Search for Supersymmetry in Opposite-sign Dilepton Final States with the CMS Experiment

Von der Fakultät für Mathematik, Informatik und Naturwissenschaften der RWTH Aachen University zur Erlangung des akademischen Grades eines Doktors der Naturwissenschaften genehmigte Dissertation

vorgelegt von

Diplom-Physiker Daniel Sprenger
aus Leverkusen

Berichter:

Prof. Dr. Lutz Feld
Prof. Dr. Michael Krämer


Diese Dissertation ist auf den Internetseiten der Hochschulbibliothek online verfügbar.
Abstract

For centuries, men’s curiosity drove him to explore the secrets of nature and the essential principles and constituents of matter. Today, the Standard Model of particle physics successfully describes a significant part of the known universe and most experimental observations. Still several experimental and theoretical findings indicate that the Standard Model is only valid up to an energy scale of a few TeV.

The Large Hadron Collider is, together with its detector experiments, the largest and most sophisticated experiment ever constructed and conducted in the history of mankind. It is designed to collide protons at a centre-of-mass energy of up to 14 TeV, and its purpose is to test the current Standard Model of particle physics and search for physics beyond this model. One promising new-physics theory is Supersymmetry, which extends the Standard Model by supersymmetric partner particles for the particles in the Standard Model.

This work uses proton-proton collision data at a centre-of-mass energy of 7 TeV taken by the Compact Muon Solenoid experiment at the Large Hadron Collider to search for signs of Supersymmetry. A special decay mode of a neutralino particle is targeted, which produces two flavour-correlated leptons and a stable supersymmetric particle that escapes the experiment undetected. The decay mode results in a characteristic edge in the invariant-mass distribution of the two produced leptons. This characteristic mass edge is used as a search criterion for physics beyond the Standard Model, and a way to separate this edge from background produced by Standard Model processes is presented. Separation is achieved by performing a fit to the invariant-mass distribution of events with two leptons. The fit shape is constructed using a background model, which is determined from data, and a signal hypothesis based on the mass edge that is produced by the target decay. This approach does not rely on simulated data, all backgrounds are estimated directly from data.

The search method is tested on simulated data with and without a hypothetical signal, resulting in a good performance. Afterwards, the method is performed on data. The dataset, recorded in the year 2011, corresponds to an integrated luminosity of 4.98 fb$^{-1}$. No signs of physics beyond the Standard Model are found. Therefore, exclusion limits on non-Standard-Model physics are calculated. These limits are interpreted in the context of a simplified model that incorporates the targeted decay and the constrained minimal supersymmetric extension of the Standard Model.

This work is structured in the following way: an introduction into the theoretical foundations on which this work is based is given in Chapter 1. Chapter 2 describes the Compact Muon Solenoid experiment, whose data serves as input to this analysis. In the following chapter, Chapter 3, the data taking and production of simulated data is described as well as the procedure for reconstruction of usable physics objects from data and simulated data. In Chapter 4, the search method is detailed. This includes the chosen event selection, a description of the expected Standard Model backgrounds, methods to estimate the various background contributions, the performance of the method on simulated data and finally the search results on actual data. These results are interpreted in Chapter 5.
no indication of a signal is found, exclusion limits on theoretical models are estimated. Chapter 6 concludes this work with a summary.
# Contents

## Title

### i

## Abstract

### iii

## 1 Theoretical Foundations

### 1.1 Standard Model of Particle Physics

#### 1.1.1 Natural Units

### 1.2 Reasons for Dissatisfaction

#### 1.2.1 Hierarchy Problem

#### 1.2.2 Unification of Couplings

#### 1.2.3 Astrophysical Observations

### 1.3 Supersymmetry

#### 1.3.1 Minimal Supersymmetric Extension of the Standard Model

#### 1.3.2 Symmetry Breaking

#### 1.3.3 Constrained MSSM

### 1.4 Dileptonic Final States

#### 1.4.1 Sparticle Production and Decay

#### 1.4.2 Kinematic Edges

### 1.5 Benchmark Scenarios

#### 1.5.1 Mass Spectra

#### 1.5.2 Cross Sections and Branching Ratios

#### 1.5.3 Invariant-Mass Distributions

### 1.6 Simplified Models

#### 1.6.1 T3lh Model

## 2 Experimental Setup

### 2.1 Large Hadron Collider

### 2.2 Compact Muon Solenoid

#### 2.2.1 Coordinate System

#### 2.2.2 Silicon Tracker

#### 2.2.3 Electromagnetic Calorimeter

#### 2.2.4 Hadronic Calorimeter

#### 2.2.5 Muon Chambers

#### 2.2.6 Data Acquisition and Computing

### 19

### 20

### 21

### 22

### 23

### 24

### 25
3 Data Analysis 29
  3.1 Datasets . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 29
  3.2 Monte Carlo Simulation . . . . . . . . . . . . . . . . . . . . . . . . . . . . 30
    3.2.1 Hard Scattering and Parton Distribution Functions . . . . . . . . . . 30
    3.2.2 Hadronisation . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 31
    3.2.3 Detector Simulation . . . . . . . . . . . . . . . . . . . . . . . . . . . 32
    3.2.4 Fast Simulation . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 32
    3.2.5 Pile-up Simulation . . . . . . . . . . . . . . . . . . . . . . . . . . . . 33
    3.2.6 Background Samples . . . . . . . . . . . . . . . . . . . . . . . . . . . 33
    3.2.7 Signal Samples . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 33
    3.2.8 Scaling . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 34
  3.3 Software Versions . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 34
  3.4 Object Selection . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 35
    3.4.1 Muons . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 35
    3.4.2 Electrons . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 36
    3.4.3 Taus . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 37
    3.4.4 Jets . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 38
      Jet Finding with the anti-$k_t$ Algorithm . . . . . . . . . . . . . . . . . 38
      Jet Selection . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 39
      Jet Energy Correction and Uncertainty . . . . . . . . . . . . . . . . . . 39
      Jet Tagging . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 40
    3.4.5 Transverse Hadronic Activity . . . . . . . . . . . . . . . . . . . . . . 41
    3.4.6 Missing Transverse Energy . . . . . . . . . . . . . . . . . . . . . . . 41
    3.4.7 Lepton Isolation . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 42
  3.5 Event Selection . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 43
    3.5.1 Rejection of non-Collision Events . . . . . . . . . . . . . . . . . . . . 43
    3.5.2 Trigger . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 43
  3.6 Pile-up . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 44
    3.6.1 Pile-up Reweighting of Monte Carlo Simulation . . . . . . . . . . . . 45
    3.6.2 Hadronic Pile-up Subtraction . . . . . . . . . . . . . . . . . . . . . . 46
    3.6.3 Pile-up Dependency of Lepton Reconstruction . . . . . . . . . . . . . 47
  3.7 Efficiencies . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 48
    3.7.1 Electrons and Muons . . . . . . . . . . . . . . . . . . . . . . . . . . . 48
    3.7.2 Lepton Efficiency Ratio . . . . . . . . . . . . . . . . . . . . . . . . . 49
    3.7.3 Trigger . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 51

4 Search for a Kinematic Edge 53
  4.1 Motivation . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 53
  4.2 Event Selection . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 53
    4.2.1 Preselection . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 53
      Important Quantities . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 54
    4.2.2 Signal Region . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 54
    4.2.3 Control Region . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 59
  4.3 Background Estimation . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 59
    4.3.1 Uncorrelated-Flavour Subtraction . . . . . . . . . . . . . . . . . . . . 61
      Consideration of Trigger Efficiencies . . . . . . . . . . . . . . . . . . . . 61
    4.3.2 Shape-Based Subtraction . . . . . . . . . . . . . . . . . . . . . . . . . 62
    4.3.3 Simultaneous Fit . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 64
4.3.4 Leptons From Misidentified Jets ............................................. 66
  Measurement of the Tight-to-Loose Ratio ..................................... 66
  Estimated Contribution to Signal Region .................................... 67

4.4 Performance on Monte Carlo Simulation ........................................ 68
  4.4.1 Adding CMSSM Benchmarks ............................................... 70
  4.4.2 T3lh Simplified Model ...................................................... 72

4.5 Results ...................................................................................... 74
  4.5.1 Control Region ..................................................................... 74
  4.5.2 Signal Region ........................................................................ 75
  4.5.3 Variation of the Signal Shape .............................................. 77

5 Interpretation of Results .................................................................. 81
  5.1 Statistical Procedure ............................................................... 81
    5.1.1 Observed Limits .............................................................. 81
    5.1.2 Expected Limit ............................................................... 82
  5.2 Constraining the Signal Yield ..................................................... 82
  5.3 Limit Calculation ...................................................................... 83
  5.4 Systematic Uncertainties ............................................................ 84
  5.5 Interpretation within the T3 lh Simplified Model ............................. 84
    5.5.1 Efficiency and Acceptance ............................................... 85
    5.5.2 Fit Efficiency .................................................................... 86
    5.5.3 Exclusion Limit .................................................................. 87
  5.6 Interpretation within the CMSSM ................................................ 87
    5.6.1 Efficiency and Acceptance ............................................... 90
    5.6.2 Fit Efficiency .................................................................... 90
    5.6.3 Exclusion .......................................................................... 92
    5.6.4 Possibilities of Improvement .............................................. 96

6 Summary ....................................................................................... 97

A Datasets ......................................................................................... 99

B CMS Physics Analysis Toolkit ......................................................... 101
  B.1 Software Packages ............................................................... 101
  B.2 Object Selection ...................................................................... 101

Bibliography .................................................................................... 103
1. Theoretical Foundations

The theoretical foundations on which this work is based are summarised in this chapter. First the current Standard Model of particle physics, and afterwards Supersymmetry including the minimal supersymmetric extension of the Standard Model and the signature that is looked for in this analysis is presented.

1.1 Standard Model of Particle Physics

The Standard Model of particle physics (SM) is a renormalizable quantum field theory that describes elementary particles and their interactions [1–5]. A brief overview on the SM is given in the following.

Particles are described by quantum fields in the SM. They are divided into three groups: leptons, quarks and mediators. As of today, six lepton flavours and six quarks flavours are known. The six leptons are grouped in three generations of increasing mass. Each generation holds a charged lepton and an associated, electrically neutral neutrino. The charged leptons, *electron*, *muon* and *tau*, hold one negative electrical elementary charge.

Similarly to the leptons, the six flavours of quarks are grouped into three generations of increasing mass. Each generation holds a positively charged *up*-type quark and a negatively charged *down*-type quark. The charge of all *up*-type quarks is $+2/3$ of the elementary charge, $e$. *Down*-type quarks hold a charge of $-1/3e$.

Leptons and quarks have half-integer spin and are therefore called fermions. Table 1.1 illustrates the three generations of leptons and quarks.

The four known interactions are the electromagnetic coupling, the weak and the strong coupling and gravitation. These interactions are mediated by mediator particles, also called bosons because of their integer spin. Bosons couple to the charge that is associated with its interaction.

The electromagnetic interaction couples to electrical charge and is mediated by photons. The weak force is the only interaction that allows (to a certain extent) the transformation of lepton and quark flavours and is the cause for nuclear decays. It is mediated by the massive $W$ and $Z$ bosons and couples to the weak charge. The strong interaction couples to color charge. It holds protons and neutrons in atomic nuclei together. The strong interaction is mediated by eight color-charged gluons. Gravity is extremely weak, compared to the


1. Theoretical Foundations

| Table 1.1: The three generations of fermions in the Standard Model. |
|------------------------|------------------|------------------|
|                        | Leptons          | Quarks           |
| 1st generation         | (e) electron     | (u) up           |
|                        | (νe) electron neutrino | (d) down       |
| 2nd generation         | (μ) muon         | (c) charm        |
|                        | (νμ) muon neutrino | (s) strange      |
| 3rd generation         | (τ) tau          | (t) top          |
|                        | (ντ) tau neutrino | (b) bottom       |

other three forces at the energy scales that are reachable at particle colliders. It is not included in the Standard Model.

The Standard Model is described by the gauge group $SU(3)_C \times SU(2)_L \times U(1)_Y$ with the color charge, $C$, the weak isospin, $T_3$, and the weak hypercharge, $Y$. The unification of electromagnetic and weak interaction into the electroweak interaction is represented by the gauge subgroup $SU(2)_L \times U(1)_Y$. Table 1.2 displays the four interactions and their gauge properties.

| Table 1.2: The four interactions of the SM, their gauge symmetries and bosons. |
|------------------------|------------------|------------------|
| Interaction            | Gauge Symmetry   | Boson            | Mass   | Range   |
| electromagnetic        | $SU(2)_L \times U(1)_Y$ | Photon          | 0      | $\infty$ |
| weak                   | $SU(3)_C$        | $Z^0$            | 91 GeV | $10^{-18}$ m |
|                        |                  | $W^\pm$          | 80 GeV | $10^{-15}$ m |
| strong                 | $SU(3)_C$        | eight gluons     | 0      | $\approx 10^{-15}$ m |
| gravity                |                  | graviton?        | 0      | $\infty$ |

The origin of particle masses is still unknown. Particle masses are introduced in the Standard Model using the Higgs mechanism [6, 7], which also proposes an additional particle, the Higgs boson. This particle has not been found so far. Recent analyses of the ATLAS and the CMS experiment indicate that the Higgs boson might be existent with a mass of about 125 GeV [8, 9].

The Standard Model is, except for the Higgs boson, experimentally well confirmed to a very high level of precision [10]. It was introduced in the 1960s with a prediction of the massive gauge bosons of the weak interaction. These were discovered about 20 years later at CERN. Another major experimental confirmation of the Standard Model was the discovery of the top quark, which was postulated in 1973, in the year 1995 at the Tevatron.

### 1.1.1 Natural Units

Throughout this work, the system of natural units is used. The speed of light, $c$, and the reduced Planck constant, $\hbar$, are set to unity:

$$c = \hbar = 1.$$  (1.1)
Energies and momenta are measured in GeV, lengths in GeV\(^{-1}\). For convenience, lengths are usually still given in cm in this work. The relation of cm and GeV is

\[ 1 \text{ cm} = \frac{1 \text{ cm}}{h c} \approx 5.1 \cdot 10^{13} \text{ GeV}^{-1}. \tag{1.2} \]

Cross sections are expressed in barn: 1b = 10\(^{-24}\) cm\(^2\).

If clear from the context, particles and antiparticles are referred to only by their symbols, and their charge-specification indices are dropped. The reaction \(Z^0 \rightarrow \ell^+ \ell^-\), for example, is therefore shortened to \(Z \rightarrow \ell \ell\).

### 1.2 Reasons for Dissatisfaction

The Standard Model is a very successful theory, which predicts current experimental results with a great precision. However, it is clear that the Standard Model is not complete, and therefore just an effective theory, because it does not include gravity. At low energy scales, gravity is weaker than the other forces and can be neglected. At the Planck scale (\(\sim 10^{19} \text{ GeV}\)) however, gravity reaches a strength comparable to the other forces, and the Standard Model cannot be valid any more.

Notwithstanding the above, there are further indications for physics beyond the Standard Model, even below the Planck scale.

#### 1.2.1 Hierarchy Problem

One indication for physics beyond the Standard Model is the hierarchy problem. It arises from the fact that the mass scale of loop corrections to the Higgs boson mass is significantly higher than the Higgs boson mass itself [11–14].

The electroweak sector of the SM contains a parameter representing the energy scale of electronweak interactions, the weak scale:

\[ v \approx 264 \text{ GeV}. \tag{1.3} \]

This parameter is related to the vacuum expectation value of the neutral Higgs field, \(v/\sqrt{2}\), the mass of the Higgs boson, \(m_H\), and the mass of the W boson, \(m_W\):

\[ m_W = \frac{g v}{2} \approx 80 \text{ GeV}, \tag{1.4} \]

\[ m_H = v \sqrt{\frac{\lambda}{2}}. \tag{1.5} \]

In this relation, \(g\) is the electroweak gauge coupling constant and \(\lambda\) the strength of the Higgs self-interaction in a Higgs potential given by

\[ V = m_H^2 |H|^2 + \lambda |H|^4. \tag{1.6} \]

Fermion-loop corrections (see Fig. 1.1a) produce a quadratically diverging correction to the Higgs-boson mass:

\[ \Delta m_H^2 = -\frac{|\lambda_f|^2}{8\pi^2} \Lambda_{UV}^2 + \ldots, \tag{1.7} \]

with the coupling of the fermion to the Higgs boson, \(\lambda_f\), and the cut-off scale, \(\Lambda_{UV}\). If the Standard Model shall be valid until the Planck scale, with \(\Lambda_{UV} = \Lambda_{Pl} \sim 10^{19} \text{ GeV}\), these corrections are many orders of magnitude larger than \(m_H\) itself, which is expected to be
of the order of a few hundred GeV. This difference in the mass hierarchy is considered to be “unnatural”, and hence ways to avoid it are investigated.

The scale difference of loop corrections and the Higgs mass would be smaller if new physics appeared at a scale, \( \Lambda \), that was smaller than the Planck scale. This way however, the problem is not solved; just the severity of the problem is reduced. A different approach to this problem is the introduction of a symmetry that groups scalar particles with fermions [12] [15–17]. Supersymmetry is such a symmetry, which introduces a scalar degree of freedom for each fermion degree of freedom in the SM. These new degrees of freedom add another contribution to the Higgs mass (see diagram in Fig. 1.1b) of the form

\[
\Delta m_H^2 = \frac{\lambda_S}{16\pi^2}(\Lambda_{UV}^2 - 2m_S^2 \ln(\Lambda_{UV}/m_S) + ...),
\]

with \( \lambda_S \) being the coupling to the Higgs boson and \( m_S \) the mass of the scalar.

In a theory with two scalars for each fermion (one scalar for each fermion degree of freedom), the quadratic divergences cancel out if the couplings to the Higgs boson of both particles are the same (\( \lambda_f = \lambda_S \)). This way, Supersymmetry provides an elegant way of solving the hierarchy problem.

1.2.2 Unification of Couplings

Extrapolation of the three coupling constants of the Standard Model to higher energies shows no common intersection point. An intersection of the coupling constants is however desired, because this point would mark the scale at which a unification of the three couplings is possible. In the Minimal Supersymmetric extension of the Standard Model (MSSM), the unification of the three couplings is possible as displayed in Figure 1.2.

1.2.3 Astrophysical Observations

Measurements by the Wilkinson Microwave Anisotropy Probe [19] (WMAP) implicate that the universe consists only to 4% of baryonic matter. The remaining part comprises dark matter (23%) and dark energy (73%). Dark Matter (DM) cannot directly be observed, but affects baryonic matter indirectly by weak or gravitational interactions.

Another indication for gravitational interacting, but not visible, hence dark, matter was found in the analysis of galaxy rotation curves [20].

Two hypothetical types of dark matter are considered: hot dark matter (ultra-relativistic particles, e.g. neutrinos) and cold dark matter (slower-moving particles). The structure of the universe indicates that the observed dark matter must be cold.

The Standard Model does in general not provide any good explanation for observation of dark matter. Supersymmetric models, on the other hand, provide a candidate for dark matter with the lightest supersymmetric particle in the particle spectrum.
1.3 Supersymmetry

Supersymmetry (SUSY) relates fermionic and bosonic degrees of freedom [12, 13, 21–23]. It is generated by an operator, $Q$, that transforms bosonic into fermionic, and fermionic into bosonic states:

$$Q|\text{fermion}\rangle = |\text{boson}\rangle, \quad Q|\text{boson}\rangle = |\text{fermion}\rangle. \quad (1.9)$$

Following this relation, superpartners can be assigned to SM particles. SUSY is constructed such that these superpartners have the same quantum numbers as their symmetry partners, except for the spin, which differs by $\frac{1}{2}$. While in principle more than one superpartner per SM particle is possible, just the minimal extension of the Standard Model, the MSSM, is considered here.

1.3.1 Minimal Supersymmetric Extension of the Standard Model

In the Minimal Supersymmetric extension of the Standard Model (MSSM), one superpartner field is introduced for each degree of freedom in the Standard Model. Fermion superpartners are usually named by their SM partners and the prefix “s”, e.g. squarks or sleptons. In general, they are also referred to as sfermions. Superpartners of bosons are indicated with the suffix “-ino”, e.g. wino or gluino, and are called gauginos.

Two independent complex Higgs doublets are needed, because Yukawa interactions involving a complex scalar field and its hermitian conjugate are forbidden in a supersymmetric lagrangian. Therefore, two complex Higgs doublets, corresponding to four Higgs fields, are used in the MSSM:

$$H_1 = \begin{pmatrix} H_1^0 \\ H_1^- \end{pmatrix}, \quad H_2 = \begin{pmatrix} H_2^0 \\ H_2^+ \end{pmatrix}. \quad (1.10)$$

$H_1$ is used to introduce masses of down-type quarks and charged leptons, $H_2$ creates up-type quark masses. Thus, $H_1$ is sometimes also referred to as $H_d$, and $H_2$ as $H_u$, in

Figure 1.2: Evolution of the three SM coupling constants with rising energy in the SM (left) and the minimal supersymmetric extension of the SM (MSSM) (right) [18]. $\alpha_1$, $\alpha_2$ and $\alpha_3$ are the electromagnetic, the weak and the strong coupling constant, respectively. In the MSSM, unification of the three couplings is possible. Contribution of SUSY particles is assumed above an effective SUSY energy scale of 1 TeV. Uncertainties are indicated by the thickness of the lines.
textbooks. For each of the four Higgs particle, one supersymmetric partner (higgsino) is introduced.

Table 1.3 displays the SM particles and their SUSY partners. While the listed states are gauge eigenstates of the SUSY particles, the experimentally accessible eigenstates are the mass eigenstates. These are obtained as a mixture of gauge eigenstates. Gauge and mass eigenstates of the SUSY particles in the MSSM are listed in Table 1.4.

The four gauge eigenstates of neutral SUSY gauge particles (bino, wino and neutral higgsinos) mix into four neutral gauginos, the neutralinos, \( \tilde{\chi}_1^0, \tilde{\chi}_2^0, \tilde{\chi}_3^0 \) and \( \tilde{\chi}_4^0 \). Analogously, the mixing of the four charged SUSY gauge particles results in four charged gauginos, the charginos, \( \tilde{\chi}_1^\pm \) and \( \tilde{\chi}_2^\pm \).

**Table 1.3:** Standard Model particles and their SUSY partners in the MSSM.

<table>
<thead>
<tr>
<th>SM particle</th>
<th>Spin</th>
<th>SUSY particle</th>
<th>Spin</th>
</tr>
</thead>
<tbody>
<tr>
<td>( q )</td>
<td>quark</td>
<td>( \tilde{q} )</td>
<td>squark</td>
</tr>
<tr>
<td>( \ell )</td>
<td>lepton</td>
<td>( \tilde{\ell} )</td>
<td>slepton</td>
</tr>
<tr>
<td>( g )</td>
<td>gluon</td>
<td>( \tilde{g} )</td>
<td>gluino</td>
</tr>
<tr>
<td>( W^\pm, W^0 )</td>
<td>( W ) boson</td>
<td>( \tilde{W}^\pm, \tilde{W}^0 )</td>
<td>wino</td>
</tr>
<tr>
<td>( B^0 )</td>
<td>( B ) boson</td>
<td>( \tilde{B}^0 )</td>
<td>bino</td>
</tr>
<tr>
<td>( 4 \times H )</td>
<td>Higgs boson</td>
<td>( 4 \times \tilde{H} )</td>
<td>higgsino</td>
</tr>
<tr>
<td>( (G) )</td>
<td>Graviton</td>
<td>( \tilde{G} )</td>
<td>gravitino</td>
</tr>
</tbody>
</table>

**Table 1.4:** Gauge and mass eigenstates of SUSY particles in the MSSM.

<table>
<thead>
<tr>
<th>Particle</th>
<th>Gauge eigenstates</th>
<th>Mass eigenstates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Higgs bosons</td>
<td>( H_1^0, H_2^0, H_1^-, H_2^+ )</td>
<td>( h^0, A^0, H_0^0, H^\pm )</td>
</tr>
<tr>
<td>squark</td>
<td>( \tilde{q} )</td>
<td>( \tilde{q} )</td>
</tr>
<tr>
<td>slepton</td>
<td>( \tilde{\ell} )</td>
<td>( \tilde{\ell} )</td>
</tr>
<tr>
<td>neutralinos</td>
<td>( \tilde{B}^0, \tilde{W}^0, \tilde{H}_1^0, \tilde{H}_2^0 )</td>
<td>( \tilde{\chi}_1^0, \tilde{\chi}_2^0, \tilde{\chi}_3^0, \tilde{\chi}_4^0 )</td>
</tr>
<tr>
<td>charginos</td>
<td>( \tilde{W}^\pm, \tilde{H}_1^+, \tilde{H}_2^+ )</td>
<td>( \tilde{\chi}_1^\pm, \tilde{\chi}_2^\pm )</td>
</tr>
<tr>
<td>gluino</td>
<td>( \tilde{g} )</td>
<td>( \tilde{g} )</td>
</tr>
<tr>
<td>gravitino</td>
<td>( \tilde{G} )</td>
<td>( \tilde{G} )</td>
</tr>
</tbody>
</table>

**R-Parity**

In principle, SUSY introduces terms that allow lepton and baryon number violation into the MSSM lagrangian. Since no such processes have ever been experimentally observed so far, an additional symmetry is introduced to protect lepton and baryon number conservation: the \( R \)-parity. \( R \) is defined by

\[
R = (-1)^{3B+L+2s}
\]

with the baryon number, \( B \), the lepton number, \( L \), and the spin, \( s \). \( R \) has the value +1 for all SM particles and −1 for all SUSY particles. Assuming \( R \)-parity conservation, \( R \) is multiplicatively conserved. In this case, every vertex contains an even number of SUSY particles.
Conservation of $R$-parity has important phenomenological consequences:

- The Lightest Supersymmetric Particle (LSP) is stable. If not electrically charged, it is an attractive dark matter candidate.
- All SUSY particles decay into an odd number of LSPs.
- SUSY particles can only be produced in pairs.

$R$-parity violating SUSY models are also possible, but these are not further considered in this work.

**Higgs Sector and Gaugino Mixing**

The two introduced Higgs doublets correspond to eight real scalar degrees of freedom. Three of these degrees of freedom are used to give mass to $W$ and $Z$ bosons. The remaining five degrees of freedom result in five massive Higgs states: two neutral scalars, $h^0$ and $H^0$, a neutral pseudo-scalar, $A^0$, and two charged scalars, $H^\pm$. By convention, $h^0$ is lighter than $H^0$.

The ratio of the Higgs vacuum expectation values, also called vevs, is defined as $\tan \beta$:

$$\tan \beta = \frac{v_2}{v_1} = \frac{\langle H_2^0 \rangle}{\langle H_1^0 \rangle}.$$  

Superpartners to the four SM Higgs fields are the four higgsinos, $\tilde{H}_1^0$, $\tilde{H}_1^1$, $\tilde{H}_2^0$, and $\tilde{H}_2^1$. The gauge eigenstates of the two neutral higgsinos, the wino and the bino mix into the four neutralino mass eigenstates. The corresponding mass term in the lagrangian can be written as

$$\mathcal{L}_{\tilde{\chi}^0} = -\frac{1}{2} \cdot \begin{pmatrix} \tilde{H}_1^0 & \tilde{H}_1^1 & \tilde{H}_2^0 & \tilde{H}_2^1 \end{pmatrix} \begin{pmatrix} M_1 & 0 & -c_\beta s_W \cdot m_Z & s_\beta s_W \cdot m_Z \\ 0 & M_2 & c_\beta c_W \cdot m_Z & -s_\beta c_W \cdot m_Z \\ -c_\beta s_W \cdot m_Z & c_\beta c_W \cdot m_Z & 0 & -\mu \\ s_\beta s_W \cdot m_Z & -s_\beta c_W \cdot m_Z & -\mu & 0 \end{pmatrix} \begin{pmatrix} \tilde{H}_1^0 \\ \tilde{H}_1^1 \\ \tilde{H}_2^0 \\ \tilde{H}_2^1 \end{pmatrix} + \text{h.c.}$$  

with the the neutralino mass matrix, $M_{\tilde{\chi}^0}$. The neutralino mass matrix is given by

$$M_{\tilde{\chi}^0} = \begin{pmatrix} M_1 & 0 & -c_\beta s_W \cdot m_Z & s_\beta s_W \cdot m_Z \\ 0 & M_2 & c_\beta c_W \cdot m_Z & -s_\beta c_W \cdot m_Z \\ -c_\beta s_W \cdot m_Z & c_\beta c_W \cdot m_Z & 0 & -\mu \\ s_\beta s_W \cdot m_Z & -s_\beta c_W \cdot m_Z & -\mu & 0 \end{pmatrix}$$  

with the abbreviations, $s_\beta \equiv \sin \beta$, $c_\beta \equiv \cos \beta$, $s_W \equiv \sin \theta_W$ and $c_W \equiv \cos \theta_W$, the mass of the Z boson, $m_Z$, and the electroweak mixing angle, $\theta_W$. $M_1$ and $M_2$ are gaugino mass parameters, and $\mu$ is a higgsino mass parameter. $\mu$ is real, but the sign of $\mu$ is not fixed by the Higgs sector.

The neutralino sector is determined by the (real) SUSY parameters, $M_1$, $M_2$, $\tan \beta$, and $\mu$. Under the assumption of gaugino-mass unification, e.g. in the constrained MSSM (see Sec. 1.3.3), one of the parameters $M_1$ and $M_2$ can be expressed in terms of the other.

