

## **Part 1**

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



## 1.1 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 ??). While supersymmetric models, especially the Minimal Supersymmetric Standard Model (MSSM) (Section ??), have been studied at previous particle colliders including Tevatron and LEP [1, 2], 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. [3–5]). 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 ??). 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 [6, 7]. While it is not possible to explain the full relic density with thermally produced wino-like neutralinos for  $m_{\tilde{\chi}_1^0} \lesssim 3 \text{ TeV}$  [8], neutralinos can still be the dominant part if they are non-thermally produced via the decay of an almost decoupled particle [9, 10].

Additionally, explorations of the MSSM parameter space reveal that many models that are consistent with current observations and theoretical constraints and that offer a neu-

tralino as Dark Matter candidate include a metastable chargino with a mass gap of the order of  $\sim 160$  MeV with respect to the neutralino [11].

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 MSSM (see Section ?? for an introduction to the pMSSM). The results of such a study which has been performed in [12] are shown in Figure 1.1. 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 [13] and the search for disappearing tracks [12] 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 [13] and the search for disappearing tracks [12]. While in [13], the high ionisation losses of hypothetical new massive particles is exploited, it does not take into account whether its reconstructed track is disappearing. In [12], 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

---

<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.

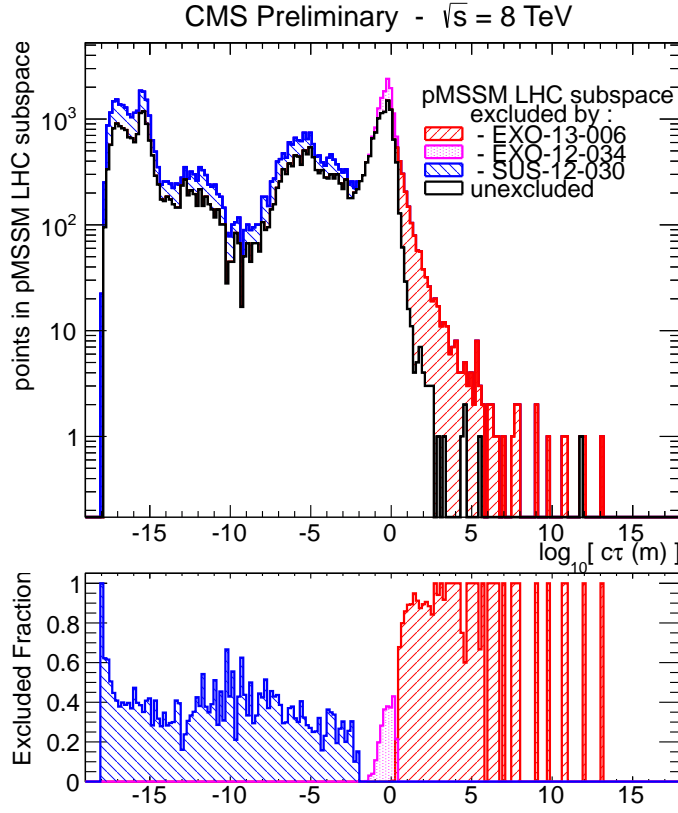


Figure 1.1: The number of excluded pMSSM points at 95% C.L. (upper part) and the fraction of excluded pMSSM points (bottom part) vs. the chargino lifetime for different CMS searches. Red area: the search for long-lived charged particles [13], Purple area: the search for disappearing tracks [12], Blue area: a collection of various general SUSY searches [14] 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 [15] for further details). Taken from: [16].

properties that are highly distinctive for charginos with intermediate lifetimes: first, the 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.

## 1.2 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. 1.2. Other possible chargino pair production channels include the exchange of a supersymmetric Higgs boson or a  $t$ -channel squark exchange (Fig. 1.3).

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. 1.4).

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

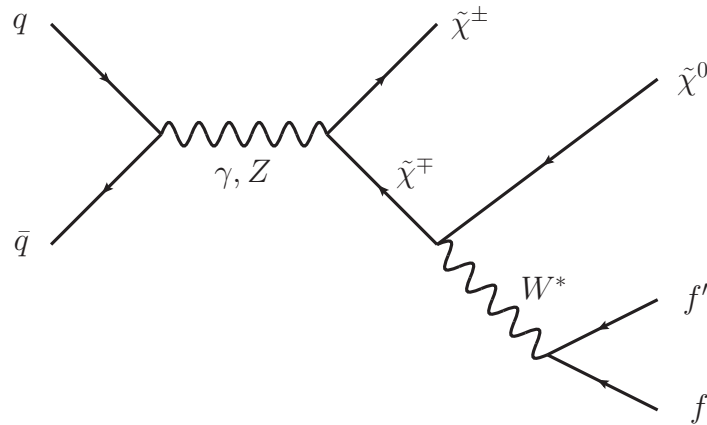


Figure 1.2: Feynman diagram of chargino pair production via gamma or  $Z$ -boson exchange and the subsequent decay via a virtual  $W$ -boson.

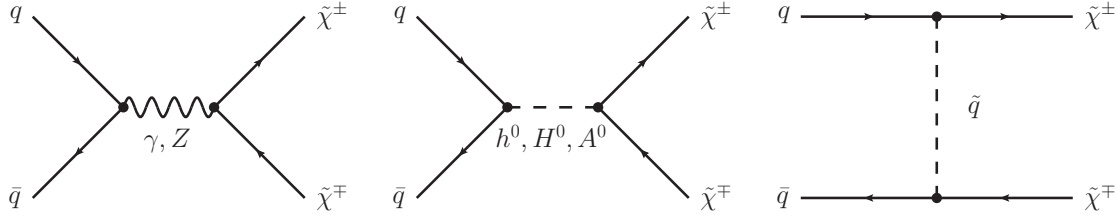


Figure 1.3: Main tree-level diagrams for chargino pair production.

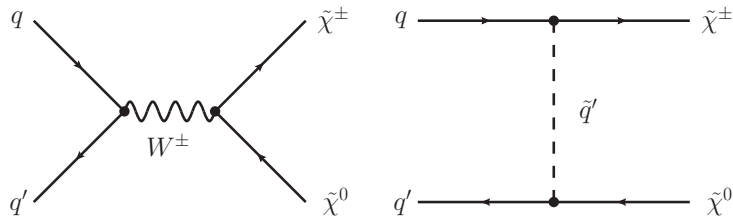


Figure 1.4: Main tree-level diagrams for chargino neutralino production.

can reach the muon chambers and can be reconstructed as a muon even despite a longer time-of-flight [17]. 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. 1.5, where simulated chargino-chargino events are shown in a cross-sectional view of the CMS detector. In the left picture of Fig. 1.5, 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  $c\tau \approx 1 \text{ cm} - 10 \text{ cm}$ , 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.

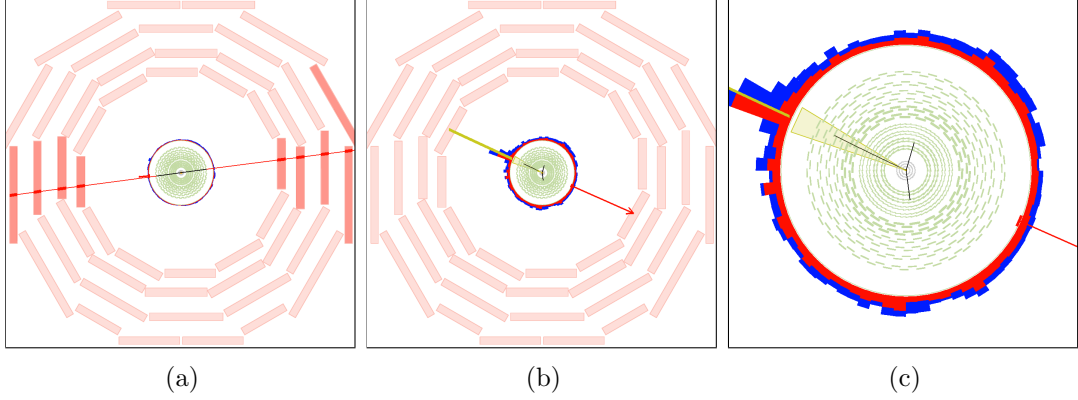


Figure 1.5: Visualisation of possible signatures of a chargino pair produced with a lifetime of  $c\tau = 10$  m (a) and a lifetime of  $c\tau = 0.5$  m (b and c). The muon chambers are the outer layers of the detector and are depicted as red boxes. The black lines represent the reconstructed chargino 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 arrow shows the missing transverse energy in the event. 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

$$p_\pi \sim \sqrt{\left(m_{\tilde{\chi}_1^\pm} - m_{\tilde{\chi}_1^0}\right)^2 - m_\pi^2}. \quad (1.1)$$

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. 1.6).

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. 1.7 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 [18]. 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,



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. 1.3 and 1.4, 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

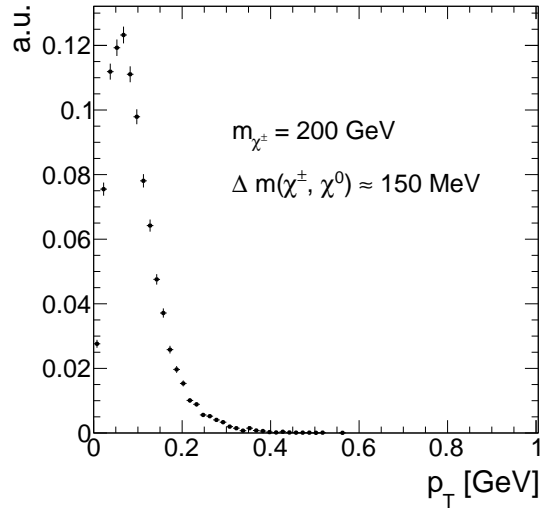


Figure 1.6: Transverse momentum distribution of pions coming from charginos decay into a neutralino with a mass gap of 150 MeV.

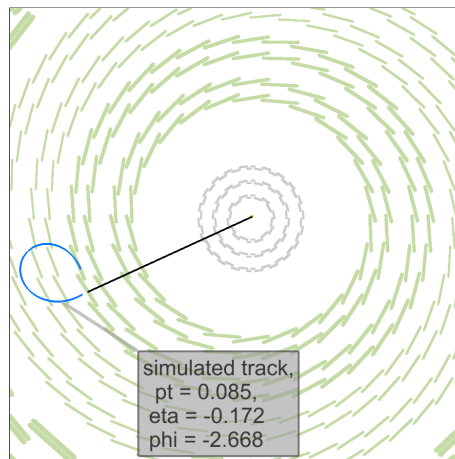


Figure 1.7: Cross-sectional view of the silicon strip tracker (green lines) and silicon pixel tracker (grey lines). A simulated chargino track (black line) decays to a pion (bended blue line) with a  $p_T$  of  $\sim 85$  MeV and a neutralino (not visible).

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 due to their high transverse momenta (Fig. 1.8). Furthermore, the energy loss per path length ( $dE/dx$ ) depends quadratically on the particle mass for low velocities ( $0.2 < \beta\gamma < 0.9$ ):

$$\left\langle \frac{dE}{dx} \right\rangle = K \frac{m^2}{p^2} + C \quad (1.2)$$

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 is required since defining a helical path requires five parameters (see [18]). 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

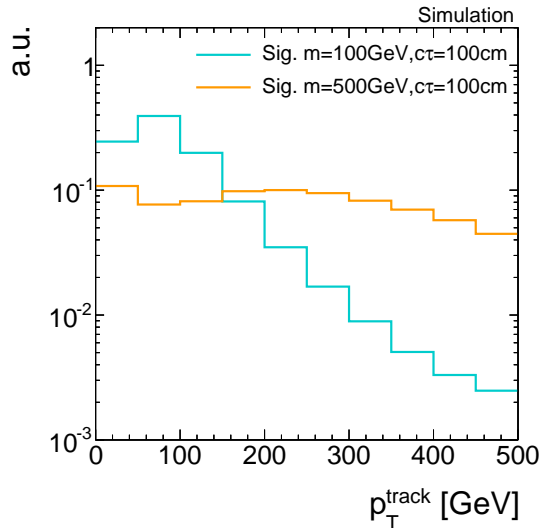


Figure 1.8: Normalised distribution of the transverse momentum of chargino tracks for two different signal models. No selection is applied.

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.

### 1.2.1 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 [13] and the search for disappearing tracks [12]. 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 [12].

In [13], a minimum number of eight hits were required for every track, whereas in [12] 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. 1.9 (left), the normalised distribution of the number of measurements ( $N_{\text{hits}}$ ) of reconstructed chargino tracks is shown. Since a minimum number of three hits is required for track reconstruction, there are no tracks with  $N_{\text{hits}} < 3$ . It can be seen, that the distribution peaks at the minimal possible value of  $N_{\text{hits}} = 3$  for lower lifetimes. For a lifetime of  $c\tau = 100$  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 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. 1.9 (right). For very short tracks ( $N_{\text{hits}} = 3$ ) the efficiency is still around 20%.

Additionally, the search for disappearing tracks [12] 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 [13], 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. 1.1).

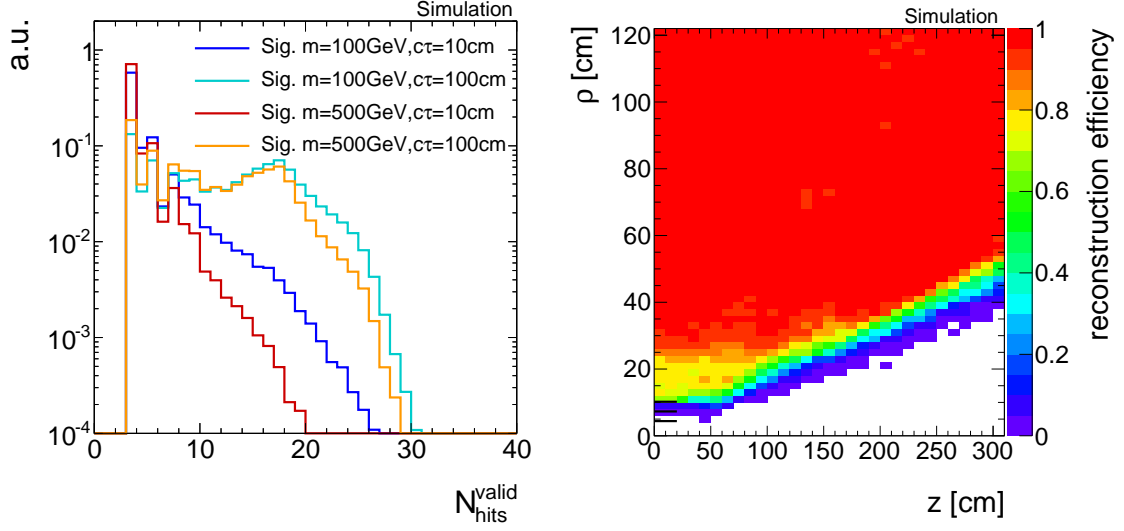


Figure 1.9: Left: Number of measurements in the tracker system  $N_{\text{hits}}$  for four different signal lifetimes. Right: Probability to reconstruct a chargino track ( $z$ -axis) in dependency of the chargino decay point in  $z$ -direction and radial direction  $\rho = \sqrt{x^2 + y^2}$  ( $x$ - and  $y$ -axis). More information on the generation of the simulated signal samples can be found in Section 1.4.2.

To conclude, the general search strategy of the here presented analysis is to unite the strategies of [13] and [12] and to lower the strong selection on the number of hits in these analyses in order to get an optimised selection for lifetimes around  $1\text{ cm} \lesssim c\tau \lesssim 30\text{ cm}$ .

### 1.3 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 [13], 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. This is done by measuring the current induced by the electron hole

pairs produced during the ionisation of the silicon. A detailed introduction to the CMS tracker system and the energy measurement can be found in Section ??.

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 1.3.1. The pixel energy calibration is then described in Section 1.3.2. How to discriminate SM particles and beyond SM particles with the help of a  $dE/dx$  measurement is discussed in Section 1.3.3, 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 1.3.4).

### 1.3.1 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 [19]:

$$\langle \frac{dE}{dx} \rangle = K z^2 \frac{Z}{A} \frac{1}{\beta^2} \left[ \frac{1}{2} \ln \frac{2m_e c^2 \beta^2 \gamma^2 T_{\max}}{I^2} - \beta^2 - \frac{\delta(\beta\gamma)}{2} \right]. \quad (1.3)$$

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 [20].  $T_{\max}$  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  $\text{MeV mol}^{-1} \text{cm}^2$ . 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” [21]. 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 [22]

$$p(x) = \frac{1}{\pi} \int_0^\infty e^{-t \log t - xt} \sin(\pi t) dt. \quad (1.4)$$

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. 1.10). Theoretically it extends to infinite energies, however in nature the maximal deposited energy is of course

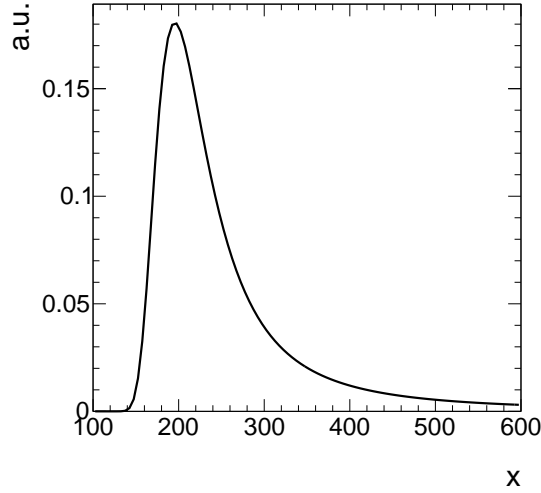


Figure 1.10: Illustration of the shape of a Landau distribution. Parameters were chosen as  $\mu = 200$  and  $\sigma = 20$ .

limited by the particle's full energy.

