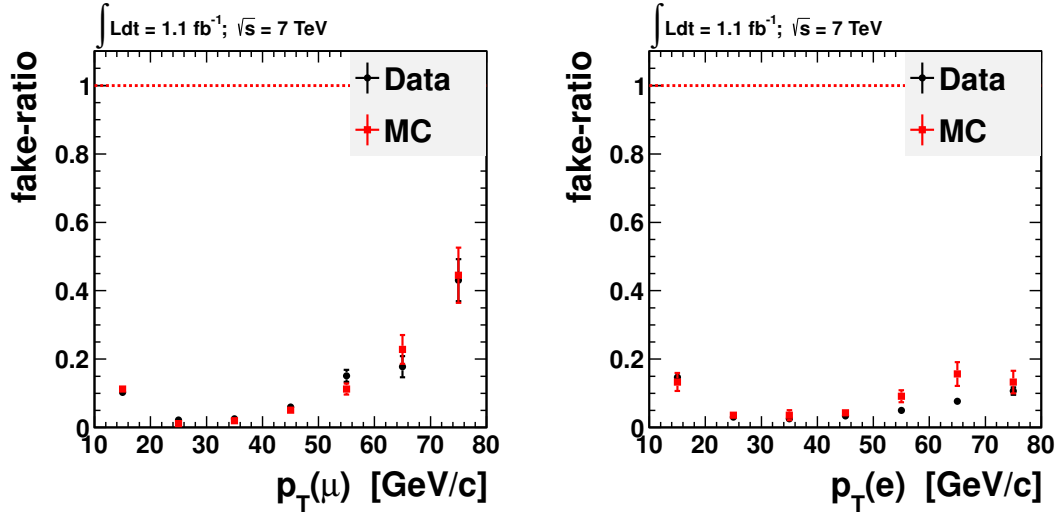


Figure C.3: Fake-ratio for muons (*left*) and electrons (*right*) with an increased range in p_T . For the investigation of the systematic uncertainty, the fake-ratio is applied up to 80 GeV or 40 GeV and compared to the default range of 60 GeV. *Black circles*: ratios estimated from data, *red squares*: ratios estimated from simulation.

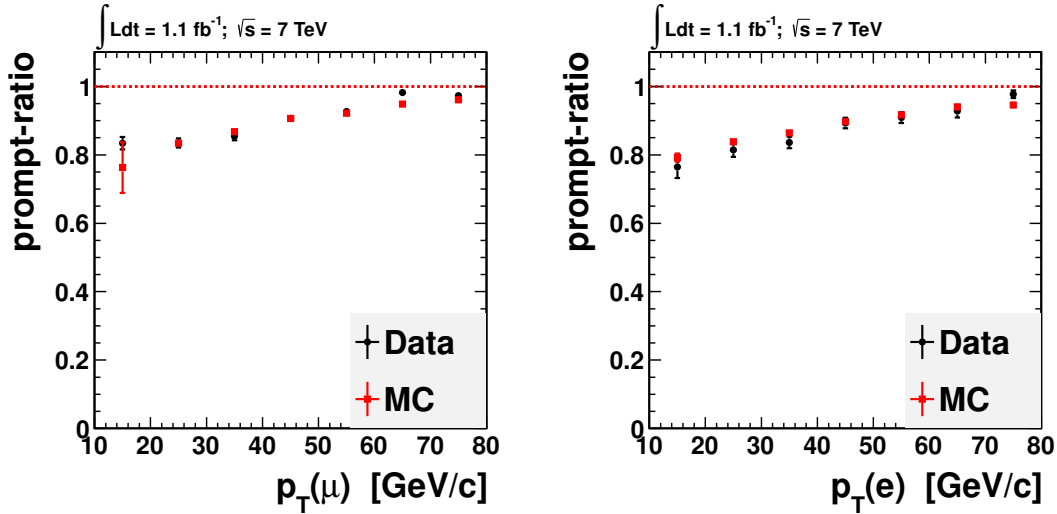


Figure C.4: Prompt-ratio for muons (*left*) and electrons (*right*) with an increased range in p_T . For the investigation of the systematic uncertainty, the prompt-ratio is applied up to 80 GeV or 40 GeV and compared to the default range of 60 GeV. *Black circles*: ratios estimated from data, *red squares*: ratios estimated from simulation.

