

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>The Standard Model and its supersymmetric extension</b>	<b>7</b>
2.1.	The Standard Model of particle physics . . . . .	9
2.1.1	The particle content . . . . .	9
2.1.2	The Lagrangian density . . . . .	10
2.1.3	The Brout-Englert-Higgs mechanism . . . . .	12
2.1.4	Limitations of the Standard Model . . . . .	14
2.2.	Supersymmetry . . . . .	15
2.2.1	The MSSM . . . . .	18
2.2.1.1	The particle content of the MSSM . . . . .	19
2.2.1.2	The Lagrangian density of the MSSM . . . . .	20
2.2.1.3	The phenomenological MSSM . . . . .	21
2.2.2	Supersymmetry breaking . . . . .	22
2.3.	Long-lived particles in the MSSM . . . . .	22
2.3.1	Previous searches and constraints from indirect searches . . . . .	24
<b>3</b>	<b>Experimental setup: Collider, detector and algorithms</b>	<b>27</b>
3.1.	The Large Hadron Collider . . . . .	29
3.2.	The CMS detector . . . . .	31
3.2.1	The tracking system . . . . .	32
3.2.2	The electromagnetic calorimeter . . . . .	34
3.2.3	The hadronic calorimeter . . . . .	35
3.2.4	The muon system . . . . .	37
3.2.5	The trigger system . . . . .	38
3.3.	Event reconstruction and particle identification . . . . .	38
3.3.1	The particle-flow algorithm . . . . .	39
3.3.2	Object reconstruction . . . . .	40
3.3.2.1	Reconstruction of primary vertices . . . . .	40
3.3.2.2	Reconstruction of tracks . . . . .	41
3.3.2.3	Reconstruction of jets . . . . .	43

---

3.3.2.4	Reconstruction of photons . . . . .	46
3.3.2.5	Reconstruction of muons . . . . .	47
3.3.2.6	Reconstruction of electrons . . . . .	48
3.3.2.7	Reconstruction of taus . . . . .	49
3.3.2.8	Reconstruction of missing transverse energy . . . . .	50
3.3.3	Event cleaning . . . . .	50
<b>3.4.</b>	<b>Simulation of events . . . . .</b>	<b>51</b>
<b>4</b>	<b>A search for highly ionising, short tracks at the CMS detector</b>	<b>53</b>
<b>4.1.</b>	<b>Motivation . . . . .</b>	<b>55</b>
<b>4.2.</b>	<b>General search strategy . . . . .</b>	<b>57</b>
4.2.1	Comparison to earlier searches . . . . .	62
<b>4.3.</b>	<b>Improved dE/dx measurement for short tracks . . . . .</b>	<b>64</b>
4.3.1	Estimation of the ionisation loss of charged particles . . . . .	65
4.3.2	Energy calibration of the silicon pixel tracker . . . . .	67
4.3.3	Discrimination of highly-ionising particles . . . . .	72
4.3.4	Discrimination improvements . . . . .	74
<b>4.4.</b>	<b>Simulated samples . . . . .</b>	<b>78</b>
4.4.1	Standard Model background samples . . . . .	78
4.4.2	Signal samples . . . . .	79
<b>4.5.</b>	<b>Event selection . . . . .</b>	<b>81</b>
4.5.1	Datasets and triggers . . . . .	81
4.5.2	Selection of signal candidate events . . . . .	83
4.5.2.1	Event-based selection . . . . .	83
4.5.2.2	Candidate track selection . . . . .	85
<b>4.6.</b>	<b>Characterisation and estimation of the Standard Model backgrounds</b>	<b>95</b>
4.6.1	Fake background . . . . .	96
4.6.1.1	Inclusive fake background estimation . . . . .	97
4.6.1.2	dE/dx shape of fake background . . . . .	102
4.6.2	Leptonic background . . . . .	104
4.6.2.1	Inclusive leptonic background estimation . . . . .	107
4.6.2.2	dE/dx shape of leptonic background . . . . .	109
4.6.3	Background estimation validation . . . . .	110
4.6.4	Systematic uncertainties . . . . .	112
4.6.4.1	Uncertainty on the fake rate . . . . .	112
4.6.4.2	Uncertainty on the dE/dx shape of fake tracks . . . . .	113

4.6.4.3 Uncertainty on the leptonic scale factor . . . . .	114
4.6.4.4 Uncertainty on the leptonic dE/dx shape . . . . .	116
<b>4.7. Optimisation of the search sensitivity . . . . .</b>	<b>116</b>
<b>4.8. Results . . . . .</b>	<b>120</b>
<b>4.9. Interpretation . . . . .</b>	<b>122</b>
4.9.1 Systematic uncertainties of simulated signal samples . . . . .	122
4.9.2 Statistical Methods/ Limit setting . . . . .	128
4.9.3 Exclusion limits . . . . .	131
4.9.4 Comparison to the search for disappearing tracks . . . . .	134
<b>4.10. Conclusion and outlook . . . . .</b>	<b>135</b>
<b>5 Measurement of the jet transverse-momentum resolution . . . . .</b>	<b>139</b>
5.1. Introduction . . . . .	141
5.2. General approach of the resolution measurement using photon+jet events . . . . .	141
5.3. Datasets and event selection . . . . .	144
5.3.1 Datasets and triggers . . . . .	145
5.3.2 Simulated sample . . . . .	146
5.3.3 Event selection . . . . .	147
5.4. Methodology of the measurement . . . . .	150
5.4.1 Validation of the method . . . . .	156
5.5. Systematic uncertainties . . . . .	158
5.6. Results . . . . .	164
5.6.1 Comparison to 2011 measurement . . . . .	166
5.6.2 Comparison to 2012 dijet measurement . . . . .	168
5.7. Discussion and conclusion . . . . .	169
<b>6 Summary . . . . .</b>	<b>171</b>



# **Part 1**

## **Introduction**



---

With the discovery of the Higgs boson at the LHC in the year 2012, the last missing piece of the Standard Model of particle physics was found [?, ?]. Thus, all particles contained in the Standard Model are discovered and all of its parameters are measured, many of them with accuracies at the per-mille level. Up to now, the Standard Model has been tested at many particle physics experiments and has proven its ability to explain - and even predict - experimental results in a remarkable way.

Nonetheless, there are strong reasons to believe that the Standard Model is not the ultimate theory of particle physics. Experimental observations as well as theoretical considerations have led to the belief that there exists physics beyond the Standard Model. For instance, the observation of Dark Matter cannot be explained within the Standard Model since no suitable Dark Matter candidate is contained. From a theoretical point of view, a major concern is related to the occurrence of quadratic divergencies in the calculation of the Higgs boson mass at higher radiative orders. The Higgs boson mass is measured at a value of around  $125\text{ GeV}^1$ , which is considered very low regarding the huge radiative corrections at the Planck scale ( $\sim 10^{19}\text{ GeV}$ ). This raises the question of what kind of mechanism is responsible for the stabilisation of the Higgs boson mass at the electroweak scale. Among others, these shortcomings of the Standard Model have led to strong efforts to develop theories that go beyond the Standard Model of particle physics.

One of these theories is able to solve the above mentioned problems by imposing a new symmetry into the Lagrangian formulation of particle physics, a so-called supersymmetry (SUSY). This symmetry relates bosons and fermions by new fermionic generators and leads to the prediction of a supersymmetric partner particle for each of the particles contained in the Standard Model. This could have drastic implications for the phenomenology of particle physics, since a doubling of the particle content is predicted. Therefore, a variety of searches for supersymmetric particles has been performed at many particle physics experiments.

This PhD thesis presents a search for supersymmetric particles in  $19.7\text{ fb}^{-1}$  of data, taken in the year 2012 at a centre-of-mass energy of  $8\text{ TeV}$  at the CMS detector. The search is motivated by supersymmetric models with nearly mass-degenerate lightest ( $\tilde{\chi}_1^0$ ) and next-to-lightest ( $\tilde{\chi}_1^\pm$ ) supersymmetric particles that have not yet been targeted by existing SUSY searches. A small mass splitting between the two particles can lead to a long lifetime of the next-to-lightest supersymmetric particle  $\tilde{\chi}_1^\pm$  because of phase space suppression. The charged  $\tilde{\chi}_1^\pm$  can therefore appear as a reconstructed track in the inner tracking system of the CMS detector. At rather low  $\tilde{\chi}_1^\pm$  lifetimes, the  $\tilde{\chi}_1^\pm$  potentially decays inside the tracker and the reconstructed track can be very short. Furthermore, since the masses of supersymmetric particles are in general higher than their Standard Model

---

<sup>1</sup>Throughout this thesis, natural units ( $\hbar = c = 1$ ) are used.

partners,  $\tilde{\chi}_1^\pm$  can be heavy and can therefore deposit much higher energies in the tracker compared to minimally ionising Standard Model particles. Therefore, the analysis strategy of the here presented analysis is to search for highly ionising, short tracks. It is the first analysis at CMS that incorporates tracks with down to three measurement and that makes use of the energy information of the silicon pixel tracker, which has been subject to an energy calibration within this thesis.

The second research objective of this thesis is a measurement of the jet transverse-momentum resolution at a centre-of-mass energy of 8 TeV at CMS. The knowledge of the jet  $p_T$  resolution is a crucial ingredient for many analyses at CMS, e. g. the measurement of the dijet cross section [?] and searches for physics beyond the Standard Model that rely on a good understanding of missing energy originating from wrongly measured jets [?].

In order to exploit the good energy resolution of the electromagnetic calorimeter at the CMS experiment, the measurement is performed using  $\gamma + \text{jet}$  events. Due to the transverse momentum balance of  $\gamma + \text{jet}$  events in the absence of further jet activity, the photon energy can be used as a measure for the true jet transverse momentum. The applied method is based on earlier measurements [?, ?] but is further developed within this thesis in order to consistently account for the influence of additional jet activity on the jet transverse-momentum response.

The thesis is structured into six main parts.

**Part 2:** This part summarises the theoretical foundations, comprising an introduction to the Standard Model of particle physics as well as to its supersymmetric extensions. A special focus is on the theoretical description and phenomenology of long-lived particles in supersymmetric models.

**Part 3:** Within this part, the experimental setup is presented, including an introduction to the Large Hadron Collider and the CMS experiment as well as a description of the algorithms used for event reconstruction and particle identification at CMS. Finally, a short introduction into the techniques of event simulation is given.

**Part 4:** In this part, the search for highly ionising, short tracks is presented. It starts with a motivation and an outline of the general search strategy. Afterwards, the calibration of the silicon pixel tracker is described and its impact on the search is discussed. Subsequently, the event selection is described and the background estimation methods are introduced. Finally, the results are presented and interpreted in the context of supersymmetric models with long-lived  $\tilde{\chi}_1^\pm$ . The last chapter of this part is devoted to a conclusion and discussion of the most important findings.

**Part 5:** This part presents the measurement of the jet transverse-momentum resolution

in  $\gamma + \text{jet}$  events recorded at CMS at  $\sqrt{s} = 8 \text{ TeV}$ . It starts with a motivation and a presentation of the general approach of the measurement. The introduction of the event selection is followed by a thorough description of the methodology. Afterwards, the systematic uncertainties are discussed. Finally, the results are presented, followed by a conclusion and discussion.

**Part 6:** This part concludes and summarises the most important results of this thesis.



## **Part 2**

# **The Standard Model and its supersymmetric extension**



## 2.1 The Standard Model of particle physics

The formulation of a relativistic quantum field theory and of spontaneous symmetry breaking (SSB) by the Brout-Englert-Higgs mechanism, allowed to build a theory which is capable of explaining almost all observations of particle physics at colliders until today. This theory is known as the Standard Model (SM) of particle physics. Its last missing piece, the Higgs boson, was found at the LHC in the year 2012 [?, ?].

The Standard Model is a  $SU(3)_C \times SU(2)_L \times U(1)_Y$  non-abelian gauge theory. “After” spontaneous symmetry breaking, its symmetries are reduced to  $SU(3)_C \times U(1)_{EM}$ . All particles that were found until today are contained in it<sup>1</sup>. Furthermore, it is able to describe three of the four fundamental forces: the strong, weak and electromagnetic force.

Despite its great success, there are many open questions that cannot be addressed within the Standard Model. These “shortcomings” of the Standard Model will be discussed in Section 2.1.4. Before, however, a small introduction to the theory (Sections 2.1.1-2.1.3) of the Standard Model is given. It is not meant as a complete description. For a thorough and extensive introduction, the reader is referred to [?, ?, ?].

### 2.1.1 The particle content

First, it should be noted, that since the Standard Model is a quantum field theory, each type of field corresponds to a different particle type and vice versa.

The Standard Model of particle physics contains three different particle types, or three different types of fields. First, there are the so-called “matter particles”, which are all spin 1/2 particles in the SM. Second, the forces are described by spin 1 vector bosons. And finally, in order to give masses to all particles, the Standard Model embeds the Higgs boson, a scalar spin 0 particle.

#### Fermions in the Standard Model

The fermionic content can be further subdivided into leptons and quarks. In contrast to quarks, leptons are not strongly interacting, thus they only couple weakly and/or electromagnetically to other particles. Both, the quarks and the leptons are ordered into three

---

<sup>1</sup>One can argue, that the right-handed neutrino, which is proven to exist, is not contained. But as at least the left-handed neutrino is embedded, we want to ignore that for a moment.

different families. Across these families, all quantum numbers are conserved. They only differ by their mass.

All left-handed particles of each family form a  $SU(2)_L$  doublet, which causes the coupling via the weak force. The right-handed partners form  $SU(2)$  singlets, thus, don't couple via the weak interaction. As quarks carry one further quantum number, the colour, they are additionally grouped into  $SU(3)_C$  triplets.

### Vector bosons in the Standard Model

As mentioned before, the vector bosons describe three of the four fundamental forces. There is one gauge boson corresponding to every generator of the above mentioned gauge groups. For  $U(1)_Y$ , it is the  $B$ -boson, for  $SU(2)_L$ , there are three gauge bosons  $W^{1,2,3}$  and finally eight gauge bosons  $G^{1\dots 8}$  for  $SU(3)_C$ , which are called gluons. As the  $B$ -field and the neutral  $W^3$ -field can mix, a change in the basis is possible “after” SSB and leads to the well known photon and  $Z$ -boson.

### The Higgs boson

The Higgs boson which was already predicted 50 years ago by Peter Higgs [?, ?] and was found by the LHC experiments CMS and ATLAS in 2012 [?, ?] plays a somehow extraordinary role. This particle is a consequence of spontaneous symmetry breaking when three degrees of freedom of the Higgs field are absorbed by the  $W$ -and  $Z$ -bosons. It is the only known fundamental scalar particle.

An overview of all Standard Model particles and their transformation properties are shown in Table 2.1. For the  $SU(3)_C$  and  $SU(2)_L$  gauge groups, the corresponding representations are given for each of the particle types. For the  $U(1)_Y$  gauge group, the corresponding quantum number, the so-called hypercharge, is depicted. If particles transform as singlets under  $SU(2)_L$  or  $SU(3)_C$ , they don't couple via the corresponding force. The hypercharges  $Y$  are determined by  $Q = Y + I_3$ , where  $Q$  is the electric charge and  $I_3$  is the third component of the weak isospin with  $I^a = \sigma^a/2$ ,  $\sigma^a$  being the Pauli matrices.

### 2.1.2 The Lagrangian density

In particle physics, the dynamics of a particle system is described by the Lagrangian density. The Lagrangian density of the Standard Model is the most general set of Lagrangian terms, that are renormalisable and contain all up to date known particles as well as the

Table 2.1: All particles contained in the Standard Model and their transformation properties under  $SU(3)_C \times SU(2)_L \times U(1)_Y$ . For the gauge groups  $SU(3)_C$  and  $SU(2)_L$ , the representations are listed whereas for  $U(1)_Y$  the hypercharge is given.

	$SU(3)_C$	$SU(2)_L$	$U(1)_Y$
Fermions:			
$(\nu_L, e_L)^T$	<b>1</b>	<b>2</b>	-1
$e_R$	<b>1</b>	<b>1</b>	-2
$(u_L, d_L)^T$	<b>3</b>	<b>2</b>	$+\frac{1}{3}$
$u_R$	<b>3</b>	<b>1</b>	$+\frac{4}{3}$
$d_R$	<b>3</b>	<b>1</b>	$-\frac{2}{3}$
Vector bosons:			
$B_\mu$	<b>1</b>	<b>1</b>	0
$W_\mu^a$	<b>1</b>	<b>3</b>	0
$G_\mu^a$	<b>8</b>	<b>1</b>	0
Higgs boson: $H$	<b>1</b>	<b>2</b>	-1

above mentioned symmetries. The full SM Lagrangian density reads as follows:

$$\begin{aligned} \mathcal{L} = & (D_\mu \Phi)^\dagger (D^\mu \Phi) - \mu^2 \Phi^\dagger \Phi - \frac{\lambda}{4} (\Phi^\dagger \Phi)^2 \\ & + \bar{L}_i^L i \not{D} L_i^L + \bar{e}_i^R i \not{D} e_i^R + \bar{Q}_{ib}^L i \not{D} Q_{ib}^L + \bar{u}_{ib}^R i \not{D} u_{ib}^R + \bar{d}_{ib}^R i \not{D} d_{ib}^R \\ & - \left( Y_{ij}^e \bar{L}_i^L \Phi e_j^R + Y_{ij}^u \bar{Q}_{ib}^L \Phi^c u_{jb}^R + Y_{ij}^d \bar{Q}_{ib}^L \Phi d_{jb}^R + h.c. \right) \\ & - \frac{1}{4} (B_{\mu\nu} B^{\mu\nu} + W_{\mu\nu}^a W^{a\mu\nu} + G_{\mu\nu}^a G^{a\mu\nu}), \end{aligned} \quad (2.1)$$

with  $\not{D} = \gamma_\mu D^\mu$  and the covariant derivative  $D^\mu = \partial^\mu + ig' Y_W B^\mu - ig C_1 I^a W_a^\mu - ig_S C_2 T^a G_a^\mu$  including the three gauge couplings  $g$ ,  $g'$  and  $g_S$ .  $I^a$  and  $T^a$  denote the generators of the  $SU(2)_L$  and  $SU(3)_C$ , respectively. They are connected to the three Pauli matrices and the eight Gell-Mann matrices by  $I^a = \frac{\sigma^a}{2}$  and  $T^a = \frac{\lambda^a}{2}$ . The sum of the hypercharge  $Y_W$  and the third component of the weak isospin yields the electrical charge  $Q = Y_W + I_3$ . Furthermore, it is  $C_1 = 1$  for doublets and  $C_1 = 0$  for singlets under  $SU(2)_L$ ,  $C_2 = 1$  for

triplets and  $C_2 = 0$  for singlets under  $SU(3)_C$ .

The first line in Eq. (2.1) corresponds to the kinetic term of the Higgs field and its potential. Via this Higgs field, it is possible to give masses to the  $Z$ -and  $W^\pm$ -bosons as well as the fermions. This will be explained in detail in the following Section 2.1.3. The second line describes the kinetic terms of the leptons and quarks. The index  $i$  represents the three different families ( $i = 1, 2, 3$ ). Since they are spin 1/2 particles, they can be described with the help of Dirac spinors. The left-handed leptons and quarks are described as  $SU(2)_L$  doublets,  $L_I^L = (\nu_{eL}, e_L)_i$ ,  $Q_I^L = (u_L, d_L)_i$ , the right-handed as singlets under  $SU(2)_L$   $e_i^R$ ,  $u_i^R$ ,  $d_i^R$ . Quarks carry a further quantum number, the colour, which is indicated by the index  $b$  with  $b = 1, 2, 3$ . Quarks transform as triplets under the  $SU(3)_C$  gauge group. The third line contains the interaction terms between the fermions and the Higgs boson, called Yukawa interactions. After SSB, these terms lead to the fermion mass terms as can be seen later. The last line correspond to the kinetic terms of the gauge fields. These are connected to the field strength tensors by

$$\begin{aligned} B^{\mu\nu} &\equiv \partial^\mu B^\nu - \partial^\nu B^\mu \\ W^{\mu\nu} &\equiv \partial^\mu W^\nu - \partial^\nu W^\mu - ig [W^\mu, W^\nu] \\ &= \left( \partial^\mu W_i^\nu - \partial^\nu W_i^\mu + g \epsilon_{ijk} W_j^\mu W_k^\nu \right) \frac{\sigma_i}{2} \equiv \frac{\sigma_i}{2} W_a^{\mu\nu} \\ G^{\mu\nu} &\equiv \partial^\mu G^\nu - \partial^\nu G^\mu - ig_S [G^\mu, G^\nu] \\ &= \left( \partial^\mu G_a^\nu - \partial^\nu G_a^\mu + g_S f_{abc} G_b^\mu G_b^\nu \right) \frac{\lambda_a}{2} \equiv \frac{\lambda_a}{2} G_a^{\mu\nu}. \end{aligned} \quad (2.2)$$

The factors  $\epsilon_{ijk}$  and  $f_{abc}$  are the structure constants of the corresponding Lie groups. The notation implies a summation over all indices that appear twice.

### 2.1.3 The Brout-Englert-Higgs mechanism

An essential ingredient of the Standard Model is the Brout-Englert-Higgs mechanism (BEH mechanism), earlier also called Higgs mechanism. It was developed by Peter Higgs, Robert Brout and François Englert in the 1960s [?, ?, ?, ?, ?, ?, ?]. Based on work from Sheldon Glashow [?], the BEH mechanism was later applied to a  $SU(2) \times U(1)$  gauge theory by Steven Weinberg and Abdus Salam [?, ?]. By this, a renormalisable theory of the weak and the electromagnetic interactions was born.

#### Mass terms of the gauge bosons

Due to the BEH mechanism, it is possible to give masses to the  $W^\pm$ -and  $Z$ -bosons. A scalar field  $\Phi$  (Higgs field) is required, which has a non-zero vacuum expectation value. This is possible, if the mass parameter  $\mu$  in front of the bilinear term in line one of Eq. (2.1)

is smaller than zero and  $\lambda > 0$  at the same time.

The resulting potential of the Higgs field is then the famous ‘‘Mexican hat’’ potential. Expanding the Lagrangian density around the minimum of  $\Phi = (0, v)$ , the gauge symmetries of  $SU(2)_L \times U(1)_Y$  are spontaneously broken and only the electrical charge conserving symmetry  $U(1)_{EM}$  remains. After a unitary transformation, three of the four degrees of freedom of the Higgs field are absorbed by the gauge fields. Thus, ‘‘after’’ SSB, the part of the Lagrangian containing the scalar field is as follows

$$\mathcal{L}_{\text{Higgs}} = \frac{1}{2} (\partial_\mu h^0)^\dagger (\partial^\mu h^0) - \mu^2 (h^0)^2 + \frac{1}{2} v^2 g^2 W_\mu^- W^{+\mu} + \frac{1}{4} v^2 (g^2 + g'^2) Z_\mu Z^\mu + \text{interaction terms} \quad (2.3)$$

One kinetic and one mass term for one of the degrees of freedom of the Higgs fields remains, which is the Higgs boson ( $h^0$ ). Furthermore, three of the four gauge bosons acquire a mass. The remaining gauge boson, the photon, remains massless because of the conserved  $U(1)_{EM}$  gauge symmetry.

The mass eigenstates of the gauge bosons in Eq. (2.3) are obviously different to the interaction eigenstates in Eq. (2.1). The diagonalisation of the neutral mass matrices is described by the Weinberg angle  $\theta_W$

$$\begin{pmatrix} A_\mu \\ Z_\mu \end{pmatrix} = \begin{pmatrix} \cos \theta_W & -\sin \theta_W \\ \sin \theta_W & \cos \theta_W \end{pmatrix} \begin{pmatrix} B_\mu \\ W_\mu^3 \end{pmatrix}. \quad (2.4)$$

For the charged gauge bosons, the relation between  $W_{1,2}$  and  $W^\pm$  is the following

$$W_\mu^\pm = \frac{1}{\sqrt{2}} [W_\mu^1 \mp W_\mu^2]. \quad (2.5)$$

Consequently, the masses of the gauge bosons are the following (with  $\sin \theta_W = \frac{g'}{\sqrt{g^2 + g'^2}}$ )

$$\begin{aligned} M_H &= \sqrt{2}\mu \\ M_W &= \frac{g}{\sqrt{2}} v \\ M_Z &= \frac{1}{\sqrt{2}} v \sqrt{g^2 + g'^2} \\ M_\gamma &= 0. \end{aligned} \quad (2.6)$$

The first direct observation of the  $Z$ -and  $W^\pm$ -bosons was made in  $p\bar{p}$ -collisions in the year 1983 at the Super Proton Synchrotron (SPS) at CERN [?, ?]. The experimental values of the masses are  $m_Z = 91.1876 \pm 0.0021$  GeV and  $m_{W^\pm} = 80.385 \pm 0.015$  GeV [?]. Finally, as mentioned several times before, the Higgs boson was found at the LHC in the year

2012 [?, ?]. The mass is measured to  $m_{h^0} = 125.09 \pm 0.21(\text{stat.}) \pm 0.11(\text{sys.}) \text{ GeV}$  [?].

### Mass terms of the fermions

Fermion mass terms cannot be easily inserted into the Lagrangian density in Eq. (2.1), since they would violate the imposed gauge symmetries. With the help of the BEH mechanism it is possible to generate fermion mass terms via the Yukawa interaction terms (line three of Eq. (2.1)). After spontaneous symmetry breaking, the Yukawa interactions lead to the following mass terms (colour indices are suppressed)

$$\begin{aligned} \mathcal{L}_{\text{Yukawa}} = & - \left( Y_{ij}^e v \bar{e}_i^L e_j^R + Y_{ij}^u v \bar{u}_i^L u_j^R + Y_{ij}^d v \bar{d}_i^L d_j^R + h.c. \right) \\ & + \text{interaction terms} \end{aligned} \quad (2.7)$$

The fermion masses are thus described by the following mass matrices

$$M_{ij}^e = Y_{ij}^e v \quad M_{ij}^u = Y_{ij}^u v \quad M_{ij}^d = Y_{ij}^d v \quad (2.8)$$

Since the Standard Model does not contain right-handed neutrinos, there are no gauge invariant Yukawa interactions, that could produce mass terms for neutrinos.

#### 2.1.4 Limitations of the Standard Model

Despite the great success of the Standard Model, there remain observations and theoretical considerations that cannot be answered within the SM. In the following, the most important of such “limitations” shall be reviewed.

First of all, the Standard Model suffers from the so-called hierarchy problem. It is caused by the occurrence of quadratic divergencies in the calculation of the Higgs mass. The appearance of infinities is not uncommon in higher order calculations and happens for all particles. Still, for scalar particles the infinite term is quadratically divergent, which makes a huge difference compared to the logarithmic divergencies for fermion self-energies. When considering the Standard Model valid up to the Planck scale, an extraordinary fine-tuning would be needed to cancel a large bare mass with large counter terms

$$m_{h^0}^{\text{ren}\,2} = m_{h^0}^{\text{bare}\,2} + \Delta m_{h^0}^2 \quad (2.9)$$

to end up with a Higgs mass of about 125 GeV. Thus, this renormalisation procedure, even if mathematically possible, is regarded as highly unnatural. This raises the question why the Higgs mass is so small, despite the presence of such massive radiative corrections to the bare mass. A formulation of naturalness was given by t’Hooft in 1977 [?]. He stated,

that a small parameter can only be regarded natural, if the symmetries of the theory are enhanced by setting this parameter to zero. In the Standard Model, though, there is no enhancement of the symmetries of the Lagrangian by setting  $\mu = 0$ . Thus, the small mass of the Higgs boson compared to the Planck scale is considered as highly unnatural.

A further and probably the most striking shortcoming of the Standard Model is the missing fourth fundamental force, the gravitational force. Within the SM, it is not possible to add renormalisable terms, that can describe the gravitational force. Although, gravity is not important for particle physics at energies that are accessible at current particle colliders (it only becomes important at the Planck scale  $\sim 10^{19}$  GeV), the fact that it cannot be embedded into the Standard Model leads to an understanding of the SM as an effective theory, only valid for low energies. Thus, it is obviously not an ultimate theory and something must be beyond.

Furthermore, in particle physics there is always the wish to describe nature with a theory as simple as possible. This usually implies the effort to embed the Standard Model into a higher symmetry group. To achieve a simplification by unifying the three fundamental forces is usually done within so-called Grand-Unified-Theories (GUTs). Calculating the running of the coupling constants in the Standard Model, the couplings seem to meet at a scale of  $M_{\text{GUT}} \sim 10^{15}$  GeV. Unfortunately, they don't meet exactly. Therefore, a unification is not achievable in the Standard Model under the assumption that there are no new particles up to the GUT scale.

Finally, there is experimental evidence that cannot be explained within the Standard Model. Astrophysical observations suggest that there is a large amount of Dark Matter (DM) in the universe, that cannot be explained with the particle content of the Standard Model. Measurements of the velocity curves of galaxies, e.g. M33 [?] show discrepancies between the observed velocities and the velocities predicted from visible matter. Furthermore, the measurement of the cosmic microwave background [?] together with type Ia supernovae measurements [?] hint at a large DM fraction. The share of non-visible matter to the total amount of matter in the universe is estimated to be  $\sim 84\%$  [?]. Unfortunately, there is no suitable (only weakly interacting) candidate within the SM, that can make up the full DM contribution.

In the following Chapter 2.2, a theory is introduced, that can address most of the above mentioned problems. This theory is called supersymmetry.

## 2.2 Supersymmetry

As noted in the last chapter, the Higgs boson mass suffers from quadratic divergencies through radiative corrections. The reason for the quadratic divergencies is due to the fact, that the Lagrangian density does not contain further symmetries for  $\mu \rightarrow 0$ . This behaviour is typical for scalar particles. For fermions, on the other hand, there is a further symmetry for  $m_f = 0$ . The Lagrangian density becomes invariant under chiral transformations of the form  $\Psi \rightarrow e^{i\bar{\alpha}\frac{\tilde{\sigma}}{2}\gamma_5}\Psi$ . Although the mass terms of the fermions break this symmetry, it protects the fermions against large radiative corrections.

Due to these considerations, it seems natural to transfer the mechanism of protecting fermion masses by a chiral symmetry to scalar masses as well. This can only be done by connecting fermions with bosons and vice versa. In the 1970s, exactly this linking between the different particle types was achieved. A work from Gol'fand and Likhtman in the year 1971 stated that an extension of the Poincaré algebra is possible via fermionic generators [?]. R. Haag, J. Lopuszanski and M. Sohnius finalised these considerations by showing that with the help of fermionic generators a connection between space-time symmetries and internal symmetries is possible [?]. The extensions of a symmetry group by fermionic generators is called supersymmetry (SUSY). These were the foundation of supersymmetric theories.

In the following, the most important aspects of supersymmetric theories are discussed. For a detailed introduction the reader is referred to [?, ?, ?].

In the subsequent sections, the description is restricted to the case of  $\mathcal{N} = 1$  supersymmetry, i. e. there is only one supersymmetric generator and thus only one supersymmetric partner for every particle. A supersymmetric transformation transfers every bosonic state into a fermionic state and vice versa

$$\begin{aligned} Q |\text{boson}\rangle &= |\text{fermion}\rangle \\ Q |\text{fermion}\rangle &= |\text{boson}\rangle. \end{aligned} \tag{2.10}$$

The most important (anti-) commutation relations for SUSY algebra with fermionic gen-

erators  $Q$  are

$$\begin{aligned}\{Q_\alpha, Q_\beta\} &= \{Q^\dagger_{\dot{\alpha}}, Q^\dagger_{\dot{\beta}}\} = 0, \\ [Q_\alpha, P_\mu] &= [Q^\dagger_{\dot{\alpha}}, P_\mu] = 0, \\ \{Q_\alpha, Q^\dagger_{\dot{\alpha}}\} &= 2\sigma^\mu_{\alpha\dot{\alpha}}P_\mu.\end{aligned}\tag{2.11}$$

Here,  $P_\mu$  denotes the four component generator of translations and  $\sigma^\mu = (\mathbb{1}, \sigma_i)$  with the Pauli matrices  $\sigma_i$ . From the second relation in (2.11) follows

$$[Q_\alpha, P^2] = [Q^\dagger_{\dot{\alpha}}, P^2] = 0.\tag{2.12}$$

Equation 2.12 implies, that particles that are transformed into each other with the generator  $Q$  need to have the same eigenvalues of  $P^2$ , thus, also the same masses. This obviously imposes a problem, since no “scalar electron” was ever found with a mass of 0.51 MeV. Therefore, SUSY must be broken. There are many ideas how the breaking can actually happen. However, since up to now, only little is known about the breaking mechanism, usually further terms which break SUSY explicitly, are added by hand to the Lagrangian density. These terms can parametrise SUSY breaking without any knowledge about the breaking mechanism. One condition is however imposed on the supersymmetry breaking terms: they should not spoil the naturalness of the new theory, i. e. no new quadratic divergencies shall occur due to these terms. Therefore, they are called soft-breaking terms. How they actually look, will be explained in the Section 2.2.1.

Finally, it shall be discussed how a supersymmetric extension of the Standard Model can give possible answers to the shortcomings of the Standard Model, discussed in Section 2.1.4.

Radiative corrections by fermions always have a factor  $-1$  compared to bosonic corrections. Thus, calculating radiative corrections of the Higgs boson mass in a supersymmetric theory leads in addition to the corrections by SM particles to further corrections by SUSY particles. If SUSY were exact, the quadratic divergencies  $\Delta m_h^2$  would exactly cancel. However, as argued, SUSY must be broken. The cancellation of quadratic divergencies can therefore only be assured, if the breaking is “soft”, i. e. only logarithmic divergencies remain.

Embedding gravity into a four dimensional quantum field theory proves difficult because of the non-renormalisability of the resulting theory. A possible answer, beyond quantum field theory, are string theories [?], where the existence of a fundamental length - that of the strings - regularises the ultraviolet behaviour and allows for a quantum description of objects in curved space. Yet, the only known procedure to incorporate fermions within a string Lagrangian density consists in adding supersymmetry as a condition. Supersym-

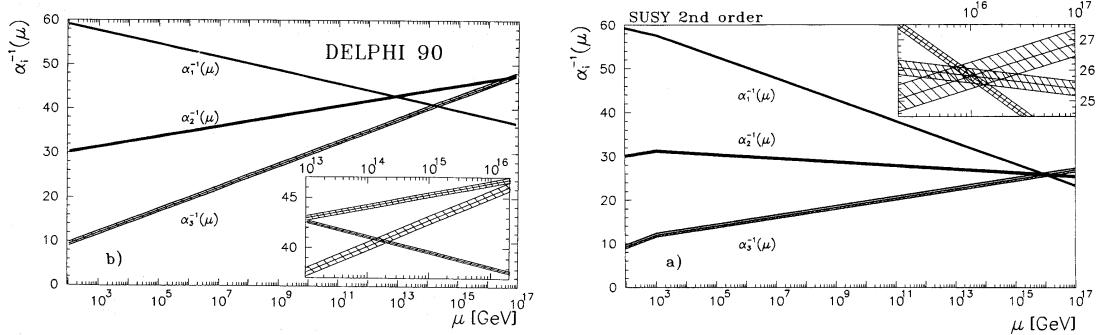


Figure 2.1: The running of the gauge couplings in the Standard Model (left) and in the minimal supersymmetric extension of the SM (right). Taken from [?].

metry thus appears as a fundamental ingredient for a description of quantum gravity.

The renormalisation group equations change under a supersymmetric extension of the Standard Model. By this, a unification of the gauge couplings at a GUT scale of about  $10^{16}$  GeV is possible, as can be seen in Fig. 2.1. It can be nicely seen, that all three gauge couplings cross each other within the uncertainties.

Besides these arguments, SUSY can also provide an answer to the problem of non-visible matter in the universe. If the conservation of the so-called R-parity is required, the lightest supersymmetric particle (LSP) is stable. If this particle is only weakly interacting, it can serve as a good candidate to explain fully or partially the sources of the relic density. R-parity is a multiplicative quantum number with

$$\begin{aligned} P_R &= 1 && \text{SM particles} \\ P_R &= -1 && \text{SUSY particles.} \end{aligned} \tag{2.13}$$

If R-parity is conserved, only terms are allowed in the Lagrangian density, that contain an even number of supersymmetric particles. Therefore, no single SUSY particle can decay into only SM particles and thus, the LSP is stable. The following discussions are restricted to R-parity conserving supersymmetric models.

### 2.2.1 The MSSM

The supersymmetric extension of the Standard Model with a minimal particle content is called the Minimal Supersymmetric Standard Model (MSSM). In the following section, the particle content of the MSSM is introduced.

### 2.2.1.1 The particle content of the MSSM

In  $\mathcal{N} = 1$  supersymmetry, every SM particle has exactly one supersymmetric partner particle, which leads to a doubling of the particle content in the MSSM with respect to the SM<sup>1</sup>. Additionally, it is necessary to introduce a second Higgs doublet to ensure the holomorphicity of the superpotential in the presence of mass terms for the up-type particles. Furthermore, the MSSM only stays free from anomalies if there is a further Higgs doublet [?]. This leads to the fact, that in the MSSM, there are five Higgs bosons instead of only one as in the SM.

In supersymmetry, all particles and their partner particles are described by so-called supermultiplets. Since the generators of the gauge group commute with the generators of supersymmetry, all particles within one supermultiplet have same quantum numbers, besides the spin. In a renormalisable theory, there are two different types of supermultiplets: chiral multiplets, which contain a two-component Weyl spinor describing the fermionic degrees of freedom and a complex scalar field for the bosonic degrees of freedom; vector multiplets containing a vector field and a two-component Weyl spinor. The complete particle content of the MSSM is depicted in Tables 2.2 and 2.3. Since in supersymmetric theories only left-handed Weyl spinors appear in the Lagrangian density, the right-handed particles are described as charge conjugated spinors of the left-handed spinors.

Table 2.2: Chiral supermultiplets in the MSSM.

	spin 0	spin $\frac{1}{2}$	$SU(3)_C, SU(2)_L, U(1)_Y$
squarks/quarks	$(\tilde{u}_L, \tilde{d}_L)$	$(u_L, d_L)$	<b>3, 2, <math>+\frac{1}{3}</math></b>
	$\tilde{\bar{u}}_L = \tilde{u}_R^\dagger$	$\bar{u}_L = (u_R)^c$	<b><math>\bar{3}, 1, -\frac{4}{3}</math></b>
	$\tilde{\bar{d}}_L = \tilde{d}_R^\dagger$	$\bar{d}_L = (d_R)^c$	<b><math>\bar{3}, 1, +\frac{2}{3}</math></b>
sleptons/leptons	$(\tilde{\nu}_{eL}, \tilde{e}_L)$	$(\nu_{eL}, e_L)$	<b>1, 2, -1</b>
	$\tilde{\bar{e}}_L = \tilde{e}_R^\dagger$	$\bar{e}_L = (e_R)^c$	<b><math>\bar{1}, 1, +2</math></b>
Higgs/higgsinos	$(H_u^+, H_u^0)$	$(\tilde{H}_u^+, \tilde{H}_u^0)$	<b>1, 2, +1</b>
	$(H_d^0, H_d^-)$	$(\tilde{H}_d^0, \tilde{H}_d^-)$	<b>1, 2, -1</b>

<sup>1</sup>The supersymmetric partner particles of the fermions are called sfermions, whereas the partner particles of the gauge (Higgs) bosons are referred to as gauginos (higgsinos).

Table 2.3: Vector supermultiplets in the MSSM.

	spin $\frac{1}{2}$	spin 1	$SU(3)_C, SU(2)_L, U(1)_Y$
gluinos/gluons	$\tilde{g}$	$g$	<b>8, 1, 0</b>
winos/ $W$ -bosons	$\tilde{W}^\pm, \tilde{W}^0$	$W^\pm, W^0$	<b>1, 3, 0</b>
bino/ $B$ -boson	$\tilde{B}$	$B$	<b>1, 1, 0</b>

### 2.2.1.2 The Lagrangian density of the MSSM

In the following, only the most important parts of the MSSM Lagrangian density will be described. For a complete description of the Lagrangian density, the reader is again referred to [?].

#### The superpotential

The superpotential of the MSSM contains the self interaction terms of the Higgs bosons and generates the interaction terms of the Higgs bosons with the fermions and their superpartners. As already noted, it is very common to assume R-parity conservation. Hence, no terms appear in the Lagrangian that would violate lepton or baryon number conservation and the lightest supersymmetric particle is stable. Thus, all possible terms are

$$W_{\text{MSSM}} = \mu H_u \cdot H_d - Y_u^{ij} H_u \cdot Q_L^i u_R^{cj} + Y_d^{ij} H_d \cdot Q_L^i d_R^{cj} + Y_e^{ij} H_d \cdot L_L^i e_R^{cj}, \quad (2.14)$$

with the dot product defined as in [?]

$$A \cdot B = \epsilon^{\alpha\beta} A_\alpha B_\beta = A_1 B_2 - A_2 B_1. \quad (2.15)$$

#### The soft-breaking Lagrangian density

Since supersymmetry is broken, explicit SUSY breaking terms are added to the Lagrangian density. In order not to introduce new sources of quadratic divergencies, only bilinear and

trilinear terms appear in the soft-breaking Lagrangian

$$\begin{aligned}
-\mathcal{L}_{soft}^{MSSM} = & m_{H_u}^2 H_u^\dagger \cdot H_u + m_{H_d}^2 H_d^\dagger \cdot H_d + (B\mu H_u \cdot H_d + h.c.) \\
& + m_{\tilde{Q}}^2_{ij} \tilde{Q}_{Li}^\dagger \cdot \tilde{Q}_{Lj} + m_{\tilde{u}}^2_{ij} \tilde{u}_{Ri}^{c\dagger} \cdot \tilde{u}_{Rj}^c + m_{\tilde{d}}^2_{ij} \tilde{d}_{Ri}^{c\dagger} \cdot \tilde{d}_{Rj}^c \\
& + m_{\tilde{L}}^2_{ij} \tilde{L}_{Li}^\dagger \cdot \tilde{L}_{Lj} + m_{\tilde{e}}^2_{ij} \tilde{e}_{Ri}^{c\dagger} \cdot \tilde{e}_{Rj}^c \\
& + \left( - (A_u Y_u)_{ij} H_u \cdot \tilde{Q}_{Li} \tilde{u}_{Rj}^c + (A_d Y_d)_{ij} H_d \cdot \tilde{Q}_{Li} \tilde{d}_{Rj}^c \right. \\
& \left. + (A_e Y_e)_{ij} H_d \cdot \tilde{L}_{Li} \tilde{e}_{Rj}^c + h.c. \right) \\
& + \left( M_1 \tilde{B} \tilde{B} + M_2 \tilde{W}_a \tilde{W}_a + M_3 \tilde{g}_i \tilde{g}_i + h.c. \right)
\end{aligned} \tag{2.16}$$

The first line contains mass terms for the Higgs bosons, the second and third line for the sfermions. In the fourth and fifth line the trilinear couplings between the Higgs bosons and the sfermions appear. Finally, the last line gives rise to mass terms for the gauginos (gluinos, winos, bino).

Because of the soft-breaking terms, the MSSM contains more than 100 free parameters. Constraining the MSSM is thus a difficult task and usually in experimental particle physics, constrained versions of the MSSM or assumptions at the GUT scale are used to report the impact of searches on SUSY. In the following a short introduction of the phenomenological MSSM is given. With its reduced parameter space, it allows to elaborate on long-lived particles in the MSSM in a much easier way.

### 2.2.1.3 The phenomenological MSSM

The phenomenological MSSM (pMSSM) imposes constraints that are reasonable in the sense that the pMSSM fulfils current observations and still keeps the phenomenological variety of the MSSM [?]. The following assumptions are imposed (in [?] more detailed information about these assumptions can be found):

- No new sources of CP violation,
- No flavour changing neutral currents,
- First and second generation universality.

These assumption reduce the number of SUSY parameters to only 19. The remaining free parameters are the following:

- $\tan \beta$  (the ratio of the vacuum expectation values of the two Higgs doublets)
- $M_A$  (the mass of the pseudo-scalar Higgs boson)
- $\mu$  (the Higgs mass parameter)

- $M_1, M_2, M_3$  (bino, wino and gluino mass parameters, respectively)
- $m_{\tilde{q}}, m_{\tilde{l}}, m_{\tilde{u}}, m_{\tilde{d}}$  and  $m_{\tilde{e}}$  (the first and second generation mass parameters)
- $m_{\tilde{Q}}, m_{\tilde{L}}, m_{\tilde{t}}, m_{\tilde{b}}$  and  $m_{\tilde{\tau}}$  (the third generation mass parameters)
- $A_t, A_b$  and  $A_\tau$  (third generation trilinear couplings).

## 2.2.2 Supersymmetry breaking

As already noted, the mechanism of supersymmetry breaking is unknown. There exist, however, several ideas how to spontaneously break supersymmetry. All mechanisms have in common that they need to happen at high energies in a hidden sector. ‘‘Messenger’’ particles are introduced which mediate the breaking to the TeV scale. This, however, implies that supersymmetry breaking is a question of extraordinary high energies and one can parametrise the breaking by the soft breaking terms introduced in Section 2.2.1.2. The most popular breaking mediation mechanisms are gravity-mediated supersymmetry breaking [?] and gauge-mediated supersymmetry breaking [?].

## 2.3 Long-lived particles in the MSSM

There are various mechanisms how particles can be long-lived, such as small couplings or (almost) conserved quantum numbers. For a comprehensive review, the reader is referred to [?].

In this thesis, the focus is set on particles that have a long lifetime due to a small decay phase space. A phase space suppression is possible when the mass splitting between the decaying particle and one of the decay products is very small. In Part 4, a search for highly ionising, short tracks is presented. This search is motivated by long-lived charginos, that are nearly mass-degenerate with the lightest supersymmetric particle, the neutralino. The underlying mechanism of this mass-degeneracy in the MSSM will be addressed in the next paragraphs.

In the MSSM, the lightest chargino ( $\tilde{\chi}_1^\pm$ ) and the lightest neutralino ( $\tilde{\chi}_1^0$ ) can be almost mass-degenerate, if the wino mass parameter ( $M_2$ ) is smaller than the bino ( $M_1$ ) and higgsino ( $\mu$ ) mass parameters. This can be seen from the chargino and neutralino mass matrices. The chargino mass matrix in the basis  $\Psi_i^+ = \begin{pmatrix} -i\tilde{W}^+, \tilde{h}_u^+ \end{pmatrix}$  and  $\Psi_i^- = \begin{pmatrix} -i\tilde{W}^-, \tilde{h}_d^- \end{pmatrix}$

is given by

$$\mathcal{M}_{\Psi^\pm} = \begin{pmatrix} M_2 & gv_d \\ gv_u & \mu \end{pmatrix}. \quad (2.17)$$

The mass eigenstates can be deduced with the help of orthogonal matrices  $V$  and  $U$ ,  $\chi_i^+ = V_{ij}\Psi_j^+$  and  $\chi_i^- = U_{ij}\Psi_j^-$ .

The neutralino mass matrix in the basis  $\Psi_i^0 = (-i\tilde{B}, -i\tilde{W}^0, \tilde{h}_u^0, \tilde{h}_d^0)$  is

$$\mathcal{M}_{\Psi^0} = \begin{pmatrix} M_1 & 0 & \frac{g'v_u}{\sqrt{2}} & -\frac{g'v_d}{\sqrt{2}} \\ 0 & M_2 & -\frac{gv_u}{\sqrt{2}} & \frac{gv_d}{\sqrt{2}} \\ \frac{g'v_u}{\sqrt{2}} & -\frac{gv_u}{\sqrt{2}} & 0 & -\mu \\ -\frac{g'v_d}{\sqrt{2}} & \frac{gv_d}{\sqrt{2}} & -\mu & 0 \end{pmatrix}. \quad (2.18)$$

The mass matrix can be diagonalised with an orthogonal matrix  $N$  leading to four different mass eigenstates of  $\tilde{\chi}_i^0 = N_{ij}\Psi_j^0$ .

It can be easily seen from the mass matrices (2.17) and (2.18), that in first order approximation - neglecting the off-diagonal elements which are of electroweak order - the lightest chargino and the lightest neutralino are both wino-like for  $M_2 < M_1, \mu$  with a mass of  $m_\chi \simeq M_2$ . Thus, the mass difference between  $\tilde{\chi}_1^\pm$  and  $\tilde{\chi}_1^0$  is only determined by higher order corrections: radiative corrections as well as tree-level mixing with other states.

Furthermore, recent analyses of the pMSSM parameter space [?, ?] show, that models with almost pure wino-like neutralinos as LSPs mostly come along with wino-like charginos being the next-to-lightest supersymmetric particle (NLSP). In [?], a parameter scan in the pMSSM parameter space is performed, flat in the 19 different SUSY parameters. Afterwards, the generated 3 million pMSSM models are confronted with theoretical constraints as well as experimental observations. Theoretical constraints are e.g. requiring stable vacua and no colour- or charge breaking minima. Furthermore, the agreement with precision electroweak data, heavy flavour physics and collider results from LEP, Tevatron and LHC is required. The accordance with relic density data is only implemented as an upper bound. Phenomenological MSSM models that survive these constraints and have a wino-like neutralino as lightest supersymmetric particle, do frequently contain a metastable chargino. In a fraction of  $\sim 25\%$  of these models, the metastable chargino decays inside the tracker, calorimeter or muon chamber. The mass splitting between chargino and neutralino in these scenarios is typically of the order of  $\sim 160$  MeV [?].

Furthermore, within [?] a study has been performed which interprets the results of various beyond Standard Model searches in terms of the fraction of excluded parameter points in the pMSSM. This study shows that models with chargino lifetimes between  $1\text{ cm} \lesssim c\tau \lesssim 30\text{ cm}$  are not excluded by any of the existing searches (cf. Fig. 4.1 in Section 4.1).

### 2.3.1 Previous searches and constraints from indirect searches

Several previous searches are sensitive on SUSY scenarios with almost mass-degenerate wino-like charginos and neutralinos. In the following an overview about these previous searches will be given.

#### Searches at LEP

Several searches at LEP were hunting for almost mass-degenerate neutralino-chargino scenarios [?, ?, ?, ?]. These searches were looking for events with a high-energetic initial state radiated photon leading to missing energy in events with chargino-pair production and invisible decay products. The excluded parameter regions by these searches can be found in [?] and are depicted in Fig. 2.2. The searches were interpreted for  $M_1$  and  $M_2$  almost degenerate and with a large unified scalar mass  $m_0$  leading to sneutrino masses larger than 500 GeV. Charginos are excluded up to a mass of 92.4 GeV [?].

#### Searches at ATLAS at 7 and 8 TeV

At the ATLAS experiment at the LHC, searches for events with a disappearing track signature were performed at  $\sqrt{s} = 7\text{ TeV}$  [?] as well as at  $\sqrt{s} = 8\text{ TeV}$  [?]. Furthermore, a search for metastable particles with high ionisation losses was performed with  $\sqrt{s} = 8\text{ TeV}$  data [?]. These searches were interpreted within an anomaly-mediated supersymmetry breaking model [?] with  $\tan\beta = 5$  and  $\mu > 0$ . The excluded parameter space by these searches is shown in Fig. 2.3. Models with charginos down to lifetimes of 0.06 ns could be excluded.

#### Searches at CMS at 7 and 8 TeV

There are several searches at the CMS experiment at the LHC that are sensitive to long-lived wino-like charginos. Among them is the search for long-lived charged particles [?], which searched for heavy particles with large energy deposits in the tracker at  $\sqrt{s} = 7\text{ TeV}$  and  $\sqrt{s} = 8\text{ TeV}$ . Furthermore, there is the search for disappearing tracks [?] which analysed events with disappearing tracks in the tracker with respect to wino-like charginos almost mass-degenerate with the lightest neutralino. This search was performed at the

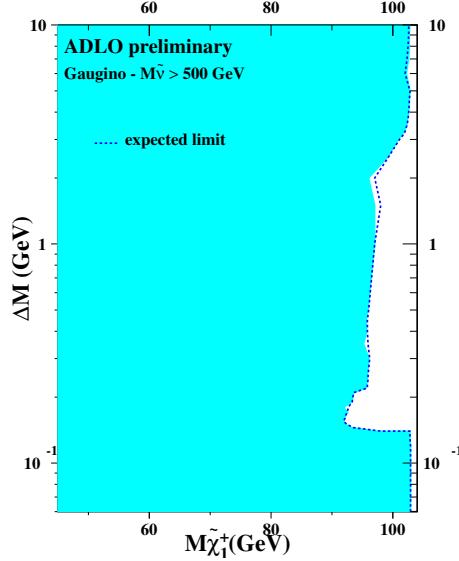


Figure 2.2: Observed and expected exclusion limits by LEP searches in the  $m_{\tilde{\chi}_1^\pm}$  –  $\Delta m (\tilde{\chi}_1^\pm, \tilde{\chi}_1^0)$  plane for almost degenerate  $M_1$  and  $M_2$  and a large unified scalar mass  $m_0$ . Taken from [?].

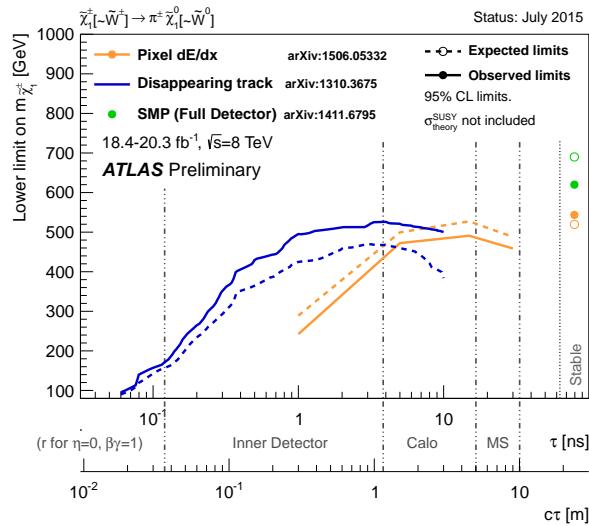


Figure 2.3: Excluded parameter space by ATLAS searches in the  $m_{\tilde{\chi}_1^\pm}$  –  $\tau_{\tilde{\chi}_1^\pm}$  plane for an AMSB model with  $\tan \beta = 5$  and  $\mu > 0$ . Only chargino pair production is taken into account. The area below the curves is excluded. Taken from [?].

