UNIVERSIDAD DE CANTABRIA
DEPARTAMENTO DE FÍSICA MODERNA
INSTITUTO DE FÍSICA DE CANTABRIA (UC-CSIC)

Search for long-lived particles decaying into displaced dilepton vertices in proton-proton collisions at $\sqrt{s} = 13$ TeV with the CMS detector

Búsqueda de partículas de alta vida media en su desintegración a vértices dileptónicos desplazados en colisiones protón-protón a $\sqrt{s} = 13$ TeV con el detector CMS

Memoria presentada para optar al título de Doctora por la Universidad de Cantabria. Doctorado en Ciencia y Tecnología.

Por:

Celia Fernández Madrazo

Bajo la supervisión de:

Pablo Martínez Ruiz del Árbol
Celso Martínez Rivero

Mayo 2023

Pablo Martínez Ruiz del Árbol, Doctor en Ciencias Físicas y Profesor Titular de la Universidad de Cantabria

y

Celso Martínez Rivero, Doctor en Ciencias Físicas e Investigador Científico del Consejo Superior de Investigaciones Científicas

CERTIFICAN que la presente memoria

Búsqueda de partículas de alta vida media en su desintegración a vértices dileptónicos desplazados en colisiones protón-protón a $\sqrt{s} = 13$ TeV con el detector CMS

ha sido realizada, bajo su dirección, por Celia Fernández Madrazo, y constituye su Tesis para optar al grado de Doctor por la Universidad de Cantabria. Asimismo, emiten su conformidad para que la presente Memoria sea depositada y se celebre, ulteriormente, la correspondiente Lectura y Defensa.

En Santander, a 25 de Mayo de 2023,

Fdo.: Dr. Pablo Martínez Ruiz del Árbol   Fdo.: Dr. Celso Martínez Rivero

Acknowledgements

It's been five years already since I started to work in which now it has become my thesis. In the process, I had the luck to share my time with wonderful people. I will try to thank some of the many of them that have helped me, in one way or another, during this time.

Pablo, creo que podría llenar esta página de agradecimientos dedicados a ti y aún así me quedaría corta. Intentaré, por lo tanto, ser breve y destacar algunas de las muchas cosas que me vienen a la cabeza. En primer lugar, quiero darte las gracias por haberme dado la oportunidad de hacer el análisis, ya que sin ti, ni tus ideas, no hubiese sido posible siquiera llevarlo a cabo. Dentro de eso quiero también agradecerte todo lo que he aprendido, todo lo que me has enseñado y por la paciencia que has tenido conmigo. Desde luego, si hay algún físico al que yo quiera parecerme mínimamente algún día, ese eres tú. Dejando de lado lo académico, te agradezco la capacidad que has tenido de hacer de este proceso algo divertido y llevadero. Para mí, el charlar tomando un pincho para que yo estuviera tranquila antes de dar mi primera charla en CMS ha sido mil veces más valioso que cien rehearsals (que también hicimos). Gracias por las recomendaciones literarias, las historias increíbles, las charlas, los debates y las muchas muchas risas.

Celso, muchas gracias por haberme dado la posibilidad de hacer el doctorado en el IFCA, por el respaldo y orientación recibida durante el desarrollo de la tesis. Gracias por todas las oportunidades y por todos los medios con los que he contado.

A Lara, gracias por todo lo que me has enseñado, académico o no académico. Pero por encima de todo: gracias por el surrealismo, ya que ha hecho de esta experiencia una de las más divertidas.

Me gustaría agradecer enormemente a todo el grupo de Altas Energías del IFCA, con el que he trabajado y compartido un café al día, todas las mañanas. Muchas gracias por haberme hecho sentir como en casa estos cinco años y por la ayuda brindada cuando ha sido necesaria. En especial, me gustaría dar las gracias a Chus, con el que he podido trabajar conjuntamente en el análisis que ahora se ha convertido en mi tesis; y a Alicia, con la que he tenido la oportunidad de trabajar primero en muones, y después al final juntas como validadoras.

A los ya doctores Pedro, Nicolò y Cedric: Gracias por la tan necesaria introducción al mundo académico que me habéis dado y por la gran cantidad de risas que han hecho de comienzo a final todo el proceso más ameno. Gracias también a los futuros doctores Agustín, Clara, Pablo, Paula, Eloy, Andrés y Efrén. De forma especial, quiero dar las gracias a Sergio y a Rubén. Ha sido toda una alegría compartir despacho con vosotros. Gracias por haber sido un apoyo cada vez que lo he necesitado. En especial contigo, Rubén, he tenido el inmenso placer de trabajar codo con codo en un análisis y, a día de hoy escribiendo esto, solo puedo decir que espero que no sea el último.

Todo esto no habría sido lo mismo sin los estudiantes del IFCA con los que he podido compartir algo de este tiempo (no exagerando si digo que al menos el 80% ha sido en el comedor). Gracias a Iván y a Jorge, por las lecciones de acuarela y guitarra; a Chema, por ser un gran compañero de batallas y gimnasio; a Mikel, por su "cinismo necesario" y los intensos pero muy interesantes debates; a Guille, por su admirable capacidad de hacerme llorar a carcajadas; a Christian, por los momentazos que nos ha dado, y a Elena, que me ha acompañado ya tantos años que no puedo ni contarlos con los dedos de las manos. Gracias, en general, a toda la piña del comedor.

También dentro del IFCA, me gustaría agradecer a Silvia toda la ayuda (y toda la paciencia que ha tenido conmigo) para cada trámite burocrático que he tenido que hacer.

Gracias a Jesús Marco y a Teresa Rodrigo, por apostar por mí desde el principio.

Muchas gracias también a todo el Equipo Oviedo; a Santi y Alejandro, con los que he tenido la suerte de poder trabajar y discutir sobre el análisis; a Sergio, Carlos, Xuan, Pietro y Víctor, de los cuales a día de hoy todavía sigo aprendiendo cosas; a Clara, Andrea y Carlos, amigos con los que también yo he ido creciendo y espero seguir coincidiendo dentro y fuera del campo.

Gracias también a todos los estudiantes de doctorado que he ido conociendo a lo largo del proceso y que han hecho de las conferencias y del CERN algo muy divertido. Emanuela, Quique, Lorenzo, Gorka, Naseem y Lourdes, entre muchos otros. Gracias especiales a Gonzalo, que probablemente sea la persona a día de hoy con la que más tiempo he pasado en el CERN.

I would like to say a big thank you to all the people in the Muon POG. It has been a pleasure to work with you, I really enjoyed my time there. A special mention goes to Daniele, who had the patience to teach me during my first steps in muon reconstruction and from whom I have learned quite a lot.

Thank you to the people from the long-lived community, with whom I had the pleasure to share very interesting conversations that definitely enriched my experience during this time. Thank you Alberto, Slava, Ivan, Alice, Andre, Mangesh, Ansar, Juliette, Zeynep and Livia.

Gracias a la gente que me ha acompañado todos estos años en Santander. Aquí termina una etapa que gracias a vosotros ha sido maravillosa. Oliver, Celia, Dani, Mateo y Mario, voy a echar mucho de menos las *Lupe Sessions*. Mis canallitas Darío y Manu, gracias por todas las recomendaciones gastronómicas. A Javi, Raquel, Eli, Unai, Cristina, Fran y Fer, mis físicos favoritos. A Eva, que me ha aguantado al final de la tesis.

A mis compañeras de piso, Ana y Marta: gracias por haber sido mis mediohermanas estos años y por poder contar con vosotras para cualquier cosa. Quién diría que varios años después, hemos conseguido no quemar ninguna casa.

A mi piña, testigo de todas mis *celiadas*. Gracias por los viajes, las comidas, las cañas, el parchís en covid, las nocheviejas, el amigo invisible, la patrona, la torre de chatos... y un largo etc. Carol, Diego, María, Sara, Carmen, Marti, Claudia y Ruth, os quiero mucho.

Elena, gracias por estar ahí, y por ser una persona con la que siempre he podido contar, para absolutamente todo. Para mí, el saber que tengo eso, es de lo más valioso que hay.

Quiero también agradecer el apoyo a toda mi familia, que me ha acompañado en el durante, pero también mucho antes de que esto empezase. Un muy cariñoso gracias a mis abuelos. En especial, gracias a la división internacional Madrazo con la que he explorado Londres, Oporto, Milán, Burdeos y París (y sumando).

Todos los agradecimientos son pocos para mis padres, Menchu y Carlos. Gracias por todo el apoyo incondicional que me dais, para esto y para todo. Soy muy consciente de que todo el esfuerzo que está plasmado en esta tesis es también vuestro. Habéis sido y estado, siempre que lo he necesitado, y de la forma que lo he necesitado. Y yo no os puedo querer más. A mi hermana y ejemplo a seguir: Carmen, tú eres el mayor apoyo de todos, te quiero un mundo. Gracias por mucho.

Abstract

Some of the most popular extensions of the Standard Model of Particle Physics contain long-lived particles, which can be framed in a wide variety of theories, covering different regions of the phase space. Long-lived particles manifest in the particle detectors in the form of non-conventional signatures that often require the development and optimization of new analysis techniques, different to those that are commonly used in other searches where the particles have smaller lifetimes.

This thesis presents a search for long-lived particles decaying to final states with a displaced dilepton vertex. It analyzes proton-proton collisions at a center-of-mass energy of 13 TeV that were produced by the CERN LHC and collected by the CMS detector in 2016, 2017 and 2018. Leptons in this analysis appear as displaced from the interaction point and need to be reconstructed with special algorithms. The collision events are required to have at least two opposite charge, same flavor leptons fitted to a common displaced vertex in the tracker, which is the main handle of the analysis. The results of the search are interpreted in the context of an extended Higgs-sector model with a non-SM Higgs boson decaying to two spinless long-lived bosons. The search is designed to cover a wide mass range, and be sensitive to small and medium lifetimes from 0.1 cm to 1000 cm.

Observed data in the analysis is found to be in good agreement with the Standard Model expectations. Upper limits have been derived for a Higgs model decaying to two scalar long-lived particles decaying into displaced leptons. The limits are estimated separately for the muon and electron channels, and then both combined. These results are competitive with similar results shown by the CMS and ATLAS collaborations and in particular they provided the most stringent limits for the electron channel.

Resumen

Una de las extensiones más populares del Modelo Estándar de la Física de Partículas son las partículas de alta vida media, que pueden proponerse en una gran variedad de teorias apareciendo en diferentes regiones del espacio de fases. Las partículas de alta vida media se pueden manifestar en los detectores de partículas en forma de signaturas experimentales poco convencionales, que a menudo requieren del desarrollo y optimización de nuevas técnicas de análisis, diferentes de las cuales se utilizan en otras búsquedas donde la vida media es más pequeña.

El trabajo presentado en esta tesis consiste en una búsqueda de partículas de alta vida media en su canal de desintegración a estados finales con un vértice dileptónico desplazado. En esta búsqueda se analizan colisiones protón-protón a una energía del centro de masas de 13 TeV, producidas en el LHC del CERN y detectadas en el detector CMS en 2016, 2017 y 2018. Los leptones en este análisis aparecen desplazados del punto de interacción, por lo que se reconstruyen con algoritmos especializados para ello. Se requiere a los eventos que tengan al menos dos leptones del mismo tipo y carga opuesta que produzcan un ajuste válido a un vértice desplazado en el tracker, siendo este vértice el objeto principal del análisis. Los resultados de la búsqueda se interpretan en el contexto de un modelo con un sector de Higgs extendido, que predice la existencia de un bosón de Higgs, diferente al del Modelo Estándar. Este bosón se desintegra en dos bosones escalares de alta vida media que a su vez se desintegran en dos leptones desplazados. La búsqueda se ha diseñado para cubrir un amplio rango de masa y obtener resultados para un rango de vidas medias de 0.1 cm a 1000 cm.

Los datos observados son compatibles con las predicciones del Modelo Estándar. Se han derivado límites superiores para el modelo de bosones de Higgs en su desintegración a bosones escalares desintegrándose a leptones desplazados. Los límites se estiman por separado para los canales de muones y electrones, que son posteriormente combinados. Los resultados son competitivos con resultados similares obtenidos por CMS y ATLAS y, en particular, representan los mejores límites para el canal de electrones.

Contents

List of Figures

Introduction

The Standard Model (SM) of particle physics is the current accepted theory to describe the matter and their interactions. It has provided successful predictions for the different phenomena observed in the experiments of particle physics. Despite its enormous success, it is already known that the SM is not a complete theory and an extension would be needed to provide a fundamental basis on some of the experimental results that remain unexplained. This is why the theoretical community has proposed several *beyond the Standard Model* (BSM) theories to extend the current theoretical framework.

The CERN Large Hadron Collider (LHC) is the most powerful particle accelerator and the largest physics machine ever built. It started to operate in 2010, accelerating particles close to the speed of light to produce high energetic collisions where new and exotic physics could manifest. The most recent success of the LHC was the experimental observation of the Higgs boson in 2012.

Some of the new exotic theories propose the existence of particles with relatively longer lifetimes that the fundamental particles of the SM. These so-called long-lived particles (LLP) naturally arise in a wide diversity of theoretical scenarios. Their macroscopic decay length can emerge from the presence of small couplings, heavy off-shell mediators or compressed mass spectra, among others. From the experimental point of view, large lifetimes allows long-lived particles to travel through the detector a certain time before decaying into other particles. This results in challenging signatures, motivating the use of non-standard tools and innovative techniques in the searches. Within this document some of the most popular models containing LLPs are briefly introduced and the most typical experimental signatures are also presented.

The core of my thesis is a search for long-lived particles (LLP) decaying to final states that contain displaced dilepton vertices using pp collisions collected by the CMS detector during the full LHC Run 2, in 2016, 2017 and 2018. In the models considered in this thesis, the LLPs do not interact with the detector material and are not directly observable. Their existence can be, instead, inferred by studying their decays. This search targets long-lived particles decaying into a pair of oppositely charged, same flavor leptons in the tracker volume (radius up to 100 cm). The decay vertex for these LLPs is found displaced from the interaction point and its position, which is measured by fitting the two lepton tracks, is the main handle of the search. The analysis is sensitive to LL models with at least one displaced dilepton vertex in the final state. Two decay channels are considered, dielectron and dimuon channels, depending whether the LLP decays into an electron-positron or a muon-antimuon pair.

The search is intended to be as agnostic as possible and could be interpreted in any model with long-lived particles decaying to lepton pairs in the tracker volume. However, one BSM simplified model with an extended Higgs sector is used to guide the search and derive interpretations. In this scenario, an exotic Higgs bosons is produced through gluon-gluon fusion. This boson decays into a pair of long-lived spinless bosons S. After a certain time, driven by the S boson lifetime τ_S, the bosons decay into lepton-antilepton pairs, produced in displaced dilepton vertices: H $\rightarrow$ SS $\rightarrow 2\ell^+2\ell^-$.

Standard CMS algorithms were initially optimized to reconstruct leptons originated from the primary vertex of the pp collisions, strongly constraining the pattern recognition and the parameter fitting of the tracks. Displaced tracks are frequently missed by the standard algorithms, leading to a poor reconstruction efficiency for displacements beyond one or two centimeters. Furthermore, the use of constraints in the parameter fitting of the tracks results in worse p_T resolution for the leptons whose origin is not the primary vertex. The search implements specific electron and muon reconstruction algorithms to enhance the efficiency at large displacements and have better momentum resolution. Their usage and main properties are described in detail in Chapter 4 of this document, together with the methodology followed to reconstruct the displaced dilepton vertices.

Most of SM processes produce particles at the center of the detector, so the analysis is relatively background free. However, some instrumental background coming from resolution or misreconstruction effects arise at high displacement. The Drell-Yan process will be the main instrumental background due to its large cross-section. Other secondary backgrounds include $t\bar{t}$ and diboson (WW, WZ, ZZ) production. QCD processes, with leptons produced in the core of the jets or misindentified pions, will also populate the tails of the distribution at high displacement. In addition, this analysis also takes into account two potential displaced background sources. Only for the dimuon channel, cosmic muons travelling through the detector can be reconstructed as back-to-back muons whose tracks may be ultimately fitted into a common displaced vertex. The other displaced background source are displaced vertices produced by nuclear interactions in the detector material. The contribution of these two backgrounds is strongly mitigated and their impact in the Signal Region found to be negligible. Since the instrumental background contribution that remains is usually hard to reproduce with the Monte Carlo simulation, the background in this analysis is completely estimated from data control samples. However, we use Monte Carlo simulation of the main background contributions to cross-check and validate the background estimation method.

The reconstructed vertices are required to pass a baseline selection to remove as many background vertices as possible and optimize the search. Finally, two variables are used to measure the displacement and discriminate between vertices produced through LLP decays and the ones that arise in instrumental backgrounds. These are the significance of the minimum transverse impact parameter of the two tracks used to fit the vertices $|d_0|/\sigma_{d_0}$ and the transversal distance L_{xy} between the dilepton vertex and the primary vertex reconstructed in the collision.

The results of the search are used to set 95% confidence upper limits on the production cross-section $\sigma \times \mathcal{B}(\mathrm{H} \to \mathrm{SS})$ of the BSM Higgs model. The two channels of the analysis are used to interpret the results assuming three different scenarios. First, pure muon decays are assumed, with branching fraction of the LLP of $\mathcal{B}(\mathrm{S} \to \mu\mu) = 100\%$. Second, a total decay ratio to electrons with LLPs fractions of $\mathcal{B}(\mathrm{S} \to ee) = 100\%$ is also considered. Finally, both channels are combined to under the interpretation of having democratic electron and muon decays with a $50 - 50\%$ rate.

In addition to the development of this analysis, during my thesis I have worked in other central aspects of the CMS Collaboration connected to LLP searches. In particular, during this period, I have been responsible of the validation of the objects utilized by the Exotica Physics Analysis Group (PAG). Also, I have served as L3 muon reconstruction convener since 2021 and I have been involved in multiple activities performed within the Muon Particle Object Group (POG), specially in aspects related to the reconstruction of displaced muons.

Chapter 1

The Standard Model of particle physics

This chapter gives an overall description of the Standard Model of particle physics (SM), the theory that currently describes three of the four fundamental forces in the universe and the matter content in terms of fundamental particles. This thesis will focus on a search of new exotic physics that is proposed to solve some of the already known existing open problems of the SM, that are also introduced in Section 1.6 of this chapter.

1.1. The fundamental particles

The Standard Model of particle physics (SM) [1] is the currently accepted theory in which the particles and their interactions are described. The SM is a Quantum Field Theory (QFT) [2], a theoretical framework which combines quantum mechanics and special relativity. Within a QFT, the particles are described by excitations of quantum fields and their interactions arise from establishing accomplishment of the internal symmetries that are present in the system. The dynamics of the fields are expressed in terms of a certain Lagrangian density $\mathcal{L}$ which is required to be invariant under the local gauge transformations that corresponds to each one of these symmetries. $\mathcal{L}$ determines, in turn, the field equations that the particles have to satisfy. In specific, the SM contains the internal symmetries of the group $\mathrm{U}(1)_Y \times \mathrm{SU}(2)_L \times \mathrm{SU}(3)_C$.

These symmetries correspond with the electromagnetic, weak and strong forces, for which the SM provides an effective mathematical description. Gravity is not yet included in the Standard Model. Instead, the basics of gravity are described by the theory of *General Relativity*. In general, a certain particle can interact with others by means of one or more forces, but not necessarily all of them. In the SM the type of each particle is determined by the forces that can act over it, allowing for a particle classification.

Within the SM, the matter is composed of spin-half particles, called *fermions*, which act as fundamental constituents of the Universe. The matter content of the Universe is entirely described with a total of twelve fermions and their antiparticles, which have exactly the same mass but opposite electric charge. These fermions interact with other fermions through the four fundamental forces previously mentioned and are classified into several categories accordingly. The twelve particles undergo through weak interactions while only the ones that carry electric charge suffer the electromagnetic force. The strong force only acts on particles which carry the strong interaction equivalent of electric charge, known as *color charge*. There are six different particles that carry color charge and have fractional electric charge, referred as *quarks*. These are named as up-quark u, down-quark d, charm-quark c, strange-quark s, top-quark t and bottom-quark b. Due to the nature of the strong interactions, quarks cannot be found in nature as single particles but in bound states called *hadrons*. For example, the proton is a hadron consisting of two up-quarks and one down-quark i.e. $p(uud)$. The other six fundamental particles are collectively referred as *leptons*, which are grouped in three distinct families. Each one of these families is represented by a doublet composed of a lepton with negative charge and a neutral partner. The charged leptons of the three families are the electron e^-, the muon μ^- and the tau-lepton τ^-. The neutral lepton partners associated with them are called *neutrinos* ν.

In the SM the forces are interactions mediated by several integer-spin particles called *gauge bosons*. The boson that mediates electromagnetic interactions between electrically charged particles is the photon γ. Similarly, the force-carring boson for the strong interaction is called the *gluon* g, which acts on particles with color charge. The weak force is divided into neutral-current interactions, that are mediated by the electrically neutral Z boson, and charged-current interactions, mediated by the charged W^+ and W^- bosons. Both photon and gluon are massless particles while the weak gauge bosons Z and $W^\pm$ were observed to have mass. Their values were measured to be $m_Z = 91.1876 \pm 0.0021$ GeV and $m_W = 80.379 \pm 0.012$ GeV [1] which is known as the electroweak scale.

The SM picture is completed by the Higgs boson H, a spin-0 scalar particle introduced in the theory to provide the mechanism by which all other massive particles acquire their mass. Its existence was postulated in 1964 by Peter W. Higgs [3] and it was discovered in 2012 at the LHC by the ATLAS and CMS experiments with a mass of $m_H = 125$ GeV [4, 5]. It is considered to be an excitation of the Higgs field.

All fundamental particles encompassed in the SM are summarized in Figure 1.1. The twelve fermions are the fundamental constituents of matter, which are divided into six quarks and six leptons, being the latter split into three charged leptons and three neutrinos. For the force carriers, we distinguish between five g, γ, Z and $W^\pm$ vector bosons that mediate the interactions. Finally, the H scalar boson is included to provide mass to the other particles.

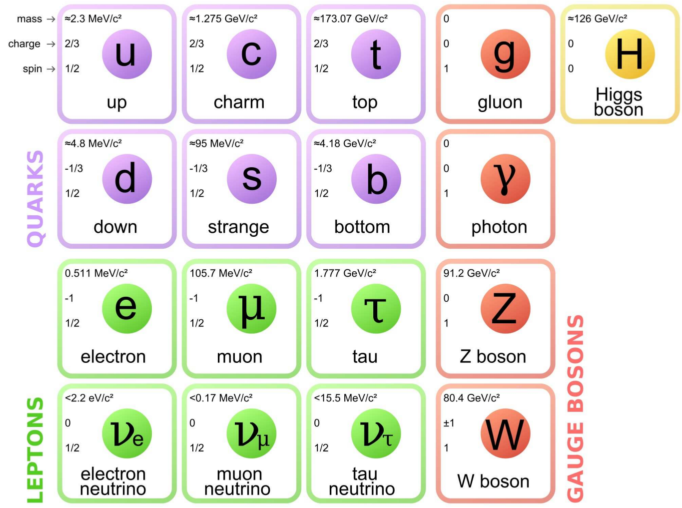


Figure 1.1: Fundamental particles of the Standard Model of Particle Physics [6].

The main aspects of the SM interactions are described in the following sections. The fundamental theory for the strong force is known as Quantum Chromodynamics (QCD) which will be briefly introduced in Section 1.2. After the masses of the $Z/W^{\pm}$ bosons were introduced in the SM, the electromagnetic and weak interactions are unified and known as Electroweak (EW) interaction which will be introduced in Section 1.3. Finally, the main aspects of the Higgs mechanism that is introduced in the EW theory is described in a separate Section 1.4.

1.2. Quantum Chromodynamics

The strong interaction is described by its associated QFT, known as Quantum Chromodynamics (QCD). Strong interactions are invariant under the non-Abelian SU$(3)_C$ local gauge symmetry group which has eight generators $\hat{\mathbf{T}} = \{T^a\}$ better known as *Gell-Mann* matrices. This symmetry also implies a conserved color charge. The QCD Lagrangian $\mathcal{L}_{\text{QCD}}$ can be written as

$$\mathcal{L}_{\text{QCD}} = \bar{\psi}(i\gamma^{\mu}\partial_{\mu} - g_S\gamma^{\mu}T^aG^a_{\mu} - m)\psi - \frac{1}{4}G_{\mu_\nu}G^{\mu\nu} \tag{1.1}$$

where ψ refers to the quark field which includes the new color state with three additional degrees of freedom. This Lagrangian must be invariant under SU$(3)_C$ local phase transformations given by

$$\psi(x) \rightarrow \exp[ig_S\hat{\mathbf{T}} \cdot \boldsymbol{\alpha}(x)]\psi(x) \tag{1.2}$$

being $\boldsymbol{\alpha}(x)$ the eight rotation angles of the transformation.

The term G^a_{μ} in Eq(1.1) refers the new eight fields that correspond to the gluon.

The corresponding strength field tensor $G_{\mu\nu}$ of this field is given by

$$G_{\mu\nu} = \partial_\mu G^a_\nu - \partial_\nu G^a_\mu - g_S f^{abs} G^b_\mu G^c_\nu \tag{1.3}$$

where $g_S = \sqrt{8\pi\alpha_S}$ stands for the coupling constant of the strong interaction. f^{abc} are the structure constants of the SU$(3)_C$ group which determine the commutation relations of the T^a generators

$$[T_b, T_c] = i f_{abc} \lambda_q \tag{1.4}$$

The value of the strong coupling α_S is not constant and depends on the energy of the interaction. While at low-enery scales its value is large ($\mathcal{O}(1)$) and the theory is non-perturbative, at high energies α_S decreases and its value is small enough for perturbation theory to be applied again. This property is known as *asymptotic freedom*.

1.3. The Electroweak interaction

In the 1960s, Glashow, Salam and Weinberg [7–9] developed a theory where the electromagnetic and weak interactions are unified, known as Electroweak theory. This theory predicts a weak neutral-current mediated by the Z and γ bosons apart from the weak charged-current, mediated by the $W^\pm$ bosons, that had already been described within the weak interaction at that time. Electroweak interactions are invariant under transformations of the SU$(2)_L \times$U$(1)_Y$ local gauge symmetry group.

Within the Electroweak theory, the weak charged-current is associated with invariance under SU$(2)_L$ local gauge transformations given by

$$\psi(x) \to \exp[i g_W \boldsymbol{\alpha}(x) \mathbf{T}] \psi(x) \tag{1.5}$$

where g_W is known as the weak coupling constant of the SU$(2)_L$ gauge interaction and $\mathbf{T}$ are the three generators of the SU$(2)_L$ group, identified with the well-known Pauli spin matrices $\boldsymbol{\sigma}$

$$\mathbf{T} = \frac{1}{2}\boldsymbol{\sigma} \tag{1.6}$$

Three gauge fields W^k_μ arise from this local invariance requirement and a new charge is to be conserved in weak processes, called weak isospin I_W.

Charged-currents violate parity, and this violation should be satisfied by their interaction vertex which is of the following form

$$\frac{-i g_W}{2\sqrt{2}} \gamma^\mu (1 - \gamma^5) \tag{1.7}$$

In general, each particle spinor can be decomposed into its left and right chiral components through the chiral projection operators

$$\psi_L = \frac{1}{2}(1-\gamma^5)\psi, \quad \psi_R = \frac{1}{2}(1+\gamma^5)\psi \tag{1.8}$$

and, as it can be seen from Eq(1.7) and Eq(1.8), the right-handed component ψ_R will be suppressed and consequently, only the left-handed component ψ_L will interact in the weak charged-currents. This suppression has important consequences for relativistic particles whose helicity and chiral states are the same, as it is the example of the neutrinos ν, which won't be observed unless they have left chirality.

These requirements leave the twelve fermions grouped in weak isospin doublets with total isospin $I_W = 1/2$ and left chirality. Their third component of the weak isospin will be $I_W^{(3)} = +1/2$ for the first doublet component and $I_W^{(3)} = -1/2$ for the second:

$$\begin{pmatrix}\nu_e \\ e^-\end{pmatrix}_L, \quad \begin{pmatrix}\nu_\mu \\ \mu^-\end{pmatrix}_L, \quad \begin{pmatrix}\nu_\tau \\ \tau^-\end{pmatrix}_L, \quad \begin{pmatrix}u \\ d\end{pmatrix}_L, \quad \begin{pmatrix}c \\ s\end{pmatrix}_L, \quad \begin{pmatrix}t \\ b\end{pmatrix}_L \tag{1.9}$$

On the other hand, the non-interacting right-handed particles correspond to weak singlets with $I_W^{(3)} = 0$

$$e_R^-, \quad \mu_R^-, \quad \tau_R^-, \quad u_R, \quad d_R, \quad c_R, \quad s_R, \quad t_R, \quad b_R \tag{1.10}$$

For the mediators of the charged-currents, the gauge fields $W_\mu^\pm$ of the physical $W^\pm$ bosons are identified with the linear combinations of the $W_\mu^{(1)}$ and $W_\mu^{(2)}$ new gauge fields

$$W_\mu^\pm = \frac{1}{\sqrt{2}}\left(W_\mu^{(1)} \pm iW_\mu^{(2)}\right) \tag{1.11}$$

Finally, to unify weak and electromagnetic interaction the electroweak theory replaced the U(1) symmetry of QED with a new U$(1)_Y$ local gauge symmetry where ψ transforms as

$$\psi(x) \to \exp[ig'\frac{Y}{2}\xi(x)]\psi(x) \tag{1.12}$$

being g' the neutral-charge current coupling and Y the associated conserved charge, known as weak hypercharge.

A new field B_μ arises from imposing invariance under U$(1)_Y$ which couples with the remaining $W_\mu^{(3)}$ weak field through the weak mixing angle θ_W. The fields A_μ and Z_μ, that are the gauge fields of the neutral photon and Z boson mediators, arise from this mixing

$$\begin{pmatrix} A_\mu \\ Z_\mu \end{pmatrix} = \begin{pmatrix} \cos\theta_W & \sin\theta_W \\ -\sin\theta_W & \cos\theta_W \end{pmatrix} \begin{pmatrix} B_\mu \\ W_\mu^{(3} \end{pmatrix} \tag{1.13}$$

In the same way, the weak hypercharge is a combination of the weak isospin and the electromagnetic Q charge components

$$Y = \alpha Q + \beta I_W^{(3)} \tag{1.14}$$

The Lagrangian density associated with the weak theory is not invariant under the aforementioned gauge transformations, Eq.(1.5) and Eq.(1.12), if the W and Z bosons are massive. To overcome this problem the Higgs mechanism, explained in the next section, was proposed. Starting from a QFT with massless gauge bosons B_μ, $W_\mu^{(1)}$, $W_\mu^{(2)}$ and $W_\mu^{(3)}$, and symmetry $\mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$, the Higgs mechanism leads to the appearance of the new mass states introduced before, corresponding with $W_\mu^{\pm}$, A_μ and Z_μ.

1.4. The Higgs Mechanism

For the SM to be consistent with the experimental results some of the particles must be massive, as it is the case of the weak gauge $W^\pm$ bosons. However, mass terms for the W and Z bosons would violate $\mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$ local gauge invariance if the Higgs mechanism were not included in the SM. The Higgs mechanism is introduced in the theory to generate a mass for the electroweak gauge bosons $W^\pm$, Z and the fermions by adding at least one new massive scalar particle, which is identified with the Higgs boson.

The SM Higgs mechanism predicts the existence of two complex scalar fields placed in a weak isospin doublet $\phi(x)$, which written in the unitary gauge has the form of

$$\phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix} \tag{1.15}$$

and its corresponding Lagrangian with the Higgs potential $V(\phi)$ defined as

$$\mathcal{L} = (D_\mu\phi)^\dagger(D^\mu\phi) - V(\phi), \quad V(\phi) = \mu^2\phi^\dagger\phi + \lambda(\phi^\dagger\phi)^2 \tag{1.16}$$

Provided that $\mu^2 < 0$, the Higgs potential has a set of degenerate minima resulting in a non-zero *vacuum expectation value* v. The choice of a physical vacuum state breaks the symmetry, a process which is called *spontaneous symmetry breaking*. Once the symmetry is broken, perturbations of $\phi(x)$ around this physical vacuum state are identified with the physical Higgs field $h(x)$. For the Higgs mechanism to the embedded in the $\mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$ local gauge symmetry, the covariant derivative D_μ is defined in terms of the generators and electroweak fields defined in Section 1.3 as

$$D_\mu = \partial_\mu + ig_W \mathbf{T} \cdot \mathbf{W}_\mu + ig' \frac{Y}{2} B_\mu \tag{1.17}$$

The masses of the gauge bosons are determined by the terms in the $(D_\mu \phi)^\dagger (D^\mu \phi)$ component of the Lagrangian that are quadratic in the corresponding gauge boson fields. For the $W^\pm$ bosons, the mass is identified with the quadratic term of the $W^{(1)}$ and $W^{(2)}$ fields which has the value

$$m_W = \frac{1}{2} g_W v \tag{1.18}$$

determined by both the weak coupling constant g_W and the vacuum expectation value v. In the same way, for the gauge bosons of the neutral currents the masses are obtained from the quadratic terms in the neutral $W^{(3)}$ and B_μ fields. The mass terms in the corresponding physical boson field basis, A_μ and Z_μ, are assembled in a diagonal mass matrix of the form

$$\frac{1}{8} v^2 \begin{pmatrix} A_\mu & Z_\mu \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & g_w^2 + g'^2 \end{pmatrix} \begin{pmatrix} A^\mu \\ Z^\mu \end{pmatrix} \tag{1.19}$$

where the mass of the Z boson depends on g_W, g' and v, and the photon remains massless as expected:

$$m_A = 0, \qquad m_Z = \frac{1}{2} v \sqrt{g_W^2 + g'^2} \tag{1.20}$$

The Higgs boson H itself acquires a mass whose value is determined by the vacuum expectation value v and the λ parameter of the Higgs boson potential previously introduced in Eq.(1.16)

$$m_\mathrm{H} = \sqrt{2\lambda} v \tag{1.21}$$

The spontaneous symmetry breaking mechanism also generates the masses for the fermions of the SM. Due to the different behavior that the left- and right-handed chiral states present under the electroweak interactions, the mass terms $m\psi\bar{\psi}$ are not invariant under the $\mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$ gauge symmetry. The Yukawa terms that are added in the Lagrangian in the case of the leptons only generate the mass for the lower component of the $\mathrm{SU}(2)_L$ doublet, leaving the neutrino massless. The Yukawa term for the electron-neutrino case, shown as an example, has the form

$$\mathcal{L}_e = -g_e \left[\begin{pmatrix} \bar{\nu}_e & \bar{e} \end{pmatrix}_L \phi e_R + \bar{e}_R \phi^\dagger \begin{pmatrix} \bar{\nu}_e \\ \bar{e} \end{pmatrix}_L \right] \tag{1.22}$$

In the case of the quarks the picture is different as both the upper and lower member of the weak doublet need to acquire mass. For up-type quarks to have mass, the conjugate doublet ϕ_c is constructed and an additional hermitian conjugate of the scalar field is added to the Yukawa term. The Yukawa term for up-down quarks, shown as an example, has the form

$$\mathcal{L}_{u,d} = -g_u \begin{pmatrix} \bar{u} & \bar{d} \end{pmatrix}_L \phi_c u_R - g_d \begin{pmatrix} \bar{u} & \bar{d} \end{pmatrix}_L \phi d_R + h.c. \tag{1.23}$$

The fermion masses are determined by the Yukawa couplings and the non vacuum expectation value

$$m_f = \frac{v}{\sqrt{2}} g_f \tag{1.24}$$

1.5. Scattering theory

Feynman developed a simplified way to represent the physical processes in the form of diagrams. These so-called Feynman diagrams are not proper spacial-time representations of the particles. Instead, they are a pictorial representation of the involved equations which allows to derive the properties of the processes by relating the form of the diagram with the Feynman rules that arise from the Lagrangian densities.

The mathematical treatment of the SM allows to estimate the interaction probabilities based on a perturbative expansion provided that the coupling involved in the considered processes is small enough, as it is the case of colliding experiments. In general, the limit where the interaction is weak is referred as perturbative regime. Within this perturbative regime, one single process, defined by an initial and final state, has multiple contributions. In a colliding experiment, where the time in which the interaction occurs is short, the initial and final states, given at times $t \to -\infty$ and $t \to \infty$ respectively, can be considered as free particle states. The quantum probability of a given transition between the initial state $|i>$ and the final state $|f>$, corresponds with the projection of the final state over the time evolution of the initial state from a moment well before and well after the interaction. If the time evolution operator of the initial state is denoted as $U(-\infty, \infty)$, then, this probability is:

$$S_{if} =< f|U(-\infty, \infty)|i > \tag{1.25}$$

This equation defines the so-called S-Matrix that contains the probabilities between all possible initial and final states. This bracket is usually solved using the interaction picture for the fields and the operator. In this representation, the time operator, which depends on the total Hamiltonian of the system, can be expanded as a perturbative series on the interaction Hamiltonian. The first term in this expansion is known as the leading-order (LO) contribution, with is associated to the LO diagram. The higher order corrections terms are known as the next-to leading order (NLO), the next-to-next-to leading order (NNLO), etc.

One of the most important observables in collision experiments is the *cross-section* σ, which is defined as the number of interactions per unit time and target over the incident flux. In general terms, the cross-section is proportional to the probability that a certain process has to occur. The cross-section for a given process can be expressed in terms of the S-matrix.

In high energy collisions with composite particles such as protons, the interaction takes place between the fundamental particles composing the inner structure, called *partons*. In this situation the momentum that governs the process is not the one of the proton p_p, but the fraction taken by the partons that participate in the interaction, namely $x_1 = p_1/p_p$ and $x_2 = p_2/p_p$. The cross-section of the process will be therefore determined by the distributions of momenta of these partons, known as the *Parton Distribution Functions* (PDFs). The resulting cross-section $\sigma(pp \to X)$ for a certain final state X in a proton-proton collision is thus given by the *master equation* at the LHC [10] and the factorization theorem

$$\sigma(pp \to X) = \sum_{a,b} \int dx_1 dx_2 f_a(x_1, \mu_F) f_b(x_2, \mu_F) \hat{\sigma}_{ab\to X}(x_1 x_2 s; \mu_F, \mu_R) \quad (1.26)$$

where $\hat{\sigma}_{ab\to X}$ is the partonic cross-section evaluated from the theory, μ_R the renormalization scale and μ_F the energy scale that governs the process. Eq(1.26) runs over the partons a and b present in the two protons.

1.6. Open problems of the Standard Model

The Standard Model is one of the triumphs of modern physics as it is capable of accurately describe the numerous experimental measurements done in HEP over the last decades. However, despite its success, the SM has a total of 26 parameters that are assigned by the experiment and there is not an underlying higher theoretical principle providing a fundamental explanation for them. There are also patterns between these parameters that may suggest the existence of an additional unknown symmetry principle that remains undiscovered, a Grand Unified Theory (GUT) of the forces [11, 12].

One of the most popular open questions of the SM is why the theory does not include gravity, the fourth force, in the same way as the electromagnetic, weak and strong force. For the gravity to be included in the SM, a new particle, commonly referred as *gravitino* G, would need to be found. Quantum theories of gravity [13] have been proposed, but there is currently no experimental evidence for their existence.

In the same way, while the SM provides the theoretical mechanisms to describe the interactions of baryonic matter, the nature of dark matter and dark energy remains unsolved. Early experiments in the 1930s already proved that baryonic matter is not enough to properly describe the velocity distributions of the stars that orbit around the galactic centres. The observations suggest that there is a non-luminous component of mass responsible for the measured velocities. Unfortunately, there is no experimental evidence for a proper dark matter candidate that could be added to the SM particle family.

One of the most popular unsolved problems of the SM is related to the mass of the Higgs boson and goes under the name of Hierarchy Problem. The mass of the Higgs boson was already measured to be $m_H = 125$ GeV. Through the theory, this value is determined by high order corrections Δm_H that arise from the coupling between the Higgs boson and the massive particles. These higher order correction terms depend on the ultraviolet cutoff Λ_{UV} which is used to regular the loop integral of the higher order Feynman diagrams:

$$\Delta m_H^2 \propto \frac{g_f^2}{8\pi^2}\Lambda_{\mathrm{UV}}^2 \tag{1.27}$$

The energy scales in particle physics are very different and specially at the Planck scale these corrections can become very large for the Higgs mass to remain low. A natural solution to this problem would come from a suppression provided by additional corrections arising from couplings with new particles. In this context, the so-called Naturalness theories predict the existence of exotic particles whose couplings with the Higgs boson naturally introduce the necessary high order terms to control its mass.

Neutrinos are responsible for another inconsistency within the SM. It is known through the observation of the neutrino oscillations, that they have small a mass that, although almost negligible in comparison with other particles, remains non-zero [14]. However, in the current formalism of the electroweak theory and the Higgs mechanism the neutrinos are assumed to be massless.

Additional tensions in the SM are posed by two recent experimental results. The first one is the measurement of the positive muon anomalous magnetic moment by the Fermi National Accelerator Laboratory (FNAL), found to be 0.46 ppm [15]. This result is 3.3 standard deviations greater than the SM prediction, and shows good agreement with the measurement performed at the Brookhaven laboratory [16]. The second one is the measurement of the W boson mass by the CDF II detector [17], also in FNAL. In this case the measured mass is higher than the predicted by the SM with a discrepancy of 7 standard deviations.

1.7. Theories beyond the Standard Model

The fact that there are several unanswered questions that remain open in the SM points out that the current theoretical framework is not complete. A wide variety of new theories have emerged over the last decades to provide a solution to one or some of the problems already presented in Section 1.6. These are referred as theories *beyond the SM* (BSM) and this sections summarizes the most popular ones. Unfortunately and despite the huge effort made by the HEP community, no evidence of the existence of any of them has been found so far. This motivates the evolution of such theories and also the proposal of new ones. This is the case of long-lived particles, whose study is the focus of this thesis and that will be widely described in Chapter 3. Specifically, Section 3.2 will focus on some specific models in which long lifetimes may arise.

Some compilations of BSM theories can be found in the literature [18–20]. This section summarizes the details of the most popular ones, such as Supersymmetry, Extra dimensions, Hidden Valley theories, Little Higgs models and Technicolor, although it is important to note that these are by no means the only ones proposed.

Supersymmetry

One of the most popular theories is Supersymmetry (SUSY) [21], which predicts that every SM particle has a particle, referred as *superpartner*, with a $1/2$ difference in spin but keeping the other quantum numbers to be the same. It is then implied by the theory that the superpartner of every fermion is a boson and vice versa. The operator Q that generates the transformation between superpartners must be an anti-commuting spinor such that

$$Q|\text{Boson}\rangle = |\text{Fermion}\rangle \qquad\qquad Q|\text{Fermion}\rangle = |\text{Boson}\rangle \tag{1.28}$$

Both fermion and boson states are contained in a irreducible representation of the supersymmetry algebra which is known as *supermultiplet*.

The boson superpartners of the SM leptons are called *sleptons* and their corresponding gauge eigenstates are labelled as ($\tilde{e}_R$, $\tilde{e}_L$, $\tilde{\nu}_e$) for the first generation, ($\tilde{\mu}_R$, $\tilde{\mu}_L$, $\tilde{\nu}_\mu$) for the second one, and ($\tilde{\tau}_R$, $\tilde{\tau}_L$, $\tilde{\nu}_\tau$) for the third. In the same way, the boson superpartners of the SM quarks are called *squarks* and their gauge eigenstates are dennoted as ($\tilde{u}_R$, $\tilde{u}_L$, $\tilde{d}_R$, $\tilde{d}_L$), ($\tilde{c}_R$, $\tilde{c}_L$, $\tilde{s}_R$, $\tilde{s}_L$) and ($\tilde{t}_R$, $\tilde{t}_L$, $\tilde{b}_R$, $\tilde{b}_L$) for the three generations, respectively. The gauge eigenstates mix to form mass eigenstates with a mixing proportional to the Yukawa couplings between the scalar field of the Higgs boson and the Dirac field of each fermion. Since this mixing is proportional to the masses of the different generations of fermions, the mass eigenstates of the first two generations turn out to be the same as the masses are small. However, for the third one, where the masses are higher and the Yukawa couplings non-negligible, the mass eigenstates are different. The mass eigenstates for *staus*, *stops* and *sbottoms* are denoted as $(\tilde{\tau}_1, \tilde{\tau}_2, \tilde{\nu}_\tau)$, $(\tilde{t}_1, \tilde{t}_2)$ $(\tilde{b}_1, \tilde{b}_2)$

The vector bosons of the Standard Models have fermionic superpartners called *gauginos*. For the color gauge interactions that are mediated by the gluon g the fermionic superpartner $\tilde{g}$ is called *gluino*. For the electroweak gauge symmetry associated with the gauge bosons W^+, W^0, W^- and B^0 we have the superpartners $\tilde{W}^+$, $\tilde{W}^0$, $\tilde{W}^-$ and $\tilde{B}^0$ which are known as *winos* and *bino*. When electroweak symmetry is broken the W^0 and B^0 are mixed into the Z^0 and γ gauge eigenstates and the supersymmetric parterns W^0 and B^0 mix as well into the *zino* $\tilde{Z}^0$ and *photino* $\tilde{\gamma}$ gauge eigenstates. For the Higgs boson, a single chiral supermultiplet is not enough, because if it were, the electroweak gauge symmetry would be inconsistent as a quantum theory [21]. The partners of the Higgs field are called *higgsinos* and they are a weak isospin doublet of charged $\tilde{H}^{\pm}$ and neutral $\tilde{H}^0_{1,2}$ spin-half eigenstates. The neutral higgsinos and the neutral gauginos combine to form mass eigenstates called *neutralinos* $\tilde{\chi}^0$. On the other hand, charged higgsinos and winos mix to form two charged mass eigenstates called *charginos* $\tilde{\chi}^{\pm}$.

There is a new symmetry that is introduced in some SUSY models called R-parity that assigns a positive value $R = +1$ to the SM particles and a negative one $R = -1$ to the superpartners. R-parity allows for baryon B and lepton L number conservation. If R-parity is conserved the decay of every supersymmetric particle to any SM particle would be forbidden and consequently the lightest supersymmetric particle (LSP) will be neutral and stable i.e. a good dark matter candidate.

SUSY particles have not been observed in the experiments so far and they are assumed to have extremely high masses, much larger than the ones of their SM partners. To account for the difference in mass between SM particles and superpartners there should be a spontaneous symmetry breaking of the theory. This breaking can be mediated by different mechanisms.

Extra dimensions

Extra dimension theories are those which add extra dimensions to the spacetime. They were first proposed by Kaluza and Klein [22–24] in a theory with five spacetime dimensions. Now, there are several types of models regarding how many dimensions are added, how they are included and the way in which the fields are incorporated in them.

Some of the models incorporate new flat dimensions which are separable to the existing ones. These can be accessible to gravity only, as it is the case of the *large extra dimensions* models, referred as ADD [25], which consist on models with $(4 + n)$ compact spacetime dimensions with relatively large radius. In this case, the SM fields are confined in a 4-dimensional subsurface, called *brane*. Other models, like the *universal extra dimension* (UED) models [26, 27], add additional spatial dimensions where all fields, including the SM ones, are incorporated.

Other models, called *warped extra dimensions*, incorporate extra dimensions which have an intrinsic curvature. This kind of model was first proposed by Randall and Sundrum in the RS theory [28], which considered a 5-dimensional Anti-de Sitter spacetime. The spactime in this case contains two branes, one where the gravity is relatively strong and other where the SM fields are contained. The gravity in the second brane is weak as its strength decreases exponentially between branes.

Hidden Valley theories

Hidden Valley (HV) theories [29] are theories which add a hidden sector of generally light particles that interact only weakly with the SM particles. These theories generally add a new confining non-Abelian gauge group G_ν with new conservation laws that extends the SM gauge group G_{SM}. All SM are neutral under G_ν and the new particles charged under G_ν are neutral under G_{SM}. The interaction between the SM fields and the particles of the hidden sector would take place by quantum mixing effects with new heavy particles that are also introduced. These heavy particles would have both G_{SM} and G_ν charges.

HV models contain therefore a visible sector with the SM fields, the hidden sector with suppressed coupling with SM and a set of heavy mediators that couple to both of them.

Compositeness and technicolor

Compositeness are models which in general propose to solve the hierarchy problem by replacing the fundamental Higgs scalar with the vacuum expectation value of a composite scalar operator constructed with fundamental fermion fields [20].

Specificaly, Technicolor models (TC) [30] postulate new gauge interactions which break the chiral symmetry of some new fermions, called *techni-fermions*. TC is assumed to be a confining theory consisting on TC singlets which are not found as stable states in nature. The intrinsic confinement mass scale must be of the order of the weak scale. These singlets would decay into the light observed quarks and leptons.

Little Higgs theories

Little Higgs (LH) theories [31] are proposed as a solution to the hierarchy problem which state that our SM Higgs boson is a pseudo-Nambu-Goldstone boson arising from the breaking of a new strong continuous global symmetry. In these models, the new strong coupling scale is of the order of 10 TeV and no fine-tuning is required to control the Higgs mass high order terms. LH models include new particles which cancel the one-loop quadratic divergent contributions to the Higgs mass from the SM loops.

Chapter 2

The LHC and the CMS detector

The analyses presented in this thesis were performed with collision data collected by the CMS Collaboration in the CERN Large Hadron Collider (LHC). In this chapter, a brief description of the LHC and the CMS detector is provided, together with the introduction of some physics concepts that are referred over the whole thesis. First, the LHC is described in Section 2.1 and then the CMS detector and its different subsystems are described in Section 2.2.

2.1. The Large Hadron Collider

The Large Hadron Collider (LHC) [32] is a circular particle accelerator that was built at CERN and operates since 2010. It is 27 km ring structure of superconducting magnets placed at 100 m underground with several accelerating systems that gradually increase the energy of the particles that travel along the accelerator. Within the ring, beams of hadrons are accelerated until they reach a centre-of-mass energy of $\sqrt{s} = 13$ TeV and collide in one of the four detectors that are placed around the collision points, namely CMS, ATLAS [33], LHCb [34] and ALICE [35]. Although the analysis presented in this thesis studies proton-proton collisions, other LHC physics programs also involve the collision of proton-lead and lead-lead.

The protons are extracted from hydrogen gas which is ionized by means of an electric field which strips the atoms off their electrons. Once ionized, the hydrogen atoms are first injected into a Radio Frecuence Quadrupole where they are squeezed into beams. The beams enter the chain of accelerators to increase their energy before being injected into the LHC pipes. The first accelerator is LINAC2, a linear accelerator that increase their energy up to 50 MeV. After they reach this threshold, the protons are injected into the Proton Synchrotron

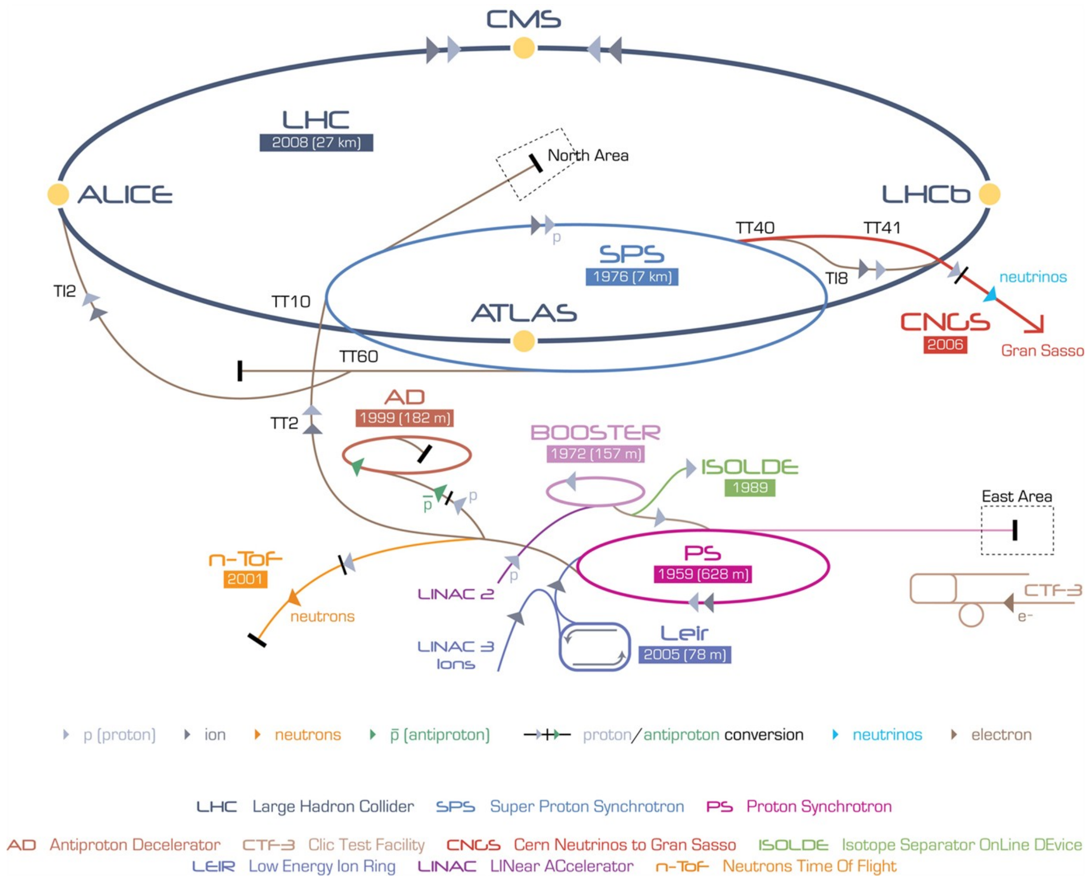


Figure 2.1: Scheme of the LHC accelerators chain [36].

Booster where they are accelerated up to 1.4 GeV and they are split into different bunches of nearly $2 \cdot 10^{11}$ protons. The bunches are injected in the Proton Synchrotron where they reach an energy of 26 GeV. Finally, the bunches enter the Super Proton Synchrotron where they are accelerated until they have 450 GeV and are injected into the LHC pipes with a bunch separation of 25 ns.

The proton bunches increase their energy as they pass through a chain of 16 superconducting radio frequency cavities located at one point of the accelerator circumference. The cavities operate at a frequency of 400 MHz and provide a 2 MV acceleration voltage which oscillates synchronized with the arrival time of the protons. The direction of the protons is set by a set of superconducting magnets which bend the trajectory with a magnetic field of 8.3 T. The temperature for these magnets to operate at the LHC is set to 1.9 K by cooling the system with liquid helium. Beam losses are avoided by means of quadrupole magnets to collimate the beam as it traverses the bending magnets.

When the bunches reach an energy of $\sqrt{s} = 13$ TeV[1] they collide at the interaction points inside the particle detectors. A scheme of the LHC acceletaror

[1]The $\sqrt{s} = 13$ TeV energy is achieved in the Run 2 physics program of the LHC. In Run 3 data taking era, the energy was leveled up to $\sqrt{s} = 13.6$ TeV.

chain with the four detectors is shown in Figure 2.1.

In collider physics, the instantaneous luminosity $\mathcal{L}$ is defined as the number of proton-proton collision per unit area and unit time t. When a collision takes place in the detector, the probability that a certain process has to occur is given by the cross-section σ. The event rate dN/dt is proportional to the instantaneous luminosity by means of the interaction cross-section

$$\frac{dN}{dt} = \sigma\mathcal{L}(t) \tag{2.1}$$

The cross-section is a fundamental physics quantity that measures the quantum mechanical probability of the interaction and is derived from the corresponding matrix elements of the Feynman diagrams involved in the process. The instantaneous luminosity, on the other hand, depends exclusively on the collider and the beam parameters. It can be approximated as a Gaussian beam distribution of the form [32]

$$\mathcal{L} = \frac{N_b^2 n_b f_{\text{rev}} \gamma_r}{4\pi\epsilon_n\beta^*} F \tag{2.2}$$

where N_b is the number of particles per bunch, n_b the number of bunches, f_{rev} the revolution frequency, γ_r the gamma factor, ϵ_n the normalized transverse beam emittance, β^* the amplitude function and F the so-called geometric luminosity reduction factor due to the crossing angle of the colliding beams at the point of interaction.

The magnitude representative of the amount of data collected for the analysis is called integrated luminosity L (or luminosity, for simplification) which is obtained from the instantaneous luminosity

$$L = \int \mathcal{L} dt \tag{2.3}$$

The bunches are composed by many protons and there is a certain number of proton-proton collisions that occur simultaneously at every bunch crossing. For this reason, several interaction vertices with outgoing particles can be reconstructed in every collision event. This effect is known as pile-up and may hinder the event construction or impact the identification criteria, which makes it a limiting factor for the data analyses.

Both the luminosity integrated from 2015 and 2018 and the corresponding average number of pile-up interactions are shown in Figure 2.2.

2.2. The CMS Detector

The data used in the analysis of this thesis was collected by the Compact Muon Solenoid (CMS) [38], a multipurpose hermetic detector which operates at the

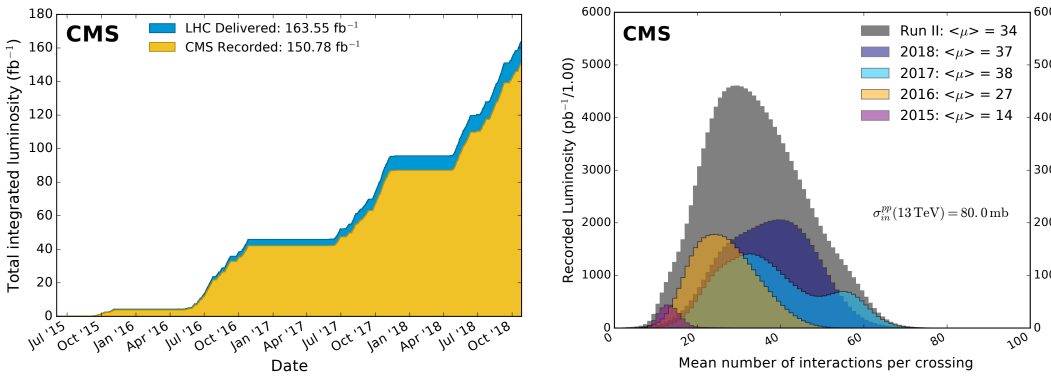


Figure 2.2: Run 2 Integrated Luminosity delivered by the LHC and recorded by the CMS experiment during Run 2 (left) and the distribution of the average number of interactions per bunch crossing in the same data taking period (right) [37].

LHC at CERN. The CMS detector counts with a high-field solenoid to bend the trajectory of charged particles, a synchronized set of sub-detectors designed to both reconstruct and identify the different types of outgoing particles produced in the pp collisions, and a sophisticated data acquisition system. The latter includes the trigger, a powerful mechanism with selects the most relevant collisions among the vast amount of interactions that occur in the bunches collision. The collected data is processed through the Worldwide LHC Computing Grid (WLCG), which provides the necessary global computing resources for the storage, distribution and analysis of the data generated by the LHC.

The CMS detector has cylindrical geometry and the beam is directed along its axis. The whole structure is divided into three regions: one middle central region called barrel and two endcaps coupled at both sides. A scheme of the CMS detector is shown in Figure 2.3.

The coordinate system (ϕ, η) used to define the direction of outgoing particles is adjusted to the geometry of the LHC and the detector. The $\hat{z}$ axis of the detector is aligned with the beam and is therefore parallel to the LHC circumference. The $\hat{x}$ axis points towards the center of the accelerator, while the $\hat{y}$ axis is parallel to the vertical. The angle ϕ is the usual azimuthal angle in cylindrical coordinates, while the pseudorapidity is defined in terms of the θ angle between the direction of the particle and the beam $\hat{z}$ axis

$$\eta = -\ln\left(\tan\frac{\theta}{2}\right) \tag{2.4}$$

Particle momenta are denoted as p and its projection over the transverse plane formed by the axes $\hat{x}$ and $\hat{y}$ is denoted as p_T. It is worth noticing that the total component of the transverse momentum $\vec{p}_T$ of colliding hadrons is null by

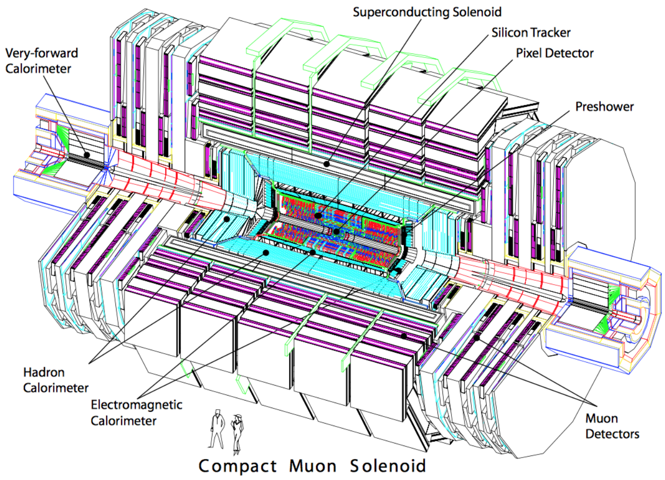


Figure 2.3: Scheme of the CMS detector [38].

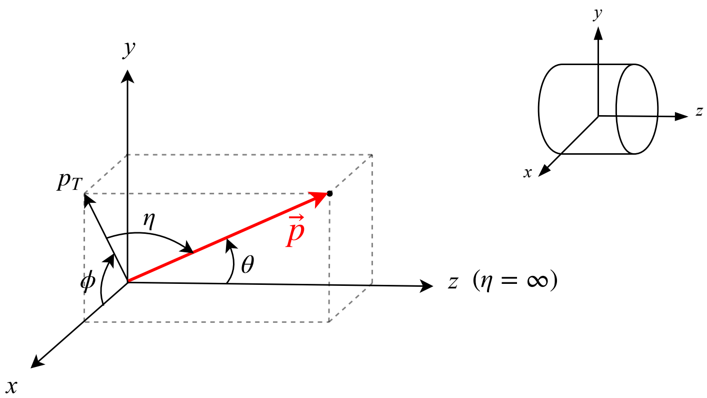


Figure 2.4: Coordinate system of the CMS detector.

definition. The coordinate system is presented in Figure 2.4.

In the following sections, a brief description of the subsystems of the CMS detector is provided.

2.2.1. The Superconducting Solenoid

The CMS detector is embedded in a magnetic field that bends the trajectory of the particles emerging from the pp collisions. The curvature of the reconstructed paths is measured to provide a measurement of the transverse momentum of charged particles.

The magnetic field is generated by the CMS magnet [39], a superconducting solenoid of 6 m of diameter that has a length of 12.5 m. It is made of NbTi, cooled with liquid helium, and able to generate up to 3.8 T magnetic field thanks to the 19 kA that run through it. This device is found in between the hadronic calorimeters and the muon system.

2.2.2. The Inner Tracker

The CMS silicon tracker [38, 40] is the subdetector that reconstructs the trajectory of the charged particles that traverse the CMS detector. It also provides a precise reconstruction of the primary and secondary vertices of the collision, enabling the identification of the particles that were produced in the primary process, particles from the pileup vertices and also heavy-flavor decays.

It is a cylindrical structure with a length of 5.8 m and a diameter of 2.5 m which has an active area of nearly 200 m^2 of silicon. It is subdivided in two separate subsystems: the inner pixel tracker and the silicon strip tracker. The pixel tracker is found in the innermost part of the detector and it was originally composed of three barrel layers at radii of 4.4 cm, 7.3 cm and 10.2 cm. It is completed by two disks of pixel modules in the endcaps, on each side. For 2017 the original pixel tracker was upgraded to the Phase-1 pixel tracker [41], where one barrel layer and one endcap disk were added to the original design. The innermost radius of the pixel tracker was extended to 2.9 cm with the new layer. Figure 2.5 shows a scheme of the Phase-1 pixel tracker compared with the original design. The silicon strip tracker deploys 10 layers in the barrel that extend the tracker structure from 20 cm up to a radius of 1.16 m. It is composed of the Tracker Inner Barrels and Disks, referred as TIB/TID, which are four barrel layers placed up to a radius of 55 cm and 3 endcaps disks. The outermost barrel region is the Tracker Outer Barrel (TOB), which is composed of 6 layers that cover a radius of 116 cm and a longitude along the z axis of ± 118 cm. The tracker is finally completed by two Tracker Endcaps, composed by 9 disks on each side, extend the coverage in $|z|$ from 124 cm to 282 cm. The one in the positive direction labeled as TEC+ and the one in the negative direction as TEC-. Each TEC radii range between 22.5 to 113.5 cm.

The whole tracker system provides spacial coverage of $|\eta| < 2.5$. A layered scheme of the whole inner tracker is shown in Figure 2.6.

2.2.3. The Electromagnetic Calorimeter

The system dedicated to the reconstruction and identification of electrons and photons is called Electromagnetic Calorimeter (ECAL) [38, 43] and it is a hermetic homogeneous calorimeter designed to precisely measure their energy. The ECAL system is located after the inner tracking system and divided into a central region, called the ECAL barrel (EB), and two ECAL endcaps (EE). While the former covers pseudorapidity values up to $|\eta| < 1.48$, the endcaps extend the coverage up to $|\eta| < 3$. The region where these two subsystems are coupled, called barrel-endcap transition region, covers the pseudorapidity range of $1.4442 \leq |\eta| \leq 1.566$, and it is known to have limited efficiency.

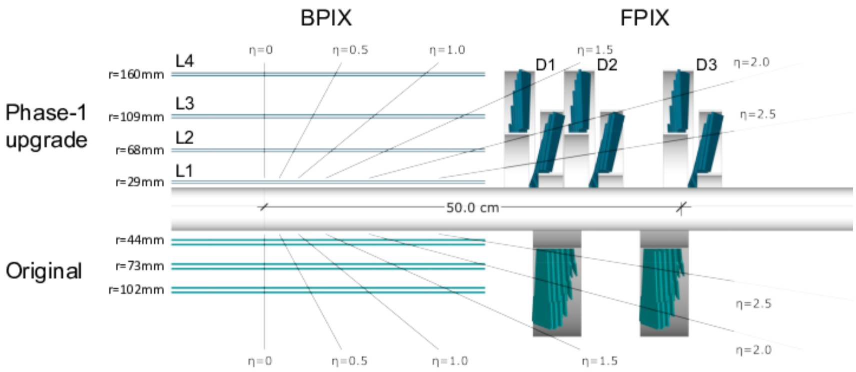


Figure 2.5: Layout of the CMS Phase-1 pixel tracker and the original detector layout, in longitudinal view [41].

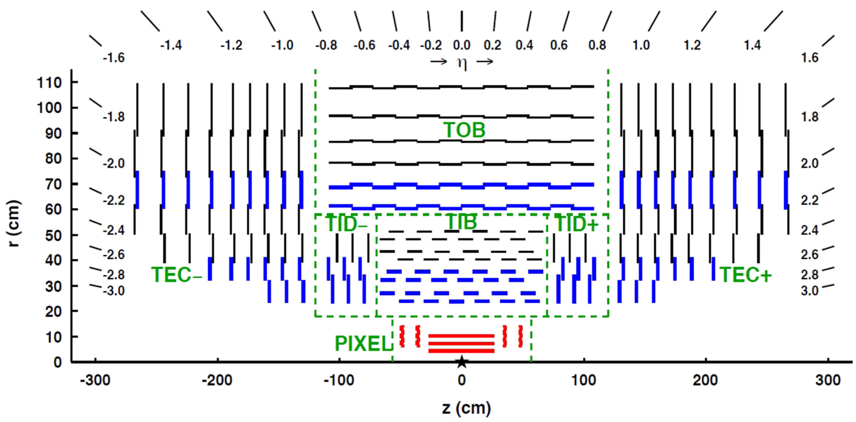


Figure 2.6: Scheme of the CMS tracker slice in the r-z plane. Pixel modules are shown in red, single-sided strip modules are represented as black thin lines and strip tracker modules are shown in blue lines. [42].

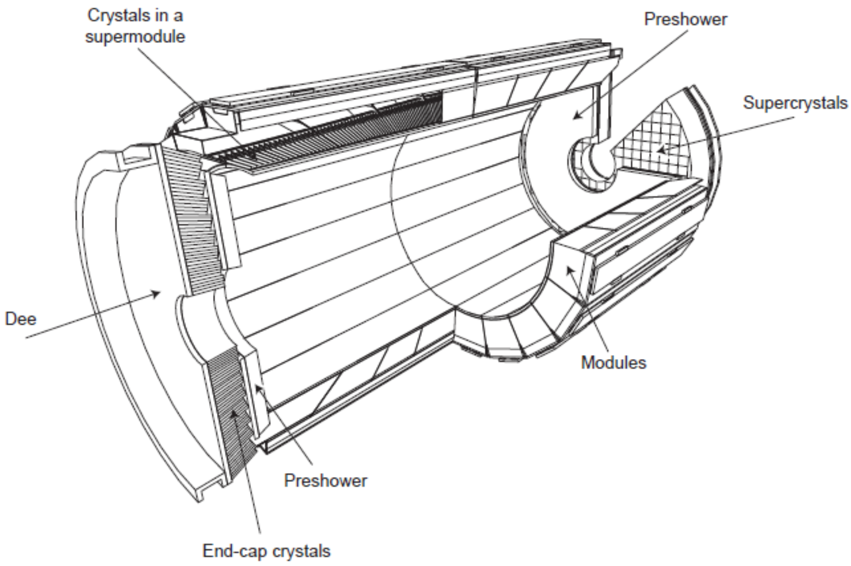


Figure 2.7: Scheme of the CMS ECAL detector including the barrel EB, endcaps EE and the ECAL preshower [43].

The ECAL is made of 61 200 $PbWO_4$ scintillating crystals which is a material with high density, short radiation length and small Molière radius. Whenever an electron or a photon hits a crystal surface, it starts a showering process based on chains of electron-positron pair production and photon emission. The light produced in the scintillating crystals is collected by Avalanche Photodiodes (APDs) in the barrel and vacuum phototriodes (VPTs) in the endcaps. There is an additional sub-system placed at $1.65 \leq |\eta| \leq 2.6$ called the ECAL preshower. It is a sampling calorimeter with two layers made of lead absorbers and silicon strip sensors. This detector helps to improve photon identification and to identify neutral pions in the endcaps. A sketch of the different ECAL subsystems is shown in Figure 2.7.

2.2.4. The Hadronic Calorimeter

There is a second calorimeter called Hadronic Calorimeter (HCAL) [38, 44], designed to measure the energy of hadrons produced in the hadronization of quarks and gluons. The overall structure is made of brass, which acts as absorber material, and plastic scintillators which compose the active medium of the HCAL detector.

The HCAL is composed of several subregions optimized to operate in different parts of the detector under specific conditions. There is a central region in the barrel (HB) covering pseudorapity values up to $|\eta| < 1.3$. It is divided into two half-barrel sections that are made of flat brass absorber plates which are assembled parallel to the beam axis. Two hadron calorimeter endcaps (HE) are placed at both sides of the HB to cover the pseudorapidity range of $1.3 < |\eta| < 3$. The HE are made of C26000 cartridge brass, a non-

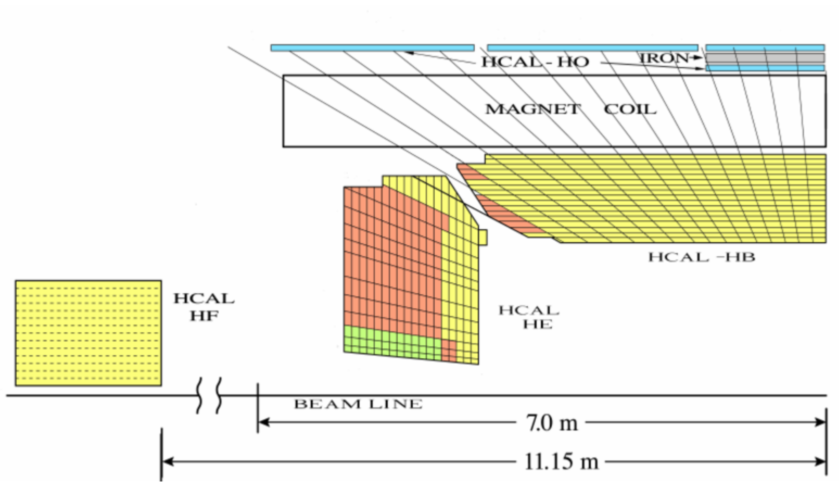


Figure 2.8: Scheme of a quarter view of the CMS hadron calorimeter in the $R - z$ plane [45].

magnetic material which allows to contain hadronic showers in the endcaps. Both the HB and HE subsystems occupy the limited space available in between the electromagnetic calorimeter at $R > 1.77$ m and before the magnet at $R < 2.95$ m As in the central region the stopping power of the combined HB+HE system is not enough, an outer calorimeter (HO) is placed out of the solenoid at $|\eta| < 1.3$ complementing the barrel region. Finally, the forward hadron calorimeter (HF) is placed at 11.2 m from the interaction point to extend the pseudorapidity coverage beyond $|\eta| > 3$ down to $|\eta| < 5.2$. It uses a Cherenkow-based radiation-hard based technology to cope with the extremely high hadron rates and the hostile environment that is present in this part of the detector.

A scheme of the CMS hadron calorimeters is shown in Figure 2.8 indicating the locations of the HB, HE, HO and HF subsystems in the $R - z$ plane.

2.2.5. The Muon System

The muon system [38, 46] is placed in the outermost part of the CMS detector and it is of key importance as it provides efficient and robust measurement of muons. It works as a tracking detector that identifies muons, reconstructs their momenta and triggers on their experimental signature. The solenoid magnet and its flux-return iron yoke enable to achieve good muon momentum resolution and also high trigger capability. In addition, the iron yoke also reduces the amount of background in the muon system as it absorbs part of the hadrons that are produced in the collisions.