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 by the Landau-Vavilov-Bichsel equation [23]:

$$\Delta_p = \xi \left[ \ln \frac{2m_e c^2 \beta^2 \gamma^2}{I} + \ln \frac{\xi}{I} + j - \beta^2 - \delta(\beta\gamma) \right], \quad (1.5)$$

with  $\xi = (K/Z)\langle Z/A \rangle(x/\beta^2)$ . 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. 1.11. The restricted energy loss refers to the energy loss when the maximal energy transfer  $T_{\max}$  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  $\beta\gamma \sim 100$ .

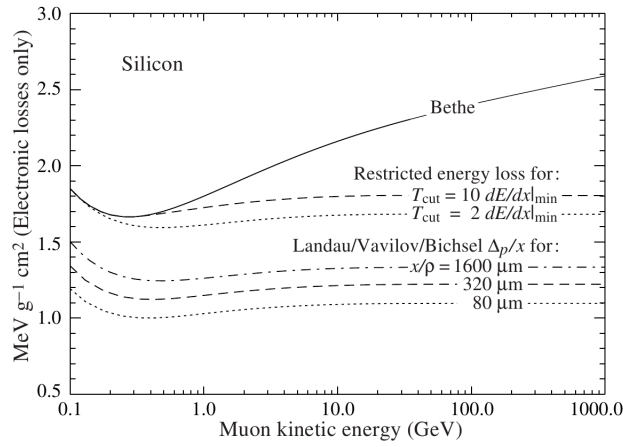


Figure 1.11: Comparison between the Bethe mean energy loss, restricted energy loss and the most probable energy loss described by the Landau-Vavilov-Bichsel function for muons for different values of absorber thickness of silicon.  $T_{\text{cut}}$  refers to the maximal energy transfer for the restricted energy loss calculation. The absorber thickness is normalised to the absorber density  $\rho$ , thus  $[x] = g \cdot \text{cm}^{-2}$ . Taken from [21].

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, several “estimators” can be used that suppress a potential bias towards the high end without introducing a bias towards lower values [24]. One of the estimators for determining a track’s  $dE/dx$  is the harmonic-2 estimator

$$I_{\text{h2}} = \left( \frac{1}{N} \sum_{i=1}^N (\Delta E_i / \Delta x_i)^{-2} \right)^{-1/2}, \quad (1.6)$$

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 [24].

A further estimator of  $dE/dx$  used for the discrimination of highly ionising particles

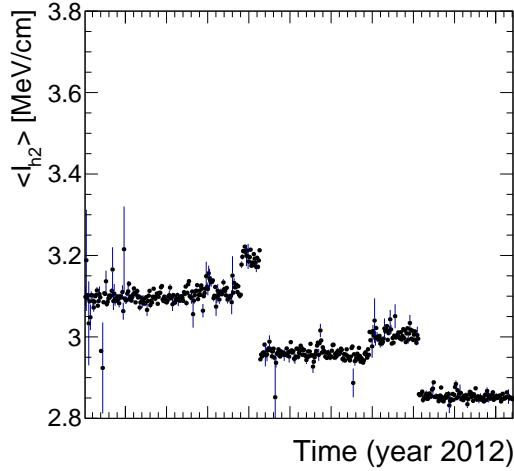


Figure 1.12: Mean of all track's  $dE/dx$  (harmonic-2 estimator) over the full year 2012. Only pixel hits are taken into account. Every data point corresponds to one run.

will be introduced in Section 1.3.3.

### 1.3.2 Energy calibration of the silicon pixel tracker

During Run I in 2012, the pixel silicon detector was continuously subject 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 [25]. 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.

This imperfection of the gain calibration can be seen in Fig. 1.12, where the mean of the harmonic-2 estimator  $\langle I_{h2} \rangle$  for high-quality muon tracks with  $p_T > 45 \text{ GeV}$  over the full data-taking period in 2012 is shown. 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. 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 [24].

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. 1.11). 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  $p > 2$  GeV. Additionally, only tracks with at least eight hits and a good quality trajectory fit ( $\chi^2/n_{\text{dof}} < 3$ ) 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. For this sample, events are randomly recorded with a basic requirement of a signal in the Beam Pick-up Timing for experiments (BPTX). BPTX refers to two devices  $\sim 170$  m from the interaction point that allow for identifying a bunch crossing.

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

$$\Delta x = \frac{d_{\text{module}_i}}{\cos(\phi_{\text{track}})}, \quad (1.7)$$

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 1.13 shows an example distribution for one module. To extract the MPV for every

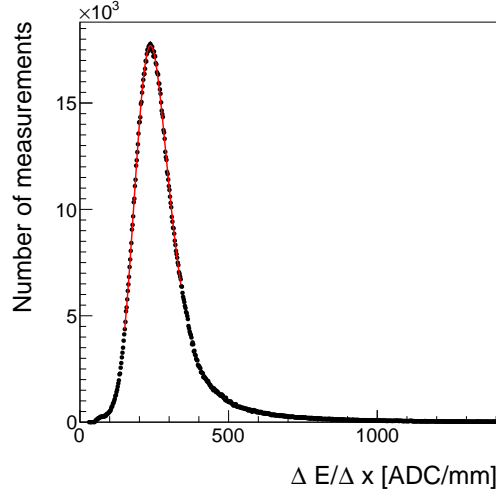


Figure 1.13: An example of the  $\Delta E/\Delta x$  distribution measured in ADC count per mm for one module of the CMS pixel tracker within one time interval. A Landau convoluted with a Gaussian is fitted to the core of the distribution.

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

$$c_{\text{inter}} = \frac{\text{MPV}_{\text{target}} [\text{ADC/mm}]}{\text{MPV} [\text{ADC/mm}]} = \frac{300 \cdot 265 \text{ ADC/mm}}{\text{MPV} [\text{ADC/mm}]} \quad (1.8)$$

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 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

$$\left( \frac{\Delta E}{\Delta x} \right)_{\text{calibrated}} = c_{\text{inter}} \cdot \left( \frac{\Delta E}{\Delta x} \right)_{\text{uncalibrated}} \quad (1.9)$$

The determination of the calibration factor is done for every of the five time intervals, shown in Fig. 1.12 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. 1.14. 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.  $\text{MeV/cm}$ ). 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

$$\Delta E [\text{eV}] = c_{\text{inter}} \cdot \Delta E [\text{ADC}] \cdot \frac{N_e}{\text{ADC}} \cdot 3.61 \text{ eV}, \quad (1.10)$$

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 [24]). The absolute calibration of the silicon pixel tracker can rely on the existing absolute calibration of the silicon strip detector. In [24], 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. (1.5). 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

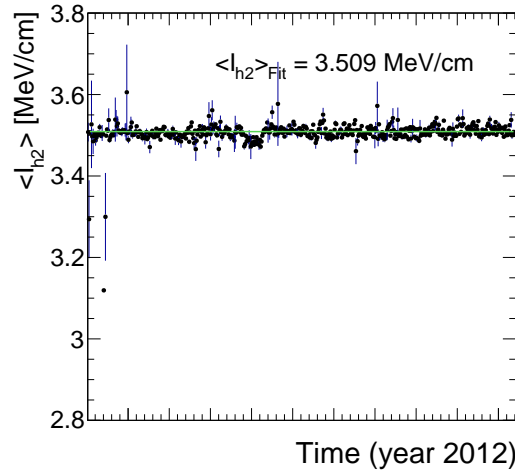


Figure 1.14: Mean of all track's  $dE/dx$  (harmonic-2 estimator) over the full year 2012 after applying the calibration factors, resulting in an average  $dE/dx$  of 3.51  $\text{MeV/cm}$ . Only pixel hits are taken into account. Every data point corresponds to one run.

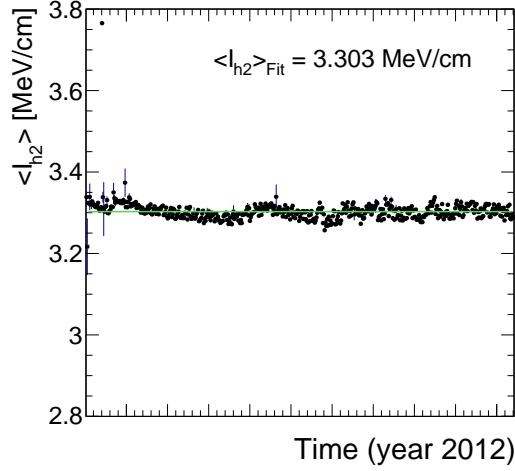


Figure 1.15: Mean of all track's  $dE/dx$  (harmonic-2 estimator) measured in the silicon strip detector over the full year 2012. The average most probable  $dE/dx$  is  $I_{h2} = 3.303 \text{ MeV/cm}$ . Every data point corresponds to one run.

in Fig. 1.14 to the average measured  $dE/dx$  in the strip tracker, shown in Fig. 1.15 by

$$c_{\text{absolute}} = \frac{\langle dE/dx_{\text{strip}} \rangle}{\langle dE/dx_{\text{pixel}} \rangle} = \frac{3.303 \text{ MeV/cm}}{3.509 \text{ MeV/cm}} = 0.941. \quad (1.11)$$

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.

### 1.3.3 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 1.3.1 in Eq. (1.6). It is known to be a robust estimator not easily affected by large fluctuations in  $\Delta E/\Delta x$ . However, it was shown in [24] 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 [26, 27]. 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 [24]

$$I_s = \frac{3}{N} \cdot \left( \frac{1}{12N} + \sum_{i=1}^N \left[ P_i - \frac{2i-1}{2N} \right]^2 \right), \quad (1.12)$$

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 incompatibilities 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 [24] which is only sensitive to incompatibilities to the MIP hypothesis towards higher energy depositions

$$I_{as} = \frac{3}{N} \cdot \left( \frac{1}{12N} + \sum_{i=1}^N \left[ P_i \cdot \left( P_i - \frac{2i-1}{2N} \right)^2 \right] \right). \quad (1.13)$$

A value of  $I_{as}$  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. 1.16 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 [24].

A comparison between the energy release by MIPs ( $I_{as}$ ) in data and simulation for high-quality tracks with  $p > 5$  GeV and  $|\eta| < 2.1$  can be found in Fig. 1.17. The energy loss per path length,  $dE/dx$ , shows good agreement in data and simulation for  $I_{as} < 0.1$ . For larger values,  $I_{as}$  shows a larger decrease in simulation than in measured data. For this reason a data-based approach for analyses exploiting  $dE/dx$  information is needed.

### 1.3.4 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. 1.18

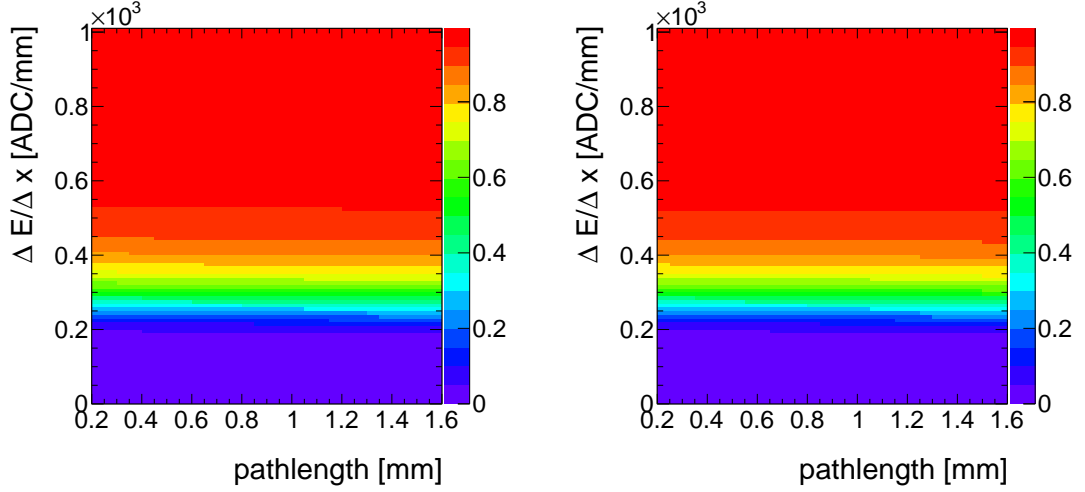


Figure 1.16: Cumulative probability for a MIP to release a  $\Delta E/\Delta x$  (y-axis) vs. the path length (x-axis) in data (left) and simulation (right) for the pixel tracker based on the “minimum bias” sample.

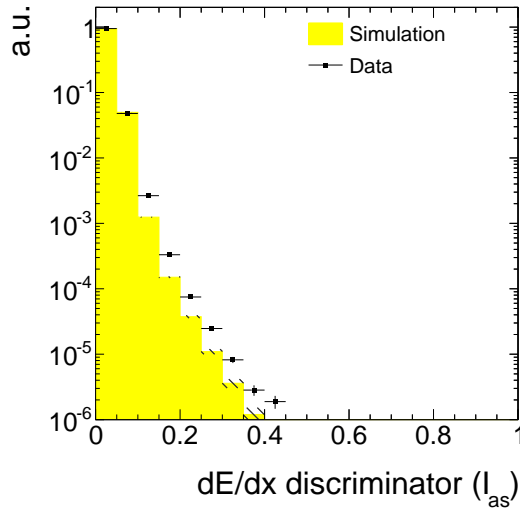


Figure 1.17: Normalised  $I_{\text{as}}$  distribution for MIPs from the minimum bias sample in data and simulation for high-quality (high purity as defined in [28], a minimum number of eight hits and no missing inner and middle hits) tracks with  $p > 5$  GeV and  $|\eta| < 2.1$ .

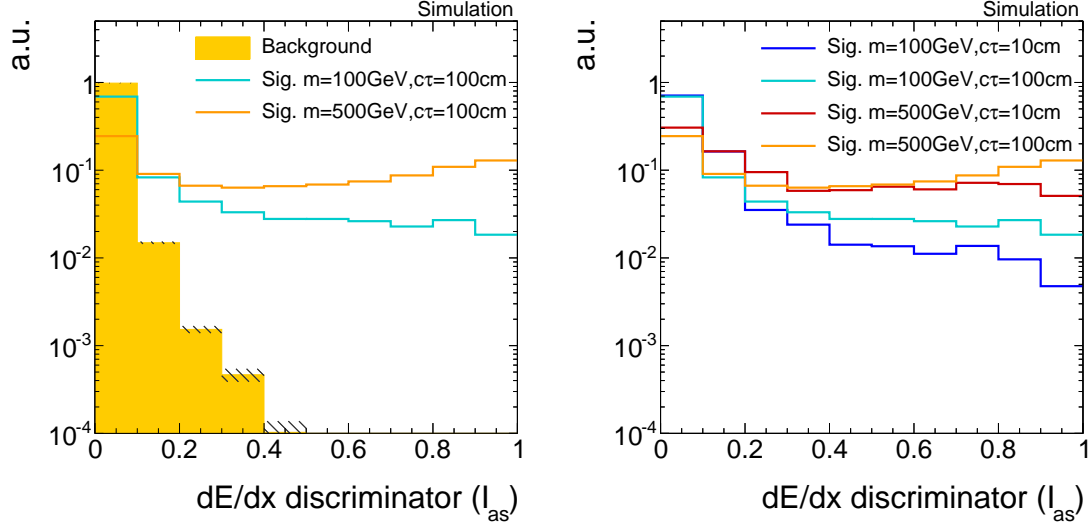


Figure 1.18: Normalised  $I_{as}$  distribution for simulated background and signal tracks (left) and for four different signal models (right) for high-purity tracks (as defined in [28]) with  $p_T > 10$  GeV and  $|\eta| < 2.1$ . For the illustration of the background track spectrum simulated  $t\bar{t}$ +jets events are used (more information about this sample is given in Chapter 1.4).

(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 1.4.2). It can be seen, that the  $I_{as}$  distributions of the signal models show a larger tail towards  $I_{as} = 1$ , whereas the  $I_{as}$  of the background is rapidly falling.

In the right part of Fig. 1.18, a comparison of the  $I_{as}$  distributions of four different signal models is shown. Charginos with longer lifetimes have a more pronounced tail toward  $I_{as} = 1$ . This can be understood with the help of Eq. (1.5), 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.  $c\tau = 10$  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 the survival probability

$$P(t) = e^{-\frac{t}{\gamma\tau}}. \quad (1.14)$$

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  $I_{\text{as}}$  discriminator, because of the long right tail of the Landau distribution. A low number of hits, therefore, leads to higher  $I_{\text{as}}$  values. This effect is also visible in Fig. 1.18 (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 1.19 shows the signal selection efficiency against the background selection efficiency for different selection cuts in  $I_{\text{as}}$ , 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-multijet events (further information about the simulated samples can be found in the next Chapter 1.4). 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).