The two charged wino states and the two charged higgsino states mix into four chargino mass eigenstates. The mass term in the lagrangian reads

$$\mathcal{L}_{\tilde{\chi}^{\pm}} = -\frac{1}{2} \cdot \begin{pmatrix} \tilde{W}_1^+ & \tilde{H}_1^+ \\ \tilde{W}_2^- & \tilde{H}_2^- \end{pmatrix} \begin{pmatrix} \tilde{W}_1^- & \tilde{H}_1^- \\ \tilde{W}_2^+ & \tilde{H}_2^+ \end{pmatrix} + \text{h.c.}$$

with the chargino mass matrix, $\mathcal{M}_{\tilde{\chi}^\pm}$, given by

$$\mathcal{M}_{\tilde{\chi}^\pm} = \begin{pmatrix} M_2 & \sqrt{2}s_\beta \cdot m_W \\ \sqrt{2}c_\beta \cdot m_W & \mu \end{pmatrix},$$  

(1.16)

where $m_W$ denotes the mass of the $W$ boson.

The couplings of neutralinos and charginos strongly depend on the gaugino mixing scheme and hence, the parameters that were introduced in this section.

### 1.3.2 Symmetry Breaking

Since no SUSY particles have been discovered so far, they have to have higher masses than their SM partners. Therefore, Supersymmetry must be broken. Two ways of symmetry breaking are known: explicit and spontaneous symmetry breaking. Spontaneous breaking is theoretically favoured because it preserves renormalisability in the Standard Model. However, the MSSM has to be extended in order to get a phenomenologically acceptable model including spontaneous SUSY breaking. In practice, SUSY breaking is realised by introducing SUSY-breaking terms, which are constrained in a way that they do not re-introduce quadratic divergences (see Section 1.2.1, the hierarchy problem), into the lagrangian. Such terms are called soft breaking terms.

Theoretically, SUSY breaking can be mediated from a hidden sector, in which SUSY is broken (see Fig. 1.3). If this hidden sector carries no SM quantum numbers, it can be assumed not to show any other interaction with the MSSM. The symmetry breaking is then mediated into the visible MSSM by flavour-blind interactions, for example gravity.

![Figure 1.3: Presumed SUSY breaking scheme, in which SUSY breaking is mediated from a hidden sector.](image)

The soft SUSY-breaking terms that are introduced in the lagrangian are expressed using chiral supermultiplets. The chiral supermultiplet fields of the MSSM are listed in Table 1.5. The contribution of the soft breaking terms to the lagrangian is

$$\mathcal{L}_{\text{soft}} = -\frac{1}{2} \left( M_3 \tilde{g} \tilde{g} + M_2 \tilde{W} \tilde{W} + M_1 \tilde{B} \tilde{B} + \text{h.c.} \right)$$

(1.17)

$$- \tilde{Q}^\dagger \cdot m_Q^2 \cdot \tilde{Q} - \tilde{u}_L^\dagger \cdot m_u^2 \cdot \tilde{u}_L - \tilde{d}_L^\dagger \cdot m_d^2 \cdot \tilde{d}_L$$

(1.18)

$$- \tilde{L}^\dagger \cdot m_L^2 \cdot \tilde{L} - \tilde{e}_L^\dagger \cdot m_e^2 \cdot \tilde{e}_L$$

(1.19)

$$- m_H^2 H_1^\dagger H_1 - m_H^2 H_2^\dagger H_2 - (b H_1 H_2 + \text{h.c.})$$

(1.20)

$$+ \tilde{a}_d \cdot \tilde{Q} H_1 + \tilde{e}_L \cdot a_e \cdot \tilde{L} \cdot H_1 - \tilde{u}_L \cdot a_u \cdot \tilde{Q} H_2 + \text{h.c.}$$

(1.21)

with the gluino, wino and bino mass parameters, $M_3$, $M_2$ and $M_1$. Line 1.17 introduces squark and Line 1.18 introduces slepton mass terms. $m_Q^2$, $m_U^2$, $m_D^2$, $m_L^2$ and $m_e^2$ are $3 \times 3$ mass matrices. The contribution to the Higgs potential is shown in Line 1.20 including three Higgs mass parameters. The last line, Line 1.21, contains triple scalar couplings of the Higgs bosons, where $a_d$, $a_e$ and $a_u$ are complex $3 \times 3$ matrices.

Many additional parameters are introduced by these soft SUSY-breaking terms. Summing up the introduced mass parameters, coupling parameters and phases, the MSSM extends the Standard Model by 105 additional parameters.
1.3.3 Constrained MSSM

The large amount of additional parameters in the MSSM is reduced in the constrained MSSM, or CMSSM. In this model, symmetry breaking is assumed to be mediated by gravity from a hidden sector.

The parameters take a particularly simple form at energies of the GUT scale ($\sim 10^{16}$ GeV) in the CMSSM:

\begin{align*}
M_3 &= M_2 = M_1 = m_{1/2}, \\
m_Q^2 &= m_u^2 = m_d^2 = m_e^2 = m_{\tilde{e}}^2 = m_{\tilde{\nu}}^2 = m_{\tilde{H}_2}^2, \\
\frac{m_{\tilde{H}_1}^2}{m_{\tilde{H}_2}^2} &= m_{\tilde{H}_1}^2 = m_{\tilde{H}_2}^2, \\
a_u &= A_0 y_u, \\
a_d &= A_0 y_d, \\
a_e &= A_0 y_e.
\end{align*}

The $y_x$ matrices are the same Yukawa matrices that describe fermion mixing in the SM, and $\mathbf{1}$ denotes the $3 \times 3$ unity matrix. Using these constraints, the only remaining parameters are:

- $m_0$: common mass of scalars at the GUT scale.
- $m_{1/2}$: common mass of gauginos at the GUT scale.
- $A_0$: trilinear coupling at the GUT scale.
- $\tan \beta$: ratio of Higgs vacuum expectation values.
- $\text{sgn}(\mu)$: sign of higgsino mass parameter at the GUT scale.

Particle masses at lower energies than the unification scale can be obtained using renormalisation group (RG) equations. An exemplary evolution of scalar and gluino masses with the energy scale in the CMSSM is shown in Figure 1.4. The unification is enforced at an energy scale of $2 \cdot 10^{16}$ GeV. At low energies, the mass parameter $\mu^2 + m_{\tilde{H}_2}^2$ runs negative and hence provokes electroweak symmetry breaking (EWSB). This symmetry breaking is necessary, but not realised in all regions of the CMSSM parameter space. The mass configuration at lower energies, where SUSY particles might be observed in particle collider experiments, depends strongly on the chosen parameters.

It is practically impossible to scan the complete 105-parameter space that is created by the MSSM. Therefore the CMSSM and similar constrained models with a reduced amount of free parameters are favoured for the interpretation of experimental results. Despite their sometimes firm constraints, these models share many phenomenological consequences, such as sparticle decay chains, with more general MSSM models.
1. Theoretical Foundations

1.4 Dileptonic Final States

This analysis concentrates on events with two light leptons (electrons and muons) in the final state. Leptons are usually easy to identify and provide an efficient suppression against many Standard Model background processes. Especially QCD processes, which have a huge cross section at hadron colliders, are efficiently discriminated. In the following, SUSY particle production mechanisms and the decay mechanisms that lead to dileptonic final states are introduced.

1.4.1 Sparticle Production and Decay

Assuming R-parity conservation, SUSY particles have to be produced in pairs. The production of SUSY particles can happen via electroweak coupling (charginos and neutralinos) or via strong coupling (squarks and gluinos). At the LHC, the strong production of SUSY particles is favoured. Therefore $\tilde{g}\tilde{g}$, $\tilde{q}\tilde{q}$ and $\tilde{g}\tilde{q}$ are the dominant production channels.

After their production, SUSY particles decay in cascades into SM particles and the LSP. Since strong production of SUSY particles is favoured at the LHC, at least two quarks or gluons are produced in these cascades. In addition to the hadronic energy of these quarks and gluons, the LSPs escape the detector and result in a significant momentum imbalance in the transverse detector plane.

The branching fraction into leptons is in general low compared to decays into hadronic products. A dilepton signature has still been chosen, because leptons provide a good possibility to suppress many SM backgrounds. The most important cascade decay for dilepton signatures involves the decay of a $\tilde{\chi}_0^0$ into a $\tilde{\chi}_1^0$: 

$$\tilde{g}/\tilde{q} \rightarrow g/q + \tilde{\chi}_2^0 \rightarrow g/q + \tilde{\chi}_1^0 + \ell^+\ell^-.$$  \hspace{1cm} (1.26)

Only electrons and muons are considered for $\ell$, because these are easier to identify and provide a better discrimination against QCD background than taus. The total branching fraction of such a cascade into electrons or muons is of the order of 1%.

Figure 1.4: Evolution of scalar and gaugino masses in the CMSSM with unification at $2 \cdot 10^{16}$ GeV [13].
Lepton production in SUSY cascades other than this is also possible via $W$ bosons. These leptons are however not correlated in flavour, and the corresponding amount of background events can therefore be predicted and subtracted using methods described later on.

### 1.4.2 Kinematic Edges

Two flavour-correlated leptons can be produced in the decay of a $\tilde{\chi}_2^0$ into a $\tilde{\chi}_1^0$. This neutralino decay can be performed via an intermediate virtual or real slepton or by radiation of a $Z$ boson that decays leptonically (see Fig. 1.5). The invariant-mass distribution of the two produced leptons is defined by the kinematics of the decay and often contains characteristic edge structures.

![Figure 1.5: Lepton-pair production in the decay of a $\tilde{\chi}_2^0$ into a $\tilde{\chi}_1^0$ via (virtual or real) slepton (a) and via $Z$-boson radiation (b).](image)

If the neutralino mass difference is smaller than the $Z$-boson mass, and the neutralino cannot decay into a slepton, then the decay $\tilde{\chi}_2^0 \rightarrow \tilde{\chi}_1^0 \ell \ell$. In this case, the invariant-mass distribution endpoint for the lepton pair is given by the neutralino mass difference:

$$m_{\text{max}} = m_{\tilde{\chi}_2^0} - m_{\tilde{\chi}_1^0}. \quad (1.27)$$

If the slepton mass is between the two neutralino masses and hence, the decay $\tilde{\chi}_2^0 \rightarrow \tilde{\ell} \ell$ is possible, the neutralino decays via two subsequent two-body decays. The maximum invariant mass of the lepton pair is then given by

$$m_{\text{max}}^2 = \frac{(m_{\tilde{\chi}_2^0}^2 - m_{\tilde{\ell}}^2)(m_{\tilde{\ell}}^2 - m_{\tilde{\chi}_1^0}^2)}{m_{\tilde{\ell}}^2} \quad (1.28)$$

with the slepton mass, $m_{\tilde{\ell}}$.

The shape of the distribution depends strongly on the parameters of the SUSY model (see Fig. 1.6). For heavy sleptons and large neutralino mass differences, the decay $\tilde{\chi}_2^0 \rightarrow Z \tilde{\chi}_1^0$ becomes more important, and the distribution is peaked towards the $Z$ mass. For light sleptons, the decay amplitudes of a decay via $Z$ boson and via slepton show negative interference, which enhances the general dependence of the invariant-mass distribution on the model parameters. If two subsequent two-body decays are the dominant decay mechanism ($\tilde{\chi}_2^0 \rightarrow \tilde{\ell} \ell \rightarrow \ell \ell \tilde{\chi}_1^0$), the shape of the lepton-pair invariant-mass distribution is mainly triangular with its maximum at the highest possible mass, $m_{\text{max}}$.

### 1.5 Benchmark Scenarios

Multiple benchmark scenarios are defined within the CMSSM. These scenarios were chosen to give a broad coverage of the various regions of different SUSY characteristics that are
possible within the CMSSM. Ten low-mass (LM) points and four high-mass (HM) points were defined [25]. The LM points were designed as benchmark scenarios in regions that are accessible by SUSY searches during the first years of LHC running. The HM points are dedicated for searches after a couple of years of data taking. Figure 1.7 displays the location of the various benchmark points in the $m_0$-$m_1/2$ plane. At all benchmark points, the $\tilde{\chi}_1^0$ is the LSP and a good candidate for dark matter.

Three of these CMSSM benchmark scenarios, namely LM1, LM3 and LM6, are paid special interest in this work and are discussed in the following. The CMSSM parameters of these points can be found in Table 1.6. It has to be noted though, that in principle, LM1 and LM3 have already been excluded by the LEP Higgs-mass limit of 114 GeV [26]. Nevertheless, they can serve as representative benchmarks for certain SUSY signatures.

**Table 1.6:** Benchmark points, their CMSSM parameters and the maximum invariant mass, $m_{\text{max}}$, of lepton pairs that are produced in the decay of a $\tilde{\chi}_0^0$.

<table>
<thead>
<tr>
<th>Name</th>
<th>$m_0$ [GeV]</th>
<th>$m_{1/2}$ [GeV]</th>
<th>$A_0$ [GeV]</th>
<th>$\tan \beta$</th>
<th>sgn $\mu$</th>
<th>$m_{\text{max}}$ [GeV]</th>
</tr>
</thead>
<tbody>
<tr>
<td>LM1</td>
<td>60</td>
<td>250</td>
<td>0</td>
<td>10</td>
<td>+1</td>
<td>78.2</td>
</tr>
<tr>
<td>LM3</td>
<td>330</td>
<td>240</td>
<td>0</td>
<td>10</td>
<td>+1</td>
<td>79.5</td>
</tr>
<tr>
<td>LM6</td>
<td>85</td>
<td>400</td>
<td>0</td>
<td>10</td>
<td>+1</td>
<td>105</td>
</tr>
</tbody>
</table>

### 1.5.1 Mass Spectra

The SUSY particle mass spectra in the various benchmark scenarios are derived using SOFTSUSY [27]. For LM1, LM3 and LM6, they are illustrated in Figure 1.8.

At LM1 and LM3, the masses of gluinos and squarks are about equal, yielding approximately 600 GeV. The gluino is the heaviest particle at LM1; therefore the production of squarks is favoured. Compared to these two points, LM6 has a remarkably high squark mass spectrum. Squark masses range up to about 850 GeV. The gluino is even heavier, again the heaviest SUSY particle in the spectrum, with about 950 GeV. Thus, also at this point, squark production is favoured.
Figure 1.7: Position of the CMS benchmark points in the $m_0$-$m_{1/2}$ plane of the CMSSM [25]. For all points, $\tan \beta = 10$, $A_0 = 0$ and $\text{sgn} \mu = +1$ are fixed. Regions in which no electroweak symmetry breaking (EWSB) occurs are indicated in yellow. Region in which the stau is the lightest supersymmetric particle (LSP) are shown in teal. Changes in mass relations are displayed as green lines. The region excluded by the Tevatron is indicated by the pink curve. Exclusion limits corresponding to exclusions from direct Higgs searches are shown as red, dashed lines for various Higgs boson masses.

The lightest two neutralinos have masses between 100 GeV and 200 GeV at LM1 and LM3 and between 150 GeV and 300 GeV at LM6. LM1 has a light slepton spectrum, which ranges from 100 GeV to 200 GeV, and it has a slepton placed in between the two lightest neutralino masses. Therefore the neutralino decay is performed by two subsequent two-body decays leading to a characteristic triangular mass edge in the invariant-mass distribution of the produced lepton pairs. At LM6, the mass spectrum globally is higher, but a neutralino decay via slepton is also possible. The sleptons at LM3 have masses of around 350 GeV. Therefore, the neutralino cannot decay via real slepton and a virtual slepton or a $Z$ boson is used instead.

1.5.2 Cross Sections and Branching Ratios

Cross sections of the SUSY processes are calculated using Prospino [28, 29] at next-to-leading order (NLO) accuracy. The $(\tilde{u}, \tilde{d}, \tilde{c}, \tilde{s}, \tilde{b})_{L/R}$ squarks are virtually mass-degenerate in the considered scenarios. Therefore, it is possible to simply sum over the cross sections of these ten squarks. The production cross sections for the three benchmark scenarios, LM1, LM3 and LM6, at a centre-of-mass energy of 7 TeV yield $\sigma_{LM1} = 6.55 \text{ pb}$, $\sigma_{LM3} = 4.81 \text{ pb}$ and $\sigma_{LM6} = 0.40 \text{ pb}$.

All branching ratios for the decaying SUSY particles are estimated with SUSYHIT [30].

1.5.3 Invariant-Mass Distributions

Figure 1.9 shows the invariant-mass distribution of electron and muon pairs for the three benchmark scenarios, LM1, LM3 and LM6. Lepton pairs that do not emerge from $\tilde{\chi}_2^0$ decays have been subtracted.
Figure 1.8: SUSY mass spectra for the benchmark scenarios LM1 (a), LM3 (b) and LM6 (c).
1.6 Simplified Models

In the LM1 scenario, a mass edge following the kinematics of two subsequent two-body decays is visible. At LM3, the decay via Z boson plays a significant role. Thus, the mass distribution peaks at the Z mass. Like LM1, LM6 has the characteristics of two-body decays. At this point however, a variation of the mass edge due to interference of the decay amplitudes is visible.

![Invariant-mass distribution of electron and muon pairs in benchmark scenarios LM1 (a), LM3 (b) and LM6 (c). Contributions not coming from the decay \( \tilde{\chi}_2^0 \to \tilde{\chi}_1^0 + \ell^+ \ell^- \) have been subtracted.](image)

**Figure 1.9:** Invariant-mass distribution of electron and muon pairs in benchmark scenarios LM1 (a), LM3 (b) and LM6 (c). Contributions not coming from the decay \( \tilde{\chi}_2^0 \to \tilde{\chi}_1^0 + \ell^+ \ell^- \) have been subtracted.

1.6 Simplified Models

New physics models usually introduce a full set of new particles and their interactions, which makes the interpretation of experimental results in the context of the new model a complicated matter. Simplified models were introduced to reduce the effort of interpretation. Designed to contain only a few particles and interactions, they can be
Another topology, T3w, is constructed by setting one gluino to decay directly to the LSP, and mass is adjusted such that the mass splitting is close to the case when the intermediate neutralino 2 decays to a pair of SM particles: gluino, neutralino 2 and neutralino 1, which is the LSP. In this model, gluinos and neutralinos that are part of these models might be very simplistic replicas of the original SUSY particles. Still they incorporate specific properties of their originals that lead to interaction processes, which can be characteristic for certain parameter regions of SUSY models.

### 1.6.1 T3lh Model

Of special interest for this analysis is the T3lh model \[34\] (Fig. 1.10). It involves three non-SM particles: gluino, neutralino 2 and neutralino 1, which is the LSP. In this model, gluinos decay either into the neutralino 2 and two quarks or into the LSP using a similar three-body decay. The neutralino 2 produces a pair of flavour-correlated, oppositely charged leptons when it decays into the LSP.

The seemingly cryptic name of the model, T3lh, implies the production and decay mechanism and the final states of the modeled events. In each event, a pair of gluinos is produced, and one gluino decays in a cascade decay. “lh” in the model name indicates that one gluino results in a leptonic final state, and the other gluino decays pure hadronically.

![Diagram of the T3lh simplified model](image)

The three-body decay of the neutralino 2 into the LSP produces a kinematic edge in the invariant-mass spectrum of the lepton pair. The position of the mass edge depends on the masses of the neutralinos and the gluino. In this model, the neutralino-2 mass is positioned exactly underway between the gluino and the LSP mass. Therefore the mass-edge position depends on the remaining two model parameters, \(m_\tilde{g}\) and \(m_\chi_1^0\):

\[
m_{\text{edge}} = m_{\tilde{g}} - m_{\chi_1^0} = \frac{1}{2} \left( m_{\tilde{g}} - m_{\chi_1^0} \right)
\]

(1.29)

\(m_{\chi_1^0}\) is also called \(m_{\text{LSP}}\). The difference between the gluino and the neutralino mass is the most important quantity in this model, because it defines the maximum invariant mass of the produced lepton pair and hence the endpoint of the mass edge. The difference is also called the mass splitting.

Figure 1.11 shows an example of the kinematic edge in the invariant-mass spectrum of lepton pairs produced in the T3lh model.
Figure 1.11: Invariant-mass spectrum of electron and muon pairs for an example configuration ($m_{\tilde{g}} = 500 \text{ GeV}, m_{\tilde{\chi}_1^0} = 200 \text{ GeV}$) of the T3lh simplified model.
2. Experimental Setup

In the following, an overview on the Large Hadron Collider, the currently most sophisticated particle accelerator, and one of its experiments, the CMS detector, is given.

2.1 Large Hadron Collider

The Large Hadron Collider (LHC) \cite{35} is a synchrotron at the European Organization for Nuclear Research (CERN), designed to create proton-proton collisions at a centre-of-mass energy of 14 TeV as well as heavy ion collisions. It is located in a 27 km long tunnel about 100 m below the border of France and Switzerland north west of Geneva.

Several experiments are located at the beam interaction points of the LHC. The four main experiments are ALICE \cite{36}, ATLAS \cite{37}, CMS \cite{38}, and LHCb \cite{39} (Fig. 2.1). ATLAS and CMS are multi-purpose detectors, designed for a broad programme of physics studies. The ALICE experiment investigates predominantly heavy ion collisions, and LHCb concentrates on studying b-quark physics and CP violation.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure2.1.png}
\caption{Overview on the Large Hadron Collider and the location of the four main experiments, ALICE, ATLAS, CMS and LHCb \cite{40}.}
\end{figure}
Protons are pre-accelerated to 450 GeV by an injector chain consisting of the Linac 2, the Proton Synchrotron Booster (PSB), the Proton Synchrotron (PS) and the Super Proton Synchrotron (SPS) before they are injected into the LHC. The LHC can be filled with up to 2808 proton bunches with a bunch spacing of 25 ns. 1232 superconducting dipole and 858 superconducting quadrupole magnets, which are operated at a temperature of 1.9 K, are installed to keep and focus the proton beam. The design luminosity of the LHC is $10^{34} \text{cm}^{-2}\text{s}^{-1}$.

In September 2008, the first beam circulated in the LHC. After a severe technical incident [41] that followed nine days later and was caused by a malfunctioning superconducting connection, several magnets had to be repaired or replaced. In November 2009, re-commissioning of the accelerator started with pilot runs first at a centre-of-mass energy of 900 GeV and then at world-record centre-of-mass energy of 2.36 TeV.

In 2010, the centre-of-mass energy was increased to 7 TeV and the LHC delivered an integrated luminosity of 47 pb$^{-1}$ of proton-proton collisions between March and November. The maximum instantaneous luminosity during this period was about 0.2 nb$^{-1}$s$^{-1}$.

In the following year, 2011, the LHC continued running at the same centre-of-mass energy. The amount of delivered integrated luminosity was multiplied by about two orders of magnitude. A total integrated luminosity of 6.1 fb$^{-1}$ was accumulated. The maximum instantaneous luminosity that was reached in 2011 was about 4.0 nb$^{-1}$s$^{-1}$ [42]. Figure 2.2 displays peak instantaneous and integrated luminosity delivered by the LHC as well as the integrated luminosity recorded by the CMS experiment.

For 2012, it is planned to increase the LHC centre-of-mass energy to 8 TeV. A total integrated luminosity delivery of about 15 fb$^{-1}$ is targeted during this year.

### 2.2 Compact Muon Solenoid

The Compact Muon Solenoid (CMS) is a multi-purpose detector at the LHC. A detailed description of CMS and all of its subdetectors can be found in [38].

The CMS detector consists of multiple layers of subdetectors arranged in a barrel region and two endcaps (Fig. 2.3). The innermost detector is a pixel detector, which is used to determine the origin of charged particles. The tracks of these particles are precisely measured inside the silicon strip tracker, which comprises the next layer. Particle energies...
2.2. Compact Muon Solenoid

are measured with the electromagnetic and the hadronic calorimeters, which represent the next two detector layers. The calorimeters are placed inside a superconducting solenoid coil, which creates an almost homogeneous magnetic field of 3.8 T over a cylindrical volume of 6 m diameter and 12.5 m length. The outermost detector layers consist of four muon stations interlaced in iron flux return yokes.

![Figure 2.3: The CMS detector](image)

2.2.1 Coordinate System

CMS has agreed on the following convention of defining coordinate systems: The $y$-axis points upwards, and the $x$-axis points radially to the centre of the LHC ring. The $z$-axis points in direction of the beam line, completing a right-handed coordinate system.

The azimuthal angle, $\phi$, describes rotations around the $z$-axis and specifies the angle to the $x$-axis in the $x$-$y$-plane. The polar angle, $\theta$, describes the angle to the $z$-axis.

The polar angle can also be specified using the pseudo-rapidity, $\eta$, which is defined as

$$\eta = -\ln \left( \tan \frac{\theta}{2} \right). \quad (2.1)$$

The pseudo-rapidity is positive at the plus side ($\theta < \frac{\pi}{2}$).

Angular distances can be described using the azimuthal angle and the pseudo-rapidity in the $\eta$-$\phi$ plane:

$$\Delta R = \sqrt{\Delta \eta^2 + \Delta \phi^2}. \quad (2.2)$$

The transverse components of vector quantities play an important role because events contain an unknown boost in $z$ direction (along the beam line). The projection of a measured vector quantity (e.g. momentum) on the $x$-$y$ plane is called *transverse component* of the quantity, e.g. the transverse momentum:

$$p_\perp = |p_\perp| = p \cdot |\sin \theta| = \begin{vmatrix} p_x \\ p_y \\ 0 \end{vmatrix}. \quad (2.3)$$
2.2.2 Silicon Tracker

The CMS tracker is entirely based on silicon detector technology. It covers the range of \(|\eta| < 2.5\) and is divided into five subsystems (Fig. 2.4): Pixel detector (PIXEL), Tracker Inner Barrel (TIB), Tracker Outer Barrel (TOB), Tracker Inner Discs (TID) and Tracker Endcaps (TEC).

![Figure 2.4](image)

**Figure 2.4:** View of the CMS tracker in the \(rz\)-plane [38]. Each line represents a silicon strip or pixel detector. Stereo modules are indicated by double lines.

Pixel Detector

The pixel system is essential for secondary-vertex and impact-parameter reconstruction as well as track seeding and counting. It surrounds the interaction point with hybrid pixel detectors, which are arranged in three cylindrical layers in the barrel region and in two discs in the forward and in the backward region, each. The cylinders have radii of 4.4, 7.3, and 10.2 cm and a length of 53 cm. The discs are positioned at \(z = \pm 34.5\) cm and \(z = \pm 46.5\) cm and extend from 6 to 15 cm in radius.

The individual pixels are realised as \(n^+\) implants on \(n\) substrates with pixel dimensions of \(100 \times 150 \mu m^2\). This size leads to a spatial resolution of \(15 - 20 \mu m\).

In total, the pixel detector consists of about 66 million pixels covering an active silicon area of about \(1 m^2\).

Strip Tracker

The remaining tracker subsystems form the silicon strip tracker. Its inner part consists of the TIB with four barrel layers of silicon modules and the TID representing two endcaps, which contain three discs each. This inner part is surrounded by the TOB containing six cylindrical detector layers. The inner strip tracker and the TOB are complemented in the forward regions by the TECs, two endcaps, which are each formed by nine discs with up to seven rings of silicon modules on them. In total, the strip tracker has an active silicon area of 198 \(m^2\) covered by 9.3 million strips.

Detector modules in the strip tracker are single-sided \(p\text{-on-}n\) type silicon micro-strip sensors made of 6-inch wafers. The modules are rectangular in the barrel part and wedge-shaped in the endcaps. On the wafer front sides, strip-shaped \(p^+\) implantations into the \(n\) bulk form diodes. There is a uniform \(n^+\) implantation, covered by aluminium, on the back sides to create ohmic contact to the positive voltage. The CMS tracker is designed to operate at a temperature of \(-10^\circ C\) to reduce the effects of radiation damage to the silicon detector modules.
To allow 2-dimensional hit measurements, stereo modules are formed by mounting additional modules back-to-back onto some of the detector modules. These extra modules are rotated by 100 mrad with respect to the original module. This allows a hit-position determination in $z$ direction with a precision of 230 $\mu$m in the TIB and 530 $\mu$m in the TOB. The accuracy of $r$ measurements varies in the endcap regions. Stereo modules are placed in layers 1 and 2 of the TIB and TOB, rings 1 and 2 of each disc in the TID and rings 1, 2 and 5 in the TECs.

For high-momentum tracks (100 GeV) in the range of $|\eta| < 1.6$, the $p_{T}$ resolution of the CMS tracker varies between 1% and 2%. The transverse impact parameter resolution of these tracks is about 10 $\mu$m \[38\]. Measurements with early collision data confirm these resolution estimates \[43\].

Tracking efficiency is degraded by multiple scattering, which is caused by the material inside the tracker. The material budget of the tracker has therefore been kept as low as possible. A non-negligible amount of support material is however necessary to provide electrical powering of about 60 kW, to cool the resulting heat and to give mechanical stability. Beginning at $|\eta| = 0$, the material budget of the tracker increases from 0.4 radiation lengths, $X_0$, to its maximum of about 1.8 $X_0$ at $|\eta| \approx 1.4$. Beyond the barrel-endcap transition region it decreases again to about 1.0 $X_0$ at $|\eta| \approx 2.5$ \[38\]. Using collision data, the tracking efficiency has been measured to 98.8% for muons emerging from $J/\Psi$ decays \[44\].

### 2.2.3 Electromagnetic Calorimeter

The CMS electromagnetic calorimeter \[45\] is a high-resolution, high-granularity detector made of lead tungstate (PbWO$_4$) crystals. Lead tungstate is a fast scintillator, which provides a small Molière radius and short radiation length.

The geometrical coverage of the electromagnetic calorimeter extends up to $|\eta| = 3.0$. The calorimeter is divided into a barrel, which covers the range of $|\eta| < 1.479$, and two endcaps, which cover the rapidity range of 1.479 < $|\eta|$ < 3.0. The crystal front-face area of $22 \times 22 \text{mm}^2$ matches the Molière radius of lead tungstate (21.9 mm). This area corresponds to a granularity of $\Delta \eta \times \Delta \phi = 0.0175 \times 0.0175$ in the barrel region and 0.0175 $\times$ 0.0175 up to 0.05 $\times$ 0.05 in the endcaps.

