SAPIENZA
UNIVERSITÀ DI ROMA

Searches for Heavy Neutral Leptons in B Meson Decays with the CMS Experiment

Scuola di Dottorato in Scienze Astronomiche, Chimiche, Fisiche e Matematiche 'Vito Volterra'
Physics (XXXV cycle)

Raffaella Tramontano
ID number 1652729

Advisors
Prof. Shahram Rahatlou
Dr. Francesco Pandolfi

Academic Year February 2023

Searches for Heavy Neutral Leptons in B Meson Decays with the CMS Experiment

Sapienza University of Rome

This thesis has been typeset by LaTeX and the Sapthesis class.

Version: February 2023

Author's email: raffaella.tramontano@roma1.infn.it

Contents

Introduction

The Standard Model of elementary particles is one of the most successful scientific theories ever conceived. It describes the elementary building blocks of matter and their interactions within the self-consistent mathematical environment of quantum field theory. Countless experiments have confirmed the Standard Model predictions with astounding precision over the last century, crowning the theory as an unprecedented tool for the description of the strong and electro-weak interactions.

The Standard Model is, however, far from being an exhaustive portrayal of the Universe dynamics. The model does not include a description of gravitational interactions, nor explains why gravity is so much weaker than the electro-weak and strong forces. It shows theoretical drawbacks, as its capability of encompassing phenomena at very distant energy scales derives mainly by fine-tuned relations among the free parameters of the model. Finally, it lacks a correct description of several experimental pieces of evidence. The Standard Model does not provide a suitable particle candidate for dark matter, nor justifies its abundance in the Universe; it does not explain the matter anti-matter asymmetry and does not account for neutrino flavor oscillations and masses.

Several extensions to the Standard Model have been developed to encompass solutions to these open questions. New physics candidates usually consist in particles or interactions lying on high energy scales with respect to the ones experimentally available to probe, or in objects and forces that are weakly coupled to the Standard Model: they yield small effects at the energy scales for which the theory already provides precise predictions. In the scenario of new physics weakly coupled to the SM, evidence is sought either as a significant discrepancy with respect to the predicted value over a high-precision observable, or as appearing in rare Standard Model processes: small cross sections have enhanced sensitivity to possible anomalies and new resonances.

Following the latter approach, the search presented in this thesis targets right-handed Heavy Majorana Neutral Leptons in B meson decays. Heavy Neutral Leptons with masses $m_N < m_B$ can be sought through this signature. The Heavy Neutral Lepton is expected to be weakly coupled to the SM, therefore its lifetime can be non-negligible: lifetimes in the $c\tau$ $[10^{-4};1]$ m range are investigated in this search. The search finds motivation in the framework of neutrino ν Minimal Standard Model extension (νMSM), where the introduction of right-handed heavy neutral leptons allows for the construction of neutrino mass terms, thus justifying the observed neutrino flavor oscillations.

The search is performed with the detector of the Compact Muon Solenoid (CMS) experiment at CERN, analyzing data from proton-proton collisions provided by the

Large Hadron Collider (LHC), a superconducting proton collider with a center of mass energy up to 13.6 TeV. This search is made possible by the B-Parking dataset, a novel sample of 10^{10} high-purity $B\bar{B}$ events, collected in 2018 with innovative trigger and data-taking strategies.

The analysis configures as a search for a peak over a smooth background in the invariant mass system of the heavy neutrino decay products (a charged lepton and a charged pion). The main challenge for this analysis is the correct reconstruction and identification of these decay products, as they are typically soft and displaced, while standard CMS particle reconstruction is optimized for hard particles produced close to the interaction region.

The first Chapter of this thesis offers a brief review of the theoretical foundation of the Standard Model, with a particular focus on the mixing phenomena arising through the Higgs Mechanism and on the neutrino mass puzzle. An introduction to νMSM is presented, together with a summary of the existing constraints on Heavy Neutral Leptons.

Chapter 2 details the experimental apparatus and the data collection strategy that make the search possible. The Large Hadron Collider is introduced, and a detailed description of the CMS experiment is presented. Finally, a review of CMS trigger and reconstruction strategies is reported, with a particular focus on the novel data-taking techniques designed for the 2018 B-Parking dataset.

The challenges of electron reconstruction and identification in CMS are presented in Chapter 3. After an overview of the general aspects of electron reconstruction in CMS, the energy measurement calibration procedure is detailed, together with the energy linearity measurement for electrons in the $Z \to e^+e^-$ and $J/\psi \to e^+e^-$ resonances. Finally, the dedicated reconstruction and identification algorithms developed for low-energy electrons in the framework of the B-Parking dataset are described.

The analysis event selection is reported in Chapter 4 together with a summary of the data samples used in the analysis and a detailed account of the generation procedure designed for the signal events simulation. The analysis is performed in six exclusive categories, designed to cover different possible neutrino lifetimes and decay signatures. A Paramatric Neural Network is employed for signal-over-background discrimination over the several neutrino masses hypotheses.

Signal and background modeling are presented in Chapter 5, while possible sources of systematic uncertainties are detailed in Chapter 6. In Chapter 7, the statistical tools used to derive the exclusion limits on the Heavy Neutral Leptons couplings for different mass hypotheses are described and the expected results are presented.

Chapter 1

The Standard Model, the Rare and the Unknown

In this section a review of the Standard Model is presented, together with the open questions left unsolved by the model. A particular focus will be set on the physics sectors related to the search presented in this work: neutrino physics and the neutrino masses puzzle. Standard Model extensions which account for neutrino masses will be presented, with particular focus on νMSM with the introduction of an arbitrary numbers of sterile neutrinos within the SM.

1.1 The Standard Model of Particle Physics

The Standard Model is the most complete theory to date describing interactions among fundamental constituents of ordinary matter. It encompasses most of the phenomena caused by three of the four fundamental interactions (the strong, electromagnetic and weak interactions) and accounts for the existence of massive objects. The first building block of the Standard Model dates back to 1961 and it is to be found in Sheldon Glashow's unification of the electromagnetic and weak interactions [32]. The electroweak interaction was then coupled to the Brout-Englert-Higgs [28] [38] Mechanism by Weinberg [62] and Salam [54] a few years later, shaping the model in its modern form. Throughout the following years, over countless experiments, the SM has established itself as a thoroughly tested and astoundingly precise theory [37].

The Standard Model describes sub-atomic interactions within the self-consistent framework of quantum field theory. It consists of a renormalizable gauge non-abelian field theory based on the symmetry group

$$SU(3) \times SU(2)_L \times U(1). \tag{1.1}$$

The symmetry group is non abelian, meaning that commutation rules among the group generators are non trivial, and *gauge invariant*: the interaction terms are asked to be locally invariant under the group transformations.

The SM describes the interaction of spin 1/2 particles, fermions. Interaction terms among fermion fields in the SM Lagrangian have to be invariant under the symmetry group reported in Equation 1.1. Coupling this requirement with the

Noether's theorem [47], a conserved quantity arises for each preserved symmetry in the system. The electroweak symmetry group $SU(2) \times U(1)$ is linked to the weak isospin (T_3) and the hypercharge Y, while $SU(3)$ yields the *color* quantum number.

The L subscript for SU(2) points to the chiral features of the electroweak transformations. Fermion quantum fields are a reducible representations of the Lorentz group. The chirality operator , γ^5, distinguishes two irreducible representations of the Lorentz group that are used to describe spin 1/2 states. Chirality coincides, for massless objects, with a particle helicity or headedness. It is defined by the projection of the spin on the direction of motion of a particle, If the particle spin projection is opposed to the particle direction, the particle is left-handed, otherwise it is right-handed. In the SM, only the left-handed fields are involved in weak interactions.

SM fermions are naturally split in two groups, based on whether they participate to the strong interaction. The quarks, are coloured particles, subjected to strong interaction. Leptons, on the other hand, do not carry any color quantum number and interact only electro-weakly.

A representation of Standard Model fermion fields is reported below:

$$l_{L,i} = \begin{pmatrix} \nu_i \\ \ell_i \end{pmatrix}_L , \quad \ell_{R,i}\; q_i^\alpha = \begin{pmatrix} u_{L,i}^\alpha \\ d_{L,i}^\alpha \end{pmatrix}_L \; u_{R,i}^\alpha \; d_{R,i}^\alpha \tag{1.2}$$

where i is the flavor index. Quark and lepton flavor is grouped in three families: $i \in$ e, μ, τ for the leptons and $i \in \binom{u}{d}, \binom{c}{s}, \binom{t}{b}$ for the quarks. Each flavor family carries a left-handed doublet and right-handed singlets, which yield identical transformation properties under Equation 1.1.

The L and R subscripts in Equation 1.2 explicit the chirality of the multiplets. Left ended states transform as doublets under $SU(2)_L$ with I = 1/2, while right-handed particles transform as singlets with I = 0 under the electroweak symmetry. Leptons, l in Equation 1.2, are distinguished in charged and neutral leptons, or neutrinos, based on their electromagnetic charge. Notice that no right-handed neutrino lepton is present in this summary, as no process involving right-handed neutral leptons has been ever observed up to this moment. Quarks, q in Equation 1.2, interact both strongly and electro-weakly: they act as triplets under SU(3) and therefore carry an α threefold color charge, and are equipped with fractional electromagnetic charges.

The presence of colored and non colored states in the SM naturally splits the model in two sectors: Quantum Chromodynamics (QCD), describing the strong interactions, and the electroweak sector. The SM Lagrangian can be therefore written as:

$$\mathcal{L} = \mathcal{L}_{QCD} + \mathcal{L}_{EW}. \tag{1.3}$$

The gauge invariance requirement over the two sectors leads to the introduction of a number of spin-1 bosons equal to the number of generators of the considered symmetry group. These gauge vector bosons configure as mediators to the interactions described in the SM.

Spin-1 fields enter the model through a process defined as "Lagrangian gauging": it consists in the substitution of the Standard directional derivative ∂ with a generalized *covariant derivative*, D. The covariant derivative is built by applying

a local transformation to the Lagrangian, and absorbing all the terms that would break the invariance in the new derivative definition.

The 8 mediators for the strong interaction are the gluons, G_c^α, color charged, electrically neutral massless particles. They appear within the $SU(3)_C$ covariant derivative D, defined as:

$$D^\alpha_{\mu\beta} = \partial_\mu \delta^\alpha_\beta - i g_F G_{\mu i} \frac{\lambda^{i\alpha}_\beta \gamma}{2}, \tag{1.4}$$

where λ_c are SU(3) generators and g_F is the strong coupling constant. The quark Lagrangian $\bar{q}(\gamma^\mu \partial_\mu)q$ results invariant under SU(3) local transformations when $(\partial \rightarrow D)$. Moreover, the gluon-gluon interaction term $G^\alpha_{\mu\nu} G^{\alpha\mu\nu}$ results as well invariant under local SU(3) transformations , and can enter in the QCD Lagrangian $\mathcal{L}_{QCD}$, which therefore results:

$$\mathcal{L}_{\mathcal{QCD}} = -\frac{1}{4} G^\alpha_{\mu\nu} G^{\alpha\mu\nu} + i \sum_i \bar{q}^\alpha_i \gamma^\mu D_{\mu c} q^c_i \tag{1.5}$$

Four electroweak bosons, $\{W^{1,2,3}_\mu, B_\mu\}$ are introduced to the model when requiring gauge invariance under $SU(2)_L \times U(1)$. SU(2) and U(1) covariant derivative D is of the form:

$$D_\mu = \partial_\mu - i g_W \sum_{i=1,2,3} W^{i\nu} \lambda^i_{\nu\mu} - i g'_B B^\nu \lambda_{\nu\mu}, \tag{1.6}$$

where g_W is the weak interaction coupling constant, g_B is the electromagnetic interaction coupling constant, and λs are the generators to the SU(2) and U(1) symmetry groups. Analogously to what has been described for gluons, the electroweak fermions Lagrangian becomes invariant under $SU(2)_L \times U(1)$ transformations when sending $(\partial \rightarrow D)$, and the mediators self interaction terms, defined as $F^\alpha_{\mu\nu} F^{\alpha\mu\nu}$, comply as well with the gauge local invariance. The electroweak Lagrangian can be written as:

$$\mathcal{L}_{EW} = -\frac{1}{4} F^\alpha_{\mu\nu} F^{\alpha\mu\nu} + i \sum_i \bar{f}^\alpha_i \gamma^\mu D_{\mu c} f^c_i \tag{1.7}$$

where f stands for a generic fermionic field, as electroweak interactions involve both quarks and leptons.

The electroweak physical mediators, $W^\pm$, Z^0 and the photon A_μ, configure as a linear combination of the SU(2) $\times$ U(1) gauge bosons:

$$W^\pm_\mu = \frac{W^1_\mu \mp i W^1_\mu}{\sqrt{2}} \tag{1.8}$$

$$Z^0_\mu = cos\theta_G W^3_\mu - sen\theta_G B_\mu \tag{1.9}$$

$$A_\mu = sin\theta_G W^3_\mu + cos\theta_G B_\mu. \tag{1.10}$$

Notice that a single mixing parameter, the weak mixing angle θ_G, regulates the boson mixing.

The $W^{\pm}$ and Z^0 weak bosons are self- interacting, massive particles; the $W^{\pm}$ carry an electromagnetic charge, while the Z boson is neutral. The photon, A_μ is the electromagnetic mediator: a massless, neutral particle, which does not interact with itself.

The SM dynamics description presented above does not naturally allow for the arisal of mass terms for either the fermionic and bosonic fields, as they would break gauge invariance. The mass content of the SM and mixing of electroweak bosons are in contrast with this constraint: in the next section, the mechanism which allows for the introduction of mass terms in the gauge invariant SM, is presented.

1.1.1 The Higgs Mechanism and Flavor Mixing

A spontaneous symmetry breaking mechanism allows for mass generation without breaking the gauge invariance of the SM Lagrangian. The model was published by Brout, Englert [28] and, independently, Higgs [38] in 1964. A new $SU(2)_L$ doublet, the Higgs field

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix}, \tag{1.11}$$

where ϕ^+ and ϕ^0 are respectively the doublet charged and neutral components, is introduced in the theory. The doublet acts as a Lorentz scalar and has Y = +1 and Q = 0. The Higgs scalar field is a self-interacting scalar field, hence its Lagrangian can be written as:

$$\mathcal{L}_H = (D^\mu \phi)^\dagger (D_\mu \phi) - V(\phi), \tag{1.12}$$

where the covariant derivative D is obtained by requiring gauge invariance over the ful symmetry group reported in Equation 1.1. It takes the form:

$$D_\mu = \partial_\mu + ig\sigma_j W_\mu^j + ig' Y B_\mu. \tag{1.13}$$

where g and g' are the coupling constant of fermions to $W^{1,2,3}$ and Z respectively, σ_j are the Pauli matrixes and Y is the weak hypercharge.

Interaction terms among the Higgs bosons and the fermionic fields can be added to the SM Lagrangian, provided that they preserve its gauge invariance. The Yukawa interaction terms couple left-handed fermion doublets to right-handed singlets through the Higgs field doublet, with a structure:

$$\mathcal{L}_{Yukawa} = \bar{\psi_i^L} Y_{ij} \psi_j^R \phi + h.c., \tag{1.14}$$

where ψ_i^L is a left-handed weak isospin doublet, ψ_j^R is a right-handed weak isospin singlet and Y_{ij} are the Yukawa couplings among fermions.

The Higgs scalar potential takes the form:

$$V(\phi) = \mu^2 \phi^\dagger \phi + \lambda (\phi^\dagger \phi)^2. \tag{1.15}$$

It depends on two parameters: the requirement $\lambda > 0$ ensures that the potential has a lower bound, while if $\mu^2 < 0$, the potential has a minimum:

$$\phi^\dagger \phi = \frac{\mu^2}{\lambda} = \frac{v^2}{2} \tag{1.16}$$

where the parameter v, also referred to as vacuum expectation value, *v.e.v*, might be in general different from 0. Spontaneous symmetry breaking is thought to have happened [32] at a critical, very high temperature in the primordial universe, when the *v.e.v* shifted from 0 to its current value ~ 246 GeV.