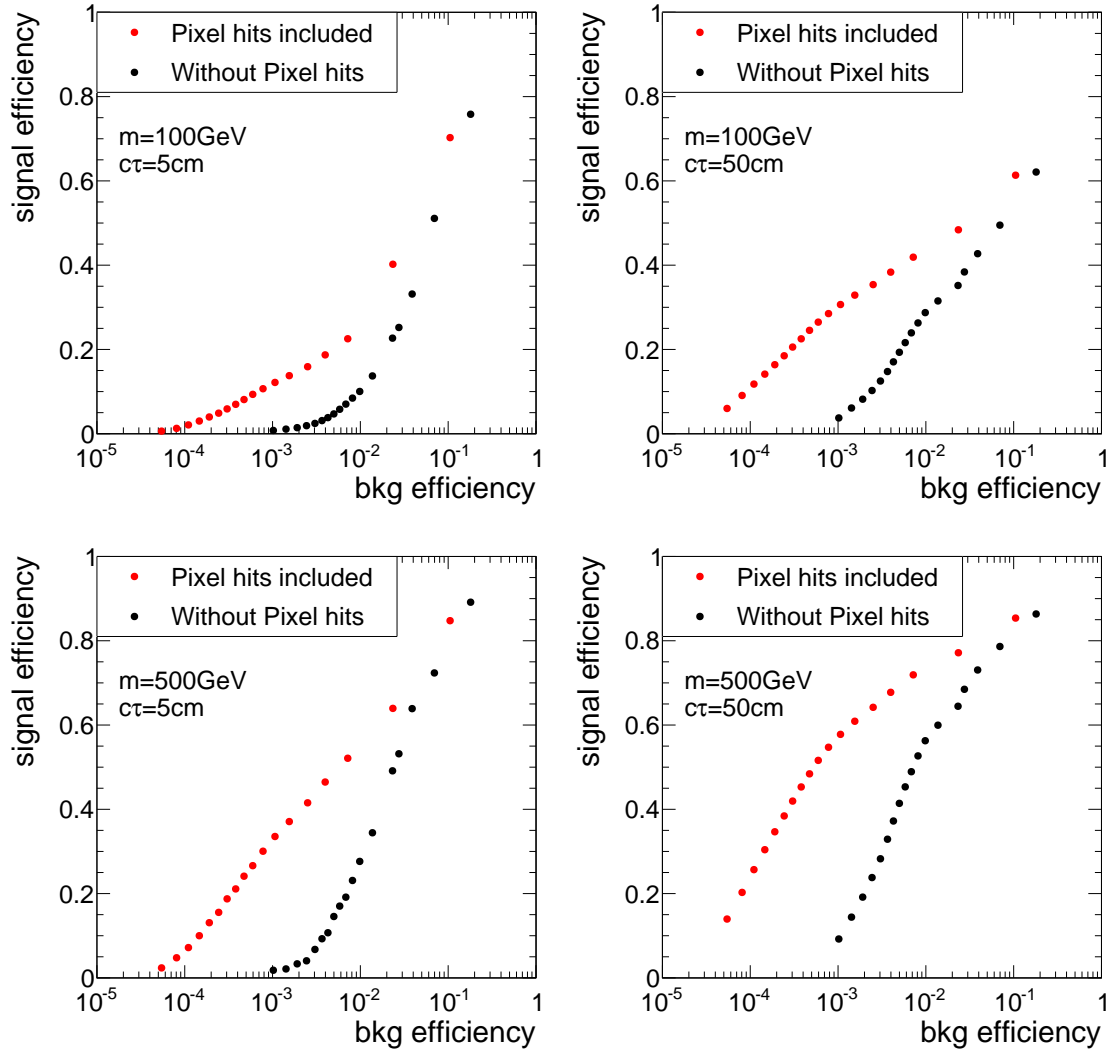


Figure 1.19: Signal selection efficiency vs. background selection efficiency with (red) and without (black) pixel information. Each point correspond to one selection cut in  $I_{\text{as}}$ . The figure is based on a simulated  $W + \text{jets}$  sample and a simulated signal sample with chargino-chargino production, both subject to a selection of high-quality tracks (without a selection on  $N_{\text{hits}}$ ) with  $p_T > 10 \text{ GeV}$ .

## 1.4 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 ??.

The following two sections present an overview of the SM (Section 1.4.1) and SUSY samples (Section 1.4.2) 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.

### 1.4.1 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 [29]. 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  $W$  + jets events is available. Due to the intrinsic missing energy in  $W$  + jets events it constitutes the major background to the presented search (see Section 1.6 for further details on the backgrounds).

In Table 1.1 all available SM samples used in this analysis are listed. The matrix-elements of the  $W$  + jets,  $t\bar{t}$  + jets and  $Z \rightarrow \ell\bar{\ell}$  + jets samples are generated using MADGRAPH 5 [30]. For the QCD sample PYTHIA6 [31] 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 [32,33].

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  $Z \rightarrow \nu\bar{\nu}$  + 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  $W$  + jets sample.

which contain at least one jet with a minimum transverse momentum of  $p_T > 60$  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 1.1: 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 [34]
$t\bar{t} + \text{jets}$	MADGRAPH 5	245.8	NNLO [35]
$Z \rightarrow \ell\bar{\ell} + \text{jets}$ ( $\ell = e, \mu, \tau$ )	MADGRAPH 5	3531.9	NNLO [34]
QCD ( $50 \text{ GeV} < \hat{p}_T < 1400 \text{ GeV}$ )	PYTHIA 6	9374794.2	LO

### 1.4.2 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 1.4.1 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 ( $c\tau = 1 \text{ cm}, 5 \text{ cm}, 10 \text{ cm}, 50 \text{ cm}, 100 \text{ cm}, 1\,000 \text{ cm}$  and  $10\,000 \text{ cm}$ ). 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

$$w = \prod_{i=1}^n \frac{\tau^{\text{gen}}}{\tau^{\text{target}}} \cdot \exp \left[ t_i \cdot \left( \frac{1}{\tau^{\text{gen}}} - \frac{1}{\tau^{\text{target}}} \right) \right], \quad (1.15)$$

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 1.20 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 1.2, 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 [36, 37]. The cross section does not depend on the lifetime of the chargino.

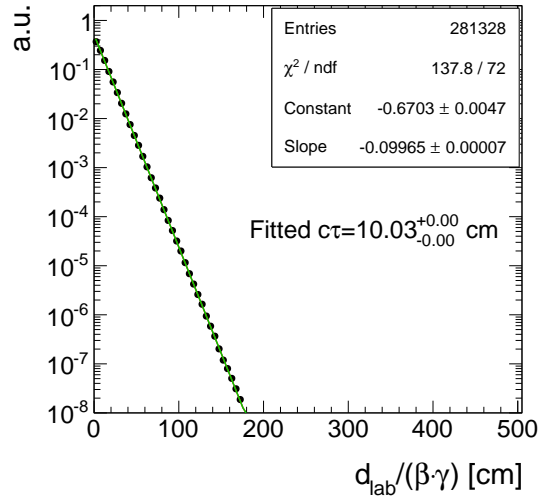


Figure 1.20: Normalised distribution of the proper individual lifetime  $d_{\text{lab}} / (\beta\gamma)$  of all charginos contained in a signal sample with a generated lifetime of  $c\tau^{\text{gen}} = 50 \text{ cm}$  reweighted to a lifetime of  $c\tau^{\text{target}} = 10 \text{ cm}$ . Fitting an exponential curve  $a \cdot \exp\left[\frac{1}{c\tau} ct_i\right]$  yields  $c\tau = \text{slope}^{-1} = 10 \text{ cm}$ .

Table 1.2: 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 [36, 37].

$m_{\tilde{\chi}_1^\pm}$ [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

## 1.5 Event selection

### 1.5.1 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 1.2, 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 ( $p_T^{\text{1st jet}}$ ) 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 1.3 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 1.3:  $\cancel{E}_T$  and  $\cancel{E}_T + \text{jet}$  triggers used in this analysis together with the corresponding recorded integrated luminosity during the time when they were in place.

Trigger	Integrated luminosity [ $\text{fb}^{-1}$ ]
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 1.4. 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 1.4: MET data samples used in the search with the contained integrated luminosity.

Dataset	Integrated luminosity [ $\text{fb}^{-1}$ ]
/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.

## 1.5.2 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 [38, 39]. It relies on event-based and track-based variables as described in the following two sections.

### 1.5.2.1 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 [40] 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 1.5.1). In Fig. 1.21, the distributions of  $\cancel{E}_T$  and the transverse momentum of the leading jet,  $p_T^{1st\text{ jet}}$ , are shown for different signal models. The leading jet has to be centrally produced,  $|\eta_{1st\text{ jet}}| < 2.4$ , and to fulfil the following criteria:

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

These additional jet quality criteria ensure that noise from cosmic and beam halo muons and high- $p_T$  photons and electrons is suppressed [41].

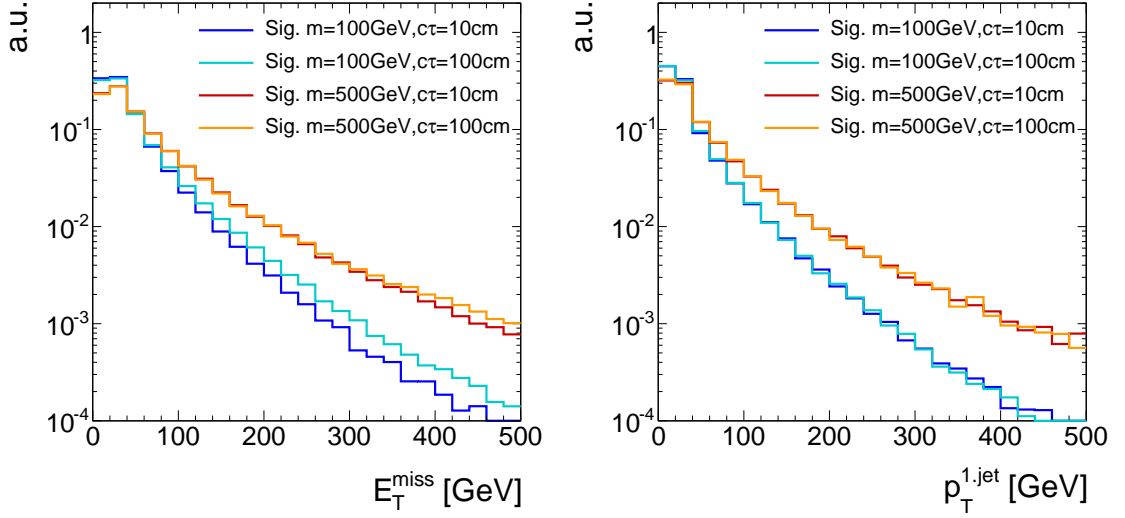


Figure 1.21: Normalised distributions of the missing transverse momentum (left) and the transverse momentum of the leading jet (right) for four different signal models.

The trigger efficiency as a function of  $\cancel{E}_T$  and  $p_T^{1\text{st jet}}$  was determined within [42] with a single-muon reference sample. The trigger paths become fully efficient for  $p_T^{1\text{st jet}} \gtrsim 110$  GeV and  $\cancel{E}_T \gtrsim 220$  GeV [41]. However, it can be seen in Fig. 1.21 that for a selection of  $\cancel{E}_T > 220$  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 GeV which fulfils the above mentioned jet noise cleaning criteria:  $p_T^{1\text{st jet}} > 110$  GeV.
- ❖ The missing transverse momentum must be larger than 100 GeV:  $\cancel{E}_T > 100$  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  $\cancel{E}_T$ , at the  $\cancel{E}_T$  thresholds [41]. 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



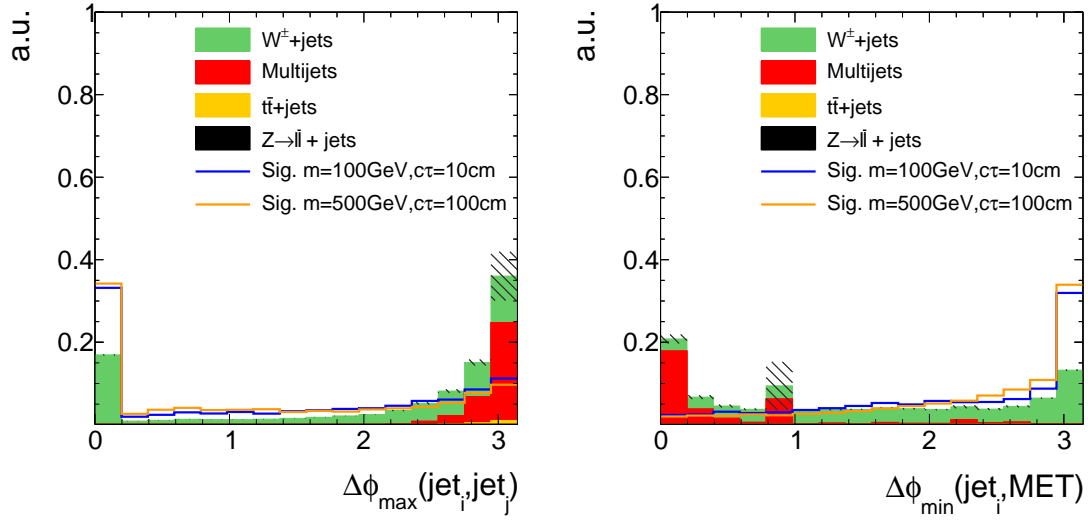


Figure 1.22: Maximum  $\Delta\phi$  between any of two jets (left) and the minimum  $\Delta\phi$  between the  $\cancel{E}_T$  vector and any of the two leading jets (right) normalised to unit area after trigger selection. Only jets with  $p_T > 20$  GeV and  $|\eta| < 4.5$  are considered.

event. Figure 1.22 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.

### 1.5.2.2 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 [28].
- ❖ 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:  $|d0| < 0.02$  cm,  $|dz| < 0.5$  cm.

In Figs. 1.23 and 1.24, the power of the latter two quality selection cuts is shown.

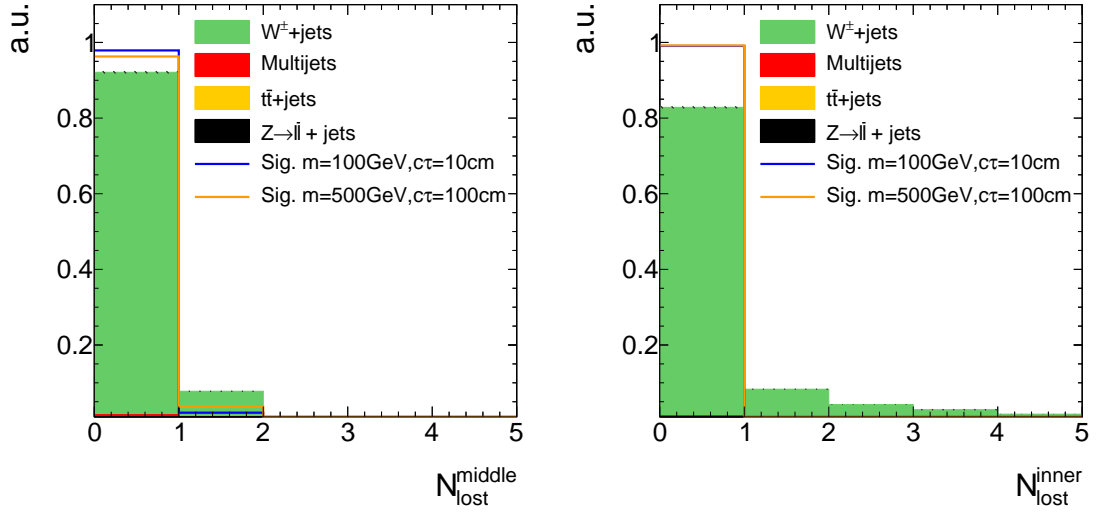


Figure 1.23: Normalised number of missing middle (left) and inner (right) hits of background and signal tracks after trigger selection and QCD suppression cuts.

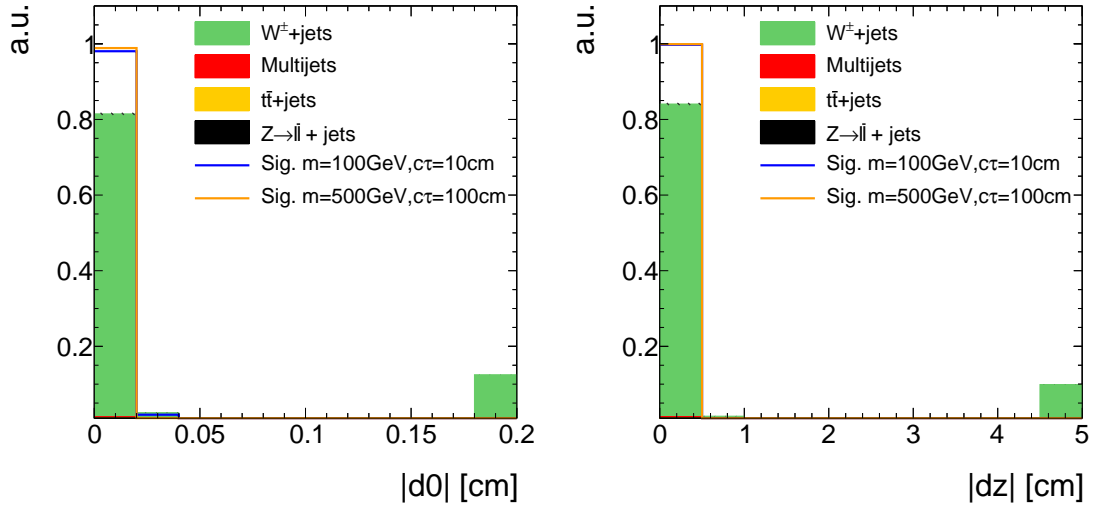


Figure 1.24: Absolute value of the radial (left) and longitudinal (right) distance between the track and the primary vertex after trigger selection and QCD-multijet suppression cuts. All events with a candidate track with a radial (longitudinal) distance larger than 0.2 cm (5 cm) are contained in the 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 ?? 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 ?? 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 ?? 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. 1.25 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 [38, 39] 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. 1.26 showing a map of all ECAL channels not considered in the search.