D Examples for Rare Standard Model Processes

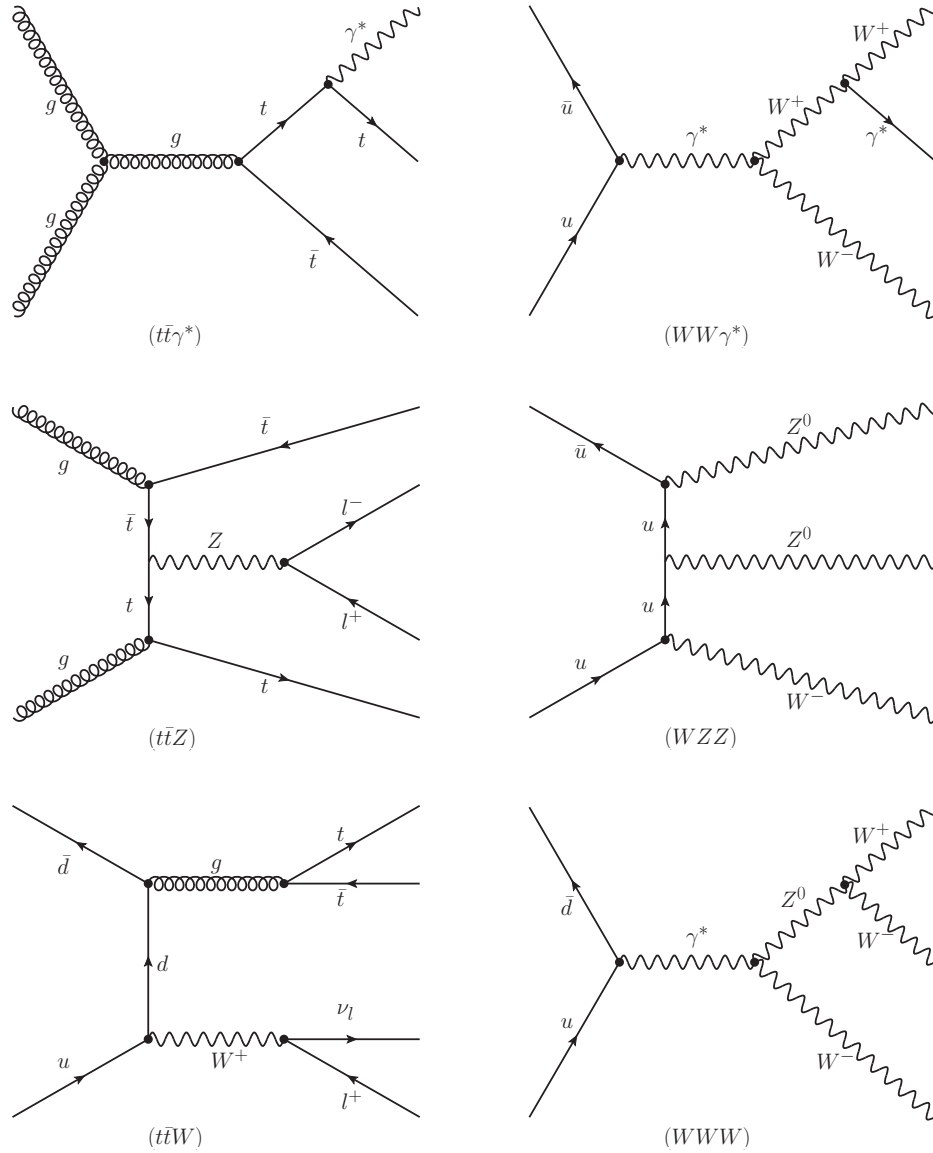


Figure D.1: Feynman diagrams of some rare SM processes leading to a prompt-prompt same-sign di-lepton signature.

E Consistency Check of the Background Prediction

Here, a consistency test of the background prediction is performed, in order to show that the background prediction is (almost) independent from a SUSY signal (this is from importance for the model limit setting, compare Section 8.2).

The irreducible background contribution does not depend on the signal, since it is taken directly from simulation. However, it is a priori not obvious if the prompt-fake background estimation (via the TL method) is also independent from a SUSY signal. In principle, a signal can influence the prompt-fake background estimation in two ways:

- Bias in the estimation of the prompt- and fake-ratio.
- Bias in the background prediction (from the measured numbers N_{TT}, N_{TL} etc.).

From Fig. 6.7 and 6.9 it can be seen, that a SUSY signal does *not* contribute to the estimation of the prompt- and fake-ratio (the contribution of the displayed SUSY signal to the control samples is at least eight orders of magnitudes smaller than the SM contribution).

A consistency check is performed in order to illustrate the performance of the TL method for the prediction of a signal. The TL method is applied to several SUSY signal benchmark points and the result is summarized in Table E.1. The prediction of prompt-prompt di-lepton sources (here: SUSY signal) agrees very well with the event yield after the signal selection, which is already a good consistency check of the method. The other sources (prompt-fake and fake-fake) are negligible: In the examples from Table E.1, the prompt-fake and fake-fake background is consistent with zero. Only for the LM3 benchmark point a small contribution to the prompt-fake background prediction is obtained.¹ However, investigations with artificially increased background estimations in the SUSY scan indicate, that such a small contribution barely affects the setting of exclusion limits.

Hence, as an approximation, the background prediction is assumed to be independent from a SUSY signal.

¹Note, that in principle, it is not forbidden that also a SUSY signal also contains sources of prompt-fake or a fake-fake same-sign di-leptons.

Table E.1: Signal event yield for several benchmark points in the m_0 - $m_{1/2}$ -plane in comparison with the prediction from the TL method. “Yield” indicates the event yield after the whole signal selection, “p-p” (“p-f” and “f-f”) refers to the prediction of events with two prompt-prompt (prompt-fake and fake-fake) leptons. The prediction is in good agreement with the simulated event yield and gives a good consistency check of the method. Numbers are summed up before they are rounded.

sample	$\mu\mu$	$e\mu$	ee	sum
LM1 (yield)	13.3 ± 1.7	20.6 ± 2.1	7.9 ± 1.3	41.8 ± 3.0
LM1 (p-p)	13.3 ± 1.8	20.5 ± 2.2	8.0 ± 1.4	41.7 ± 3.1
LM1 (p-f)	0	0.2 ± 0.1	-0.1 ± 0.1	0.1 ± 0.2
LM1 (f-f)	0	0	0	0
LM3 (yield)	4.2 ± 0.9	9.8 ± 1.6	5.9 ± 1.3	19.9 ± 2.3
LM3 (p-p)	3.9 ± 1.0	9.3 ± 1.6	5.4 ± 1.3	18.7 ± 2.3
LM3 (p-f)	0.2 ± 0.1	0.5 ± 0.1	0.2 ± 0.1	0.9 ± 0.2
LM3 (f-f)	0	0	0	0
LM8 (yield)	1.8 ± 0.6	5.5 ± 1.1	1.6 ± 0.6	8.9 ± 1.4
LM8 (p-p)	1.7 ± 0.7	5.4 ± 1.1	1.6 ± 0.6	8.7 ± 1.4
LM8 (p-f)	0.1 ± 0.1	0.1 ± 0.1	0.1 ± 0.1	0.2 ± 0.1
LM8 (f-f)	0	0	0	0