The effect of symmetry breaking in the system can be evaluated expanding the Higgs potential around its minimum. An expansion around one of the states satisfying Equation 1.16 and realizing the potential miminum can be performed to gauge the effect of a non null minimum on the model. The ground state choice is in principle arbitrary, yet a $v \neq 0$ value breaks the Lagrangian rotational invariance. As the photon is a massless mediator in the SM, the V potential ground state and expansion term are chosen in order to preserve $U(1)_{em}$ symmetry, and take the form:

$$\frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix}. \tag{1.17}$$

The Higgs kinetic term reported in Equation 1.12 carries interaction terms among the mediators and the Higgs fields. When the symmetry is broken with a non zero v value, the v dependent terms take a form:

$$\propto \mathcal{C}(v) G_i^{\mu} G_{\mu j}, \tag{1.18}$$

where $\mathcal{C}(v)$ holds all the scalar constants. Such terms are not diagonal in the SU(2) gauge bosons basis: the mass eigenstates are the electroweak mediators presented in Section 1.8. Notice that weak bosons are spin-1 particles, equipped with only two possible polarization directions, while a mass-given particle has to have three polarization degrees of freedom. Three scalar Goldstone [33] bosons arise from SU(2) symmetry breaking: they are respectively absorbed by the 3 electroweak bosons as a third polarization degree, which allows for the $W^{\pm}$ and Z bosons to acquire mass as well.

The masses of the weak bosons can be expressed as a function of the vacuum expectation value and the electroweak coupligs:

$$m_W = \frac{v}{2} g m_Z = \frac{v}{2} \sqrt{g^2 + g'^2} \tag{1.19}$$

The Yukawa interaction term is similarly projected in an interaction term depending on $h(x)$ and mass terms of the form

$$\propto \mathcal{C}(v)_f \bar{u}_i^L d_j^R + h.c. \tag{1.20}$$

where u and d here represent the upper and lower components of a fermion left-handed doublet, and i and j run on the flavor indexes of the two fields.

The mass terms in Equation 1.20 are in general not diagonal in the u and d fields. A rotation has to be applied on the fields to diagonalize the fermions mass terms: such operation acts on the whole Lagrangian, transforming the $W^{\pm}$ - fermions interaction terms in a non diagonal structure in the flavor basis. Each interaction term between two quarks of flavor *i,j* is equipped with an additional coefficient V_{ij}. This feature allows, in the quark sector, for the description of flavor changing interactions mediated by the weak charged bosons. The rotation of

the quark basis, however, does not spoil the interaction diagonal structure for the photon and the Z boson. Flavor changing weak interactions mediated by a neutral boson are therefore forbidden at tree level: they appear at loop level and are heavily suppressed by the GIM mechanism [31].

The V_{ij} coefficients brought to the electroweak interaction by the mass term diagonalization are summarizable in the Cabibbo-Kobayashi-Maskawa (CKM) matrix:

$$V = V_{u_L}^{\dagger} V_{d_L} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \tag{1.21}$$

The CKM coefficients are fully determined, excluding a single phase, which has to be included to allow for CP symmetry violation in weak interactions.

In the lepton sector, the absence of a right-handed neutrino has two crucial consequences: SM neutrinos are massless particles and no flavor mixing mechanism arises from the mass term diagonalization procedure. No mass term for the neutral lepton can be in fact obtained without either including a right-handed singlet or breaking gauge invariance. On the other hand, the absence of such a singlet allows for the definition of a lepton basis under which both the mass and interaction terms are diagonal, therefore no lepton flavor mixing arises in the model.

Once the diagonalization of the fermions mass terms is performed, it is possible to extract the fermion mass value as a function of the $v.e.v.$ Fermion masses take the form:

$$m_f = \frac{v}{\sqrt{2}} Y_f \tag{1.22}$$

where Y_f is the Yukawa coupling between the fermion f and the Higgs boson. The relation states that the Higgs couples more to heavier fermions, or inversely, that strongest couplings to the Higgs boson lead to higher masses.

Through the Higgs mechanism the gauge symmetry group is spontaneously broken, with a pattern:

$$SU(3)_C \times SU(2)_L \times U(1)_Y \rightarrow SU(3)_C \times U(1)_{em}, \tag{1.23}$$

The symmetry structure of the Standard Model does not feature gauge symmetries alone: the internal symmetry group yields a set of global symmetries, which are referred to as accidental symmetries. Accidental symmetries arise in models when the terms which might break the symmetry and preserve local gauge symmetries are too high in dimension to be included in the model.

The SM carries a global symmetry $U(3)^5$, which is broken, within the Higgs Mechanism, in four independent U(1) global symmetries over the electroweak and strong Lagrangians. The electroweak Lagrangian introduced in Equation 1.7, in presence of non null Yukawa couplings, is invariant under U(1) : applying Noether's theorem [47], a L conserved quantity is yielded by this invariance. Since the charged lepton basis diagonalizes the Higgs mass terms, the symmetry holds for the total number of leptons and for leptons of specific flavors, for which three conserved quantities are defined as $\{L_\mu, L_e, L_\tau\}$.
Analogously, theQCD Lagrangian in Equation 1.5, equipped with its own mass terms, results invariant under global U(1) transformations, yielding through the

Noether theorem a conserved quantity defined as barion number B. In the case of quark, flavor mixing prevents from defining a per flavor preserved number.

1.2 Open Questions in the Standard Model

The Standard Model is a mathematical solid theory and a successful method for predictions of the electroweak and strong interaction dynamics. The discovery, in 2012 by the ATLAS [1] and CMS [15] experiments at CERN, of a particle with characteristics comparable to the Higgs boson crowned the success of the SM in its current formulation.

The Standard Model is, however, far from being a complete or exhaustive theory in the description of our Universe. Multiple issues have to be accounted for, both within the model itself and concerning unpredicted evidence.

The SM theory holds 18 free parameters, which span over a vast range of energy scales. This feature does not align with the naturalness principle, which demands that the free parameters of a theory fall in the same order of magnitude.
Patterns appear in the SM free parameters: the lepton and quark masses show hierarchies in flavor which are not justified in the theory. In particular, the u,c,t quark masses are separated by at least 2 orders of magnitudes each, with the up quark u being the lightest at $m_u = 2.2$ MeV and the top quark t reaching $m_t = 173.1$ GeV. An analogous hierarchy occurs in the low components of the quark doublets and in the masses of the charged leptons: the down, strange and bottom quarks differ by at least to orders of magnitude among the three flavors, while the charged leptons masses go from the lightest 0.5 MeV electron, through the 105 MeV muon up to the 1.77 GeV tauon. Moreover, the loop contributions to the Higgs mass are of the order of 10^{18} GeV. They come in two, competitive contributions and appear to be precisely *fine tuned* to allow for the Higgs mass to be extremely small with respect to the value of the correction terms themselves.

Several pieces of evidence hint at the non-completeness of the SM theory: a brief summary of phenomena non explained by the SM is reported below.

- The SM is fully compatible with special relativity, yet it fails to embed general relativity and embrace all the four fundamental interactions in a single theory.

- Only 4% of the energy content of our universe is accounted for within the SM: evidence of galaxy rotation curve plateaus and enhanced gravitational lensing point to the existence of non-luminous, very weakly interacting matter, the dark matter, which might account for an additional 27% of the universe energy content. The remaining ∼69% can be framed by setting a constant energy density for the vacuum, the dark energy. Such an energy density is needed to account for cosmological observation showing the Universe expansion is accelerating when its mass content and global gravitational pull are expected to slow the expansion down and eventually stop it.

- The SM interactions are mostly symmetrical between matter and antimatter, yet the universe is dominated by matter. No exhaustive explanation for such a pronounced asymmetry is available within the SM.

- Finally, experimental observation of neutrino flavor oscillation prove that SM neutrinos are massive particles, while no mass term can be accounted for neutral leptons in the SM.

The last part of this chapter will be dedicated to the description of the neutrino mass puzzle and possible solutions. SM neutrino physics will be detailed, together with a review of the principal pieces of evidence for neutrino oscillation and mass given nature. A review of Standard Model extension accounting for neutrino masses by the inclusion of a right-handed neutral lepton will be presented, with particular focus on the description of the properties of right-handed massive neutrinos. Finally, a brief review of the searches targeting such right-handed neutral leptons is presented, together with the main features of the signatures investigated in this work.

1.3 The Neutrino Mass Puzzle

Neutrinos are among the most elusive components of the Standard Model. Neutral leptons were brought to the particle physics scenario by Wolfgang Pauli, as *desperate remedy* to the apparent energy conservation violation in heavy radioactive nuclei β^- decays. The particle proposed by Pauli would carry a part of the energy in the decay and escape standard means of detection. A novel experimental technique had to be developed for neutrino detection: the first neutrino was observed in 1956, by Cowan and Reines, in inverse β decays. The Homestake experiment [16], in the 1960s exploited the inverse β decay technique for the first-ever measurement of the solar neutrino flux. The measured flux resulted roughly a third of the expected one: the discrepancy opened a 20 years puzzle in neutrino physics. The reason for the Homestake experiment to measure such a deficit is to be found in neutrino oscillations, which had been proposed in a work of 1957, by Bruno Pontecorvo.

The SuperKamiokande experiment presented in 1998 the very first evidence of neutrino oscillations [40]. The experiment was designed to measure muonic and electronic fluxes for atmospheric neutrinos, thus relating a deficit in one of the flavors in excesses of the other one. SuperKamiokande results have been confirmed in 2001 by solar neutrino oscillation measurements published by the Subdury Neutrino Observatory [7]. In the later years, several experiments have measured neutrino oscillations amplitudes in solar [40], atmospheric [5], accelerator [3] and reactor [10] neutrinos.

The state of the art knowledge of neutrinos frames these particles as optimal probes to beyond the SM mechanisms. According to the SM, neutrinos are massless particles, as discussed in Section 1.1.1. However, the phenomenon of neutrino oscillations has provided compelling evidence that neutrinos have nonzero masses. Knowledge of the neutrino amplitudes is essential for exploring the possible mechanisms that could explain the origin of neutrino masses beyond the current formulation of the SM.

Results from the combination of the solar, atmospheric and accelerator neutrino oscillation measurements provided by the NuFit collaboration [29] point to a minimal mixing scenario between the three flavour neutrinos of the standard

model $\{\nu_e, \nu_\mu, \nu_\tau\}$ in at least three distinct mass eigenstates $\{\nu_1, \nu_2, \nu_3\}$. Estimates of the mass splitting among the three generations' masses have been derived by the neutrino oscillation amplitudes: neutrino masses are expected to be in the sub eV range. A hierarchy has been measured in the mass splittings, with $\Delta m^2_{21} << |\Delta m^2_{31}| \simeq |\Delta m^2_{32}|$ and $\Delta m^2_{ij} = m^2_i - m^2_j$, where the {1,2,3} indexes refer to the three neutrino mass eigenstates. The neutrino mass spectrum is usually presented in two different configurations:

- Normal Ordering (NO) with $m_1 < m_2 < m_3$ which parallels SM flavor hierarchy
- Inverted Ordering (IO) with $m_3 < m_1 < m_2$.

Coupling the hierarchy and the mass splitting relations, the neutrino mass spectrum can be linked to proper mass scales:

- Normal Hierarchical Spectrum (NH): $m_1 << m_2 < m_3$:
  $m_2 \simeq \sqrt{\Delta m^2_{21}} \sim 8.6\ 10^{-3}$eV, $m_3 = \sqrt{\Delta m^2_{32} + \Delta m^2_{21}} \sim 0.05$ eV
- Inverted Hierarchical Spectrum (IH): $m_3 << m_1 < m_2$:
  $m_1 = \sqrt{|\Delta m^2_{32} + \Delta m^2_{21}|} \sim 0.0492$ eV, $m_3 = \sqrt{\Delta m^2_{32}} \sim 0.05$ eV
- Quasidegenerate Spectrum (QD): $m_1 \simeq m_2 \simeq m_3 >> \sqrt{\Delta m^2_{32}}$.

In the next section, a review of minimal extensions to the Standard Model allowing for SM neutrino masses will be presented and current constraints and exclusions for the presented models will be discussed.

1.4 Neutrino Minimal Standard Model (νMSM)

The discussion reported in Section 1.1.1 highlights the impossibility of building a neutrino mass term with the current SM particle content, preserving gauge invariance and the model renormalizability. In principle, models including neutrino mass terms might be built loosening up either one of these conditions, or more than one at the same time.

A multitude of SM extensions has been produced throughout the years to explore possible neutrino masses arisal mechanisms: in this work, the focus is set on an SM extension referred to as Neutrino Minimal Standard Model (νMSM) [34] [35].

In νMSM, the SM particle content is extended by a set of three or more right-handed sterile neutrinos $\nu_{s,i=1,..,m}$, singlets under $SU(2)_L$. Right-handed neutrinos introduction allows the building of a gauge invariant, renormalizable mass term, of the form:

$$-\mathcal{L}_{m_\nu} = M_D^{ij} \bar{\nu}_{si} \nu_{Lj} + h.c. + \frac{1}{2} M_N^{ij} \bar{\nu}_{si} \nu^c_{sj} + \frac{1}{2} M_N^{ij} \bar{\nu}_{Li} \nu^c_{Lj} \tag{1.24}$$

where $\nu^c_{sj} = C\nu_{sj}$ and C is the charge conjugation operator, M_D is a complex $3 \times m$ matrix, while M_N is a symmetric $m \times m$.

The first neutrino mass term in Equation 1.24 is a Higgs mechanism driven term, and it is the direct consequence of the inclusion of a right-handed neutral particle to

the SM lepton content.The D subscript for this mass term points to the *Dirac* nature of the particles involved in the mass term, which necessarily need to be a particle and an antiparticle to let the mass term be a singlet under $SU(2)_L$. Neutrino mass values are linked to the Yukawa couplings as in Equation 1.22, therefore the Yukawa interaction between SM and right-handed neutrinos are expected to be mediated by minuscule Y_{ij} couplings.

The second and third term in Equation 1.24 have a more peculiar structure: $\bar{\nu_i}\nu_j^c$ is already a singlet under the SM gauge, therefore it can appear as a "bare" mass term. Moreover, the third term, which hosts two SM neutrinos, violating the global SM lepton number by two units.

The neutrino mass term Lagrangian can be rewritten as:

$$-\mathcal{L}_{m_\nu} = \frac{1}{2}(\overline{\vec{\nu}_L^c}, \overline{\vec{\nu}_s}) \begin{pmatrix} 0 & M_D^T \\ M_D & M_N \end{pmatrix} \begin{pmatrix} \vec{\nu}_L \\ \vec{\nu}_s^c \end{pmatrix} + h.c. = \frac{1}{2}\overline{\vec{\nu}^c} M_\nu \vec{\nu} + h.c. \tag{1.25}$$

where the SM and ν_s neutrinos have been condensed in a basis $\vec{\nu} = (\vec{\nu}_L, \vec{\nu}_s^c)^T$ of dimension (3+m). The M_ν matrix is complex and symmetric, and in general not diagonal in the $\vec{\nu}$ basis. M_ν is diagonalized by a V^ν matrix, such that the k = (3+m) mass eigenstates with eigenvalues m_k can be obtained by:

$$\vec{\nu}_{mass} = V^{\nu\,\dagger}\vec{\nu}. \tag{1.26}$$

The Lagrangian is now diagonal in the mass eigenstates:

$$-\mathcal{L}_{m_\nu} = \frac{1}{2}\sum_{k=0}^{3+m} m_k(\bar{\nu}^c_{mass,k}\nu_{mass,k} + \bar{\nu}_{mass,k}\nu^c_{mass,k}) = \frac{1}{2}\sum_{k=0}^{3+m} m_k \bar{\nu}_{M_k}\nu_{M_k}, \tag{1.27}$$

where ν_{M_k} is defined as:

$$\nu_{M_k} = \nu_{mass,k} + \nu^c_{mass,k}. \tag{1.28}$$

ν_{M_k} are referred to as Majorana Neutrinos, as they obey the Majorana condition

$$\nu = \nu^c. \tag{1.29}$$

Majorana neutrinos are described by a single field which represents both the particles and the antiparticle, and can therefore be represented by a two component spinor, unlike the charged fermions. For Dirac fields, as the charged leptons, particle antiparticle are represented by two different fields, and therefore a 4 component spinor is needed.

Two limits scenarios for the mass term presented in Equation 1.25 are described in the next two subsections.