The scintillator crystals in the barrel part of the detector are 23 cm thick, which corresponds to 26 radiation lengths. In the endcap region, 22 cm thick crystals are complemented by preshower detectors. These consist of lead absorber and silicon detector layers, and they cover a range of 1.65 < $|\eta|$ < 2.6. The preshower detectors allow the identification of neutral pions and improve the position measurement of electrons and photons.

In total, the electromagnetic calorimeter consists of 82 728 crystals, 61 200 of them in the barrel part, and sums up to a volume of 11.18 m$^3$.

Scintillation signals are detected by photodetectors. Avalanche photodiodes, which provide gain even within high transverse magnetic fields, are used in the barrel region of the calorimeter. Vacuum phototriodes have been chosen for the endcaps, since these are able to resist the higher integrated radiation dose in these regions.

The photodetectors are readout by a 16-bit readout chain. The corresponding energy range extends from about 30 MeV in the barrel part and about 150 MeV in the endcaps up to about 2 TeV per crystal.

The energy resolution can be parameterised by

$$
(\sigma_E/E)^2 = \left(\frac{0.8}{\sqrt{E}}\right)^2 + \left(\frac{0.12}{E}\right)^2 + (0.5\%)^2 \quad (E \text{ in GeV})
$$

(2.4)
In this parametrisation, $a = 2.8\%$ is a stochastic term including fluctuations in the shower containment and photostatistics, $\sigma_\eta = 0.12$ describes noise caused by electronics or pileup energy, and $c = 0.30\%$ is a constant that characterises energy leakage and intercalibration errors. These coefficients were found to be typical values for electrons with momenta between 20 GeV and 250 GeV during a test beam study in 2004 \[38\]. During the first proton collisions it was found that the electromagnetic calorimeter is about 99\% functioning and that this estimated resolution is actually being achieved \[10\].

2.2.4 Hadronic Calorimeter

The CMS hadronic calorimeter is a sampling calorimeter \[47\] that consists of brass and stainless steel absorbers and plastic scintillators. Its dynamic energy range of 5 MeV to 3 TeV allows the observation of single muons as well as the measurement of the highest possible particle energies.

The hadronic calorimeter is arranged in a central calorimeter, which covers the range of $|\eta| < 3$ and contains the Hadronic Barrel (HB), the Hadronic Endcaps (HE), and the Outer Hadronic calorimeter (HO), complemented by forward and backward calorimeters (HF) in the range of $3 < |\eta| < 5$.

The HB is divided into two half barrels, each containing 18 identical wedges. The wedges are made of absorber plates, complemented by 17 layers of plastic scintillators. The absorber plates consist of brass to maximise the hadronic interaction length, except for the innermost and the outermost plates, which consist of stainless steel for stability reasons. The thickness of the HB is restricted to 100 cm because it is located within the magnet coil. In the HB, the granularity of the scintillators is $\Delta \eta \times \Delta \phi = 0.087 \times 0.087$.

The Hadronic Endcaps cover the range of $1.3 < |\eta| < 3.0$. Each endcap consists of absorber plates and scintillators arranged in sectors of $10^\circ$ angular size. There are 18 absorber layers, each 80 mm thick. The innermost and outermost layer is, as in the barrel part, made of stainless steel, while the other layers consist of brass. The scintillator granularity is the same as in the barrel part, except for the highest $\eta$ regions. In these regions the granularity matches the granularity of the electromagnetic calorimeter.

The HO is placed outside the magnetic coil and envelopes the first iron layer of the magnetic-flux return yoke. It contains one sampling layer in the endcap region, two layers in the barrel region and an additional layer in the central region of $|\eta| < 0.4$. The Outer Hadronic Calorimeter is essential for the full containment of hadron showers.

The total absorber thickness of the hadronic calorimeter corresponds to 5.15 hadronic interaction lengths, $\lambda$, at $\eta = 0$ up to 9.1 $\lambda$ at $\eta = 1.3$ in the barrel region. In the endcaps it averages at about 10.5 $\lambda$.

The HF calorimeters are positioned at a distance of about 11 m from the interaction point and are needed for identification and reconstruction of very forward jets. Due to the high radiation field in this region, quartz fibres were chosen as active material. When hadrons traverse the quartz material, it emits Cherenkov light, which is afterwards detected by photodiodes. The fibres are placed between 5 mm thick steel absorber plates with a total thickness of 10 hadronic interaction lengths.

During test beam studies, the energy resolution of the CMS calorimeter system (including the electromagnetic calorimeter) was found to be

$$\left(\frac{\sigma_E}{E}\right)^2 = \left(\frac{100\%}{\sqrt{E}}\right)^2 + (4.5\%)^2 \quad (E \text{ in GeV}) \quad (2.5)$$

for hadrons in the energy range of 30 GeV $< E < 1$ TeV \[38\]. At the time of first collisions, more than 99\% of the HCAL channels were found to be functional. Measurements with
first collision and cosmic muon data indicate that the HCAL is functioning with the expected performance [48].

2.2.5 Muon Chambers

The purpose of the muon system [49] is the identification, triggering and momentum measurement of muons. To ensure the hermeticity of the detector, an acceptance of nearly 100% is demanded. Therefore, the system is designed with high redundancy, and it complements high precision muon detectors with sensors dedicated for fast muon triggering.

Drift Tubes

Four layers of aluminium Drift Tube (DT) chambers are used for precision measurements in the barrel part of the CMS muon system, covering the range of $|\eta| < 1.3$. The barrel muon system consists of 250 DT chambers, which are 2.5 m long. There are 12 layers of drift tubes inside each chamber, four of them measuring the $z$ coordinates of crossing muons, and eight layers that are sensitive in $\phi$ direction. The $r\phi$ precision of a single DT layer is about 250 $\mu$m providing a total accuracy of about 100 $\mu$m per chamber.

Cathode Strip Chambers

The magnetic field and the occupancy is expected to be higher in the endcap part of the CMS muon system than in the barrel region. Therefore, Cathode Strip Chambers (CSC), which allow space and time precision measurements in the presence of a high and varying magnetic field, are used here. They cover the region between $|\eta| = 0.9$ and $|\eta| = 2.4$. These chambers have a faster response time than drift tubes. Separated by iron discs of the flux return yoke, there are four CSC stations consisting of trapezoidal shaped chambers that are arranged in concentric rings in each endcap. The chambers contain seven cathode strip panels with six layers of wires in between. A robust pattern recognition is used to match detected muons with tracks and to reject non-muon background.

The CSCs are designed to fit accuracy requirements of 75 $\mu$m in the innermost and 150 $\mu$m in the remaining stations. In total, the endcaps are made up of 540 CSC chambers with about 1000 readout channels each.

Resistive Plate Chambers

In addition to precision-measurement muon detectors (DTs and CSCs), Resistive Plate Chambers (RPC) are integrated both into the barrel and endcaps of the CMS detector. They cover a range of $|\eta| < 2.1$. There are six RPC layers in the barrel region and four layers in each endcap. The RPCs are dedicated muon trigger detectors with a very fast response time of about 3 ns, which allows unambiguous assignment of a detected muon to a bunch crossing. An RPC consists of parallel plates of phenolic resin, which is coated with conductive graphite paint, with a distance of about 3 mm to each other. The sensors are readout by aluminium strips outside the resin plates.

Due to high segmentation, the RPCs can provide measurement of muon transverse moments at trigger time.

Combined Resolution

The best momentum resolution is achieved if measurements of the silicon tracker and the muon system are combined. At low muon momenta, the resolution is dominated by multiple scattering effects in the tracker. For muons with $p_\perp = 10$ GeV, a resolution of $\Delta p_\perp / p_\perp \approx 1\%$ is achieved in the central region of $|\eta| < 0.8$. At high muon momenta, the resolution of the muon system is dominant. For muons with a momentum of 1 TeV,
the momentum resolution yields about 4% [38]. Measurements with 2010 collision data indicate that this resolution can actually be achieved [50].

To monitor positions and deformations of the muon chambers, an optical alignment system is installed. It also determines the position of the silicon tracker in relation to the muon systems. The precision of the alignment system is comparable to the resolution of the muon chambers.

2.2.6 Data Acquisition and Computing

At LHC design luminosity ($10^{34}$ cm$^{-2}$s$^{-1}$), an event rate of about $10^7$ Hz is expected. The mass-storage systems of the CMS experiment however are able to write data at a maximum sustained event rate of about 300 Hz [51] to tape. This corresponds to about 450 MB/s. Therefore, a trigger system [52] is needed to filter events down to this rate.

CMS filters events in two steps: the Level-1 Trigger (L1), implemented by custom electronics, reduces the event rate to a maximum rate of 100 kHz, which can be handled by the second trigger level. To provide a safety margin, an L1 output rate of 30 kHz is targeted. The remaining filtering is done by the High Level Trigger (HLT), which is a farm of commercial processors.

The data acquisition, transfer and analysis is done by the CMS Data Acquisition (DAQ) system. Figure 2.5 gives an overview on the DAQ architecture.

**Figure 2.5:** Schematic overview of the CMS Trigger and DAQ system [38]. The boxes labeled “Filter Systems” represent HLT computing nodes.

### Level 1 Trigger

The Level-1 trigger has to reduce the event rate by a factor of about $10^3$. It is realised as deadtime-less electronic pipelines, which perform all calculations for each bunch crossing.

The event data are held in pipelines for 3.2 µs (128 bunch crossings at 25 ns bunch spacing). After this time, the trigger decision whether or not to keep the event and to transport its data to the HLT, has to be available. Therefore only promptly available data can be used for the L1 decision. This excludes information from the tracker and the preshower detectors. The trigger decisions are based upon pre-defined thresholds for, for instance, particle counts, transverse momentum or missing transverse energy, $E_T$.

### High-Level Trigger

Data belonging to events that have been selected by the Level-1 trigger are read from the front-end readout buffers and put together by an event builder network. Afterwards the events are distributed to the HLT computer farm to determine HLT decisions. A series of filters is applied to select interesting events. These HLT filters have access to the full
event data including all subdetector measurements. The HLT filter processing takes about 50 ms per event on a single CPU.

The HLT reduces the event rate by a factor of about $10^3$ and writes the remaining 150 Hz-event stream (corresponding to 225 MB/s) to a mass-storage system. During 2010 and 2011 data taking, the HLT filters were continuously adjusted to match the needs of the rising instantaneous luminosity and event rates.

The performance of the DAQ system was tested during cosmic muon data taking in 2009 and matched all expectations. Using random L1 triggering, the HLT input rate was adjusted to 60 kHz. The DAQ system performed stably with an output rate of up to 600 MB/s. An average HLT decision time of 20 ms was observed due to the simple event topology.

**Computing**

After the HLT decision, the remaining event stream is promptly reconstructed at the Tier 0 computing centre at CERN. From there, copies of the data (as well as simulation) are distributed to several Tier 1 and Tier 2 computing centres, which are cross-linked and together form the Worldwide LHC Computing Grid. Using the CMS Remote Analysis Builder (CRAB) tool, the hardware of these centres and the stored data can be accessed to perform data analysis.
3. Data Analysis

In this chapter, an overview over the used data and simulation samples is given, together with details on the event reconstruction and generation processes. Furthermore, the reconstruction of physics objects and the applied quality criteria are described.

3.1 Datasets

This analysis uses the CMS data taken in 2011. The 2011 data was taken during two data-taking periods, Run2011A and Run2011B. The instantaneous luminosity was slowly raised during the data-taking periods themselves and was significantly enlarged between the two periods. Therefore, the average number of interactions per bunch crossing and hence, the amount of pileup, differs distinctly in Run2011A and Run2011B.

Data taking is performed in consecutively numbered runs. These can extend up to the total duration of an LHC fill, which usually is about 12 h. Runs are divided into luminosity sections of about 20 s of data taking. Collision events can be uniquely identified with their run number, luminosity section and event number within the luminosity section.

The integrated luminosity of the recorded data is measured by counting the number of clusters reconstructed in the silicon pixel detector. A bunch-by-bunch luminosity estimate is derived from these measurements \[56, 57\]. Results are corrected for afterglow effects, which are caused by late-arriving particles and activated detector material. An absolute calibration of the pixel cluster cross section is performed using Van der Meer scans. The systematic uncertainty on the luminosity measurement is 2.2\%.

Data are certified for analysis use only if all CMS subdetectors performed properly during the data taking. This quality control is carried out by luminosity section and rejects about 5\% of the recorded data. The total integrated luminosity of analysis-certified data is

\[
L_{\text{int}} = (4.98 \pm 0.21) \text{ fb}^{-1}.
\]  

(3.1)

After the data taking, events are sorted into data streams by the CMS trigger system. Four data streams are used by this analysis. The DoubleElectron data stream comprises events that were triggered by a double electron trigger. The DoubleMu, MuEG and HT data streams similarly contain the results of double muon triggers, muon electron triggers and \(H_T\) triggers, respectively.
Table 3.1 lists the collection of datasets used in this analysis. This collection consists of several datasets that have been reconstructed promptly during the data taking (PromptReco) and other datasets that have been re-reconstructed afterwards (ReReco) to correct for errors during the prompt reconstruction. The datasets are stored in the CMS bookkeeping system DBS \cite{DBS} and are identified and accessed using DBS datasetpaths.

Table 3.1: Datasets used in this analysis.

<table>
<thead>
<tr>
<th>DBS datasetpath</th>
<th>run range</th>
<th>luminosity [fb$^{-1}$]</th>
</tr>
</thead>
<tbody>
<tr>
<td>/DoubleElectron/Run2011A-May10ReReco-v1/AOD</td>
<td>160329 – 163869</td>
<td>0.22</td>
</tr>
<tr>
<td>/DoubleElectron/Run2011A-PromptReco-v4/AOD</td>
<td>165071 – 168437</td>
<td>0.96</td>
</tr>
<tr>
<td>/DoubleElectron/Run2011A-PromptReco-v6/AOD</td>
<td>172620 – 175770</td>
<td>0.71</td>
</tr>
<tr>
<td>/DoubleElectron/Run2011B-PromptReco-v1/AOD</td>
<td>175832 – 180296</td>
<td>2.71</td>
</tr>
<tr>
<td>/DoubleMu/Run2011A-May10ReReco-v1/AOD</td>
<td>160329 – 163869</td>
<td>0.22</td>
</tr>
<tr>
<td>/DoubleMu/Run2011A-PromptReco-v4/AOD</td>
<td>165071 – 168437</td>
<td>0.96</td>
</tr>
<tr>
<td>/DoubleMu/Run2011A-05Aug2011-v1/AOD</td>
<td>170053 – 172619</td>
<td>0.39</td>
</tr>
<tr>
<td>/DoubleMu/Run2011A-PromptReco-v6/AOD</td>
<td>172620 – 175770</td>
<td>0.71</td>
</tr>
<tr>
<td>/DoubleMu/Run2011B-PromptReco-v1/AOD</td>
<td>175832 – 180296</td>
<td>2.71</td>
</tr>
<tr>
<td>/MuEG/Run2011A-May10ReReco-v1/AOD</td>
<td>160329 – 163869</td>
<td>0.22</td>
</tr>
<tr>
<td>/MuEG/Run2011A-PromptReco-v4/AOD</td>
<td>165071 – 168437</td>
<td>0.96</td>
</tr>
<tr>
<td>/MuEG/Run2011A-05Aug2011-v1/AOD</td>
<td>170053 – 172619</td>
<td>0.39</td>
</tr>
<tr>
<td>/MuEG/Run2011A-PromptReco-v6/AOD</td>
<td>172620 – 175770</td>
<td>0.71</td>
</tr>
<tr>
<td>/MuEG/Run2011B-PromptReco-v1/AOD*</td>
<td>175832 – 180296</td>
<td>2.71</td>
</tr>
<tr>
<td>/HT/Run2011B-PromptReco-v1/AOD*</td>
<td>175832 – 180296</td>
<td>0.96</td>
</tr>
</tbody>
</table>

While the double lepton triggers are used to select the signal sample, the HT data stream only provides control samples to measure trigger efficiencies. Therefore, the completeness of the HT datasets is not needed in this analysis.

3.2 Monte Carlo Simulation

Monte Carlo event generation is a common technique in particle physics to simulate collisions \cite{MC}. Collision events with predefined underlying physics processes (Standard Model and non Standard Model) can be simulated, including detector response. The event simulation is performed in three steps, hard scattering process, hadronisation and detector simulation, which are detailed in the following.

3.2.1 Hard Scattering and Parton Distribution Functions

The cross section of a scattering process of two elementary particles (hard scattering process) can be determined by calculating the process matrix element in perturbation theory (see e.g. \cite{HSPDF}). For proton collisions, and hadron collisions in general, this is more complex, because the inner hadron structure has to be considered. The inner structure of protons is described by Parton Distribution Functions (PDFs). These characterise the probability that the proton constituents (three valence quarks, sea-quarks and gluons) take part in a scattering process. Following the factorisation theorem, proton-proton collision cross sections are given by the cross section of hard processes convoluted with the proton PDFs, $f_i^p$:

$$
\sigma(pp \rightarrow C) = \sum_{i,j} \int dx_1 dx_2 \cdot f_i^p(x_1, Q^2) \cdot f_j^p(x_2, Q^2) \cdot \hat{\sigma}(ij \rightarrow C). \quad (3.2)
$$

$x_1$ and $x_2$ denote the momentum fraction of the partons $i$ and $j$, $Q^2$ the momentum scale of the interaction and $\hat{\sigma}$ the cross section of the hard scattering process between the two partons.
Parton distribution functions have been extensively measured in electron-proton deep-inelastic scattering at the HERA collider [61]. Due to PDF universality, these measurements can be used to estimate cross sections at proton-proton and proton-antiproton colliders like LHC and Tevatron. The PDF momentum scale dependence is described by the DGLAP (Dokshitzer, Gribov, Lipatov, Altarelli, Parisi) equations [62–64]. Following these, measurements can be extrapolated into kinematic regimes of other colliders.

Figure 3.1 shows parton distribution functions deduced from HERA measurements for a momentum scale of $Q^2 = 10 \text{ GeV}^2$. The HERA measurements have been fitted independently by several groups, e.g. HERAPDF [61], CTEQ [65] and MRST [66]. CMS uses the CTEQ6.1 PDFs [67] for the production of its Monte Carlo simulation.

**Figure 3.1:** Parton distribution functions from HERAPDF1.0 for $u$ and $d$ valence quark, sea quark and gluon contribution [61]. Sea quark and gluon distributions are scaled.

The hard scattering process itself is calculated using an event generator, such as Pythia [68], Madgraph [69] or Powheg [70].

In the following parton shower step, initial and final state gluon radiation is simulated in addition to the outgoing particles of the leading-order interaction. Soft radiation is non-perturbative and cannot be analytically calculated. Several phenomenological models have been developed to describe soft gluon radiation and parton showers and have been implemented in Pythia. Details can be found in [68].

A slightly different approach is the matrix element and parton shower matching. Hard, large angle emissions are calculated here using tree-level matrix elements. Soft emissions are still modeled by a full parton shower. A matching algorithm then assures that the two types of emissions blend into each other. Madgraph performs matrix element and parton shower matching using the MLM algorithm [71] to match the calculated hard emissions to the parton shower modeled with Pythia. Powheg follows a similar strategy and allows the implementation of specifically process-tuned radiative corrections before the application of the parton shower [72, 73].

### 3.2.2 Hadronisation

After the scattering, the produced partons fragment into hadrons, which contribute to the final state of the collision. This fragmentation process is called hadronisation. It is
described by non-perturbative QCD and can only be simulated using phenomenological models. Two common approaches are the string model and the cluster model.

In this analysis, Pythia is used for the hadronisation process, which implements the string model. Following this model, color-charged particles that emerge from the parton shower are connected with strings that incorporate the color-potential and can also contain intermediate gluons. These strings break by quark-antiquark production as they grow and release the contained color-charged particles. In a final step, the produced particles are clustered into non-colored hadron multiplets.

3.2.3 Detector Simulation

Using Geant4 [74], the detector and interactions of the produced particles with detector material, e.g. multiple scattering, photon conversions and bremsstrahlung, are simulated. A detailed geometrical model of the CMS detector has been created for this procedure.

From interactions of the particles that are produced during hard-scattering and hadronisation with the detector material, hits in the detector are modeled. In a following step, called digitisation, the response of the detector electronics to these hits is simulated. Afterwards, all information that is provided by reading out the detector during real collisions is available for simulated events.

3.2.4 Fast Simulation

To generate large amounts of simulated collision samples, e.g. parameter scans of SUSY signals, CMS has developed the Fast Simulation framework [75]. Fast Simulation uses simplified geometrical models of the subdetectors and parametrisations, which are tuned to match the full simulation, for the detector response. It allows about 100 times higher production rates than the Geant4-based (full) simulation while it still provides comparable accuracy.

The tracker is modeled using several nested cylinders of active detector area instead of thousands of individual sensor modules. Simulated hits are generated along particle trajectories at intersections with the active area and kept with an efficiency measured in full simulation.

For electrons and photons, the calorimeters are modeled as homogenous masses in the simulation of particle showers. The deposited energy is distributed according to the actual crystal geometry. Afterwards HCAL leakage, noise hits and zero suppression are simulated. For charged and neutral hadrons, the calorimeter energy response is taken from pions in full simulation, and the energy is smeared across the crystals following parametrised shower profiles.

Muons are propagated through the tracker, the calorimeters and the solenoid into the muon chambers taking the magnetic field, multiple scattering and ionisation energy loss into account. The same muon chamber geometry is used to simulate detector hits as in the full simulation. A lack of low-$p_T$ muons is observed compared to full simulation, because Fast Simulation does currently not simulate hadron in-flight decays and hadron punch-through. This difference can be neglected after the application of loose muon ID cuts.

The final reconstructed physics objects have the same format in Fast Simulation and the full simulation.
3.2.5 Pile-up Simulation

During 2011 running, the number of average proton-proton interactions per LHC bunch crossing rose to about ten interactions. Thus, each recorded collision event is overlayed with a couple of additional collision events, called pile-up. Pile-up effects can result from pile-up collisions in the same bunch crossing (in-time pile-up) or in the previous bunch crossings (out-of-time pile-up).

In order to reproduce conditions in data as best as possible, pile-up is also implemented in simulated collision events. After an event has been generated, it is overlayed with a number of minimum-bias events. For the number of added pile-up events, a probability density function is chosen that is expected to match the final distribution of pile-up collisions in the data.

It is impossible to guess the exact pile-up distribution in data beforehand, at production time of the MC samples. Also the pile-up conditions change during data-taking with each adjustment of the beam intensities. Therefore, the simulated samples are re-weighted afterwards to match the collected data. The re-weighting procedure is detailed in Section 3.6.1.

3.2.6 Background Samples

All Standard Model background samples used in this analysis were produced in the CMS Fall11 simulation campaign. For generation, digitisation and event reconstruction the CMS software framework release CMSSW.4_2_9_HLT1_patch1 was used.

To generate $t\bar{t} + jets$, $DY + jets$ (Drell-Yan) and diboson processes, Madgraph was used. The $t\bar{t} + jets$ processes were generated with up to three additional jets in the matrix element calculation, the $DY + jets$ processes with up to four additional jets and the diboson processes with up to one additional jet. $DY + jets$ processes were generated separately for centre-of-mass energies of the colliding partons, $\sqrt{s}$, lower and higher than 50 GeV. All Madgraph generated samples were interfaced to Pythia for a full parton shower modeling using the MLM matching algorithm. Single top processes were generated using Powheg including radiative corrections specifically for these processes [76, 77]. A summary of the used software versions is given in Section 3.3.

All simulated samples are scaled to the correct integrated luminosity using NLO cross sections [78–80]. Table 3.2 gives an overview on the used background simulation samples. A list of the full DBS dataset paths can be found in Appendix A.

3.2.7 Signal Samples

The signal samples were produced in the CMS Summer11 simulation campaign. SUSY events were generated with SOFTSUSY [27], SUSYHIT [30] and Pythia. Simplified models were generated using just Pythia. All samples are scaled to NLO cross sections [81], which are calculated using Prospino [28, 29].

Table 3.3 gives an overview on the used signal samples. A detailed list of the used datasets including their DBS paths can be found in Appendix A. The used software versions are summarised in Section 3.3.

The LMX benchmark scenarios were generated, digitised and reconstructed with the CMS software framework release CMSSW.4_2_3_patch3. For interpretation of results within the T3lh simplified model and the CMSSM, parameter scans were produced using the software release CMSSW.4_2_5 and fast simulation.

The T3lh model scan comprises a grid of parameter points with 10,000 generated events, each. The two varied parameters are the gluino mass, $m_{gl}$, and the LSP mass, $m_{LSP}$.
Table 3.2: Background simulation samples used in this analysis.

<table>
<thead>
<tr>
<th>Process</th>
<th>Description</th>
<th># events</th>
<th>NLO cross section</th>
</tr>
</thead>
<tbody>
<tr>
<td>$tt + jets$</td>
<td>Madgraph up to 3 jets</td>
<td>12,805,954</td>
<td>157.5 pb</td>
</tr>
<tr>
<td>$t\bar{t}$</td>
<td>Powheg</td>
<td>7,637,322</td>
<td>85 pb</td>
</tr>
<tr>
<td>$DY + jets$</td>
<td>Madgraph up to 4 jets lepton</td>
<td>28,231,804</td>
<td>3048 pb</td>
</tr>
<tr>
<td></td>
<td>$m(\ell\ell) &gt; 50$ GeV</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$DY + jets$</td>
<td>Madgraph up to 4 jets lepton</td>
<td>11,812,623</td>
<td>9611 pb</td>
</tr>
<tr>
<td></td>
<td>$m(\ell\ell) \leq 50$ GeV</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$WW$, $WZ$ and $ZZ$</td>
<td>Madgraph up to 1 jet leptonic or semi-lepton decays</td>
<td>8,197,426</td>
<td>8.8 pb</td>
</tr>
</tbody>
</table>

The gluino mass is varied between 100 GeV and 1200 GeV in steps of 25 GeV. LSP masses of 50 GeV up to 50 GeV below the gluino mass are simulated, also with a step width of 25 GeV.

The CMSSM parameter space is scanned in $m_0$ and $m_{1/2}$. The remaining parameters are fixed to $A_0 = 0$, $\tan \beta = 10$ and $\text{sgn} \mu = +1$. For each parameter point, 10,000 events were generated. $m_0$ varies between 40 GeV and 3000 GeV in steps of 20 GeV, and $m_{1/2}$ ranges from 100 GeV to 1000 GeV with the same step size.

Table 3.3: Signal simulation samples used in this analysis.

<table>
<thead>
<tr>
<th>Process</th>
<th>Description</th>
<th># events</th>
</tr>
</thead>
<tbody>
<tr>
<td>LMX Summer11</td>
<td>SOFTSUSY $\rightarrow$ SUSYHIT $\rightarrow$ Pythia CMSSW_4_2_3_patch3</td>
<td>8 – 80 k</td>
</tr>
<tr>
<td>T3lh scan</td>
<td>Pythia CMSSW_4_2_5 (FASTSIM)</td>
<td>$\sim$ 10 M</td>
</tr>
<tr>
<td>CMSSM scan</td>
<td>SOFTSUSY $\rightarrow$ SUSYHIT $\rightarrow$ Pythia CMSSW_4_2_5 (FASTSIM)</td>
<td>$\sim$ 65 M</td>
</tr>
</tbody>
</table>

3.2.8 Scaling

The luminosity of all simulated samples is scaled to match the total integrated luminosity of the analysed data. For each simulated process, $X$, the corresponding scaling factor, $f_{\text{scale}}(X)$, is calculated by

$$f_{\text{scale}}(X) = \frac{L_{\text{int}} \cdot \sigma_X}{N_{\text{total}}},$$

with the total integrated luminosity of the recorded data, $L_{\text{int}}$, the process cross section, $\sigma_X$, and the total number of generated events for this process, $N_{\text{total}}$.

3.3 Software Versions

A summary of the software packages used for this analysis is given in Table 3.4.
For generation of SUSY signal events, first the SUSY mass spectra were calculated by SOFTSUSY 3.1.6 [27]. Decay widths and branching fractions were determined using SUSYHIT 1.3 [30]. This information was transferred to and read-in by Pythia 6.4.25 (version 6.4.25) in form of SUSY Les Houches Accord [52] (SLHA) files. Pythia then generated SUSY events using the calculated particle masses and branching fractions.

Standard model processes were generated using Madgraph 5.1.1.0 [69], Powheg 301 [70] and Pythia 6.4.25.

CMSSW versions from 4.2.3_patch3 up to 4.2.5 were used for detector simulation, digitisation and reconstruction of events.

Reconstructed events were processed with the CMS Physics Analysis Toolkit (PAT) [83] into PAT tuples. These served as input for the final end-user analysis that was carried out with the SuSyAachen software package. The exact tags that were used for the PAT and SuSyAachen packages can be found in Appendix B. Histogramming and fitting of the data was done using the ROOT framework [84] in version 5.30.

Table 3.4: Software versions.