F Limit Setting Procedure

As mentioned in Section 8.2, an upper limit for the cross-section of a SUSY signal has been set in the m_0 - $m_{1/2}$ -plane at a 95 % Credibility Level (CL) utilizing Bayesian statistics. The cross-section σ_{max} (the cross-section for which the signal added to the background prediction is still in agreement with the measurement at a 95 % CL) is estimated from the probability density $p(\sigma|N)$, given the number of measured events N from data

$$0.95 = \int_0^{\sigma_{\text{UL}}^{95\%}} p(\sigma|N) d\sigma, \quad (\text{F.1})$$

where $\sigma_{\text{UL}}^{95\%}$ is the upper limit of the cross-section at a 95 % CL. The probability density is a concatenation of the signal efficiency ϵ , the luminosity L , and the data-driven background prediction b

$$p(\sigma|N) = \int p(\sigma\epsilon Lb|N) d\epsilon dL db, \quad (\text{F.2})$$

which can be rewritten using Bayes theorem

$$p(\sigma|N) = \int \frac{p(N|\sigma\epsilon Lb) \cdot p(\sigma\epsilon Lb)}{p(N)} d\epsilon dL db. \quad (\text{F.3})$$

Assuming that the quantities σ , ϵ , L , and b are independent from each other, the prior can be factorized as

$$p(\sigma\epsilon Lb) = \underbrace{p(\sigma)}_{\text{flat}} \cdot \underbrace{p(\epsilon) \cdot p(L) \cdot p(b)}_{\text{log normal}}, \quad (\text{F.4})$$

where for σ a flat distribution is assumed (since, a priori, no value for σ is favored), and for ϵ , L , and b a log-normal distribution is assumed,¹ since they are error-prone measured quantities. The probability function $p(N|\sigma\epsilon Lb)$ is given in terms of a Likelihood function $\mathcal{L}$, assuming a Poisson distribution

$$\mathcal{L} = \frac{e^{-\mu} \mu^N}{N!}, \quad (\text{F.5})$$

where N is the number of measured events from data ($N = 7$), and $\mu = b + L \cdot \sigma \cdot \epsilon$. The calculation is performed numerically using ROOSTATS [126].

¹For the luminosity e.g. 1.1 fb^{-1} is taken as mean of the distribution and the uncertainty of 4.5 % as its width.

List of Figures

2.1	Representation of the SM particles	4
2.2	LO and NLO Feynman diagrams for electron-positron scattering	6
2.3	Higgs mass as a function of the top-quark and W boson mass	8
2.4	Running behavior of the interactions as a function of energy	10
2.5	One-loop quantum corrections to the Higgs squared mass parameter	10
2.6	Some LO Feynman diagrams of SUSY particle production	14
2.7	Dominant weak production mechanisms of SUSY particles	15
2.8	Possible decay of gluinos and squarks	15
2.9	SUSY decay cascade leading to the signature of a same-sign di-lepton pair	16
3.1	Areal view of the LHC ring in the region of Geneva	17
3.2	Schematic view of the LHC	19
3.3	Schematic view of the CMS detector	21
3.4	Layout of the STS of the CMS detector	23
3.5	Lead tungstate crystal of the ECAL barrel	25
3.6	Simulation of an electromagnetic shower in lead tungstate	25
3.7	y - z -view of a quadrant of the CMS ECAL	25
3.8	y - z -view of a quadrant of the CMS HCAL	27
3.9	y - z -view of a quadrant of the CMS HCAL after a possible upgrade	27
3.10	Drawing of the CMS solenoid coil	28
3.11	y - z -view of a quadrant of the CMS MS	30
3.12	Muon momentum resolution as a function of p_T	30
4.1	Schematic illustration of the HCAL readout before and after the HCAL Phase II upgrade	34
4.2	Illustration of WFs	35
4.3	Geometry of the GEANT3 standalone simulation	36
4.4	Illustration of the investigated readout schemes of the HCAL	38
4.5	Calibration of the simulated CMS HCAL	38
4.6	Scatter plots of the simulated energies of the HCAL in single readout channels	40
4.7	Profile plots of the WFs	41
4.8	Average energy density of simulated pions in the HCAL with a “1-4-4-8” readout design	42
4.9	Example of energy distributions of simulated pions in the HCAL	43
4.10	Energy resolution and linearity before and after the energy weighting for the “1-4-4-8” readout design	44
4.11	Energy resolution and linearity before and after weighting for the “1-4-4-8” readout design including a consistency check	45

5.1	Pile-up distributions for data and simulated SM processes without a weighting procedure.	51
5.2	Pile-up distributions for data and simulated SM processes including a weighting procedure.	51
5.3	Sketch illustrating the kinematics of a leptonic SUSY decay cascade	53
5.4	Transverse slice through the CMS detector	54
5.5	Illustration of muon isolation	56
5.6	Illustration of an electron radiating a photon in the CMS detector	57
5.7	Event selection flow for the $\mu\mu$ -channel.	68
6.1	Sketch illustrating the background contribution from the semi-leptonic $t\bar{t}$ -decay	73
6.2	Illustration of the extrapolation method	75
6.3	Illustration of the jet multiplicity and lepton p_T for muons of a simulated $t\bar{t}$ and QCD sample	76
6.4	Sketch to illustrate the workflow of the b-T&P method	77
6.5	Illustration of the relative muon isolation in single kinematic bins	78
6.6	Sketch to illustrate the workflow of the TL method	79
6.7	Event selection flow of the QCD-enriched control sample	80
6.8	Fake-ratio for muons and electrons	82
6.9	Event selection flow of the Z -enriched control sample	83
6.10	Prompt-ratio for muons and electrons	84
7.1	Possible decay of gluinos and squarks	96
7.2	Simulation study for the lepton p_T in the QCD-enriched control sample . .	98
7.3	Simulation study for the lepton p_T in the Z -enriched control sample	99
8.1	Event Yield after the whole signal selection	101
8.2	Exclusion line (<i>blue</i>) in the m_0 - $m_{1/2}$ -plane of the cMSSM.	104
A.1	Shower profiles	108
A.2	Energy resolution of channel breakdown	109
A.3	Channel breakdown for the “1-4-4-8” design	110
A.4	Channel breakdown for the “1-3-3-3-4” design	111
A.5	Energy distribution with and without first layer	112
A.6	Scatter plots of weighting factors	113
B.1	Event selection flow for the $e\mu$ -channel.	115
B.2	Event selection flow for the ee -channel.	116
C.1	Fake-ratio for muons and electrons with and without an additional b-Tag requirement	126
C.2	Prompt-ratio for muons and electrons with and without an additional b-Tag requirement	127
C.3	Fake-ratio for muons and electrons with an increased range in p_T	128
C.4	Prompt-ratio for muons and electrons with an increased range in p_T	128