1.4.1 Massless Scenario ($M_N = 0$)

The condition $M_N = 0$ is equivalent to imposing the lepton number conservation over the model. In this scenario, only the first, Standard Model like, term of Equation 1.24 is allowed. If sterile neutrinos are three, they can be identified with the

right-handed component of a four-spinor neutrino field. The Dirac mass term can be diagonalized through two unitary 3×3 matrices V^{ν} and V^{ν}_{R} as:

$$V_R^{\nu\,\dagger} M_D V^{\nu}. \tag{1.30}$$

The eigenstates for the diagonal matrix take the form:

$$\nu_{D\,k} = (V^{\nu\dagger}\vec{\nu}_L)_k + (V_R^{\nu\,\dagger}\vec{\nu}_s)_k. \tag{1.31}$$

The weak-doublet neutrino is then obtained by isolating the left-handed component of such eigenstate through the $\mathrm{P}_L = \frac{1-\gamma^5}{2}$:

$$\nu_L^i = P_L \sum_{j=1}^{3} V_{\nu}^{ij} \nu_{D\,j}. \tag{1.32}$$

Experimental observations constrain this mass term to very low energy and mass content with respect to the charged leptons mass terms. This features does not find any explanation in the SM framework.

1.4.2 Heavy Sterile Neutrinos ($\mathrm{M}_N \gg 0$)

If the mass eigenvalues to M_N are much larger than the scale of the electroweak symmetry breaking, the diagonalization M_μ leads to three light neutrinos ν and m heavy neutrinos N. The Lagrangian in diagonal form can be therefore split in two terms:

$$-\mathcal{L}_{m_\nu} = \frac{1}{2}\bar{\nu}_l M_l \nu_l + \frac{1}{2}\bar{N} M_h N, \tag{1.33}$$

where

$$M_l \simeq -V_l^T M_D^T M_N^{-1} M_D V_l \qquad M_h \simeq V_h^T M_N V_h \tag{1.34}$$

and V_l and V_h are 3×3 and $m \times m$ matrices. The lighter mass eigenvalues are proportional to M_N^{-1} while the heavy ones follow M_N: this is known as the see-saw mechanism.The heavy neutrinos result dominated by right-handed components, while the lighter ones are mostly left-handed, and both the states are Majorana particles.

Such a mechanism preserves SM symmetries and accounts for the low energy content of neutrino mass terms. The see-saw mechanism configures as a full theory whose low energy effective theory realization coincides with the SM with three light Majorana neutrinos.

The νMSM targeted by this work sees the introduction of $m = 3$ sterile right-handed neutrinos: the lightest N_1, has mass in the keV range and serves as possible Dark Matter candidate, while N_2 and N_3 are quasi-degenerate heavy neutral leptons with masses in an interval ranging from around one to several hundreds of GeV. The search described in the next chapters targets the heavy N_i candidates. From the experimental point of view N_2 and N_3 are undistinguishable and will be therefore generally referred to as a single type of particle, the heavy neutral lepton N.

1.4.3 Neutrino Flavor Mixing

As discussed in Section 1.1.1, the electroweak Lagrangian can be written in a diagonal form for both the electroweak interaction terms and the mass terms, due to the absence of right-handed neutral leptons. The νMSM exercise, however, deals with the introduction of right-handed sterile neutrinos to the model, and therefore spoils the freedom of the defining a basis for the neutral leptons under which both the electro-weak terms and mass terms result diagonal.

The electroweak Lagrangian charged interaction term can be written in this general scenario as:

$$-\mathcal{L}_{EW} = \frac{g}{\sqrt{2}}(\bar{e}_L, \bar{\mu}_L, \bar{\tau}_L)\gamma^\mu U \begin{pmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \\ : \\ \nu_n \end{pmatrix} W_\mu^+ + h.c. \tag{1.35}$$

where U is a 3 $\times n$ unitary matrix mixing the SM lepton doublets to the neutrino mass basis defined in 1.25. The mass term for the electroweak sector becomes

$$-\mathcal{L}_{M_\ell} = [(\bar{e}_L, \bar{\mu}_L, \bar{\tau}_L)M_\ell \begin{pmatrix} \bar{e}_R \\ \bar{\mu}_R \\ \bar{\tau}_R \end{pmatrix} + h.c.] - \mathcal{L}_{M_\nu} \tag{1.36}$$

where M_ℓ can be diagonalized through two unitary V^ℓ and V_R^ℓ matrix. The diagonal Lagrangian will then assume the form:

$$-\mathcal{L}_{M_\ell} = \sum_{k=1}^{3} m_{\ell_k} \bar{\ell}_k \ell_k, \tag{1.37}$$

where

$$\ell_k = (V^{\ell\,\dagger}\ell_L)_k + (V_R^{\ell\,\dagger}\ell_R)_k \tag{1.38}$$

are the mass eigenstates for the charged leptons. The weak doublet charged lepton can be extracted from Equation 1.38, inverting the expression and projecting on left-handed eigestates only:

$$\ell_{Li} = P_L \sum_{k=1}^{3} V_{ij}^\ell \ell_j, \tag{1.39}$$

where the i index runs over the leptonic flavors.

The structure for the U matrix can be now derived from Equations 1.39 and 1.32 and can be written as:

$$U_{ij} = \mathcal{P}_{\ell,ii} V_{ik}^{\dagger\,\ell} V_{kj}^\nu (\mathcal{P}_{\ell,jj}), \tag{1.40}$$

where the $\mathcal{P}$ matrixes are phase matrixes. For n $>$ 3 Majorana [Dirac] neutrinos the U matrix contains a total of $6(n-2)$ $[5n-11]$ real parameters, of which $3(n-2)$ are angles and $3(n-2)$ $[2n-5]$ can be interpreted as physical phases.

In the $m = 3$ scenario, U degenerates in the Pontecorvo-Maki-Nakagawa-Sakata (PMNS) [44]:

$$U = \begin{pmatrix} U_{e1} & U_{e2} & U_{e3} \\ U_{\mu 1} & U_{\mu 2} & U_{\mu 3} \\ U_{\tau 1} & U_{\tau 2} & U_{\tau 3} \end{pmatrix} = \tag{1.41}$$

$$\begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta_{CP}} \\ -s_{12}c_{23} - c_{12}s_{13}s_{23}e^{i\delta_{CP}} & c_{12}c_{23}s_{12}s_{13}s_{23}e^{i\delta_{CP}} & c_{13}s_{23} \\ s_{12}s_{23} - c_{12}s_{13}c_{23}e^{i\delta_{CP}} & c_{12}s_{23}s_{12}s_{13}c_{23}e^{i\delta_{CP}} & c_{13}c_{23} \end{pmatrix}$$

The matrix has an analogous structure to the CKM quark mixing matrix introduced in Section 1.1.1 but, due to the Majorana nature of neutrinos, it is equipped with 6 phases rather than 3. As in the CKM matrix parametrization, c_{ij} and s_{ij} stand for $cos\theta_{ij}$ and $sin\theta_{ij}$, the θ_{ij} lie in the first quadrant $\theta_{ij} \in [0, \frac{\pi}{2}]$ and the phases δ_{CP}, η_{ij} fall in $[0, 2\pi]$.

The general expression in Equation 1.4.3 includes flavor mixing for charged leptons, yet if such particles do not have any other interaction than the SM ones, their interactions and mass eigenstates differ only by a phase. In this scenario, U is a 3 $\times n$ matrix describing flavor mixing among neutral leptons only.

The parametrization introduced in Equation 1.4.3 holds in the hypothesis of three Dirac light neutrinos, for which U is a unitary matrix rotating the light states in linear combinations of each other. Under the Majorana hypothesis, when three light neutrinos are linked by the see-saw mechanism to three heavy neutral lepton sterile counterparts, the unitarity is violated by factors of the order of $\mathcal{O}(M_D/M_N)$. The violation is very small and experimentally constrained: for this reason, U is usually considered unitary independently from the nature of neutrinos.

1.5 Experimental Limits on Heavy Neutral Leptons

In this section a review of the currents exclusion limits on Heavy Neutral Leptons is presented. HNLs can be described of four parameters: the heavy neutrino mass, m_{HNL} and the couplings to the three flavor families $V_e = U_{eN}$, $V_\mu = U_{\mu N}$, $V_\tau = U_{\tau N}$, where U_{ij} are the couplings introduced in the PMNS matrix reported in Equation 1.4.3, for the $M_N \gg 0$ scenario described in Section 1.4.2.

Exclusion limits for the existence of HNL candidates are set in the coupling V_f^2 versus HNL mass m_{HNL} plane. Most of the experimental results set limits on heavy neutral leptons coupling exclusively to a single SM flavor. Summary plots for current limits [24] are reported, for each exclusive flavor coupling scenario, in Figure 1.1

Limits are currently set, for all the SM flavor exclusive couplings, down to $|V|_f^2 \sim 10^{-4}$ for all the explored mass range. Muon and electron flavor exclusive couplings have limits reaching $\sim 10^{-8}$ for the very low mass range, $m_N \in [0.1, \sim 0.3]$ GeV.

Exclusive coupling scenarios are experimentally easier to probe, and provide clean benchmarks for comparison among experiments. Yet, recent studies on neutrino flavor oscillations seem to favor scenarios with spurious couplings to at least two SM flavors.

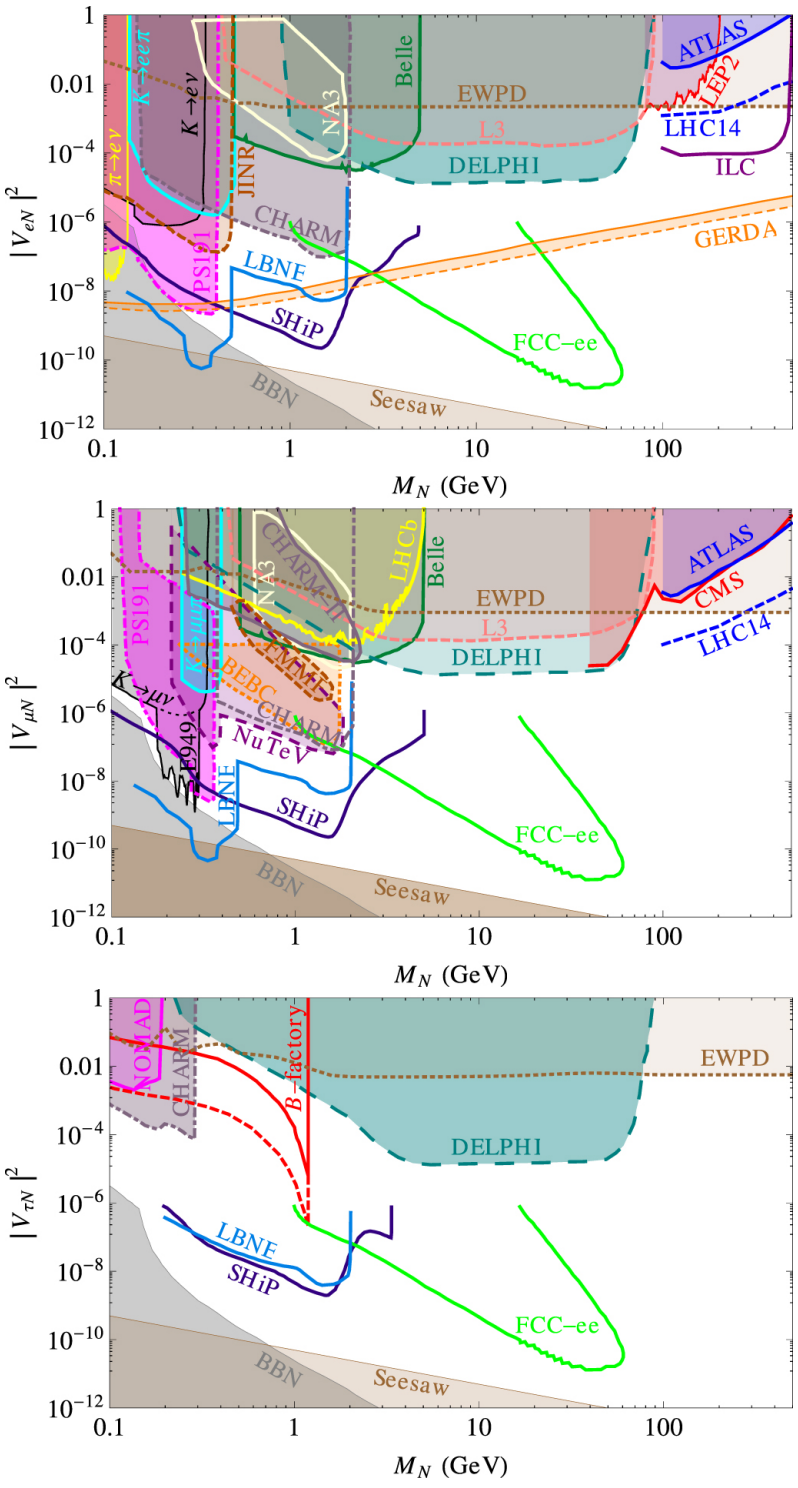


Figure 1.1. Current limits [24] on Heavy Neutral Leptons coupling exclusively to electron (top), muon (center) and tau (bottom) neutrinos as a function of the coupling and the HNL mass m_{HNL}.

In Figure 1.2 current probability contours for U^2_{if}/U^2, where f runs over the lepton flavor index, are reported for a model with $m = 2$ heavy neutrinos and normal (top) or inverted (bottom) mass ordering are reported [27]. The 1σ (darkest), 2σ and 3σ contours are obtained by a global fit on present neutrino oscillation data performed with NuFit 3.1 [29]. The contours for both the mass ordering do not currently point, for a minimal $m = 2$ νMSM, to exclusive couplings to a single SM flavor.

These results are extendable to νMSM such as the ones targeted by this work, as the HNLs introduced in Section 1.4.2 are, from a phenomenological point of view, indistinguishable from the two heavy neutrinos of the minimal $m = 2$ hypothesis and very small mixing is expected for the light right-handed neutrino candidate.

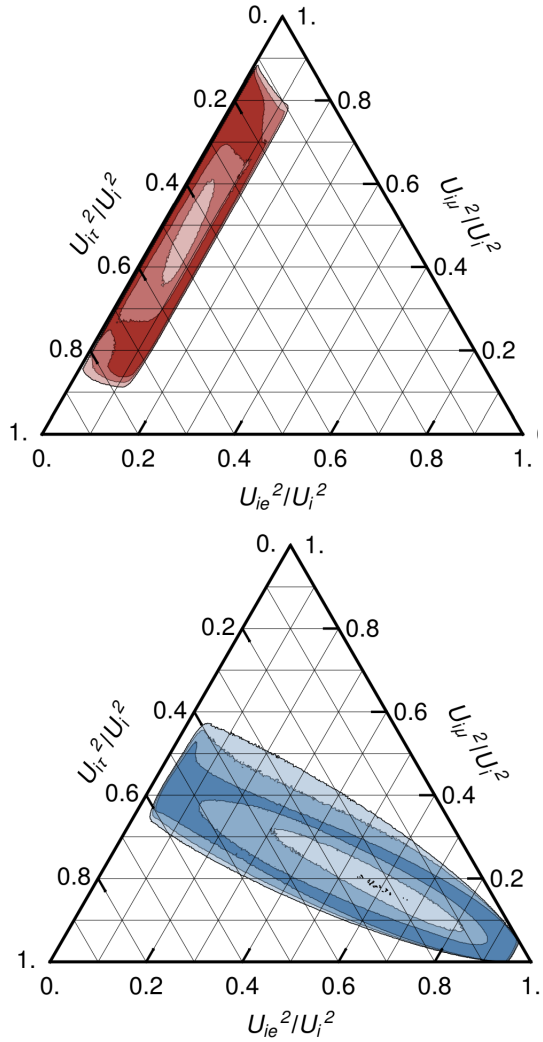


Figure 1.2. 1σ (darkest), 2σ and 3σ probability contours for U^2_a/U^2 as extracted by global fit to present neutrino oscillation data performed with NuFit 3.1.[27][29]. The top plot shows contours for normal mass ordering, while the bottom plot represents contours for the inverted one.

1.6 Heavy Neutral Leptons in B Meson Decays

The search presented in this work targets HNLs produced in B meson decays. This choice limits the range of possible HNL masses investigated through this channel to $m_{HNL} < m_B$, yet provides a clear experimental signature.