<table>
<thead>
<tr>
<th>Software</th>
<th>Version</th>
<th>Task</th>
</tr>
</thead>
<tbody>
<tr>
<td>SOFTSUSY</td>
<td>3.1.6</td>
<td>Determination of SUSY mass spectrum</td>
</tr>
<tr>
<td>SUSYHIT</td>
<td>1.3</td>
<td>Calculation of decay widths and branching fractions</td>
</tr>
<tr>
<td>Madgraph</td>
<td>5.1.1.0</td>
<td>Matrix element</td>
</tr>
<tr>
<td>Powheg</td>
<td>301</td>
<td>Matrix element</td>
</tr>
<tr>
<td>Pythia</td>
<td>6.4.25</td>
<td>Matrix element, parton shower and hadronisation</td>
</tr>
<tr>
<td>CMSSW</td>
<td>4.2_X</td>
<td>Detector simulation</td>
</tr>
<tr>
<td>CMSSW</td>
<td>4.2_X</td>
<td>Digitisation and reconstruction</td>
</tr>
<tr>
<td>CMSSW</td>
<td>4.2_X</td>
<td>Trigger (L1 and HLT)</td>
</tr>
<tr>
<td>CMSSW</td>
<td>4.2_X</td>
<td>PAT</td>
</tr>
<tr>
<td>SuSyAachen</td>
<td>V00-04-57</td>
<td>End user analysis</td>
</tr>
</tbody>
</table>

3.4 Object Selection

In the following, the reconstruction methods along with the selection criteria of the objects that are used in this analysis are described. Technical details on the implementation of the object selection using the CMS Physics Analysis Toolkit can be found in Appendix B.

3.4.1 Muons

The CMS muon reconstruction [50] [85] uses a Kalman Filter algorithm [86] to reconstruct muon tracks from measurements in the muon detectors. To improve the momentum resolution, the track is constrained to origin at the beam spot. These reconstructed muons, relying only on the muon detectors, are called standalone muons.

The muon momentum resolution for muons with a transverse momentum below 200 GeV is driven by measurements in the silicon tracker. Each standalone muon is matched to tracks reconstructed from tracker measurements. With a Kalman Filter algorithm a global muon is reconstructed from the standalone muon and the best-matching tracker track (outside-in reconstruction). Energy loss, multiple scattering and the magnetic field are taken into account.
A second approach starts from tracks in the silicon tracker and extrapolates them into the muon system under consideration of energy loss, multiple scattering and the magnetic field. If at least a short track stub that matches this extrapolation is found in the muon system, the track fulfils the tracker muon criterion (inside-out reconstruction).

For this analysis, global muons that also fulfil the tracker-muon requirement are selected. The geometrical acceptance of the muon chambers limits the muon reconstruction to $|\eta| \leq 2.4$. A transverse momentum of at least 10 GeV is required. The underlying track in the silicon tracker has to have at least 11 hits, and the relative transverse-momentum uncertainty must not exceed 10%. The impact parameter of the muon track with respect to the primary vertex (see Sec. 3.5.1) must not exceed 200 $\mu$m in the transverse plane and 1 cm along the beamline. Table 3.5 gives an overview on the muon selection.

**Table 3.5: Muon selection.**

<table>
<thead>
<tr>
<th>Description</th>
<th>Cut</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acceptance</td>
<td>$p_\perp &gt; 10$ GeV</td>
</tr>
<tr>
<td></td>
<td>$</td>
</tr>
<tr>
<td>Muon ID</td>
<td>global muon</td>
</tr>
<tr>
<td></td>
<td>tracker muon</td>
</tr>
<tr>
<td>Tracker track</td>
<td>$n_{Hits} \geq 11$</td>
</tr>
<tr>
<td></td>
<td>$\sigma_{p_\perp}/p_\perp &lt; 0.1$</td>
</tr>
<tr>
<td>Impact parameter</td>
<td>$</td>
</tr>
<tr>
<td></td>
<td>$</td>
</tr>
</tbody>
</table>

**3.4.2 Electrons**

Two approaches are used to reconstruct electrons: tracker driven and ECAL driven reconstruction[87]. Tracker driven reconstruction is seeded by trajectories in the tracker and searches for ECAL clusters that match the end of the trajectories [88]. ECAL driven reconstruction starts from collections of clusters, so-called superclusters, in the ECAL and searches for matching track seeds in the pixel detector [89]. While tracker-seeded electron reconstruction performs better for low-momentum electrons and electrons inside jets, ECAL seeded reconstruction is optimised for reconstruction of isolated electrons with a momentum of about 5–70 GeV. After the seeding step, the electron track is reconstructed from tracker and ECAL information using a Gaussian Sum Filter algorithm [90], which takes energy loss inside the tracker volume into account.

For this analysis, electrons inside the tracker volume ($|\eta| < 2.5$) with a transverse momentum of at least 10 GeV are selected. The ECAL barrel-to-endcap transition region ($1.4442 \leq |\eta| < 1.566$) is excluded.

Several ID criteria are used to discriminate prompt electrons from fake electrons, electrons emerging from semi-leptonic $b$- or $c$-quark decays and electrons produced in photon conversions [89]. A geometrical matching in $\phi$ and $\eta$ of the supercluster position and an extrapolation from the interaction point using the track parameters is performed, and the differences, $\Delta \phi$ and $\Delta \eta$, must not exceed predefined tolerance levels. Furthermore, the fraction of energy deposited in the HCAL behind the ECAL seed cluster, $H/E$, has to be small. Finally a shower-shape variable is used to suppress non-prompt and fake electrons:

$$
\sigma_{ir\eta} = \sum_{\text{crystals}} (\eta_i - \eta_{seed})^2 \cdot \frac{E_i}{E_{seed}}.
$$

(3.4)
It builds an energy-weighted sum of the quadratic distance in $\eta$ to the seed cluster, considering all ECAL clusters in a $5 \times 5$ grid around the seed cluster. Thus, $\sigma_{i\eta i\eta}$ is a measure for the $\eta$ extent of the electron-candidate energy deposit. Energy deposits of prompt electrons tend to result in small values of $\sigma_{i\eta i\eta}$.

To reject electrons coming from photon conversions, at most one hit may be missing in the innermost detector layers. Further conversion rejection is achieved by searching for a partner track to the electron trajectory. All tracks within $\Delta R = 0.3$ of the electron are selected and compared to the electron track. If a track is found that has the opposite charge, a similar polar angle ($|\Delta \cot \Theta| \leq 0.02$) and a small distance to the electron track it fulfills the requirements for a partner track. The distance is measured in the $x$-$y$ plane at the position in which the (extrapolated) tracks are parallel, and has to be smaller than 0.02 cm for a partner track. The electron is discarded if at least one partner track is found.

To avoid double-counting muons also as electrons, a minimum distance of $\Delta R = 0.1$ to all reconstructed standalone muons is required. The impact parameter of the electron track with respect to the primary vertex must not exceed 400 $\mu$m in the transverse plane and 1 cm along the beamline. An overview on the electron selection including all acceptance and quality cuts is given in Table 3.6.

<table>
<thead>
<tr>
<th>Description</th>
<th>Cut</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acceptance</td>
<td>$p_{\perp} &gt; 10 \text{ GeV}$</td>
</tr>
<tr>
<td></td>
<td>$</td>
</tr>
<tr>
<td>Fiducial volume</td>
<td>$</td>
</tr>
<tr>
<td></td>
<td>$</td>
</tr>
<tr>
<td>Electron ID (barrel)</td>
<td>$</td>
</tr>
<tr>
<td></td>
<td>$</td>
</tr>
<tr>
<td></td>
<td>$H/E \leq 0.1$</td>
</tr>
<tr>
<td></td>
<td>$\sigma_{i\eta i\eta} \leq 0.01$</td>
</tr>
<tr>
<td>Electron ID (endcaps)</td>
<td>$</td>
</tr>
<tr>
<td></td>
<td>$</td>
</tr>
<tr>
<td></td>
<td>$H/E \leq 0.075$</td>
</tr>
<tr>
<td></td>
<td>$\sigma_{i\eta i\eta} \leq 0.03$</td>
</tr>
<tr>
<td>Conversion rejection</td>
<td>$n_{\text{lost hits}} \leq 1$</td>
</tr>
<tr>
<td></td>
<td>No partner track with $</td>
</tr>
<tr>
<td></td>
<td>$\text{dist} \leq 0.02 \text{ cm}$</td>
</tr>
<tr>
<td>Impact parameter</td>
<td>$</td>
</tr>
<tr>
<td></td>
<td>$</td>
</tr>
</tbody>
</table>

### 3.4.3 Taus

Tau leptons have an average life time of only 0.29 ps. They can decay leptonically into a light lepton (electron or muon) and two neutrinos (branching fraction about 35%) or hadronically into one or more mesons and a neutrino [59].

Leptonically decaying taus appear as light leptons accompanied by missing transverse energy (see Sec. 3.4.6) in the detector. In events with already high genuine missing transverse energy, it is basically impossible to distinguish light leptons from leptonically decaying taus. This search selects events with two light leptons and missing transverse energy in
3. Data Analysis

the final state. This selection also covers produced taus with an efficiency similar to the branching fraction of leptonic tau decays.

Hadronically decaying taus are more difficult to identify than light leptons. CMS has developed two main algorithms to identify and reconstruct hadronically decaying taus \cite{91}: the Hadron Plus Strips (HPS) algorithm and the Tau Neural Classifier (TaNC).

The hadronic tau reconstruction efficiency is lower than the reconstruction efficiency for light leptons, and background due to lepton misidentification plays a more important role than for light leptons. Accordingly, hadronically decaying taus are not selected in this analysis. Therefore, in the following, lepton usually means light lepton, which should also be clear from the context.

A dilepton search extended also to hadronic-tau channels can be found here \cite{92} and here \cite{93}.

3.4.4 Jets

Hadronising partons produce jets, collimated showers of hadrons that emerge from the interaction point. To cluster all hadrons that belong to the same jet, a jet finding algorithm is used. This analysis uses a fast implementation \cite{94} of the anti-\kt\ algorithm \cite{95} within the Particle Flow event reconstruction method \cite{96}.

The Particle Flow event reconstruction combines the measurements of all CMS subdetectors to identify and reconstruct all stable particles in an event. An iterative tracking strategy \cite{97} is performed to reconstruct directions and momenta of all charged particles from silicon tracker measurements. Neutral particles are afterwards identified from remaining calorimeter deposits on top of the expected charged-particle energy deposits. The reconstructed particles are referred to as particle candidates, and the final collection of all particle candidates is used as input for the jet finding algorithm.

Jet Finding with the anti-\kt\ Algorithm

The anti-\kt\ algorithm is based on the \kt\ algorithm \cite{98}. Geometrical stability is enhanced with respect to the \kt\ algorithm, so that soft QCD radiation does not disrupt the conical geometry of the reconstructed jets.

The algorithm uses identified particle candidates of the particle flow event reconstruction as input and considers these as starting pseudo jets. In order to cluster these to jets, for a pseudo jet, \(i\), the value \(d_i\) and the distance to another pseudo jet, \(d_{ij}\), is defined:

\[
d_i = p_{\perp}^{-2} \cdot i, \tag{3.5}
\]

\[
d_{ij} = \min(p_{\perp}^{-2} \cdot i, p_{\perp}^{-2} \cdot j) \cdot \frac{(\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2}{R^2}. \tag{3.6}
\]

\(R\) denotes a geometrical parameter describing the radius of the final jets and is 0.5 here.

For a given set of pseudo jets, the minimum of all possible distances is calculated. If \(d_i\) for any pseudo jet is smaller than this minimum, the corresponding pseudo jet is considered a final jet and is removed from the event. Otherwise the two pseudo jets with the minimal distance are merged into one pseudo jet:
\[ p_{\perp k} = p_{\perp i} + p_{\perp j}, \]
\[ \eta_k = \frac{\eta_i \cdot p_{\perp i} + \eta_j \cdot p_{\perp j}}{p_{\perp k}}, \]
\[ \phi_k = \frac{\phi_i \cdot p_{\perp i} + \phi_j \cdot p_{\perp j}}{p_{\perp k}}. \]

This procedure is repeated until all pseudo jets have been removed from the event.

Jet Selection

Clustered jets with a transverse momentum of at least 30 GeV are kept. Jets are restricted to the \( \eta \) acceptance of ECAL and HCAL. Therefore, the jet axis has to be within the range of \(|\eta| < 3.0\).

Jets have to fulfil the FIRSTDATA LOOSE particle flow ID requirement [99] to reject jets that are not physical, but caused by detector malfunction and electronic noise. The FIRSTDATA LOOSE requirement is based on several jet ID variables, which mostly ensure that a jet is detected in more than one subdetector, for example the fraction of the jet energy that is deposited in the ECAL, \( f_{EM} \).

Particle Flow jets are clustered from all identified particle candidates, including leptons. To avoid overlaps between leptons and jets, jets are discarded if they are within \( \Delta R \leq 0.4 \) of a lepton that passes full lepton ID criteria.

Jet Energy Correction and Uncertainty

Measured jet energies typically differ from the corresponding true particle jet energies. The main reason for such deviations is the non-uniform and non-linear calorimeter response. Energy corrections are applied to compensate this effect. In this analysis, the JEC11\_v1 corrections are applied.

The CMS jet energy correction strategy forsees three steps [100]. A MC correction factor, \( C_{MC} \), compensates the bulk of the non-uniformity in \( \eta \) and the non-linearity in \( p_{\perp} \) of the calorimeter response. Residual corrections of the relative and absolute energy scale, \( C_{rel} \) and \( C_{abs} \), account for the remaining small differences between simulation and data. The full transformation of the raw jet four-momentum, \( p_\mu \), to the corrected jet four-momentum, \( p'_\mu \), is given by

\[ p'_\mu = C(p_T, \eta) \cdot p_\mu \]

with

\[ C(p_T, \eta) = C_{MC}(p_{\perp}, \eta) \cdot C_{rel}(\eta) \cdot C_{abs}(p_{\perp} \cdot C_{MC} \cdot C_{rel}). \]

The MC calibration factor, \( C_{MC} \), is derived on a QCD sample simulated with Pythia. Reconstructed jets are spatially matched with MC particle jets in the \( \eta \)-\( \phi \)-plane. From differences between generated and reconstructed jet energies, correction factors for various values of \( p_{\perp} \) and \( \eta \) are calculated.

A correction factor for the relative jet energy scale, \( C_{rel} \), is estimated using a dijet \( p_{\perp} \) balance technique. Dijet events with one jet in the central region (\(|\eta| < 1.3\)) are selected. From the assumption that the event is balanced in \( p_{\perp} \), energy corrections for jets in non-central regions with respect to the central region are derived.

\(^1\)The additional offset correction that is mentioned in [100] can be – and is in this analysis – replaced by a hadronic pile-up subtraction (see Sec. 3.6.2).
The absolute jet energy response is measured using the Missing transverse energy Projection Fraction (MPF) method. For this, $\gamma/Z + \text{jets}$ events are selected. The energy of photons or lepton $Z$ decay products can be accurately measured using the tracker and the ECAL. Assuming that these events are balanced in the transverse plane, the absolute jet energy can be extracted as the recoil of the photon or the decay products of the $Z$.

The $\text{JEC11}_v1$ corrections do not include residual corrections These were not yet available for the 2011 data-taking, when this analysis was set up. The correction factors, $C_{\text{rel}}$ and $C_{\text{abs}}$, are hence set to one in this analysis.

Figure 3.2 displays the MC jet energy correction factor in dependence of $\eta$ and $p_T$. For the particle flow jets that are used in this analysis, the total correction factor averages at about 1.1, independently of the jet momentum, in the region of up to $|\eta| = 2.5$. The jet energy determination relies to a great extent on tracker measurements in this region. Outside this region, the total correction factor yields a maximum of about 1.2 at $|\eta| = 3.0$ for low-momentum jets with $p_T = 50\text{ GeV}$.

![Jet Energy Correction Factor](image)

\textbf{Figure 3.2:} $\eta$ dependence of the MC jet correction factor, $C_{\text{MC}}$, for jet $p_T$ of 50 GeV \text{[a]} and of 200 GeV \text{[b]} [100]. Correction factors for three different jet reconstruction algorithms are shown including the particle flow (PF) jets used in this analysis.

The jet energy scale (JES) uncertainty is also measured using the MPF method [101]. It is determined to be about 7.5% and is in the following treated as a systematic uncertainty. Evaluation of the JES uncertainty is performed by scaling relevant quantities ($H_T$, $E_T$ and jet $p_T$) and selection thresholds for these quantities up and down by the uncertainty amount.

\textbf{Jet Tagging}

Because of the relatively long lifetime of $b$ quarks, $b$ mesons move a measurable distance from the interaction point before they decay ($c\tau \sim 450\mu m$ [39]). Jets emerging from the decay of $b$ mesons can therefore be distinguished from light-quark jets using so-called $b$-tagging algorithms.

In this analysis, the Track Counting (TC) algorithm [102] is used. It relies on the typically large impact parameters of tracks from which $b$-jets are clustered. The algorithm tags a jet as $b$-jet if it contains at least $N$ tracks with an impact parameter significance, $d_0/\sigma_{d_0}$,
larger than a given threshold. \( N = 2 \) results in the high-efficiency mode of the algorithm, which is used here. Setting \( N = 3 \) switches the algorithm into high-purity mode.

The efficiency and impurity of the track counting algorithm has been measured on simulation and data \([102, 103]\), and agreement was found. Figure 3.3 displays the mis-tag rate versus the \( b \)-tag efficiency for various jet types. The significance threshold chosen for this analysis is 3.3, which corresponds to the medium working point. Using this parameter, the \( b \)-tag efficiency is 63\% and the average mis-tag rate 7\% \([102]\).

3.4.5 Transverse Hadronic Activity

An important variable to quantify the jet activity in an event is the transverse hadronic activity, \( H_T \). It is defined as the scalar sum of all jet momenta:

\[
H_T = \sum_{\text{jets}} p_{\perp}.
\]  

(3.12)

3.4.6 Missing Transverse Energy

The initial state of a collision process is balanced in the transverse plane. If the reconstructed final state is not balanced, the amount of energy needed to achieve a balance in the transverse plane, the Missing Transverse Energy (MET or \( \not{E}_T \)), is calculated:

\[
\not{E}_T = |\not{E}_T| = \sum_{i \in \text{particles}} (-\vec{E}_{i\perp}).
\]  

(3.13)

A large amount of missing transverse energy is a hint for particles that left the detector without being detected.

CMS uses three methods for the reconstruction of missing transverse energy: Calo MET, Track-Corrected (TC) MET and Particle Flow (PF) MET \([104]\). While Calo MET relies solely on calorimeter measurements, TC MET adds measurements of the tracker to achieve a more precise measurement than Calo MET. PF MET uses all particle candidates of the particle flow event reconstruction (see Sec. 3.4.4) as input of the calculation.

The \( \not{E}_T \) resolution is estimated using gaussian fits to the \( x \) and \( y \) components of the missing transverse energy, \( \not{E}_{x,y} \), in multi-jet events without any real \( \not{E}_T \). Figure 3.4 shows the resolution as function of the transverse energy, \( E_{\perp} \) for all three reconstruction methods in data and simulation. The resolution in data is slightly worse than the resolution in MC. PF MET has the best resolution and is therefore used in this analysis.
6.5 Effect of multiple interactions

jets to PF calorimetric activity (parametrized by CaloMET) indicates that PF MET above a threshold of 20 GeV. The good agreement of the normalized shapes in Fig. 14 in these events is zero (e.g., the transverse energy resolution compared to the CaloMET, TC MET and PF MET in data and simulation [104]. PF MET has the best resolution.

3.4.7 Lepton Isolation

Only leptons coming from the hard interaction process, called prompt leptons, are of interest for this analysis. Leptons that are produced inside jets or in photon conversions as well as mesons misidentified as leptons are background to these prompt leptons and will be called fake leptons in the following.

To discriminate prompt leptons from fake leptons, a combined relative isolation is used. The transverse momenta of all tracks and the energy deposits in both calorimeters within a cone of $\Delta R = \Delta \eta \times \Delta \phi < 0.3$ are summed up and divided by the transverse momentum of the lepton:

$$I_{so} = \frac{1}{p_{\perp,\text{lepton}}} \left[ \sum_{\text{tracks}} p_{\perp} + \sum_{\text{ECAL}} E_{\perp} + \sum_{\text{HCAL}} E_{\perp} \right]_{\Delta R < 0.3}$$  \hspace{1cm} (3.14)

A pedestal of 1 GeV is subtracted from energy deposits in the ECAL barrel to reduce noise effects. Energy deposits are however restricted to be non-negative.

A tighter isolation cut increases the rejection against fake leptons, but also decreases the amount of accepted prompt leptons. To find the best trade-off, the efficiency of the isolation cut is investigated on a $t\bar{t} +\text{jets}$ MC sample.

The isolation cut efficiency of prompt leptons, $\varepsilon_{iso}$, is determined depending on the isolation cut value. The isolation cut efficiency of leptons coming from heavy-quark decays, $\varepsilon_{HF}$, is also measured in dependency on the isolation cut value. By combination of both measurements, a point in the $\varepsilon_{iso}$-$\varepsilon_{HF}$-plane can be associated with every isolation cut value. By varying this cut value, the isolation performance curve is obtained. The isolation performance curves for electrons and muons are shown in Figure 3.5.

An isolation cut of 0.15 is chosen for both electrons and muons. This results in a prompt-lepton isolation efficiency of about 85% for both lepton flavours. The efficiency of leptons from heavy-flavour decays is about 0.9% for muons and 2.1% for electrons.
3.5. Event Selection

3.5.1 Rejection of non-Collision Events

To suppress background from non-collision events, the reconstruction of at least one vertex with quality criteria as presented in the following is demanded in each event. The vertex reconstruction is based on an adaptive vertex fitting algorithm [105] and deterministic annealing clustering [106] [107]. The deterministic annealing is introduced by a temperature parameter and allows to cope with noisy (e.g. pile-up) environments.

The reconstructed vertex position has to be within 24 cm from the interaction point along the beamline and not farther away from it than 2 cm in the transverse plane. The number of degrees of freedom of the vertex has to exceed four. It is defined as

\[
n_{dof} = \sum_{i \in \text{tracks}} 2 \cdot w_i - 3, \tag{3.15}
\]

where \(w_i\) represents a weight associated with each track. The weights range from zero to one depending on the compatibility of the track with this vertex. The number of degrees of freedom is a good measure to select real proton-proton interactions since it is strongly correlated to the number of tracks that are compatible with the primary interaction region [108].

If more than one vertex is reconstructed per event, the vertex with the highest quadratically summed track momentum is chosen as primary vertex.

The just described requirement of at least one well reconstructed vertex is applied in all event selections that are presented in this work.

3.5.2 Trigger

This analysis focusses on dilepton final states. Dilepton triggers are used to select the signal sample and also most control samples. Table 3.7 lists the used HLT trigger paths for the initial dilepton selection.

The HLT path names contain information on the triggered lepton kind, the minimum required transverse momentum and the applied ID criteria. Some selection criteria were
tightened in later stages of the LHC running to cope with rising event rates due to higher instantaneous luminosity. The tightest transverse-momentum requirements that were applied at trigger level are 17 GeV for the first lepton and 8 GeV for the second lepton, the leptons being sorted by momentum. The used trigger selection is not affected by trigger prescales over the whole run range.

Table 3.7: HLT trigger paths and selection criteria of triggers used for dilepton-event selection.

<table>
<thead>
<tr>
<th>HLT path</th>
<th>Topology</th>
<th>Lepton 1</th>
<th>Lepton 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>HLT_Ele17_CaloIdL_CaloIsoVL_(\rightarrow) ...Ele8_CaloIdL_CaloIsoVL_(\ast)</td>
<td>ee</td>
<td>(p_{\perp}^{\ell_1} &gt; 17) GeV</td>
<td>(p_{\perp}^{\ell_2} &gt; 8) GeV</td>
</tr>
<tr>
<td>HLT_Ele17_CaloIdT_TrkIdVL_CaloIsoVL_TrkIsoVL_(\leftarrow) ...Ele8_CaloIdT_TrkIdVL_CaloIsoVL_TrkIsoVL_(\ast)</td>
<td>ee</td>
<td>(p_{\perp}^{\ell_1} &gt; 17) GeV</td>
<td>(p_{\perp}^{\ell_2} &gt; 8) GeV</td>
</tr>
<tr>
<td>HLT_Ele17_CaloIdT_CaloIsoVL_TrkIdVL_TrkIsoVL_(\leftarrow) ...Ele8_CaloIdT_CaloIsoVL_TrkIdVL_TrkIsoVL_(\ast)</td>
<td>ee</td>
<td>(p_{\perp}^{\ell_1} &gt; 17) GeV</td>
<td>(p_{\perp}^{\ell_2} &gt; 8) GeV</td>
</tr>
<tr>
<td>HLT_DoubleMu6*</td>
<td>(\mu\mu)</td>
<td>(p_{\perp}^{\mu_1} &gt; 6) GeV</td>
<td>(p_{\perp}^{\mu_1} &gt; 6) GeV</td>
</tr>
<tr>
<td>HLT_DoubleMu7_(\ast)</td>
<td>(\mu\mu)</td>
<td>(p_{\perp}^{\mu_1} &gt; 7) GeV</td>
<td>(p_{\perp}^{\mu_1} &gt; 7) GeV</td>
</tr>
<tr>
<td>HLT_Mu13_Mu7_(\ast)</td>
<td>(\mu\mu)</td>
<td>(p_{\perp}^{\mu_1} &gt; 13) GeV</td>
<td>(p_{\perp}^{\mu_2} &gt; 7) GeV</td>
</tr>
<tr>
<td>HLT_Mu13_Mu8_(\ast)</td>
<td>(\mu\mu)</td>
<td>(p_{\perp}^{\mu_1} &gt; 13) GeV</td>
<td>(p_{\perp}^{\mu_2} &gt; 8) GeV</td>
</tr>
<tr>
<td>HLT_Mu17_Mu8_(\ast)</td>
<td>(\mu\mu)</td>
<td>(p_{\perp}^{\mu_1} &gt; 17) GeV</td>
<td>(p_{\perp}^{\mu_2} &gt; 8) GeV</td>
</tr>
<tr>
<td>HLT_Mu8_Ele17_CaloIdL_(\ast)</td>
<td>(\mu e)</td>
<td>(p_{\perp}^{\mu} &gt; 8) GeV</td>
<td>(p_{\perp}^{e_1} &gt; 17) GeV</td>
</tr>
<tr>
<td>HLT_Mu17_Ele8_CaloIdL_(\ast)</td>
<td>(\mu e)</td>
<td>(p_{\perp}^{\mu} &gt; 17) GeV</td>
<td>(p_{\perp}^{e_1} &gt; 8) GeV</td>
</tr>
<tr>
<td>HLT_Mu10_Ele10_CaloIdL_(\ast)</td>
<td>(\mu e)</td>
<td>(p_{\perp}^{\mu} &gt; 10) GeV</td>
<td>(p_{\perp}^{e_1} &gt; 10) GeV</td>
</tr>
<tr>
<td>HLT_Mu8_Ele17_CaloIdT_CaloIsoVL_(\ast)</td>
<td>(\mu e)</td>
<td>(p_{\perp}^{\mu} &gt; 8) GeV</td>
<td>(p_{\perp}^{e_1} &gt; 17) GeV</td>
</tr>
<tr>
<td>HLT_Mu17_Ele8_CaloIdT_CaloIsoVL_(\ast)</td>
<td>(\mu e)</td>
<td>(p_{\perp}^{\mu} &gt; 17) GeV</td>
<td>(p_{\perp}^{e_1} &gt; 8) GeV</td>
</tr>
</tbody>
</table>

For the measurement of lepton fake rates, support triggers are needed, which select single-lepton events with the same lepton-ID criteria that are used for the dilepton triggers. These support triggers are prescaled and may also include a jet requirement to ensure a moderate trigger rate. A list of the support-trigger HLT paths can be found in Table 3.8.

Table 3.8: HLT trigger paths and selection criteria of triggers used for selection of a control sample to determine lepton fake rates.

<table>
<thead>
<tr>
<th>HLT path</th>
<th>Topology</th>
<th>Selection</th>
</tr>
</thead>
<tbody>
<tr>
<td>HLT_Ele8_CaloIdL_TrkIdVL_(\ast)</td>
<td>e</td>
<td>(p_{\perp}^{\ell} &gt; 8) GeV</td>
</tr>
<tr>
<td>HLT_Ele8_CaloIdL_CaloIsoVL_Jet40_(\ast)</td>
<td>e + jet</td>
<td>(p_{\perp}^{\ell} &gt; 8) GeV, (p_{\perp}^{jet} &gt; 40) GeV</td>
</tr>
<tr>
<td>HLT_Mu8_Jet40_(\ast)</td>
<td>(\mu + jet)</td>
<td>(p_{\perp}^{\mu} &gt; 8) GeV, (p_{\perp}^{jet} &gt; 40) GeV</td>
</tr>
</tbody>
</table>

Trigger efficiencies are measured on an event selection triggered by other, largely independent trigger paths. The efficiencies of the dilepton triggers are measured on an \(H_{T}\)-triggered event selection. Table 3.9 shows a list of the used \(H_{T}\) triggers. The trigger prescales of these \(H_{T}\) triggers varied during the data taking.

### 3.6 Pile-up

With rising beam intensity, more and more proton-proton collisions happen during a bunch crossing. Usually only one of these interactions triggers the readout of the event; it is called the primary interaction. Additional interactions are called pile-up.
3.6. Pile-up

Table 3.9: HLT trigger paths and selection criteria of triggers used for selection of a control sample to measure dilepton trigger efficiencies.