D.1 Feynman diagrams of some rare SM processes 129

List of Tables

2.1	The fundamental forces, their mediators, and some of their properties	5
2.2	Summary of SUSY particles in the MSSM and the extended Higgs sector .	12
2.3	“Low Mass” (LM) benchmark points of the cMSSM of the CMS collaboration	13
4.1	Single layers of the GEANT3 standalone simulation of the HCAL	37
4.2	Number of simulated pions for different energies and remaining pions after applied cuts	39
4.3	Energy resolution after weighting from Gaussian fit	44
4.4	Energy resolution after weighting from RMS	45
5.1	Simulated SM processes used in this analysis	48
5.2	QCD samples	49
5.3	Simulated rare SM processes used in this analysis	49
5.4	SUSY signal processes of the CMS benchmark points of the cMSSM	50
5.5	Used high-level triggers for the different decay channels	53
5.6	Muon selection requirements	57
5.7	Electron shower shape requirements	59
5.8	Loose electron identification requirements	59
5.9	Electron conversion rejection requirements	60
5.10	Electron selection requirements	61
5.11	Loose Jet ID requirements	64
5.12	Jet selection requirements	64
5.13	Vertex selection requirements	64
5.14	Preselection requirements	66
5.15	Event selection requirements	67
5.16	Summary of the event yield	69
6.1	Event selection requirements in order to select a QCD-enriched control sample	80
6.2	Mean of the fake-ratio for muons and electrons.	81
6.3	Event selection requirements in order to select a Z -enriched control sample	83
6.4	Mean of the prompt-ratio for muons and electrons	85
6.5	Background prediction of the TL method for the <i>prompt-fake</i> and <i>fake-fake</i> contribution to the signal region	87
6.6	Estimation of the electron charge flip rate	89
6.7	Background prediction due to an electron charge flip in the signal region .	90
6.8	Summary of the event yield of simulated irreducible SM background processes	90
7.1	Summary of the relevant uncertainties for the simulated SM processes	92

7.2	Uncertainties relevant for the background prediction	92
7.3	JES variation for the $\mu\mu$ -channel and its impact on the event yield of the $t\bar{t}$ sample	93
7.4	JES variation for the $\mu\mu$ -channel and its impact on the event yield of the $t\bar{t}$ sample with a changed order of the selection steps	93
7.5	PDF variation for the $\mu\mu$ -channel and its impact on the event yield of the $t\bar{t}$ sample	95
7.6	Pile-up variation for the $\mu\mu$ -channel and its impact on the event yield of the $t\bar{t}$ sample	97
7.7	Variation of the prompt-fake background prediction due to systematic studies of the TL method.	99
7.8	Uncertainty due to a change of the charge flip rate estimation	100
8.1	Summary of the event yield	102
A.1	Number of simulated pions which have been generated for different energies	108
B.1	Event Yield of the event selection flow for the $\mu\mu$ -channel.	117
B.2	Event Yield of the event selection flow for the $e\mu$ -channel.	118
B.3	Event Yield of the event selection flow for the ee -channel.	119
B.4	Summary of the event yield	120
C.1	Simple test of the trigger efficiency	121
C.2	JES variation for the $e\mu$ -channel and its impact on the event yield of the $t\bar{t}$ sample	122
C.3	JES variation for the $e\mu$ -channel and its impact on the event yield of the $t\bar{t}$ sample with a changed order of the selection steps	122
C.4	JES variation for the ee -channel and its impact on the event yield of the $t\bar{t}$ sample	123
C.5	JES variation for the ee -channel and its impact on the event yield of the $t\bar{t}$ sample with a changed order of the selection steps	123
C.6	PDF variation for the $e\mu$ -channel and its impact on the event yield of the $t\bar{t}$ sample	124
C.7	PDF variation for the ee -channel and its impact on the event yield of the $t\bar{t}$ sample	124
C.8	Pile-up variation for the $e\mu$ -channel and its impact on the event yield of the $t\bar{t}$ sample	125
C.9	Pile-up variation for the ee -channel and its impact on the event yield of the $t\bar{t}$ sample	125
E.1	Signal event yield for several benchmark points in the m_0 - $m_{1/2}$ -plane in comparison with the prediction from the TL method.	132

Bibliography

- [1] ATLAS and CMS Collaboration, *Combined Standard Model Higgs Boson Searches with up to 2.3 Inverse Femtobarns of pp Collision Data at $\sqrt{s} = 7$ TeV at the LHC*, CMS-PAS-HIG-11-023, ATLAS-CONF-2011-157.
- [2] *Combination of Higgs Boson Searches with up to 4.9 fb^{-1} of pp Collisions Data Taken at a Center-of-mass Energy of 7 TeV with the ATLAS Experiment at the LHC*, Tech. Rep. ATLAS-CONF-2011-163, CERN, Geneva, (Dec, 2011).
- [3] CMS Collaboration, S. Chatrchyan et al., *Combined Results of Searches for the Standard Model Higgs Boson in pp Collisions at $\sqrt{s} = 7$ TeV*, Phys. Lett. **B710** (2012) 26-48, [arXiv:1202.1488](#).
- [4] H.-C. Cheng, K. T. Matchev, and M. Schmaltz, *Bosonic Supersymmetry? Getting Fooled at the CERN LHC*, <http://link.aps.org/doi/10.1103/PhysRevD.66.056006>, Phys. Rev. D (Sep, 2002) 056006.
- [5] S. Jung, H. Murayama, A. Pierce, and J. D. Wells, *Top Quark Forward-Backward Asymmetry from New t-Channel Physics*, Phys. Rev. **D81** (2010) 015004, [arXiv:0907.4112](#).
- [6] E. L. Berger, Q.-H. Cao, C.-R. Chen, C. S. Li, and H. Zhang, *Top Quark Forward-Backward Asymmetry and Same-Sign Top Quark Pairs*, Phys. Rev. Lett. **106** (2011) 201801, [arXiv:1101.5625](#).
- [7] R. Contino and G. Servant, *Discovering the Top Partners at the LHC Using Same-Sign Dilepton Final States*, JHEP **06** (2008) 026, [arXiv:0801.1679](#).
- [8] Almeida, F. M. L. , Jr. and Coutinho, Yara Do Amaral and Martins Simoes, Jose Antonio and Queiroz Filho, P. P. and Porto, C. M., *Same-Sign Dileptons as a Signature for Heavy Majorana Neutrinos in Hadron-Hadron Collisions*, Phys. Lett. **B400** (1997) 331-334, [arXiv:hep-ph/9703441](#).
- [9] Y. Bai and Z. Han, *Top-Antitop and Top-Top Resonances in the Dilepton Channel at the CERN LHC*, JHEP **04** (2009) 056, [arXiv:0809.4487](#).
- [10] S. L. Glashow, *Partial Symmetries of Weak Interactions*, Nucl. Phys. **22** (1961) 579-588.
- [11] S. Weinberg, *A Model of Leptons*, Phys. Rev. Lett. **19** (1967) 1264-1266.
- [12] D. H. Perkins, *Introduction to High-Energy Physics; 4th Ed.*, ISBN-9780521621960.