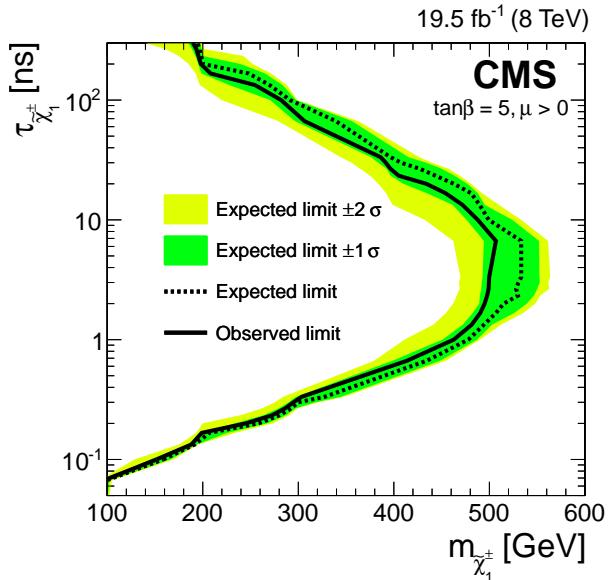


Figure 2.4: Excluded parameter space by the Disappearing track search of CMS in the  $\tau_{\tilde{\chi}_1^\pm} - m_{\tilde{\chi}_1^\pm}$  plane for wino-like charginos. The region left to the curve is excluded. Taken from [?].

CMS experiment at a centre-of mass energy of  $\sqrt{s} = 8$  TeV. Since the latter search is more sensitive to shorter lifetime, only the exclusion limits derived by this search are shown in Fig. 2.4. The disappearing track search by CMS shows a very similar sensitivity as the searches done at the ATLAS experiment.

### Indirect searches

Finally, also results from indirect Dark Matter searches constrain the parameter region of SUSY models with wino-like charginos and neutralinos. The most stringent limits are due to results by the Fermi Gamma-Ray Space Telescope (Fermi) [?] and the High Energy Spectroscopic System (H.E.S.S.) [?].

By the comparison of the observed gamma-ray signal to the theoretical prediction, Fermi sets upper limits on the Dark Matter annihilation cross-section considering six different decay channels [?].

H.E.S.S. sets upper limits on the DM annihilation cross section by the observation of the  $\gamma$ -ray line which is expected near the DM mass [?].

Recent interpretations of the Fermi and H.E.S.S. data [?, ?, ?, ?] show that thermally produced wino-like neutralinos can only account for the full relic density for  $\sim 3.1$  TeV [?], while DM masses between  $1.6 - 3.0$  TeV are excluded by Fermi and H.E.S.S. observations. Lower masses are still allowed, however, the neutralino cannot make up the full relic density.

## **Part 3**

### **Experimental setup: Collider, detector and algorithms**



## 3.1 The Large Hadron Collider

The Large Hadron Collider (LHC) [?, ?] is a particle accelerator installed in the former LEP [?] tunnel at CERN<sup>1</sup>. It is 26.7 km in circumference and consists of two separate rings, which are, in periods of operation, inhabited by two counter-circulating beams. At the interaction points of the two beams, either proton-proton collisions or heavy ion collisions take place. In this thesis, only  $pp$ -collision data from the year 2012 is analysed. Thus, all machine values cited in the following chapters and paragraphs refer to the setup for  $pp$  collisions in 2012 if not stated otherwise.

The beams are separated into bunches which rotate with a bunch spacing of 50 ns corresponding to a collision frequency of 20 MHz. Before the bunches are actually filled into the LHC ring they are pre-accelerated in other accelerators, which are in the order they are actually passed by the protons: Linac2, Proton Synchrotron Booster (PSB), Proton Synchrotron (PS), Super Proton Synchrotron (SPS). The beam energy is 450 GeV when the beams enter the LHC ring. The injector chain and the LHC ring with its experiments are visualised in Fig 3.1.

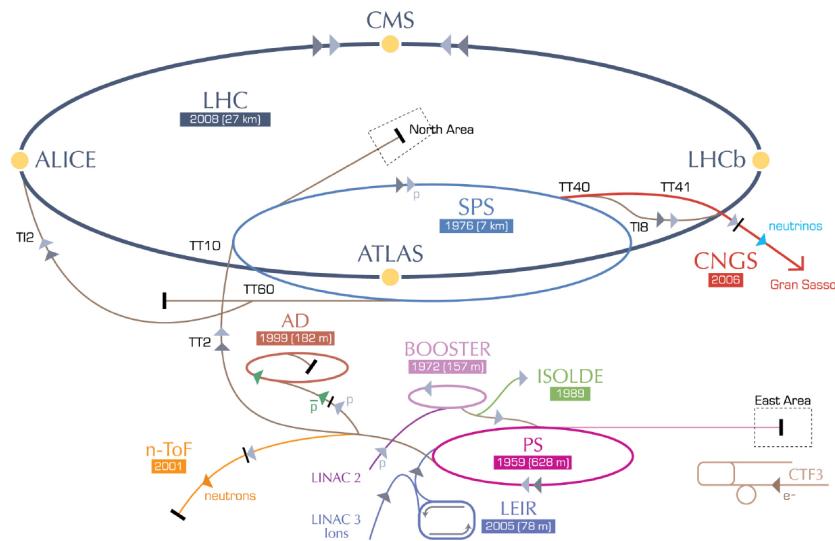


Figure 3.1: Visualisation of the LHC with its experiments and the injector chain. Taken from [?].

<sup>1</sup>European Organisation for Nuclear Research

In the LHC, the beams are kept on their circular path with the help of a magnetic dipole field of 4.76 T. Further quadrupole and sextupole magnets squeeze and focus the bunches resulting in a bunch spread of roughly 8 cm length and a Gaussian shape radius of 20  $\mu\text{m}$  RMS at the interaction point. The number of protons contained in each bunch is of the order  $10^{11}$ . The LHC hosts four main particle physics experiments: the CMS, ATLAS, LHCb and ALICE experiments. CMS [?, ?] and ATLAS [?, ?, ?] are so-called “general purpose experiments”, that are used for a variety of different physics analyses. In contrast, the LHCb [?] and ALICE [?] experiments are designed with an emphasis on CP-violation measurements and heavy ion collisions, respectively.

The number of expected events  $N$  for a given process can be expressed in terms of the corresponding cross section  $\sigma$  times the integrated luminosity  $L$

$$N = L \cdot \sigma, \quad (3.1)$$

with an integrated luminosity of  $L = \int \mathcal{L} dt$ , where  $\mathcal{L}$  is the instantaneous luminosity. The instantaneous luminosity  $\mathcal{L}$  depends on several machine parameters, such as the collision frequency  $f$ , the number of particles in the bunches  $n_1$  and  $n_2$ , the spread in the transverse plane of the bunches  $\sigma_x$  and  $\sigma_y$ , and a geometrical correction parameter  $F$  due to the crossing angle of the two bunches at the interaction point:

$$\mathcal{L} = \frac{fn_1n_2}{4\pi\sigma_x\sigma_y} \cdot F. \quad (3.2)$$

In 2012, the peak luminosity was  $7.7 \cdot 10^{33} \frac{1}{\text{cm}^2\text{s}}$ . The total integrated luminosity of  $pp$  collisions over time recorded at the CMS experiment is shown in Fig. 3.2. The recording of

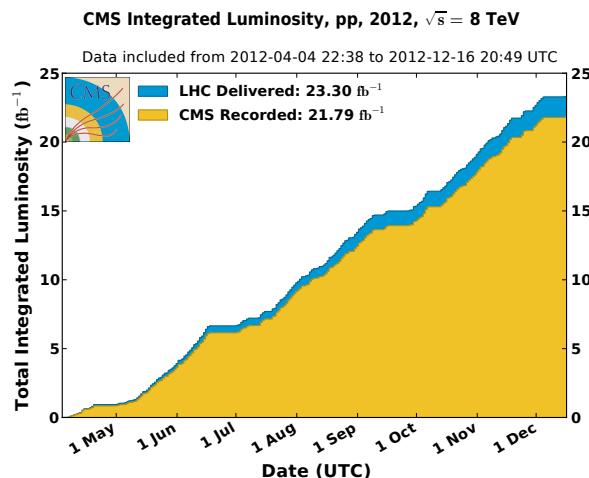


Figure 3.2: Integrated luminosity delivered by LHC (blue) and recorded by CMS (orange) in the year 2012. Taken from [?].

events at the CMS experiment is devided into so-called “runs”, where one run refers to a period of data taking in which the detector is working under stable conditions. These runs are furthermore subdevided into “luminosity blocks”. They refer to short intervals of data taking ( $< 1$  min) in which the instantaneous luminosity is stable.

## 3.2 The CMS detector

The Compact Muon Solenoid (CMS) detector [?, ?] is a general purpose detector, designed to explore particle physics phenomena up to the multi-TeV scale. The detector concept is an onion-like structure of different layers, each one made up of a different type of detector. The CMS detector measures 21.6 m in length and 14.6 m in diameter with a total weight of 12 500 tons. In Fig. 3.3, a perspective view of the CMS detector is depicted.

The coordinate system used at CMS consists of the pseudorapidity  $\eta = -\ln [\tan(\theta/2)]$  and the azimuthal angle  $\phi$ . The advantage of the pseudorapidity  $\eta$  is the Lorentz invariance with respect to the z-axis (beam axis). The angle  $\phi$  covers the direction in the  $x - y$  plane (orthogonal to the beam axis). Since the CMS experiment utilises proton-proton collisions, where the inelastic scattering happens on parton level, the initial total momentum of the

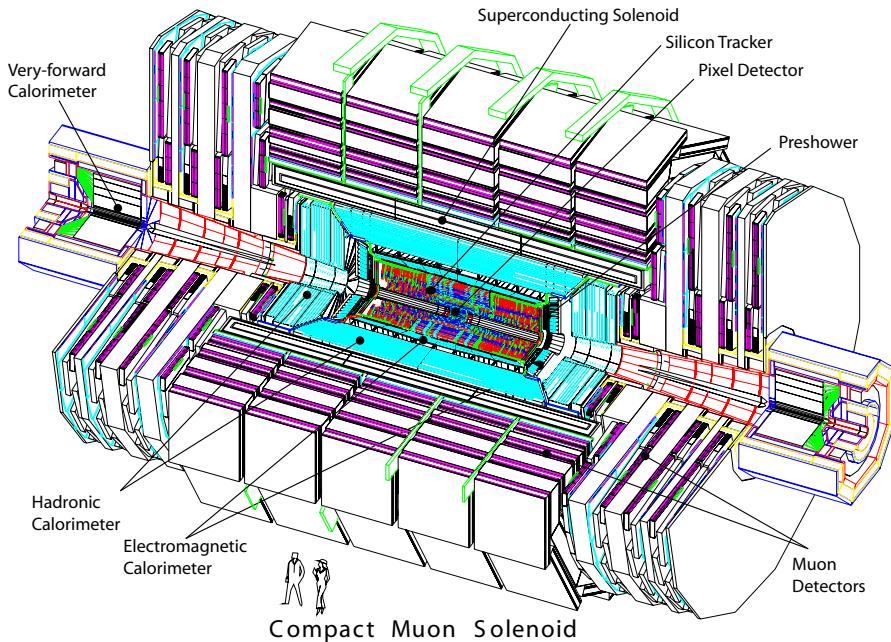


Figure 3.3: A perspective view of the CMS detector. Taken from [?]

corresponding partons are unknown. Thus, instead of exploiting the conservation of total momentum, only the conservation of transverse momentum is usually used. This can be done because the partons have no (or only a negligible amount) of initial momentum in the transverse plane. Thus the initial state is characterised by  $\sum p_T = 0$ , with  $p_T = \sqrt{p_x^2 + p_y^2}$ .

In order to bend the particle trajectories and thus to measure the momentum of charged particles a superconducting solenoid is incorporated between the calorimeter system and the muon system (see Fig. 3.3) providing a uniform axial magnet field of 3.8 T. Iron yokes contained within the muon system ensure the return of the magnetic flux.

In the following, the various detector components of the CMS detector from the inside to the outside as well as the trigger system will be explained.

### 3.2.1 The tracking system

The tracking detector [?, ?, ?] is the innermost detector at CMS. *FIXME:* It is a silicon semiconductor detector and is included for the tasks of vertex and track reconstruction by the measurement of particle energy losses. A schematic sketch of the tracker at CMS is depicted in Fig. 3.4. The tracking system is divided into two parts: the innermost tracker is a silicon pixel detector surrounded by a silicon strip detector. Since the silicon pixel tracker is closest to the interaction point and has thus to deal with much higher particle fluxes, its granularity is higher compared to the silicon strip detector.

Both parts will be explained in detail in the upcoming sections. Because, the variable

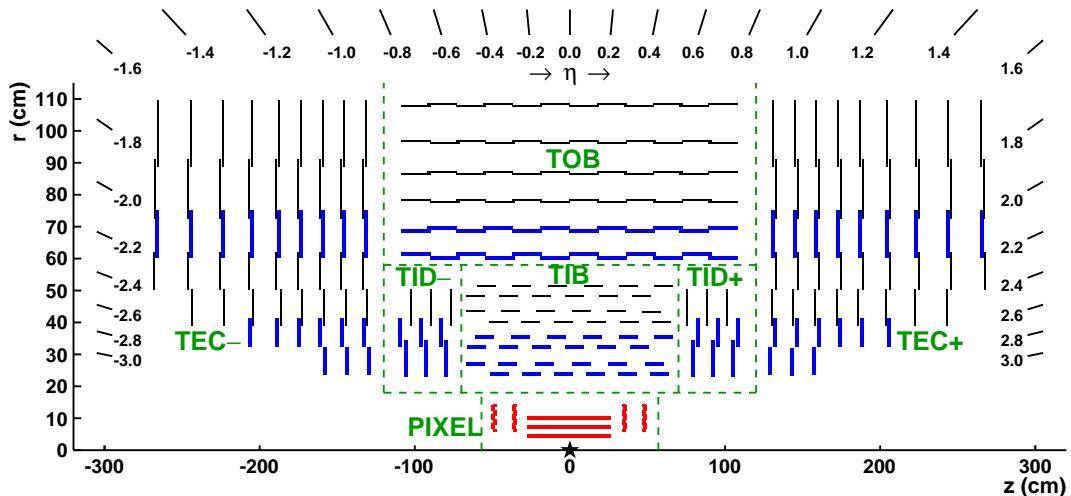


Figure 3.4: Schematic sketch of the silicon tracker at CMS in the  $r - z$  plane including the silicon pixel detector (PIXEL) as well as the different components of the silicon strip detector: tracker inner barrel (TIB), tracker outer barrel (TOB), tracker endcap (TEC), and tracker inner disk (TID). Taken from [?].

$dE/dx$ , the energy loss per path length, is used in this thesis for the search for highly ionising, short tracks (Part 4), the last part of this section is dedicated to a short description of how the energy of a traversing particle is measured with the silicon sensors. Furthermore, since a calibration of the silicon pixel detector was performed within this PhD thesis (see Section 4.3.2), an emphasis will be put on the pixel detector.

### The silicon pixel tracker

The silicon pixel detector consists of three different cylindrical layers in the barrel region at radii of 4.4 cm, 7.3 cm and 10.2 cm and two discs in the endcaps at  $z$ -distances of 34.5 cm and 46.5 cm. It is made up of 1440 modules in total (barrel + endcaps), each module comprising 8 or 16 read-out-chips (ROCs). The read-out-chips are bump bonded [?] to a pixel system of  $52 \times 80$  pixels, which are read out in double columns (see [?] on detailed information of the readout electronics). A visualisation of a part of a pixel module is shown in Fig. 3.5. In total, there are 65 million pixels comprised in the pixel detector. The large number of pixels and their small size ensure a low occupancy close to the vertex of around 0.002 – 0.02% [?] and a high hit efficiency of around 99% [?].

The silicon pixel detector is very important for the reconstruction of primary and secondary vertices as well as the reconstruction of particle tracks. Therefore, a high spatial resolution is needed. This is achieved by the small size of the pixels ( $150 \times 100 \mu\text{m}^2$ ) and the exploitation of the spread of the energy deposition across several pixels, called charge sharing (in average the energy is deposited across 3-5 pixels [?]). Exploiting the energy spread across pixels, a spatial resolution in the barrel region of  $9.4 \mu\text{m}$  in the  $r - \phi$  plane and - dependent on the incident angle of a track - a hit resolution between  $20 - 45 \mu\text{m}$  in the  $z$ -direction is achieved [?]. The spatial resolution of the primary vertex depends on the number of tracks taken into account for the reconstruction of the primary ver-

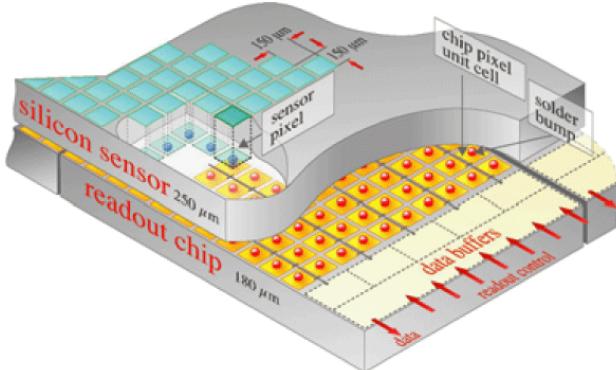


Figure 3.5: Schematic sketch of a part of a silicon pixel tracker module including the silicon sensors and the read-out-chip (ROC). Taken from [?].

tex. For more than 50 tracks originating from the primary vertex the spatial resolution is around  $10 - 12 \mu\text{m}$  for each of the three spatial dimensions [?]. The reconstruction efficiency of primary vertices is close to 100% if more than two tracks are used for the vertex reconstruction [?].

### The silicon strip tracker

The silicon strip tracker is the next-to innermost detector of the CMS detector and ranges up to a radius of 1.1 m. The barrel region consists of a tracker inner barrel (TIB) and a tracker outer barrel (TOB). The TIB has four layers with two layers equipped with so-called “stereo” modules that can measure the hit position additionally in the  $r - z$  plane. The silicon sensors in the TIB are of  $320 \mu\text{m}$  thickness with a strip pitch varying between  $80 - 120 \mu\text{m}$ . The TOB has six different layers (two layers of stereo modules) with silicon sensors of  $500 \mu\text{m}$  thickness and strip pitches between 120 and  $180 \mu\text{m}$ .

The endcaps are subdivided into a tracker endcap (TEC) and a tracker inner disk (TID). They ensure a coverage of a pseudorapidity up to  $|\eta| = 2.5$ . In each TEC, 9 disks between a  $z$ -position of  $120 \text{ cm} < z < 280 \text{ cm}$  are contained. Each of the TID comprises three disks between  $60 \text{ cm} < z < 110 \text{ cm}$ .

In the barrel region, a single-point resolution between  $23 - 52 \mu\text{m}$  in the  $r - \phi$  plane and  $230 - 530 \mu\text{m}$  in the  $z$ -direction is achieved.

### Energy measurements in the tracking system

A charged particle traversing the above mentioned silicon detectors produces electron hole pairs in the semiconducting material along its trajectory, thus loosing energy due to ionisation. For silicon, the mean energy to create an electron-hole pair is 3.61 eV at  $-10^\circ\text{C}$ . Minimally ionising particles produce an average of 22 000 electron hole pairs in silicon sensors [?]. Electrons that are subject to a hard collision with the incoming particle (so-called delta-rays), produce further ionisation and can thus lead to much higher energy deposits in the silicon. Because of the applied bias voltage at the sensors (for the creation of a depletion zone), the released electrons (holes) travel to the n-contacts (p-contacts), thereby inducing a current proportional the released energy, which is measured by the readout electronics. A more detailed description of the energy measurement in silicon sensors can be found in [?].

### 3.2.2 The electromagnetic calorimeter

The electromagnetic calorimeter (ECAL) [?, ?] encloses the tracking system and starts at a radius of 129 cm in the barrel region. It measures the energy and position of a traversing particle. It is divided into a barrel part and two endcaps, which are at a distance of 314 cm

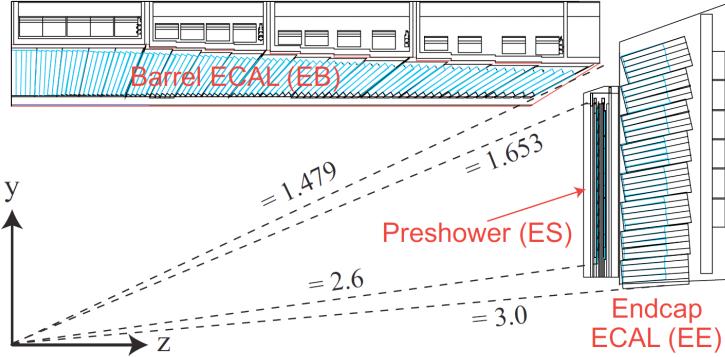


Figure 3.6: A quarter section of the ECAL in the  $r - z$  plane. Taken from [?].

from the vertex. Figure 3.6 depicts a schematic sketch of the electromagnetic calorimeter system in the transverse plane. It can be seen, that the ECAL barrel (EB) covers a pseudorapidity region up to  $|\eta| = 1.479$ . The ECAL endcaps (EE) start at  $|\eta| = 1.653$  and reach up to  $|\eta| = 3.0$ . In front of the endcaps, a preshower detector ( $1.653 < |\eta| < 2.6$ ) is installed with the main task to identify neutral pions. It additionally improves the position measurement of electrons and photons.

The EB and EE consist of lead tungstate ( $\text{PbWO}_4$ ) scintillating crystals, 61200 in the barrel region and 7324 in each of the endcaps. Their advantage is the short radiation length ( $X_0 = 0.89 \text{ cm}$ ) and a small Molière radius (2.2 cm). Thus, particles deposit their energy on rather short distances and a compact design is possible. Since they lack a fine longitudinal segmentation, the information on the shower development is limited. To detect the rather low light yield (30  $\gamma/\text{MeV}$ ) of a traversing particle, silicon avalanche photodiodes (APDs) are used in the barrel region and vacuum phototriodes (VPTs) in the endcaps. For information on the readout electronics, the reader is referred to [?].

The resolution of an energy measurement in the calorimeter can be expressed by

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{S}{\sqrt{E}}\right)^2 + \left(\frac{N}{E}\right)^2 + C^2, \quad (3.3)$$

with  $S$  referring to the stochastic term,  $N$  to the noise term, and  $C$  to a constant term. For the ECAL at CMS the parameters of Eq. (3.3) are measured to  $S = 3.63 \sqrt{\text{GeV}}$ ,  $N = 0.124 \text{ GeV}$ , and  $C = 0.26$  [?]. These numbers lead to an energy resolution of around 0.4% for an electron with  $E \approx 200 \text{ GeV}$  and around 0.6% for an electron with  $E \approx 50 \text{ GeV}$ .

### 3.2.3 The hadronic calorimeter

The hadronic calorimeter (HCAL) [?,?] of the CMS detector is splitted into four different detector modules: the hadron barrel (HB), the hadron outer (HO), the hadron endcap

(HE) and the hadron forward (HF) calorimeters. A schematic sketch is depicted in Fig. 3.7.

The HCAL is dedicated to measuring the energy of hadrons as well as improving the estimate of the missing energy in the event because of its large pseudorapidity coverage. The latter one is achieved by the high pseudorapidity coverage ( $|\eta| < 5.0$ ) that assures the detection of most of the visible particles.

The HCAL is a so-called sampling calorimeter which consists of brass absorber material, initiating the hadronic shower, as well as active plastic scintillators. The emitted photons are read out with wavelength-shifting (WLS) fibres which are embedded into the scintillators. These in turn are connected to fibres that transfer the light to the readout system.

The hadron barrel (HB) covers the pseudorapidity range between  $-1.4 < \eta < 1.4$ . It is composed of 17 layers of absorber material (15 brass and 2 steel layers) interleaved with scintillator layers. The scintillator layers are divided into 2304 towers with a size of  $\Delta\eta \times \Delta\phi = 0.087 \times 0.087$ .

The hadron outer (HO) covers a pseudorapidity range up to  $|\eta| = 1.26$ . It is located between the solenoid and the barrel detector of the muon system. The HO is dedicated to measuring the energy of the shower leakage of hadrons. Its thickness corresponds to over ten interaction lengths.

The hadron endcap (HE) extends the pseudorapidity coverage of the HCAL up to  $|\eta| = 3.0$  and starts at  $|\eta| = 1.3$ . It consists of 2304 towers in total, which vary in size between  $5 - 10^\circ$  in the  $\phi$  direction and  $0.087 - 0.35$  in  $\eta$  direction.

Finally, the hadron forward (HF) calorimeter extends the pseudorapidity range up to  $|\eta| = 5.0$ , starting from  $|\eta| = 3.0$ . It is build out of steel plates with  $1 \text{ mm}^2$  grooves contain-



Figure 3.7: A quarter section of the CMS detector including the HCAL in a transverse view. Taken from [?].

ing quartz fibres. The emitted light by the quartz fibres is transferred to photomultipliers. The HF is divided into 13 towers where almost all towers have a spread of  $\Delta\eta \approx 0.175$  in  $\eta$  direction and  $\sim 10^\circ$  in  $\phi$  direction.

### 3.2.4 The muon system

The muon system [?, ?] is the outermost component of the CMS detector. It comprises three different types of gaseous detectors, mounted into the iron return yokes: drift tube (DT) chambers in the barrel region ( $|\eta| < 1.2$ ), cathode strip chambers (CSC) in the endcap region ( $1.04 < |\eta| < 2.4$ ) and resistive plate chambers (RPC) in the barrel as well as the endcap region ( $|\eta| < 1.6$ ) (see Fig. 3.8 for a schematic sketch of the muon system). In the barrel part of the muon system, four layers (so-called stations) of drift-tube chambers are assembled inside the iron return yoke layers at radii of 4.0, 4.9, 5.9 and 7.0 m from the beam axis. The position of a muon traversing these layers can be measured with a precision of  $\approx 100 \mu\text{m}$  in radial direction and  $\approx 1 \text{ mrad}$  in  $\phi$  direction.

The four endcap disks are made up of 468 cathode strip chambers in total. By measuring the centre-of-gravity, they achieve a spatial resolution of  $\approx 100 - 200 \mu\text{m}$  and an angular resolution of  $\approx 10 \text{ mrad}$  in  $\phi$  direction. They are designed in order to cope with a high particle flux of about  $1\text{kHz}/\text{cm}^2$  and a non-uniform magnetic field. As signals can be transferred very fast, they are used for the level-1 trigger system (see Section 3.2.5).

Finally, the resistive plate chambers cover the barrel as well as the endcap region up to a pseudorapidity of  $\eta = 1.6$ . They provide a fast response with a good time resolution enabling the exact identification of the respective bunch-crossing. It is used for the level-1

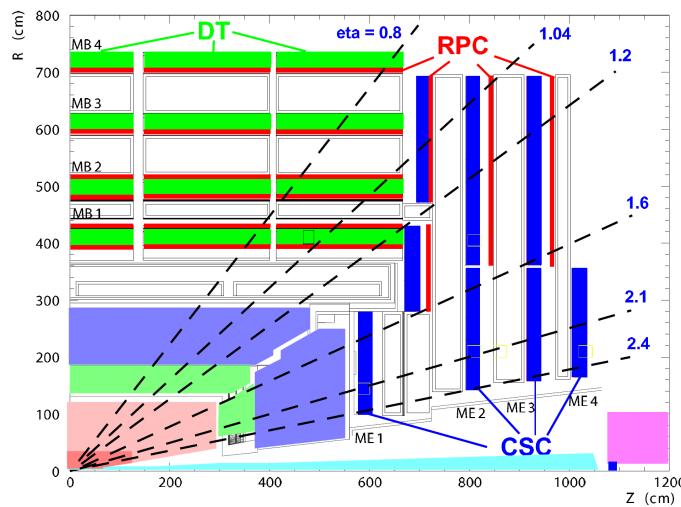


Figure 3.8: A quarter section of the CMS detector in the transverse plane with a detailed view on the muon system. Taken from [?].

trigger system as well.

### 3.2.5 The trigger system

Because of the impossibility of storing each event occurring at the CMS experiment (cf. bunch crossing rate of 20 MHz), a multistage trigger system [?] is used to achieve a drastic reduction of recorded events by nearly six orders of magnitude. It comprises two main parts: a so-called level-1 (L1) trigger system and a high-level trigger (HLT) system.

The L1 triggers need to provide a very fast decision ( $3.2\ \mu\text{s}$ , where around  $1\ \mu\text{s}$  is allocated to the actual L1 trigger calculations) whether an event shall be recorded or not. During this time, the recorded event data is held in buffers located close to the single detector components. Information from the muon system and the calorimeters is used for the L1-trigger decisions. Objects used for these decisions are so-called “trigger primitive” objects: photons, electrons, muons, jets above certain  $E_T$  and  $p_T$  thresholds and global variables like missing transverse energy,  $\cancel{E}_T$ . The design value of the number of events per second that pass this trigger stage is 100 kHz.

After a time of  $3.2\ \mu\text{s}$ , the stored data in the buffers close to the single detector components are transferred to the front-end readout buffers in case the event passed the L1-trigger requirements. By partial event reconstruction and the use of various detector signals (calorimeter, muon information followed by pixel information and full event reconstruction), higher event objects can be used to check whether HLT-trigger requirements are fulfilled. On HLT level, the decision time amounts to 50 ms and a reduction from 100 kHz to 100 Hz of event recording is achieved.

## 3.3 Event reconstruction and particle identification

A crucial ingredient of data analysis at the CMS experiment is the translation of the signal measurements in the various sub-detector components into physical objects, like muons, electrons, etc.. For this translation, i. e. for particle identification, information from all detector components is used. This is known as the particle-flow event reconstruction algorithm [?]. In Fig. 3.9, a slice through the CMS detector is shown with the signatures of different particles indicated as coloured lines.

In the next section an introduction to this algorithm is given, followed by the definition and classification of physics objects at the CMS experiment.

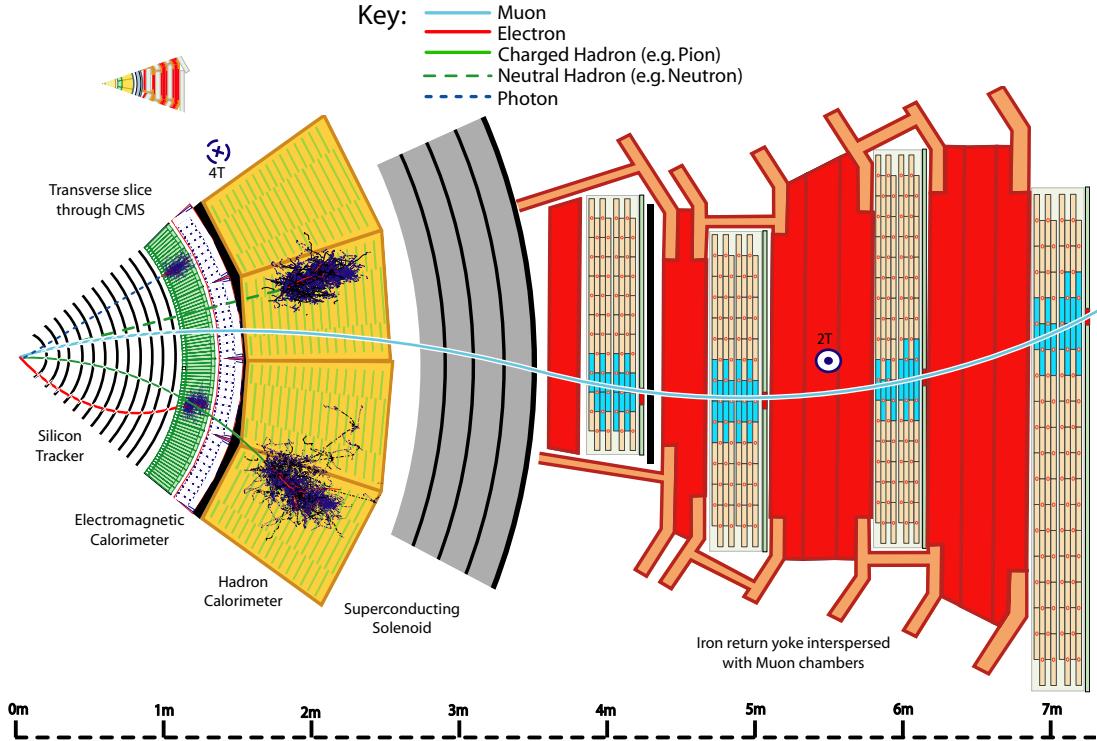


Figure 3.9: A radial slice through the CMS detector with several particle signatures indicated as coloured lines. Taken from [?].

### 3.3.1 The particle-flow algorithm

The particle-flow (PF) event description [?] aims at optimising particle identification and reconstruction by the usage of all sub-detector components of the CMS detector. There are three main building bricks used for global event description: reconstructed charged-particle tracks, calorimeter clusters, and muon tracks. The main requirements for these building bricks are a high reconstruction efficiency and a small fake rate. Therefore, a special emphasis was put on developing a very efficient tracking algorithm (see Section 3.3.2) and a well performing calorimeter clustering algorithm [?].

The particle-flow algorithm proceeds as follows for each event (the description is based on [?]):

1. For each pair of detected building bricks, a “link distance” is calculated in order to quantify the quality of their link, i. e. whether the two building bricks stem from the same particle.
2. “Blocks” are produced from the building bricks that are linked together (with a typical number of one, two or three building bricks contained in a block) using different

algorithms for different sets of building bricks (see [?] for detailed information about these algorithms). Blocks comprise either charged-particle tracks and calorimeter clusters, or several calorimeter clusters, or a charged-particle track in the tracker and a muon track in the muon system. The latter is called a global muon.

3. For each block the following steps are performed:
  - a) Each global muon where the  $p_T$  measured in both sub-detectors is compatible with the  $p_T$  measurement in the tracker is defined as a particle-flow muon and the track in both sub-detectors is removed from the block.
  - b) Next, electrons are reconstructed and identified using tracker hits and ECAL clusters. For an identified particle-flow electron, the corresponding tracker hits and the ECAL clusters (including energy deposits from Bremsstrahlung photons) are removed from the block.
  - c) Tighter track quality criteria are applied.
  - d) The compatibility of the remaining ECAL and HCAL energy deposits to the transverse momentum of the reconstructed tracks within a block is checked. This allows for the identification of particle-flow charged hadrons with a momentum estimate using tracker and calorimeter information. If the energy deposits in the ECAL or HCAL are much larger than the corresponding track  $p_T$ , the signature is identified as a particle-flow photon or particle-flow neutral hadron, respectively. All ECAL and HCAL clusters used for the identification as well as the reconstructed tracks are removed from the block.
  - e) Finally, the remaining ECAL and HCAL clusters (which are all not linked to any other building block) are identified as particle-flow photons or particle-flow neutral hadrons, respectively.

For the final identification of particles and other objects further reconstruction criteria are applied (see Section 3.3.2).

### 3.3.2 Object reconstruction

In this section, an overview of the required criteria for the identification of particles and other physics objects is given.

#### 3.3.2.1 Reconstruction of primary vertices

The reconstruction of primary vertices (PVs) is important in order to determine the location of the interaction point and to get an estimate of the corresponding uncertainty. At the CMS experiment, the reconstruction of vertices proceeds as follows [?]: First, a

track selection is performed which depends on the number of hits in the tracker, on the significance of the transverse distance to the beam spot ( $|d_0^{\text{BS}}|/\delta d_0$ ), and the normalised  $\chi^2$  of the track fit (see the subsequent Section 3.3.2.2 for more information on the track reconstruction algorithms). All selected tracks are afterwards clustered to several vertices based on their point of closest approach in  $z$ -direction to the beam spot. The clustering first identifies candidate vertices using a so-called deterministic annealing (DA) algorithm [?]. Subsequently, the candidate vertices are evaluated with the adaptive vertex fitter [?], which estimates the location of the vertices as well as performance parameters of the fit. A weight is assigned to each track that is close to one in case for a good compatibility and close to zero for a bad compatibility with the vertex. The track with the largest sum of all squared track transverse momenta,  $\sum p_T^2$ , is referred to as the primary vertex. All other vertices during a bunch-crossing are so-called pileup interactions.

Further quality criteria of the primary vertex are required in the search for highly ionising, short tracks (Part 4) and the measurement of the jet transverse-momentum resolution (Part 5). For this purpose, the “number of degrees of freedom” of the primary vertex is introduced:

$$n_{\text{dof}} = -3 + 2 \sum_{\text{tracks}} w_i. \quad (3.4)$$

The requirement of a well reconstructed primary vertex is  $n_{\text{dof}} > 4$ . Furthermore, the vertex is required to be within 24 cm in  $z$ - and 2 cm in  $r$ -direction with respect to the nominal interaction point.

### 3.3.2.2 Reconstruction of tracks

The reconstruction of tracks aims at linking several hits in the tracking system to one reconstructed track that matches the original trajectory of the particle with a high probability. With the track reconstruction, an estimate of the particle momentum as well as its position can be achieved. Track reconstruction is challenging because the large number of hits<sup>1</sup>, especially in the layers close to the interaction vertex, leads to a high combinatorial complexity. In the following, an overview of the tracking algorithm used at CMS is given. It is based on [?] and the reader is referred to this reference for more information on the reconstruction of tracks at CMS.

The tracking software used at CMS is usually referred to as the Combinatorial Track Finder (CTF). It is based on the so-called combinatorial Kalman filter [?, ?, ?] which in turn is based on the Kalman filter [?]. The Kalman filter is an algorithm that allows for the estimation of parameters of interest based on a set of observations that are subject

---

<sup>1</sup>At design luminosity around 1000 charged particles are expected to traverse the tracker in each bunch-crossing [?]

to noise or other inaccuracies. It is mathematically equivalent to a global least square minimisation for linear models with Gaussian noise.

The basic idea of the tracking algorithm at CMS is to avoid applying the combinatorial Kalman filter on all hits in one step by using an iterative procedure (called iterative tracking). A reduction of complexity can be achieved by first reconstructing tracks that are easy to identify because of e.g. a relatively high  $p_T$ . These tracks are removed afterwards and the remaining tracker hits are subject to further reconstruction iterations. The following iterations are performed (these steps refer to the setting from May to August in 2011 but are in their basic structure retained in the year 2012):

- Iteration 0: Tracks near the  $pp$ -interaction point that have three pixel hits and a  $p_T > 0.8 \text{ GeV}$  are reconstructed.
- Iteration 1: Tracks with only two pixel hits and  $p_T > 0.8 \text{ GeV}$  are reconstructed.
- Iteration 2: Low  $p_T$  tracks from the  $pp$ -interaction point are reconstructed.
- Iteration 3-5: Reconstruction of tracks that are not originating from the primary vertex or that were not found by previous iterations.

Within these iterations, the reconstruction is subdivided into four different steps:

- Seed generation: Only 2-3 hits are used to define track candidates.
- Extrapolation: Based on the expected flight path, additional hits are assigned to the candidate track using a combinatorial Kalman filter.
- Track fitting: With the usage of the Kalman filter and a smoother, the trajectory is fitted in order to estimate the track parameters.
- Setting of quality flags: Quality flags are assigned to all tracks and tracks that fail certain quality criteria are discarded.

A special task when reconstruction tracks consists of the suppression of fake tracks, i.e. tracks that are not associated with a charged particle. Therefore, the requirement of fulfilling certain quality criteria (fourth step) is crucial to substantially reduce the contamination of fake tracks. For this purpose, requirements on the following variables are imposed (the values of the parameters  $\alpha$  and  $\beta$  vary across iterations):

- A minimum number of layers in which the track has an associated hit.
- A minimum number of layers in which the track has an associated 3D hit<sup>2</sup>.

---

<sup>2</sup>A 3D hit refers to a measurement that provides 3D position information, such as a pixel hit or a strip hit in a stereo module.

- A maximum number of layers in which the track has no associated hit.
- A high quality of the trajectory fit:  $\chi^2/\text{ndof} < \alpha_0 N_{\text{layers}}$ .
- A low transverse impact parameter significance:  $|d_0^{\text{BS}}|/\delta d_0 < (\alpha_3 N_{\text{layers}})^\beta$ ,  
 $|d_0^{\text{BS}}|/\sigma_{d_0}(p_T) < (\alpha_1 N_{\text{layers}})^\beta$ .
- A low longitudinal impact parameter significance:  $|z_0^{\text{PV}}|/\delta z_0 < (\alpha_4 N_{\text{layers}})^\beta$   
 $|z_0^{\text{PV}}|/\sigma_{z_0}(p_T, \eta) < (\alpha_2 N_{\text{layers}})^\beta$ .

The variable  $d_0^{\text{BS}}$  is the distance of the track to the centre of the beam spot in the transverse plane to the beam line and  $z_0^{\text{PV}}$  is the distance along the beam line to the closest pixel. The variables  $\delta d_0$  and  $\delta z_0$  are the uncertainties on the distance (impact) parameters and  $\sigma$  is the standard deviation corresponding to the length of the beam spot in z-direction.

### 3.3.2.3 Reconstruction of jets

In Section 3.3.1, the reconstruction of particle-flow neutral and charged hadrons was already introduced. Charged and neutral hadrons are usually formed in the hadronisation process of final state quarks and gluons. Since the interest at particle colliders concerns the sum of all particles arising from one final state gluon or quark, which is referred to as a jet, the reconstruction of jets is a very important task at particle colliders. Foremost, the clustering of the particle-flow candidates to a jet is a crucial ingredient for the reconstruction of jets. Furthermore, the calibration of the measured jet transverse momentum as well as the subtraction of energy due to pileup interactions is important to ensure a good quality  $p_T$  measurement of jets at CMS. All three subjects are explained in detail in the following sections.

#### Clustering of jets

The clustering algorithm used at the CMS experiment belongs to the class of sequential recombination jet algorithms. In sequential recombination algorithms, a distance measure between two particles - or more generally objects - is used to cluster them into jet cones. For this purposes, two quantities are defined

$$d_{ij} = \min(p_{T,i}^{2k}, p_{T,j}^{2k}) \frac{\Delta_{ij}^2}{R^2}, \quad (3.5)$$

$$d_{iB} = p_{T,i}^{2k},$$

with  $\Delta_{ij}$  referring to the distance between particles  $i$  and  $j$  in the  $\eta - \phi$  plane and  $R$  denoting the radius parameter of the algorithm which determines the cluster scale. The quantity  $k$  can be any integer and specifies the behaviour of the algorithm, as will be seen later. The algorithm is an iterative procedure comprising the following steps: First, the

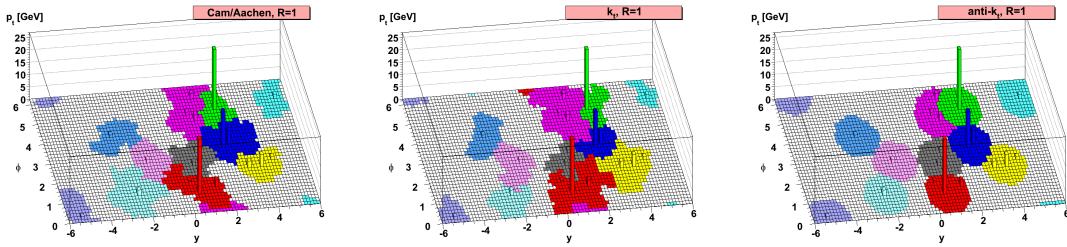


Figure 3.10: Parton-level event in the  $\eta - \phi$  plane with  $\sim 1000$  soft particles added. Three different algorithms are used to cluster the particles into jet cones (coloured areas): the Cambridge/Aachen algorithm (left), the  $k_T$  algorithm (middle), and the anti- $k_T$  algorithm (right). Taken from [?].

quantities  $d_{iB}$  and  $d_{ij}$  are calculated for every single and every pair of “objects” in the event, respectively. An object can be a particle or a set of already clustered particles. If the smallest quantity is a distance between a pair of objects, these objects are clustered together. If the smallest quantity is  $d_{iB}$ , the corresponding object is considered a jet and removed from the object list. If all objects are removed from the input list of the algorithm, the clustering comes to an end.

As noted before, the integer  $k$  in Eq. (3.5) defines different types of algorithms. For the choice of  $k = 0$ , the algorithm is referred to as Cambridge/Aachen algorithm [?, ?]. It doesn’t take the momentum of particles into account for the clustering but is only based on spatial information. The  $k_T$  algorithm [?, ?] uses the configuration with  $k = 1$ , thus starting the clustering with low momentum particles. At CMS, the anti- $k_T$  algorithm is used with  $k = -1$ . It begins with clustering high momentum particles and leads to very regular shapes of the jet cones. The behaviour of the three different clustering algorithm is illustrated in Fig. 3.10. One single parton-level event was subject to the three different jet clustering algorithms. Roughly 1000 soft particles are added to the event in order to test the adaptiveness of the jet algorithms to soft particles. It can be seen, that the Cambridge/Aachen and the  $k_T$  algorithm are both more irregular in shape, while jets clustered with the anti- $k_T$  method result in circular shapes. Furthermore, jets clustered with sequential recombination jet algorithms are infrared and collinear safe.

Throughout this thesis jets clustered with the anti- $k_T$  algorithm with a radius parameter of  $R = 0.5$  are used.

### Jet energy calibration

In order to account for possible mismeasurements of the jet transverse momentum, all jets at CMS are subject to a calibration procedure. This calibration aims at eliminating discrepancies between the measured jet  $p_T$  and the true jet  $p_T$  because of pileup interac-

tions (level-1 corrections) as well as discrepancies depending on jet  $p_T$  and  $\eta$  (level-2 and level-3 corrections). The calibration factors of the level-1, level-2 and level-3 corrections are determined fully in simulation. To ensure also a complete calibration of jets in data, additional correction factors are determined which are only applied on the jet  $p_T$  measured in data (level-2 and level-3 residual corrections). These calibration factors are determined with the help of reference objects in data, such as photons. An overview of the applied correction factors in simulation and data is shown in Fig. 3.11.

In the following, a more detailed description of the determination of the factorised calibration factors will be given. For an extensive description of the calibration approach, the reader is referred to [?]

**L1 pileup offset corrections:** In order to eliminate a mismeasurement of the jet  $p_T$  because of pileup interactions, correction factors ( $C_{\text{offset}}$ ) are determined that are supposed to restore the true jet  $p_T$ . This correction is determined in bins of jet  $p_T$ , jet  $\eta$ , the jet area  $A$ , and the event energy density  $\rho$ . The correction factors are determined with simulated samples and are applied to measured data as well as simulated events.

**L2 relative + L3 absolute corrections:** These two corrections are devoted to correct the measured jet  $p_T$  in order to ensure a uniform response ( $p_T^{\text{reco,jet}}/p_T^{\text{true,jet}}$ ) vs.  $\eta$  (L2) ( $C_{\text{rel}}$ ) and  $p_T$  (L3) ( $C_{\text{abs}}$ ). They are also determined in simulated samples and applied to simulated and measured events.

**L2L3 residual corrections:** Finally, correction factors ( $C_{\text{res}}$ ) are determined in order to remove any residual bias in data. They are estimated in data with the help of

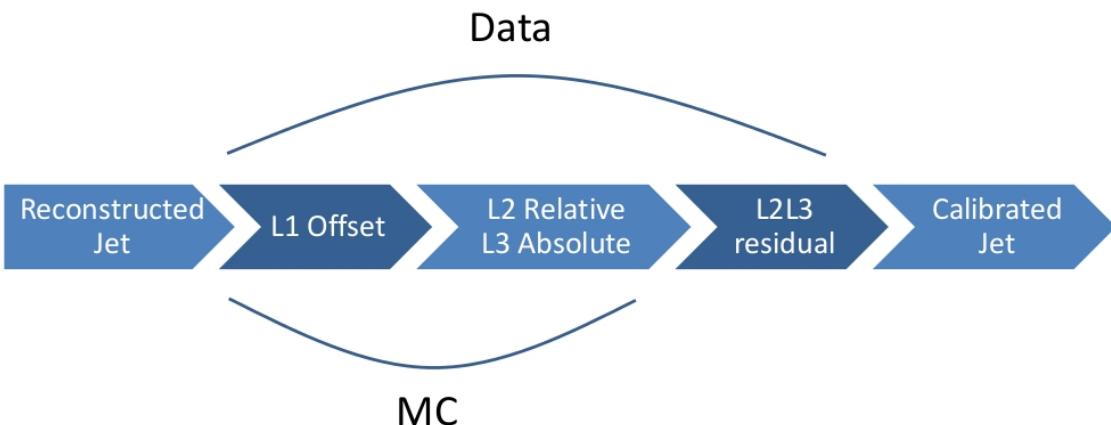


Figure 3.11: Visualisation of the various steps of the jet energy calibration at the CMS experiment. The lines embrace the applied calibration steps for simulation (MC) and data. Taken from [?].

QCD-multijet events to ensure a uniform response vs.  $\eta$  (L2) and with  $\gamma/Z +$  jets events in order to ensure a uniform response vs.  $p_T$  (L3). These factors are not applied to simulation.

The derived correction factors are applied on the measured transverse momentum in a factorised approach according to the following equation:

$$p_T^{\text{cor}} = C_{\text{res}}(p_T'', \eta) \cdot C_{\text{rel}}(\eta) \cdot C_{\text{abs}}(p_T') \cdot C_{\text{offset}}(p_T^{\text{raw}}, \eta) \cdot p_T^{\text{raw}}, \quad (3.6)$$

where  $p_T^{\text{cor}}$  is the corrected jet  $p_T$  and  $p_T^{\text{raw}}$  is the fully uncorrected jet  $p_T$ . By the application of the various correction factors, the jet transverse momentum is calibrated in a factorised manner.

Throughout this thesis, all jets are calibrated by the application of the jet transverse momentum correction factors, referred to as jet energy corrections.

### **Charged hadron subtraction**

Particles originating from pileup vertices (additional vertices during a bunch crossing) can be clustered into jets leading to higher energies of the reconstructed jet. Therefore a procedure called charged hadron subtraction (CHS) is applied. It removes reconstructed charged particles that do not originate from the primary vertex (see [?] for more details).

In this thesis, all jets are subject to charged hadron subtraction.

#### **3.3.2.4 Reconstruction of photons**

The following description of the photon reconstruction is based on [?], where also a more detailed explanation of the photon reconstruction algorithms can be found.

Photons are reconstructed from energy deposits in the ECAL. The clustering algorithms of the ECAL energy deposits do not differentiate between electrons and photons. Thus, they are the same as used for electron identification (Section 3.3.2.6). So-called “super-clusters” are formed from a “seed crystal” which has an energy deposit greater than all of its neighbouring crystals and above a certain threshold. In the barrel region a so-called “hybrid” algorithm is used which proceeds by adding in  $\phi$  direction fixed arrays of  $5 \times 1$  crystals in  $\eta \times \phi$  if their energy deposits are larger than a certain minimal energy. In the endcap region, the so-called “multi  $5 \times 5$ ” clustering algorithm is used. It proceeds by adding fixed arrays of  $5 \times 5$  crystals if their energy exceeds a certain threshold. These clustering procedures collect energy from radiating electrons as well as converted photons.

After the clustering the measured energy in a supercluster is subject to an energy correction procedure [?]. The energy estimate is finally based on the variable  $R_9$ . It is the ratio of the energy deposit within a  $3 \times 3$  array of crystals around the seed crystal

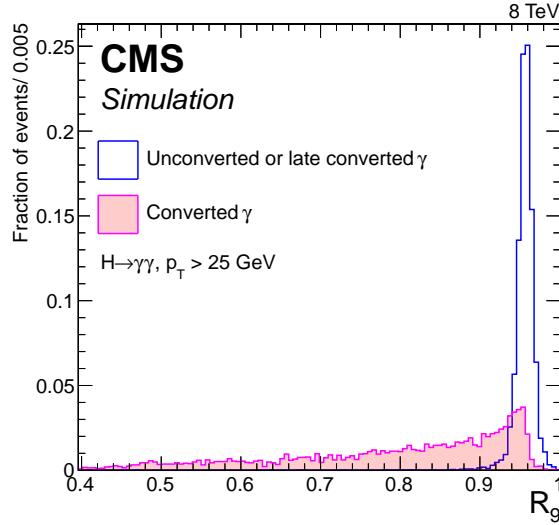


Figure 3.12: Normalised distribution of the  $R_9$  variable for converted and unconverted (or late converted) photons. Taken from [?]

divided by the total energy in the supercluster. This variable nicely discriminates between converted and unconverted photons as can be seen in Fig. 3.12. The energy spread is wider for converted photons than for unconverted photons. Thus, the energy is estimated as the sum of energy deposits in the superclusters for converted photons ( $R_9 < 0.94$  in the barrel and  $R_9 < 0.94$  in the endcap region) and as the sum of energy deposits in the  $5 \times 5$  crystal matrix for the unconverted photons.

### 3.3.2.5 Reconstruction of muons

There are three different muon definitions at CMS [?]: global muons, tracker muons, and standalone muons. They all have in common that they require energy deposits in the muon system. The reconstruction of each of the three muon types is explained in the following.

**Standalone muons:** For the reconstruction of standalone muons, all reconstructed segments in the muon system are utilised. Similar to the track reconstruction in the tracking system, Kalman filter techniques [?] are exploited to reconstruct muon trajectories in the muon chambers. A compatibility to the interaction point is imposed to reconstruct only muons produced at the LHC (no cosmic muons). Further details about the reconstruction of standalone muons can be found in [?, ?]

**Tracker muons:** To reconstruct so-called tracker muons, all tracker tracks with a  $p_T > 0.5 \text{ GeV}$  and  $p > 2.5 \text{ GeV}$  are extrapolated to the muon system. If at least one muon

segment is matched to a reconstructed track fulfilling certain quality criteria, the trajectory is considered as a tracker muon (see [?] for more detailed information).

**Global muons:** For the reconstruction of global muons an outside-in approach is utilised. For each reconstructed standalone muon, the compatibility to the reconstructed tracks in the tracking system is checked. If compatible, a global muon track is reconstructed using Kalman filter techniques. For high- $p_T$  muons, the momentum resolution can be increased this way compared to the momentum estimated using tracker information only [?].

### 3.3.2.6 Reconstruction of electrons

The reconstruction of electrons at the CMS experiment is based on a mixture of the particle-flow algorithm explained in Section 3.3.1 and a standalone approach [?]. Thus, it is a very complex procedure and only the basic approach shall be explained here. The reader is referred to [?] for a complete description of the reconstruction procedure.

Electrons are identified using clustered energy deposits in the ECAL and reconstructed tracks in the tracker system. Both inputs are tested on compatibility and finally used to estimate the electron properties.