The Feynman diagram for the process is reported in Figure 1.3:

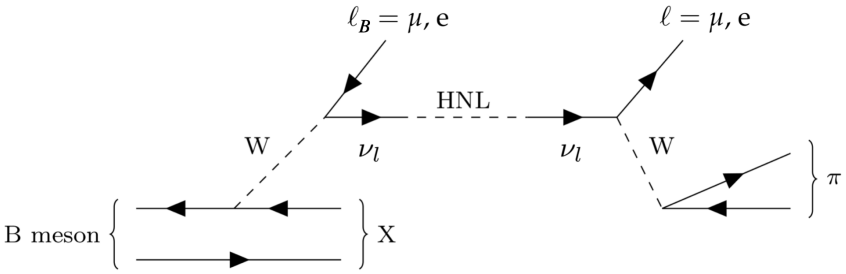


Figure 1.3. Feynman diagram of a semi-leptonic B meson decay to an N labeled as HNL in the diagram and a charged lepton (μ or e). The heavy neutrino N decays weakly to $\mu^{\pm}\pi^{\mp}$ or $e^{\pm}\pi^{\mp}$. At least one muon must be present in the process.

The decay chains considered as signal signatures are:

$$B \to \ell_B N + X, \;\; N \to \ell^{\pm}\pi^{\mp}, \tag{1.42}$$

where:

- the HNL is labeled as N;
- B is one of the following mesons: B^+, B^0, B_s^+ or their charge conjugates;
- in general, ℓ_B and ℓ can be an electron or a muon.

This search is made possible by the B-Parking dataset, collected by the CMS experiment in 2018, which requires the presence of at least one muon in the event with a production vertex that is not compatible with the primary proton-proton collision, as will be detailed in Section 2.2.4. Because of the B meson lifetime ($c\tau \sim 0.5$ mm), the ℓ_B production vertex can be easily distinguished from the primary proton-proton interaction vertex. Furthermore the N can be long-lived, therefore the lepton ℓ produced in its decay can be even more displaced. These displacement-related features provide a powerful tool for the rejection of background sources.

In light of these considerations, three decay modes are going to be considered in the analysis:

- B $\to \mu$N+X, N $\to \mu^{\pm}\pi^{\mp}$
- B $\to \mu$N+X, N $\to e^{\pm}\pi^{\mp}$
- B $\to$ eN+X, N $\to \mu^{\pm}\pi^{\mp}$

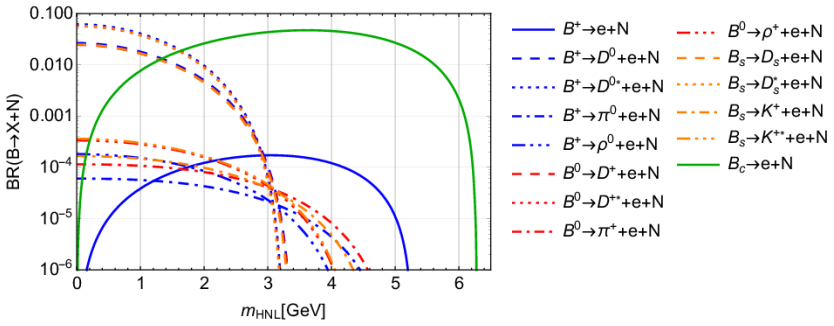


Figure 1.4. Branching ratios for different $B \to \ell N + X$ channels. Here the case $\ell = e$ is shown as an example. The $\ell = \mu$ and $\ell = \tau$ scenarios show similar behavior.

The first process contains exactly two muons and will be referred to as the $\mu\mu$-*channel* while the other two processes, containing exactly one muon and one electron, will constitute the μe-*channel*. The μe-*channel* is the main focus of the work presented in this thesis.

In general, three categories of processes can contribute to heavy neutrino production in B meson decays:

- fully-leptonic decays ($B \to \ell N$);
- semi-leptonic decays ($B \to \ell N + X$) involving a pseudo-scalar meson ($X = \pi, K, D$);
- semi-leptonic decays ($B \to \ell N + X$) involving a vector meson ($X = \rho, K^*, D^*$).

The amplitudes of such processes are obtained in the Fermi theory framework, following the calculations reported in [13]. Examples of branching ratios for different $B \to \ell N + X$ decays are reported in Figure 1.4 [13] as a function of m_N.

Due to CKM suppression, the branching ratio for $B^\pm \to N\,\ell^\pm$ is significantly smaller than $B_c \to N\,\ell$. The fragmentation fractions for the production of different B meson species at LHC energies are $f_u \sim f_d = 0.4$ [52], for the $B^\pm$ and B^0 respectively, while the fragmentation fraction for B_c production, as measured in [2], results $f_c = 2.61 \cdot 10^{-3}$. B meson species production at LHC favors the charged B meson over the B_c by a factor of ~ 150. For this reason, the B_c production channel contribution becomes relevant in the analysis context only for $m_N > 3$ GeV.

The decay widths of the N are calculated in electroweak theory following [13]. The targeted signature sees the N decay hadronically through a charged current in the process

$$N \to \ell_\alpha^+ u\bar{d} \tag{1.43}$$

where α runs over the electron and muon flavors, and the up and down quark are bound in a pion in the final state. N is assumed to be a Majorana particle in the search, therefore the charge conjugate of these processes has to be considered as well.

The process shown in Figure 1.3 involves a lepton in both the heavy neutrino production and decay vertexes. Such vertexes are accessible to the heavy neutrino

through mixing with SM neutrinos. The mixing is parameterized by the HNL to ν_{SM} couplings $V_e = U_{eN}$, $V_\mu = U_{\mu N}$, $V_\tau = U_{\tau N}$, where U_{ij} are the PMNS matrix elements in Equation 1.4.3.

The total coupling of the heavy neutrino to SM neutrinos can be then defined as:

$$|V|^2 = |V_e|^2 + |V|^2_\mu + |V|^2_\tau \tag{1.44}$$

and coupling fractions can be defined as functions of the total coupling as:

$$f_\ell = \frac{|V_\ell|^2}{|V|^2}, \tag{1.45}$$

with $\ell = e, \mu, \tau$ and $f_e + f_\mu + f_\tau = 1$.

The heavy neutrino lifetime can be written as a function of the total coupling, the Γ_ℓ decay widths to a lepton of flavor ℓ and the coupling fractions through the relation:

$$\begin{aligned} \frac{1}{\tau} &= |V|^2\Gamma & (1.46)\\ &= |V_e|^2\Gamma_e + |V_\mu|^2\Gamma_\mu + |V_\tau|^2\Gamma_\tau & (1.47)\\ &= |V|^2 \left(\Gamma_e f_e + \Gamma_\mu f_\mu + \Gamma_\tau f_\tau\right). & (1.48) \end{aligned}$$

The targeted signature requires the presence of a muon in the final state, therefore $V_\mu \neq 0$. The *$\mu\mu$-channel* can be used to probe the pure muon coupling scenario and to test possible spurious couplings by assigning a $f_\mu \neq 1$ value to the muon coupling fraction. On the another hand, the presence of a muon and an electron is required in the *μe-channel*: this enforces $V_e \neq 0$ and lepton flavor violation in the final state. The channel allows exclusively for spurious ($V_\mu \neq 0$ and $V_e \neq 0$) coupling tests. The combination of the two channels allows probing of spurious coupling scenarios, where in general $V_e, V_\mu, V_\tau \neq 0$, with enhanced sensitivity with respect to searches focusing on mono-flavor signatures only.

Chapter 2

The CMS Experiment at the Large Hadron Collider

This chapter is dedicated to the description of the experimental apparatus which made the search presented in this work possible. The Large Hadron Collider complex, which provided the collision data used in this thesis is described in Section 2.1, together with a brief description of its future prospects. The CMS experiment, a multi-purpose detector which collected and analyzed those data, is detailed in Section 2.2. In particular, a dedicated discussion on the novel trigger and data-taking strategies exploited in B-Parking dataset is presented in Section 2.2.4.

2.1 The Large Hadron Collider

The Large Hadron Collider (LHC) is the largest particle collider ever built. It is a proton-proton collider based at CERN, near Geneva, Switzerland. It consists of a 27 km main ring of superconducting magnets, which bend the protons to follow the ring trajectory, lined with accelerating radiofrequency cavities. In Figure 2.1 a schematic description of the CERN accelerating complex and of the primary and secondary experiments and beams is presented.

Protons undergo a process of collimation and acceleration before collision: they are extracted from an hydrogen bottle and accelerated in small groups, "bunche", in the linear accelerator Linac2 until they reach 50 MeV. They are then injected in the Proton Synchrotron Booster (PSB), a 4 rings accelerator, in which they reach the energy of 1.4 GeV (28 times their initial energy). The PSB provides the protons with enough energy to be injected in the Proton Synchrotron (PS), an accelerating ring of 628 m, in which the protons are brought to 25 GeV and the bunches are spaced so that there is a 25 ns time interval among them.
The bunches are then sent into the Super Proton Synchrotron (SPS), where they are accelerated to 450 GeV; finally, the beams are injected into the main ring in opposite directions and in separate ultrahigh vacuum beam pipes. Here, they are bent to follow the ring by the magnets and accelerated, through the radiofrequency cavities. When they reach their maximum energy, they are brought to collision in four points, where the four main experiments are located: ATLAS (*A Toroidal LHC Apparatus*), CMS (*Compact Muon Solenoid*), ALICE (*A Large Ion Collider*

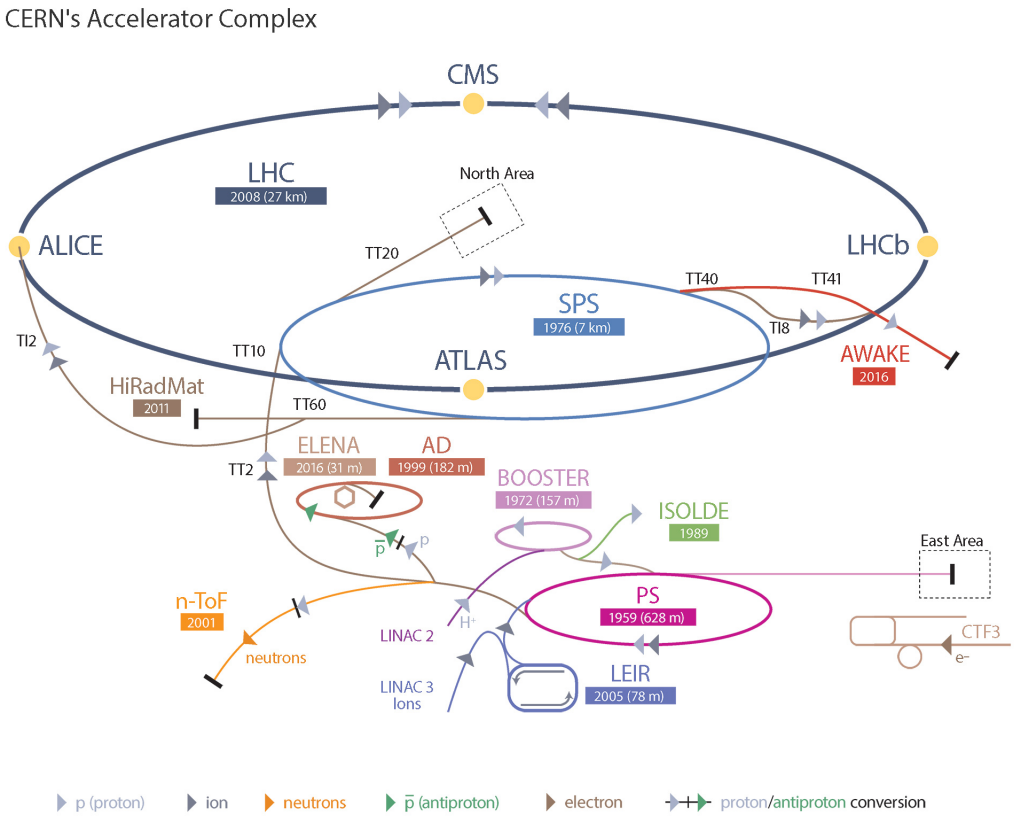


Figure 2.1. Structure of the CERN accelerating complex. Linac2, Booster, SPS and LHC represent the acceleration chain protons undergo before collision; interaction points are shown along the LHC ring, together with on beam facilities and heavy ions experiments.

Experiment) and LHCb (*LHC beauty*).

LHC is a hadronic collider: hadrons are complex structures, formed by a system of smaller objects, quarks, held together by the strong force mediating particles, the gluons. When accelerated at energies much higher then their own masses, hadrons do not behave as a coherent system anymore: the collisions happen among their constituents, which carry a fraction of the proton momenta, so that the effective center of mass energy varies from event to event. This feature makes LHC a "discovery collide", as it allows to probe a wide range of energies.

An outline of LHC program is shown in Figure 2.2. The collider has been fully operational since 2009 and will continue its activity at least until 2037.

The collider life is articulated in Runs, for collisions and data taking, and Shutdowns, for maintenance and upgrades.

Runs are characterized by stable values of two main collider variables: instantaneous peak luminosity and centre of mass energy ($\sqrt{s}$). During the the Long Shutdowns, which coincide with the experiment upgrade Phases, major interventions on the machine and the accelerating complex allow for modification and stabilization of these parameters for the future run.

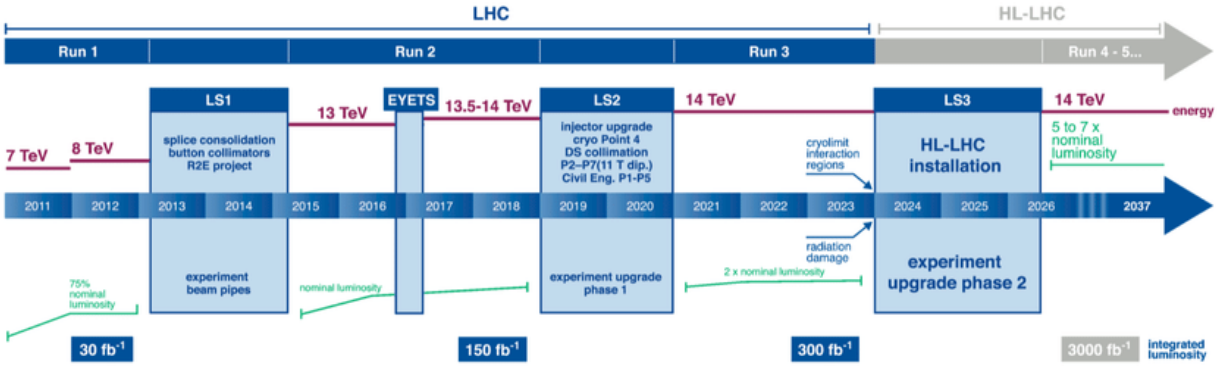


Figure 2.2. A timeline for LHC operations until 2037. Center of mass energies are indicated in red and luminosity growth in green. Nominal LHC luminosity is $L = 1{\cdot}10^{34}$ cm^{-2} s^{-1}. The integrated luminosity for each Run is also presented.

The instantaneous peak luminosity is defined as:

$$L = \gamma \frac{n_b N^2 f_{rev}}{4\pi\beta^* \epsilon_n} R, \tag{2.1}$$

where γ is the relativistic gamma factor, n_b is the number of bunches colliding at interaction point (IP), N is the number of protons per bunch, f_{rev} is the bunches revolution frequency in the ring, β^* is the beam focal length and ϵ_n is the beam transverse normalized emittance and R is a luminosity geometrical reduction factor. The peak luminosity was gradually risen, as can be seen in the green line in Figure 2.2, through LHC operating years by tuning the beams focus in proximity of the interaction regions.

The center of mass energy, $\sqrt{s}$, in red in Figure 2.2, was set to 7 TeV at LHC start, and brought to 13 TeV when passing from Run1 to Run 2, by the upgrade of the accelerating system and the introduction of more powerful superconducting magnet for beam bending. The record energy of 13.6 TeV has been reached in 2022 Run3 collisions.