<table>
<thead>
<tr>
<th>HLT path</th>
<th>Selection</th>
</tr>
</thead>
<tbody>
<tr>
<td>HLT_HT160_v*</td>
<td>$H_T &gt; 160 \text{ GeV}$</td>
</tr>
<tr>
<td>HLT_HT200_v*</td>
<td>$H_T &gt; 200 \text{ GeV}$</td>
</tr>
<tr>
<td>HLT_HT240_v*</td>
<td>$H_T &gt; 240 \text{ GeV}$</td>
</tr>
<tr>
<td>HLT_HT250_v*</td>
<td>$H_T &gt; 250 \text{ GeV}$</td>
</tr>
<tr>
<td>HLT_HT260_v*</td>
<td>$H_T &gt; 260 \text{ GeV}$</td>
</tr>
<tr>
<td>HLT_HT300_v*</td>
<td>$H_T &gt; 300 \text{ GeV}$</td>
</tr>
<tr>
<td>HLT_HT350_v*</td>
<td>$H_T &gt; 350 \text{ GeV}$</td>
</tr>
<tr>
<td>HLT_HT360_v*</td>
<td>$H_T &gt; 360 \text{ GeV}$</td>
</tr>
<tr>
<td>HLT_HT400_v*</td>
<td>$H_T &gt; 400 \text{ GeV}$</td>
</tr>
<tr>
<td>HLT_HT440_v*</td>
<td>$H_T &gt; 440 \text{ GeV}$</td>
</tr>
<tr>
<td>HLT_HT450_v*</td>
<td>$H_T &gt; 450 \text{ GeV}$</td>
</tr>
<tr>
<td>HLT_HT500_v*</td>
<td>$H_T &gt; 500 \text{ GeV}$</td>
</tr>
<tr>
<td>HLT_HT520_v*</td>
<td>$H_T &gt; 520 \text{ GeV}$</td>
</tr>
<tr>
<td>HLT_HT550_v*</td>
<td>$H_T &gt; 550 \text{ GeV}$</td>
</tr>
</tbody>
</table>

Pile-up is an important issue for the analysis of collision data as it can affect the object and event reconstruction. In this section, the handling of pile-up in this analysis is detailed. Using a re-weighting technique, the pile-up conditions in all MC datasets are matched to the one observed in data. An area-based subtraction method reduces the effect of pile-up to jet reconstruction. Finally the impact of pile-up on the lepton reconstruction is investigated.

3.6.1 Pile-up Reweighting of Monte Carlo Simulation

While the true amount of pile-up collisions in the event is unknown, the amount of reconstructed vertices gives a good indication of the amount of pile-up. Figure 3.6a shows the number of reconstructed vertices against the number of interactions in simulated $t\bar{t} + jets$ events. It can be seen that the reconstruction efficiency of pile-up vertices is not 100\%. However, there is still a strong correlation of the number of reconstructed vertices and the number of interactions in the event.

During the data-taking period, beam intensity and hence pile-up configuration changed multiple times. In order to adjust the MC to the final pile-up configuration in the data, all simulated samples are re-weighted to match the distribution of the number of reconstructed vertices in data. The pile-up re-weighting factor, $f_{pu}$, for a generated event with $\hat{n}_{vertices}$ reconstructed vertices evaluates to

$$f_{pu} = \frac{N_{data}(n_{vertices} = \hat{n}_{vertices})}{N_{MC}(n_{vertices} = \hat{n}_{vertices})}.$$  

(3.16)

$N_{data}(n_{vertices})$ denotes the total number of data events with the given number of vertices and $N_{MC}(n_{vertices})$ the total number of events in the MC sample with the specified number of reconstructed vertices.

To accurately account for out-of-time pile-up, the re-weighting is extended to include information from neighbouring bunch crossings. This so-called 3d re-weighting procedure modifies Equation 3.16 such that MC events are not only matched to data in bins of $n_{vertices}$, but in bins of the tuple $(n_{vertices}^{-1}, n_{vertices}^{0}, n_{vertices}^{+1})$. Here $n_{vertices}^{-1}$, $n_{vertices}^{0}$ and $n_{vertices}^{+1}$ stand for the number of reconstructed vertices in the previous bunch crossing, in the current bunch crossing and in the following bunch crossing, respectively.

Figure 3.6b shows the number of reconstructed vertices in data and MC after the re-weighting procedure. The bulk of the distribution shows very good agreement of data and
3.6 Data Analysis

Some differences are visible in the tail of the distribution. Small differences are possible in regions of low statistics, because the normalisation distribution for the re-weighting is calculated globally for all samples and not individually for each simulated dataset.

![Figure 3.6: Number of reconstructed vertices versus number of interactions in the event for simulated $t\bar{t} + \text{jets}$ events (a). Number of reconstructed vertices in data and the re-weighted Monte Carlo datasets for dilepton events (b).](image)

### 3.6.2 Hadronic Pile-up Subtraction

To reduce the impact of pile-up on jet energy measurements, a jet-area based pile-up subtraction technique is applied. This purely data-driven technique is performed on an event-by-event basis and provides a very effective correction of jet energies for pile-up effects [109]. First an average pile-up energy density, $\rho$, is calculated. Then for each jet a pile-up energy estimate, based on $\rho$ and the jet area, is subtracted.

To perform the subtraction, the effective area of jets are needed as input. Since jets comprise a collection of point-like particles, they have no intrinsic area. The sensible area of a jet can however be estimated by distributing so-called ghost particles uniformly in the $\eta$-$\phi$ plane of the event. The ghost particles carry an infinitesimal momentum and should therefore not alter the jet reconstruction in any way. However, it is important that an infrared-safe jet algorithm is used. Counting the number of ghost particles that have been clustered into a jet then yields a measure of the sensitive jet area. The jet areas are measured in the $\eta$-$\phi$ plane and are hence dimensionless.

The pile-up energy density is calculated as the median of all jet energy densities:

$$\rho = \text{median} \left( \frac{p_{\perp i}}{A_i} \right)_{i \in \text{jets}}.$$  \hspace{1cm} (3.17)

Jet reconstruction with the $k_T$ (and anti-$k_T$) algorithm leads to a large sample of regular soft pile-up jets for each event. Therefore a representative pile-up jet for the calculation of $\rho$ is selected with the median of the jet collection.
The measured energy of a jet belonging to the hard interaction can afterwards be corrected using its sensitive jet area and the pile-up energy density of the event:

$$p_{\perp}^{\text{corr}} = p_{\perp} - A \cdot \rho.$$  \hspace{1cm} (3.18)

This correction is applied directly on the raw momentum of the jet, before the energy corrections described in Section 3.4.4 are applied.

### 3.6.3 Pile-up Dependency of Lepton Reconstruction

While lepton reconstruction and identification is not expected to be severely affected by additional particles in the same event, lepton isolation is. The pile-up dependency of the lepton performance curves (see Sec. 3.4.7) is calculated to estimate the influence of pile-up on lepton isolation. Figure 3.7 displays the lepton performance curves separately for events with less than ten reconstructed vertices and for events with ten or more vertices.

**Figure 3.7:** Isolation performance curve for electrons (a) and muons (b) in events with less than ten and ten or more reconstructed vertices. Markers are placed in isolation cut value intervals of 0.025. The point corresponding to a cut value of 0.15 is indicated with a circle. Statistical errors are shown for the encircled point and are smaller than the marker size.

An increase of pile-up leads to a lower isolation efficiency for prompt leptons as well as for leptons emerging heavy-flavour decays. This is expected since pile-up generally increases the amount of energy that is deposited in the isolation cone of the lepton, while the lepton isolation cut is constant.

The performance curves themselves do not change significantly. The electron performance curve is shifted about 1% point towards lower isolation efficiencies in the high-pile-up bin compared to the low-pile-up bin. The muon performance curve changes only marginally.

The main influence of pile-up on the lepton reconstruction is a shift of the working point along the performance curve. Comparison of the two vertex-count samples shows an isolation efficiency drop of about 8% points for prompt leptons of both flavours. The isolation efficiency of leptons from heavy-flavour decays drops about 1.0% point for electrons and about 0.5% points for muons.

Pile-up affects the lepton isolation, and it results in changes of the isolation efficiency. These changes are of a tolerable amount for this analysis. For higher instantaneous luminosities however, as they are expected for coming LHC runnings, a modification of the lepton isolation calculation is suggested. One possibility would be to calculate the amount of energy inside the isolation cone after indentification and subtraction of pile-up contribution from the event on a particle-based level, which is possible e.g. using the Particle Flow algorithm [88].
3.7 Efficiencies

3.7.1 Electrons and Muons

Lepton efficiencies are measured on a simulated $t\bar{t} + jets$ sample. Simulated efficiencies have been shown to be modeled quite accurately in MC and agree within 2 % with the measurements on data for transverse lepton momenta above 15 GeV. For lower lepton momenta, simulated electron efficiencies agree within 7 % with measurements on data and simulated muon efficiencies within 5 GeV [87, 92].

The electron efficiency is shown in Figure 3.8a against $p_\perp$ and in Figure 3.8b against $\eta$. The muon efficiency is shown in Figures 3.8c and 3.8d. Lepton reconstruction efficiencies after application of the ID criteria, the efficiency of the isolation cut and the efficiency of the final lepton selection, which is the combination of both, are displayed.

![Graphs showing lepton reconstruction efficiency for electrons and muons](image)

**Figure 3.8:** Lepton reconstruction efficiency for electrons (top) and muons (bottom) versus lepton $p_\perp$ (left) and $\eta$ (right) on a $t\bar{t} + jets$ MC sample. The efficiency of the reconstruction with full ID criteria is shown in red, the isolation efficiency in blue and the combination of both in black.

The isolation efficiency for electrons and muons with low momentum is significantly lower than for high-momentum leptons, because the applied relative isolation cut affects low-momentum leptons stronger than high-momentum leptons. For electrons, the reconstruction efficiency also drops for low transverse momenta.

The calorimeter barrel-endcap transition region is excluded for electrons, which results in an efficiency drop at $|\eta| \approx 1.5$. Another efficiency drop is visible for electrons and muons at high $\eta$ values of 2.4. The drop in this region is caused by the end of the pixel detector acceptance.
The isolation efficiency depends strongly on the event topology and drops for more crowded events. $t\bar{t}+jets$ events provide a significant amount of hadronic activity and $E_T$ in addition to promptly produced leptons. Therefore, of all Standard Model processes, they offer the most similar event topology to SUSY events.

### 3.7.2 Lepton Efficiency Ratio

For this analysis, not the individual lepton efficiencies, but the efficiency ratio for muons and electrons is needed (see Sec. 4.3.1). The efficiency ratio can easily be determined from the ratio of events on the $Z$ peak.

A $Z$-dominated event sample is selected by requiring dilepton events with an invariant mass between 60 GeV and 120 GeV. The same lepton requirements as for the signal selection (see Sec. 4.2.1) are applied. To suppress contamination by top processes, a jet veto is applied. The contamination by other processes (mainly diboson) to this selection are less than 1% according to MC simulation. Figure 3.9 shows the invariant-mass distribution of electron and muon pairs for this selection. The invariant-mass distribution of dimuon events is very accurately modeled in the simulation. The dielectron $Z$ peak is, compared with the simulation, somewhat broader and shows a small shift towards lower energies. This results from a remaining mis-calibration of the ECAL-cluster energy correction.

![Figure 3.9: Dilepton mass of dielectron (a) and dimuon (b) events in a jet-veto selection.](image)

On this sample, the number of dielectron events, $n_{ee}$, and the number of dimuon events, $n_{\mu\mu}$, is determined. Using these values, the efficiency ratio, $r_{\mu e}$, can be deduced:

\[
\begin{align*}
    n_{ee} &= (1.226 \pm 0.001 \text{ (stat)}) \cdot 10^6 \\
    n_{\mu\mu} &= (1.564 \pm 0.001 \text{ (stat)}) \cdot 10^6 \\
    r_{\mu e} &= \sqrt{\frac{n_{\mu\mu}}{n_{ee}}} = 1.129 \pm 0.001 \text{ (stat)} \pm 0.113 \text{ (sys)}.
\end{align*}
\]
As a cross check, $r_{\mu e}$ is also determined with the same event selection on $DY + jets$ MC:

$$n_{ee}^{MC} = (1.224 \pm 0.001 \text{ (stat)}) \cdot 10^6, \quad \text{(3.22)}$$

$$n_{\mu\mu}^{MC} = (1.558 \pm 0.001 \text{ (stat)}) \cdot 10^6, \quad \text{(3.23)}$$

$$r_{\mu e}^{MC} = \sqrt{\frac{n_{\mu\mu}^{MC}}{n_{ee}^{MC}}} = 1.128 \pm 0.001 \text{ (stat)} \pm 0.113 \text{ (sys)}. \quad \text{(3.24)}$$

The ratios determined on data and on MC agree very well with each other. Systematic uncertainties on MC event yields due to lepton efficiencies, trigger efficiencies and luminosity measurement cancel out in the ratio calculation.

The $r_{\mu e}$ dependence on the number of jets in the event and the lepton $p_T$ is investigated on data (see Fig. 3.10). For the lepton $p_T$-dependence, the transverse momentum of the second-leading lepton is chosen. The transverse momentum of the leading lepton is required to be larger than 50 GeV to avoid a bias by this lepton.

While the efficiency ratio does not seem to be influenced by the amount of hadronic activity in the event, a variation with the lepton $p_T$ is visible. This effect is caused by the low efficiency of soft electrons (see Sec. 3.7.1). A systematic uncertainty of 10% is assigned to $r_{\mu e}$ to account for this variation. Since leptons with a transverse momentum lower than 20 GeV make only a small contribution to the final event selection, this uncertainty is considered to be conservative.

Finally, the dependence of the efficiency ratio on the event topology is investigated. Instead of a $DY + jets$ MC, the ratio is determined on a $t\bar{t} + jets$ MC. The jet veto is removed, otherwise the same event selection as before is used. The efficiency ratio is determined to

$$r_{\mu e, t\bar{t}}^{MC} = 1.108 \pm 0.007 \text{ (stat)} \pm 0.111 \text{ (sys)}. \quad \text{(3.25)}$$

It slightly differs from the ratios that were determined on data and the $DY + jets$ MC sample, but the deviation is easily covered by the combined uncertainties.

Since $r_{\mu e}$ does not show significant dependence on the event topology and on hadronic activity, it is therefore also valid in other event selections, especially in the control region and the signal region that are defined later in Chapter 4.2. The impact of possible variations in the lepton momentum spectrum are absorbed in the systematic uncertainty that is assigned to this value.

\begin{figure}[h]
\centering
\begin{subfigure}{0.45\textwidth}
\centering
\includegraphics[width=0.9\textwidth]{a.png}
\caption{(a)}
\end{subfigure} \hspace{1cm}
\begin{subfigure}{0.45\textwidth}
\centering
\includegraphics[width=0.9\textwidth]{b.png}
\caption{(b)}
\end{subfigure}
\caption{Muon-to-electron efficiency ratio, $r_{\mu e}$, on data versus number of jets in the event (a) and $p_T$ of the second-leading lepton (b) for the invariant-mass window of 60 GeV to 120 GeV. An additional cut of 50 GeV is applied on the transverse momentum of the leading lepton in (b) to avoid a bias by the leading lepton.}
\end{figure}
3.7.3 Trigger

Knowing the efficiencies of the trigger for the final event selection is important for the uncorrelated-flavour subtraction method (see Sec. 4.3.1), for the determination of limits and for the correct scaling of simulated datasets. The efficiencies of the dilepton trigger selections are measured using an $H_T$-triggered event sample. $H_T$ and dilepton triggers should be mostly uncorrelated.

A control sample is created by selecting events that have triggered at least one of the $H_T$ triggers listed in Table 3.9. From this control sample, dilepton events are selected using the same criteria that are used later for signal selection: both leptons have to have a transverse momentum of at least 10 GeV. Additionally, one of the two leptons has to exceed a transverse momentum of 20 GeV. The trigger efficiency of the dilepton trigger selection can now be measured as the quantile of events in this sample that pass this trigger selection.

This procedure is performed for each of the three light lepton combinations ($ee$, $e\mu$ and $\mu\mu$) separately. Table 3.10 lists the measured trigger efficiencies for the three dilepton channels. The $ee$, $e\mu$, $\mu\mu$ trigger selection is found to be 99%, 94%, 92% efficient, respectively. All Monte Carlo simulated datasets are scaled according to these trigger efficiencies.

Table 3.10: Efficiencies of the dilepton trigger selection.

<table>
<thead>
<tr>
<th>Trigger selection</th>
<th>$\epsilon$ [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>$ee$ paths</td>
<td>$99^{+1}_{-1}$</td>
</tr>
<tr>
<td>$e\mu$ paths</td>
<td>$94^{+2}_{-5}$</td>
</tr>
<tr>
<td>$\mu\mu$ paths</td>
<td>$92^{+2}_{-4}$</td>
</tr>
</tbody>
</table>
4. Search for a Kinematic Edge

A search for kinematic edges in the dilepton invariant mass spectrum is presented. This search has been published in [92] and is an update and an extension of the works by N. Mohr [110].

4.1 Motivation

This search method aims at models with flavour-correlated lepton production from a decaying particle. In the context of SUSY, the neutralino decay, $\tilde{\chi}_2^0 \rightarrow \ell \tilde{\ell} \rightarrow \ell^\pm \ell'^\mp \tilde{\chi}_1^0$, is targeted. This decay results in a characteristic kinematic edge in the invariant-mass spectrum of the lepton pair (see Sec. 1.4.2), which serves as indicator for a potential signal in this search.

Using the full shape information of the lepton-pair invariant mass, the signal of flavour-correlated lepton production is separated from SM background with a fit method that is described in Section 4.3.2. The shape information allows – and in fact this search needs – a larger amount of background in the signal region than counting experiments. Thus, the sensitivity of this search extends to regions in the parameter space that are not reachable with counting experiments due to tighter event selections. As, on the other hand, a specific signal process is targeted, the field of application of this search is not as broad as it is for counting experiments. Therefore, for the analysis of the 2011 CMS data in the opposite-sign dilepton channel, a counting experiment was also developed as complementary analysis to this search [92].

4.2 Event Selection

4.2.1 Preselection

As a preselection for the following analysis, events with two oppositely-charged leptons and at least two jets are selected. If more than two leptons are present in the event, the lepton pair with the highest scalar transverse-momentum sum is considered the primary lepton pair. The transverse momentum of both leptons has to exceed 10 GeV and at least one lepton has to fulfil $p_\perp > 20$ GeV. Furthermore, each dielectron event has to be selected by a dielectron trigger, each dimuon event by a dimuon trigger and each electron-muon event by an electron-muon cross trigger.
Important Quantities

In the following, distributions of relevant quantities for this analysis are investigated for events passing the preselection, and compared to expectations based on MC simulation.

Figure 4.1 shows the invariant mass of leptons pairs, in the following also named dilepton mass, for events in the preselection region. The dilepton mass is displayed for $ee$, $\mu\mu$ and $e\mu$ pairs separately and shows general agreement with the MC expectation. Due to the remaining ECAL-cluster energy mis-calibration, a small shift of the $Z$ peak is visible for $ee$ events. Some further differences of MC and data are expected because of the remaining jet energy mis-calibration. At least two jets are required in the preselection, and the number of jets is affected by the jet energy scale due to the minimum-$p_\perp$ requirement of 30 GeV.

$H_T$ and $E_T$ are the most important hadronic quantities for this analysis, because they define the signal region of this search. Figure 4.2 shows the $H_T$ and $E_T$ distributions for data and the MC expectation after the preselection.

The jet requirement implies an $H_T$ greater than 60 GeV. In general, less data is observed than expected from MC. This effect increases with rising $H_T$ and could be caused by a slight mis-calibration of the jet energy. The deviations are compatible with the JES uncertainty. The $E_T$ distributions in data and simulation also agree within statistical and systematic uncertainties with each other.

Figure 4.3 shows the number of jets and $b$-jets in events after the preselection. Jet multiplicities are well modeled in simulation for low jet and $b$-jet multiplicities. For jet multiplicities larger than six and for $b$-jet multiplicities larger than four, less jets are observed in data than expected from MC. This observation is consistent with the lower $H_T$ observed in data compared with MC.

The $p_\perp$ spectra of the four leading jets are shown in Figure 4.4. Events without a third or fourth jet are placed at 0 GeV in the momentum distribution for the corresponding jet. All measured distributions agree well with the simulation considering the systematic uncertainty due to JES.

4.2.2 Signal Region

The signal region for this analysis is defined by the following $H_T$ and $E_T$ requirements:

**Signal region** $H_T > 300$ GeV, $E_T > 150$ GeV.

Figure 4.5 shows the $H_T$ and $E_T$ distribution for data and MC simulation in the signal region. No significant differences of data and MC expectation outside statistical and systematic uncertainties are found in this region.

The signal region event yields on data and MC can be found in Table 4.1. The main SM contribution in this region is top-quark production, dominated by dileptonic top pair production. Top pair production makes up about 75% of the total yield, single top production another 7%. The remaining about 18% comprise diboson production and Drell-Yan processes. For reference, the hypothetical event yields of the three SUSY benchmark scenarios LM1, LM3 and LM6 are also shown in the table.

In all three dilepton channels, $ee$, $\mu\mu$ and $e\mu$, less data is observed than expected from MC simulation. The largest discrepancy is seen in dimuon events, in which 65 events are observed and 118 events are predicted by simulation. This discrepancy is barely compatible with the uncertainties on the MC prediction: the systematic uncertainties due to jet energy scale, cross section, trigger and lepton efficiencies, and luminosity add up to
4.2. Event Selection

Figure 4.1: Invariant mass of lepton pairs after dilepton preselection in the $ee$ channel (a), $\mu\mu$ channel (b) and $e\mu$ channel (c) for data and MC expectation. The green band in the ratio view represents the combined systematic lepton efficiency, trigger efficiency, luminosity and cross-section uncertainties.
Figure 4.2: $H_T$ (a) and $E_T$ (b) after preselection. The green band in the ratio view represents the combined systematic lepton efficiency, trigger efficiency, luminosity and cross-section uncertainties. JES uncertainties are shown as red, hatched band.

Figure 4.3: Number of jets (a) and number of b-jets (b) in events after preselection. The green band in the ratio view represents the combined systematic lepton efficiency, trigger efficiency, luminosity and cross-section uncertainties.
Figure 4.4: Transverse momentum of the leading jet (a), the second-leading jet (b), the third-leading jet (c) and fourth-leading jet (d) for data and MC in the preselection. The green band in the ratio view represents the combined systematic lepton efficiency, trigger efficiency, luminosity and cross-section uncertainties. JES uncertainties are shown as red, hatched band.
Table 4.1: Data and MC yields in the signal region defined by \( H_T > 300 \) GeV and \( \mathcal{E}_T > 150 \) GeV. Displayed uncertainties on the MC yields are statistical uncertainties and combined systematic uncertainties due to jet energy scale, cross section, trigger and lepton efficiencies, and luminosity.

<table>
<thead>
<tr>
<th>Process</th>
<th>( ee )</th>
<th>( \mu\mu )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( Z + ) jets</td>
<td>8.4 ± 2.1 ± 2.2</td>
<td>10.1 ± 2.3 ± 4.3</td>
</tr>
<tr>
<td>WW</td>
<td>5.0 ± 0.3 ± 1.2</td>
<td>6.2 ± 0.4 ± 1.8</td>
</tr>
<tr>
<td>WZ</td>
<td>3.2 ± 0.1 ± 0.9</td>
<td>4.2 ± 0.1 ± 1.1</td>
</tr>
<tr>
<td>ZZ</td>
<td>0.7 ± 0.1 ± 0.3</td>
<td>0.9 ± 0.1 ± 0.1</td>
</tr>
<tr>
<td>( tt + ) jets</td>
<td>71.4 ± 2.1 ± 26.5</td>
<td>89.1 ± 2.3 ± 35.1</td>
</tr>
<tr>
<td>( t/\bar{t} + ) jets</td>
<td>6.2 ± 0.6 ± 1.9</td>
<td>7.6 ± 0.6 ± 2.1</td>
</tr>
<tr>
<td>Total MC</td>
<td>94.9 ± 2.1 ± 32.9</td>
<td>118.0 ± 2.4 ± 44.3</td>
</tr>
</tbody>
</table>

\( \mathcal{E}_T > 300 \) GeV requirement (a) and \( \mathcal{E}_T < 150 \) GeV requirement (b). The green band in the ratio view represents the combined systematic lepton efficiency, trigger efficiency, luminosity and cross-section uncertainties. JES uncertainties are shown as red, hatched band.

**Figure 4.5:** \( H_T \) in the signal region before applying the \( H_T > 300 \) GeV requirement (a) and \( \mathcal{E}_T \) before the \( \mathcal{E}_T > 150 \) GeV requirement (b).
a total of 44 events. In both same-flavour channels together, 141 events are observed. The MC prediction is $212 \pm 3 \text{ (stat)} \pm 77 \text{ (sys)}$ and is also just compatible with the observation.

In Figure 4.2a, a discrepancy between data and MC expectation is visible, and it increases with rising $H_T$. This discrepancy could be the effect of a jet energy mis-calibration in a way that jet energies are systematically measured too low. Moreover, this kind of mis-calibration would result in a lack of events observed in the signal region, because an $H_T$ of 300 GeV is required in this region, and could hence explain the observed discrepancy.

The expected signal yields in the same-flavour channel for the benchmark scenarios, LM1, LM3 and LM6, are 431, 208 and 41, respectively, in addition to the SM model processes. However, no sign of an excess over the SM expectation is seen.

### 4.2.3 Control Region

A top-dominated control region is defined to test the correct functioning of the kinematic fit:

Control region \(100 \text{ GeV} < H_T < 300 \text{ GeV}, 100 \text{ GeV} < \slashed{E}_T < 150 \text{ GeV}, n_{b\text{Jets}} \geq 1\).

Loosened $H_T$ and $\slashed{E}_T$ cut compared to signal region. $b$-tag requirement to suppress background processes other than top quark production.

The $H_T$ and $\slashed{E}_T$ distributions before applying the selection on these variables can be seen in Figure 4.6 for data and MC expectation in the control region. Some discrepancies of data and MC are visible in the $H_T$ distribution. The low-$H_T$ region below 100 GeV seems not very accurately modeled in the simulation for this selection. The rest of the distribution and the $\slashed{E}_T$ distribution show reasonable agreement of data and MC considering statistical and systematic uncertainties.

The event yields in the control region can be found in Table 4.2. The total MC yield agrees with the observation. Simulation indicates that the SM background consists almost exclusively of top processes. The amount of signal that is expected in this region in the benchmark scenarios LM1, LM3 and LM6 is negligible.

**Table 4.2:** Data and MC yields in the control region defined by \(100 \text{ GeV} < H_T < 300 \text{ GeV}, 100 \text{ GeV} < \slashed{E}_T < 150 \text{ GeV} \) and \( n_{b\text{Jets}} \geq 1 \). Displayed uncertainties on the MC yields are statistical uncertainties and combined systematic uncertainties due to jet energy scale, cross section, trigger and lepton efficiencies, and luminosity.

<table>
<thead>
<tr>
<th>Process</th>
<th>$ee$ (Stat)</th>
<th>$\mu\mu$ (Stat)</th>
<th>$e\mu$ (Stat)</th>
<th>$ee + \mu\mu$ (Stat)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Z + \text{jets}$</td>
<td>2.0 $\pm$ 1.0 $\pm$ 1.4</td>
<td>5.4 $\pm$ 1.7 $\pm$ 1.7</td>
<td>3.1 $\pm$ 1.3 $\pm$ 0.7</td>
<td>7.4 $\pm$ 2.0 $\pm$ 1.6</td>
</tr>
<tr>
<td>$WW$</td>
<td>1.4 $\pm$ 0.2 $\pm$ 0.1</td>
<td>2.0 $\pm$ 0.2 $\pm$ 0.3</td>
<td>3.6 $\pm$ 0.3 $\pm$ 0.1</td>
<td>3.4 $\pm$ 0.3 $\pm$ 0.2</td>
</tr>
<tr>
<td>$WZ$</td>
<td>0.7 $\pm$ 0.1 $\pm$ 0.1</td>
<td>0.8 $\pm$ 0.1 $\pm$ 0.1</td>
<td>0.4 $\pm$ 0.1 $\pm$ 0.0</td>
<td>1.5 $\pm$ 0.1 $\pm$ 0.1</td>
</tr>
<tr>
<td>$ZZ$</td>
<td>0.3 $\pm$ 0.1 $\pm$ 0.0</td>
<td>0.2 $\pm$ 0.1 $\pm$ 0.1</td>
<td>0.1 $\pm$ 0.0 $\pm$ 0.0</td>
<td>0.6 $\pm$ 0.1 $\pm$ 0.0</td>
</tr>
<tr>
<td>$t\bar{t} + \text{jets}$</td>
<td>400.4 $\pm$ 4.9 $\pm$ 97.1</td>
<td>509.9 $\pm$ 5.5 $\pm$ 130.1</td>
<td>910.9 $\pm$ 7.4 $\pm$ 227.0</td>
<td>910.3 $\pm$ 7.4 $\pm$ 227.2</td>
</tr>
<tr>
<td>$t\bar{t} + \text{jets}$</td>
<td>17.0 $\pm$ 0.9 $\pm$ 3.2</td>
<td>22.1 $\pm$ 1.1 $\pm$ 4.7</td>
<td>42.2 $\pm$ 1.5 $\pm$ 8.0</td>
<td>39.1 $\pm$ 1.4 $\pm$ 7.9</td>
</tr>
<tr>
<td><strong>Total observed</strong></td>
<td>421.8 $\pm$ 5.0 $\pm$ 100.9</td>
<td>540.4 $\pm$ 5.6 $\pm$ 135.5</td>
<td>960.3 $\pm$ 7.5 $\pm$ 234.3</td>
<td>962.3 $\pm$ 7.5 $\pm$ 236.3</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Total observed</strong></th>
<th>438</th>
<th>562</th>
<th>1005</th>
<th>1000</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>LM1</strong></td>
<td>2.7 $\pm$ 1.3 $\pm$ 1.6</td>
<td>4.3 $\pm$ 1.6 $\pm$ 2.0</td>
<td>2.2 $\pm$ 1.1 $\pm$ 1.6</td>
<td>7.0 $\pm$ 2.0 $\pm$ 3.6</td>
</tr>
<tr>
<td><strong>LM3</strong></td>
<td>1.0 $\pm$ 0.8 $\pm$ 4.1</td>
<td>3.3 $\pm$ 1.5 $\pm$ 2.2</td>
<td>1.2 $\pm$ 0.9 $\pm$ 2.4</td>
<td>4.3 $\pm$ 1.7 $\pm$ 4.2</td>
</tr>
<tr>
<td><strong>LM6</strong></td>
<td>0.9 $\pm$ 0.5 $\pm$ 0.2</td>
<td>0.3 $\pm$ 0.2 $\pm$ 0.6</td>
<td>0.3 $\pm$ 0.3 $\pm$ 0.3</td>
<td>1.2 $\pm$ 0.5 $\pm$ 0.9</td>
</tr>
</tbody>
</table>

### 4.3 Background Estimation

The following SM processes are background to this search and have to be considered:
Figure 4.6: $H_T$ before applying $H_T$ cut (a) and $E_T$ before applying $E_T$ cut (b) in the control region defined by $100 \text{GeV} < H_T < 300 \text{GeV}, 100 \text{GeV} < E_T < 150 \text{GeV}$, $n_{\text{bjets}} \geq 1$. The green band in the ratio view represents the combined systematic lepton efficiency, trigger efficiency, luminosity and cross-section uncertainties. JES uncertainties are shown as red, hatched band.