The difficulty of the electron reconstruction lies in the possibly large energy losses due to bremsstrahlung. This can change the direction of the electron significantly and lead to a reduced efficiency of the standard track reconstruction as well as a reduced efficiency of the association of ECAL clusters to the reconstructed track in the inner tracking system. Therefore, an optimised track reconstruction for electrons is additionally performed in order to account for direction changes due to the radiation of photons. Because the dedicated electron track reconstruction can be very time consuming, the seeding of tracker hits relies already on ECAL information to reduce the number of candidate tracks (ECAL-based seeding).

Therefore, the electron reconstruction proceeds in detail as follows. First, ECAL clusters, so-called superclusters, are formed using a “hybrid” algorithm in the barrel region and a “multi- $5 \times 5$ ” algorithm in the endcaps. It is the same clustering used for photons (see Section 3.3.2.4 where the more information about these algorithms can be found). Superclusters are used, since due to bremsstrahlung, the energy deposits by electrons are usually spread out over several crystals. Additionally, so-called PF clusters are formed that add all neighbouring crystals to the above mentioned superclusters. Afterwards, the standard track reconstruction is performed which is able to reconstruct electron tracks when bremsstrahlung is negligible. Very short reconstructed tracks or tracks with a bad  $\chi^2$  are refitted using a so-called Gaussian sum filter (GSF) [?] to account for electrons with high radiative energy losses. As mentioned above, these tracks use ECAL-based seeding to reduce the number of candidate tracks. To identify these refitted tracks as candidate

electron tracks, a multivariate analysis is performed using tracker and ECAL information. All ECAL PF clusters tangent to the electron track are added in order to capture the energy of bremsstrahlung photons. Finally, track-cluster association criteria are applied in order to reduce misidentification while preserving a high reconstruction efficiency. The association criteria rely on spatial distances between the superclusters and the tracker track for ECAL-seeded tracks (GSF tracks) and rely on multivariate techniques for the tracker-seeded (standard track reconstruction) electrons.

### 3.3.2.7 Reconstruction of taus

The tau reconstruction at CMS refers to the reconstruction of hadronically decaying taus (a fraction of around 65% of taus decays hadronically). Leptonically decaying taus can be identified through the reconstruction of the muon or the electron in the final state. In the following a tau refers therefore to hadronically decaying taus only. At CMS, the main algorithm used to identify and reconstruct taus is the so-called hadron plus strip algorithm (HPS) [?, ?], which is also used in this thesis. The description of the tau reconstruction algorithm within this thesis follows [?].

The HPS algorithm consists of two separate steps. First, reconstructed PF particles are combined in a way that they are compatible with any of the tau decay modes and the four-momentum of this combined object is estimated. Second, multivariate techniques are used in order to discriminate the candidate taus against decays from gluon or quark jets and from muons and electrons. The HPS algorithm starts from reconstructed jets with a radius parameter of  $R = 0.5$  and  $p_T > 14 \text{ GeV}$  and  $|\eta| < 2.5$ .

The successful identification of a tau requires the reconstruction of a neutral pion. A neutral pion is present in most of the hadronic decays and decays into two photons with a probability of  $\sim 99\%$ . The photons in turn convert into electron pairs with a high probability inside the tracker volume. Therefore, photons and electrons are clustered into strips in the  $\eta - \phi$  plane in order to identify possible  $\pi^0$ . The initial position of the strip equals the position of the seed electron or photon, i.e. the photon or electron with the highest  $p_T$  not yet clustered into any strip. The next-highest  $p_T$  electron/photon within a certain distance ( $\eta \times \phi = 0.05 \times 0.20$ ) to the original strip is merged into the strip. Afterwards, the strip position is recalculated and the step is repeated. The clustering stops if no more electrons/photons are found.

After the reconstruction of the neutral pion, multiple decay mode hypotheses are tested by combining up to two clustered strips with one or three charged PF hadrons. For all hypotheses, several restrictions on the invariant mass of the hadronically decaying tau are imposed. Finally, the hypothesis fulfilling the invariant mass requirement and - in case this happens for more than one hypothesis - the one with the highest  $p_T$  tau is chosen.

Furthermore, for all taus in this thesis, the discriminators used for the reduction of jet,

electron and muon contamination are required to be zero (see [?] for further details on the discriminators).

Finally, a loose isolation criterion is imposed, which is calculated by the pileup corrected scalar sum of the transverse momentum of all charged particles and photons with a  $p_T$  larger than 0.5 GeV in a cone of  $R = 0.5$  around the tau.

### 3.3.2.8 Reconstruction of missing transverse energy

The particle-flow missing transverse energy, PF  $\cancel{E}_T$ , in an event is defined as the negative vectorial sum of all particle-flow particles'  $p_T$ . Thus - neglecting momentum mismeasurements - it refers to the transverse momentum of all only weakly interacting particles in the event. In order to avoid negative effects from e.g. tracker inefficiencies, non-linear calorimeter responses, etc.,  $\cancel{E}_T$  can be measured more precisely by correcting the transverse momentum of all jets contained in an event with the jet energy corrections (see Section 3.3.2.3).

For a more detailed explanation of  $\cancel{E}_T$  as well as for information on the performance of the reconstruction, the reader is referred to [?]

## 3.3.3 Event cleaning

The two analyses described in this thesis are subject to an event cleaning procedure as recommended by the CMS collaboration [?]. It refers to the application of several filters to remove events with detector signals not caused by particles from the hard interaction. In the following, the filters are listed and a short description will be given.

**HBHE noise filter:** This filter rejects anomalous HCAL noise in the barrel (HB) and in the endcap (HE). The noise in the HCAL is mainly related to the hybrid photodiodes (HPD) and readout boxes (RBX). It rejects events considering pulse shape information and anomalous high hit multiplicities in an HPD or RBX.

**CSC beam halo filter:** Particles contained in the beam halo can induce signals in the detector. Therefore, events with a muon moving parallel to the beam are rejected.

**HCAL laser event filter:** This filter provides a list of events, where the laser used for the HCAL calibration fired when a bunch crossing occurred. Detector signals from the laser can be misinterpreted as particle signals and, therefore, these events are removed.

**ECAL dead cell filter:** Events with jets pointing towards a non-working ECAL cell are rejected. This filter relies on information of the surrounding crystals as well as on “trigger-primitive” information.

**EE bad supercrystal filter:** In the ECAL endcap, two supercrystals were not working properly during the data taking period in 2012. Therefore, events where at least one of the two crystals detected unnaturally high energy deposits are removed.

**Tracking failure filter:** Sometimes, unnaturally high energy deposits are recorded in the calorimeters without corresponding tracks in the inner tracking system. The reconstruction of tracks can fail, because of too many tracker seeds or because the hard interaction was too far away from the interaction point. This filter rejects therefore events, where less than 10% of the energy in the event was measured by the tracker.

**Scraping filter (data only):** When there are at least ten tracks in the event, at least 25% of them need to be flagged as high-purity tracks as defined in [?].

**Tracking POG filter:** Events with aborted track reconstruction and/or strip tracker noise are rejected.

## 3.4 Simulation of events

A very important input for physics analyses at particle colliders is the use of simulated collision events. They serve as a comparison to measured data and can give an estimation of the expected number of background and signal events for a given selection. Furthermore, they are used to study event kinematics for Standard Model (SM) and beyond-SM processes.

The generation of simulated samples is done with the help of Monte Carlo (MC) methods (see e.g. [?]). MC techniques are used to randomly sample points according to a specified probability distribution. The underlying probability distributions for the event generation are the transition probabilities from the initial into all possible final states.

Besides the simulation of the hard interaction of the event, also the subsequent showering and hadronisation as well as the detector response needs to be simulated. In the following sections, a short introduction into the various steps of event simulation is given. For more information, the reader is referred to [?].

The first step of the generation of  $pp$ -collision events consists of the simulation of the hard process, i.e. the primary interaction of two partons of the two colliding protons. Which types of partons are involved in the primary interaction is specified by so-called

parton distribution functions (PDFs). They describe the probability of a parton (up-quarks, down-quarks, gluons, sea quarks) to carry a certain energy fraction of the proton and thus define the probability for each parton to take part in the hard interaction process.

After the simulation of the hard interaction between the selected partons, the showering and hadronisation of the particles is simulated. Throughout this thesis, the matrix element generator MADGRAPH [?] plus the event generator PYTHIA [?] or the event generator PYTHIA only are used. Since MADGRAPH only provides the generation of matrix elements, the subsequent showering and hadronisation is done by PYTHIA. For this task, particles produced in the hard interaction are matched to the particles produced in the showers in order to avoid a double-counting of emissions in multijet processes. Since MADGRAPH is responsible for the simulation of well separated partons whereas PYTHIA is designed for simulating the soft and collinear showers, the event is rejected if a parton from the matrix element generator cannot be matched within a certain distance  $\Delta R$  to any of the simulated showers (MLM matching scheme [?]). The showers are clustered with the  $k_T$  method (see Section 3.3.2.3). Additionally, events are rejected if there are more showers than partons in the event.

The formation of colour neutral particles, called hadronisation, and the subsequent decay of unstable particles is simulated with PYTHIA for all samples used in this thesis. A visualisation of the various steps performed for the generation of events is shown in Fig. ??.

In order to account for signals caused by the interaction of the non-primary partons of the interacting protons - called the underlying event (UE) - various models (tunes) are exploited. They rely on the measured signals in data in proton collisions without any hard interaction. Detector signals caused by additional interactions (pileup interactions) during a bunch-crossing, are also mixed into every simulated sample.

Finally, the interaction of the traversing particles with the detector material as well as the response and the readout electronics is simulated with GEANT4 [?, ?] throughout this thesis.

## **Part 4**

**A search for highly ionising, short tracks at  
the CMS detector**



## 4.1 Motivation

R-parity conserving supersymmetric models are able to offer solutions to many unexplained phenomena in astrophysics and can solve many of the shortcomings of the Standard Model of particle physics (see Section 2.2). While supersymmetric models, especially the Minimal Supersymmetric Standard Model (MSSM) (Section 2.2.1), have been studied at previous particle colliders including Tevatron and LEP [?, ?], the LHC with its high centre-of-mass energy offers a unique opportunity to investigate SUSY models with high particle masses that were not accessible in previous experiments.

Therefore, a variety of searches were hunting for SUSY during Run I of the LHC from 2010 to 2012. Proton-proton collision data from the CMS and ATLAS experiments were analysed with a strong focus on the search for SUSY in production channels via the strong interaction (e.g. [?, ?, ?]). As a consequence, wide, previously unexplored regions of the MSSM parameter space are already excluded. However, due to the unknown mechanism of supersymmetry breaking, the most general parametrisation of the MSSM introduces over 100 new parameters and thus opens up an incredibly large phenomenological space. Therefore, SUSY models can lead to a plethora of possible signatures at particle colliders, many of which could not - or not fully - be explored.

A very interesting signature occurs when particles live long enough to travel through a part or the whole detector before decaying. This is possible for SUSY models with compressed spectra, in which a particle can be long-lived because of phase-space suppression. In the MSSM, such a mass-degeneracy naturally occurs if the wino mass parameter ( $M_2$ ) is smaller than the bino ( $M_1$ ) and higgsino ( $\mu$ ) mass parameters. In this case, the lightest chargino ( $\tilde{\chi}_1^\pm$ ) and the lightest neutralino ( $\tilde{\chi}_1^0$ ) are both wino-like and their mass gap is fully determined by higher order corrections (see Section 2.3). Therefore, they are almost mass-degenerate and the chargino is long-lived.

Such scenarios can be very interesting from a cosmological perspective as the wino-like lightest supersymmetric particle, the neutralino  $\tilde{\chi}_1^0$ , can serve as a plausible Dark Matter candidate [?, ?]. While it is not possible to explain the full relic density with thermally produced wino-like neutralinos for  $m_{\tilde{\chi}_1^0} \lesssim 3$  TeV [?], neutralinos can still be the dominant part if they are non-thermally produced via the decay of an almost decoupled particle [?, ?].

Additionally, explorations of the MSSM parameter space reveal that many models that are consistent with current observations and theoretical constraints and that offer a neutralino as Dark Matter candidate include a metastable chargino with a mass gap of the

order of  $\sim 160$  MeV with respect to the neutralino [?].

SUSY scenarios with nearly mass-degenerate particles have two distinctive phenomenological properties that require a very different search strategy compared to general SUSY searches. First, because of the mass-degeneracy, the remaining decay product (e.g. a pion) is very soft in  $p_T$ , making it hard to detect. Since the other decay product, the neutralino, is only weakly interacting, it is very difficult to identify charginos via their decay products. Second, as the chargino is long-lived, it may traverse several detector layers before decaying. Thus, there is the possibility of reconstructing the chargino itself, e.g. as a reconstructed track in the tracker system.

Despite the exotic signatures of supersymmetric models with nearly mass-degenerate  $\tilde{\chi}_1^\pm$  and  $\tilde{\chi}_1^0$ , current CMS searches are already sensitive to a very broad range of lifetimes. The exclusion power of existing SUSY searches can be assessed by interpreting their results in terms of the fraction of excluded parameter points in the phenomenological pMSSM (see Section 2.2.1.3 for an introduction to the pMSSM). The results of such a study which has been performed in [?] are shown in Figure 4.1. It can be seen that general SUSY searches (blue area) are sensitive to shorter chargino lifetimes ( $c\tau \lesssim 1$  cm).<sup>1</sup> Two existing searches, the search for long-lived charged particles [?] and the search for disappearing tracks [?] focus on long and intermediate chargino lifetimes, respectively. These two searches (purple and red areas) are sensitive to chargino lifetimes of  $c\tau \gtrsim 10$  cm. Taken together, the existing searches exclude a large fraction of pMSSM points at different chargino lifetimes. However, there is a gap between the general SUSY searches and the search for disappearing tracks that is not accessible by any of the existing searches.

The here presented analysis aims at targeting this gap by optimising the search strategy for charginos with intermediate lifetimes of  $1 \text{ cm} \lesssim c\tau \lesssim 30 \text{ cm}$ . The targeted optimisation strategy is a combination of the strategies used in the search for long-lived charged particles [?] and the search for disappearing tracks [?]. While in [?], the high ionisation losses of hypothetical new massive particles is exploited, it does not take into account whether its reconstructed track is disappearing. In [?], the disappearance of the track is utilised but it does not incorporate the large ionisation losses into the search. Additionally, neither of the search does take into account the possibly very short tracks of early decaying charginos.

Thus, the here presented search is the first analysis at CMS combining the two signature properties that are highly distinctive for charginos with intermediate lifetimes: first, the

---

<sup>1</sup>Since the pMSSM interpretation relied on the use of fast simulation techniques which are not capable of simulating charginos with lifetimes  $c\tau > 1$  cm, the general SUSY searches were never interpreted in the context of SUSY models with longer chargino lifetimes.

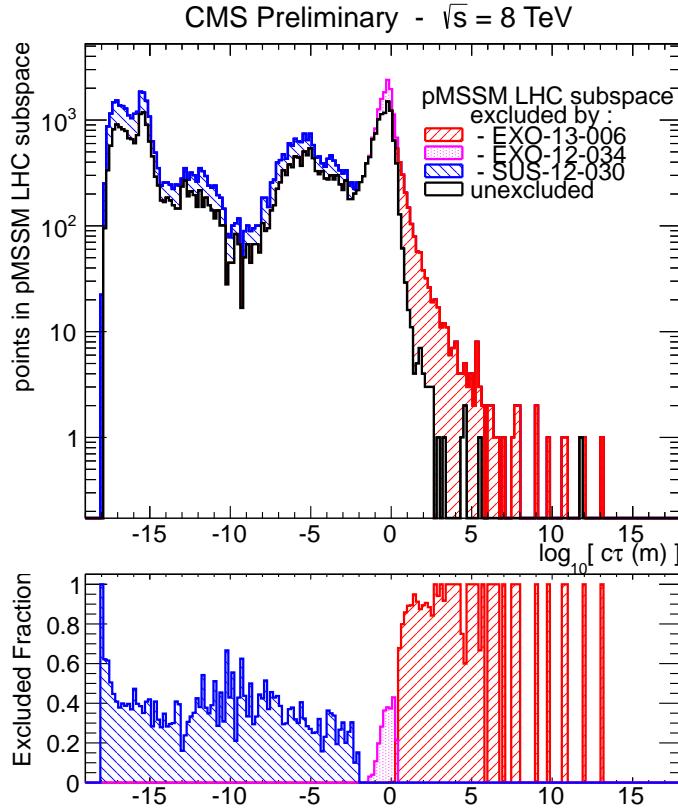


Figure 4.1: The number of excluded pMSSM points at 95% C.L. (upper part) and the fraction of excluded pMSSM points (bottom part) vs. the chargino lifetime for different CMS searches. Red area: the search for long-lived charged particles [?], Purple area: the search for disappearing tracks [?], Blue area: a collection of various general SUSY searches [?]. The black line indicates the unexcluded pMSSM parameter points. The sampling of the parameter space points was done according to a prior probability density function which takes pre-LHC data and results from indirect SUSY searches into account (see [?] for further details). Taken from: [?].

characteristically high ionisation losses of heavy charginos; second, short reconstructed tracks due to chargino decays early in the detector.

The associated challenges and the general search strategy of this analysis will be presented in the next section.

## 4.2 General search strategy

At the LHC, there are several possible chargino production channels. Chargino pairs can be produced through a photon or a  $Z$ -boson exchange. The chargino then decays via a virtual  $W$ -boson to the lightest neutralino and a fermion pair (e.g. a pion). This process is illustrated in the Feynman diagram in Fig. 4.2. Other possible chargino pair production channels include the exchange of a supersymmetric Higgs boson or a t-channel squark exchange (Fig. 4.3).

Apart from pair production, charginos can be produced via the chargino-neutralino channel. On tree-level, there exist two production mechanisms: the s-channel  $W$ -boson exchange and the t-channel squark exchange (Fig. 4.4).

Alternatively, charginos can be produced via strong production modes, i.e. in cascade decays of new heavy particles, such as gluinos or squarks. In the here presented search, the focus is, however, put on the electroweak production channels: chargino pair and chargino-neutralino production.

When searching for supersymmetric models with long-lived  $\tilde{\chi}_1^\pm$ , the strategy is of course highly dependent on the actual lifetime of the chargino. For long lifetimes, the chargino can

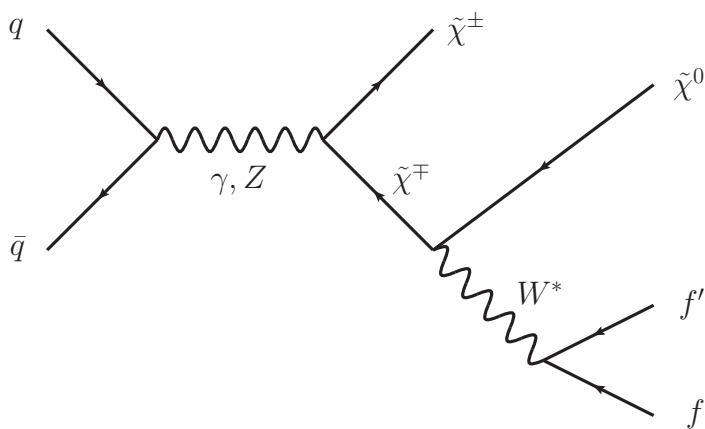


Figure 4.2: Feynman diagram of chargino pair production via gamma or  $Z$ -boson exchange and the subsequent decay via a virtual  $W$ -boson.

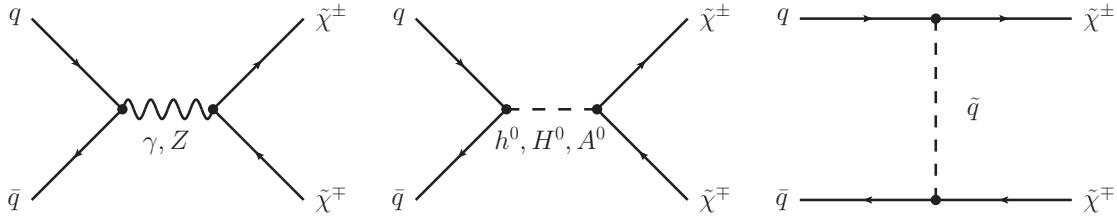


Figure 4.3: Main tree-level diagrams for chargino pair production.

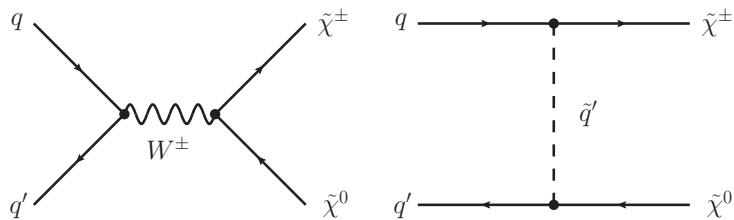


Figure 4.4: Main tree-level diagrams for chargino neutralino production.

reach the muon chambers and can be reconstructed as a muon even despite a longer time-of-flight [?]. For lower lifetimes, the chargino can already decay inside the detector (e.g. the tracker), and hence can not be reconstructed as a muon but leads to an isolated, potentially disappearing track in the tracker. The detector signatures of these two scenarios are visualised in Fig. 4.5, where simulated chargino-chargino events are shown in a cross-sectional view of the CMS detector. In the left picture of Fig. 4.5, both charginos are reconstructed as muons, which can be seen by the energy deposition in the muon chambers. In the middle and right pictures both charginos have a lower lifetime of  $c\tau = 0.5$  m and thus are only visible as tracks in the tracker, where both trajectories end inside the silicon strip tracker (by coincidence the tracks are equally long). Since this analysis targets a search for Supersymmetry with charginos of lifetimes between  $c\tau \approx 1$  cm – 10 cm, the charginos decay rather early in the detector, possibly even in the inner layers of the tracker. Thus, the signature of chargino events consists of isolated, short tracks and the signatures of the decay products, i.e. of a neutralino and a fermion pair.

In case of R-parity conservation, one of the chargino decay products, the neutralino, is stable and weakly interacting, thus traversing the detector without leaving any further signature.

The signature of the other decay product, the fermion pair, can in principle be used to select chargino events. However, for mass-degenerate charginos, it can be very hard or even impossible to detect these fermions as will be explained in detail in the next paragraph.

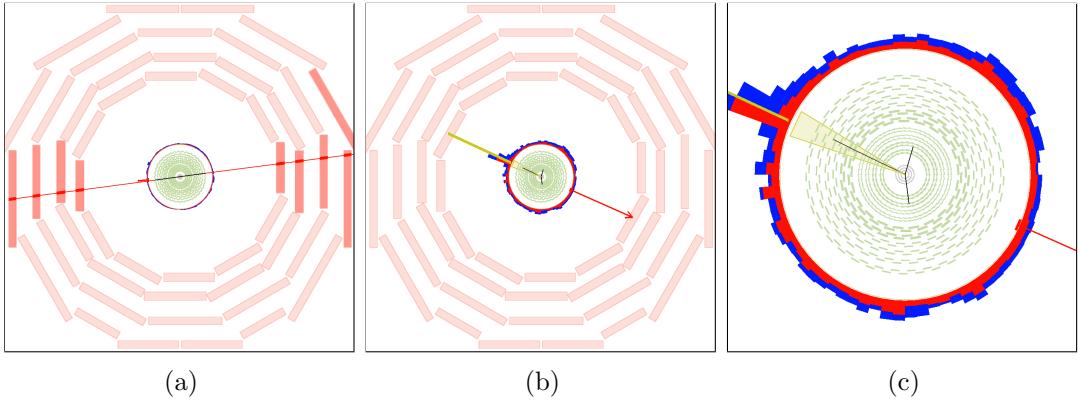


Figure 4.5: Visualisation of possible signatures of a chargino pair produced with a lifetime of  $c\tau = 10$  m (a) and a lifetime of  $c\tau = 0.5$  m (b and c). The muon chambers are the outer layers of the detector and are depicted as red boxes. The black lines represent the reconstructed chargino tracks. The right picture (c) is a zoom of the middle picture (b). Here, only the cross-section of the tracker (green wavy lines for the strip and grey lines for the pixel) is displayed. The red (blue) towers correspond to the energy deposition in the ECAL (HCAL). The ISR jet in the middle and right picture is indicated as a yellow line.

First of all, the fermionic decay product (e.g. a pion) can usually not be reconstructed because it does not originate from the primary vertex. Secondly, it is very low in momentum because of the mass-degeneracy between  $\tilde{\chi}_1^\pm$  and  $\tilde{\chi}_1^0$ . The typical momentum of a pion originating from a chargino to neutralino decay in the  $\tilde{\chi}_1^\pm$  rest frame is of the order

$$p_\pi \sim \sqrt{\left(m_{\tilde{\chi}_1^\pm} - m_{\tilde{\chi}_1^0}\right)^2 - m_\pi^2}. \quad (4.1)$$

For a mass gap between  $\tilde{\chi}_1^\pm$  and  $\tilde{\chi}_1^0$  of 150 MeV, the  $p_T$  distribution of the resulting pion peaks at  $\sim 100$  MeV and ends at  $p_T \sim 400$  MeV in the laboratory frame (Fig. 4.6).

If the transverse momentum of a particle is very low, the particle trajectory is much more bended compared to a particle with higher  $p_T$  (see Fig. 4.7 for illustration). Due to this bending, the track reconstruction efficiency of particles with a transverse momentum below 1 GeV decreases rapidly, reaching around 40% for isolated pions with a  $p_T$  of 100 MeV [?]. Furthermore, for pions that are not produced in the primary vertex, this reconstruction efficiency will be even smaller. It is therefore impossible to rely on a reconstruction of the fermionic chargino decay products in this analysis.

In summary, since an early decaying chargino is not reconstructed as a PF particle, the event signature of a chargino-pair or a chargino-neutralino event consists only of one (or two) - potentially - disappearing track. Such a signature is very difficult to detect,

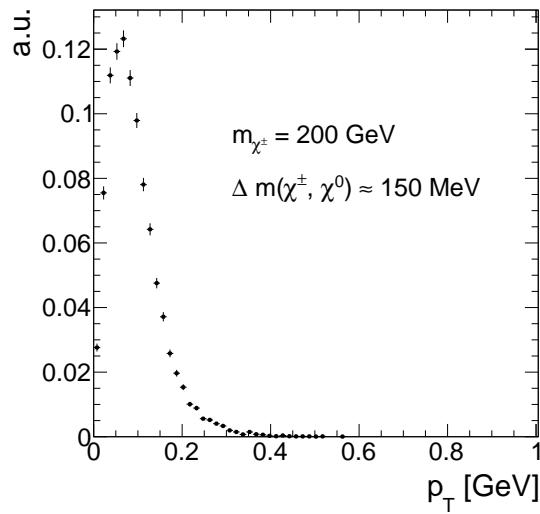


Figure 4.6: Transverse momentum distribution of pions coming from charginos decay into a neutralino with a mass gap of 150 MeV.

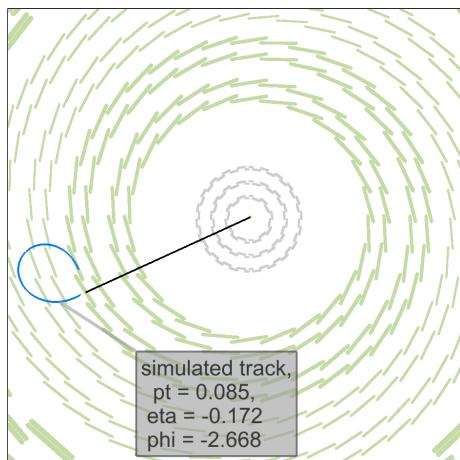


Figure 4.7: Cross-sectional view of the silicon strip tracker (green lines) and silicon pixel tracker (grey lines). A simulated chargino track (black line) decays to a pion (bended blue line) with a  $p_T$  of  $\sim 85$  MeV and a neutralino (not visible).

especially since CMS doesn't offer a dedicated track trigger so that triggering on the chargino track is impossible.

In order to search for such signatures, one therefore needs to trigger on other, less obvious properties of chargino events. This analysis takes advantage of higher order contributions to the Feynman diagrams shown in Figs. 4.3 and 4.4, resulting in initial state radiation (ISR). If the initial quarks radiate a high  $p_T$  gluon, the resulting jet can be detected and can offer a possibility to search for events with - apart from the ISR jet - nothing more than isolated tracks. Furthermore, the non-detection of the chargino's decay products plus a high  $p_T$  ISR jet leads to missing transverse energy (MET) in the event. Exploiting these two circumstances, it is possible to detect chargino-pair or chargino-neutralino events with the help of Jet+MET triggers.

Since Jet+MET triggers are not very specific for chargino events, it is important to identify further track properties that can be used to select chargino candidates. One distinctive property of charginos compared to SM particles is their high mass. Therefore, charginos can be identified by selecting high  $p_T$  tracks. Furthermore, the energy loss per path length ( $dE/dx$ ) depends quadratically on the particle mass for low velocities ( $0.2 < \beta\gamma < 0.9$ ):

$$\langle \frac{dE}{dx} \rangle = K \frac{m^2}{p^2} + C \quad (4.2)$$

Therefore,  $dE/dx$  constitutes a very nice discriminating variable for massive particles like charginos against SM particles. The selection of chargino events in this analysis thus relies on the selection of isolated high  $p_T$  tracks with high  $dE/dx$  values.

If the chargino decays before it has crossed the full pixel and strip detector, the associated track is disappearing. For low lifetimes, the tracks can be very short and can have only a few hits in the detector. In order to reconstruct a particle trajectory, a minimum of three hits are required since defining a helical path requires five parameters (see [?]). A specific challenge for this analysis is hence the combination of searching for short tracks and utilising the measurement of the energy deposition of the chargino. For very short tracks, eventually only passing the first couple of layers of the whole tracker system, the pixel tracker information becomes very important. Therefore, an accurate energy loss measurement in the pixel system is of great importance to this analysis. However, no other CMS analysis has used the energy information of the pixel tracker so far. This analysis thus requires a thorough study of the quality of the pixel energy calibration and, potentially, a recalibration in case the pixel energy calibration is not sufficient.

### 4.2.1 Comparison to earlier searches

As already mentioned before, there are two analyses at CMS at  $\sqrt{s} = 8$  TeV with  $20\text{ fb}^{-1}$  data that search for intermediate lifetime charginos: the search for long-lived charged

particles [?] and the search for disappearing tracks [?]. The here presented analysis aims at achieving an increase in sensitivity towards shorter lifetimes compared to the earlier analyses in a twofold way. First, the selection is optimised for the inclusion of very short tracks. Second, the inclusion of the variable  $dE/dx$  is used to increase the search sensitivity compared to [?].

In [?], a minimum number of eight hits were required for every track, whereas in [?] a minimum of seven hits are required. This can be very inefficient for shorter lifetimes, where most of the charginos already decay shortly after the pixel tracker. In Fig. 4.8 (left), the normalised distribution of the number of measurements ( $N_{\text{hits}}$ ) of chargino tracks is shown. FIXME: Since a minimum number of three hits are required to reconstruct a particle trajectory, there are no tracks with a lower number of measurements in the tracker system. It can be seen, that  $N_{\text{hits}}$  peaks at the minimal possible value needed for track reconstruction of  $N_{\text{hits}} = 3$  for lower lifetimes. For a lifetime of  $c\tau = 100\text{ cm}$ , a second peak at  $\sim 17$  hits appears corresponding to the number of measurements when crossing all pixel barrel (3) and strip inner and outer barrel (6 from stereo and 8 from normal) layers. However, a notable fraction of  $\sim 40\%$  of chargino tracks still has a number of measurements of  $N_{\text{hits}} < 8$ .

It should also be mentioned that the track reconstruction efficiency is sufficient for short

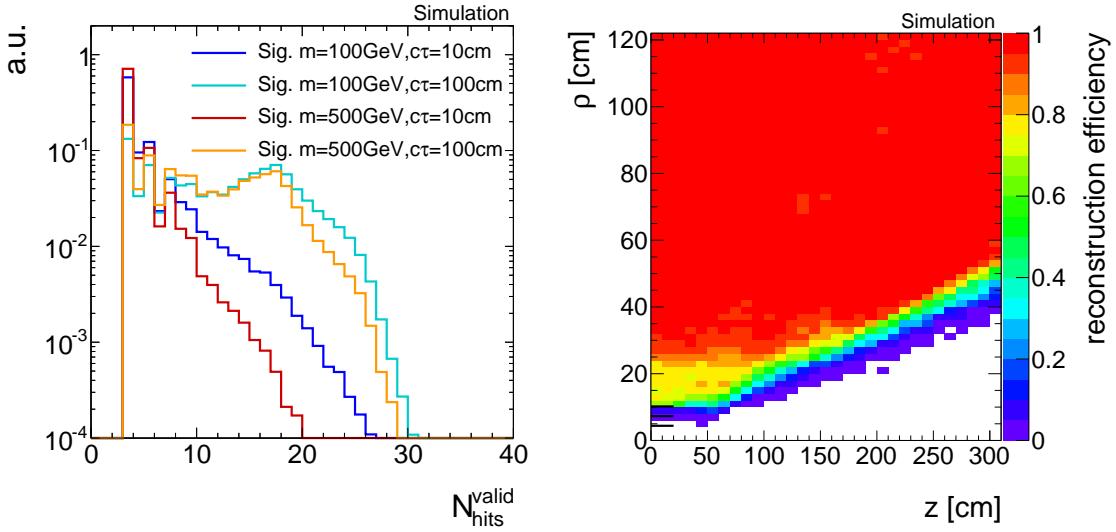


Figure 4.8: FIXME: Left: Number of measurements in the tracker system  $N_{\text{hits}}$  for four different signal lifetimes. Right: Probability to reconstruct a track ( $z$ -axis) in dependency of the chargino's decay point in  $z$ -direction and radial direction  $\rho = \sqrt{x^2 + y^2}$  ( $x$ - and  $y$ -axis). More information on the generation of the simulated signal samples can be found in Section 4.4.2.

chargino tracks so that loosening the  $N_{\text{hits}}$  requirement is expected to really improve the signal acceptance. The track reconstruction efficiency for different chargino decay points is depicted in Fig. 4.8 (right). For very short tracks ( $N_{\text{hits}} = 3$ ) the efficiency is still around 20%.

Additionally, the search for disappearing tracks [?] which targets models with charginos decaying inside the tracker did not make use of the high energy deposition of heavy particles. Although this variable was indeed used in the search for long-lived charged particles [?], this search was not optimised for intermediate lifetimes (e.g. no explicit muon veto on the selected tracks was required). Thus, it shows less sensitivity compared to the disappearing track search in the lifetime region between  $10 \text{ cm} \lesssim c\tau \lesssim 100 \text{ cm}$  (see Fig. 4.1).

To conclude, the general search strategy of the here presented analysis is to unite the strategies of [?] and [?] and to lower the strong selection on the number of hits in these analyses in order to get an optimised selection for lifetimes around  $1 \text{ cm} \lesssim c\tau \lesssim 30 \text{ cm}$ .

## 4.3 Improved $dE/dx$ measurement for short tracks

As already pointed out in the previous chapter, the inclusion of the pixel energy measurements can increase the sensitivity when searching for short and highly ionising tracks. While the energy measurements in the silicon strip detector have already been calibrated as part of the search for long-lived charged particles [?], no complete calibration has been done for the pixel silicon tracker so far. To increase the discrimination power of  $dE/dx$  for short tracks, such a calibration procedure has therefore been performed within this PhD thesis.

The CMS tracker system provides a measurement of the particle’s energy loss for each hit in the tracker. *FIXME:* This is done by the measurement of the charge induced by the electron hole pairs produced by the ionisation of the silicon. A detailed introduction to the CMS tracker system and the energy measurement can be found in Section 3.2.1.

The procedure to combine single energy measurements for each tracker hit into one track  $dE/dx$  estimator that can be used for analysis purposes will be explained in the following Section 4.3.1. The pixel energy calibration is then described in Section 4.3.2. How to discriminate SM particles and beyond SM particles with the help of a  $dE/dx$  measurement is discussed in Section 4.3.3, followed by an exploration of how the inclusion of the pixel energy measurements in the  $dE/dx$  estimates leads to a better discrimination between Standard Model particles and long-lived charginos (Section 4.3.4).

### 4.3.1 Estimation of the ionisation loss of charged particles

Energy losses for moderately relativistic charged particles travelling through matter are mostly caused by ionisation effects. The mean energy loss per path length can be described with the Bethe formula [?]:

$$\langle \frac{dE}{dx} \rangle = K z^2 \frac{Z}{A} \frac{1}{\beta^2} \left[ \frac{1}{2} \ln \frac{2m_e c^2 \beta^2 \gamma^2 T_{\max}}{I^2} - \beta^2 - \frac{\delta(\beta\gamma)}{2} \right]. \quad (4.3)$$

It is a function of the atomic number ( $Z$ ), the atomic mass ( $A$ ) of the absorber, and the mean excitation energy ( $I$ ) which is 173 eV for silicon [?].  $T_{\max}$  represents the maximum energy transfer in a single collision. The relevant particle's properties are the velocity ( $\beta$ ), the Lorentz factor ( $\gamma$ ) and the charge ( $z$ ) of the incident particle. The density correction  $\delta(\beta\gamma)$  reduces the mean energy loss at high energies because of polarisation effects of the material. The factor  $K$  is constant and is 0.307 in units of MeV mol<sup>-1</sup>cm<sup>2</sup>. The Bethe formula is valid if the main energy loss originates from ionisation effects, i. e. in a region between  $0.1 \lesssim \beta\gamma \lesssim 1000$ .

Even if widely used, the mean energy loss is a quantity which is “ill-defined experimentally and is not useful for describing energy loss by single particles” [?]. The problem is caused by the underlying probability distribution of one single  $dE/dx$  measurement (this will be named  $\Delta E/\Delta x$  throughout the following sections), which can be parametrised by a Landau distribution [?]

$$p(x) = \frac{1}{\pi} \int_0^\infty e^{-t \log t - xt} \sin(\pi t) dt. \quad (4.4)$$

The Landau distribution has no free parameters. Its most probable value is around 0.222. However, it is possible to introduce artificially a different most probable value and a width (at half maximum) with  $x \rightarrow \frac{x - \text{MPV}}{\sigma} - 0.222$ . The Landau distribution is a highly asymmetric distribution with a long tail towards large  $x$  values (see Fig. 4.9). Theoretically it extends to infinite energies, however in nature the maximal deposited energy is of course limited by the particle's full energy.

**FIXME:** Because of its strong asymmetry, measurements of the mean energy loss per path length  $\langle dE/dx \rangle$ , consisting of only a few single measurements  $\Delta E/\Delta x$ , are easily fluctuating towards high values. This makes the use of the mean energy loss described by the Bethe formula for the discrimination of new heavy particles problematic, because fluctuations to high values reduce the discrimination power against massive particles which release in general higher amounts of energy in matter.

A much better observable is the most probable value (MPV) of the Landau distribution. The MPV is much more stable compared to the mean and is not subject to high  $dE/dx$  fluctuations. The most probable energy loss of a charged particle,  $\Delta_p$ , can be described

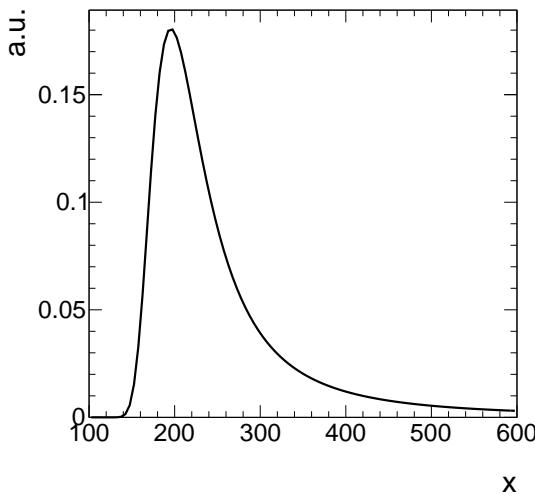


Figure 4.9: Illustration of the shape of a Landau distribution. Parameters were chosen as  $\mu = 200$  and  $\sigma = 20$ .

by the Landau-Vavilov-Bichsel equation [?]:

$$\Delta_p = \xi \left[ \ln \frac{2m_e c^2 \beta^2 \gamma^2}{I} + \ln \frac{\xi}{I} + j - \beta^2 - \delta(\beta\gamma) \right], \quad (4.5)$$

with  $\xi = (K/Z)(Z/A)(x/\beta^2)$ . The thickness of the absorber  $x$  appears explicitly in the Landau-Vavilov-Bichsel equation making the most probable energy loss per path length  $\Delta_p/dx$  logarithmically dependent on  $x$ . A comparison between the Bethe mean energy loss  $\langle dE/dx \rangle$  and the most probable energy loss  $\Delta_p/dx$  for muons is shown in Fig. 4.10. The restricted energy loss refers to the energy loss when the maximal transfer energy  $T_{\max}$  is restricted to a certain value  $T_{\text{cut}}$ .

Particles such as muons are minimally ionising in silicon for  $\beta\gamma \sim 3 - 4$ . For higher momenta the deposited energies increase again reaching a plateau at around  $\beta\gamma \sim 100$ . However, new heavy charged particles would mainly be non-relativistic because of their high mass and would therefore deposit much higher energies in the detector. This makes  $dE/dx$  a very well discriminating variable. Thus, the energy loss per path length can be used to discriminate between SM particles and new heavy charged particles due to the different velocity distributions.

As said before, the most probable energy loss is much more stable compared to the Bethe mean energy loss. Still, combining only a few measurements of  $\Delta E/\Delta x$  can also lead for  $\Delta_p/dx$  to large fluctuations towards high  $dE/dx$  values. In order to estimate experimentally the most probable  $dE/dx$  value from only a few energy measurements,

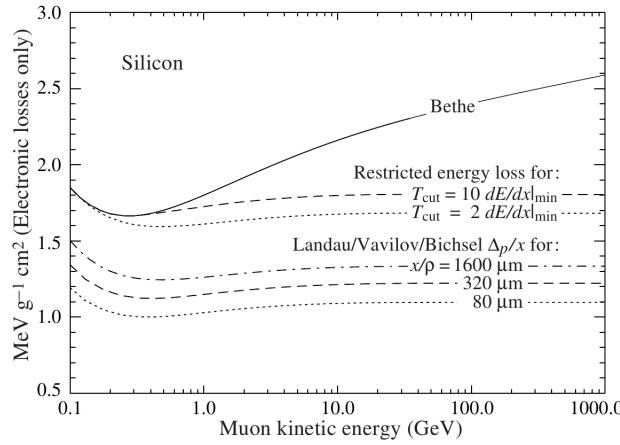


Figure 4.10: FIXME: Comparison between the Bethe mean energy loss, restricted energy loss and the most probable energy loss described by the Landau-Vavilov-Bichsel function for muons for different values of absorber thickness of silicon. The absorber thickness is normalised to the absorber density  $\rho$ , thus  $[x] = g \cdot cm^{-2}$ . Taken from [?].

several “estimators” can be used that suppress a potential bias towards the high end without introducing a bias towards lower values [?]. One of the estimators for determining a track’s  $dE/dx$  is the harmonic-2 estimator

$$I_{h2} = \left( \frac{1}{N} \sum_{i=1}^N (\Delta E_i / \Delta x_i)^{-2} \right)^{-1/2}, \quad (4.6)$$

where  $\Delta E_i / \Delta x_i$  corresponds to the  $\Delta E$  and  $\Delta x$  measurement in the  $i$ th hit of the track. This estimator is known to be robust and is not easily biased by large fluctuations in  $\Delta E / \Delta x$  because of the suppression by the power of minus two [?].

A further estimator of  $dE/dx$  used for the discrimination of highly ionising particles will be introduced in Section 4.3.3.

### 4.3.2 Energy calibration of the silicon pixel tracker

During Run I in 2012, the pixel silicon detector was continuously subjected to an energy calibration, a so-called gain calibration. Every pixel was calibrated to the same response, so that the whole pixel tracker should have been well inter-calibrated [?]. Unfortunately, due to various reasons, such as the imperfect constancy of the reference signal, or radiation and temperature induced changes, the energy calibration could not ensure a fully calibrated pixel tracker.

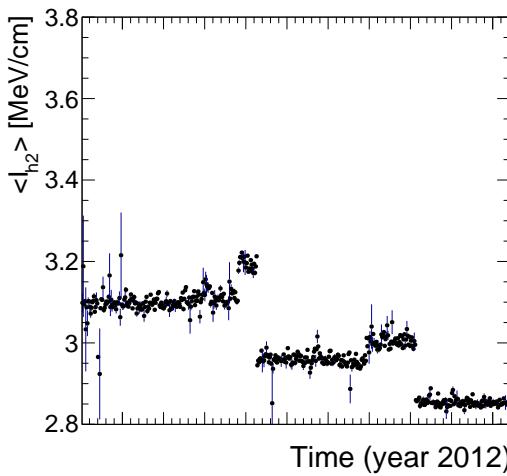


Figure 4.11: Mean of all track's  $dE/dx$  (harmonic-2 estimator) over the full year 2012. Only pixel hits are taken into account. Every data point corresponds to one run.

This imperfection of the gain calibration can be seen in Fig. 4.11, where the mean of the harmonic-2 estimator for all tracks  $\langle I_{h2} \rangle$  over the full data-taking period in 2012 is shown. FIXME Four different steps can be spotted. The first and the third steps correspond to changes in the settings of the tracker due to irradiation. The second and fourth step are induced by associated adjustments in the online gain calibration. Unfortunately, although the gain calibration was adjusted (even with some delay), it was not able to ensure a constant energy response of the pixel tracker over time. The variations of the  $dE/dx$  measurement over time of around 15% are too large to use  $dE/dx$  without a further calibration.

The following sections explain the method of the gain calibration of the pixel silicon tracker which is performed for this analysis. It is splitted into two sections. The first section is dedicated to the gain inter-calibration of the pixel tracker which ensures a homogeneous energy response of all tracker modules. In the second section, the absolute gain calibration is discussed. FIXME: This calibration step is needed to ensure that the measurement of the energy release of a particle can actually be translated to a correct physical value.

### Inter-calibration of gain

The main goal of the gain calibration is to get a uniform response in the ionisation energy loss  $dE/dx$  over the full data taking period in 2012. To also ensure a uniform response over all modules within one time period, an additional inter-calibration on module level is

carried out. The inter-calibration can in principle be done on various levels: the highest granularity would be a calibration on pixel level, followed by a calibration on read-out-chip (ROC) level and then on module level. Lower granularities in descending order are rings (modules with same z-position) and finally layers (3 layers in the barrel and 4 disks in the endcap). Within this analysis the calibration is performed on module level since the response variations between all ROCs on one module of  $\sim 5\%$  are acceptable (see Appendix ??). The applied method for the gain calibration of the pixel tracker closely follows the method in [?].

The gain calibration of the pixel silicon tracker is carried out with the help of minimally ionising particles (MIPs). MIPs in this context are not defined as particles at the minimum of the Bethe formula, but more generally as particles located at or near the plateau of the  $dE/dx$  distribution vs. momentum (see Fig. 4.10). This approach ensures that all particles deposit similar amounts of energy so that the variation due to different momenta is minimised. MIPs are selected by a momentum selection of  $p > 2 \text{ GeV}$ . *FIXME:* Additionally, only tracks with at least eight hits and a quality of the trajectory fit of  $\chi^2/\text{ndof} < 3$  are used to ensure a high-quality track reconstruction. A sample containing around 50 million “minimum bias” events is used for calibration. The “minimum bias” sample was specifically recorded for tracker calibration purposes. *FIXME:* Its distinctive property is that only a minimum of detector activity is required to record an event, e.g. pixel hits in the tracking system.

For every module in the pixel tracker (there are 1440 modules in total), a distribution of the energy loss per path length  $\Delta E/\Delta x$  is built. The path length  $\Delta x$  is calculated with

$$\Delta x = \frac{d_{\text{module}_i}}{\cos(\phi_{\text{track}})}, \quad (4.7)$$

where  $d_{\text{module}_i}$  is the thickness of module  $i$  and  $\phi_{\text{track}}$  is the relative angle of the particle trajectory to the normal axis of the module. The measurement of  $\Delta E/\Delta x$  is done in ADC counts per mm. ADC counts are a measure for the deposited charge after digitisation. Figure 4.12 shows an example distribution for one module. To extract the MPV for every module a fit to the core distribution is performed. The fit is not only done with a Landau but a Landau convoluted with a Gaussian function to be closer to the experimentally observed energy spectrum. This also improves the fit performance and the stability of the fit. With the measured MPV extracted from the fit, an inter-calibration factor is calculated for every module

$$c_{\text{inter}} = \frac{\text{MPV}_{\text{target}} [\text{ADC/mm}]}{\text{MPV} [\text{ADC/mm}]} = \frac{300 \cdot 265 \text{ ADC/mm}}{\text{MPV} [\text{ADC/mm}]} \quad (4.8)$$

The factor  $300 \cdot 265 \text{ ADC/mm}$  is in principal an arbitrary number since the final response is adjusted by the absolute gain calibration described in the next section. However, it is

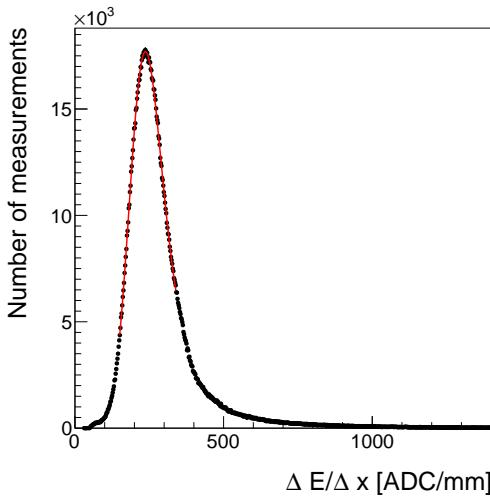


Figure 4.12: An example of the  $\Delta E/\Delta x$  distribution measured in ADC count per mm for one module of the CMS pixel tracker within one time interval. A Landau convoluted with a Gaussian is fitted to the core of the distribution.

chosen such that the measured calibration factors are close to one. The calibration factor can then be used to scale every single measurement in a module to a calibrated  $\Delta E/\Delta x$  measurement

$$\left(\frac{\Delta E}{\Delta x}\right)_{\text{calibrated}} = c_{\text{inter}} \cdot \left(\frac{\Delta E}{\Delta x}\right)_{\text{uncalibrated}} \quad (4.9)$$

The determination of the calibration factor is done for every of the five time intervals, shown in Fig. 4.11 independently, in order to get rid of the time dependency. The outcome of the application of the calibration factors to the single energy measurements in the pixel tracker can be seen in Fig. 4.13. The variation over time is indeed eliminated, resulting in a maximal time variation of less than  $\sim 1\%$ .

Additionally, the same procedure is carried out for a corresponding simulated data sample to ensure the inter-calibration of the pixel modules on all simulated samples.

### Absolute calibration of gain

As a final step, the targeted MPV being  $\text{MPV}_{\text{target}} = 300 \cdot 265 \text{ ADC/mm}$  needs to be translated to a meaningful physical quantity given in physical units (e.g. MeV/cm). That means, that the charge measurement in ADC counts needs to be converted to the real energy release from a particle. The relation between  $\Delta E$  in ADC counts and the energy loss in eV is given by

$$\Delta E [\text{eV}] = c_{\text{inter}} \cdot \Delta E [\text{ADC}] \cdot \frac{N_e}{\text{ADC}} \cdot 3.61 \text{ eV}, \quad (4.10)$$

where  $N_e/\text{ADC}$  is the number of electrons which correspond to one calibrated ADC count and 3.61 eV is the mean energy needed to create one electron-hole pair in silicon at  $-10^\circ\text{C}$ . Such an absolute gain calibration can be done with the help of several methods (all explained in [?]). The absolute calibration of the silicon pixel tracker can rely on the existing absolute calibration of the silicon strip detector. In [?], the absolute gain calibration was done with the help of the most probable energy release per path length of muons, theoretically described by the Landau-Vavilov-Bichsel formula in Eq. (4.5). To calibrate the pixel tracker to the correct energy loss per path length it is therefore sufficient to determine one calibration factor to relate the average  $dE/dx$  of all tracks in the pixel tracker as shown in Fig. 4.13 to the average measured  $dE/dx$  in the strip tracker, shown in Fig. 4.14 by

$$c_{\text{absolute}} = \frac{\langle dE/dx_{\text{strip}} \rangle}{\langle dE/dx_{\text{pixel}} \rangle} = \frac{3.303 \text{ MeV/cm}}{3.509 \text{ MeV/cm}} = 0.941. \quad (4.11)$$

This factor is then applied on top of  $c_{\text{inter}}$  for all pixel modules.

Finally, an absolute calibration factor needs to be determined for the simulated samples, where the simulated pixel tracker is calibrated to the average  $dE/dx$  of the silicon strip measured in data.

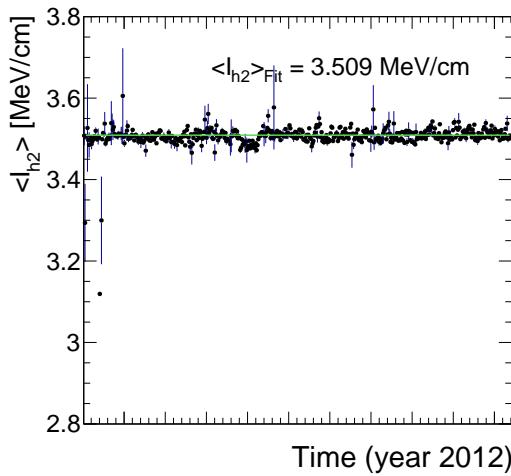


Figure 4.13: Mean of all track's  $dE/dx$  (harmonic-2 estimator) over the full year 2012 after applying the calibration factors, resulting in an average  $dE/dx$  of 3.51 MeV/cm. Only pixel hits are taken into account. Every data point corresponds to one run.

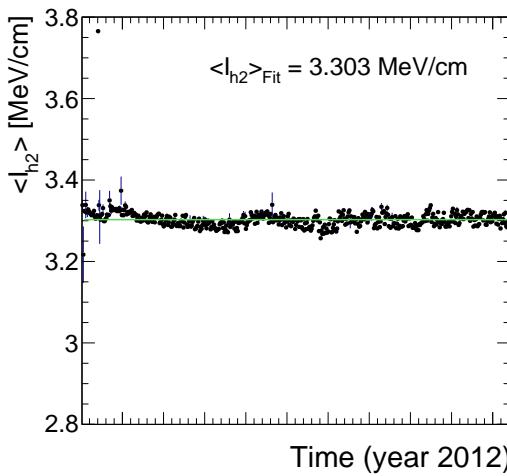


Figure 4.14: Mean of all track’s  $dE/dx$  (harmonic-2 estimator) measured in the silicon strip detector over the full year 2012. The average most probable  $dE/dx$  is  $I_{h2} = 3.303 \text{ MeV/cm}$ . Every data point corresponds to one run.