- [13] P. Jordan and E. P. Wigner, *About the Pauli Exclusion Principle*, Zeit. Phys. **47** (1928) 631-651.
- [14] Wikipedia, *Standard Model*, <http://en.wikipedia.org/wiki/>, (Nov, 2011).
- [15] T. Aoyama, M. Hayakawa, T. Kinoshita, and M. Nio, *Revised Value of the Eighth-Order QED Contribution to the Anomalous Magnetic Moment of the Electron*, Phys. Rev. **D77** (2008) 053012, [arXiv:0712.2607](#).
- [16] D. Hanneke, S. F. Hoogerheide, and G. Gabrielse, *Cavity Control of a Single-Electron Quantum Cyclotron: Measuring the Electron Magnetic Moment*, (2010), [arXiv:1009.4831](#).
- [17] Particle Data Group, K. Nakamura et al., *Review of Particle Physics*, J. Phys. G **G37** (2010) 075021.
- [18] C. Wu, E. Ambler, R. Hayward, D. Hoppes, and R. Hudson, *Experimental Test of Parity Conservation in Beta Decay*, Phys. Rev. **105** (1957) 1413-1414.
- [19] Gargamelle Neutrino Collaboration, F. Hasert et al., *Observation of Neutrino Like Interactions Without Muon or Electron in the Gargamelle Neutrino Experiment*, Phys. Lett. **B46** (1973) 138-140.
- [20] H. Shibuya, K. Hoshino, M. Komatsu, K. Niwa, S. Buontempo, et al., *The OPERA Emulsion Detector for a Long Baseline Neutrino Oscillation Experiment*, (1997), Letter of intent.
- [21] B. Cleveland, T. Daily, J. Davis, Raymond, J. R. Distel, K. Lande, et al., *Measurement of the Solar Electron Neutrino Flux with the Homestake Chlorine Detector*, Astrophys. J. **496** (1998) 505-526.
- [22] K. Scholberg, *Neutrino Mass and Oscillations*, J. Phys. Conf. Ser. **312** (2011) 072002.
- [23] TEVNPH (Tevatron New Phenomina and Higgs Working Group), *Combined CDF and D0 Search for Standard Model Higgs Boson Production with up to 10.0 fb^{-1} of Data*, [arXiv:1203.3774](#).
- [24] CDF Collaboration, *Effect on Global Electroweak Fit and SM Higgs Boson*, http://www-cdf.fnal.gov/physics/new/top/public_mass.html, (Nov, 2011).
- [25] Gerard't Hooft (ed.) et al., *Recent Developments in Gauge Theories. Proceedings, Nato Advanced Study Institute, Cargese, France, August 26 - September 8, 1979*, NATO Adv. Study Inst. Ser. B Phys. **59** (1980) 1-438.
- [26] G. F. Hinshaw, *What is the universe made of?*, http://map.gsfc.nasa.gov/universe/uni_matter.html, Universe 101. NASA website (Jan, 2010).

- [27] W. de Boer, *University Personal Web Page*,
<http://www-ekp.physik.uni-karlsruhe.de/deboer/html/Forschung/>, (Nov, 2011).
- [28] S. P. Martin, *A Supersymmetry Primer*, **arXiv:hep-ph/9709356**.
- [29] M. Battaglia et al., *Updated Post-WMAP Benchmarks for Supersymmetry*, Eur. Phys. J. **C33** (2004) 273-296, **arXiv:hep-ph/0306219**.
- [30] H. Nishino et al., *Search for Proton Decay via $p \rightarrow e^+\pi^0$ and $p \rightarrow \mu^+\pi^0$ in a Large Water Cherenkov Detector*, Physical Review Letters **102(14)** (Apr, 2009) 141801, **arXiv:0903.0676**.
- [31] R. M. Barnett, J. F. Gunion, and H. E. Haber, *Discovering Supersymmetry with Like Sign Dileptons*, Phys. Lett. **B315** (1993) 349-354, **arXiv:hep-ph/9306204**.
- [32] M. Guchait and D. P. Roy, *Like-Sign Dilepton Signature for Gluino Production at the CERN LHC Including Top Quark and Higgs Boson Effects*,
<http://link.aps.org/doi/10.1103/PhysRevD.52.133>, Phys. Rev. D **52** (Jul, 1995) 133–141.
- [33] H. Baer, C.-h. Chen, F. Paige, and X. Tata, *Signals for Minimal Supergravity at the CERN Large Hadron Collider. II. Multilepton Channels*,
<http://link.aps.org/doi/10.1103/PhysRevD.53.6241>, Phys. Rev. D **53** (Jun, 1996) 6241–6264.
- [34] L. Evans and P. Bryant (editors), *LHC Machine*, JINST **3** (2008) S08001.
- [35] O. S. Bruning (ed.) et al., *LHC Design Report. Vol. 1: The LHC Main Ring*, CERN-2004-003-V-1.
- [36] O. S. Bruning (ed.) et al., *LHC Design Report. Vol. 2: The LHC Infrastructure and General Services*, CERN-2004-003-V-2.
- [37] M. Benedikt, P. Collier, V. Mertens, J. Poole, and K. e. Schindl (editors), *LHC Design Report. 3. The LHC Injector Chain*, CERN-2004-003-V-3.
- [38] Science Discovery, *Large Hadron Collider*,
<http://science.discovery.com/slideshows/lhc/large-hadron-collider.html>, (Nov, 2011).
- [39] LHC Outreach, *Standard Model*,
<http://lhc-machine-outreach.web.cern.ch/lhc-machine-outreach/images/>, (Nov, 2011).
- [40] ATLAS Collaboration, G. Aad et al., *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08003.
- [41] ATLAS Collaboration, *ATLAS Detector and Physics Performance Technical Design Report – Volume I*, (1999), CERN/LHCC 99-14.