**Top-quark pair production** is the main background. Two $W$ bosons that are produced in the top quark decays can produce two prompt leptons and two neutrinos, which cause a large amount of real $E_T$. At least two jets are produced and provide a large amount of hadronic activity. Additionally, in rare cases, a contribution is possible if just one $W$ boson decays leptonically, and one jet in the event is mistaken for a lepton. Top-quark pair production can be estimated using uncorrelated-flavour subtraction since there is no flavour correlation of the two leptons.

**$Z$-bosons decaying into light leptons** produce two flavour-correlated, prompt leptons. Accompanied by jets, e.g. by initial-state gluon radiation, also hadronic activity can be present. However, since there is no genuine $E_T$ in these processes, they can be strongly suppressed by a $E_T$ cut. The remaining yield of these processes is fitted in this analysis from the invariant-mass spectrum.

**$Z$-bosons decaying into tau-pairs** can produce pairs of light leptons if both taus decay leptonically. These decays also provide a medium amount of genuine $E_T$, and hadronic activity can be present due to initial-state radiation. The two light leptons are uncorrelated, hence this background can be targeted by uncorrelated-flavour subtraction.

**Diboson ($WW$, $WZ$, $ZZ$) production** can provide two or more prompt leptons and genuine $E_T$ by neutrinos. Flavour-correlated lepton production is possible in these processes by decaying $Z$ bosons. The dilepton mass should accordingly be close to the $Z$ mass, and hence, this contribution is estimated along with the single-$Z$ production. Other diboson processes produce uncorrelated leptons and can be estimated using uncorrelated-flavour subtraction.

**$W$-boson production** provides one prompt lepton and a neutrino causing $E_T$. Similar to $Z$-boson production it can also be accompanied by jets. This background is
4.3. Background Estimation

strongly suppressed by the lepton-pair requirement. There is however a chance of contribution if a jet is mistaken for a lepton.

4.3.1 Uncorrelated-Flavour Subtraction

Assuming a flavour-uncorrelated production of dilepton events, $e\mu$ events can be used to predict the yield that is caused by the same process in the same-flavour channels. If differences in trigger and object reconstruction efficiencies are not taken into account, the yields in the $ee$ and $\mu\mu$ channel can both be estimated by dividing the yield in the $e\mu$ channel by a factor of two (combinatorial factor for this channel).

Trigger and reconstruction efficiencies differ, however, significantly for the three channels and hence have to be measured and considered in this background prediction. As a first step, electron and muon reconstruction efficiencies are taken into account in this calculation, while trigger efficiencies are neglected. The impact of trigger inefficiencies on the derived results is investigated afterwards in a second step.

Using the ratio of muon efficiency to electron efficiency, which is measured using the number of dielectron and dimuon events in the Z peak (see Sec. 3.7.2),

\[ r_{\mu e} = \sqrt{\frac{n_{\mu\mu}}{n_{ee}}} \]  

(4.1)

the extrapolation of the event yields in the same-flavour channels from the $e\mu$ channel is

\[ n_{ee} = \frac{1}{2} \cdot \frac{n_{e\mu}}{r_{\mu e}} \]  

(4.2)

and

\[ n_{\mu\mu} = \frac{1}{2} \cdot r_{\mu e} n_{e\mu} \]  

(4.3)

Only the ratio of the efficiencies is needed here; the absolute reconstruction efficiencies do not have to be measured to apply this method.

The total same-flavour event yield can be predicted using

\[ n_{ll} = n_{ee} + n_{\mu\mu} = \frac{n_{e\mu}}{2} \left( r_{\mu e} + \frac{1}{r_{\mu e}} \right) \]  

(4.4)

Using Equations 4.2 and 4.3, the contribution of $ee$ and $\mu\mu$ events to the total same-flavour yield can be deduced:

\[ n_{ee} = \frac{1}{1 + r_{\mu e}^2} \cdot n_{ll} \]  

(4.5)

\[ n_{\mu\mu} = \frac{r_{\mu e}^2}{1 + r_{\mu e}^2} \cdot n_{ll} \]  

(4.6)

Consideration of Trigger Efficiencies

All three light-lepton channels are affected by different trigger efficiencies (see Sec. 3.7.3), which have to be taken into account in this background-estimation method.

Be $\epsilon_{ee}$, $\epsilon_{e\mu}$ and $\epsilon_{\mu\mu}$ the trigger efficiency for the $ee$, $e\mu$ and $\mu\mu$ channel, respectively. The relation of the measured event yield in each channel, $n_{ee}$, $n_{e\mu}$ and $n_{\mu\mu}$, and the yield before the application of the trigger, $n^*_{ee}$, $n^*_{e\mu}$ and $n^*_{\mu\mu}$, is then given by

\[ n_{ee} = \epsilon_{ee} \cdot n^*_{ee} \]  

(4.7)

\[ n_{e\mu} = \epsilon_{e\mu} \cdot n^*_{e\mu} \]  

(4.8)

\[ n_{\mu\mu} = \epsilon_{\mu\mu} \cdot n^*_{\mu\mu} \]  

(4.9)
Accordingly, the extrapolation from the different-flavour channel into the same-flavour channels changes to

\[ n_{ee} = \frac{1}{2} \cdot \epsilon_{ee} \cdot \frac{n_{e\mu}^*}{r_{\mu e}^*} = \frac{1}{2} \cdot \epsilon_{ee} \cdot \frac{n_{\mu \mu}^*}{r_{\mu e}^*}, \] (4.10)

\[ n_{\mu \mu} = \frac{1}{2} \cdot \epsilon_{\mu \mu} \cdot r_{\mu e}^* n_{e\mu}^* = \frac{1}{2} \cdot \epsilon_{\mu \mu} \cdot r_{\mu e}^* n_{e\mu}, \] (4.11)

with the efficiency ratio before trigger application, \( r_{\mu e}^* \).

The measured efficiency ratio, \( r_{\mu e} \), is also affected by the trigger efficiencies, and is related to \( r_{\mu e}^* \) by

\[ r_{\mu e}^2 = \frac{n_{\mu \mu}}{n_{ee}} = \frac{\epsilon_{\mu \mu}}{\epsilon_{ee}} \cdot \frac{n_{e\mu}}{n_{ee}} = \frac{\epsilon_{\mu \mu}}{\epsilon_{ee}} \cdot r_{\mu e}^* \cdot r_{\mu e}^*, \] (4.12)

Thus, the transformation between the measured efficiency ratio and the ratio before trigger application is given by

\[ r_{\mu e} = \sqrt{\frac{\epsilon_{\mu \mu}}{\epsilon_{ee}} \cdot r_{\mu e}^*}, \] (4.13)

\[ r_{\mu e}^* = \sqrt{\frac{\epsilon_{ee}}{\epsilon_{\mu \mu}} \cdot r_{\mu e}}. \] (4.14)

Using this transformation, Equations 4.10 and 4.11 change to

\[ n_{ee} = \frac{1}{2} \cdot \sqrt{\epsilon_{ee} \epsilon_{\mu \mu}} \cdot \frac{n_{e\mu}^*}{r_{\mu e}^*}, \] (4.15)

\[ n_{\mu \mu} = \frac{1}{2} \cdot \sqrt{\epsilon_{ee} \epsilon_{\mu \mu}} \cdot r_{\mu e} n_{e\mu}. \] (4.16)

The relation between \( n_{ll} \) and \( n_{ee} \) (\( n_{\mu \mu} \)) does not change from Equation 4.5 (Eqn. 4.6): using

\[ n_{ll} = n_{ee} + n_{\mu \mu} = \frac{n_{e\mu}}{2} \cdot \frac{\sqrt{\epsilon_{ee} \epsilon_{\mu \mu}}}{\epsilon_{ee}} \cdot (r_{\mu e} + \frac{1}{r_{\mu e}}) \] (4.17)

it can be seen that the efficiencies cancel against each other after inserting \( n_{e\mu} \) from Equation 4.15 (Eqn. 4.16).

4.3.2 Shape-Based Subtraction

In this search, the dilepton invariant-mass distribution is searched for a kinematic edge. To distinguish signal and background components in this distribution, the uncorrelated-flavour subtraction is extended to a shape subtraction.

If leptons are produced uncorrelatedly, they follow the same kinematic distributions independent of the lepton flavour. Therefore, not only the event yield, but also the shape of the invariant-mass distribution in same-flavour dilepton events can be predicted from different-flavour events. Figure 4.7 shows the invariant-mass distribution of same-flavour and different-flavour lepton pairs in a \( t\bar{t} + jets \) MC sample. The two shapes agree very well with each other.

Figure 4.8 displays the invariant mass of same-flavour and different-flavour lepton pairs in the control region. The control region is dominated by top processes, hence a good agreement of same-flavour and different-flavour shapes is expected. The two observed shapes are quite well described by the simulation, and they agree well with each other.
Figure 4.7: Invariant mass of $ee$ and $\mu\mu$, and invariant mass of $e\mu$ lepton pairs in a simulated top sample.

Figure 4.8: Data and MC estimation of same-flavour ($ee$ and $\mu\mu$) (a) and different-flavour ($e\mu$) (b) events in the control region. The green band in the ratio view represents the combined systematic lepton efficiency, trigger efficiency, luminosity and cross-section uncertainties. JES uncertainties are shown as red, hatched band.
This provides evidence that the shape-based background prediction can be applied on this and similar event selections.

The $b$-jet requirement provides a very strong suppression against $DY + jets$ events in the control region. In the signal region, these are suppressed by tighter $H_T$ and $E_T$ cuts. Still, the contribution of Drell-Yan (and also diboson) processes is higher than in the control region. Therefore, in the signal region, the dilepton-mass shapes are expected to show differences from each other. The background contribution by Drell-Yan processes cannot be predicted from the $e\mu$ dilepton-mass shape and has to be estimated separately.

The invariant-mass distribution for same- and different-flavour lepton pairs in the signal region is shown in Figure 4.9. The observed event yield is lower than expected from MC, the difference is however covered by the JES uncertainty. The dilepton-mass shapes on the other hand, are well described in the simulation.

No unexpected discrepancies are observed between the distributions of same-flavour and different-flavour events. Therefore, as a next step, the simultaneous fit, which incorporates the shape-based subtraction method, is presented.

**Figure 4.9:** Data and MC estimation of same-flavour ($ee$ and $\mu\mu$) (a) and different-flavour ($e\mu$) (b) events in the signal region. The green band in the ratio view represents the combined systematic lepton efficiency, trigger efficiency, luminosity and cross-section uncertainties. JES uncertainties are shown as red, hatched band.

### 4.3.3 Simultaneous Fit

The shape-based uncorrelated-flavour subtraction is performed as a simultaneous, unbinned maximum likelihood fit to the invariant mass distributions of $ee$, $\mu\mu$ and $e\mu$ pairs. This fit comprises the complete background estimation of this search and does not depend on MC.

The fit consists of three components: a triangular shaped signal component, a component to model $Z$ background and a component to model background by uncorrelated dilepton production.
4.3. Background Estimation

As model for a potential signal, an edge model for two subsequent two-body decays is used, given by

\[ S(m_{\ell\ell}) = \frac{1}{\sqrt{2\pi}\sigma} \int_{0}^{m_{\text{max}}} dy \cdot ye^{-\frac{(m_{\ell\ell} - y)^2}{2\sigma^2}}. \] (4.18)

It describes a triangular mass edge with endpoint, \( m_{\text{max}} \), assuming a mass resolution of \( \sigma \), which can be independent for \( ee \) and \( \mu\mu \) events. A variation of the signal shape and its impact on the results is studied later in Section 4.5.3.

For same-flavour background resulting from \( Z \) decays, a Breit-Wigner function convolved with a gaussian is used. The \( Z \) mass and width is fixed to the current world average \([59]\), and the width of the gaussian is set independently to the mass resolution of the corresponding channel. For the \( ee \) channel a mass resolution of \( 2 \pm 1.0 \text{ GeV} \) is assumed and for the \( \mu\mu \) channel a resolution of \( 1 \pm 0.5 \text{ GeV} \) \([110]\).

The uncorrelated-flavour background component is modeled by a power function and an exponential:

\[ B(m_{\ell\ell}) = m_{\ell\ell}^a \cdot e^{-b \cdot m_{\ell\ell}}. \] (4.19)

This function is motivated empirically and has been found to describe the background distribution well. It has a simple form and depends, apart from \( m_{\ell\ell} \), on just the two parameters \( a \) and \( b \).

The total likelihood of the invariant-mass distribution of electron pairs and muon pairs is

\[ L = \frac{(N_{S}^{\ell\ell} + N_{B}^{\ell\ell} + N_{Z}^{\ell\ell}) N \cdot e^{-(N_{S}^{\ell\ell} + N_{B}^{\ell\ell} + N_{Z}^{\ell\ell})}}{(N_{S}^{\ell\ell} + N_{B}^{\ell\ell} + N_{Z}^{\ell\ell})!} \times \prod_{i} \frac{(N_{S}^{\ell\ell} P_{S}(m_{\ell\ell,i}) + N_{B}^{\ell\ell} P_{B}(m_{\ell\ell,i}) + N_{Z}^{\ell\ell} P_{Z}(m_{\ell\ell,i}))}{N_{S}^{\ell\ell} + N_{B}^{\ell\ell} + N_{Z}^{\ell\ell}} \] (4.20)

with the probability density function \( P_{S} = S \) for the signal, \( P_{B} = B \) for the uncorrelated-flavour background and the \( Z \) lineshape, \( P_{Z} \).

For \( e\mu \) events, the likelihood is given by

\[ L = \frac{(N_{B}^{e\mu}) N \cdot e^{-(N_{B}^{e\mu})}}{(N_{B}^{e\mu})!} \times \prod_{i} \frac{N_{B}^{e\mu} P_{B}(m_{\ell\ell,i})}{N_{B}^{e\mu}}. \] (4.21)

The fit is performed, using the RooFit package \([111]\), simultaneously to the invariant-mass distribution of \( ee \), \( \mu\mu \) and \( e\mu \) events. For this, the background yields in the three channels, \( N_{ee}^{B} \), \( N_{\mu\mu}^{B} \) and \( N_{e\mu}^{B} \), are linked using Equations 4.15 and 4.16. The yields of the same-flavour components in the \( ee \) and \( \mu\mu \) channel, \( N_{S}^{ee} \), \( N_{S}^{\mu\mu} \), \( N_{Z}^{ee} \) and \( N_{Z}^{\mu\mu} \), are linked to each other and the total yields, \( N_{S} \) and \( N_{Z} \), following Equations 4.5 and 4.6.

The fit provides the total yields \( N_{S} \), \( N_{B} \) and \( N_{Z} \), which are implemented as floating parameters. To improve convergence of the fit algorithm, \( N_{S} \) is constrained such that the combined shape \( (N_{S}^{\ell\ell} P_{S} + N_{B}^{\ell\ell} P_{B} + N_{Z}^{\ell\ell} P_{Z}) \) is non-negative over the whole dilepton mass range.

The fit is performed including all events with a dilepton mass between 0 GeV and 450 GeV. This fit window is chosen such that all events in the signal region are included.

Before the application of the fit method is tested, a background that is potentially not covered with this subtraction is investigated.
4.3.4 Leptons From Misidentified Jets

Jets can be misidentified as leptons. The lepton-ID requirements (see Sec. 3.4.1 and Sec. 3.4.2) provide a good suppression against such fake leptons. Nevertheless, the impact of misidentifications has to be investigated, because misidentifications are not described correctly by the uncorrelated-flavour subtraction method if they occur strongly flavour-dependent.

Due to the low rate of jets mimicking leptons, only events with one mimicked lepton and another prompt lepton are considered. The contribution to the event yield in the signal region by this type of events is estimated using a tight-to-loose ratio method. The amount of events with two fake leptons in the final events selection is neglected.

The tight-to-loose method is basically an isolation sideband extrapolation. First, a loose set of lepton identification criteria is defined by loosening the isolation requirement. The so-called tight-to-loose ratio, $\epsilon_{TL}$, is determined on data as the ratio of the number of leptons that pass the tight identification criteria to the number of leptons that pass the loose criteria. The ratio needs to be determined on a data sample that is dominated by fake leptons, e.g. a QCD dominated region. Afterwards, the tight-to-loose ratio can be used to estimate the number of expected events in the signal region. This is done as extrapolation from a loosened event selection, in which one of the two leptons just needs to satisfy the loose identification criteria.

**Measurement of the Tight-to-Loose Ratio**

The loose identification criteria are defined by loosening the isolation cut of the default lepton selection from 0.15 to 1.00. All other identification requirements remain unchanged.

The tight-to-loose ratio is determined on a QCD dominated selection defined by $H_T > 150 \text{ GeV}$ and $\not{E}_T < 20 \text{ GeV}$. The anti-$\not{E}_T$ cut suppresses contamination by top processes. Exactly one lepton is required to reject leptons from $Z$ decays. With a transverse-mass requirement of $M_T < 25 \text{ GeV}$, leptons from $W$ decays are rejected. The selection is triggered by prescaled lepton-support triggers that require one lepton with the same selection criteria as the dilepton triggers and one additional jet to reduce the trigger rate (see Tab. 3.8).

The ratio, $\epsilon_{TL}$, is obtained as the number of leptons passing the tight selection criteria, $N_{\text{tight}}$, divided by the number of leptons passing the loose selection criteria, $N_{\text{loose}}$:

$$\epsilon_{TL} = \frac{N_{\text{tight}}}{N_{\text{loose}}}.$$  \hspace{1cm} (4.22)

$\epsilon_{TL}$ is expected to show variation with lepton $p_{\perp}$ and $\eta$, because lepton identification and isolation efficiencies are not constant for these variables. Therefore the ratio is measured for several bins of lepton $p_{\perp}$ and $\eta$. Figure 4.10 shows the electron and muon tight-to-loose ratio versus $p_{\perp}$ for the barrel and the endcap region.

For electrons, the tight-to-loose ratio yields about 8% in the barrel region and about 25% in the endcaps. A slight decrease with rising lepton $p_{\perp}$ is visible. Since identification and isolation requirements are different for electrons in the barrel and in the endcaps (see Sec. 3.4.2), the tight-to-loose ratio is also expected to be different for electrons in these two $\eta$ regions. For muons, $\epsilon_{TL}$ is about 5% in the barrel and about 7% in the endcaps. Also a slight decrease of the ratio with rising lepton $p_{\perp}$ can be observed. Contrary to the electrons, there is no strong dependence on $\eta$.

---

1The transverse mass of a lepton, $\ell$, and the missing transverse energy in an event, $\not{E}_T$, is defined as $M_T = \sqrt{(E_\ell + \not{E}_T)^2 - (p_{\perp,\ell} + \not{E}_T)^2}$ and represents the invariant mass of a particle that decayed into $\ell$ and a neutrino if $z$ momentum components are neglected.
Estimated Contribution to Signal Region

To estimate the amount of events with fake leptons in the signal region, a loose signal region is defined. Events with one lepton passing the tight identification criteria, and one lepton passing the loose identification criteria and failing the tight ones, are selected. The second lepton is called fake lepton candidate. Apart from the lepton requirements, the signal region selection remains unchanged.

Events in this loose signal region are now weighted according to $\epsilon_{TL}(p_{\perp}, \eta)$ of the fake lepton candidate:

$$W = \frac{\epsilon_{TL}(p_{\perp}, \eta)}{1 - \epsilon_{TL}(p_{\perp}, \eta)}.$$  \hfill (4.23)

The weight sum, $\sum W_i$, over all events $i$ in the loose signal region then yields the predicted contribution of events with fake leptons in the signal region. Figure 4.11 displays the estimated background prediction. The predicted yield due to fake leptons in the three dilepton channels is

$$n_{ee, \text{fake}} = 4.9 \pm 1.0 \,(\text{stat}) \pm 0.7 \,(\text{sys})$$  \hfill (4.24)

$$n_{\mu\mu, \text{fake}} = 3.1 \pm 0.4 \,(\text{stat}) \pm 0.5 \,(\text{sys})$$  \hfill (4.25)

$$n_{e\mu, \text{fake}} = 8.3 \pm 1.1 \,(\text{stat}) \pm 1.2 \,(\text{sys}),$$  \hfill (4.26)

which corresponds to about 5% of the total observed yield in these channels. Systematic uncertainties represent uncertainties due to the statistical uncertainties of the tight-to-loose ratio.

In this analysis, all flavour-symmetric backgrounds are subtracted by the shape-based uncorrelated-flavour method. Since the background caused by fake leptons does not show signs of significantly flavour-asymmetric behaviour, this background is almost completely dealt with by the mentioned subtraction method. The remaining non-symmetric contribution of this background is not larger than 1% of the total yield per channel and is neglected.
Figure 4.11: Data and MC estimation of same-flavour ($ee$ and $\mu\mu$) (a) and different-flavour ($e\mu$) (b) events in the signal region as in Figure 4.9. In addition, the expected contribution by fake leptons estimated using the tight-to-loose method is shown as a red line.

4.4 Performance on Monte Carlo Simulation

Before performing the fit procedure on data, it is tested on MC to ensure correct functionality. The SM background is described by adding up $t\bar{t} + jets$, $t/\bar{t} + jets$, $DY + jets$ and diboson samples.

As a first step, the fit procedure is performed on only SM background. No simulated signal is added. Figure 4.12 displays the invariant-mass distributions of same-flavour and different-flavour lepton pairs and the resulting fit to these distributions. The background-component of the fit is shown in Figure 4.12b, and it fits the $e\mu$ data well. Figure 4.12a displays the final combined fit, the individual fit components and the same-flavour data. An example value of 77.8 GeV is chosen for $m_{max}$. This is the theoretical mass-edge position in the LM1 benchmark scenario.

As expected, the signal contribution obtained from the fit is compatible with zero. The best-fit estimate is $N_S = 5.7 \pm 11.5$. This means that the dilepton-mass shape is mainly compatible with the background estimation. The flavour-symmetric background is estimated to be $N_B = 206.4 \pm 14.4$ and the $Z$ contribution $N_Z = 12.8 \pm 5.7$.

The best-fit values of the shape parameters, $a$ and $b$ (see Eqn. 4.19), are given in Table 4.3 together with the $\chi^2/n_{dof}$ values of the same-flavour and the different-flavour distribution. The $\chi^2/n_{dof}$ values are between one and ten and are therefore not optimal. Considering however the high statistics of the MC sample, the resulting small statistical uncertainties and the simple background shape, which uses just two parameters, these values seem reasonable.

Next, the fit procedure is repeated for different mass-edge positions. $m_{max}$ is varied between 20 GeV and 300 GeV in steps of 10 GeV. Mass-edge positions beyond 300 GeV are not tested, because very few events are observed in data in the signal region beyond this mass. In total, four events, two same-flavour events and two $e\mu$ events, are observed in this region.
4.4. Performance on Monte Carlo Simulation

Figure 4.12: Simultaneous fit to the $ee$ and $\mu\mu$ dilepton-mass distribution (a) and to the $e\mu$ dilepton-mass distribution (b) in the signal region, $H_T > 300$ GeV and $E_T > 150$ GeV, for simulated Standard-Model background. The mass-edge position of the signal hypothesis is $m_{max} = 77.8$ GeV. The combined fit shape is displayed as blue, solid line, the individual fit contributions as dashed lines. The fit uncertainty on the different-flavour background component is indicated by the green, hatched band.

Table 4.3: Best-fit values of background-shape parameters, $a$ and $b$, and $\chi^2/n_{dof}$ of the final fit with respect to the same-flavour (SF) and the different-flavour (DF) dilepton-mass distributions in the signal region on SM simulation.

<table>
<thead>
<tr>
<th>Region</th>
<th>$a$</th>
<th>$b$</th>
<th>SF $\chi^2/n_{dof}$</th>
<th>DF $\chi^2/n_{dof}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Signal region (MC)</td>
<td>1.186 ± 0.147</td>
<td>0.021 ± 0.002</td>
<td>8.65</td>
<td>1.99</td>
</tr>
</tbody>
</table>
Figure 4.13 displays the estimated signal yield for the various tested mass-edge hypotheses. No deviation from the background-only hypothesis beyond statistical uncertainties is seen. Around the $Z$-boson mass, positive signal yields of up to $9 \pm 15$ events are estimated. The lowest signal yields are observed at $m_{\text{max}} = 220$ GeV with $-6 \pm 10$ events.

**Figure 4.13:** Fitted signal yield, $N_S$, from SM simulation in the signal region versus $m_{\text{max}}$ of the signal hypothesis.

This result indicates, that in general, the chosen background shape fits the invariant-mass distributions well. The observed, small deviations around 100 GeV and 200 GeV can be the result of statistical fluctuations in the MC or indications of a non-perfect agreement of the background shape with the dilepton-mass distribution. In the latter case, this effect will have to be considered if this analysis is performed in regions with event yields higher than the MC event count in this region (about $5 \cdot 10^3 \ t\bar{t} + \text{jets}$ event).

### 4.4.1 Adding CMSSM Benchmarks

As a next step, the fit procedure is tested on a simulated signal that is added to the SM simulation. The CMSSM benchmark scenarios LM1, LM3 and LM6 are tested. The fit results for the scenarios LM1 and LM3 are shown in Figure 4.14, the results for LM6 in Figure 4.15. For each benchmark scenario, $m_{\text{max}}$ is set to the theoretical position of the mass-edge endpoint. LM1 is scaled to 20% of its original cross section to adapt the signal yield to the region of interest and allow better comparison with the other benchmark scenarios.

For LM1, LM3 and LM6, a signal yield of $70.3 \pm 15.0$, $102.5 \pm 17.2$ and $27.9 \pm 16.8$ is extracted, respectively. The fitted shapes match the simulated data reasonably well. LM6 is on the verge of sensitivity of this analysis. The expected same-flavour signal yield for these scenarios after subtraction of the $e\mu$ contribution (compare Tab. 4.1) is about 62, 93 and 23 events for LM1 (scaled by 0.2), LM3 and LM6, respectively. The yields that are extracted by the fit agree within their uncertainty with these expectations, but are all three larger than the expectation. It has to be noted though, that the three yields are not uncorrelated, because the position of the mass-edge is similar in the three scenarios. On the background-only MC, a small, but positive signal of $N_S = 5.7 \pm 11.5$ was calculated for a mass-edge position of 78 GeV. This offset covers the main part of the difference between fitted yield and expectation for LM1 and LM3.

Neutralino decays via $Z$ boson make an important contribution to lepton pair production in the LM3 scenario. Therefore, in addition to the mass-edge, an amplification of the $Z$ peak can be observed. This additional yield is ignored and not treated as signal in this analysis because the $Z$ contribution is not fixed or measured in any other way.
Figure 4.14: Fit of the $ee$ and $\mu\mu$ dilepton-mass distributions (left) and of the $e\mu$ dilepton-mass distributions (right) in the signal region, $H_T > 300$ GeV and $E_T > 150$ GeV, for simulated Standard-Model background together with simulated LM1 (scaled) (a)-(b) and LM3 (c)-(d) signal. $m_{max}$ is set to the theoretical position of the mass edge in the benchmark scenarios. The combined fit shape is displayed as blue, solid line, the individual fit contributions as dashed lines. The fit uncertainty on the different-flavour background component is indicated by the green, hatched band.
Figure 4.15: Fit of the $ee$ and $\mu\mu$ dilepton-mass distribution (a) and of the $e\mu$ dilepton-mass distribution (b) in the signal region, $H_T > 300 \text{ GeV}$ and $E_T > 150 \text{ GeV}$, for simulated Standard-Model background together with simulated LM6 signal. $m_{\text{max}}$ is set to the theoretical position of the mass edge in the benchmark scenario. The combined fit shape is displayed as blue, solid line, the individual fit contributions as dashed lines. The fit uncertainty on the different-flavour background component is indicated by the green, hatched band.

4.4.2 T3lh Simplified Model

Lastly, a signal according to the T3lh simplified model (see Sec. 1.6.1) is added to SM MC and used as input to the fit procedure. Two example configurations of the model are shown in Figure 4.16. The configuration $m_{\text{gluino}} = 600 \text{ GeV}$, $m_{\text{LSP}} = 450 \text{ GeV}$ results in a low mass splitting and a mass-edge position of 75 GeV. The fit results are shown in Figures 4.16a and 4.16b. The other example configuration, $m_{\text{gluino}} = 800 \text{ GeV}$ and $m_{\text{LSP}} = 400 \text{ GeV}$, yields a high mass splitting and a mass-edge position at 200 GeV. The fit results are displayed in Figures 4.16c and 4.16d. In both cases, $m_{\text{max}}$ is fixed to the theoretical mass-edge position.