### 4.3.3 Discrimination of highly-ionising particles

As mentioned before, it is difficult to find a robust estimator for the most probable energy loss of a particle, if only a few measurements of  $\Delta E/\Delta x$  along the particle’s trajectory are available. The harmonic-2 estimator  $I_{h2}$  was already introduced in Section 4.3.1 in Eq. (4.6). It is known to be a robust estimator not easily affected by large fluctuations in  $\Delta E/\Delta x$ . However, it was shown in [?] that a better discrimination between SM particles and possible new heavy particles can be achieved when using likelihood techniques, i. e. determining the probability that the set of all  $\Delta E/\Delta x$  belonging to one track is actually compatible with the hypothetical probability distribution of a MIP.

That a measured sample has been drawn from a specific distribution can be tested with the co-called Smirnov-Cramér-von Mises test [?, ?]. It is deduced from the integral of the squared difference of a measured distribution to a hypothesis distribution, and leads to a test statistics of [?]

$$I_s = \frac{3}{N} \cdot \left( \frac{1}{12N} + \sum_{i=1}^N \left[ P_i - \frac{2i-1}{2N} \right]^2 \right), \quad (4.12)$$

where  $N$  is the total number of energy measurements and  $P_i$  is the cumulative probability that a MIP would release a  $\Delta E/\Delta x$  equal or smaller than the measured  $\Delta E/\Delta x$  with all  $P_i$  arranged in increasing order.

However, this test statistics is not sensitive to the sign of the difference between the measured and the theoretical distribution. It can therefore not distinguish between in-

compatibilities due to variations towards higher or lower energy deposits compared to the hypothesis distribution. Thus it is not optimal for the discrimination between MIPs and heavy new particles by  $dE/dx$ . A so-called Asymmetric Smirnov-Cramér-von Mises discriminator was developed in [?] which is only sensitive to incompatibilities to the MIP hypothesis towards higher energy depositions

$$I_{\text{as}} = \frac{3}{N} \cdot \left( \frac{1}{12N} + \sum_{i=1}^N \left[ P_i \cdot \left( P_i - \frac{2i-1}{2N} \right)^2 \right] \right). \quad (4.13)$$

A value of  $I_{\text{as}}$  close to zero indicates good compatibility with the MIP hypothesis, whereas a value close to one indicates bad compatibility because of unexpectedly high energy losses.

The underlying probability distribution of the energy release for a given path length in the pixel tracker, which is used for the determination of  $P_i$ , is extracted from the same “minimum bias” sample used for the pixel energy calibration. In total 28 different templates each for a different given path length are created. In Fig. 4.15 the probability distribution template for the pixel tracker in data and simulation is shown. The corresponding templates for the energy release in the silicon strip detector were already built by [?].

A comparison between the energy release by MIPs ( $I_{\text{as}}$ ) in data and simulation for high-quality tracks with  $p > 5 \text{ GeV}$  and  $|\eta| < 2.1$  can be found in Fig. 4.16. The energy loss per path length,  $dE/dx$ , shows good agreement in data and simulation for  $I_{\text{as}} < 0.1$ . For larger values,  $I_{\text{as}}$  shows a larger decrease in simulation than in measured data. For this

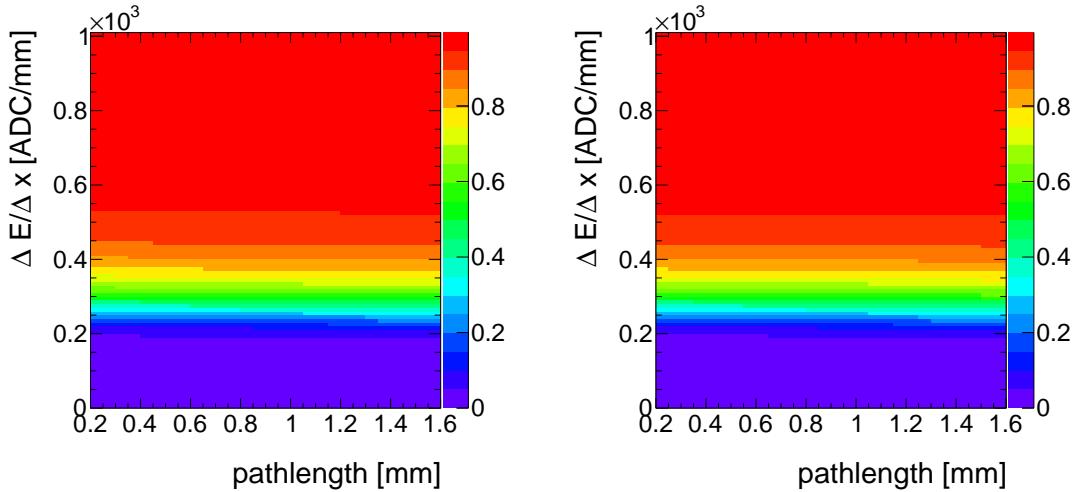


Figure 4.15: Cumulative probability for a MIP to release a  $\Delta E/\Delta x$  (y-axis) vs. the path length (x-axis) in data (left) and simulation (right) for the pixel tracker based on the “minimum bias” sample.

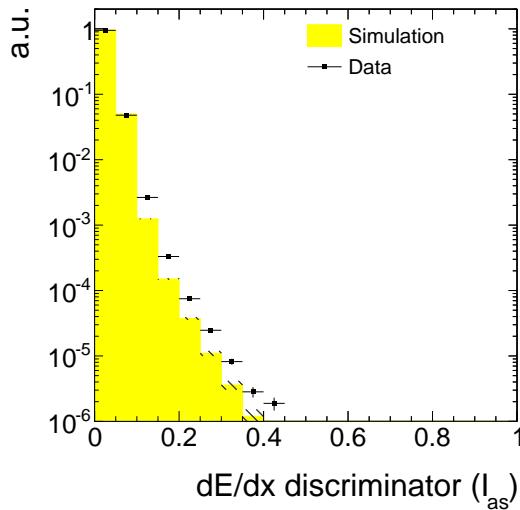


Figure 4.16: Normalised  $I_{\text{as}}$  distribution for MIPs from the minimum bias sample in data and simulation for high-quality (high purity as defined in [?], a minimum number of eight hits and no missing inner and middle hits) tracks with  $p > 5 \text{ GeV}$  and  $|\eta| < 2.1$ .

reason a data-based approach for analyses exploiting  $dE/dx$  information is needed.

#### 4.3.4 Discrimination improvements

The goal of including the pixel energy information is to increase the discrimination power of  $I_{\text{as}}$  between background and signal tracks, especially for shorter lifetimes. In Fig. 4.17 (left), a comparison of the shapes of the energy release by MIPs and by signal tracks in simulation is shown (details about the simulated samples can be found in the next section Section 4.4.2). It can be seen, that the  $I_{\text{as}}$  distributions of the signal models show a larger tail towards  $I_{\text{as}} = 1$ , whereas the  $I_{\text{as}}$  of the background is rapidly falling.

In the right part of Fig. 4.17, a comparison of the  $I_{\text{as}}$  distributions of four different signal models is shown. Charginos with longer lifetimes have a more pronounced tail toward  $I_{\text{as}} = 1$ . This can be understood with the help of Eq. (4.5), where the influence of the velocity ( $\beta$ ) on the ionisation loss can be seen. The velocity distribution of the charginos is mostly affected by the mass of the chargino. However, also for charginos with same mass, the velocity is higher on average for shorter lifetimes. This is caused by the fact, that for shorter lifetimes (e.g.  $c\tau = 10 \text{ cm}$ ), already a sizable fraction of the charginos decay before reaching the tracker system. The probability of reaching the detector increases for higher velocities because of the boost, which can be clearly seen at

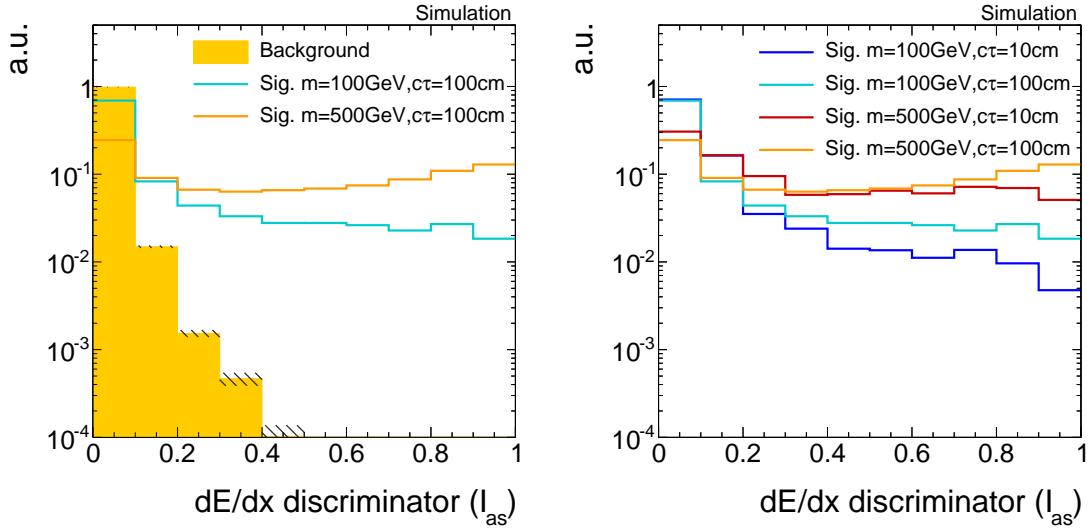


Figure 4.17: Normalised  $I_{as}$  distribution for simulated background and signal tracks (left) and for four different signal models (right) for high-purity tracks (as defined in [?]) with  $p_T > 10$  GeV and  $|\eta| < 2.1$ . For the illustration of the background track spectrum simulated  $t\bar{t}$ +jets events are used (more information about this sample is given in Chapter 4.4).

the survival probability

$$P(t) = e^{-\frac{t}{\gamma\tau}}. \quad (4.14)$$

This means that the track reconstruction/selection lead to a biased average  $\beta$  for shorter lifetimes which in turn lead to lower values of  $I_{as}$ .

The  $I_{as}$  distribution is not only influenced by the velocity of a particle but also by the number of hits of a track. The number of measurements in the tracker system defines the influence of single fluctuations in  $\Delta E/\Delta x$  on the  $I_{as}$  discriminator, because of the long right tail of the Landau distribution. A low number of hits, therefore, leads to higher  $I_{as}$  values. This effect is also visible in Fig. 4.17 (right). The small surplus for lower lifetimes between 0.1 and 0.2 is caused by the smaller number of measurements for earlier decaying charginos.

Finally, the impact of the additional  $\Delta E/\Delta x$  information from the pixel tracker on the selection efficiency of signal and background tracks is quantified. Figure 4.18 shows the signal selection efficiency against the background selection efficiency for different selection cuts in  $I_{as}$ , once including the pixel information and once without it. The background selection efficiency is estimated with simulated  $W+jets$  events but was additionally checked on simulated  $t\bar{t}+jets$  and QCD-multiparticle events (further information about the simulated samples can be found in the next Chapter 4.4). No significant difference between these processes in the background selection efficiency was observed.

The signal selection efficiency and the background suppression depend on the mass and the lifetime of the charginos. The improvement of the discriminating power is much more pronounced for higher chargino masses.

It can be seen that the inclusion of the pixel information increases the background suppression for a given signal efficiency throughout the investigated signal models. This background suppression improvement is most pronounced for very tight cuts on  $I_{\text{as}}$  up to a factor of 20 and even more and still considerable for looser selections with signal efficiencies of around 40% (factor of 10).

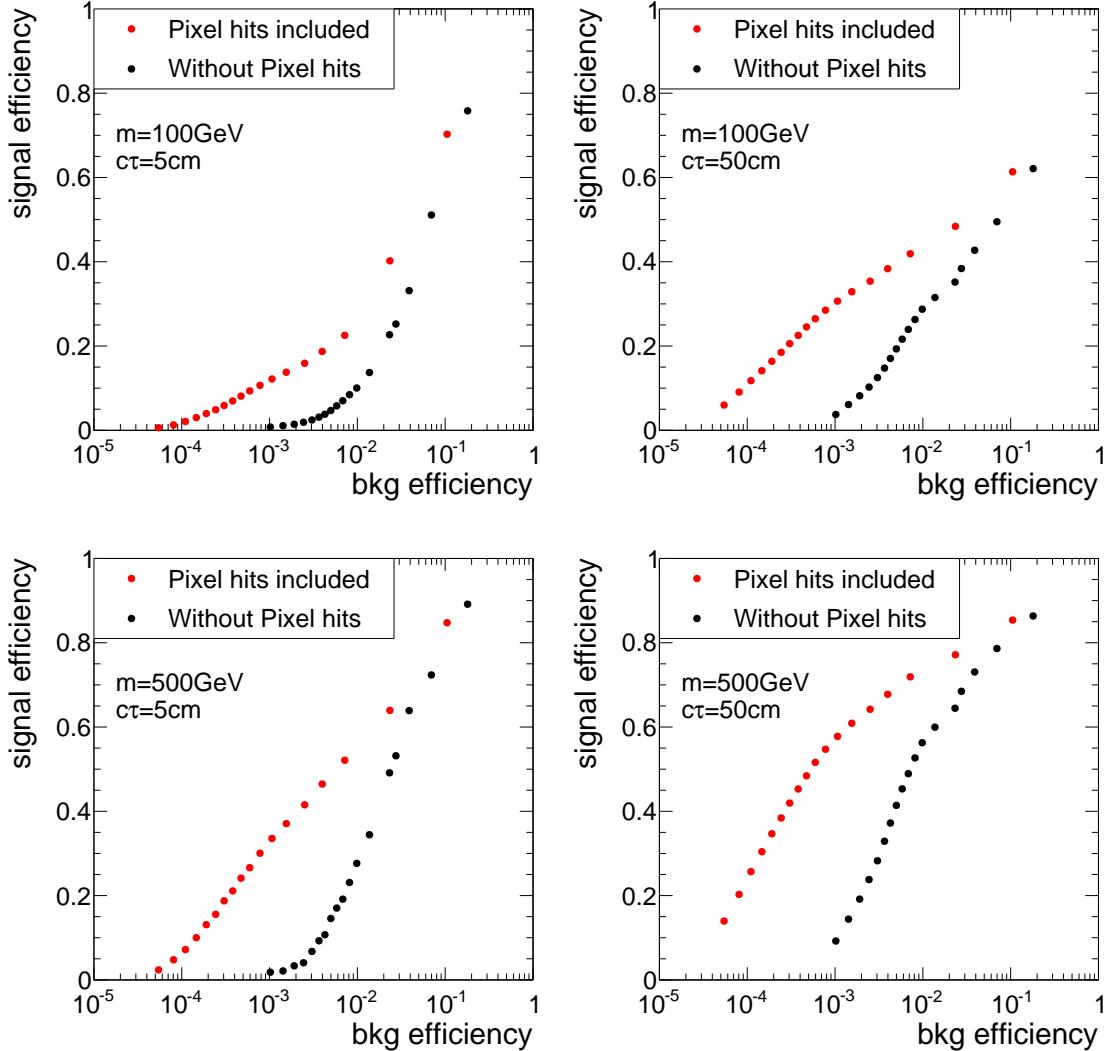


Figure 4.18: Signal selection efficiency vs. background selection efficiency with (red) and without (black) pixel information. Each point correspond to one selection cut in  $I_{as}$ . The figure is based on a simulated  $W + \text{jets}$  sample and a simulated signal sample with chargino-chargino production, both subject to a selection of high-quality tracks (without a selection on  $N_{\text{hits}}$ ) with  $p_T > 10\text{ GeV}$ .

## 4.4 Simulated samples

In order to design the search and to study background and signal characteristics, this analysis relies on simulated SM and SUSY datasets. An introduction to the techniques and tools required for the simulation of SM and beyond SM processes can be found in Section 3.4.

The following two sections present an overview of the SM (Section 4.4.1) and SUSY samples (Section 4.4.2) used in this search. All samples are reweighted to match the measured distribution of pileup interactions per event in data. Additionally, event weights are applied to ensure the same ISR spectrum in simulation as in data.

### 4.4.1 Standard Model background samples

To investigate the sources of background, various simulated SM samples are used. Since this analysis aims at making use of  $dE/dx$ , a data format is required that includes tracker hit information, the so-called RECO format [?]. Unfortunately, not all SM processes are available in this format making it impossible to compare the total number of events in simulation and real data. This, however, does not constitute a serious problem since this analysis will finally use data-based background estimation methods. The simulated SM datasets can still be used to compare the shapes of important distributions in simulation and data.<sup>1</sup> Still, the most important SM background sample including  $W + \text{jets}$  events is available. Due to the intrinsic missing energy in  $W + \text{jets}$  events it constitutes the major background to the presented search (see Section 4.6 for further details on the backgrounds).

In Table 4.1 all available SM samples used in this analysis are listed. The matrix-elements of the  $W + \text{jets}$ ,  $t\bar{t} + \text{jets}$  and  $Z \rightarrow \ell\bar{\ell} + \text{jets}$  samples are generated using MADGRAPH 5 [?]. For the QCD sample PYTHIA6 [?] is used for generation. All samples are then passed to PYTHIA 6 to simulate the hadronisation and the showering. The interactions between the particles and the detector material is simulated using GEANT4 [?, ?].

Due to the size of the samples (between 5 and 70 TB per sample) a reduction is required in order to limit the storage space requirements. This is achieved by selecting only events

---

<sup>1</sup>For example, the simulated  $Z \rightarrow \nu\bar{\nu} + \text{jets}$  sample that can contribute to the background of this search via fake tracks is not available in RECO format. However, as the shape of important observables of fake tracks is independent of the underlying process, this background can be studied with a simulated  $W + \text{jets}$  sample.

which contain at least one jet with a minimum transverse momentum of  $p_T > 60 \text{ GeV}$ .

In addition, further simulated samples not containing  $dE/dx$  information are used (so-called AOD samples). Because of their much smaller size, these samples are available in full size. They are needed to study the background inclusively in the variable  $dE/dx$ .

Table 4.1: Available Standard Model background samples containing  $\Delta E/\Delta x$  information that are used for background estimation studies.

Process	Generator	Cross section [ pb ]	$\mathcal{O}_{\text{cross section calculation}}$
$W + \text{jets}$	MADGRAPH 5	36703.2	NNLO [?]
$t\bar{t} + \text{jets}$	MADGRAPH 5	245.8	NNLO [?]
$Z \rightarrow \ell\bar{\ell} + \text{jets}$ ( $\ell = e, \mu, \tau$ )	MADGRAPH 5	3531.9	NNLO [?]
QCD ( $50 \text{ GeV} < \hat{p}_T < 1400 \text{ GeV}$ )	PYTHIA 6	9374794.2	LO

#### 4.4.2 Signal samples

For the investigation of a possible SUSY signal, events containing either chargino pair production  $q\bar{q} \rightarrow \tilde{\chi}_1^\pm \tilde{\chi}_1^\mp$  or chargino neutralino production  $q\bar{q} \rightarrow \tilde{\chi}_1^\pm \tilde{\chi}_1^0$  are simulated within this thesis. The simulation of the samples is done as described in Section 4.4.1 for the  $W + \text{jets}$  sample. However, a special treatment for long-lived particles is required for this analysis. In order to get a correct detector simulation of the energy loss of long-lived particles that decay after the beam pipe, the decay of the chargino cannot be simulated within MADGRAPH or PYTHIA but needs to be simulated within GEANT4. The decay mode of the chargino is also specified within GEANT4 to a neutralino plus pion decay,  $\tilde{\chi}_1^\pm \rightarrow \tilde{\chi}_1^0 \pi^\pm$ .

To reduce the required computing sources, the simulation is only done for a few lifetimes ( $c\tau = 1 \text{ cm}, 5 \text{ cm}, 10 \text{ cm}, 50 \text{ cm}, 100 \text{ cm}, 1000 \text{ cm}$  and  $10000 \text{ cm}$ ). The lifetime is hereby not controlled by changing the mass gap between the chargino and the neutralino but is independently specified within GEANT4. In order to scan in a high resolution over the lifetime space, other lifetimes are generated using lifetime reweighting. The weight for each event depends on the individual proper lifetime of the chargino and is given by

$$w = \prod_{i=1}^n \frac{\tau^{\text{gen}}}{\tau^{\text{target}}} \cdot \exp \left[ t_i \cdot \left( \frac{1}{\tau^{\text{gen}}} - \frac{1}{\tau^{\text{target}}} \right) \right], \quad (4.15)$$

where  $n$  is the number of charginos in the event,  $\tau^{\text{gen}}$  is the generated mean lifetime in the particle's rest frame and  $t_i$  is the individual proper lifetime of the chargino. The targeted mean lifetime is given by  $\tau^{\text{target}}$ . A derivation of this formula can be found in Appendix ???. Using this reweighting procedure a good coverage of the lifetime space can be achieved with lifetimes of  $1 \text{ cm} \leq c\tau \leq 10^4 \text{ cm}$ . Figure 4.19 shows the exponential distribution of the individual proper lifetime of the charginos after the reweighting of a simulated sample with  $c\tau^{\text{gen}} = 50 \text{ cm}$  to a lifetime of  $c\tau^{\text{target}} = 10 \text{ cm}$ . It can be seen that the reweighting procedure does indeed reproduce the targeted lifetime of 10 cm.

All samples are generated for different masses of the chargino, but always almost mass-degenerate to the lightest neutralino. The mass gap between chargino and neutralino is set to 150 MeV and, as said before, is hereby disentangled to the chargino lifetime for the simulation of the signal samples. However, since this analysis does not make use of the decay products of the chargino and the mass gap is for all simulated lifetimes of a similar small size, the choice of the mass gap does not affect the signal prediction. Six different masses from 100 GeV to 600 GeV are simulated. This leads to a total number of 42 signal samples. In Table 4.2, the cross sections at  $\sqrt{s} = 8 \text{ TeV}$  for  $\tilde{\chi}_1^\pm \tilde{\chi}_1^\mp$  and  $\tilde{\chi}_1^\pm \tilde{\chi}_1^0$  production for wino-like charginos and neutralinos are listed [?, ?]. The cross section does not depend on the lifetime of the chargino.

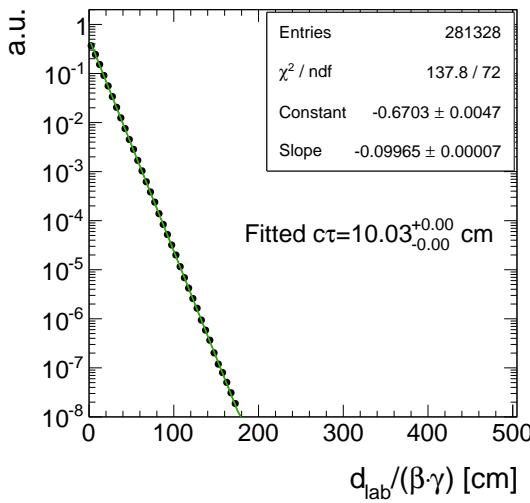


Figure 4.19: Normalised distribution of the proper individual lifetime  $d_{\text{lab}}/(\beta\gamma)$  of all charginos contained in a signal sample with a generated lifetime of  $c\tau^{\text{gen}} = 50 \text{ cm}$  reweighted to a lifetime of  $c\tau^{\text{target}} = 10 \text{ cm}$ . Fitting an exponential curve  $a \cdot \exp\left[\frac{1}{c\tau}ct_i\right]$  yields  $c\tau = \text{slope}^{-1} = 10 \text{ cm}$ .

Table 4.2: Simulated signal mass points with corresponding cross sections at NLO-NLL (NLO: next-to-leading order, NLL: next-to-leading logarithmic) accuracy for wino-like charginos [?, ?].

$m_{\tilde{\chi}_1^\pm}$ [GeV]	$\sigma_{\tilde{\chi}_1^\pm \tilde{\chi}_1^\mp}$ [pb]	$\sigma_{\tilde{\chi}_1^0 \tilde{\chi}_1^\mp}$ [pb]
100	5.8234	11.5132
200	0.37924	0.77661
300	0.06751	0.14176
400	0.01751	0.03758
500	0.00553	0.01205
600	0.00196	0.00431

## 4.5 Event selection

### 4.5.1 Datasets and triggers

The analysis is performed on  $pp$ -collision data recorded in the year 2012 by the CMS experiment at a centre-of-mass energy of  $\sqrt{s} = 8$  TeV. In total an integrated luminosity of  $19.7\text{ fb}^{-1}$  was recorded in 2012.

As outlined in Section 4.2, the detection of chargino tracks is a challenging task already on trigger level. Direct triggering of events containing chargino-like tracks is not possible because in 2012 there was no dedicated track trigger available. Furthermore, there is no intrinsic missing transverse energy in the event if the chargino is not reconstructed as a PF particle, e.g. when it decays inside the tracker. Therefore, this analysis uses initial state radiation for the detection of chargino events. If ISR occurs, it is possible to trigger on a high- $p_T$  jet ( $p_T^{1\text{st jet}}$ ) and missing transverse momentum ( $\cancel{E}_T$ ).

For this purpose, several triggers are utilised in this analysis. An event is selected, if at least one of the three triggers in Table 4.3 fired. The HLTMonocentralPFJet80\_PFMETnoMu95\_NHEF0p95 and HLTMonocentralPFJet80\_PFMETnoMu105\_NHEF0p95 triggers both rely on the L1 ETM40 trigger which requires the missing energy to be larger than 40 GeV. On HLT level, they further require at least one particle-flow jet within

Table 4.3:  $\cancel{E}_T$  and  $\cancel{E}_T +$  jet triggers used in this analysis together with the corresponding recorded integrated luminosity during the time when they were in place.

Trigger	Integrated luminosity [fb $^{-1}$ ]
HLTMonoCentralPFJet80_PFMETnoMu95_NHEF0p95	5.3
HLTMonoCentralPFJet80_PFMETnoMu105_NHEF0p95	14.4
HLT_MET120_HBHENoiseCleaned	19.7

the pseudorapidity range of  $|\eta| < 2.6$  with  $p_T > 80\text{ GeV}$  and the PF missing transverse momentum (not taking into account the  $p_T$  of muons) to be larger than 95 GeV or 105 GeV, respectively. Finally, no more than 95% of the jet energy must be carried by neutral hadrons. The HLTMonoCentralPFJet80\_PFMETnoMu95\_NHEF0p95 trigger was active during Run A and Run B in 2012 data taking, whereas HLTMonoCentralPFJet80\_PFMETnoMu105\_NHEF0p95 was in place during Run C and Run D in 2012.

The HLT\_MET120.HBHENoiseCleaned trigger is based on the L1 trigger ETM36. On HLT level, the trigger requires that the missing energy measured from calorimeter energy deposits is larger than 120 GeV. The HBHENoise-filter reduces background from electronic noise in the HCAL. This trigger was active during the full data taking period in 2012.

The events that were selected by the described triggers are available in the MET datasets listed in Table 4.4. Again, because of the size of the datasets ( $\sim 150\text{ TB}$  in total), a reduction of the size is achieved by selecting only events where one of the used triggers

Table 4.4: MET data samples used in the search with the contained integrated luminosity.

Dataset	Integrated luminosity [fb $^{-1}$ ]
/MET/Run2012A-22Jan2013-v1/RECO	0.876
/MET/Run2012B-22Jan2013-v1/RECO	4.412
/MET/Run2012C-22Jan2013-v1/RECO	7.055
/METParked/Run2012D-22Jan2013-v1/RECO	7.354

fires and that contains at least one jet with a minimum  $p_T$  of 50 GeV. This reduction in size of almost 80% is necessary to make this analysis technically feasible.

## 4.5.2 Selection of signal candidate events

In order to suppress events originating from Standard Model processes such as QCD-multijet events,  $W + \text{jets}$ , etc., a selection favouring signal-like tracks is applied. The signal candidate selection closely follows the selection required in the disappearing track search [?, ?]. It relies on event-based and track-based variables as described in the following two sections.

### 4.5.2.1 Event-based selection

First, a selection on the quality of the primary vertex is applied in order to suppress cosmic events and noise from the beam halo. This selection includes requirements on the position of the vertex with respect to the beam axes and the number of degrees of freedom of the vertex which is strongly correlated to the number of tracks originating from the vertex (see [?] for further details):

- ❖ The vertex must have at least four degrees of freedom:  $vtx$  with  $\geq 4$  d.o.f.
- ❖ The position of the vertex along the beam line must be within 24 cm with respect to the nominal interaction point:  $|dz| \leq 24$  cm.
- ❖ The position in the transverse direction must be within 2 cm with respect to the nominal interaction point:  $|d0| \leq 2$  cm.

To maximise the signal acceptance, the trigger related selection cuts are chosen close to the trigger thresholds (see Section 4.5.1). In Fig. 4.20, the distributions of  $\cancel{E}_T$  and the transverse momentum of the leading jet,  $p_T^{1^{\text{st}}\text{jet}}$ , are shown for different signal models. The leading jet has to be centrally produced,  $|\eta_{1^{\text{st}}\text{jet}}| < 2.4$ , and to fulfil the following criteria:

- Charged hadron energy fraction ( $\text{CHF}_{1^{\text{st}}\text{jet}}$ )  $> 0.2$
- Charged electromagnetic energy fraction ( $\text{CEF}_{1^{\text{st}}\text{jet}}$ )  $< 0.5$
- Neutral hadron energy fraction ( $\text{NHF}_{1^{\text{st}}\text{jet}}$ )  $< 0.7$
- Neutral electromagnetic energy fraction ( $\text{NEF}_{1^{\text{st}}\text{jet}}$ )  $< 0.7$ .

These additional jet quality criteria ensure that noise from cosmic and beam halo muons and high- $p_T$  photons and electrons is suppressed [?].

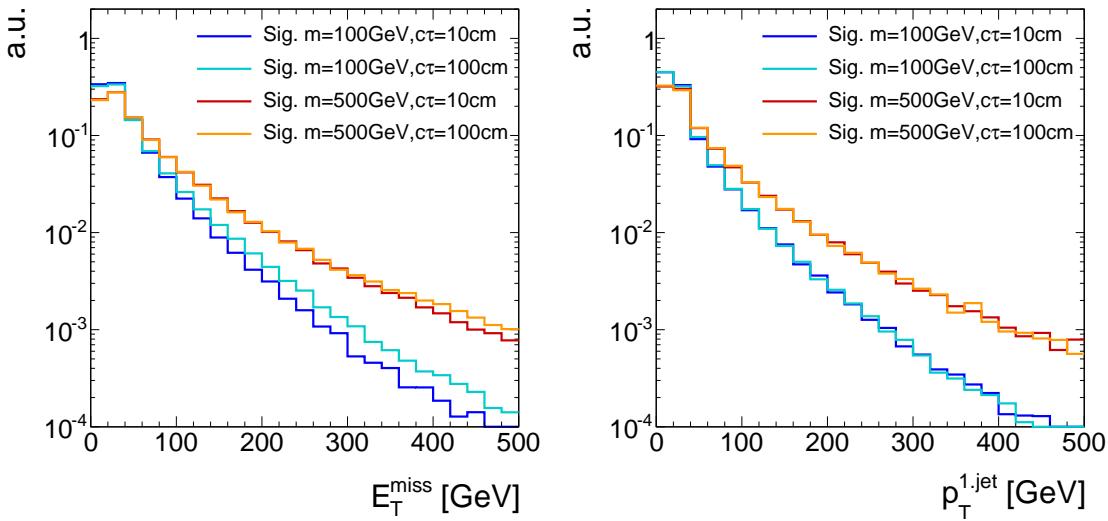


Figure 4.20: Normalised distributions of the missing transverse momentum (left) and the transverse momentum of the leading jet (right) for four different signal models.

The trigger efficiency as a function of  $E_T$  and  $p_T^{1\text{-jet}}$  was determined within [?] with a single-muon reference sample. The trigger paths become fully efficient for  $p_T^{1\text{-jet}} \gtrsim 110\text{ GeV}$  and  $E_T \gtrsim 220\text{ GeV}$  [?]. However, it can be seen in Fig. 4.20 that for a selection of  $E_T > 220\text{ GeV}$  more than 99% of the signal events are rejected.

In order to achieve a reasonable signal acceptance, this search imposes, therefore, a trigger selection closer to the intrinsic trigger thresholds. The trigger selection is as follows:

- ❖ There is at least one jet within  $|\eta| < 2.4$  with transverse momentum larger than  $110\text{ GeV}$  which fulfils the above mentioned jet noise cleaning criteria:  $p_T^{1\text{-jet}} > 110\text{ GeV}$ .
- ❖ The missing transverse momentum must be larger than  $100\text{ GeV}$ :  $E_T > 100\text{ GeV}$

These requirements result in an efficiency of 100% for the trigger requirements on the jet  $p_T$  and an efficiency of  $\sim 5 - 20\%$  for the trigger requirement on  $E_T$ , at the  $E_T$  thresholds [?]. Throughout the following sections, these trigger related requirements will be referred to as “trigger selection”.

Because of the huge cross section, QCD-multijet events are frequently produced at the LHC. Due to jet energy mismeasurements, they can also contribute to data samples recorded with MET triggers. Therefore, special requirements are enforced in order to suppress events emerging from QCD-multijet processes. QCD-multijet events can be characterised by topologies where two jets are almost back-to-back. Additionally, in QCD-multijet events the missing energy is usually aligned with one of the leading jets in the

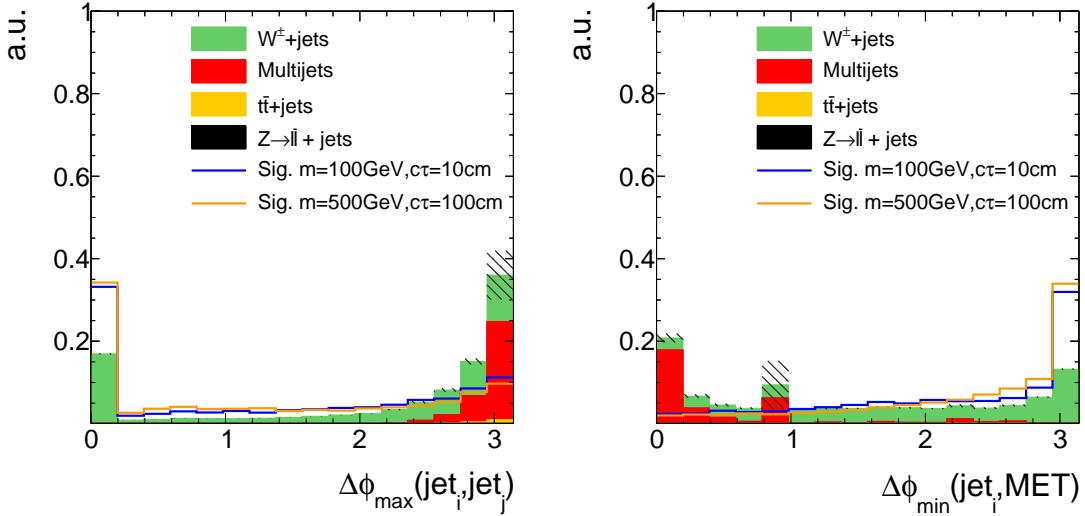


Figure 4.21: Maximum  $\Delta\phi$  between any of two jets (left) and the minimum  $\Delta\phi$  between the  $\cancel{E}_T$  vector and any of the two leading jets (right) normalised to unit area after trigger selection. Only jets with  $p_T > 20$  GeV and  $|\eta| < 4.5$  are considered.

event. Figure 4.21 shows the maximum  $\Delta\phi$  of any of two jets ( $p_T > 20$  GeV,  $|\eta| < 4.5$ ) and the minimum  $\Delta\phi$  between the  $\cancel{E}_T$  vector and any of the two leading jets for the SM background and two different signal datasets.

The following two requirements are sufficient to suppress QCD-multijet events efficiently:

- ❖  $\Delta\phi$  between any of two jets (with  $p_T > 20$  GeV and  $|\eta| < 4.5$ ) in the event must be smaller than 2.5.
- ❖  $\Delta\phi$  between any of the two leading jets (with  $p_T > 20$  GeV and  $|\eta| < 4.5$ ) and the  $\cancel{E}_T$  must be larger than 0.5.

#### 4.5.2.2 Candidate track selection

After the reduction of background processes with event-based variables, a track-based selection is carried out. To get an optimised selection for possible chargino tracks several signal candidate track characteristics are exploited.

First, a selection of high-quality tracks is enforced:

- ❖ The track must be classified as “high purity” as defined in [?].
- ❖ The track is required to have no missing middle or inner hits:  $N_{\text{miss}}^{\text{middle/inner}} = 0$

- ❖ The radial and longitudinal distance of the track to the primary vertex must be small:  $|d0| < 0.02 \text{ cm}$ ,  $|dz| < 0.5 \text{ cm}$ .

In Figs. 4.22 and 4.23, the power of the latter two quality selection cuts is shown.

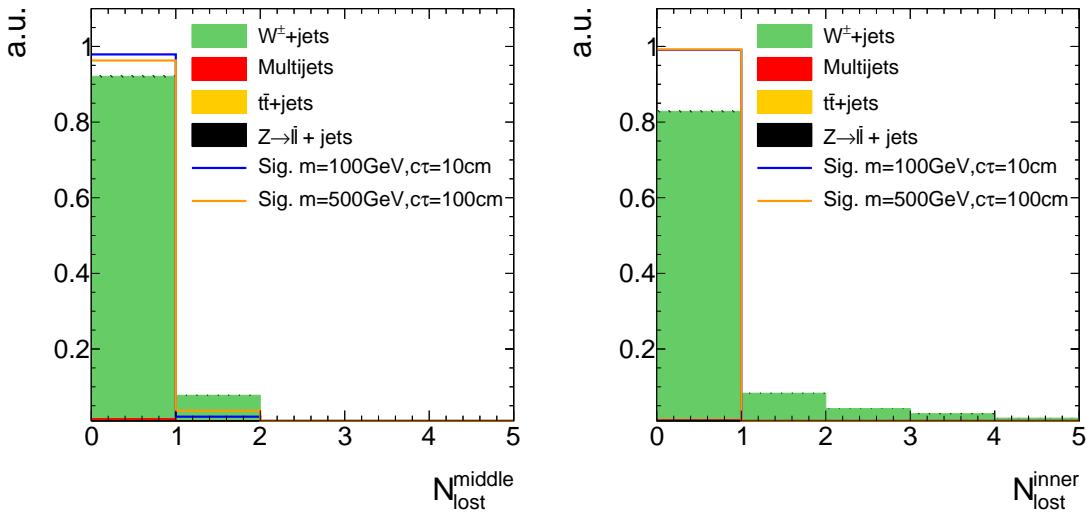


Figure 4.22: Normalised number of missing middle (left) and inner (right) hits of background and signal tracks after trigger selection and QCD suppression cuts.

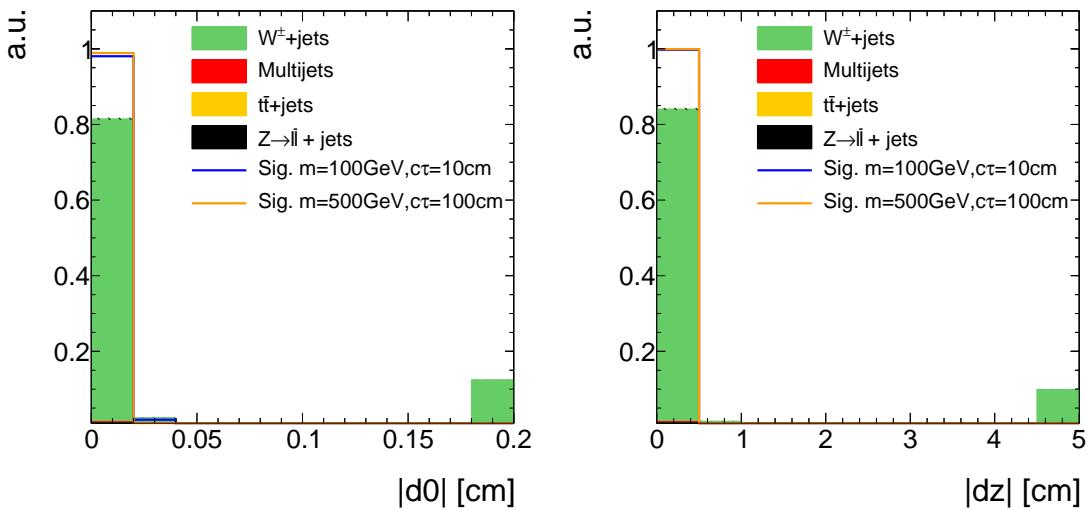


Figure 4.23: Absolute value of the radial (left) and longitudinal (right) distance between the track and the primary vertex after trigger selection and QCD-multijet suppression cuts. All events with a candidate track with a radial (longitudinal) distance larger than 0.2 cm (5 cm) are contained in the last bin.

Furthermore, a first kinematic preselection is applied:

- ❖ Only tracks in the central region are considered :  $|\eta| < 2.1$ .
- ❖ Only tracks with a minimum transverse momentum of 20 GeV are considered:  
 $p_T > 20 \text{ GeV}$ .

In order to suppress background tracks emerging from SM processes, an electron, muon and tau veto is applied. This rejects tracks that are close to a reconstructed electron, muon or tau. Additionally, the candidate track must not be close to a jet ( $p_T > 20 \text{ GeV}$  and  $|\eta| < 4.5$ ):

- ❖ The track must not be within a cone of  $\Delta R < 0.15$  to a reconstructed standalone, tracker or global muon with a transverse momentum larger than 10 GeV (see Section 3.3.2.5 for details on the different muon definitions).
- ❖ The track must not be within a cone of  $\Delta R < 0.15$  to a reconstructed electron with a transverse momentum larger than 10 GeV (see Section 3.3.2.6 for details on the electron reconstruction).
- ❖ The track must not be within a cone of  $\Delta R < 0.15$  to a reconstructed tau with  $p_T > 20 \text{ GeV}$  and  $|\eta| < 2.3$  (see Section 3.3.2.7 for details on the tau reconstruction). Some loose isolation requirements are enforced to protect the tau reconstruction from jet contamination.
- ❖ The track must not be within a cone of  $\Delta R < 0.5$  to a reconstructed jet ( $p_T > 20 \text{ GeV}$  and  $|\eta| < 4.5$ ).

These lepton and jet veto selection requirements are highly suppressing the background emerging from real lepton/jet production like in  $W + \text{jets}$  events. The discrimination power of the lepton and jet vetos is shown in Fig. 4.24 where the minimum  $\Delta R$  between the candidate track and a reconstructed electron, muon, tau or jet is shown.

Unfortunately, the lepton veto selection cuts are less effective in some of the detector directions. For example, the reconstruction of an electron easily fails in the direction of a dead ECAL cell. This reduces the discrimination power of the electron veto. For this reason, tracks that point towards dead or noisy ECAL cells are rejected. A general list of dead and noisy ECAL cells is provided centrally by the CMS collaboration. Further dead cells were identified within a study in [?, ?] resulting in a total number of 1234 dead or noisy ECAL channels (which is about 1.6% of all ECAL crystals). These are illustrated in Fig. 4.25 showing a map of all ECAL channels not considered in the search.

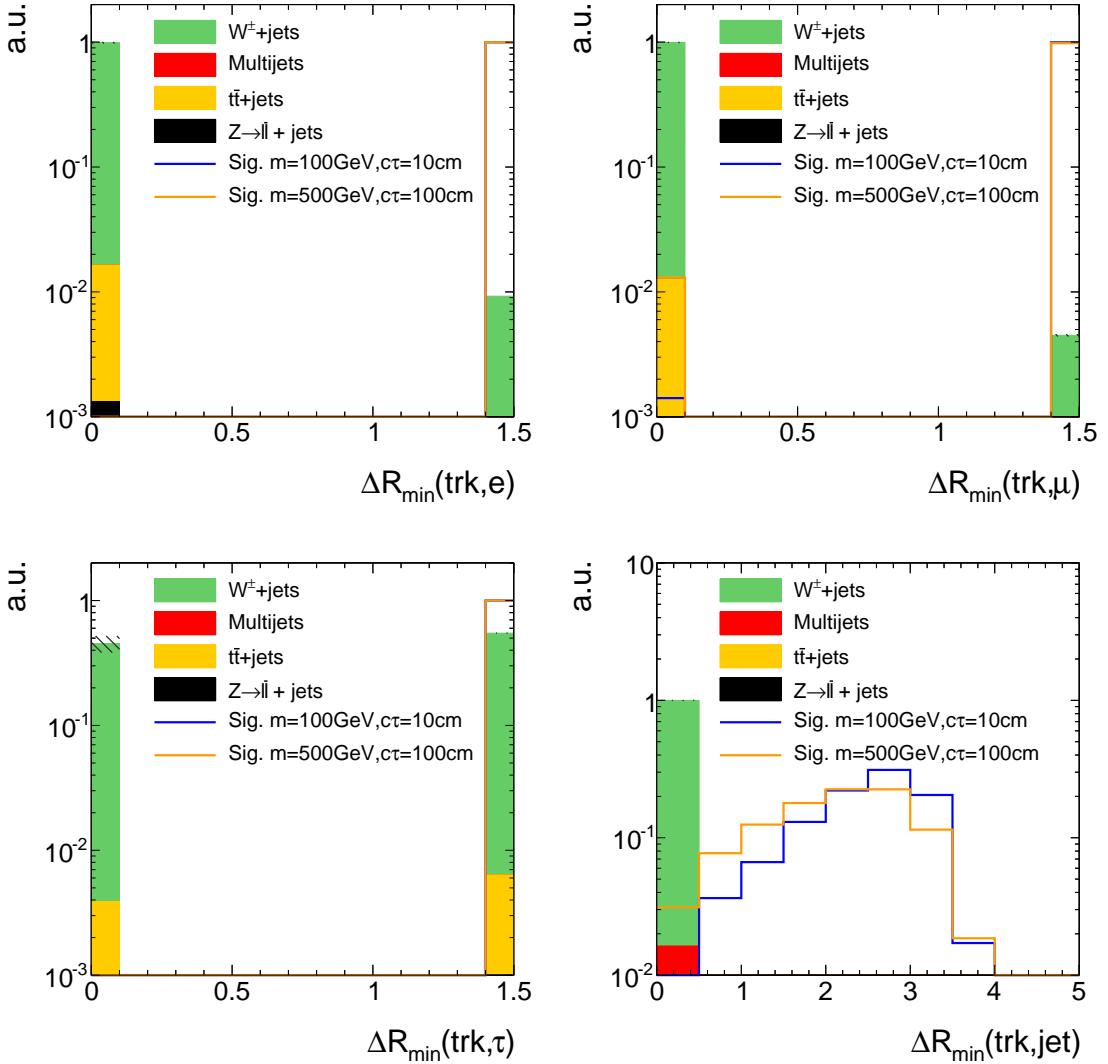


Figure 4.24: The minimum  $\Delta R$  between the candidate track and a reconstructed electron (top left), muon (top right), tau (bottom left) or jet (bottom right) after the event-based selection and the high-quality, kinematic and lepton/jet veto selection of the candidate track selection but without the one shown in the corresponding plot (“N-1 plot”). The last bin contains all events where the candidate track has a  $\Delta R_{\min}$  larger than 1.5 or 5.0 to the next lepton or jet respectively. Events with no respective lepton or jet are also contained in the last bin.

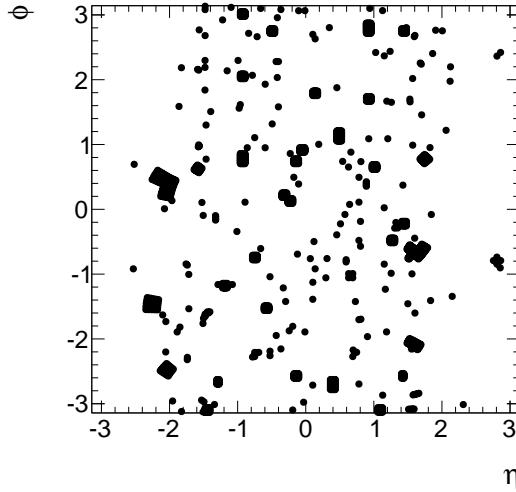


Figure 4.25: Visualisation of dead and noisy ECAL cells in the detector's  $\phi - \eta$  plane according to [?, ?]. The radius of the dots correspond to  $\Delta R = 0.05$ .

Additionally, tracks that point towards intermodule gaps of ECAL cells or to the ECAL barrel endcap gap at  $1.42 < |\eta| < 1.65$  are rejected. A list of the ECAL intermodule gaps, which is supplied centrally by CMS, is given in Table 4.5.

The muon reconstruction is less efficient for muons in detector regions with bad cathode strip chambers (CSC). These bad chambers are also identified centrally by the CMS collaboration and their  $\eta$  and  $\phi$  values are visualised in Fig. 4.26. Thus, also tracks pointing towards these regions within a distance of  $\Delta R < 0.25$  are rejected.

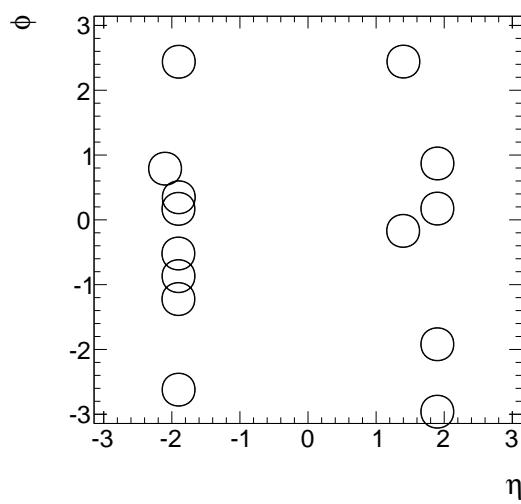
To summarise, tracks pointing towards detector regions, that are not working properly or where the lepton reconstruction efficiencies are reduced, are vetoed as follows:

- ❖ Veto tracks within a cone of  $\Delta R < 0.05$  to a dead or noisy ECAL cell (visualised in Fig. 4.25).
- ❖ Veto tracks that point towards the direction of the ECAL intermodule gap listed in Table 4.5.
- ❖ Veto tracks that point towards a bad CSC within  $\Delta R < 0.25$  (visualised in Fig. 4.26).
- ❖ Veto tracks that point towards the region between ECAL barrel and endcap at  $1.42 < |\eta| < 1.65$

Finally, two further characteristics of chargino tracks are exploited. As the chargino is produced in a very clean environment (no further particles around the chargino is ex-

Table 4.5: Intermodule ECAL gaps.

$\eta$ -ranges
$-1.14018 < \eta < -1.1439$
$-0.791884 < \eta < -0.796051$
$-0.44356 < \eta < -0.447911$
$0.00238527 < \eta < -0.00330793$
$0.446183 < \eta < 0.441949$
$0.793955 < \eta < 0.789963$
$1.14164 < \eta < 1.13812$

Figure 4.26: Visualisation of the excluded region by the bad cathode strip chamber veto in the detector's  $\phi - \eta$ .

pected), the isolation of the track can discriminate signal against background events. Furthermore, for charginos decaying inside the tracker there is no associated energy deposition in the calorimeters in the direction of the track. This is a very pronounced characteristics of signal tracks for short chargino lifetimes. The resulting selection cuts are as follows

- ❖ No further substantial track activity is allowed in a cone of  $\Delta R < 0.3$  around the candidate track:  $\sum_{\Delta R < 0.3} p_T^{\text{trk}}/p_T^{\text{cand}} - 1 < 0.1$  (the subtraction by 1 corresponds to the contribution of the candidate track itself)
- ❖ No large calorimeter energy deposits (ECAL+HCAL) in a cone of  $\Delta R < 0.5$  around the track:  $E_{\text{calo}}^{\Delta R < 0.5} < 5 \text{ GeV}$ .

The discrimination power of these two variables is shown in Fig. 4.27. For even higher chargino lifetimes  $c\tau > 100 \text{ cm}$ , the calorimeter isolation requirement starts rejecting signal events because the charginos reach the calorimeters.

As emphasised before, this analysis aims at being sensitive especially on shorter lifetimes. Still, in order to allow for charginos decaying at any position in the tracker, no explicit selection cut on the number of missing outer hits is required.

Events are selected if they at least contain one track fulfilling all candidate track selection requirements. An overview over the full analysis preselection is given in Table 4.6. The event yields after the selections of each of the categories from Table 4.6 are listed in

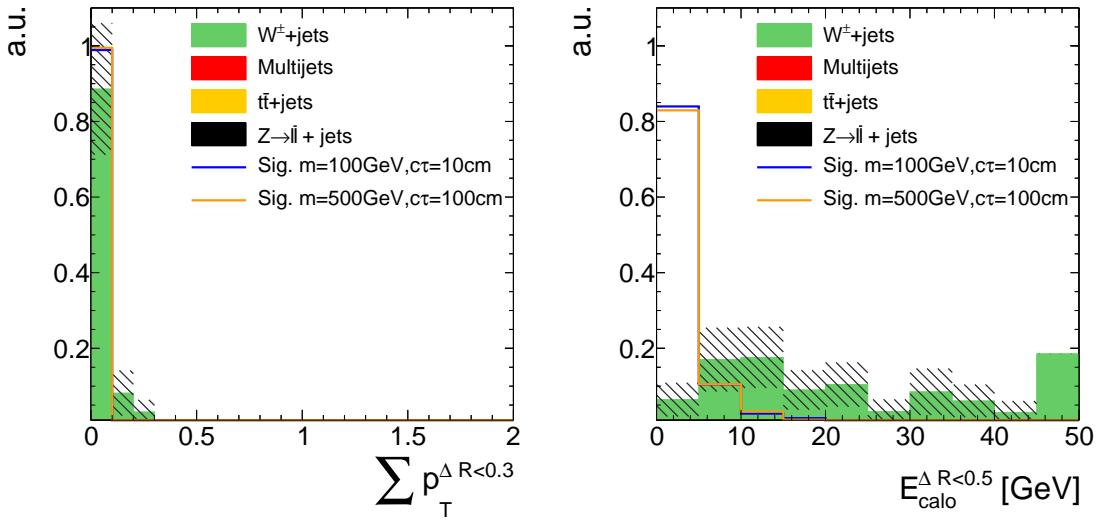


Figure 4.27: Track isolation (left) and calorimeter energy deposits (right) of the candidate track after the full previous selection. All events with a track isolation or a calorimeter energy deposit larger than the range shown in the figures are contained in the last bin.