- [42] ATLAS Collaboration, *ATLAS Detector and Physics Performance Technical Design Report – Volume II*, (1999), CERN/LHCC 99-15.
- [43] ALICE Collaboration, K. Aamodt et al., *The ALICE Experiment at the CERN LHC*, JINST **3** (2008) S08002.
- [44] *ALICE TDR Technical Design Report*, (Oct, 2001), CERN/LHCC 2001-021.
- [45] LHCb Collaboration, A. Alves et al., *The LHCb Detector at the LHC*, JINST **3** (2008) S08005.
- [46] *LHCb: Technical Proposal*, CERN, Geneva, (1998), CERN/LHCC 98-004.
- [47] LHCf Collaboration, O. Adriani et al., *The LHCf Detector at the CERN Large Hadron Collider*, JINST **3** (2008) S08006.
- [48] *LHCf – Technical Design Report of the LHCf Experiment*, (Feb, 2006), CERN/LHCC 2006-004.
- [49] TOTEM Collaboration, G. Anelli et al., *The TOTEM Experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08007.
- [50] *TOTEM – Technical Design Report*, TOTEM-TDR-001, (Jan, 2004), ISBN 92-9083-219-3, CERN/LHCC 2004-02.
- [51] M. Oriunno et al., *The Roman Pot for LHC*, Prepared for European Particle Accelerator Conference (EPAC 06), Edinburgh, Scotland, 26-30 June 2006.
- [52] CMS Collaboration, M. Della Negra et al., *CMS: The Compact Muon Solenoid: Letter of Intent for a General Purpose Detector at the LHC*, CERN-LHCC-92-03.
- [53] CMS Collaboration, R. Adolphi et al., *The CMS Experiment at the CERN LHC*, JINST **0803** (2008) S08004.
- [54] CMS Collaboration, G. Bayatian et al., *CMS Physics Technical Design Report Volume I: Detector Performance and Software*, CERN/LHCC **2006-001** (2006).
- [55] CMS Collaboration, G. Bayatian et al., *CMS Technical Design Report, Volume II: Physics Performance*, J. Phys. G **G34** (2007) 995-1579.
- [56] V. Andreev et al., *Performance Studies of a Full-Length Prototype for the CASTOR Forward Calorimeter at the CMS Experiment*, Eur. Phys. J. **67** (2010) 601-615.
- [57] CMS-CASTOR Collaboration, P. Gottlicher, *Design and Test Beam Studies for the CASTOR Calorimeter of the CMS Experiment*, Nucl. Instrum. Meth. **A623** (2010) 225-227.
- [58] CMS Collaboration, O. A. Grachov et al., *Status of Zero Degree Calorimeter for CMS Experiment*, AIP Conf. Proc. **867** (2006) 258-265, [arXiv:nuc1-ex/0608052](https://arxiv.org/abs/nuc1-ex/0608052).

- [59] CMS Collaboration, *CMS, Tracker Technical Design Report*, 15 April 1998.
- [60] CMS Collaboration, *Measurement of Momentum Scale and Resolution of the CMS Detector using Low-mass Resonances and Cosmic Ray Muons*, CMS Physics Analysis Summary (2010), CMS-PAS-TRK-10-004.
- [61] CMS Collaboration, *Tracking and Primary Vertex Results in First 7 TeV Collisions*, CMS Physics Analysis Summary (2010), CMS-PAS-TRK-10-005.
- [62] CMS Collaboration, S. Chatrchyan et al., *Alignment of the CMS Silicon Tracker during Commissioning with Cosmic Rays*, JINST **5** (2010) T03009, [arXiv:0910.2505](https://arxiv.org/abs/0910.2505).
- [63] *The CMS Electromagnetic Calorimeter Project: Technical Design Report*, CERN, Geneva, (1997).
- [64] D. Berry, *Notre Dame extended Research Community*, http://farm3.static.flickr.com/2482/3742463006_ba6dce47c9.jpg, (Nov, 2011).
- [65] *Technology for Particle Physics*, http://hepwww.rl.ac.uk/OpenDays97/CMS_ECAL_Shower.htm, (Nov, 2011).
- [66] *The CMS Hadron Calorimeter Project: Technical Design Report*, CERN, Geneva, (1997).
- [67] A. Bodek, *Performance of a Prototype CMS Hadron Barrel Calorimeter in a Test Beam*, IEEE Trans. Nucl. Sci. **46** (1999) 407-409, [arXiv:hep-ex/9810033](https://arxiv.org/abs/hep-ex/9810033).
- [68] CMS Collaboration, *Technical Proposal for the Upgrade of the CMS Detector Through 2020*, (CERN-LHCC-2011-006. LHCC-P-004) (Jun, 2011).
- [69] CMS Collaboration, G. Acquistapace et al., *CMS, the Magnet Project: Technical Design Report*, (1997).
- [70] Wikipedia, *CMS Solenoid*, <http://it.wikipedia.org/wiki/>, (Nov, 2011).
- [71] CMS Collaboration, S. Dasu et al., *CMS. The TriDAS Project. Technical Design Report, Vol. 1: The Trigger Systems*, (2000).
- [72] CMS Collaboration, P. Sphicas (editor), *CMS: The TriDAS project. Technical Design Report, Vol. 2: Data Acquisition and High-Level Trigger*, (2002).
- [73] V. Andreev, K. Borras, D. Krucker, I. Melzer-Pellmann, M. Stein, et al., *Energy Weighting for the Upgrade of the Hadronic Calorimeter of CMS*, Proceedings of the Lepton Photon 2009 Conference (LP09), Hamburg, Germany, 17-22, Aug. 2009, (2010) 463-465.
- [74] To be published: K. Borras, D. Krücker, I. Melzer-Pellmann, M. Stein, and P. Schleper, *Energy Weighting for the Upgrade of the CMS HCAL*, Physics of Atomic Nuclei **75**, Issue 9 (Sep, 2012).