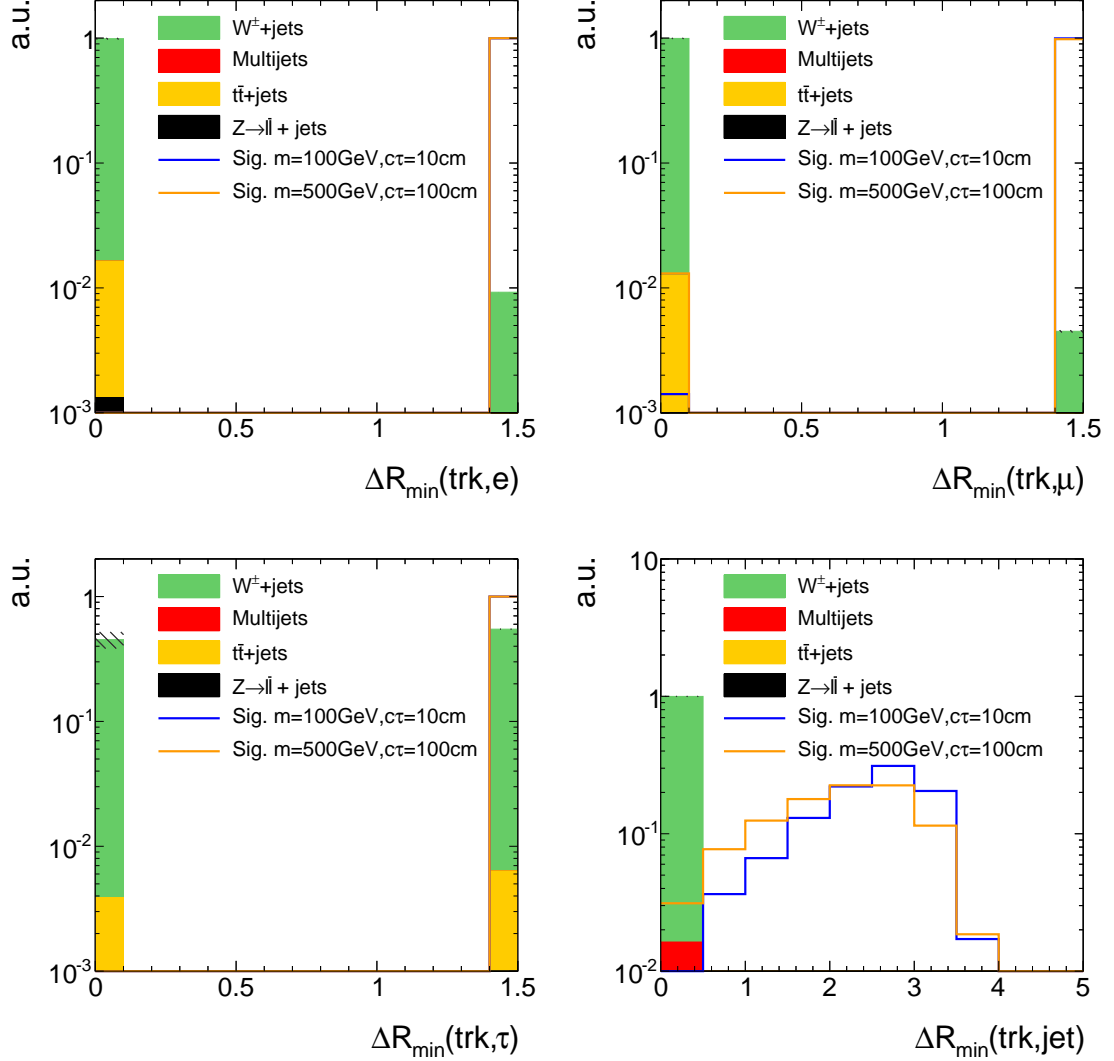


Figure 1.25: The minimum  $\Delta R$  between the candidate track and a reconstructed electron (top left), muon (top right), tau (bottom left) or jet (bottom right) after the event-based selection and the high-quality, kinematic and lepton/jet veto selection of the candidate track selection but without 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.

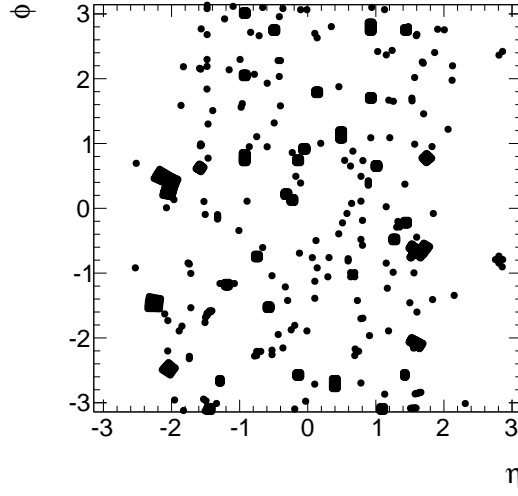


Figure 1.26: Visualisation of dead and noisy ECAL cells in the detector's  $\phi - \eta$  plane according to [38,39]. The radius of the dots correspond to  $\Delta R = 0.05$ .

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 1.5.

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. 1.27. Thus, also tracks pointing towards these regions within a distance of  $\Delta R < 0.25$  are rejected.

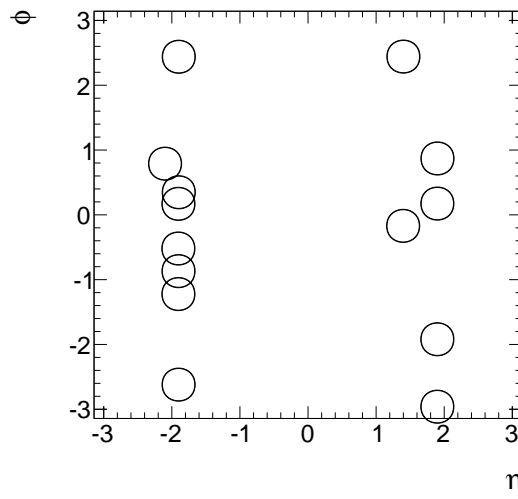
To summarise, tracks pointing towards detector regions, that are not working properly or where the lepton reconstruction efficiencies are reduced, are vetoed as follows:

- ❖ Veto tracks within a cone of  $\Delta R < 0.05$  to a dead or noisy ECAL cell (visualised in Fig. 1.26).
- ❖ Veto tracks that point towards the direction of the ECAL intermodule gap listed in Table 1.5.
- ❖ Veto tracks that point towards a bad CSC within  $\Delta R < 0.25$  (visualised in Fig. 1.27).
- ❖ 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 1.5: 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 1.27: 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. 1.28. 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 1.6. The event yields after the selections of each of the categories from Table 1.6 are listed in

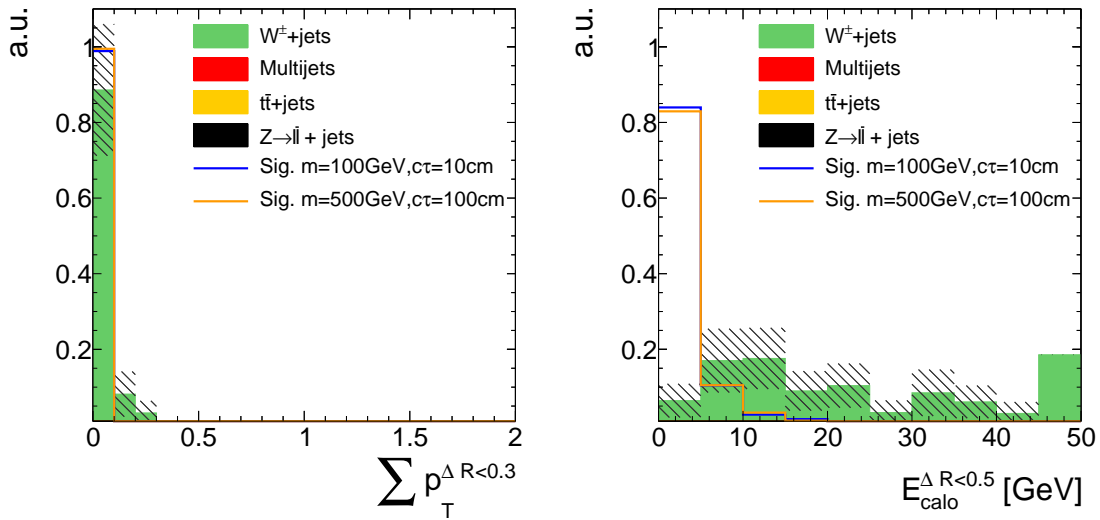


Figure 1.28: Track isolation (left) and calorimeter energy deposits (right) of the candidate track after the full previous selection. All events with a track isolation or a calorimeter energy deposit larger than the range shown in the figures are contained in the last bin.

Table 1.6: 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	$\Delta\phi_{\text{max}}(\text{jet}_i, \text{jet}_j) < 2.7$ for all jets with $p_T > 20 \text{ GeV}$ , $ \eta  < 4.5$ $\Delta\phi_{\text{max}}(\text{jet}_i, \cancel{E}_T) > 0.5$ for two leading jets
Candidate track selection	$\geq 1$ track that fulfils the following criteria:	
	Good quality selection	high-purity as defined in [28] $N_{\text{miss}}^{\text{middle/inner}} = 0$ $ d0  < 0.02 \text{ cm}$ $ dz  < 0.5 \text{ cm}$
	Kinematic selection	$ \eta  < 2.1$ $p_T > 20 \text{ GeV}$
	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 1.7 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  $I_{as}$  is used to enhance the discriminating power of  $dE/dx$ . See Section 1.3.3 for the definition and a detailed explanation of  $I_{as}$ . In Fig. 1.29, 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 1.7. 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.

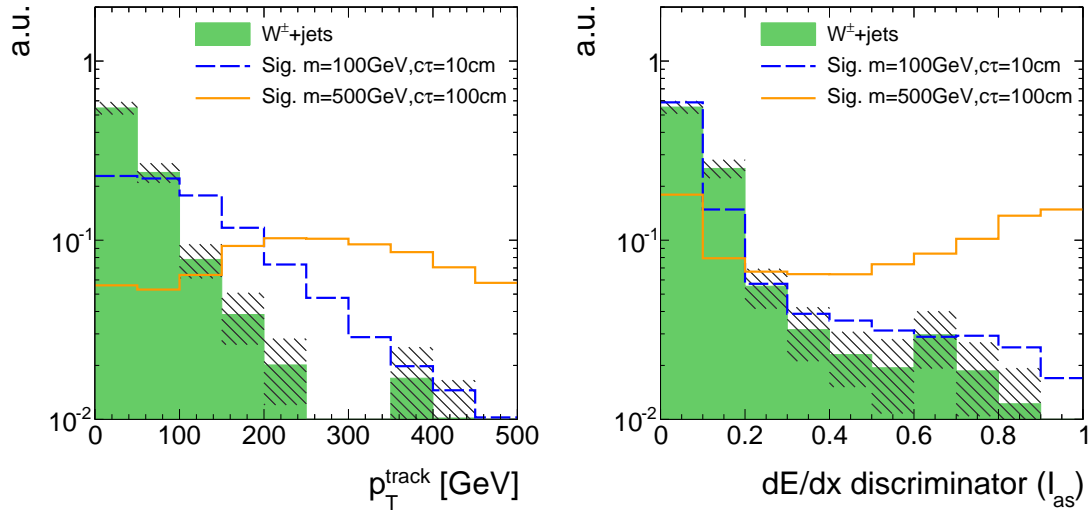


Figure 1.29: Candidate track  $p_T$  (left) and  $I_{as}$  (right) after the full signal candidate selection for signal and  $W + \text{jets}$  events. Because of the low statistical precision of the  $W + \text{jets}$  sample, the trigger selection is not applied.

Table 1.7: Event yields in simulation and data after the selections of each of the categories from Table 1.6

Selection	Simulated background samples				Simulated signal samples				Data
	$W + \text{jets}$	$t\bar{t} + \text{jets}$	$Z \rightarrow \ell\bar{\ell}$	Multijet	$m=100\text{GeV}$ $c\tau=10\text{ cm}$	$m=100\text{GeV}$ $c\tau=100\text{ cm}$	$m=500\text{GeV}$ $c\tau=10\text{ cm}$	$m=500\text{GeV}$ $c\tau=100\text{ cm}$	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	35.7	$3.94 \cdot 10^6$
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	27.1	$1.38 \cdot 10^6$
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

## 1.6 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 vetoes 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 1.6.2.

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 1.6.1. 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 1.6).

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 1.7) to ensure an ideal selection for the search for short and highly ionising tracks.

### 1.6.1 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 ??.

Reconstructed tracks that are fake tracks consist in general only out of a few hits. This can be seen in Fig. 1.30, 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 1.7).

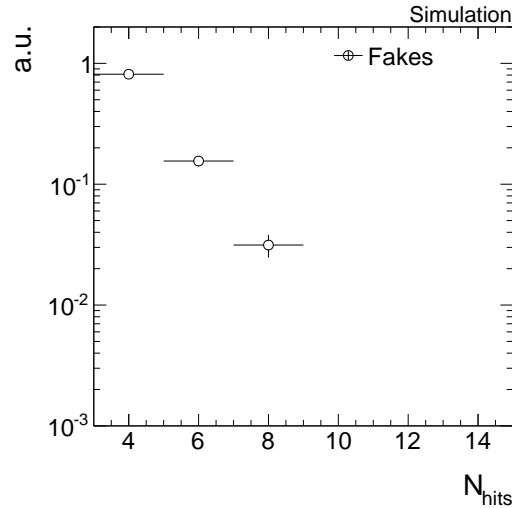


Figure 1.30: Normalised distribution of the number of hits for fake tracks in the simulated  $W + \text{jets}$  sample. To increase the statistical precision, only the candidate track selection from Table 1.6 is applied.

### 1.6.1.1 Inclusive fake background estimation

The inclusive background estimation closely follows the background estimation method done in [38, 39]. It aims at determining the probability of having a fake track in an event that passes the full signal candidate selection (Table 1.6) plus a potential additional  $p_T$  selection cut that is determined in an optimisation procedure (Section 1.7). This probability will be called the fake rate  $\rho_{\text{fake}}$ . Within [38, 39], 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 1.8. These datasets contain at least one muon or one electron in every event.

Table 1.8: 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  $p_T > 25$  GeV and  $|\eta| < 2.4$ . To suppress background from cosmic muons, the distance from the primary vertex must be less than  $|d0| < 0.2$  cm in radial and  $|dz| < 0.5$  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  $Z \rightarrow \mu\bar{\mu}$  + fake track selection is summarised in Table 1.9.

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

Table 1.9: Event selection cuts for the  $Z \rightarrow \mu\bar{\mu}$  + fake control sample to estimate the inclusive fake background.

Event-based selection	Two global muons with $p_T > 25$ 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$ cm $ dz  < 0.5$ cm $\geq 1$ hit in the muon detector considered in global fit $\geq 2$ hits in different muon stations $\geq 1$ hit in the pixel detector $\geq 6$ 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 [43] 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 [44] 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 1.10.

In Fig. 1.31, 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 1.10 and 1.9).

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. 1.32, 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 1.10: 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$
Event-based selection		$\sum_{\Delta R < 0.4} p_T^{\text{PF particle}} / p_T(e) < 0.15$
		pass conversion veto [43]
		no missing inner tracker hits
		good MVA electron as defined in [44]
	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	

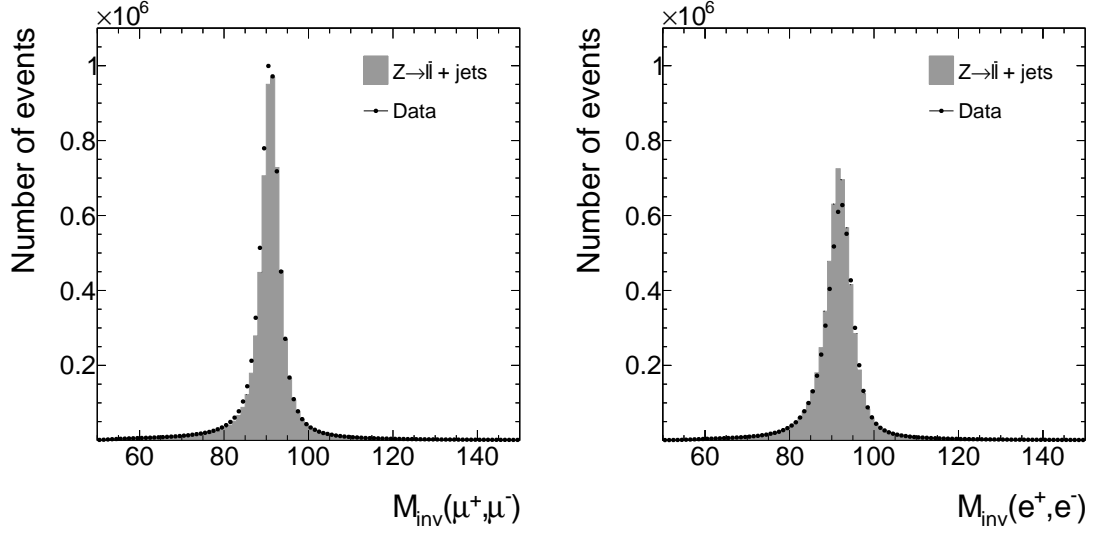


Figure 1.31: The invariant mass of the two selected muons (left) and the two selected electrons (right) after the event-based selection from Tables 1.10 and 1.9, respectively.