Another key parameter is the integrated luminosity $\int L(t)dt$: the peak luminosity varies in time due to continue beam focusing and to the progressive degradation the beams suffer at each interaction. The integrated luminosity is a useful value to estimate the number of collisions happening at LHC, through

$$N = \sigma_{pp} \int dt L(t), \tag{2.2}$$

where N is the number of collisions, σ_{pp} is the cross section for a proton-proton interaction, and the integral is the integrated luminosity. The number of events for a certain final state can be similarly retrieved, as

$$\sigma_{pp} = \sum_{process}^{i} \sigma_i \qquad\qquad N_i = \sigma_i \int dt L(t). \tag{2.3}$$

where i is the final state of interest, and N and σ are the number of events and the expected cross section for that process.

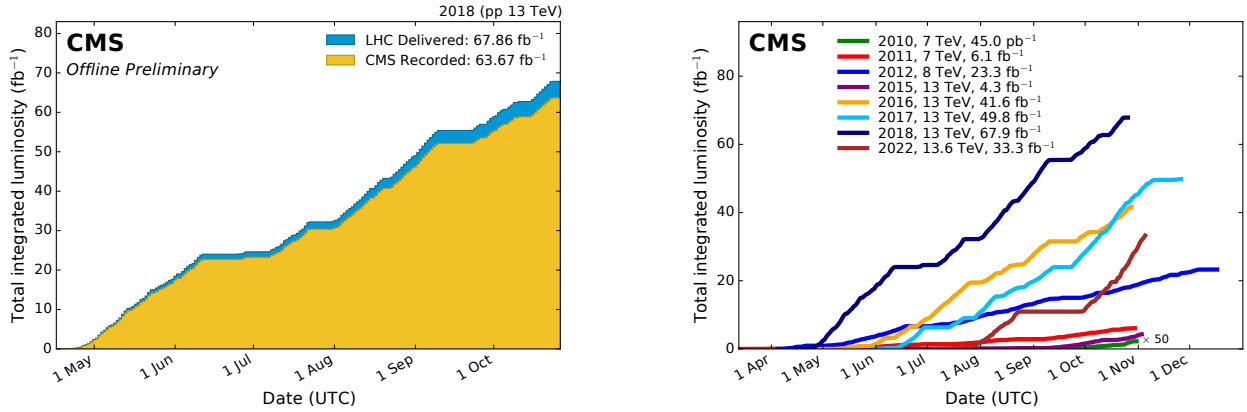


Figure 2.3. Left: integrated luminosity recorded ar CMS from May 2018 to November 2018. Right: CMS integrated luminosity data per year from 2010 to 2018

Figure 2.3 right plot shows the integrated luminosity delivered to CMS through LHC operating time in a year, for different years. The integrated luminosity has grown from year to year, nearly doubling between 2016 and 2018.

The first LHC operational run, Run1, consisted of two years (2009-2011) of beam energies ramping: during its first months of activity, LHC beat Tevatron's 0.98 TeV per beam energy record, becoming the world's highest-energy particle accelerator. The beam energies were then further increased, reaching 3.5 TeV in 2011 and 4 TeV in 2012. July 2012 coincides with the discovery of a new particle of 125 GeV in mass, identified as the Higgs boson, one of the main aims and greatest achievements of the LHC project.

The first collider shutdown (LS1) came after reaching 30 fb^{-1} of integrated luminosity, with peak luminosity reaching the $\sim 75\%$ of the design value. During LS1, the acceleration cycle was optimized to let each beam reach 7 TeV, while the ring magnet were trained to bend 6.5 TeV beams. Run2, which followed LS1, started in 2015, delivering 13 TeV center of mass energy, and reaching and doubling the design luminosity of the collider. Run2 provided the experiments with five times the integrated luminosity acquired during Run1. It has ended in December 2018, delivering a peak luminosity of $2.14 \cdot 10^{34}$ cm^{-2} s^{-1} at $\sqrt{s} = 13$ TeV and exceeding the LHC design luminosity $L = 1 \cdot 10^{34}$ $\text{cm}^{-2}\text{s}^{-1}$. 2018 integrated luminosity is reported in Figure 2.3 left plot.

LHC experiments have delivered, during their activity, precision measurements that are in fair agreement with the SM predictions or have set limits to Beyond the Standard Model (BSM) phenomena.

The collider data taking activities has resumed for Run 3, in July 2022.

LHC will be upgraded during another long shutdown, LS3, to reach peak luminosities a factor 5 to 7.5 larger than the nominal LHC luminosity, in the High Luminosity program (HL-LHC) [9]. By 2023 the beam focusing quadrupoles around the interaction points of ATLAS and CMS will be at the end of their lives, due to radiation damage. They will be replaced with new quadrupole triplets; additional crab cavities will be added to optimize bunch overlaps at crossing. These changes,

together with a new scheme for bright bunch trains in the PS (expected to take place during LS2), will allow to obtain brighter beams and smaller beam sections at collision and, thus, to significantly increase LHC luminosity.

2.2 The Compact Muon Solenoid

The Compact Muon Solenoid experiment is a general purpose detector. It is formed by different detector layers, concentric and roughly cylindrical, in a typical "onion" shape structure. Different layers are specialized in different measurements or sensitive to a specific category of particles.

2.2.1 Coordinate System

The conventional reference frame for CMS is right-handed and centered in the interaction point. It sees the longitudinal z axis parallel to the beam line with counter clockwise direction, the y axis directed towards the surface and the x axis pointing to LHC ring center. The angular variables are, with respect to the beamline, φ, azimuth angle, for the xy-plane and θ, polar angle, for the zy-plane.
The pseudorapidity, η, defined as

$$\eta = -\ln \tan \frac{\theta}{2}, \tag{2.4}$$

is preferred to θ because, in the massless particle approximation, which is applicable in most of LHC hard interactions, it coincides to the rapidity

$$y = \frac{1}{2} \ln \frac{E - p_z}{E + p_z}, \tag{2.5}$$

where c = 1. The rapidity goes to 0 when the longitudinal component of the momentum goes to 0, and to $-\infty$ when the momentum is purely longitudinal. Differences in rapidity are an invariant under Lorentz boosts along the beam axis: they transform linearly so that Δy, or for massless particles, $\Delta\eta$, does not depend on the longitudinal boost of the reference frame. This is a fundamental feature for hadron collider physics, as the partons inside the hadrons carry different fractions of the original longitudinal momentum, leading to collisions with different boosts along the beam axis.

2.2.2 Sub-detectors

Most detectors are naturally split in a central, barrel section, which has a tubular geometry around the beam line, and two forward regions, the endcaps, that increase the detector acceptance to the regions in the immediate surroundings of the beam line far from the interaction region. A schematic view of the CMS detector and a section with particles passing through it are in Figure 2.4.

The main layers of CMS detector are [17]:

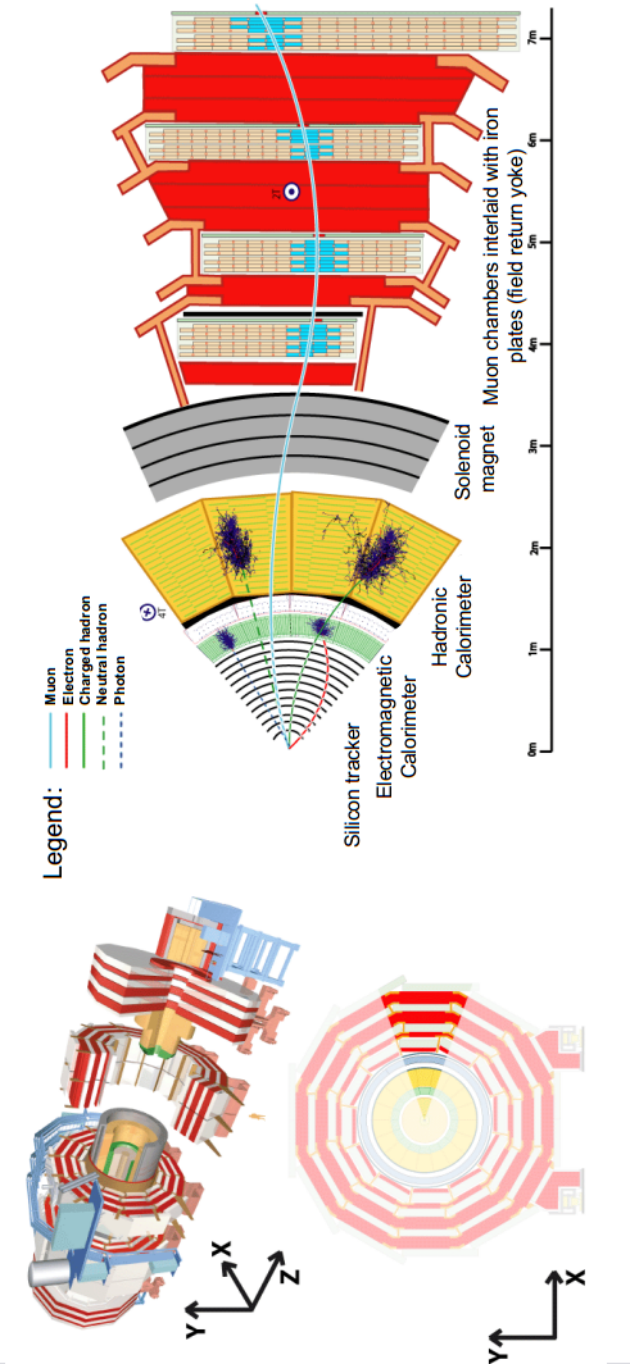


Figure 2.4. Views of the CMS detector. Top left: exploded view of CMS detector from one of the endcap; bottom left: CMS section; right: radial view of the subdetectors with particles stopping in different layers.

Solenoid Magnet

A coil of superconducting wires that creates a magnetic field when current runs through it. It has an overall length of 12.5 m, a diameter of 6 m, and it is provided with a return yoke to control the field and bend particles in the detector peripheral regions. It contains the tracker and the calorimeters, and generates a magnetic field directed along the beam axis, of 3.8 T, uniform inside the magnet coil. This intense magnetic field is necessary to bend charged particles in the detectors, as in Lorenz law, to identify the sign of their charge and measure their transverse momentum with respect to the field direction.

Tracker

The innermost detector, it encloses the beamline, has an overall radius of 130 cm and receives the higher particle fluxes. Its main aim is to provide information on charged particle paths with excellent resolution and without sensibly perturbing their kinematic parameters. The tracker is formed by silicon detectors and divided into two sections: the Pixel detector and the Silicon Strips detectors.
The Pixel is formed by four cylindrical barrel layers and three endcap disks of hybrid pixel detectors, 100×150 mm^2 in size; it has maximum transverse radius of 16 cm, while the endcap disks are positioned within 50 cm from the interaction point. The Silicon Strips detector reaches a transverse radius of 1.2 m, covering a region $|\eta|< 2.5$. Each strip is between 320 μm and 500 μm in thickness with a pitch that varies from 122 μm to 205 μm . The high segmentation and response of silicon provides the position measurement in layers with a $\sim$ 10 μm resolution, allowing for precise tracks reconstruction through the different tracker layers.

Electromagnetic Calorimeter (ECAL)

A homogeneous calorimeter hermetically enclosing the tracker and specifically built to measure electron and photon energies with high precision. It is composed of roughly 75 000 crystals, divided into a barrel section (EB), covering the pseudo-rapidity region $|\eta| < 1.479$, and an endcap section (EE) consisting of two disks, covering the $1.479 < |\eta| < 3.0$ region. Lead tungstate is characterized by small radiation length, $X_0 \sim 0.89$ cm, and Molière radius, $R_M = 1.96$ cm, providing optimal electromagnetic shower containment, has fast response (15 ns) and it is highly radiation tolerant.
ECAL energy resolution for electrons and photons varies with the particle energy through the expression

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{S}{\sqrt{E}}\right)^2 + \left(\frac{N}{E}\right)^2 + C^2 \tag{2.6}$$

where $S = 2.8\%$ is the stochastic term, $N = 12\%$ the noise and $C = 0.3\%$ the constant term [46]. The relative resolution with respect to the particle energy goes from 1.5% to less than 0.4%. ECAL deposits are also processed to obtain information on the centroid position.The electromagnetic showers in ECAL deposit different amount of energies in the crystals, the centroid position is the average of the position of the crystals, weighted by the energy deposited in each crystal. A detailed

discussion of ECAL deposit reconstruction is reported in Chapter 3.

Hadronic Calorimeter (HCAL)

A sampling, hermetic calorimeter, specifically suited for measurements of hadrons. It is formed by alternating layers of brass absorber and fluorescent scintillating material and it is highly hermetic. The hermeticity is crucial in order to account for the total energy brought by hard components of the interactions and provide a good resolution for the missing transverse energy.
The energy resolution in HCAL depends on the objects showering and it is function of the deposited energy. For hadronic jet showers

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{S}{\sqrt{E}}\right)^2 + C^2 \tag{2.7}$$

with S = 125% stochastic term and C= 5% constant term [20].

Muon System

Formed by four layers of drift cambers interlined with the magnet return yoke in the barrel, while composed by Cathode Strips Chambers (CSC) and Resistive Plate Chambers (RPC) in the endcaps. In the barrel region, the magnetic field provided by the return yoke is $\sim$ 1.5 T, so that the muon trajectories are bent, and muon momentum can be retrieved measuring the curvature radius from the particle path through the drift chambers. The use of both CSCs and RPCs in the endcaps provides for fast and precise position measurements and a redundant trigger for rapid data selection.

Upgrade

CMS has planned an upgrade for its detectors for LHC High Luminosity program. HL-LHC will provide the experiments with greater amounts of collisions and, therefore, data; the increase in data acquisition is a breach towards new physics, but represents also a major challenge for the detectors. The main aim of the CMS upgrade is to preserve the experiment current performance, in terms of background rejection and particle identification, in the more challenging HL-LHC frame.

2.2.3 Trigger

The LHC delivers 40 million proton-proton collisions per second. A typical raw collision event size is of the order of 1 MB, hence it is not possible to record all the events. The vast majority of collision events consists, on the other hand, of soft interactions, which are not interesting for the CMS physics program. The CMS trigger system has the scope of selecting interesting events while maintaining the rate of recorded collisions below 400 Hz. It is a two level trigger system formed by a Level 1 (L1) handware system and High Level Trigger (HLT) running high-level physics algorithms.

L1 [58] is a fixed latency hardware trigger. Within 4 μs, it determines whether an event should be accepted or rejected based on information coming from the calorimeters and the muon system. These subdetectors Trigger Primitives (TP), which consists on information on deposits passing energy thresholds for an high energy event and quality flags, are collected and combined to serve a tentative reconstruction of the particles collected in the event. Such information is finally combined to check wheter the event satisfies a set of selections required by the HLT algorithms to meet the physics data-taking objectives.

The second level, High Level Trigger (HLT) [56], is an array of computers running high-level physics algorithms. The data processing of the HLT is structured around the concept of a HLT path, which is a set of steps run in a predefined order to both reconstructs physics objects and makes selections on these objects. Each HLT path is implemented as a sequence of steps of increasing complexity, reconstruction refinement, and physics sophistication.

The HLT performs a "prompt event reconstruction" on the particles passing its paths. Such a prompt reconstruction is available within ∼48 hours after data taking, can be already used for physics analysis and serves as basis for each subdetector calibration and refinement algorithms. The L1 trigger thresholds are adjusted to provide an output of 100 kHz, while the HLT delivers a 1 kHz rate to the storage stations.

CMS trigger system is optimized for full exploitation of the bandwidths available for on line data collection. CMS physics program has expanded through the years to include processes with lower and lower cross sections, which can be studied thanks to the high luminosity delivered from LHC and collected at the experiment.

Both L1 and HLT paths are equipped with trigger prescales: a prescale sets the number of events that are required to satisfy a trigger path requirement, for a single event to be actually recorded by that trigger path. Trigger paths dedicated to detection of rare events, such as the Higgs boson production during Run 1, would have a prescale of 1 (unprescaled trigger). On the other hand, the trigger rate of higher cross sections processes might be modulated, using a high prescale factor, in order to avoid for a single, dominating physics process to saturate the full trigger bandwidth.

Two main novel data taking streams have been put in place to collect and fully exploit the collision information potential: data scouting and data parking.

Data scouting was introduced in Run 1 and allows to take data that would otherwise be rejected by the normal trigger filters. It is based on event-size reduction rather than event filtering and it is useful, for instance, to search for low mass resonances. Multiple scouting streams, reconstructing specific physics objects with basic kinematics information, are currently collected in CMS.