For Run 2, this subsystem deployed three different types of gaseous detectors that are placed in different regions of the muon system. These are called drift tube (DT) chambers, cathode strip chambers (CSC) and resistive plate chambers (RPC) All of these together cover $|\eta|$ values from 0 to 2.4, being optimized to identify the muons and measure their momenta under the different conditions that may be present along this range.

The central region is called the barrel muon system and it uses Drift Tube (DT) chambers, that are arranged into 4 stations, placed at different r values. The DT chambers are composed of rectangular drift cells that cover a range of $|\eta| < 1.2$. The drift cells are rectangular cells with two electrode strips at the top and bottom, two cathode strips placed at both sides and a wire anode of 50-μm-diameter crossing the cell which embeds the whole structure into a non-uniform electric field. The cells contain a mixture of 85% Ar + 15% Co_2 that is ionized when the muons pass through. The drift cells are staggered by half a cell into layers. The smallest independent unit of the barrel central region is called superlayer (SL) and it is composed by 4 layers of cells. Then, a DT chamber is made from 2 or 3 SL. The wires of the 2 outer SLs are aligned with the beam line, so they provide a measurement of the ϕ coordinate of the muon. The inner SL wire is orthogonal to the beam line and therefore measures the z coordinate. It is only present in the first 3 stations, so the fourth one just provides a ϕ-only measurement.

The CMS Endcap Muon system is composed by Cathode Strip Chambers (CSC) which are arranged in rings which are labelled in blocks named ME1/1, ME1/2, ME1/3, ME2/1, ME2/2, ME3/1, ME3/2 and ME4/1, where the first number sorts the position of the rings in $|z|$ and the second one in r. The CSC chambers are trapezoidal detectors that split the ϕ angle in sections that can cover either $10°$ or $20°$, and provide continuous ϕ-coverage. Their pseudorapidity range goes from 0.9 to 2.4 and they overlap with the DTs in the range $0.9 < |\eta| < 1.2$ and with the RPCs in $0.9 < |\eta| < 2.1$. Each CSC chamber is made of 7 trapezoidal cathode panels which form 6 gas gaps where 6 anode wire planes are comprised. The gas is a mixture of 40% Ar + 50% CO_2 + 10%CF_4 which is ionized when traversed by a muon. The wires of the same plane run over ϕ and give a measurement on the radial coordinate r. The cathode panels are divided in strips with constant $\Delta\phi$ width and the ϕ value of the muons is obtained by interpolating charges induced in the different strips.

The Resistive Plate Chambers (RPC) are gaseous parallel-plate detectors with a good time resolution that allows to tag the time of an ionizing event. RPCs provide the time and position of muon hits with high enough accuracy to identify the LHC bunch crossing (BX) to which a muon is associated. An RPC chamber consists of two gaps named as *upper gap* and *lower gap* that operate in avalanche mode with common pick-up read-out strips in between. Both gaps contribute to the total signal induced in the read-out strips. RPC chambers are arranged both in the barrel and in the endcaps in two different structures. In the barrel, they are assembled with the DT chambers forming 6 coaxial sensitive cylinders. In the endcaps, the RPCs are alternated with the three iron disks that constitute the iron yoke.

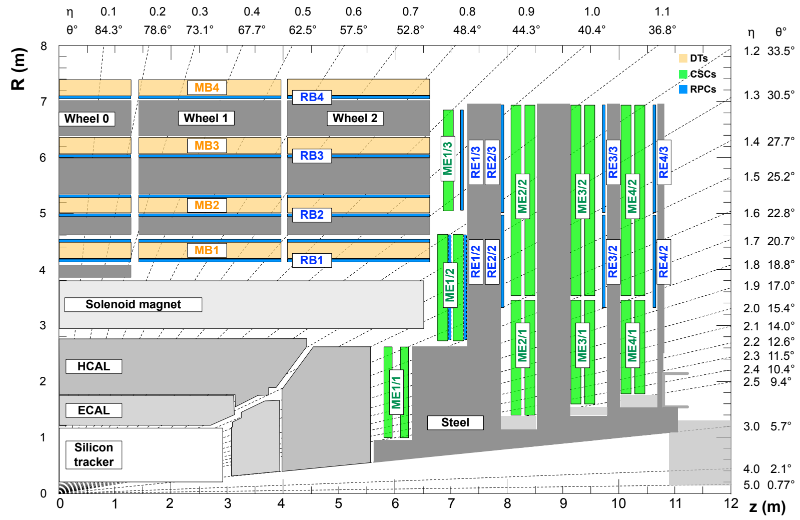


Figure 2.9: Layout of the quadrant of the CMS detector in its Run 2 configuration (from 2015) [47]. Muon chambers of the different subdetectors (DT/CSC/RPC) are colored.

2.2.6. The Trigger System

The LHC produces hadron collisions at an initial rate of 40 MHz. It is not possible for the particle detectors to cope with this amount of data and process every event that is produced. The Trigger system of CMS was designed to identify the most interesting events from all the collisions that occur at the bunch crossings and effectively filter the ones that are to be stored and processed. The trigger manages to deliver a final rate of 1 kHz, that is, it reduces the produced data four orders of magnitude. The CMS trigger is composed of two parts: The Level-1 (L1) trigger and the High-Level Trigger (HLT).

The L1 trigger operates at hardware level and makes use of the electronics of both calorimeters, ECAL and HCAL, and the muon system. It runs a fast, very simple particle candidate reconstruction from the detector information provided by these subsystems. This reconstruction takes around $\sim$4 μs.

Each of the aforementioned subsystems has its local L1 trigger. Energy superclusters are reconstructed in the calorimeters that further serve to reconstruct electrons, jets and hadronic taus. Muon hits are collected in the muon system and potential muon candidates are reconstructed by means of three track finder algorithms called BMTF, OMTF and EMTF [2] [48], which sort the muon candidates to the Global Muon Trigger (μGMT). Once both calorimeter and muon objects are processed, they are combined into a Global Trigger system that makes the decision if the event passes the L1 filter or not. The L1 algorithms identify very simple topologies, containing single or double objects with very low p_T thresholds and passing loose conditions. The final rate de-

[2]BMTF: *Barrel Muon Track Finder*, OMTF: *Overlap Muon Track Finder* and EMTF: *Endcap Muon Track Finder*

livered by the events accepted at L1 level has an upper limit of 100 kHz. These events serve as seeds for the HLT reconstruction that only runs on top of L1 filtered collisions.

Unlike the L1 trigger, the HLT is implemented in software and profits from the information provided by the whole CMS detector to perform a more precise reconstruction of the particle candidates and reduce the event rate even more. Each HLT algorithm is commonly referred as HLT path and it is characterized by running over the events that were previously filtered by one (or more) L1 triggers. The candidate reconstruction is more similar to that of the offline reconstruction, resulting in a better estimation of the particle parameters and higher identification efficiency. The HLT event reconstruction is remarkably slower than the L1 one, but it is structured in such a way that the particles are reconstructed and filtered in layers that become more complex and restrictive. Therefore, the majority of the events are rejected in the early stages of the HLT reconstruction and only the most likely particle candidates are processed by the most complex algorithms. The paths are optimized to target specific topologies that may suit for the future analyses. The final rate delivered by the HLT is limited to about 1 kHz.

2.2.7. Event processing and the Grid

The large amount of data collected by the trigger is stored and further analyzed by the CMS offline computing system [49]. The processing of all the luminosity delivered by the detector requires both computing and storage resources, which are provided by the Worldwide LHC Computing Grid (WLCG). The Grid is composed by a series of infrastructures which are geographically separated, established in different researching centers and laboratories that are connected by means of an extensive internet network.

The WLCG is staggered in structures that are known as Tier. The first structure that processes the raw data is a farm of CPU based at CERN, named Tier-0, which also performs additional monitoring tasks and fast calibration of the detector. The events are reconstructed by the next computing layer, called Tier-1, that encompasses a total of 11 computing centers which process and store the data. Finally, the last layer is known as Tier-2 and it is composed of nearly 50 physics institutes, where the datasets are finally stored. This thesis has made extensively use of the local Tier2 infrastructure at IFCA

There are successive stages of data processing to produce the higher-level physics objects and apply the necessary calibrations to perform the data analyses. The event information in every processing step is written in a supported *data format* that can be read by the analyzers. In general, each one of these steps produces more specific information while discarding part of the event content to reduce the size.

The collisions collected in the detector are first stored in the so-called `RAW` data format, which contains the primary record of the physics events. At this point, the data has been formatted together with the L1 and HLT selections and some of the high level quantities derived from HLT reconstruction. `RAW` data is processed to create the `RECO` format, which contains all reconstructed objects as described in the previous sections. It is then processed into `AOD` format, which is a subset with the reconstructed objects where some hits and other detector-level info that could be useful to characterize the objects is kept. The `AOD` format is further processed by slimming the physics objects. They undergo through a more restrictive selection and other algorithms are run to discard more information and reduce the event size. The resulting data format is called `MiniAOD` [50] and it is based on CMS's *Physics Analysis Toolkit* (PAT) object classes. Finally, the `MiniAOD` samples are processed to produce a "flat" format, called `NanoAOD`.

The analysis presented in this thesis essentially uses a custom extended version of the `MiniAOD` format, processed from `AOD`, as the muon physics objects were not available in standard `MiniAOD` samples.

2.3. Event reconstruction in CMS

The signals measured in the different subdetectors are used to reconstruct the different elements that are produced in the proton collisions. This process is commonly referred as *event reconstruction* and it involves the reconstruction of the different objects and observables that characterize the event.

2.3.1. The Particle Flow algorithm

The Particle Flow (PF) [51] algorithm reconstructs the event by combining the information provided by all the subdetectors to reconstruct and identify the particles. This information encompasses the tracks from the inner tracker, the energy superclusters reconstructed in the calorimeters and the signals detected in the muon system.

In the local reconstruction, tracker tracks associated to signals in the muon system are identified as muons while tracks compatible with ECAL superclusters are associated to electrons. Charged hadrons are made from inner tracks linked with HCAL depositions. Neutral particles, as neutral hadrons and photons, do not interact with the tracker and they are associated with unlinked HCAL and ECAL superclusters. The PF algorithm acts after local reconstruction, deploying a linking algorithm that connects all these elements by means of geometrical criteria. It combines the information of the object specific algorithms to produce the most accurate description possible

2.3.2. Tracking and vertexing

Whenever a charged particle pass through the tracker, it is expected to leave hits in the traversed silicon layers that can be combined to reconstruct their trajectories. The CMS tracker deploys an iterative algorithm [52] which runs over multiple iterations until all potential tracks are identified. This iterative process works in such a way that the initial iterations search for tight tracks, that are easy to identify, while the most challenging ones are the last of the loop. The last iterations are dedicated to non-standard topologies, like displaced tracks, which will be of great relevance for the work presented in this thesis. More details will be given in Chapter 4.

Each track reconstruction starts with an initial *seed*, that is, a pairs or triplet of hits that serve as an initial estimate for the track parameters and act as a proto-track. The hit pattern recognition and track fitting is done by means of a Kalman Filter (KF) [53]. To build the track, the trajectory is propagated iteratively through the tracker layers with the KF to find compatible hits. If a hit is found to be compatible with the track in a given layer, it is added to the track and its parameters are updated for the next extrapolation. The process is repeated in the successive layers until the extrapolation reaches the end of the tracker. After all associated hits are combined, the track is fitted to extract its parameters and its hits are removed for subsequent iterations.

The reconstructed inner tracks are fitted to reconstruct the primary vertices that correspond to several interactions in the same event. An Adaptive Vertex Fitter is used for this purpose [54]. Then, the primary vertex of the hard interaction is identified with the vertex with the highest quadratic sum of the p_T of the tracks that were used to fit it.

2.3.3. Muon reconstruction

The CMS detector deploys a powerful muon system and a sophisticated tracker which are exploited to reconstruct muons in a wide range of energies, even under very high background conditions. Muons are expected to leave a track in the inner tracker and also DT/CSC/RPC hits in the muon system. These DT and CSC hits are combined into *segments*, which are pairs of hits that are reconstructed locally. The detection of the muon in the muon system provides an efficient identification while the high resolution of the tracker provides a very precise measurement of the muon properties. The combined operation of the two systems together make the muons a very robust object to use both for discoveries of new physics and precision measurements.

There are three standard algorithms [55] to reconstruct muons in CMS, which are optimized to reconstruct muons produced in the neighborhood of the beam spot:

- **Standalone muons** are muon tracks reconstructed only in the muon system starting from a initial state, called *seed*, which is made from DT or CSC segments. This state is extended by using a Kalman Filter algorithm that adds more hits compatible with the state and updates the track parameters. The standalone tracks are extrapolated to the closest point of the beam line and a beam spot constraint is applied to improve the momentum resolution.

- **Global muons** are built by extrapolating the standalone tracks to the inner tracker and finding a compatible inner track. The track matching is performed in a Region-Of-Interest close to the beam spot defined by the standalone track parameters. If the track matching succeeds, a global track is fitted from the hits of both the inner and the standalone track. The momentum resolution of the Global muons improves with respect to the initial Standalone muon.

- **Tracker muons** are made by taking the tracks reconstructed in the iterative tracking sequence and extrapolating them to the muon system. If a geometrically compatible DT/CSC segment or RPC hit is found, the inner track is identified as Tracker muon. Tracker muons recover some of the efficiency that is lost at low p_T, where the signal that the muons leave in the muon chambers is not enough to properly reconstruct a standalone track.

These are the main algorithms that are used in the majority of the analyses. However, CMS also deploys other algorithms, like the dedicated displaced muon algorithms which will be widely discussed in Chapter 4 as they are the ones used for the analysis presented in this thesis.

2.3.4. Photon reconstruction

Photons are reconstructed [56] as ECAL energy superclusters (SC) which do not have an associated tracker track. When an electron or a photon hits the ECAL surface its interaction with the dense material starts a cascade of secondary particles, commonly known as electromagnetic shower, whose energy is deposited in the nearby crystals. As photons are neutral particles, they are not expected to leave hits in the tracker layers and their reconstruction consist in building the SC from energy deposits in a process which is known as *superclustering*.

The ECAL supercluster reconstruction starts with a cluster seed made from an energy deposit whose energy is found over a certain threshold [57]. The energy threshold for the ECAL barrel is set to $E_{\text{seed}} > 230$ MeV while in the endcaps also the transverse energy is considered, being the thresholds $E_{\text{seed}} > 600$ MeV and $E^{\text{T}}_{\text{seed}} > 150$ MeV. Clusters within a geometrical window are merged into SCs by combining the energy deposited in the contiguous crystals of the initial seed which have an energy excess that doubles the noise level. In addition, electrons and photons may interact with the detector material and radiate bremsstrahlung photons, being these converted into electron-positron pairs. Superclustering is designed to include such photon conversion and bremsstrahlung losses.

In the analysis presented in this thesis, the SCs objects will be manually combined with tracker tracks to construct efficient displaced electrons from LLP decays as described in Chapter 4.

2.3.5. Electron reconstruction

Electrons are reconstructed in the CMS detector by combining the signals from the ECAL and the tracker [57]. Within the object reconstruction, superclusters belonging to electrons and photons are practically indistinguishable and are processed together. Nevertheless, electrons not only leave an energy supercluster in the ECAL but also hits in the tracker layers, allowing for their track reconstruction.

In order to reconstruct the track, the supercluster position is back-propagated to the nominal interaction point of the collision. Compatible hit pairs or triplets are used to seed the track reconstruction. The already mentioned radiative effects have important consequences for their reconstruction, as the electrons loss part of their energy before being detected and the photon emission results in an additional bending of the track curvature. The trajectory is built by means of a combinatorial KF and fitted with a Gaussian Sum Filter (GSF) [58] which takes into account the potential energy losses that the electron may have suffered.

Electrons made from compatible superclusters and GSF tracks are called GSF electrons. Their energy resolution is improved by combining the tracker momentum measurement provided by the track and the energy measurement of the ECAL supercluster.

2.3.6. Jet reconstruction

Quarks and gluons are produced in the hard interactions, but the nature of the strong interactions do not allow them to remain as singlet states. Whenever they are far enough from the quarks they undergo a process called *hadronization* [14]. The energy density between separated quarks starts to increase and new quarks and gluons are constantly emitted. This process continues until all the partons are gradually combined into colorless hadrons that travel through the detector in the form of jets of particles which conserve the direction of the initial quark/gluon.

Since jets are composite objects, CMS has deployed an algorithm which is optimized to gather the detected particles that are likely to proceed from the same original quark/gluon. This algorithm is called anti-k_T [59] and it collects particle candidates reconstructed by the PF algorithm. It iteratively groups the particle candidates until they are clustered into jets for a given radius parameter R. The algorithm works in such a way that a distance d_{ij} is defined for two different objects that may be particles i, j or pseudo-jets in construction:

$$d_{ij} = \min\left(\frac{1}{p_{Ti}^2}, \frac{1}{p_{Tj}^2}\right)\frac{\Delta R_{ij}^2}{R^2} \tag{2.5}$$

where ΔR_{ij} is the angular distance between both objects. If a given object i has $1/p_{Ti}^2$ smaller than d_{ij} with another object j, the two objects are merged into a single object by summing their momenta. If an object cannot be merged with other objects, it is excluded from the iterative process and promoted to jet. The algorithm gradually merges the remaining objects until all jets are reconstructed.

In CMS, the radius parameter is usually set to $R = 0.4$ and the resulting jets are named AK4 jets. Wider jets are also reconstructed with a radius parameter of $R = 0.8$, named AK8 (or fatjets) which may be used for specific searches.

B-tagging

Jets are produced from hadronization of gluons and quarks excepting the top quark which decays before developing any parton shower. Heavy-flavour jet identification techniques [60] allow to discriminate between jets originating from b or c quarks from those originated from light-flavour quarks or gluons. These jets are commonly referred as heavy-flavour jets or b-tagged jets.

B-tagging algorithms exploit the properties of heavy-flavour hadrons present in heavy-flavour jets, such as the relatively longer lifetime of b-hadrons, which leads to a measurable displacement for the b-hadron decay involved in the parton shower. For this reason, heavy-flavour jets are prone to contain displaced tracks arising from a secondary vertex (SV) that can be reconstructed.

B-tagged jets in Run 2 are characterized by the jet probability (JP) and the combined secondary vertex (CSV) taggers that were already deployed in Run 1 [61]. Run 2 improvements [60] also include a re-training of the combined multivariate analysis (cMVA) and the development of two versions of the CSV algorithm: The CSVv2 optimization and the new DeepCSV, which uses machine learning (ML) techniques.

2.3.7. Missing transverse momentum

As mentioned in Section 2.2, colliding protons only carry momentum in their direction of propagation through the beam pipe. This means that the total amount of momentum in the transverse plane is null. Therefore, due to conservation of momenta, the vector sum of the transverse momentum of all the particles present in the same event should be zero i.e.

$$\sum_i \vec{p}_{Ti} = 0 \tag{2.6}$$

In practice, this is not usually the case for several reasons. First, the resolution of the particle properties is limited and some unbalance in the vectorial sum is expected due to instrumental effects. On the other hand, some particles may escape out of the detector's acceptance and therefore they would not be reconstructed and thus excluded from the sum. Other processes may have the presence of neutrinos which, as they interact very weakly with matter, would remain undetected. Finally, some proposed exotic particles may not interact with the electromagnetic field nor with the detector material and therefore appear as a momentum imbalance.

The negative sum of the transverse momentum of all the reconstructed particles in the event is called *Missing Transverse Momentum* $\vec{p}_T^{\text{miss}}$:

$$p_T^{\text{miss}} = -\sum_{i,} \vec{p}_T(i) \tag{2.7}$$

and it is often used in the searches for undetectable particles such as dark matter.

2.4. Goals of the LHC and the CMS detector

Up to now, the LHC is the world's largest and most powerful particle accelerator, being the latest addition to the CERN's accelerator complex. It started to operate in 2010 with the initial goal of studying the electroweak symmetry breaking mechanism, through the discovery of a Higgs boson. In parallel, an intense program of searches for new physics beyond the Standard Model was deployed by the LHC experiments. The Standard Model was also extensively tested and studied, especially regarding multi-jet QCD, the physics of the top quark, or the study of the CP violation on the B-sector in order to perform precise measurements of the Cabibbo-Kobayashi-Maskawa (CKM) mixing matrix. Finally, the properties of the quark-gluon plasma were studied in heavy ion collisions.

The greatest success of the LHC and the CMS and ATLAS experiments was the discovery of the Higgs boson in 2012, where both collaborations announced the observation of a new particle with properties compatible with the predicted SM-like Higgs boson [4, 5]. After this success, the LHC experiments intensified their BSM searches program especially concerning Supersymmetry and exotic-particle searches, constraining the parameters of many BSM theories.

Since no evidence of deviation with respect to the SM was found, the LHC experiments started to pay attention to unconventional signatures of new physics that might have escaped detection due to inefficiencies in the reconstruction process, optimized for prompt signatures. One of these topologies that has become very popular over the last years is that emerging from long-lived particles (LLPs), which because of their longer lifetimes can decay away from the interaction point of the detector. The motivation of this thesis is to perform an analysis to study the existence of one specific signature for LLPs. Chapter 3 of this document is exclusively dedicated to introduce the long-lived physics.

Chapter 3

Long-lived particles and their interactions

Long-lived particles (LLPs) have gained popularity over the last decades, becoming the focus of attention for a good part of the HEP community. They are well motivated by the theory, as they can arise naturally in a wide variety of exotic BSM models, some of which are presented in Section 3.2 of this chapter.

From the experimental point of view, LLPs exhibit a different phenomenology compared to the SM particles due to their relatively large lifetimes. They provide unique experimental signatures, summarized in Section 3.3, that may have not been observed before in a colliding experiment. These non-conventional signatures pose significant challenges when studying long-lived particles in data analyses. If new physics is realized in a model with long-lived particles, the sensitivity of the search could be significantly reduced. This fact would explain why they have not yet been observed, and also illustrates the need to develop new dedicated analysis techniques.

The core of this thesis is a search for long-lived particles decaying into displaced dilepton vertices. Section 3.4 focuses on describing the main aspects of this experimental signature and the most important properties of their reconstructed physics objects.

3.1. Long-lived particles

Unstable particles are expected to undergo a spontaneous decay process in which they decay and transform into other different particles. The decay chain of unstable particles continues until the final decay products are stable. This is the case of the majority of the particles of the SM.

The probability that a particle will decay is constant in time. This means that the probability that a particle has not decayed after a certain time t is given by

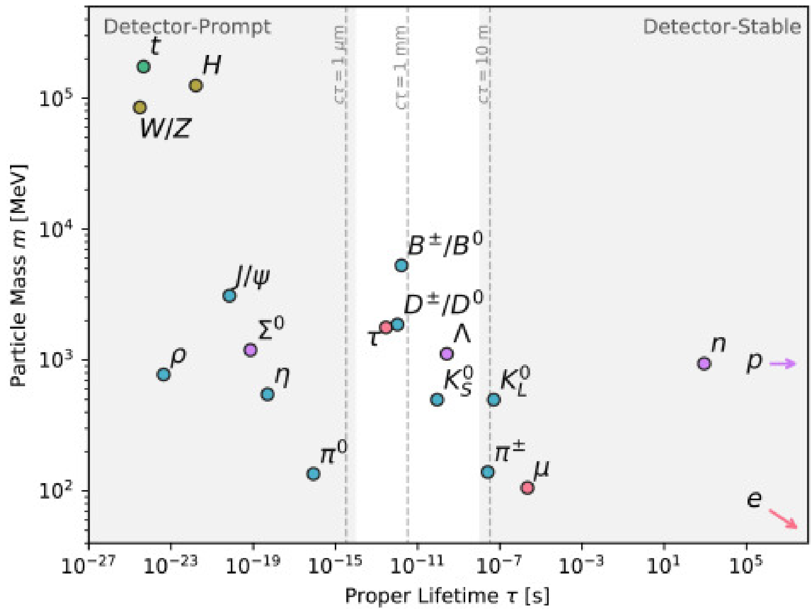


Figure 3.1: Standard Model particles as a function of their lifetime τ and mass [62].

$$P(t) = \frac{1}{\tau} e^{-\frac{t}{\tau}} \tag{3.1}$$

where τ is identified with the proper lifetime of the particle. Particle lifetimes in the SM span a wide range of orders of magnitude, as it can be seen in Figure 3.1, where the lifetime and mass of several SM particles are shown. Particles with long lifetimes, which will result in macroscopically measurable decay lengths, are commonly known as long-lived particles (LLP).

The inverse of the lifetime provides the decay rate Γ of the decay, and its value can be calculated using perturbative theory in the context of QFT. Large lifetimes arise naturally in the theory through three distinct mechanisms: the presence of feeble couplings, heavy off-shell mediators or compressed spectra.

One example of a long-lived SM particle that combines these three mechanisms is the charged pion $\pi^{\pm}$ in its decay into a muon and its neutrino partner:

$$\pi^{\pm} \to \mu^{\pm} \nu_{\mu}$$

The lifetime of the charged pion, when computed in its rest frame, is approximately $c\tau \sim 7.8$ m. This value is derived and found to be proportional to a term of the form

$$\frac{1}{\tau} \propto \left[\frac{g^2}{M_W^2} \frac{m_\mu}{m_\pi} \left(m_\pi^2 - m_\mu^2 \right) \right]^2 \tag{3.2}$$

In this specific case, the decay is determined by a small coupling term g^2, a heavy massive mediator $1/M_W^2$ and also by a phase space suppression given by a small mass difference $\left(m_\pi^2 - m_\mu^2\right)$ between the pion and the final state muon.

The lifetime of the particles will determine the way in which they are studied in the experiments. A LLP produced at the interaction point will travel a given length through the detector until it will eventually decay. The SM products of this decay will no longer emerge from the detector center, but from a displaced origin. On the contrary, the majority of SM particles such as the Z, W and Higgs bosons, whose lifetimes are of the order of $\mathcal{O}(10^{-25})$ s, will decay at the detector center. As a consequence, both the detection and analysis techniques developed during Run 1 and Run 2 at the LHC experiments, have been optimized to reconstruct and identify particles that decay promptly in the detectors.

Despite the high expectations generated by the prospects of BSM theories, no clear sign of new physics has been provided by the experiments since the discovery of the Higgs boson. These results suggest the possibility that new physics may appear in unexpected forms. In this context, exotic LLPs are one of the proposals that are currently in the spotlight of the HEP community.

LLPs pose significant challenges that the searches have to overcome. One of the most critical is the triggering, which in general is not optimized to reconstruct particles that were not produced promptly. In addition, the standard algorithms used to reconstruct the event information were not initially designed to work with non-conventional signatures that LLPs may leave in the detector. Furthermore, these searches usually have to cope with non-standard backgrounds that are hard (or even impossible) to reproduce by the simulation. The reduced sensitivity that may result from all or some of these effects could explain why long-lived particles have not been extensively searched for, and motivate the need to develop new fashion analysis techniques to provide a more robust insight.

3.2. Exotic long-lived particles in BSM theories

As introduced in the previous section, there are several mechanisms by which particles can acquire long lifetimes and, as a consequence, LLPs can be framed in a wide variety of BSM theories. It is important to remark that it is not easy to provide a categorization of such theories. On the one hand, models from the same category can predict very similar types of LLPs even though the details of the model itself are very different. On the other hand, several models can be encompassed by more than one category, making them non-mutually exclusive. Some model compilations can be found in the literature [63] and there is also a recent effort to provide a suitable model categorization to assess the coverage of future LLP searches [64]. This section gives a brief overview of theories that may include LLPs.

Supersymmetry

Supersymmetry is one of the most appealing BSM theories. It covers a wide range of different models with a very rich phenomenology. LLPs in SUSY models can emerge according to the three proposed mechanisms: feeble coupling, compressed spectra and scale suppression.

In this section, some of the most common examples are presented:

- The *Minimal Supersymmetric Standard Model* (MSSM) is the minimal extension to the SM and its particle content is the one described above. The LSP of the MSSM is often identified with the neutralino $\tilde{\chi}_1^0$ and the next-to the lightest supersymmetric particle NLSP can be long-lived provided that the difference in mass between the NLSP and the LSP is small.
- The RPV-SUSY models are non-minimally extensions of SUSY in which R-parity violating interactions are introduced. In such models B and L could be violated and LSP decays to SM particles would be allowed as long as the RPV decays are hierarchically small. The presence of such small RPV couplings would introduce macroscopic lifetimes for the LSP, making it long-lived. In principle, any superpartner could be the LSP and the largest RPV coupling with the SM particle would determine the final state of the decay.
- In *Gauge Mediated Supersymmetry Breaking* (GMSB) theories, the SUSY breaking is mediated by gauge interactions. GMSB models include the SM particles, their superpartners but also a *goldstino* that is absorbed by a *gravitino* $\tilde{G}$, the superpartner of the graviton and the LSP of the model. GMSB models are characterized by the effective mass scale of SUSY breaking Λ, which is proportional to the light gravitino mass. The SM superpartners decay to SM particles and $\tilde{G}$ via high dimensional operators and long lifetimes may arise for the NLSP if the SUSY breaking scale is small enough for the operators to be suppressed.

Higgs-portal theories

Within Higgs-portal theories, the SM-like Higgs boson couples to a new sector of particles non-connected to the SM gauge group. The coupling to this sector is expected to be very small, making possible the existence of new particles with long lifetimes. These kinds of models are well motivated as the small Higgs boson width provides the possibility to have large BSM fractions. In addition, the Higgs portal is a renormalizable portal where new particles can easily couple to the SM. These theories also provide a natural solution to the Hierarchy problem as higher order corrections would have the possibility to arise thanks to the new couplings that are introduced. Exotic Higgs decays may also exist in some *Hidden-Valley models* (HV) scenarios [29]. Specifically, the phenomenology of a HV model in which a Higgs boson decays into two neutral LLPs is studied in [65]. In this example, the LLPs decay into two oppositely charged leptons in the same displaced vertex, which is the theoretical scenario considered in this thesis.

In general, for Higgs portal models, each Higgs decay may produce two or more long-lived resonances that may lead to different final states. Furthermore, the Higgs associated production modes (gluon-gluon fusion, VBF, VH) may give additional prompt objects that would enrich the experimental signature.

Gauge-portal theories

Gauge-portal theories include scenarios where there are new vector mediators that couple to LLPs [64]. New vector mediators can arise in models with a new charge associated with a new gauge boson Z' [66] to which both SM and long-lived particles couple. Abelian or non-Abelian models predicting dark photons or dark Zs which couple to the SM particles by kinetic mixing [67] are also included in this category.

Heavy neutral leptons

Neutrino portal theories introduce the existence of heavy neutral leptons (HNLs) with right-handed chirality [68]. Their existence would imply the addition of new gauge-invariant mass terms for the SM neutrinos by means of a See-Saw mechanism [69] that could explain why the neutrino masses are extraordinarily small in comparison with other SM particles and therefore neutrino oscillations. If stable, an HNL could also be a good dark matter candidate.

Typically, an HNL is produced through its mixing with an SM neutrino with the corresponding matrix elements denoted as $V_{N\ell}$. Then, its decay is mediated by either a W^* or a Z^* boson. The lifetime of the HNL τ_N is strongly dependent on its mass m_N and the coupling with the neutrino $|V_{N\ell}|^2$ [70]. For small masses and low values of the mixing parameter, the HNL can become long-lived.

Long-lived particles in other dark matter theories

Other theories with proposed cosmological dark matter candidates provide more examples where particles with long lifetimes may arise.

Some theories under this umbrella predict new electroweak multiplets. An example is discussed in [71] in a model which proposes new massive vectorlike lepton doublets $(L^0, L^\pm)$. Provided that the mixing with SM fermions is suppressed, the neutral state of the representation L^0 will be effectively stable while the charged $L^\pm$, on the other hand, is expected to have a decay length of the order of $\mathcal{O}(\text{cm})$.

Other theories within this category are inelastic dark matter models (iDM) [72, 73] which postulate the existence of at least two inelastically coupled states in the dark sector χ and χ_*. A small mass gap $\Delta = m_{\chi_*} - m_\chi$ between heavy and light dark matter states leads to macroscopic lifetimes for the former.

There are also models that introduce a new particle in addition to the DM candidate, a co-annihilation partner (CAP) that can also be long-lived [74]. The proposed DM candidate χ and the CAP annihilate into the SM particles in a decay that is determined by the mass splitting $\Delta = m_{\text{CAP}} - m_\chi$. A small mass splitting results into long lifetimes for the CAP.

In general, as it is the case of the examples that were given, these models postulate the existence of a DM candidate and a LLP. This is a distinct feature of this category, as it also introduces a detector-level signature of a DM candidate in the experiment phenomenology.

3.3. The long-lived experimental signature

Due to their long lifetimes, which are large in comparison with those of the particles that are often studied in the colliders, LLPs can manifest in the form of non-conventional experimental signatures. The kind of signal expected strongly depends on the model, the theory parameters, the final state of the process and the kinematic properties of the LLPs. Nevertheless, while two theories may be very different from each other, they may give rise to the same signature from the experimental point of view, providing the opportunity to perform inclusive model-independent searches. There is a wide variety of different long-lived signatures arising from long lifetimes and Figure 3.2 shows a summary scheme of the most typical ones that are studied in the analyses.

In general, long lifetimes have been proved to be a powerful discriminator against standard SM processes as the non-conventionality of its signature introduces clear differences between the SM background and the expected signal. In fact, the kind of backgrounds that may be mistaken with the signal are often introduced by non-standard effects capable to mimic that signature.

The first thing that determines the experimental signature of a LLP is the kind of interaction (or no interaction) that it is expected to have with the detector material. Within this categorization, we can distinguish between direct detection and indirect detection of LLPs.

Signatures of direct detection are dominated by electric charged LLPs, which are expected to deposit hits in the silicon layers of the tracker. If the lifetime

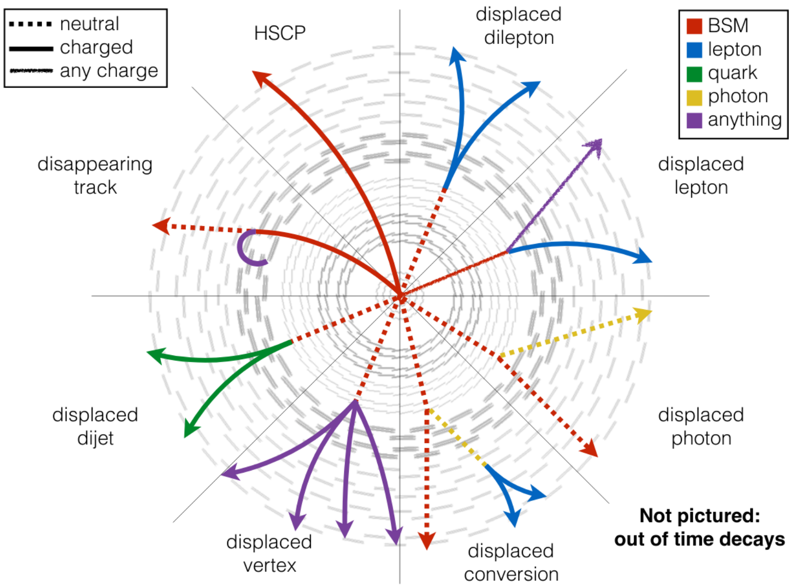


Figure 3.2: Scheme of the most popular experimental signatures arising from decays of long-lived particles [75].

is long enough for the LLP to travel through several layers before decaying, there is the possibility to reconstruct its inner track from the deposited hits. For LLPs that decay into undetected decay products before going through all the tracker, the experimental signature appears as a disappearing track i.e. a track that "seems to vanish" [76]. In the opposite case, for extremely high lifetimes or stable LLPs, it would be also possible to reconstruct a track in the muon chambers and have a global track fitted in the whole detector. This is the case of Heavy Stable Charged Particles (HSCPs) that due to their high masses are characterized to have small velocity $\beta = v/c$ and appear as high momentum global tracks with large rate of energy loss through ionization (dE/dx) [77]. Other stable (or quasi-stable) heavy LLPs which have fractional electric charge, result in global tracks but with anomalously low dE/dx instead [78].

Within indirect detection, LLPs are not expected to leave any energy deposit in the detector and they are inferred from their decay products. Therefore, the kind of experimental topology will strongly depend on the nature of the decay and the particles involved in the final state of the process. In general, particles arising from the LLPs decay vertex will appear as displaced from the interaction point instead of being produced in the centre of the detector as it is commonly expected from most of SM processes.

In cases where the LLPs decay hadronically the final state quarks will hadronize at a certain distance from the beamspot. This effect will result in displaced jets made from tracks that seem to emerge from a displaced vertex that could be also fitted. For large displacements, some of the tracks do not get to be reconstructed, leading to incomplete displaced jets with relatively low number of tracks. In extreme cases here the LLPs decay out of the tracker, the "jets" could also appear as anomalous clusters of muon hits. LLPs decaying into tau leptons may also manifest as displaced jets, provided that the tau decays hadronically.

Similarly, whenever a LLP decays leptonically into either electrons or muons, the outgoing leptons may be reconstructed displaced from the interaction point. Their tracks usually have fewer hits, and do not necessarily point to the primary vertex. If produced in pairs, both displaced lepton tracks could also be fitted to identify the displaced dilepton vertex. As the search presented in this thesis considers LLP decays into displaced dileptons, this signature will be widely described in the next section.

Finally, long lifetimes do not only affect the geometrical properties of the reconstructed objects, but also the time at which they are detected. High mass LLPs with relatively large lifetimes will spend a reasonable amount of time travelling through the detector before decaying. The corresponding decay products will be consequently detected later than usual, being known as delayed objects.

3.4. Displaced leptons and displaced vertices

The displaced lepton signature is a specific experimental topology in which LLPs decaying to leptons within the detector can manifest. In general, leptons are identified by reconstructing their trajectories in tracks that are built and fitted from the hits they left as they travel through the detector.

Leptons coming from promptly-produced particles, as from Z or W bosons, will travel through the whole tracker, being expected to leave hits in every layer that intercepts their trajectory. Trajectories built from their tracks naturally emerge from the primary vertex of the collision and any deviation from this point is usually attributed to instrumental effects or miss-measurements of the track parameters. Similarly, the number of layers without hits intercepting the track are labelled as *missing hits*, being often used with identification purposes as it has been proved to reject fake tracks made from spurious hit combinations.

LLPs, on the other hand, may decay in the middle of the detector provided that their lifetimes are long enough. Leptons arising from such decays will conserve the momenta of the decaying LLPs and travel inside-out leaving hits in the outermost layers to the decay vertex. On the contrary, the innermost layers to this point are not expected to have lepton activity. While for standard prompt leptons these would be labelled as missing hits and would indicate poor quality of the lepton track, for leptons produced in LLPs is a characteristic of their experimental signature. The fitted trajectory in this case does not necessarily pass by the primary vertex anymore, but it is expected to show some deviation and appear as "displaced".

The displacement of a track can be measured from the deviation of the trajectory with respect to a reference point (x_0, y_0, z_0), that in many cases is identified with the primary vertex. The main observables that are representative of this displacement are the so-called *impact parameters*, which are also part of the standard parameterization of the track.

In general, a track is parametrized by a reference position of the track (v_x, v_y, v_z) and 5 additional parameters. In the following definitions, (v_x, v_y, v_z) is identified with the point of closest approach (PCA) to the center of CMS i.e. the point of the trajectory that minimizes the transverse distance with the beam axis. The remaining five parameters are defined as follows:

- The signed inverse momentum q/p computed from the charge q and the total momentum p fitted from the track.
- The λ angle, which measures the angle of the track and the transverse plane at the reference point. It relates with the polar angle θ as:

$$\lambda = \pi/2 - \theta \tag{3.3}$$

- The azimuth angle ϕ at the given point.
- The transverse impact parameter d_{xy}, which is the signed distance in the XY plane between the straight line passing through (v_x, v_y) with angle ϕ and the reference (x_0, y_0). It can be computed as

$$d_{xy} = -(v_x - x_0)\sin(\phi) + (v_y - y_0)\cos(\phi) \tag{3.4}$$

- The d_{sz} parameter, which is defined by means of an additional S axis, that corresponds with the projection of the straight line passing through (v_x, v_y, v_z) with angles (ϕ, λ) into the XY plane. d_{sz} measures the signed distance in the SZ plane between this straight line and the (v_x, v_y, z_0) point. It is computed as

$$d_{sz} = (v_z - z_0)\cos(\lambda) - ((v_x - x_0)\cos(\phi) + (v_y - y_0)\sin(\phi))\sin(\lambda) \tag{3.5}$$

The parameters d_{xy} and d_{sz} are called impact parameters and are estimates of the displacement of track provided that (v_x, v_y, v_z) is the PCA as stated above. The distance d_{xy} is referred as transverse impact parameter and it can be also referred in the so-called perigee convention i.e. $d_0 = -d_{xy}$ [1]. For the displacement along the longitudinal direction, it is more common to use the so-called longitudinal impact parameter d_z which relates with d_{sz} as

$$d_z = \frac{d_{sz}}{\cos(\lambda)} \tag{3.6}$$

and coincides with $d_z \equiv v_z - z_0$[2]. Figure 3.3 shows a simplified graphical representation of the impact parameters in Cartesian coordinates as defined in the CMS detector frame.

[1] In this thesis the impact parameter is referred as d_{xy} or d_0 indistinctly. In all cases, the absolute value is used, which is equivalent for both, that is $|d_{xy}| \equiv |d_0|$.

[2] This definition holds as long as (v_x, v_y, v_z) is the PCA to (x_0, y_0, z_0).

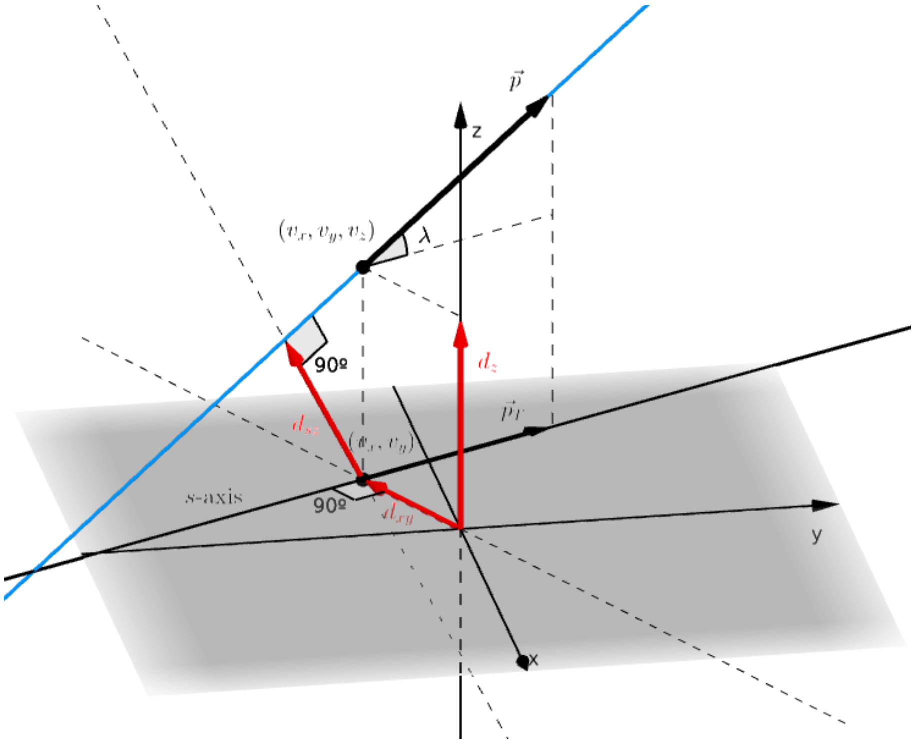


Figure 3.3: Graphical representation of the impact parameters d_{xy}, d_{sz} and d_z. The blue straight line represents the tangent of the trajectory in the PCA i.e. (v_x, v_y, v_z). For high p_T tracks where small or no curvature is appreciated, it would coincide with the trajectory itself. The axes are defined in the reference frame of the CMS detector, being the z axis aligned with the beam pipe. The center of the coordinate frame can be identified with either the center of the detector or with any order reference point (x_0, y_0, z_0) provided that the axes choice remains unchanged.

The absolute values of the transverse $|d_0|$ and longitudinal $|d_z|$ impact parameters are powerful discriminators of long-lived signals against the vast prompt background that is encountered in proton-proton collisions.

Leptonically decaying LLPs can decay either into a single lepton (and another particle) or into pairs of displaced leptons. In the latter case, not only will their tracks be displaced from the interaction point (IP), but they are also expected to emerge from the same point in the space i.e. the decay vertex. In practice, the two tracks are fitted to a common dilepton vertex (DV) which is ultimately identified with this decay vertex. Once it is reconstructed, it is also possible to measure the distance travelled by the long-lived particle.

In such cases it is common to use the *transverse decay length* L_{xy} that is defined as the distance travelled by the LLP in the XY plane before decaying. As the LLP is assumed to be produced in the hard interaction, this distance is measured with respect to the primary vertex of the collision:

$$L_{xy} = |\vec{x}_{\text{DV}} - \vec{x}_{\text{PV}}| \tag{3.7}$$

where here $\vec{x}_{\text{DV}}$ and $\vec{x}_{\text{PV}}$ denote the positions of the DV and the PV in the transverse plane, respectively.

Many searches targeting displaced leptons e.g. [79, 80], as it is the case of the analysis presented in this thesis, make use of the aforementioned observables to discriminate LLPs from prompt background. In general, it is more common to rely on the distanced measured in the XY plane and use the lepton $|d_0|$ and/or the L_{xy} distance. In some cases, the *significance* of $|d_0|$ i.e. its value relative to the uncertainty σ_d of the measurement $|d_0|/\sigma_d$, has proven to be a powerful discriminator.

3.5. Current status of long-lived particles searches

Over the last years, LHC experiments have been exploring the different long-lived particle topologies to determine if the data observations would be compatible with a BSM signature involving long-lived particles. In this context, ATLAS and CMS detectors have provided substancial insight in a wide variety of them. Although the current observations do not show any sign of long-lived new physics, the searches have constraint the most popular theoretical models predicting particles with long lifetimes.

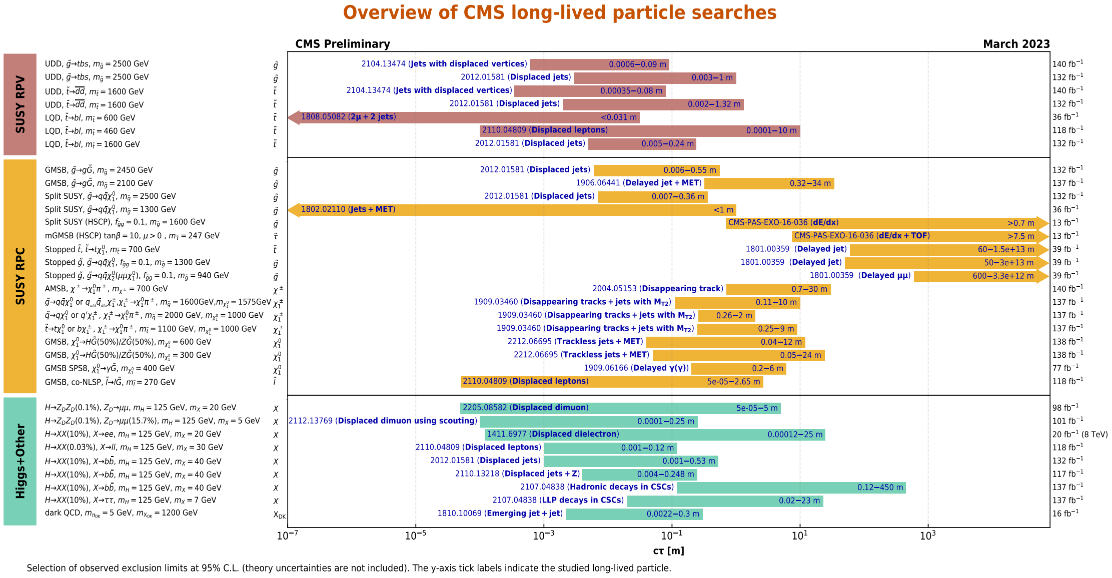
The long-lived experimental picture was last reviewed in the 57^{th} Rencontres de Moriond 2023 in La Thuille, where the most recent experimental results involving LLPs were presented to the HEP community [81].

On the hadronically decaying long-lived particles side, CMS recently released a new search for trackless jets [82] which manifest as high missing transverse momentum and the reconstruction of jets without tracks. Other presented analyses studying hadronic final states considered the displaced vertices topology instead. While the most recent results from CMS where presented in 2022 [83], constraining long-lived neutralinos/gluinos/stops, ATLAS presented in 2023 an analysis [84] studying this topology in the context of RPV long-lived neutralinos/charginos.

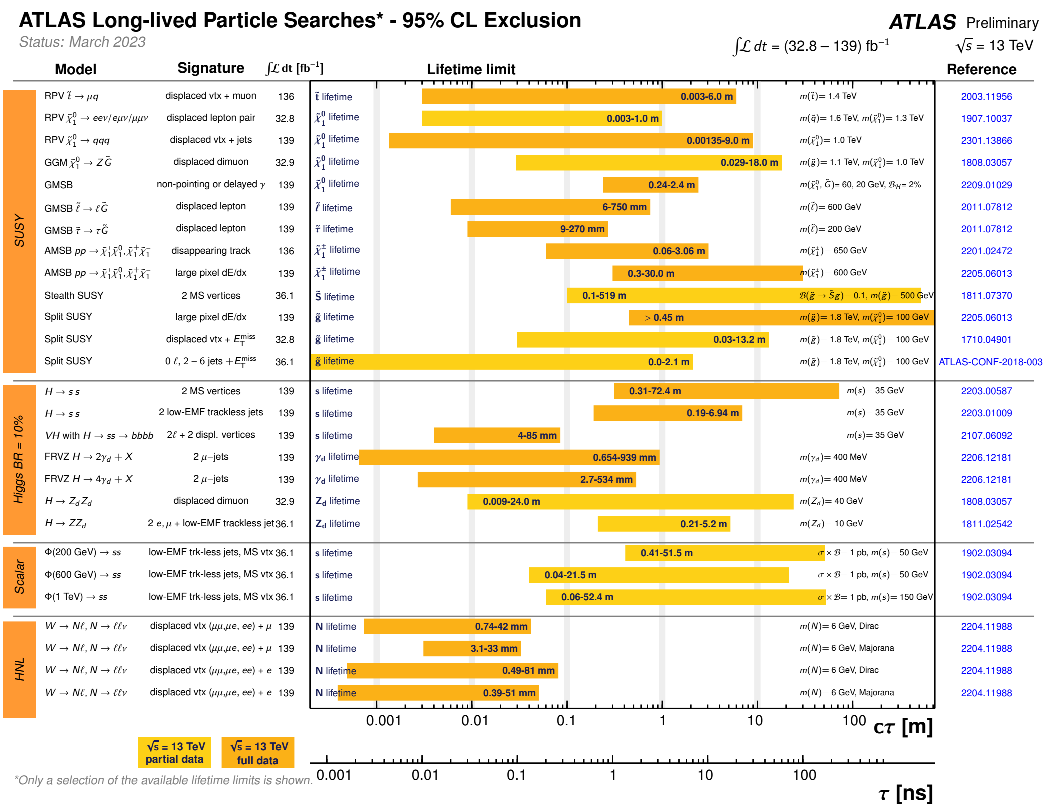
New results were also presented by ATLAS in Moriond 2023 studying displaced muons arising from LLPs with small lifetimes [85]. The analysis constraints on GMSB $\tilde{\mu}$ production complete the existing ones presented by the same collaboration in 2021 [86] increasing the sensitivity to lifetimes of the order of 0.01 cm. The parameter space probed by this analysis was already probed by the CMS Collaboration in [79].

Other recent results include the first search for inelastic dark matter at a collider [87] and results on HNLs in displaced leptons (including taus) [88] by the CMS experiment, and a search of heavy particles [89] yielding displaced $H \to \gamma\gamma$ or $Z \to ee$ by ATLAS.

Figure 3.4 collects an overview of the long-lived particles searches performed by the CMS and ATLAS Collaborations showing the lifetime range and the models excluded by each analysis.



(a) Overview of CMS long-lived searches performed until March 2023.



(b) Overview of ATLAS long-lived searches performed until March 2023.

Figure 3.4: Overview of the long-lived particles searches performed by the CMS and ATLAS Collaborations until March 2023 [90, 91].

Chapter 4

Displaced lepton and displaced dilepton vertex reconstruction

The standard CMS reconstruction algorithms that were presented in Section 2.3, which are the ones used by the majority of the analyses in CMS, are optimized to reconstruct leptons that are produced promptly. However, these algorithms have limited applicability in the case of long-lived particle decays, where the leptons are typically displaced from the interaction point. For this reason, using such reconstructions in analyses like the search presented in this thesis may lead to a significant loss of signal events and sensitivity, especially for LLP decays with a long lifetime. Furthermore, the typical identification criteria for standard electron and muon objects, also known as working ID, are optimized for prompt physics and impose certain requirements that may reject displaced candidates. These factors motivate the use of alternative reconstruction methods to recover part of the efficiency that is lost at high displacement and also the development of dedicated working IDs to ensure the quality of the reconstructed displaced leptons.

This chapter describes how electron and muon objects used in this thesis are reconstructed, together with the identification criteria that was optimized in each case. In addition, these objects are calibrated with both data and simulation to derive proper corrections that are later applied in the presented search.

4.1. Displaced muon reconstruction

As introduced in Section 2.3.3, standard muon algorithms do offer the possibility to reconstruct muons in the muon chambers only (Standalone muons), but also in combination with the tracker by joining together the tracker tracks and the signals detected in the muon system (Global and/or Tracker muons). However, their use is not optimal to reconstruct muons that were produced displaced from the interaction point.

In general, all these standard algorithms, regardless if they have a tracker track or not, have some constraints to the beam spot in both the pattern recognition and the fit to extract the muon parameters. While these constraints highly improve the momentum resolution for prompt muons, they turn out to be a strong limiting factor for displaced muons where the beam spot does not necessarily be a point compatible with the muon trajectory. Even if standalone muons could be a better handle for displaced muon searches, rather than Global/Tracker muons which require the existence of a tracker track, their momentum resolution is very limited.

Additionally, dedicated cosmic muon reconstruction methods [92] exist, removing the beam spot constraints. Nevertheless, their application in a LLP search is not feasible as they are algorithms strongly optimized for the cosmic muon signature. In fact, they are only available in dedicated cosmic runs and their routines are not even included in the standard event processing for events collected in the collision runs.

To overcome the lost of efficiency in displaced muons searches, the CMS Collaboration has developed dedicated displaced muon algorithms [93]. These are variations of the standard routines where the beam spot constraints were either relaxed or removed:

- **Refitted Standalone Muons (rSA)**: It is a variant of the Standalone muon reconstruction where the muon tracks are seeded by standard muon seeds but whose tracks are fitted without assuming any constraints to the interaction point.
- **Displaced Standalone Muons (dSA)**: It is a second variant of the Standalone muon reconstruction where the seeds used to build the tracks are dedicated cosmic seeds, made from pairs of DT/CSC segments where all possible bending directions are accepted. The dSA is then reconstructed by iterating over all compatible segments in successive stations. As in the case of rSA muons, the fit used to extract the track parameters does not assume that the muon trajectory is initiated at the beam spot.
- **Displaced Global Muons (dGlobal)**: It is a variant of the Global muon reconstruction. The seeds are dSA muons which are extrapolated to the tracker to find compatible tracker tracks. These tracker tracks are reconstructed with a dedicated muon tracking iteration. No vertex constraints are used either in the pattern recognition algorithm nor in the fit to extract the muon p_T.

The analysis presented in this thesis uses dGlobal muons to reconstruct the muons produced in LLPs decays.

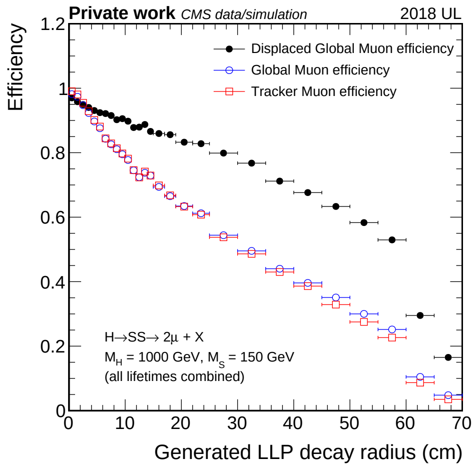


Figure 4.1: Reconstruction efficiency of dGlobal, Global and Tracker muon reconstruction algorithms as a function of the decay radius measured in simulated H decays into long-lived S bosons with masses $m_H = 1000$ GeV and $m_S = 150$ GeV.

In order to quantify the gain that these objects bring at high displacement, their performance is studied in simulated signal long-lived particle decays into muons. The simulated sample taken for the studies is the reference model with BSM Higgs decays into long-lived scalars S, which is the model in which the results of the search presented in this thesis have been interpreted. More details about the simulation will be given in Section 6. The masses of the two bosons are $m_H = 1000$ GeV and $m_S = 150$ GeV for lifetimes ranging as $c\tau_S = 1, 10, 100, 1000$ mm of the S boson.

The dGlobal muon performance is compared with the standard reconstruction algorithms that make use of tracker tracks. The efficiency as a function of the decay radius of the simulated S bosons is shown in Figure 4.1, where dGlobal muons exhibit a better performance compared to the one measured for both Tracker and Global muons. The different simulated lifetimes $c\tau = 1, 10, 100, 1000$ mm of the S boson are integrated to measure the efficiency.

In addition, the p_T resolution of dGlobal muons is measured in simulation and compared with the one provided by Global and Tracker muons. The figure of merit is the so-called p_T-residual, which measures the relative deviation of the reconstructed p_T^{reco} with respect to the generated p_T^{gen} for the simulated particle:

$$p_T\text{-residual} = \frac{p_T^{\text{reco}} - p_T^{\text{gen}}}{p_T^{\text{gen}}} \tag{4.1}$$

Figure 4.2 shows the normalized distributions of the measured p_T-residuals for the three reconstructions: Global, Tracker and Displaced Global muons. The p_T-residual distributions are found to be very similar for the muons reconstructed by the three reconstructions, within a resolution of $\sim 5\%$.

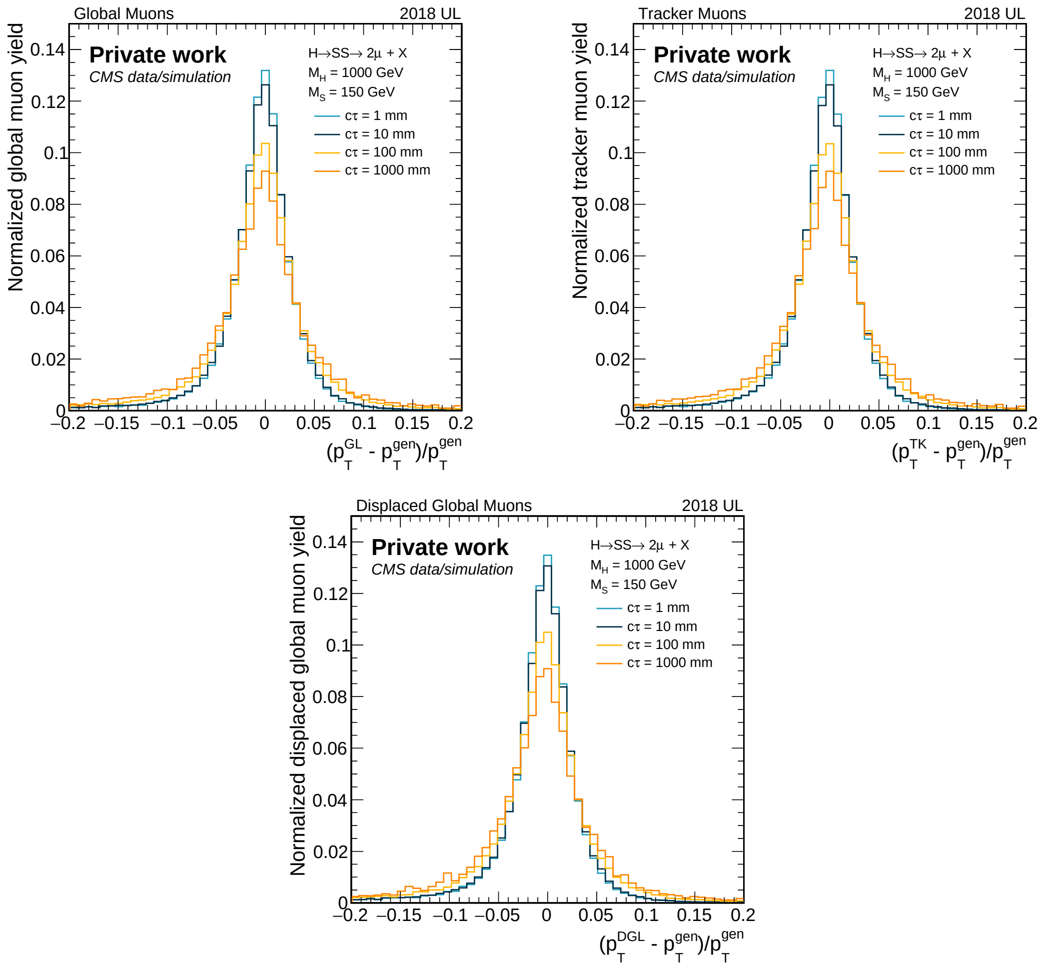


Figure 4.2: Transverse momentum resolution of muons reconstructed with the Global (left), Tracker (right) and Displaced Global (bottom) algorithms. It is computed as the p_T-residual with respect to the generated p_T^{gen} of simulated muons in several samples with long-lived S bosons whose lifetimes range from 0.1 to 100 cm.

These two results show that the dGlobal reconstruction is able to reconstruct a higher number of displaced muons with a p_T resolution comparable to the one provided by the standard algorithms.

4.1.1. Displaced Global muon identification

As dGlobal muons have not been used in a CMS search before, there is no standard criteria to select good quality muons reconstructed with this algorithm. A new dedicated identification selection (ID) was designed to target well reconstructed dGlobal muons produced in LLP decays and reject fake candidates. This so-called dGlobal ID has been optimized considering muons having $p_T > 10$ GeV and $|\eta| < 2.4$ but as it will be introduced in Section 7, the selection criteria applied to select the LLP candidates of the search is more

dGlobal working ID	
Variable	**Selection**
Transverse momentum p_T	> 10 GeV
Pseudorapidity $\|\eta\|$	< 2.4
Transverse momentum resolution σ_{p_T}/p_T	< 0.3
Number of hits in strip tracker layers N_{strip}	> 5
Number of hits in muon chambes N_{muon}	> 12

Table 4.1: Working ID applied to the dGlobal muons.

restrictive.

A maximum value in the transverse momentum resolution of $\sigma_{p_T}/p_T < 0.3$ is set to restrict the use of dGlobal with good p_T resolution. Additionally, a minimum number of hits is also required, which is applied for hits in the tracker and muon chamber separately. The tracker track must have a minimum number of hits in the strip layers $N_{\text{strip}} > 5$ while no requirement is imposed on the pixel hits to be sensitive to LLPs decaying beyond the pixel tracker volume. Tighter requirements on the number of hits in the tracker track are expected to have a negative impact in the dGlobal efficiency at high displacements. Finally, the muon track reconstructed in the chambers must have a minimum number of hits $N_{\text{muon}} > 12$. This requirement assures that the muons considered have hits in more than one DT station and it is chosen to suppress hadron punch-through.

The dGlobal ID is summarized in Table 4.1. Its efficiency is measured in simulated long-lived S boson decays into muons samples of the H→SS model with masses $m_H = 1000$ GeV and $m_S = 150$ GeV. The efficiency is shown as a function of muon p_T, η and the decay radius of the S boson in Figure 4.3. Identification efficiency is over 97% along the whole considered p_T range, reaching nearly 100% in the central parts of the barrel and gradually decreasing as muons are reconstructed towards higher $|\eta|$ values. The efficiency is kept over 95$\%$ up to 50 cm of displacement.

4.2. Displaced electron reconstruction

Standard CMS electrons are identified by pairing tracks reconstructed in the tracker with superclusters (SC) detected in the Electromagnetic Calorimeter (ECAL). However, it is known that electrons suffer from bremsstrahlung effects that make them loose a fraction of their initial energy before reaching the ECAL surface. These radiative effects usually affects to their pass through the tracker system making the standard tracking routines not optimal for their reconstruction. The standard electron reconstruction takes care of radiative affects by reconstructing the electrons with the specific algorithm that was described in Section 2.3.5. However, due to the fact that these algorithms are com-

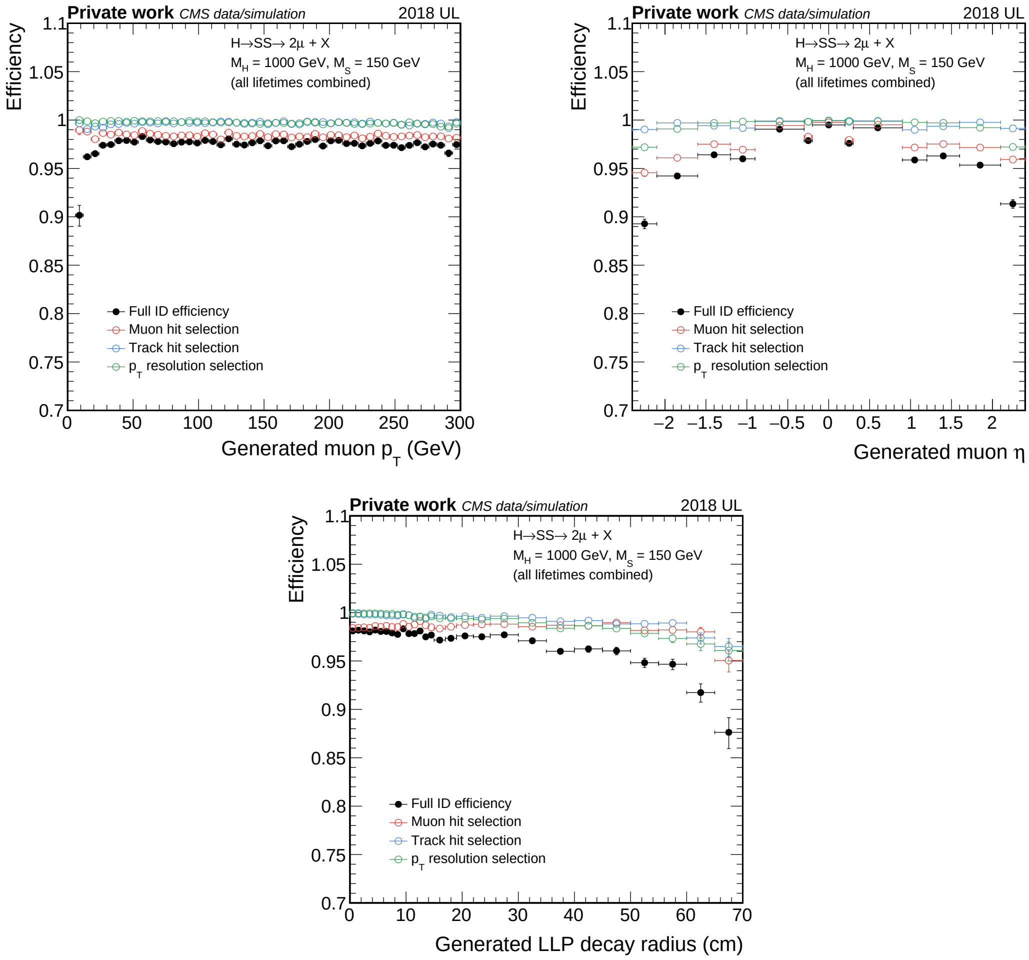


Figure 4.3: Efficiency of the dGlobal ID criteria applied on dGlobal muons reconstructed in a sample of simulated long-lived S decays. The results are shown as a function of the generated p_T (left), η and S decay radius (bottom). Single efficiencies of the hit and p_T resolution requirements are also represented.

putationally expensive, some of the latest iterations of the tracking sequence are excluded, including the ones that were optimized for displaced track reconstruction. Electrons reconstructed in this thesis aim to recover part of the efficiency that may be lost at high displacement by deploying an alternative electron object reconstruction.

The alternative electron reconstruction used in this search takes as starting point a subset of the full set of tracks built by the standard tracking sequence. This subset is called *isolated tracks* and contains track candidates reconstructed by all the tracking iterations provided that they pass a soft selection. All tracks with transverse momentum over $p_T > 20$ GeV are included while the ones below this threshold are only added if they have $p_T > 5$ and pass minimum isolation criteria. They are written in `MiniAOD` samples in their own `pat::IsolateTrack` class format of `CMSSW`.

To find which of the isolated tracks belong to electrons, their parameters are extrapolated to the ECAL surface in order to find compatible SCs that could be deposited by the same particle that left the track. The SCs are identified as offline reconstructed photons written in their own `pat::Photon` class format which contains ECAL SCs that without additional identification criteria could belong to electrons or photons indistinctly. An extrapolated isolated track and a SC are considered to be matched if the angular distance between their parameters is below $\Delta R(\text{track}, \text{SC}) < 0.1$, being this angular distance defined as

$$\Delta R(\text{track}, \text{SC}) = \sqrt{\left(\eta_\gamma - \eta^{\text{ext}}_{\text{track}}\right)^2 + \left(\phi_\gamma - \phi^{\text{ext}}_{\text{track}}\right)^2} \tag{4.2}$$

where the $\phi^{\text{ext}}_{\text{track}}$ and $\eta^{\text{ext}}_{\text{track}}$ values of the isolated track are extrapolated to the ECAL surface and $(\phi_\gamma, \eta_\gamma)$ are the angular coordinates of the SC (offline photon), measured in the ECAL surface. Figure 4.4 shows an schematic representation of this track-SC matching.

The degeneration that may result from multiple valid track-SC associations fulfilling the presented criteria is suppressed as follows. If two (or more) isolated tracks could be matched to the same SC object, the pair with the minimum $\Delta R(\text{track}, \text{SC})$ value is kept.

Electrons reconstructed according to these criteria will be referred as *displaced electrons* in the rest of the document. These displaced electrons are, together with the dGlobal muons described in Section 4.1, the main objects used in the search presented in this thesis.

The performance of the displaced electron reconstruction is measured in simulated signal long-lived particle decays into electrons. The simulated sample taken for the studies is the same $\text{H} \to \text{SS}$ simulation that was used to evaluate the dGlobal muon efficiency and resolution.

The minimum $\Delta R(\text{track}, \text{SC})$ distance between the reconstructed tracks of the

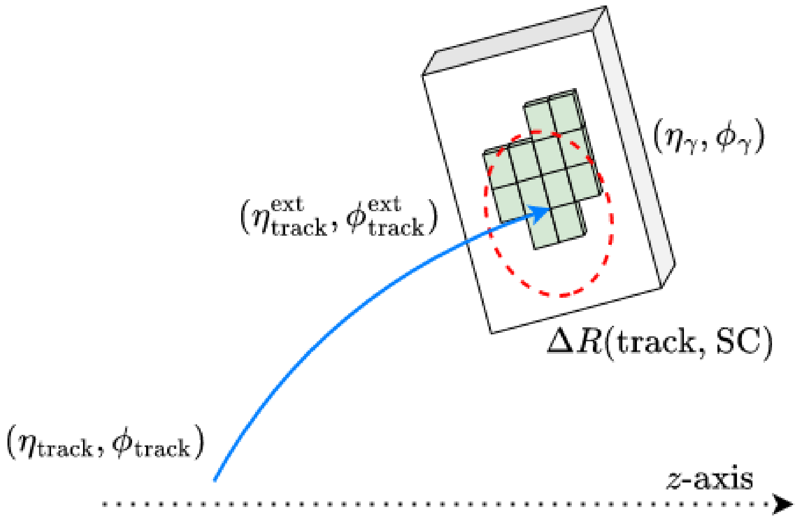


Figure 4.4: Schematic representation of the reconstruction of displaced electrons. The algorithm extrapolates tracker tracks to the ECAL surface to find superclusters that are compatible with the extrapolated track parameters ϕ^{ext}_{track} and η^{ext}_{track}. The extrapolated track is shown in blue, while the matched ECAL supercluster is shown in green.