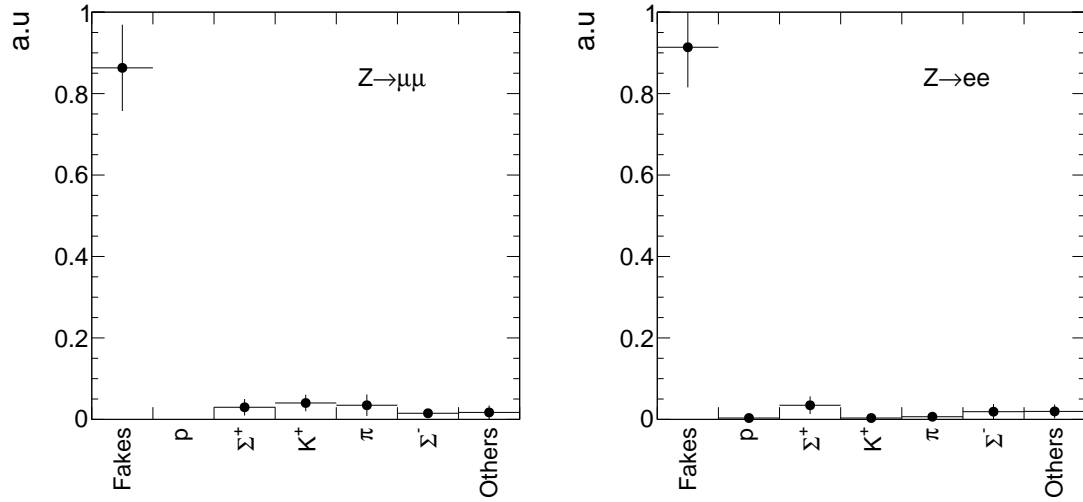


Figure 1.32: Corresponding generator-level particles of all tracks within  $Z \rightarrow \ell\bar{\ell} + \text{fake}$  that were selected according to the candidate track selection. The full selection for tracks in  $Z \rightarrow \mu\bar{\mu}$  events (left) is given in Table 1.9. The full selection for tracks in  $Z \rightarrow e\bar{e}$  events (right) is given in Table 1.10. “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 1.9 (Table 1.10) divided by the number of events that pass only the event-based selection in Table 1.9 (Table 1.10)

$$\rho_{\text{fake}} = \frac{N_{Z \rightarrow \ell\ell}^{\text{cand trk selection}}}{N_{Z \rightarrow \ell\ell}} \quad (1.16)$$

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 1.11 the results of  $N_{Z \rightarrow \ell\ell}^{\text{cand trk selection}}$ ,  $N_{Z \rightarrow \ell\ell}$  and the resulting fake rate for the candidate track selection given in Table 1.6 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 [38, 39] that the fake rate is constant for different Standard Model processes. This is shown in Fig. 1.33 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 1.4) by applying the event-based signal candidate requirements from Table 1.6

$$N_{\text{bkg}}^{\text{fake, inclusive in } I_{\text{as}}} = \rho_{\text{fake}} \cdot N_{\text{event-based selection}}^{\text{MET}} \quad (1.17)$$

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 1.11: Results of  $N_{Z \rightarrow \ell\ell}^{\text{cand trk selection}}$ ,  $N_{Z \rightarrow \ell\ell}$  and  $\rho_{\text{fake}}$  for the  $Z \rightarrow \mu\bar{\mu} + \text{fake}$  (Table 1.9) and  $Z \rightarrow e\bar{e} + \text{fake}$  (Table 1.10) selection.

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

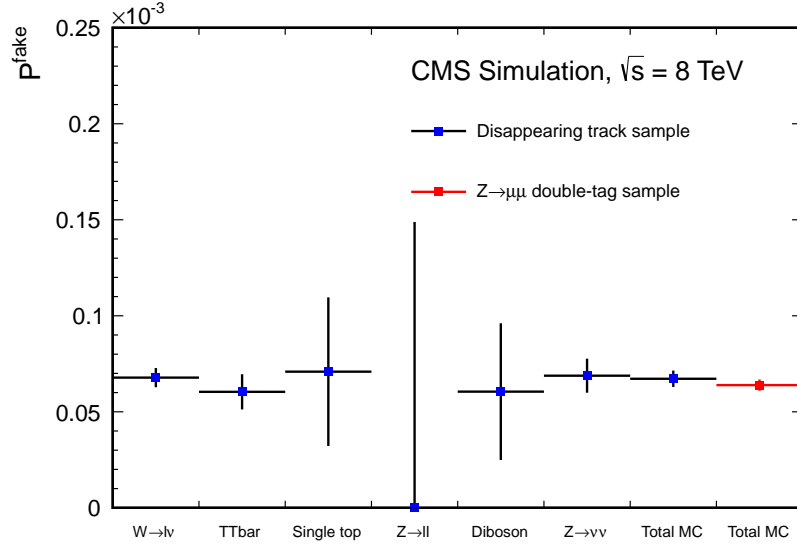


Figure 1.33: Fake track rate estimated in [38,39] for tracks with four hits. Taken from [39]

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 1.6.4.1).

### 1.6.1.2 $dE/dx$ 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. 1.17). Within this analysis the Asymmetric Smirnov discriminator  $I_{as}$  is used to discriminate signal against background with respect to  $dE/dx$  (see Section 1.3.3). In order to estimate the  $I_{as}$  shape of fake tracks, a control region  $CR_{I_{as}}^{fake}$  is defined that is enriched with fakes and shows the same  $I_{as}$  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 ( $N_{miss}^{inner} + N_{miss}^{middle} > 0$ ). Figure 1.34 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  $CR_{I_{as}}^{fake}$  is about 98% (see Fig. 1.35).

Additionally, it must be checked whether the  $I_{as}$  shape in  $CR_{I_{as}}^{fake}$  is representative for the  $I_{as}$  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  $I_{as}$  shape of fake tracks can only be done with simulated events.

Thus, simulated  $W$ +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. 1.35.

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

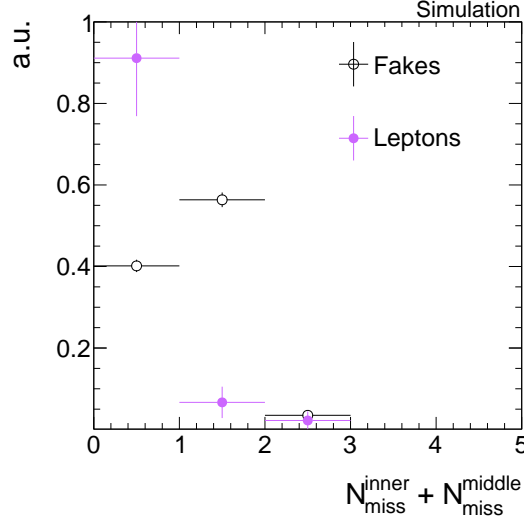


Figure 1.34: Normalised number of missing inner plus missing middle hits for fake and leptonic tracks for the full candidate track selection with the selection requirements on  $N_{\text{miss}}^{\text{inner}}$  and  $N_{\text{miss}}^{\text{middle}}$  removed. Trigger selection and QCD suppression cuts were removed to enhance the statistical precision.

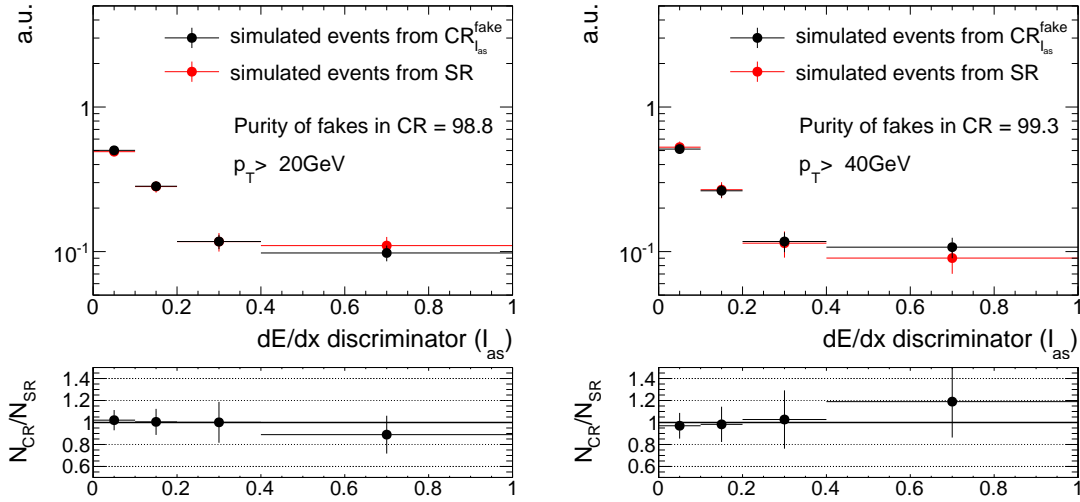


Figure 1.35: Comparison of the  $I_{\text{as}}$  shape between  $\text{CR}_{I_{\text{as}}}^{\text{fake}}$  and the signal region for two different track  $p_T$  selections of  $p_T > 20$  GeV (left) and  $p_T > 40$  GeV (right). To enhance the statistical precision only the track-based selection is applied.

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 1.6.4.2).

## 1.6.2 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 lose 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 1.5.2. 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. 1.36 (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. 1.36 (right)). An ISR jet ( $p_{\text{T}} \simeq 230 \text{ GeV}$ ) causes the  $\cancel{E}_{\text{T}}$  in the event.

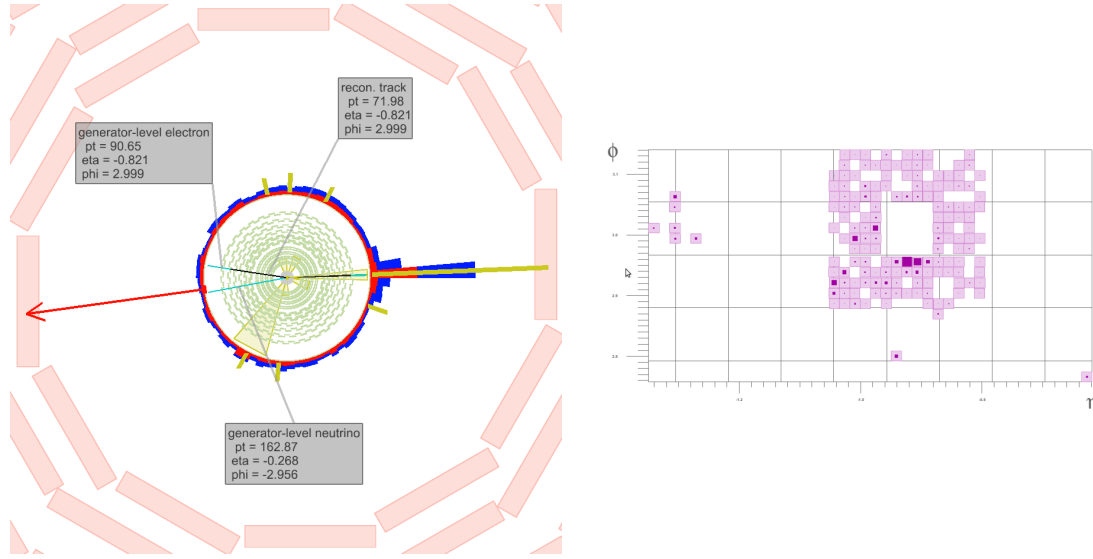


Figure 1.36: Left: Visualisation of a  $W \rightarrow e\nu_e$  event contributing to the SM background. In light blue, generator-level particles including  $e$  and  $\nu_e$  of the  $W$ -boson decay are shown. Black lines represent reconstructed tracks and the red arrow indicates the missing transverse 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. 1.37. 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  $\cancel{E}_T$  in the event.

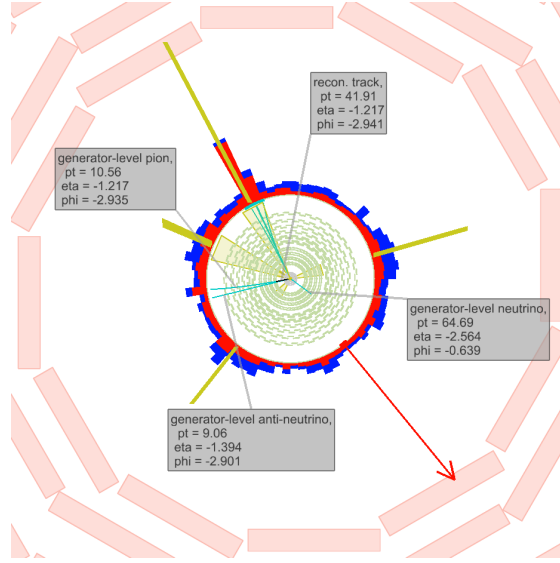


Figure 1.37: Visualisation of a  $W^+ \rightarrow \tau^+ \nu_\tau \rightarrow \pi^+ \bar{\nu}_\tau \nu_\tau$  event contributing to the SM background. In light blue, the generator-level particles including  $\pi^+$ ,  $\bar{\nu}_\tau$  and  $\nu_\tau$  are shown. The black line represents the reconstructed 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 [38, 39]. This possibility is illustrated in a simulated event shown in Fig. 1.38. 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 [38, 39] 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  $I_{\text{as}}$ , muons can be efficiently suppressed. E.g. in the event shown in Fig. 1.38, the muon has an  $I_{\text{as}}$  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  $I_{\text{as}}$  information. Second, the estimation of the  $I_{\text{as}}$  shape for all three leptonic background sources. To have the possibility to make an optimisation in the two main discriminating variables  $p_T$  and  $I_{\text{as}}$ , the background estimation methods are designed to work for all different  $p_T$  and  $I_{\text{as}}$  selection cuts.

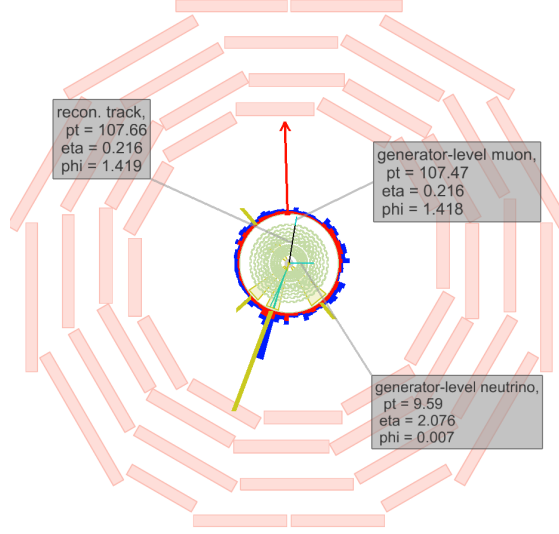


Figure 1.38: Visualisation of an  $W \rightarrow \mu\nu_\mu$  event contributing to the SM background. In light blue, the generator-level particles including  $\mu$  and  $\nu_\mu$  of the  $W$  decay are shown.

### 1.6.2.1 Inclusive leptonic background estimation

The inclusive (without  $dE/dx$  information) lepton background estimation method is similar to the background estimation method used in [38, 39].

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  $W$  + jets events, the ratio  $\rho_{MC}^{\text{lep}_i}$  between the number of events in the signal region with the selected track matched to a generator-level lepton  $N_{SR}^{\text{trk matched to lepton}_i}$  and the number of events in a control region  $N_{CR}^{\text{lepton}_i \text{ veto inverted}}$  with a inverted lepton veto is determined. For muons, this lead to the following expression

$$\rho_{MC}^\mu = \frac{N_{SR,MC}^{\text{trk matched to } \mu}}{N_{CR,MC}^{\mu \text{ veto inverted}}}. \quad (1.18)$$

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

$$\rho_{MC}^{e,\tau} = \frac{N_{SR,MC}^{\text{trk matched to } e,\tau}}{N_{CR,MC}^{e,\tau \text{ veto inverted, } E_{\text{calo}}^{\Delta R < 0.5} < 5 \text{ GeV}}}. \quad (1.19)$$

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

$$N_{\text{predicted}}^{\mu, \text{ inclusive in } I_{\text{as}}} = \rho_{\text{MC}}^{\mu} \cdot N_{\text{CR, data}}^{\mu \text{ veto inverted}}. \quad (1.20)$$

for muons, and

$$N_{\text{predicted}}^{e, \tau, \text{ inclusive in } I_{\text{as}}} = \rho_{\text{MC}}^{e, \tau} \cdot N_{\text{CR, data}}^{e, \tau \text{ veto inverted}, E_{\text{calo}}^{\Delta R < 0.5} < 5 \text{ GeV}}. \quad (1.21)$$

for electrons and taus.

This method relies on the simulation of the lepton reconstruction efficiencies which is known to be reasonably accurate [45–47]. 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 1.6.4.3).

To reduce the statistical uncertainty, the scale factor is calculated without applying the QCD suppression cuts. After the signal candidate selection described in Section 1.5.2, only one event remains in the simulated  $W$  + 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 [21, 48]. Table 1.12 gives the result for the prediction of the inclusive leptonic background for the signal candidate selection from Section 1.5.2.

Table 1.12: 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	$1.25_{-0.77}^{+1.70} \cdot 10^{-4}$	60067	$7.49_{-4.63}^{+10.19}$
muons	$2.17_{-0.93}^{+1.65} \cdot 10^{-4}$	76664	$16.64_{-7.12}^{+12.64}$
taus	$< 2.13 \cdot 10^{-2}$	445	$< 9.46$



### 1.6.2.2 $dE/dx$ shape of leptonic background

In order to get information about the  $I_{as}$  (see Section 1.3.3) 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. 1.39. 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. 1.39 and Fig. 1.40.

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

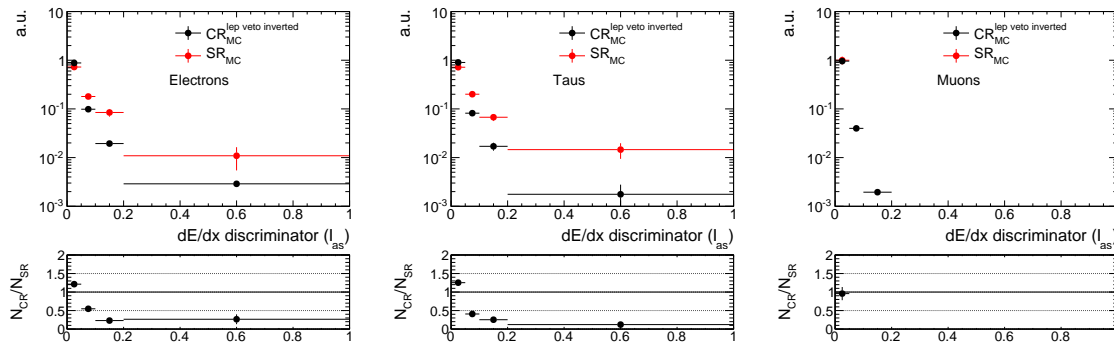


Figure 1.39: Normalised  $I_{as}$  distribution for electrons (left), pions from the tau decay (middle) and muons (right) in the signal region (red) and the lepton veto inverted control region (black).