The second strategy is referred to as Data Parking, and aims at overcoming the main limitation in the CMS data taking, which is the computing power involved in the prompt reconstruction. The 2018 B-Parking dataset was collected through the data parking strategy, and it is exploited in the analysis presented in this work. In the next section, details on the data taking procedure and on the resulting dataset are described.

2.2.4 The B-Parking Data Stream

In 2018, CMS collected a 10^{10} $B\bar{B}$ sample of events, exploiting a novel trigger and data processing strategy. The main motivation for the collection of a B meson sample relies on the B decays rareness and therefore, enhanced sensibility to possible anomalies and new resonances. A B meson rich dataset can be used to perform Lepton Flavor Universality (LFU) tests in semileptonic decays and searches for new physics in the soft region below the B meson mass.

CMS parking strategy is rooted on the possibility of exploiting the L1 bandwidth which becomes available with decreasing luminosity in a single LHC fill. The instantaneous luminosity delivered by LHC decreases the more the beams are circulating and colliding in the accelerator, as the proton bunches become less and less populated and the beam transverse section is spread out through collisions and proton scattering. This results in a decreasing L1 trigger rate, as L1 menus are designed to output $\sim$ 100 kHz at full instantaneuous luminosity, and a smaller event size. The CMS L1 standard trigger rate in a 14 hours data taking period in Run 2 2017 is reported in Figure 2.5 left plot. The trigger rate is around $\sim$ 100 kHz in the first data taking minutes, and decreased through the hours, down to less than $\sim$30 kHz.

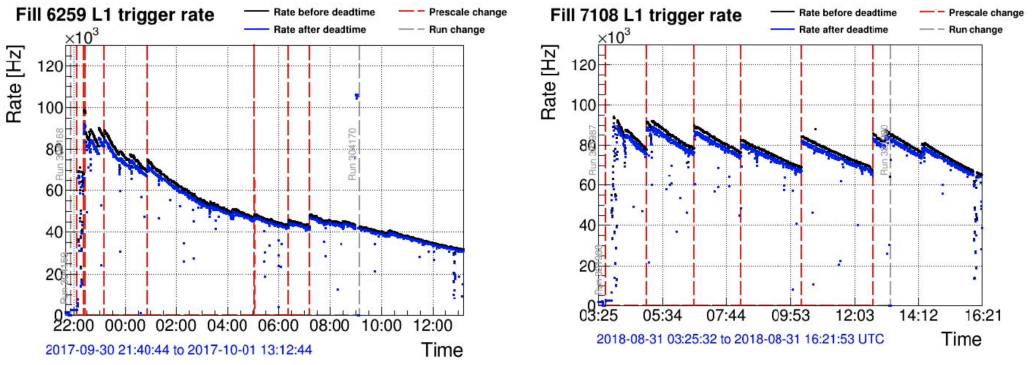


Figure 2.5. Behavior of L1 trigger rates in a $\sim$14 hours data taking period in single fill in 2017 (left) and 2018 (right). The red lines indicates the prescale changes, which in 2018 concided with the parking trigger thresholds loosening. The 2017 L1 trigger rate decreases from $\sim$ 100 kHz to $\sim$ 30 kHz, while it is maintained in a $\sim$ 100 kHz to $\sim$ 60 kHz for 2018 data taking.

A non standard physics stream can be designed to top up the L1 available bandwidth throughout the data taking. In 2018, this was realized with the implementation of a dedicated muon trigger for the selection of $B\bar{B}$ events.

The B-Parking trigger aims at the collection of an unbiased sample of B hadron decays. Its logic exploits the fact that, due to the absence of valence b quarks in the colliding protons, B mesons are produced in pairs in LHC collisions. The trigger selects events where one of the two B hadrons decays semi-leptonically, emitting a low transverse momentum muon in the CMS barrel. Since B hadrons have a non-negligible proper decay length within the CMS detector, the muon can be further characterized through the significance of its distance from the interaction vertex, which will be referred to as impact parameter (IP) significance. The B semileptonic

decay identified through this trigger is defined as B-Parking "tag-side", while no particle content nor kinematic constraint bias is posed by the trigger on the B meson not decaying to the triggering muon. The unbiased B is referred to as B-parking "probe-side".

A set of triggers with different transverse momentum and IP significance thresholds has been used for the 2018 B-Parking data collection. The transverse momentum thresholds were decreased from $\sim$ 12 GeV to $\sim$ 7 GeV throughout data taking to allow for full exploitation of the L1 bandwidth freed by the decreasing instantaneous luminosity. Several trigger paths were activated at the same time during data taking, and control over the rate of each path was achieved by the use of prescales. The CMS L1 trigger with B-Parking physics stream for a 13 hours fill in 2018 is reported in Figure 2.5 right plot. The red lines represent the B-parking prescale and transverse momentum threshold changes to top up L1 bandwidth. The L1 trigger rate is maintained between $\sim$ 100 kHz and $\sim$ 60 kHz throughout the whole data-taking period.

HLT physics stream input rates are limited by the necessity of having a 1 kHz output flux. Two separate streams are put in place for standard physics and parking data and independently fed to HLT algorithms. The B-Parking data stream goes through HLT selection yet it is not promptly reconstructed as the standard physics stream. Parking selected data are "parked" to tape, and the reconstruction algorithm is run on the parked dataset when enough computing resources are available, usually after the end of the yearly data-taking period. The HLT rates for the physics and data parking streams are reported in Figure2.6 for a single fill data taking in 2018. The HLT trigger rates for the standard physics stream are

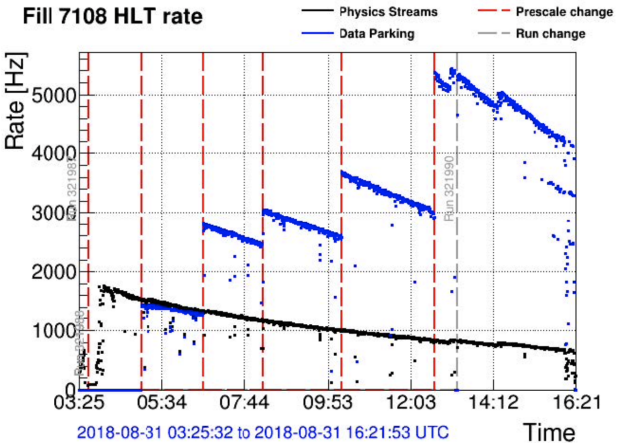


Figure 2.6. Behavior of HLT trigger rates for the standard physics stream (black) and the data parking stream (blue) in a $\sim$14 hours data taking period over a single fill in 2018

kept around 1kHz at the start of the fill and decrease in time, while the amount of HLT selected and parked data increases in steps of prescales and run changes. The collected integrated luminosity amounts to 41.6 fb^{-1}, with a B meson purity of $\sim$ 75%. The dataset purity is measured as the fraction of events triggered by

muons and coming from a B decay:

$$P = \frac{N(B \to X\mu)}{N(\mu)}. \tag{2.8}$$

It has been measured on the $B^0 \to D^{*+}\mu^-\nu_\mu$, which have a high branching fraction, where the D^{*+} is asked to follow a fully visible decay $D^{*+} \to D^0\pi^+ \to K^-\pi^+\pi^+$.

The B-Parking dataset features 12 billions B meson decays, providing CMS with new potential to explore the B sector physics measurements.

2.2.5 Particle Flow Event Reconstruction

The CMS offline reconstruction algorithm, Particle Flow [12], has the aim of reconstructing and identifying all the stable particles produced in a proton-proton collision. It runs on an event basis to match groups of tracks and deposits and identify the particles that have caused them. To do so, it mainly relies on an efficient and pure track reconstruction, clustering algorithms which are able to distinguish overlapping showers, and on a linking algorithm to build relations among different deposits.

Tracks are reconstructed starting from a seeding algorithm in the internal layers of the tracker pixel. Their trajectories are then reconstructed through topical algorithms which, given a seed and its preliminary transverse momentum measurement, search for a hit in specific regions of the outer layers.

The tracks are then extrapolated to the calorimeters: if they fall near a cluster, they are associated with it. The HCAL cluster+track candidates are labeled as charged hadrons; the ECAL cluster+track candidates should be electrons, but bremsstrahlung emission complicates these associations which therefore undergo a dedicated reconstruction. Such reconstruction aims at recovery of the full electron energy by searching for photon deposits in ECAL (clusters without an associated track) tangent to the electron trajectory, and will be described in details in Section 3.

The muons are identified by simple matching between a track and muon system deposit; the remaining clusters in ECAL and HCAL, which are not linked to tracks, are reconstructed as photons and neutral hadrons respectively.

The particle flow algorithm, specifically tuned to CMS detectors parameters, has a fundamental role in particle identification and reconstruction of complex objects, such as jets, and it serves as main tool for pile up mitigation. The algorithm builds for each event the list of reconstructed objects, with their relations and kinematic parameters. This list represents a global event description, it allows the rejection of soft contributions and the identification of relations among particles.

Reconstruction and identification algorithms evolve with the ever changing needs of the CMS physics program. The collection of a dataset dedicated to B physics posed several challenges to CMS reconstruction algorithm, as B meson decay products from LHC collisions are very soft. A particular effort was dedicated to the development and optimization of novel algorithms for the reconstruction of low energy electrons. Next chapter is dedicated to the description of the motivation, development and performance of such novel reconstruction and identification approaches in CMS.

Chapter 3

Electron Reconstruction and Identification in CMS

Electrons are essential objects to CMS event reconstruction, as they are yielded in the weak and electromagnetic decays of the bound states produced within proton-proton collision. Together with the other charged leptons, they are crucial in triggering interesting collision events and can be reconstructed with high precision, resulting essential for both searches of narrow resonances and SM precision measurements. Electrons reconstruction can be challenging, as these particles are the lightest detectable objects in the detector, with $m_e = 0.511$ MeV, and they are significantly subjected to energy loss due to bremsstrahlung radiation.

In this chapter, a review of electron reconstruction and identification algorithms currently in use in CMS is described in Section 3.1. In particular, a review of the ECAL energy measurement correction and calibration techniques is reported in Section 3.1.3, while results on the ECAL linearity measurements for electrons from the $Z \to e^+e^-$ and $J/\psi \to e^+e^-$ with Run2 data are described in 3.1.4. Reconstruction and identification techniques for low-energy electrons are presented in Section 3.3, together with a novel technique for the selection of a sample of particles misreconstructed as electron candidates in data, to be used as control sample to validate identification requirements. Such developments are essential to the full exploitation of the B-Parking dataset physics potential, as the B meson decay products hold energies of the order of a few GeV, while the standard CMS objects reconstruction is optimized for deposits higher than tenths of GeV in energy.

3.1 Electron Reconstruction

Electrons in CMS produce hits in the tracker layers by ionization, are bent by the solenoidal magnetic field in the transverse plane, and produce electromagnetic showers in ECAL, where they are absorbed. Therefore, two CMS subdetectors are crucial to electron reconstruction: the tracker for transverse momentum measurements and ECAL for energy measurements.

A schematic representation of an electron propagating through the tracker and to the electromagnetic calorimeter is presented in Figure 3.1. The electron trajectory is bent by the solenoidal magnetic field, while the particle emits two photons via

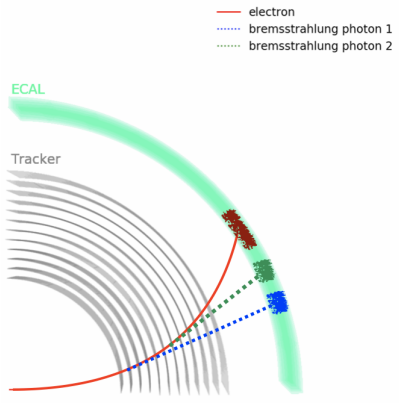


Figure 3.1. Representation of an electron, in red, propagating through the CMS tracker and in the electromagnetic calorimeter. The electron radiates two bremsstrahlung photons in tangent directions with respect to its travel path, in blue and green. The total ECAL topology for the electron is therefore represented by the electron cluster, in dark red, plus the two secondary clusters from bremsstrahlung photons.

bremsstrahlung radiation in its path, producing a complex topological pattern in ECAL. An electron loses through radiation of photons an average 33% of energy when the intervening material is minimal ($|\eta| \sim 0$), and about 86% of its energy where the intervening material is the largest ($|\eta| \sim 1.4$). To correctly measure the total electron energy, the track and all the three ECAL deposits have to be correctly measured and identified as yielded by same particle.

The trajectory of an electron in the CMS tracker and through the magnetic field provides, through the Lorentz law, a measurement of the particle's transverse momentum. The electron then reaches the electromagnetic calorimeter, where is absorbed via the development of an electromagnetic shower within ECAL high density lead-tungstate crystals, providing an energy measurement for the impinging particle.

The total electron reconstruction efficiency is reported in η bins and for different p_T ranges for simulated Drell-Yan processes samples and 2017 data in Figure 3.2 [57]. The ratio of data and simulation efficiencies are reported in the lower pad. The efficiency is higher than 95% in all the η regions for $p_T > 20$ GeV, while simulation efficiencies are compatible with the ones in data within 2%.

3.1.1 Electron Track Reconstruction

The reconstruction of electron candidates starts with the search of "seeds", topological patterns in the detector which might hint at an electron deposit. The seeding procedure is fundamental, as electron-dedicated track fitting algorithms are resource intensive and cannot be run on all combinations of hits in the CMS tracker.

The electron dedicated track reconstruction algorithm is known as Gaussian Sum Filter (GSF). It is a generalization of the standard track reconstruction filter, the Kalman Filter algorithm. Both the Kalman Filter (KF) and the GSF are path finder, recursive algorithms by which the track fit parameter are updated at each new layer measurement of the track [21]. The algorithms are meant to follow

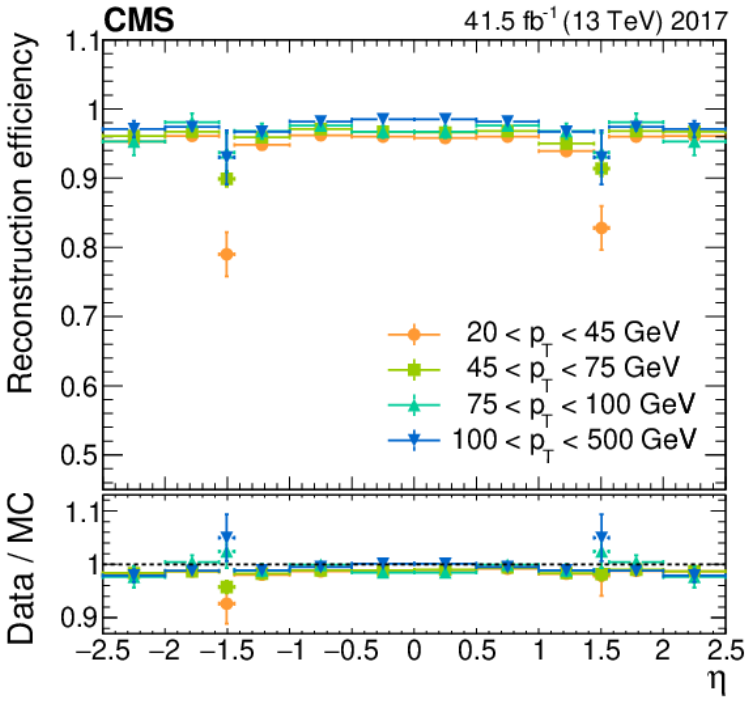


Figure 3.2. Electron reconstruction efficiency versus η in data (upper panel) and data-to-simulation efficiency ratios (lower panel) for the 2017 data taking period. The vertical bars on the markers represent the combined statistical and systematic uncertainties. The region $1.44 < |\eta| < 1.57$ corresponds to the transition between the barrel and endcap regions of ECAL and is not considered in physics analyses

particles bending in the CMS magnetic fields and loosing energy via ionization in the tracker layer, and is applied in the transverse plane only: the transverse radius ρ and the azimuthal angle ϕ are free parameters.