In the two shown example configurations, signal yields of $N_S = 87.0 \pm 15.0$ and $N_S = 55.9 \pm 15.5$ are determined. The true MC yields are 87.8 and 81.6 events. While in the configuration with low mass splitting, the extracted yield almost exactly matched the true signal yield of the scenario, a lower signal yield is determined in the signal configuration with high mass splitting. Here, a discrepancy between the fitted signal shape and the invariant-mass distribution of the simulated signal can be seen (Fig. 4.16c).

The simplified model incorporates neutralinos that decay in a three-body decay, while the triangular shape of the signal hypothesis assumes two consecutive two-body decays. Therefore, the signal hypothesis does not match the simulated signal shape exactly. For low mass splitting, the differences are barely noticeable, but they increase with rising mass splitting and hence growing mass-edge position. Even if the fitted signal shape does not perfectly match the simulated signal distribution, the extracted signal yield still gives a good impression of the amount of signal that is overlaid on the SM simulation. A detailed study of the signal response of the fit procedure in the T3lh model is given in Section 5.5 in the context of result interpretation and limit setting.
4.4. Performance on Monte Carlo Simulation

Figure 4.16: Fit of the $ee$ and $\mu\mu$ dilepton-mass distributions (left) and of the $e\mu$ dilepton-mass distributions (right) in the signal region, $H_T > 300$ GeV and $E_T > 150$ GeV, for simulated Standard-Model background together with a T3lh model signal with low mass splitting ($m_{\text{gluino}} = 600$ GeV, $m_{\text{LSP}} = 450$ GeV) (a)(b) and with high mass splitting (c)(d) ($m_{\text{gluino}} = 800$ GeV, $m_{\text{LSP}} = 400$ GeV). $m_{\text{max}}$ is set to the theoretical mass-edge position of the signal. The combined fit shape is displayed as blue, solid line, the individual fit contributions as dashed lines. The fit uncertainty on the different-flavour background component is indicated by the green, hatched band.
4.5 Results

The simultaneous fit shows the expected results on plain SM simulation as input and also in case simulated signal samples are added. Therefore, it is proceeded to the analysis of data in the control region and in the signal region.

4.5.1 Control Region

As a next step, the fit is performed on data in the control region. \( m_{\text{max}} \) is scanned in steps of 10 GeV over the range of 20 GeV to 300 GeV, and the fit is performed for each of these signal hypotheses. The best-fit value for the number of signal events versus \( m_{\text{max}} \) is shown in Figure 4.17.

![Figure 4.17: Fitted signal yield, \( N_S \), from data in the control region versus \( m_{\text{max}} \) of the signal hypothesis.](image_url)

No significant deviations from the background-only hypothesis are observed. The largest deviations are a maximum of \( N_S \) around the \( Z \) boson mass and a minimum at around \( m_{\text{max}} = 180 \text{ GeV} \). The final fit parameters of the background component for these two signal configurations are shown in Table 4.4. As expected, the fit parameters are compatible with each other for the two configurations. \( \chi^2/\text{n}_{\text{dof}} \) values of around 0.8 are obtained. A good fit convergence and performance is concluded.

**Table 4.4:** Best-fit values of background-shape parameters, \( a \) and \( b \), and \( \chi^2/\text{n}_{\text{dof}} \) of the final fit with respect to the same-flavour (SF) and different-flavour (DF) dilepton-mass distributions in the control region.

<table>
<thead>
<tr>
<th>Region</th>
<th>( a )</th>
<th>( b )</th>
<th>SF ( \chi^2/\text{n}_{\text{dof}} )</th>
<th>DF ( \chi^2/\text{n}_{\text{dof}} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control region, ( m_{\text{max}} = 90 \text{ GeV} )</td>
<td>1.622 ± 0.081</td>
<td>0.029 ± 0.001</td>
<td>0.79</td>
<td>0.83</td>
</tr>
<tr>
<td>Control region, ( m_{\text{max}} = 180 \text{ GeV} )</td>
<td>1.646 ± 0.079</td>
<td>0.030 ± 0.001</td>
<td>0.77</td>
<td>0.83</td>
</tr>
</tbody>
</table>

Figure 4.18 displays the fit results for \( m_{\text{max}} = 90 \text{ GeV} \) and 180 GeV. The signal yield maximum at \( m_{\text{max}} = 90 \text{ GeV} \) seems to be caused by the proximity to the \( Z \) peak and the steep decline that follows at around 100 GeV. A higher signal yield is fitted here, because a mass edge at this position enhances the following decline and results in better agreement with the data. The fitted signal yield is however still compatible with zero within its statistical uncertainty.
A signal hypothesis with $m_{\text{max}} = 180 \text{ GeV}$ results in a best-fit estimate of $N_S = -44.1 \pm 22.2$. A negative signal on top of the background shape improves the agreements of the fitted shape with the data with respect to the background-only model. The position of this signal-yield minimum coincides approximately with the one that was determined on SM simulation (see Fig. 4.13), but does not match it perfectly. This indicates that the effect might be partly, but not completely, caused by the chosen background shape. The main reason is for the non-perfect agreement of the background shape with the dilepton-mass distribution is assumed to be a statistical fluctuation of the data in this region.

Figure 4.18: Fit of the same-flavour (left) and the opposite-flavour (right) dilepton-mass distributions of events in the control region. $m_{\text{max}}$ is 90 GeV in (a) and (b) and 180 GeV in (c) and (d). The combined fit shape is displayed as blue, solid line, the individual fit contributions as dashed lines. The fit uncertainty on the different-flavour background component is indicated by the green, hatched band.

### 4.5.2 Signal Region

Similarly to the control region, the fit is performed for various values of $m_{\text{max}}$ in the signal region. Figure 4.19 shows the best-fit values of $N_S$ for the investigated $m_{\text{max}}$ positions. In
general, most of the fitted yields are negative with a minimum at around $m_{\text{max}} = 210$ GeV. This is consistent with the observation of less same-flavour than different-flavour events in the signal region (see Tab. 4.1). A negative interference of a SUSY signal with SM background is however not possible. Therefore, these negative yields are considered an effect of statistical fluctuations in combination with the subtraction strategy. For further interpretation of the results and limit setting, a negative signal yield is regarded as a yield of zero.

![Figure 4.19: Fitted signal yield, $N_S$, from data in the signal region versus $m_{\text{max}}$ of the signal hypothesis.](image)

The most significant positive signal yield is at 280 GeV. The fit results for $m_{\text{max}} = 280$ GeV are displayed in Figure 4.20. A signal yield of $N_S = 8.9 \pm 5.9$ is extracted. The uncertainty is actually non-gaussian and also not symmetric. The symmetric uncertainty that is quoted at this point is a loose approximation based on the assumption of a parabolic minimisation problem. Therefore, it cannot directly be translated into a significance. The significance of this fluctuation is derived in Section 5.3. In any case, a fluctuation of this magnitude is not unreasonable considering that a collection of about 30 signal hypotheses is tested.

The final fit parameters for this signal configuration are shown in Table 4.5. A $\chi^2/n_{\text{dof}}$ of about 0.4 is achieved, which indicates correct convergence and performance of the fit. Since the event yield in the signal region quite low, and hence statistical uncertainties are relatively large, such a low $\chi^2/n_{\text{dof}}$ is not unplausible.

### Table 4.5: Best-fit values of background-shape parameters, $a$ and $b$, and $\chi^2/n_{\text{dof}}$ of final fit with respect to same-flavour and different-flavour dilepton-mass distribution in the signal region.

<table>
<thead>
<tr>
<th>Region</th>
<th>$a$</th>
<th>$b$</th>
<th>SF $\chi^2/n_{\text{dof}}$</th>
<th>DF $\chi^2/n_{\text{dof}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Signal region</td>
<td>$1.424 \pm 0.196$</td>
<td>$0.025 \pm 0.002$</td>
<td>0.45</td>
<td>0.42</td>
</tr>
</tbody>
</table>
4.5. Results

Figure 4.20: Fit of the same-flavour (a) and opposite-flavour (b) dilepton-mass distributions of events in the signal region. \( m_{\text{max}} \) is set to 280 GeV, the position in which the most significant positive signal is fitted. The combined fit shape is displayed as blue, solid line, the individual fit contributions as dashed lines. The fit uncertainty on the different-flavour background component is indicated by the green, hatched band.

4.5.3 Variation of the Signal Shape

In order to get an estimate of the dependency of the result on the signal shape, two other signal shapes are investigated. The first shape models the signal using a quartic instead of a linear function:

\[
S^{+4}(m_{\ell\ell}) = \frac{1}{\sqrt{2\pi}\sigma} \int_{0}^{m_{\text{max}}} dy \cdot y^4 e^{-\frac{(m_{\ell\ell}-y)^2}{2\sigma^2}}. \tag{4.27}
\]

Following this description, the signal edge becomes concave. The second shape describes a convex signal shape and also uses a quartic function, but with a negative coefficient:

\[
S^{-4}(m_{\ell\ell}) = \frac{1}{\sqrt{2\pi}\sigma} \int_{0}^{m_{\text{max}}} dy \cdot (m_{\text{max}}^4 - (y-m_{\text{max}})^4) e^{-\frac{(m_{\ell\ell}-y)^2}{2\sigma^2}}. \tag{4.28}
\]

The concave shape, \( S^{+4} \), is constructed to show a more distinct difference to the background shape than the triangular signal shape, \( S \). Therefore a better distinction of signal and background is expected with this shape, and the fitted signal yield in absence of a signal should be closer to zero. The convex shape on the other hand, is more similar to the background shape, hence it should tend to yield a larger signal than the other shapes.

The effect of the two alternative signal shapes is shown in Figure 4.21, in which the number of extracted signal events, \( N_S \), is displayed against the mass-edge position, \( m_{\text{max}} \), of the signal hypothesis. The absolute amount of extracted signal events is, as expected, smaller for the concave signal shape, \( S^{+4} \), and larger for the convex shape, \( S^{-4} \). The dependence of \( N_S \) on the mass-edge position, \( m_{\text{max}} \), shows in both cases the same behaviour as for the triangular signal shape (see Fig. 4.19).

The two signal shapes, \( S^{+4} \) and \( S^{-4} \), are demonstrated in Figure 4.22. The endpoint of the mass edge is set to the position in which the highest signal yield is fitted: \( m_{\text{max}} = 280 \) GeV.
Figure 4.21: Fitted signal yield, $N_S$, from data in the signal region versus $m_{\text{max}}$ for alternative signal shapes: the concave signal shape, $S^{+4}$, (a) and the convex signal shape, $S^{-4}$, (b).

The fitted signal yield with the concave $S^{+4}$ shape is $8.0 \pm 4.0$, which is lower than the signal yield with a triangular shape of $8.9 \pm 5.9$. The convex $S^{-4}$ shape results in a signal yield of $8.4 \pm 7.5$, which is comparable to the yield with a triangular shape. The uncertainty on the fitted yield increases however with the convex shape, because signal and background components are harder to distinguish with this signal hypothesis.

These alternative signal shapes are later used to estimate the dependency of the set limit on the signal shape.
Figure 4.22: Fit of same-flavour (left) and opposite-flavour (right) contribution for events in the signal region with convex signal shape, \( S^{-4} \), (a) (b), and with concave signal shape, \( S^{+4} \), (c) (d). \( m_{\text{max}} \) is 280 GeV. The combined fit shape is displayed as blue, solid line, the individual fit contributions as dashed lines. The fit uncertainty on the different-flavour background component is indicated by the green, hatched band.
4. Search for a Kinematic Edge
5. Interpretation of Results

No sign of physics beyond the Standard Model has been observed. Therefore, limits on non-SM physics are set. In this chapter, first the statistical procedures that are used to calculate such limits are introduced. Afterwards, limits on new physics are derived, and systematic uncertainties are pointed out. Finally, the limits are discussed and interpreted in detail within the T3lh simplified model and the CMSSM. An interpretation of this search within the T3lh simplified model and comparison to other CMS searches has been published in [34].

5.1 Statistical Procedure

Exclusion limits are calculated following the Higgs analysis limit setting procedure performed by ATLAS and CMS in 2011 [112]. The method is based on the modified frequentists method, $CL_s$ [113, 114], and is briefly described in the following.

Be $s$ the signal and $b$ the background yields. Now, the signal strength modifier, $\mu$, is defined. Using this signal strength modifier, the signal cross section can be scaled with respect to the background in order to find the exact exclusion limit on the number of signal events.

Systematic uncertainties (e.g. the lepton efficiency ratio, $r_{\mu e}$, see Sec. 5.4) are handled by the introduction of nuisance parameters, $\theta$. The yields become functions of these parameters: $s(\theta)$, $b(\theta)$. Now, be $p(\tilde{\theta}|\theta)$ the pdf (probability density function) of the nuisance parameters. $p(\tilde{\theta}|\theta)$ describes the probability to measure the values $\theta$ for the nuisance parameters, given that their true values are $\tilde{\theta}$.

5.1.1 Observed Limits

Observed limits are derived using the following steps [112]:

1. First, the likelihood function is constructed [115]:

   \[ L(\text{data}|\mu, \theta) = \text{Poisson}(\text{data}|\mu \cdot s(\theta) + b(\theta)) \cdot p(\tilde{\theta}|\theta), \]

   (5.1)

   with data representing the observation on data (or simulated pseudo-data). Poisson is the probability to observe data under the assumption of a given model. It can stand for a product of poisson probabilities or, as in this case, for an unbinned likelihood (see Eqn. 4.20).
2. The test statistic, \( q_\mu \), is constructed as profile likelihood ratio:

\[
q_\mu = -2 \ln \frac{L(\text{data}|\mu, \hat{\theta}_\mu)}{L(\text{data}|\hat{\mu}, \hat{\theta})} \quad \text{with } 0 \leq \hat{\mu} \leq \mu.
\] (5.2)

It is used to evaluate the compatibility of the data and a signal with a given strength, \( \mu \). Here, \( \hat{\theta}_\mu \) maximises the likelihood for the given signal strength. The variables \( \hat{\mu} \) and \( \hat{\theta} \) represent the values for which the likelihood reaches its global maximum.

The parameter \( \mu \) is constrained to be positive because a negative signal cross section is not allowed. In order to force a one-sided limit, the constraint, \( \hat{\mu} \leq \mu \), is applied.

3. Two tail probabilities are calculated by generation of toy-MC:

\[
CL_{s+b}(\mu) = P(q_\mu \geq q^{\text{obs}}_\mu | \text{signal + background}),
\] (5.3)

\[
CL_b(\mu) = P(q_\mu \geq q^{\text{obs}}_\mu | \text{background-only}).
\] (5.4)

These are the probabilities to obtain a test statistic value, \( q_\mu \), larger than the observed test statistic value, \( q^{\text{obs}}_\mu \), under the assumption of the signal+background hypothesis and the background-only hypothesis.

From these values, \( CL_s(\mu) \) is calculated:

\[
CL_s(\mu) = \frac{CL_{s+b}}{CL_b}.
\] (5.5)

4. If \( CL_s(\mu = 1) \leq \alpha \), then the investigated signal hypothesis is excluded with \( (1 - \alpha) \) CL confidence level (C.L.).

5. To calculate 95% confidence upper limits, \( \mu \) is adjusted until \( CL_s(\mu) = 0.05 \) is reached.

### 5.1.2 Expected Limit

For the calculation of expected limits, the profile likelihood asymptotic approximation [112] is used, because it is less computing intensive than the standard limit calculation. In this method, the asymptotic behaviour of the test statistic distribution is approximated such, that the signal strength corresponding to a \( CL_s = 0.05 \) probability can be extrapolated without generation of toy MC in the limit-calculation step.

To calculate the median expected limit and the \( 1\sigma \) and \( 2\sigma \) bands around it, a set of 100 MC pseudo-data samples is used. These are generated following the pdf of the fit to the SM simulation, and the exclusion limit based on each of these samples is derived. From this distribution of limits, the median and the gaussian \( 1\sigma \) and \( 2\sigma \) quantiles are calculated.

### 5.2 Constraining the Signal Yield

For the estimation of the best-fit value, the signal yield, \( N_S \), has been left floating freely. The SUSY signal, which is the target of this analysis, can however not interfere with the SM background. Therefore a negative signal yield is not physical.

Furthermore, a negative signal yield artificially increases the background yield, because the fit procedure applies an implicit constraint for the fit integral to match the data yield. In order to avoid a resulting bias of the background-only hypothesis, it is chosen to constrain the signal yield for the following limit calculations to be non-negative. This is a conservative approach, because the allowance of a negative signal yield would improve the resulting upper limit on non-SM processes.
5.3 Limit Calculation

From the previously obtained fit results, upper limits on the number of signal events in the signal region are estimated. These limits vary with the signal model and are therefore calculated in dependence on the mass-edge position, $m_{\text{max}}$.

The observed limit is calculated for signal models with $m_{\text{max}}$ in the range of 20 GeV to 300 GeV using the $CL_s$ method. The median expected limit is derived using the profile likelihood asymptotic approximation. For each evaluation, 100 toy MC samples are generated. Observed and expected limit including the 1σ and 2σ bands are shown in Figure 5.1. For reference, the LM1 benchmark scenario scaled to 20% of its original cross section, and the LM6 benchmark scenario are also displayed in the figure. The expected limit is derived for two other signal shapes – a concave and a convex shape (see Sec. 4.5.3) – in addition to the default triangular shape. The variation of the median expected limit with these two other shapes is indicated by the red, hatched band.

![Figure 5.1: Observed and expected 95% confidence level upper limit on the signal yield in the signal region, $H_T > 300\text{ GeV}$ and $E_T > 150\text{ GeV}$, versus mass-edge position, $m_{\text{max}}$. The 1σ and 2σ bands of the expected limit are shown in green and yellow. Variation of the expected limit with other signal shapes is indicated by the red, hatched band. For comparison, the LM1 benchmark scenario, scaled to 20% of its nominal cross section, and the LM6 benchmark scenario are displayed as red cross and red square. The LM3 scenario is outside the visible y-axis range.](image)

The observed limit curve follows, as expected, the trend of the signal yield that was extracted for the various values of $m_{\text{max}}$ (see Fig. 4.19). The upper limit on the number of signal events varies between 3 and 30 over the scanned $m_{\text{max}}$ range and shows reasonable fluctuations around the expected limit. The best limits are obtained for mass-edge positions below 50 GeV and between 160 GeV and 250 GeV.
The two alternative signal shapes result in a decrease of the derived upper limit of up to 10 events and an increase of up to 5 events. These numbers translate into relative variations of about 50% downwards and 25% upwards, which is comparable to or less than the statistical 1 $\sigma$ variation.

5.4 Systematic Uncertainties

A measurement of the trigger efficiencies has been presented in Section 3.7.3. From the statistical uncertainties on these measurements, an overall systematical uncertainty of 4% is deduced on the correct scaling of signal samples.

Lepton efficiencies have been found to be measured quite accurately in Monte Carlo and match actual efficiencies within 2% for leptons with $p_\perp > 15$ GeV [92]. Efficiencies for electrons with $10$ GeV < $p_\perp < 15$ GeV are consistent within 7%, and efficiencies for muons in this momentum range are consistent within 5% with measurements on data. However, these low-$p_\perp$ leptons affect only a very small fraction of signal events (less than 4% in case of the T3lh simplified model). Therefore, a general 2% systematic uncertainty is assigned on signal yields from simulation.

The uncertainty of the luminosity measurement is 2.2% [56].

The uncertainties on trigger efficiencies, lepton efficiencies and luminosity measurement can be combined to a global MC scaling uncertainty of 5.0%.

The lepton efficiency ratio, $r_{\mu e}$, is used to constrain the signal, background and Drell-Yan yields in $ee$, $e\mu$ and $\mu\mu$ events with respect to each other during the fit procedure. It has been measured to be 1.13 ± 0.11 (see Sec. 3.7.2). Further input to the fit procedure are the $ee$ and $\mu\mu$ mass resolutions, $\sigma_{ee}$ and $\sigma_{\mu\mu}$. The uncertainties on these three values are treated as nuisance parameter in the limit calculation.

Measurements of the hadronic energy scale yield an uncertainty of 7.5% (see Sec. 3.4.4). This affects the measurement of $H_T$ and $E_T$, which are important quantities for background rejection and have been used to define the signal region. Therefore, a mis-calibration of the hadronic energy scale directly affects the signal acceptance. The impact of the uncertainty on the signal yield is estimated by a correlated scaling of $H_T$ and $E_T$ selections by 7.5% in both directions. The resulting signal yield uncertainty strongly depends on the final states of the signal and has to be determined separately for every signal configuration.

5.5 Interpretation within the T3lh Simplified Model

After an upper limit on the number of signal events for a given mass-edge position is set, the search results can be interpreted in the context of specific new-physics models. In the following, the interpretation of the limit in the context of the T3lh simplified model is detailed.

For the analysis of results within this model, a simulated signal sample (including full CMS detector simulation) is used that scans the two-dimensional parameter space of the model. The gluino mass is varied between 100 GeV and 1200 GeV in steps of 25 GeV. LSP masses in the range of 50 GeV and $(m_{\text{gluino}} - 50$ GeV) have been modeled, also in steps of 25 GeV.

The gluino-pair production cross section for the T3lh signal is calculated with Prospino [28] and depends only on the gluino mass, $m_{\text{gluino}}$. Figure 5.2 displays the cross section for each point in the T3lh model parameter scan.
5.5. Interpretation within the T3lh Simplified Model

Figure 5.2: Production cross section for the T3lh simplified model signal versus gluino and LSP mass.

5.5.1 Efficiency and Acceptance

Efficiency and acceptance of a hypothetical signal define the reach of a search. The product of efficiency and acceptance is shown in Figure 5.3 for each point of the T3lh model parameter scan. It is calculated as the fraction of signal events that end up in the signal region, and is affected by the definition of the signal region including the boundaries of the fit window.

Figure 5.3: Product of efficiency and acceptance versus gluino and LSP mass for the signal region defined by $H_T > 300\,\text{GeV}$ and $E_T > 150\,\text{GeV}$, and the T3lh simplified model signal.
The boundaries of the invariant mass fit window have been chosen to be 0 and 450 GeV. Therefore, the signal acceptance drops if leptons are produced with a higher invariant mass than 450 GeV. In this model, lepton pairs are produced with a dilepton mass up to the mass difference of neutralino 2 and the LSP in a three-body decay. The mass of the neutralino 2 is placed halfway between the gluino and the LSP mass. Thus, the acceptance should in principle begin to decrease for mass splittings of gluino and LSP larger than 900 GeV. This drop however, is broadly smeared, because the mass-edge shape is not exactly triangular, especially for large mass splittings, and is hence not visible in Figure 5.3.

Only opposite-charge same-flavour lepton pairs are selected as signal in this search. In the considered model, no other lepton pair configuration is produced. Due to charge misidentification and fake leptons, there still is the possibility that events fail this selection. These effects cause an overall decrease of the signal acceptance of about 2 %.

Apart from the dilepton selection, the signal region is defined by an $H_T$ requirement of 300 GeV and a $E_T$ requirement of 150 GeV. In the T3lh model, jets are produced in the decay of a gluino into the intermediate neutralino 2 and in the decay of a gluino into the LSP. Therefore, the produced hadronic energy is proportional to the mass splitting of gluino and LSP. Accordingly, the acceptance of the selection rises steadily from the diagonal (with no mass splitting) to higher mass splittings. The $E_T$ requirement, in addition, causes a minor drop of the acceptance for the very low LSP masses.

### 5.5.2 Fit Efficiency

This search uses a fit to distinguish signal and background events in the signal region. Thus, also the efficiency of the fit on the simulated signal has to be investigated.

The fit efficiency is defined as the best-fit estimate of the signal yield divided by the actual yield in the signal region. Figure 5.4 shows the best-fit estimate of the signal yield and the fit efficiency for the T3lh parameter scan. For modest gluino and LSP masses, especially the whole exclusion region (see. Sec. 5.5.3), fit efficiencies between 70 % and 110 % are observed. The fit efficiency tends to be lower than one, because the fitted signal shape does not perfectly match the invariant-mass distribution of the simulated three-body decay in this model.

*Figure 5.4: Fitted signal yield (a) and resulting fit efficiency (b) for the T3lh simplified model signal versus gluino and LSP mass. Only parameter points for which the relative uncertainty on the number of signal events is less than 150 % are shown in (b).*

For high gluino masses, deviations from these values are visible. These are especially present in regions of small signal cross section or acceptance, where the actual signal yield is
of the order of few events. Therefore, small fluctuations within the statistical uncertainties are amplified to seemingly large distortions in the fit efficiency. Coherent regions of small and large fit efficiencies are caused by fluctuations in the Standard Model simulation. In the absence of a notable signal, such fluctuations in the dilepton-mass distribution at the mass-edge position migrate into the fitted signal yield. For mass splittings of about 180 GeV, the fit tends to overestimate the signal yield, because the mass edge collides with the Z peak in this region.

Additionally, a drop in the fit efficiency for very large mass splittings is visible. A maximum of $m_{\text{max}} = 300$ GeV is used for the signal hypothesis. Therefore, the compatibility of signal hypothesis and simulated signal distribution worsens in this region and results in a lower fit efficiency.

### 5.5.3 Exclusion Limit

A point in the scanned parameter space is excluded if the fitted signal yield is higher than the signal-yield upper limit that was calculated for the corresponding mass-edge position (shown in Fig. 5.1). Each point in the $m_{\text{gluino}}$-$m_{\text{LSP}}$ plane is tested for exclusion. Figure 5.5a displays the excluded region.

To demonstrate the dependency of the derived limit on the signal cross section, the exclusion for three times and a third of the nominal cross section is calculated (Fig. 5.5b and Fig. 5.5c). This also gives the possibility to approximately estimate the effect of the 5% total scale uncertainty on the limit: it should be negligible. Finally, the impact of the JES on the exclusion is estimated by scaling $H_T$ and $E_T$ up and down by 7.5% JES uncertainty (Fig. 5.5d and Fig. 5.5e).

The JES has a comparably small effect on the exclusion region. The largest impact is visible at gluino-LSP mass splittings below 300 GeV. This region is characteristic for dilepton masses below 150 GeV and for low hadronic activity, because gluino and LSP masses are close together. Therefore, the scale of the $H_T$ cut has a significant impact on the selected signal in this region. A shift of the excluded region of about 100 GeV in $m_{\text{gluino}}$ and $m_{\text{LSP}}$ is observed within the JES uncertainty range.

All derived exclusion contours are summarised in Figure 5.6 overlaid on the 95% confidence level upper limit on the signal cross section. The T3lh simplified model is excluded for gluino masses up to about 900 GeV if the mass of the LSP is lower than about 600 GeV. Assuming a signal cross section of three times the nominal cross section, the excluded range extends to $m_{\text{gluino}} \approx 1000$ GeV and $m_{\text{LSP}} \approx 700$ GeV. For a third of the nominal cross section, it drops to $m_{\text{gluino}} \approx 800$ GeV and $m_{\text{LSP}} \approx 500$ GeV.

### 5.6 Interpretation within the CMSSM

Finally, the set limits are interpreted in the context of the CMSSM, which is a popular choice for interpretation of SUSY search results. The interpretation is done using a CMSSM parameter scan, which contains simulated signal (including full CMS detector simulation) for various choices of $m_0$ and $m_{1/2}$. The remaining parameters of the model are fixed: $\tan \beta = 10$, $A_0 = 0$ and $\text{sgn} \mu = +1$. The simulated values of $m_0$ range from 40 GeV to 3000 GeV in steps of 20 GeV. $m_{1/2}$ is varied between 100 GeV and 1000 GeV, also in steps of 20 GeV. Regions in the parameter space that are theoretically excluded, because the LSP is charged or no electroweak symmetry breaking is realised, are not simulated. Furthermore, points with $m_{1/2} > 500$ GeV are considered to be beyond the reach of this search and are not tested for exclusion.

The position of the targeted dilepton-mass edge cannot be calculated from chosen parameter values as easily for the CMSSM as it was done for the simplified model. Therefore, the
Figure 5.5: Excluded region in the T3lh simplified model for nominal (a), three times (b) and a third of the nominal signal cross section (c). The excluded region with variation of the JES by 7.5\% is shown in (d) and (e).
5.6. Interpretation within the CMSSM

Figure 5.6: Cross section upper limit for the T3lh simplified model signal versus gluino and LSP mass. The dotted and dashed lines show the excluded region if the signal cross section is scaled to a third and three times the nominal signal cross section. The contribution of the JES uncertainty to the exclusion region is indicated as red, hatched band.

The position of the mass-edge, \( m_{\text{max}} \), is left floating for this interpretation and is determined by the minimisation procedure. A starting value of 78 GeV is chosen for \( m_{\text{max}} \). The result of the minimisation procedure is however largely independent from the starting position as long as values are chosen that are not close to the parameter boundaries, 0 GeV and 450 GeV.

The NLO cross section of the CMSSM signal depending on \( m_0 \) and \( m_{1/2} \) is displayed in Figure 5.7. It is of the order of 100 pb at about \( m_{1/2} = 100 \) GeV and drops exponentially with rising \( m_{1/2} \).

Figure 5.7: Cross section of a CMSSM signal versus \( m_0 \) and \( m_{1/2} \).
5.6.1 Efficiency and Acceptance

Similar to the interpretation within the simplified model, first the efficiency times acceptance of the search is calculated for the CMSSM. Figure 5.8 displays this product as function of $m_0$ and $m_{1/2}$. The value is significantly lower for the CMSSM than for the T3lh simplified model because other than leptonic decays are favoured in many regions of the CMSSM parameter space. In addition, the uncorrelated production of leptons, and hence different-flavour lepton pairs, is also possible at many parameter points.