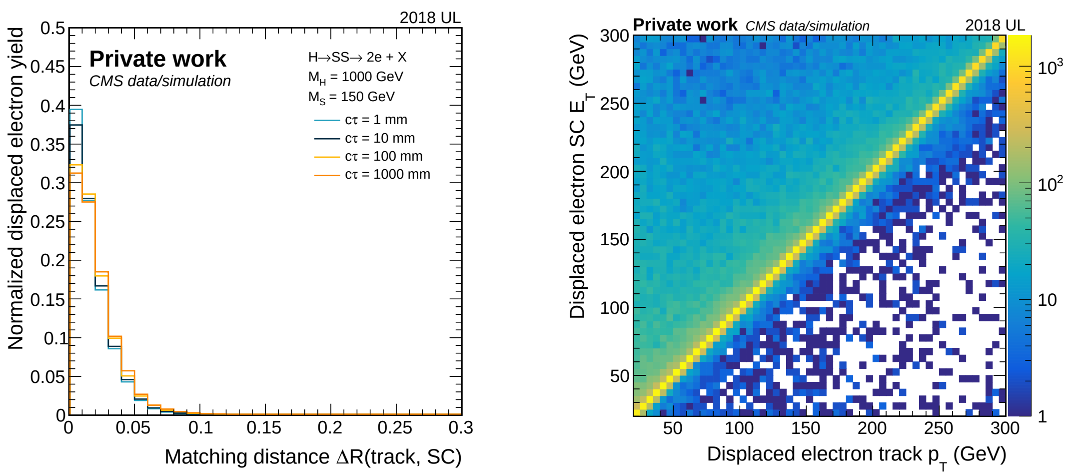


Figure 4.5: Left: Minimum $\Delta R(\text{track}, \text{SC})$ between matched tracks and superclusters corresponding to simulated electrons produced in S decays. Right: Energy E_T^{SC} of the ECAL supercluster as a function of the isolated track momentum p_T^{track} of the displaced electrons reconstructed within $\Delta R(\text{track}, \text{SC}) < 0.1$. The electrons are taken from a sample of long-lived S bosons with masses $m_H = 1000$ GeV and $m_S = 150$ GeV simulated under 2018 detector conditions and pile-up. The lifetimes $c\tau$ are shown separately for the $\Delta R(\text{track}, \text{SC})$ distributions while they are integrated for the E_T^{SC} vs p_T^{track} comparison.

leptons produced in the long-lived S boson decay, and the nearest SC is shown in Figure 4.5 for different lifetimes of the S boson. The minimum $\Delta R(\text{track}, \text{SC})$ is seen to increase with the lifetime, indicating that the displacement impacts the performance of the track extrapolation to the ECAL, and therefore the track-SC association. Despite the effect the displacement may have on the measured value of $\Delta R(\text{track}, \text{SC})$, the majority of the displaced electrons are found within the chosen $\Delta R(\text{track}, \text{SC}) < 0.1$ threshold even for the most displaced sample.

Matched tracks and SCs were reconstructed by different subsystems that provide different measurements for the p_T of the reconstructed displaced electrons, referred as p_T^{track} and E_T^{SC}, respectively. The tracker extracts the p_T^{track} value from a fit to the hits used to build the particle track while the ECAL measures E_T^{SC} from the amount energy deposited by the electron in the reconstructed SC. The comparison between E_T^{SC} vs p_T^{track}, is also shown in Figure 4.5 for the displaced electrons reconstructed from the displaced leptons of the long-lived S boson decays. The majority of the studied displaced electrons populate the diagonal region indicating that both measurements tend to be consistent with each other. However it is also seen that in a non-negligible amount of cases the p_T^{track} measurement of the track is underestimated with respect to the E_T^{SC} measurement provided by the ECAL supercluster. This result suggests that the E_T^{SC} estimation is more stable with displacement than the track p_T^{track} which suffers a small degradation. In light of the presented behavior, the angular variables η and ϕ of the electrons are taken from the track while the estimation of the p_T is taken from the SC, that is, identified with E_T^{SC}.

The reconstruction efficiency is studied in the simulated S decays of the same sample. Figure 4.6 collects the efficiencies computed as a function of the transverse decay length of the S boson. These efficiencies in Figure 4.6 (left) include the tracking efficiency, defined as the fraction of generated electron reconstructed as isolated tracks; the supercluster efficiency, which is the fraction of generated electron reconstructed as ECAL superclusters; and the displaced electron efficiency, defined as the fraction of times that a generated electron is reconstructed from the matching of both as described above. The reconstruction efficiency of electron objects as reconstructed with the standard CMS algorithms is shown for comparison. Figure 4.6 (right) shows the supercluster efficiency measured on electrons that are reconstructed as isolated tracks. This efficiency measures the additional efficiency loss on top of track reconstruction which is found to be around $\sim 93\%$ in the considered displacement ranges.

4.2.1. Displaced electron identification

A set of identification (ID) criteria was defined to identify the displaced electron suitable for analysis and assure their quality. This ID selection includes requirements imposed on the isolated track and the supercluster matched to construct the electron, separately. Unless specified, all requirements described in

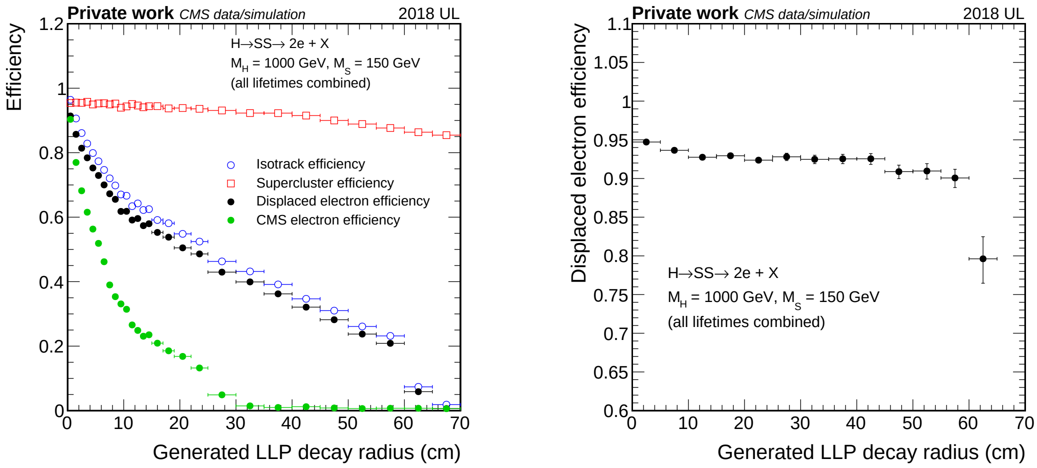


Figure 4.6: Left: Reconstruction efficiency of isolated tracks (blue circles), superclusters (red squares), reconstructed displaced electrons (black points) and standard CMS electron reconstruction (green points). Right: Reconstruction efficiency of the supercluster measured over simulated electrons that are required to have their isolated track reconstructed. The electrons are taken from a sample of long-lived S bosons with masses $m_H = 1000$ GeV and $m_S = 150$ GeV simulated under 2018 detector conditions and pile-up.

this section are applied before the isolated tracks and superclusters are paired together. Applying the selection before the pairing helps to reduce the number of extreme cases where a random isolated track is associated to a true supercluster that could be associated to the correct track, and viceversa.

First, minimum values of $p_T^{\text{track}} > 15$ GeV and $E_T^{\text{SC}} > 15$ GeV are imposed in both track and supercluster. However, no requirement on the p_T^{track}-E_T^{SC} agreement is applied as the track often underestimates the displaced electrons p_T with respect to the measurement provided by the ECAL supercluster (see Figure 4.5).

The isolated tracks must be contained within the geometrical acceptance set by the tracker volume, $|\eta^{\text{track}}| < 2.4$, rejecting candidates that could lie on the tracker boundaries. Isolated tracks are also required to have a minimum number of hits higher than $N_{\text{hits}} > 6$. This requirement was chosen to be loose enough to not reduce the tracking reconstruction efficiency when the electrons are produced in the outer layers of the tracker.

Displaced electrons whose supercluster is reconstructed within the narrow Barrel-Endcap region, in $|\eta^{\text{SC}}| \in [1.4442, 1.566]$, are rejected since they can lead to misintificated electrons or fakes. The ratio of the energy in the HCAL tower behind the seed crystal of the supercluster and the energy of the supercluster is required to have a maximum value of $H/E < 0.05$.

Finally, there is a cut on the measured spread $\sigma_{i\eta i\eta}$ of the supercluster. $\sigma_{i\eta i\eta}$

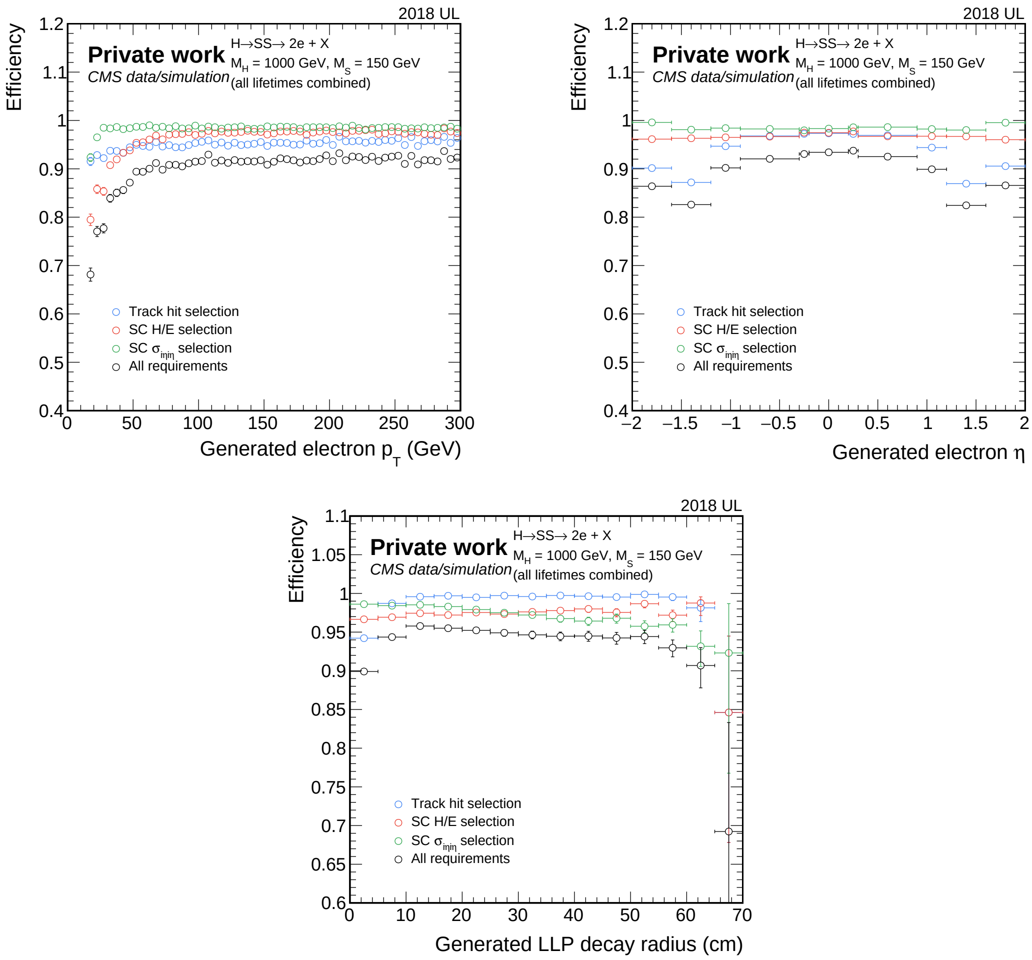


Figure 4.7: Efficiency of the ID selection applied on displaced electrons reconstructed in a sample of simulated long-lived S decays. The results are shown as a function of the generated electron p_T (left), η and the S decay radius (bottom). Single efficiencies of the track hit, SC H/E fraction and $\sigma_{i\eta i\eta}$requirements are also represented separately.

is a lateral shower-shape variable which provides a measure of the spread of an electromagnetic shower in the η direction. It is defined as the logarithm of the energy-weighted RMS of the electromagnetic shower in units of crystals. For the selection applied in this ID it must be below $\sigma_{i\eta i\eta} < 0.012(0.0425)$ in the Barrel (Endcaps) which is the same requirement imposed in the central cut-based "loose" ID for electrons reconstructed with standard algorithms, recommended when the backgrounds levels are rather low [94].

The displaced electron ID is summarized in Table 4.2 and the efficiency of the requirements, measured in simulated S $\to e^+e^-$ decays, is shown in Figure 4.7.

Displaced electron identification ID criteria	
Variable	**Requirement**
Isolated track p_T	> 15 GeV
Isolated track $\|\eta\|$	< 2.4
Number of isolated track hits	> 6
ECAL supercluster E_T	> 15 GeV
ECAL supercluster $\|\eta\|$	$\notin [1.4442, 1566]$
ECAL supercluster H/E	< 0.05
ECAL supercluster $\sigma_{i\eta i\eta}$	$< 0.012(0.0425)$ in Barrel (Endcaps)

Table 4.2: Working ID applied to displaced electrons.

4.3. Displaced vertex reconstruction

As previously introduced in Section 3.4, long-lived particles decaying to a pair of leptons can lead to a pair of displaced leptons. This signature consists on a pair of tracks emerging displaced from the interaction point. When both leptons are produced by the same LLP decay, the tracks emerge from a so-called displaced vertex.

In this search, LLPs are assumed to decay at least to a couple of leptons, in a such a way that the signal topology can be identified with the presence of reconstructed displaced dilepton vertices in the event. To reconstruct them, electrons and muons passing ID cuts (of each channel, respectively) are paired and their tracks are fitted to 2-track secondary vertices. The fitting process is done by means of a Kalman Vertex Fitter [95].

Among all the possible 2-track vertices that can result from the fitting of the reconstructed leptons of a given type, only the vertices with the lowest χ^2/ndof fit values are retained. Therefore, if one track can be paired with two different other tracks and be fitted to two distinct 2-track vertices, the vertex with the minimum χ^2/ndof is kept. The routine to reconstruct all possible vertices from the reconstructed leptons according to this criteria works as follows:

1. Iterate over the tracks i of a given lepton type (passing ID).
2. For each track i, iterate over the other tracks of the same type j $(i \neq j)$.
3. Fit all possible vertices V_{ij} that can be made from the fit of all i and j track combinations.
4. From all V_{ij} vertices, keep the one with lowest χ^2/ndof fit value, fitted from tracks i_{min} and j_{min}.
5. Steps 1, 2, 3 and 4 are repeated excluding already fitted tracks i_{min} and j_{min}.
6. Stop when there are not more tracks to fit or when the tracks that remain

cannot be fitted to a new vertex.

This procedure does not allow two distinct vertices to share the same track, and the set of vertices that is obtained is made from the track combinations that minimize the χ^2/ndof fit value.

Only in the case of dimuon vertices, it is additionally required that the muons have $\Delta R > 0.1$ before the pairing is done and the vertex is fitted. The reason for that is that is pure technical and depends on the format in which the muons are stored in the samples. The dGlobal muons are available only as raw `reco::Track` objects, which is the standard format for CMS tracks. These objects, unfortunately, have not run over the standard muon processing sequences of CMSSW. Apart from processing the information in a friendly format, these muon sequences apply additional processing on the muon tracks, such as the *duplicate removal*. Sometimes, when the muon trajectory is built, it may happen that the algorithms fail to merge all segments/hits from the same muon. Failures of these kind in the muon trajectory building may give rise to muon track duplicates. Since dGlobal muons are not processed by standard muon sequences that take care to clean muons sharing tracks or hits, there are some cases of dGlobal pairs where both muons are actually reconstructing the same physical muon. In such cases the muon tracks share a fraction of their measurements and are found very close to each other, with very low ΔR values. By applying the $\Delta R > 0.1$ requirement before the muon pairing is done, these muons are not fitted into spurious fake vertices, and they are given the chance to be paired to another genuine muon reconstructed in the event.

The vertex reconstruction efficiency was measured in simulated LLP decays of the H→SS model with masses $m_H = 1000$ GeV and $m_S = 150$ GeV. Figure 4.8 shows the efficiency as a function of the S decay length where electron and muon decays are studied separately. The denominator represents the number of generated lepton pairs that are also reconstructed as displaced electrons or displaced global muons as defined in previous sections. The numerator is defined as the number of these reconstructed lepton pairs that are fitted to a common vertex with the Kalman Vertex Fitter. The efficiency is measured for the full set of lepton pairs and additionally for the pairs in events with four generated leptons of the same flavor separately (both S bosons decay to the same lepton type). In general, the efficiency is higher for muons than for electrons, being the former close to the 100% and the second nearly $\sim 96\%$. In events with four reconstructed electrons it can be observed a degradation in the efficiency at low displacement, which may indicate the presence of cases where the vertex is fitted from the wrong electron pair.

The impact of lepton mismatch is evaluated by studying all the reconstructed vertices and checking the generated leptons that are matched to the reconstructed leptons used to perform the fit. Four different vertex types can be distinguished:

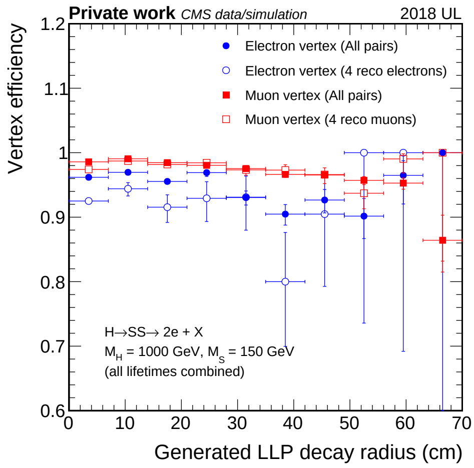


Figure 4.8: Efficiency to fit a signal displaced vertex when both generated leptons of the same simulated S boson are reconstructed. Measured in simulated LL S decays of the H→SS model with masses $m_H = 1000$ GeV and $m_S = 150$ GeV.

- Vertices fitted from leptons matched to generated leptons that were produced in the same S boson decay.
- Vertices fitted from leptons matched to generated leptons produced in two different S bosons decays (only in events where both S bosons decay to the same lepton type).
- Vertices fitted from one single lepton matched to a generated lepton produced in a certain S decay and from an unrelated lepton.
- Vertices fitted from leptons that are not matched to generated leptons.

The invariant mass of these reconstructed vertices is shown in Figure 4.9 for both lepton types. The vertices that are matched to the correct generated leptons manifest as a resonance peak around the mass of the simulated S boson while fake vertices fitted from one or two unrelated leptons accumulate at lower invariant mass values. The special case of vertices fitted from generated leptons from different S decays tend to populate higher invariant mass values, as expected from leptons with high momentum values as the ones simulated in this sample. Regarding these special cases, we note that they are more prone to occur at low displacement when the decay vertices from different S bosons may be found close to each other. This effect is also enhanced when only one lepton of the generated pair is successfully reconstructed (note this contribution is not taken into account in Figure 4.8 where we explicitly require that both generated leptons are reconstructed). As the reconstruction efficiency is lower for electrons than for muons, lepton mismatch have a higher impact in the former channel. Additionally, the lower resolution of electron track reconstruction may also enhance the wrong lepton pairing.

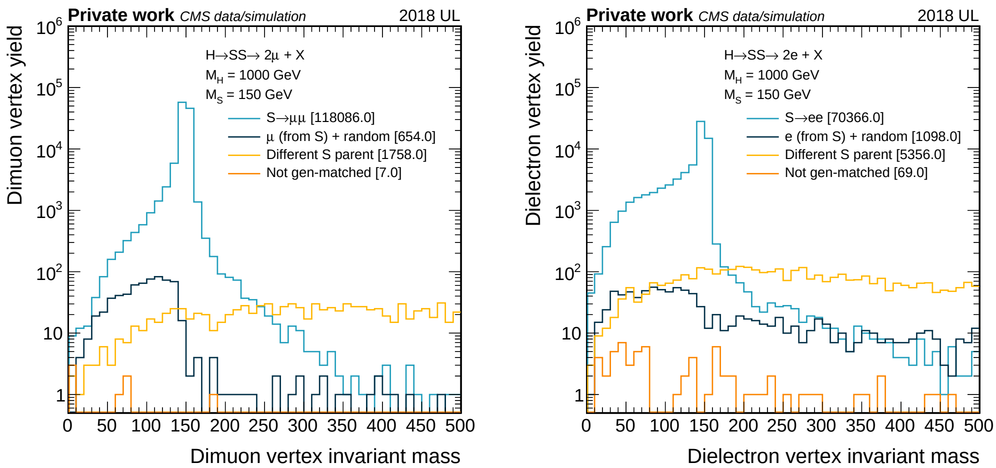


Figure 4.9: Invariant mass of the vertex reconstructed in simulated LL S decays of the H→SS model with masses $m_H = 1000$ GeV and $m_S = 150$ GeV. Vertices are splitted in four separate categories: 1) fitted from leptons matched to generated leptons that were produced in the same S boson decay; 2) fitted from leptons matched to generated leptons produced in two different S bosons decays; 3) fitted from one single lepton matched to a generated lepton produced in a certain S decay and from an unrelated lepton and 4) fitted from leptons that are not matched to generated leptons.

Chapter 5

Analysis strategy

The displaced lepton algorithms are applied in the analysis presented in this thesis, which is a search for long-lived particles (LLP) decaying to final states that contain displaced dilepton vertices. The analysis uses pp collisions collected by the CMS detector during the full LHC Run 2, in 2016, 2017 and 2018.

The search is intended to be as much agnostic as possible and could be interpreted in any model with long-lived particles decaying to lepton pairs within the tracker volume. In this case, one benchmark model is used to optimize the search strategy and interpret the final results. This is a simplified model which considers non-SM Higgs boson production and can be framed in the context of an extended Higgs sector. The exotic Higgs boson decays to a pair of long-lived spinless bosons S with neutral electric charge. After a certain time, driven by their lifetimes $c\tau_S$, both S bosons decay into lepton-antilepton pairs, produced in displaced dilepton vertices:

$$\mathrm{H} \rightarrow \mathrm{SS} \rightarrow 2\ell^{+}2\ell^{-}$$

The feynman diagram of this process is shown in 5.1.

Two final states are considered, depending on whether the long-lived S decays into an electron-positron or a muon-antimuon pair. In the search, two distinct channels are defined to study each lepton type: dielectron channel and dimuon channel. They are optimized separately and are treated independently over the entire search.

Each lepton channel is initially defined by means of the trigger used to collect the events. For the muon channel, events are triggered by a set of dimuon trigger paths which are fed from muons reconstructed only in the muon system. For the electron channel, the events are collected by a set of diphoton triggers where a tracker track is not required. The only exception to this definition occurs in 2017 data, where the absence of an efficient displaced dimuon trigger limits the sensitivity to long-lived particle decays into muons. As a

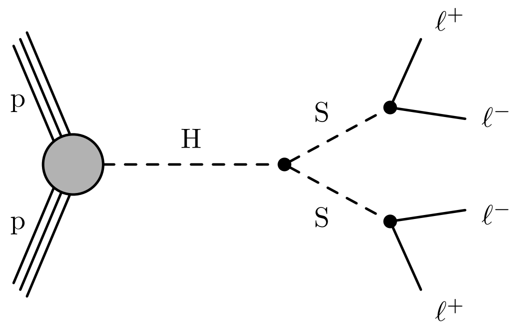


Figure 5.1: Benchmark model used as a reference for the analysis. It encompasses a BSM Higgs H decaying into two long-lived spinless neutral bosons S, which further decay into a pair of leptons each.

consequence, only dielectron decays are considered in 2017. The luminosity collected by the set of triggers in the full analysis sum up to 96 fb^{-1} in the muon channel and 112 fb^{-1} in the electron channel.

Within each channel both electron and muon objects are reconstructed with the reconstructions presented in Chapter 4 with the objective of recovering as much efficiency as possible. Additionally, this analysis is optimized for the identification requirements presented in the same chapter. In this way, the quality of the leptons reconstructed by these dedicated displaced algorithms is ensured.

Once leptons are reconstructed, they are paired and fitted to common dilepton vertices (DV). These vertices, which are the main objects of the search, are used to reconstruct the decay vertex of the LLP and may be displaced provided that its lifetime is long enough. To discriminate potential signal vertices from those produced in SM background processes, a dedicated selection is optimized. This selection is used to select potential long-lived candidates in the events that are going to define the Signal Region. It includes kinematic cuts, isolation, vertex quality, charge and some angular variables among others. The details are presented in Section 7.

The most powerful discriminants to identify the long-lived signal are the geometrical properties of the vertex. In this search, several distance variables are employed to determine how much displaced they are. These were already introduced in Section 3.4, and are:

- Transverse impact parameter $|d_0|$ and its significance $|d_0|/\sigma_d$: It corresponds with the measured $|d_0|$ and $|d_0|/\sigma_d$ values of the lepton with minimum $|d_0|/\sigma_d$ of the pair.
- Vertex transverse displacement L_{xy} and its significance L_{xy}/σ_L: Defined as the transverse distance between the PV of the collision and the position of the DV that is reconstructed.

Events are required to have at least one dilepton vertex passing the baseline selection, being this vertex fitted from muons for dimuon channel events and fitted from electrons for dielectron channel ones. If there are more than one vertex fulfilling the full set of criteria, the one with the largest $|d_0|/\sigma_d$ is chosen as the long-lived candidate of the event. For this reason, this analysis would also be expected to be sensitive to LLPs produced alone.

Since the properties of the LLP decay are strongly dependent on the model parameters, such as the mass of the particles involved, the boosting and of course, the lifetime, the Signal Region is divided into several sub-regions to increase the sensitivity. Although only one vertex is selected as the LLP candidate, the number of fitted vertices is used to define the splitting, together with its measured invariant mass, $|d_0|/\sigma_d$ and L_{xy}.

The background contribution in the Signal Region is carefully studied. While there are some processes that may introduce displaced background vertices, most of SM phenomena are prompt, so this search is relatively background free. Even if highly reduced, some instrumental background coming from resolution or misreconstruction effects is expected to arise at high displacement. Therefore, this analysis considers two different background sources: Truly displaced SM processes and instrumental background with virtual displacement due to resolution effects. The former is totally rejected by means of the discrimination done when applying the baseline selection. The remaining instrumental background cannot be completely removed from the Signal Region and its contribution is estimated entirely from data contained in an orthogonal control region. The background estimation method of the analysis profits from the behavior of an angular variable called collinearity that measures how collinear are the transverse momentum of the dilepton system and the displacement of the selected DV. Dedicated tests and validation in control regions are provided to assure the validity of the background estimation method. In addition, other dedicated studies are presented to evaluate the impact of the displaced background processes.

The results of the search are obtained by measuring the final signal and background yields in the different signal sub-regions. Upper limits on the production cross-sections times the branching fraction $\sigma \times \mathcal{B}(\mathrm{H} \to \mathrm{SS})$ are computed for the $\mathrm{H} \to \mathrm{SS}$ scenario in the context of three different interpretations by considering:

- Pure $\mathrm{S} \to \mu\mu$ decays where the branching fraction of the long-lived particle is assumed to be $\mathcal{B}(\mathrm{S} \to \mu\mu) = 100\%$.
- Pure $\mathrm{S} \to ee$ decays where the branching fraction of the long-lived particle is assumed to be $\mathcal{B}(\mathrm{S} \to ee) = 100\%$.
- Democratic $\mathrm{S} \to \mu\mu$ and $\mathrm{S} \to ee$ decays with a branching fraction of $\mathcal{B}(\mathrm{S} \to \mu\mu) = \mathcal{B}(\mathrm{S} \to ee) = 50\%$.

Chapter 6

Datasets, samples and triggers

6.1. Datasets

The datasets used in this analysis are the di-leptonic streams of 13 TeV p-p collision data recorded by CMS at a bunch spacing of 25 ns. The datasets analyzed corresponds to the 2016, 2017 and 2018 years of the Run 2 of the LHC.

The `DoubleMuon` primary dataset is used in the $\mu\mu$ channel. For the ee channel, `DoubleEG` dataset is used to analyze collisions collected in 2016 and 2017 while `EGamma` dataset is used for 2018 data. All events used for the final event counts were made sure to be triggered by the corresponding di-leptonic triggers chosen for each channel. The trigger path defined by these triggers will be widely described later in Section 6.4

The analyzed datasets contain an integrated luminosity of 35.9 fb^{-1} in 2016, 41.5 fb^{-1} in 2017 and 59.8 fb^{-1} in 2018 data-taking period. A fraction of the 2016 dataset was affected by the so-called APV saturation effect leading to a hit efficiency loss in the tracker. Due to this effect only 16.2 fb^{-1} from 2016 have been considered for the ee channel of this analysis. More details about this are provided later. On the other hand, the 2017 dataset is excluded from the $\mu\mu$ channel since the triggers used in this analysis were not available for that year and sensitivity could not be recovered at high displacement by other triggers. Finally, about 5 fb^{-1} are excluded for the ee channel in 2018 since the main HLT path was not present. Integrated luminosity totals to 95.6 fb^{-1} in the $\mu\mu$ channel and 112.4 fb^{-1} in the ee channel.

Data analyzed in this search were reconstructed in the Ultra Legacy (UL) reprocessing campaigns `21Feb2020`, `09Aug2019` and `12Nov2019` in 2016, 2017 and 2018 respectively with software version `CMSSW_10_6_20`. Different data formats are used for the dielectron and dimuon channels. Data events with electrons are taken from the MiniAOD samples. For the muon channel, `AOD` was used in order to utilize the Displaced Global Muon (dGlobal) object, not available in the `MiniAOD` format, nor in further processings of the data. Monte Carlo samples were all produced starting from the corresponding `AOD` samples, with the exception of 2017 ones where dimuon channel is not analyzed and `MiniAOD` format is used instead.

The list of the analyzed collision samples is given below in Table 6.1.

6.2. Background Monte Carlo simulation

The background of this analysis originates from SM processes with one or two leptons reconstructed to have a large impact parameter as a consequence of instrumental or resolution effects. The dominant background in this analysis consists of Drell-Yan production of dileptons, in the three decay modes e^+e^-, $\mu^+\mu^-$ and $\tau^+\tau^-$. Other processes contributing to the dilepton final state such as $t\bar{t}$ production and diboson (i.e. WW, WZ, ZZ) processes are considered. The W+jets production is also considered to take into account cases where one jet is misidentified and reconstructed as a lepton. Finally, a contribution from QCD processes is also expected to account for pion misidentification and leptons emerging from jet hadronization. Monte Carlo simulation is used to study the Drell-Yan, $t\bar{t}$, diboson and W+jets contributions. As W+jets and QCD simulation do not provide enough statistics to proper study the instrumental effects they are studied in two data control regions.

This analysis estimates the background contribution in the signal regions entirely from data with a data-driven background estimation method as described in section 8.1. However, Data-MC comparisons are performed in control regions to understand the background behavior and guide the search. In addition closure tests of background prediction are performed with simulation to validate the background estimation method. CMS central Monte Carlo samples are used in these two cases.

Two different samples are used for Drell-Yan. These are not exclusive but different parts of the analysis use one or the other depending on the nature of the studies that are done. One sample is produced at NLO with Madgraph (FXFX) [96] and Pythia 8 [97], which is proved to reproduce better the distributions observed in data and therefore is the preferred one to do the data vs Monte Carlo comparisons of the analysis. The other is produced at LO with Pythia 8 and Madgraph and it is used for the rest of the studies. The $t\bar{t}$ sample is produced with powheg [98] and Pythia 8 and diboson WW, WZ and ZZ samples with Pythia 8 only. Tune CP5 is set for sample generation in the three years. The campaigns to produce the background Monte Carlo samples are the ones corresponding to the Ultra Legacy reprocessing `RunIISummer20UL16`, `RunIISummer20UL17` and `RunIISummer20UL18` for 2016, 2017 and 2018. Table 6.2 list the most important SM Monte-Carlo samples used in this analysis in order

2016		
Dataset name	Run-range	L (fb^{-1})
/DoubleEG/Run2016B-21Feb2020_ver1_UL2016_HIPM-v1/MINIAOD /DoubleEG/Run2016B-21Feb2020_ver2_UL2016_HIPM-v1/MINIAOD	273150–275376	5.75
/DoubleEG/Run2016C-21Feb2020_UL2016_HIPM-v1/MINIAOD	275656–276283	2.57
/DoubleEG/Run2016D-21Feb2020_UL2016_HIPM-v1/MINIAOD	276315–276811	4.24
/DoubleEG/Run2016E-21Feb2020_UL2016_HIPM-v1/MINIAOD	276831–277420	4.02
/DoubleEG/Run2016F-21Feb2020_UL2016_HIPM-v1/MINIAOD /DoubleEG/Run2016F-21Feb2020_UL2016-v1/MINIAOD	277932–278808	3.10
/DoubleEG/Run2016G-21Feb2020_UL2016-v1/MINIAOD	278820–280385	7.58
/DoubleEG/Run2016H-21Feb2020_UL2016-v1/MINIAOD	281613–284044	8.65
/DoubleMuon/Run2016B-21Feb2020_ver1_UL2016_HIPM-v1/AOD /DoubleMuon/Run2016B-21Feb2020_ver2_UL2016_HIPM-v2/AOD	273150–275376	5.75
/DoubleMuon/Run2016C-21Feb2020_UL2016_HIPM-v1/AOD	275656–276283	2.57
/DoubleMuon/Run2016D-21Feb2020_UL2016_HIPM-v1/AOD	276315–276811	4.24
/DoubleMuon/Run2016E-21Feb2020_UL2016_HIPM-v1/AOD	276831–277420	4.02
/DoubleMuon/Run2016F-21Feb2020_UL2016_HIPM-v1/AOD /DoubleMuon/Run2016F-21Feb2020_UL2016-v1/AOD	277932–278808	3.10
/DoubleMuon/Run2016G-21Feb2020_UL2016-v1/AOD	278820–280385	7.58
/DoubleMuon/Run2016H-21Feb2020_UL2016-v1/AOD	281613–284044	8.65

2017		
Dataset name	Run-range	L (fb^{-1})
/DoubleEG/Run2017B-09Aug2019_UL2017-v1/MINIAOD	297047–299329	4.79
/DoubleEG/Run2017C-09Aug2019_UL2017-v1/MINIAOD	299368–302029	9.63
/DoubleEG/Run2017D-09Aug2019_UL2017-v1/MINIAOD	302030–302663	4.25
/DoubleEG/Run2017E-09Aug2019_UL2017-v1/MINIAOD	303818–304797	9.31
/DoubleEG/Run2017F-09Aug2019_UL2017-v1/MINIAOD	305040–306460	13.54

2018		
Dataset name	Run-range	L (fb^{-1})
/EGamma/Run2018A-12Nov2019_UL2018-v2/MINIAOD	315974–316995	8.7
/EGamma/Run2018B-12Nov2019_UL2018-v2/MINIAOD	317080–319310	7.06
/EGamma/Run2018C-12Nov2019_UL2018-v2/MINIAOD	319337–320065	6.90
/EGamma/Run2018D-12Nov2019_UL2018-v8/MINIAOD	320413–325175	31.74
/DoubleMuon/Run2018A-12Nov2019_UL2018-v2/AOD	315257–316995	14.03
/DoubleMuon/Run2018B-12Nov2019_UL2018-v2/AOD	317080–319310	7.06
/DoubleMuon/Run2018C-12Nov2019_UL2018-v2/AOD	319337–320065	6.90
/DoubleMuon/Run2018D-12Nov2019_UL2018-v3/AOD	320500–325175	31.74

Table 6.1: List of 2016, 2017 and 2018 collision datasets analyzed.

to design the signal region definitions.

6.3. Long-lived signal Monte Carlo simulation

Samples of the BSM Higgs model are produced centrally. The BSM Higgs decay process is simulated with Powheg and Pythia 8 with the same tune definitions than the background Monte Carlo (`RunIISummer20UL*` campaigns). Each simulated Higgs boson H decays into two long-lived spinless S bosons that further decays into lepton pairs. In these samples, the branching fraction of the H particles is fixed to $\mathcal{B}(\mathrm{H} \rightarrow \mathrm{SS}) = 100\%$. The branching fraction of the S bosons is set to 50% into electron decays and the other 50% into muon decays[1]. The lifetime of the S bosons is defined by setting their decay width to cover proper decay lengths (particle's reference frame) from 1 to 10^4 mm. All mass points and lifetimes considered in this thesis are summarized in Table 6.3.

6.4. Triggers

In order to capture the displaced dilepton signature, this analysis uses triggers with objects that are as independent of the displacement as possible. These objects feeding the trigger paths are not required to emerge from the primary vertex and their vertex constraints have been removed for their p_T computation. The muon trigger objects are reconstructed in the muon chambers using dedicated algorithms that reduce the vertex constraints, making them sensitive to the presence of muons produced outside the tracker. Displaced electrons, on the other hand, are expected to leave an energy deposit in the ECAL provided that they are produced within the tracker volume. Consequently, the electrons of this search are triggered using superclusters in the ECAL i.e. photon objects reconstructed at HLT level.

The trigger paths which are made from these objects and that are used in this search are listed in Table 6.4. Detailed information about these triggers paths and further discussion about its use in this search is provided over this section.

6.4.1. Trigger path for dimuon channel

In 2016 data and simulation, the events of the dimuon channel are required to pass one single path which makes use of the *Refitted StandAlone muon* (rSA) algorithm[2] to reconstruct the HLT muon objets. These rSA HLT muons are reconstructed by combining pairs of DT and/or CSC segments to make seeds

[1] In order to make interpretations with other scenarios e.g. $\mathcal{B}(\mathrm{S} \rightarrow \mu\mu) = 100\%$ or $\mathcal{B}(\mathrm{S} \rightarrow]]) = 100\%$, the corresponding subset of events is taken to match the desired branching ratio.

[2] This muon reconstruction algorithm is briefly described in Section 4.1

Background Monte Carlo samples
2016 (preVFP)
/DYJetsToLL_M-50_TuneCP5_13TeV-madgraphMLM-pythia8/RunIISummer20UL16RECOAPV/
/DYJetsToLL_M-50_TuneCP5_13TeV-amcatnloFXFX-pythia8/RunIISummer20UL16RECOAPV/
/DYJetsToLL_M-10to50_TuneCP5_13TeV-madgraphMLM-pythia8/RunIISummer20UL16RECOAPV/
/TTTo2L2Nu_TuneCP5_13TeV-powheg-pythia8/RunIISummer20UL16RECOAPV/
/WW_TuneCP5_13TeV-pythia8/RunIISummer20UL16RECOAPV/
/WZ_TuneCP5_13TeV-pythia8/RunIISummer20UL16RECOAPV/
/ZZ_TuneCP5_13TeV-pythia8/RunIISummer20UL16RECOAPV/
/WJetsToLNu_TuneCP5_13TeV-madgraphMLM-pythia8/RunIISummer20UL16RECOAPV/
2016 (postVFP)
/DYJetsToLL_M-50_TuneCP5_13TeV-madgraphMLM-pythia8/RunIISummer20UL16RECO/
/DYJetsToLL_M-50_TuneCP5_13TeV-amcatnloFXFX-pythia8/RunIISummer20UL16RECO/
/DYJetsToLL_M-10to50_TuneCP5_13TeV-madgraphMLM-pythia8/RunIISummer20UL16RECO/
/TTTo2L2Nu_TuneCP5_13TeV-powheg-pythia8/RunIISummer20UL16RECO/
/WW_TuneCP5_13TeV-pythia8/RunIISummer20UL16RECO/
/WZ_TuneCP5_13TeV-pythia8/RunIISummer20UL16RECO/
/ZZ_TuneCP5_13TeV-pythia8/RunIISummer20UL16RECO/
/WJetsToLNu_TuneCP5_13TeV-madgraphMLM-pythia8/RunIISummer20UL16RECO/
2017
/DYJetsToLL_M-50_TuneCP5_13TeV-madgraphMLM-pythia8/RunIISummer20UL17MiniAODv2/
/DYJetsToLL_M-50_TuneCP5_13TeV-amcatnloFXFX-pythia8/RunIISummer20UL17MiniAODv2/
/DYJetsToLL_M-10to50_TuneCP5_13TeV-madgraphMLM-pythia8/RunIISummer20UL17MiniAODv2/
/TTTo2L2Nu_TuneCP5_13TeV-powheg-pythia8/RunIISummer20UL17MiniAODv2/
/WW_TuneCP5_13TeV-pythia8/RunIISummer20UL17MiniAODv2/
/WZ_TuneCP5_13TeV-pythia8/RunIISummer20UL17MiniAODv2/
/ZZ_TuneCP5_13TeV-pythia8/RunIISummer20UL17MiniAODv2/
/WJetsToLNu_TuneCP5_13TeV-madgraphMLM-pythia8/RunIISummer20UL17MiniAODv2/
2018
/DYJetsToLL_M-50_TuneCP5_13TeV-madgraphMLM-pythia8/RunIISummer20UL18RECO/
/DYJetsToLL_M-50_TuneCP5_13TeV-amcatnloFXFX-pythia8/RunIISummer20UL18RECO/
/DYJetsToLL_M-10to50_TuneCP5_13TeV-madgraphMLM-pythia8/RunIISummer20UL18RECO/
/TTTo2L2Nu_TuneCP5_13TeV-powheg-pythia8/RunIISummer20UL18RECO/
/WW_TuneCP5_13TeV-pythia8/RunIISummer20UL18RECO/
/WZ_TuneCP5_13TeV-pythia8/RunIISummer20UL18RECO/
/ZZ_TuneCP5_13TeV-pythia8/RunIISummer20UL18RECO/
/WJetsToLNu_TuneCP5_13TeV-madgraphMLM-pythia8/RunIISummer20UL18RECO/

Table 6.2: Simulated SM datasets analyzed in the course of the analysis. Both preVFP and postVFP 2016 samples are of the `AOD` data format from the `RunIISummer20UL16` campaign. Labels `RunIISummer20UL16RECOAPV` and `RunIISummer20UL16RECO` are abbreviations for `RunIISummer20UL16RECOAPV-106X_mcRun2_asymptotic_preVFP_v8-v*` and `RunIISummer20UL16RECO-106X_mcRun2_asymptotic_v13-v*`. 2017 samples are of the `MiniAOD` data format from the `RunIISummer20UL17` campaign. `RunIISummer20UL17MiniAODv2` is an abbreviation for `RunIISummer20UL17MiniAODv2-106X_mc2017_realistic_v9-v*`. 2018 samples are of the `AOD` data format from the `RunIISummer20UL18` campaign. `RunIISummer20UL18RECO` is an abbreviation for `RunIISummer20UL18RECO-106X_upgrade2018_realistic_v11_L1v1-v*`

BSM Higgs signal Monte Carlo samples			
m_H (GeV/c^2)	m_S (GeV/c^2)	$c\tau$ (mm)	σ (pb)
125	50	1, 10, 10^2, 10^3, 10^4	21.460
300	50		4.938
400	150		5.040
500	50, 150		2.588
600	50, 150		1.1785
800	50, 250, 350		0.3025
1000	150, 250, 350		0.10770

Table 6.3: Mass points studied for the H $\rightarrow$ SS model. Production cross-sections σ are taken from the production gridpacks for H generation and correspond with the estimation provided by powheg.

whose direction is required to point from the beam spot. However, the fit to get the p_T estimate of the rSA muons removes the beam spot constraint to improve the resolution of displaced muons that are not expected to be originated near the interaction point. Muons of this path are required to have two reconstructed segments in two chambers of the muon system and be measured within $|\eta| < 2$. This HLT path requires to have at least two rSA HLT muons with $p_T > 28$ GeV. In addition, their 3D apperture angle between must be lower than $\alpha < 2.5$ radians and their invariant mass must be over $m_{\mu\mu} > 10$ GeV.

The equivalent HLT path failed to enter into the 2017 trigger menu and no displaced muon trigger with HLT rSA muons is available in 2017. Dimuon triggers of the 2017 menu, which use muons reconstructed both in the tracker and in the chambers with higher p_T thresholds, have proven not to be very efficient for this search. For this reason, the dimuon channel is excluded from 2017 analysis.

For 2018 data, events are required to pass the logical OR of two different paths. The first one employs the same rSA HLT reconstruction as the 2016 path, with at least two segments in two muon chambers and $|\eta| < 2$. The muon p_T thresholds are lowered to 23 GeV and the muon apperture angle and invariant mass requirements are removed to gain sensitivity. The second one, labelled with CosmicSeed, makes use of the *Displaced StandAlone muon* (dSA) reconstruction. These dSA HLT muons are reconstructed from cosmic like seeds which accept all possible bending directions of the segment combinations to improve the reconstruction efficiency of displaced muons. The beam spot constraint is also removed in the p_T fit so these muons do not have vertex constraints in the pattern recognition nor in the track fit to estimate the momentum. This HLT path requires at least two dSA HLT muons, with two segments in two chambers, reconstructed within $|\eta| < 2$ and no angle nor invariant mass requirements.

The HLT path `HLT_L2DoubleMu28_NoVertex_2Cha_Angle2p5_Mass10` in 2016 was

Trigger paths of the muon channel
2016
HLT_L2DoubleMu28_NoVertex_2Cha_Angle2p5_Mass10
2018
HLT_DoubleL2Mu23NoVtx_2Cha
HLT_DoubleL2Mu23NoVtx_2Cha_CosmicSeed

Trigger paths of the electron channel
2016
HLT_Photon42_R9Id85_OR_CaloId24b40e_Iso50T80L_Photon25_AND_HE10_R9Id65_Eta2_Mass15
2017
HLT_Diphoton30_22_R9Id_OR_IsoCaloId_AND_HE_R9Id_Mass90
HLT_DoublePhoton70
2018
HLT_Diphoton30_18_R9IdL_AND_HE_AND_IsoCaloId_NoPixelVeto
HLT_Diphoton30_22_R9Id_OR_IsoCaloId_AND_HE_R9Id_Mass90
HLT_DoublePhoton70

Table 6.4: List of triggers corresponding to the analysis of the muon and electron channels

seeded by the L1 seeds L1_DoubleMu11_4 and L1_DoubleMu12_5. While the latter was kept unprescaled during all the data taking, the former was prescaled in about 4% of the lumi sections. This effect was studied in other CMS analysis [80] concluding that the effect was negligible and in any case much smaller than the systematic uncertainty assigned to the trigger scale factors. On the other hand, the discrepancies were more important for the 2018 year, where the main analysis trigger was seeded by a logical OR of L1_DoubleMu_12_5, L1_DoubleMu_15_5, L1DoubleMu_15_7, and L1_TripleMu_5_3_3. There were differences in some of these L1 seeds in the signal simulation, affecting especially large lifetime signals and very low momentum muons. This effect has been checked to be negligible for this analysis where the minimum offline p_T muon threshold is $p_T > 25$ GeV. This test was done by checking that the higher momenta L1 seed was fully efficient for the analysis offline selection.

6.4.2. Trigger path for dielectron channel

In order to gain sensitivity to reconstructed displaced electrons the events of the dielectron channel are triggered with diphoton triggers. Electrons are expected to leave an energy cluster in the ECAL as long as they are produced within the tracker volume. This makes diphoton triggers a good handle to target displaced electrons produced in LLP decays with short to medium lifetimes as the ones targeted in this search.

In the three years the search either uses a single HLT path or profit from the logical OR combination of different ones with the objective of recovering as much efficiency as possible at high displacement. The diphoton HLT paths are also summarized in Table 6.4.

2016 data and simulation make use a double photon path which is only available for this year and that requires a leading (subleading) photon with energy $E_T > 42(25)$ GeV. General requirements are imposed on R_9, calorimeter identification, isolation and H/E. Both reconstructed photons are restricted to have $|\eta| < 2$ and the diphoton mass must be over $m_{\gamma\gamma} > 15$. This path was seeded by the OR combination of a large set of L1 seeds including single EGamma seeds, double EGamma seeds, and double EGamma seeds with eta restriction. The pattern of L1 seeds used for the HLT path and/or the corresponding prescales changed frequently along the 2016 run. The MC simulation samples were produced with a particular choice of L1 seeds corresponding to a particular L1 menu. Large differences among the seeds used in data and in MC could introduce a bias in the data/MC agreement. In order to check the impact of this effect, the studies profit from the fact that there were three seeds, `L1_DoubleEG_23_10`, `L1_DoubleEG_24_17`, and `L1_SingleEG40`, that were present without prescaling the whole 2016 data taking period. The efficiency of these L1 seeds was estimated for events passing the analysis offline selection (electrons with $p_T > 45$ GeV and $p_T > 30$ GeV), for MC and for data in periods where L1 seeds with the lowest p_T thresholds were present. It was observed that this efficiency is always higher than 99% with differences smaller than 1% for the different cases. This effect was concluded to be negligible for the analysis.

The `HLT_DiphotonX_Y_R9Id_OR_IsoCaloId_AND_HE_R9Id_Mass90` path is used in 2017 and 2018. Both photons are required to pass basic requirements on R_9, calorimeter identification, isolation and H/E. The invariant mass of the photons is required to be over $m_{\gamma\gamma} > 90$ GeV to reduce the rate which makes this path sensitive to high m_{LLP} resonances. As it will be seen in Section 7, the offline selection is also consistent with this mass requirement.

The HLT path `HLT_Diphoton30_18_R9IdL_AND_HE_AND_IsoCaloId_NoPixelVeto` that is used in 2018 is an improvement of the previous paths. The filters are reorganized in such a way that the rate is controlled and the invariant mass requirement and PixelVeto requirements are removed. This trigger increase the sensitivity to low mass LLP resonances and small lifetimes with respect to the 2016 and 2017 versions. This trigger was not active in the first $5\ \text{fb}^{-1}$ of the 2018 data taking. For this reason, these runs were excluded from the search in the electron channel.

Finally the `HLT_DoublePhoton70` path is added to 2017 and 2018, which requires at least two photons with $E_T > 70$ GeV. This path is expected to recover some efficiency at low m_{LLP} mass with high p_T leptons e.g. Long-Lived S bosons produced from high mass Higgs boson decays.

6.4.3. Trigger sensitivity to displaced leptons

The sensitivity of the presented trigger paths is measured for several mass points and lifetimes of the H $\rightarrow$ SS model simulated sample. It is done separately for events triggered with dimuon and diphoton triggers. The figure of merit is the acceptance of the trigger paths, defined as the ratio of the events that pass the path over the total number of generated events:

$$\text{Trigger acceptance} = \frac{\text{Number of events passing full HLT path}}{\text{Number of generated events}} \tag{6.1}$$

Since the simulated LLP are split into dimuon and dielectron decays in equal proportion ($\mathcal{B}(\mathrm{S} \rightarrow ee) = \mathcal{B}(\mathrm{S} \rightarrow \mu\mu) = 50\%$) the trigger acceptance will depend on the number of the selected flavor leptons in the final state. To measure the sensitivity of the trigger for every case considered, the trigger acceptance defined in Eq. 6.1 by splitting the number of generated events in three bins:

- SS $\rightarrow 4\ell$: Events with both S particles decaying to the same lepton flavor
- SS $\rightarrow 2e + 2\mu$: Events with each S particle decaying to one different lepton flavor
- SS $\rightarrow 2\ell + X$: Combination of final states with at least one of the S particles decaying to the selected lepton flavor.

The full trigger path acceptance, i.e. the OR combination of selected paths for every year, is computed according to this definition as a function of the Higgs mass m_{H}, the LL S scalar m_{S} and the lifetime of the particle $c\tau_{\mathrm{S}}$ (which is expressed in length units).

The full dimuon trigger path acceptance for 2016 and 2018 is shown in Figure 6.1. In general, low acceptance is found for low masses that gradually increases as both the m_{S} and m_{H} masses increase. This dependence with the mass of the particles is driven by the p_T thresholds. The efficiency as a function of the lifetime remains stable while $c\tau_{\mathrm{S}} \leq 10^2$ mm, starts to decrease at $c\tau_{\mathrm{S}} = 10^3$ mm and is finally lost for lifetimes higher than $c\tau_{\mathrm{S}} \geq 10^4$ mm. There is a general increase in the sensitivity of dimuon triggers in 2018 with respect to 2016 for every mass point considered, which is due to the decrease of the p_T thresholds of the trigger paths included in the 2018 menu. In addition, the 2018 trigger path recovers some efficiency for displaced muons, which can be deduced from the increase in the trigger acceptance for simulated S bosons with $c\tau_{\mathrm{S}} \geq 10^3$ mm. This increase is due to the inclusion of the CosmicSeed path which enables muon seeds with multiple bendings to reconstruct the L2 dSA objects.

The full diphoton trigger path acceptance for 2016, 2017 and 2018 is shown in Figure 6.2. Like dimuon paths, lower efficiencies are found for low masses which increase as the m_{S} and m_{H} masses increase. A decrease of acceptance is measured in the $m_{\mathrm{H}} = 125$ GeV, $m_{\mathrm{S}} = 30$ GeV point in both 2016 and 2017

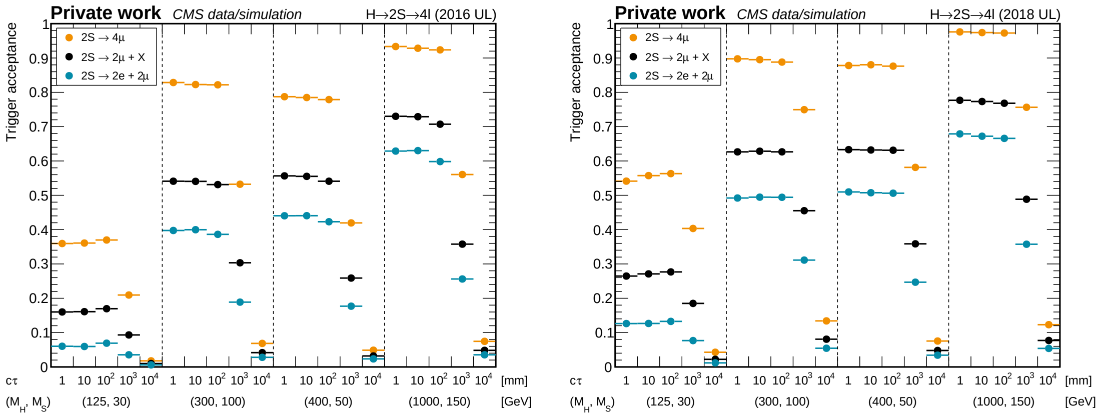


Figure 6.1: Trigger acceptance of the OR combination of dimuon trigger paths used in 2016 (left) and 2018 (right) computed in simulated H $\rightarrow$ SS decays. This acceptance in given as a function of the mass of the H boson, the LL S scalar and the simulated $c\tau_{\mathrm{S}}$. The acceptance is represented separately for simulated events with four (orange), two (blue) muons in the final state and the combination of both (black).

as the lifetime decreases for $c\tau_{\mathrm{S}} \leq 10^2$ mm which is not observed in 2018. This decrease is caused by the absence of low p_T paths in 2016 and 2017. At low mass regimes, specially when leptons have low momentum, the electrons are lost by the triggers with invariant mass requirements and/or with high p_T thresholds.

Lower efficiency is measured in 2017 for the $m_{\mathrm{H}} = 400$ GeV, $m_{\mathrm{S}} = 50$ GeV point with only one pair of electrons in the final state. The efficiency of this point is higher in 2016 and 2018 where trigger paths with softer or none $m_{\gamma\gamma}$ requirements are present.

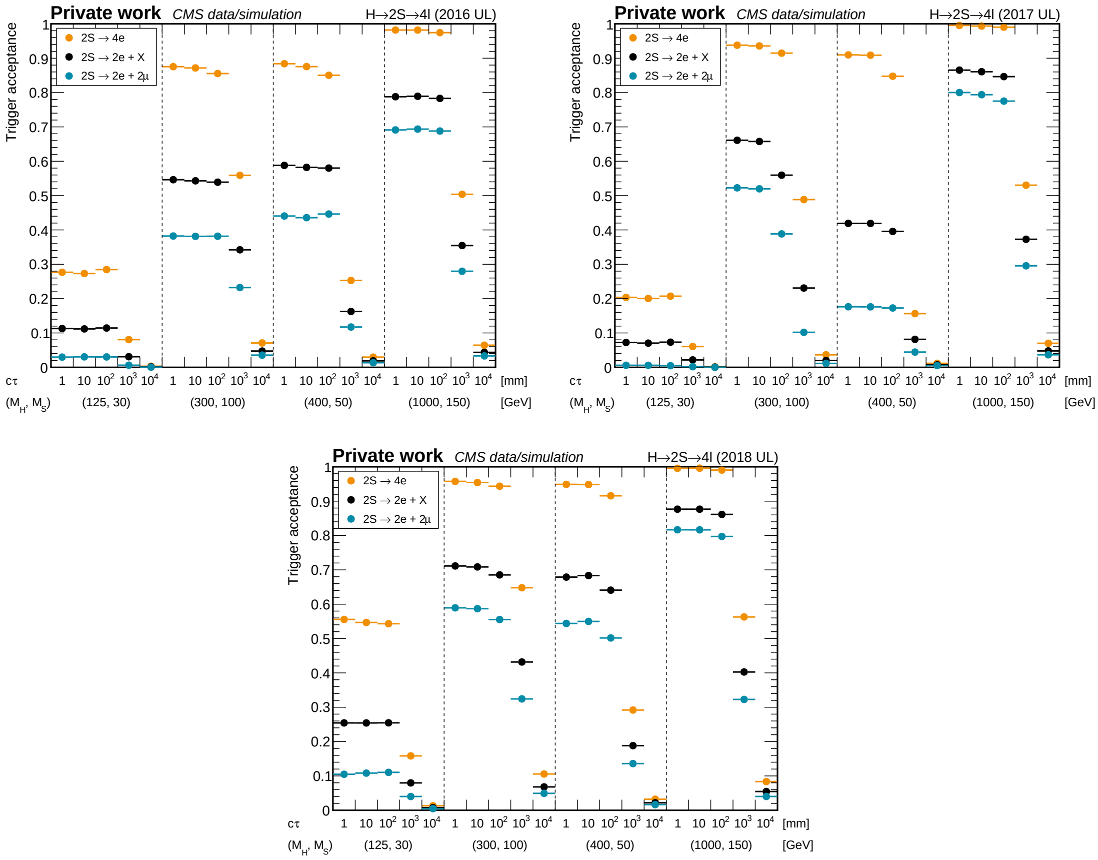


Figure 6.2: Trigger acceptance of the OR combination of diphoton trigger paths used in 2016 (left), 2017 (right) and 2018 (bottom) computed in simulated $H \to SS$. This acceptance in given as a function of the mass of the H boson, the LL S scalar and the simulated $c\tau_{\mathrm{S}}$. The acceptance is represented separately for simulated events with four (orange), two (blue) electrons in the final state and the combination of both (black).

Chapter 7

Event selection and signal regions

The dilepton vertices fitted from the tracks of the reconstructed leptons are the main objects of the search. A dedicated selection is optimized to eliminate background vertices and isolate a sample with selected vertices that can be identified with long-lived particle candidates decaying leptonically.

The following sub-sections present this selection criteria and how they were optimized. The distribution of the various variables comprising this selection are compared for the signal and the background. The background distributions shown in this chapter correspond to the background vertices predicted using the method described in the background section of this thesis, in Section 8.

The selected sample of long-lived particle candidates constitutes the Signal Region of the search. To enhance the sensitivity to different LLP decay scenarios, this Signal Region is further divided into several sub-regions, which are also defined at the end of this section.

7.1. Baseline event selection

In this analysis, the selection of dimuon and dielectron vertices is optimized separately. Events are splitted into two distinct channels: dimuon channel, when the LLP decays to muon pairs, and dielectron channel, when it decays to electron pairs. Events passing the muon trigger requirements are contained in the former while the ones firing the photon trigger requirements compound the latter as previously described in Section 6.4.

Events of both channels are required to contain a primary vertex PV fulfilling a set of standard quality requirements. The PV must be a non-fake vertex i. e. a genuine vertex obtained from the fit of a set of tracks. The number of degrees of freedom of the vertex must be ndof > 4, that is, it must have more than 4 associated tracks. In addition, the PV position must be placed inside the bunch crossing: its position can be displaced from the interaction point by no more than $\rho < 2$ cm in the direction transverse to the proton beam, and no more than $|z| < 25$ cm in the longitudinal direction.

Events are selected to be in the trigger p_T plateau by requiring the existence of offline reconstructed lepton tracks above the given p_T/E_T thresholds of each year. For muon events, it is required at least two dGlobal muon tracks having transverse momentum values of $p_T > 30$ GeV (25 GeV) if they were collected in 2016 (2018). The p_T requirement was relaxed from one year to the other due to the lower trigger p_T in the 2018 path. For electron events, the leading electron of the event is required to have transverse energy values over $E_T > 45$ GeV in 2016 and $E_T > 40$ GeV in 2017 and 2018. In the same way, the subleading electron must have transverse energy values over $E_T > 30$ GeV in 2016 and $E_T > 25$ GeV in 2017 and 2018. In this channel, the E_T is taken from the ECAL supercluster used to reconstruct the electron and no cuts are applied on electron track p_T for two distinct reasons. The first one is that the trigger used for this channel is a photon trigger where the E_T selection is applied in the reconstructed ECAL superclusters. The second is that the isolated tracks used to reconstruct the displaced electrons do not account for bremsstrahlung and radiative effects and the track p_T is often underestimated, leading to a significant decrease in electron efficiency for displaced decays as it was studied in Section 4.2.

7.2. Dilepton vertices selection

LLP candidates are reconstructed from lepton pairs fitted to common vertices following the procedure presented in Section 4.3. Events of the dimuon and dielectron channels must have at least one valid dimuon or dielectron vertex reconstructed, respectively. Several criteria are applied to select a potential long-lived candidate among the N_{DV} vertices that are fitted in the event. This LLP candidate is identified with the vertex with the largest $|d_0|/\sigma$ that fulfills all the requirements.

This section revisits the different studied requirements and compares the distributions for the long-lived candidates of several simulated displaced signals and the expected background. The background yields, unless specified, are obtained by following the background estimation method of the search, that will be later presented in Section 8.

Kinematic selection

No additional p_T or E_T requirements apart from those applied at preselection level (see Section 4.3) are required for the fitted vertices. The p_T spectra of the leading and subleading muon of the selected vertices are shown in Figures 7.1

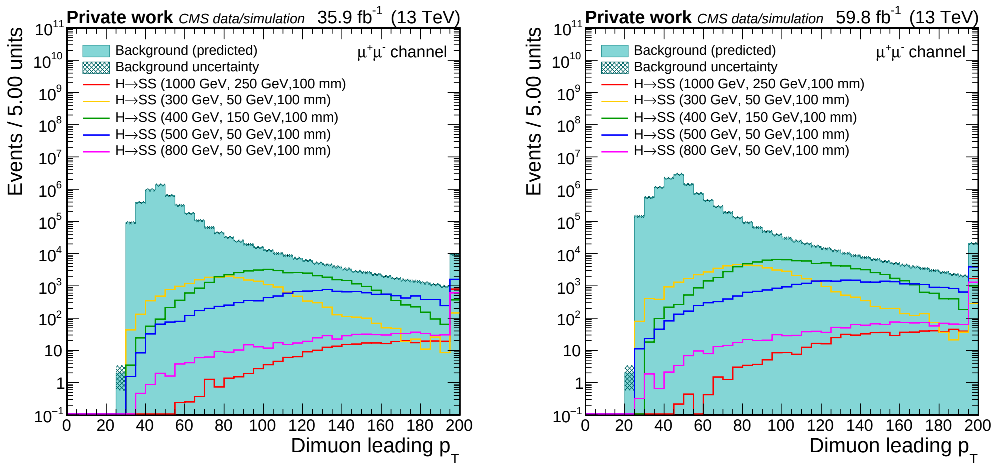


Figure 7.1: Leading muon p_T spectra in 2016 (left) and 2018 (right). All cuts of the baseline selection are applied.

and 7.2. The E_T spectra of the leading and subleading electron of the selected vertices are also shown in Figures 7.3 and 7.4. While the events are requested to have leptons on the plateau of the corresponding triggers, these are not required to be fitted to any vertex since the presence of a vertex is expected to be uncorrelated from the requirements applied on the leptons at HLT level.

To reject background events coming from low mass resonances, both dielectron and dimuon pairs are required to have reconstructed invariant mass values over $m_{\ell\ell} > 15$ GeV. For dielectron vertices in 2017 this requirement is increased to $m_{\ell\ell} > 100$ GeV due to the mass requirement present in the 2017 trigger which impacts the background estimation method. More details will be provided in Section 8.6.1. The distribution of invariant mass is shown in Figures 7.5 and 7.6 for dimuon and dielectron vertices, respectively.

Only in the muon channel, dGlobal muons of the same pair are required to be separated by at least $\Delta R > 0.3$. A preliminary requirement $\Delta R > 0.1$ was already introduced before fitting the muon tracks because the dGlobal reconstruction used in this search was not processed by standard duplicate cleaning algorithms. This results not only in the reconstruction of spurious vertices fitted from the same muon track reconstructed twice, but also on the impediment of associating each muon with its corresponding partner. By requiring the muons to be found over this conservative $\Delta R > 0.3$ we reject spurious vertices fitted from dGlobal muon duplicates with bad quality that may have survived this requirement.

Vertex quality

The quality of the reconstructed displaced vertices must be high for genuine displaced decays. Specially, when the reconstructed vertex position is dis-

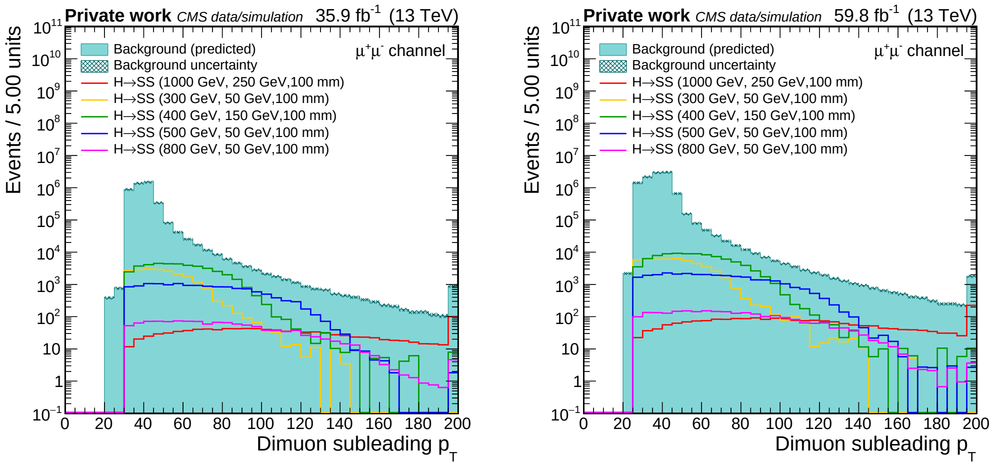


Figure 7.2: Subleading muon p_T spectra in 2016 (left) and 2018 (right). All cuts of the baseline selection are applied.

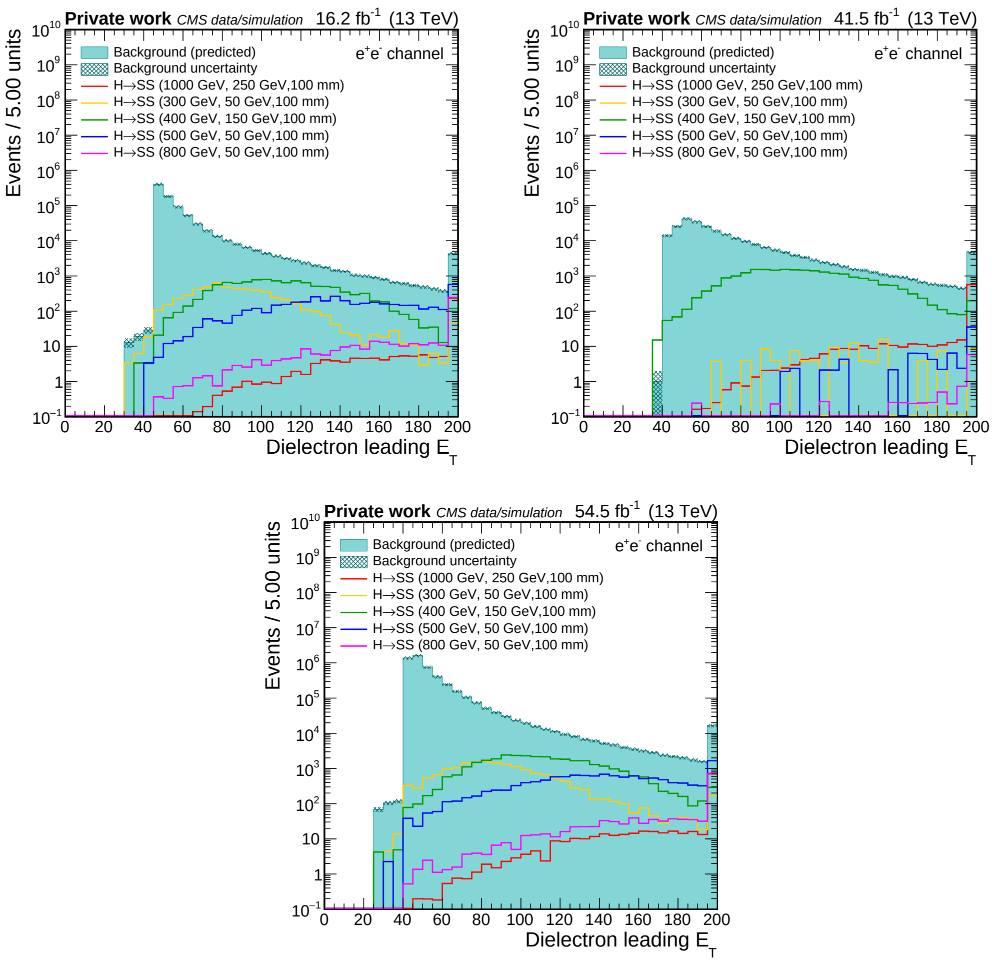


Figure 7.3: Leading electron E_T spectra in 2016 (top-left), 2017 (top-right) and 2018 (bottom). All cuts of the baseline selection are applied.

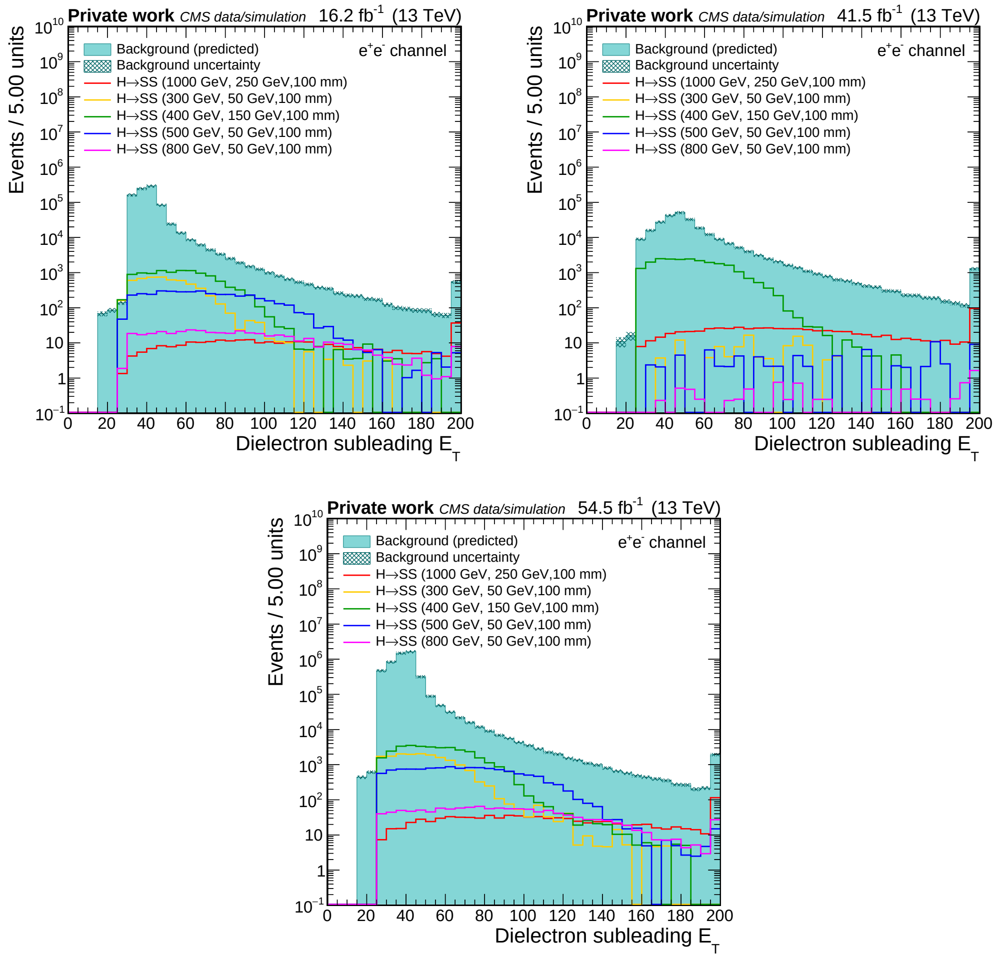


Figure 7.4: Subleading electron E_T spectra in 2016 (top-left), 2017 (top-right) and 2018 (bottom). All cuts of the baseline selection are applied.

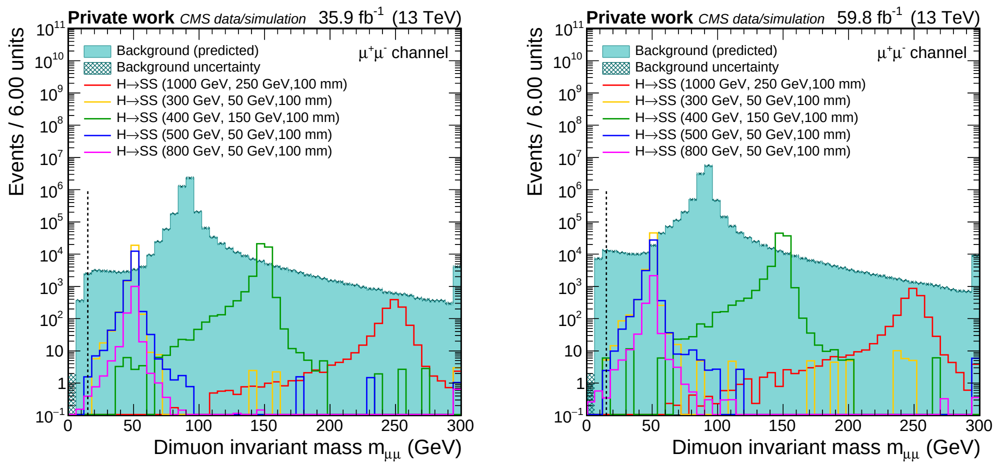


Figure 7.5: Invariant mass distribution of the dimuon candidates in 2016 (left) and 2018 (right). All cuts of the baseline selection are applied except for the invariant mass requirements.