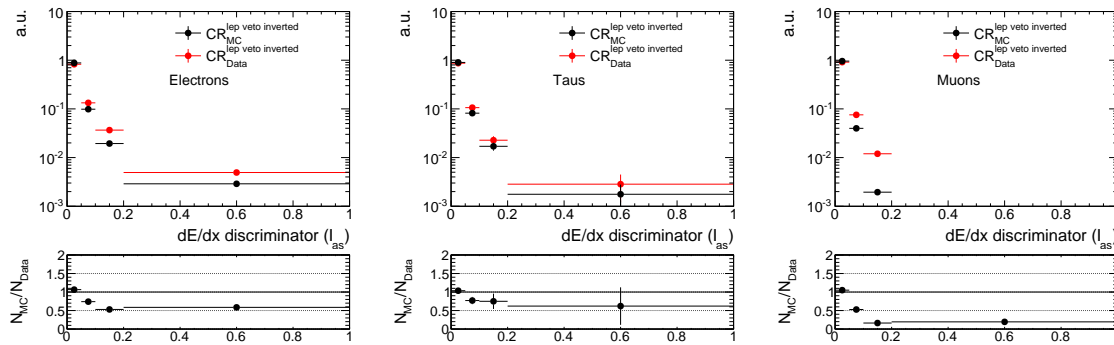


Figure 1.40: Normalised  $I_{as}$  distribution for electrons (left), pions from the tau decay (middle) and muons (right) in the lepton veto inverted control region from simulated events (black) and data (red).

uncertainty is estimated that addresses simulation-data differences of the  $I_{\text{as}}$  distributions. This systematic uncertainty is discussed in Section 1.6.4.4.

### 1.6.3 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 [12] (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. 1.30). 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 1.13. The predicted number of events by the leptonic background estimation is compatible with the observed data yield.

Table 1.13: 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 isola-

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 1.6.4.2. 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. 1.41, 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 1.14, 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  $I_{\text{as}}$

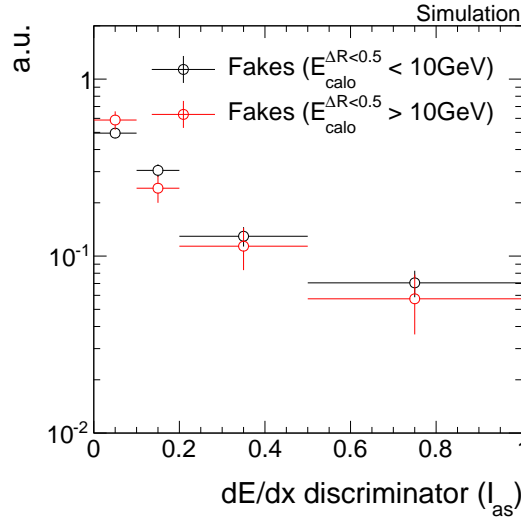


Figure 1.41: Comparison of the  $I_{\text{as}}$  shape of fake tracks with high calorimeter energy deposits (red) and low calorimeter energy deposits (black) in simulation. To enhance the statistical precision, no trigger selection is applied.

Table 1.14: 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.

## 1.6.4 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.

### 1.6.4.1 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 1.9 and 1.10) and the fake rate in the signal candidate selection from Table 1.6 in  $W + \text{jets}$  events are compared.

Unfortunately, the statistical precision of the simulated  $W$ +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  $p_T^{1\text{st jet}}$  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 1.15.

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 1.15: Fake rates in simulated  $W$ +jets and  $Z \rightarrow \ell\bar{\ell}$ +jets events for different event-based selections of the  $W$ +jets sample. The track-based selection is the candidate track selection from Table 1.6.

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

#### 1.6.4.2 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  $W$ +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. 1.42. 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  $I_{\text{as}} > 0.2$  around 25%.

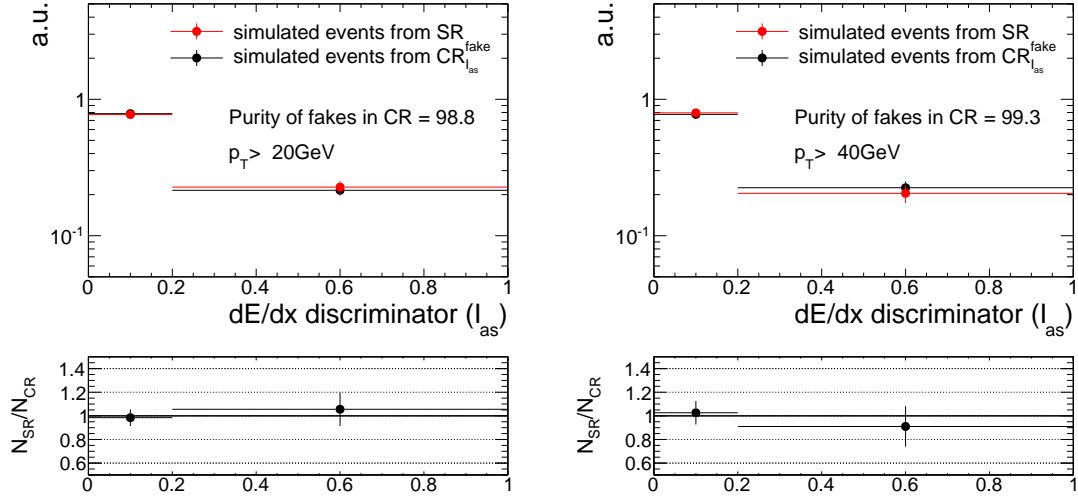


Figure 1.42: Normalised distributions of the  $I_{as}$  shape of fake tracks in the signal and control region of simulated  $W + \text{jets}$  events with a  $p_T$  selection of 20 GeV (left) and a 40 GeV (right).

#### 1.6.4.3 Uncertainty on the leptonic scale factor

The leptonic scale factor  $\rho_{MC}^{\text{lep}_i}$  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  $80 \text{ GeV} < M_{\text{inv}}(\text{lepton, cand. trk}) < 100 \text{ GeV}$  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  $m_T(\mu, \cancel{E}_T) < 40 \text{ GeV}$  [38, 39]. 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  $N_{SR}^{\text{T\&P}}$ , and once inverting the lepton veto selection requirement which gives the number of events in the “tag-and-probe” lepton inverted control region  $N_{CR, \text{lepton veto inverted}}^{\text{T\&P}}$ . 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

$$\rho^\mu = \frac{N_{\text{SR}}^{\text{T\&P}\mu}}{N_{\text{CR}, \mu \text{ veto inverted}}^{\text{T\&P}}}. \quad (1.22)$$

and for electrons and taus

$$\rho^{e,\tau} = \frac{N_{\text{SR}}^{\text{T\&P}e,\tau}}{N_{\text{CR}, e,\tau \text{ veto inverted}}^{\text{T\&P}} \frac{E_{\text{eato}}^{\Delta R < 0.5}}{< 5 \text{ GeV}}}. \quad (1.23)$$

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 1.8. 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. (1.22) and (1.23). In Table 1.16, 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 1.16: 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	$N_{\text{SR}}^{\text{T\&Plep}_i}$	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	$N_{\text{SR}}^{\text{T\&Plep}_i}$	$153.9 \pm 15.4$	$125.1 \pm 15.8$	$9.1 \pm 4.0$
	$N_{\text{CR}, \text{lep}_i \text{ veto inverted}}^{\text{T\&P}}$	$(4.284 \pm 0.003) \cdot 10^6$	$(4.112 \pm 0.003) \cdot 10^6$	$30.9 \pm 7.8$
	$\rho^{\text{lep}_i}$	$(3.59 \pm 0.36) \cdot 10^{-5}$	$(3.04 \pm 0.39) \cdot 10^{-5}$	$(2.95 \pm 1.49) \cdot 10^{-1}$

### 1.6.4.4 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 1.43 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  $I_{as} > 0.2$ .

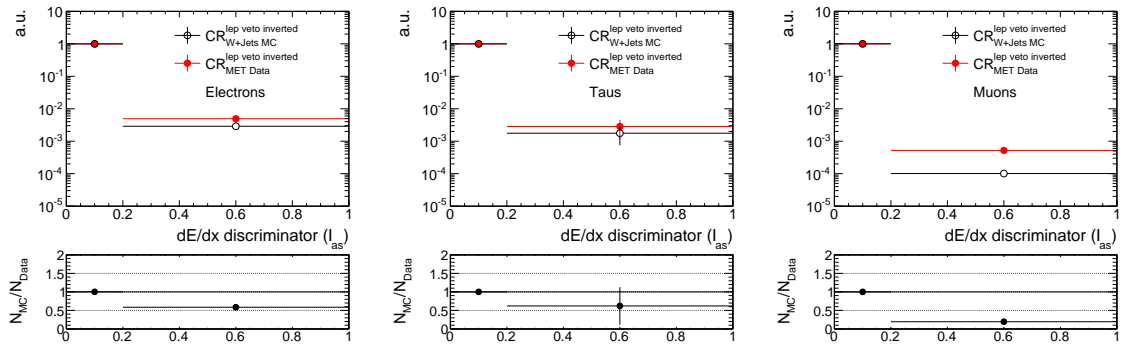


Figure 1.43: Normalised distributions of the lepton  $I_{as}$  distributions in the lepton veto inverted control region for data (red) and simulation (black) for all three lepton types. The event-based selection requirements and the calorimeter isolation requirement are removed to enhance the statistical precision.

## 1.7 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 1.4.2. The optimisation is done in the most sensitive variables,  $p_T$  and  $I_{as}$  (see Section 1.3.3 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

$$\kappa = \frac{\alpha_{\min} \cdot N_S(\text{mass}, c\tau, p_T^{\text{cut}}, I_{as}^{\text{cut}})}{\Delta B(p_T^{\text{cut}}, I_{as}^{\text{cut}})} = 5. \quad \text{with } \alpha_{\min} = \frac{\sigma_{\min}}{\sigma_S}. \quad (1.24)$$

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 [21, 48]. The systematic uncertainty on the background prediction includes systematic uncertainties as described in Section 1.6.4, 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:  $c\tau = 1$  cm, 10 cm, 50 cm. These lifetimes are further suitable as they lie at the edge of the sensitivity of the search for disappearing tracks [12]. 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 1.17. It can be seen that the optimal selection is highly dependent on the signal models. The best sensitivity for low masses ( $\leq 200$  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. 1.18 (right)). For high masses the highest search sensitivity is always achieved by a high  $I_{as}$  selection cut of 0.3.

Table 1.17: Optimal  $p_T$  and  $I_{as}$  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 $I_{as}$ 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  $I_{as}$  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. 1.44, where the minimum cross section that is possible to discover is shown in the  $p_T - I_{as}$  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. (1.24). Similar to the full optimisation, it can be seen that for low masses and low lifetimes, the highest search

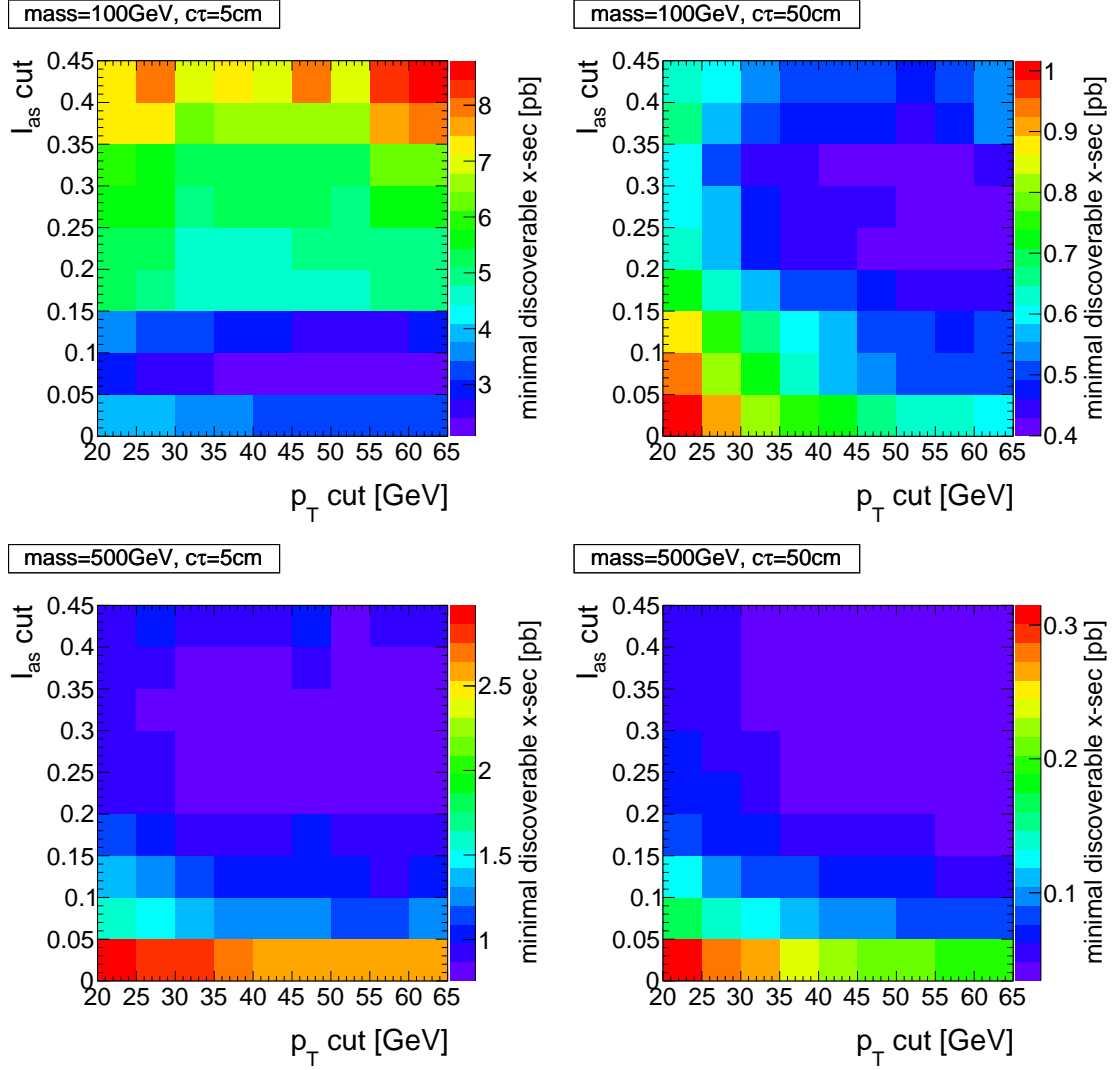


Figure 1.44: Minimum possible cross section that can be discovered with  $5\sigma$  significance as a function of minimum  $p_T$  and  $I_{as}$  requirements for four different signal models. The systematic uncertainties are taken to be 20% and 100% for the fake and the leptonic background respectively. 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  $I_{\text{as}} > 0.3$ .

## 1.8 Results

After developing the methods of the background estimation for all different background sources and their corresponding systematic uncertainties (all explained in Section 1.6), 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 1.18. 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. 1.45, 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 1.18: 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

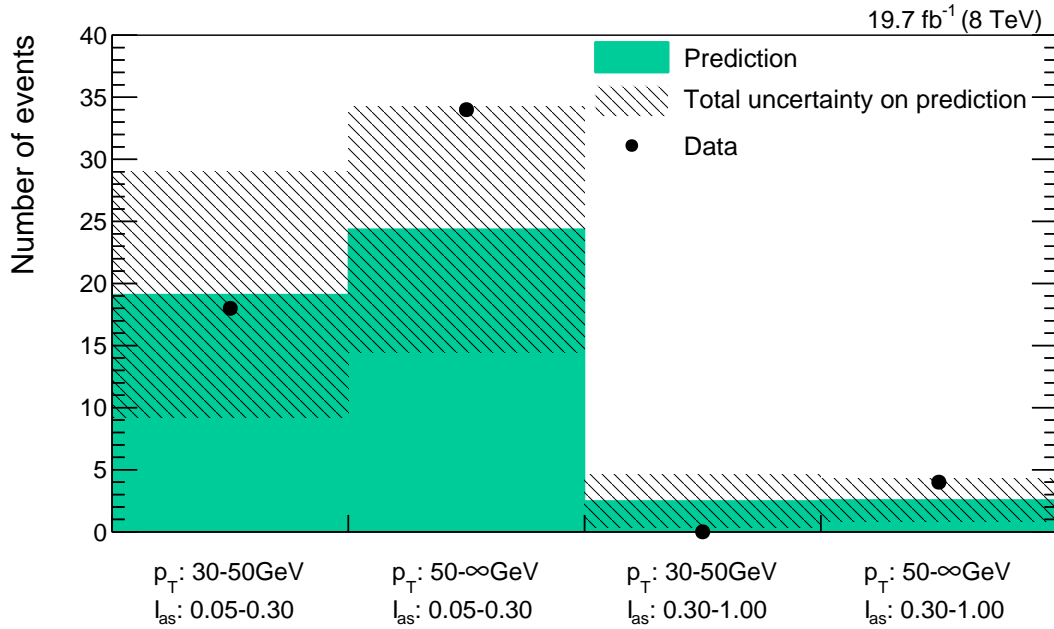


Figure 1.45: Number of predicted (green area) and observed (black dots) events for the four different signal regions. The hashed area represents the total uncertainty on the background prediction.

## 1.9 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 1.4.2.

### 1.9.1 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 1.4.2) 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 [36, 37]. 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 [49].

#### 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 [50] 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 [50] 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 [38,39] 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. 1.46, 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. 1.46 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

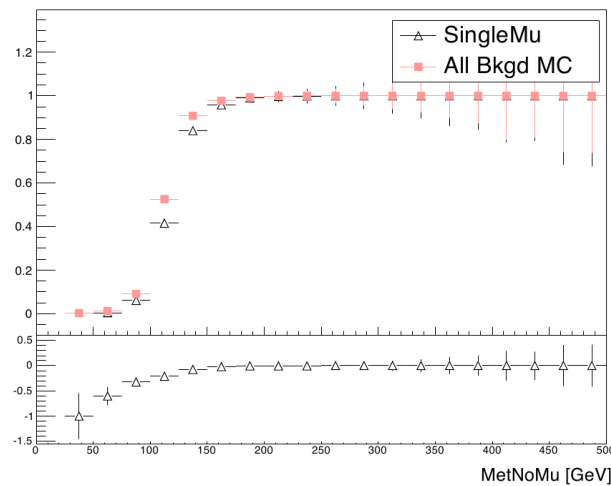


Figure 1.46: Trigger turn-on curves in data (black triangles) and simulation (red squares) as a function of PF  $\cancel{E}_T$  (not taking into account the energy of the muons).