Table 4.6: Summary and categorisation of the analysis selection.

Trigger		HLTMonocentralPFJet80_PFMETnoMu95_NHEF0p95 HLTMonocentralPFJet80_PFMETnoMu105_NHEF0p95 HLT_MET120_HBHENoiseCleaned
Event-based selection	Trigger selection	$p_T^{1^{\text{st}} \text{ jet}} > 110 \text{ GeV}$ with $ \eta_{1^{\text{st}} \text{ jet}}  < 2.4$ , $\text{CHF}_{1^{\text{st}} \text{ jet}} > 0.2$ , $\text{CEF}_{1^{\text{st}} \text{ jet}} < 0.5$ , $\text{NHF}_{1^{\text{st}} \text{ jet}} < 0.7$ , $\text{NEF}_{1^{\text{st}} \text{ jet}} < 0.7$ $\cancel{E}_T > 100 \text{ GeV}$
	QCD suppression	$\Delta\phi_{\max}(\text{jet}_i, \text{jet}_j) < 2.7$ for all jets with $p_T > 20 \text{ GeV},  \eta  < 4.5$ $\Delta\phi_{\max}(\text{jet}_i, \cancel{E}_T) > 0.5$ for two leading jets
		$\geq 1$ track that fulfils the following criteria:
		Good quality selection high-purity as defined in [?] $N_{\text{miss}}^{\text{middle/inner}} = 0$ $ d0  < 0.02 \text{ cm}$ $ dz  < 0.5 \text{ cm}$
		Kinematic selection $ \eta  < 2.1$ $p_T > 20 \text{ GeV}$
Candidate track selection	Lepton/jet veto	No muon within $\Delta R < 0.15$ No electron within $\Delta R < 0.15$ No tau within $\Delta R < 0.15$ No jet within $\Delta R < 0.5$ No dead/noisy ECAL cell within $\Delta R < 0.05$ Not within an ECAL intermodule gap Not within $1.42 <  \eta  < 1.65$ Not within $\Delta R < 0.25$ to a bad CSC
		Isolation selection $\sum_{\Delta R < 0.3} p_T^{\text{trk}} / p_T^{\text{cand}} - 1 < 0.1$ $E_{\text{calo}}^{\Delta R < 0.5} < 5 \text{ GeV}$

Table 4.7 for the available simulated background samples, some exemplary simulated signal models and for observed data. The discrepancies between data and simulation after the full preselection is stemming from three effects. First, the simulated yield is highly uncertain because of large event weights ( $\sim 15$ ). Second, not all relevant background samples were available (e.g.  $Z \rightarrow \nu\bar{\nu} + \text{jets}$  sample), leading to a lower prediction in simulation. Third, simulation is not expected to describe the here selected events (lepton/jet veto) very well. This emphasises once more the need for data-driven background estimation methods. Detailed event yield tables can be found in Appendix ??.

Given the presented signal candidate selection, two variables remain that are highly discriminating: The transverse momentum  $p_T$  and the energy release per path length  $dE/dx$  of the candidate track. In this analysis, the Asymmetric Smirnov discriminator  $I_{\text{as}}$  is used to enhance the discriminating power of  $dE/dx$ . See Section 4.3.3 for the definition and a detailed explanation of  $I_{\text{as}}$ . In Fig. 4.28, the distribution of the remaining two variables are shown after the selection of signal candidate events. These variables are used to optimise the sensitivity of the search. The optimisation process will be explained in Section 4.7. However, before the optimisation can be accomplished, a characterisation and estimation of the background is needed. This topic will be discussed in the following chapter.

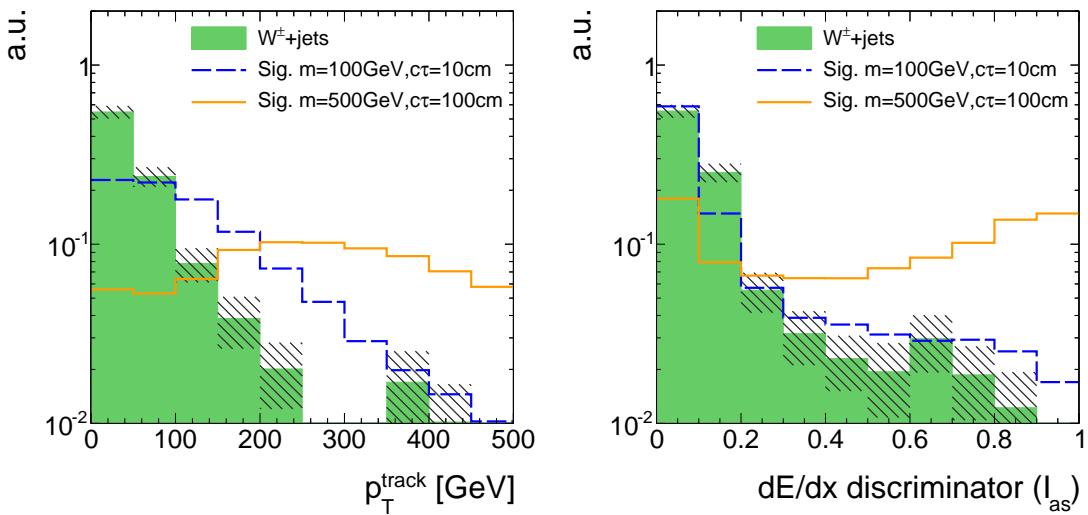


Figure 4.28: Candidate track  $p_T$  (left) and  $I_{\text{as}}$  (right) after the full signal candidate selection for signal and  $W + \text{jets}$  events. Because of the low statistical precision of the  $W + \text{jets}$  sample, the trigger selection is not applied.

Table 4.7: Event yields in simulation and data after the selections of each of the categories from Table 4.6

Selection	Simulated background samples			Simulated signal samples			Data		
	$W + \text{jets}$	$t\bar{t} + \text{jets}$	$Z \rightarrow \ell\bar{\ell}$	Multijet	m=100GeV $c\tau=10\text{ cm}$	m=100GeV $c\tau=100\text{ cm}$	m=500GeV $c\tau=10\text{ cm}$	m=500GeV $c\tau=100\text{ cm}$	MET data
After skim	$9.16 \cdot 10^7$	$1.04 \cdot 10^6$	$2.21 \cdot 10^7$	$1.38 \cdot 10^{11}$	$3.41 \cdot 10^5$	$3.41 \cdot 10^5$	$3.46 \cdot 10^2$	$3.46 \cdot 10^2$	$1.07 \cdot 10^7$
Event-based selection:									
Trigger	$4.31 \cdot 10^6$	$1.15 \cdot 10^5$	$4.23 \cdot 10^3$	$4.32 \cdot 10^6$	$1.55 \cdot 10^4$	$1.49 \cdot 10^4$	$46.2$	$46.2$	$1.07 \cdot 10^7$
Trigger selection	$1.89 \cdot 10^6$	$5.31 \cdot 10^4$	$6.26 \cdot 10^2$	$9.63 \cdot 10^5$	$1.09 \cdot 10^4$	$9.83 \cdot 10^3$	$36.3$	$36.3$	$3.94 \cdot 10^6$
QCD suppression	$1.11 \cdot 10^6$	$6.76 \cdot 10^3$	$1.32 \cdot 10^2$	$9.55 \cdot 10^3$	$7.90 \cdot 10^3$	$6.98 \cdot 10^3$	$27.6$	$27.6$	$1.38 \cdot 10^6$
Track-based selection:									
Good quality selection	$1.07 \cdot 10^6$	$6.63 \cdot 10^3$	$1.32 \cdot 10^2$	$9.55 \cdot 10^3$	$2.80 \cdot 10^3$	$5.38 \cdot 10^3$	$5.07$	$20.0$	$1.30 \cdot 10^6$
Kinematic selection	$8.14 \cdot 10^5$	$5.63 \cdot 10^3$	$1.32 \cdot 10^2$	$5.48 \cdot 10^3$	$2.54 \cdot 10^3$	$4.93 \cdot 10^3$	$4.73$	$18.9$	$9.51 \cdot 10^5$
Lepton/jet veto	$5.02 \cdot 10^2$	$5.88$	$0$	$0$	$1.99 \cdot 10^3$	$3.67 \cdot 10^3$	$3.83$	$15.0$	$616$
Isolation selection	$31.9$	$0.67$	$0$	$0$	$1.67 \cdot 10^3$	$3.04 \cdot 10^3$	$3.39$	$12.6$	$119$

## 4.6 Characterisation and estimation of the Standard Model backgrounds

After the application of the signal candidate selection, explained in the previous chapter, the background arising from Standard Model processes is dramatically reduced. Only two events in the simulated  $W + \text{jets}$  sample remain. One of these originates from an unreconstructed muon, the other one from an unreconstructed electron, both passing the lepton vetoes. This implies, that the electron, muon, and tau vetos cannot reject all leptons because some are not properly reconstructed. Due to the limited size of the simulated  $W + \text{jets}$  dataset (15 times smaller than the number of events expected from  $W + \text{jets}$  processes during 2012 data taking), it is not possible to rely on a full simulation-based estimation of the leptonic background. The underlying mechanism that a lepton can pass the lepton veto and the corresponding methods to estimate the leptonic background will be explained in detail in Section 4.6.2.

Furthermore, there is the possibility that a track is reconstructed out of a set of hits that do not originate from only one single particle. Such tracks are called “fake tracks”. Background tracks arising from a combination of unrelated hits will be explained in the following Section 4.6.1. It should be noted that the fake background is contributing through all SM processes, not only via  $W + \text{jets}$ . Still, as the characteristics of fake tracks are independent of the underlying process, this background can also be studied on simulation using  $W + \text{jets}$  events only.

There is no contribution of jets to the background, because jets are efficiently suppressed by the jet veto and the track isolation requirement of the signal candidate selection (see Table 4.6).

The reader will recognise, that the importance of the two contributions to the background, the fake and leptonic background, is very different to this search. It will be seen, that the leptonic background is of negligible size.

However, both background are estimated in a similar approach. First, an inclusive background estimation (without the use of  $dE/dx$  information) is estimated. Afterwards, the efficiency of a  $dE/dx$  selection for the fake and leptonic background is determined.

Finally, the final signal regions are determined within an optimisation procedure. This optimisation is carried out in the track variables  $p_T$  and  $dE/dx$  (see Chapter 4.7) to ensure an ideal selection for the search for short and highly ionising tracks.

### 4.6.1 Fake background

Fake tracks are reconstructed out of the tracker hits of more than one particle. The rate at which this false reconstruction occurs is highly restrained by the quality cuts on  $\chi^2$  and the vertex compatibility of the track reconstruction algorithm. Details on the reconstruction algorithm of tracks at CMS can be found in Section 3.3.2.2.

Reconstructed tracks that are fake tracks consist in general only out of a few hits. This can be seen in Fig. 4.29, where the normalised distribution of the number of hits from fake tracks in simulated  $W + \text{jets}$  events is depicted. There are almost no fakes with more than seven hits. In simulation, fake tracks are defined as tracks that cannot be matched to a generator-level particle within a distance of  $\Delta R < 0.01$ .

Fakes are efficiently suppressed by the requirements of no missing middle or inner hits and the compatibility with the primary vertex. Unfortunately, wrongly reconstructed tracks which pass these criteria, do also easily pass the  $E_{\text{calo}}^{\Delta R < 0.5} < 5 \text{ GeV}$  requirement with high efficiency, as fake tracks cannot be correlated to energy deposits in the calorimeters.

Fake tracks are mainly caused by the wrong combination of pileup tracks and electronic noise. This leads to the fact, that the occurrence of fake tracks across various background processes is stable, as will be seen later.

In this analysis, the estimation of the fake background is split into two parts. First, the background is determined inclusively in  $dE/dx$ . Second, the  $dE/dx$  ( $I_{\text{as}}$ ) distribution is estimated with the help of a fake enriched control region. This second step is needed to enable an optimisation in  $dE/dx$  (see Section 4.7).

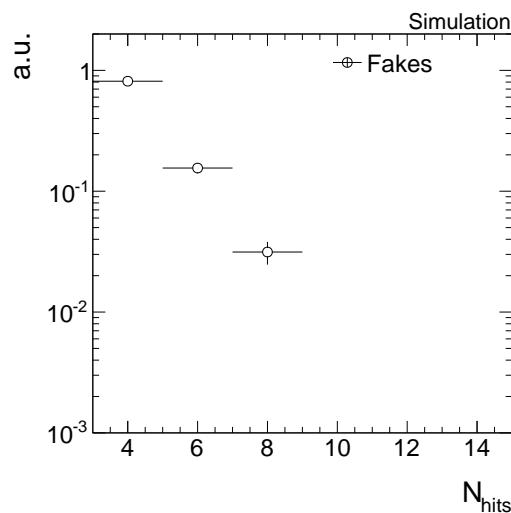


Figure 4.29: Normalised distribution of the number of hits for fake tracks in the simulated  $W + \text{jets}$  sample. To increase the statistical precision, only the candidate track selection from Table 4.6 is applied.

#### 4.6.1.1 Inclusive fake background estimation

The inclusive background estimation closely follows the background estimation method done in [?, ?]. It aims at determining the probability of having a fake track in an event that passes the full signal candidate selection (Table 4.6) plus a potential additional  $p_T$  selection cut that is determined in an optimisation procedure (Section 4.7). This probability will be called the fake rate  $\rho_{\text{fake}}$ . Within [?, ?], it was checked that the fake rate is constant for different processes. Thus, it is possible to determine  $\rho_{\text{fake}}$  with the help of one SM process and then generalise it to all SM background processes.

The inclusive fake background is estimated with the help of  $Z \rightarrow \mu\bar{\mu}$  and  $Z \rightarrow e\bar{e}$  events from data.  $Z \rightarrow \ell\bar{\ell}$  events can be selected with high purity by requiring two well reconstructed muons or electrons that are opposite in charge and for which the invariant mass is around the  $Z$ -boson mass of  $\sim 90$  GeV. As these events do not contain further leptons from the hard interaction (processes with a further lepton are negligible), any additional track is either a constituent of an ISR jet, a soft particle from pileup events or is a fake. Since the candidate track selection requires a track with a  $p_T > 20$  GeV that is no lepton or jet, it suppresses ISR jets and soft tracks from the underlying event. Thus, applying the track-based signal candidate selection on  $Z \rightarrow \ell\bar{\ell}$  events selects fake tracks with high purity.

The selection of two well reconstructed muons and electrons is done with the single-muon and single-electron datasets listed in Table 4.8. These datasets contain at least one muon or one electron in every event.

Table 4.8: Datasets used for the determination of the fake rate.

Dataset	Integrated luminosity [ $\text{fb}^{-1}$ ]
/SingleMu/Run2012A-22Jan2013-v1/AOD	0.876
/SingleMu/Run2012B-22Jan2013-v1/AOD	4.405
/SingleMu/Run2012C-22Jan2013-v1/AOD	7.040
/SingleMu/Run2012D-22Jan2013-v1/AOD	7.369
/SingleElectron/Run2012A-22Jan2013-v1/AOD	0.876
/SingleElectron/Run2012B-22Jan2013-v1/AOD	4.412
/SingleElectron/Run2012C-22Jan2013-v1/AOD	7.050
/SingleElectron/Run2012D-22Jan2013-v1/AOD	7.368

For the  $Z \rightarrow \mu\bar{\mu}$  selection, an event is required to have two muons with  $p_T > 25 \text{ GeV}$  and  $|\eta| < 2.4$ . To suppress background from cosmic muons, the distance from the primary vertex must be less than  $|d0| < 0.2 \text{ cm}$  in radial and  $|dz| < 0.5 \text{ cm}$  in longitudinal direction. In order to suppress background arising from jets that fake muons, various quality criteria are applied: it is required that there is at least one hit in the muon detector that is considered in the global muon fit, and that at least two measurements are from different muon detector stations. Concerning the track of the muon in the silicon tracker system, at least six hits in the full tracker system of which at least one pixel hit is required. An isolation criterion is applied that requires the sum of transverse momenta of all particle-flow particles in a cone of  $\Delta R < 0.4$  around the muon to be less than 12% of the muon  $p_T$ . Finally, the muons are required to be opposite in charge and to have an invariant mass between 80 to 100 GeV. The  $Z \rightarrow \mu\bar{\mu} + \text{fake track}$  selection is summarised in Table 4.9.

In order to select  $Z \rightarrow e\bar{e}$  events in data, the two electrons are required to have

Table 4.9: Event selection cuts for the  $Z \rightarrow \mu\bar{\mu} + \text{fake}$  control sample to estimate the inclusive fake background.

Event-based selection	Two global muons with $p_T > 25 \text{ GeV}$ $ \eta  < 2.4$ $\sum_{\Delta R < 0.4} p_T^{\text{PF particle}} / p_T(\mu) < 0.12$ $\frac{\chi^2}{ndof} \Big _{\text{global track}} < 10$ $ d0  < 0.2 \text{ cm}$ $ dz  < 0.5 \text{ cm}$ $\geq 1 \text{ hit in the muon detector considered in global fit}$ $\geq 2 \text{ hits in different muon stations}$ $\geq 1 \text{ hit in the pixel detector}$ $\geq 6 \text{ hits in the tracker system}$
	Muons opposite in charge $80 \text{ GeV} < M_{\text{inv}}(\mu_1, \mu_2) < 100 \text{ GeV}$
Candidate track selection	Good quality selection Kinematic selection Lepton/jet veto Isolation selection

$p_T > 25 \text{ GeV}$ ,  $|\eta| < 2.5$  and no missing hits in the inner layers of the tracker. Furthermore, the electrons need to pass a conversion veto as described in [?] in order to reduce background arising from photon conversions. An isolation requirement similar to the muon isolation criterion is applied with an increased threshold of 15%. The electron identification is further based on a multivariate technique developed within [?] that exploits electron characteristics concerning the track quality, the ECAL cluster shapes, and the combination of the measurements in the tracker and in the ECAL. Again, the two electrons must be opposite in charge and their invariant mass must be between  $80 - 100 \text{ GeV}$ . A summary of the  $Z \rightarrow e\bar{e} + \text{fake track}$  event selection can be found in Table 4.10.

In Fig. 4.30, the invariant mass distribution  $M_{\text{inv}}$  is shown for  $Z \rightarrow \mu\bar{\mu}$  and  $Z \rightarrow e\bar{e}$  events in simulation and data after the event-based selections described above (Tables 4.10 and 4.9).

When applying a  $Z \rightarrow \ell\bar{\ell}$  selection plus the candidate track selection, the selected tracks should be mostly fakes. Whether this is indeed the case can be tested on simulated  $Z \rightarrow \ell\bar{\ell}$  events. As can be seen in Fig. 4.31, a reasonable purity in fake tracks can be achieved by applying the candidate track selection on top of the  $Z \rightarrow \ell\bar{\ell}$  selection. In simulated  $Z \rightarrow \mu\bar{\mu}$  events, a purity of 88% is achieved, whereas in simulated  $Z \rightarrow e\bar{e}$  events a purity of 92% of fake tracks is achieved.

Table 4.10: Event selection cuts for the  $Z \rightarrow e\bar{e} + \text{fake}$  control sample to estimate the inclusive fake background.

Event-based selection	Two Electrons with
	$p_T > 25 \text{ GeV}$ $ \eta  < 2.5$ $\sum_{\Delta R < 0.4} p_T^{\text{PF particle}} / p_T(e) < 0.15$ pass conversion veto [?] no missing inner tracker hits good MVA electron as defined in [?]
	Electrons opposite in charge $80 \text{ GeV} < M_{\text{inv}}(e_1, e_2) < 100 \text{ GeV}$
Candidate track selection	Good quality selection Kinematic selection Lepton/jet veto Isolation selection

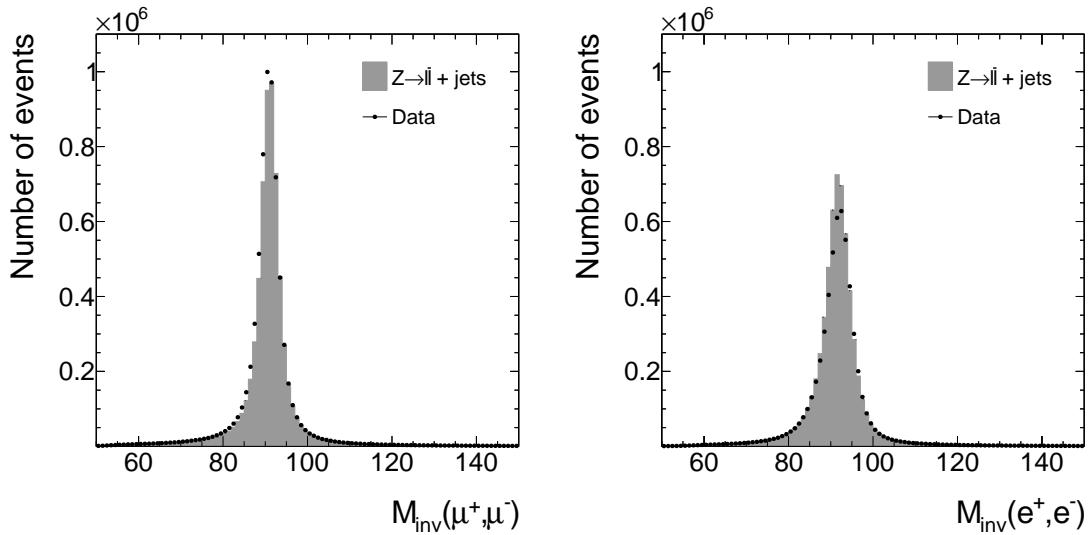


Figure 4.30: The invariant mass of the two selected muons (left) and the two selected electrons (right) after the event-based selection from Tables 4.10 and 4.9, respectively.

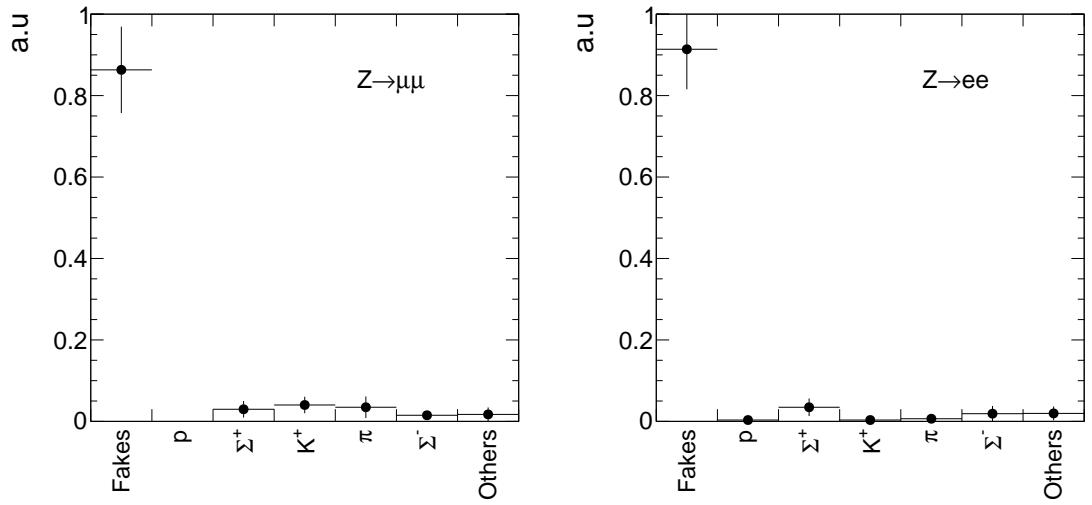


Figure 4.31: Corresponding generator-level particles of all tracks within  $Z \rightarrow \ell\bar{\ell} + \text{fake}$  that were selected according to the candidate track selection. The full selection for tracks in  $Z \rightarrow \mu\bar{\mu}$  events (left) is given in Table 4.9. The full selection for tracks in  $Z \rightarrow e\bar{e}$  events (right) is given in Table 4.10. “Fake” means that no corresponding generator-level particle is found.

As already mentioned, the fake rate is defined as the probability that an event contains a fake track that fulfils the candidate track selection. Thus, for the  $Z \rightarrow \ell\bar{\ell}$  datasets it is defined as the number of events passing the full selection described in Table 4.9 (Table 4.10) divided by the number of events that pass only the event-based selection in Table 4.9 (Table 4.10)

$$\rho_{\text{fake}} = \frac{N_{Z \rightarrow ll}^{\text{cand trk selection}}}{N_{Z \rightarrow ll}} \quad (4.16)$$

Fake rates are determined independently for the  $Z \rightarrow \mu\bar{\mu} + \text{fake}$  and  $Z \rightarrow e\bar{e} + \text{fake}$  event selection and then averaged to obtain the final fake rate. In Table 4.11 the results of  $N_{Z \rightarrow ll}^{\text{cand trk selection}}$ ,  $N_{Z \rightarrow ll}$  and the resulting fake rate for the candidate track selection given in Table 4.6 are presented. The averaged fake rate is thus  $(6.86 \pm 0.25) \cdot 10^{-5}$ . This is not the final result as the optimisation in  $p_T$  will add an additional  $p_T$  selection cut to the candidate track selection.

As mentioned before, it was checked within [?, ?] that the fake rate is constant for different Standard Model processes. This is shown in Fig. 4.32 where the fake rate is depicted for the most important SM processes. Since the fake rate is constant for different SM processes, the fake rate determined on the  $Z \rightarrow \ell\bar{\ell}$  dataset can be generalised for all SM background possibly contributing to this search. Thus, the inclusive fake background can be estimated by multiplying the fake rate with the number of events selected from the MET dataset (Table 4.4) by applying the event-based signal candidate requirements from Table 4.6

$$N_{\text{bkg}}^{\text{fake, inclusive in } I_{\text{as}}} = \rho_{\text{fake}} \cdot N_{\text{event-based selection}}^{\text{MET}} \quad (4.17)$$

Given the number of events after the event-based selection of  $N_{\text{event-based selection}}^{\text{MET}} = 1.38 \cdot 10^6$  and the fake rate cited above, the inclusive fake background can be estimated to  $94.7 \pm 3.4$  for the candidate track selection.

It should be noted again that the inclusive fake background estimation will be only inclusive in  $I_{\text{as}}$  not in  $p_T$ . That means that after the definition of the signal region,  $N_{\text{bkg}}^{\text{fake, inclusive in } I_{\text{as}}}$  is determined with the additional optimal  $p_T$  selection.

Table 4.11: Results of  $N_{Z \rightarrow ll}^{\text{cand trk selection}}$ ,  $N_{Z \rightarrow ll}$  and  $\rho_{\text{fake}}$  for the  $Z \rightarrow \mu\bar{\mu} + \text{fake}$  (Table 4.9) and  $Z \rightarrow e\bar{e} + \text{fake}$  (Table 4.10) selection.

Channel	$N_{Z \rightarrow ll}^{\text{cand trk selection}}$	$N_{Z \rightarrow ll}$	$\rho_{\text{fake}}$
$Z \rightarrow \mu\bar{\mu}$	403	$6.17 \cdot 10^6$	$(6.53 \pm 0.33) \cdot 10^{-5}$
$Z \rightarrow e\bar{e}$	369	$5.08 \cdot 10^6$	$(7.26 \pm 0.38) \cdot 10^{-5}$

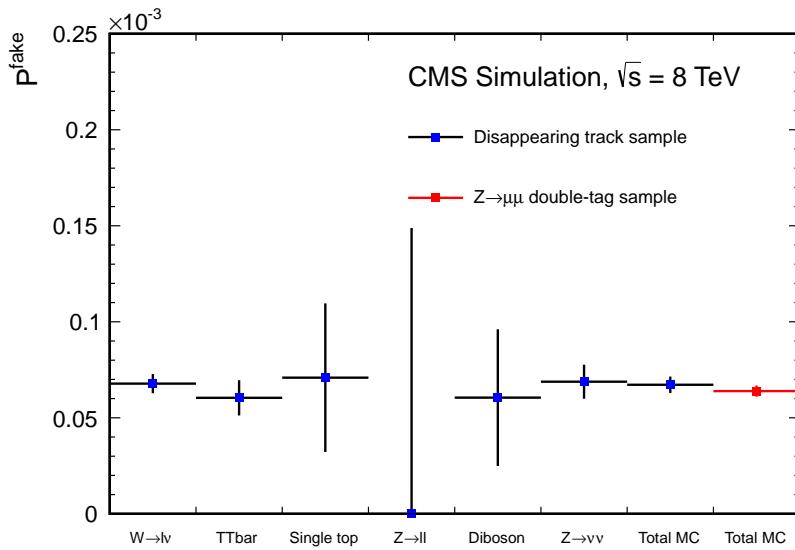


Figure 4.32: Fake track rate estimated in [?, ?] for tracks with four hits. Taken from [?]

Possible differences between the fake rate in  $Z \rightarrow \ell\bar{\ell}$  events and other SM processes are estimated on simulated events and taken into account as a systematic uncertainty (see Section 4.6.4.1).

#### 4.6.1.2 $dE/dx$ shape of fake background

The information about the energy release per path length for fake tracks should not be taken from simulated samples as the simulation of  $dE/dx$  is not reliable (cf. Fig. 4.16). Within this analysis the Asymmetric Smirnov discriminator  $I_{\text{as}}$  is used to discriminate signal against background with respect to  $dE/dx$  (see Section 4.3.3). In order to estimate the  $I_{\text{as}}$  shape of fake tracks, a control region  $\text{CR}_{I_{\text{as}}}^{\text{fake}}$  is defined that is enriched with fakes and shows the same  $I_{\text{as}}$  distribution as fake tracks in the signal region.

To enrich fake tracks, it is possible to invert the selection cuts on the number of missing middle and inner hits, i.e. requiring at least one missing inner or middle hit ( $N_{\text{miss}}^{\text{inner}} + N_{\text{miss}}^{\text{middle}} > 0$ ). Figure 4.33 shows the distribution of the number of missing inner plus missing middle hits for fake and leptonic tracks in simulated  $W + \text{jets}$  events. It can be seen that this selection is enriched by fakes. The resulting purity of fakes in  $\text{CR}_{I_{\text{as}}}^{\text{fake}}$  is about 98% (see Fig. 4.34).

Additionally, it must be checked whether the  $I_{\text{as}}$  shape in  $\text{CR}_{I_{\text{as}}}^{\text{fake}}$  is representative for the  $I_{\text{as}}$  shape in the signal region. As the exact definition of the signal region will be addressed during optimisation, this test is done for various  $p_T$  selection cuts.

The comparison of the  $I_{\text{as}}$  shape of fake tracks can only be done with simulated events.

Thus, simulated  $W + \text{jets}$  events are used to select fake tracks in both regions. A comparison of the shape for the candidate track selection and the  $\text{CR}_{I_{\text{as}}}^{\text{fake}}$  is shown in Fig. 4.34.

The  $I_{\text{as}}$  shape is almost identical in the signal and in the control region which makes the

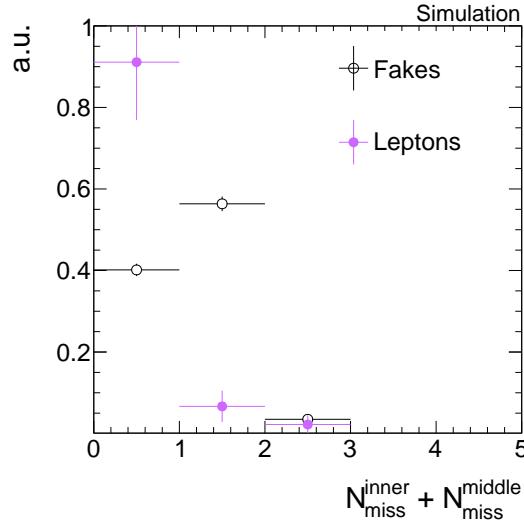


Figure 4.33: Normalised number of missing inner plus missing middle hits for fake and leptonic tracks for the full candidate track selection with the selection requirements on  $N_{\text{miss}}^{\text{inner}}$  and  $N_{\text{miss}}^{\text{middle}}$  removed. Trigger selection and QCD suppression cuts were removed to enhance the statistical precision.

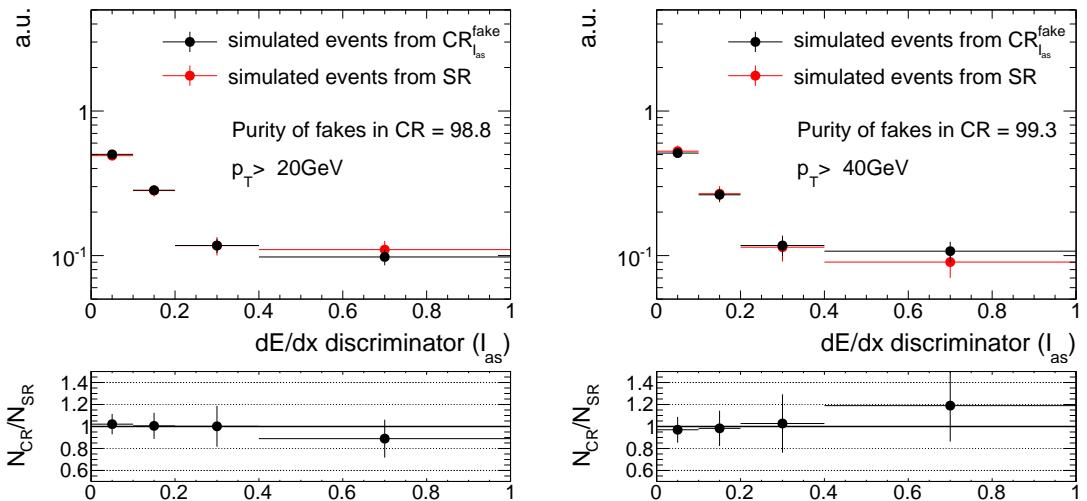


Figure 4.34: Comparison of the  $I_{\text{as}}$  shape between  $\text{CR}_{I_{\text{as}}}^{\text{fake}}$  and the signal region for two different track  $p_T$  selections of  $p_T > 20 \text{ GeV}$  (left) and  $p_T > 40 \text{ GeV}$  (right). To enhance the statistical precision only the track-based selection is applied.

definition of the control region perfectly suited for estimating the  $I_{\text{as}}$  shape from  $\text{CR}_{I_{\text{as}}}^{\text{fake}}$  in data. The remaining shape differences are taken into account as a systematic uncertainty (discussed in Section 4.6.4.2).

## 4.6.2 Leptonic background

The leptonic background of the here presented search is caused by non-reconstructed leptons that circumvent the lepton veto selection. However, at least non-reconstructed electrons or taus should in principle deposit enough energy in the calorimeters such that they can still be vetoed by the calorimeter isolation requirement  $E_{\text{calo}}^{\Delta R < 0.5} < 5 \text{ GeV}$ . As muons don't deposit much energy in the calorimeters, this reasoning does not apply to them. All of the three lepton types behave like MIPs. Thus, they loose much less energy than hypothetical new heavy particles and can therefore be further discriminated by their ionisation loss in the tracker system.

In the following, the sources of the three different leptonic backgrounds are characterised.

### Electrons

To reject unreconstructed electrons, all tracks pointing to a dead or noisy ECAL cell, to an ECAL intermodule gap, or to the region between ECAL barrel and endcap at  $1.42 < |\eta| < 1.65$  are vetoed, as described in Section 4.5.2. By this selection, almost all electrons are efficiently rejected.

However, there is still the possibility that an electron fails reconstruction and pass the signal candidate selection. This can happen either, if an electron do bremsstrahlung and the direction of the electron is significantly changed. Thus, the energy deposits in the ECAL can possibly not be matched to the original electron. Alternatively, there is also the possibility that an electron track is pointing towards a non-working ECAL cell, that is not included in the dead and noisy ECAL cell veto.

This possibility can be seen in the single event in the  $W + \text{jets}$  sample that pass the full signal candidate selection and where the candidate track can be matched to a generator-level electron. This event is visualised in Fig. 4.35 (left). The neutrino, only weakly interacting does not show any signature in the detector, whereas the electron ( $p_{\text{T}}^{\text{gen}} \simeq 90 \text{ GeV}$ ) leaves a track with  $p_{\text{T}} \simeq 70 \text{ GeV}$  in the tracker. Only little ECAL energy deposits in the direction of the electron are visible. This is caused by the fact that one of the corresponding ECAL crystal is not working properly and thus no energy deposition can be recorded (cf. Fig. 4.35 (right)). An ISR jet ( $p_{\text{T}} \simeq 230 \text{ GeV}$ ) causes the  $\cancel{E}_{\text{T}}$  in the event.

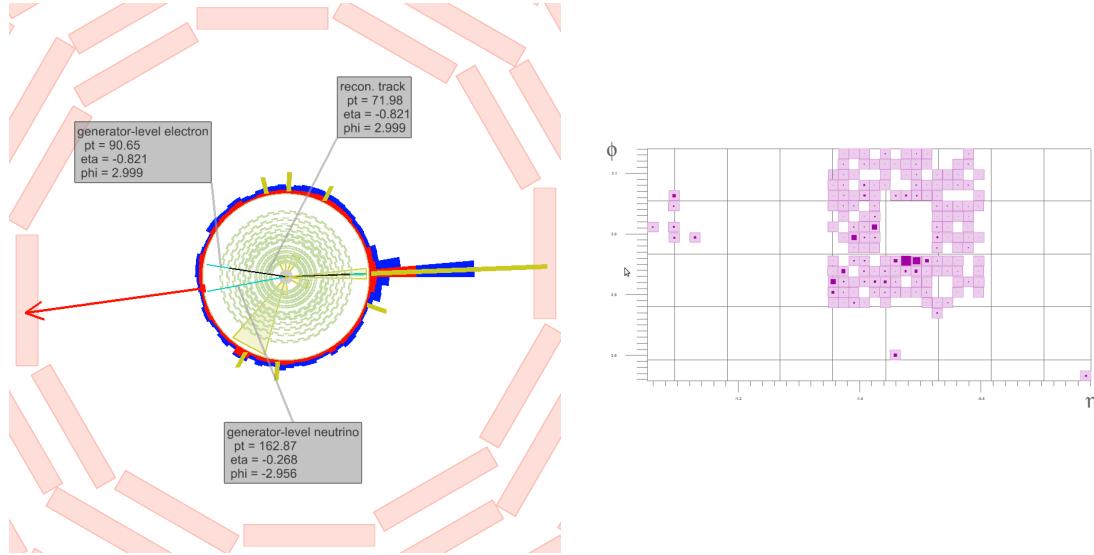


Figure 4.35: Left: Visualisation of a  $W \rightarrow e\nu_e$  event contributing to the SM background. In light blue, generator-level particles including  $e$  and  $\nu_e$  of the  $W$ -boson decay are shown. Black lines represent reconstructed tracks and the red arrow indicates the missing transverse energy in the event. Right: Detailed view of the energy deposition in the ECAL towers in the direction of the electron track ( $\phi - \eta$  plane).

## Taus

Taus that decay hadronically are contributing to the leptonic background through the decay of a tau lepton to one charged pion  $\tau \rightarrow \pi^\pm \nu_\tau$ . Other hadronic decay modes of the tau lepton are suppressed by the track isolation criterion. Taus can fail reconstruction if they only deposit little energy in the HCAL or ECAL. This is usually due to a low energetic pion from the tau decay. Unreconstructed taus can therefore also easily bypass the calorimeter isolation criterion. Because of nuclear interactions in the tracker, pions often result in short reconstructed tracks that can easily be highly mismeasured in  $p_T$ . Thus, taus can contribute to the background even if imposing a tight selection in the transverse momentum.

Such an event is shown in Fig. 4.36. The transverse momentum of the generator-level pion is only  $p_T \sim 10$  GeV, but because the reconstructed track is very short, it leads to a high mismeasurement of the track  $p_T$  of  $\sim 40$  GeV. The shortness of the track is caused by nuclear interactions of the pion. As only little corresponding ECAL or HCAL energy deposits are measured, the reconstruction of the pion fails. The ISR jet causes the  $E_T$  in the event.

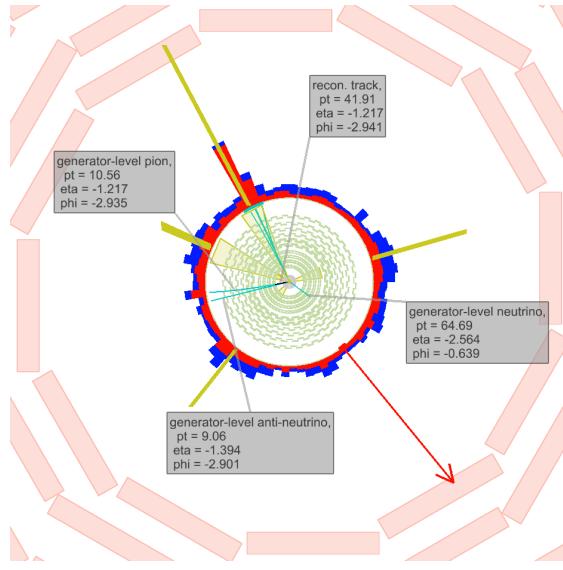


Figure 4.36: Visualisation of a  $W^+ \rightarrow \tau^+\nu_\tau \rightarrow \pi^+\bar{\nu}_\tau\nu_\tau$  event contributing to the SM background. In light blue, the generator-level particles including  $\pi^+$ ,  $\bar{\nu}_\tau$  and  $\nu_\tau$  are shown. The black line represents the reconstructed pion track and the red arrow indicates the missing transverse energy in the event.

## Muons

Muons can fail reconstruction if they point towards a bad cathode strip chamber. This is taken into account in the candidate track selection. However, some of the muons still fail reconstruction if they fall within the gap between stations 0 and 1 of the drift tube system at  $|\eta| = 0.25$ . The muon reconstruction efficiency drops from around 99% to a value of around 94%, as shown in [?, ?]. This possibility is illustrated in a simulated event shown in Fig. 4.37. There, the muon is pointing to the  $\eta$ -region between stations 0 and 1 of the DT system. No signal in the muon chambers is visible. Therefore the muon could not be reconstructed.

In [?, ?] events are rejected if the track is pointing in a region of  $0.15 < |\eta| < 0.35$ . In this search, this cut was omitted to maximise signal acceptance. Due to the additional selection in  $I_{as}$ , muons can be efficiently suppressed. E.g. in the event shown in Fig. 4.37, the muon has an  $I_{as}$  value of about 0.007.

As for the fakes, the leptonic background estimation is splitted into two parts. First, the estimation of the inclusive background without  $I_{as}$  information. Second, the estimation of the  $I_{as}$  shape for all three leptonic background sources. To have the possibility to make an optimisation in the two main discriminating variables  $p_T$  and  $I_{as}$ , the background estimation methods are designed to work for all different  $p_T$  and  $I_{as}$  selection cuts.

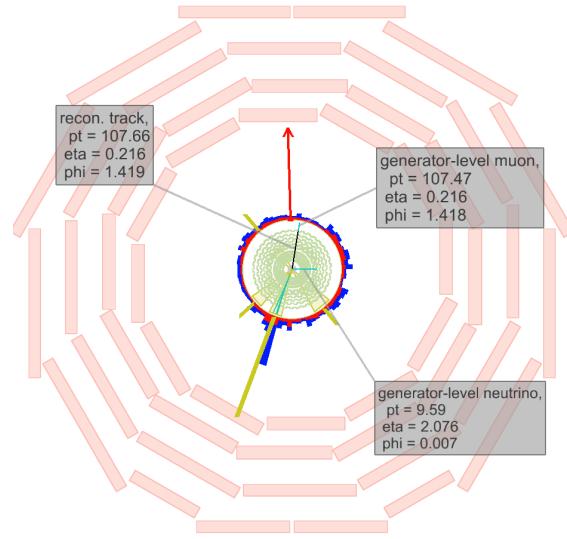


Figure 4.37: Visualisation of an  $W \rightarrow \mu\nu_\mu$  event contributing to the SM background. In light blue, the generator-level particles including  $\mu$  and  $\nu_\mu$  of the  $W$  decay are shown.

#### 4.6.2.1 Inclusive leptonic background estimation

The inclusive (without  $dE/dx$  information) lepton background estimation method is similar to the background estimation method used in [?, ?].

In order to estimate the number of events in the signal region originating from leptons that pass the lepton veto, information from simulated events is used. With the help of simulated  $W + \text{jets}$  events, the ratio  $\rho_{\text{MC}}^{\text{lept}_i}$  between the number of events in the signal region with the selected track matched to a generator-level lepton  $N_{\text{SR}}^{\text{trk matched to lepton}_i}$  and the number of events in a control region  $N_{\text{CR}}^{\text{lepton}_i \text{ veto inverted}}$  with a inverted lepton veto is determined. For muons, this lead to the following expression

$$\rho_{\text{MC}}^\mu = \frac{N_{\text{SR},\text{MC}}^{\text{trk matched to } \mu}}{N_{\text{CR},\text{MC}}^{\mu \text{ veto inverted}}}. \quad (4.18)$$

Since for electrons and taus the reconstruction efficiency is highly correlated with the  $E_{\text{calo}}^{\Delta R < 0.5}$  selection requirement, the  $E_{\text{calo}}^{\Delta R < 0.5}$  requirement is additionally removed in the control regions for these two lepton types

$$\rho_{\text{MC}}^{e,\tau} = \frac{N_{\text{SR},\text{MC}}^{\text{trk matched to } e,\tau}}{N_{\text{CR},\text{MC}}^{e,\tau \text{ veto inverted, } E_{\text{calo}}^{\Delta R < 0.5} < 5 \text{ GeV}}}. \quad (4.19)$$

In order to estimate the inclusive background for all three lepton types, the scale factor

$\rho_{\text{MC}}^{\text{lep}_i}$  is applied to the number of events in the lepton veto inverted control region measured in data. Also in data the control region for electrons and taus is defined with the  $E_{\text{calo}}^{\Delta R < 0.5}$  requirement removed. Thus, the inclusive number of predicted background events can be estimated with

$$N_{\text{predicted}}^{\mu, \text{ inclusive in } I_{\text{as}}} = \rho_{\text{MC}}^{\mu} \cdot N_{\text{CR,data}}^{\mu \text{ veto inverted}}. \quad (4.20)$$

for muons, and

$$N_{\text{predicted}}^{e,\tau, \text{ inclusive in } I_{\text{as}}} = \rho_{\text{MC}}^{e,\tau} \cdot N_{\text{CR,data}}^{e,\tau \text{ veto inverted}, \cancel{E_{\text{calo}}^{\Delta R < 0.5} < 5 \text{ GeV}}}. \quad (4.21)$$

for electrons and taus.

This method relies on the simulation of the lepton reconstruction efficiencies which is known to be reasonably accurate [?, ?, ?]. For electrons and taus the simulation of the calorimeter isolation is utilised as well. Possible discrepancies between simulation and data are taken into account as a systematic uncertainty via a comparison of the lepton reconstruction efficiencies in data and simulation in  $Z \rightarrow \ell\bar{\ell}$  events (see Section 4.6.4.3).

To reduce the statistical uncertainty, the scale factor is calculated without applying the QCD suppression cuts. After the signal candidate selection described in Section 4.5.2, only one event remains in the simulated  $W + \text{jets}$  sample where the candidate track can be matched to an electron. There are five events with a track candidate that can be matched to a muon, and no selected events have tracks that can be matched to a pion from a tau decay. The statistical uncertainties are calculated as the 68% upper and lower limits on the inclusive background with the Neyman procedure [?, ?]. Table 4.12 gives the result for the prediction of the inclusive leptonic background for the signal candidate selection from Section 4.5.2.

Table 4.12: Scale factor  $\rho_{\text{MC}}^{\text{lep}_i}$ , number of events in the data control region  $N_{\text{CR,data}}$  and the resulting inclusive estimation  $N_{\text{predicted}}$  after the candidate track selection.

	$\rho_{\text{MC}}^{\text{lep}_i}$	$N_{\text{CR,data}}^{\text{veto inverted}}$	$N_{\text{predicted}}^{\text{inclusive in } I_{\text{as}}}$
electrons	$1.25^{+1.70}_{-0.77} \cdot 10^{-4}$	60067	$7.49^{+10.19}_{-4.63}$
muons	$2.17^{+1.65}_{-0.93} \cdot 10^{-4}$	76664	$16.64^{+12.64}_{-7.12}$
taus	$< 2.13 \cdot 10^{-2}$	445	$< 9.46$

### 4.6.2.2 $dE/dx$ shape of leptonic background

In order to get information about the  $I_{as}$  (see Section 4.3.3) shape in the signal region of electrons, muons and taus, a control region should be found where the shape of the observable is at least similar to that in the signal region. The most natural control region, being the lepton veto inverted control region, cannot be used because the variable  $I_{as}$  differs between the signal and the control region, as can be seen in Fig. 4.38. The discrepancies reach factors up to an order of magnitude.

As this control region is not suitable, it is decided to use the  $I_{as}$  information from simulation. This introduces a large bias since  $dE/dx$  (and therefore  $I_{as}$ ) is not simulated well. However, the corresponding bias is still smaller than the differences of the  $I_{as}$  shape between the signal and a control region: compare Fig. 4.38 and Fig. 4.39.

In order to take into account the bias when using  $I_{as}$  from simulation, a systematic

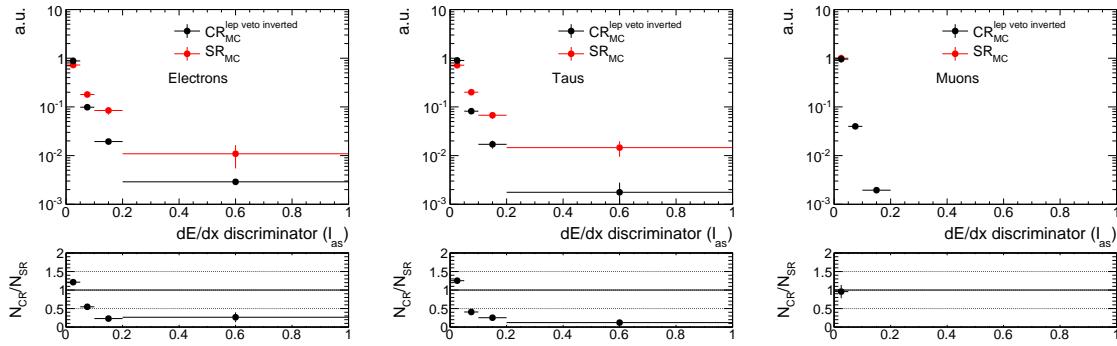


Figure 4.38: Normalised  $I_{as}$  distribution for electrons (left), pions from the tau decay (middle) and muons (right) in the signal region (red) and the lepton veto inverted control region (black).

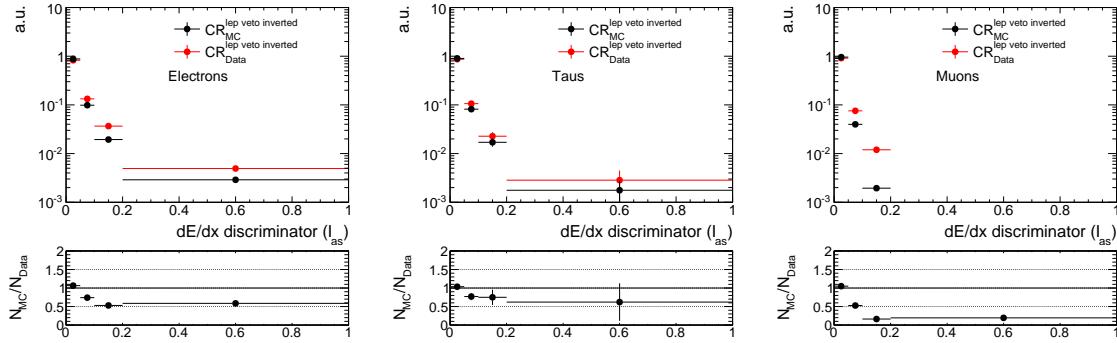


Figure 4.39: Normalised  $I_{as}$  distribution for electrons (left), pions from the tau decay (middle) and muons (right) in the lepton veto inverted control region from simulated events (black) and data (red).

uncertainty is estimated that addresses simulation-data differences of the  $I_{\text{as}}$  distributions. This systematic uncertainty is discussed in Section 4.6.4.4.

### 4.6.3 Background estimation validation

The background estimation methods are exhaustively validated in signal depleted control regions. Various control regions are used for validation. For each control region it has been verified that the signal contamination is less than the statistical uncertainty of the background prediction. For some of the models the expected number of events exceeds this limit. However, these models are already ruled out by the search for disappearing tracks [?] (see Appendix ??).

To validate the estimation method of the leptonic background, a leptonic control region is defined by selecting only tracks with a minimum number of seven hits in the tracker. This reduces the fake contribution to a negligible level (cf. Fig. 4.29). Additionally in order to minimise signal contamination, the calorimeter isolation requirement is inverted to  $E_{\text{calo}}^{\Delta R < 0.5} > 10 \text{ GeV}$ . This requirement ensures no overlap to the signal region.

The validation test for the control region with  $E_{\text{calo}}^{\Delta R < 0.5} > 10 \text{ GeV}$  and  $N_{\text{hits}} > 6$  is shown in Table 4.13. The predicted number of events by the leptonic background estimation is compatible with the observed data yield.

Table 4.13: Validation test of leptonic background estimation. Left:  $E_{\text{calo}}^{\Delta R < 0.5} > 10 \text{ GeV}$  and  $N_{\text{hits}} > 6$ . Right:  $E_{\text{calo}}^{\Delta R < 0.5} > 10 \text{ GeV}$ ,  $N_{\text{hits}} > 6$  and  $I_{\text{as}} > 0.2$ . Only statistical uncertainties are included.