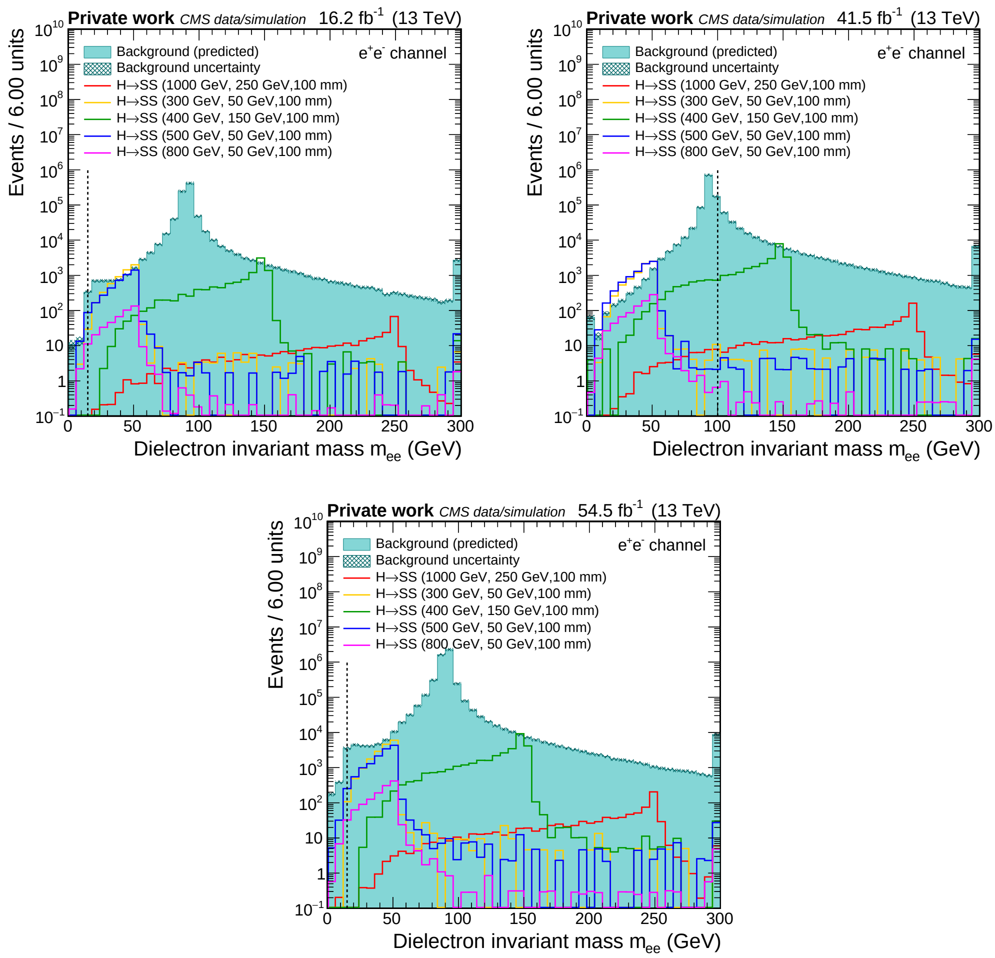


Figure 7.6: Invariant mass distribution of the dielectron candidates in 2016 (top-left), 2017 (top-right) and 2018 (bottom). All cuts of the baseline selection are applied except for the invariant mass requirements.

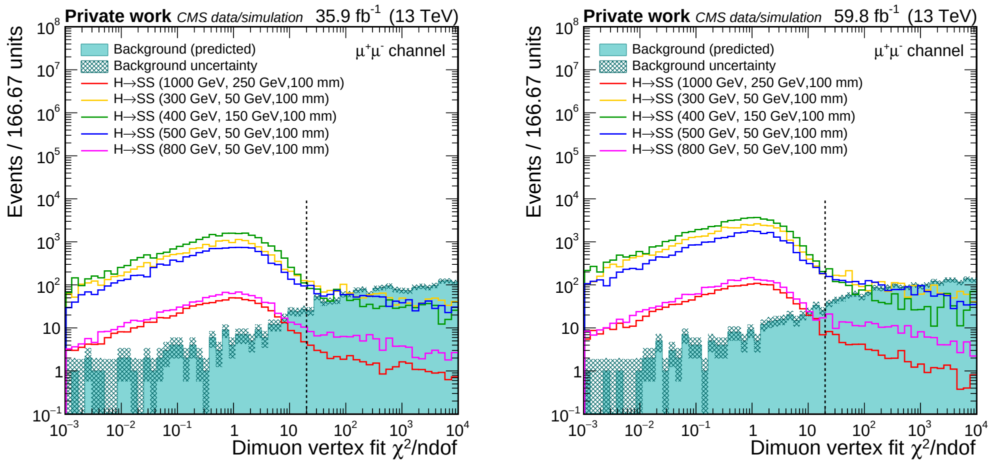


Figure 7.7: Distribution of the fit chi-squared per degree of freedom χ^2/ndof of dimuon vertices in 2016 (left) and 2018 (right). All cuts of the baseline selection are applied except for the χ^2/ndof one. An additional cut of $|L_{xy}| > 2.5$ is also applied.

placed from the interaction point, their vertex quality is expected to be better when compared with background prompt processes (where the displacement could be due to instrumental effects). Dilepton vertices of both lepton types are required to have a fit chi-squared per degree of freedom χ^2/ndof. This cut value was optimized by evaluating its signal efficiency and background rejection rate in dilepton events displaced by at least $|L_{xy}| > 2.5$. Keeping as many signal vertices as possible is prioritized and this cut value was chosen to retain at least 93% of the simulated signal events. Distributions of χ^2/ndof are shown for selected dimuon and dielectron vertices in Figures 7.7 and 7.8.

Cosmic muon rejection

Cosmic muons may be reconstructed as back-to-back displaced tracks found in the upper and lower hemispheres of the detector. There is a possibility that these tracks belonging to the same comic muon could be paired and fitted to a common displaced vertex becoming an important source of background. To reject these potential background vertices, a cut of $\cos(\alpha) > -0.8$ (-0.9) is applied only to the reconstructed dimuon candidates in 2016 (2018), where the α angle is the opening angle between the muon tracks. The distribution of this variable is shown in Figure 7.9 for vertices passing the whole selection except this cut.

The minimum value of -0.8 in 2016 was imposed by the dimuon trigger as the chosen path only accepts dimuons having $\alpha < 2.5$ rads. As a consequence, dimuon vertices which fall below $\cos(\alpha) > -0.8$ are strongly suppressed in 2016. This angle requirement was removed for the 2018 path and the vertex selection is loosened to -0.9 as it is proved to still reject the entire cosmic back-

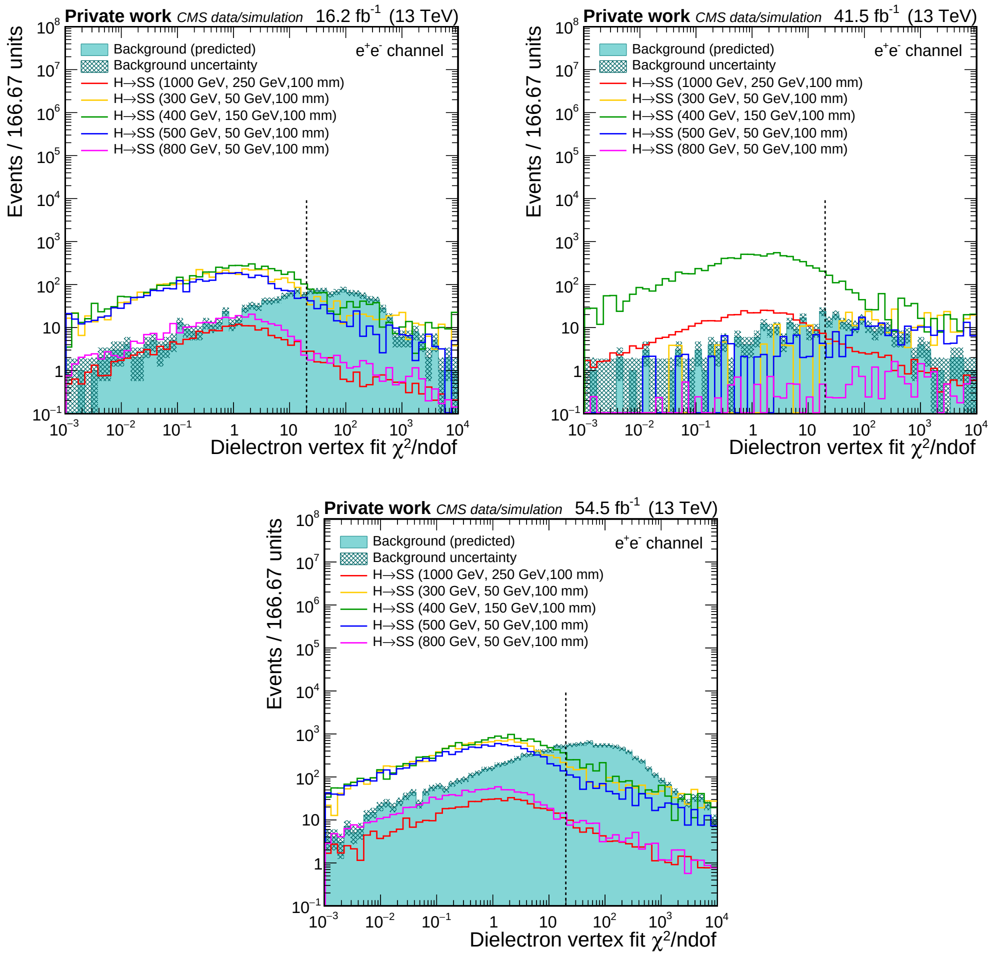


Figure 7.8: Distribution of the fit chi-squared per degree of freedom χ^2/ndof of dielectron vertices in 2016 (top-left), 2017 (top-right) and 2018 (bottom). All cuts of the baseline selection are applied except for the χ^2/ndof one. An additional cut of $|L_{xy}| > 2.5$ is also applied.

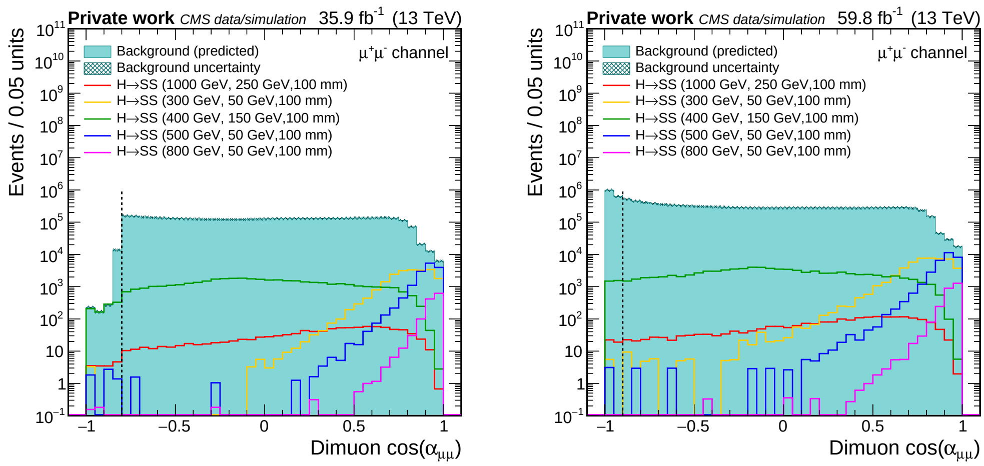


Figure 7.9: Cosine of the opening angle $\cos(\alpha)$ distribution of selected dimuon candidates in 2016 (left) and 2018 (right). All cuts of the baseline selection are applied except for the one in $\cos(\alpha)$.

ground. A deeper study of the impact of this cut in real cosmic muon data is presented in Section 8.8.1 where this conclusion is derived. The effect of the 2016 trigger can be seen in Figure 7.9 where the distribution of background vertices below this threshold is suppressed for 2016 with respect to 2018.

The loosen selection in 2018 with respect to 2016 aims to retain more efficiency in the less boosted signals with high mass long-lived resonances that may decay into back-to-back muons. This cut is expected to be sensitive to the opening angle of the muons and therefore, to some model parameters such as m_{LLP} or/and its mother particle mass. In order to check the dependence of this selection with the model the efficiency of the $\cos\alpha$ requirement in several mass points of the H→SS simulation was estimated. It was seen that most of dimuon vertices pass this selection, with efficiencies over 95% in every case studied, without having a clear dependence with the lifetime of the LLP.

Isolation

The main handle for rejecting electrons and muons produced in jets through the decay of heavy-flavour hadrons and other particles is isolation. Isolation is generally defined as the particle activity contained in a geometrical cone $\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2}$ surrounding each lepton. Evaluated relative to the p_T of the lepton, it is known as *relative isolation* and used to identify prompt leptons that come from the decay of massive particles. Leptons produced in the core of jets hadronization are rejected if they have relatively large isolation.

There are several definitions for the lepton isolation depending on which signals are included in the cone sum. These are computed by standard routines which run for each lepton reconstructed in the event. Unfortunately, as dGlobal muons are not included in the standard muon sequences, their isolation is not included in these either. For the analyses presented in this thesis, the value of their isolation is manually computed by considering all the PF candidates available in the `MiniAOD` format, which are reconstructed with the routines presented in Section 2.3.

The relative isolation definition used for dGlobal muons corresponds with the relative *Particle-Flow* (PF) isolation. This value is defined as the ratio between the PF isolation of each lepton divided by the p_T value of its track. The computation of the PF isolation is done by summing together the p_T of charged hadrons originating from the primary vertex with the energy sum of all neutral particles (hadrons and photons) within $0.03 < \Delta R < 0.3$. The minimum value $0.03 < \Delta R$ is fixed to avoid including the own lepton's track in the isolation computation. The contribution from pileup to the neutral particles is corrected by computing the sum of charged hadron deposits originating from pileup vertices, scaling it by a factor of 0.5. The corrected energy sum from neutral particles is computed by subtracting this scaled value from the neutral hadron and photon sums. Finally, the expression for the relative PF isolation is:

$$\text{rel. PF isolation} = \frac{1}{p_T^{\text{(lepton)}}}\left(\sum_{h^{\pm}} p_T + \max\left[0, \sum_{h^0,\gamma} p_T - 0.5 \sum_{\text{pileup}} p_T\right]\right) \quad (7.1)$$

The isolation also needs to be manually implemented for displaced electrons as these are not standard electron CMS objects. Unfortunately, the performance of the PF isolation defined above is highly reduced when applied to displaced electrons produced from a LLP decay. If the electron radiates one photon before hitting the ECAL there is a chance that this photon would enter inside the geometrical cone and its energy be counted in the PF isolation, overestimating its value. While this effect is accounted for standard CMS electrons, when using just isolated track objects it is possible that the matched supercluster fails to be removed from the isolation sum. In these cases, the electron would be probably rejected when the isolation is evaluated. The possibility of having such effect is small but it increases as the decay is more displaced, making the PF isolation definition inefficient when studying highly displaced electrons at large lifetimes.

Displaced electrons used in this thesis utilise another definition for the isolation, called relative *Tracker isolation*. It is computed by summing the p_T of the tracks that are associated with the primary vertex of the collision within $0.03 < \Delta R < 0.3$, that is:

$$\text{rel. Tracker isolation} = \frac{1}{p_T^{\text{(lepton)}}} \sum_{\text{tracks}} p_T^{\text{(track)}} \quad (7.2)$$

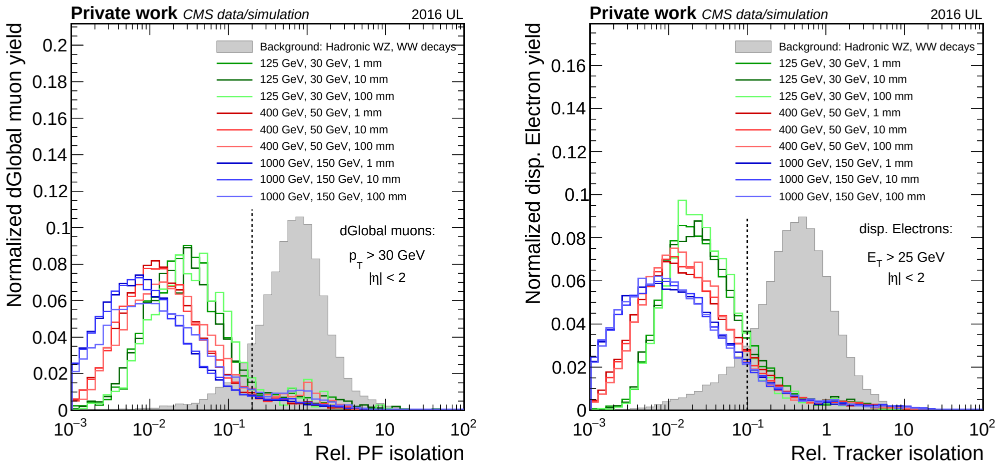


Figure 7.10: Normalized distributions of relative PF isolation for simulated muons (left) and relative Tracker isolation for simulated electrons (right). Hadronic decays of WW and WZ processes are chosen as a representative background for non-isolated leptons. Reconstructed muons (electrons) must fulfill $p_T > 30$ ($E_T < 25$) GeV and $|\eta| < 2$.

As in this case the ECAL superclusters are not included in the sum, the risk of having an overestimation of the isolation due to the own electron cluster is totally mitigated.

Muon tracks fitting the vertices are required to have a value of the relative PF isolation < 0.2. while tracks used to fit the dielectron vertices are required to have a value of the relative Tracker isolation < 0.1. These two requirements were optimized by comparing the relative isolation distributions in leptons from simulated H $\rightarrow$ SS decays against the distribution measured in the simulation of background processes with hadronic activity. In this study, simulated hadronic diboson decays are used as a representative hadronic background. The two comparisons are presented in Figure 7.10.

Charge

The LLPs considered in this search are neutral and consequently the pair of leptons produced from their decays must conserve charge neutrality. In the analysis it is required that the leptons of the fitted vertices have opposite sign (OS) electric charge.

Collinearity

The collinearity $|\Delta\Phi|$ is defined as the difference in azimuthal angle between the dilepton transverse momentum vector $\vec{p}_T^{\,\ell\ell}$ and the $\overrightarrow{DV}$ vector that joins the Primary Vertex PV of the collision and the displaced dilepton vertex DV reconstructed from the lepton pair. A graphical representation of this variable, to-

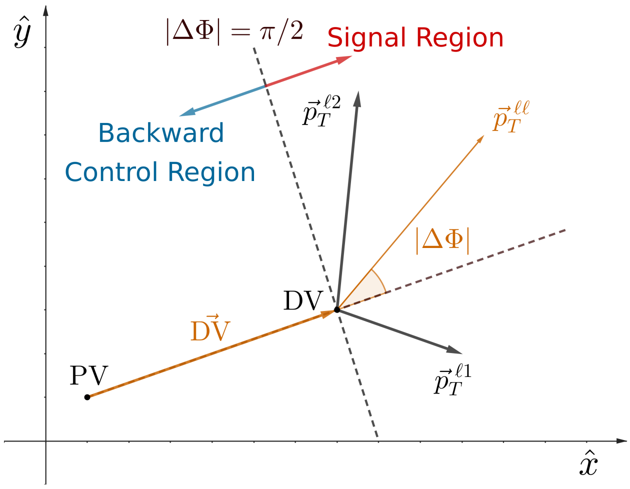


Figure 7.11: Scheme of the vertex collinearity $|\Delta\Phi|$ in a displaced LLP decay. The $\vec{DV}$ is the vector that joins the primary vertex (PV) and the fitted dilepton vertex (DV). The collinearity $|\Delta\Phi|$ corresponds with the difference in azimuthal angle between the dilepton transverse momentum $\vec{p}_T^{\,\ell\ell}$ and the vector $\vec{DV}$.

gether with some of the definitions that are introduced in the next paragraphs, is shown in Figure 7.11.

For real lepton pairs originated in a displaced LL particle decay, the $\overrightarrow{DV}$ vector is representative of the transversal distance travelled by the LL particle since this particle is supposed to be generated at the PV and move to the DV were the decay occurs. Therefore, the direction of $\overrightarrow{DV}$ is identified with the direction in which the LLP was moving. Once the LLP decays and the leptons are produced, they inherit its transverse momentum due to conservation of momenta that must hold in any physical process. As a result, the leptons are likely to be thrown in this direction and the reconstructed $|\Delta\Phi|$ should take small values for displaced LLP decays. In specific cases where the LL particle decays only to leptons (e.g. the S$\rightarrow$ $\ell\ell$ model), the dilepton transverse momentum $\vec{p}_T^{\,\ell\ell}$ and the $\overrightarrow{DV}$ vector should be collinear and ideally $|\Delta\Phi|$ should be zero. This analysis selects dilepton vertices with collinearity values below $|\Delta\Phi| < \pi/2$ as potential long-lived signal candidates.

Leptons from LLP decays would need to move in the opposite direction in which the LLP was moving to fail the $|\Delta\Phi| < \pi/2$ requirement. The efficiency of this cut is really high even for the less boosted decays, as long as the decay vertex is displaced enough for a proper $\overrightarrow{DV}$ vector to be reconstructed. The relative efficiency of the cut is measured in signal dilepton vertices passing the other requirements of the baseline selection. This efficiency is shown for several mass points and lifetimes of H$\rightarrow$SS Monte Carlo simulated samples in Figure 7.12. These efficiencies were measured in 2018 signal simulation, just for ilustration.

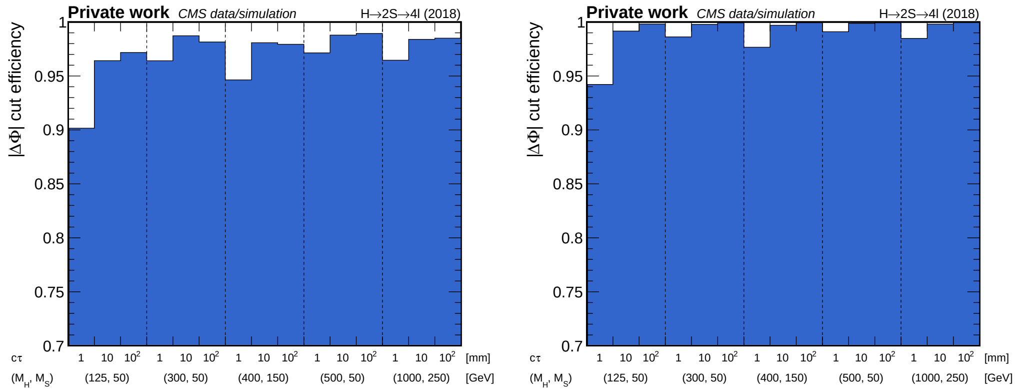


Figure 7.12: Relative efficiency of the $\Delta\Phi$ requirement in electron (left) and muon (right) vertices of 2018 $H \rightarrow SS$ simulation. This efficiency is defined as the fraction of dimuon vertices that fulfill $\Delta\Phi < \pi/2$. All the other cuts in the baseline vertex selection are applied in both denominator and numerator.

Inefficiencies in the $|\Delta\Phi|$ measurement come mainly from limitations in the reconstruction of the dilepton vertex and the $\overrightarrow{DV}$ vector. This has several implications which impact the efficiency of the $|\Delta\Phi| < \pi/2$ requirement and that are reflected in Figure 7.12. The first one is that the overall efficiency is higher for dimuon vertices than for dielectron vertices. This is a consequence of the better resolution of the dGlobal muon reconstruction with respect to the one provided by the electron isotracks. Better resolution in dimuon vertices reconstruction implies better reconstruction of the $\overrightarrow{DV}$ muon vectors, better $|\Delta\Phi|$ determination and higher selection efficiency. Other implication is that the efficiency is lower for LLPs with low lifetimes as dilepton vertices close to the PV may not be resolved. In these cases, the $\overrightarrow{DV}$ is not properly measured and the computed $|\Delta\Phi|$ may be overestimated or even meaningless.

Apart from selecting the LLP candidates of the search, the events which fail the collinearity requirement also used to define a *signal-free* control region where instrumental background can be estimated. The control region for the background prediction is defined by inverting the collinearity requirement i.e. the background prediction corresponds with data events with vertices that fulfill the opposite $|\Delta\Phi| > \pi/2$ requirement. The methodology will be widely described in the dedicated background section of this thesis.

Selection summary

A summary of the selection requirements applied in the dilepton vertices for the three years is presented in Table 7.1. The cuts are applied sequentially on the events and lepton pairs. Events with at least one dilepton vertex passing the aforementioned selection define the Signal Region (SR) of the search. Ultimately, only one vertex is identified with the long-lived particle candidate of each event. If more than one vertex survives the presented selection, the one with the largest $|d_0|/\sigma_d$ value is chosen i.e. the vertex that is more displaced.

Both the isolation and charge requirements are also inverted to define dedicated data control regions to study the QCD and W+jets background contributions, as it will be seen later in Section 8.2.

Event selection					
Requirement	**Dimuon channel**		**Dielectron channel**		
	2016	**2018**	**2016**	**2017**	**2018**
Primary vertex quality	Number of tracks > 4, $\rho < 2$ cm, $|z| < 25$ cm				
Leading muon p_T (GeV)	> 30	> 25	-		
Subleading muon p_T (GeV)	> 30	> 25	-		
Leading electron E_T (GeV)	-	-	> 45	> 40	
Subleading electron E_T (GeV)	-	-	> 30	> 25	

Vertex selection					
Variable	**Dimuon vertices**		**Dielectron vertices**		
	2016	**2018**	**2016**	**2017**	**2018**
Pseudorapidity $|\eta|$	< 2		< 2		
Inv. mass $m_{\ell\ell}$ (GeV)	> 15		> 15	> 100	> 15
Angular distance ΔR	> 0.3		-		
$\cos(\alpha)$	> -0.8	> -0.9	-		
Vertex χ^2/ndof	< 20		< 20		
Rel. Isolation	< 0.2		< 0.1		
Dilepton charge	OS-sign		OS-sign		
Collinearity $|\Delta\Phi|$	$< \pi/2$		$< \pi/2$		

Table 7.1: Summary of the full selection applied in the events and the selection of the displaced dilepton vertex of each channel.

7.3. Signal Region definition

Events with at least one dilepton vertex passing selection described above define the Signal Region (SR) of the search. The LLP candidate of each event contained in this region is identified with the selected vertex that has higher $|d_0|/\sigma_d$ value i.e. the vertex that is more displaced. In addition , the Signal Region is further splitted into several sub-regions to reduce the background and increase the sensitivity to the signal as much as possible.

High vertex multiplicity sub-region

Although ultimately one dilepton vertex is selected as the LLP candidate per event, the number of reconstructed dilepton vertices is used to apply an ad-

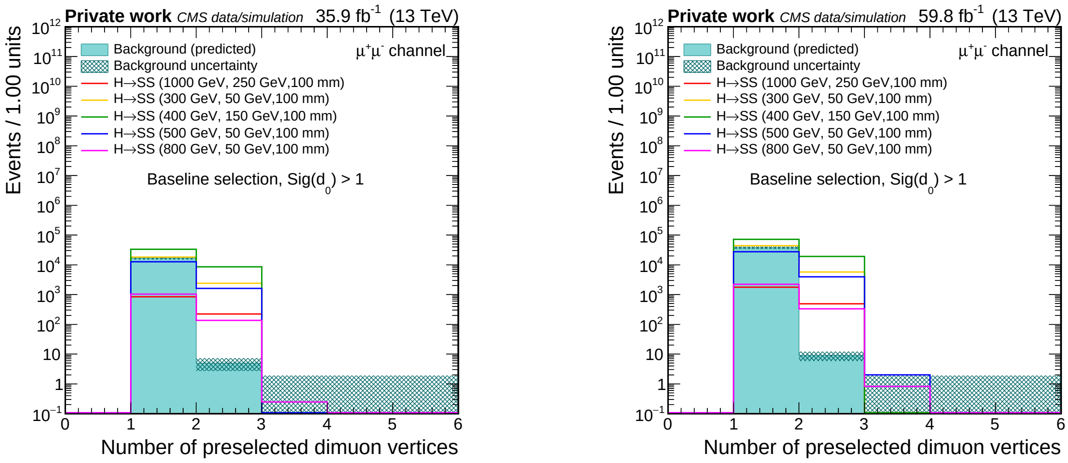


Figure 7.13: Number of reconstructed dimuon vertices passing preselection cuts in 2016 (left) and 2018 (right) in dimuon channel. Events selected are required to have a valid LLP candidate passing the selection criteria as described in Section 7 and be displaced by $\min(|d_0|/\sigma_d) > 1$.

ditional splitting of the SR and increase the sensitivity at low to medium displacement. The variable N_{DV} is defined as the number of reconstructed vertices (dielectron or dimuon, in each case) fitted from leptons that pass ID and the preselection p_T and η cuts as described in the previous sections. Selected events are divided into two orthogonal regions as a function of N_{DV} which are referred as SRI and SRII. Events with exactly $N_{\mathrm{DV}} = 1$ vertices are included in sub-region SRI and the remaining events with $N_{\mathrm{DV}} > 1$ vertices define the sub-region SRII, also referred as the *high multiplicity* sub-region. The distribution of N_{DV} is shown for events collected in Figure 7.13 for dimuon channel and in Figure 7.14 for dielectron channel.

There are several SM processes which have more than one lepton vertex in the final state but when one of the fitted vertices is required to be displaced, the number of events with more than one vertex gets strongly reduced. Signal events where both LLPs decay to the same lepton flavor and with small enough displacement will have more probability of having two vertices reconstructed instead of one. This makes SRII a sub-region with a just few background events and highly sensitive to models where the LLPs are produced in pairs with small to medium $c\tau$ values.

The SRI sub-region, on the other hand, will contain most of the background but also a good portion of signal events as SRI will feed from scenarios where one LLP decays to muons and the other to electrons. SRI will also retain some events in cases were only one vertex managed to be reconstructed, despite of the nature of its decay. This last scenario is likely to happen in models predicting LLPs with high lifetimes where the efficiency to reconstruct the vertices is so small that it is unlikely to have both vertices reconstructed.

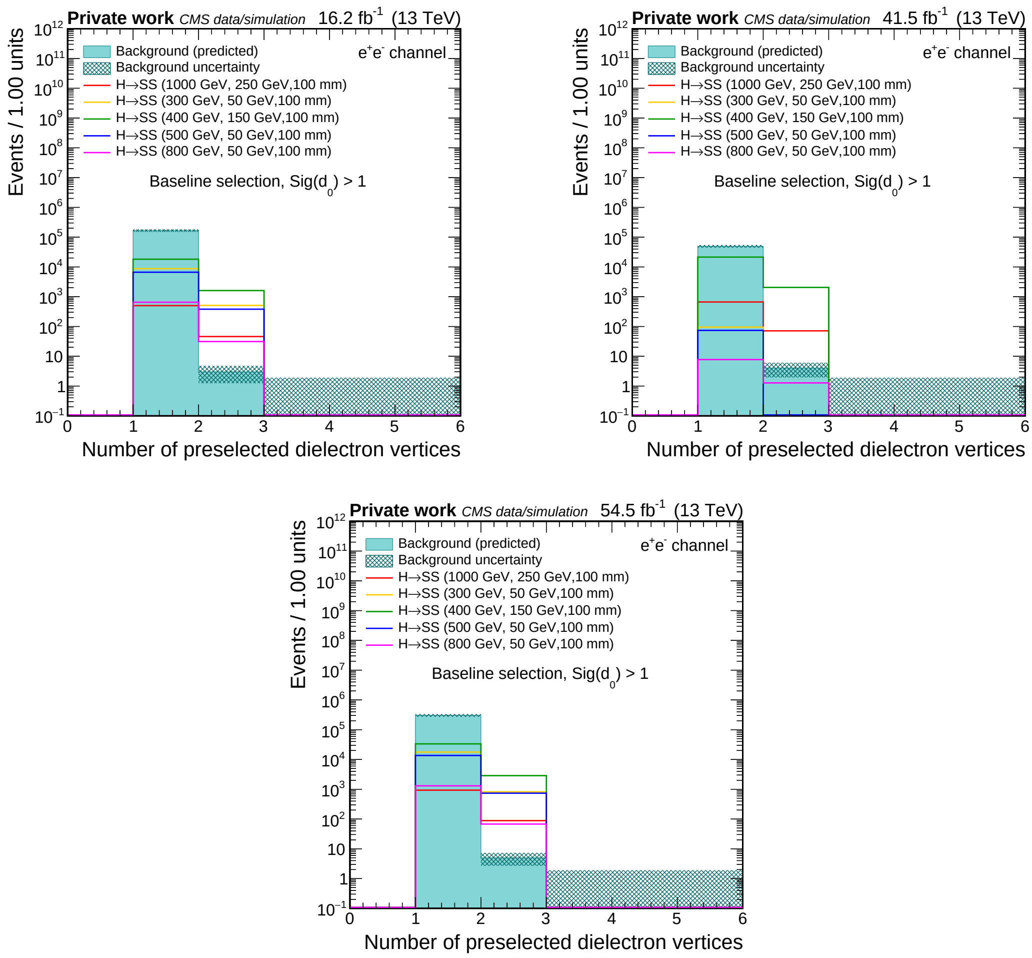


Figure 7.14: Number of reconstructed dielectron vertices passing preselection cuts in 2016 (left), 2017 (right) and 2018 (bottom) in dielectron channel. Events selected are required to have a valid LLP candidate passing the selection criteria as described in Section 7 and be displaced by $\min(|d_0|/\sigma_d) > 1$.

This search is sensitive to LLPs that are produced alone as only the reconstruction of just one dilepton vertex is required. However, sensitivity to LLPs produced in pairs is incremented with the N_{DV} splitting and the inclusion of the high multiplicity region SRII.

The minimum $|d_0|/\sigma_d$ of the LLP candidate in SRII is optimized as a function of the channel and year. For dimuon channel we have required $\min(|d_0|/\sigma_d) > 2$ in both 2016 and 2018 and for dielectron channel the requirement is $\min(|d_0|/\sigma_d) > 1(1.5)$ in 2016 and 2017 (2018). The $\min(|d_0|/\sigma_d)$ is tightened in 2018 because the background yield is found to be higher for this year.

Invariant mass selection in single vertex region

To cope with the large amount of background in the SRI sub-region it is necessary a more restrictive selection applied to the dilepton vertices. An additional splitting of the resulting phase space is also performed to increase the sensitivity to the different models considered.

The major background contribution in SRI comes from mismeasurement of Drell-Yan events due to the large cross section of this process. For this reason, a cut rejecting vertices around the Z resonance peak is applied to reduce this background contribution. The selected LLP candidates are required to have $|m_{\ell\ell} - m_Z| > 10$ GeV, being m_Z the mass of the Z boson. The exclusion of the $[81, 101]$ mass range was proved to not decrease the sensitivity to the studied signals as their contribution is expected to be negligible in comparison to the high amount of background measured in this region. The predicted background yields within this range was measured to be at least 10-20 times higher than in the Z peak vicinity (depending on the lepton type and year). In additionm, this requirement also enables the possibility to use the excluded $|m_{\ell\ell} - m_Z| < 10$ GeV region as a control region to validate the background estimation method, as it is done in Section 8.6.2.

Since the number of background vertices decreases as we go to higher invariant mass values, the sensitivity of the analysis is increased by defining two subregions at low and high mass. The SRI region is splitted into two sub-regions below and above the already excluded Z peak:

- Low mass sub-region (SRIa): $m_{\ell\ell} < 81$ GeV
- High mass sub-region (SRIb): $m_{\ell\ell} > 101$ GeV

SRIa and SRIb are expected to provide sensitivity to LLPs with masses below and above the chosen threshold. It is worth noting that electron vertices collected in 2017 will only contribute to the high mass SRIb sub-region as a selection of $m_{ee} > 100$ GeV is applied to account for the diphoton trigger mass turn-on.

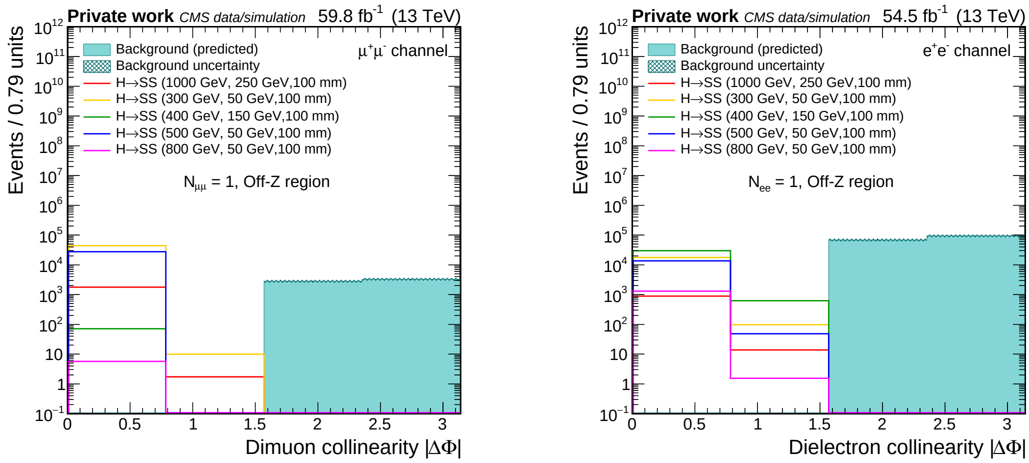


Figure 7.15: Binned $|\Delta\Phi|$ distribution for predicted background and simulated signal for dimuon (left) and dielectron (right) vertices contained in region SRI. The cut in $m_{\ell\ell}$ is also applied so the events shown are out of the Z peak vecinity. Distributions correspond to data collected in 2018. As the background prediction is estimated by inverting the $|\Delta\Phi|$ requirement, the predicted background yields correspond to data measured in the mirrored $[\pi/2, 3\pi/4, \pi]$. The data observation in the signal region bins $[0, \pi/4, \pi/2]$ is not shown.

Collinearity selection in single vertex region

After the cut in $m_{\ell\ell}$ is applied, the behavior of $|\Delta\Phi|$ is further exploited as this variable is expected to have a great part of the signal close to $|\Delta\Phi| \sim 0$ while being much more distributed in the range $[0, \pi/2]$ for the background. It was first tried to split SRI in two bins divided by $|\Delta\Phi| = \pi/4$, that is

$$|\Delta\Phi| = \left[0, \frac{\pi}{4}, \frac{\pi}{2}\right]$$

The resulting binned distribution of $|\Delta\Phi|$ is shown for both for signal and background estimation in Figure 7.15 for 2018 just to ilustrate the splitting. In this representation, the Z peak is already excluded, but events for both low and high mass regions are included together. As it can be seen in these plots, the predicted background is distributed in both sub-regions while nearly the 100% of the simulated H→SS events are contained in the former. Unfortunately, the portion of signal events in the bin $[\pi/4, \pi/2]$ is so small that its addition does not add sensitivity with respect to consider just the first $[\pi/4, \pi/2]$ range. For this reason, the $|\Delta\Phi|$ selection was tightened to $|\Delta\Phi| < \pi/4$ in order to get rid of the large background contribution present in the range $[\pi/4, \pi/2]$, which was found to be a more efficient strategy.

Final splitting of single vertex region

The final splitting of low and high mass sub-regions (SRIa and SRIb) is done by making use of $\min(|d_0|/\sigma_d)$ and L_{xy} as these are the main discriminants of the search.

To split the dimuon vertices in SRIa and SRIb the phase space is divided by splitting the events below and over $\min(|d_0|/\sigma_d)$ and L_{xy} thresholds that are optimized to provide the best sensitivity. SRIa and SRIb are divided into four sub-regions A, B, C and D each that are defined as:

- **IaA**: $3 < \min(|d_0|/\sigma_d) < 9$ and $0.02 < L_{xy} < 0.2$
- **IaB**: $3 < \min(|d_0|/\sigma_d) < 9$ and $L_{xy} > 0.2$
- **IaC**: $\min(|d_0|/\sigma_d) > 9$ and $0.02 < L_{xy} < 0.2$
- **IaD**: $\min(|d_0|/\sigma_d) > 9$ and $L_{xy} > 0.2$
- **IbA**: $3 < \min(|d_0|/\sigma_d) < 9$ and $0.02 < L_{xy} < 0.07$
- **IbB**: $3 < \min(|d_0|/\sigma_d) < 9$ and $L_{xy} > 0.07$
- **IbC**: $\min(|d_0|/\sigma_d) > 9$ and $0.02 < L_{xy} < 0.07$
- **IbD**: $\min(|d_0|/\sigma_d) > 9$ and $L_{xy} > 0.07$

The combination of sub-regions A, B, C and D is expected to cover the sensitivity at different lifetime, mass and boosting regimes. Sub-region D contains the less amount of background while B and C are expected to keep a higher amount of it. Sub-region A has most of the background but still can provide sensitivity to LLP decays that are poorly displaced.

Figures 7.16 and 7.17 collect the 2-dimensional $\min(|d_0|/\sigma_d)$ vs L_{xy} distributions of the background dimuon vertices estimated for low and high mas sub-regions.

In general for the signal, the higher the lifetime of the LLP, the higher the values of $\min(|d_0|/\sigma_d)$ and L_{xy} will be. At a second order, the vertices will have a lower or higher L_{xy} and $\min(|d_0|/\sigma_d)$ depending on the kinematics of the LLP decay. For example, the same lifetime will result in a lower L_{xy} for high masses and/or low LLP p_T, as in these scenarios the LLP is expected to move slowly. To ilustrate this effect, the distributions of simulated H→SS decays with masses $m_H = 500$ GeV and $m_S = 50$ GeV for region SRIa, where $c\tau = 1, 10,$ and 100 cm lifetimes are shown separately. The same lifetimes for H→SS decays with masses $m_H = 500$ GeV and $m_S = 150$ GeV in SRIb are also presented. Both mass points are collected in Figure 7.18. By comparing the plots of the same $c\tau$ for the two models one can conclude that the way in which the subregions of the $\min(|d_0|/\sigma_d)$ vs L_{xy} plane are populated has a high dependence on the masses involved in the process.

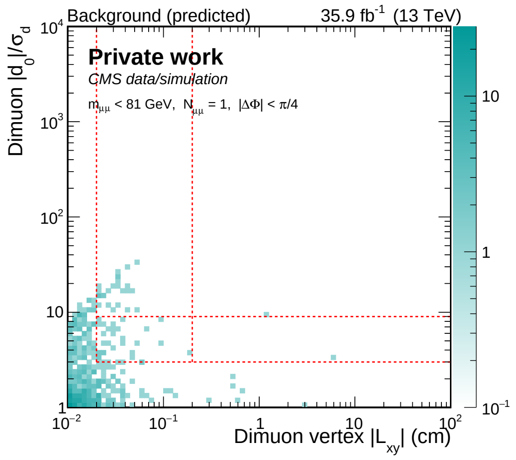


Figure 7.16: $\min(|d_0|/\sigma_d)$ vs L_{xy} distribution of the background dimuon vertices estimated in the low mass sub-region (SRIa) in 2016 (left) and 2018 (right).

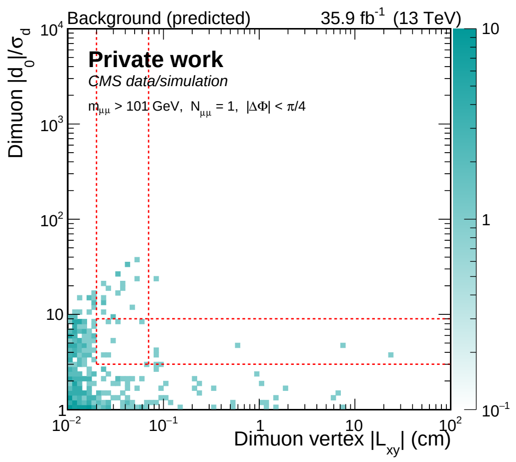


Figure 7.17: $\min(|d_0|/\sigma_d)$ vs L_{xy} distribution of the background dimuon vertices estimated in the high mass sub-region (SRIb) in 2016 (left) and 2018 (right).

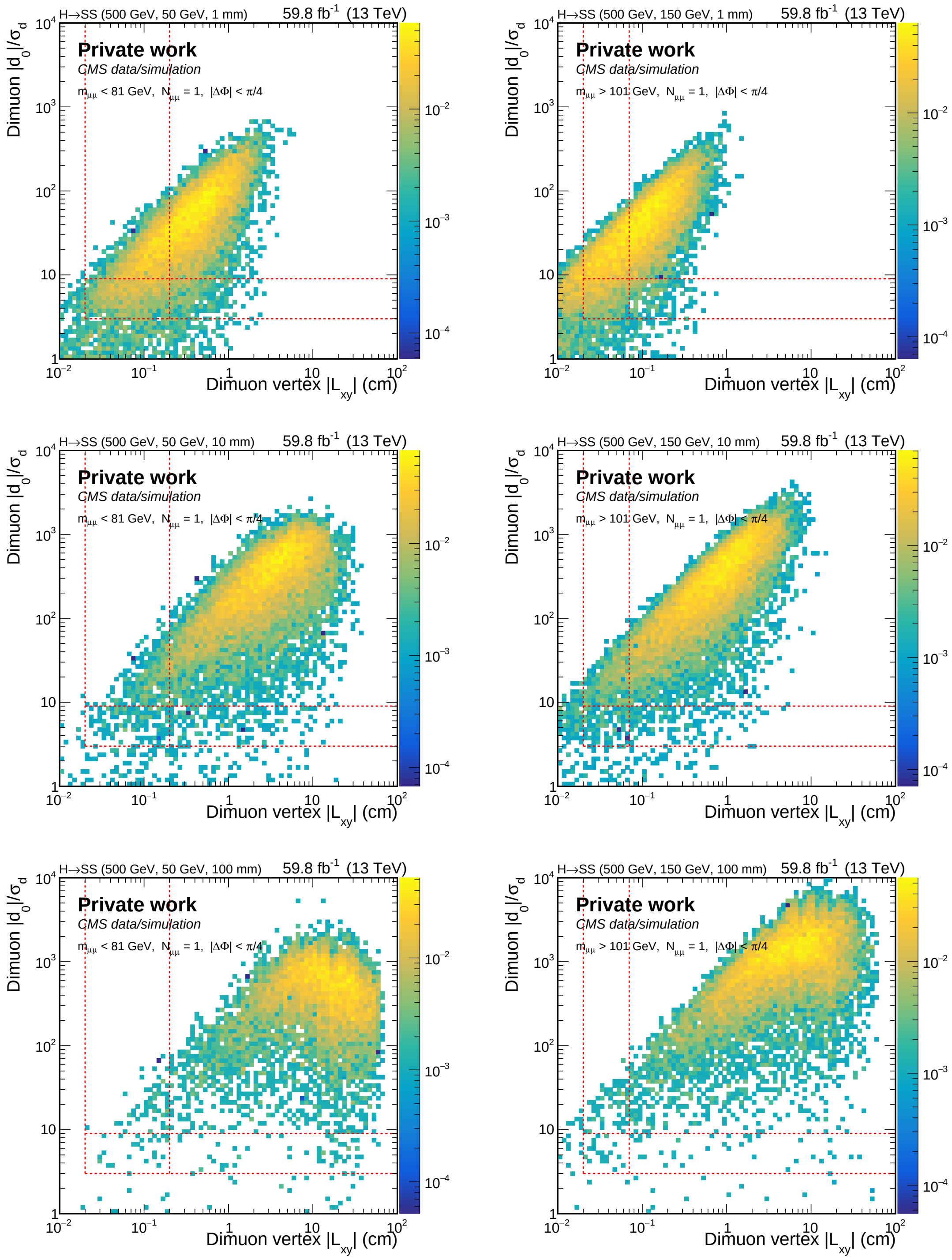


Figure 7.18: $\min(|d_0|/\sigma_d)$ vs L_{xy} distribution of the simulated dimuon vertices of the H→SS model with masses $m_H = 500$ GeV, $m_S = 50$ GeV and $m_H = 500$ GeV, $m_S = 150$ in SRIa (left) and SRIb (right), respectively. 2018 simulation is shown.

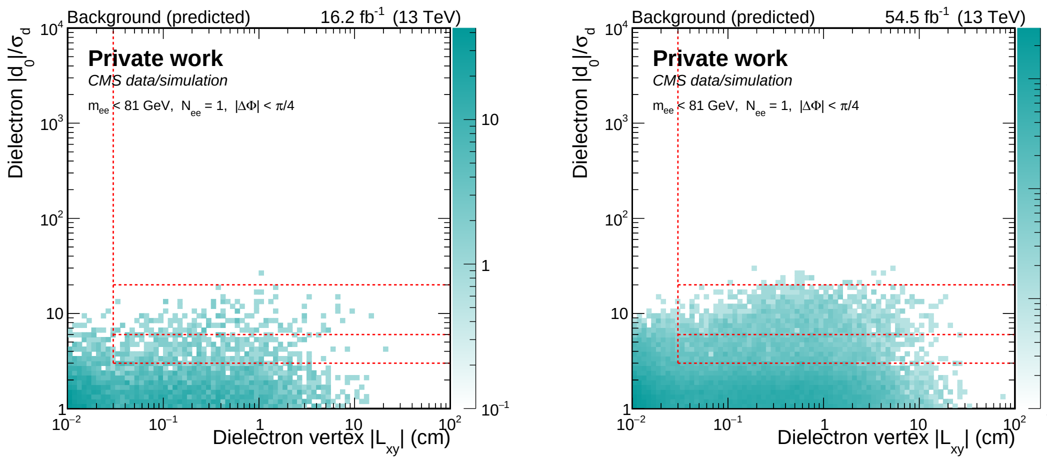


Figure 7.19: $\min(|d_0|/\sigma_d)$ vs L_{xy} distribution of the background dielectron vertices estimated in the low mass sub-region (SRIa) in 2016 (left), 2017 (right) and 2018 (bottom).

Due to the limited resolution in the position of the dielectron vertices the spread in L_{xy} is larger for the background and the L_{xy} separation is not found as efficient as it is for dimuon vertices. Therefore the dielectron SRIa and SRIb regions are subdivided in three subregions each with the following $\min(|d_0|/\sigma_d)$ bins:

- **Ia[A,B,C]**: $\min(|d_0|/\sigma_d) \in [3, 6, 20, \infty)$
- **Ib[A,B,C]**: $\min(|d_0|/\sigma_d) \in [3, 6, 15, \infty)$

Figures 7.19 and 7.20 present the 2-dimensional $\min(|d_0|/\sigma_d)$ vs L_{xy} distributions of the background estimated for all years of data taking for SRIa and SRIb, respectively. The plots clearly show a larger spreading in L_{xy} for dielectron vertices compared to dimuon vertices. Simulated H→SS dielectron vertices with a mass of the H of $m_H = 500$ GeV are chosen again as representative signals. The plots of the low mass SRIa sub-region are shown for LLP mass of $m_S = 50$ GeV and those of the high mass SRIb sub-region are shown for a mass of $m_S = 150$ GeV. The distribution is shown separately for the different lifetimes available in Figure 7.21.

Summary of signal region splitting

The aforementioned selections define a set of sub-regions that are combined to extract the final results of the search. This set is constituted of 18 orthogonal sub-regions for 2016 and 2018 data analyzed in the dimuon channel and other 18 orthogonal sub-regions analyzed in the dielectron channel for 2016, 2017 and 2018 data. Table 7.2 collects all the requirements applied to the events passing baseline selection that define each one of these regions.

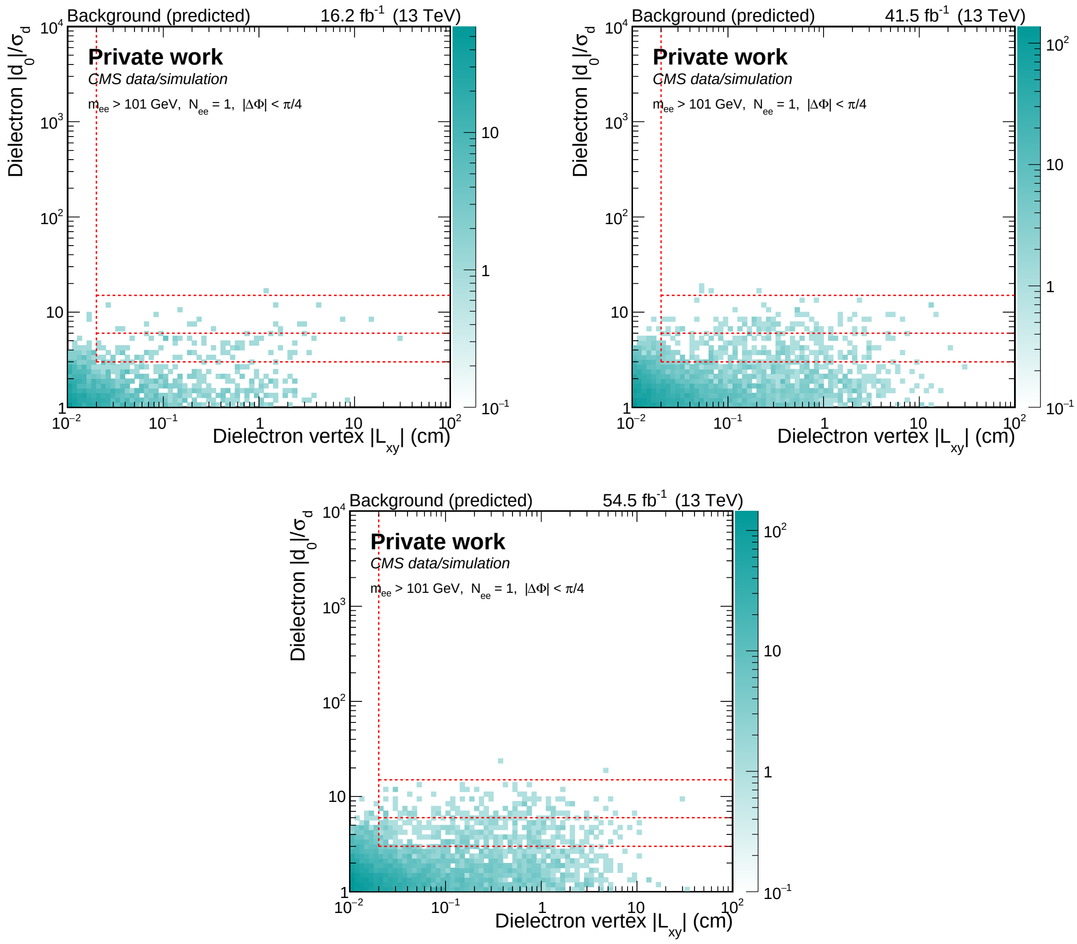


Figure 7.20: $\min(|d_0|/\sigma_d)$ vs L_{xy} distribution of the background dielectron vertices estimated in the high mass sub-region (SRIb) in 2016 (left), 2017 (right) and 2018 (bottom).

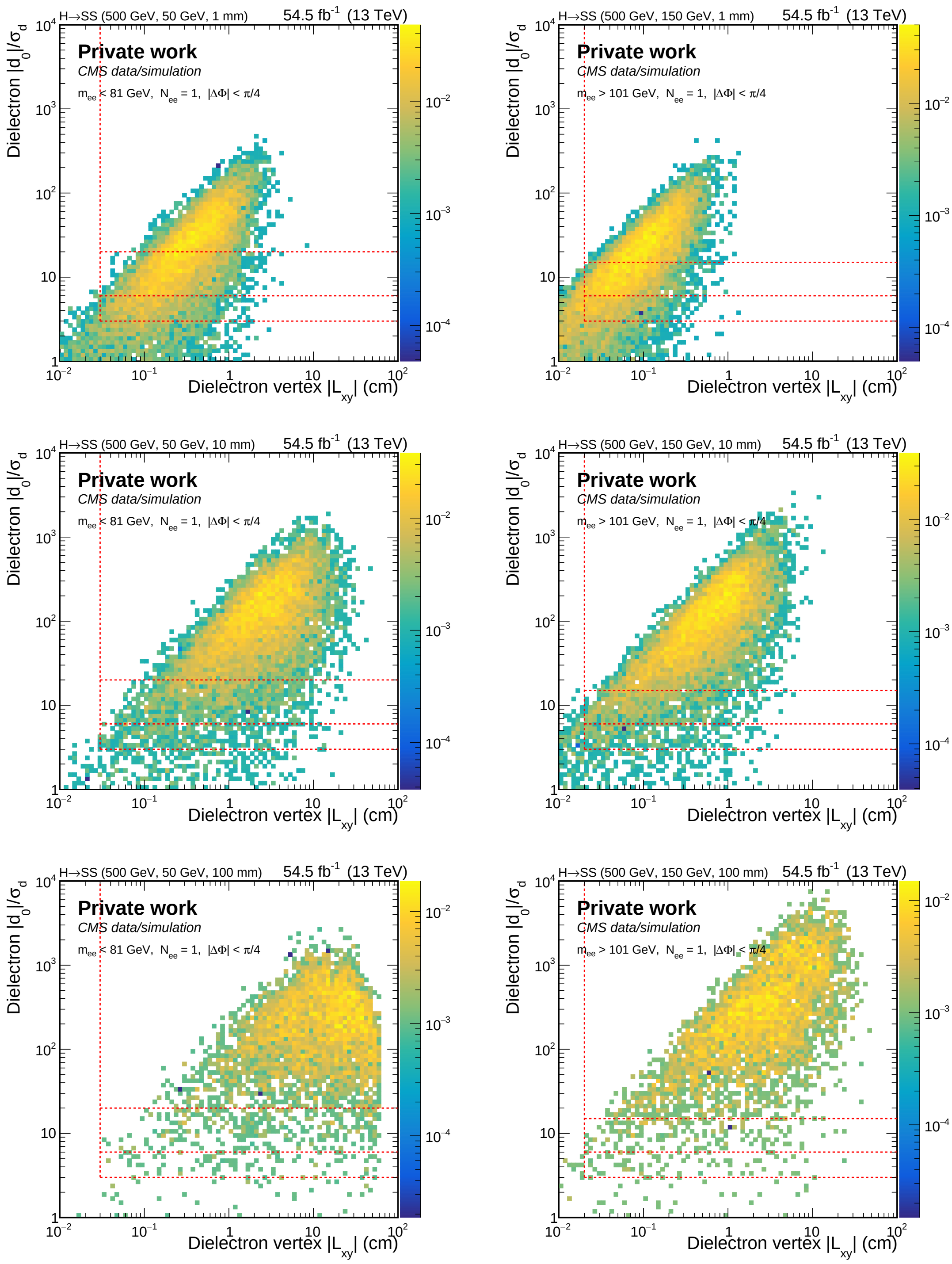


Figure 7.21: $\min(|d_0|/\sigma_d)$ vs L_{xy} distribution of the simulated dielectron vertices of the H→SS model with masses $m_H = 500$ GeV, $m_S = 50$ GeV and $m_H = 500$ GeV, $m_S = 150$ in SRIa (left) and SRIb (right), respectively. 2018 simulation is shown.

Muon channel sub-regions					
Region	$N_{\mu\mu}$	$\|\Delta\Phi\|$	$m_{\mu\mu}$ (GeV)	$\min(\|d_0\|/\sigma_d)$	L_{xy} (cm)
SRIaA	$= 1$	$< \pi/4$	< 81	$[3, 9)$	$[0.02, 0.2)$
SRIaB					$[0.2, \infty)$
SRIaC				$[9, \infty)$	$[0.02, 0.2)$
SRIaD					$[0.2, \infty)$
SRIbA			> 101	$[3, 9)$	$[0.02, 0.07)$
SRIbB					$[0.07, \infty)$
SRIbC				$[9, \infty)$	$[0.02, 0.07)$
SRIbD					$[0.07, \infty)$
SRII	> 1	$< \pi/2$	-	> 2	-

Electron channel sub-regions					
Region	N_{ee}	$\|\Delta\Phi\|$	m_{ee} (GeV)	$\min(\|d_0\|/\sigma_d)$	L_{xy} (cm)
SRIaA	$= 1$	$< \pi/4$	< 81	$[3, 6)$	> 0.03
SRIaB				$[6, 20)$	> 0.03
SRIaC				$[20, \infty)$	> 0.03
SRIbA			> 101	$[3, 6)$	> 0.02
SRIbB				$[6, 15)$	> 0.02
SRIb3				$[15, \infty)$	> 0.02
SRII	> 1	$< \pi/2$	-	$> 1(1.5)$ (2018)	-

Table 7.2: Defition of the signal regions for dimuon (top) and dielectron (bottom) channels.

Chapter 8

Background

This section describes the sources of background processes that can be reconstructed as signal events and the strategies followed to mitigate them or to estimate the irreducible component.

In general, the background sources of the search comprise all the potential processes involving dilepton vertices which have the possibility to be displaced from the interaction point. In this analysis, the background is divided into two separate categories depending on whether the displacement of the vertices is genuine or an instrumental effect:

- **Displaced background**: Conformed by processes that could end up with dilepton vertices that are truly displaced, and be consequently counted in the signal region as signal events. This background source is entirely rejected by means of the dilepton vertex selection presented in Section 7, which was optimized to that purpose. Two main sources can be distinguished within this category: Cosmic muons and nuclear interactions with CMS material. Section 8.8 of this chapter is dedicated to their study.
- **Instrumental background**: It contains SM processes where the displacement is not genuine or physical but caused by resolution effect that may shift the dilepton vertex position. This is the main background of the search and, unlike the displaced background, it has an irreducible component. It cannot be entirely removed without reducing the signal sensitivity and its contribution needs to be estimated.

Since most Standard Model processes are prompt and the so-called displaced background is rejected by the analysis selection, background in this analysis is considered to be purely instrumental. The displacement observed in this background events is not physical but caused by resolution effects which may shift the transverse impact parameter or the vertex position to higher values. These resolution effects are usually hard to emulate by the simulation and could result in a critical loss of precision in the tails of the $|d_0|/\sigma_{d_0}$ distribution where the highly displaced regime is studied. In addition, the statistics of the Monte Carlo samples are limited and, while the prompt regime contains an abundance of statistics, the tails are often populated by just a few events, making this regions prone to fluctuations.

To provide a realistic estimate of the background, this search estimates the final background contribution exclusively from data. The method used for this estimation is described in Section 8.1 and several studies for its optimization and validation are presented in the following sections.

8.1. Background estimation method

The background estimation method makes use of the collinearity variable $|\Delta\Phi|$, which was one of the variables used to select the dilepton vertices. As it will be seen in this section, $|\Delta\Phi|$ has a well-defined behavior for both signal and background dilepton vertices, and its properties are exploited in the background estimation method.

The $|\Delta\Phi|$ variable measures the difference in azimuthal angle between the dilepton position vector $\overrightarrow{DV}$ and the dilepton momentum $\vec{p}_T^{\ell\ell}$. For a true LLP $\overrightarrow{DV}$ is representative of the distance traveled by the LLP and these two vectors should be collinear due to the conservation of momentum between the LLP and the leptons. That is, for signal LLP decays the collinearity should be essentially $|\Delta\Phi| \sim 0$. Hence, the Signal Region is defined as the region with $|\Delta\Phi| < \pi/2$[1], which is also called the Forward Region (FR).

The signal efficiency of this selection is very high, and therefore, the orthogonal region with $|\Delta\Phi| > \pi/2$ vertices is expected to have none or negligible contribution from displaced LLP decays. This second orthogonal region is called the *Backward Control Region* (BCR) and it is expected to be signal-free.

For instrumental background vertices $|\Delta\Phi|$ has a very complex behavior which is strongly correlated with the lepton kinematics. In these events, there is no genuinely displaced dilepton vertex that can be reconstructed because Standard Model particles decay so fast that the decay occurs at the interaction point. Ideally, the reconstructed dilepton vertex DV should match the original primary vertex PV. However, in practice, the fitted DV position is found slightly displaced from the PV due to resolution effects and the fact that both fitting methods were performed differently (e.g. PV is fitted with all tracks associated to it while the DV is fitted with just two lepton tracks).

Despite this displacement, the instrumental effects that cause it are not expected to have a privileged direction to occur, and $|\Delta\Phi|$ is expected to remain symmetric with respect to $\pi/2$ for any instrumental background source. This property is known as *forward-backward symmetry* and it implies that the SR, with

[1] Although it is further categorized in $|\Delta\Phi| < \pi/4$ or $\pi/4 < |\Delta\Phi| < \pi/2$

$|\Delta\Phi| < \pi/2$, and the BCR, with $|\Delta\Phi| > \pi/2$, are expected to be equally populated by instrumental background vertices. Consequently, the BCR is expected to have no signal vertices and contain the same number of background events as the SR.

The background estimation method consists of identifying the instrumental background contribution in the Signal Regions bins with the data measured in the BCR regions that would be obtained by inverting the $|\Delta\Phi|$ selection with respect to $\pi/2$.

8.2. Control regions for instrumental background studies

Standard Model processes with at least two leptons in the final state will have some dilepton candidates that will contribute to the background yield. Drell-Yan contribution will be especially relevant as it has two opposite charged leptons coming from a resonance with a genuine vertex decay, matching the search signature, and its cross section is relatively high in comparison with other SM processes. This is the main background of the search. To a lesser extent, dileptonic final states of $t\bar{t}$+jets and diboson (ZZ, WZ, WW) processes are prone to have dilepton candidates passing the baseline selection. The former will not only include leptons from W decays, but also leptons produced in semileptonic B decays. In addition, although W bosons from W+jets processes decay in single-lepton final states, there is a chance that the final single lepton could be paired with leptons produced from jets hadronization or jet misidentification. This is also the case for QCD interactions, where pions produced in the core of the jet hadronisation could be reconstructed at a certain distance from the interaction point.

All these processes are expected to contribute to the instrumental background component that needs to be estimated. Several control regions are defined to study these background sources and to validate the background estimation method. These regions are defined by splitting the dilepton vertices in four regions depending on whether they pass the isolation and opposite charge requirements. No collinearity requirement is imposed on dilepton vertices since its behavior is what it is under study. This region splitting is shown in the scheme of Figure 8.1 and it can be summarized as follows:

- **Signal-like Region**: It matches the analysis vertex selection with the exception of the $|\Delta\Phi| < \pi/2$ requirement. Only Monte Carlo simulation was studied in this region as data was blinded until the final stages of the analysis. Contributions from Drell-Yan, $t\bar{t}$-jets and diboson processes are studied in this region.
- **QCD Control Region**: It contains all dilepton vertices passing baseline selection except charge, but fitted with both leptons failing the cut in isola-

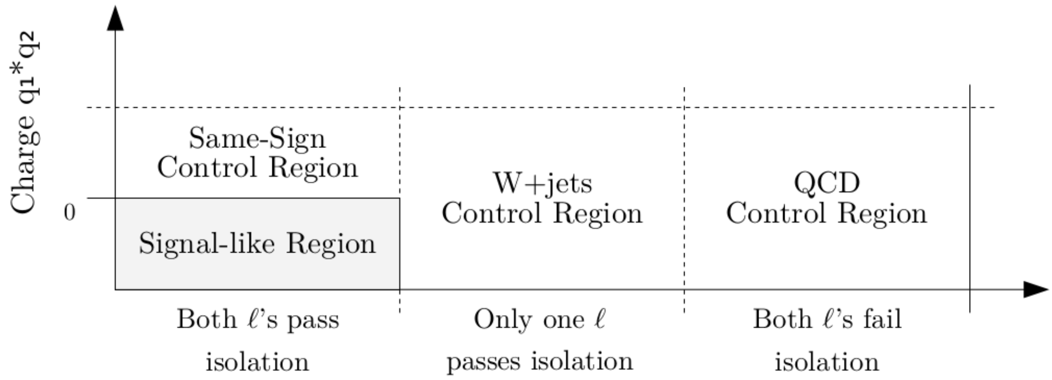


Figure 8.1: Regions defined to study the potential processes contribution to the instrumental background.

tion. Since this background major contribution arises from dilepton pairs produced in the core of jet hadronization or pion misindentification we expect this region to be enriched in QCD.

- **W+jets Control Region**: It contains dilepton vertices passing baseline selection except charge, where only one of the leptons (either negatively or positively charged) is isolated and the other one is non-isolated. This region matches the W+jets signature, so we would expect it to be enriched in W+jets events.
- **Same-Sign Control Region**: It contains dilepton vertices that are isolated but have same sign charge. This region is left for additional studies to reinforce the conclusions derived in the other regions.

8.3. The background collinearity distribution

As previously introduced, the collinearity is a variable that can serve a dual purpose of not only facilitating the rejection of notable background contributions but also enabling the estimation of any remaining instrumental background that persists after the vertex selection process. This subsection is entirely devoted to conducting an in-depth examination of the properties exhibited by the collinearity in instrumental background vertices.

The main feature that is studied is the isotropy of the displacement of the dilepton vertices introduced by instrumental effects. It is checked by studying the population of instrumental displaced vertices in the transverse plane of the detector, where it is observed that the DV position spreads from the original PV position in an isotropic way along the inner tracker volume. This isotropic spreading is shown for simulated dileptonic $t\bar{t}$ vertices (2018 Monte Carlo) in a PV-centered frame (x', y') in Figure 8.2, where the (x', y') coordinates refer

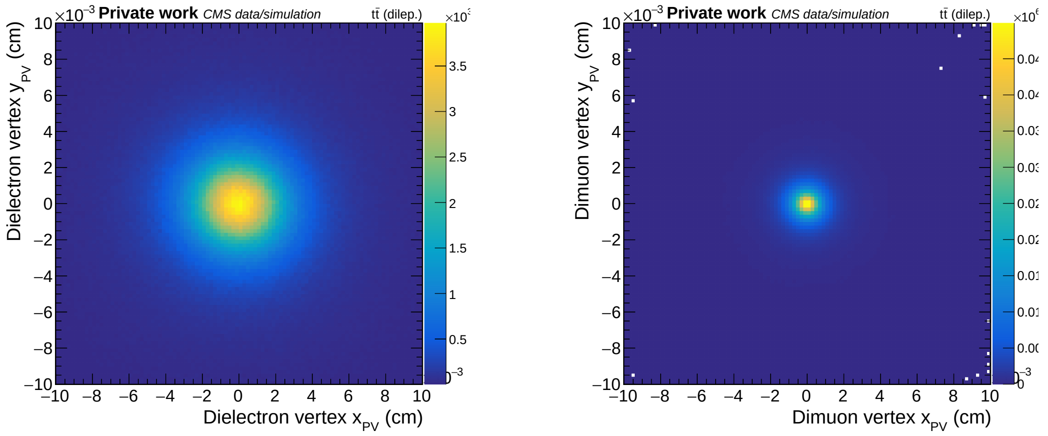


Figure 8.2: Fitted dielectron (left) and dimuon (right) vertices position from Monte Carlo simulated $t\bar{t}$ decays expressed in PV-centered coordinates, defined with Eq.(8.1).

to the translation from the detector frame to the PV-centered frame:

$$\begin{cases} x' = x - x_{\text{PV}} \\ y' = y - y_{\text{PV}} \end{cases} \tag{8.1}$$

From the vertex patterns of Figure 8.2 it can be seen how the dilepton vertices populate the vicinity of the primary vertex in prompt processes. It is observed that dielectron vertices spread more than dimuon vertices, revealing a lower accuracy when reconstructing the vertex position. Instrumental background is therefore expected to have a higher impact in this channel.

The isotropic spreading along the detector, where no direction seems to be privileged, would suggest that the $\overrightarrow{DV}$ vector would be totally uncorrelated with $\vec{p}_T^{\,\ell\ell}$. Under such circumstances one may expect that the collinearity distribution would be flat for prompt background events. However, this is not the case, as it turns out that the reconstructed dilepton vertex position has a small correlation with the dilepton transverse momentum. The effect can be seen in Figures 8.3 and 8.4 where the $|\Delta\Phi|$ normalized distribution is shown for every background source presented in the previous section. Every one of them presents a very well defined non-flat shape revealing the existence of a correlation between the reconstructed DV and the dileptonic transverse momentum.