- [75] C. Issever, K. Borras, and D. Wegener, *An Improved Weighting Algorithm to Achieve Software Compensation in a Fine Grained LAr Calorimeter*, Nucl. Instrum. Meth. **A545:803-812** (2005), DESY-04-127, Aug 2004., 18pp.
- [76] P. Loch, *Calibration of the H1 Liquid Argon Calorimeter Using the Signal Weighting Method for Particle Jets*, DESY-FH1K-92-02.
- [77] M. Goossens, *GEANT: Detector Description and Simulation Tool*, long writeup W5013; March 1994, CERN Program Library, CERN, Geneva, (1993).
- [78] CMS-HCAL, V. V. Abramov et al., *Studies of the Response of the Prototype CMS Hadron Calorimeter, Including Magnetic Field Effects, to Pion, Electron, and Muon Beams*, Nucl. Instrum. Meth. **A457** (2001) 75-100, [arXiv:hep-ex/0007045](#).
- [79] S. Banerjee, *Material Budget in the Hadron Calorimeter*, (2009), CMS-IN 2009/004.
- [80] J. B. Birks, *The Theory and Practice of Scintillation Counting*, International series of Monographs on Electronics and Instrumentation, v. 27 Macmillan, New York.
- [81] H. Saka and C. Tully, *CMS Barrel Hadron Calorimeter Test Beam 2009 Analysis with Layer Weighting Technique*, (2010), CMS-IN 2010/016.
- [82] J. Alwall et al., *MadGraph/MadEvent v4: The New Web Generation*, JHEP **09** (2007) 028, [arXiv:0706.2334](#).
- [83] T. Sjostrand, S. Mrenna, and P. Z. Skands, *PYTHIA 6.4 Physics and Manual*, JHEP **05** (2006) 026, [arXiv:hep-ph/0603175](#).
- [84] M. L. Mangano, M. Moretti, F. Piccinini, and M. Treccani, *Matching Matrix Elements and Shower Evolution for Top- Quark Production in Hadronic Collisions*, JHEP **01** (2007) 013, [arXiv:hep-ph/0611129](#).
- [85] J. Allison et al., *Geant4 Developments and Applications*, IEEE Trans. Nucl. Sci. **53** (2006) 270.
- [86] J. M. Campbell and R. K. Ellis, *MCFM for the Tevatron and the LHC*, Nucl. Phys. Proc. Suppl. **205-206** (2010) 10-15, [arXiv:1007.3492](#).
- [87] K. Melnikov and F. Petriello, *Electroweak Gauge Boson Production at Hadron Colliders Through $\mathcal{O}(\alpha(s)^2)$* , Phys. Rev. **D74** (2006) 114017, [arXiv:hep-ph/0609070](#).
- [88] CMS Collaboration, S. Chatrchyan et al., *Search for new Physics with Same-Sign Isolated Dilepton Events with Jets and Missing Transverse Energy at the LHC*, JHEP **1106** (2011) 077, [arXiv:1104.3168](#).
- [89] CMS Collaboration, *Search for New Physics with Same-Sign Isolated Dilepton Events with Jets and Missing Energy*, (2011), CMS-PAS-SUS-11-010.

- [90] CMS Collaboration, *Search for New Physics with Same-Sign Di-leptons, Jets, and Missing Transverse Energy with 4.7 fb^{-1} of Data*, (Dec, 2011), CMS AN-2011/437.
- [91] CMS Collaboration, *Multileptonic SUSY Searches with Jets and MET*, (Feb, 2012), CMS AN-2011/288.
- [92] Wikipedia, *Compact Muon Solenoid*,
http://en.wikipedia.org/wiki/File:CMS_Slice.gif, (Nov, 2011).
- [93] H. Bethe, *Zur Theorie des Durchgangs Schneller Korpuskularstrahlen Durch Materie*, Annalen der Physik **397** (1930) 325-400.
- [94] H. Bichsel, D. Groom, and S. Klein, *Passage of Particles Through Matter*, Phys. Lett. (2008) 267-280.
- [95] CMS Collaboration, *Performance of Muon Identification in pp Collisions at $\sqrt{s} = 7\text{ TeV}$* , (2010), CMS-PAS-MUO-10-002.
- [96] CMS Collaboration, *Performance of muon identification in pp collisions at $\sqrt{s} = 7\text{ TeV}$* , (2010), CMS-PAS-MUO-10-002.
- [97] CMS Collaboration, *Search for New Physics with Same-Sign Dileptons Using the 2011 Dataset of CMS*, (Dec, 2011), CMS AN-2011/468.
- [98] CMS Collaboration, *Electron Reconstruction within the Particle Flow Algorithm*, (Feb, 2010), CMS AN-2010/034.
- [99] CMS Collaboration, *Electron Reconstruction and Identification at $\sqrt{s} = 7\text{ TeV}$* , (2010), CMS-PAS-EGM-10-004.
- [100] S. Cucciarelli, *Track and Vertex Reconstruction with the CMS Detector at LHC*, Nucl. Phys. Proc. Suppl. **156** (2006) 207-212.
- [101] S. Baffioni, C. Charlot, F. Ferri, D. Futyan, P. Meridiani, et al., *Electron reconstruction in CMS*, Eur. Phys. J. **C49** (2007) 1099-1116.
- [102] CMS Collaboration, *Electron Reconstruction and Identification at $\sqrt{s} = 7\text{ TeV}$* , (2010), CMS-PAS-EGM-10-004.
- [103] CMS Collaboration, *Performance of Jet Algorithms in CMS*, (2009), CMS-PAS-JME-07-003.
- [104] G. P. Salam and G. Soyez, *A Practical Seedless Infrared-Safe Cone Jet Algorithm*, JHEP **0705** (2007) 086, [arXiv:0704.0292](#).
- [105] CMS Collaboration, *Performance of Jet Algorithms in CMS*, (Mar, 2007), CMS-PAS-JME-07-003.
- [106] S. D. Ellis and D. E. Soper, *Successive Combination Jet Algorithm for Hadron Collisions*, Phys. Rev. **D48** (1993) 3160-3166, [arXiv:hep-ph/9305266](#).