	Predicted Yield	Data Yield		Predicted Yield	Data Yield
Total bkg	$131.70^{+26.30}_{-18.42}$	156	Total bkg	$0.0^{+0.50}_{-0.0}$	1
Electrons	$14.67^{+11.16}_{-6.29}$		Electrons	$0.0^{+0.07}_{-0.0}$	
Muons	$7.99^{+10.90}_{-5.00}$		Muons	$0.0^{+0.32}_{-0.0}$	
Taus	$109.04^{+21.18}_{-16.58}$		Taus	$0.0^{+0.38}_{-0.0}$	

In order to validate the fake background, the same selection as for the leptonic background validation is applied except for the number of tracker hits. The fake rate can only be estimated within the low calorimeter isolation region ( $E_{\text{calo}}^{\Delta R < 0.5} < 10 \text{ GeV}$ ) to ensure high fake purity. To be able to validate the method in the high calorimeter isolat-

tion region ( $E_{\text{calo}}^{\Delta R < 0.5} > 10 \text{ GeV}$ ), a translation factor from the low to the high calorimeter isolation region for the number of fake tracks is determined in the fake enriched control region  $\text{CR}_{I_{\text{as}}}^{\text{fake}}$  defined in Section 4.6.4.2. In this control region, the ratio of  $N_{E_{\text{calo}}^{\Delta R < 0.5} > 10 \text{ GeV}} / N_{E_{\text{calo}}^{\Delta R < 0.5} < 10 \text{ GeV}}$  is estimated and taken as a multiplicative factor to the number of events predicted from the  $E_{\text{calo}}^{\Delta R < 0.5} < 10 \text{ GeV}$  region. Furthermore, also the  $I_{\text{as}}$  shape is taken from the low calorimeter region ( $E_{\text{calo}}^{\Delta R < 0.5} < 10 \text{ GeV}$ ) to ensure a sufficient number of events in the  $I_{\text{as}}$  distribution. Thus, this method is exposed to differences in the  $I_{\text{as}}$  shape of the two  $E_{\text{calo}}^{\Delta R < 0.5}$  regions. In Fig. 4.40, a comparison between the  $I_{\text{as}}$  shape of fake tracks with low calorimeter energy deposits and high calorimeter energy deposits in the fake enriched control region is depicted. It shows, that fakes with higher associated calorimeter energy have typically lower  $I_{\text{as}}$  values. This can thus lead to an overprediction of the fake contribution in the high  $E_{\text{calo}}^{\Delta R < 0.5}$  region. However, the differences are covered by statistical uncertainties that are propagated to the final background estimate and can therefore be considered as accounted for. Additionally, it has to be emphasised that only the validation region with  $E_{\text{calo}}^{\Delta R < 0.5} > 10 \text{ GeV}$  is affected by the  $I_{\text{as}}$  shape difference and that this is not of concern for the signal regions with  $E_{\text{calo}}^{\Delta R < 0.5} < 5 \text{ GeV}$ . In Table 4.14, two different validation tests are shown, once an inclusive validation in  $I_{\text{as}}$  and once with an  $I_{\text{as}}$  selection of 0.2. Again, the predicted background events is in agreement with the number of observed events.

The whole validation is done for different selections in  $p_{\text{T}}$  and  $I_{\text{as}}$ . All validation tests show good agreement. Results of a variety of validation tests with different  $p_{\text{T}}$  and  $I_{\text{as}}$

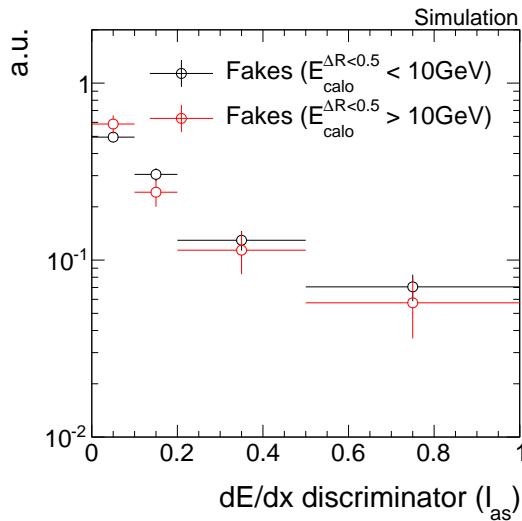


Figure 4.40: Comparison of the  $I_{\text{as}}$  shape of fake tracks with high calorimeter energy deposits (red) and low calorimeter energy deposits (black) in simulation. To enhance the statistical precision, no trigger selection is applied.

Table 4.14: Validation test of fake and leptonic background estimation methods. Left:  $E_{\text{calo}}^{\Delta R < 0.5} > 10 \text{ GeV}$ . Right:  $E_{\text{calo}}^{\Delta R < 0.5} > 10 \text{ GeV}$  and  $I_{\text{as}} > 0.2$ . Only statistical uncertainties are included.

	Predicted Yield	Data Yield		Predicted Yield	Data Yield
Total bkg	$308.35^{+33.48}_{-26.64}$	324	Total bkg	$14.69^{+2.91}_{-2.84}$	16
Electrons	$59.92^{+16.11}_{-11.85}$		Electrons	$0.75^{+0.36}_{-0.25}$	
Muons	$8.04^{+10.97}_{-5.03}$		Muons	$0.00^{+0.32}_{-0.00}$	
Taus	$173.06^{+24.62}_{-20.23}$		Taus	$2.33^{+0.74}_{-0.55}$	
Fakes	$67.34^{+11.61}_{-11.61}$		Fakes	$11.61^{+2.78}_{-2.78}$	

selections can be found in Appendix ??.

Still, systematic uncertainties need to be estimated. The sources of systematic uncertainties and how they are estimated will be explained in the following section.

## 4.6.4 Systematic uncertainties

Systematic uncertainties on the background estimation include:

- the uncertainty on the fake rate  $\rho_{\text{fake}}$ ;
- the uncertainty on the  $I_{\text{as}}$  shape of fake tracks predicted from a control region;
- the uncertainty on the leptonic scale factor  $\rho_{\text{MC}}^{\text{lep}_i}$  determined with simulated events;
- the uncertainty on the  $I_{\text{as}}$  shape of the leptonic background.

### 4.6.4.1 Uncertainty on the fake rate

The fake rate  $\rho_{\text{fake}}$  is determined with the help of observed  $Z \rightarrow \ell\bar{\ell}$  events. To estimate the uncertainty on this fake rate caused by differences in the fake rate between different underlying processes, a comparison between the fake rate in simulated  $Z \rightarrow \ell\bar{\ell} + \text{jets}$  and simulated  $W + \text{jets}$  events is done. The fake rate in the  $Z \rightarrow \ell\bar{\ell} + \text{fake track control samples}$  (see Tables 4.9 and 4.10) and the fake rate in the signal candidate selection from Table 4.6 in  $W + \text{jets}$  events are compared.

Unfortunately, the statistical precision of the simulated  $W + \text{jets}$  dataset is limited. Thus, the estimation of the systematic uncertainty is mainly driven by statistical uncertainties. In order to enhance the statistical precision of the estimation, the selection requirements on  $\cancel{E}_T$  and  $p_T^{\text{1st jet}}$  are loosened and the QCD suppression requirements are removed. These variables are not expected to be correlated with the fake rate and thus should not affect it. That this is indeed the case, can be seen in Table 4.15.

The systematic uncertainty is estimated as the largest difference from one of the ratio  $\rho_{\text{fake}}^{W+\text{jets}} / \rho_{\text{fake}}^{Z \rightarrow \ell\bar{\ell}}$  and its statistical uncertainty. For the candidate track selection, this is estimated to  $\rho_{\text{fake}}^{W+\text{jets}} / \rho_{\text{fake}}^{Z \rightarrow \ell\bar{\ell}} = 0.96 \pm 0.16$  leading to a systematic uncertainty on the fake rate of 20%.

Table 4.15: Fake rates in simulated  $W + \text{jets}$  and  $Z \rightarrow \ell\bar{\ell} + \text{jets}$  events for different event-based selections of the  $W + \text{jets}$  sample. The track-based selection is the candidate track selection from Table 4.6.

$W + \text{jets}$ selection	$\rho_{\text{fake}}^{W+\text{jets}}$	$\rho_{\text{fake}}^{Z \rightarrow \ell\bar{\ell}}$
$\cancel{E}_T > 100 \text{ GeV}, p_T^{\text{1st jet}} > 110 \text{ GeV}$	$(3.16^{+4.26}_{-1.94}) \cdot 10^{-5}$	$(3.17 \pm 0.21) \cdot 10^{-5}$
$\cancel{E}_T > 0 \text{ GeV}, p_T^{\text{1st jet}} > 70 \text{ GeV}$	$(3.03 \pm 0.68) \cdot 10^{-5}$	$(3.17 \pm 0.21) \cdot 10^{-5}$
$\cancel{E}_T > 0 \text{ GeV}, p_T^{\text{1st jet}} > 70 \text{ GeV}, \text{ no QCD cuts}$	$(3.05 \pm 0.44) \cdot 10^{-5}$	$(3.17 \pm 0.21) \cdot 10^{-5}$

#### 4.6.4.2 Uncertainty on the $dE/dx$ shape of fake tracks

The systematic uncertainty on the shape of the  $I_{\text{as}}$  distribution takes into account the differences between the  $I_{\text{as}}$  shape in the fake control region  $\text{CR}_{I_{\text{as}}}^{\text{fake}}$  and in the signal region. For the estimation, information from simulated  $W + \text{jets}$  events is used. A comparison between the simulated  $I_{\text{as}}$  shape in the signal and in the control region can be seen in Fig. 4.41. To enhance the statistical precision only track-based selection cuts are applied.

The largest deviation from one of the ratio of the number of events in the signal region and the control region with its 1-sigma statistical uncertainty is taken as systematic uncertainty. For a signal region definition with  $p_T > 20 \text{ GeV}$  and  $I_{\text{as}} > 0.2$  this corresponds to an uncertainty of around 21% and for a definition with  $p_T > 40 \text{ GeV}$  and  $I_{\text{as}} > 0.2$  of around 25%.

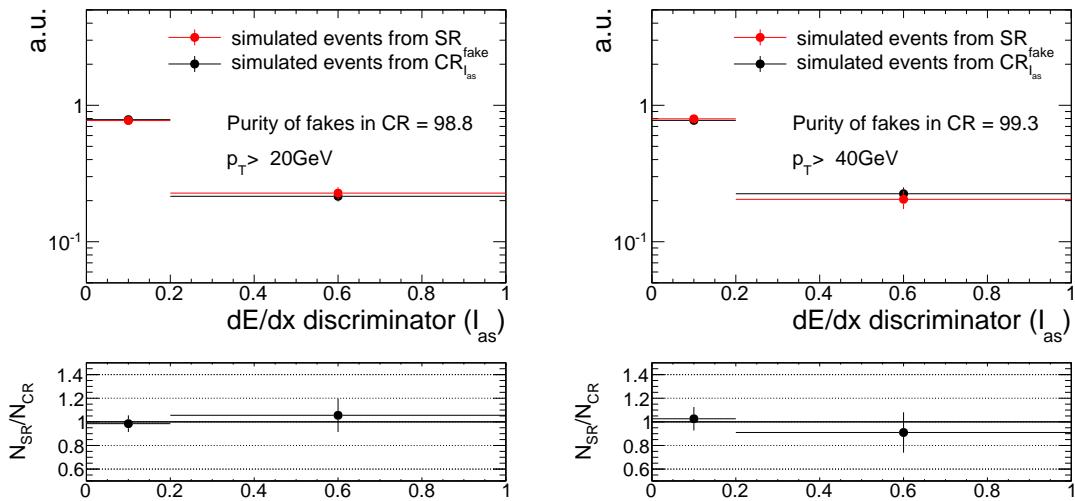


Figure 4.41: Normalised distributions of the  $I_{\text{as}}$  shape of fake tracks in the signal and control region of simulated  $W + \text{jets}$  events with a  $p_T$  selection of 20 GeV (left) and a 40 GeV (right).

#### 4.6.4.3 Uncertainty on the leptonic scale factor

The leptonic scale factor  $\rho_{\text{MC}}^{\text{lep}_i}$  is estimated on simulated  $W + \text{jets}$  events. The corresponding systematic uncertainty that addresses the use of information from simulation is derived by a “tag-and-probe” method performed on real data and simulated events.

For this method a selection of  $Z \rightarrow \ell\bar{\ell}$  events is done with one “tagged” well reconstructed lepton and one “probed” candidate track. To ensure a selection of  $Z \rightarrow \ell\bar{\ell}$  events, a selection on the invariant mass of the reconstructed lepton and the candidate track is applied with  $80 \text{ GeV} < M_{\text{inv}}(\text{lepton}, \text{cand. trk}) < 100 \text{ GeV}$  for muons and electrons. For taus, a muon from a  $\tau \rightarrow \mu\nu\nu$  decay is selected with  $40 \text{ GeV} < M_{\text{inv}}(\mu, \text{cand. trk}) < 75 \text{ GeV}$  and  $m_T(\mu, \cancel{E}_T) < 40 \text{ GeV}$  [?, ?]. Furthermore, the candidate track and the lepton are required to be opposite in charge. In order to reduce the contamination of fakes in the “tag-and-probe” samples an additional selection on the number of hits of  $N_{\text{hits}} > 5$  is required.

The “tag-and-probe” selection is done for each lepton type separately. In order to determine the leptonic scale factors, the number of events is once estimated for the candidate track selection including the corresponding lepton veto which gives the number of events in the “tag-and-probe” signal region  $N_{\text{SR}}^{\text{T\&P}}$ , and once inverting the lepton veto selection requirement which gives the number of events in the “tag-and-probe” lepton inverted control region  $N_{\text{CR, lepton veto inverted}}^{\text{T\&P}}$ . As for the determination of the tau and electron scale factor with simulated  $W + \text{jets}$  events, no requirement on the calorimeter isolation is applied in the lepton veto inverted control region for taus and electrons. This leads to the

following expression of the lepton scale factor for muons

$$\rho^\mu = \frac{N_{\text{SR}}^{\text{T&P}\mu}}{N_{\text{CR}, \mu \text{ veto inverted}}^{\text{T&P}}}. \quad (4.22)$$

and for electrons and taus

$$\rho^{e,\tau} = \frac{N_{\text{SR}}^{\text{T&P}e,\tau}}{N_{\text{CR}, e,\tau \text{ veto inverted}}^{\text{T&P}} \cdot \cancel{E_{\text{lepto}}^{\Delta R < 0.5} < 5 \text{ GeV}}}. \quad (4.23)$$

The selection requirements for the three tag-and-probe samples are listed in Tables ??, ?? and ?? in Appendix ??.

The leptonic scale factors are calculated using simulated  $Z \rightarrow \ell\bar{\ell}$  events and real data from the single-muon and single-electron samples listed in Table 4.8. Since only scaling factors are derived within simulation and data, no trigger is applied. This does not change the scaling factors since the requirements on the well reconstructed lepton are exactly the same for the numerator and denominator of Eqs. (4.22) and (4.23). In Table 4.16, the results of the event yields in the ‘‘tag-and-probe’’ control and signal regions as well as the corresponding lepton scale factor are depicted. The largest difference from unity of the ratio  $\rho_{\text{MC}}^{\text{lep}_i}/\rho_{\text{Data}}^{\text{lep}_i}$  and its statistical uncertainty is taken as systematic uncertainty. This results for the signal candidate selection in an uncertainty of 69% for the electron, 39% for the muon and 79% for the tau scale factor.

Table 4.16: Event yields in the tag-and-probe signal region and control region with the resulting scale factors in simulation and data.

		Muons	Electrons	Taus
Data	$N_{\text{SR}}^{\text{T&P}\text{lep}_i}$	211	319	19
	$N_{\text{CR}, \text{lep}_i \text{ veto inverted}}^{\text{T&P}}$	$4.10 \cdot 10^6$	$3.74 \cdot 10^6$	33
	$\rho^{\text{lep}_i}$	$(5.14 \pm 0.35) \cdot 10^{-5}$	$(8.52 \pm 0.48) \cdot 10^{-5}$	$(5.76 \pm 1.66) \cdot 10^{-1}$
Simulation	$N_{\text{SR}}^{\text{T&P}\text{lep}_i}$	$153.9 \pm 15.4$	$125.1 \pm 15.8$	$9.1 \pm 4.0$
	$N_{\text{CR}, \text{lep}_i \text{ veto inverted}}^{\text{T&P}}$	$(4.284 \pm 0.003) \cdot 10^6$	$(4.112 \pm 0.003) \cdot 10^6$	$30.9 \pm 7.8$
	$\rho^{\text{lep}_i}$	$(3.59 \pm 0.36) \cdot 10^{-5}$	$(3.04 \pm 0.39) \cdot 10^{-5}$	$(2.95 \pm 1.49) \cdot 10^{-1}$

#### 4.6.4.4 Uncertainty on the leptonic $dE/dx$ shape

The uncertainty on lepton  $I_{as}$  shape is estimated by a comparison of the  $I_{as}$  shape in data and simulation in the lepton veto inverted control region. Figure 4.42 shows the leptonic  $I_{as}$  distributions for all three lepton types in the lepton veto inverted control region in data and simulation. The largest difference from one of the ratio (and its statistical uncertainty) of the number of events in the control region in data and simulation is taken as systematic uncertainty. This leads for example to uncertainties between 37% – 81% for the signal candidate selection plus a selection requirement of  $I_{as} > 0.2$ .

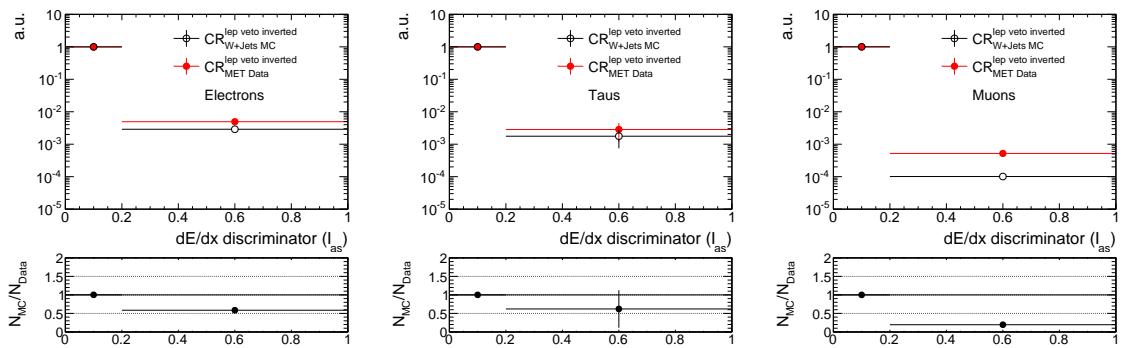


Figure 4.42: Normalised distributions of the lepton  $I_{as}$  distributions in the lepton veto inverted control region for data (red) and simulation (black) for all three lepton types. The event-based selection requirements and the calorimeter isolation requirement are removed to enhance the statistical precision.

## 4.7 Optimisation of the search sensitivity

Finally, having all background estimation methods in place, an optimisation procedure is performed in order to increase the search sensitivity with respect to different signal models as introduced in Section 4.4.2. The optimisation is done in the most sensitive variables,  $p_T$  and  $I_{as}$  (see Section 4.3.3 for a definition and explanation of the Asymmetric Smirnov discriminator  $I_{as}$ ). A potential additional discriminating variable is the number of missing outer hits  $N_{lost}^{outer}$  in the tracker system. This variable is, however, not considered in this analysis because the discriminating potential for this search is limited, as shown in Appendix ??.

SUSY models with different chargino lifetimes and masses are characterised by different

$p_T$  and  $I_{as}$  distributions as well as different theoretical cross sections. Therefore, the usual search optimisation strategy that maximises  $N_S/\Delta B$  ( $N_S$  = number of signal events of model  $S$ ,  $\Delta B$  = background uncertainty) implies a potential fine-tuning on the specific SUSY cross sections. In order to keep the search as general as possible, a cross section independent optimisation is performed. This is achieved by a minimisation of the cross section for which a  $5\sigma$ -discovery ( $\kappa = 5$ ) of the corresponding signal model is expected, i.e. finding the optimal selection cuts for  $p_T$  and  $I_{as}$  for which the lowest possible cross section,  $\sigma_{min}$ , can be discovered

$$\kappa = \frac{\alpha_{min} \cdot N_S(\text{mass}, c\tau, p_T^{\text{cut}}, I_{as}^{\text{cut}})}{\Delta B(p_T^{\text{cut}}, I_{as}^{\text{cut}})} = 5. \quad \text{with } \alpha_{min} = \frac{\sigma_{min}}{\sigma_S}. \quad (4.24)$$

The number of expected events  $N_S$  of the signal model  $S$  depends on the  $p_T$  and  $I_{as}$  selection cut as well as the mass and the lifetime of the chargino. The uncertainty on the background  $\Delta B$  is dependent on the  $p_T$  and  $I_{as}$  cut, and takes into account the full systematic uncertainty as well as the statistical uncertainty on the background prediction which is defined as the 68% one sided upper limit of a Poisson distribution with  $\mu = N_B$  estimated with the Neyman construction [?, ?]. The systematic uncertainty on the background prediction includes systematic uncertainties as described in Section 4.6.4, and statistical uncertainties arising from limited statistical precision of the control regions and simulated samples used in the background estimation. The factor  $\alpha_{min}$  that is minimised is the ratio of the minimum cross section  $\sigma_{min}$  divided by the nominal cross section  $\sigma_S$  of the signal model  $S$ .

As this analysis focuses on short tracks, rather low lifetimes are considered in the optimisation procedure:  $c\tau = 1\text{ cm}, 10\text{ cm}, 50\text{ cm}$ . These lifetimes are further suitable as they lie at the edge of the sensitivity of the search for disappearing tracks [?]. To cover the full mass space, the optimisation is done for masses between 100 GeV and 500 GeV in steps of 100 GeV.

The corresponding results are shown in Table 4.17. It can be seen that the optimal selection is highly dependent on the signal models. The best sensitivity for low masses ( $\leq 200\text{ GeV}$ ) is mainly achieved by soft selection cuts in  $p_T$  between 20 to 30 GeV, while models with higher chargino masses require tighter  $p_T$  selections of around 50 GeV. The optimal  $I_{as}$  selection is mostly dependent on the mass of the chargino. For low masses and low lifetimes a soft selection in  $I_{as} > 0.05$  is preferred. Since for longer lifetimes more charginos are able to reach the tracking system, a tighter selection in  $I_{as}$  of 0.3 is preferable. Additionally, signal models with longer chargino lifetimes have a more pronounced right tail in the  $I_{as}$  distribution (cf. Fig. 4.17 (right)). For high masses the highest search sensitivity is always achieved by a high  $I_{as}$  selection cut of 0.3.

Table 4.17: Optimal  $p_T$  and  $I_{\text{as}}$  selection cuts and the corresponding minimum cross section  $\sigma_{\min}$  that can be discovered with  $5\sigma$  significance for different signal models. For some signal samples, an optimisation result is not available due to the limited size of these samples.

Mass [GeV]	Lifetime $c\tau$ [cm]	Optimal $p_T$ cut	Optimal $I_{\text{as}}$ cut	$\sigma_{\min}$
100	1	30	0.05	61.596
200	1	20	0.05	43.414
300	1	n/a	n/a	n/a
400	1	n/a	n/a	n/a
500	1	n/a	n/a	n/a
100	10	30	0.05	1.531
200	10	30	0.30	0.561
300	10	30	0.30	0.354
400	10	30	0.30	0.238
500	10	50	0.30	0.201
100	50	50	0.30	0.435
200	50	50	0.30	0.110
300	50	50	0.30	0.063
400	50	50	0.30	0.045
500	50	50	0.30	0.037

In order to visualise the mass and  $c\tau$  dependence of the optimal  $p_T$  and  $I_{\text{as}}$  selection, the optimisation results for two very different lifetimes (5 cm and 50 cm) and masses (100 GeV and 500 GeV) are shown in Fig. 4.43, where the minimum cross section that is possible to discover is shown in the  $p_T - I_{\text{as}}$  plane. For simplicity, general systematic uncertainties on the leptonic and the fake background of 100% and 20% respectively are imposed in the visualisation. Uncertainties arising from limited statistical precision of the samples used for the background estimation are propagated consistently into Eq. (4.24). Similar to the full optimisation, it can be seen that for low masses and low lifetimes, the highest search

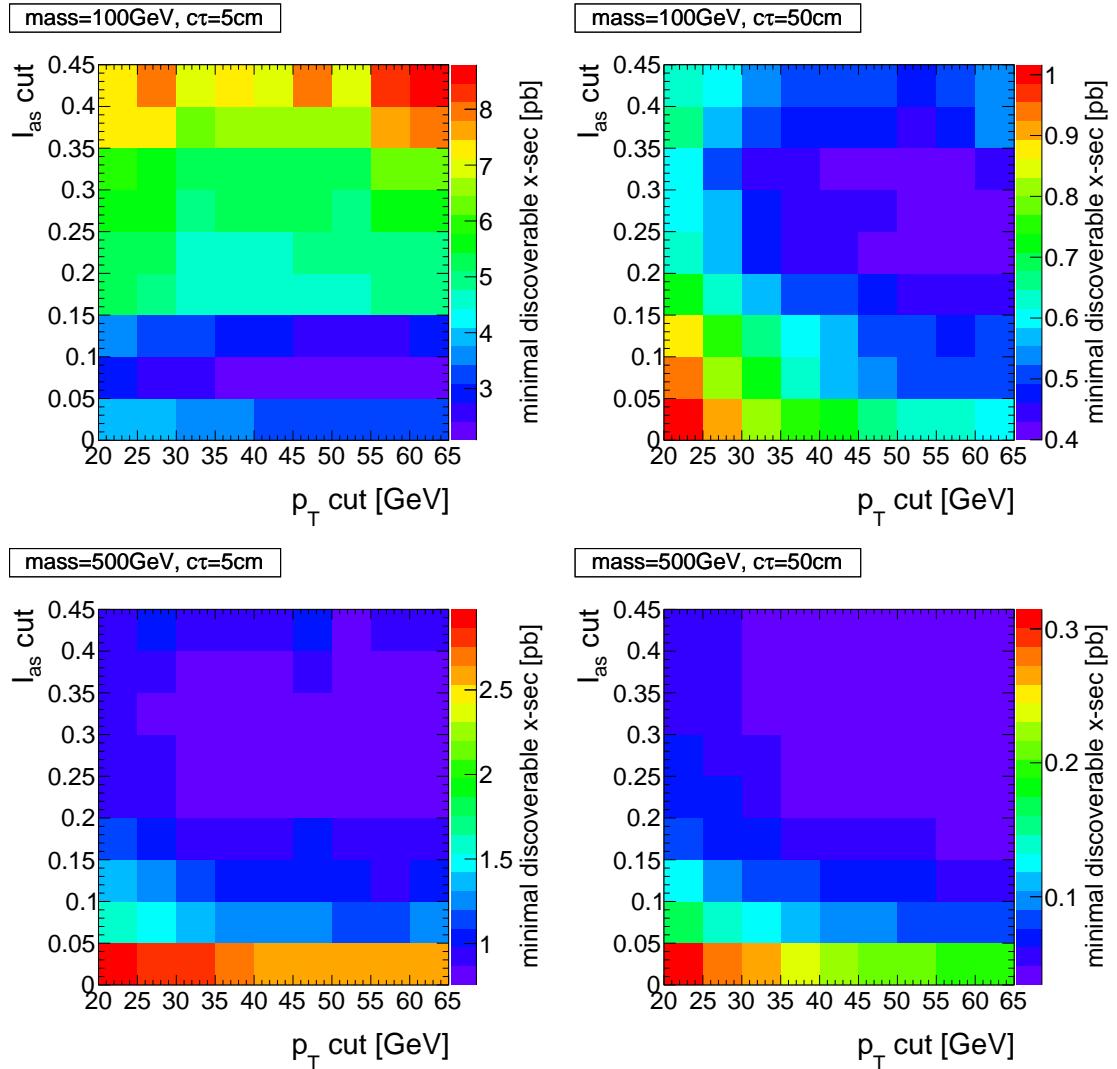


Figure 4.43: Minimum possible cross section that can be discovered with  $5\sigma$  significance as a function of minimum  $p_T$  and  $I_{as}$  requirements for four different signal models. The systematic uncertainties are taken to be 20% and 100% for the fake and the leptonic background respectively. The uncertainty on the background arising from the limited size of the used samples are propagated consistently to the search optimisation. In Table ?? of Appendix ??, the corresponding histograms of the background yield, the background uncertainty and the signal yield for the four signal models can be found.

sensitivity is achieved by imposing rather soft selection cuts on  $I_{\text{as}}$  and  $p_{\text{T}}$ . Optimising for higher lifetime pushes the optimal selection in  $p_{\text{T}}$  and  $I_{\text{as}}$  to larger values, where signal models with higher masses prefer even tighter  $I_{\text{as}}$  selection cuts than the corresponding lower mass signal model. It can also be seen, that for low lifetimes, the  $p_{\text{T}}$  dependence of the search sensitivity is less pronounced than for long lifetimes.

Based on the optimisation, four different exclusive signal regions are defined in order to achieve an optimal coverage over a wide mass space and a high sensitivity for different lifetimes:

- 1.)  $30 \text{ GeV} < p_{\text{T}} < 50 \text{ GeV}$  and  $0.05 < I_{\text{as}} < 0.3$
- 2.)  $p_{\text{T}} > 50 \text{ GeV}$  and  $0.05 < I_{\text{as}} < 0.3$
- 3.)  $30 \text{ GeV} < p_{\text{T}} < 50 \text{ GeV}$  and  $I_{\text{as}} > 0.3$
- 4.)  $p_{\text{T}} > 50 \text{ GeV}$  and  $I_{\text{as}} > 0.3$ .

## 4.8 Results

After developing the methods of the background estimation for all different background sources and their corresponding systematic uncertainties (all explained in Section 4.6), the search is performed in four exclusive signal regions with  $19.7 \text{ fb}^{-1}$  of data collected at a centre-of-mass energy of  $\sqrt{s} = 8 \text{ TeV}$  at the CMS experiment. The predicted numbers of events for the fake and the leptonic background in the four signal regions, as well as the number of observed events are listed in Table 4.18. It can be seen, that fake tracks are by far the dominant background to this search. The leptonic background contributes only in one signal region to the total background with a share of about 10%.

Furthermore, the observations are compatible with the Standard Model background within  $1\sigma$  uncertainties in all four signal regions. This is also visualised in Fig. 4.44, where a comparison of the total background prediction to the number of observed events is shown. No excess above the SM prediction is observed in either of the four signal regions. Thus, no evidence for physics beyond the Standard Model could be found.

Therefore, in the following section these results will be used to constrain the parameter space of supersymmetric models with almost mass degenerate charginos and neutralinos.

Table 4.18: Number of predicted (fake, leptonic and total) and observed events for the four different signal regions.

Signal region		Fake Bkg			Leptonic Bkg			Total Bkg			Data
$p_T$ [ GeV ]	$I_{as}$	pred	stat	sys	pred	stat	sys	pred	stat	sys	
30-50	0.05-0.30	19.11	$^{+2.61}_{-2.61}$	$\pm 9.35$	0.00	$^{+2.58}_{-0.00}$	$\pm 0.00$	19.11	$^{+3.67}_{-2.61}$	$\pm 9.35$	18
50- $\infty$	0.05-0.30	22.21	$^{+3.60}_{-3.60}$	$\pm 8.78$	2.17	$^{+2.99}_{-1.34}$	$\pm 1.65$	24.38	$^{+4.68}_{-3.84}$	$\pm 8.93$	34
30-50	0.30-1.00	2.49	$^{+0.85}_{-0.85}$	$\pm 1.98$	0.00	$^{+0.22}_{-0.00}$	$\pm 0.00$	2.49	$^{+0.87}_{-0.85}$	$\pm 1.98$	0
50- $\infty$	0.30-1.00	2.52	$^{+1.14}_{-1.14}$	$\pm 1.27$	0.04	$^{+0.30}_{-0.03}$	$\pm 0.03$	2.57	$^{+1.18}_{-1.14}$	$\pm 1.27$	4

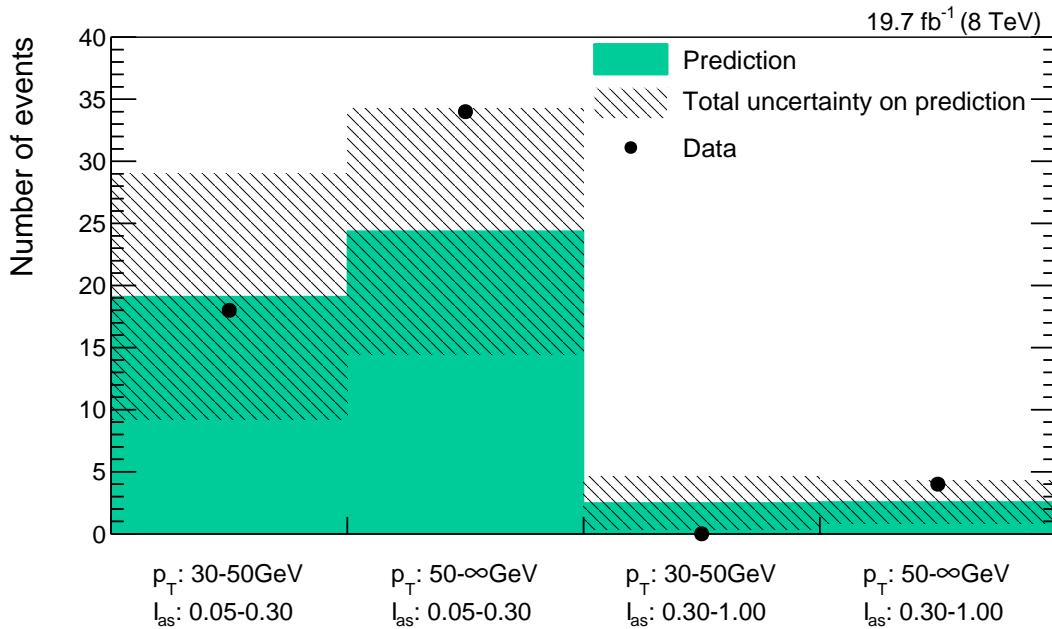


Figure 4.44: Number of predicted (green area) and observed (black dots) events for the four different signal regions. The hashed area represents the total uncertainty on the background prediction.

## 4.9 Interpretation

In order to interpret the result of the search in the context of supersymmetric models with almost mass degenerate charginos and neutralinos, sources of systematic uncertainties on the number of selected signal events must be identified and quantified. The interpretation will then be done with statistical methods that allow for the exclusion of parts of the supersymmetric parameter space on a 95% confidence level. The signal models used for this interpretation are listed in Section 4.4.2.

### 4.9.1 Systematic uncertainties of simulated signal samples

The systematic uncertainties on the number of signal events in the four signal regions are caused by uncertainties in the generation and simulation of signal events and the integrated luminosity of the considered data.

All systematic uncertainties are estimated for each signal model (cf. Section 4.4.2) and each search bin separately. In the following, the sources of systematic uncertainties are discussed and the range of the corresponding uncertainty is given.

#### Theoretical cross section

The theoretical cross sections of  $\tilde{\chi}_1^\pm \tilde{\chi}_1^\mp$  and  $\tilde{\chi}_1^\pm \tilde{\chi}_1^0$  production at a centre-of-mass energy of 8 TeV are taken from [?, ?]. The corresponding theoretical uncertainties range between 4.5 – 12.1%.

#### Luminosity

The integrated luminosity recorded at CMS during the year 2012 is measured by counting of pixel clusters during the crossing of two bunches (zero-bias event). A detailed explanation of this method and the corresponding total uncertainty of 2.6% can be found in [?].

#### Simulation of initial state radiation

Initial state radiation (ISR) affects the transverse momentum distribution of the 2-particle system,  $p_T(p_1^\mu + p_2^\mu)$ , in a 2-body decay. Differences between data and simulation of ISR

are taken into account by reweighting the simulated events, such that the simulated transverse momentum distribution matches the measured distribution in data. The weights and associated systematic uncertainties are determined in [?] by comparing simulated and observed  $p_T$  distributions of  $Z$  and  $t\bar{t}$  events. These weights are applied to the simulated  $\tilde{\chi}_1^\pm \tilde{\chi}_1^\mp$  and  $\tilde{\chi}_1^\pm \tilde{\chi}_1^0$  events. To account for the systematic uncertainties on the reweighting procedure, the event weights are varied up and down by up to 25% according to [?] depending on the transverse momentum of the  $\chi_1 \chi_2$  system. The resulting uncertainty on the ISR simulation is between 9.2 – 12.6%.

### Simulation of the trigger efficiency

The HLTMonocentralPFJet80\_PFMETnoMu105\_NHEF0p95 trigger with the higher MET threshold of 105 GeV active in Run C and Run D during 2012 was not available in the simulated signal samples. It is therefore emulated using HLT trigger information. More details on the emulation of this trigger can be found in Appendix ??.

The trigger uncertainty is assessed by comparing data-simulation differences of the trigger efficiency. This uncertainty has been quantified within [?, ?] by comparing simulated and measured trigger turn-on curves and determining weights for simulated events such that simulated and observed turn-on curves are compatible. In Fig. 4.45, the trigger turn-on curves, measured in data and in simulation are depicted. It can be seen, that the turn-on in data is slower compared to simulation. The ratio of Fig. 4.45 determines the event weights for the estimation of the systematic uncertainties. These event weights are applied on the simulated signal samples in this analysis and lead to changes in the signal

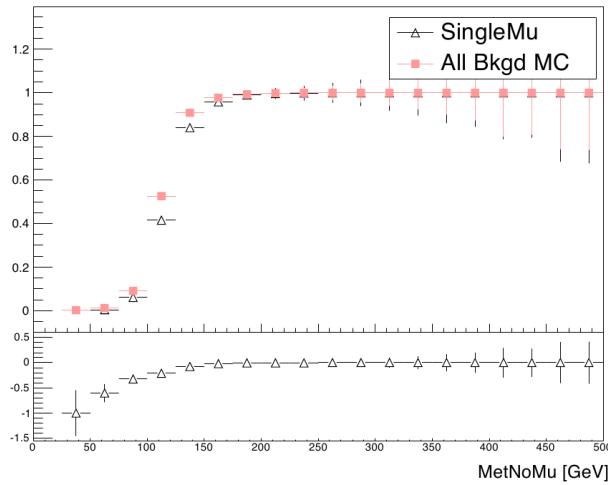


Figure 4.45: Trigger turn-on curves in data (black triangles) and simulation (red squares) as a function of PF  $\cancel{E}_T$  (not taking into account the energy of the muons).

prediction of 1.9 to 4.4%, which are taken into account as systematic uncertainties.

### Jet energy scale

The transverse momentum of all jets is corrected for non-uniformities in the energy response as a function of the jet  $\eta$  and  $p_T$  and for data-simulation differences. The uncertainty on the jet energy scale (JES) is neatly described and quantified in [?]. It arises from uncertainties on the measured jet response in data including jet fragmentation, jet flavor composition, etc.. The JES correction is applied as a multiplicative factor on each jet’s transverse momentum contained in an event. The corresponding systematic uncertainty is assessed by an up- and downward variation of the correction factor within  $1\sigma$ . The resulting uncertainties are of minor importance and range between 0.4 – 3.1%.

### Jet transverse-momentum resolution

The jet transverse-momentum resolution (JER) is smaller in simulation than in measured data (see Part 5). In order to take these differences into account, the simulated jet transverse-momentum response is smeared to match the measured response. The systematic uncertainty on the smearing factors is estimated in [?, ?]. It covers the uncertainty on JER in data, including the JES uncertainty, uncertainties arising from out-of-cone showering etc. [?, ?]. The resulting uncertainty on the signal efficiency in this study is between 0.1 – 2.0% and therefore almost negligible.

### Simulation of the parton distribution functions

The parton distribution function (PDF) used for the simulation of proton-proton collisions is provided by the CTEQ group [?] (see Section 3.4 for more information about PDFs). In [?], a detailed description of the determination of a parton distribution function and its uncertainties is given. Practically, the estimation of the PDF uncertainty is done by the application of 44 different sets of event weights which take into account 22 different sources of uncertainties [?, ?] (up and down variations lead to a factor of 2). The sources correspond inter alia to uncertainties in the single distributions of gluons, up/down-quarks, etc, with the gluon distribution being by far the largest source of uncertainty. The resulting uncertainties on the signal efficiency for this search are between 2.6 – 6.8%.

### Pileup reweighting

The distribution of the number of pileup interactions in simulation is reweighted to match the measured distribution in data. The distribution of the number of pileup interactions in data is estimated by the luminosity of each bunch-crossing times the proton-proton inelastic cross section which is 69.4 mb [?]. The uncertainty on the number of interactions

thus consists of the uncertainty on the luminosity and the uncertainty on the cross section. To cover both sources, a variation of the inelastic cross section by plus/minus 5% is done according to the recommendation by [?].

For most of the signal models and signal regions, the signal efficiency is only affected by less than 1% by the pileup reweighting uncertainty. If the statistical precision of the signal prediction in a specific search bin is low, the uncertainty can become significantly larger. However, the search sensitivity is always driven by search bins with high signal content so that large values of this uncertainty have no effect on the overall search sensitivity.

### Simulation of the calorimeter isolation

The uncertainty on the simulation of the calorimeter isolation  $E_{\text{calo}}^{\Delta R < 0.5}$  is estimated by comparing simulated and measured selection efficiencies of  $E_{\text{calo}}^{\Delta R < 0.5} < 5 \text{ GeV}$  in the fake enriched control sample  $\text{CR}_{I_{\text{as}}}^{\text{fake}}$ . The fake enriched control region is well suited for this estimation, because fake tracks are not correlated to the energy deposits in the calorimeters. The selection efficiency in data is higher than in simulation in both  $p_T$  bins of  $30 - 50 \text{ GeV}$  and  $50 - \infty \text{ GeV}$ . This difference between data and simulation is taken into account as systematic uncertainty, resulting in uncertainties of 12.1% and 3.0%.

### Simulation of missing middle/inner hits

The uncertainty on the simulation of the number of missing inner and middle hits is assessed by comparing the probability in simulation and data of passing the selection requirements of  $N_{\text{miss}}^{\text{middle/inner}} = 0$  of a candidate track in the muon-veto inverted control region. This control region is particularly suitable because muons are not expected to have intrinsic sources of missing hits, as e.g. pions or electrons have. Pions can interact nuclearly with the tracker material and electrons can have sizable radiative losses, such that both can change direction or don't deposit energy in a tracker layer. For muons, on the other hand, sources of missing inner and middle hits are mainly algorithmic [?, ?], making them very similar to the algorithmic sources of missing inner/middle hits for chargino tracks.

The uncertainty is estimated as the observed difference of the cut selection efficiency of  $N_{\text{miss}}^{\text{middle/inner}} = 0$  in data and simulation. The selection efficiency is always higher in simulation, resulting in systematic uncertainties of around 3.5% for the simulation of  $N_{\text{miss}}^{\text{inner}} = 0$  and around 2.2% for  $N_{\text{miss}}^{\text{middle}} = 0$ . The uncertainties are of very similar size in the signal regions with different  $p_T$ . No  $I_{\text{as}}$  dependence is considered.

### Simulation of $I_{\text{as}}$

An uncertainty on the simulation of  $I_{\text{as}}$  needs to be estimated in order to account for possible data-simulation differences for highly ionising particles. The estimation of the  $I_{\text{as}}$  uncertainty is done following the methodology in [?, ?]. The  $I_{\text{as}}$  uncertainty can be assessed by comparing data and simulation differences of slow protons. Slow protons are highly ionising and can thus be used to determine the uncertainty in the high  $I_{\text{as}}$  region.

In order to select slow protons, high quality tracks with a momentum smaller than 2.5 GeV are selected. The  $I_{\text{as}}$  versus momentum distribution for the selected tracks is shown in Fig. 4.46. The kaon and proton line are visible in both datasets. The deuteron line is only visible in data, as deuteron's are not simulated. Two different slices in the momentum are extracted where the proton line is contained:  $p$  between 0.80 – 0.85 GeV and 0.95 – 1.00 GeV. A Gaussian function is fitted to the proton peak and the maximum difference of the mean of the fitted Gaussian between simulation and observed data is taken as systematic uncertainty. The  $I_{\text{as}}$  distribution for the two momentum ranges with the Gaussian fit is depicted in Fig. 4.47. The systematic uncertainty is estimated to a value of 6%.

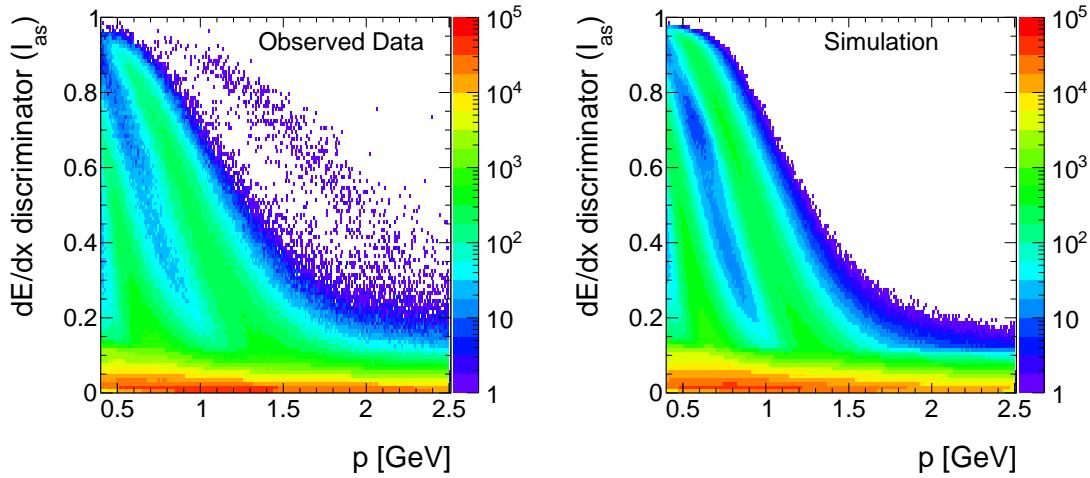


Figure 4.46:  $I_{\text{as}}$  versus momentum for good quality tracks with at least eight hits in observed data (left) and simulation (right).

### Simulation of the track reconstruction efficiency

One final source of uncertainty is the simulation of the track reconstruction efficiency. Possible differences of the reconstruction efficiency in simulation and data can lead to a different signal acceptance. Differences in the track reconstruction efficiency are especially

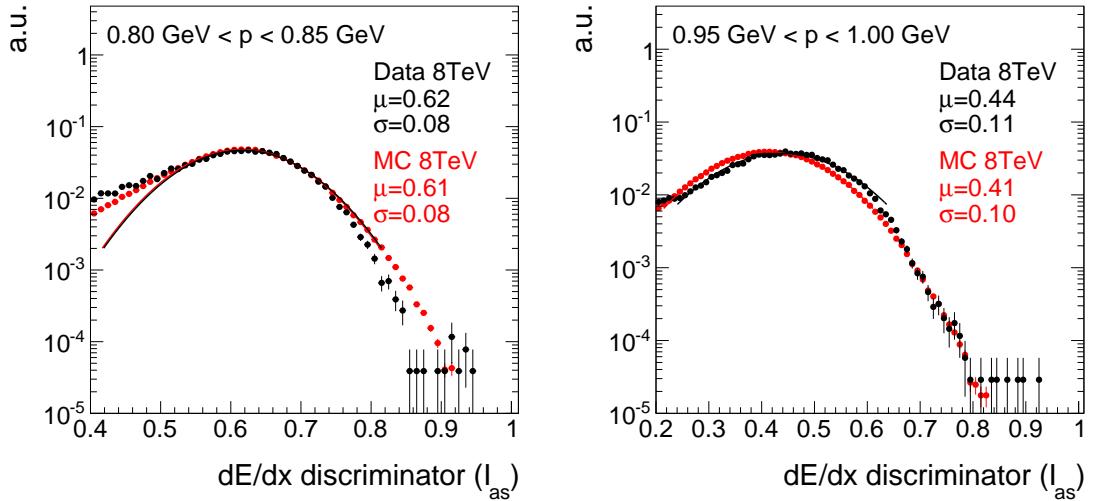


Figure 4.47:  $I_{as}$  distribution for slow protons in simulation and observed data for a momentum range of  $0.80 - 0.85 \text{ GeV}$  (left) and  $0.95 - 1.00 \text{ GeV}$  (right). For the momentum range of  $0.80 - 0.85 \text{ GeV}$ , the proton line is contained between  $I_{as}$  values of  $0.4 - 0.8$ , whereas for the momentum range of  $0.95 - 1.00 \text{ GeV}$ , the proton line  $I_{as}$  lies between  $0.2 - 0.6$ .

expected for short tracks. Therefore, a worst case estimation is done, comparing the track reconstruction efficiency in data and simulation for tracks with only three hits.

In simulation and observed data, well reconstructed muon tracks are selected and all hits after the third hit are removed. Afterwards the full track reconstruction is performed again. The relative difference of this track reconstruction efficiency in data and simulation is taken as systematic uncertainty. The track reconstruction efficiency is higher in simulation than in data and results in uncertainties between  $4.6 - 6.0\%$ .

### Summary of systematic uncertainties on the simulated signal samples

All systematic uncertainties are estimated for all simulated signal samples and in each of the four signal regions. An overview of the range of the uncertainties is given in Table 4.19.

In order to avoid an overestimation of the systematic uncertainties due to limited sizes of the samples (especially for low lifetimes like 1 cm), the corresponding signal sample with longer lifetime (100 cm) is used instead for determining the systematic uncertainty. This is possible for uncertainty sources, where the size is not affected by the lifetime of the chargino, including ISR, trigger efficiency, JES, JER, and PDF uncertainties.

It can be seen, that major uncertainties are the simulation of the initial state radiation, of the calorimeter isolation, and of  $I_{as}$ . The high maximum value of the pileup uncertainty

is caused by limited statistical precision.

The systematic uncertainties on the simulated signal samples are considered as fully correlated among the four signal regions.

Table 4.19: Ranges of systematic uncertainties on the simulated signal samples. Min and Max correspond to variations between different signal samples and search bins.

Uncertainty	Min [%]	Max [%]
Theoretical x-section	4.5	12.1
Luminosity	2.6	2.6
Simulation of ISR	9.2	12.6
Simulation of trigger efficiency	1.9	4.4
JES	0.4	3.1
JER	0.1	2.0
Simulation of PDF	2.6	6.8
Pileup reweighting	0.0	16.0
Simulation of calorimeter isolation	3.0	12.1
Simulation of missing middle hits	2.2	2.2
Simulation of missing inner hits	3.3	3.7
Simulation of $I_{\text{as}}$	6.0	6.0
Simulation of track reconstruction efficiency	4.6	6.0

## 4.9.2 Statistical Methods/ Limit setting

This section is a small interlude to briefly introduce the methods and techniques that are used to exclude beyond-SM models with the results of this search. For a detailed and pedagogical introduction to the methods, the reader is referred to [?].

In this analysis, the exclusion of the underlying theoretical model is achieved with the  $\text{CL}_s$  method [?, ?, ?]. A model is considered as excluded at a 95% confidence level if  $\text{CL}_s$  is smaller than 5%. The  $\text{CL}_s$  method was developed for the Higgs searches at LEP in order not to overestimate the exclusion power of a result if an underfluctuation of the background expectation occurs.  $\text{CL}_s$  is defined as the confidence level of the background

plus signal hypothesis divided by the confidence level of the background only hypothesis

$$\text{CL}_s = \frac{\text{CL}_{s+b}}{\text{CL}_b}. \quad (4.25)$$

The confidence level CL is defined as the probability of obtaining less than or equal the number of observed events  $P(n \leq n_{\text{obs}})$  for a given background (or background+signal) hypothesis. In general, the considered distribution can refer to any test statistics  $q$  and is not necessarily the distribution of the number of events. For Poissonian statistics it leads to the following expressions for  $\text{CL}_{s+b}$  and  $\text{CL}_b$  for one signal region

$$\begin{aligned} \text{CL}_{s+b} &= \sum_0^{n_{\text{obs}}} q_\mu = \text{Poisson}(n \leq n_{\text{obs}} | \lambda = b + \mu \cdot s), \\ \text{CL}_b &= \sum_0^{n_{\text{obs}}} q_{\mu=0} = \text{Poisson}(n \leq n_{\text{obs}} | \lambda = b), \end{aligned} \quad (4.26)$$

where  $\lambda$  is the mean of the Poisson distribution and the signal strength  $\mu$  is the measure for the size of the signal cross section.

Systematic uncertainties are included by varying the background expectation  $b$  and the signal expectation  $\mu \cdot s$  according to a predefined probability density function (pdf). For one Gaussian distributed source of systematic uncertainty on the background, this leads to the following expressions for the test statistics of  $q_\mu$  and  $q_{\mu=0}$

$$\begin{aligned} q_\mu &= \text{Poisson}(n | \lambda = b \cdot (1 + \delta_b) + \mu \cdot s) \cdot \text{Gauss}(\delta_b | \text{mean} = 0, \sigma = \sigma_b), \\ q_{\mu=0} &= \text{Poisson}(n | \lambda = b \cdot (1 + \delta_b)) \cdot \text{Gauss}(\delta_b | \text{mean} = 0, \sigma = \sigma_b), \end{aligned} \quad (4.27)$$

These expressions can be generalised for more than one signal region and more than one systematic uncertainty [?]. In case of multiple signal regions, the distribution of the systematic uncertainties becomes a multi-dimensional probability density function that takes the covariance matrix of the systematic uncertainties in different signal regions into account. In order to get rid of the nuisance parameters  $\delta_b$  and  $\delta_s$  either an integration over  $\delta_b$  and  $\delta_s$  or a maximisation with respect to  $\delta_b$  and  $\delta_s$  can be performed.

In this analysis, the procedure for limit setting follows [?]. Instead of the number of observed events, the profile likelihood ratio

$$q(n, \mu) = -2 \ln \frac{\mathcal{L}(n | \mu, \hat{\delta}_b^\mu, \hat{\delta}_s^\mu)}{\mathcal{L}(n | \hat{\mu}, \hat{\delta}_b, \hat{\delta}_s)}, \quad \text{with the constraint } 0 \leq \hat{\mu} \leq \mu \quad (4.28)$$

is used as the test statistics for determining  $\text{CL}_{s+b}$  and  $\text{CL}_b$  ( $\mu = 0$ ).  $\mathcal{L}$  refers to the

likelihood function

$$\mathcal{L}(n|b, s, \delta_b, \delta_s) = \text{Poisson}(n|\lambda = b \cdot (1 + \delta_b) + \mu \cdot s(1 + \delta_s)) p(\delta_b) p(\delta_s). \quad (4.29)$$

that already appeared in Eq. (4.26). The parameters  $\hat{\mu}$ ,  $\hat{\delta}_b$  and  $\hat{\delta}_s$  are fixed to the values that maximise the likelihood with respect to the observed data. Determining  $\text{CL}_{s+b}$  and  $\text{CL}_b$  requires knowing the probability distributions of  $q$  for the signal+background hypothesis (depending on  $\mu$ ) and the background-only hypothesis ( $\mu = 0$ ). These distributions are determined by toy pseudo data with fixed values for  $\delta_s$  and  $\delta_b$ . Finally, the signal strength  $\mu$  is adjusted until  $\text{CL}_{s^+}$  calculated based on  $q$  - equals 0.05%. Signal models are considered as excluded on a 95% confidence level, if the signal strength is larger than the one found by the limit setting procedure.