From the track seed, which is usually formed by doublets or triplets of hits in the Inner Tracker layers, the algorithm produces an initial estimate of the track direction and propagates it to the subsequent layer. Hits compatible with the track in the subsequent layer are searched within a window around the expected track direction. The width and shape of such window is what distinguishes the KF and GSF algorithms. The distribution of the expected hit at the i-th tracker layer for KF tracks is assumed Gaussian $G(\mu_i, \sigma_i(\sigma_{exp}))$: the σ_{exp} depends on the uncertainties on the direction parameters and decreases at each tracker layer. As the energy loss for electrons in a given amount of tracker material is expected to have a distribution described by the Bethe-Heitler formula, the distribution of the expected hits in the i-th tracker layer is not a simple Gaussian. In the GSF, it is approximated by a sum of Gaussians $\sum_j G(\mu_{i,j}, \sigma_{i,j}(\sigma_{exp}))$: the number of Gaussian components is chosen by compromising between the predictive layer to layer power and the CPU time needed to compute the layer prediction, which increases with each Gaussian added to the model.

Electron seeding is performed in two different ways: from the inside out, searching for patterns of hits in the pixel inner layers ("tracker-driven" seeding), or from the outside in, matching the ECAL super-cluster deposits to doublets or triplets of hits in the Inner Tracker ("ECAL-driven" seeding).

Tracker driven seeding has an efficiency of $\sim 50\%$ for electrons from Z decay with $p_T \sim 3$ GeV and drops to less than 5% for $p_T > 10$ GeV. The algorithm iterates over all the KF tracks. It checks compatibility with ECAL SCs by the logical OR of a cut based selection and the output of a BDT exploiting track quality and track-cluster matching variables as inputs. This seeding is computationally expensive, and is therefore run only in off-line reconstruction and not at HLT.

The ECAL-driven seeding selects SCs with $E_T > 4$ GeV and H/E_{SC} <0.15, where E_{SC} and H are the SC energy and the sum of the energy deposits in the HCAL towers within a cone of $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$ centered on the SC position. Helical trajectories for the positive and negative charge hypotheses are back propagated, from the SC position and with E_T^{SC} into the inner tracker. The relative positions of the back-propagated trajectories and of the track seed are compared in ϕ and z, and if at least two track hits are matched to the SC trajectory, under any charge hypothesis, the seed is accepted as input to GSF tracking. Note that the procedure does not account for bremsstrahlung energy losses.

3.1.2 ECAL Clustering

The localization and correct measurement of energy deposits in ECAL is a complex operation. The experimental signature of an electron in the electromagnetic calorimeter can be complex, as the radiated photons yield several satellite clusters which have to be linked to the one induced by the electron itself. Shower shapes in ECAL crystals vary depending on the particle incident angle and total energy. Moreover, as multiple signals reach the calorimeter, electromagnetic showers developing in neighboring crystals might yield overlapping signals, which have to be detangled from one another to correctly extract energy measurements. A brief review of the main algorithms for ECAL super-clustering is presented below, while extensive information on clustering procedures can be found in [57]

Two main algorithms are currently used in CMS for crystal super clustering. The "mustache" algorithm is mostly used to measure low energy deposits. The algorithm adds to the ECAL seed cluster all the cluster that fall in a "mustache" region around the central seed in the transverse plane $\Delta\eta = \eta_{seed-cluster} - \eta_{cluster}$ versus $\Delta\eta = \eta_{seed-cluster} - \eta_{cluster}$. The "mustache" region around a seed cluster is shown in Figure 3.3. It configures in this characteristic shape as the CMS solenoidal magnetic field tends to spread radiated energy along the ϕ rather then in η. The dimension of the mustache region depends on the total E_T, as high energy particles are less bent by the magnetic field.

The second clustering algorithm is known as the “refine” algorithm, and makes use of all the information coming from the mustache algorithm, the photon conversions and track to cluster compatibility to refine the super cluster shape. It combines the output of a BDT reconstructing single and double legged photon conversions and of an algorithm searching for bremsstralhung clusters tangent to the tracks with the information coming from the GSF tracks and the ECAL SCs. It

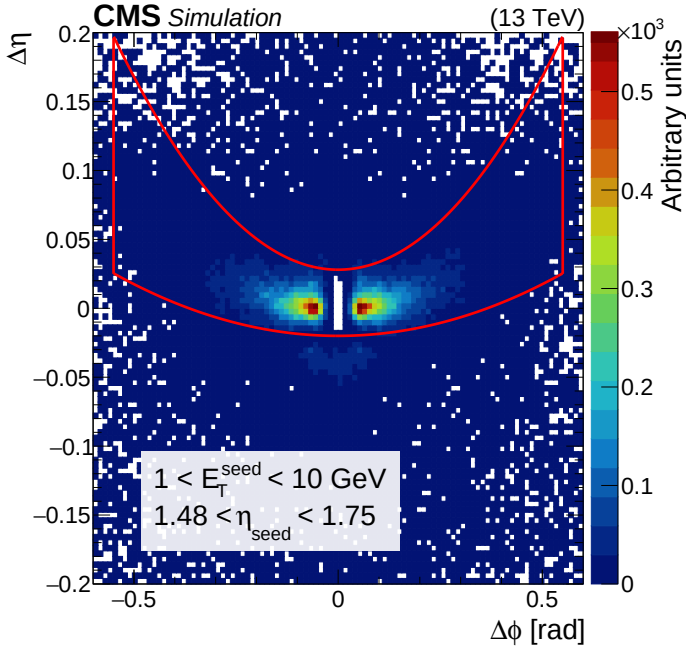


Figure 3.3. Distribution of $\Delta\eta = \eta_{seed-cluster} - \eta_{cluster}$ versus $\Delta\eta = \eta_{seed-cluster} - \eta_{cluster}$ for simulated electrons with $1 < E_T^{seed} < 10$ GeV and $1.48 < \eta_{seed} < 1.75$. The z axis represents the occupancy of the number of PF clusters matched with the simulation (requiring to share at least 1% of the simulated electron energy) around the seed. The red line contains approximately the set of clusters selected by the mustache algorithm. The white region at the centre of the plot represents the $\eta - \phi$ footprint of the seed cluster.

removes and adds satellite clusters to the energy estimate in ECAL depending on the energy measurement compatibility with the transverse momentum measurement coming from the tracker.

3.1.3 Energy Reconstruction and Calibration

The energy deposited by an electromagnetic shower in ECAL is reconstructed according to (3.1)

$$E_{e,\gamma} = \sum_i [A_i(t) \cdot L_i(t) \cdot C_i(t)] \cdot G(\eta) \cdot F_{e,\gamma} + E_{preshower} \tag{3.1}$$

The sum runs over a set of i crystals involved in a SC and $A_i(t)$ is the signal amplitude, extracted by the Avalanche Photodiodes ADC. The number of extracted counts as a function of the released energy is proportional to the High Voltage (HV) of the APDs. High voltage stability measurements for ECAL APDs represent a fundamental step, as variations in the HV impact the energy resolution of the electromagnetic calorimeter. HV stability is required to be within 0.01% and needs to be continuously checked before and within data taking. A description of the calibration procedure performed for the Run 3 is reported in Appendix **??**.

The amplitude measurement is corrected to account for multiple environmental effects. Crystal-to-crystal corrections, $L_i(t)$ and $C_i(t)$ in Equation 3.1, are applied to

respectively correct for radiation driven loss of crystal transparency and to ensure a homogeneous energy measurement output through the detector. The corrected per crystal amplitudes are then summed in an electromagnetic deposit energy estimate. The $G(\eta)$ term is used to convert ADC counts in energy estimates, and is derived by calibration of ECAL energy measurements over the π_0 and Z Standard Model candles. Finally, refinement corrections depending on the reconstructed type of particle, $F_{e,\gamma}$ are applied to account for imperfect clustering procedure, geometric and material effects. The preshower energy is finally added to the estimate.

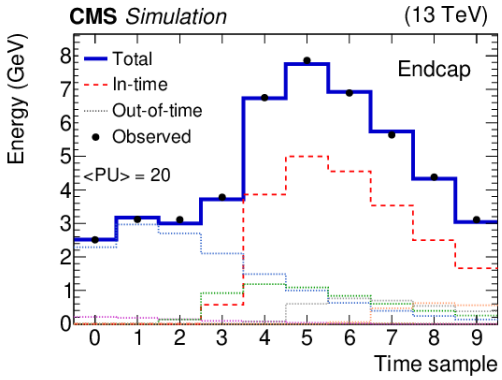
In the following subsections, a description of the strategies put in place for each calibration step is presented.

Signal amplitude reconstruction

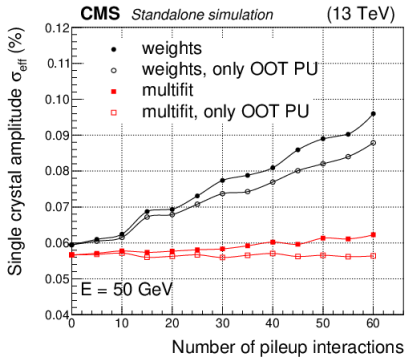
ECAL single crystal energy measurement is performed by extracting the APD signal pulse amplitude. The signal shape and pedestal information are fundamental for the amplitude extraction procedure: they are continuously monitored and corrected to ensure proper modeling.

Crystals light transmission decreases with radiation damage, leading to changes in the shape of the collected pulse. During Run 2, pulse shapes for signal extractions were recomputed every 3-4 fb^{-1} to accommodate for shape evolution with increasing radiation.

Amplitude pedestals, on the other hand, depend mostly on APDs behavior: a ∼40 MeV/year increase in pedestal absolute value has been observed in average throughout ECAL operational years. Monitoring of the pedestals value and spread has been performed in Run 2 by collecting pedestal runs every 40 minutes, using the same laser system presented in Figure 3.1.3 for radiation damage monitoring.



(a) Multifit fitting procedure example. An in time pulse in red is fitted together with a late pulse (grey) and several early pulses



(b) Comparison of the multifit (red) and weights (black) algorithms single crystal amplitude σ_{eff} as a function of the number of overlapping interactions (pile up).

Figure 3.4

The ECAL amplitude reconstruction changed from Run 1 to Run 2, adapting to

the higher pile up conditions of LHC collisions. The Multifit [59] algorithm specifically treats overlapping signal amplitudes in the same crystal by fitting the pulse shape with a single in time pulse and up to 9 out of time pulses. An example of the fitted signal pulse and its components is presented in Figure 3.4a. The time offset of the 9 out of time pulses are left floating, together with the signal amplitudes, while the pulse shape is fixed.
The performance in single crystal effective signal width, σ_{eff}, of the Multifit algorithm with respect to Run 1 weights algorithm for APD signal extraction is presented in Figure 3.4b, as function of the number of simultaneous interactions. The multifit performance results stable with increasing pile up, whereas the weights algorithm shows worsening performances in σ_{eff} with increasing simultaneous interactions.

Crystal Ageing Corrections

LHC proton-proton collisions struck the detector with extreme radiation rates. Crystal light transmission and, consequently, light output, are heavily influenced by the radiation damage. Transparency loss impacts pulse shapes on very short times scales and it is heavier at high η, where crystals are subjected to higher radiation rates. A moderate transparency recovery can be obtained in lead tungstate crystals by annealing at room temperature.

ECAL is equipped with a dedicated laser + LED monitoring system [8] which is used to asses the transparency loss within each crystal. A crystal-by-crystal laser scan is performed every 40 minutes during data-taking, to derive time dependent crystal transparency corrections. The relation between crystal response to laser light and to an electromagnetic shower signal can be modeled as:

$$\frac{S}{S_0} = \left(\frac{L}{L_0}\right)^{\alpha} \tag{3.2}$$

where L_0 and S_0 are reference responses to laser light and electromagnetic shower measured at the beginning of each data taking and the exponent α is measured to be $\sim$1.5 in the barrel and between 0.6 and 1.1 in the endcaps. The normalized π_0 mass of two photon deposits in ECAL is reported in Figure 3.5, with (green) and without (red) laser corrections, as a function of time. The impact of laser corrections is impressive, as it allows for full recovery of the correct measurements for the photon energies and, consequently, π_0 mass. The uncorrected π_0 mass absolute value and resolution worsens over a few hours time range, thus deeming a fine-grained in time correction necessary for optimal energy measurement recovery.

Residual response losses were observed on yearly time scales through Run 2, due to radiation damage on laser reference diodes and transmission fibers. The effect is corrected exploiting the E/p ratio of the electromagnetic energy measurement over the particle transverse momentum as measured in CMS tracker, for electrons coming from W and Z bosons decays.

Crystal Inter-Calibrations

Inter-calibration coefficients are used to correct the single crystal output to ensure a homogeneous response behavior over the detector. CMS currently exploits three

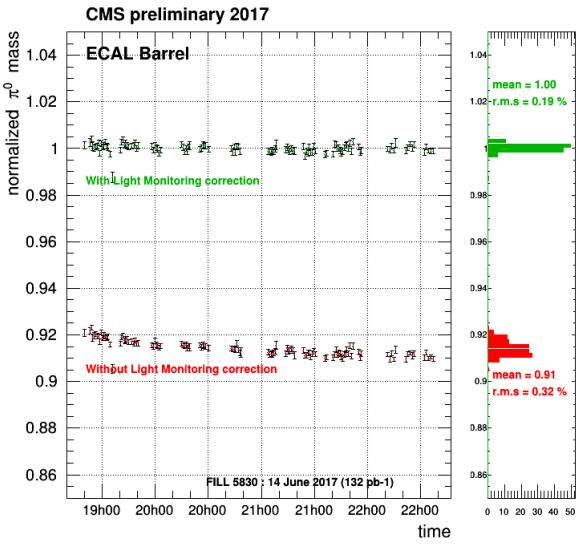


Figure 3.5. Normalized π_0 mass behavior over a few hours data-taking period with (green) and without (red) laser corrections applied.

independent methods for inter-calibration coefficient derivation: $\pi_0 \to \gamma\gamma$ and $Z \to e^+e^-$ based calibrations and E/p ratios with p transverse momentum as measured from the tracker . The inter-calibration outputs from those techniques are then combined in a single, per crystal, inter-calibration correction.

$\pi_0 \to \gamma\gamma$ *inter-calibrations:* a π_0 invariant mass spectrum is built considering all the $\gamma\gamma$ candidates for which one of the photons has deposited a fraction of its energy in a crystal. The mass shift between the measured peak position and the PDG [36] mass value for π_0 is used to derive an energy correction to be applied to the crystal. The calibration algorithm proceeds iteratively correcting each crystal output by the observed mass peak shift, equalizing the channel response in rings of fixed azimuth.

E/p calibrations exploit the independent transverse momentum measurement for electrons in the CMS tracker. A set of high energy electrons from $W^{\pm}$ and Z decays is selected through a combination of kinematic cuts, identification and isolation criteria. The calibration algorithm iteratively assigns a coefficient to each crystal, such that the average E/p ratio for such high purity electron dataset is constrained to 1.

$Z \to e^+e^-$ *inter-calibrations*: the method has been introduced in Run 2 thanks to the high per year integrated available luminosity, which allowed to performed per crystal calibrations on an high energy, smaller cross section resonance. The crystal by crystal calibration consists in the maximization of a likelihood comparing the reconstructed mass distribution with that predicted by Monte Carlo simulation. The three calibration methods are combined and their performances is assessed in term of the calibration method precision. The precision of each method and of their combination is reported in Figure 3.6. The overall inter-calibration precision

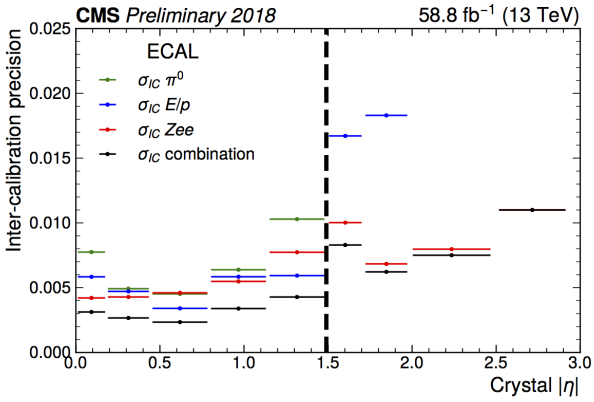


Figure 3.6. Intercalibration precision as a function of pseudo-rapidity with the 2018 dataset. The red, blue and green points represent performances for three different methods performance, while their combination is reported in the black points .

is smaller than 1% and decreases to less than 0.5% in the barrel region. The novel $Z \to e^+e^-$ calibration method drives the inter-calibration precision in the high η region. The $2.5 < |\eta| < 3.0$ region in particular is out of tracker coverage, hence the ability to calibrate ECAL energy response becomes crucial for very forward jets energy reconstruction.