The existing correlation between $|\Delta\Phi|$ and $\vec{p}_T^{\,\ell\ell}$ is due to the fact that a small shift in the DV position is introduced by the track fitting procedure itself. This effect was also observed in the similar displaced lepton analysis that was done in the CMS Collaboration in 2012 [99]. When fitting the vertex, each track constrains the fitted vertex position more strongly in the direction perpendicular to its transverse momentum. As a result, the DV position is likely to be shifted to-

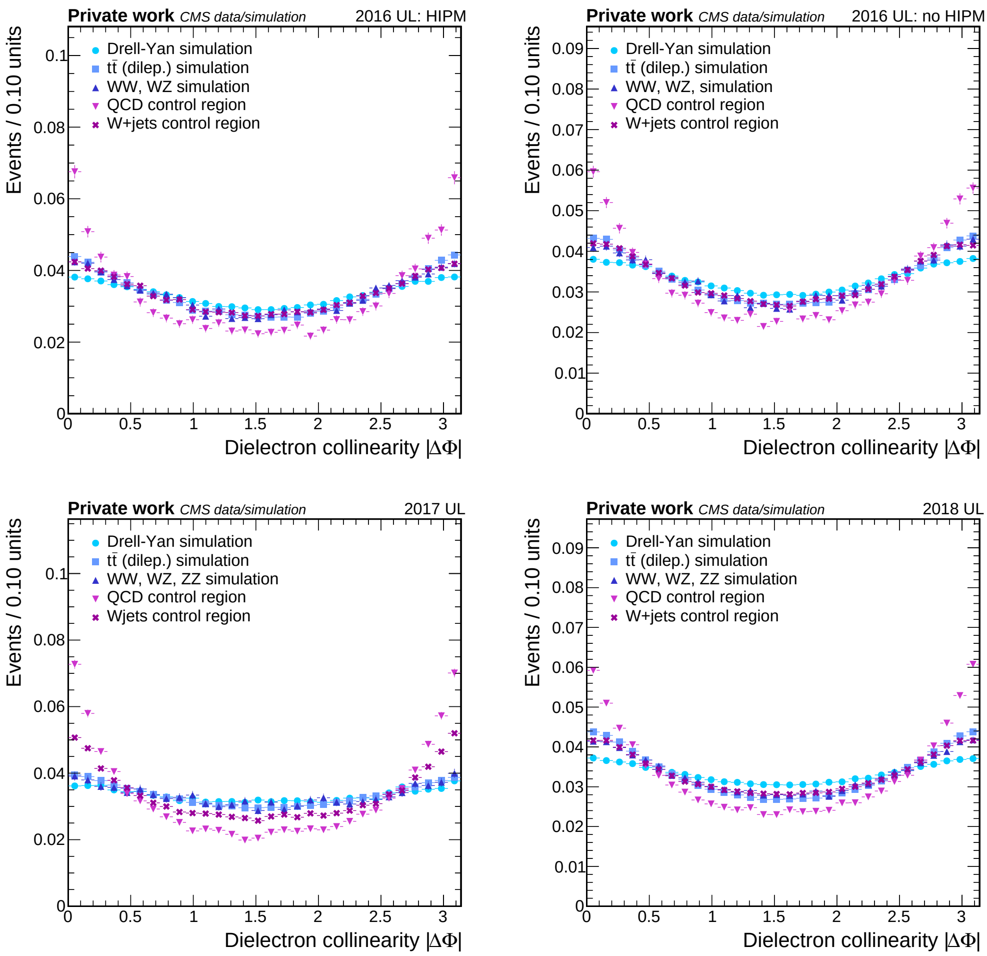


Figure 8.3: Collinearity normalized distributions of dielectron vertices for each considered background source.

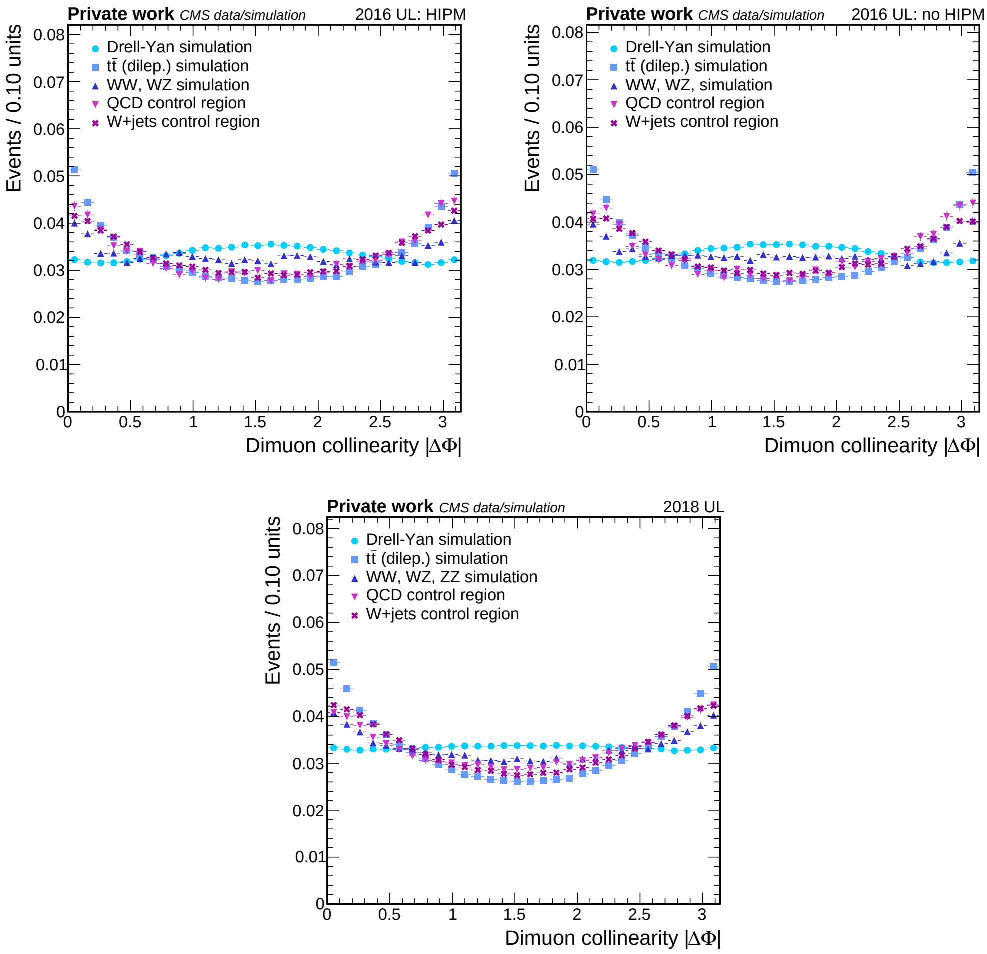


Figure 8.4: Collinearity normalized distributions of dimuon vertices for each considered background source.

wards or backwards each track direction and, ultimately, the final DV position will depend on the effects introduced by the two tracks jointly.

This shifting can be noticed by studying how the fitted DV position is dispersed from the original PV position along these specific directions. Several rotations are applied to the $(x,\, y)$ coordinates of the fitted vertices to that purpose. To appreciate each track effect, the x axis is aligned with the higher and lower p_T tracks directions of the vertices. The global impact introduced by both tracks is studied by aligning the x axis with the total $\vec{p}_T^{\ell\ell}$ transverse momentum vector. The reference frames that result from such transformations are denoted as $(x_1,\, y_1)$, $(x_2,\, y_2)$ and $(x'',\, y'')$, respectively, and they relate with the detector reference frame as:

$$\begin{cases} x_1 = (x - x_{\mathrm{PV}})\cos(\phi_{\ell 1}) + (y - y_{\mathrm{PV}})\sin(\phi_{\ell 1}) \\ y_1 = -(x - x_{\mathrm{PV}})\sin(\phi_{\ell 1}) + (y - y_{\mathrm{PV}})\cos(\phi_{\ell 1}) \end{cases} \tag{8.2}$$

$$\begin{cases} x_2 = (x - x_{\mathrm{PV}})\cos(\phi_{\ell 2}) + (y - y_{\mathrm{PV}})\sin(\phi_{\ell 2}) \\ y_2 = -(x - x_{\mathrm{PV}})\sin(\phi_{\ell 2}) + (y - y_{\mathrm{PV}})\cos(\phi_{\ell 2}) \end{cases} \tag{8.3}$$

$$\begin{cases} x'' = (x - x_{\mathrm{PV}})\cos(\phi_{\ell\ell}) + (y - y_{\mathrm{PV}})\sin(\phi_{\ell\ell}) \\ y'' = -(x - x_{\mathrm{PV}})\sin(\phi_{\ell\ell}) + (y - y_{\mathrm{PV}})\cos(\phi_{\ell\ell}) \end{cases} \tag{8.4}$$

where $\phi_{\ell 1}$, $\phi_{\ell 2}$ and $\phi_{\ell\ell}$ are the azimuthal angles of both tracks and the $\vec{p}_T^{\ell\ell}$ vector.

The DV position is shown for the three new coordinate frames in Figure 8.5 for displaced dilepton vertices reconstructed in the 2018 $t\bar{t}$ Monte Carlo simulation. This sample is chosen to avoid additional correlations between the leptons as in the case of Drell-Yan processes where the lepton pair comes from the same particle decay. As it can be seen these plots, the spread of the fitted DV position elongates along the x axis in all cases i.e. towards and backwards the two tracks $\vec{p}_T$ and the overall $\vec{p}_T^{\ell\ell}$ directions. The correlation between $\overrightarrow{DV}$ and $\vec{p}_T^{\ell\ell}$ manifests itself in this lack of isotropicity. The fact that the spreading is enhanced in the $\vec{p}_T^{\ell\ell}$ direction results in more dilepton vertices populating $|\Delta\Phi|$ values close to 0 and $\pi/2$.

The collinearity is built up from the fitted $\overrightarrow{DV}$, that is correlated with the tracks directions, and $\vec{p}_T^{\ell\ell}$, which is reconstructed by summing up the two leptons contributions. All of this results in underlying relations between these two variables that will determine the shape of the collinearity distribution. At the end, it can be concluded that the collinearity is a very complex variable that is strongly correlated with the kinematics, specially the leptons p_T and aperture angle $|\Delta\Phi(\ell_1, \ell_2)|$ between the leptons. Since these variables are model-dependent (i.e. both the angular separation and lepton p_T are distributed differently for different SM processes) it may exist some differences between the different background sources considered in the search. Additional correlations

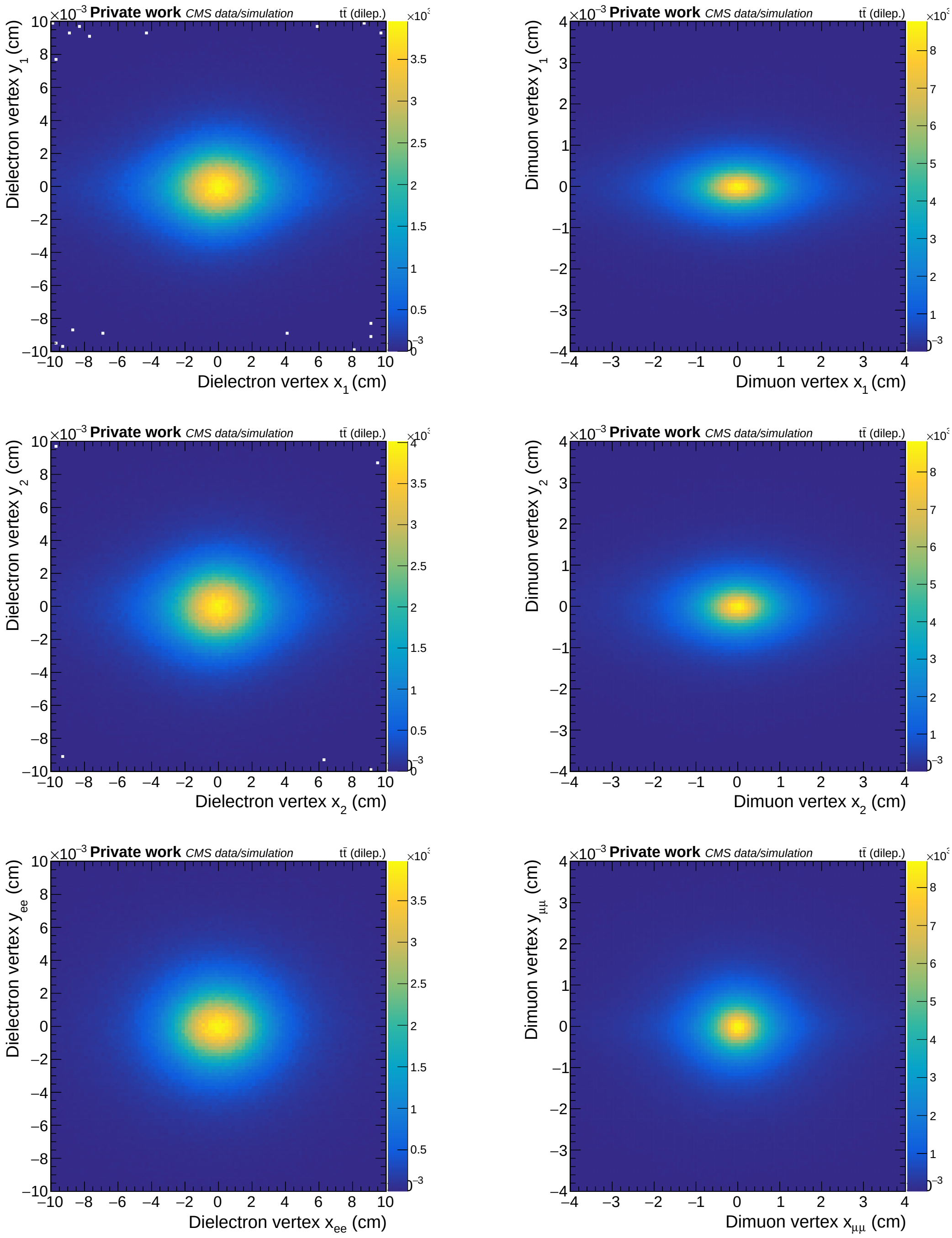


Figure 8.5: Fitted dielectron (left) and dimuon (right) vertices position from Monte Carlo simulated $t\bar{t}$ decays. Top: Using the high-p_T aligned coordinates defined by Eq.(8.2); Medium: Using the lower-p_T aligned coordinates defined by Eq.(8.3); Bottom: Using the dilepton aligned coordinates defined by Eq.(8.4).

will naturally arise if the leptons are related e.g. coming from the same particle decay as in Drell-Yan. Moreover, different cuts applied on electrons and muons may introduce differences in the collinearity shape between dielectron and dimuon channels as they can increase/decrease the difference in the p_T of the two tracks that are fitting the vertex.

The most representative case of this dependence if the one of the dimuon Drell-Yan $|\Delta\Phi|$ distribution in Figure 8.4 which is different from the others. The combination of two tracks with a similar p_T and the fact that great part of them will be back-to-back makes a small peak arise around $|\Delta\Phi| \sim \pi/2$. This effect is cross-checked by studying the $|\Delta\Phi|$ shape that is measured when fitting both electron and muon track pairs passing different p_T cuts. The $|\Delta\Phi|$ distribution of vertices with both leptons passing the same cut $p_T > 40$ GeV is compared with the distribution obtained when the leptons are requested to pass different p_T thresholds. In the later, the leading (subleading) lepton must fulfill $p_T > 40(25)$ GeV, emulating the selection of the dielectron channel. This comparison is shown in Figure 8.6 for 2018 Drell-Yan simulation.

As observed, the shape of the $|\Delta\Phi|$ distribution is different for the two situations considered. In both scenarios the leptons are mostly back-to-back, so the vertex will tend to be shifted towards the tracks direction. When the selection is asymmetric i.e. different p_T cuts for leading and subleading leptons, $|\Delta\Phi|$ is softly distributed and tends to populate more the 0 and π values. The dilepton momentum, which in this case is dominated by the leading lepton, is expected to point close to its track direction and somewhat collinear to the $\vec{DV}$ vector. However, when the selection is symmetric and both leptons are allowed to have the same momentum, the scenario changes and a peak at $\pi/2$ arises in the distribution. This peak comes from lepton tracks that are close to be back-to-back but also have the same momentum, therefore the dilepton $\vec{p}_T^{\ell\ell}$ vector points perpendicularly to the tracks direction. In this case $\vec{DV}$ direction is also perpendicular to $\vec{p}_T^{\ell\ell}$, and $|\Delta\Phi| \sim \pi/2$.

8.4. Background tests of collinearity symmetry

As the basis of the background estimation method is the the forward-backward symmetry of the collinearity distribution, this feature is studied through this section. The tests presented in this section evaluate the agreement between the $|\Delta\Phi|$ distribution and the distribution that results from applying a reflection with respect to $\pi/2$. The later will be referred as inverted-$|\Delta\Phi|$ distribution and is extracted from the $|\Delta\Phi|$ one by applying a reflection as:

$$\text{inverted-}|\Delta\Phi| = \pi - |\Delta\Phi|$$

If the background processes are forward-backward symmetric, the $|\Delta\Phi|$ and inverted-$|\Delta\Phi|$ distributions should be in agreement with each other.

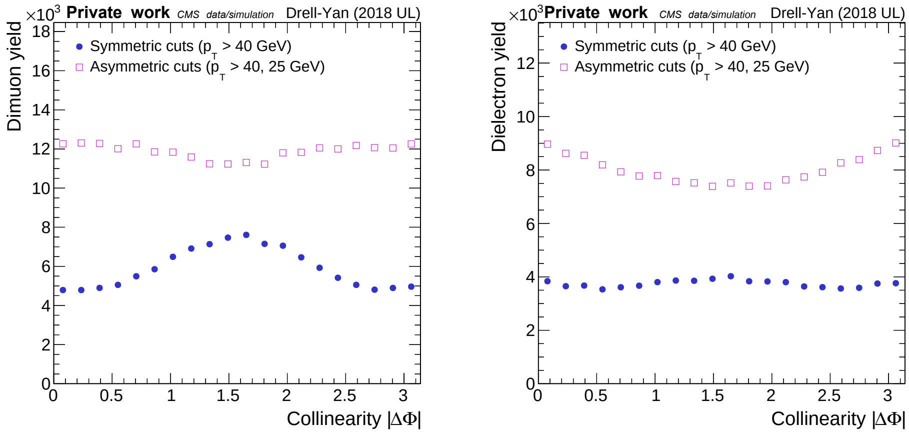


Figure 8.6: Collinearity distributions for dimuon (left) and dielectron (vertices) comparing two different p_T selection scenarios. In the symmetric scenario (solid circular markers), both leptons are required to pass $p_T > 40(25)$ GeV. In the asymmetric scenario (empty square markers) the leading lepton must fulfill $p_T > 40$ GeV and the subleading one $p_T > 25$ GeV.

The symmetric behavior is checked for the aforementioned background sources in their corresponding control regions. Both the collinearity and its inverted distributions are built for selected dilepton vertices passing the analysis baseline selection. An additional cut on the transverse impact parameter significance of $|d_0|/\sigma > 2.5$ is applied to better study the trends in processes that could have a certain displacement.

Tests done in Monte Carlo simulation for Drell-Yan, dileptonic $t\bar{t}$ and diboson processes are presented in Figure 8.7. No major disagreement is observed between the trends of the collinearity and inverted distribution proving the symmetric behavior of $|\Delta\Phi|$.

Tests done with data in the dedicated control regions are collected in Figure 8.8. Vertices contained in the W+jets enriched region again show good agreement. Deviations are observed for events in the QCD-enriched control region, where the inverted distribution seems to have more events, this effect being stronger for dielectron vertices. This would imply that these events, which fail the isolation cuts, tend to populate the backward region more than the forward region. In this case the estimated background yield would overestimate the background contribution in the signal region. The final comparison with the data, done in the Same-Sign control region, shows that this trend disappears for dielectron vertices and its effect is reduced for dimuon vertices.

In the light of the tests performed, the estimated background yield is assigned a systematic of 10%, corresponding to the largest deviation observed in the different background sources, i.e. that in the QCD enriched region. Since this deviation leads to a more conservative approach and is covered by the systematic,

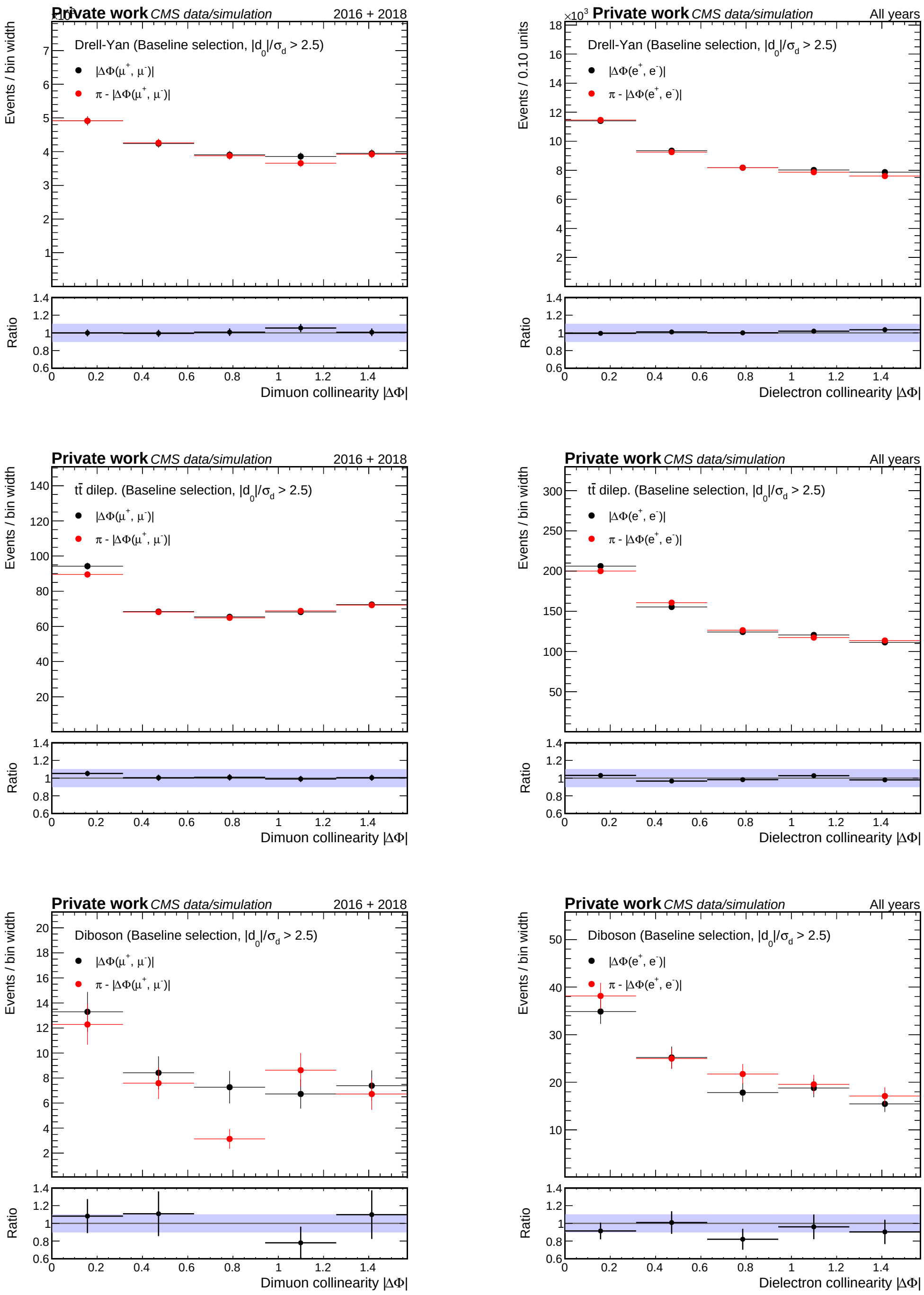


Figure 8.7: Symmetry tests performed in Drell-Yan (top), dileptonic $t\bar{t}$ (middle) and diboson (bottom) Monte Carlo simulation. The tests compare the behavior of the collinearity in black, with its inverted value in red, in dimuon (left) and dielectron (right) vertices. The grey band represents a deviation of 10%.

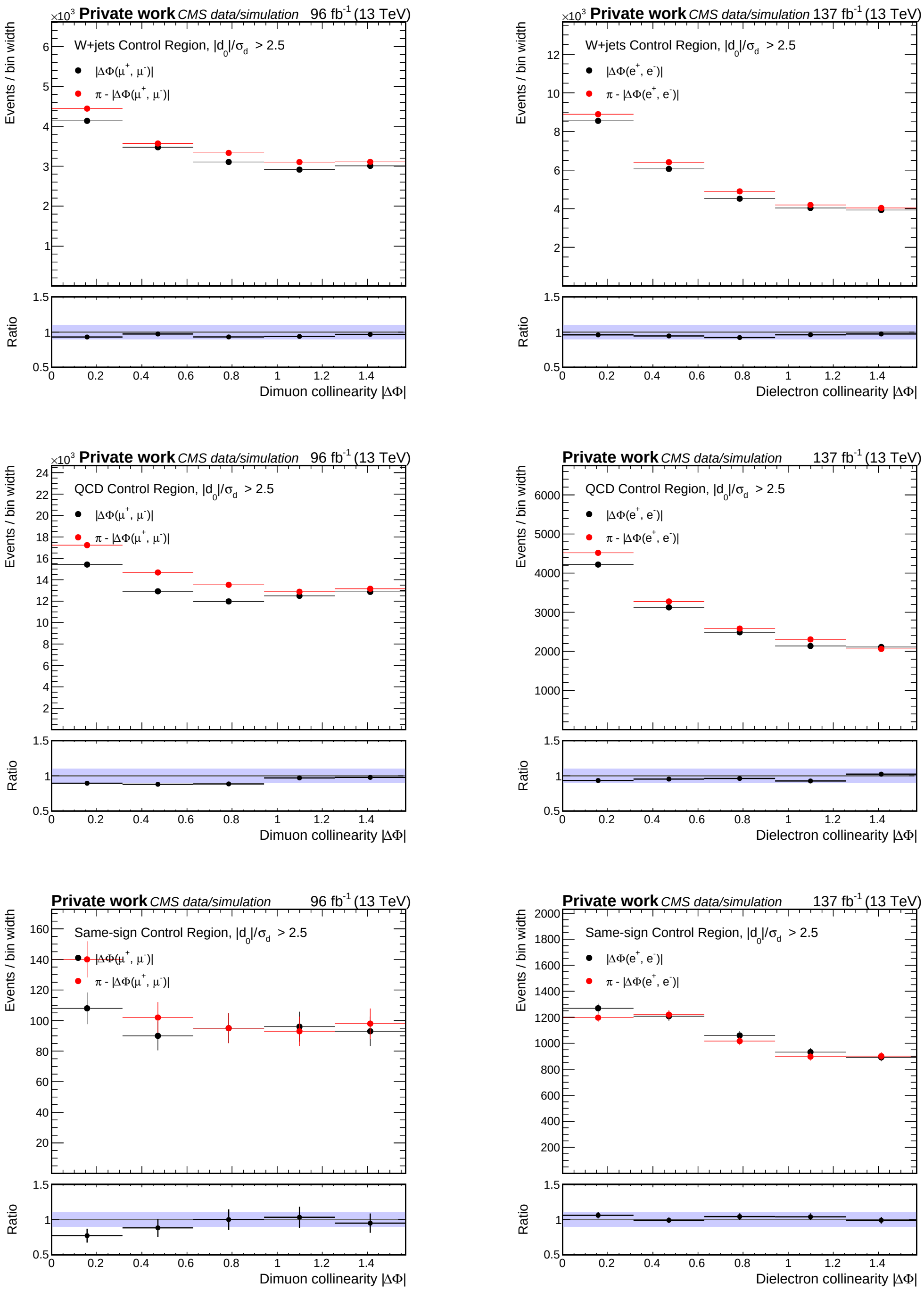


Figure 8.8: Symmetry tests performed in the W+jets (top), QCD (middle) and Same-Sign (bottom) data control regions. The tests compare the behavior of the collinearity in black, with its inverted value in red, in dimuon (left) and dielectron (right) vertices. The grey band represents a deviation of 10%.

no further action/correction is considered necessary to apply the background estimation method.

8.5. Background tests of collinearity symmetry at high displacement

The background estimation method is tested by checking the background yields in the Forward Region ($|\Delta\Phi| < \pi/2$), simplified as Signal Region (SR), and the Backward Control Region ($|\Delta\Phi| > \pi/2$), where background is estimated, at different displacement regimes. Asymmetries occurring at high displacement may be unnoticed when looking at the whole background distribution together as the background mainly populates the prompt regime. By making this comparison as a function of the displacement, we rule out the possibility of having a non-prompt contribution that could induce asymmetry in the collinearity distribution.

These tests compound the closure tests of the background estimation method. They are first made by comparing the ratio between the displaced vertices yield reconstructed in the SR and BCR regions as a function of the maximum transverse impact parameter significance $|d_0|/\sigma_{d_0}$ of the dilepton pair. Two comparisons are made in this study. The first one includes all dilepton vertices in the SR and BCR as they were defined previously i.e. it compares dilepton vertices with $|\Delta\Phi| < \pi/2$ vs dilepton vertices with $|\Delta\Phi| > \pi/2$. The different eras of the analysis are integrated in the comparison to profit from more statistics in the tails of the $|d_0|/\sigma_{d_0}$. The comparison is shown for dimuon vertices in Drell-Yan, $t\bar{t}$ and diboson simulations, together with the data in the W+jets and QCD control regions, in Figure 8.9. The equivalent plots for the dielectron vertices are collected in Figure 8.10. For completeness, the tests were also done by comparing the distribution of vertices with $|\Delta\Phi| < \pi/4$ against the distribution of vertices in the symmetric region $|\Delta\Phi| > 3\pi/4$, as this is the selection applied in some of the Signal Region sub-regions. The resulting plots are not included in this document as the conclusions were equivalent to the ones already presented and no significant effects manifested when restricting the $|\Delta\Phi|$ range.

8.6. Background estimation validation

The background estimation method is validated in validation regions where the presence of long-lived signals considered in this search is expected to be absent or negligible. We compare the observed data with dilepton vertices with $|\Delta\Phi| < \pi/2$ and the estimated background yield from dilepton vertices with $|\Delta\Phi| > \pi/2$. Since these regions are essentially signal-free, observed data must be background and should agree with our background estimation. The main objective is to validate the background estimation as a function of the displacement and the invariant mass of the leptons used to fit the vertex, as

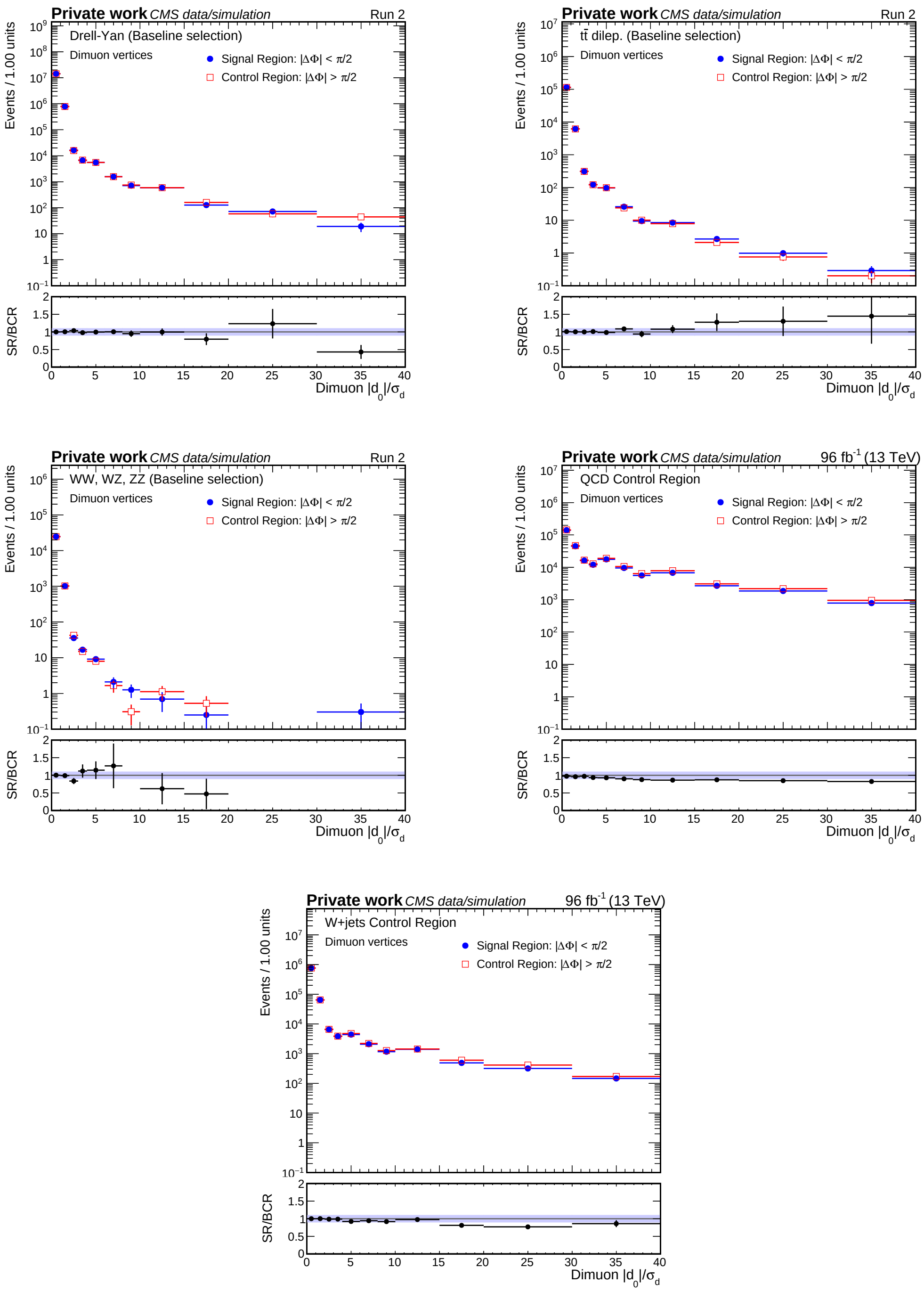


Figure 8.9: Closure tests of dimuon vertices for several background sources which include Drell-Yan, $t\bar{t}$, diboson, QCD and W+jets. The test compare the background yield for vertices with $|\Delta\Phi| < \pi/2$ and $|\Delta\Phi| > \pi/2$ in different $|d_0|/\sigma_{d_0}$ bins.

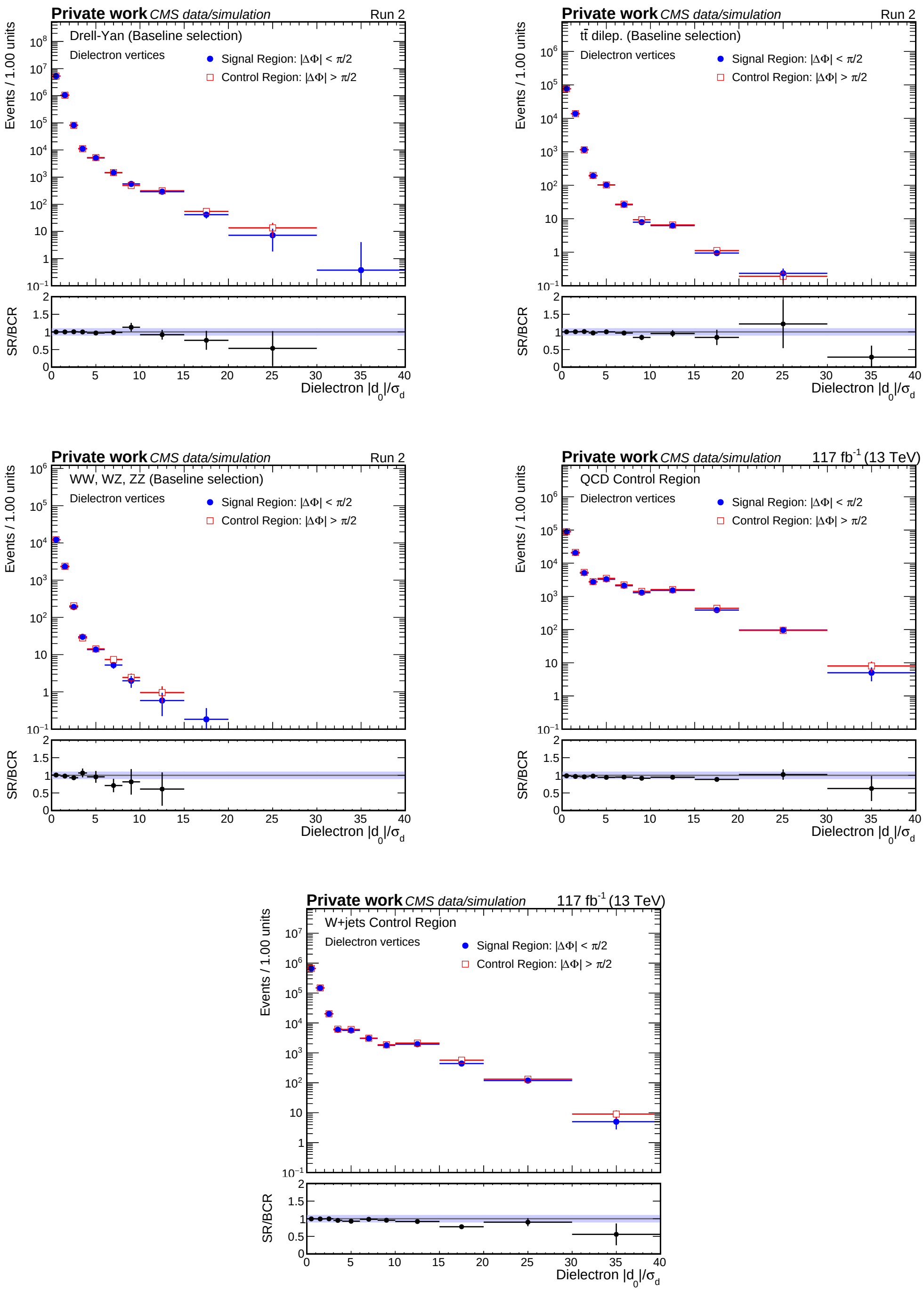


Figure 8.10: Closure tests of dielectron vertices for several background sources which include Drell-Yan, $t\bar{t}$, diboson, QCD and W+jets. The test compare the background yield for vertices with $|\Delta\Phi| < \pi/2$ and $|\Delta\Phi| > \pi/2$ in different $|d_0|/\sigma_{d_0}$ bins.

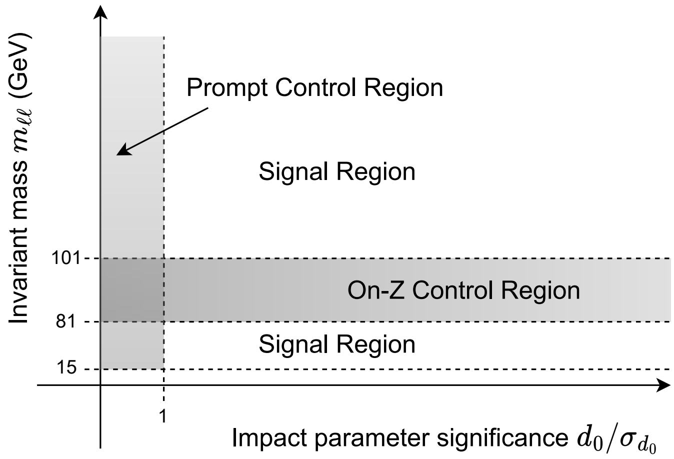


Figure 8.11: Schematic representation of the space defined by the transverse impact paremeter significance $|d_0|/\sigma_d < 1$ and the invariant mass $m_{\ell\ell}$ of the long-lived candidates. Prompt and On-Z regions are indicated, as well as the Signal Region of the search.

these are the variables that are used to split the signal region (see Section 7.3). Two distinct validation regions are defined to that purpose:

- **Prompt Validation Region**: Composed of events with long-lived candidates passing the baseline selection and fulfilling $|d_0|/\sigma_d < 1$ so they are not found much displaced from the PV. Additionally, events should have no more than one reconstructed vertex passing baseline selection. This region is expected to be enriched in SM prompt processes and allows to validate the agreement as a function of the displacement variables.
- **On-Z Validation Region**: Composed of events with long-lived candidates passing the baseline selection and falling inside the On-Z window. That is, these dilepton vertices should satisfy $|m_Z - m_{\ell\ell}| < 10$ GeV. This region is expected to be Drell-Yan enriched and since it does not have any requirements on displacement, allows to evaluate our background estimation method along the full range of displacement that may arise from resolution effects or PU. Events falling in this region are also required to have no more than one reconstructed vertex passing baseline selection.

The requirements on $|d_0|/\sigma_d$, $m_{\ell\ell}$ and the number of reconstructed vertices make these two regions orthogonal to the final Signal Region that will be used to evaluate the signal presence. An schematic representation of the phase space with the Prompt and On-Z control regions is shown in Figure 8.11.

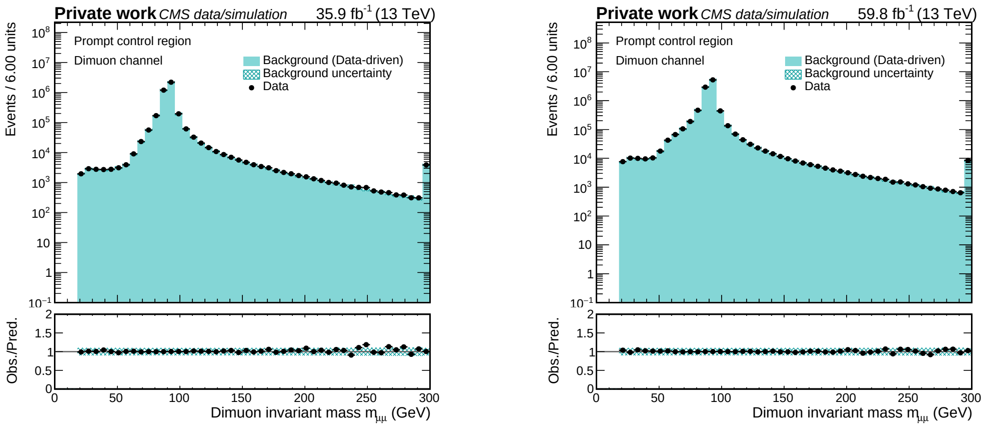


Figure 8.12: Background validation in Prompt Control Region for invariant mass spectra of dimuon long-lived candidates.

8.6.1. Validation in the Prompt Region

Agreement between observed background and estimated background is studied in the Prompt Validation Region. In this validation we compare the yields of the selected long-lived candidates with $|d_0|/\sigma_d < 1$ in events with only one dilepton vertex passing the baseline selection that are reconstructed with $|\Delta\Phi| < \pi/2$ and $|\Delta\Phi| > \pi/2$. The observable that is studied is the invariant mass of the candidates. The invariant mass serves to define the other On-Z validation region and also will be used to define the binning in the Signal Region. In this way, we cross-check the background validation and the Signal Region definition.

The validation is shown in Figure 8.12 for selected dimuon candidates in this region for 2016 and 2018 data. Both 2016 and 2018 show good agreement over the whole invariant mass range. Deviations due to statistic fluctuations are found in the tails of the mass distributions over $m_{\mu\mu} > 250$ GeV where the number of background events decreases.

Invariant mass spectra of observed and estimated background is shown in Figure 8.13 for dielectron long-lived candidates in 2016, 2017 and 2018 data separately. As in the dimuon vertices validation, there are fluctuations appearing over $m_{ee} > 250$ GeV where statistics is limited. Despite this, good agreement between observed and estimated background is found in both 2016 and 2018 data.

On the other hand, dielectron candidates show a deviation at low mass in 2017 data, where we have also removed the $m_{ee} > 100$ GeV requirequirement of the 2017 selection, just to ilustrate the behavior below the trigger threshold. The estimated background yield increases for masses below $m_{ee} < 90$ GeV producing a discrepancy of the 20% with the observed background. Then, the estimated dielectron vertices slightly increase from 90 GeV up to 100 GeV, when the

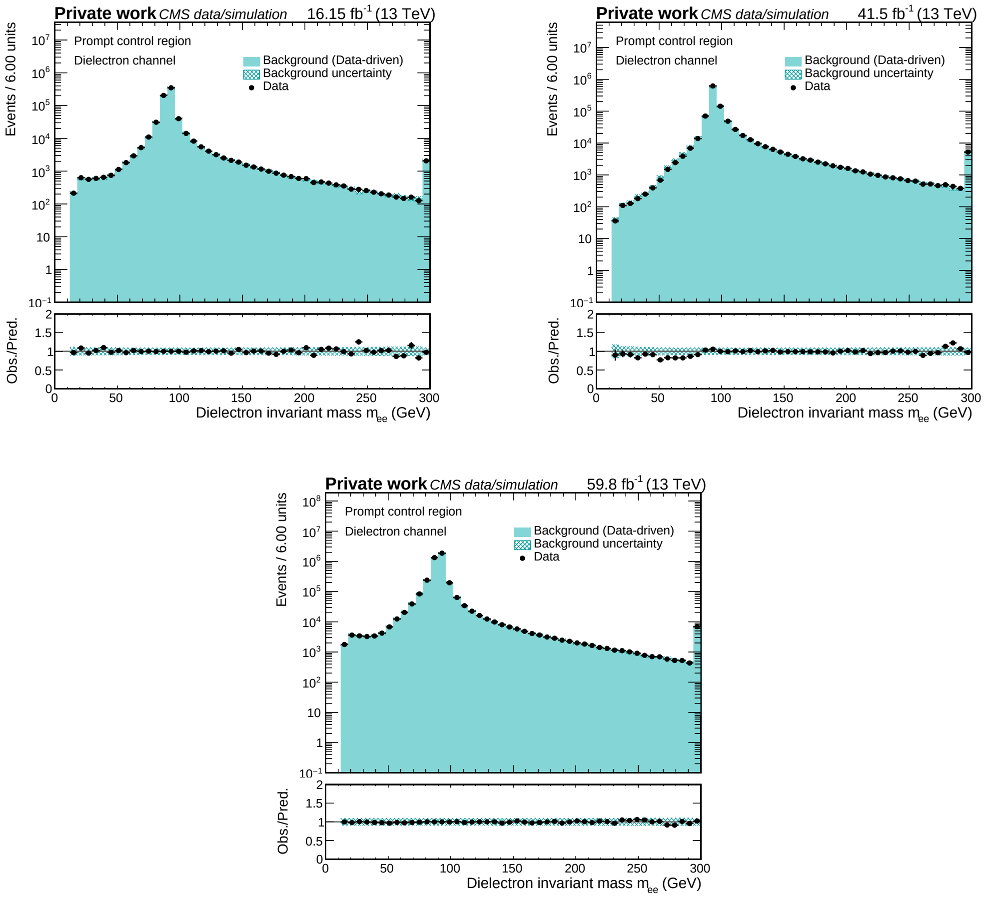


Figure 8.13: Background validation in Prompt Control Region for invariant mass spectra of dielectron long-lived candidates.

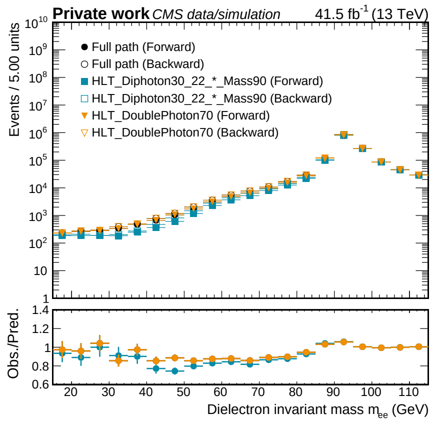


Figure 8.14: Background validation in Prompt Control Region for invariant mass spectra of dielectron long-lived candidates for the individual trigger paths of 2017.

agreement is finally recovered. For this reason, the event selection described in Chapter 7 includes a cut of $m_{ee} > 100$ GeV for 2017 dielectrons, opposed to the $m_{ee} > 15$ GeV cut for 2016 and 2018.

Since the main difference between 2017 data and the other years is the trigger path, the behavior of the observed and estimated background is studied as a function of this trigger choice. The discrepancy is likely to come from the diphoton trigger that is used in 2017. The diphoton trigger path in 2017 is the OR combination of two different trigger paths:

- `HLT_Diphoton30_22_R9Id_OR_IsoCaloId_AND_HE_R9Id_Mass90`
- `HLT_DoublePhoton70`

Both of them impose direct and indirect requirements in the invariant mass of the two photons. The first path requires $m_{\gamma\gamma} > 90$ GeV while the second sets a tight $E_T > 70$ GeV requirement in the two photons that would strongly suppress low mass diphoton pairs. This suppression manifests in the plots of Figure 8.13 where it can be seen how the background yields at low mass decrease by almost an order of magnitud below 90 GeV. The discrepancy occurs at this low mass range that is suppressed by the trigger.

To check the effect of the three paths separately, the observed and estimated background yields are compared for events fired for each one of them.

The `DoublePhoton70` and `Diphoton*Mass90` paths have very similar behavior at low mass. From their distributions we can see that although the selected dielectron vertex has invariant mass below 90 GeV, the two paths seem to recover some of the events below this value. This is likely due to the fact that even if we are selecting two clusters to reconstruct the dielectron vertex there may be other photons in the event that may fire the diphoton triggers. Events falling at this low mass values have probably fired the trigger with a photon out of the selected pair.

The background estimation method relies on the assumption that we should have the same background dielectron vertices reconstructed with $|\Delta\Phi| < \pi/2$ and $|\Delta\Phi| > \pi/2$. Since the events with dielectron vertices that are included in the analysis may have triggered with a photon independent of this $|\Delta\Phi|$ measurement there may be some external factors that affect this agreement.

From the studies presented above, it can be concluded that the background estimation is not biased in the invariant mass regimes that are considered per year and channel.

8.6.2. Validation in the On-Z Region

Validation in the On-Z region allows to evaluate the agreement between observed background and the predictions as a function of the distance variables that will be used to define the Signal Region of the search. The On-Z region is expected to be enriched in Drell-Yan events, which is the main instrumental background due to its large cross section and used as a representative background in this validation.

The agreement is studied as a function of distance variables derived from the dilepton vertices that are used to define the Signal Region of the search, and therefore, provide the main signal vs background discrimination. These variables were introduced in Section 5 and are the transverse impact parameter significance $|d_0|/\sigma_{d_0}$ assigned to the dilepton vertices and the the transverse distance between the PV and the vertex L_{xy}.

The comparison done for the transversal distance L_{xy} is shown in Figure 8.15 for dimuon vertices and in Figure 8.16 for dielectron vertices.

The validation is performed as a function of the transverse impact parameter significance $|d_0|/\sigma_d$ and included in Figures 8.17 and 8.18 for dimuon and dielectron vertices.

The plots do not show any systematic deviation that could indicate that the background events in the studied regions. Therefore, it can be concluded that there is no significative effect as a function of the displacement that may invalidate the methodology follow to estimate the background.

8.7. Background validation in Signal Region

Background estimation method should be tested in the final Signal Region, considering the splitting described in Table 7.2. This is done by comparing the background yields in each sub-region of the the Signal Region and the corresponding mirrored sub-region of the Background Control Region (by inverting

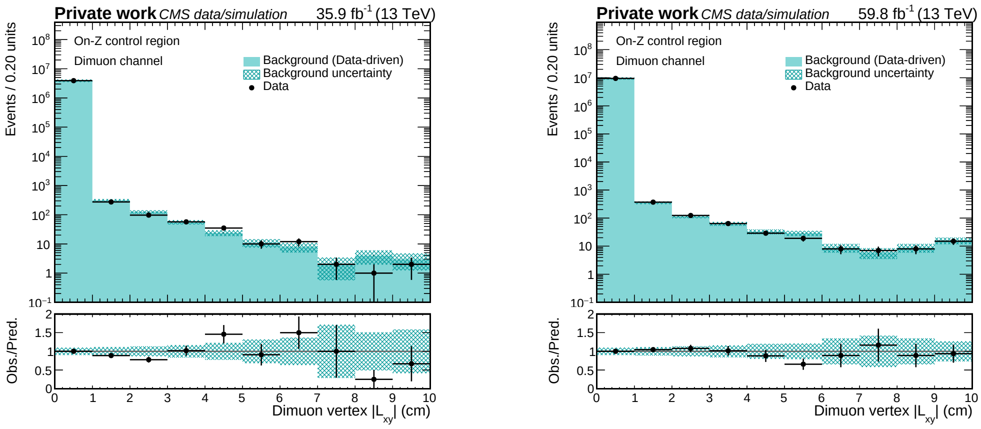


Figure 8.15: Background validation in On-Z Control Region for vertex distance L_{xy} of dimuon long-lived candidates.

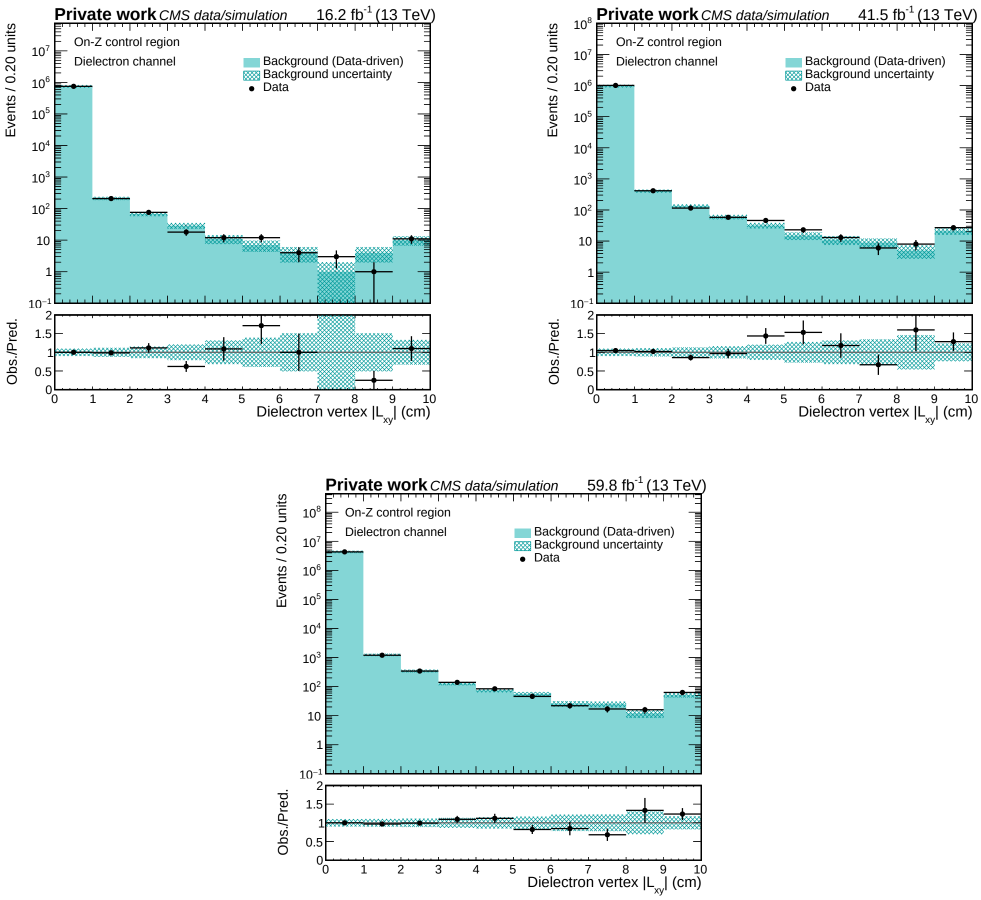


Figure 8.16: Background validation in On-Z Control Region for vertex distance L_{xy} of dielectron long-lived candidates.

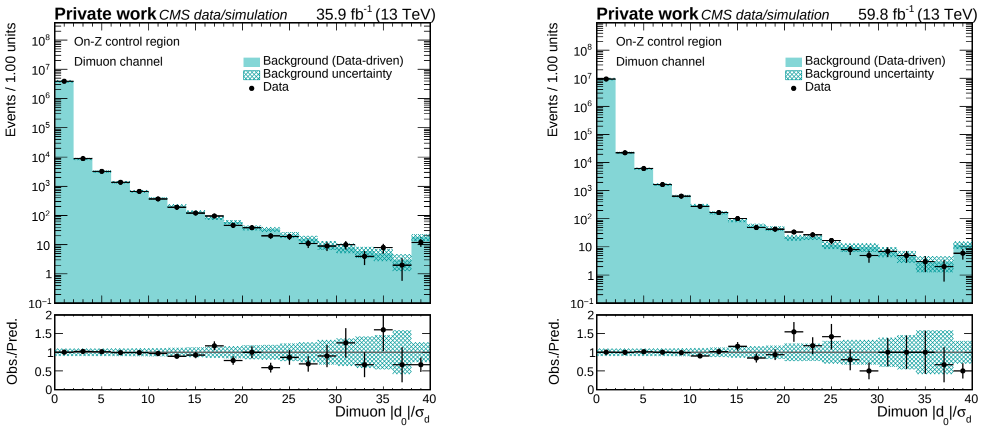


Figure 8.17: Background validation in On-Z Control Region for the maximum transverse impact parameter $|d_0|/\sigma_d$ of dimuon long-lived candidates.

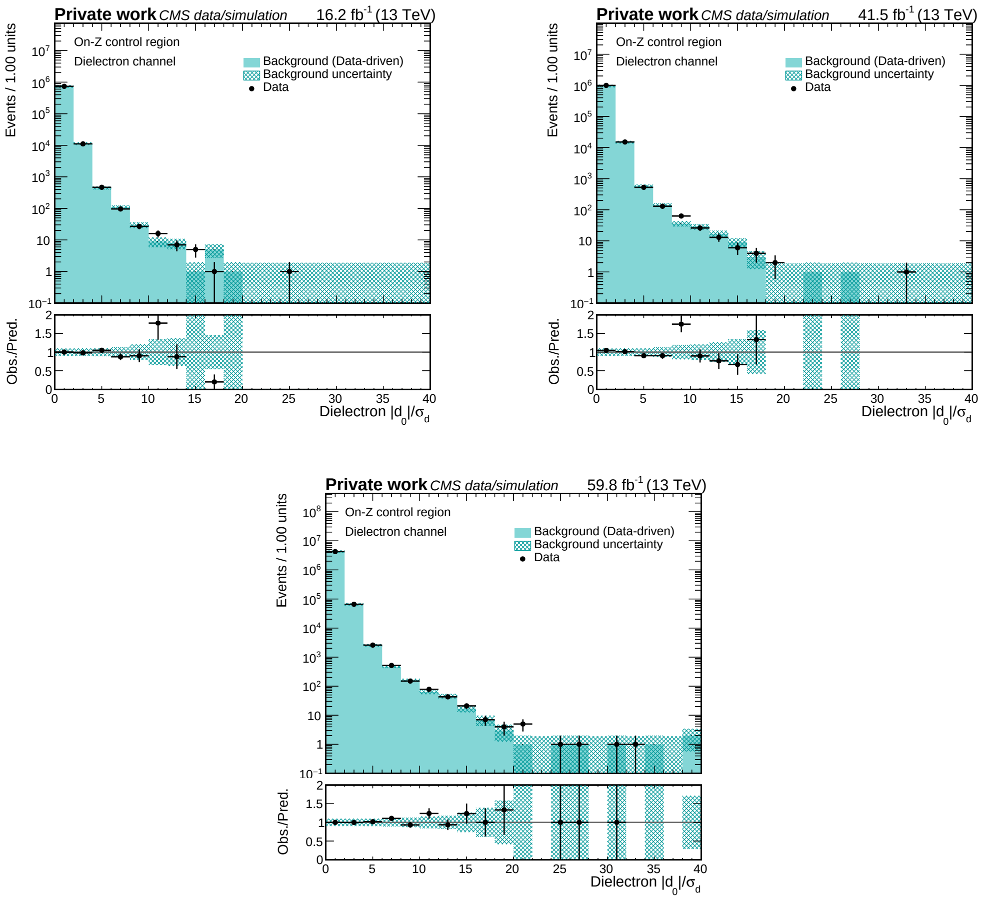


Figure 8.18: Background validation in On-Z Control Region for the maximum transverse impact parameter $|d_0|/\sigma_d$ of dielectron long-lived candidates.

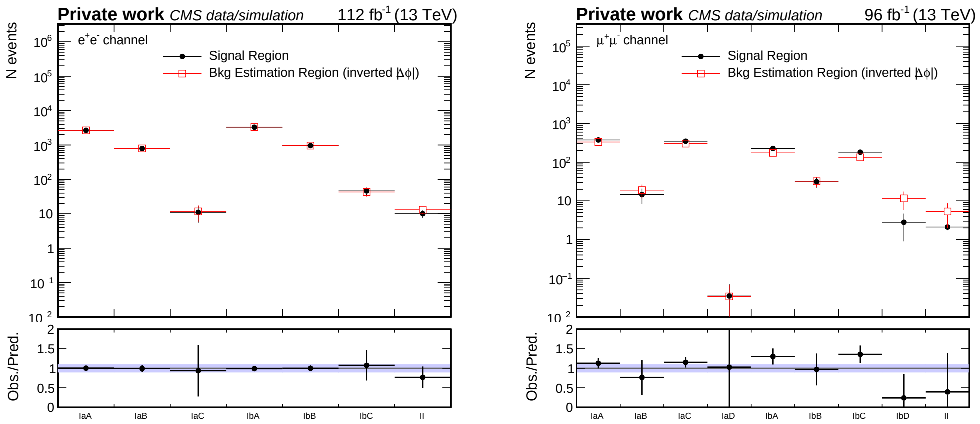


Figure 8.19: Validation tests for background estimation method for dielectron (left) and dimuon (right) vertices. SR and CR background yields comparison in each splitting region. Grey band represents the systematic uncertainty in the background estimation.

the selection on $|\Delta\Phi|$). However, there is a lack of background events in some of these sub-regions, since they are optimized to have maximal sensitivity. Particularly, the main source of background that is missing is Drell-Yan events, as the splitting in the SR rejects the Z peak. For these reason, in order to increase the statistics, the tests are performed using the same splitting that is described in Section 7.3 but removing the mass restrictions. This can be done without biasing the background estimation method, since the symmetrical behaviour of the collinearity with respect to $\pi/2$ does not depend either on the mass or the displacement, as it was studied in Section 8.6.1.

The first set of tests are performed using Monte Carlo simulation events, as this study was done in the analysis stages when data is still blinded. Monte Carlo samples used are the ones described in Section 6.2. Figure 8.19 shows the background yields for each subregion of the SR, and the corresponding mirror subregion in the CR which is defined by inverting the $|\Delta\Phi|$ selection that is applied in each case. The plots combine all years (2016 postVFP, 2017 and 2018 for dielectrons; 2016 and 2018 for dimuons) to increase the statistics in each region.

The agreement is almost optimal for dielectron vertices. For dimuon vertices, we can observe a lower agreement in regions IbD and II. This is because the measurement in these regions is based on a very few amount of Drell-Yan MC events (4 for IbD and 1 for II).

The second set of tests are devoted to study QCD-like events in the different signal region bins. In order to study this contribution, the signal region bins defined in Table 7.2 are filled with events where the LLP fails the isolation cut in both leptons (all the other requirements of the vertex selection are ap-

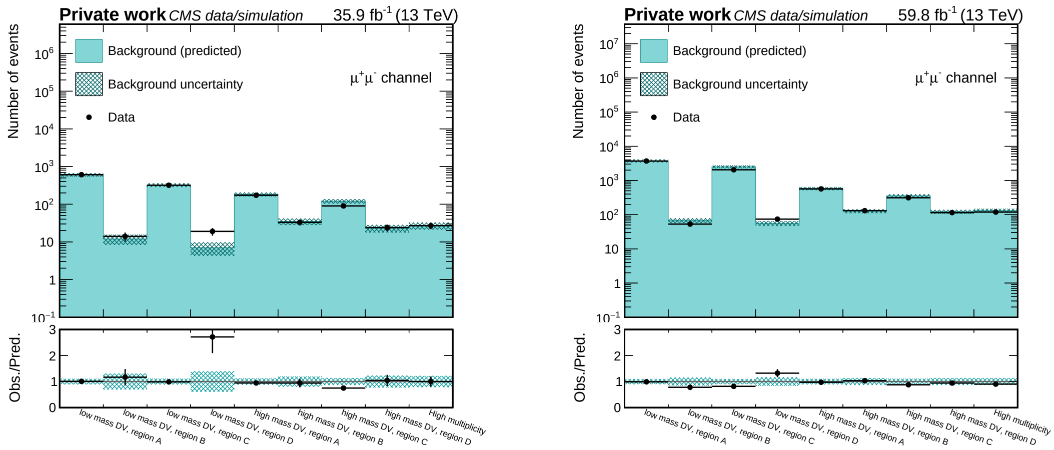


Figure 8.20: Observed and predicted yields for QCD-like dimuon events in the signal region bins. These events are defined by vertices splitted in sub-regions as defined in Table 7.2 which fail the isolation requirements imposed on the muons. Predicted yields are measured in the $|\Delta\Phi|$-mirrored sub-regions.

plied). The obtained yields are compared with the ones measured in the $|\Delta\Phi|$-mirrored regions with the same selection. The results are shown per channel and studied years in Figures 8.20 and 8.21. The predicted and observed yields for QCD-like events are in agreement in all the sub-regions considered for the electron channel. For the muon channel, a 2-σ deviation is found in the low mass region D (SRIaD), only for 2016 data. No such deviation is observed for the equivalent region when studying 2018 dimuon data. A deviation of this magnitude could be expected when considering 36 orthogonal regions as it is the case of this analysis.

This effect has been investigated further by repeating the same comparison with a subset of QCD-like events with properties more similar to the ones expected in the signal region. To do so, instead of completely inverting the isolation requirements on the leptons, an upper cut on the isolation has been required. In this subset, the considered QCD-like vertices are fitted from leptons failing the isolation requirements but with relative isolation values below 1.5. In this way, events falling in the tails of the isolation distribution are removed from the comparison. Results are shown in Figures 8.22 and 8.23, where it can be seen that the fluctuation in the SRIaD bin for 2016 is no longer visible.

8.8. Displaced background rejection

8.8.1. Cosmic muons

Cosmic muons are muons that pass through the entire CMS detector. One single cosmic muon could be reconstructed by the dGlobal algorithm in both

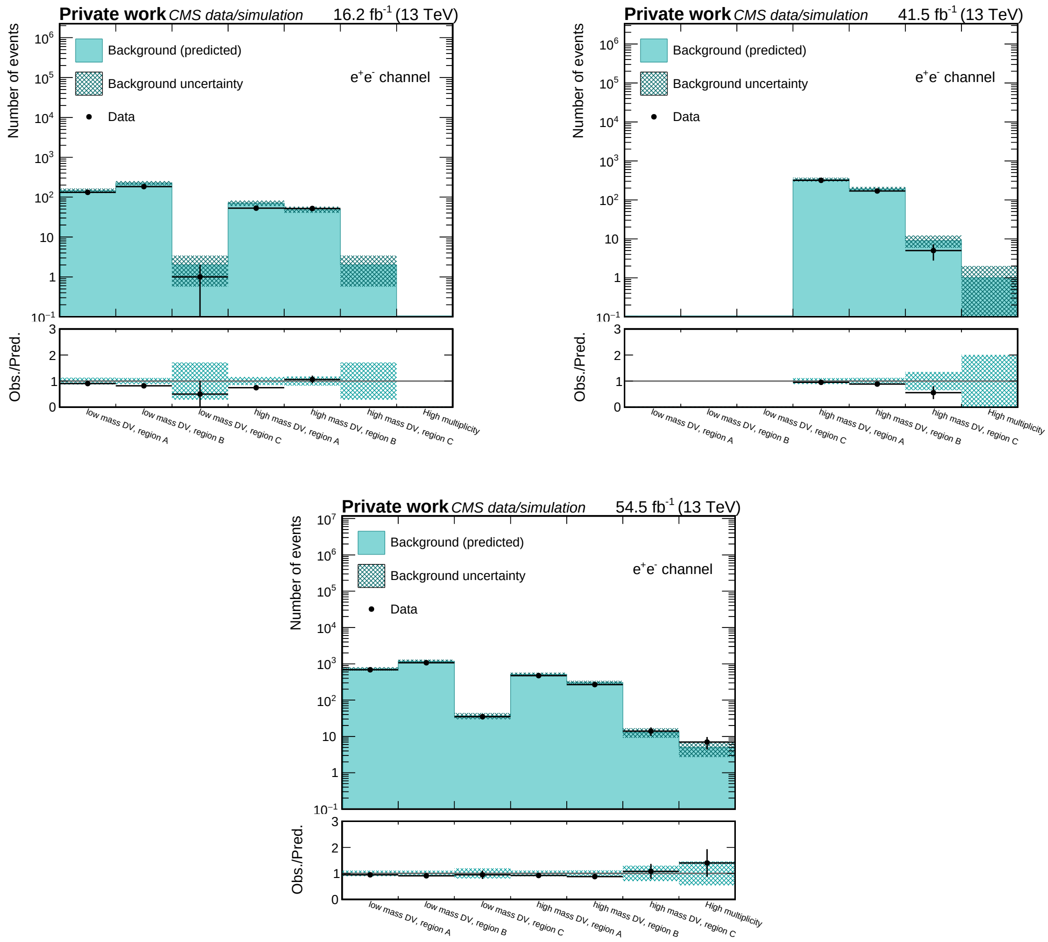


Figure 8.21: Observed and predicted yields for QCD-like dielectron events in the signal region bins. These events are defined by vertices splitted in sub-regions as defined in Table 7.2 which fail the isolation requirements imposed on the electrons. Predicted yields are measured in the $|\Delta\Phi|$-mirrored sub-regions.

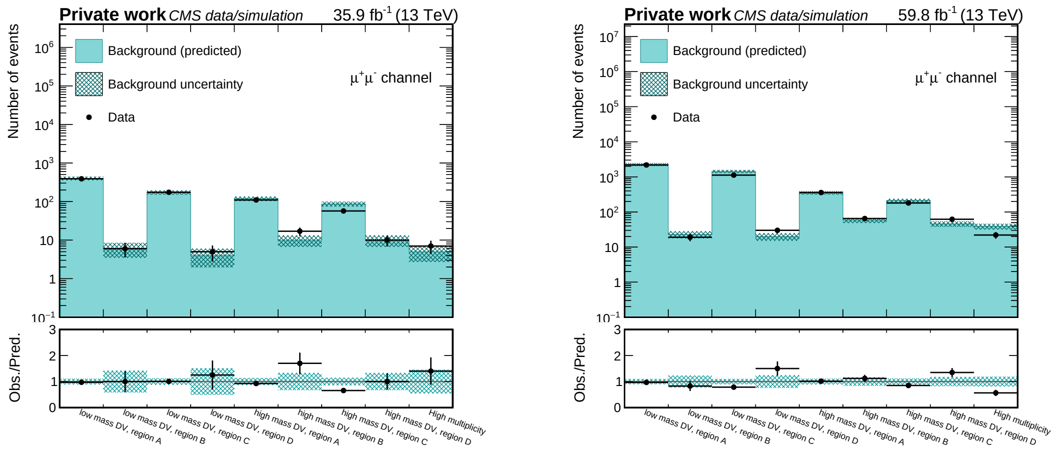


Figure 8.22: Observed and predicted yields for QCD-like dimuon events in the signal region bins. These events are defined by vertices splitted in sub-regions as defined in Table 7.2 which fail the isolation requirements imposed on the muons but with an upper cut on the relative isolation of 1.5. Predicted yields are measured in the $|\Delta\Phi|$-mirrored sub-regions.

hemispheres as two separate back-to-back dGlobal muons. It may happen that the tracks of these two muons could be extrapolated and fitted to a secondary vertex that ultimately could also be displaced and counted in the signal region. Since the two fitted muons are not two genuine muons, the behavior of the collinearity is not physical and difficult to predict in such vertices. This cosmic vertices could introduce an asymmetry in the collinearity distribution that would make this background source impossible to predict with our background estimation method. Instead, the $\cos(\alpha_{\mu\mu}) > -0.8$ (-0.9) requirements applied in 2016 (2018) in the dimuon selection are proven to be enough to reject the vertices produced from cosmic muons.

In this section, an evaluation is conducted on the impact of the cosmic background in a sample of cosmic muons. The focus of these studies is on the examination of the displaced vertices that could be reconstructed from this particular sample. Cosmic muons are taken from the `NoBPTX` dataset, which contains cosmic rays from the dedicated cosmic runs (`Cosmics` dataset) and adittional cosmic rays from *Parasitic cosmic runs*, collected during pp collision running. Events are triggered with `HLT_L2Mu10_NoVertex_NoBPTX3BX` trigger, which selects events out of the pp bunch crossing.

The studied sample contains a total of 65886568 events from which 26692600 of them pass the trigger. 35178 of these events have one secondary vertex fitted from two separate dGlobal tracks passing the preselection and ID cuts. Finally, 32592 of these vertices pass the selection criteria defined in the event selection of the analysis, defined in Section 7.2, except from the one in the collinearity (not applied to check the collinearity behavior) and the $\cos(\alpha_{\mu\mu})$ require-

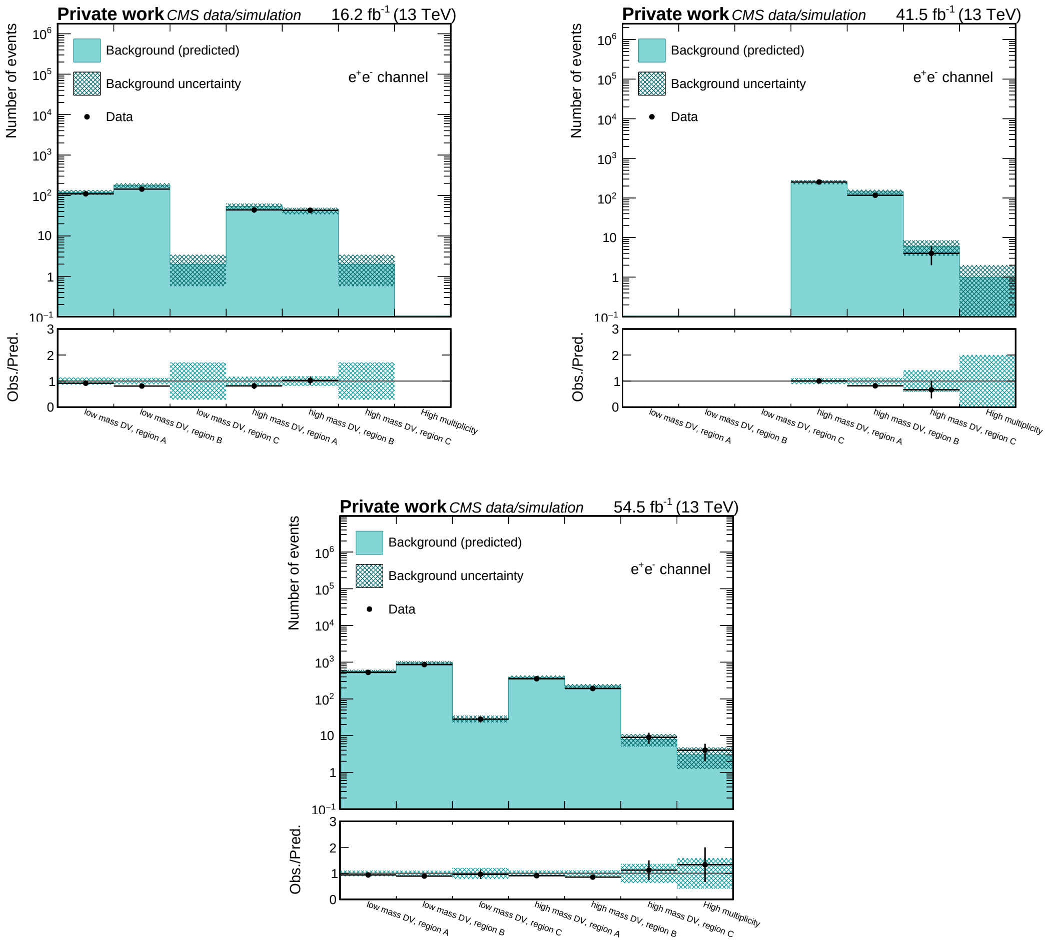


Figure 8.23: Observed and predicted yields for QCD-like dielectron events in the signal region bins. These events are defined by vertices splitted in sub-regions as defined in Table 7.2 which fail the isolation requirements imposed on the electrons but with an upper cut on the relative tracker isolation of 1.5. Predicted yields are measured in the $|\Delta\Phi|$-mirrored sub-regions.

ment, which is evaluated. The distributions of these selected cosmic vertices are shown in Figure 8.24.