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 [51]. 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 ??). 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 [51, 52]. It covers the uncertainty on JER in data, including the JES uncertainty, uncertainties arising from out-of-cone showering etc. [51, 52]. 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 [53] (see Section ?? for more information about PDFs). In [53], 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 [54, 55] (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 [56]. 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 [57].

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  $p_{\text{T}}$  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 [38, 39], 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  $p_{\text{T}}$ . 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 [13, 58]. 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. 1.47. The kaon and proton line are visible in both datasets. The deuteron line is only visible in data, as deuteron's are not simulated. All of these particles are highly ionising only at low momentum but become minimally ionising for higher momenta. Thus, they don't contribute to the SM background after the final analysis selection with  $p_T > 30$  GeV.

In order to estimate data to simulation differences in  $I_{\text{as}}$ , 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. 1.48. The systematic uncertainty is estimated to a value of 6%.

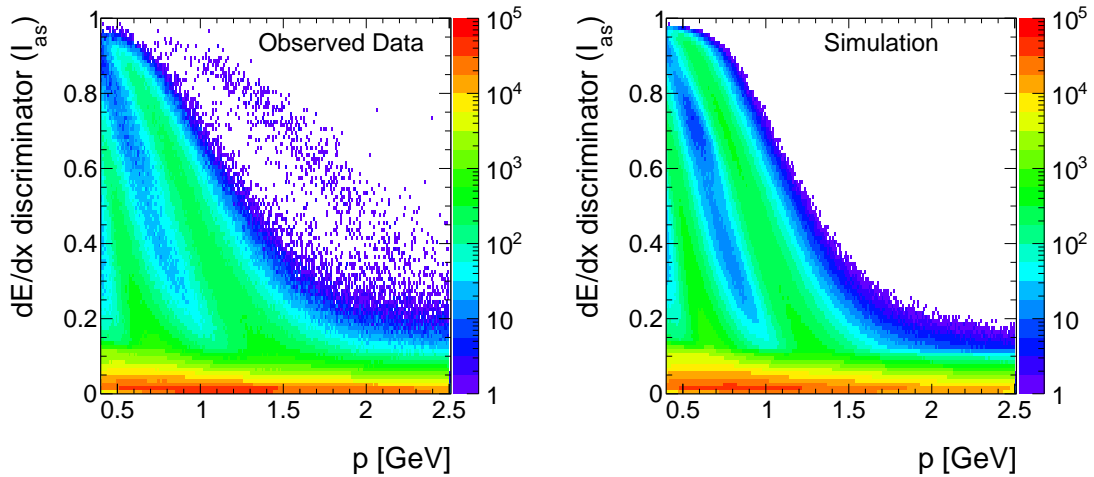


Figure 1.47:  $I_{\text{as}}$  versus momentum for good quality tracks with at least eight hits in observed data (left) and simulation (right).

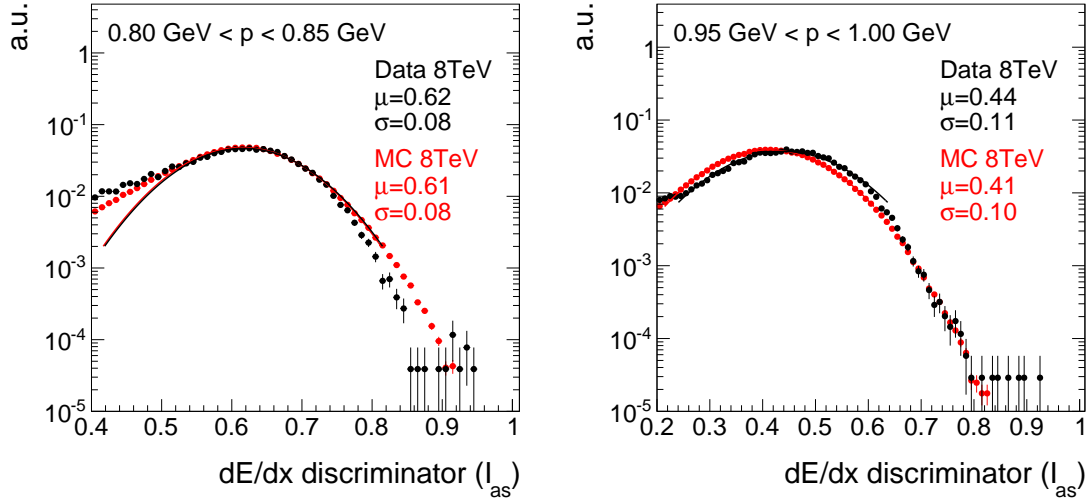


Figure 1.48:  $I_{as}$  distribution for slow protons in simulation and observed data for a momentum range of 0.80 – 0.85 GeV (left) and 0.95 – 1.00 GeV (left). For the momentum range of 0.80 – 0.85 GeV, the proton line is contained between  $I_{as}$  values of 0.4 – 0.8, whereas for the momentum range of 0.95 – 1.00 GeV, the proton line  $I_{as}$  lies between 0.2 – 0.6.

### 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 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 1.19.

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  $I_{\text{as}}$ . 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 1.19: 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

### 1.9.2 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 [59].

In this analysis, the exclusion of the underlying theoretical model is achieved with the

$\text{CL}_s$  method [60–62]. 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

$$\text{CL}_s = \frac{\text{CL}_{s+b}}{\text{CL}_b}. \quad (1.25)$$

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

$$\begin{aligned} \text{CL}_{s+b} &= \sum_0^{n_{\text{obs}}} q_\mu = \text{Poisson}(n \leq n_{\text{obs}} | \lambda = b + \mu \cdot s), \\ \text{CL}_b &= \sum_0^{n_{\text{obs}}} q_{\mu=0} = \text{Poisson}(n \leq n_{\text{obs}} | \lambda = b), \end{aligned} \quad (1.26)$$

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  $q_{\mu=0}$

$$\begin{aligned} q_\mu &= \text{Poisson}(n | \lambda = b \cdot (1 + \delta_b) + \mu \cdot s) \cdot \text{Gauss}(\delta_b | \text{mean} = 0, \sigma = \sigma_b), \\ q_{\mu=0} &= \text{Poisson}(n | \lambda = b \cdot (1 + \delta_b)) \cdot \text{Gauss}(\delta_b | \text{mean} = 0, \sigma = \sigma_b), \end{aligned} \quad (1.27)$$

These expressions can be generalised for more than one signal region and more than one systematic uncertainty [59]. 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 [63]. Instead of the number of

observed events, the profile likelihood ratio

$$q(n, \mu) = -2 \ln \frac{\mathcal{L}(n|\mu, \hat{\delta}_b^\mu, \hat{\delta}_s^\mu)}{\mathcal{L}(n|\hat{\mu}, \hat{\delta}_b, \hat{\delta}_s)}, \quad \text{with the constraint } 0 \leq \hat{\mu} \leq \mu \quad (1.28)$$

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

$$\mathcal{L}(n|b, s, \delta_b, \delta_s) = \text{Poisson}(n|\lambda = b \cdot (1 + \delta_b) + \mu \cdot s(1 + \delta_s)) p(\delta_b) p(\delta_s). \quad (1.29)$$

that already appeared in Eq. (1.26). 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 [64].

Correlations between systematic uncertainties on the background expectation in different search bins are assumed as shown in Table 1.20. The systematic uncertainties on the 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 [64] which was developed for the Higgs searches at CMS.

Table 1.20: 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 $p_{\text{T}}$	100% for bins with same $p_{\text{T}}$	100% for bins with same $p_{\text{T}}$

### 1.9.3 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 1.2.

In total, 37 different lifetimes from  $c\tau = 1 - 10000$  cm for each mass point (100–600 GeV in steps of 100 GeV) are considered. Four exemplary exclusion limits are shown in Fig. 1.49, 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 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. 1.50.

Charginos with masses of 100 GeV can be excluded down to a lifetime of  $c\tau = 2$  cm. Charginos with a higher mass of 500 GeV are excluded for lifetimes between  $c\tau = 70 - 500$  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 [65], 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  $\sim 140$  MeV. The corresponding 2d exclusion limit can be found in Fig. 1.51. It can be seen that this search is sensitive to mass splittings between  $\sim 140$  MeV – 210 MeV.



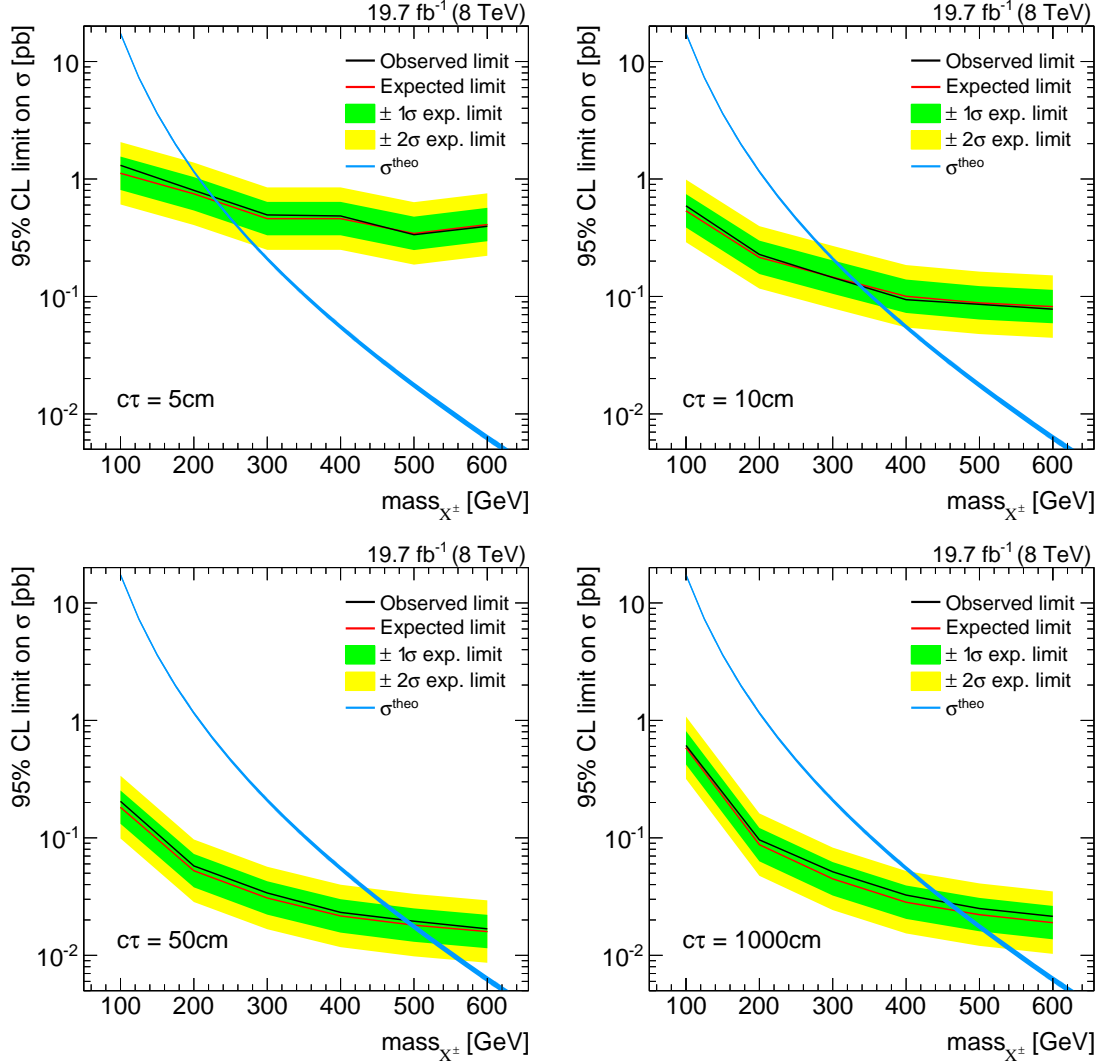


Figure 1.49: Four different  $\text{CL}_s$  exclusion limits for charginos with mean lifetimes of 5 cm (top left), 10 cm (top right), 50 cm (bottom left), 1000 cm (bottom right). The red line depicts the expected 95% confidence level (CL) upper cross-section limit with the 1- $\sigma$  (green band) and 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.

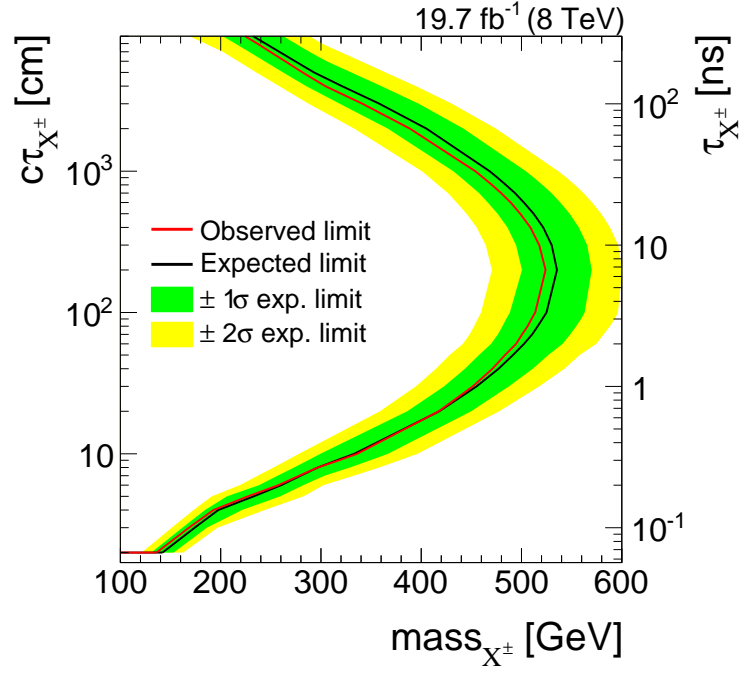


Figure 1.50: Excluded regions in the mass versus lifetime space. All excluded models are located left of the contour line. The red line depicts the expected 95% CL upper cross-section limit with the 1- $\sigma$  (green band) and 2- $\sigma$  (yellow band) intervals. The black line is the observed limit.

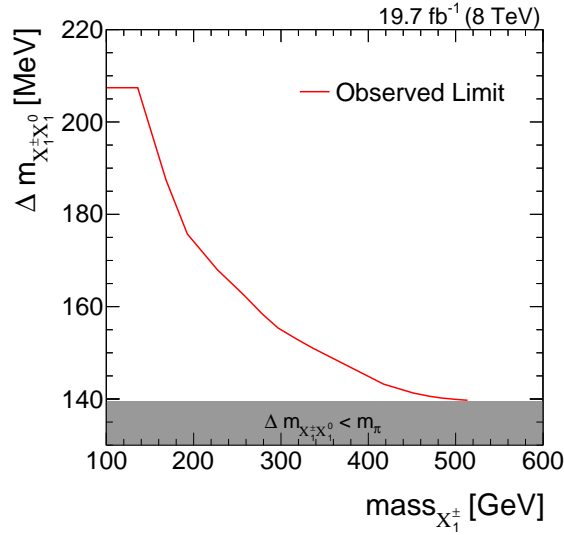


Figure 1.51: Excluded parameter region at 95% CL for wino-like charginos and neutralinos depending on the chargino mass and the mass splitting between  $\tilde{\chi}_1^\pm$  and  $\tilde{\chi}_1^0$ ,  $\Delta m_{\tilde{\chi}_1^\pm \tilde{\chi}_1^0}$ . All SUSY models between the red line and the grey area are excluded.

### 1.9.4 Comparison to the search for disappearing tracks

The presented exclusion limits confirm the exclusion from the search for disappearing tracks [12] with slight improvements in the low lifetime region. The comparison of the two searches is shown in Fig. 1.52.

For charginos with a lifetime of  $\tau = 0.07 \text{ ns}$  ( $c\tau = 2.1 \text{ cm}$ ), the observed limit of this search improves the limits derived in [12] by  $\sim 35 \text{ GeV}$  in chargino mass, for a lifetime of  $\tau = 0.4 \text{ ns}$  ( $c\tau = 12.0 \text{ cm}$ ) by  $\sim 25 \text{ GeV}$ . For SUSY models with long chargino lifetimes the here presented search shows a higher exclusion power. The weaker exclusion for long lifetimes in [12] 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  $c\tau = 5 \text{ cm}$ . The event overlap between simulated signal events, that pass the selection from [12],  $N_A$ , and the selection used in this analysis,  $N_B$ , can be estimated by

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

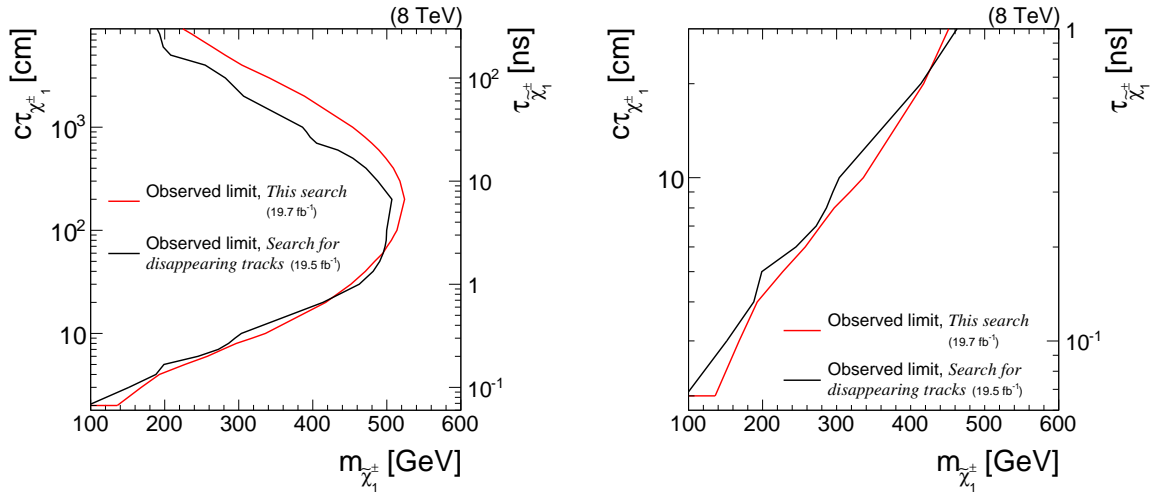


Figure 1.52: Comparison of the excluded regions in the mass versus lifetime space in this analysis (red line) and the search for disappearing tracks [12] (black line). The right figure is a zoom on the low lifetime region. All SUSY models left of the lines are excluded.