![Figure 5.8: Product of efficiency and acceptance versus $m_0$ and $m_{1/2}$ for the signal region defined by $H_T > 300$ GeV and $\not{E}_T > 150$ GeV, and a CMSSM signal.](image)

For $m_0$ around 400 GeV and $m_{1/2}$ around 300 GeV, a maximum is visible. In this region, the slepton mass is lower than the mass of the neutralino 2, and a neutralino decay via slepton is possible. For lower $m_0$, the acceptance drops significantly at the border to a region in which the slepton becomes heavier than the neutralino 2 (diagonal through $m_0 \approx 200$ GeV and $m_{1/2} \approx 300$ GeV). For low values of $m_{1/2}$ (region already excluded), the $H_T$ and $\not{E}_T$ spectra become very soft, and as a result the acceptance also drops in this region.

5.6.2 Fit Efficiency

The number of fitted signal events and the resulting fit efficiency, defined as fitted number of signal events divided by the simulated number of SUSY events in the event selection, is displayed in Figure 5.9. The fit efficiency is only shown for parameter points for which the relative uncertainty on the signal yield is smaller than 100%. These are the points at which a significant mass edge is found.

A fit efficiency of about 50% is achieved in the central region of the investigated parameter space. The extracted signal yield ranges from thirty up to a few hundred events. Above $m_{1/2} \approx 260$ GeV, the number of extracted signal events, and as a result also the fit efficiency, drops significantly. Only for very low $m_0$ (smaller than 200 GeV), higher event yields are obtained at $m_{1/2}$ values above 260 GeV. The fit efficiency drop is displayed only for $m_0$ between 200 GeV and 700 GeV, since for larger $m_0$, the found mass edges become insignificant.

In contrast to the T3lh simplified model, the flavour-correlated production of lepton pairs in neutralino decays is not the only source of dilepton events in the CMSSM. Therefore, in general, a lower efficiency of this fit method is expected in this model.
Figure 5.9: Fitted number of signal events (a) and fit efficiency (b) versus $m_0$ and $m_{1/2}$ for a CMSSM signal. Only parameter points for which the relative uncertainty on the number of signal events is less than 100% are shown in (b).
The steep efficiency drop at \( m_{1/2} \approx 260 \text{ GeV} \) marks the sensitivity limit of this analysis. The reason for this drop is the rising contribution of a neutralino decay via real \( Z \) bosons, which is the dominant decay mode for high \( m_{1/2} \) values. In this case, an enhancement of the \( Z \) peak can be observed in the dilepton-mass spectrum. This enhancement, however, is not interpreted as signal in this analysis, because the \( Z \) background yield has no constraints in the fit.

Figure 5.10 displays one example signal configuration with \( m_{1/2} = 260 \text{ GeV} \), which is positioned just below the observed fit-efficiency drop. In this signal scenario, neutralinos can decay via a slepton and also via a \( Z \) boson. Thus, a mass edge as well as an enhanced \( Z \) peak are visible. The figure also shows a signal configuration with \( m_{1/2} = 280 \text{ GeV} \), which is positioned just above the observed drop in fit efficiency. Here, the neutralino decay via on-shell \( Z \) boson is the dominant decay mode. Therefore, a mass-edge is only barely visible. Most of the signal events are located at the \( Z \)-boson mass in this scenario, and they are hence not recognised as signal in this analysis.

Figure 5.11 shows the best-fit estimate of the mass-edge position, \( m_{\text{max}} \), versus the CMSSM parameters \( m_0 \) and \( m_{1/2} \). Only parameter points for which the relative uncertainty on the signal yield is smaller than 100\% are shown. A mass-edge with an endpoint around 40 GeV is visible for \( m_{1/2} \) of about 150 GeV. With rising \( m_{1/2} \), also the position of the mass edge increases until it reaches the \( Z \)-boson mass. At this point, the decay via \( Z \) boson becomes the dominant neutralino decay mode.

### 5.6.3 Exclusion

Similar to the interpretation within the simplified model, which was described in Section 5.5.3, the parameter points in the CMSSM \( m_0 - m_{1/2} \) plane are tested for exclusion. A point is tested by applying the fit procedure on the SM simulation together with the signal sample corresponding to the parameter point under investigation. If a signal yield is extracted that exceeds the derived upper limit (see Sec. 5.3), the parameter point is excluded.

Parton distribution function uncertainties are evaluated for the CTEQ 6.6 PDF set [116]. Each eigenvector of the PDF set is scaled up and down within its 90\% uncertainty contour, and the resulting variations are summed quadratically. The uncertainty due to the choice of the renormalisation and factorisation scale is evaluated by a variation of the scale to half and twice the nominal scale. PDF and scale uncertainties are calculated for each signal event separately. Then, for every CMSSM parameter point, the mean values of the PDF uncertainty and the scale uncertainty are determined from all signal events in the signal region. The total theory uncertainty is calculated afterwards as quadratic sum of these uncertainty components.

The impact of the JES on the exclusion limit is investigated, and the most conservative limit compatible with the JES uncertainty is chosen. Figure 5.12 shows the expected and observed exclusion in the CMSSM \( m_0-m_{1/2} \) plane. The figure includes the variation of the observed limit due to theoretical uncertainties, which include the uncertainties due to renormalisation and factorisation scale, and parton distribution functions.

Parameter points with \( m_{1/2} < 260 \text{ GeV} \) are excluded for all tested values of \( m_0 \) (up to 3000 GeV). For lower \( m_0 \) than 500 GeV, the excluded range extends to even higher values of \( m_{1/2} \) and reaches a maximum of 420 GeV for \( m_0 = 100 \text{ GeV} \). A translation of the exclusion into sparticle masses yields that parameter configurations with gluino masses below 600 GeV and squark masses up to 3000 GeV are excluded within the CMSSM for \( \tan \beta = 10, A_0 = 0 \) and \( \mu > 0 \).
5.6. Interpretation within the CMSSM

Figure 5.10: Fit of the $ee$ and $\mu\mu$ dilepton-mass distributions (left) and of the $e\mu$ dilepton-mass distributions (right) in the signal region for SM simulation together with a simulated CMSSM signal. A CMSSM parameter configuration for which both the neutralino decay via slepton and the decay via on-shell $Z$ boson is possible ($m_0 = 900\text{ GeV}$ and $m_{1/2} = 260\text{ GeV}$) (a)(b) and a parameter configuration for which the decay via on-shell $Z$ bosons is the dominant decay ($m_0 = 900\text{ GeV}$ and $m_{1/2} = 280\text{ GeV}$) (c)(d) are shown. The combined fit shape is displayed as blue, solid line, the individual fit contributions as dashed lines. The fit uncertainty on the different-flavour background component is indicated by the green, hatched band.
Figure 5.11: Best-fit estimate of $m_{\text{max}}$ versus $m_0$ and $m_{1/2}$ for a CMSSM signal. Only parameter points for which the relative uncertainty on the number of signal events is less than 100% are shown.

Figure 5.12: Observed (red) and expected (blue) exclusion limit in the CMSSM parameter space for $\tan \beta = 10$, $A_0 = 0$ and $\mu > 0$. The impact of theoretical uncertainties on the observed limit is shown by the dashed lines. The 1 $\sigma$ expected band is indicated by the dotted lines. Regions of the parameter space that have already been excluded by LEP measurements are indicated in green and yellow. Theoretically inaccessible regions because of a charged LSP or lack of electroweak symmetry breaking are marked in grey [117].
A counting experiment was developed by CMS for the analysis of opposite-sign dilepton events as a general, complementary analysis to this search [92]. It is based on an event selection using $E_T$ and $H_T$, and a data-driven estimation of the top-quark background using similarities in the $E_T$ distribution and the transverse-momentum distribution of lepton pairs. A veto against lepton pairs compatible with the $Z$ mass is applied. The CMSSM exclusion limits derived with the counting experiment are shown in Figure 5.13.

**Figure 5.13:** Observed (red) and expected (blue) exclusion limit of the complementary counting experiment in the CMSSM parameter space for $\tan \beta = 10$, $A_0 = 0$ and $\mu > 0$ [92]. The impact of theoretical uncertainties on the limits is shown by the dashed lines. The $1\sigma$ expected band is indicated by the dotted lines. The exclusion based on 34 pb$^{-1}$ of 2010 data with a similar analysis is shown as solid, purple area. Regions of the parameter space that have already been excluded by LEP measurements are indicated in green and yellow. Theoretically inaccessible regions because of a charged LSP or lack of electroweak symmetry breaking are marked in grey.

In comparison to the edge-fit analysis described in this thesis, the counting experiment has a slightly higher sensitivity in the region of $m_0$ lower than about 700 GeV. It does not target a specific decay mode and is therefore able to extend the exclusion limit to parameter regions, in which the decay $\tilde{\chi}_2^0 \rightarrow \ell\ell\tilde{\chi}_1^0$ is not dominant. In this low-$m_0$ region, the exclusion limit derived with the counting experiment extends up to $m_{1/2}$ of 450 GeV (at $m_0 = 100$ GeV).

For $m_0$ higher than 1000 GeV, the exclusion drops to $m_{1/2} = 200$ GeV. In this region of high $m_0$, the sensitivity of the counting experiment is worse than the sensitivity of the edge-fit analysis, because the higher $E_T$ and $H_T$ selections of the counting experiment reduce its signal efficiency. Here, the exclusion derived with the edge-fit extends up to $m_{1/2} = 260$ GeV.

Overall, the two analyses provide sensitivity to different regions of the CMSSM parameter space. Therefore, they complement each other very well.
5.6.4 Possibilities of Improvement

The neutralino decay via an on-shell $Z$ boson is dominant in regions of higher $m_{1/2}$ in the CMSSM parameter space. This results especially in an enhancement of the $Z$ peak instead of a mass edge. A $Z$ contribution that exceeds the Standard Model yield is, however, not recognised as signal in this search method, because no prediction method for $Z$ background is included.

The data-MC comparison shows that no sign of non-SM $Z$ production is observed in the signal region. Still, extending this analysis by a data-driven background prediction for events on the $Z$ peak, e.g. using a jet-$Z$ balance technique as described in \cite{118}, would add additional sensitivity to this kind of signal and improve the derived exclusion limit. This extension however, is considered to be beyond the scope of this thesis.
6. Summary

A search for Supersymmetry in events with two light, oppositely charged leptons was performed on 4.98 fb$^{-1}$ of proton-proton collision data taken in 2011 by the CMS experiment. The flavour-correlated production of leptons in a neutralino decay, $\tilde{\chi}_0^2 \rightarrow \ell\ell\tilde{\chi}_1^0$, was targeted. This decay results in a characteristic edge in the invariant-mass spectrum of the lepton pair.

The search was performed using a simultaneous combined fit to the invariant-mass distributions of $ee$, $e\mu$ and $\mu\mu$ dilepton events. The mass edge corresponding to a potential signal was expected in the same-flavour channels, $ee$ and $\mu\mu$, and the $e\mu$ channel was used to determine the flavour-uncorrelated (main) background directly from data. The fit consisted of three components: one component describing the flavour-uncorrelated background, another component for the $Z$ background contribution and one component representing the signal hypothesis. As signal hypothesis, a triangularly shaped mass edge, as expected from the targeted signal decay, was used.

The background estimation based on the $e\mu$ dilepton event yield was corrected for differences in lepton and trigger efficiencies. Therefore, the efficiencies of the lepton reconstruction and identification, and the used dilepton triggers were measured. The most important correction factor is the lepton efficiency ratio, $r_{\mu e}$, which was calculated and shown to be largely independent from other kinematic variables.

The search method was tested on simulated data and in a dedicated control region on data, and good performance was observed. Afterwards, the signal region, defined by hadronic activity of $H_T > 300$ GeV and missing transverse energy of $E_T > 150$ GeV, was investigated. No signs for physics beyond the Standard Model were found. Thus, limits on a hypothetical signal yield were set. These limits were calculated in dependence of various mass-edge positions of the signal hypothesis. Using these limits, exclusions in the parameter space of a simplified signal model for flavour-correlated lepton production and of the constrained minimal supersymmetric extension of the Standard Model were derived.

The considered simplified model is parameterised by two mass values: the gluino and the LSP mass. Using this analysis, the model is excluded for gluino masses up to 900 GeV if the LSP is lighter than 600 GeV.

Using the derived limits, the constrained MSSM is excluded for values of $m_{1/2}$ greater than 260 GeV in the range of up to $m_0 = 3000$ GeV ($\tan\beta = 10, A_0 = 0$ and $\mu > 0$). For low values of $m_0$, the excluded range extends up to $m_{1/2} = 420$ GeV.
The exclusion reach within the CMSSM was compared to a counting experiment that was performed by CMS using opposite-sign dilepton events as complementary analysis. A higher sensitivity of the edge-fit analysis compared to this counting experiment was found in the region of $m_0$ larger than 700 GeV. Main reason for this sensitivity difference is the looser event selection of the edge-fit analysis in terms of $H_T$ and $E_T$. For $m_0$ lower than 700 GeV, the sensitivity was found to be slightly lower than the counting experiment sensitivity, due to the more general approach of the counting experiment. It was concluded that the presented edge-fit analysis complements the counting experiment very well, and it represents a valuable component of the CMS SUSY searches.
## A. Datasets

The DBS dataset paths of all used background simulation samples are listed in Table A.1 and of all used signal simulation samples in Table A.2. The NLO cross sections used to scale the datasets \[78–81\] are also given in both tables.

Remark: the T3lh simplified model has originally been named “T1lh”. The DBS dataset path of this sample still uses the deprecated description T1lh.

### Table A.1: DBS paths of the individual background simulation samples.

<table>
<thead>
<tr>
<th>DBS path</th>
<th>NLO cross section</th>
</tr>
</thead>
<tbody>
<tr>
<td>/TTJets_TuneZ2_7TeV-madgraph-tauola/Fall11-PU_S6_START42_V14B-v2/AODSIM</td>
<td>157.5 pb</td>
</tr>
<tr>
<td>/Tbar_TuneZ2_j-channel_7TeV-powheg-tauola/Fall11-PU_S6_START42_V14B-v1/AODSIM</td>
<td>22.65 pb</td>
</tr>
<tr>
<td>/Tbar_TuneZ2_s-channel_7TeV-powheg-tauola/Fall11-PU_S6_START42_V14B-v1/AODSIM</td>
<td>1.44 pb</td>
</tr>
<tr>
<td>/Tbar_TuneZ2_tW-channel-DR_7TeV-powheg-tauola/Fall11-PU_S6_START42_V14B-v2/AODSIM</td>
<td>7.87 pb</td>
</tr>
<tr>
<td>/T_TuneZ2_t-channel_7TeV-powheg-tauola/Fall11-PU_S6_START42_V14B-v1/AODSIM</td>
<td>41.92 pb</td>
</tr>
<tr>
<td>/T_TuneZ2_s-channel_7TeV-powheg-tauola/Fall11-PU_S6_START42_V14B-v1/AODSIM</td>
<td>3.19 pb</td>
</tr>
<tr>
<td>/T_TuneZ2_tW-channel-DR_7TeV-powheg-tauola/Fall11-PU_S6_START42_V14B-v1/AODSIM</td>
<td>7.87 pb</td>
</tr>
<tr>
<td>/DYJetsToLL_TuneZ2_M-50_7TeV-madgraph-tauola/Fall11-PU_S6_START42_V14B-v1/AODSIM</td>
<td>3048 pb</td>
</tr>
<tr>
<td>/DYJetsToLL_M-10To50_TuneZ2_7TeV-madgraph/Fall11-PU_S6_START42_V14B-v1/AODSIM</td>
<td>9611 pb</td>
</tr>
<tr>
<td>/ZZJetsTo2L2Nu_TuneZ2_7TeV-madgraph-tauola/Fall11-PU_S6_START42_V14B-v1/AODSIM</td>
<td>0.300 pb</td>
</tr>
<tr>
<td>/ZZJetsTo2L2Q_TuneZ2_7TeV-madgraph-tauola/Fall11-PU_S6_START42_V14B-v1/AODSIM</td>
<td>1.000 pb</td>
</tr>
<tr>
<td>/ZZJetsTo4L_TuneZ2_7TeV-madgraph-tauola/Fall11-PU_S6_START42_V14B-v1/AODSIM</td>
<td>0.076 pb</td>
</tr>
<tr>
<td>/WZJetsTo3LNu_TuneZ2_7TeV-madgraph-tauola/Fall11-PU_S6_START42_V14B-v1/AODSIM</td>
<td>0.856 pb</td>
</tr>
<tr>
<td>/WZJetsTo2L2Q_TuneZ2_7TeV-madgraph-tauola/Fall11-PU_S6_START42_V14B-v1/AODSIM</td>
<td>1.786 pb</td>
</tr>
<tr>
<td>/WWJetsTo2L2Nu_TuneZ2_7TeV-madgraph-tauola/Fall11-PU_S6_START42_V14B-v1/AODSIM</td>
<td>4.783 pb</td>
</tr>
</tbody>
</table>
Table A.2: DBS paths of the individual signal simulation samples.

<table>
<thead>
<tr>
<th>DBS path</th>
<th>NLO cross section</th>
</tr>
</thead>
<tbody>
<tr>
<td>/LM1_SUSY_sfsbht_7TeV-pythia6/Summer11-PU_S4START42_V11-v1/AODSIM</td>
<td>6.5 pb</td>
</tr>
<tr>
<td>/LM3_SUSY_sfsbht_7TeV-pythia6/Summer11-PU_S4START42_V11-v1/AODSIM</td>
<td>4.8 pb</td>
</tr>
<tr>
<td>/LM6_SUSY_sfsbht_7TeV-pythia6/Summer11-PU_S4START42_V11-v1/AODSIM</td>
<td>0.4 pb</td>
</tr>
<tr>
<td>/LM13_SUSY_sfsbht_7TeV-pythia6/Summer11-PU_S4START42_V11-v1/AODSIM</td>
<td>9.9 pb</td>
</tr>
<tr>
<td>/SMS-T1Lh_Mgluino=100to1200_mLSP=50to1150_7TeV-Pythia6Z/</td>
<td>various</td>
</tr>
<tr>
<td>... Summer11-PU_START42_V11_FastSim-v2/AODSIM</td>
<td></td>
</tr>
<tr>
<td>/mSUGRA_dilepton_m0-220to3000_m12-100to1000_tanb-10andA0-0_7TeV-Pythia6Z/</td>
<td>various</td>
</tr>
<tr>
<td>... StoreResults-PU_START42_V11_FastSim-v6/USER</td>
<td></td>
</tr>
</tbody>
</table>
B. CMS Physics Analysis Toolkit

This analysis has been carried out using the CMS Physics Analysis Toolkit (PAT) \cite{pat}. Technical details on the used software packages and the implementation of the analysis are given in the following.

B.1 Software Packages

The following shell-command listing describes the complete setup of the used CMSSW environment. This includes all version tags of used software packages that were changed from the official CMSSW release.

```bash
cmsrel CMSSW_4_2_5
cd CMSSW_4_2_5/src
cmsenv
addpkg DataFormats/PatCandidates V06-04-18
addpkg PhysicsTools/PatAlgos V08-06-36
addpkg PhysicsTools/SelectorUtils V00-03-17
addpkg RecoJets/Configuration V02-04-17
addpkg FWCore/GuiBrowsers V00-00-56
addpkg MuonAnalysis/MuonAssociators V01-13-00
addpkg PhysicsTools/Configuration V00-10-15
addpkg RecoTauTag/Configuration V01-02-03
addpkg RecoTauTag/RecoTau V01-02-07
addpkg RecoTauTag/TauTagTools V01-02-00
cvs co -rV00-04-53 -dSuSyAachen UserCode/SuSyAachen
```

B.2 Object Selection

Table \ref{tab:muon_selection} lists the implementation of the muon selection using PAT. Table \ref{tab:electron_selection} lists the implementation of the electron selection using PAT.
Table B.1: Muon selection using the PAT::Muon class.

<table>
<thead>
<tr>
<th>Description</th>
<th>PAT::Muon memberfunction / expression</th>
<th>Cut</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acceptance</td>
<td>pt</td>
<td>&gt; 10.</td>
</tr>
<tr>
<td></td>
<td>abs(eta)</td>
<td>≤ 2.4</td>
</tr>
<tr>
<td>Muon ID</td>
<td>muonID(&quot;GlobalMuonPromptTight&quot;)</td>
<td>True</td>
</tr>
<tr>
<td></td>
<td>isTrackerMuon</td>
<td>True</td>
</tr>
<tr>
<td></td>
<td>track.numberOfOfValidHits</td>
<td>≥ 11</td>
</tr>
<tr>
<td></td>
<td>track.ptError / track.pt</td>
<td>≤ 0.1</td>
</tr>
<tr>
<td>Impact parameter</td>
<td>abs(dxy(pv))</td>
<td>≤ 0.02</td>
</tr>
<tr>
<td></td>
<td>abs(dz(pv))</td>
<td>≤ 1.0</td>
</tr>
<tr>
<td>Isolation</td>
<td>(isolationR03.hadEt + isolationR03.emEt + isolationR03.sumPt) / pt</td>
<td>≤ 0.15</td>
</tr>
</tbody>
</table>

Table B.2: Electron selection using the PAT::Electron class.

<table>
<thead>
<tr>
<th>Description</th>
<th>PAT::Electron memberfunction / expression</th>
<th>Cut</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acceptance</td>
<td>pt</td>
<td>&gt; 10.</td>
</tr>
<tr>
<td></td>
<td>abs(eta)</td>
<td>≤ 2.5</td>
</tr>
<tr>
<td>Fiducial volume</td>
<td>abs(eta)</td>
<td>≤ 1.4442</td>
</tr>
<tr>
<td></td>
<td>abs(eta)</td>
<td>≥ 1.566</td>
</tr>
<tr>
<td>Electron ID (barrel)</td>
<td>abs(deltaPhiSuperClusterTrackAtVtx)</td>
<td>≤ 0.15</td>
</tr>
<tr>
<td></td>
<td>abs(deltaEtaSuperClusterTrackAtVtx)</td>
<td>≤ 0.007</td>
</tr>
<tr>
<td></td>
<td>hadronicOverEm</td>
<td>≤ 0.1</td>
</tr>
<tr>
<td></td>
<td>sigmaIetaIeta</td>
<td>≤ 0.01</td>
</tr>
<tr>
<td>Electron ID (endcaps)</td>
<td>abs(deltaPhiSuperClusterTrackAtVtx)</td>
<td>≤ 0.10</td>
</tr>
<tr>
<td></td>
<td>abs(deltaEtaSuperClusterTrackAtVtx)</td>
<td>≤ 0.009</td>
</tr>
<tr>
<td></td>
<td>hadronicOverEm</td>
<td>≤ 0.075</td>
</tr>
<tr>
<td></td>
<td>sigmaIetaIeta</td>
<td>≤ 0.03</td>
</tr>
<tr>
<td>Conversion rejection</td>
<td>gsfTrack-&gt;trackerExpectedHitsInner.numberOfHits</td>
<td>≤ 1</td>
</tr>
<tr>
<td>Partner track finding</td>
<td>abs(convDcot) ≤ 0.02 &amp;&amp; abs(convDist) ≤ 0.02</td>
<td>False</td>
</tr>
<tr>
<td>Impact parameter</td>
<td>abs(dxy(pv))</td>
<td>≤ 0.04</td>
</tr>
<tr>
<td></td>
<td>abs(dz(pv))</td>
<td>≤ 1.0</td>
</tr>
<tr>
<td>Isolation (barrel)</td>
<td>(dr03HcalTowerSumEt + max(0., dr03EcalRecHitSumEt-1.) ... + dr03TkSumPt) / pt</td>
<td>≤ 0.15</td>
</tr>
<tr>
<td>Isolation (endcaps)</td>
<td>(dr03HcalTowerSumEt + dr03EcalRecHitSumEt ... + dr03TkSumPt) / pt</td>
<td>≤ 0.15</td>
</tr>
</tbody>
</table>
Bibliography


Acknowledgement

First of all, I would like to thank prof. Lutz Feld for not only providing me the opportunity to carry out this analysis and supervising this thesis, but also giving me guidance, advice and the possibility to present my work at workshops and conferences!

Thanks also to prof. Michael Krämer for discussions on the theoretical aspects of supersymmetry, and for being the co-referee of this work.

I would like to explicitly thank Matthias Edelhoff and Niklas Mohr for several years of very cooperative work in our common goal of finding supersymmetry.

Matthias, Niklas and my other office colleagues, Martin, Knut, Jakob, Daniel, Melissa, Stefanie and Jan, always provided an enjoyable, yet productive office environment during my PhD years. Thanks! Moreover, I would like to thank these and Prof. Dr. Lutz Feld, Dr. Katja Klein, Jennifer Merz, Jan Sammet, Jakob Wehner and all other members (and former members) of the Feld working group for the helpful atmosphere in the group. Thanks also to our secretaries, Natalie Drießen, Tanja Bingler and Daniela Gorissen, for their tireless efforts!

Within the CMS collaboration, I would like to especially thank the SUSY group and its conveners and leptonic sub-group conveners Jeff Richman, Oliver Buchmüller, Alex Tapper, David Stuart, Eva Halkiadakis, Frederic Ronga, Didar Dobur, Filip Moortgat and Ben Hooberman for their support and the constant review of this analysis. Thanks also to the analysis review committee for providing valuable criticism during the review process. Furthermore, I would like thank Marco-Andrea Buchmann, Pablo Martinez, Frederic Ronga and Konstantinos Theofilatos for their collaboration during the continuation of this analysis in the year 2012. Thanks to Henning Flächer and the other members of the Susy-CAF team for the interesting discussion we had on the planning and performing of prompt data validation after the LHC startup.

I am grateful to the teams of the Aachen and DESY computing clusters, especially Thomas Kress, Achim Burdziak, Andreas Nowack and Hartmut Stadie for their support and their patience when I exceeded my quota.

CMS is an amazing experiment. Thanks to all people who were involved in planning and constructing this detector! Thanks also to the LHC crew, who did the marvellous job of delivering over 5 fb⁻¹ of data in the past year!

I am grateful to Stefan Kawalla for discussions on the English language.

Thanks to all my friends! Your company has made my studies a lot easier, and you helped me to make the best of everything. In particular, I would like to thank Aaron, Konstantin and especially Matthias for the joint efforts of managing our studies, and for all the discussions about the world of physics and everything beyond.

My family has taken special part in who I am. I would like to especially thank my parents, Ina and Bernd, for their support in all matters, in particular for making my studies possible, and my two brothers, Christoph and Georg.
Curriculum Vitae

General Information

Name: Daniel Sprenger  
Address: Achterstr. 18  
52062 Aachen  
Germany  
Date of birth: July 8th, 1984  
Place of birth: Leverkusen, Germany  
Email: daniel.sprenger@rwth-aachen.de

Education

Since October 2008 PhD student, Rheinisch-Westfälische Technische Hochschule, Aachen. 
PhD degree presumably in November 2012, Thesis title “Search for Supersymmetry in Opposite-sign Dilepton Final States with the CMS Experiment”.


August 2008 Diploma in physics, Rheinisch-Westfälische Technische Hochschule, Aachen. 
Thesis title “Track-based Alignment of a CMS Tracker Endcap” (CMS TS-2008/025), Grade: excellent

2003-2008 Undergraduate student, Rheinisch-Westfälische Technische Hochschule, Aachen. 
Fields of study: elementary particle physics (excellent), experimental physics (excellent), theoretical physics (good), materials in electrical engineering (good).

May 2003 Abitur, Konrad-Adenauer Gymnasium (Secondary school), Langenfeld. 
Grade 1.2, Main subjects: physics (excellent), mathematics (excellent), German (excellent), social studies (good)

Work experience

2009-2011 Visiting researcher, CMS Experiment, CERN. 
Statistical data analysis in searches for supersymmetry, central DCS shifts.

2011-2012 Teaching assistant, Rheinisch-Westfälische Technische Hochschule, Aachen. 
Supervision of four bachelor students.

2009-2012 Teaching assistant, Rheinisch-Westfälische Technische Hochschule, Aachen. 
Leading basic practical courses for physics students. Preparation of experiments for lectures in experimental physics. Preparation of exercise sheets and leading small groups in elementary particle physics.

2007-2009 Tracker alignment using TIF and CRAFT data, CMS Experiment, CERN.

Leading small groups and tutorial session in mathematics for computer scientists and construction engineers.
Languages

English Fluent in speech, writing and reading
German Native language

Responsibilities

2012 Analyst in SUS-12-019, CMS Experiment, CERN.
2010-2012 Maintenance of SusyPAT, CMS Experiment, CERN.
2011 Analyst in SUS-11-011, CMS Experiment, CERN.
2009-2010 Member of the SusyCAF team, CMS Experiment, CERN.

Attended international conferences and schools

August 2012 SUSY 2012: 20th International Conference on Supersymmetry and the Unification of Fundamental Interactions, Beijing, China.
August 2010 Hadron Collider Physics Summer School, Fermilab, Chicago.
March 2009 TIPP09: Technology and Instrumentation in Particle Physics, Tsukuba, Japan.

Publications

Several publications with the CMS collaboration. Listed in the following are the ones with significant contribution.

2012 “Search for new physics in events with opposite-sign dileptons and missing transverse energy with the CMS experiment”, EPJ Web of Conferences 28, 12011.


2010 “Performance of Methods for Data-Driven Background Estimation in SUSY Searches”, CMS PAS SUS-10-001.

2010 “Validation of Kalman Filter alignment algorithm with cosmic-ray data using a CMS silicon strip tracker endcap”, JINST 5 P06007.


Talks

August 2012 “Searches for SUSY in events with di-leptons at CMS”, 20th International Conference on Supersymmetry and the Unification of Fundamental Interactions, Beijing.


September 2010  “Search for SUSY in Dilepton Events”, German CMS Annual Meeting, Aachen.
November 2009  “Search for SUSY using Dilepton Events”, 3rd Annual Workshop of the Helmholtz Alliance, DESY, Hamburg.

Interests

Ju Jutsu for the body
Board games for the mind
Piano for the soul