The collinearity of the cosmic vertices breaks the forward-backward symmetry, making this source to be unpredictable by our estimation method. The opening 3D angle $\alpha_{\mu\mu}$ between the two dGlobal legs that are fitted to the vertices is also shown, where it can be checked that the tracks used are back-to-back tracks having angle values close to $\sim \pi$. The $\cos(\alpha_{\mu\mu})$ that results from those tracks is shown in the last subfigure. This last distribution has a hard peak at $\cos(\alpha_{\mu\mu}) \sim -1$ as it would be expected from cosmic vertices. From all studied vertices, only 6 are found to fail the cut $\cos(\alpha_{\mu\mu}) > -0.8$ and just 7 vertices fail the loose $\cos(\alpha_{\mu\mu}) > -0.9$ cut applied in 2018. From these numbers the fraction of cosmic muons that would be included in the Signal Region is estimated to be at most of 0.02% in both years. It is concluded by means of this study that the cosmic vertices background is negligible in comparison to the estimated instrumental background.

8.8.2. Nuclear interactions

As they travel through the detector, the different particles generated in the collisions can interact with the material present in the different CMS detectors and substructures. This process can give rise to fake displaced vertices that contribute to the background of the analysis.

The evaluation of the impact of nuclear interactions was done in a sample of data events collected in the On-Z validation region presented in Section 8.6.2 are analyzed.

The distribution of the primary vertex displacement in the xy plane was studied with the objective of identifying a potential excess of vertices at any particular value of displacement. These plots are shown in Figures 8.25 and 8.26.

The most probable locations to have an excess of vertices due to nuclear interactions are those that correspond to the main structures of the inner tracking systems, where the density of material is high. These locations are remarked in Figures 8.25 and 8.26, and are the following:

- Beam pipe: $r_{\text{bp}} = 2.2 \pm 0.2$ cm.
- Inner BPIX detector (first layer): $r_{\text{BPIX1}} = 2.95 \pm 0.15$ cm.
- Inner BPIX detector (second layer): $r_{\text{BPIX2}} = 6.8 \pm 0.2$ cm.
- Pixel detector support tube: $r_{\text{st}} = 22.0 \pm 0.5$ cm.

It can be observed that there are no effects at this particular radius values. Therefore, it is concluded that the effect of nuclear interactions is negligible and no further actions are taken regarding this background.

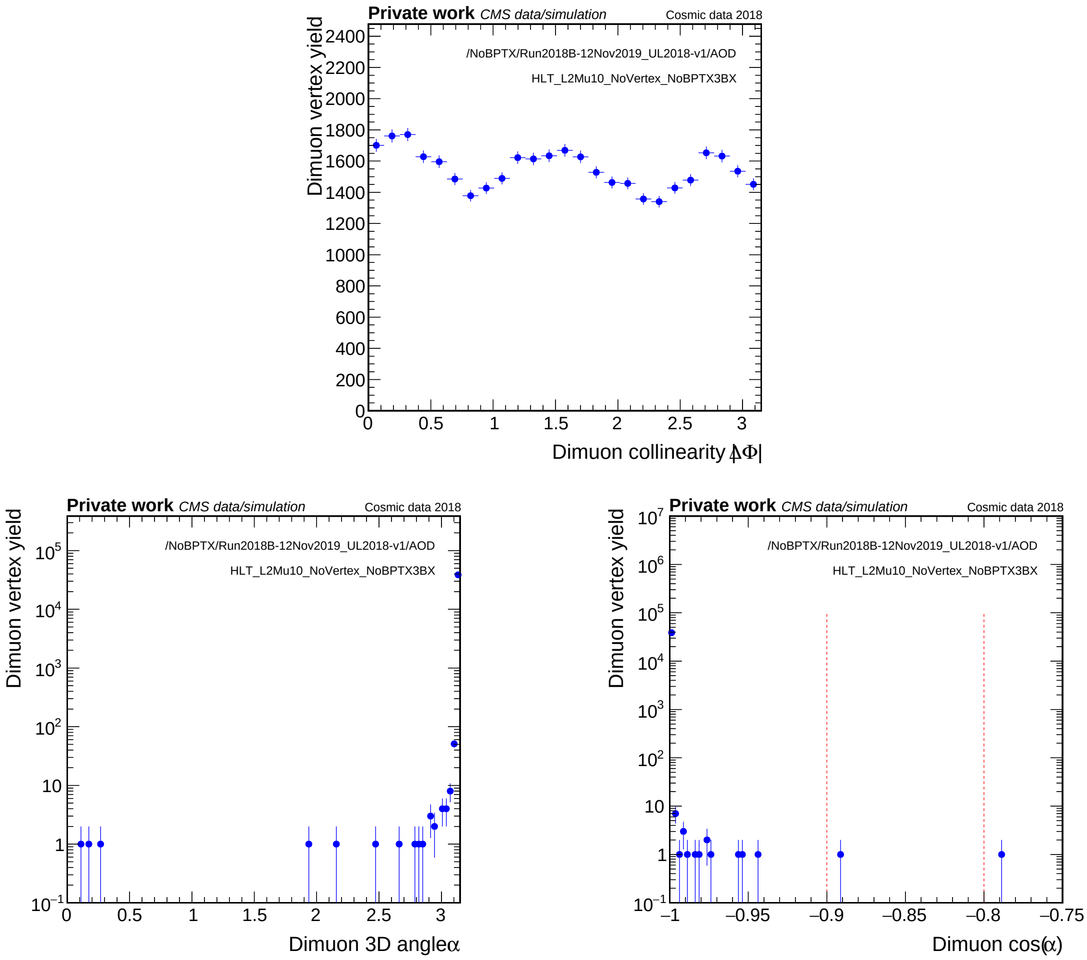


Figure 8.24: Distributions of cosmic vertices reconstructed in the `NoBPTX` dataset. Left: Collinearity distribution, middle: 3D opening angle $\alpha_{\mu\mu}$ between dGlobal tracks, right: $cos(\alpha_{\mu\mu})$ distribution

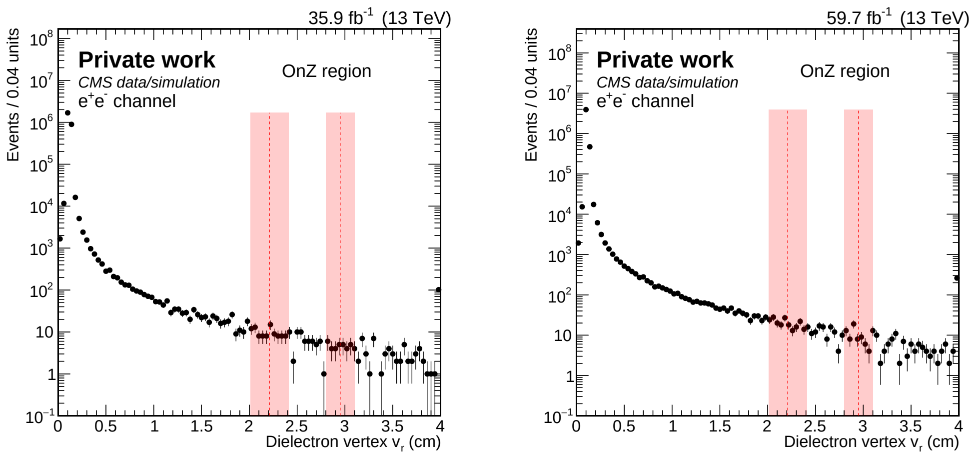


Figure 8.25: Distribution of dielectron vertex displacement for data in the On-Z region. The red bars correspond to the position of the main structures of the inner tracking system.

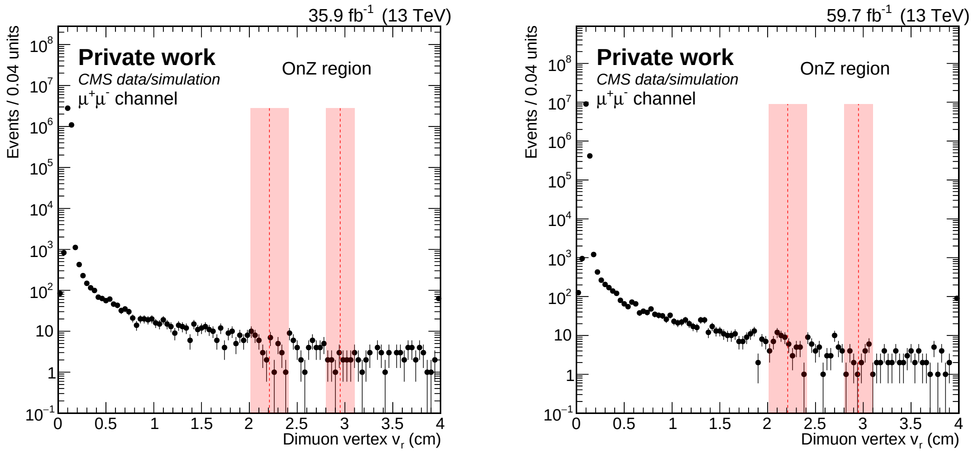


Figure 8.26: Distribution of dimuon vertex displacement for data in the On-Z region. The red bars correspond to the position of the main structures of the inner tracking system.

Chapter 9

Corrections to simulation and systematic uncertainties

Although the Monte Carlo simulation is optimized to reproduce not only the expected physics behavior, but also the expected detector response, there are several aspects of data which may be not modelled with enough precision. This is why some corrections must be applied to the simulation to accurately emulate the behavior observed in data events. The corrections generally affect the expected object yield in simulation and are commonly referred as *scale factors* (SF). In general, a SF is derived for every source of discrepancy that may exist. The methodology which was followed to derive the specific ones for the search presented in this thesis is described in the first part of this chapter.

The simulated events are first weighted by the luminosity weight w_{lumi} which normalizes the generated event yield N_{gen} to the one expected for a given integrated luminosity L in a process with cross section σ:

$$w_{\text{lumi}} = \frac{\sigma \cdot L}{N_{\text{gen}}} w_{\text{gen}} \tag{9.1}$$

where w_{gen} is an additional weight to simulate the presence of interference processes in NLO samples. For LO samples it it set to $w_{\text{gen}} = 1$.

Then, there is a set of corrections applied to the simulation that includes:

- A correction w_{PU}, called pileup re-weighting, to account for pileup modelling which is applied event-by-event (Section 9.1).
- Scale factors $\text{SF}_{\text{reco+id}}$ to account for different displaced lepton reconstruction and identification performance in data and simulation. They are applied per reconstructed lepton used to fit any vertex that pass the preselection and it is counted in N_{DV} (Section 9.3).
- Scale factors SF_{iso} to account for isolation performace. They are applied per lepton used to fit the selected displaced vertex of the event i.e. the LLP

candidate, as these are the only leptons required to pass isolation (Section 9.4).

- Scale factors $\mathrm{SF}_{\mathrm{trigger}}$ for the trigger performance which are applied to the events that have fired the analysis triggers (Section 9.5). It is derived on top of events where the presence of reconstructed isolated leptons is already required. In this way, the effects introduced by offline reconstruction or isolation performance are already absorbed by the lepton corrections described before.

The simulated events are weighted with the aforementioned corrections which are applied jointly to the events:

$$w_{\mathrm{total}} = w_{\mathrm{lumi}} \cdot w_{\mathrm{PU}} \cdot \mathrm{SF}_{\mathrm{reco+id}} \cdot \mathrm{SF}_{\mathrm{iso}} \cdot \mathrm{SF}_{\mathrm{trigger}} \tag{9.2}$$

The adjusted signal yield aims to closely approximate the expected yield that would be observed in the detector for a hypothetical signal. However, this yield inherently includes uncertainties that arise from statistical fluctuations as well as systematic effects that can potentially alter the predicted signal. This is why systematic uncertainties are estimated to account for the different sources of uncertainty in the measurement of the signal yields. They are all described over the following sections and finally summarized at the end of this chapter.

9.1. Pile-up

The Monte Carlo simulated samples were produced with a certain pileup profile that does not have to perfectly match the one measured in real data. We account for the differences between the distribution of pileup interactions in data and simulation by reweighting the latest. The pileup weight w_{PU} is applied event-by-event as a function of the number of reconstructed PU vertices N_{PU}. To derive this correction the pileup distribution in data is divided by the one used to generate our simulated samples. The CMS standard central `pileupCalc.py` script was used for the calculation. The certified runs were read from the central GOLDEN JSON files, while the pileup profile was taken from the last updated pileup json files. The recommended minimum bias cross section was 69200 pb with an uncertainty of 5%.

Uncertainty on the pile-up reweighting for the signal samples has been obtained by applying the standard recipe: pile-up weights are recalculated with a variation of the minimum bias cross-section of $\pm 5\%$. Several distributions of key observables of the search have been drawn and the differences, of the order of 2% have been used as systematic uncertainty. This value is treated uncorrelated for the different years.

9.2. Luminosity

The systematic uncertainty on the integrated luminosity has been estimated according to the standard recommendations of the CMS luminosity group. Uncertainties for the different years are treated as uncorrelated with values 1.2%, 2.3% and 2.5% for 2016, 2017 and 2018.

9.3. Calibration of displaced lepton reconstruction and identification

The displaced lepton algorithms reconstruct leptons which were produced at some distance from the interaction point of the collision. However, although the studies performed on signal simulation clearly indicate that their use is suitable in a search for LLPs decaying to leptons, the differences that may exist between the real experiment conditions and the simulation may translate into a different performance between data and simulation. For this reason, these physics objects have to be properly calibrated with real data to be used effectively in any data analysis. This calibration implies the derivation of any correction that should be applied to the simulated objects for them to reproduce the behavior of the ones reconstructed in data. This section provides the needed calibrations to assess differences arising in the lepton reconstruction and/or identification efficiencies.

While in simulation the performance of the reconstructed objects is easily compared with the expected behavior provided by generation information, in data it is required a way to identify real physical objects on which the efficiency is to be measured. The technique applied to measure the object efficiency in data is called *tag-and-probe* (TnP) and a brief introduction of the basis of the method will be given in the next section. Later, in Sections 9.3.2 and 9.3.3 its specific application for the dGlobal muon and displaced electron calibration will be presented, as well as the derived corrections to assess the differences in the object reconstruction and identification.

9.3.1. The tag and probe method

The *tag-and-probe* method (TnP) is a generic technique designed to measure physical object efficiencies in real data. In general, the purpose of the efficiency is to quantify the performance of a certain algorithm/selection. It is measured in a predefined sample of objects that would be expected to fulfill the criteria whose efficiency is to be measured. The efficiency would thus be identified with the fraction of these objects that ultimately pass the studied criteria.

The TnP method requires to define the denominator of the efficiency in a data sample in an almost unbiased way. The method exploits the presence of a physics object in the event that should be accompanied from first principles by the object whose performance is being studied. The former is known as the tag object and the latter as the probe. The efficiency associated to the defining criteria of the object under study, can be estimated by counting the total number of events containing a tag and a probe and the number of events in which the probe passes those defining criteria. If a probe fulfill the criteria is labelled as *passing probe* and *failing probe* otherwise.

The TnP efficiency is then defined as:

$$\text{TnP efficiency} = \frac{N_{\text{probe}}^{\text{pass}}}{N_{\text{probe}}^{\text{pass}} + N_{\text{probe}}^{\text{fail}}} \tag{9.3}$$

Where $N_{\text{probe}}^{\text{pass}}$ and $N_{\text{probe}}^{\text{fail}}$ are the number of passing and failing probes in the sample, respectively-

The definitions of the tags, the probes and the passing criteria depend on the kind of efficiency that is measured in each case. In the following sections two ways of applying the method are presented. One exploits the Z boson resonance and a second one the topology of cosmic muons.

9.3.2. Displaced global muon calibration

The corrections applied to dGlobal muons are scale factors derived to account for differences in dGlobal performance between data and simulation. They will be applied to simulated muons to match those in data. Their values are identified with the ratio between the muon efficiency measured in data and simulation:

$$\text{SF}_\mu = \frac{\epsilon_\mu^{\text{data}}}{\epsilon_\mu^{\text{sim}}} \tag{9.4}$$

The inefficiencies arising from reconstruction efficiency and identification selection can be considered to be uncorrelated from each other. Therefore, the muon efficiency can be factorized to account for both contributions separately. This single efficiencies for reconstruction and ID are ultimately combined to obtain the overall efficiency:

$$\epsilon = \epsilon_{\text{id}\,|\,\text{reco}} \cdot \epsilon_{\text{reco}} \tag{9.5}$$

To derive proper corrections that could be applied in a search for LLPs, the efficiency is measured in two different muon samples by means of the tag-and-probe method. The first one consist of prompt Z decays into muon pairs, where the efficiency is studied as a function of basic muon kinematics to check consistency in p_T and η. The second one contains cosmic muons which, combined with the prompt measurements, provide a powerful handle to measure and isolate the efficiency dependence with displacement. The final scale factors are derived from these efficiency measurements.

Calibration with Z resonance

For the SFs to be applied in a search for LLPs, controlling the dependence of the muon efficiency with the displacement is crucial. However, as dGlobal muons were never used in an analysis before, these studies also encompass the performance in prompt muons to assure consistency with basic kinematics and decouple it from potential effects that may appear at high displacement. Additionally, efficiency results with prompt muons also provide a correction factor to correct the efficiency measured in cosmic muons, which is biased towards lower values due to the presence of cosmic muons that do not enter in the time window of pp muon reconstruction sequences (this effect will be widely described in the next subsection).

Prompt muon efficiency is measured with $Z \to \mu\mu$ decays using the tag-and-probe method introduced in the previous section. For the efficiency measured in data, the events are taken from the standard SingleMu dataset, which contains collisions preselected with triggers requiring just one muon at HLT level. For the simulation, a Drell-Yan Monte Carlo sample with simulated $Z \to \mu\mu$ decays is chosen. Data and simulated events are required to be preselected by the `HLT_IsoMu24` trigger path, which requires one single isolated muon with $p_T > 24$ GeV reconstructed at HLT level.

The method selects muon pairs that are likely produced in Z decays where the tag muon is defined as a muon with tight identification requirements and the probe muon is defined as a loose track where tighten requirements can be applied and whose efficiency is to be measured. Probe muons that pass the evaluated criteria are the ones labelled as passing probes. The well-known distribution for the mass of the Z boson is then used to remove TnP pairs that may come from background processes. To this purpose, the invariant mass spectra of the selected muon pairs is fitted with a combination of an exponential (background model) and voight (signal model) functions. The signal yield is extracted from this fit and the efficiency is computed from the number of probes that pass $N_{\text{probe}}^{\text{pass}}$ and fail $N_{\text{probe}}^{\text{fail}}$ the criteria. The definitions for the tag and the probe are adjusted to measure reconstruction and identification efficiencies.

For both the **reconstruction** and **identification** efficiencies, the tag muon will be a good particle-flow reconstructed muon, with medium p_T selection and passing cut-based Tight ID standard criteria. It is also required to be isolated to remove the contribution of QCD background. In addition, every tag muon is required to have fired the trigger.

For the **reconstruction** efficiency, the probes are identified with Tracker muons paired with the tag muons. The passing probes are identified with the probe muons that are also reconstructed as dGlobal muons. Specifically, a probe muon is labelled as passing probe if there is a dGlobal muon in the event which is segment-matched with the Tracker muon, i.e. their tracks share at least one segment in the muon chambers.

The choice to use of Tracker muon probes instead of just inner tracks was first applied by other CMS analysis to measure the efficiency of dSA muons [87]. It allows to use the segment-matched criteria while reducing the background arising from random inner track associations. This choice inevitably assumes that the effective Tracker muon efficiency over inner tracking efficiency is ~ 1. This assumption holds to signal muons in the considered p_T range and was proved to have negligible impact in the measured efficiencies.

The tag and probe pairs are also requested to have opposite charge, matching the signature of the neutral Z resonance. The collected mass spectra is limited to the range $[60, 120]$ GeV to isolate the resonance peak. Promptness cuts in both the tag and the probe muons are applied to remove as many PU background as possible that may arise from random combinations of tag muons and PU tracks. Finally, a cut at event level on the number of tag-and-probe pairs N_{pair} passing the aforementioned criteria above is applied to remove this contribution even more. Events are only selected if they have $N_{\text{pair}} = 1$ OR $N_{\text{pair}} = 2$ pairs to evaluate the efficiency in pairs in which there is always a valid tag muon and the probe muon could play (or not) the tag role.

The summary of the different selections applied to the tag, the probe, and the selected pairs are the following:

- Tag PF muon:
  - $p_T > 26$ GeV, $|\eta| < 2.4$
  - Pass Tight ID criteria
  - Isolated
  - Matched with an HLT muon of `HLT_IsoMu24` within $\Delta R < 0.1$.
- Probe Tracker muon:
  - $p_T > 20$ GeV, $|\eta| < 2.4$
  - $|d_0| < 0.2$ cm and $|d_z| < 0.5$ cm
- Pair selection:
  - Opposite charge between tag and probe.
  - $60 < m_{\mu\mu} < 120$ GeV
  - $\Delta z < 0.5$ cm

The final reconstruction efficiency, identified with the fraction of Tracker muon probes (selected as described above) that are segment-matched with a dGlobal muon, is shown in Figure 9.1 as a function of the probe transverse momentum p_T and in Figure 9.2 as a function of probe $|\eta|$.

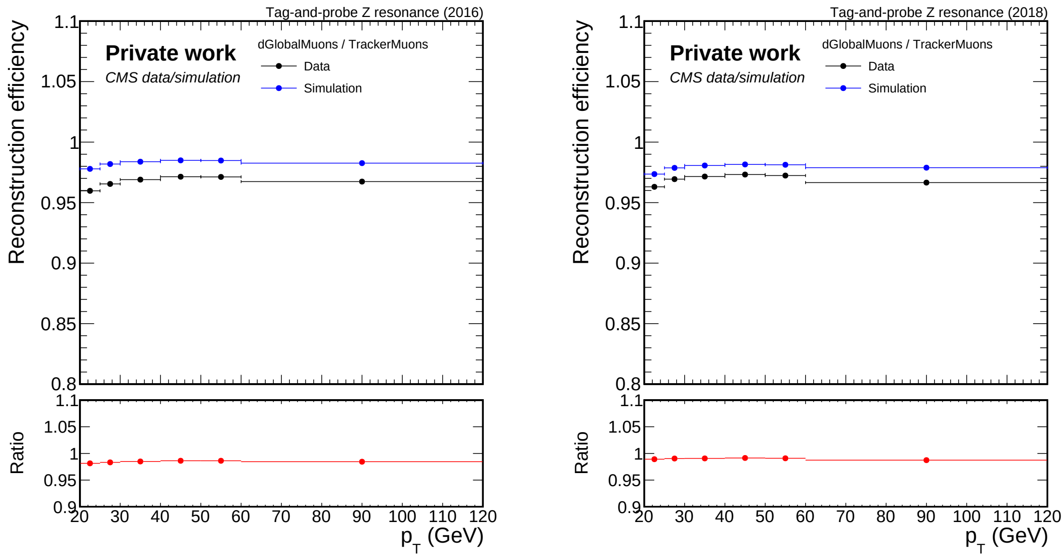


Figure 9.1: dGlobal muon reconstruction efficiency measured in $Z \to \mu\mu$ events in 2016 (left) and 2018 (right) as a function of transverse momentum p_T. Black points correspond with efficiency measured in data and blue points with efficiency measured in simulation.

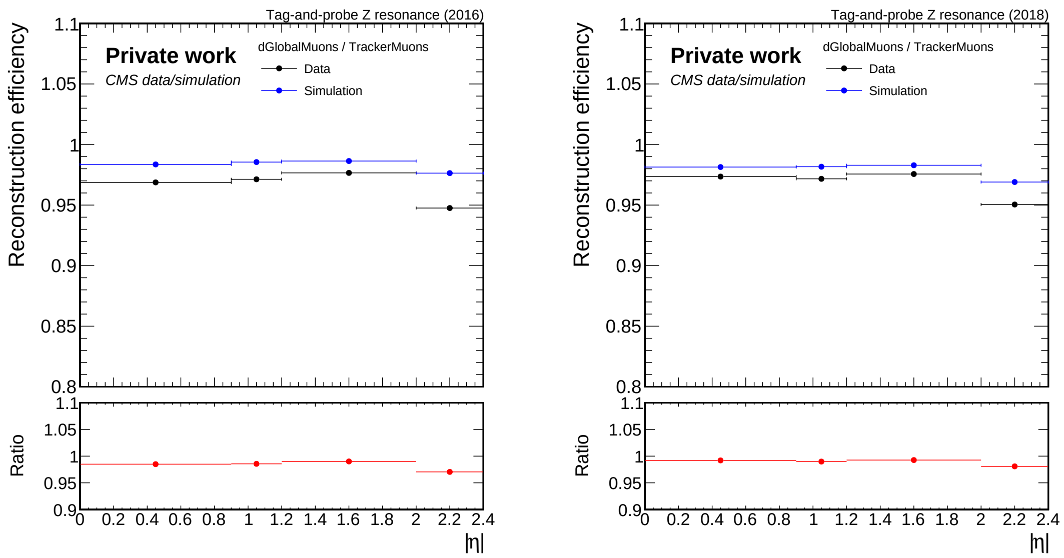


Figure 9.2: dGlobal muon reconstruction efficiency measured in $Z \to \mu\mu$ events in 2016 (left) and 2018 (right) in several bins of $|\eta|$ separating the barrel, overlap and endcaps regions of the detector. Black points correspond with efficiency measured in data and blue points with efficiency measured in simulation.

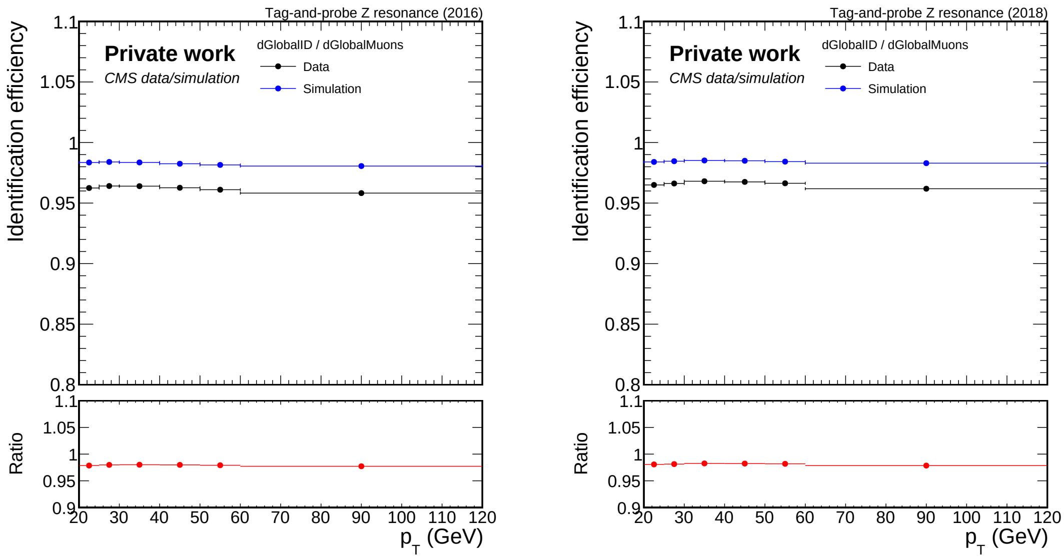


Figure 9.3: dGlobal muon ID efficiency measured in $Z \to \mu\mu$ events in 2016 (left) and 2018 (right) as a function of transverse momentum p_T. Black points correspond with efficiency measured in data and blue points with efficiency measured in simulation.

The **identification efficiency** is measured on top of reconstruction efficiency. For its efficiency to be decoupled, the identification denominator will be identified with the numerator of the reconstruction efficiency i.e. reconstructed dGlobal muons that are segment-matched with Tracker muons probes. The identification efficiency will be defined as the fraction of these already matched dGlobal muons that also pass the established ID criteria. That is, a probe dGlobal muon will be a passing probe in the identification efficiency if fulfills:

- Resolution $\Delta p_T / p_T < 0.3$
- Number of hits in strip tracker $N_{\text{strip}} > 5$
- Number of hits in muon chambers $N_{\text{muon}} > 12$

The identification efficiency is shown in Figure 9.3 as a function of transverse momentum and in Figure 9.4 in different bins of $|\eta|$.

The variation of the efficiency ratio in data and simulation along the considered p_T and η ranges will be taken as a systematic uncertainty when using the dGlobal muon objects.

Calibration with cosmic muons

The dependence of the dGlobal performance with the displacement of the muons is measured in a sample of cosmic muons. There are two main CMS standard datasets where cosmic muons are available. The first one is called `/NoBPTX/` dataset and it contains events collected during dedicated cosmic runs

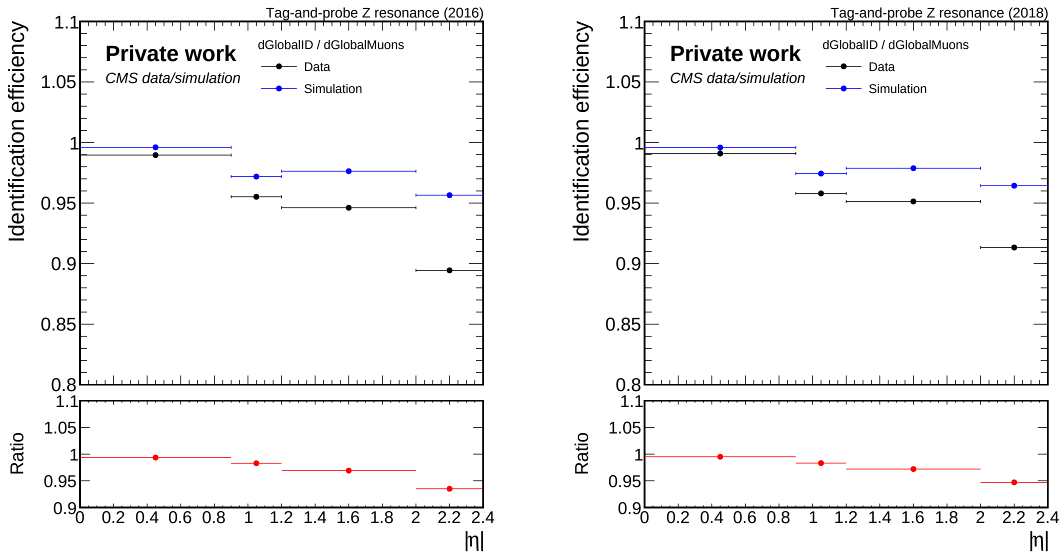


Figure 9.4: dGlobal muon ID efficiency measured in $Z \rightarrow \mu\mu$ events in 2016 (left) and 2018 (right) in several bins of $|\eta|$ separating the barrel, overlap and endcaps regions of the detector. Black points correspond with efficiency measured in data and blue points with efficiency measured in simulation.

but also events taken during normal pp collisions using a special trigger which vetoes events in colliding proton bunches. The second one is the `/Cosmics/` dataset, where only dedicated cosmic runs are contained. The main difference between both datasets, apart from the runs that are included, is that they were processed with different configurations and the event content is different. The `/Cosmics/` dataset has been processed with the reconstruction configuration optimized for cosmics muons, and essentially contains tracking, muon-related information, and additional cosmic muon reconstruction sequences. On the other hand, the `/NoBPTX/` dataset has been processed with the standard collision configuration, and it contains all information available for the collision data events. The dedicated displaced reconstructions, dSA and dGlobal muons, are included in the `/NoBPTX/` dataset but not in the `/Cosmics/` dataset, limiting their study to the runs collected in the `/NoBPTX/` dataset.

For the studies presented in this section, data events were taken from the `/NoBPTX/` dataset for the 2016 and 2018 data taking eras. The events are required to pass the `HLT_L2Mu10_NoVertex_NoBPTX3BX` trigger. This stream requires the existence of a L2 muon with $p_T > 10$ GeV with no vertex constraints apart from vetoing the events collected in colliding pp bunches.

The cosmic muon Monte Carlo was produced centrally. This sample contains simulated muons with momentum $p > 20$ GeV created at the surface of the CMS detector that travel from the top to the bottom to emulate the behavior of a genuine cosmic muon. The generated muons are filtered to have their point of closest approach within $r < 80$ cm and $|z| < 212$ cm. This filter is known as

Tracker pointing criteria and it is used to maximize the number of muons that pass close to the centre of the detector to increase the statistics when measuring tracking efficiency. In addition, the time window of the simulated muons is set to (-45, 0) ns as these values are the optimal ones to match the cosmic muon arrival time distribution in data. The cosmic trigger menu is not simulated in this sample, so simulated events are not required to pass trigger requirements.

The tag-and-probe method with cosmic muons deploys alternative definitions for the tag and the probe than those that were used to measure the efficiency with $Z \to \mu\mu$ events. As the reconstruction sequences are optimized to reconstruct muons produced in the centre of the detector, one single cosmic muon that passes through the detector may be reconstructed as two back-to-back muons instead of one. That is, for the same cosmic muon two legs can be reconstructed in an independent way in each one the hemispheres. The tag-and-probe method to measure cosmic muon efficiency relies on the assumption that for each leg reconstructed in one hemisphere and produced by a cosmic muon, there should exist another leg in the other hemisphere produced by the same cosmic muon. The efficiencies presented in this section are defined by using one muon reconstructed in one of the hemispheres as the tag, and the muons reconstructed in the other as the probe. Since in the dedicated cosmic runs the events only trigger in the lower hemisphere, muons reconstructed in this hemisphere cannot be used as the probes. This is because all events are previously required to have a muon reconstructed at HLT level and a bias in the efficiency measurement would be introduced.

Figure 9.5 collects the distributions of p_T, η and ϕ of the reconstructed dGlobal muons in both cosmic data and simulation. Cosmic data events are required to have a dGlobal muon reconstructed in the lower hemisphere ($\phi < 0$), have transverse momentum $p_T > 20$ GeV and be matched with an HLT muon of the NoBPTX trigger. Simulated cosmic events are required to pass the same selection except the trigger matching and their distributions are normalized to match the number of dGlobal muons reconstructed in the lower hemisphere in data. dGlobal muons reconstructed in the upper hemisphere in these events are also shown. The asymmetry in the η distribution that is inverted for downward and upward muons is caused by the presence of the shaft in the CMS cavern. The cosmic muons that get to the cavern through the shaft pass through less material and the flux of muons that get to reach the CMS detector is increased in this direction. The shaft is also simulated and the asymmetry also appears in the Monte Carlo simulation.

Figure 9.6 shows the distributions of the transverse $|d_0|$ and longitudinal $|d_z|$ impact parameters, measured with respect to the center of the detector. The impact parameters of the downward tag muons are a valid proxy to measure the efficiency as a function of the impact parameters of the upward probe muons.

The study of cosmic muons is limited to $p_T > 20$ GeV since cosmic muons

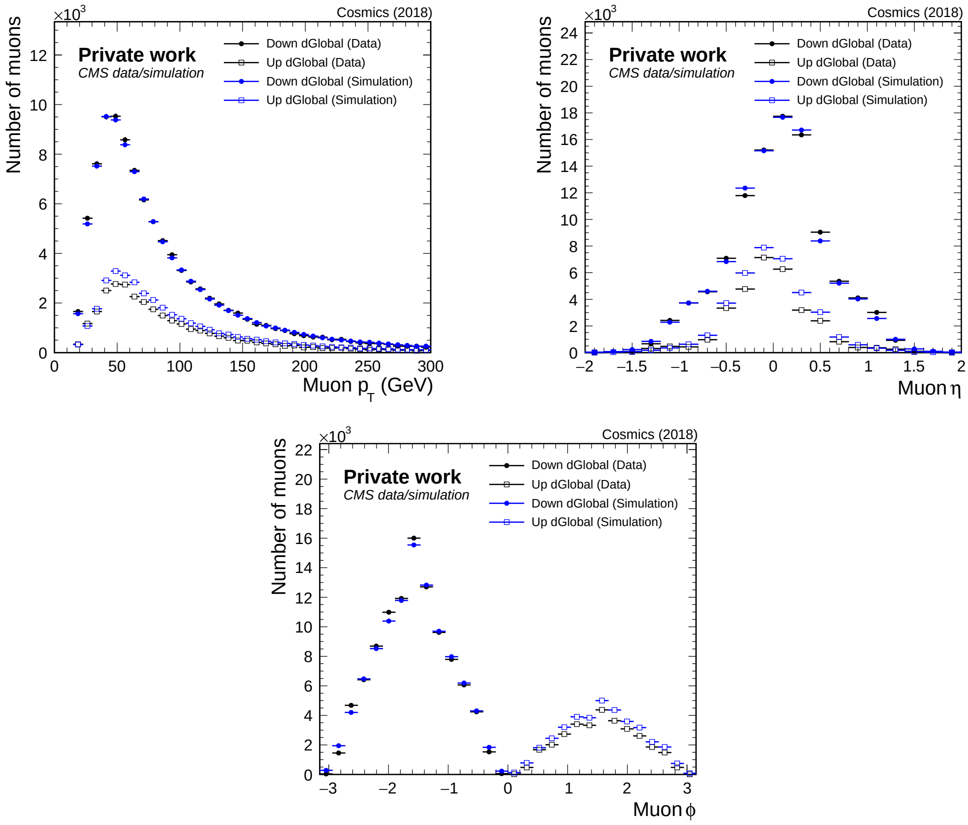


Figure 9.5: dGlobal muon p_T, η and ϕ in cosmic data (black) and simulation (blue). Events are required to have a dGlobal muon reconstructed with $p_T >$ GeV and $\phi < 0$ and be matched with an HLT muon (only in data). dGlobal muons reconstructed in the lower and upper hemispheres are shown.

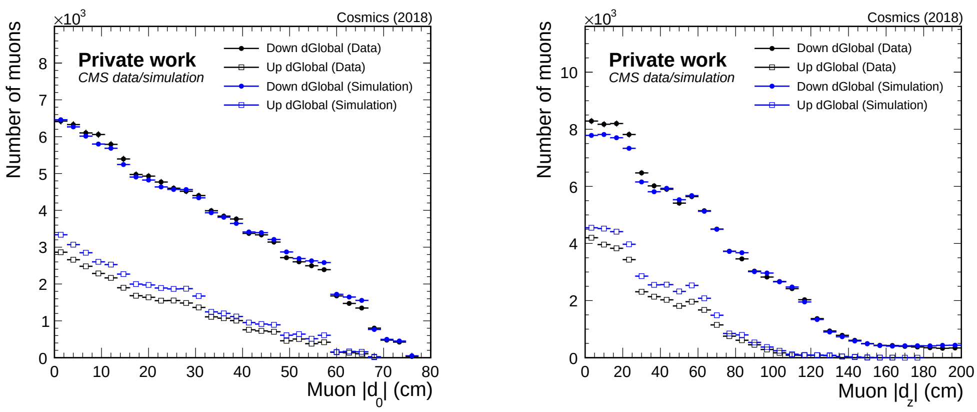


Figure 9.6: dGlobal muon $|d_0|$ and $|d_z|$ in cosmic data (black) and simulation (blue). Events are required to have a dGlobal muon reconstructed with $p_T > 20$ GeV and $\phi < 0$ and be matched with an HLT muon (only in data). dGlobal muons reconstructed in the lower and upper hemispheres are shown.

were not simulated under this threshold. Direction requirements $|\eta| < 0.9$ and $-2.6 < \phi < -0.6$ are fixed to define an acceptance region for the cosmic muons to be reconstructed in both hemispheres. With this selection, the tag and probe definitions are:

- Tag dGlobal muon:
  - Matched with an HLT muon of `HLT_L2Mu10_NoVertex_NoBPTX3BX` (only data)
  - $p_T > 20$ GeV, $|\eta| < 0.9$, $-2.6 < \phi < -0.6$
  - Working ID (as defined in Section 4.1.1)
- Probe dGlobal muon:
  - $p_T > 20$ GeV, $\phi > 0$
  - Azimuthal angle with tag muon of $|\Delta\phi(\mu_{\text{tag}}, \mu_{\text{probe}})| > 2.8$

The **reconstruction efficiency** is defined as the fraction of tag dGlobal muons reconstructed in the lower hemisphere that have a back-to-back probe dGlobal muon reconstructed in the upper hemisphere. According to this definition, the passing probes are the dGlobal muons that get to be reconstructed with the definitions stated above.

Although when tagging with the lower hemisphere we get to suppress the bias introduced by the trigger in data, the method introduces a virtual inefficiency in the results. This inefficiency is due to the fact that some of the muons the arrive at the detector are "too early in time" to be reconstructed by pp algorithms as the dGlobal reconstruction whose efficiency is measured. Studies performed by other analysis [80] reveal that the DT local reconstruction parameters of pp algorithms are configured differently for cosmics. As a consecuence, muons that do not arrive within the time window to the upper hemisphere are lost and not reconstructed as probes. The decrease in the efficiency does not have to be the same in data and simulation since in the simulation the time window for arrival muons is fixed as an additional parameter.

The reconstruction efficiency measured with the selection and definitions presented before is referred as "raw" reconstruction efficiency $\epsilon_{\text{reco}}^{\text{cosmic (raw)}}$ as it is underestimated due to the presence of early muons in the denominator. It is shown in Figure 9.7 for both data and simulation as a function of the impact parameters.

A correction factor C_{time} is computed to account for this loss in the efficiency in both data and simulation and get more realistic scale factors that could be applied to dGlobal muons used in analyses with pp collisions. This correction factor C_{time} equalizes the measured efficiency in cosmic muons that are detected prompt enough with the results obtained with $Z \to \mu\mu$ with $|\eta| < 0.9$ presented in the previous section, referred as $\epsilon_{\text{reco}}^{Z\text{ (barrel)}}$. The computation of

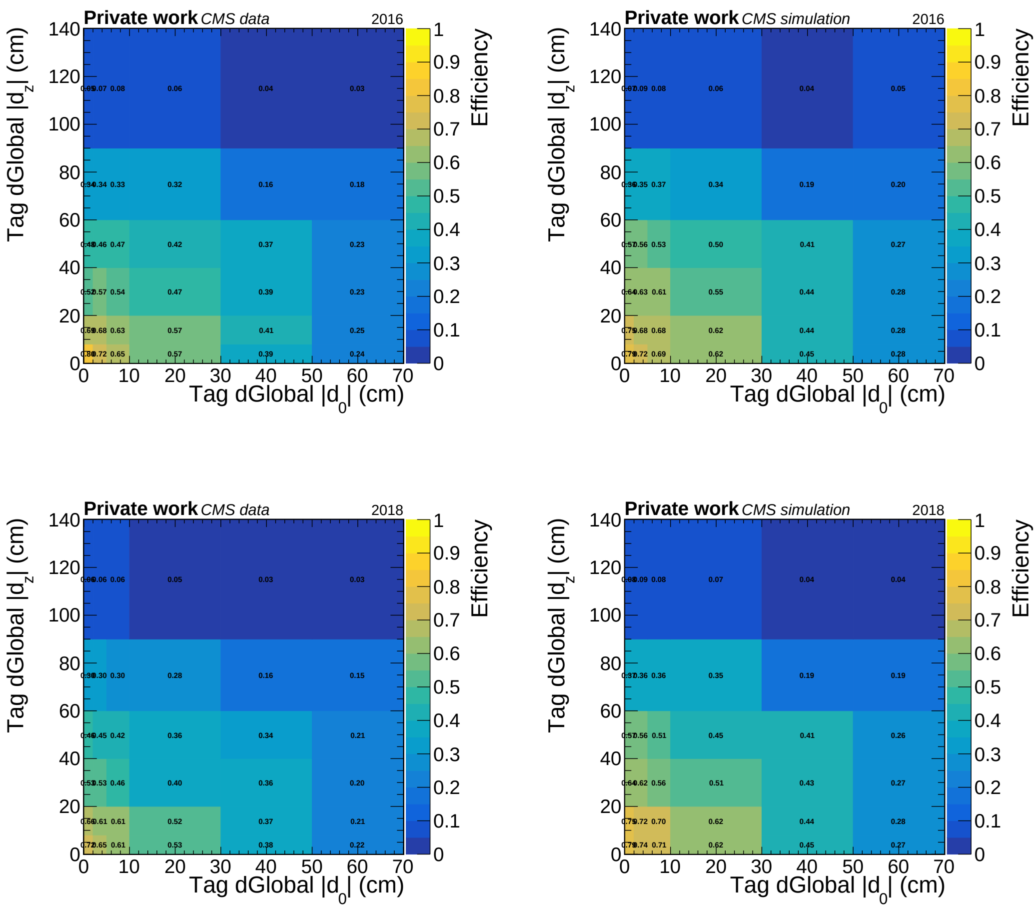


Figure 9.7: Uncorrected dGlobal reconstruction efficiency $\epsilon_{\text{reco}}^{\text{cosmic (raw)}}$ as a function of $|d_0|$ and $|d_z|$ for cosmic data (left) and cosmic simulation (right). Results are shown for runs collected in 2016 (upper) and 2018 (bottom) separately.

Year	Sample	$\epsilon_{\text{reco}}^{\text{cosmic (prompt)}}$	$\epsilon_{\text{reco}}^{Z\text{ (barrel)}}$	C_{time}
2016	Data	0.80	0.97	1.21
	Simulation	0.79	0.98	1.24
2018	Data	0.72	0.97	1.36
	Simulation	0.79	0.98	1.24

Table 9.1: dGlobal efficiency measured with cosmic muons within $|d_0| < 2$ cm and $|d_z| < 8$ cm, labelled as $\epsilon_{\text{reco}}^{\text{cosmic (prompt)}}$, with $Z \to \mu\mu$ events within $|\eta| < 0.9$, labelled as $\epsilon_{\text{reco}}^{Z\text{ (barrel)}}$, and the correction factor C_{time} computed with Eq. 9.6.

C_{time} is done in such a way that the efficiency measured with cosmic muons within $|d_0| < 2$ cm and $|d_z| < 8$ cm, denoted as $\epsilon_{\text{reco}}^{\text{cosmic (prompt)}}$, should be equal to $\epsilon_{\text{reco}}^{Z\text{ (barrel)}}$:

$$\epsilon_{\text{reco}}^{\text{cosmic}} = C_{\text{time}} \cdot \epsilon_{\text{reco}}^{\text{cosmic (raw)}} \qquad \text{with} \qquad C_{\text{time}} = \frac{\epsilon_{\text{reco}}^{Z\text{ (barrel)}}}{\epsilon_{\text{reco}}^{\text{cosmic (prompt)}}} \tag{9.6}$$

The efficiencies $\epsilon_{\text{reco}}^{\text{cosmic (prompt)}}$, $\epsilon_{\text{reco}}^{Z\text{ (barrel)}}$ and correction factors C_{time} computed for dGlobal muons measured in 2016 and 2018 are found in Table 9.1.

The corrected efficiencies in data and simulation, as defined by Eq.(9.6), are shown in Figure 9.8 as a function of the transverse impact parameters. The reconstruction scale factors are defined as the ratio between the corrected efficiency measured in data and simulation, computed as a function of $|d_0|$ and $|d_z|$. They are shown in Figure 9.9.

As in the case of the $Z \to \mu\mu$ measurement, the **identification efficiency** $\epsilon_{\text{id | reco}}^{\text{cosmic}}$ is measured on top of the reconstruction efficiency presented above. It is defined over the sample of dGlobal muons reconstructed in the upper hemisphere (properly matched with the tags) as the fraction of them that also pass the dGlobal ID criteria. It is shown in Figure 9.10 as a function of $|d_0|$ and $|d_z|$. It can be observed that the ID performance is quite similar for data and simulation and its efficiency is stable with displacement. This results in a flat, ~ 1 scale factor whose value is shown in Figure 9.11. In addition, no significant decrease is observed with respect to the results reported for $Z \to \mu\mu$ events in the previous subsection, which indicates that the selection variables for the ID are unaffected by the wrong setting of the local reconstruction parameters of the DTs.

Systematic uncertainty on muon calibration

Muon efficiency is measured in both cosmic muons and $Z \to \mu\mu$ decays. The considered systematic uncertainty cover effects in both reconstruction and identification efficiency:

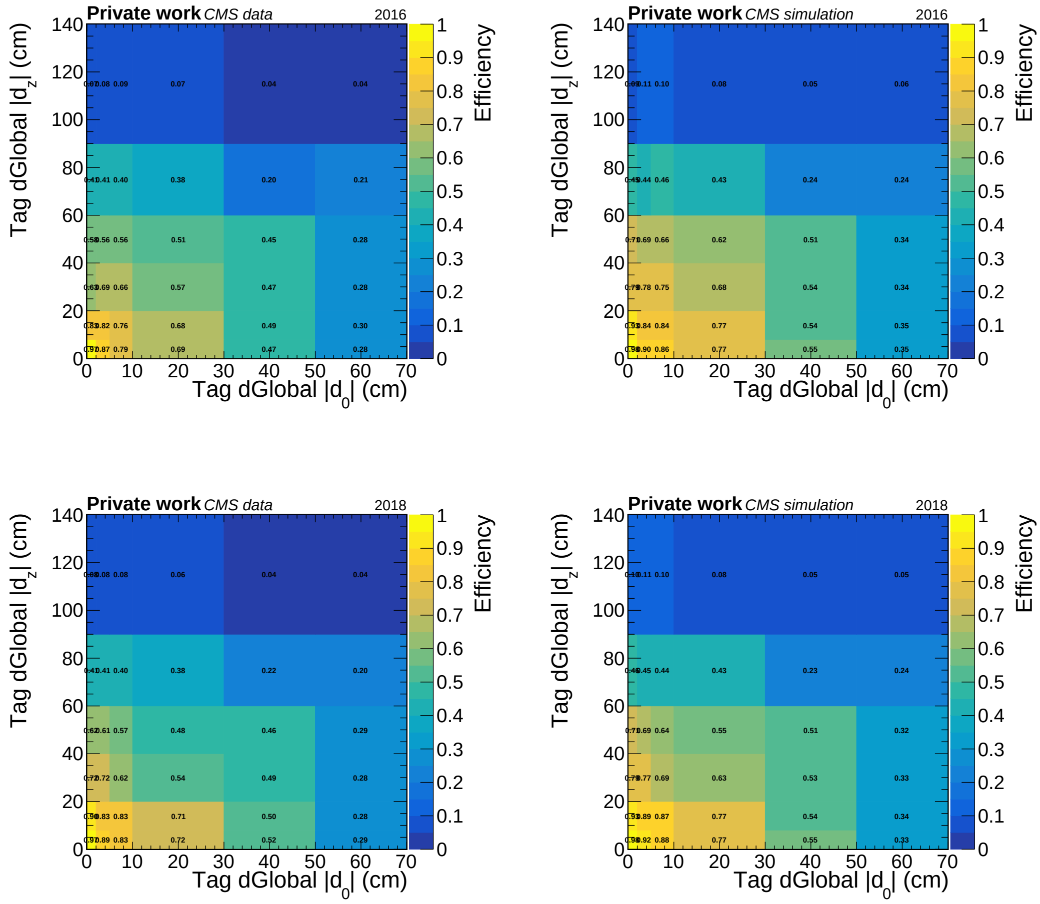


Figure 9.8: Corrected dGlobal reconstruction efficiency $\epsilon_{\text{reco}}^{\text{cosmic}}$ as a function of $|d_0|$ and $|d_z|$ for cosmic data (left) and cosmic simulation (right). Results are shown for runs collected in 2016 (upper) and 2018 (bottom) separately.

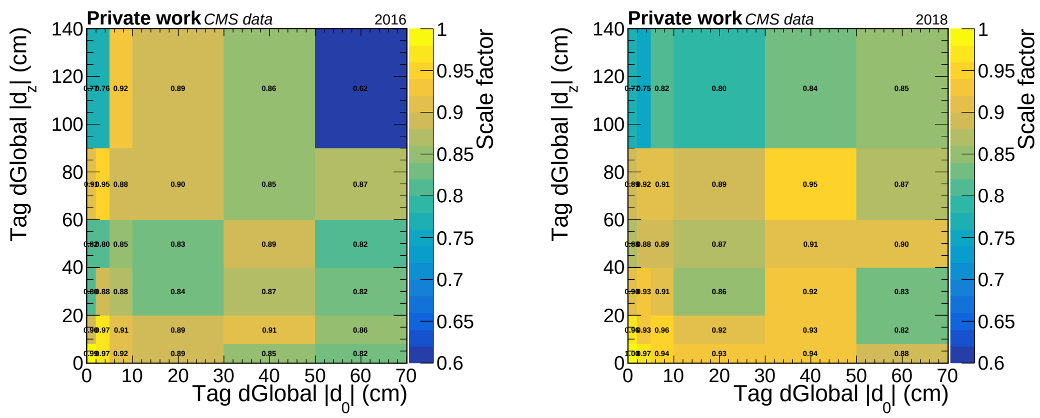


Figure 9.9: dGlobal reconstruction scale factors computed in cosmic muons collected in 2016 (left) and 2018 (right) as a function of $|d_0|$ and $|d_z|$.

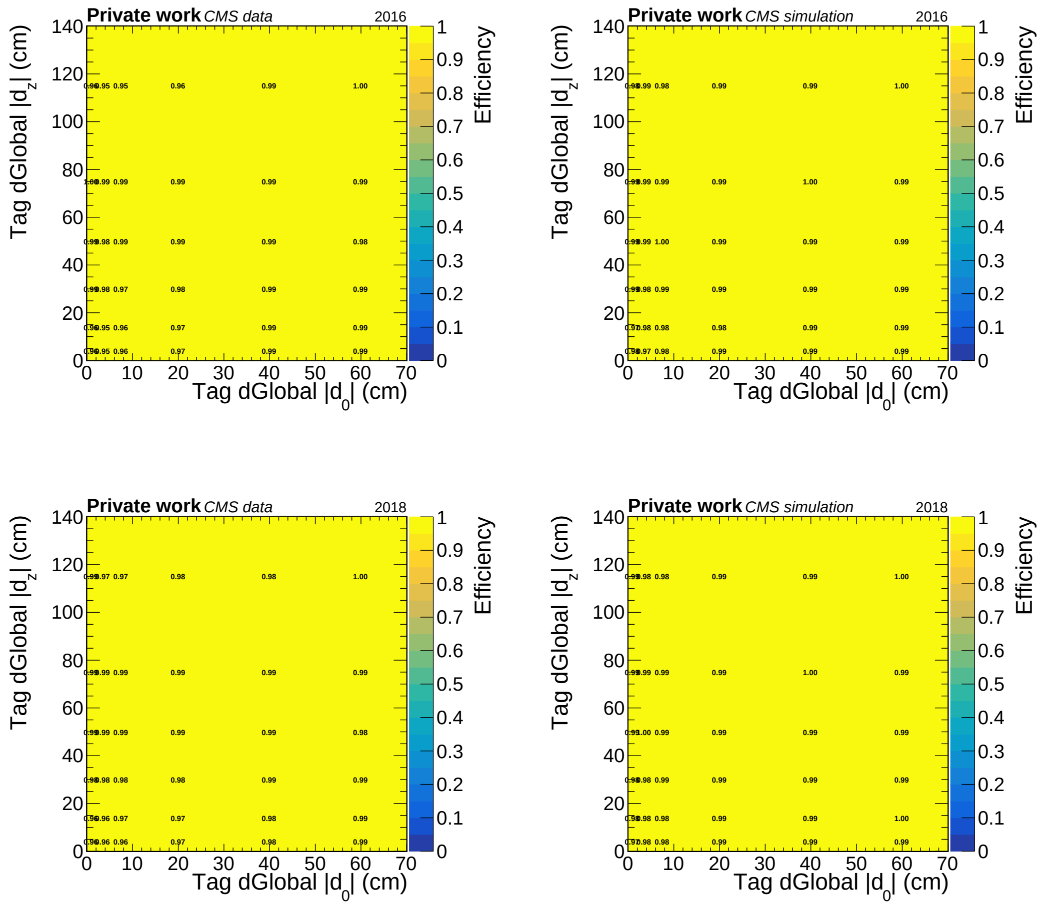


Figure 9.10: dGlobal identification efficiency $\epsilon_{\text{reco}}^{\text{cosmic}}$ as a function of $|d_0|$ and $|d_z|$ for cosmic data (left) and cosmic simulation (right). Results are shown for runs collected in 2016 (upper) and 2018 (bottom) separately.

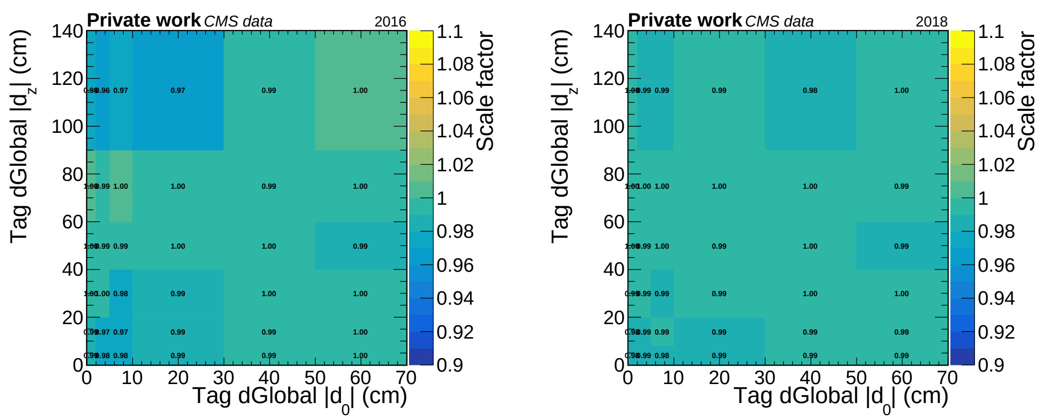


Figure 9.11: dGlobal identification scale factors computed in cosmic muons collected in 2016 (left) and 2018 (right) as a function of $|d_0|$ and $|d_z|$.

- Variations in the efficiency as a function of $|\eta|$ were considered as the cosmics only provide statistics in the central barrel region ($|\eta| < 0.9$). For this, the maximum variation that was measured in $Z \to \mu\mu$ decays was taken as an estimate for the systematic uncertainty, found to be a 4% difference between probe muons reconstructed with $|\eta| < 0.9$ and in the last bin considered $1.2 < |\eta| < 2$.
- It was identified that one of the most important sources of systematic uncertainty is the one introduced by the out-of-time muons correction C_{time}. The derivation of this factor is affected by the statistic in the first bin of the cosmic muon distribution i.e. cosmic muons with $d_0 < 2$ cm and $d_z < 8$ cm. This statistical uncertainty is found to be $\Delta\epsilon = 3.5\ (1.3)\%$ in data (simulation). The corresponding uncertainty associated to the correction factor is therefore $\Delta C_{\text{time}} = 3.5\%$.

The total uncertainty per muon is found to be 5.3%. For the events with a dimuon pair where two muons are required to be reconstructed and identified, a conservative 10.6% systematic uncertainty was assigned.

9.3.3. Displaced electron calibration

Following the same reasoning as for dGlobal muons, dedicated corrections should be derived for the simulation to match the electron performance observed in data. In this case, electron scale factors are defined as the ratio between the electron efficiency measured in data and simulation:

$$\text{SF}_e = \frac{\epsilon_e^{\text{data}}}{\epsilon_e^{\text{sim}}} \tag{9.7}$$

Displaced electrons are isolated tracks matched with superclusters in the ECAL. The efficiency of these objects depends on the efficiency of reconstructing the supercluster in the ECAL, reconstructing the track in the tracker, and performing the matching between them. The total efficiency ϵ_e can be consequently factorized to account for these steps separately as

$$\epsilon_e = \epsilon_{\text{track}} \cdot \epsilon_{\text{sc | track}} \tag{9.8}$$

where the *tracking efficiency* ϵ_{track} is the efficiency to have a track reconstructed for a certain displaced electron and the so-called *track-to-SC efficiency* $\epsilon_{\text{sc | track}}$ is the efficiency to have an ECAL SC matched for that given track.

Electrons produced in displaced LLP decays are expected to leave a supercluster in the ECAL as long as they are produced within the volume delimited by its inner surface. On the other hand, the tracking efficiency has a strong dependence on the decay distance, the number of hits associated to the track, whether the tracks points or not to the center of the detector and other effects. Both supercluster reconstruction and track-supercluster association are then expected to be highly efficient, while the tracking efficiency is expected to gradually decrease with increasing displacement. Therefore, the total displaced electron efficiency is dominated by the tracking efficiency.

In order to calibrate the displaced electron objects, these two efficiencies are measured separately. First, the tracking efficiency ϵ_{track} is measured in a sample of cosmic muons as a function of the displacement. Then, the track-to-SC efficiency $\epsilon_{\text{sc}\,|\,\text{track}}$ is measured with electrons produced in $Z \to ee$ events for tracks that are already properly reconstructed. This method assumes that the displacement dependence of the track-to-SC efficiency can be neglected compared to the strong decrease in tracking efficiency with displacement. The reasoning behind this assumption was carefully studied in signal simulation in Section 4.2 when the supercluster and track efficiency were measured in simulated signal events (see Figure 4.5).

In contrast to dGlobal muons, the efficiency measurement in this case is not divided into offline reconstruction and identification selection. Instead, both aspects are measured together. Specifically, the efficiencies ϵ_{track} and $\epsilon_{\text{sc}\,|\,\text{track}}$ are determined for tracks and SCs that already fulfill the corresponding ID requirements outlined in Table 4.2.

Track calibration

A new variation of the tag-and-probe methodology was adopted to measure the displaced tracking efficiency ϵ_{track}. In this case, the tracks reconstructed only in the muon chambers serve as a reliable tag while the tracks detected in the tracker are utilized as probes for measuring their efficiency.

This approach was developed by other analysis in CMS with displaced leptons [79] and this study adapted the efficiency definitions to the track selection defined for the displaced electron reconstruction. The basis of the method relies on the fact that a cosmic muon that crosses the entire detector can be reconstructed as a track in the muon chambers and as a track in the tracker independently. The method assumes that a cosmic muon which passes through the tracker and leaves a good quality track in the muon chambers should also leave a tracker track. The tracking efficiency ϵ_{track} is thus defined as the fraction of tight muons detected in the muon chambers (tags) that are also detected as tracker tracks (passing probes).

The displaced tracking efficiency is measured with the same cosmic data and simulated samples that were used to compute the dGlobal efficiency with cosmic muons (see Section 9.3.2). In the same way, data events are required to pass the `HLT_L2Mu10_NoVertex_NoBPTX3BX` trigger while no requirements on the trigger are imposed on simulated events.

The tag muons are reconstructed with dedicated algorithms to identify cosmic muons which are optimized to be highly efficient with displaced muons in the muon chambers. Two algorithms are used, the 1-leg algorithm which

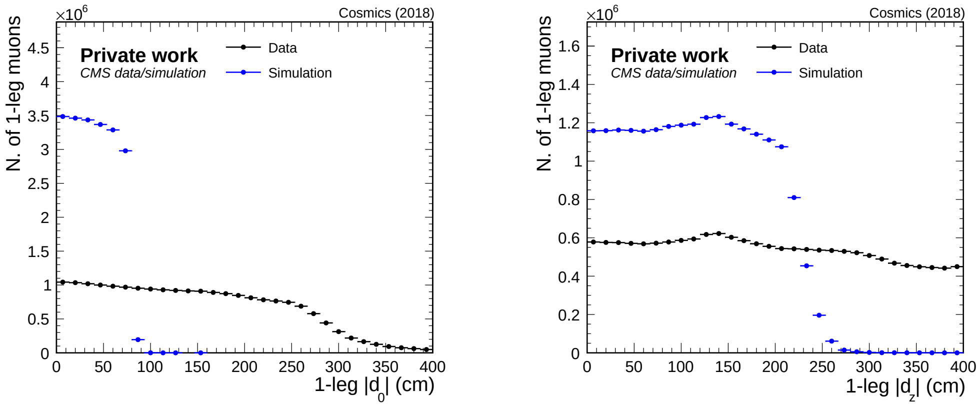


Figure 9.12: Impact parameters $|d_0|$ and $|d_z|$ of selected 1-leg muons measured in cosmic data (black) and simulation (blue). Events are required to have a 1-leg muon reconstructed with $p_T > 20$ GeV, $N_{\text{hits}} > 50$, and $\Delta t < 70$ ns which match the tag selection requirements.

reconstructs the trajectory of the cosmic muon as a single muon crossing the whole detector, and the 2-leg algorithm which reconstructs the same cosmic muons as two separate muons, one in each hemisphere [92]. The tag muons will be identified as 1-leg muons, reconstructed **only** in the muon chambers so the tracking efficiency is not biased. They are required to have $p_T > 20$ GeV, more than 50 hits in the muon system and the error of the time measured in the muon chambers should be less than $\Delta t < 70$ ns.

Figure 9.12 collects the impact parameters of the 1-leg tags where data and simulation present very different distributions. The origin of the discrepancy is a direct consequence of the aforementioned *tracking pointing* filter which is applied in the cosmic simulation. This filter, as explained in the previous section, rejects generated cosmic muons whose PCA is found at $r > 80$ cm or $|z| > 212$ cm, and therefore the distributions of $|d_0|$ and $|d_z|$ are cut off at these two values.

To assure that the tracker is working at the time the cosmic muon arrives, a cut in the time of the reconstructed tags is applied. The time is measured by matching the 1-leg muon to two reconstructed standalone 2-leg muons within $|\Delta\phi| < 0.2$, one in each half of the CMS detector. Then, the average time in the muon chambers of the two matched 2-leg muons is taken [1]. The time distribution is shown in Figure 9.13 where it can be seen how the time measured in data and in simulation differ from each other. This difference comes from the fact that the time at which the cosmic muons were generated is a free parameter that can be set.

Since the same cosmic muon can leave a tracker track in each half of the CMS

[1]The combination of the two 2-leg time measurements is proven to be a more accurate estimation of the arrival time than the measurement provided by just the 1-leg muon.

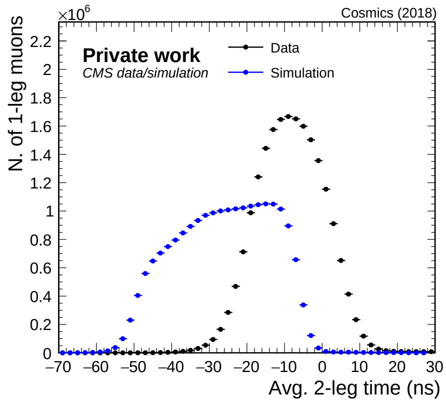


Figure 9.13: Time t of selected 1-leg muons measured in cosmic data (black) and simulation (blue). Events are required to have a 1-leg muon reconstructed with $p_T > 20$ GeV, $N_{\text{hits}} > 50$, $\Delta t < 70$ ns and be matched to two 2-leg muons. The time t is computed as the average time of the two matched 2-leg muons.

detector, two definitions can be adopted for the efficiency. In this measurement the efficiency is computed by using downward tracks since they are produced by an inside-out particle as the electrons produced from pp collision processes. The probes are tracker tracks reconstructed with the tracking pp sequence in the lower hemisphere and matched with the tag 1-leg muon within $|\Delta\phi| < 0.2$ [2]. The probes are also requested to have at least $N_{\text{hits}} > 6$ in the tracker, matching the same requirements established for the identification of the displaced electron tracks.

To identify the time intervals when the tracker is efficient, the tracking efficiency is computed as a function of the time measured in the muon chambers for the tag muons. To increase the efficiency and identify the time window correctly, the tags are required to have $|d_0| < 8$ cm and $|d_z| < 20$ cm. This efficiency is shown in Figure 9.14. The time windows are different for each year considered and found to be:

- 2016: $[-16, -4]$ ns for data and $[-38, -32]$ ns for the simulation.
- 2017: $[-17, -3]$ ns for data and $[-38, -32]$ ns for the simulation.
- 2018: $[-8, -2]$ ns for data and $[-40, -34]$ ns for the simulation.

The downward tracking efficiency is computed taking into account the selection introduced in the previous paragraphs that is summarized as follows:

- Tag 1-leg muon:
  - $p_T > 20$ GeV

[2]Analogous definition can be adopted for the upward efficiency with tracks reconstructed in the upper hemisphere. However, the $\vec{p_T}$ of the upward tracks should be flipped to perform the matching as the 1-leg muons always point downwards.

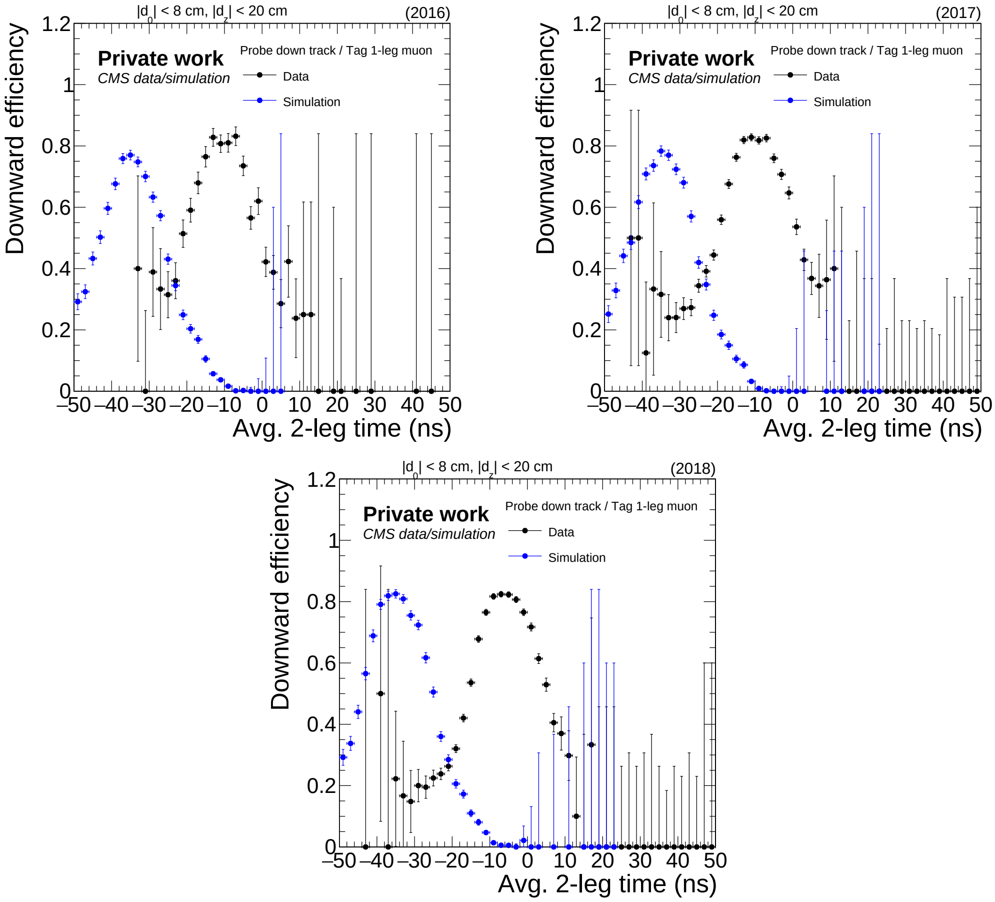


Figure 9.14: Downward efficiency as a function of the time measured in the muon chambers for the selected 1-leg muon tags. Efficiency is shown for cosmic data (black) and simulation (blue). 1-leg tags are required to be measured within $|d_0| < 8$ cm and $|d_z| < 20$ cm to increase the efficiency. The efficiency is shown for 2016, 2017 and 2018 events separately.

  - Be a muon whose track is only reconstructed in the muon chambers.
  - $N_{\text{hits}} > 50$
  - Time resolution $\Delta t < 70$ ns
  - Be matched width two 2-leg muons within $|\Delta\phi| < 0.2$
  - Average time from the 2-leg muons within tracker window.
- Probe general track:
  - $p_T > 20$ GeV, $\phi < 0$
  - Number of hits $N_{\text{hits}} > 6$
  - Matched with tag if $|\Delta\phi| < 0.2$

The downward efficiency is defined as the fraction of 1-leg tag muons that are matched with a probe downward general track in the tracker. The downward efficiency measured in data and simulation is shown in Figure 9.15 as a function of the impact parameters $|d_0|$ and $|d_z|$.