- [107] M. Cacciari, G. P. Salam, and G. Soyez, *The Anti- $k(t)$ Jet Clustering Algorithm*, JHEP **0804** (2008) 063, [arXiv:0802.1189](#).
- [108] CMS Collaboration, *Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and MET*, (Apr, 2009), CMS-PAS-PFT-09-001.
- [109] CMS Collaboration, *Jet Energy Corrections Determination at 7 TeV*, (2010), CMS-PAS-JME-10-010.
- [110] CMS Collaboration, *Jets in 0.9 and 2.36 TeV pp Collisions*, (2010), CMS-PAS-JME-10-001.
- [111] CMS Collaboration, *Calorimeter Jet Quality Criteria for the First CMS Collision Data*, CMS Physics Analysis Summary (Apr, 2009), CMS-PAS-JME-09-008.
- [112] R. Fruhwirth, W. Waltenberger, and P. Vanlaer, *Adaptive Vertex Fitting*, J. Phys. **G34** (2007) N343.
- [113] S. Chatrchyan, V. Khachatryan, and Sirunyan, *Missing Transverse Energy Performance of the CMS Detector*, (Jun, 2011).
- [114] CMS Collaboration, *MET Performance in the 2011 Data*, (Jul, 2011), CMS-DP-2011-010.
- [115] D. Dobur, *Jets and Missing Transverse Energy Reconstruction with CMS*, (2009), [arXiv:0904.0391](#).
- [116] CMS Collaboration, *Performance of Methods for Data-Driven Background Estimation in SUSY Searches*, (2010), CMS-PAS-SUS-10-001.
- [117] CMS Collaboration, V. Khachatryan et al., *First Measurement of the Cross Section for Top-Quark Pair Production in Proton-Proton Collisions at $\sqrt{s} = 7$ TeV*, Phys. Lett. **B695** (2011) 424-443, [arXiv:1010.5994](#).
- [118] CMS Collaboration, *Search for New Physics in the Same-Sign Di-Lepton Channel Using 4.7 fb^{-1} of 7 TeV pp Collisions*, (Jan, 2012), CMS AN-2011/466.
- [119] M. Weinberg, *Search for New Physics with Same-Sign Isolated Dilepton Events with Jets and Missing Transverse Energy at CMS*, [arXiv:1110.2640](#).
- [120] CMS Collaboration, S. Chatrchyan et al., *Determination of Jet Energy Calibration and Transverse Momentum Resolution in CMS*, JINST **6** (2011) P11002, [arXiv:1107.4277](#).
- [121] P. M. Nadolsky et al., *Implications of CTEQ global analysis for collider observables*, Phys. Rev. **D78** (2008) 013004, [arXiv:0802.0007](#).
- [122] G. Watt, *MSTW PDFs and Impact of PDFs on Cross Sections at Tevatron and LHC*, [arXiv:1201.1295](#).

-
- [123] M. Ubiali et al., *Combined PDF and Strong Coupling Uncertainties at the LHC with NNPDF2.0*, [arXiv:1005.0397](#).
 - [124] Van der Meer, S, *Calibration of the Effective Beam Height in the ISR.*, (1968), CERN-ISR-PO-68-31. ISR-PO-68-31.
 - [125] CMS Collaboration, *Measurement of CMS Luminosity*, (2010), CMS-PAS-EWK-10-004.
 - [126] L. Moneta, K. Belasco, K. S. Cranmer, S. Kreiss, A. Lazzaro, et al., *The RooStats Project*, PoS **ACAT2010** (2010) 057, [arXiv:1009.1003](#), 11 pages, 3 figures, ACAT2010 Conference Proceedings.

Acknowledgements

I wish to acknowledge the contribution of a number of people without whose assistance and support the completion of this PhD thesis would not have been possible.

First of all, I would like to thank Dr. Isabell Melzer-Pellmann who has been my adviser during my whole thesis and the accompanying studies at DESY. She always had an open ear and supervised me in a constructive manner.

I gratefully thank Prof. Peter Schleper for supervising this thesis. I also want to express my sincere gratitude to Dr. Kerstin Borrás for her support of the technical part of my studies.

Dirk Krückers' very useful advice concerning all technical issues is gratefully acknowledged. I would like to thank Elias Ron who never gave up struggling through a nerve-stretching synchronization effort. With him and Dean Horton I had very fruitful discussions concerning hidden details of my analysis. I also want to thank Altan Cakir for his valuable criticism during group meetings and his naive way of questioning.

I wish to thank Hannes Schettler from whom I learned many tricks in programming and who supported me by providing simulation samples. He also had good suggestions for my work. Special thanks go to Niklas Pietsch and many colleagues from the University of Hamburg (Holger Enderle, Matthias Schröder, Gordon Kaussen, etc.), as well as from the theory department of DESY (Sebastian Schmidt and Daniel Wiesler) who increased my productivity by performing various sports with me.

I cannot finish without extending my heartfelt gratitude to my family for their support. My parents made it possible for me to study physics and helped me in every conceivable way. I particularly want to thank my brother Alexander who brought me back to reality, whenever it was necessary.

Last but certainly not least, I want to thank Marion Rojahn who refreshed my mind with her unique cheekiness. She cheered me up with her peerless mischievousness when I needed it most and put my work into the context of real life.