In this search, the systematic uncertainties on the background and the signal yields as well as the statistical uncertainty on the fake background are modelled with log-normal distributions, whereas the statistical uncertainties on the leptonic background are modelled using gamma distributions. A log-normal distribution is used instead of a normal distribution to ensure that the prediction cannot become negative. The gamma distribution is well suited for statistical uncertainties arising from very limited statistical precision in control regions or in simulated samples that are used for the background estimation [?].

Correlations between systematic uncertainties on the background expectation in different search bins are assumed as shown in Table 4.20. The systematic uncertainties on the

Table 4.20: Correlation of systematic and statistical uncertainties between the four different signal regions. Statistical uncertainties include uncertainties arising from the limited size of control regions and the simulated samples.

	Fakes	Taus	Electrons	Muons
Statistical uncertainty	0% correlated	100% for bins with same $I_{\text{as}}$	0% correlated	100% for bins with same $I_{\text{as}}$
Leptonic scale factor uncertainty	-	100% for bins with same $I_{\text{as}}$	100% for bins with same $I_{\text{as}}$	100% for bins with same $I_{\text{as}}$
Fake rate uncertainty	100% for bins with same $I_{\text{as}}$	-	-	-
$I_{\text{as}}$ uncertainty	0% correlated	100% for bins with same $p_T$	100% for bins with same $p_T$	100% for bins with same $p_T$

expected signal yields are considered fully correlated across search bins.

The exclusion limits are derived according to the above presented methodology using the *Combine* framework [?] which was developed for the Higgs searches at CMS.

### 4.9.3 Exclusion limits

The presented search for highly ionising, short tracks is interpreted in the context of SUSY models with almost mass degenerate wino-like charginos and neutralinos. As explained in the previous section, the exclusion is done with the help of the  $\text{CL}_s$  method. Two direct production channels are taken into account: chargino pair production and chargino neutralino production. The corresponding cross sections can be found in Table 4.2.

In total, 37 different lifetimes from  $c\tau = 1 - 10000 \text{ cm}$  for each mass point ( $100 - 600 \text{ GeV}$  in steps of  $100 \text{ GeV}$ ) are considered. Four exemplary exclusion limits are shown in Fig. 4.48, the full set of exclusion limits can be found in Appendix ??.

The upper 95% confidence level (CL) limit on the signal cross section is strongest for lifetimes between  $10 - 100 \text{ cm}$ . For lower lifetimes a sizable fraction of the charginos already decay before reaching the tracker. For longer lifetimes, the cross section upper limit gets weaker again because the charginos start to be reconstructed as muons and do not pass the muon veto. Also, the  $E_{\text{calo}}^{\Delta R < 0.5}$  requirement rejects these charginos with higher efficiency.

Due to the falling spectrum of the chargino production cross section, charginos with lower masses are more effectively excluded than charginos with higher masses. A 2-dimensional exclusion limit in the chargino lifetime-mass parameter space is shown in Fig. 4.49.

Charginos with masses of  $100 \text{ GeV}$  can be excluded down to a lifetime of  $c\tau = 2 \text{ cm}$ . Charginos with a higher mass of  $500 \text{ GeV}$  are excluded for lifetimes between  $c\tau = 70 - 500 \text{ cm}$ .

Since the lifetime of a wino-like chargino is determined by the mass splitting between  $m_{\tilde{\chi}_1^\pm}$  and  $m_{\tilde{\chi}_1^0}$ , it is possible to express the lifetime of the chargino as a mass gap  $\Delta m_{\tilde{\chi}_1^\pm \tilde{\chi}_1^0}$  between the chargino and the lightest neutralino. The correspondence between lifetime and mass gap is taken from [?], where the decay width of  $\tilde{\chi}_1^\pm \rightarrow \tilde{\chi}_1^0 \pi^\pm$  is expressed in terms of chargino, neutralino, and pion mass. Thus, the mass gaps that are considered are bounded by the pion mass of  $\sim 140 \text{ MeV}$ . The corresponding 2d exclusion limit can be found in Fig. 4.50. It can be seen that this search is sensitive to mass splittings between  $\sim 140 \text{ MeV} - 210 \text{ MeV}$ .

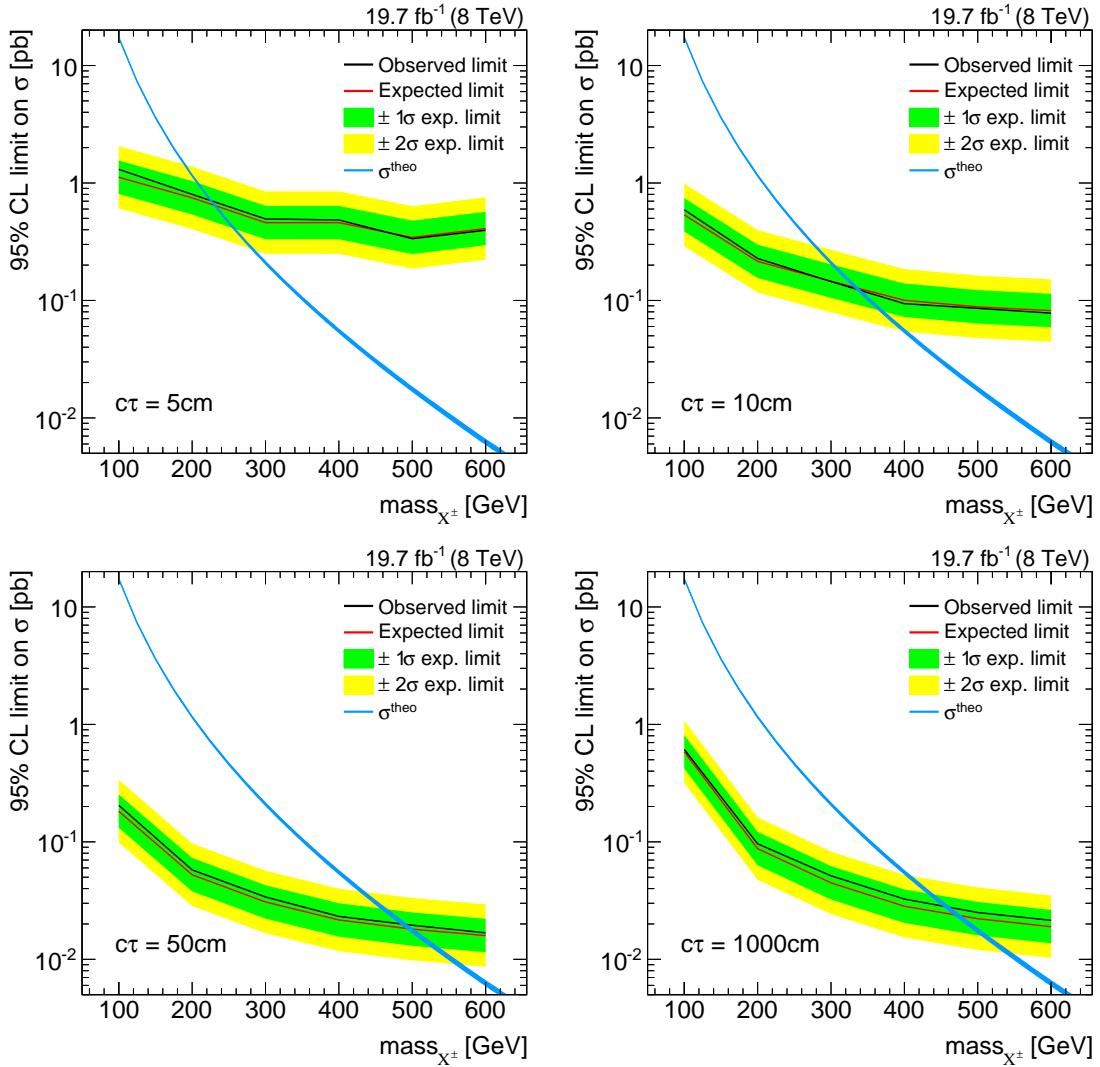


Figure 4.48: Four different  $CL_s$  exclusion limits for charginos with mean lifetimes of 5 cm (top left), 10 cm (top right), 50 cm (bottom left), 1000 cm (bottom right). The red line depicts the expected 95% confidence level (CL) upper cross-section limit with the  $1-\sigma$  (green band) and  $2-\sigma$  (yellow band) intervals. The black line is the observed limit. The signal cross section is depicted as a blue line. SUSY models can be excluded at 95% CL if the signal cross section is at least as large as the 95% CL observed upper limit on the cross section.

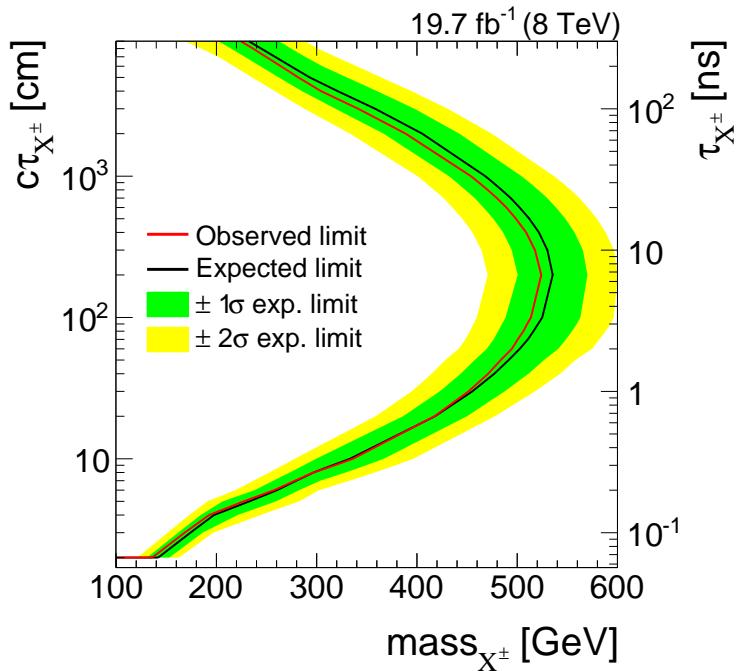


Figure 4.49: Excluded regions in the mass versus lifetime space. All excluded models are located left of the contour line. The red line depicts the expected 95% CL upper cross-section limit with the  $1-\sigma$  (green band) and  $2-\sigma$  (yellow band) intervals. The black line is the observed limit.

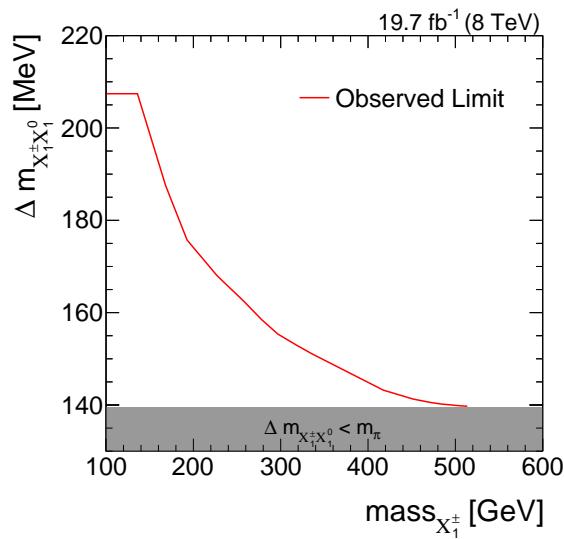


Figure 4.50: Excluded parameter region at 95% CL for wino-like charginos and neutralinos depending on the chargino mass and the mass splitting between  $\tilde{\chi}_1^\pm$  and  $\tilde{\chi}_1^0$ ,  $\Delta m_{\tilde{\chi}_1^\pm \tilde{\chi}_1^0}$ . All SUSY models between the red line and the grey area are excluded.

#### 4.9.4 Comparison to the search for disappearing tracks

The presented exclusion limits confirm the exclusion from the search for disappearing tracks [?] with slight improvements in the low lifetime region. The comparison of the two searches is shown in Fig. 4.51.

For charginos with a lifetime of  $\tau = 0.07$  ns ( $c\tau = 2.1$  cm), the observed limit of this search improves the limits derived in [?] by  $\sim 35$  GeV in chargino mass, for a lifetime of  $\tau = 0.4$  ns ( $c\tau = 12.0$  cm) by  $\sim 25$  GeV. For SUSY models with long chargino lifetimes the here presented search shows a higher exclusion power. The weaker exclusion for long lifetimes in [?] is caused by the additional selection cut on the number of missing outer hits,  $N_{\text{lost}}^{\text{outer}} \geq 3$ .

Since the disappearing track analysis and the here presented analysis use the same physics objects (tracks) and are interpreted in terms of the same signal models, one would expect a very high overlap between the selected signal events. Interestingly, the two searches show a rather small overlap of selected signal events of around 20% for short lifetimes of  $c\tau = 5$  cm. The event overlap between simulated signal events, that pass the selection from [?],  $N_A$ , and the selection used in this analysis,  $N_B$ , can be estimated by

$$\rho_{\text{corr}} = \frac{N_{A \cap B}}{N_{A \cup B}} = \frac{N_{A \cap B}}{N_A + N_B - N_{A \cap B}}.$$

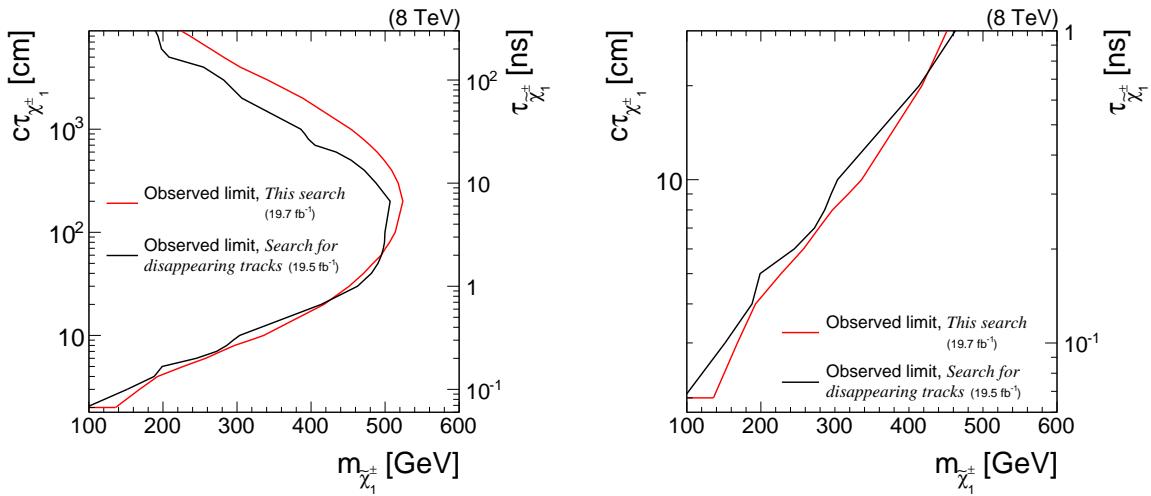


Figure 4.51: Comparison of the excluded regions in the mass versus lifetime space in this analysis (red line) and the search for disappearing tracks [?] (black line). The right figure is a zoom on the low lifetime region. All SUSY models left of the lines are excluded.

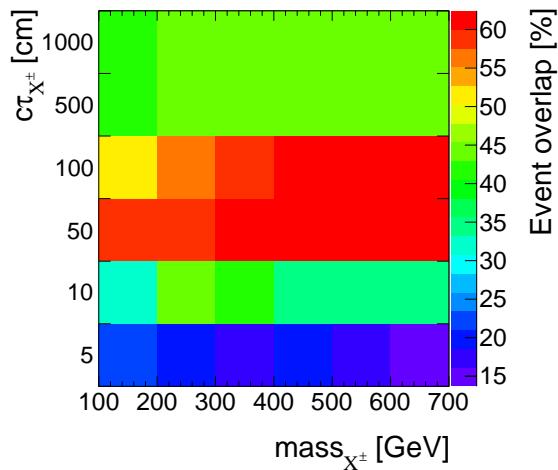


Figure 4.52: The event overlap between simulated signal events, that pass the selection from [?] and the selection used in this analysis for different signal models. The correlation is determined using only the signal region with the highest sensitivity of this analysis.

In order to avoid an over- or underestimation of the overlap, only the most sensitive signal region from this search is included in  $N_B$ . The signal event overlap for signal models with chargino masses between  $100 - 600$  GeV and lifetimes between  $5\text{ cm} - 1000\text{ cm}$  is depicted in Fig. 4.52. It can be seen that the event overlap for intermediate lifetimes of around  $100\text{ cm}$  is around 60% and decreases for shorter lifetimes to small overlaps of around 15 – 20%. Additionally, the two events that were observed in data by [?] in their signal region are not contained in any of the signal regions in the here presented analysis. Thus, this analysis constitutes a rather independent confirmation of the exclusion limits derived in [?], especially for short chargino lifetimes.

## 4.10 Conclusion and outlook

The here presented search for highly ionising, short tracks is motivated by supersymmetric models with almost mass-degenerate wino-like charginos  $\tilde{\chi}_1^\pm$  and neutralinos  $\tilde{\chi}_1^0$ . Such scenarios can have interesting astrophysical implications [?] and occur naturally in supersymmetric models, if the wino mass parameter is smaller than the bino and higgsino mass parameters.

The presented analysis targets SUSY models with intermediate chargino lifetimes. This is achieved by searching for isolated, high  $p_T$  tracks that are highly ionising. No requirement on the number of tracker hits is enforced, thus, possibly very short tracks from early decaying charginos are included in this analysis. It thus extends the search for disappearing tracks [?] by the inclusion of the variable  $dE/dx$  and the loosening of the requirement on the number of hits in the tracker ( $N_{\text{hits}} \geq 3$ ) that leads to a strong suppression of signal events for low chargino lifetimes (cf. Fig. 4.8). It is thereby the first analysis at CMS that studies disappearing tracks with down to three hits.

In order to increase the search sensitivity with respect to shorter lifetimes, energy information from the pixel silicon tracker is taken into account. For this purpose, a dedicated pixel energy calibration was carried out within this thesis to ensure stable energy measurements over time and across pixel modules. This is thus the first time that an analysis at CMS makes use of energy information from the pixel tracker. By adding pixel energy information, the discrimination power of  $dE/dx$  is substantially increased (cf. Fig. 4.18).

The Standard Model background is mainly estimated with data-based techniques. The main background to this search is arising from fake tracks, i. e. tracks that are reconstructed out of the tracker hits of more than one particle. Fake tracks are typically short and can have large values of  $I_{\text{as}}$ , thus showing a very signal-like signature in the detector. The background contribution by leptons that are passing the lepton veto is very small and in most of the signal regions almost negligible.

In the current analysis, the background is estimated at 19 and 24 events in the low  $I_{\text{as}}$  signal regions and 2.5 and 2.6 events in the high  $I_{\text{as}}$  regions. This background estimate is confronted with collision data recorded during the year 2012 at the CMS experiment at a centre-of-mass energy of 8 TeV. No evidence for physics beyond the Standard Model is found. Thus, the absence of any deviation from the Standard Model prediction is used to constrain the supersymmetric parameter space. Wino-like charginos are excluded down to lifetimes of  $c\tau = 2$  cm for  $m_{\tilde{\chi}_1^\pm} = 100$  GeV. For high mass scenarios of  $m_{\tilde{\chi}_1^\pm} = 500$  GeV, the excluded lifetime ranges between  $c\tau = 70 - 500$  cm. This confirms the parameter exclusion limits of the search for disappearing tracks [?]. Interestingly, the signal regions of the here presented search and the search from [?] show a rather small overlap for short chargino lifetimes. Therefore, this analysis yields a complementary result with respect to the search for disappearing tracks [?]. In summary, the exclusion of SUSY models with respect to earlier searches could be independently confirmed and improvements in the exclusions of around 10 – 40 GeV in chargino mass in the low lifetime region are achieved.

While this analysis is able to exclude many SUSY models with intermediate lifetime

charginos, there are several promising avenues for even enhancing the search sensitivity.   
FIXME: Maybe the most obvious is the combination of the results with the search for disappearing tracks. Since both analyses have a rather small event overlap, a combination is expected to push the exclusion further down. First, since the sensitivity of the current analysis is mainly limited by large systematic uncertainties originating from low statistical precision in the simulated datasets, simulating more events could significantly improve the search sensitivity. This strategy is however technically challenging, since storage capacity limits were already reached within the current analysis. Still, reducing this systematic uncertainty will be one of the main tasks for future research.

Second, even though this search already features low background, a further background suppression is desirable. However, the impact on the search sensitivity will be limited because of the high relative Poisson error on low background predictions. For instance - neglecting systematic uncertainties - a reduction of the number of background events by 50% from 2 to 1 reduces the signal yield required for a  $5\sigma$ -discovery by around 8%, whereas a 50% reduction of expected background events from 200 to 100 reduces the required signal yield by 26%.

Thus, in order to improve the here presented analysis, the focus should be on the other determinant of search sensitivity: the signal acceptance. First and foremost, it is important to lower the signal losses due to trigger requirements. For this purpose, a dedicated track trigger would be beneficial, especially if a future upgrade would make tracking information available on level one.

Furthermore, an implementation of a dedicated track reconstruction algorithm optimised for short tracks could increase the reconstruction efficiency of possible chargino tracks, which is currently  $\sim 20\text{--}40\%$  for chargino tracks with 3 – 4 hits. Additionally, a track reconstruction optimised for the reconstruction of soft particles that are not produced in the primary vertex could allow for a reconstruction of the Standard Model decay products of charginos, thereby enabling a better discrimination against Standard Model background.

In summary, the here presented analysis explored a new path for searching for long-lived charginos decaying early inside the detector. It is the first analysis that incorporates reconstructed tracks down to three hits. Furthermore, for the first time, ionisation losses are measured taking energy information from the pixel silicon tracker into account.

As argued, further improvements can allow for accessing new, unexplored SUSY models with long-lived charginos. Additionally, a search in collisions at a centre-of-mass energy of 13 TeV with increased cross sections makes the exploration of SUSY models with higher chargino masses possible. Since  $dE/dx$  is much more discriminating for high masses, the inclusion of  $dE/dx$  in this analysis will become even more powerful.



## **Part 5**

### **Measurement of the jet transverse-momentum resolution**



## 5.1 Introduction

The determination and quantification of the quality of the jet transverse-momentum measurement is of crucial interest for many analyses with jet final states, e. g. the measurement of the dijet cross section [?] or  $t\bar{t}$  production cross sections [?]. Also searches for physics beyond the Standard Model with missing transverse momentum,  $\cancel{p}_T$ , in the final state rely on a good understanding of  $\cancel{p}_T$  originating from wrongly measured jets [?, ?, ?]. It is therefore important to calibrate the measured jet transverse momentum in order to eliminate biases in the measured jet  $p_T$ . Thus, each jet is subject to a calibration procedure which is explained in detail in Section 3.3.2.3. FIXME Furthermore, it is necessary to determine the jet transverse-momentum resolution, i. e. the detector resolution of the jet  $p_T$  measurement. For analyses relying on information from simulation, it is necessary to correct the simulated resolution to the resolution actually present in data. Therefore, scale factors will be presented to adjust the resolution in simulation to the resolution of the real detector.

In the following sections, a data-based method to measure the jet  $p_T$  resolution in  $\gamma +$ jet events will be presented. A similar method was already applied in earlier analyses [?, ?] of 7 TeV data. It is further developed here and applied to  $19.7 \text{ fb}^{-1}$  of  $\sqrt{s} = 8 \text{ TeV}$  data.

The method is based on the transverse-momentum balance in the  $\gamma +$ jet system. It takes advantage of the high resolution of the electromagnetic calorimeter and hence the excellent measurement of the photon momentum. Without initial and final state radiation, the photon and the jet are balanced in the transverse plane. Thus, measuring the photon  $p_T$  with high accuracy leads to an accurate estimate of the true jet transverse momentum and offers a possibility to quantify the resolution of jet  $p_T$  measurements.

## 5.2 General approach of the resolution measurement using photon+jet events

The quality of the jet transverse-momentum measurement is determined by the jet transverse-momentum response  $\mathcal{R}$ , i. e. by the ratio of the reconstructed to the true jet transverse

momentum

$$\mathcal{R} = \frac{p_T^{\text{reco. jet}}}{p_T^{\text{true jet}}}.$$
 (5.1)

Because of the limited detector resolution, the reconstructed jet  $p_T$  in general differs from the true jet  $p_T$  resulting in a distribution of  $\mathcal{R}$ . In Fig 5.1, an exemplary distribution of  $\mathcal{R}$  in simulation (where the true jet  $p_T$  corresponds to the generator-level jet  $p_T$ )<sup>1</sup> is depicted (blue dots). It can be seen, that the core of the response distribution shows a typical Gaussian behaviour whereas the tails deviate from that functional form. Physical reasons for the pronounced low response tail are, inter alia, semi-leptonic decays of heavy quarks (c- and b-quarks) where the neutrino cannot be detected and the reconstructed transverse momentum of the jet is too small. This effect is visible in Fig. 5.1 which depicts the response distribution including all jet flavours (blue), as well as the contribution by c- and b-quarks (red), which is the main contribution to the pronounced left tail. Some instrumental effects, such as a non-linear response of the calorimeter, inhomogeneities of the detector material and electronic noise can contribute to both tails, others, like dead calorimeter channels only contribute to the left tail [?].

In order to avoid the coverage of non-Gaussian tails, the jet transverse-momentum resolution (JER)<sup>2</sup> is determined using only the core of the response distribution. The res-

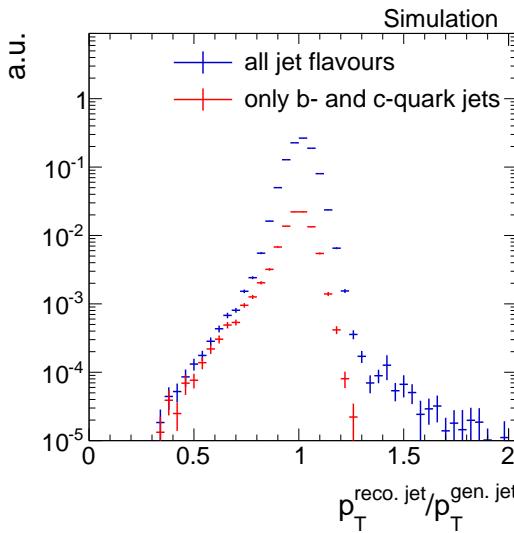


Figure 5.1: Normalised number of events over  $p_T^{\text{reco. jet}}/p_T^{\text{gen. jet}}$  from a simulated  $\gamma + \text{jet}$  sample. The response distribution containing all jet flavours is depicted in blue whereas the contribution by c- and b-quark jets is shown in red.

<sup>1</sup>For each reconstructed jet, the underlying generator-level jet is found by matching the closest generator-level jet to the reconstructed jet with a maximal value of  $\Delta R_{\max} = 0.25$ .

<sup>2</sup>This abbreviation is a historical relic from experiments where the momentum of jets were only measured in the calorimeters and therefore JER referred to jet energy resolution.

olution is thus defined as the standard deviation of the 99% truncated response histogram divided by the mean [?]:

$$\sigma_{\text{JER}} = \frac{\sigma_{99\%}}{\mu_{99\%}}.$$

The division by the mean aims at making the resolution independent of the absolute scale (= mean) of the response distribution. The mean is referred to as the jet transverse-momentum scale (JES) [?].

The determination of the 99% range of the histogram is done in several steps. First the mean of the core is found via a Gaussian fit to the histogram in a  $2\sigma$  range<sup>3</sup>. This procedure is done in three iteration steps. Then, a symmetric interval around this mean is determined with its integral equal to 99% of the integral of the full histogram.

As noted earlier, in simulated events, the response distribution can be evaluated by the ratio of the reconstructed to the generator-level jet transverse momentum. A determination of the resolution in data, however, has to rely on a different approach.

The main idea of a resolution measurement using  $\gamma + \text{jet}$  events is based on the transverse momentum balance of the  $\gamma + \text{jet}$  system and the excellent electromagnetic calorimeter resolution.

In Fig 5.2, all tree-level processes contributing to an event topology with one photon and one jet in the final state are depicted. Due to momentum conservation, the jet and the photon are back to back in the transverse plane, and therefore,  $\vec{p}_T^\gamma \approx -\vec{p}_T^{\text{true jet}}$ . Because of the high resolution of the electromagnetic calorimeter, photon energies can be measured very accurately (with a resolution between 1.4% and 3.8% in the barrel region for  $\sqrt{s} = 8 \text{ TeV}$  data [?]). Thus, they can serve as an excellent estimator for the true jet transverse-momentum.

Unfortunately, such clean events are very rare processes, and usually, the momentum balance is spoiled by further jet activity from initial and final state radiation (see Fig. 5.3). In order to select events that are balanced to a large extent, a lower bound on the angular distance in the transverse plane between the photon and the jet with the highest transverse momentum (leading jet) is required:  $\Delta\Phi(\text{1}^{\text{st}} \text{ jet}, \gamma) > 2.95 \text{ rad}$ .

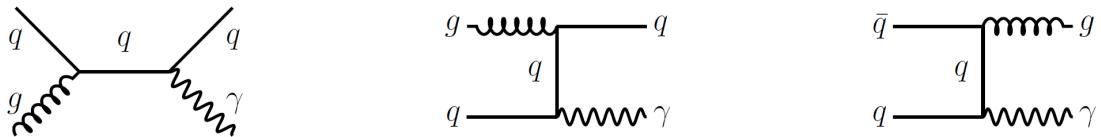


Figure 5.2: Tree-level Feynman diagrams of processes in  $pp$ -collisions with one photon and one jet in the final state.

<sup>3</sup>The  $2\sigma$  range is defined as the range  $[\mu_{100\%} - 2\sigma_{100\%}, \mu_{100\%} + 2\sigma_{100\%}]$ .



Figure 5.3: Tree-level Feynman diagrams with final (left) and initial (right) state radiation.

Additionally, the variable  $\alpha$  is defined as the ratio of the transverse momentum of the second jet to the transverse momentum of the photon

$$\alpha := \frac{p_T^{\text{2nd reco. jet}}}{p_T^\gamma}, \quad (5.2)$$

and is thus a measure of further jet activity in an event. It is, however, not sufficient to require only an upper bound on  $\alpha$ . Instead, the jet transverse-momentum resolution is measured in bins of  $\alpha$  with  $\max(\alpha) = 0.2$ , and the extrapolated value to zero further jet momentum at  $\alpha = 0$  is taken as the measured resolution of the jet  $p_T$ .

More formally, measuring the transverse momentum of the photon instead of taking the generator-level jet  $p_T$  leads to the fact that the measured resolution consists of two parts

$$\underbrace{\frac{p_T^{\text{reco. jet}}}{p_T^\gamma}}_{\text{measured}} = \underbrace{\frac{p_T^{\text{reco. jet}}}{p_T^{\text{gen. jet}}}}_{\text{intrinsic}} \cdot \underbrace{\frac{p_T^{\text{gen. jet}}}{p_T^\gamma}}_{\text{imbalance}}. \quad (5.3)$$

The intrinsic part is the resolution of interest which is independent of further jets in the event, whereas the imbalance is an artifact of further jet activity and is strongly dependent on  $\alpha$ .

To account for differences in the jet transverse-momentum resolution for different true jet  $p_T$  and different jet pseudorapidity regions, the measurement of the resolution is done in  $p_T^\gamma$  bins and bins of  $|\eta|$  of the leading jet.

As noted before, the resolution measurement is important to correct the resolution measured in simulation to the resolution measured in data. Therefore, the measurement of the jet transverse-momentum resolution is performed on data and simulation and data-to-simulation resolution scale factors ( $\rho_{\text{res}}$ ) are determined

## 5.3 Datasets and event selection

The measurement of the jet transverse-momentum resolution is carried out with  $\gamma + \text{jet}$  data recorded during the year 2012 at a centre-of-mass energy of  $\sqrt{s} = 8 \text{ TeV}$  at the CMS experiment. The datasets and triggers that are exploited for this measurement are introduced in the following Section 5.3.1. The measured resolution in data is compared to the resolution in a simulated  $\gamma + \text{jet}$  sample. Information about the simulated sample is given in Section 5.3.2. In order to select  $\gamma + \text{jet}$  events that are well suited for the resolution measurement, an event selection is applied on top. This event selection is described in Section 5.3.3.

### 5.3.1 Datasets and triggers

This analysis exploits several triggers which were active during the year 2012 at the CMS experiment. Because of the high production cross section of  $\gamma + \text{jet}$  events, especially for low photon  $p_T$ , almost all of these triggers were highly prescaled, i.e. only a fraction of events were actually recorded when the triggers fired. All triggers that are utilised in this measurement are listed in Table 5.1 together with their recorded luminosity. On

Table 5.1: Single-photon triggers together with the recorded luminosity taking the prescales of the triggers into consideration.

Trigger	Integrated luminosity [ $\text{fb}^{-1}$ ]
HLT_Photon30_CaloIdVL_IsoL	0.0029
HLT_Photon50_CaloIdVL_IsoL	0.0607
HLT_Photon75_CaloIdVL_IsoL	0.123
HLT_Photon90_CaloIdVL_IsoL	0.373
HLT_Photon135	13.77
HLT_Photon150	19.71

level-1 (L1), the triggers rely on single-photon/electron triggers, such as L1SingleEG12 and L1SingleEG30. The L1 triggers require at least one photon or electron candidate that is above a certain  $E_T$  threshold, e.g. 12 GeV or 30 GeV. The high-level triggers require a photon with a certain  $p_T$  (as indicated in the name) and, in case of thresholds below 135 GeV, additional quality and isolation criteria. All triggers with thresholds below 150 GeV were prescaled.

The events that are selected by the above mentioned triggers are contained in the datasets listed in Table 5.2.

Table 5.2: Single-photon data samples used for the resolution measurement with the contained integrated luminosity.

Dataset	Integrated luminosity [fb <sup>-1</sup> ]
/Photon/Run2012A-22Jan2013-v1/AOD	0.876
/SinglePhoton/Run2012B-22Jan2013-v1/AOD	4.412
/SinglePhoton/Run2012C-22Jan2013-v1/AOD	7.055
/SinglePhotonParked/Run2012D-22Jan2013-v1/AOD	7.354

### 5.3.2 Simulated sample

In order to compare the measured resolution in data to the resolution in simulation, a single-photon sample simulated with PYTHIA6 is used. This sample is generated with a non-physical photon  $p_T$  spectrum (slower decrease) to have a good statistical precision also for the high photon  $p_T$  region. In order to recover a physical  $p_T$  spectrum, all simulated events are reweighted. Figure 5.4 shows the photon  $p_T$  spectrum in simulation before and after reweighting.

Simulated samples come with a pileup scenario which does not necessarily match the pileup scenario in data. To match the measured distribution of interactions, the events are weighted according to their number of pileup vertices. Because almost all of the used triggers were differently prescaled over time, the distributions of pileup interactions differ among the sets of events recorded by each trigger. Thus, the weighting has to be done separately for each trigger. A comparison between the number of reconstructed vertices in data and simulation for all triggers can be found in Appendix ??.

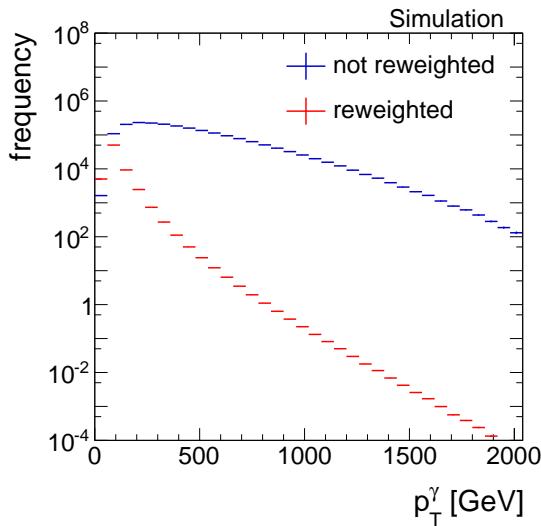


Figure 5.4: The photon  $p_T$  spectrum of the PYTHIA6 simulated single-photon sample before (blue) and after (red) reweighting.

### 5.3.3 Event selection

Events are reconstructed with the particle-flow reconstruction algorithm, which uses information of all detector components to reconstruct individual particles [?]. FIXME: Further information on the reconstruction of jets and photons at the CMS experiment can be found in Sections 3.3.2.3 and 3.3.2.4, respectively. In the following, the selection of well reconstructed jets and photons will be explained as well as the event selection of  $\gamma +$  jet events.

#### Jet selection

Particles belonging to a jet are clustered with the anti- $k_T$  jet clustering algorithm [?] with a radius parameter of  $R = 0.5$ . Furthermore, all reconstructed jets undergo a so-called “charged hadron subtraction” (CHS) which removes hadrons that are likely caused by pileup events since they do not point towards the primary vertex.

To select clean  $\gamma +$  jet events, it is required that the leading jet meets the “tight ID” selection requirements for jets [?, ?]. The tight ID selection ensures a selection efficiency of around 99% and a noise rejection efficiency of 99.98% [?]. It comprises the following selection criteria:

- Neutral hadron fraction  $< 0.90$
- Neutral electromagnetic fraction  $< 0.90$
- Number of constituents  $> 1$

And for jets in the pseudorapidity range  $|\eta| < 2.4$  :

- Charged hadron fraction  $> 0$
- Charged hadron multiplicity  $> 0$
- Charged electromagnetic fraction  $< 0.99$ .

To mitigate effects from pileup, an additional requirement on the transverse momentum is imposed:

- $p_T^{1^{\text{st}} \text{ jet}}, p_T^{2^{\text{nd}} \text{ jet}} > 10 \text{ GeV}$ .

### Photon selection

Concerning the photon, a maximum pseudorapidity of the photon of  $|\eta^\gamma| < 1.3$  is demanded to exploit the high resolution of the ECAL in the barrel region.

Furthermore, the resolution is determined for different ranges in photon  $p_T$  to avoid mixing of different prescales of the various triggers. In Table 5.3 the applied binning is shown with the respective triggers contributing to each  $p_T^\gamma$  bin.

Table 5.3: Photon  $p_T$  bins and corresponding triggers.

$p_T^\gamma$ bins	Trigger
36-60 GeV	HLT_Photon30_CaloIdVL_IsoL_v*
60-88 GeV	HLT_Photon50_CaloIdVL_IsoL_v*
88-105 GeV	HLT_Photon75_CaloIdVL_IsoL_v*
105-149 GeV	HLT_Photon90_CaloIdVL_IsoL_v*
149-165 GeV	HLT_Photon135_v*
165- $\infty$ GeV	HLT_Photon150_v*

QCD-multijet events constitute an important background to  $\gamma + \text{jet}$  events: A primary photon can be faked by a  $\pi^0$  decaying into two close-by photons. Therefore, a very clean selection of the photons is necessary to suppress this background. The following variables are used (see [?] for further explanation of the variables):

- $\frac{H}{E}$  : The ratio of the measured energy in the hadronic calorimeter over the energy measured in the electromagnetic calorimeter. For photons, this is supposed to be very small as they deposit their energy predominantly in the ECAL.

- $\sigma_{in\eta}$ : The energy weighted spatial width of the photon energy deposition. The electromagnetic shower of a photon has a small lateral size resulting in small  $\sigma_{in\eta}$  for prompt photons while showers from fake photons, e.g.  $\pi^0 \rightarrow \gamma\gamma$  have a larger lateral size.
- **ECAL isolation:** This isolation criterion uses the information of reconstructed hits “RecHits” (coming from the local reconstruction of the digital signals) in a cone around the photon supercluster of  $R = 0.4$ . Those are summed up and an upper criterion is identified to discriminate against background which is typically spatially broader.
- **Tower-based HCAL isolation:** The isolation criterion requires the energy deposited in all HCAL towers around the photon in cone of  $R = 0.4$  to be small compared to the photon energy.
- **Hollow cone track isolation:** FIXME Requires the absence of high-energetic tracks around the photon on a cone of  $R = 0.4$ . Tracks in an inner cone of  $R = 0.04$  and an  $\eta$ -slice of  $\eta = 0.015$  are not included in order to avoid a misidentification of converted photons.
- **Pixel seed veto:** In order to reduce the background from electrons and positrons further, the absence of a pixel-seed in the pixel tracker along the photon trajectory is required.

The upper bounds that are set on these observables can be found in Table 5.4.

Table 5.4: Upper bounds for all photon isolation criteria in the barrel ( $|\eta^\gamma| < 1.4442$ ).

Barrel	
$\frac{H}{E}$	< 0.05
$\sigma_{in\eta}$	< 0.013
ECAL isolation	$< 4.2 \text{ GeV} + 0.0060 \cdot p_T^\gamma$
HCAL isolation	$< 2.2 \text{ GeV} + 0.0025 \cdot p_T^\gamma$
Track Isolation	$< 2.0 \text{ GeV} + 0.0010 \cdot p_T^\gamma$
Pixel seed veto	yes

### Photon+jet event selection

Besides the mentioned requirements concerning the objects' attributes, two further criteria related to the event topology are crucial for this analysis:

A lower threshold on  $\Delta\Phi$  between the leading jet and the photon and a maximum value for  $\alpha$

- $\Delta\Phi(1^{\text{st}} \text{ jet}, \gamma) > 2.95 \text{ rad}$
- $\alpha = \frac{p_T^{2^{\text{nd}} \text{ jet}}}{p_T^\gamma} < 0.20.$

These requirements are important to suppress events with too much further hadronic activity. A summary of all selection criteria can be found in Appendix ??.

Finally, the leading jet pseudorapidity interval bounds need to be chosen in order to account for resolution differences in different detector regions. This is done according to [?] with the following binning

$$\begin{array}{ll} |\eta^{1^{\text{st}} \text{ jet}}| = 0.0 - 0.5 & |\eta^{1^{\text{st}} \text{ jet}}| = 1.1 - 1.7 \\ |\eta^{1^{\text{st}} \text{ jet}}| = 0.5 - 1.1 & |\eta^{1^{\text{st}} \text{ jet}}| = 1.7 - 2.3. \end{array}$$

The binning in the  $p_T^\gamma$  dimension is chosen according to the trigger thresholds for  $p_T^\gamma < 165 \text{ GeV}$  and ensures sufficient statistical precision for  $p_T^\gamma > 165 \text{ GeV}$

$$\begin{array}{lll} p_T^\gamma = 36 - 60 \text{ GeV} & p_T^\gamma = 149 - 165 \text{ GeV} & p_T^\gamma = 200 - 300 \text{ GeV} \\ p_T^\gamma = 60 - 88 \text{ GeV} & p_T^\gamma = 165 - 176 \text{ GeV} & p_T^\gamma = 300 - 400 \text{ GeV} \\ p_T^\gamma = 88 - 105 \text{ GeV} & p_T^\gamma = 176 - 200 \text{ GeV} & p_T^\gamma > 400 \text{ GeV} \\ p_T^\gamma = 105 - 149 \text{ GeV} & p_T^\gamma = 300 - 400 \text{ GeV}. & \end{array}$$

## 5.4 Methodology of the measurement

The basic methodology of measuring the jet transverse-momentum resolution by exploiting the  $p_T$  balance in  $\gamma + \text{jet}$  events and extrapolating the result to small  $\alpha$ , that was already used in earlier analyses [?, ?], is extended in this measurement in order to explicitly account for the influence of the direction of additional jets on the jet transverse-momentum response.

As already described in Chapter 5.2, the idea behind a resolution measurement with  $\gamma + \text{jet}$  events in data is the usage of the photon  $p_T$  instead of the true jet  $p_T$ . This results in a twofold contribution to the measured response, the intrinsic response and the imbalance:

$$\underbrace{\frac{p_T^{\text{reco.jet}}}{p_T^\gamma}}_{\text{measured}} = \underbrace{\frac{p_T^{\text{reco.jet}}}{p_T^{\text{gen.jet}}}}_{\text{intrinsic}} \cdot \underbrace{\frac{p_T^{\text{gen.jet}}}{p_T^\gamma}}_{\text{imbalance}}. \quad (5.4)$$

Taking the photon  $p_T$  as true jet  $p_T$  estimator instead of the generator-level jet  $p_T$ , and thus measuring the response defined as  $p_T^{\text{reco.jet}}/p_T^\gamma$ , results in a different shape of the response distribution compared to the intrinsic response (Fig. 5.5). The clear difference

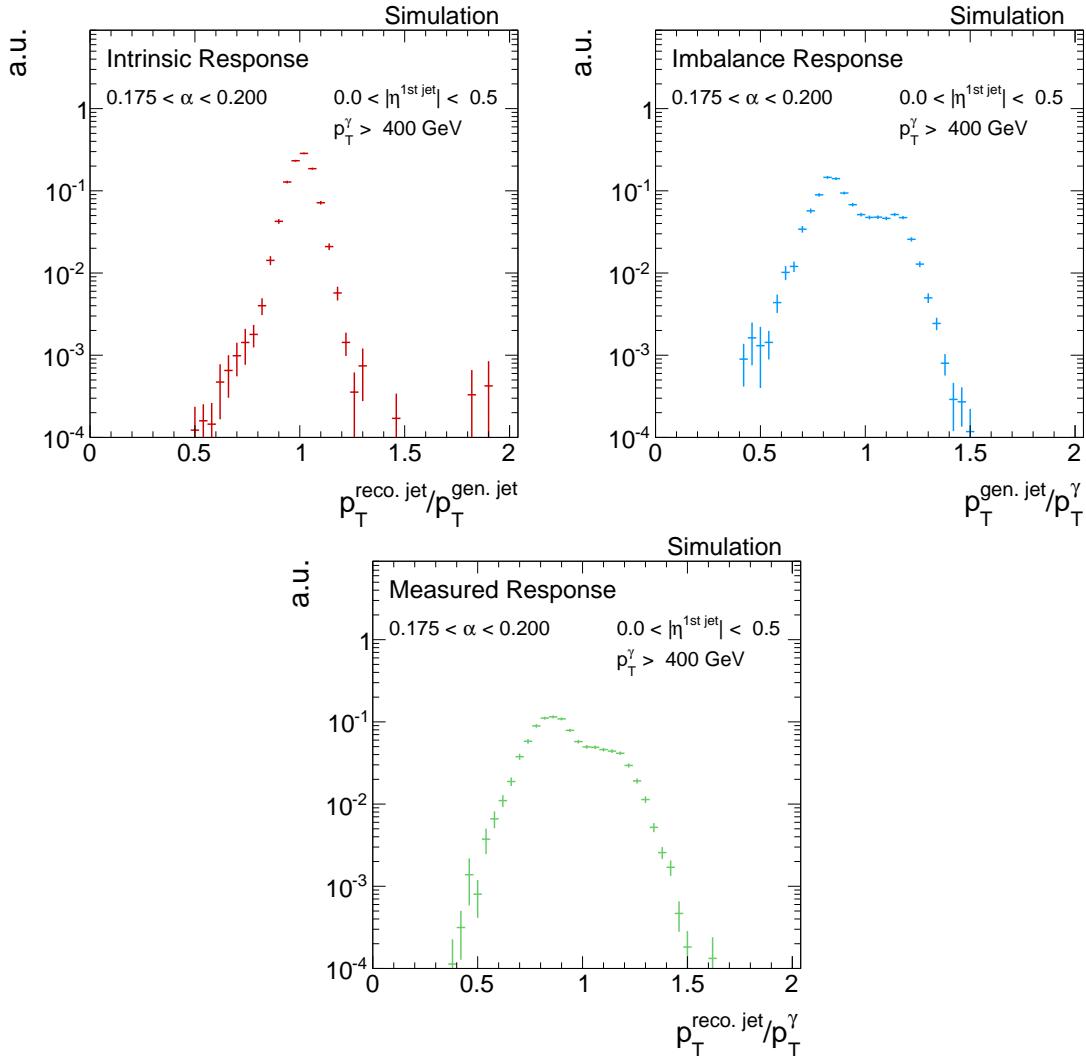


Figure 5.5: The two different contributions, intrinsic (top left) and imbalance (top right), to the measured response (bottom) (cf. Eq. (5.4)) in simulated events.

between the intrinsic response and the imbalance is the double peak structure of the latter one. The measured response is a convolution of the two contributions, where the double peak is consequentially less pronounced.

The occurrence of two peaks is caused by the hard selection in  $\Delta\Phi$  which forces the second jet to be either close to the photon or close to the leading jet due to  $p_T$  conservation. An energetic second jet perpendicular to the leading jet-photon axis would require a balancing third jet for  $p_T$  conservation. This is very unlikely due to the decreasing probability of high jet multiplicities in QCD-multijet events. Events can thus be distinguished by the relative direction of the second jet (see Fig. 5.6 for a schematic sketch), with a hemisphere definition of

$$\begin{aligned} \text{Second jet in 1<sup>st</sup> jet hemisphere:} & \quad \Delta\Phi(1^{\text{st}} \text{ jet}, 2^{\text{nd}} \text{ jet}) < \Delta\Phi(\gamma, 2^{\text{nd}} \text{ jet}), \\ \text{Second jet in photon hemisphere:} & \quad \text{else.} \end{aligned} \quad (5.5)$$

As can be seen in Fig. 5.7, events containing a second jet in the leading jet hemisphere lead to a response histogram with a mean value smaller one, while events with a second jet in photon direction result in a distribution with a mean value larger than one. The former occurs more frequently because it contains jets from final and initial state radiation, while the latter mainly consists of jets from initial state radiation.

The two response distributions coming from the different event topologies are separately evaluated. First, the resolution is determined for each of the configurations (cf. Fig. 5.7),

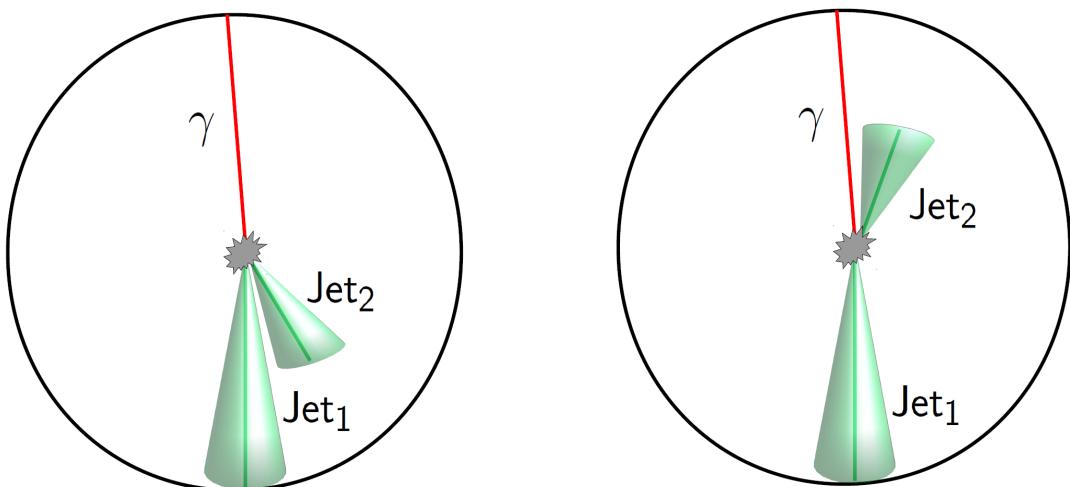


Figure 5.6: A schematic sketch of the two different event topologies where the second jet is in the leading jet (left) or the photon hemisphere (right)

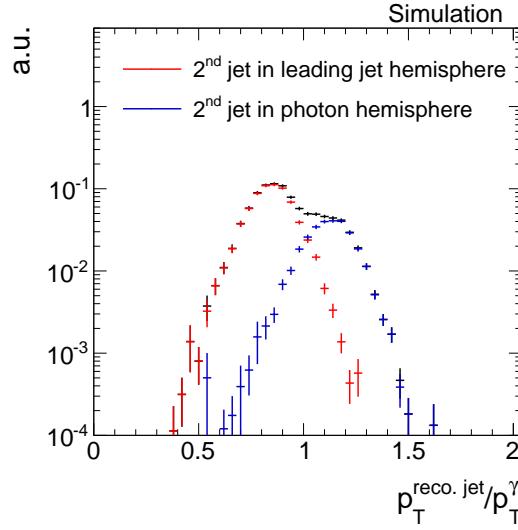


Figure 5.7: The measured jet transverse-momentum response. Events with a second jet in the photon hemisphere (red) lead to a mean response larger one, while events with a second jet in the leading jet hemisphere (blue) have a mean smaller one. The sum of both contributions is depicted in black.

and then, the weighted mean of the two resolutions is calculated.

With the above presented methodology, it is possible to measure the resolution in simulation and data for different values of  $\alpha$ . Since the quantity of interest is the jet transverse-momentum resolution independent of further jet activity, i. e. the extrapolated value of the measured resolution at  $\alpha = 0$ , the dependency of the measured resolution on alpha,  $\sigma_{\text{measured}}(\alpha)$ , is needed. Simulated events offer the possibility to study the individual contributions to the measured resolution: the intrinsic resolution  $\sigma_{\text{intrinsic}}(\alpha)$  and the imbalance  $\sigma_{\text{imbalance}}(\alpha)$ .

The intrinsic part of the resolution for a given photon  $p_T$  bin is, by definition, independent of secondary jet  $p_T$ , and can be considered as constant in terms of  $\alpha$

$$\sigma_{\text{intrinsic}}(\alpha) = c. \quad (5.6)$$

This is not true for the imbalance part. It was found empirically that the  $\alpha$  dependence of the imbalance can be described by a linear function [?]

$$\sigma_{\text{imbalance}}(\alpha) = q + m \cdot \alpha \quad (5.7)$$

Folding two independent Gaussian functions results in a quadratic addition of the corresponding standard deviations  $\sigma_{\text{intr.}} \oplus \sigma_{\text{imb.}}$  and finally yields the functional expression for

the measured resolution

$$\sigma_{\text{measured}}(\alpha) = \sqrt{c^2 + q^2 + 2q m \cdot \alpha + m^2 \cdot \alpha^2}. \quad (5.8)$$

In Fig. 5.8, the  $\alpha$  dependence of the intrinsic resolution (red dots), the imbalance (blue dots), and the measured resolution (green dots) is shown for two exemplary  $p_T^\gamma$  regions in simulated events. The intrinsic resolution is fitted with Eq. (5.6) (red line), while the imbalance is fitted with Eq. (5.7) (blue line).