The scale factors, defined as the ratio of the downward efficiency measured in data and simulation, are shown in Figure 9.16 for 2016, 2017 and 2018 data. They are applied in the signal simulation as a function of the electron $|d_z|$ and $|d_0|$.

Track-to-Supercluster calibration

The tag-and-probe method is applied to measure the track-to-SC efficiency $\epsilon_{\text{sc}\,|\,\text{track}}$ and derive a correction for the simulated displaced electrons. As introduced some paragraphs above, this efficiency quantifies the probability to have a given good quality track to be promoted to a displaced electron object.

The efficiency is measured in an unbiased sample of $Z \to ee$ decays in both data and simulation for the three years of data taking. Data events are taken from the `SingleElectron` dataset in 2016 and 2017, and from the `EGamma` dataset in 2018. For the simulation, a Drell-Yan Monte Carlo sample with generated decays over $m > 50$ GeV is used. Simulated events are strictly required to have a Z boson with pure electron decays, rejecting virtual non-resonant γ^* decays and the $Z \to \tau\tau$ leptonic contribution. In both data and simulation, events are required to pass single electron streams which require just one electron reconstructed at HLT level (that corresponds with the tag electron) [3].

The events are tagged with a good quality reconstructed displaced electron with $p_T^{\text{track}} > 15$ GeV, $E_T^{\text{SC}} > 40$ GeV and $|\eta| < 2$. This electron is also required to be isolated and prompt reconstructed with impact parameters being below

[3] Note that the use of streams requiring more electrons would risk to introduce a bias in the efficiency measurement as they would require the presence of two superclusters.

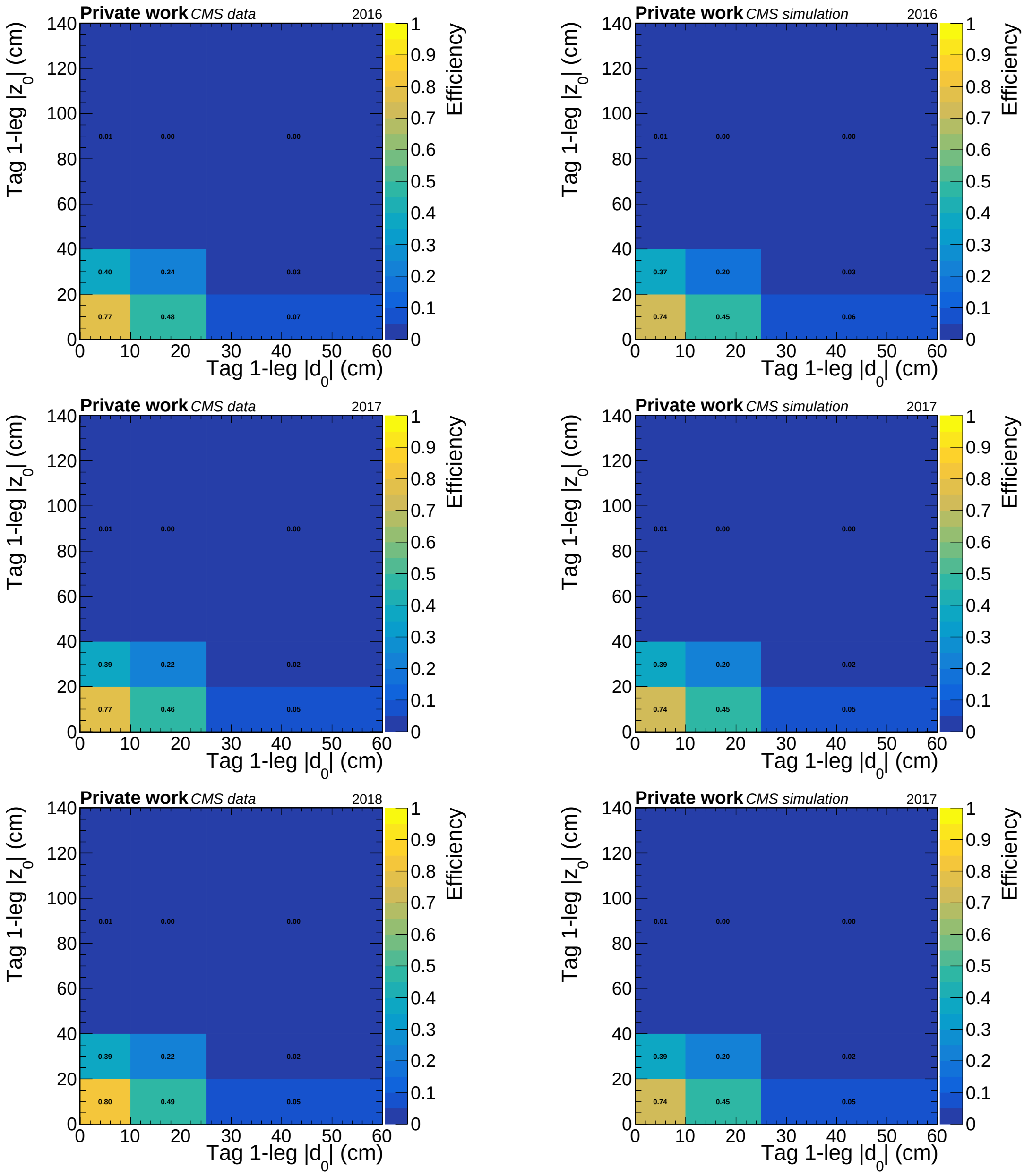


Figure 9.15: Tracking efficiency ϵ_{track} measured in cosmic muon data (left) and simulation (right) as a function of the transverse and longitudinal impact parameters $|d_0|$ and $|d_z|$. Results are shown for years 2016, 2017 and 2018.

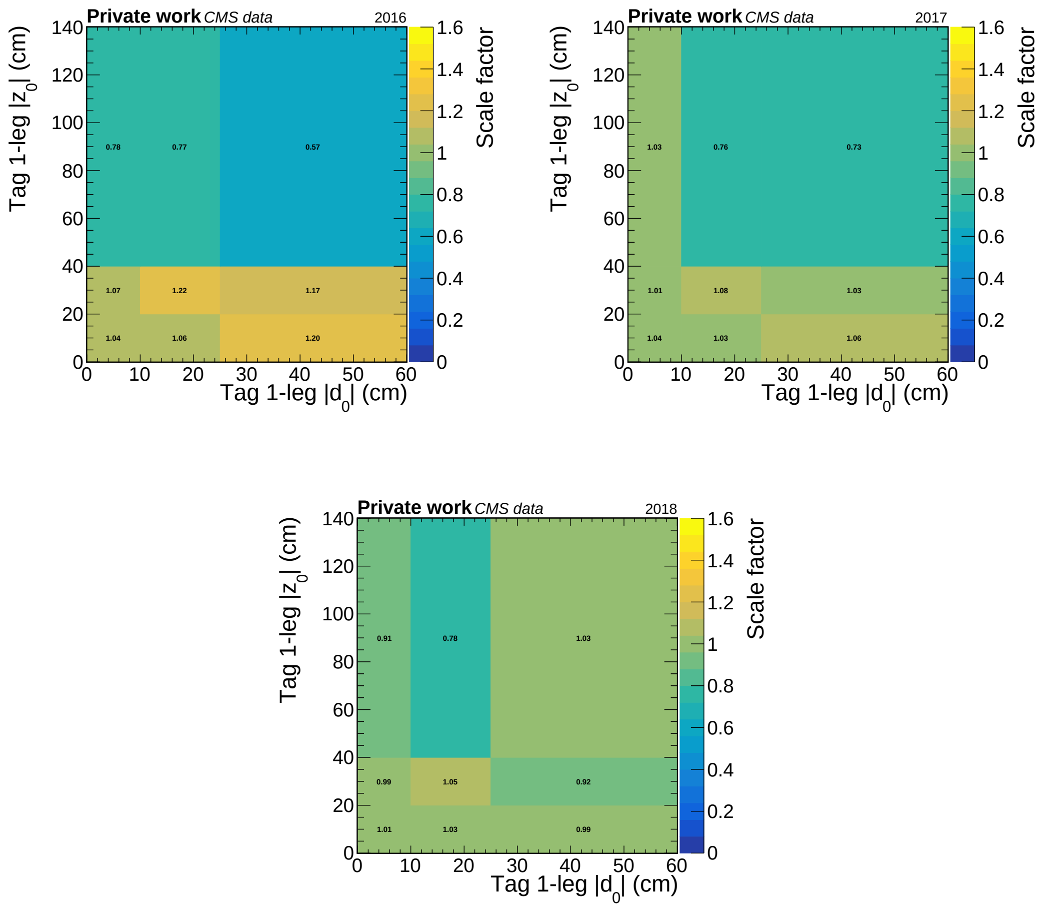


Figure 9.16: Scale factors applied on electron tracks in 2016, 2017 and 2018. They are measured as a function of the transverse and longitudinal impact parameters $|d_0|$ and $|d_z|$.

$|d_0| < 0.2$ cm and $|d_z| < 0.5$ cm. In addition, it must pass the ID selection optimized for displaced electrons.

The probes are isolated tracks with $p_T^{\text{track}} > 15$ GeV and $|\eta| < 2$. In the same way as for the tag, they are also requested to be isolated and prompt with $|d_0| < 0.2$ cm and $|d_z| < 0.5$ cm. They must also pass the track-part of the ID selection.

Tag-and-probe pairs are constructed by matching tags and probes requiring opposite charge and invariant mass values in the range $[60, 120]$ GeV. Even being strongly suppressed by promptness cuts, spurious combinations of displaced electrons are further rejected by requiring the events to have no more than two tag and probe possible pairs in the event.

The tag and probe definitions are summarized as follows:

- Tag displaced electron:
  - $p_T^{\text{track}} > 15$ GeV, $E_T^{\text{SC}} > 40$ GeV and $|\eta| < 2$
  - Isolated.
  - $|d_0| < 0.2$ cm and $|d_z| < 0.5$ cm
  - Displaced electron ID.
- Probe isolated track:
  - $p_T^{\text{track}} > 15$ GeV and $|\eta| < 2$
  - Isolated
  - $|d_0| < 0.2$ cm and $|d_z| < 0.5$ cm
  - Track selection of displaced electron ID.

Passing probes are those that are successfully matched with an ECAL SC within $\Delta R(\text{track}, \text{SC}) < 0.1$ as established for displaced electron reconstruction. The matched SC, in addition, must pass the SC ID selection completing the definition for the full displaced electron ID.

In order to measure the real probe contribution from data in $Z \to ee$ decays, the invariant mass spectra of the passing and failing tag-and-probe pairs is fitted with a signal + background analytical function. The signal is modelled with a Crystal-Ball function [100], a probability density function to model lossy processes in HEP suitable to emulate the potentially radiating e^+e^- pairs produced from Z decays. The background instead, is modelled with an exponential template.

For the simulation, $N_{\text{MC probe}}^{\text{pass}}$ and $N_{\text{MC probe}}^{\text{fail}}$ are just measured with a cut-and-count approach by simply counting the passing and failing events fulfilling the above definitions. Counting the events is considered to be an accurate estimation as the generator-level selections on the sample and the tight selection

Year	$\epsilon^{\text{data}}_{\text{sc}\,|\,\text{track}}$ (fitted)	$\epsilon^{\text{sim}}_{\text{sc}\,|\,\text{track}}$ (cut-and-count)	C_{sc}
2016	0.86	0.86	1.00
2017	0.85	0.86	0.98
2018	0.84	0.86	0.97

Table 9.2: Track-to-SC efficiencies measured in data $\epsilon^{\text{data}}_{\text{sc}\,|\,\text{track}}$ and simulation $\epsilon^{\text{sim}}_{\text{sc}\,|\,\text{track}}$ for each year of Run 2 data taking. The correction factor C_{sc} estimated as the ration between both values is also shown.

applied to the tag-and-probe have proven to remove the small background contribution that could arise in the simulation.

There is no binning applied in probe p_T nor η but just a single correction factor C_{sc} derived per year and defined as follows:

$$C_{\text{sc}} = \frac{\epsilon^{\text{data}}_{\text{sc}\,|\,\text{track}}}{\epsilon^{\text{sim}}_{\text{sc}\,|\,\text{track}}} \tag{9.9}$$

The passing and failing fits are collected in Figure 9.17 while the corresponding invariant mass distribution for the simulation are shown in Figure 9.18. The measured efficiencies in data and simulation and the corresponding correction factor is presented per year in Table 9.2.

Systematic uncertainty on electron calibration

There are two main parts that may be affected by systematic uncertainties in a separate way. First, the tracking efficiency is measured as a function of the disaplacement in a sample of cosmic muons tagging with standalone muons and probing with the tracker tracks. The displaced leptons are corrected with the derived scale factors as a function of the impact parameters. Second, the efficiency to have a supercluster associated to given reconstructed track is measured to apply an additional single correction on top of the tracking scale factor. It is measured in prompt $Z \to ee$ on top of the tag and probe method. The sources of systematic uncertainty considered for these measurements are:

- The definition of the time window to select the time at which the tracker was efficient. The width of both data and simulation time windows was varied and their position shifted by 3-4 ns. The highest variation in the SF was found to be 5% in 2016 and 2017, while in 2018 the variation was 3%. The $\Delta\phi$ between the track and the standalone muon was also changed, but the resulting effect was smaller in comparison.
- The definition of the tag and probe selection when measuring the track-to-SC correction. In this case it was changed:
  - The requirements of the tag electron p_T, E_T and isolation.

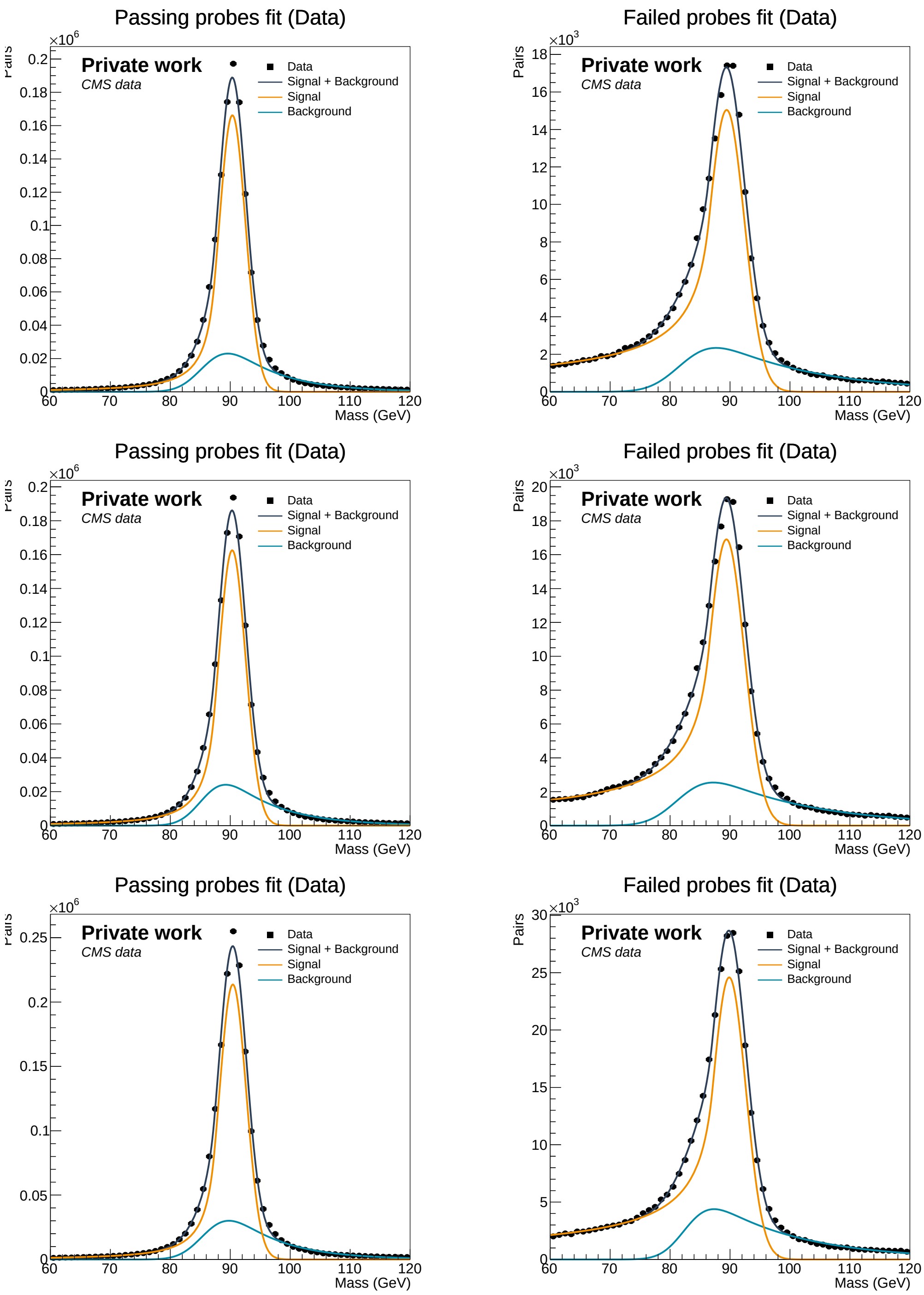


Figure 9.17: Invariant mass spectra of the tag-and-probe pairs collected in data with probes passing (left) and failing (right) the track-SC matching criteria established for the displaced electron reconstruction. The distributions are fitted with a signal + background function were the signal is modelled with a Crystal-Ball function and the background with an exponential template. The resulting fits are shown separately for 2016 (up), 2017 (middle) and 2018 (bottom) collected data.

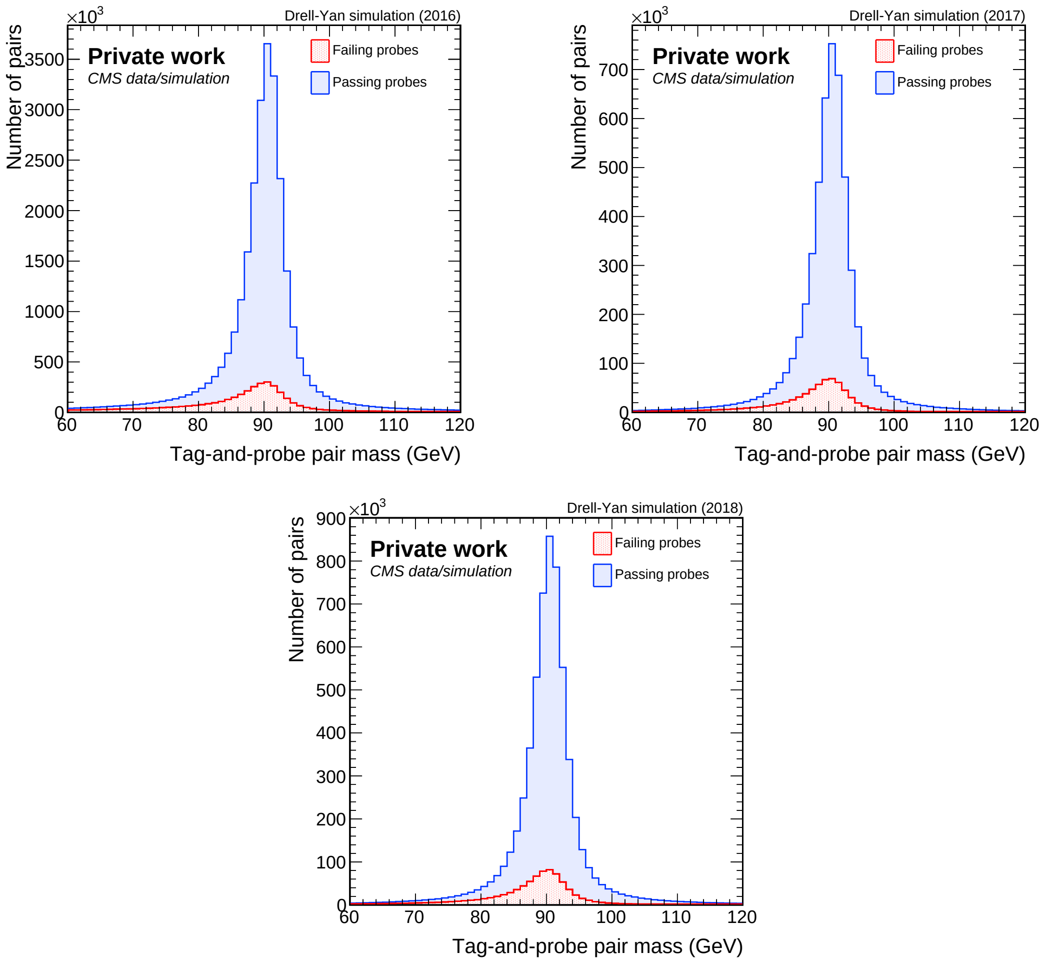


Figure 9.18: Total invariant mass spectra of the simulated tag-and-probe pairs where probes passing and failing the track-SC matching criteria established for the displaced electron reconstruction are shown separately. The distributions are shown for 2016 (up), 2017 (middle) and 2018 (bottom) simulated detector conditions.

- The $|\eta|$ bins of the probe track.
- The initial parameters of the fitting functions, including the initial normalization signal to background.

The variations of C_{SC} were no larger than 4% and this value is taken as a systematic uncertainty.

- In the same way as muons, the fact that cosmic muons just cover a limited pseudorapitity range $|\eta| < 0.9$. The same value as muons is assumed, which is 4%.

The total uncertainty per electron is found to be 7.5% in 2016 and 2017, and 6.4%. Since we require at least two electrons identified according to the presented criteria to have a valid vertex we assign a conservative 15% to these events in 2016 and 2017, and 13% in 2018.

9.4. Displaced lepton isolation correction

Since cosmic muons are isolated and there are no significant sources of non-isolated displaced leptons in collision data, this analysis measured the isolation SF in prompt events and studies their application in displaced signal simulation to account for displacement.

Regarding the kinematic variables, the bigger effect is found as expected for p_T. For this reason, p_T-dependent weights, i.e. isolation scale factors, are derived based on the comparison of data and simulation events in a Drell-Yan enriched region (dominated by prompt leptons). Then, an additional systematic uncertainty is considered to account for a potential dependence of the scale factors with the lepton displacement.

The measurement of the isolation scale factors is done in the On-Z Control Region of the search, which is completely orthogonal to the Signal Region. This Drell-Yan enriched region is defined by requiring a pair of electrons or muons with opposite charge and passing the ID selection criteria. This pair is also required to form a valid secondary vertex, have a collinearity value of $|\Delta\Phi(\ell\ell, DV)| > \pi/2$, and have an invariant mass in the range $81 < m_{\ell\ell} < 101$ GeV. The analyzed events are required to pass triggers that do not require isolation. It has been observed that the use of triggers with isolation requirements results in an efficiency bias, as the dilepton pairs are preselected based on isolation criteria at HLT level. This bias was measured to be particularly pronounced in dielectron channel for 2016, where the HLT tracker isolation is included in the trigger path requirements. The triggers available for the electron isolation measurement are[4]:

[4] `HLT_DoublePhoton33_CaloIdL` was not available in 2016. However, as the dilepton pairs are mainly prompt $Z \to ee$ decays, both options are expected to be valid.

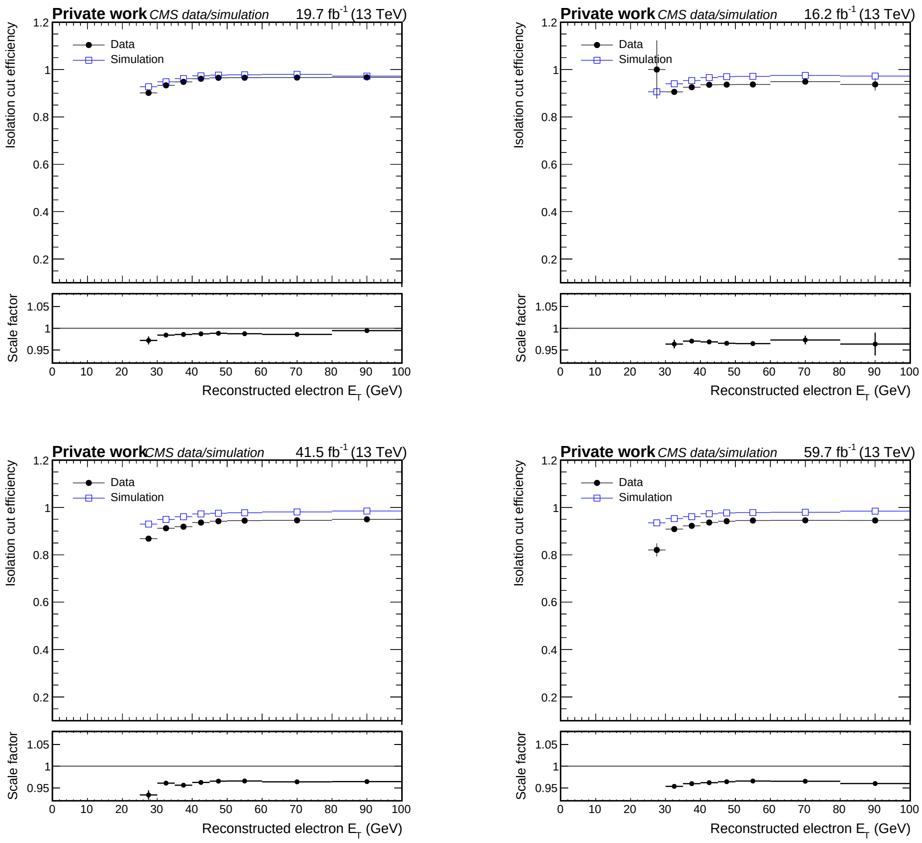


Figure 9.19: Comparison of the efficiency of the electron isolation cut vs E_T in a Drell-Yan enriched control region between data and simulation. The efficiency is shown for the different datasets: 2016 preVFP (top left), 2016 postVFP (top right), 2017 (bottom left), and 2018 (bottom right).

- `HLT_DoubleEle33_CaloIdL` for 2016
- `HLT_DoublePhoton33_CaloIdL` for 2017 and 2018

For the muon isolation, the analysis triggers are used as they do not apply any selection on isolation.

A comparison of the efficiency of the isolation criteria between data and simulation as a function of the lepton p_T for this control region is shown in Figures 9.19 and 9.20 for electrons and muons, respectively. In order to correct the simulation to match the data distribution, a weight $\epsilon_{\text{data}}/\epsilon_{\text{simulation}}$ is applied to the leptons in the signal samples passing the isolation cut and a weight $(1-\epsilon_{\text{data}})/(1-\epsilon_{\text{simulation}})$ is applied to those failing the cut. These weights are obtained in p_T bins as illustrated in Figures 9.19 and 9.20. The ratio of efficiencies between data and simulation lies approximately in the range $0.95-1.01$ $(0.99-1.01)$ for electrons (dGlobal muons) for the different years.

The relative difference between the expected signal yields obtained with and without the application of the weights is taken as the systematic uncertainty

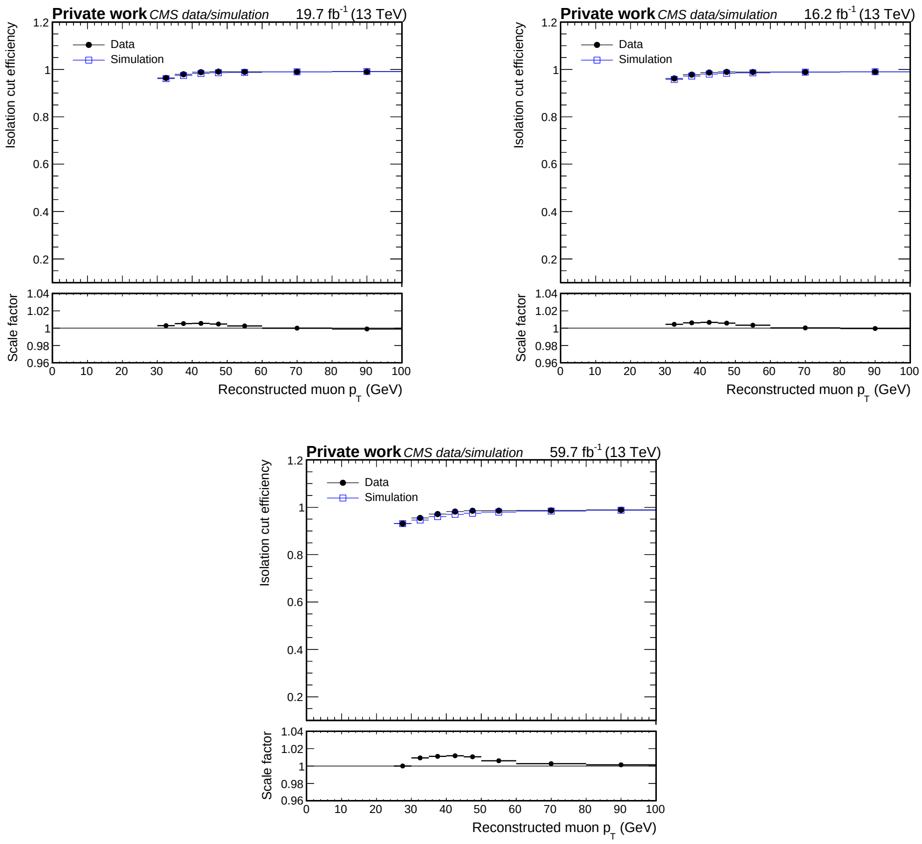


Figure 9.20: Comparison of the efficiency of the dGlobal muon isolation cut vs p_T in a Drell-Yan enriched control region between data and simulation. The efficiency is shown for the different datasets: 2016 preVFP (top left), 2016 postVFP (top right), 2018 (bottom)

associated to the isolation. The size of this uncertainty has been evaluated for different models (considering variations of the involved particle masses and of the long-lived particle $c\tau$), finding similar values as expected since the simulation isolation weights are quite flat with respect to the E_T (p_T) of the electron (muon). Therefore, for each channel and dataset, the same average value that covers for the discrepancy is assigned. For the electron channel this numbers is 7% for the 2016 dataset (pre- and postVFP) and 8% for the 2017 and 2018 datasets. For the muon channel the discrepancy is very small and a value of 2% is used for all the datasets.

9.5. Trigger efficiency corrections

Trigger scale factors $\mathrm{SF}_{\mathrm{trigger}}$ are applied to account for differences in the trigger performance between data and simulation. To derive them, the efficiency is measured in data events unbiased from the trigger used in the search. These factors are derived by measuring the efficiency in data events that remain unbiased with respect to the trigger utilized in the search. The data events are drawn from the MET datasets listed in Table 9.3, which were collected using unprescaled MET triggers specified in Table 9.4. The phase space defined by the HLT-level MET requirements results in a suitable sample for evaluating the efficiency of the triggers. All these MET paths are utilized for measuring the diphoton and dimuon trigger scale factors.

To derive corrections for the dimuon paths, events are required to have at least one vertex fitted from two muon tracks. Additionally, the events must have two reconstructed muons passing muon ID selection and p_T, η, isolation, $\cos(\alpha_{\mu\mu})$, and ΔR cuts defined in the baseline selection. The trigger efficiency is then defined as the fraction of these preselected events that successfully pass the dimuon trigger path specified in Section 6.4.1 for each year.

To obtain corrections for the photon paths, events are preselected based on the presence of at least one vertex fitted from two electron tracks, along with two reconstructed electrons satisfying specific identification criteria related to their transverse energy E_T, pseudorapidity η, and isolation, which are tailored to each year. Requirements of the baseline selection are also required. As with the muon selection, trigger efficiency is defined as the fraction of preselected events that pass the diphoton trigger paths discussed in Section 6.4.2.

The scale factors $\mathrm{SF}_{\mathrm{trigger}}$ are applied per event and are computed as the ratio of the trigger efficiency measured in data $\epsilon^{\mathrm{data}}_{\mathrm{trigger}}$ and simulation $\epsilon^{\mathrm{sim}}_{\mathrm{trigger}}$:

$$\mathrm{SF}_{\mathrm{trigger}} = \frac{\epsilon^{\mathrm{data}}_{\mathrm{trigger}}}{\epsilon^{\mathrm{sim}}_{\mathrm{trigger}}} \tag{9.10}$$

Drell-Yan simulation is used for dimuon and diphoton trigger efficiencies in

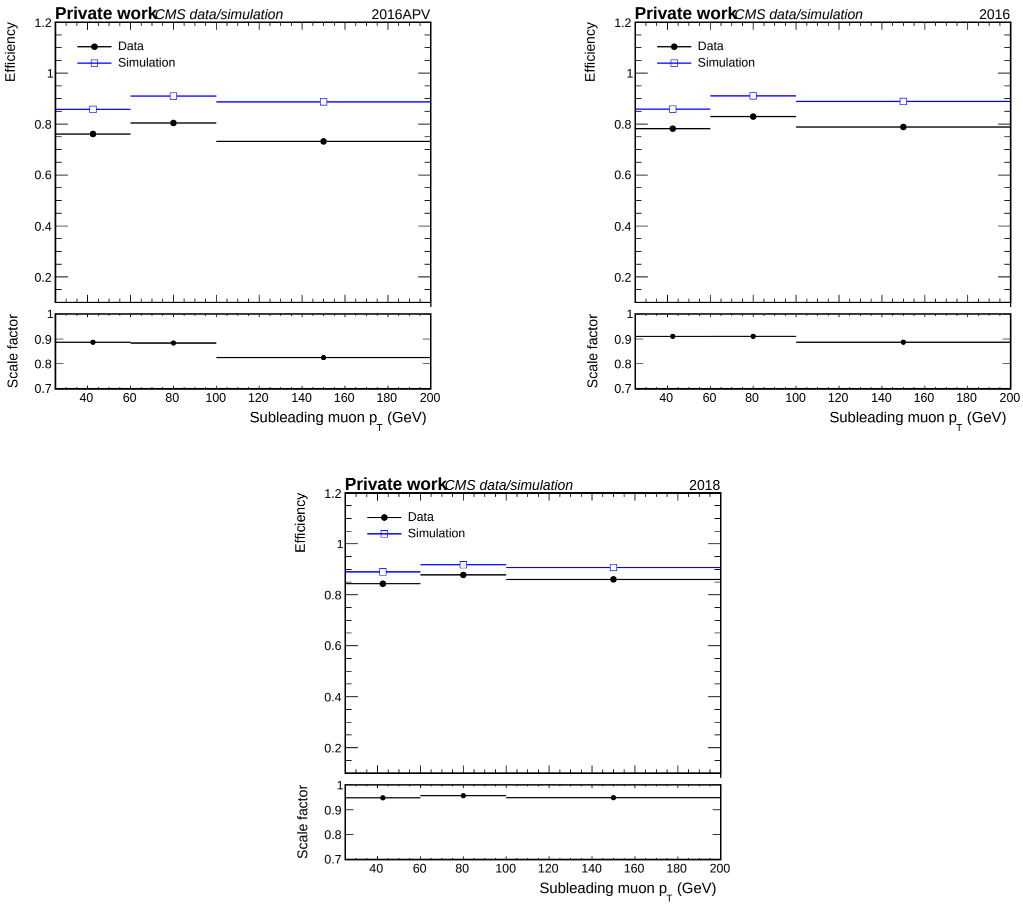


Figure 9.21: Dimuon trigger efficiency measured as a function of subleading muon p_T in data and Drell-Yan simulated events collected with MET triggers.

both 2016 and 2018, since it is the main background of the search. However, for the 2017 diphoton trigger measurement, which requires a selection of $m_{ee} > 100$ GeV and has limited statistics, the dileptonic $t\bar{t}$ simulation is used instead.

The efficiencies and corresponding scale factors are measured as a function of the subleading muon (electron) p_T (E_T) in each case, as it is the variable with the highest correlation to the trigger efficiency. Due to limited statistics, it was not feasible to bin the scale factors in both the leading and subleading lepton variables. However, the subleading p_T (E_T) was found to be sufficient in capturing the trigger dependence and providing reliable corrections. The efficiency results for the dimuon paths are presented in Figure 9.21, while the diphoton path efficiency results are shown in Figure 9.22.

The trigger scale factors are applied per event as a function of the subleading p_T or E_T (in each case) of the selected lepton pair.

Given that the trigger efficiency corrections were derived with an unbiased data sample collected with unprescaled MET triggers, the systematic uncer-

MET datasets used for trigger efficiency measurement	
Year	Datasets (processed AOD used for dimuon triggers)
2016 (with HIP)	/MET/Run2016B-21Feb2020_ver2_UL2016_HIPM-v1/AOD /MET/Run2016C-21Feb2020_UL2016_HIPM-v1/AOD /MET/Run2016D-21Feb2020_UL2016_HIPM-v1/AOD /MET/Run2016E-21Feb2020_UL2016_HIPM-v1/AOD /MET/Run2016F-21Feb2020_UL2016_HIPM-v1/AOD
2016 (without HIP)	/MET/Run2016F-21Feb2020_UL2016-v1/AOD /MET/Run2016G-21Feb2020_UL2016-v1/AOD /MET/Run2016H-21Feb2020_UL2016-v2/AOD
2018	/MET/Run2018A-12Nov2019_UL2018-v3/AOD /MET/Run2018B-12Nov2019_UL2018-v3/AOD /MET/Run2018C-12Nov2019_UL2018_rsb-v1/AOD /MET/Run2018D-12Nov2019_UL2018_rsb-v2/AOD
Year	Datasets (MiniAOD used for diphoton triggers)
2016 (with HIP)	/MET/Run2016B-ver2_HIPM_UL2016_MiniAODv2-v2/MINIAOD /MET/Run2016C-HIPM_UL2016_MiniAODv2-v2/MINIAOD /MET/Run2016D-HIPM_UL2016_MiniAODv2-v2/MINIAOD /MET/Run2016E-HIPM_UL2016_MiniAODv2-v2/MINIAOD /MET/Run2016F-HIPM_UL2016_MiniAODv2-v2/MINIAOD
2016 (without HIP)	/MET/Run2016F-UL2016_MiniAODv2-v2/MINIAOD /MET/Run2016G-UL2016_MiniAODv2-v2/MINIAOD /MET/Run2016H-UL2016_MiniAODv2-v2/MINIAOD
2017	/MET/Run2017B-UL2017_MiniAODv2-v1/MINIAOD /MET/Run2017C-UL2017_MiniAODv2-v1/MINIAOD /MET/Run2017D-UL2017_MiniAODv2-v1/MINIAOD /MET/Run2017E-UL2017_MiniAODv2-v1/MINIAOD /MET/Run2017F-UL2017_MiniAODv2-v1/MINIAOD
2018	/MET/Run2018A-UL2018_MiniAODv2-v2/MINIAOD /MET/Run2018B-UL2018_MiniAODv2-v2/MINIAOD /MET/Run2018C-UL2018_MiniAODv2-v1/MINIAOD /MET/Run2018D-UL2018_MiniAODv2-v1/MINIAOD

Table 9.3: MET datasets used to measure dimuon and diphoton trigger efficiencies and derive corrections for the simulation.

MET trigger used for trigger efficiency measurement	
Year	Triggers
2016	HLT_PFMET120_PFMHT90_IDTight HLT_PFMET120_PFMHT100_IDTight HLT_PFMET120_PFMHT110_IDTight HLT_PFMET120_PFMHT120_IDTight HLT_MET200 HLT_MonoCentralPFJet80_PFMETNoMu110_PFMHTNoMu110_IDTight HLT_PFMET170_HBHECleaned HLT_PFMET300 HLT_PFMETNoMu120_PFMHTNoMu120_IDTight
2017	HLT_PFMET120_PFMHT120_IDTight HLT_PFMET120_PFMHT120_IDTight_PFHT60 HLT_CaloMET350_HBHECleaned HLT_MonoCentralPFJet80_PFMETNoMu120_PFMHTNoMu120_IDTight HLT_PFMET250_HBHECleaned HLT_PFMETNoMu120_PFMHTNoMu120_IDTight
2018	HLT_PFMET120_PFMHT120_IDTight HLT_PFMET120_PFMHT120_IDTight_PFHT60 HLT_CaloMET350_HBHECleaned HLT_MonoCentralPFJet80_PFMETNoMu120_PFMHTNoMu120_IDTight HLT_PFMET250_HBHECleaned HLT_PFMET200_HBHE_BeamHaloCleaned HLT_PFMETNoMu120_PFMHTNoMu120_IDTight

Table 9.4: MET prescaled triggers used to select the events to measure trigger efficiencies in each year.

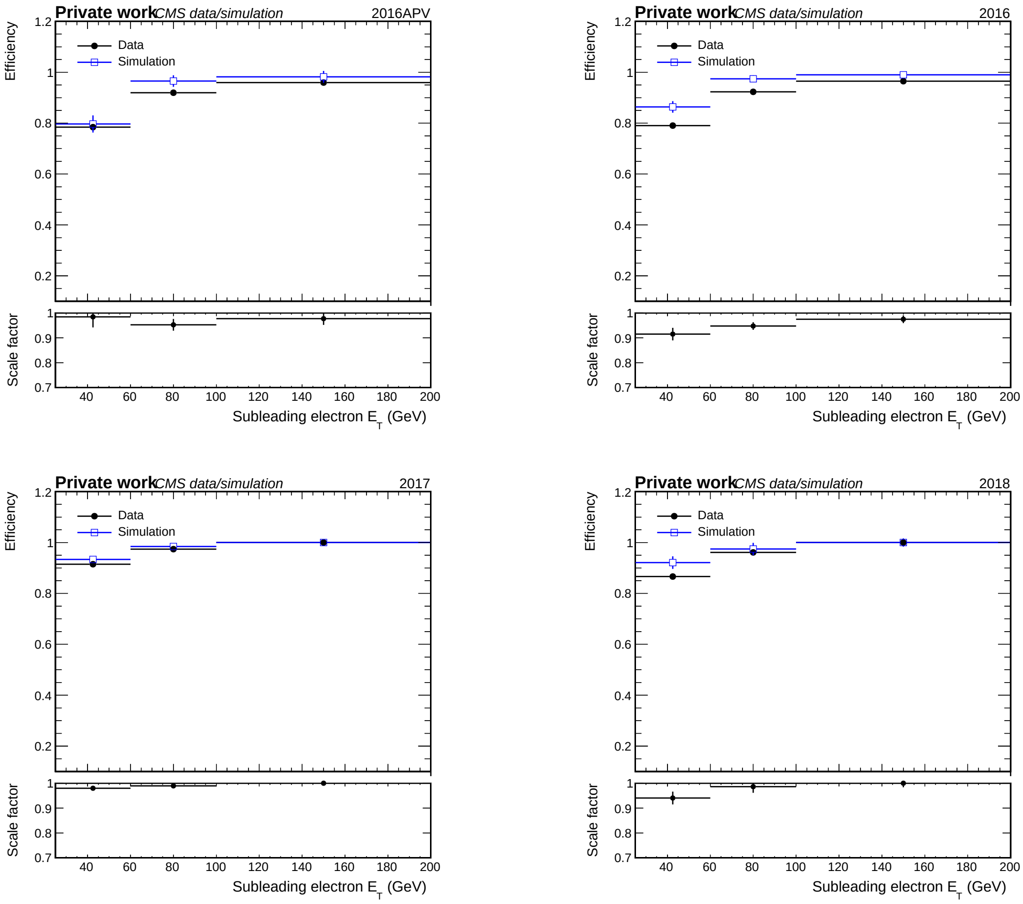


Figure 9.22: Diphoton trigger efficiency measured as a function of subleading electron E_T in data collected with MET triggers. Drell-Yan simulation was used for 2016 and 2018 while the 2017 measurement was done with dileptonic $t\bar{t}$ simulation.

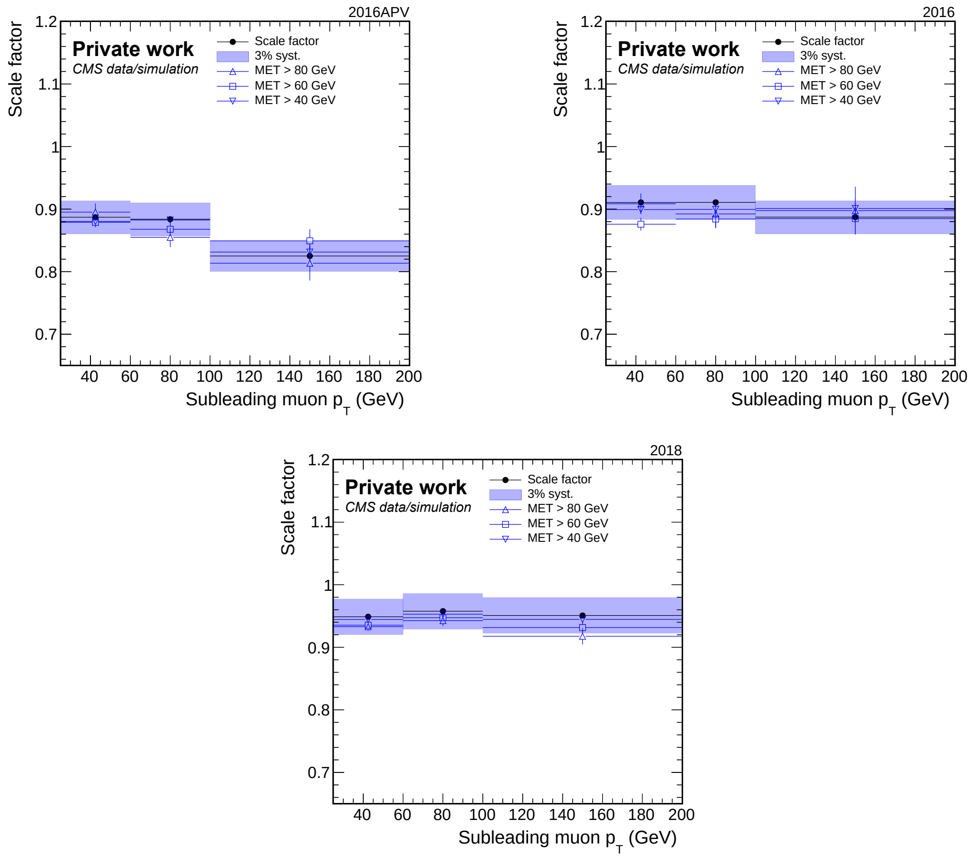


Figure 9.23: Muon trigger scale factor variations for different PFMET variations in 2016APV (left), 2016 (right) and 2018 (bottom) data.

tainty associated to this measurement is derived by applying several offline MET selections on the events and studying the variation of the derived scale factors. The values of the scale factors are found to fluctuate by 3% for the dimuon paths and by 1.5% for the diphoton paths and these two values are finally identified with the systematic uncertainties that are applied to the events of each channel. In Figures 9.23 and we 9.24 show the scale factor variations obtained after requiring the events to pass an offline MET selection of PFMET > 40, 60, 80 GeV. Additionally, the systematic uncertainties are represented with blue bands and are proven to cover the scale factor fluctuations and the robustness of the method.

There is also another source of uncertainty that comes from the fact that the trigger performance was measured in prompt simulation and data events that will likely contain prompt processes. To guarantee that our corrections are also valid in more displaced regimes, which is a more realistic scenario for this kind of search, we apply an additional systematic uncertainty that accounts for deviations in the trigger performance. We use simulated H→SS decays with several masses where the events are required to pass just the same offline selection applied to derive the scale factors i.e. the leptonic analysis selection. No

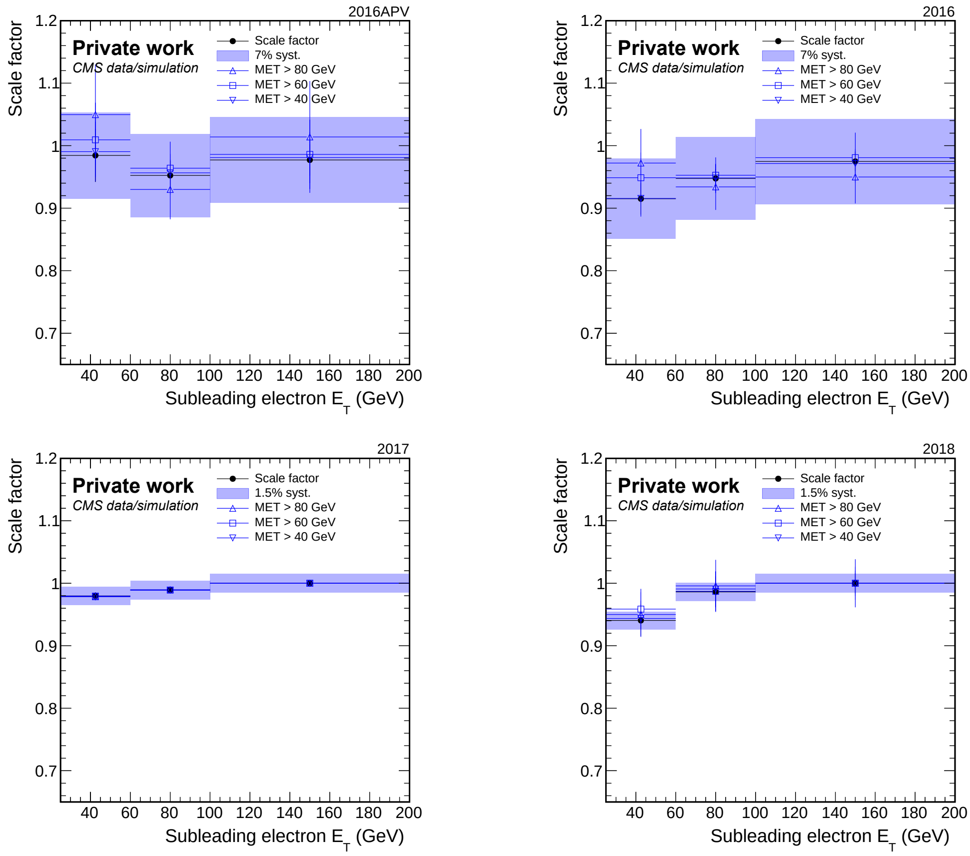


Figure 9.24: Photon trigger scale factor variations for different PFMET variations in 2016APV (upper left), 2016 (upper right), 2017 (bottom left) and 2018 (bottom right) data.

trigger path was required for this derivation. The trigger efficiency is identified as the fraction of simulated events that pass the analysis triggers and it is measured as a function of the maximum transverse impact parameter max($|d_{xy}|$) of the two leptons with the highest p_T.

It was found that for the electrons the variations in the efficiency more or less stable for the different mass points (m_H, m_S) and that can be covered within a systematic uncertainty of 1% for (800, 350) and (600, 250), and of 1.5% for (500, 150). These variations are constant over the years. The largest variation is taken as the systematic uncertainty, which is identified with 1% for the three years of data taking. The picture for the muon triggers is a little bit different as the variations are found to be higher. (800, 350) simulation fluctuates by 3% in 2016 and by 1% in 2018, (600, 250) by 3% in 2016 and 3% in 2018 and (500, 150) by 6% in 2016 and by 2% in 2018. The systematic applied in this case is identified with 6% for 2016 and 2% for 2018 data. All of these systematics are treated as correlated among the years since the HLT reconstructions used in each case are supposed to be very similar (if not the same ones in most cases).

The signal efficiencies are presented in Figures 9.25 and 9.26 for muon and electron paths, respectively. The variations for each mass point that are qualified as described in the last paragraph are represented as color bands, just for visualization.

9.6. Data vs simulation comparison

Each one of the scale factors explained in Sections 9.1-9.5 intend to account for a different effect that may be miss-modelled by the simulation. In order to evaluate the performance of the derived corrections several data distributions are compared with the distributions produced with the already corrected Monte Carlo simulation. The purpose of this comparison is to ensure that the corrections derived for the various effects are accurately correcting the simulation and improving its agreement with data in regions where the LLP signal contribution is negligible. The Backward Control Region with $|\Delta\Phi| > \pi/2$ dilepton vertices is chosen as the background control region, as it contains all the potential background contributions of the search that remain after the vertex selection.

The studied variables are the p_T of both leptons fitting the vertex, their $|\eta|$ and the dilepton invariant mass. Since the corrections are year-dependent, their distributions are shown in Figures 9.27 to 9.29 for data and simulation of the years 2016, 2017 and 2018 separately.

A subtancial disagreement was found in HIP affected runs for the electron channel. The origin is likely due to the reconstruction scale factor. This correction was measured with runs G and H of the 2016 data-taking era and cos-

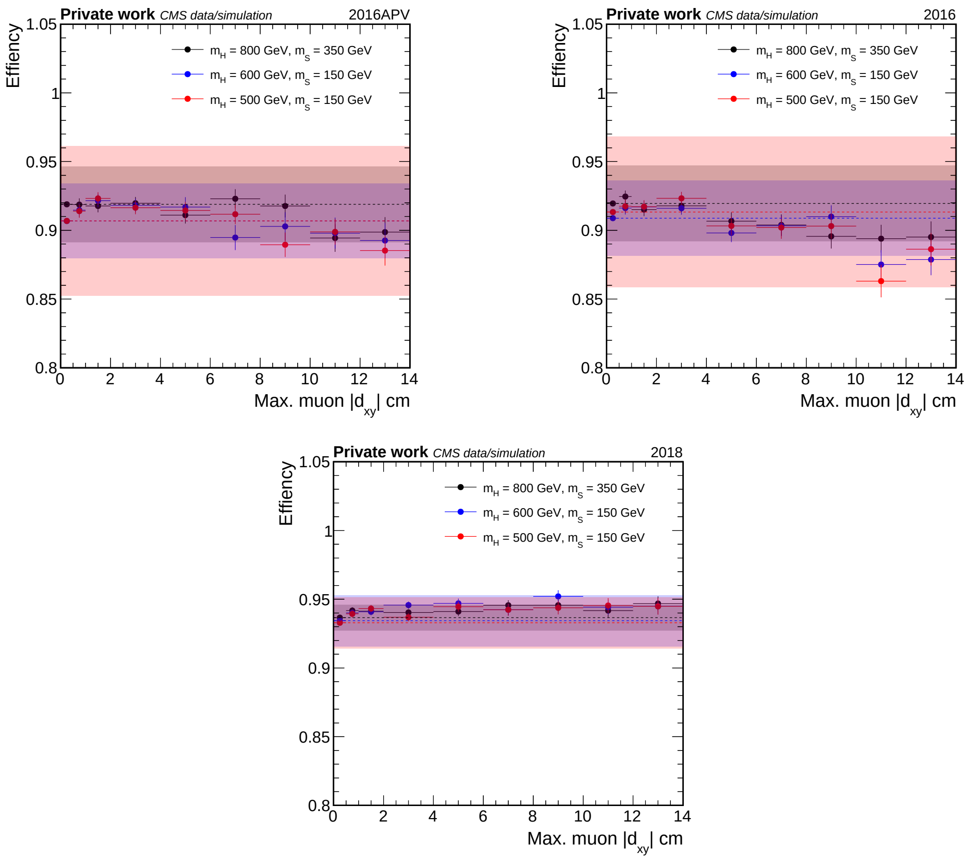


Figure 9.25: Muon trigger efficiency measured in simulated H→SS decays with several masses for 2016APV (left), 2016 (right) and 2018 (bottom) eras.

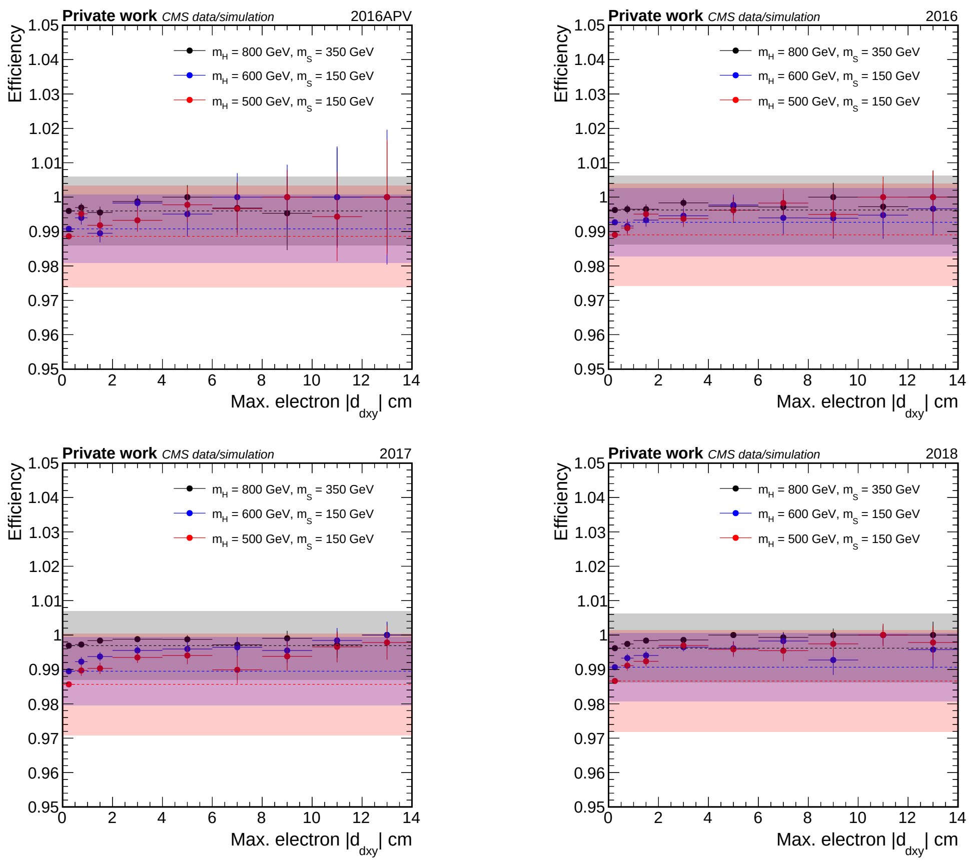


Figure 9.26: Photon trigger efficiency measured in simulated H→SS decays with several masses for 2016APV (upper left), 2016 (upper right), 2017 (bottom left) and 2018 (bottom right) eras.

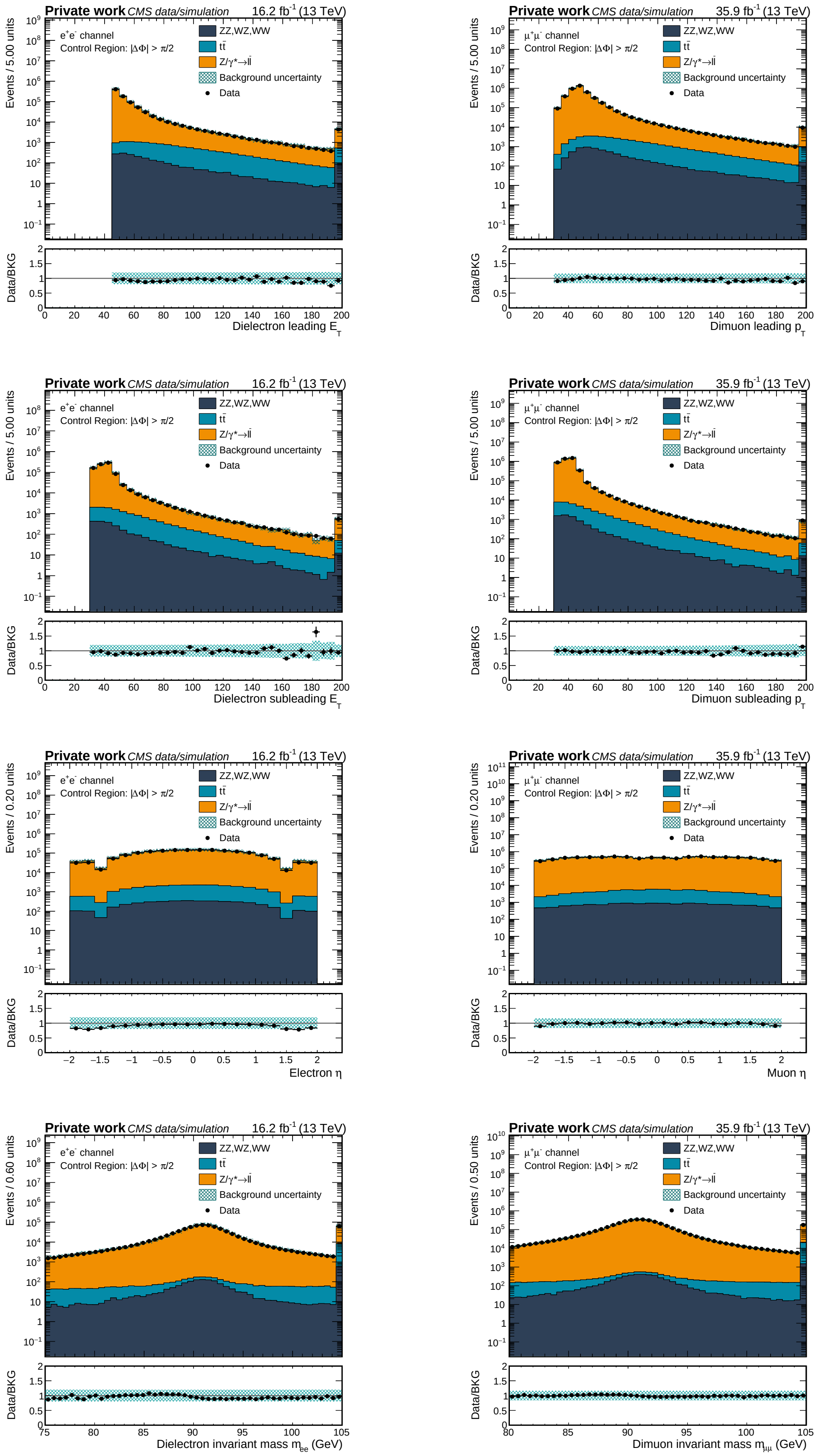


Figure 9.27: Data vs Monte Carlo simulation comparison corresponding to year 2016. Dilepton vertices reconstructed in the Backward Control Region are shown. HIP affected runs are excluded from dielectron histograms.

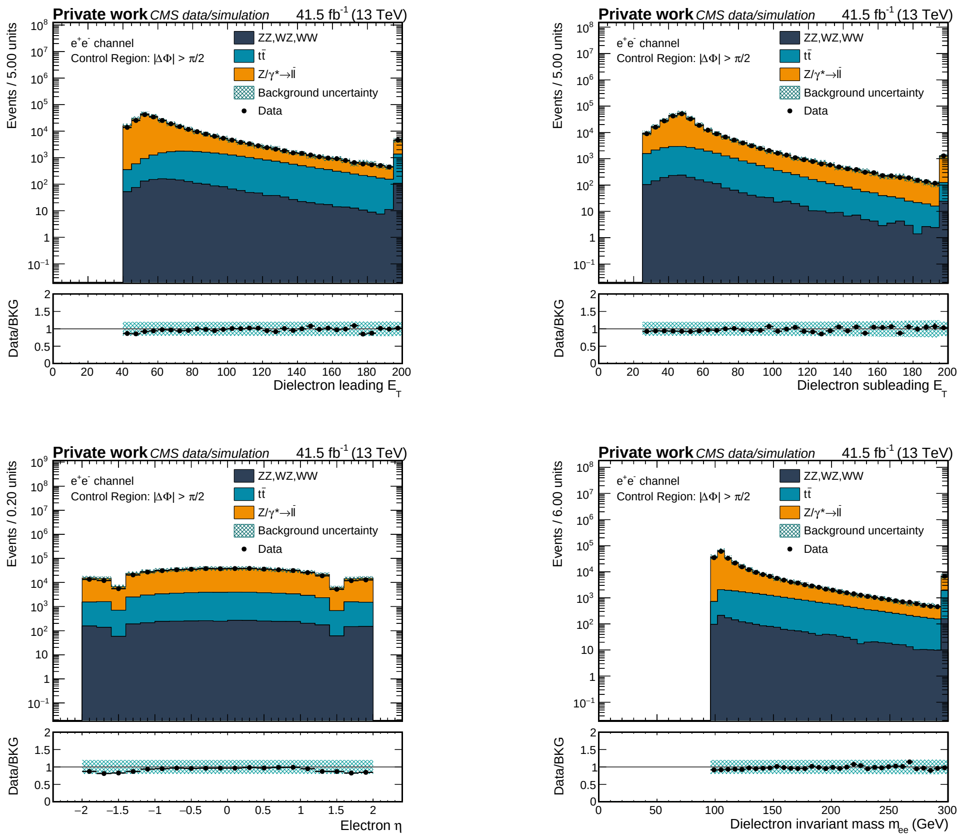


Figure 9.28: Data vs Monte Carlo simulation comparison corresponding to year 2017. Dielectron vertices reconstructed in the Backward Control Region are shown.

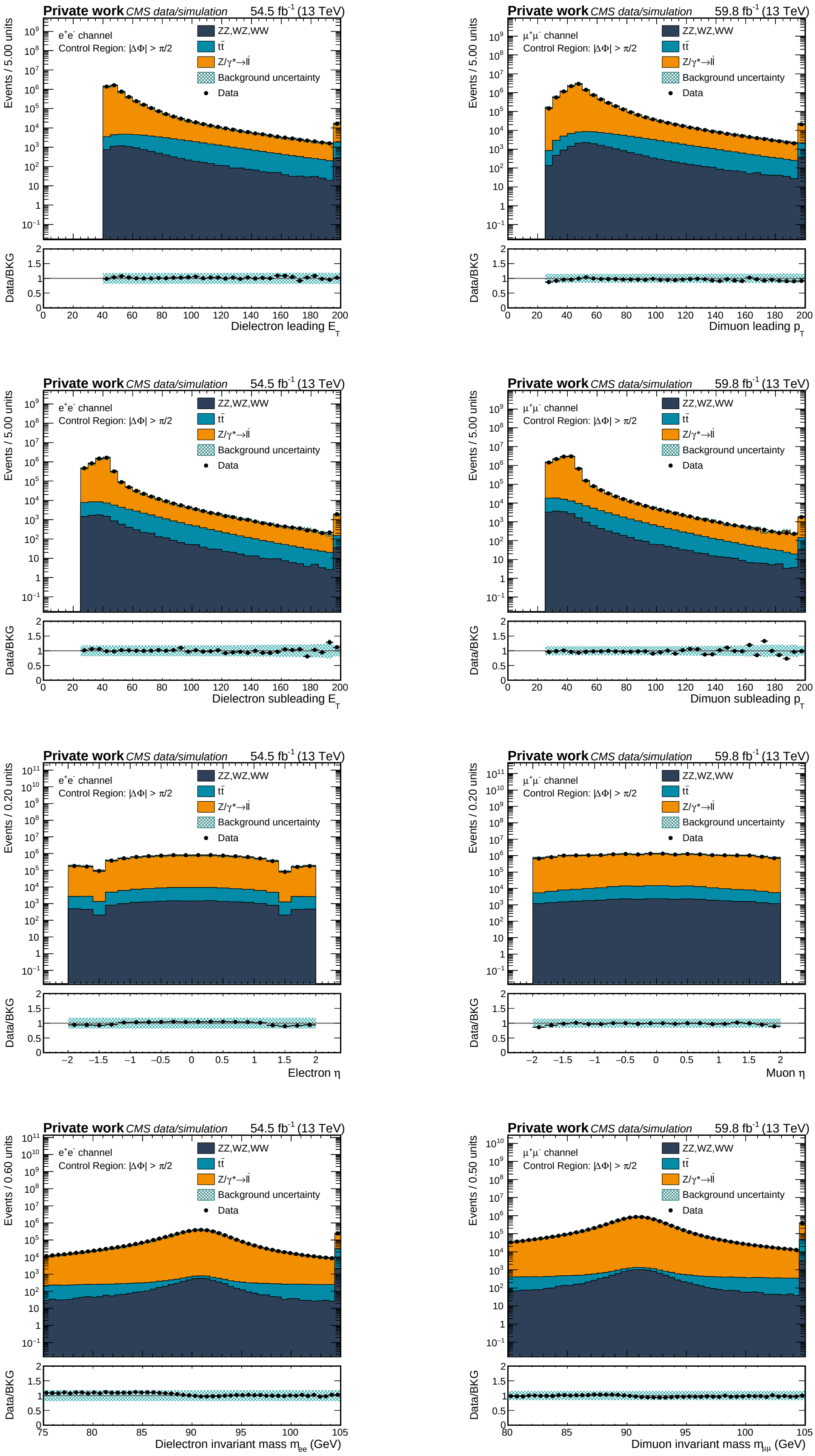


Figure 9.29: Data vs Monte Carlo simulation comparison corresponding to year 2018. Dilepton vertices reconstructed in the Backward Control Region are shown.

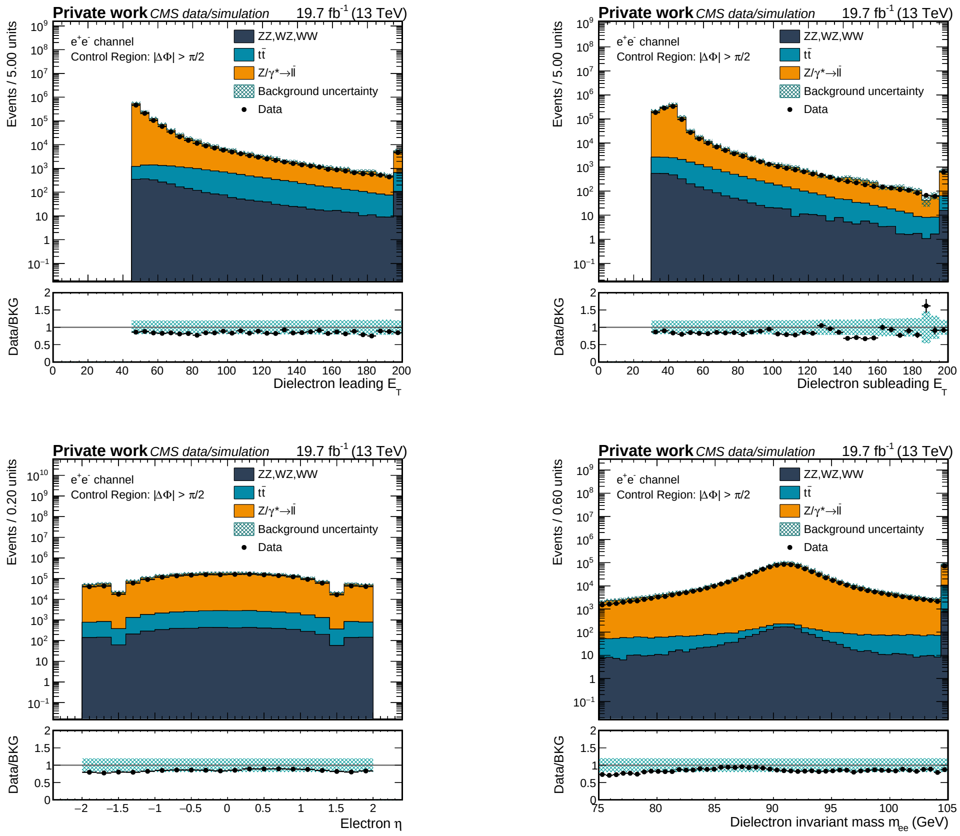


Figure 9.30: Data vs Monte Carlo simulation comparison corresponding to year 2016 HIP affected runs. Dielectron vertices reconstructed in the Backward Control Region are shown.

mic muon Monte Carlo where HIP conditions are not simulated[5]. In isotracks, where the effects on tracking may be strong, the effects observed translate into a overestimation of the efficiency in the simulation. No sample with such conditions simulated is produced for cosmic muons. Since the impact of removing these runs from the analysis is manageable and the sensitivity is not much affected, it was finally decided to exclude them when extracting the final results.

The comparisons for HIP affected runs in electron channels are presented in Figure 9.30.

9.7. Summary of systematic uncertainties

The summary of the systematic uncertainties considered in the search is presented in Table 9.5.

[5]Cosmic sample simulated under HIP conditions do not exist centrally

Systematic uncertainty	Muon channel		Electron channel		
	2016	2018	2016	2017	2018
Luminosity	1.2%	2.5%	1.2%	2.3%	2.5%
Pileup	2%				
Trigger efficiency (Method)	3%		7%	1.5%	
Trigger efficiency (Displacement)	6%	2%	1.5%		
Displaced lepton efficiency (Reco + ID)	11%		15%		13%
Isolation efficiency	2%		7%	8%	
Background estimation	10%		10%		

Table 9.5: Summary of systematic uncertainties applied to each channel for the results.

Chapter 10

Results and interpretation

10.1. Signal Region yields

Selected events of each channel populate the different bins defined for the Signal Regions that were introduced in Section 7.3. Depending on the masses, kinematics and lifetimes assumed for the exotic particles, the signal is expected to populate different bins. The estimated background and the expected yields for five different signals in each subregion of the SR are collected in Tables 10.1-10.2. These yields are also represented in Figures 10.1-10.2. The data observation has been also added to the plots once the signal region bins have been unblinded.

The presented results assume pure lepton branching ratios of $\mathcal{B}(\mathrm{S} \rightarrow ee) = 100\%$ for the yields measured in the electron channel and $\mathcal{B}(\mathrm{S} \rightarrow \mu\mu) = 100\%$ for the yields measured in the muon channel. No channel orthogonality is established for this representation as the results are presented separately for each channel.

The yields of the presented signals correspond to LLPs with lifetimes of $c\tau = 10$ cm. The chosen mass points encompass different combinations:

- H→SS ($m_H = 1000$ GeV, $m_S = 250$ GeV)
- H→SS ($m_H = 300$ GeV, $m_S = 50$ GeV)
- H→SS ($m_H = 400$ GeV, $m_S = 150$ GeV)
- H→SS ($m_H = 500$ GeV, $m_S = 50$ GeV)
- H→SS ($m_H = 800$ GeV, $m_S = 50$ GeV)

After unblinding, in general a good agreement between prediction and observation is found in all sub-regions. For the electrons, when the three years are combined, the largest deviation is found in the high multiplicity bin SRII with a significance of 1.6σ. This is mainly introduced by 2016 data and consistent with a fluctuation. For the muons, no significant excesses are observed. Instead, the observed data is below the prediction in some sub-regions where the statistics is limited. The largest deviation has a significance of 2.3σ and is compatible with a fluctuation taking into account the total number of signal regions, and the fact that is only observed in one of the years.