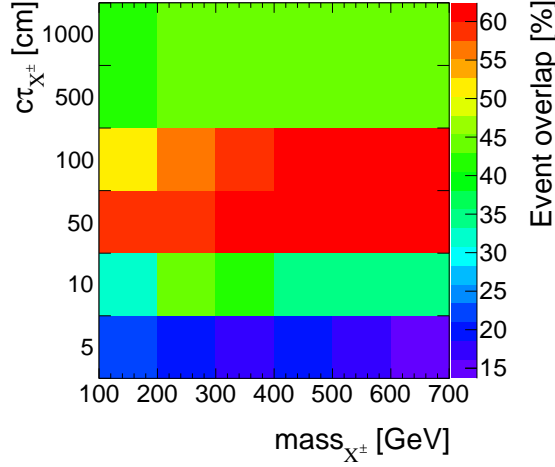


Figure 1.53: The event overlap between simulated signal events, that pass the selection from [12] and the selection used in this analysis for different signal models. The correlation is determined using only the signal region with the highest sensitivity of this analysis.

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 cm – 1000 cm is depicted in Fig. 1.53. It can be seen that the event overlap for intermediate lifetimes of around 100 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 [12] 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 [12], especially for short chargino lifetimes.

## 1.10 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 [10] 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 [12] 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. 1.9). 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. 1.19).

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. The observed number of events in data of 18 and 34 events in the low  $I_{\text{as}}$  signal regions and 0 and 4 events in the high  $I_{\text{as}}$  regions is compatible with SM expectation within  $1\sigma$ . Thus, no evidence for physics beyond the Standard Model is found. 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 [12]. Interestingly, the signal regions of the here presented search and the search from [12] show a rather small overlap for short chargino lifetimes. Therefore, this analysis yields a complementary result with respect to the search for disappearing tracks [12]. 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. First, the here presented search can be combined with the search for disappearing tracks [12]. Since both searches target similar SUSY models and show little event overlap in their final selection, a combination is expected to increase the exclusion power.

Second, 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.

Third, 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-1.

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.





## Bibliography

- [1] CDF, D0 Collaboration, T. Adams, “SUSY Searches at the Tevatron”, in *Hadron collider physics. Proceedings, 19th Symposium, HCP2008, Galena, USA, May 27-31, 2008*. 2008. [arXiv:0808.0728](#).
- [2] ALEPH, DELPHI, L3 and OPAL Collaborations, “Joint SUSY Working Group”. <http://lepsusy.web.cern.ch/lepsusy>. LEPSUSYWG (June 2004).
- [3] CMS Collaboration, “Search for new physics in the multijet and missing transverse momentum final state in proton-proton collisions at  $\sqrt{s} = 8$  TeV”, *JHEP* **06** (2014) 055, [arXiv:1402.4770](#). doi:10.1007/JHEP06(2014)055.
- [4] CMS Collaboration, “Searches for supersymmetry using the  $M_{T2}$  Variable in Hadronic Events Produced in pp Collisions at 8 TeV”, *JHEP* **05** (2015) 078, [arXiv:1502.04358](#). doi:10.1007/JHEP05(2015)078.
- [5] ATLAS Collaboration, “Search for squarks and gluinos with the ATLAS detector in final states with jets and missing transverse momentum using  $\sqrt{s} = 8$  TeV proton-proton collision data”, *JHEP* **09** (2014) 176, [arXiv:1405.7875](#). doi:10.1007/JHEP09(2014)176.
- [6] M. Ibe, S. Matsumoto, S. Shirai et al., “Wino dark matter in light of the AMS-02 2015 data”, *Phys. Rev.* **D91** (2015), no. 11, 111701, [arXiv:1504.05554](#). doi:10.1103/PhysRevD.91.111701.
- [7] J. Hisano, S. Matsumoto, M. M. Nojiri et al., “Direct detection of the Wino and Higgsino-like neutralino dark matters at one-loop level”, *Phys. Rev.* **D71** (2005) 015007, [arXiv:hep-ph/0407168](#). doi:10.1103/PhysRevD.71.015007.
- [8] T. Cohen, M. Lisanti, A. Pierce et al., “Wino dark matter under siege”, *JCAP* **1310** (2013) 061, [arXiv:1307.4082](#). doi:10.1088/1475-7516/2013/10/061.
- [9] T. Moroi and L. Randall, “Wino cold dark matter from anomaly mediated SUSY breaking”, *Nucl. Phys.* **B570** (2000) 455–472, [arXiv:hep-ph/9906527](#). doi:10.1016/S0550-3213(99)00748-8.
- [10] T. Moroi, M. Nagai, and M. Takimoto, “Non-thermal production of Wino dark matter via the decay of long-lived particles”, *JHEP* **07** (2013) 066, [arXiv:1303.0948](#). doi:10.1007/JHEP07(2013)066.

- 
- [11] M. W. Cahill-Rowley, J. L. Hewett, A. Ismail et al., “More energy, more searches, but the phenomenological MSSM lives on”, *Phys. Rev.* **D88** (2013), no. 3, 035002, [arXiv:1211.1981](#). doi:10.1103/PhysRevD.88.035002.
- [12] CMS Collaboration, “Search for disappearing tracks in proton-proton collisions at  $\sqrt{s} = 8$  TeV”, *JHEP* **01** (2015) 096, [arXiv:1411.6006](#). doi:10.1007/JHEP01(2015)096.
- [13] CMS Collaboration, “Searches for long-lived charged particles in pp collisions at  $\sqrt{s} = 7$  and 8 TeV”, *JHEP* **07** (2013) 122, [arXiv:1305.0491](#). doi:10.1007/JHEP07(2013)122.
- [14] CMS Collaboration, “Phenomenological MSSM interpretation of the CMS 2011  $5\text{ fb}^{-1}$  results”, *CMS Physics Analysis Summary CMS-PAS-SUS-12-030* (2013).
- [15] CMS Collaboration, “Reinterpreting the results of the search for long-lived charged particles in the pMSSM and other BSM scenarios”, *CMS Physics Analysis Summary CMS-PAS-EXO-13-006* (2014).
- [16] CMS Collaboration, “Search for disappearing tracks in proton-proton collisions at  $\sqrt{s} = 8$  TeV”, *Public CMS Wiki* (2015). <https://twiki.cern.ch/twiki/bin/view/CMSPublic/PhysicsResultsEXO12034>, Topic revision: r9.
- [17] CMS Collaboration, “Search for heavy long-lived charged particles in pp collisions at  $\sqrt{s} = 7$  TeV”, *Phys. Lett.* **B713** (2012) 408–433, [arXiv:1205.0272](#). doi:10.1016/j.physletb.2012.06.023.
- [18] CMS Collaboration, “Description and performance of track and primary-vertex reconstruction with the CMS tracker”, *JINST* **9** (2014), no. 10, P10009, [arXiv:1405.6569](#). doi:10.1088/1748-0221/9/10/P10009.
- [19] H. Bethe, “Theory of the passage of fast corpuscular rays through matter”, *Annalen Phys.* **5** (1930) 325–400. [Annalen Phys. 397,325(1930)]. doi:10.1002/andp.19303970303.
- [20] “National Institute of Standards and Technology”. <http://physics.nist.gov/cgi-bin/Star/compos.pl?mode=text&matno=014>. Accessed: 2015-10-21.
- [21] Particle Data Group Collaboration, “Review of Particle Physics”, *Chin. Phys.* **C38** (2014) 090001. doi:10.1088/1674-1137/38/9/090001.

- [22] L. Landau, “On the energy loss of fast particles by ionization”, *J. Phys.(USSR)* **8** (1944) 201–205.
- [23] H. Bichsel, “Straggling in thin silicon detectors”, *Rev. Mod. Phys.* **60** (1988) 663–699. doi:10.1103/RevModPhys.60.663.
- [24] L. Quertenmont, “Search for Heavy Stable Charged Particles with the CMS detector at the LHC”. PhD thesis, Louvain University, 2010.  
<https://inspirehep.net/record/1088192/files/openfile.pdf>.
- [25] D. Kotlinski. personal communication.
- [26] CMS Collaboration, “Tracking and Vertexing Results from First Collisions”, *CMS Physics Analysis Summary CMS-PAS-TRK-10-001* (2010).
- [27] T. W. Anderson, “On the Distribution of the Two-Sample Cramér-von Mises Criterion”, *The Annals of Mathematical Statistics* **33** (1962), no. 3, pp. 1148–1159.
- [28] F. James, “Statistical methods in experimental physics”. 2006.
- [29] CMS Collaboration, “Data Formats and Data Tiers”, *Public CMS Wiki*.  
<https://twiki.cern.ch/twiki/bin/view/CMSPublic/WorkBookDataFormats>, Topic revision: r30.
- [30] J. Alwall, R. Frederix, S. Frixione et al., “The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations”, *JHEP* **07** (2014) 079, arXiv:1405.0301.  
doi:10.1007/JHEP07(2014)079.
- [31] T. Sjostrand, S. Mrenna, and P. Z. Skands, “PYTHIA 6.4 Physics and Manual”, *JHEP* **05** (2006) 026, arXiv:hep-ph/0603175.  
doi:10.1088/1126-6708/2006/05/026.
- [32] GEANT4 Collaboration, “GEANT4: A Simulation toolkit”, *Nucl. Instrum. Meth.* **A506** (2003) 250–303. doi:10.1016/S0168-9002(03)01368-8.
- [33] J. Allison et al., “Geant4 developments and applications”, *IEEE Trans. Nucl. Sci.* **53** (2006) 270. doi:10.1109/TNS.2006.869826.
- [34] R. Gavin et al., “FEWZ 3.1: A User’s Guide”.  
[http://www.hep.anl.gov/fpetriello/FEWZManual\\_3.1.pdf](http://www.hep.anl.gov/fpetriello/FEWZManual_3.1.pdf). Accessed: 2015-10-27.
- [35] M. Czakon, P. Fiedler, and A. Mitov, “Total Top-Quark Pair-Production Cross Section at Hadron Colliders Through  $O(\alpha_s^4)$ ”, *Phys. Rev. Lett.* **110** (2013) 252004, arXiv:1303.6254. doi:10.1103/PhysRevLett.110.252004.

- [36] B. Fuks, M. Klasen, D. R. Lamprea et al., “Gaugino production in proton-proton collisions at a center-of-mass energy of 8 TeV”, *JHEP* **1210** (2012) 081, [arXiv:1207.2159](#). doi:10.1007/JHEP10(2012)081.
- [37] B. Fuks, M. Klasen, D. R. Lamprea et al., “Precision predictions for electroweak superpartner production at hadron colliders with Resummino”, *Eur.Phys.J.* **C73** (2013) 2480, [arXiv:1304.0790](#). doi:10.1140/epjc/s10052-013-2480-0.
- [38] J. D. Brinson, “A search for disappearing tracks in proton-proton collisions at  $\sqrt{s} = 8$  TeV”. PhD thesis, Ohio State University.  
<https://inspirehep.net/record/1381347/files/CERN-THESIS-2015-030.pdf>.
- [39] J. Brinson, C. Hill, and W. Wulsin, “Search for disappearing tracks”, *CMS Analysis Note* **CMS-AN-12-400** (2014). Internal documentation.
- [40] CMS Collaboration, “Tracking and Primary Vertex Results in First 7 TeV Collisions”, *CMS Physics Analysis Summary* **CMS-PAS-TRK-10-005** (2010).
- [41] A. Bhatti et. al., “Search for New Physics in the Monojet final state at CMS”, *CMS Analysis Note* **CMS-AN-12-421** (2012). Internal documentation.
- [42] CMS Collaboration, “Search for dark matter, extra dimensions, and unparticles in monojet events in proton-proton collisions at  $\sqrt{s} = 8$  TeV”, *Eur. Phys. J.* **C75** (2015), no. 5, 235, [arXiv:1408.3583](#). doi:10.1140/epjc/s10052-015-3451-4.
- [43] CMS Collaboration, “Studies of Tracker Material”, *CMS Physics Analysis Summary* **CMS-PAS-TRK-10-003** (2010).
- [44] CMS Collaboration, “Measurement of Higgs boson production and properties in the WW decay channel with leptonic final states”, *JHEP* **01** (2014) 096, [arXiv:1312.1129](#). doi:10.1007/JHEP01(2014)096.
- [45] CMS Collaboration, “Performance of electron reconstruction and selection with the CMS detector in proton-proton collisions at  $\sqrt{s} = 8$  TeV”, *JINST* **10** (2015), no. 06, P06005, [arXiv:1502.02701](#). doi:10.1088/1748-0221/10/06/P06005.
- [46] CMS Collaboration, “Performance of CMS muon reconstruction in pp collision events at  $\sqrt{s} = 7$  TeV”, *JINST* **7** (2012) P10002, [arXiv:1206.4071](#). doi:10.1088/1748-0221/7/10/P10002.
- [47] CMS Collaboration, “Performance of  $\tau$ -lepton reconstruction and identification in CMS”, *JINST* **7** (2012) P01001, [arXiv:1109.6034](#). doi:10.1088/1748-0221/7/01/P01001.

- [48] J. Neyman *Phil. Trans. Royal Soc. London, Series A* **236** (1937) 333–80. Reprinted in *A Selection of Early Statistical Papers on J. Neyman* (University of California Press, Berkeley, 1967).
- [49] CMS Collaboration, “CMS Luminosity Based on Pixel Cluster Counting - Summer 2013 Update”, *CMS Physics Analysis Summary* **CMS-PAS-LUM-13-001** (2013).
- [50] D. Barge et. al., “Hadronic Recoil Studies of Heavy Boosted Systems”, *CMS Analysis Note* **CMS-AN-13-059** (2013). Internal documentation.
- [51]
- [52] K. Goebel, “Probing supersymmetry based on precise jet measurements at the CMS experiment”. PhD thesis, Hamburg University, 2015.  
<http://www-library.desy.de/cgi-bin/showprep.pl?thesis15-003>.
- [53] J. Pumplin, D. R. Stump, J. Huston et al., “New generation of parton distributions with uncertainties from global QCD analysis”, *JHEP* **07** (2002) 012,  
[arXiv:hep-ph/0201195](https://arxiv.org/abs/hep-ph/0201195). doi:10.1088/1126-6708/2002/07/012.
- [54] M. Botje et al., “The PDF4LHC Working Group Interim Recommendations”,  
[arXiv:1101.0538](https://arxiv.org/abs/1101.0538).
- [55] A. Vicini, “Practical implementation of the PDF4LHC recipe”.  
[http://www.hep.ucl.ac.uk/pdf4lh/PDF4LHC\\_practical\\_guide.pdf](http://www.hep.ucl.ac.uk/pdf4lh/PDF4LHC_practical_guide.pdf). Accessed: 2015-11-27.
- [56] CMS Collaboration, “Pileup Reweighting Utilities”, *Internal CMS Wiki* (2012).  
<https://twiki.cern.ch/twiki/bin/viewauth/CMS/PileupMCReweightingUtilities>,  
Topic revision: r29.
- [57] CMS Collaboration, “Estimating Systematic Errors Due to Pileup Modeling”,  
*Internal CMS Wiki* (2013).  
<https://twiki.cern.ch/twiki/bin/view/CMS/PileupSystematicErrors>, Topic revision: r11.
- [58] L. Quertenmont et al., “Searches for Heavy Stable Charged Particles with Combined 2011 and 2012 Datasets”, *CMS Analysis Note* **CMS-AN-12-293** (2013). Internal documentation.
- [59] J. Ott, “Search for Resonant Top Quark Pair Production in the Muon+Jets Channel with the CMS Detector”. PhD thesis, Karlsruhe University, 2013.  
[https://inspirehep.net/record/1296889/files/TS2013\\_046.pdf](https://inspirehep.net/record/1296889/files/TS2013_046.pdf).

- 
- [60] T. Junk, “Confidence level computation for combining searches with small statistics”, *Nucl. Instrum. Meth.* **A434** (1999) 435–443, [arXiv:hep-ex/9902006](#). doi:10.1016/S0168-9002(99)00498-2.
- [61] A. L. Read, “Modified frequentist analysis of search results (The  $CL_s$  method)”, in *Workshop on confidence limits, CERN, Geneva, Switzerland, 17-18 Jan 2000: Proceedings*. 2000. <http://cds.cern.ch/record/451614>.
- [62] A. L. Read, “Presentation of search results: The  $CL_s$  technique”, *J. Phys.* **G28** (2002) 2693–2704. doi:10.1088/0954-3899/28/10/313.
- [63] CMS and ATLAS Collaborations, “Procedure for the LHC Higgs boson search combination in summer 2011”, *Technical Report* (2011). CMS-NOTE-2011-005, ATL-PHYS-PUB-2011-011.
- [64] CMS Collaboration, “Documentation of the RooStats-based statistics tools for Higgs PAG”, *Internal CMS Wiki* (2015). <https://twiki.cern.ch/twiki/bin/viewauth/CMS/SWGuideHiggsAnalysisCombinedLimit>, Topic rev: r124.
- [65] C. H. Chen, M. Drees, and J. F. Gunion, “Addendum/erratum for ‘searching for invisible and almost invisible particles at e+ e- colliders’ [hep-ph/9512230] and ‘a nonstandard string/SUSY scenario and its phenomenological implications’ [hep-ph/9607421]”, [arXiv:hep-ph/9902309](#).