**FIXME:** It is apparent that the imbalance is not zero for  $\alpha = 0$ . This has two reasons. First and most important, only the photon and the parton are balanced in the transverse plane. But since the transverse momentum of a jet is defined as the sum of all particle transverse momenta that are clustered within a cone of certain size, the jet  $p_T$  can differ from the  $p_T$  of the original parton due to either the clustering of noise or particles that do not originate from the parton or the non-clustering of particles that originate from the parton (out-of-cone showering). These effects can lead to a residual imbalance between the photon  $p_T$  and the generator-level jet  $p_T$  at  $\alpha = 0$  (parameter  $q$  in Eq. (5.7)). Second, the photon  $p_T$  can be wrongly measured and contribute further to (FIXME) the residual imbalance  $q$ .

Due to the existence of a residual imbalance  $q$ , it is not possible to simply use the measured resolution at  $\alpha = 0$  as an estimator of the intrinsic resolution  $c$ . This is already

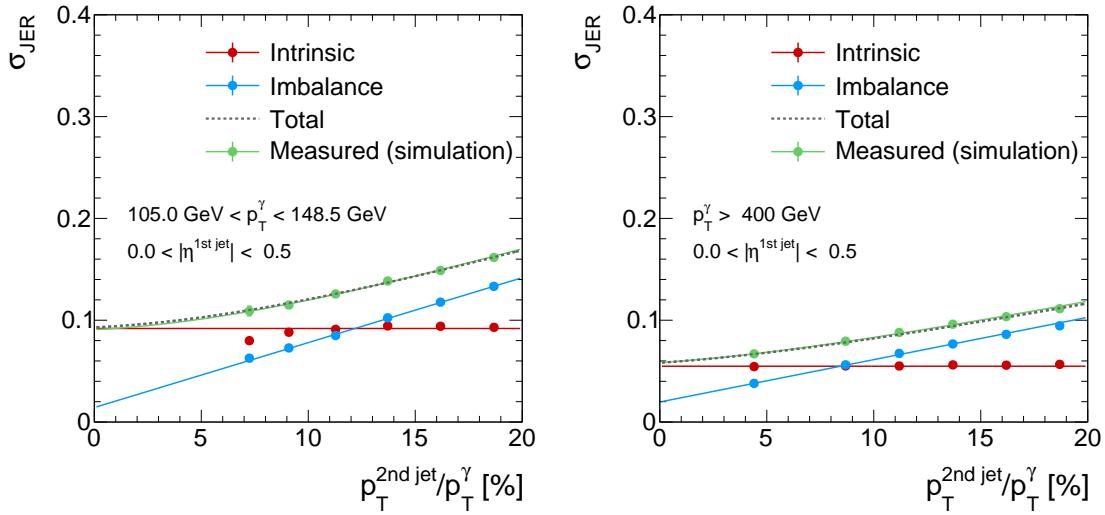


Figure 5.8: The alpha dependency of the various parts of the resolution in the simulated events for  $83 \text{ GeV} < p_T^\gamma < 99 \text{ GeV}$  (left) and  $400 \text{ GeV} < p_T^\gamma$  (right). The total resolution (grey dotted line) is the addition in quadrature of the imbalance (blue line) and the intrinsic (red line) fit functions. It can be compared to the measured resolution in simulation (green dots/line).

obvious from Eq. (5.8) which evaluates to  $\sigma_{\text{measured}} = \sqrt{c^2 + q^2}$  at  $\alpha = 0$ .

For the determination of the intrinsic resolution in data, the method relies on the residual imbalance information from simulation. Practically, when fitting the measured resolution (Eq.(5.8)) in data, the residual imbalance  $q$  is fixed to the value obtained from the imbalance fit (Eq 5.7) in simulation. The intrinsic resolution in data can then be obtained by extracting the fit parameter  $c$ . In the following the estimator of the intrinsic resolution (fit parameter  $c$  from the measured resolution) will be referred to as  $\sigma_{\text{JER}}$ .

For consistency, the same approach is used for simulated events. The fitted function of the measured resolution (green line) in Fig. 5.8 corresponds to the fit function of Eq. (5.8) with the parameter  $q$  fixed to the value obtained from the imbalance fit (blue line). Finally, the dotted grey line is the total resolution with the analytic expression of function (5.8) with the parameters set to the fit values of the intrinsic (red) and the imbalance fit (blue). It can be nicely seen that the total resolution and the fit function of the measured resolution (green line) are well in agreement. This implies, that the assumed functional forms of the resolution components and the assumed convolution leads indeed to a good description of the measured resolution. Further validations of the methodology will be described in Section 5.4.1.

In Fig. 5.9, the estimated intrinsic resolution in simulation (green) and data (black) is shown in different photon  $p_T$  bins for two different  $|\eta^{1\text{st jet}}|$  regions. The resolution improves for increasing photon  $p_T$ . For the  $|\eta^{1\text{st jet}}| < 0.5$  region, the resolution is approximately 10% for  $p_T^\gamma \approx 100 \text{ GeV}$  and decreases to values around 7% for  $p_T^\gamma \approx 300 \text{ GeV}$ .

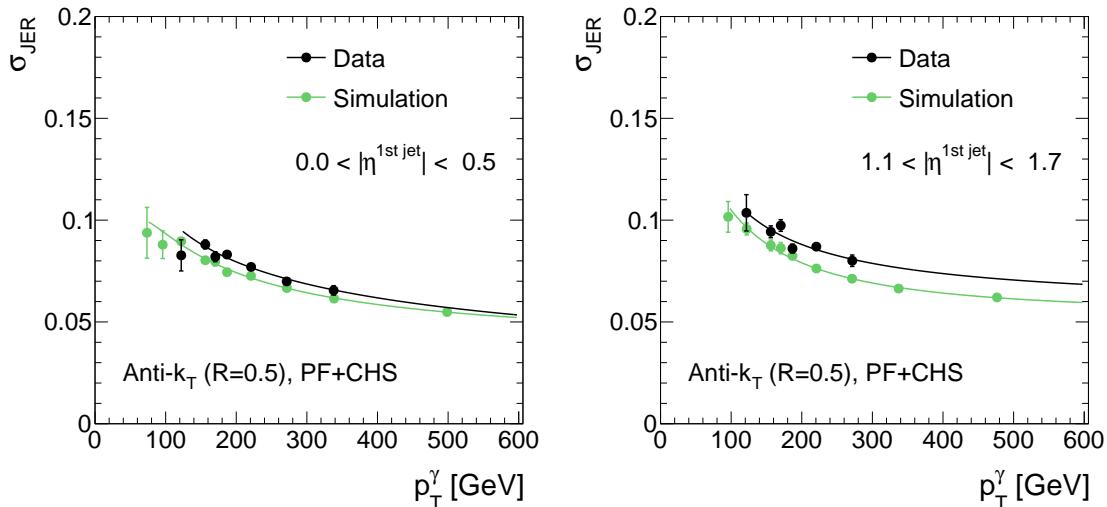


Figure 5.9: Resolution for  $|\eta^{1\text{st jet}}| < 0.5$  (left) and  $1.1 < |\eta^{1\text{st jet}}| < 1.7$  (right) in simulation (green) and data (black). Both are fitted with a particle-flow resolution function introduced in [?].

The increasing statistical uncertainties for low photon  $p_T$  arise through the requirements of  $\alpha = \frac{p_T^{2\text{nd reco. jet}}}{p_T^\gamma} < 0.2$  and  $p_T^{2\text{nd jet}} > 10 \text{ GeV}$ . This reduces the numbers of events in the low photon  $p_T$  bins. For events with  $p_T^\gamma \lesssim 50 \text{ GeV}$ , it is not possible at all to fulfil both requirements at the same time.

From Fig. 5.9, it can be seen that the resolution in data is always worse than the resolution in simulation. Thus, in order to measure data-to-simulation scale factors  $\rho_{\text{res}}$  for the jet transverse-momentum resolution, the measured resolution in data is divided by the resolution in Monte Carlo simulation (MC).

$$\rho_{\text{res}} = \frac{\sigma_{\text{JER}}^{\text{data}}}{\sigma_{\text{JER}}^{\text{MC}}} \quad (5.9)$$

The results will show (Chapter 5.6), that this ratio is independent of  $p_T^\gamma$  and can thus be fitted with a constant in terms of  $p_T^\gamma$ . The measurement of the data-to-simulation scale factors is therefore only done in bins of  $|\eta^{1\text{st jet}}|$ .

How to apply these scale factors and adjust the jet transverse-momentum resolution in simulation to the measured resolution in data is well described in [?].

### 5.4.1 Validation of the method

It has already been shown that the decomposition of the measured resolution into intrinsic resolution and imbalance works well for the two shown  $p_T^\gamma$  bins in Fig. 5.8. In order to further validate the method, a comparison between  $\sigma_{\text{JER}}$  extracted from the measured resolution and the real  $\sigma_{\text{JER}}$  of the intrinsic resolution is performed on simulated events. Figure 5.10 depicts the results of this comparison for one exemplary  $|\eta^{1\text{st jet}}|$  bin. It can be seen that the method has a bias of up to 15% towards an underestimation of the resolution for small  $p_T^\gamma$ . For  $p_T^\gamma > 100 \text{ GeV}$  the bias is less than 5% and reduces to less than 1% for  $p_T^\gamma > 200 \text{ GeV}$ .

The residual bias of the method is caused by the accumulation of high  $p_T^{1\text{st jet}}$  events in the low alpha bins and low  $p_T^{1\text{st jet}}$  events in the high alpha bins, as will be explained in the following paragraph FIXME. Since the jet transverse-momentum resolution decreases for higher  $p_T^{1\text{st jet}}$ , this leads to an upward slope of the resolution with increasing alpha (cf. red dots in Fig 5.8 (left)).

The intrinsic resolution is not affected by this  $\alpha$  dependency because it is fitted with a constant and thus corresponds to the mean resolution across all  $\alpha$  bins. In contrast, the fit of the measured resolution with an additional free parameter can adopt this increase and can therefore result in an underestimation of the y-intercept and thus of  $\sigma_{\text{JER}}$ . For high photon  $p_T$ , this effect is less pronounced, since the slope of  $\sigma_{\text{JER}}(p_T^\gamma)$  (see Fig. 5.10) flattens out. Thus the method is only biased for small  $p_T^\gamma$ .

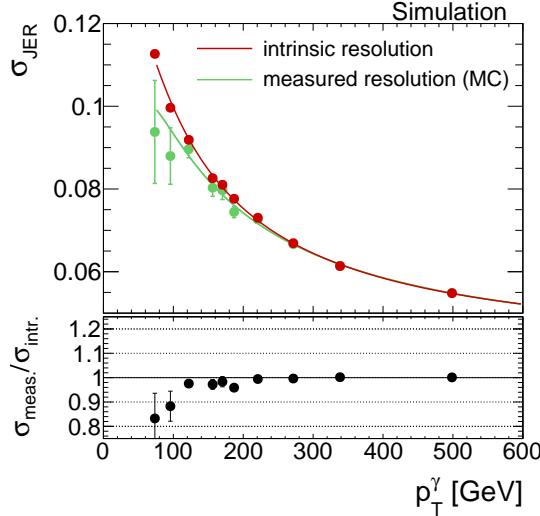


Figure 5.10: Consistency check of the method: Comparison between  $\sigma_{\text{JER}}$  obtained from the measured resolution (Eq. (5.8)) and from the intrinsic resolution (Eq. (5.6)) in simulation.

The accumulation of high  $p_{\text{T}}^{1^{\text{st}} \text{jet}}$  events in low  $\alpha$  bins is stemming from two effects. First, the binning in  $p_{\text{T}}^{\gamma}$  and the momentum balance between the photon and the first two jets lead to a dependency of the first jet  $p_{\text{T}}$  on the second jet  $p_{\text{T}}$  and therefore alpha: For a fixed  $p_{\text{T}}^{\gamma}$ , the  $p_{\text{T}}$  of the first jet gets smaller for larger  $p_{\text{T}}^{2^{\text{nd}} \text{jet}}$  in events where the second jet is in the leading jet hemisphere (see Fig. 5.6), leading to a negative correlation of  $p_{\text{T}}^{1^{\text{st}} \text{jet}}$  and  $p_{\text{T}}^{2^{\text{nd}} \text{jet}}$ . This effect is directly opposite for events with a second jet in the photon hemisphere, where the first jet  $p_{\text{T}}$  gets larger for larger  $p_{\text{T}}^{2^{\text{nd}} \text{jet}}$  and thus a positive correlation between  $p_{\text{T}}^{1^{\text{st}} \text{jet}}$  and  $p_{\text{T}}^{2^{\text{nd}} \text{jet}}$ . In principle, the effect should thus cancel out, but since the topology of a second jet in the leading jet hemisphere is much more frequent, a residual upward trend in the resolution vs.  $\alpha$  is conserved.

**FIXME:** The second source of the accumulation of high  $p_{\text{T}}^{1^{\text{st}} \text{jet}}$  events in low  $\alpha$  bins arises from the requirement of  $p_{\text{T}}^{2^{\text{nd}} \text{jet}} > 10 \text{ GeV}$ . This selection requirement induces threshold effects for low  $p_{\text{T}}^{\gamma}$  bins ( $p_{\text{T}}^{\gamma} \lesssim 150 \text{ GeV}$ ) since the tighter  $\alpha$  requirements can only be fulfilled for higher transverse momenta of the photons in the corresponding  $p_{\text{T}}^{\gamma}$  bin. Thus, high photon  $p_{\text{T}}$  events accumulate in the low alpha regions. As the selected events are almost balanced, a high  $p_{\text{T}}^{\gamma}$  is associated with a high jet  $p_{\text{T}}$ , thus also the high jet  $p_{\text{T}}$  events accumulate in the low alpha bins, leading to an upward trend in the resolution vs.  $\alpha$ .

However, since the discussed bias affects the resolution measurement in simulation and data equivalently, the residual bias for low  $p_{\text{T}}^{\gamma}$  is not of concern for the measurement of data-to-simulation resolution scale factors  $\rho_{\text{res}}$  (Eq. (5.9)). To prove the hypothesis

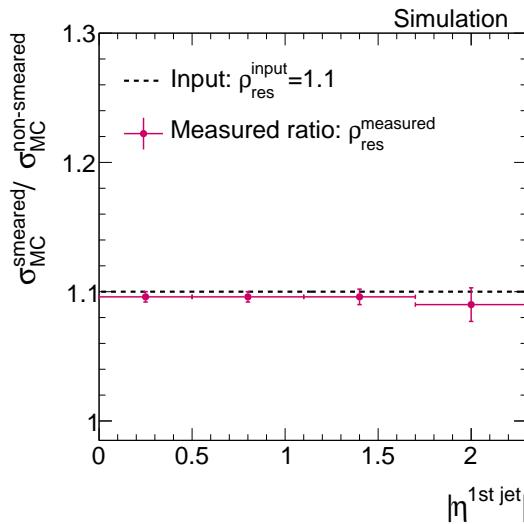


Figure 5.11: Comparison of the resolution ratio  $\sigma_{\text{JER}}^{\text{smeared MC}}/\sigma_{\text{JER}}^{\text{non-smeared MC}}$  measured in simulated events with the input smearing factor  $\rho_{\text{res}} = 1.1$  in all four  $|\eta^{1\text{st}\text{ jet}}|$  bins.

of a cancellation of the bias, a pseudo dataset is generated from the simulated dataset where the jet transverse-momentum resolution is smeared by  $\rho_{\text{res}} = 1.1$  in all  $|\eta^{1\text{st}\text{ jet}}|$  bins. Subsequently, the resolution scale factors between the smeared pseudo dataset and the original simulated dataset are determined according to the methodology presented in this thesis. The resulting scale factors  $\rho_{\text{res}}$  should reproduce the input values of  $\rho_{\text{res}} = 1.1$ . In Fig. 5.11, the result of this test is shown. In all  $|\eta^{1\text{st}\text{ jet}}|$  bins, the measurement of the scale factors reproduces the input factors within the statistical uncertainties. The deviation of the measured scale factors to the input value is less than 0.9% in all four  $|\eta^{1\text{st}\text{ jet}}|$  bins. Thus, the method is expected to also hold for the determination of the data-to-simulation ratio.

After the characterisation and estimation of the associated systematic uncertainties of the jet transverse-momentum resolution measurement with  $\gamma + \text{jet}$  events, the results of the measurement on  $\sqrt{s} = 8 \text{ TeV}$  data will be presented.

## 5.5 Systematic uncertainties

Many systematic uncertainties of the jet transverse-momentum resolution measurement cancel out when focusing on the data-to-simulation ratio  $\sigma_{\text{JER}}^{\text{data}}/\sigma_{\text{JER}}^{\text{MC}}$ . In the following sections, only uncertainties relevant for this ratio will be discussed.

For the final uncertainty, the single uncertainties are added in quadrature, resulting in relative uncertainties between 2.3% to 6.3% for the lowest and highest  $|\eta^{\text{1st jet}}|$  bin, respectively.

An overview of all systematic uncertainties can be found in Table 5.5.

Table 5.5: All relative systematic uncertainties on the data-to-simulation ratio  $\sigma_{\text{JER}}^{\text{data}}/\sigma_{\text{JER}}^{\text{MC}}$  listed by sources for the different  $|\eta^{\text{jet}}|$  bins.

	$ \eta^{\text{jet}} $			
	0.0 - 0.5	0.5 - 1.1	1.1 - 1.7	1.7 - 2.3
<b>Multijet contamination</b>	+2.0%	+2.0%	+2.3%	+2.5%
	-2.0%	-2.0%	-2.3%	-2.5%
<b>Simulation of flavour composition</b>	+0.9%	+0.9%	+0.8%	+0.6%
	-0.9%	-0.9%	-0.8%	-0.6%
<b>Simulation of out-of-cone showering</b>	+0.5%	+2.8%	+3.6%	+5.7%
	-0.5%	-2.8%	-3.6%	-5.7%
<b>Jet energy scale</b>	+0.6%	+0.6%	+0.6%	+0.7%
	-0.5%	-0.6%	-0.6%	-0.6%
<b>Pileup reweighting</b>	+0.1%	+0.1%	+0.1%	+0.2%
	-0.1%	-0.1%	-0.2%	-0.2%
<b>Total</b>	+2.3%	+3.6%	+4.4%	+6.3%
	-2.3%	-3.6%	-4.4%	-6.3%

## Uncertainty on the contamination with QCD-multijet events

Due to the huge QCD-multijet cross section, a sizable fraction of dijet events can survive the strict photon selection (see Section 5.3.3) if a jet is misidentified as a photon. This happens, when e.g. a jet hadronises to a  $\pi^0$  that decays to two close-by photons that mimic the signature of an isolated photon from a  $\gamma + \text{jet}$  event.

The two leading jets in a dijet event are balanced except for initial and final state radiation and have in principle the same topology as  $\gamma + \text{jet}$  events. Therefore, the presented method is in principle not affected by dijet contamination. Adverse effects of dijet back-

ground contamination are only expected due to a mismeasurement of the  $p_T$  of the jet misidentified as photon, because only the energy of the  $\pi^0$  is counted and not the full jet energy. Another aspect is the different flavour composition of a QCD-multijet sample. Due to the different production mechanism, QCD-dijets are predominantly initiated by gluons while the leading jet in  $\gamma +$ jet events often stems from a light quark. The number of particles after hadronisation is typically larger for gluon jets, hence the single particles are less energetic and out-of-cone showering is more pronounced. Since the residual imbalance  $q$  in this analysis is taken from simulation considering  $\gamma +$ jet events only, it is not expected to accurately describe the residual imbalance in the data sample that is contaminated with dijet events.

To investigate the impact of QCD-multijet contamination, a QCD-multijet sample, enriched in jets with a large electromagnetic fraction, is added to the  $\gamma +$ jet sample and weighted according to the cross section. Subsequently, the resolution measured using both the  $\gamma +$ jet and the QCD-multijet sample is compared to the resolution measured with the  $\gamma +$ jet sample only. Since the QCD-multijet sample has very large event weights leading to high statistical uncertainties on the measured jet transverse-momentum resolution, two selection criteria ( $\alpha < 0.4$  and  $\Delta\Phi > 2.7$ ) are relaxed in order to increase the statistical precision of the uncertainty estimation. Additionally, a rougher binning in  $\alpha$ ,  $p_T^\gamma$  and  $|\eta^{1\text{st jet}}|$  is applied. The residual imbalance  $q$  of the resolution measurement including the QCD-multijet sample is fixed to the residual imbalance determined from the  $\gamma +$ jet only analysis in order to account for a possible error in the evaluation of the  $\sigma_{\text{JER}}^{\text{data}}/\sigma_{\text{JER}}^{\text{MC}}$  ratio for which data is only compared to a  $\gamma +$ jet sample. It can be seen from Fig. 5.12 that the jet transverse-momentum resolution is worse for low  $p_T^\gamma$  when considering QCD-multijet contamination.

FIXME: The observed difference in the resolution measurement of the two approaches is taken as a systematic uncertainty on the measured resolution in data and propagated to the final data-to-simulation ratio which are assessed in different  $\eta^{\text{jet}}$  regions. The resulting uncertainties vary between 2.0 – 2.5% for the different  $|\eta^{1\text{st jet}}|$  regions (cf. Table 5.5).

## Uncertainty on the flavour composition in simulation

A possible difference between the resolution of different jet flavours (e.g. caused by more pronounced out-of-cone showering of gluon jets) should in principle not play a role for the data-to-simulation scale factors  $\rho_{\text{res}}$  as long as the flavour composition in data and simulation is the same.

To account for possible discrepancies in the flavour composition between data and simulation, the gluon and quark flavour fractions of the simulated sample (Fig 5.13) are varied by 10%. FIXME: To estimate the effect of different flavour compositions, the resolution in simulation is separately evaluated for quarks ( $\sigma_{\text{JER}}^{\text{quarks}}$ ) and gluons ( $\sigma_{\text{JER}}^{\text{gluons}}$ ) and then

weightedely added to produce a variation of the flavour composition of  $\pm 10\%$ .

There are various definitions used in the CMS collaboration how to assign the underlying generator-level quark/gluon flavour to a jet. In this measurement the so-called “algorithmic” flavour definition is used (see Appendix ?? for more details on this definition). The composition in the simulated PYTHIA  $\gamma + \text{jet}$  sample is around 60% light quarks and 20% to

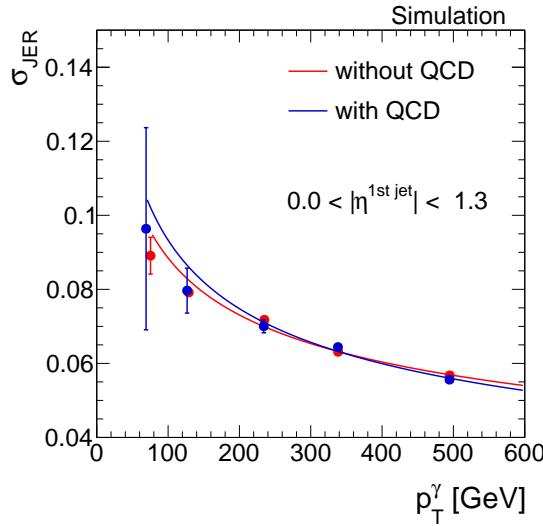


Figure 5.12: The jet transverse-momentum resolution measured in simulation for  $|\eta^{\text{1st jet}}| < 1.3$  with (blue) and without (green) a QCD-multijet sample added to the  $\gamma + \text{jet}$  sample.

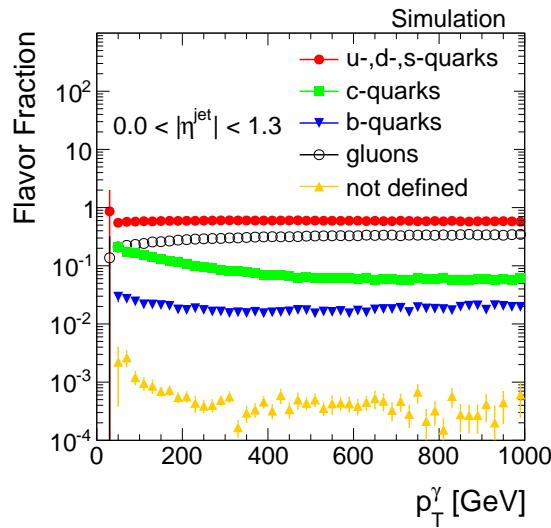


Figure 5.13: The flavour composition in the simulated  $\gamma + \text{jet}$  sample in the barrel region of the detector. The “algorithmic” flavour definition is used (see Appendix ?? for more details).

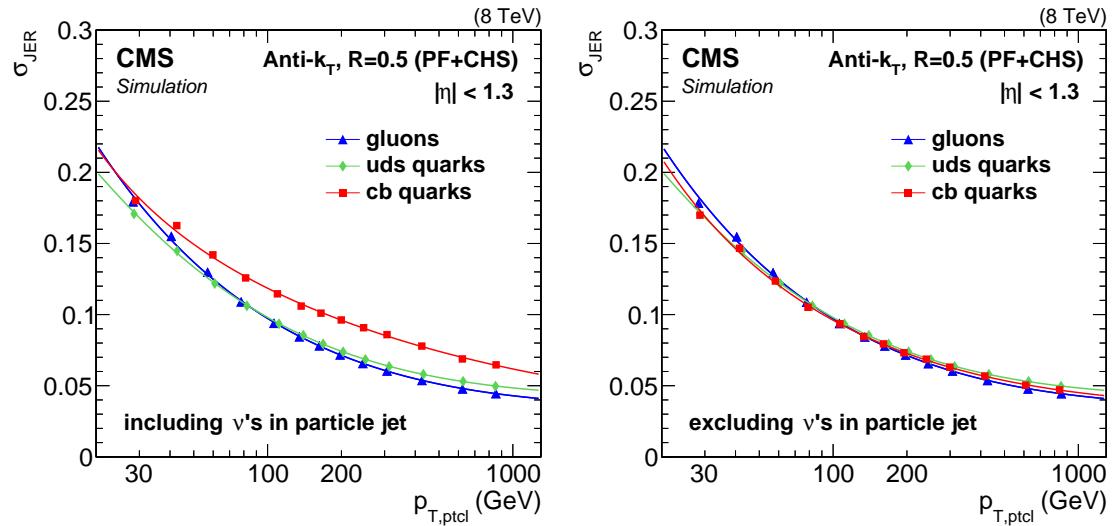


Figure 5.14: The intrinsic resolution  $\sigma_{\text{intr}}$  for  $|\eta^{\text{1st jet}}| < 1.3$  for different jet flavours with (left) and without (right) including the neutrinos in the generator-level jet. Here, JER refers to  $\sigma_{\text{intr}}$ , and  $p_{T,\text{ptcl}}$  to the generator-level jet  $p_T$ .

35% gluons (see Fig. 5.13). The missing fraction is mainly made up out of charm quarks.

The left panel of Figure 5.14 depicts the intrinsic resolution for all flavours separately for  $|\eta^{\text{1st jet}}| < 1.3$ . It can be seen that the resolution for gluon and light quark jets is comparable whereas heavier quarks have a worse resolution across the full  $p_T$  range. This is because the decay of heavy quarks can involve the production of neutrinos, whose transverse momentum is not detectable, leading to an underestimation of the reconstructed jet  $p_T$ . Since the neutrino  $p_T$  is included in the jet  $p_T$  on generator level, c- and b-jets have a more pronounced left tail of the jet transverse-momentum response, leading to a worse resolution compared to light quark jets (left part of Fig. 5.14). This is not the case, if the neutrino  $p_T$  is not added to the generator-level jet  $p_T$ , as can be seen in Fig 5.14 (right).

As already noted, the uncertainty is estimated with a variation of the flavour composition in simulation of  $\pm 10\%$ . The constant fit to the data-to-simulation ratio  $\sigma_{\text{JER}}^{\text{data}}/\sigma_{\text{JER}}^{\text{MC}}$  is reevaluated, leading to a final systematic uncertainty between 0.6 and 0.9%.

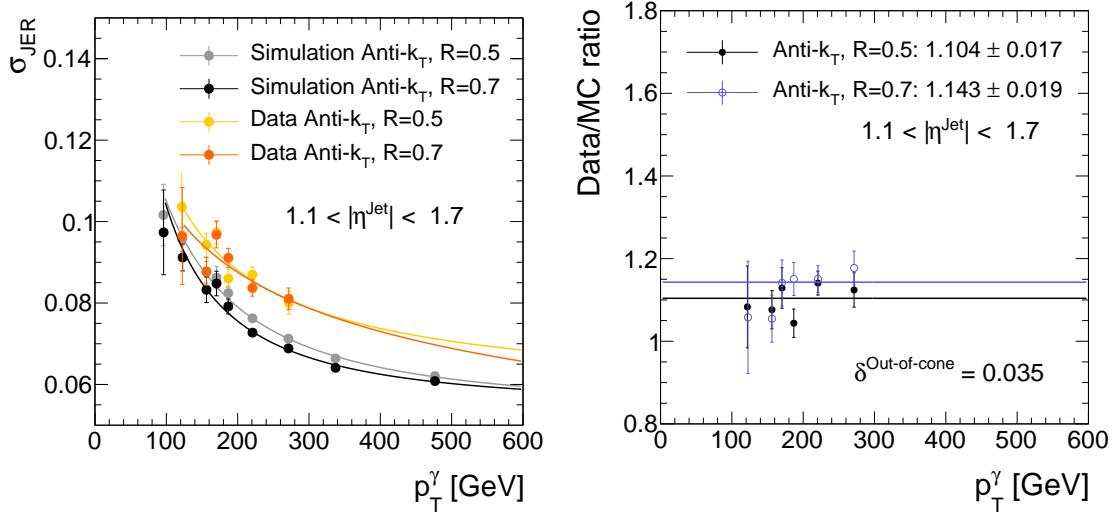


Figure 5.15: Left: The measured resolution in simulation and data with AK5-jets and AK7-jets for  $1.1 < |\eta^{1^{\text{st}} \text{ jet}}| < 1.7$ . Right: Data-to-simulation ratio for AK5- and AK7-jets for  $1.1 < |\eta^{1^{\text{st}} \text{ jet}}| < 1.7$ .

## Uncertainty on the simulation of out-of-cone showering

### Uncertainty on the parton - generator-level jet correspondence in simulation

### Uncertainty on the simulation of the residual imbalance

Another source of uncertainty is the use of information from Monte Carlo simulation for determining the residual imbalance  $q$  from the simulated  $\gamma + \text{jet}$  sample (Eq. (5.8)).

FIXME: Since it cannot be expected that the correspondence of the generator-level jet and the original parton is perfectly modelled in simulation, possible differences between real and simulated out-of-cone showering and noise or other particle clustering are estimated by the evaluation of the ratio  $\sigma_{\text{JER}}^{\text{data}}/\sigma_{\text{JER}}^{\text{MC}}$  using different jet radii in the jet reconstruction algorithm.

For the primary analysis, jets reconstructed by the anti- $k_T$  algorithm with a radius parameter of  $R = 0.5$  are used (AK5-jets). In order to evaluate the systematic uncertainty on the out-of-cone showering simulation, the measurement of  $\rho_{\text{res}}$  is redone with jets reconstructed with a jet radius of  $R = 0.7$  (AK7-jets).

The data-to-simulation ratio is in all  $|\eta^{1^{\text{st}} \text{ jet}}|$  bins larger for AK7-jets, resulting in uncertainties between 0.5 – 5.7%. The uncertainty on the simulation of out-of-cone showering is thus the largest systematic uncertainty of this measurement.

## Uncertainty on the jet energy scale

A further uncertainty arises from the correction of the jet energy scale. The transverse momentum of each jet is corrected in order to have uniform response over the full  $\eta^{\text{jet}}$ , and  $p_T^{\text{jet}}$  range [?]. In data, it is additionally corrected to account for data-simulation differences. The latter correction can impact the evaluation of the data-to-simulation ratio. To estimate this effect, the uncertainties on the correction factors that are only applied to the data samples are varied up and down within their  $1\sigma$ -uncertainties.

The effect of the jet  $p_T$  variation on the data-to-simulation ratio is of minor importance and range between  $0.5 - 0.7\%$ .

## Uncertainty on the pileup reweighting

Finally, an uncertainty due to the adjustment of the simulated events to the pileup distribution in data is evaluated.

To account for this uncertainty, the effect of a 5.0% up- and downward variation of the minimum-bias cross section (69.4 mb) on the resolution is evaluated, following the recommended procedure from [?]. The resulting uncertainties are almost negligible and ranges between  $0.1 - 0.2\%$ .

## 5.6 Results

The data-to-simulation resolution scale factors  $\rho_{\text{res}}$  are determined in  $19.7 \text{ fb}^{-1}$  of  $pp$ -collision data at  $\sqrt{s} = 8 \text{ TeV}$  with the methodology described in Chapter 5.4. In each  $\alpha$ ,  $p_T^\gamma$  and  $|\eta^{1\text{st jet}}|$  bin, the width of the 99% truncated response histogram is determined in simulation and data. Afterwards, the extrapolation to zero additional jet activity is carried out, by fixing the imbalance  $q$  in simulation and in data to the value extracted from the imbalance extrapolation in simulation. Exemplary extrapolations for the imbalance, the intrinsic and the measured resolution in simulation and the measured resolution in data are shown in Fig. 5.16. The full set of figures for each  $p_T^\gamma$  and  $|\eta^{1\text{st jet}}|$  bin can be found in Appendix ??.

Finally, the extracted resolutions in data and simulation in every  $p_T^\gamma$  bin are divided and a constant is fitted to this ratio

$$\frac{\sigma_{\text{JER}}^{\text{data}}}{\sigma_{\text{JER}}^{\text{MC}}} (p_T^\gamma). \quad (5.10)$$

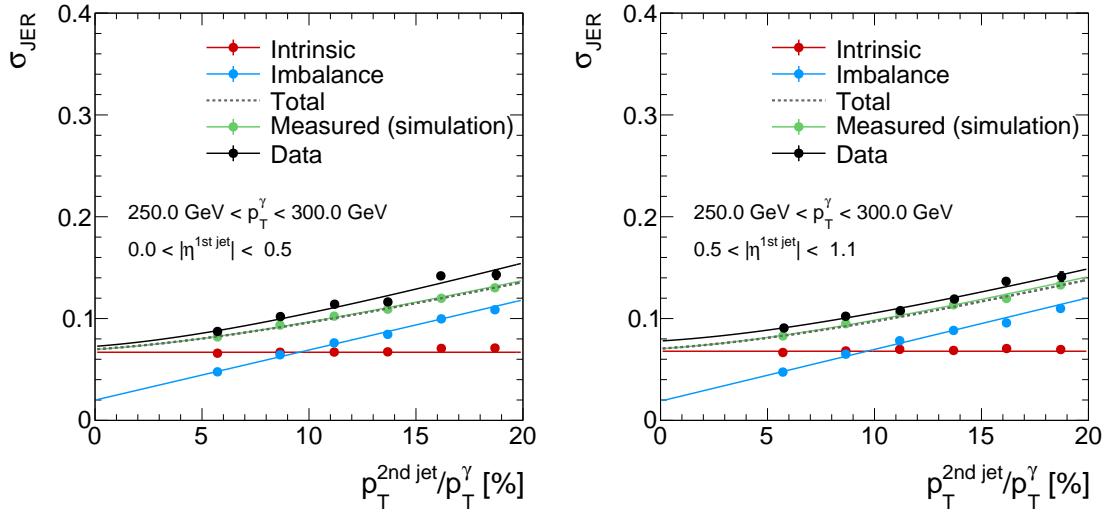


Figure 5.16: Two examples of the alpha dependency of the measured jet transverse-momentum resolution in data (black dots), in simulation (green dots) and of the intrinsic (red dots) and the imbalance (blue dots) part of the resolution in simulated events. All resolutions are fitted with the corresponding functions introduced in Chapter 5.4 (Eqs. (5.6)- (5.8)). The total resolution (grey dotted line) is the addition in quadrature of the imbalance (blue line) and the intrinsic (red line) fit functions.

In Fig. 5.17, the results for all four  $|\eta^{1\text{st jet}}|$ -ranges are depicted. The  $\chi^2/\text{NDF}$ -values for the four fits vary between 0.23 and 2.53 corresponding to fit probabilities between 0.97 and 0.06. Thus, a constant fit is justified, and one value for every  $|\eta^{1\text{st jet}}|$  bin will be reported.

The resulting  $\rho_{\text{res}}$  is always larger than one, which means that the resolution in data is always worse than the resolution in simulation. The values of the fits range from 1.067 (for the first  $\eta^{\text{jet}}$  bin) to 1.199 (for the last  $\eta^{\text{jet}}$  bin).

The systematic uncertainties are evaluated as described in the previous chapter. The single uncertainties are taken as upper and lower boundary of the 68% uncertainty band and are added in quadrature to get the total systematic uncertainty. Table 5.6 summarises the data-to-simulation ratio results determined with data collected during the year 2012 with their statistical and systematic uncertainties. The visualised result can be found in Fig. 5.18.

Though the  $\gamma + \text{jet}$  analysis is known for producing highly precise results, the systematic uncertainties are still dominating. This is mainly caused by the uncertainty on the simulation of out-of-cone showering. The statistical limitation of this analysis is due to the collected data at the CMS detector. The number of simulated events is roughly eight

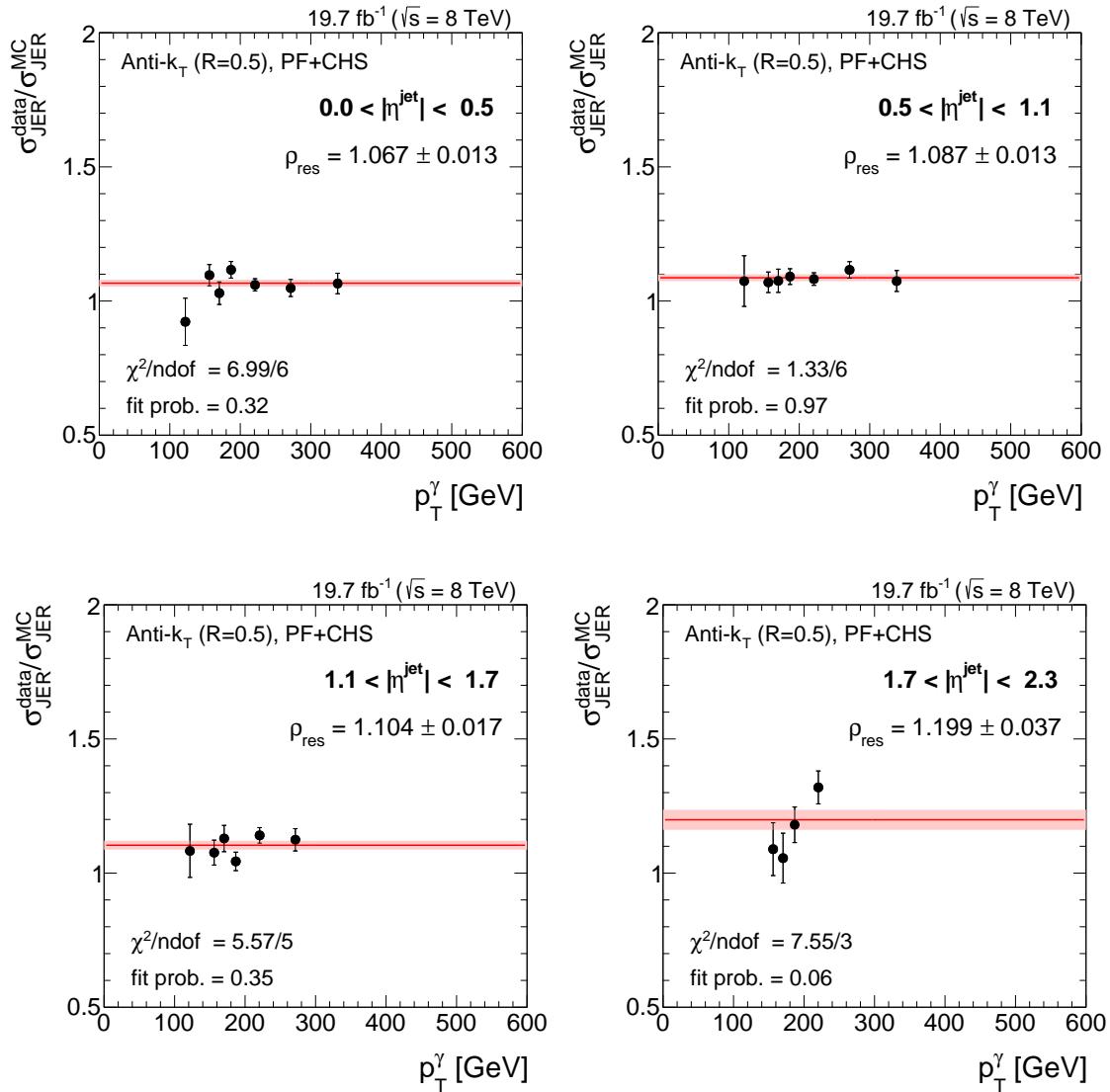


Figure 5.17: Data-to-simulation resolution ratios  $\rho_{\text{res}}$  (black dots), fitted with a horizontal line (red line) for four different  $|\eta^{\text{jet}}|$ -ranges. The fit uncertainties are depicted with light red error bands.

times larger.

### 5.6.1 Comparison to 2011 measurement

A comparison of the data-to-simulation resolution scale factors  $\rho_{\text{res}}$  between this analysis and the results of 2011 which were determined from a dijet data sample [?] can be found in Fig. 5.19. It can be seen, that throughout the whole  $\eta^{\text{jet}}$ -range the data-to-simulation resolution ratios  $\rho_{\text{res}}$  are systematically larger for the result in 2011. Comparing the

Table 5.6: Data-to-simulation resolution scale factors  $\rho_{\text{res}}$  with statistical and systematic uncertainties.

$ \eta^{\text{jet}} $	$\rho_{\text{res}}$	stat.	sys.
0.0 – 0.5	1.067	$\pm 0.013$	$+0.025$ $-0.024$
0.5 – 1.1	1.087	$\pm 0.013$	$+0.039$ $-0.039$
1.1 – 1.7	1.104	$\pm 0.017$	$+0.049$ $-0.049$
1.7 – 2.3	1.199	$\pm 0.037$	$+0.075$ $-0.075$

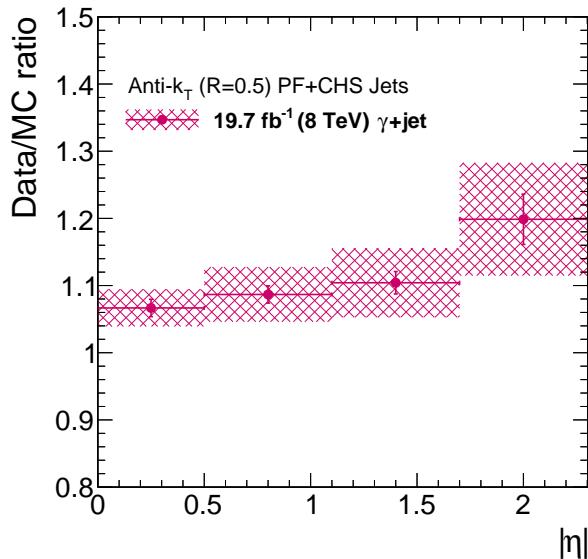


Figure 5.18: The measured data-to-simulation resolution ratio in  $\gamma + \text{jet}$  events using data recorded in the year 2012. The red band depicts the total uncertainty whereas the error bars show the statistical uncertainty only.

precision of both measurements, the  $\gamma + \text{jet}$  analysis is for all  $\eta$  bins more precise than the analysis done with dijet events. This is due to the smaller systematic uncertainties of this analysis which compensates for the better statistical precision of the QCD-multijet sample due to the large cross section.

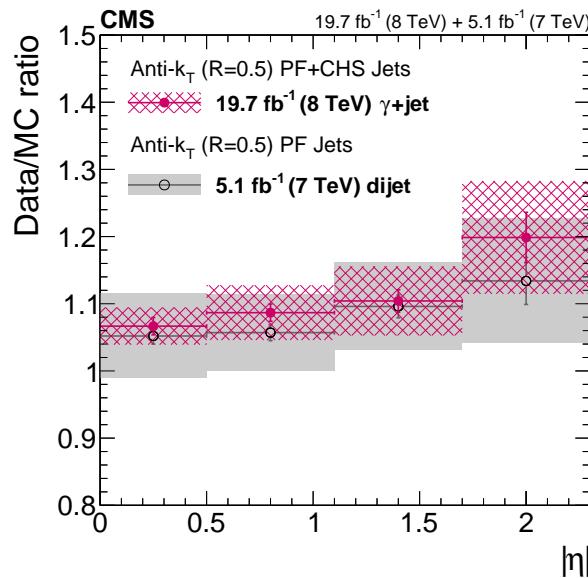


Figure 5.19: Data-to-simulation jet  $p_T$  resolution ratio determined from  $\gamma + \text{jet}$  events from 2012 compared to dijet results from 2011. The grey and magenta bands correspond to the full uncertainties, whereas the error bars depict the statistical uncertainties only.

### 5.6.2 Comparison to 2012 dijet measurement

A jet  $p_T$  resolution measurement using dijet events was also carried out in 2012 [?]. A comparison of the here presented measurement using  $\gamma + \text{jet}$  events to the measurement with dijet events can be found in Fig. 5.20. Both measurements are compatible within their uncertainties. It can be seen that the statistical precision is much better for the measurement with QCD-dijet events because of the larger cross section. The systematic uncertainties are of comparable size for small pseudorapidity bins and are larger for high  $|\eta^{\text{jet}}|$  in the here presented measurement. This is due to the uncertainty on the simulation of the out-of-cone showering. This uncertainty plays a smaller role in the dijet measurement (between 0.4% and 1.5%) as the effect by out-of-cone showering partly cancels out in case of two balanced jets.

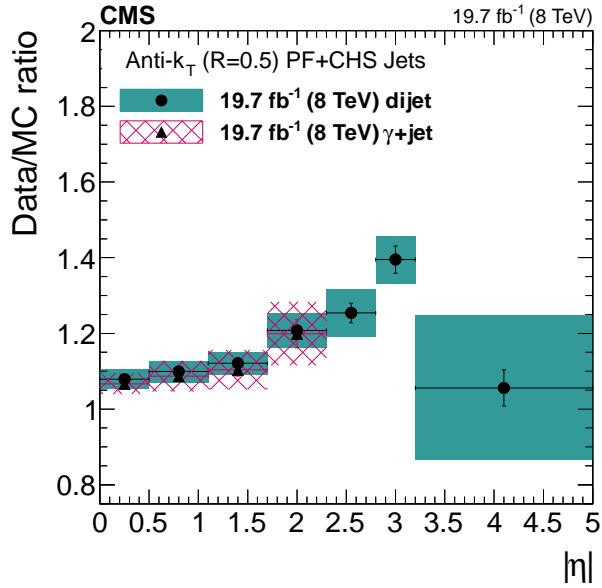


Figure 5.20: Data-to-simulation jet  $p_T$  resolution ratio determined from  $\gamma + \text{jet}$  events from 2012 compared to dijet results from 2012. The blue and magenta bands correspond to the full uncertainties, whereas the error bars depict the statistical uncertainties only.

## 5.7 Discussion and conclusion

The difference of the jet transverse-momentum resolution between data and simulation is an important input for many analyses at CMS, including searches for physics beyond the Standard Model using QCD-multijet events, e. g. [?, ?, ?].

In this analysis, the first measurement of the data-to-simulation scale factors  $\rho_{\text{res}}$  using  $\gamma + \text{jet}$  events with  $19.7 \text{ fb}^{-1}$   $pp$ -collision data at  $\sqrt{s} = 8 \text{ TeV}$  collected in 2012 at the CMS detector is presented. The analysis relies on the methodology developed in [?, ?]. During the study, it was found that the jet transverse-momentum response depends on the topology of additional jet activity in the events, leading to a double peak structure. This was not accounted for in previous analyses [?, ?]. Therefore, a new method has been developed within this thesis to consistently deal with the distortion of the response distribution due to additional jets from initial and final state radiation (see Section 5.4).

The resolution in data is found to be systematically larger than in simulation by around

7% to 20% throughout the investigated  $\eta^{\text{jet}}$  plane. A possible  $p_{\text{T}}^{\text{jet}}$  dependence of the resolution scale factors  $\rho_{\text{res}}$  is not visible. Thus, the ratio is parametrised with a constant for each  $\eta^{\text{jet}}$  bin. The accuracy of the measurement is dominated by systematic uncertainties, leading to a total uncertainty of 3% to 8%. The results of this analysis compare well to similar studies, including a dijet study at 8 TeV [?]. Both measurements are compatible within their statistical and systematic uncertainties.

For a future measurement, a reduction of the uncertainties is desirable. The main source of uncertainty is the systematic uncertainty on the simulation of out-of-cone showering. Within this analysis, this uncertainty was estimated by varying the jet clustering radius from  $R = 0.5$  to  $R = 0.7$ . This approach might overestimate the uncertainty, since the difference in  $\rho_{\text{res}}$  associated with different jet clustering radii may include other effects than out-of-cone showering. Thus, a more detailed investigation of the quality of out-of-cone showering simulation could yield a reduced uncertainty of the  $\rho_{\text{res}}$  measurement using  $\gamma + \text{jet}$  events.

Furthermore, an increase in the statistical precision is desirable to allow for the determination of resolution scale factors in the high pseudorapidity regions which have not been accessible so far.

# **Part 6**

## **Summary**



In the year 2012,  $19.7 \text{ fb}^{-1}$  of proton-proton collisions at an unprecedented centre-of-mass energy of 8 TeV were recorded at the CMS experiment. This data has been analysed with a special focus on searches for physics beyond the Standard Model and measurements of Standard Model parameters as well as important performance parameters of the CMS detector. This thesis contributes in a twofold way to the physics program of CMS. First, a search for long-lived supersymmetric particles in a parameter range to which previous analyses had no - or only limited - sensitivity is performed. Second, the jet transverse-momentum resolution is measured in  $\gamma + \text{jet}$  events at 8 TeV. The following paragraphs summarise the two presented analyses. A more detailed discussion about the most important findings and an outlook for further research can be found in the last chapters of the corresponding parts, Chapter 4.10 and Chapter 5.7, respectively.

The search for physics beyond the Standard Model that is performed within this thesis is strongly motivated by supersymmetric extensions of the Standard Model that include long-lived charginos decaying inside the tracker into the lightest supersymmetric particle, the neutralino. Because of the higher masses of supersymmetric particles, the chargino is expected to deposit much higher amounts of energy in the tracking system compared to the Standard Model background. Thus, a selection on the energy deposit per path length,  $dE/dx$ , can be highly discriminating. Furthermore, the search is designed to target supersymmetric models not yet excluded, i.e. models with chargino lifetimes of the order of  $c\tau \approx 1 - 30 \text{ cm}$  where most of the charginos decay in the first layers of the tracker. Therefore, the here presented search focuses on the selection of highly ionising, short tracks. It is the first analysis at CMS that incorporates reconstructed tracks down to three hits and includes  $dE/dx$  measurements from the innermost detector part, the silicon pixel tracker. For this purpose, a pixel tracker energy calibration was carried out that assures a uniform energy response across pixel modules and over time. By the inclusion of pixel energy information, the background suppression for a given signal efficiency is significantly increased by up to one order of magnitude (cf. Fig 4.18).

The background expectation is mainly estimated with data-based techniques and consists mostly of fake tracks, i.e. tracks that are not associated to one single particle. Fake tracks can easily mimic the signal signature because of their typically large  $dE/dx$  values. A selection is performed in four different signal regions in order to increase the search sensitivity to different chargino lifetimes and masses. The background estimation of this search is particularly challenging because of the low event yield in most of the Standard Model background enriched control regions. Therefore, the search sensitivity is mainly limited by systematic uncertainties arising from the limited size of simulated samples and control regions in data.

The results of the search are compatible with Standard Model expectations and are,

thus, used to constrain the supersymmetric parameter space with wino-like charginos. With this search, it is possible to exclude supersymmetric models with long-lived charginos down to lifetimes of  $c\tau = 2$  cm for chargino masses of 100 GeV and between  $c\tau = 70$  cm – 500 cm for masses of 500 GeV at a 95% confidence level. Current limits are confirmed and improvements of the order of 10-40 GeV in chargino mass are achieved.

The second contribution of this thesis consists of the measurement of the jet transverse-momentum resolution with  $\gamma + \text{jet}$  events at 8 TeV at the CMS detector. The jet  $p_T$  resolution is a crucial ingredient for analyses at CMS that are relying on a good understanding of the quality of the jet  $p_T$  measurement, e.g. searches for physics beyond the Standard Model where QCD-multijet background plays a major role [?, ?, ?] and Standard Model measurements such as the differential  $t\bar{t}$  production cross-section [?].

The analysis exploits the transverse momentum balance of the photon and the jet in the absence of further jet activity. Due to the excellent energy resolution of the electromagnetic calorimeter, the photon  $p_T$  can thus be used as a precise estimate of the true jet  $p_T$ . The method of the resolution measurement that is used within this thesis is based on earlier analyses [?, ?] but is the first measurement that accounts for the fundamental non-Gaussian behaviour of the measured resolution in exclusive bins of further jet activity. It is shown, that the Gaussian behaviour can be recovered when separating events by the direction of further jets in the event. By this improvement, a consistent and well performing method for the jet  $p_T$  resolution measurement is established.

Since the main application of the resolution measurement is the adjustment of the simulated resolution to the resolution measured in data, the results are presented as data-to-simulation scale factors. In four different pseudorapidity regions up to  $|\eta| = 2.3$ , the scale factors vary between 7% and 20% with uncertainties between 3% to 8%. The main uncertainty is the uncertainty on the simulation of out-of-cone showering (FIXME). The results are in agreement with the jet  $p_T$  resolution measurement performed on dijet events [?].

To conclude, this thesis contributes in several ways to the investigation of physics beyond the Standard Model and to a better understanding of the CMS detector performance. First, the search for highly ionising, short tracks helps to gain further insights into supersymmetric models with long-lived charginos and further constrains the supersymmetric parameter space. Second, the pixel energy calibration significantly improves the discrimination power of the variable  $dE/dx$  and allows for  $dE/dx$  measurements of very short tracks. Finally, an improved method is used to measure an important performance parameter of the CMS experiment, the jet transverse-momentum resolution.