Electron channel 2016							
Region	Bkg	Obs	Signal yields for (m_H, m_S) GeV with $c\tau = 10$ cm				
			$(1000, 250)$	$(300, 50)$	$(400, 150)$	$(500, 50)$	$(800, 50)$
IaA	456	458	0.3	189.5	0.0	105.4	13.5
IaB	210	171	0.8	547.7	13.0	416.8	63.4
IaC	2	3	10.4	9570.1	1019.2	6210.8	640.1
IbA	102	98	5.8	13.0	228.7	0.0	0.8
IbB	31	30	18.2	12.8	723.8	28.8	3.1
IbC	1	0	416.2	127.1	13255.9	71.0	3.7
II	3	8	115.9	1389.6	3991.2	1001.6	93.6

Electron channel 2017							
Region	Bkg	Obs	Signal yields for (m_H, m_S) GeV with $c\tau = 10$ cm				
			$(1000, 250)$	$(300, 50)$	$(400, 150)$	$(500, 50)$	$(800, 50)$
IaA	0	0	0.0	0.0	0.0	0.0	0.0
IaB	0	0	0.0	0.0	0.0	0.0	0.0
IaC	0	0	0.0	0.0	0.0	0.0	0.0
IbA	235	254	14.1	32.1	535.3	28.3	0.0
IbB	102	81	38.7	31.4	1391.7	8.6	4.9
IbC	4	1	1038.3	174.4	33151.5	147.1	8.6
II	4	4	284.7	0.0	8188.1	0.0	5.1

Electron channel 2018							
Region	Bkg	Obs	Signal yields for (m_H, m_S) GeV with $c\tau = 10$ cm				
			$(1000, 250)$	$(300, 50)$	$(400, 150)$	$(500, 50)$	$(800, 50)$
IaA	2582	2514	0.0	449.0	41.9	420.1	42.2
IaB	1598	1376	2.3	1639.9	238.4	1449.0	155.4
IaC	25	18	35.4	27071.6	3008.6	20865.0	1977.6
IbA	344	337	17.8	0.0	556.3	9.9	0.0
IbB	133	116	37.1	38.3	1730.4	59.7	8.2
IbC	2	5	1220.3	88.6	40556.4	115.0	22.4
II	5	10	322.1	2987.1	10442.6	2691.7	247.3

Table 10.1: Signal and background yields for dielectron vertices. Data from 2016, 2017 and 2018.

10.2. Interpretation

The results are interpreted in the context of the Higgs model presented in the previous sections. Maximum likelihood fits are performed under either

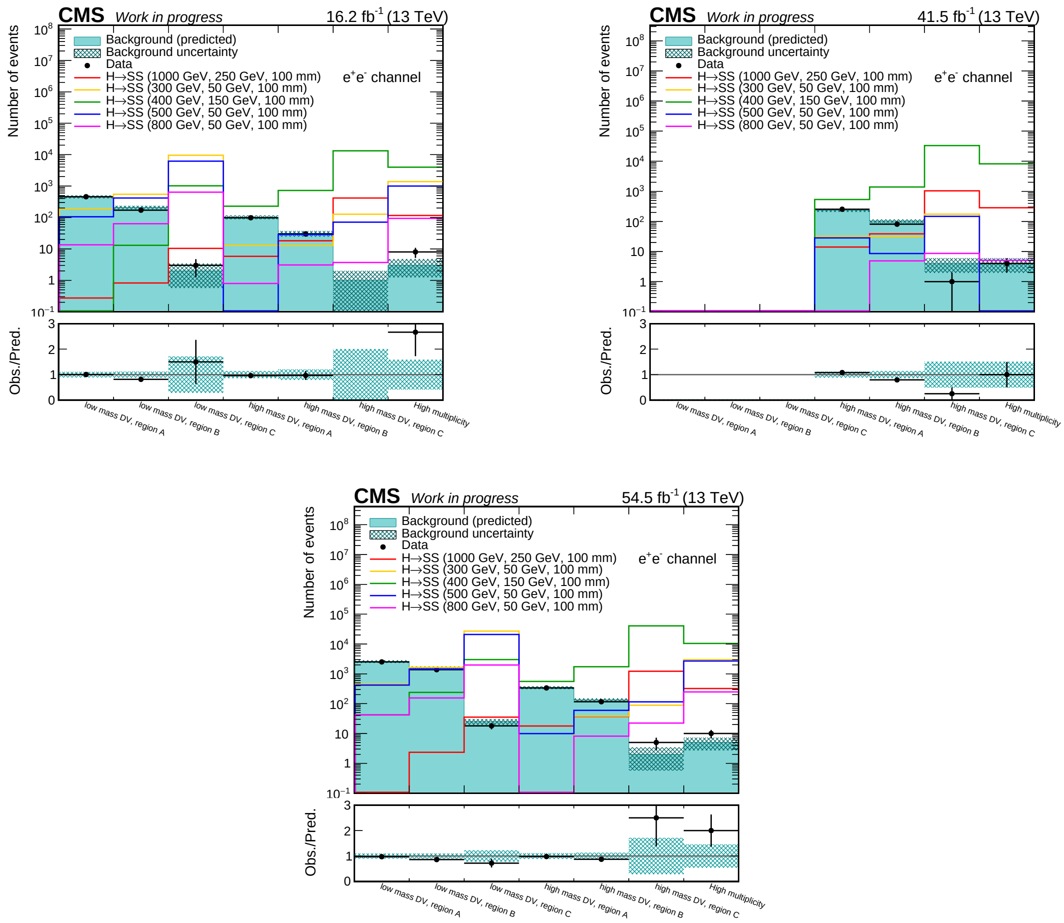


Figure 10.1: Signal region yields for dielectron vertices in 2016, 2017 and 2018. Branching ratios of $\mathcal{B}(\mathrm{S} \to ee) = 100\%$ are assumed. Channel orthogonality is not established.

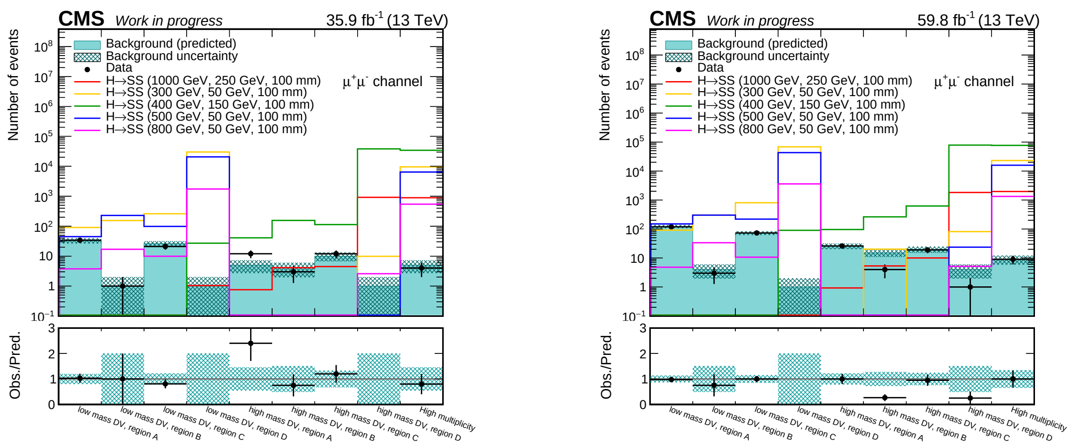


Figure 10.2: Signal region yields for dimuon vertices in 2016 and 2018. Branching ratios of $\mathcal{B}(\mathrm{S} \to \mu\mu) = 100\%$ are assumed. Channel orthogonality is not established.

Muon channel 2016							
Region	Bkg	Obs	Signal yields for (m_H, m_S) GeV with $c\tau = 10$ cm				
			$(1000, 250)$	$(300, 50)$	$(400, 150)$	$(500, 50)$	$(800, 50)$
IaA	33	34	0.0	90.4	0.0	44.7	3.8
IaB	1	1	0.0	154.4	0.0	228.9	16.9
IaC	26	21	0.0	260.1	0.0	98.4	9.9
IaD	1	0	1.0	30063.0	27.0	20662.9	1744.0
IbA	5	12	0.8	0.0	40.9	0.0	0.0
IbB	4	3	4.1	0.0	155.8	0.0	0.0
IbC	10	12	4.5	0.0	113.1	0.0	0.0
IbD	1	0	920.9	9.9	38209.7	0.0	2.6
II	5	4	891.1	9574.6	34298.2	6439.4	543.2

Muon channel 2018							
Region	Bkg	Obs	Signal yields for (m_H, m_S) GeV with $c\tau = 10$ cm				
			$(1000, 250)$	$(300, 50)$	$(400, 150)$	$(500, 50)$	$(800, 50)$
IaA	122	119	0.0	91.2	0.0	149.8	4.8
IaB	4	3	0.0	300.0	0.0	300.7	33.6
IaC	73	73	0.0	804.7	0.0	219.1	10.7
IaD	1	0	0.0	68203.4	90.1	43021.5	3589.8
IbA	26	26	0.9	0.0	96.1	0.0	0.0
IbB	15	4	5.3	20.2	262.4	0.0	0.0
IbC	20	19	10.0	0.0	621.5	0.0	0.0
IbD	4	1	1818.8	81.3	78065.0	23.6	5.2
II	9	9	1962.2	22946.9	76465.2	15808.3	1329.0

Table 10.2: Signal and background yields for dimuon vertices. Data from 2016 and 2018.

background-only or signal+background hypothesis to the data in the signal regions defined in Section 7.3. The uncertainties in the modeling of the backgrounds and on the scale factors derived for the signals are used in the fitting procedure as nuisance parameters. The likelihoods were constructed as the product of Poisson probability density functions, one for each SR, with the log-normal terms to constrain the nuisance parameters. The signal+background fits are used to set 95% confidence level (CL) upper limits on the production cross-sections of the modeled signals. A modified frequentist approach was employed, using the CLs criterion and relying on asymptotic approximations, to calculate the distribution of the profile likelihood test statistic [101–104].

The results are obtained by assuming different interpretations of the long-lived S boson decay. They are presented separately for several H and S masses combinations and for lifetimes in the range from 0.1 to 1000 cm. The figures shown combine the full luminosity studied for each channel, that is, 96 fb^{-1} for the muon channel, and 112 fb^{-1} for the electron channel. Expected results were first obtained with data blinded in the signal region of the search. Once the analysis was unblinded, the observed limits were also added. Upper limits are set on the signal production cross-section times the branching fraction to LL S i.e. $\sigma(\mathrm{H}) \times \mathcal{B}(\mathrm{H} \rightarrow \mathrm{SS})$.

Firstly, upper limits are obtained by considering only pure muon decays, that is, an assumed branching ratio of $\mathcal{B}(\mathrm{S} \rightarrow \mu\mu) = 100\%$ for the long-lived boson. In this case, only events collected in the muon channel are used to measure the signal yields. As the signal region splitting depends on the number of preselected dimuon vertices $N_{\mu\mu}$, the results are sensitive to have 4μ or $2\mu+X$ in the final state even if the reconstruction of just one vertex is required. To keep the results of the search unbiased from the final state, only 4μ events of the signal simulations are used to derive these results. They are presented for the masses and lifetimes considered in Figures 10.3-10.8. The results are compared with those presented by previous displaced leptons analyses i.e. [79] and [80].

Secondly, upper limits are derived by considering a full decay rate to electrons with a branching ratio of $\mathcal{B}(\mathrm{S} \rightarrow ee) = 100\%$. Only events of the electron channel are used to measure these results without requiring any channel orthogonality as dimuon events are not utilised. In the same way as for muons, the signal region splitting depends on the number of reconstructed dielectron vertices N_{ee} and for this reason only $4e$ signal events are used to derive the results. The upper limits for this interpretation are shown in Figures 10.9-10.14. They are compared with the upper limits of the same interpretation measured by [79]. Under this assumption, the upper limit on $\sigma(\mathrm{H}) \times \mathcal{B}(\mathrm{H} \rightarrow \mathrm{SS})$ is higher than the one obtained when assuming muon decays only. This is due to the higher reconstruction efficiency of the muons with respect to the electrons and also to the fact that the background yields are smaller for the former. However, the presented limits for $\mathcal{B}(\mathrm{S} \rightarrow ee) = 100\%$ decays are the most stringent ones obtained with Run 2.

Finally, upper limits are extracted by considering the same decay rate to electrons and muons i.e. a branching ratio of $\mathcal{B}(\mathrm{S} \rightarrow ee) = \mathcal{B}(\mathrm{S} \rightarrow \mu\mu) = 50\%$. When extracting the combined upper limit both electron and muon channel events are considered. The orthogonality between them is established by assigning the overlapping events to the muon channel. This is done by removing the events that have fired the dimuon paths and have a valid dimuon vertex from the electron channel. In this interpretation, the muon channel is prioritized in light of the results presented for the $\mathcal{B}(\mathrm{S} \rightarrow \mu\mu) = 100\%$ and $\mathcal{B}(\mathrm{S} \rightarrow ee) = 100\%$ interpretations where the muon channel proved to provide better sensitivity. The resulting upper limits are collected in Figures 10.15-10.20. They are also compared to those measured by [79]. Stringent limits for this interpretation are obtained for lifetimes over 100 cm for the majority of the masses considered.

Specifically, for the $m_{\mathrm{H}} = 125$ GeV case, upper limits on the branching fraction $\mathcal{B}(\mathrm{H} \rightarrow \mathrm{SS})$ are also derived and presented in Figure 10.21. The assumed production cross section is 21.460 ± 0.016 pb which is the value provided by Powheg for this model. The upper limits are obtained under the three different

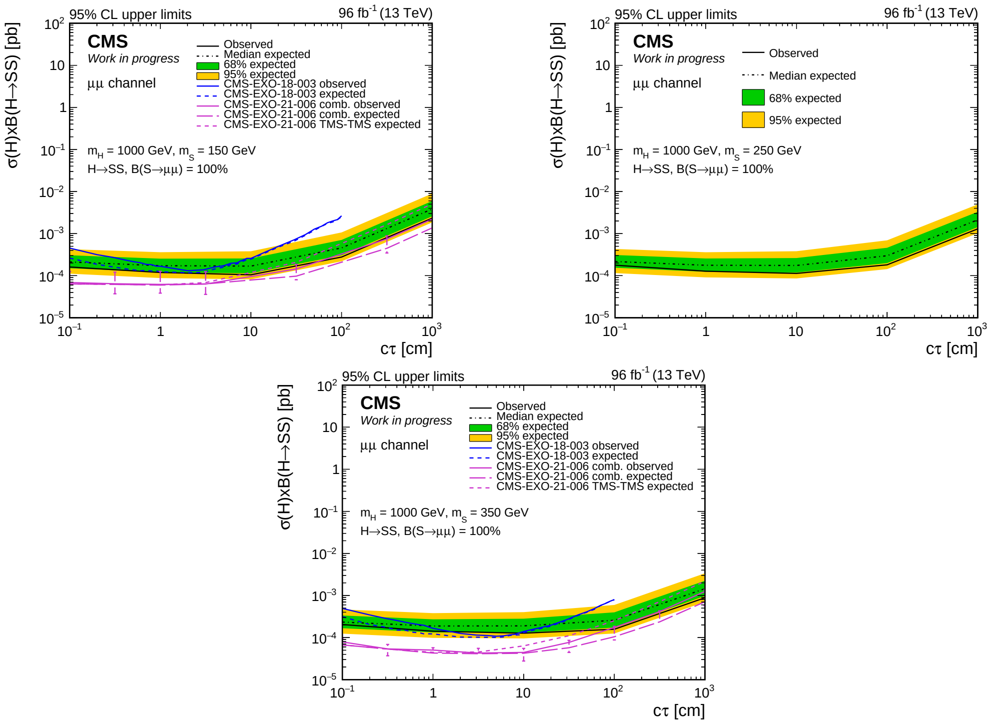


Figure 10.3: 95% CL upper limits set on $\sigma(\mathrm{H}) \times \mathcal{B}(\mathrm{H} \to \mathrm{SS})$ with masses $m_{\mathrm{H}} = 1000$ GeV and $m_{\mathrm{S}} = 150, 250, 350$ GeV (left, right, bottom). Long-lived S decays to muons with a branching fraction of $\mathcal{B}(\mathrm{S} \to \mu\mu) = 100\%$ are assumed. Only muon channel events were used.

aforementioned interpretations. In this case, the upper limits obtained when assuming pure muon decays are higher than the ones provided by [80], which remain the stringent ones presented by the CMS Collaboration. Nevertheless, the upper limits on $\mathcal{B}(\mathrm{H} \to \mathrm{SS})$ presented in this analysis are the most stringent ones up to date for both the $\mathcal{B}(\mathrm{S} \to ee) = 100\%$ and $\mathcal{B}(\mathrm{S} \to ee) = \mathcal{B}(\mathrm{S} \to \mu\mu) = 50\%$ interpretations.

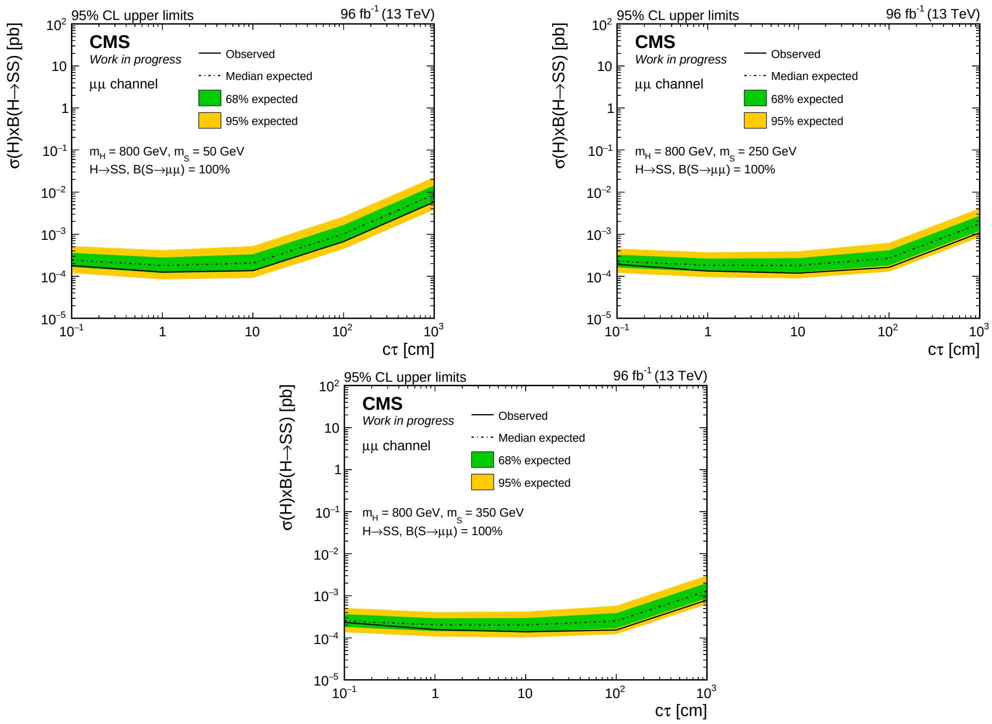


Figure 10.4: 95% CL upper limits set on $\sigma(\mathrm{H}) \times \mathcal{B}(\mathrm{H} \to \mathrm{SS})$ with masses $m_{\mathrm{H}} = 800$ GeV and $m_{\mathrm{S}} = 50, 250, 350$ GeV (left, right, bottom). Long-lived S decays to muons with a branching fraction of $\mathcal{B}(\mathrm{S} \to \mu\mu) = 100\%$ are assumed. Only muon channel events were used.

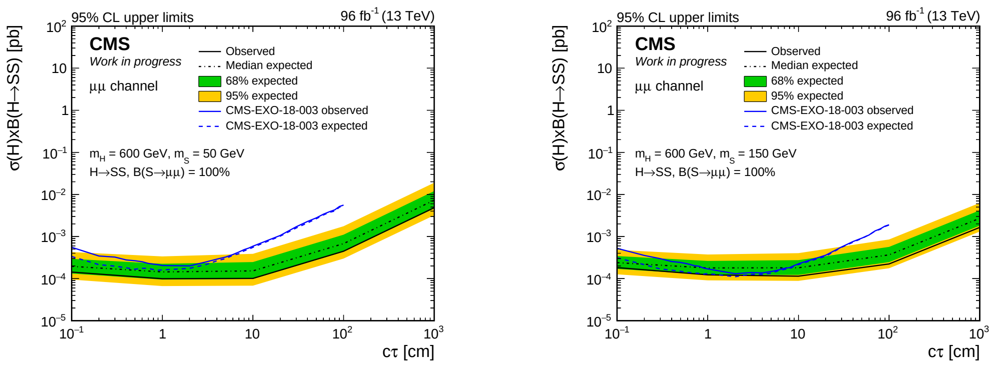


Figure 10.5: 95% CL upper limits set on $\sigma(\mathrm{H}) \times \mathcal{B}(\mathrm{H} \to \mathrm{SS})$ with masses $m_{\mathrm{H}} = 600$ GeV and $m_{\mathrm{S}} = 50, 150$ GeV (left, right). Long-lived S decays to muons with a branching fraction of $\mathcal{B}(\mathrm{S} \to \mu\mu) = 100\%$ are assumed. Only muon channel events were used.

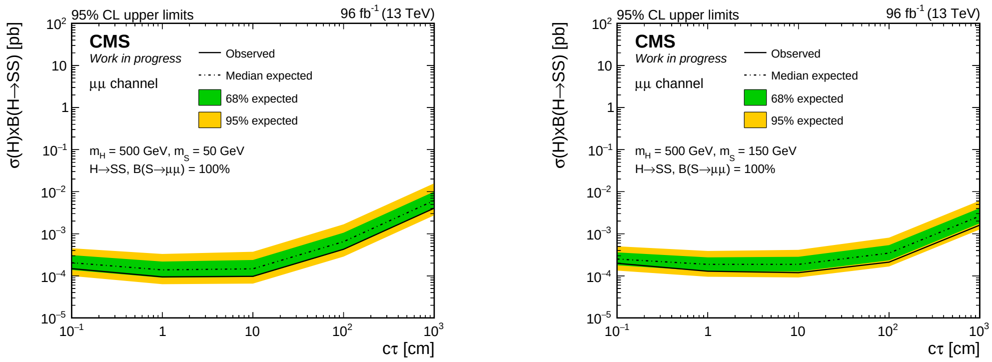


Figure 10.6: 95% CL upper limits set on $\sigma(\mathrm{H}) \times \mathcal{B}(\mathrm{H} \to \mathrm{SS})$ with masses $m_{\mathrm{H}} = 500$ GeV and $m_{\mathrm{S}} = 50, 150$ GeV (left, right). Long-lived S decays to muons with a branching fraction of $\mathcal{B}(\mathrm{S} \to \mu\mu) = 100\%$ are assumed. Only muon channel events were used.

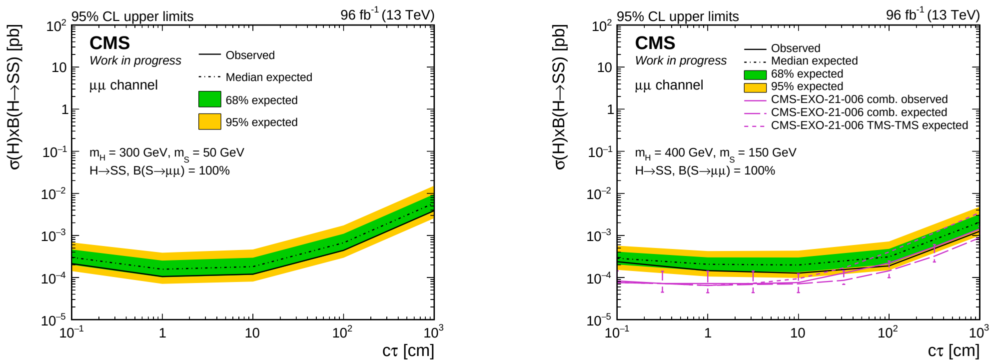


Figure 10.7: 95% CL upper limits set on $\sigma(\mathrm{H}) \times \mathcal{B}(\mathrm{H} \to \mathrm{SS})$ with masses $m_{\mathrm{H}} = 300$ GeV and $m_{\mathrm{S}} = 50$ GeV (left) and $m_{\mathrm{H}} = 400$ GeV and $m_{\mathrm{S}} = 150$ GeV (right). Long-lived S decays to muons with a branching fraction of $\mathcal{B}(\mathrm{S} \to \mu\mu) = 100\%$ are assumed. Only muon channel events were used.

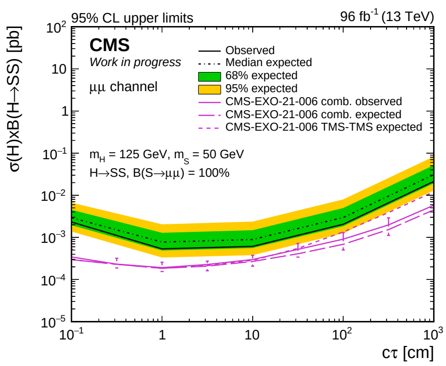


Figure 10.8: 95% CL upper limits set on $\sigma(\mathrm{H}) \times \mathcal{B}(\mathrm{H} \to \mathrm{SS})$ with masses $m_{\mathrm{H}} = 125$ GeV and $m_{\mathrm{S}} = 50$ GeV. Long-lived S decays to muons with a branching fraction of $\mathcal{B}(\mathrm{S} \to \mu\mu) = 100\%$ are assumed. Only muon channel events were used.

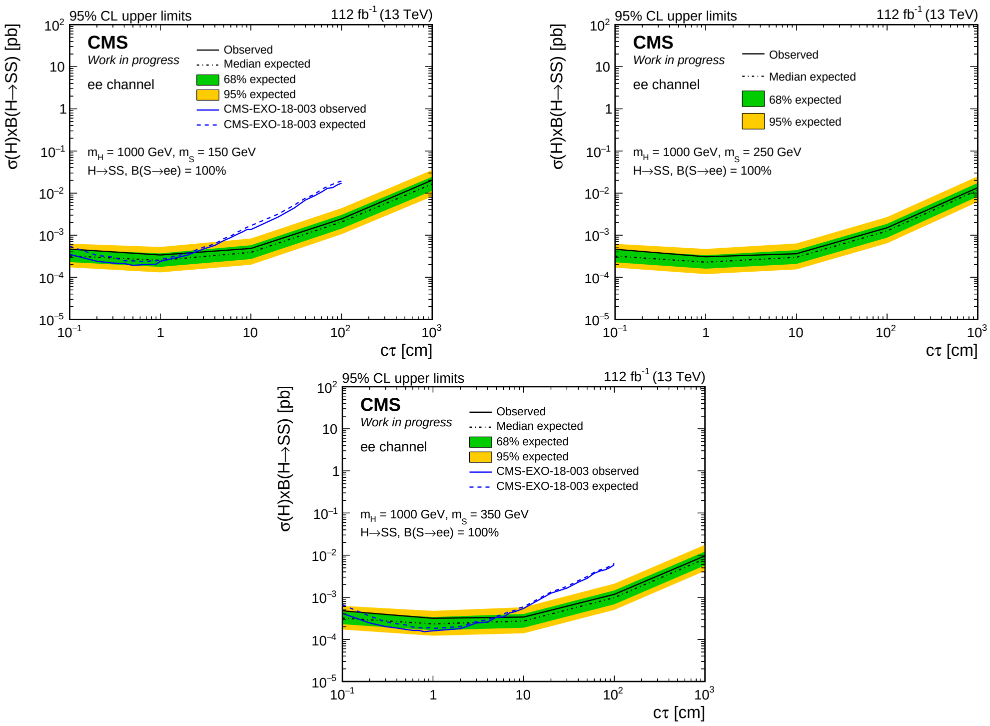


Figure 10.9: 95% CL upper limits set on $\sigma(\mathrm{H}) \times \mathcal{B}(\mathrm{H} \to \mathrm{SS})$ with masses $m_{\mathrm{H}} = 1000$ GeV and $m_{\mathrm{S}} = 150, 250, 350$ GeV (left, right, bottom). Long-lived S decays to electrons with a branching fraction of $\mathcal{B}(\mathrm{S} \to ee) = 100\%$ are assumed. Only electron channel events were used.

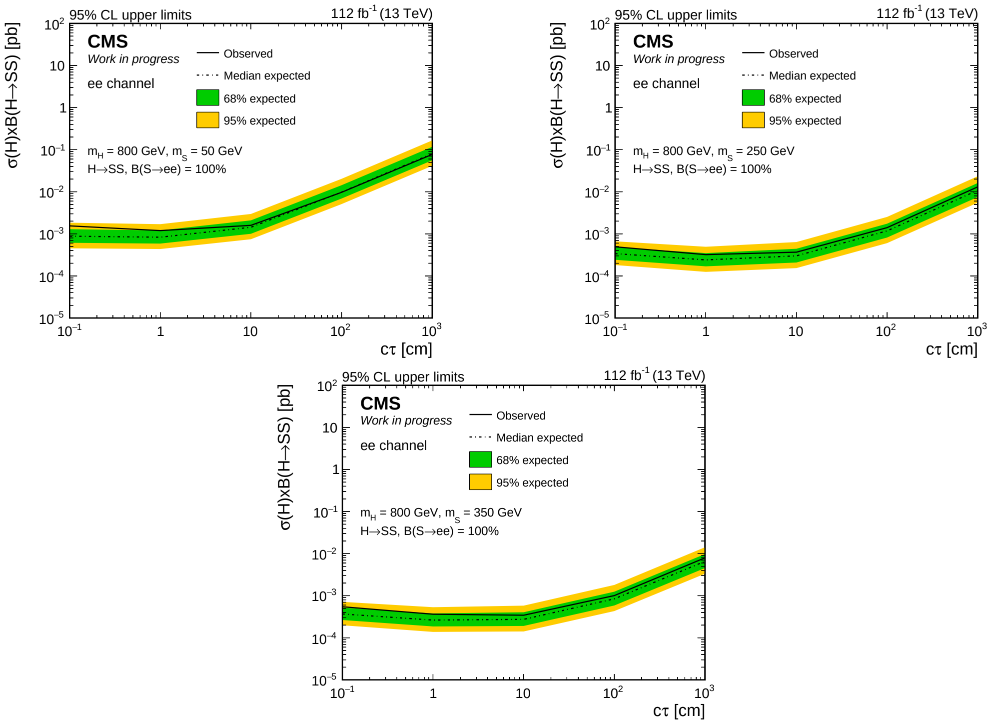


Figure 10.10: 95% CL upper limits set on $\sigma(\mathrm{H}) \times \mathcal{B}(\mathrm{H} \to \mathrm{SS})$ with masses $m_{\mathrm{H}} = 800$ GeV and $m_{\mathrm{S}} = 50, 250, 350$ GeV (left, right, bottom). Long-lived S decays to electrons with a branching fraction of $\mathcal{B}(\mathrm{S} \to ee) = 100\%$ are assumed. Only electron channel events were used.

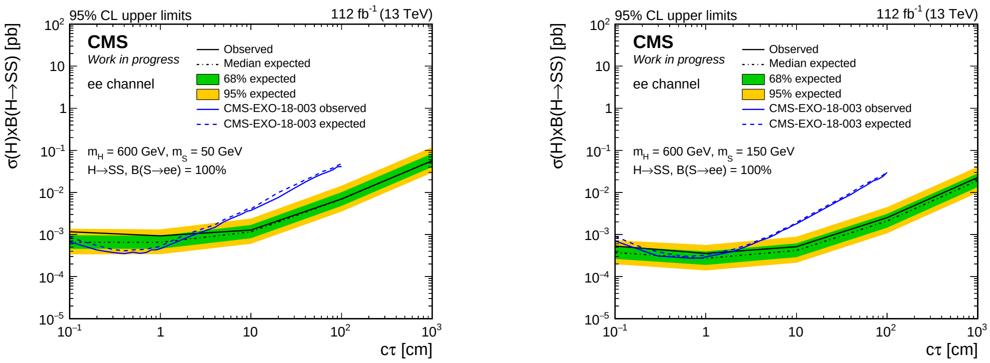


Figure 10.11: 95% CL upper limits set on $\sigma(\mathrm{H}) \times \mathcal{B}(\mathrm{H} \to \mathrm{SS})$ with masses $m_{\mathrm{H}} = 600$ GeV and $m_{\mathrm{S}} = 50, 150$ GeV (left, right). Long-lived S decays to electrons with a branching fraction of $\mathcal{B}(\mathrm{S} \to ee) = 100\%$ are assumed. Only electron channel events were used.

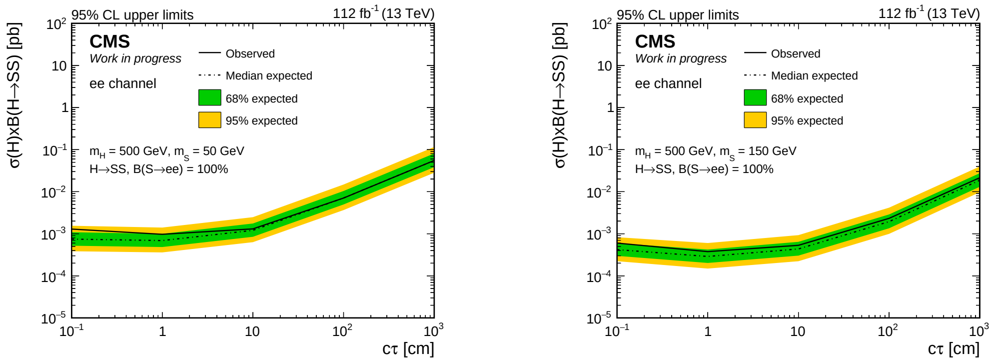


Figure 10.12: 95% CL upper limits set on $\sigma(\mathrm{H}) \times \mathcal{B}(\mathrm{H} \to \mathrm{SS})$ with masses $m_{\mathrm{H}} = 500$ GeV and $m_{\mathrm{S}} = 50, 150$ GeV (left, right). Long-lived S decays to electrons with a branching fraction of $\mathcal{B}(\mathrm{S} \to ee) = 100\%$ are assumed. Only electron channel events were used.

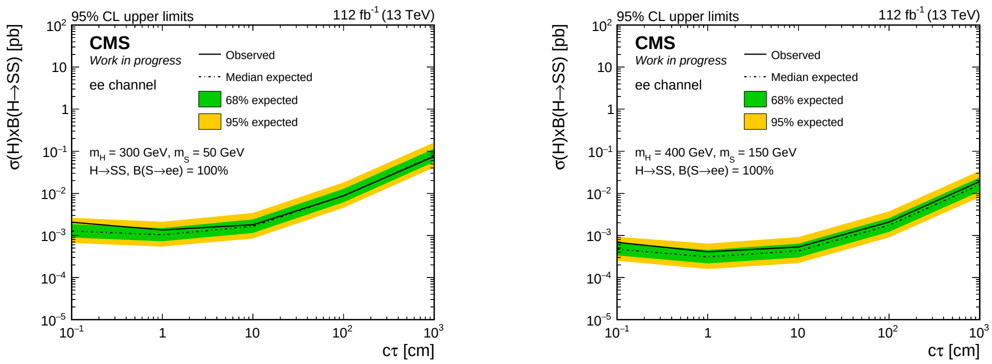


Figure 10.13: 95% CL upper limits set on $\sigma(\mathrm{H}) \times \mathcal{B}(\mathrm{H} \to \mathrm{SS})$ with masses $m_{\mathrm{H}} = 300$ GeV and $m_{\mathrm{S}} = 50$ GeV (left) and $m_{\mathrm{H}} = 400$ GeV and $m_{\mathrm{S}} = 150$ GeV (right). Long-lived S decays to electrons with a branching fraction of $\mathcal{B}(\mathrm{S} \to ee) = 100\%$ are assumed. Only electron channel events were used.

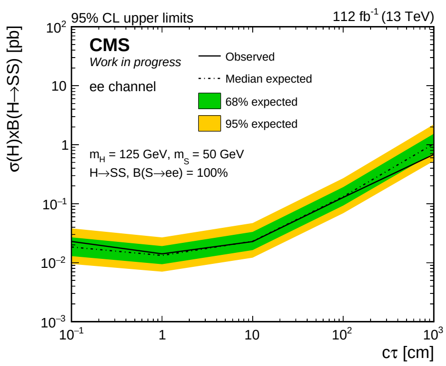


Figure 10.14: 95% CL upper limits set on $\sigma(\mathrm{H}) \times \mathcal{B}(\mathrm{H} \to \mathrm{SS})$ with masses $m_{\mathrm{H}} = 125$ GeV and $m_{\mathrm{S}} = 50$ GeV. Long-lived S decays to electrons with a branching fraction of $\mathcal{B}(\mathrm{S} \to ee) = 100\%$ are assumed. Only electron channel events were used.

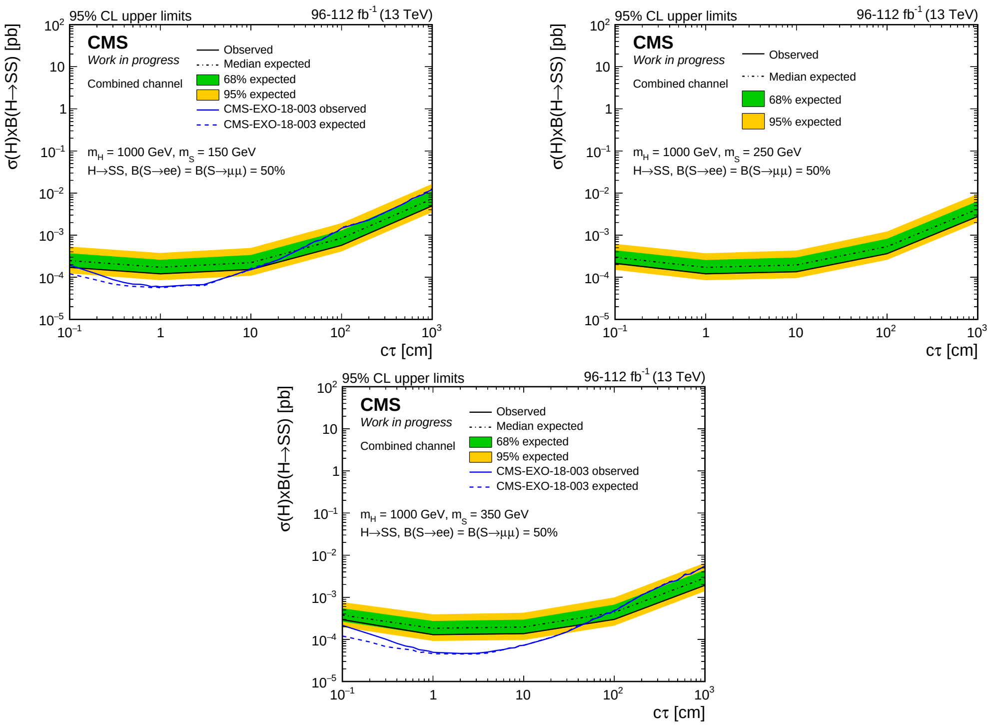


Figure 10.15: 95% CL upper limits set on $\sigma(\mathrm{H}) \times \mathcal{B}(\mathrm{H} \to \mathrm{SS})$ with masses $m_{\mathrm{H}} = 1000$ GeV and $m_{\mathrm{S}} = 150, 250, 350$ GeV (left, right, bottom). Long-lived S decays with a branching fraction of $\mathcal{B}(\mathrm{S} \to \mu\mu) = 50\%$ and $\mathcal{B}(\mathrm{S} \to ee) = 50\%$ are assumed. Only combined channel events were used.

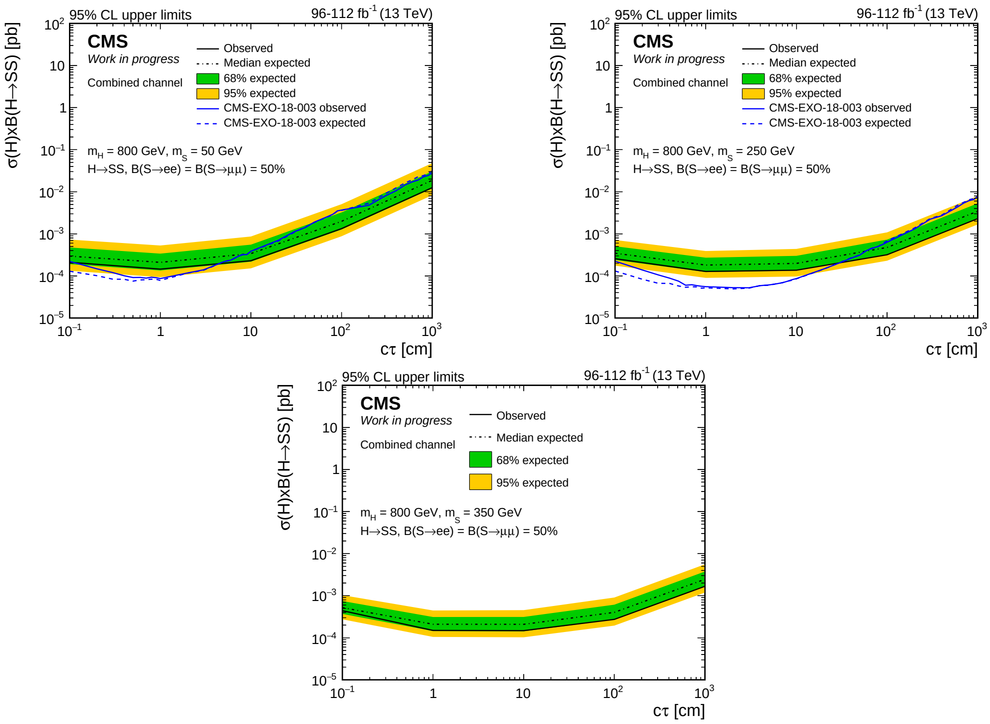


Figure 10.16: 95% CL upper limits set on $\sigma(\mathrm{H}) \times \mathcal{B}(\mathrm{H} \to \mathrm{SS})$ with masses $m_{\mathrm{H}} =$ 800 GeV and $m_{\mathrm{S}} = 50, 250, 350$ GeV (left, right, bottom). Long-lived S decays with a branching fraction of $\mathcal{B}(\mathrm{S} \to \mu\mu) = 50\%$ and $\mathcal{B}(\mathrm{S} \to ee) = 50\%$ are assumed. Only combined channel events were used.

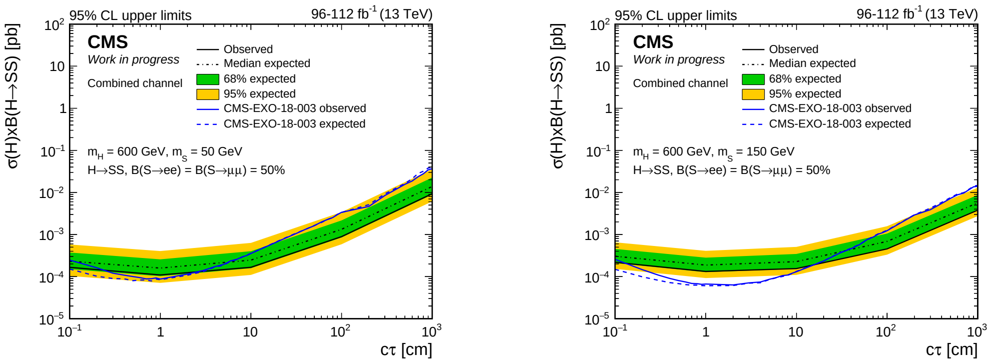


Figure 10.17: 95% CL upper limits set on $\sigma(\mathrm{H}) \times \mathcal{B}(\mathrm{H} \to \mathrm{SS})$ with masses $m_{\mathrm{H}} = 600$ GeV and $m_{\mathrm{S}} = 50, 150$ GeV (left, right). Long-lived S decays with a branching fraction of $\mathcal{B}(\mathrm{S} \to \mu\mu) = 50\%$ and $\mathcal{B}(\mathrm{S} \to ee) = 50\%$ are assumed. Only combined channel events were used.

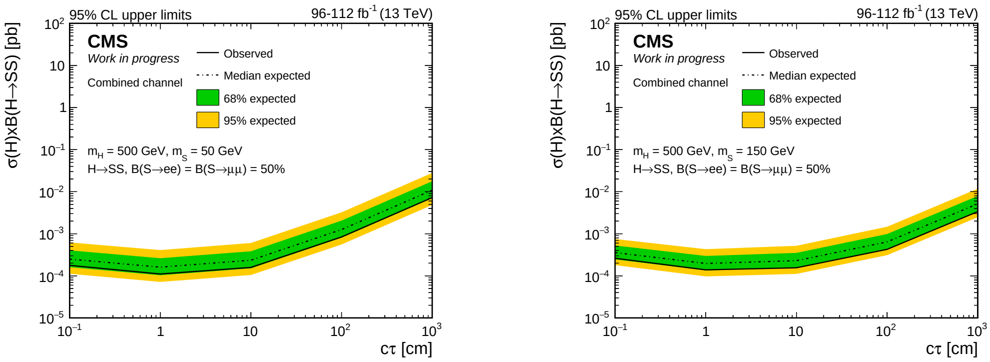


Figure 10.18: 95% CL upper limits set on $\sigma(\mathrm{H}) \times \mathcal{B}(\mathrm{H} \to \mathrm{SS})$ with masses $m_{\mathrm{H}} = 500$ GeV and $m_{\mathrm{S}} = 50, 150$ GeV (left, right). Long-lived S decays with a branching fraction of $\mathcal{B}(\mathrm{S} \to \mu\mu) = 50\%$ and $\mathcal{B}(\mathrm{S} \to ee) = 50\%$ are assumed. Only combined channel events were used.

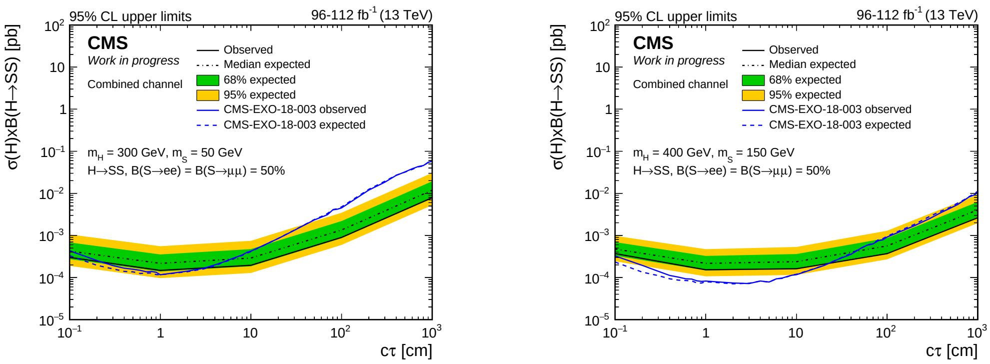


Figure 10.19: 95% CL upper limits set on $\sigma(\mathrm{H}) \times \mathcal{B}(\mathrm{H} \to \mathrm{SS})$ with masses $m_{\mathrm{H}} = 300$ GeV and $m_{\mathrm{S}} = 50$ GeV (left) and $m_{\mathrm{H}} = 400$ GeV and $m_{\mathrm{S}} = 150$ GeV (right). Long-lived S decays with a branching fraction of $\mathcal{B}(\mathrm{S} \to \mu\mu) = 50\%$ and $\mathcal{B}(\mathrm{S} \to ee) = 50\%$ are assumed. Only combined channel events were used.

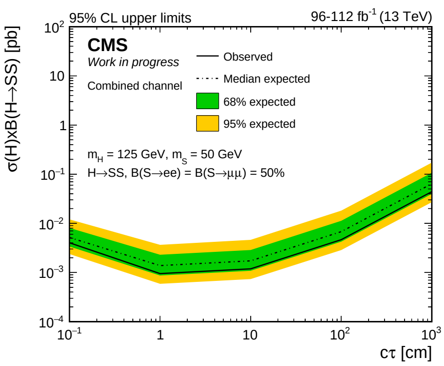


Figure 10.20: 95% CL upper limits set on $\sigma(\mathrm{H}) \times \mathcal{B}(\mathrm{H} \to \mathrm{SS})$ with masses $m_{\mathrm{H}} = 125$ GeV and $m_{\mathrm{S}} = 50$ GeV. Long-lived S decays with a branching fraction of $\mathcal{B}(\mathrm{S} \to \mu\mu) = 50\%$ and $\mathcal{B}(\mathrm{S} \to ee) = 50\%$ are assumed. Only combined channel events were used.

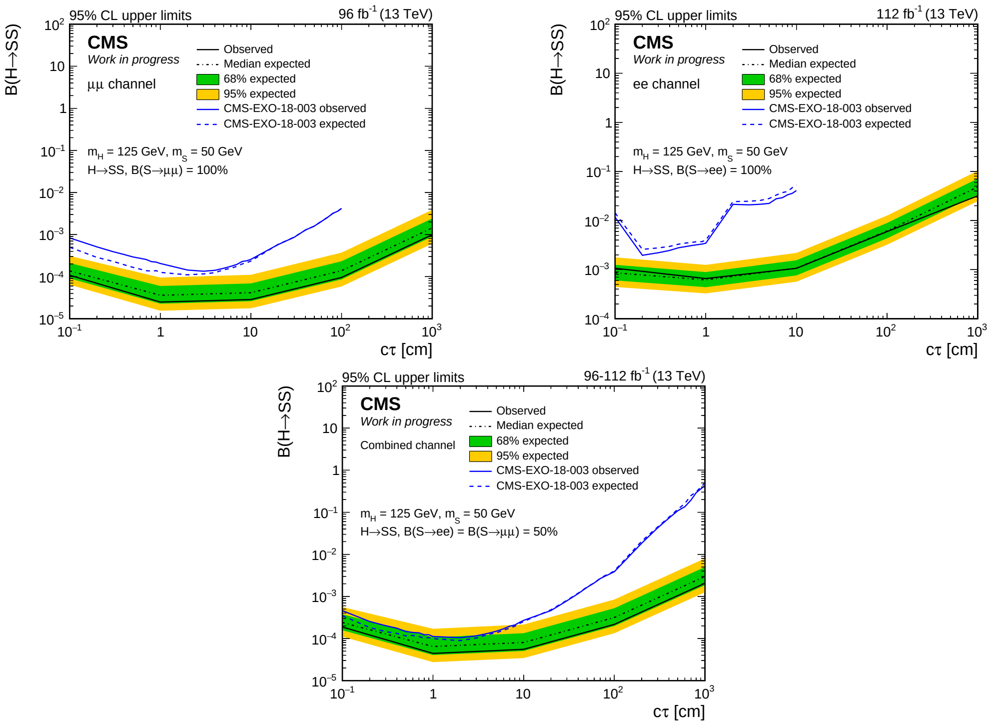


Figure 10.21: 95% CL upper limits set on $\mathcal{B}(\mathrm{H} \to \mathrm{SS})$ with masses $m_{\mathrm{H}} = 125$ GeV and $m_{\mathrm{S}} = 50$ GeV. Long-lived S decays with branching fractions of $\mathcal{B}(\mathrm{S} \to ee) = 100\%$ (left), $\mathcal{B}(\mathrm{S} \to ee) = 100\%$ (right), and $\mathcal{B}(\mathrm{S} \to ee) = \mathcal{B}(\mathrm{S} \to \mu\mu) = 50\%$ (bottom), are assumed. The assumed higgs production cross section is $\sigma(H)_{125} = 21.460 \pm 0.016$ pb.

Chapter 11

Conclusions

This thesis presents a search for long-lived particles decaying into displaced vertices with displaced leptons in pp collisions at $\sqrt{s} = 13$ TeV. The search was performed with the CMS detector analyzing LLP decays into electrons and muons with 112 fb^{-1} and 96 fb^{-1} of data, respectively, collected during the Run 2 physics program of the LHC.

Two separate channels are defined to study each lepton type contribution separately and a specific reconstruction algorithm, calibration and selection criteria have been optimized for each.

Alternative reconstruction algorithms to those used in most CMS analyses were employed, which substantially increased the signal efficiency for highly displaced LLP decays. Specific identification criteria were also optimized to select good quality leptons that could be used to reconstruct the displaced vertices used in the search. A dedicated vertex selection was designed to provide a good discrimination of the signal against the existing background.

The various background sources that could potentially manifest as displaced vertices were carefully studied. Their properties were thoroughly investigated with the objective of quantifying their impact on the analysis. The residual background contribution, due exclusively to instrumental and resolution events, was estimated using a background prediction method. This method was optimized and validated by considering different scenarios and potential background types.

The aspects that could potentially be mismodelled by the Monte Carlo risking the introduction of differences between the simulation and data were evaluated. A complete set of corrections was derived and their use was validated in the main control region of the analysis.

The search was designed to be agnostic and could be interpreted in any model with long-lived particles decaying to leptons in the same vertex within the tracker volume. The results presented are interpreted in the context of BSM

Higgs scenarios decaying to long-lived scalar bosons which can be framed in hidden sector theories.

The observed data were consistent with the background prediction and no evidence of signal was found. Upper limits on the production cross section of the process were set at the 95% confidence level as a function of the lifetime of the long-lived scalar bosons. The limits presented constrain the cross section for lifetimes from 0.1 to 1000 cm. Three different interpretations were taken for the results. The first one assumes a pure decay rate to muons while the second interprets that the LLPs decay exclusively to electrons. The third one considers democratic LLP decays into the two considered lepton types and both channels are combined to extract the results.

This thesis presents the most stringent limits on LLP decays into electrons up to date in the context of BSM Higgs theories. For the channel combination, the upper limits improve the existing results in Run 2 at the most displaced regimes over lifetimes of 100 cm.

Capítulo 11

Conclusiones

Esta tesis presenta una búsqueda de partículas de larga vida que decaen en vértices desplazados con leptones desplazados en colisiones pp a $\sqrt{s} = 13$ TeV. La búsqueda se ha realizado con el detector CMS analizando desintegraciones de particulas de alta vida media en electrones y muones con 112 fb^{-1} y 96 fb^{-1} de datos, respectivamente, recogidos durante el programa de física Run 2 del LHC.

Se han definido dos canales independientes para estudiar la contribución de cada lepton por separado y se ha optimizado un algoritmo de reconstrucción, calibración y criterios de selección específicos en cada caso.

Para la reconstrucccion de los leptones, se han empleado algoritmos alternativos a los utilizados en los análisis CMS estándar, lo que ha incrementado sustancialmente la eficiencia de la señal para las desintegraciones de partículas altamente desplazadas. También se han optimizado criterios específicos de identificación para seleccionar leptones de buena calidad que pudieran utilizarse para reconstruir los vértices desplazados utilizados en la búsqueda. Además, se ha diseñado una selección de vértices específica para proporcionar una buena discriminación de la señal frente al fondo existente.

Para el fondo, se ha llevado a cabo un estudio de procesos físicos que podrían manifestarse como vértices desplazados. Se han estudiado sus propiedades con el objetivo de evaluar su impacto en el análisis. La contribución de fondo en la región de señal, cuyo desplazamiento es introducido exclusivamente por efectos instrumentales y de resolución, se ha estimado de lo datos utilizando un método de predicción de fondo. Este método se ha optimizado y validado considerando la contribución de varios procesos al fondo a diferentes niveles de desplazamiento.

Se han estudiado los aspectos que podrían ser no reproducidos correctamente por el Monte Carlo introduciendo diferencias entre la simulación y los datos. Se han derivado correcciones para cada uno de ellos y se ha validado su aplicación en la región de control principal del análisis.

La búsqueda se ha diseñado para que sea agnóstica y se pueda interpretarse en cualquier modelo con partículas de larga vida que decaen a leptones en el mismo vértice dentro del volumen del tracker. Los resultados presentados se interpretan en el contexto de escenarios con Higgs exóticos que decaen a bosones escalares de larga vida que pueden enmarcarse en teorías de sectores ocultos.

Los datos observados son consistentes con la predicción de fondo del análisis y no se ha encontrado evidencia de nueva física. Los límites superiores de la sección transversal de producción del proceso se has establecido con un nivel de confianza del 95 % en función del tiempo de vida de los bosones escalares de larga vida media. Los límites presentados restringen la sección eficaz para tiempos de vida de 0.1 a 1000 cm. Para la interpretación de los resultados se han adoptado tres interpretaciones diferentes. La primera asume una tasa de desintegración pura a muones, mientras que la segunda interpreta que los LLP se desintegran exclusivamente a electrones. La tercera considera desintegraciones equivalentes de LLP en los dos tipos de leptones considerados, para la cual ambos canales se han combinado para extraer los resultados finales.

Esta tesis presenta los límites más estrictos sobre las desintegraciones de partículas de alta vida media en electrones hasta la fecha en el contexto de teorías con bosones de Higgs exóticos. Para la combinación de canales, los límites superiores mejoran los resultados existentes en el Run 2 en los regímenes más desplazados en vida media superiores a 100 cm.

Bibliography

1. Workman, R. L. *et al.* Review of Particle Physics. *PTEP* **2022,** 083C01 (2022).
2. Peskin, M. E. & Schroeder, D. V. *An introduction to quantum field theory* ISBN: 978-0-201-50397-5 (Addison-Wesley Pub. Co, Reading, Mass, 1995).
3. Higgs, P. W. Broken Symmetries and the Masses of Gauge Bosons. *Phys. Rev. Lett.* **13,** 508–509. link (1964).
4. Chatrchyan, S., Khachatryan, V., *et al.* Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Physics Letters B* **716,** 30–61. (2012).
5. Aad, G., Abajyan, T., *et al.* Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *Physics Letters B* **716,** 1–29. (2012).
6. Arbuzov, A. B. *Quantum Field Theory and the Electroweak Standard Model* 2018. link.
7. Salam, A. & Ward, J. C. Weak and electromagnetic interactions. *Nuovo Cim* **11,** 568–577. link (1959).
8. Weinberg, S. A Model of Leptons. *Physical Review Letters* **19,** 1264–1266. link (1967).
9. Glashow, S. L. Partial-symmetries of weak interactions. *Nuclear Physics* **22,** 579–588. link (1961).
10. Fabio Maltoni. *Basics of QCD for the LHC* Lecture. Romania, 2011.
11. Langacker, P. Grand unification. en. *Scholarpedia* **7,** 11419. link (2012).
12. Georgi, H. & Glashow, S. L. Unity of All Elementary-Particle Forces. *Phys. Rev. Lett.* **32,** 438–441. link (1974).
13. Rovelli, C. *Quantum Gravity* 1st ed. link (Cambridge University Press, 2004).
14. Thomson, M. *Modern particle physics* ISBN: 978-1-107-03426-6 (Cambridge University Press, Cambridge, United Kingdom ; New York, 2013).
15. Abi, B. *et al.* Measurement of the Positive Muon Anomalous Magnetic Moment to 0.46 ppm. *Phys. Rev. Lett.* **126,** 141801. link (2021).

16. Bennett, G. W. *et al.* Final report of the E821 muon anomalous magnetic moment measurement at BNL. *Phys. Rev. D* **73,** 072003. link (2006).

17. CDF Collaboration†‡ *et al.* High-precision measurement of the *W* boson mass with the CDF II detector. *Science* **376,** 170–176. link (2022).

18. Lykken, J. D. *Beyond the Standard Model* arXiv:1005.1676 [hep-ph]. 2011. link.

19. Dobrescu, B. A. *Beyond the Standard Model* in *2011 European School of High-Energy Physics* (CERN, Geneva, 2014), 119–149.

20. Morrissey, D. E., Plehn, T. & Tait, T. M. P. Physics searches at the LHC. *Physics Reports* **515,** 1–113. link (2012).

21. Martin, S. P. in *Advanced Series on Directions in High Energy Physics* 1–98 (WORLD SCIENTIFIC, 1998). link.

22. Kaluza, T. Zum Unitätsproblem der Physik. *Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften (Berlin,* 966–972. link (2023) (1921).

23. Klein, O. Quantentheorie und fünfdimensionale Relativitätstheorie. *Zeitschrift fur Physik* **37,** 895–906. link (1926).

24. Klein, O. The Atomicity of Electricity as a Quantum Theory Law. *Nature* **118,** 516. link (1926).

25. Arkani–Hamed, N., Dimopoulos, S. & Dvali, G. The hierarchy problem and new dimensions at a millimeter. *Physics Letters B* **429,** 263–272 (1998).

26. Antoniadis, I. A possible new dimension at a few TeV. *Physics Letters B* **246,** 377–384. link (1990).

27. Appelquist, T., Cheng, H.-C. & Dobrescu, B. A. Bounds on universal extra dimensions. *Phys. Rev. D* **64,** 035002. link (2001).

28. Randall, L. & Sundrum, R. Large Mass Hierarchy from a Small Extra Dimension. *Phys. Rev. Lett.* **83,** 3370–3373. link (1999).

29. Strassler, M. J. & Zurek, K. M. Echoes of a hidden valley at hadron colliders. *Physics Letters B* **651,** 374–379. link (2007).

30. Hill, C. T. & Simmons, E. H. Strong Dynamics and Electroweak Symmetry Breaking. link (2002).

31. Perelstein, M. Little Higgs models and their phenomenology. *Progress in Particle and Nuclear Physics* **58,** 247–291. link (Jan. 2007).

32. Evans, L. & Bryant, P. LHC Machine. *Journal of Instrumentation* **3,** S08001. link (2008).

33. The ATLAS Collaboration *et al.* The ATLAS Experiment at the CERN Large Hadron Collider. *Journal of Instrumentation* **3,** S08003. link (2008).

34. The LHCb Collaboration *et al.* The LHCb Detector at the LHC. *Journal of Instrumentation* **3,** S08005. link (2008).

35. The ALICE Collaboration *et al.* The ALICE experiment at the CERN LHC. *Journal of Instrumentation* **3,** S08002. link (2008).
36. Bulletin, C. The particle suppliers. Les fournisseurs de particules, 03. link (2010).
37. CMS Collaboration. *CMS Luminosity - Public results* 2023. link.
38. Chatrchyan, S. *et al.* The CMS Experiment at the CERN LHC. *JINST* **3,** S08004 (2008).
39. CMS Collaboration & CERN. Geneva. LHC Experiments Committee. The CMS magnet project: Technical Design Report. en. Publisher: CERN. link (1997).
40. Karimäki, V. *et al. The CMS tracker system project: Technical Design Report* link (CERN, Geneva, 1997).
41. Adam, W. *et al.* The CMS Phase-1 Pixel Detector Upgrade. *JINST* **16.** _eprint: 2012.14304, P02027 (2021).
42. Viliani, L. CMS tracker performance and readiness for LHC Run II. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **824,** 67–69. link (2016).
43. CMS Collaboration. *The CMS electromagnetic calorimeter project: Technical Design Report* link (CERN, Geneva, 1997).
44. CMS Collaboration. *The CMS hadron calorimeter project: Technical Design Report* link (CERN, Geneva, 1997).
45. CMS Collaboration. Identification and filtering of uncharacteristic noise in the CMS hadron calorimeter. *J. Inst.* **5,** T03014–T03014. link (2010).
46. CMS Collaboration. *The CMS muon project: Technical Design Report* link (CERN, Geneva, 1997).
47. CMS Collaboration. *CMS Muon DPG - Public results* 2023. link.
48. CMS Collaboration. Level-1 muon trigger performance with the full 2017 dataset. link (2018).
49. Bayatyan, G L and Della Negra, Michel and Foà, and Hervé, A and Petrilli, Achille. *CMS computing: Technical Design Report* link (CERN, Geneva, 2005).
50. Petrucciani, G., Rizzi, A. & Vuosalo, C. Mini-AOD: A New Analysis Data Format for CMS. *J. Phys.: Conf. Ser.* **664,** 072052. link (2015).
51. Sirunyan, A. M. *et al.* Particle-flow reconstruction and global event description with the CMS detector. *Journal of Instrumentation* **12.** Publisher: IOP Publishing, P10003–P10003. link (2017).
52. CMS Collaboration. Description and performance of track and primary-vertex reconstruction with the CMS tracker. *Journal of Instrumentation* **9,** P10009. link (2014).
53. Adam, W., Mangano, B., Speer, T. & Todorov, T. *Track Reconstruction in the CMS tracker* tech. rep. (CERN, Geneva, 2006). link.

54. Frühwirth, R., Waltenberger, W. & Vanlaer, P. *Adaptive Vertex Fitting* tech. rep. (CERN, Geneva, 2007). link.

55. CMS Collaboration. Performance of CMS muon reconstruction in pp collision events at $\sqrt{s} = 7$ TeV. *Journal of Instrumentation* **7,** P10002. link (2012).

56. CMS Collaboration. Electron and photon reconstruction and identification with the CMS experiment at the CERN LHC. *Journal of Instrumentation* **16.** Publisher: IOP Publishing, P05014. link (2021).

57. CMS Collaboration. Performance of electron reconstruction and selection with the CMS detector in proton-proton collisions at $\sqrt{s}$= 8 TeV. *J. Inst.* **10,** P06005–P06005. link (2015).

58. Adam, W., Frühwirth, R., Strandlie, A. & Todorov, T. Reconstruction of electrons with the Gaussian-sum filter in the CMS tracker at the LHC. *J. Phys. G: Nucl. Part. Phys.* **31,** N9–N20. link (2005).

59. Cacciari, M., Salam, G. P. & Soyez, G. The anti- k_t jet clustering algorithm. *J. High Energy Phys.* **2008,** 063–063. link (2008).

60. Sirunyan, A. M. *et al.* Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV. *Journal of Instrumentation* **13,** P05011. link (2018).

61. CMS Collaboration. Identification of b-quark jets with the CMS experiment. *J. Inst.* **8,** P04013–P04013. link (2013).

62. Lee, L., Ohm, C., Soffer, A. & Yu, T.-T. Collider searches for long-lived particles beyond the Standard Model. *Progress in Particle and Nuclear Physics* **106,** 210–255. link (2019).

63. Curtin, D. *et al.* Long-lived particles at the energy frontier: the MATHUSLA physics case. *Rep. Prog. Phys.* **82,** 116201. link (2019).

64. Alimena, J. *et al.* Searching for long-lived particles beyond the Standard Model at the Large Hadron Collider. *J. Phys. G: Nucl. Part. Phys.* **47,** 090501. link (2020).

65. Strassler, M. J. & Zurek, K. M. Discovering the Higgs through highly-displaced vertices. *Physics Letters B* **661,** 263–267. link (2008).

66. Langacker, P. The physics of heavy Z' gauge bosons. *Rev. Mod. Phys.* **81,** 1199–1228. link (2009).

67. Holdom, B. Two $U(1)$'s and ε charge shifts. *Physics Letters B* **166,** 196–198. link (1986).

68. Abdullahi, A. M. *et al.* The present and future status of heavy neutral leptons. *J. Phys. G: Nucl. Part. Phys.* **50,** 020501. link (2023).

69. Cai, Y., Han, T., Li, T. & Ruiz, R. Lepton Number Violation: Seesaw Models and Their Collider Tests. *Frontiers in Physics* **6.** link (2018).

70. Asaka, T., Blanchet, S. & Shaposhnikov, M. The νMSM, dark matter and neutrino masses. *Physics Letters B* **631,** 151–156. link (2005).

71. Thomas, S. & Wells, J. D. Phenomenology of Massive Vectorlike Doublet Leptons. *Phys. Rev. Lett.* **81,** 34–37. link (1998).

72. Bai, Y. & Tait, T. M. P. Inelastic dark matter at the LHC. *Physics Letters B* **710,** 335–338. link (2012).

73. Berlin, A. & Kling, F. Inelastic dark matter at the LHC lifetime frontier: ATLAS, CMS, LHCb, CODEX-b, FASER, and MATHUSLA. *Phys. Rev. D* **99,** 015021. link (2019).

74. Khoze, V. V., Plascencia, A. D. & Sakurai, K. Simplified models of dark matter with a long-lived co-annihilation partner. *J. High Energ. Phys.* **2017,** 41. link (2017).

75. Jamie Antonelli. *Searches for long-lived particles at CMS* Parallel talk. Chicago, 2016.

76. Sirunyan, A. M. *et al.* Search for disappearing tracks in proton-proton collisions at $\sqrt{s}$=13TeV. *Physics Letters B* **806,** 135502. link (2020).

77. CMS Collaboration. *Search for heavy stable charged particles with 12.9 fb^{-1} of 2016 data* tech. rep. CMS-PAS-EXO-16-036 (CERN, Geneva, 2016). link.

78. CMS Collaboration. *Search for fractionally charged particles in pp collisions at $\sqrt{s}$=13 TeV* tech. rep. (CERN, Geneva, 2022). link.

79. CMS Collaboration. Search for long-lived particles decaying to leptons with large impact parameter in proton-proton collisions at $\sqrt{s}$ = 13 TeV. *The European Physical Journal C* **82.** link (2022).

80. CMS Collaboration. Search for long-lived particles decaying to a pair of muons in proton-proton collisions at $\sqrt{s}$ = 13 TeV. link (2022).

81. Lowette, S. *Long-lived particle results in ATLAS and CMS* La Thuille, Mar. 2023. link.

82. CMS Collaboration. Search for long-lived particles using out-of-time trackless jets in proton-proton collisions at $\sqrt{s}$ = 13 TeV. arXiv:2212.06695 (accepted by JHEP). link (2022).

83. Sirunyan, A. M. *et al.* Search for long-lived particles decaying to jets with displaced vertices in proton-proton collisions at $\sqrt{s}$ = 13 TeV. *Phys. Rev. D* **104,** 052011. link (2021).

84. ATLAS Collaboration. Search for long-lived, massive particles in events with displaced vertices and multiple jets in *pp* collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector. arXiv:2301.13866. link (2023).

85. ATLAS Collaboration. *Search for pairs of muons with small displacements in pp collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector* arXiv:2305.02005 [hep-ex]. 2023. link.

86. Aad, G. *et al.* Search for Displaced Leptons in $\sqrt{s}$=13 TeV *pp* Collisions with the ATLAS Detector. *Phys. Rev. Lett.* **127,** 051802. link (2021).

87. CMS Collaboration. *Search for inelastic dark matter in events with two displaced muons and missing transverse momentum in proton-proton collisions at $\sqrt{s}$ = 13 TeV* tech. rep. CMS-PAS-EXO-20-010. (submitted to Phys. Rev. Lett.) (CERN, Geneva, 2023). link.
88. CMS Collaboration. *Search for long-lived heavy neutral leptons with lepton flavour conserving or violating decays to a jet and an electron, muon, or tau lepton* tech. rep. CMS-PAS-EXO-21-013 (CERN, Geneva, 2023). link.
89. ATLAS Collaboration. Search in diphoton and dielectron final states for displaced production of Higgs or Z bosons with the ATLAS detector in $\sqrt{s} = 13$ TeV pp collisions. arXiv (preprint). link (2023).
90. CMS Collaboration. *CMS Exotica Public Physics Results* tech. rep. (May 2023). link.
91. ATLAS Collaboration. *Summary Plots for Heavy Particle Searches and Long-lived Particle Searches - March 2023* tech. rep. (CERN, Geneva, 2023). link.
92. CMS Collaboration. Performance of CMS muon reconstruction in cosmic-ray events. *Journal of Instrumentation* **5,** T03022. link (Mar. 2010).
93. CMS Collaboration. Muon Reconstruction and Identification Improvements for Run-2 and First Results with 2015 Run Data. link (2015).
94. CMS Collaboration. Electron and photon reconstruction and identification with the CMS experiment at the CERN LHC. Publisher: arXiv Version Number: 2. link (2023) (2020).
95. Speer, T., Prokofiev, K., Frühwirth, R., Waltenberger, W. & Vanlaer, P. *Vertex Fitting in the CMS Tracker* tech. rep. (CERN, Geneva, 2006). link.
96. Alwall, J. *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,** 079 (2014).
97. Sjöstrand, T. *et al.* An Introduction to PYTHIA 8.2. *Computer Physics Communications* **191,** 159–177. link (June 2015).
98. Frixione, S., Nason, P. & Oleari, C. Matching NLO QCD computations with Parton Shower simulations: the POWHEG method. *JHEP* **11,** 070 (2007).
99. Khachatryan, V. *et al.* Search for long-lived particles that decay into final states containing two electrons or two muons in proton-proton collisions at \sqrts=8\text \text \mathrmTeV. *Phys. Rev. D* **91.** Publisher: American Physical Society, 052012. link (Mar. 2015).
100. Oreglia, M. J. *A study of the reactions* $\psi' \rightarrow \gamma\gamma\psi$ PhD thesis (Stanford University, Stanford, California, 1980).
101. Junk, T. Confidence level computation for combining searches with small statistics. *Nucl. Instrum. Meth. A* **434.** _eprint: hep-ex/9902006, 435 (1999).
102. Read, A. L. Presentation of search results: the *CL_s* technique. *J. Phys. G* **28,** 2693 (2002).

103. Cowan, G., Cranmer, K., Gross, E. & Vitells, O. Asymptotic formulae for likelihood-based tests of new physics. *Eur. Phys. J. C* **71.** _eprint: 1007.1727, 1554 (2011).

104. ATLAS and CMS Collaborations. *Procedure for the LHC Higgs boson search combination in summer 2011* ATLAS/CMS joint note ATL-PHYS-PUB-2011-011, CMS-NOTE-2011-005 (2011). link.