Energy Regression

A semi-parametric Boosted Decision Tree (BDT) regression is put in place to recover possible performance losses due to electromagnetic shower leakages, dead channels and material and geometric effects. The regression algorithm exploits variables related to the electromagnetic shower shape and dimension, the super-cluster isolation and, for electrons, bremsstrahlung energy loss and the angular coordinates of the track matched to the ECAL SC. A detailed description of ECAL regression algorithm is reported in [**egm**].

Scale and Smearing Corrections

Residual differences between data and simulation are found in both energy scale and resolution. They are corrected by comparing the $Z \to e^+e^-$ resolution and peak position in data and simulation.

The peak position of the data distribution is scaled to the one expected in simulation, while the simulation invariant mass resolution is corrected to match the width of the peak in data. To scale the peak position, the data and simulation distributions are fitted with convolution of a Breit-Wigner distribution and a one side crystal ball distribution, where the Breit-Wigner width is fixed to the Z resonance width, $\Gamma = 2.495$ GeV. The peak position in data is then shifted to the one in the simulation. The distribution resolution is corrected in simulation to account for effects which are not well modeled or non predictable. The correction is in this instance derived using the simulated Z boson invariant mass distribution as a probability density function in a maximum likelihood fit to the data.

Scale and smearing corrections are derived separately for low and high bremsstrahlung radiation level electrons, in the barrel and endcaps. They are computed comparing the distributions obtained with ECAL SC information only.

Scale and smearing corrections are derived on particular categories of electrons, coming from the Z boson decay. The validity of such corrections has to be checked against all the electrons categories and on different resonance. In the next subsection momentum linearity studies performed for Run 2 data are detailed.

3.1.4 Energy Linearity Studies for Run 2 Data

ECAL is a homogeneous crystal calorimeter, therefore the energy response is expected to be linear across the energy range of particles producing electromagnetic showers fully contained in the crystals. Studies on the energy linearity are therefore performed on electrons of different energy with respect to the ones used to derived the scale and smearing corrections, in order to verify their universality. They also serve as check of the overall set of corrections applied on the ECAL energy estimate.

Linearity studies have been performed in Run 1 [48] on electrons with transverse momentum greater than 7 GeV, exploiting electron candidates where both the energy and transverse momentum were determined through the combination of the tracker and ECAL measurements. At very low energies such combination is dominated by the tracker reconstruction, while is driven by ECAL SC energy measurements for electrons with $p_T > 20$ GeV.

The study presented in this subsection is performed on Run 2 data, is carried out on electrons coming from J/ψ and Z decays and exploits both the standard and B-Parking dataset. It is the first study testing scale and smearing corrections on soft electron, reaching a lower limit of 2 GeV in transverse momentum and exploiting ECAL information only.

The procedure consists in measuring the discrepancies between the peak position of the considered resonance in data and simulation, in several p_T ranges and for different η regions, after the scale and smearing corrections are applied. The measured discrepancy can be, depending on its value, corrected when building the electron candidates or assigned as a systematic uncertainty to the scale and smearing corrections.

The RunD 2018 integrated luminosity is used for the Z resonance computation in data, with 32 fb^{-1} of integrated luminosity, while the full 41.6 fb^{-1} B-Parking dataset is used for the J/ψ study. The data distribution are respectively compared in shape and peak position to resonances computed using simulated data for $Z \to e^+e^-$ Drell-Yan and $B \to KtoJ/\psi(\to e^+e^-)$ processes.

The scale and smearing corrections are automatically applied to standard datasets, such as the ones for the Z resonance. The corrections had to be applied by hand on B-Parking data, as the scale and smearing corrections have never been tested for low energy electrons: B-Parking electrons have spectra which peak around 1 - 3 GeV, and mean laying around 8.5 GeV. The corrections are applied in bins of run number, η and dividing the electrons in categories at low and high bremsstralhung emission.

The Z boson mass is reconstructed from the SC energies and the opening angles measured from the tracks, asking for both the electrons p_T to fall within the same

transverse momentum bin. The request does decrease the per p_T bin statistics and constraints the Z decay kinematics, but prevents the linearity measurement in an p_T bin from being contaminated by the linearity behavior in other p_T ranges. To ensure better peak definition over the background, the J/ψ mass is computed from one PF electron and one electron reconstructed from its SC energy and the opening angles measured from its track, binning in p_T of the latter.

The signal is fit with a convolution of a Breit-Wigner function and a Crystal Ball function for the Z resonance, and a double Crystal Ball for the J/ψ resonance. The background below the J/ψ is fitted with a 3rd grade polynomial function. Examples of the $J\Psi$ and Z fits are reported in Figure 3.7 left and right plot respectively.

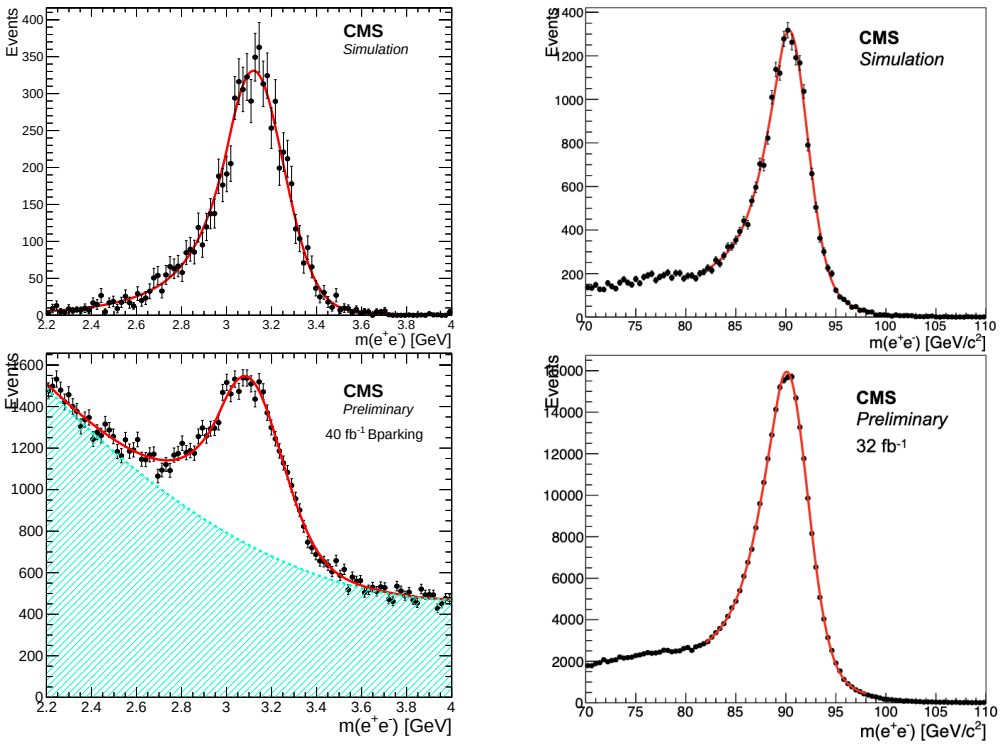


Figure 3.7. Examples of fit performed on simulation and data for extraction of peak position and computation of ECAL linearity. The left plot shows the J/ψ resonance in simulated samples on top and in data in the bottom, in the $2 < p_T < 5$ GeV bin, for $|\eta < 1|$ the signal shape is fitted with a double Crystal Ball, the background is fitted with a 3rd grade polynomial. On the right, the Z resonance as it appears in the $30 < p_T < 40$ GeV bin and for $|\eta < 1|$ in simulation (top) and data (bottom) is fitted with a convolution of a Breit-Wigner function and a Crystal Ball function.

Systematic uncertainties contributions from the fit function choice are computed in the plot. They are obtained by computing the fit parameter variation when using an exponential in the J/ψ background fit and a gaussian function in place of the nominal fit CB for the Z.

In order to perform a consistent comparison, the relative momentum scale is defined as

$$Lin_{ECAL} = \begin{cases} \frac{m^2_{data}}{m^2_{sim}} - 1 & for\ J/\psi \\ \frac{m_{Data}}{m_{sim}} - 1 & for\ Z. \end{cases} \tag{3.3}$$

The different definition ensures that the SC energy contributes the same way on the two resonances, the track momentum scale contribution being negligible. The residual scale difference between data and simulation is at most 0.4% in the barrel and 0.6% in the endcaps for the Z resonance, while it ranges from -2% to 0.5% for the J/ψ resonance. The J/ψ resonance result demonstrate that even though scale and smearing corrections computed with high energy electrons do not spoil the peak position with respect to the MC simulation, the corrections computed for the Z resonance might not be suited to account for all the possible sources of mis-modeling at low energy. Studies are currently ongoing in order to verify such hypothesis and possibly compute and test scale and smearing corrections dedicated to soft electrons. The performance of the momentum scale after scale and smearing corrections is shown in Figure 3.8, as the relative difference between data and simulation of the $J/\psi \to e^+e^-$ and $Z \to e^+e^-$ mass peaks, in several p_T and $|\eta|$ categories.

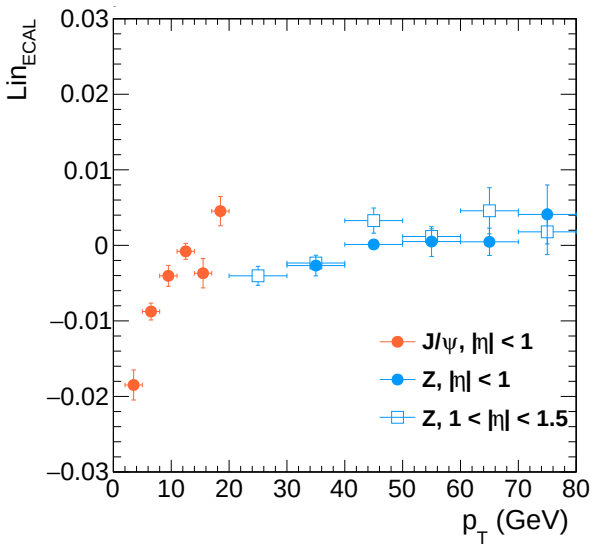


Figure 3.8. ECAL linearity as a function of the transverse momentum and in bins of η. In orange, points obtained for the J/ψ with the B-Parking dataset, in light blue the points obtained for the Z resonance with the standard CMS datasets.

An energy linearity better than 0.5% is observed for barrel electrons originating from Z boson decays in the p_T $[20, 80]$ GeV energy range. The energy linearity for J/ψ low energy electrons is better than 1% for electrons with $p_T > 5$ GeV, while it lies in the 2% range for electrons with $p_T < 5$ GeV. The linearity measurements for low energy electrons are slightly worse than the ones on high energy electrons: the residual discrepancy in linearity can be corrected at physics object.

ECAL energy measurements performance

The simulated contributions of the reconstruction and calibration steps to ECAL energy resolution are reported in Figure 4.6, for 2018 data taking conditions. The inter-calibration impact is negligible, while noise and pile up yield significant and comparable contributions to the total resolution. Unaccounted effects are also significant: they can be described through a gaussian smearing applied to simulation to match the performance in data and result stable over time.

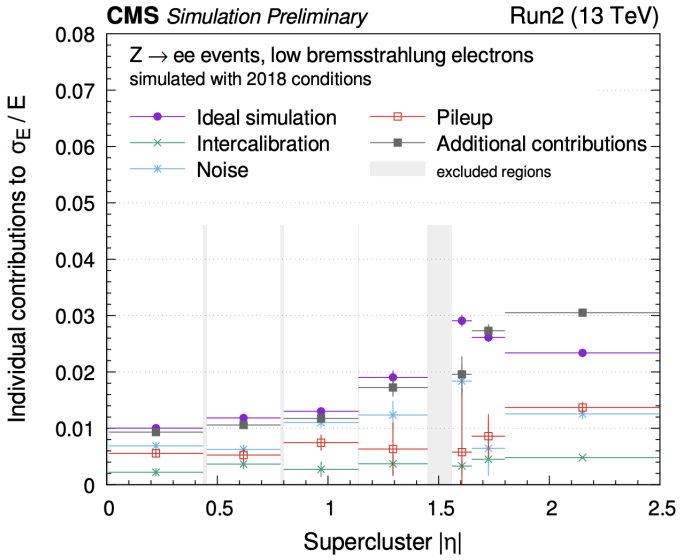


Figure 3.9. Simulated resolution contributions breakdown as a function of pseudorapidity.

3.2 Electron Identification

Electron identification strategies are essential to ensure high purity in the selection of electron candidates. Control over background sources is crucial, especially at low energies, where combinatorial association of ECAL noise clusters to spurious tracks can lead to increased, irreducible background levels.

Electron identification criteria in CMS are optimized for candidates with $p_T > 20$ GeV and are split in cut-based and multivariate techniques. Cut-based techniques consist in sets of sequential cuts on several identification variables. These include supercluster-to-track matching variables, quantities controlling the ratio of energy deposited in the hadronic and electromagnetic calorimeter and isolation variables.

A particle isolation I is computed by opening a cone of $\Delta R = 0.3$ or $\Delta R = 0.4$, around the particle and summing the energies of particles falling in the cone: the ratio I/E_T, with E_T energy of the particle, quantifies of how much the particle is isolated in the detector. Isolation variables are a powerful tool for electron identification, as they discriminate electrons from electroweak vertexes, which usually involve a small number of well-isolated particles, from electrons originating from in-flight meson decays, or pions misidentified as electrons. These are typical components of hadronic jets, and therefore have higher isolation values. More details

on identification variables can be found in [57].

To increase the identification power, especially for electrons below 40 GeV, identification variables are combined in multivariate discriminators (CMS MVA ID) based on Boosted Decision Trees (BDT) [19] The BDT input variables set is equipped with the observables used in the cut-based selection, plus cluster-shape and track quality indicators. The fractional difference between the track momentum at the innermost tracker layer and at the outermost tracker layer, f_{brem}, which quantifies the electron's bremsstrahlung energy loss, is also included.

3.3 Low-Energy Reconstruction and Identification

The electrons in the signature targeted in this work carry low energies with respect to the ones for which CMS reconstruction and identification algorithms have been optimized. The distributions of the generator-level transverse momenta for $B \to K\ell\ell$ decay products are shown in Figure 3.10 left plot: the transverse momentum for the three particles is very soft, with most probable value lying in the $1-3$ GeV region. In Figure 3.10 right plot, the reconstruction efficiency for PF electrons as a function of the p_T of the candidate is reported in blue. The algorithm for PF electrons reaches efficiencies higher than 80% for electrons with $p_T > 10$ GeV, but its efficiency for $2 < p_T < 5$ GeV lies around 40-50% in average, while candidates below 2 GeV in transverse momentum are not reconstructed.

Electron identification at low energy is a challenging task: the electron candidate is reconstructed through the association of a track to an ECAL deposit, therefore the main source of identification background, at all energies is the combinatorial association of two unrelated ECAL and Tracker deposits. The number of reconstructed deposits, and therefore, the combinatorial background, grows with lower energies, so much so that no identification criteria is available in standard PF reconstruction for electrons with $p_T < 5$ GeV. Under standard PF electrons conditions, most of B-Parking electrons would not be equipped with an identification criterium, which is crucial as such low energies to reject the background.

To fully exploit the CMS potential for B-Parking searches, two main developments on electrons reconstruction have been put in place: a dedicated reconstruction algorithm for very low energy electrons (*lowPt electrons*) and a retuning of the PF electrons multivariate identification algorithm on low energy candidates.

The *lowPt electron* algorithm extends CMS reconstruction capabilities to electron candidates with transverse momentum $p_T > 0.5$ GeV. It has been developed as a standalone algorithm with respect to PF: the reconstruction exploits the collection of the GSF tracks introduced in Section 3.1.1, with looser seeding requirements, and does not combine the p_T track measurement with the energy information of the matching deposit in the electromagnetic calorimeter. As reported in Figure 3.10, the reconstruction efficiency for *lowPt electron* candidates in the $0.5 < p_T < 2$ GeV range is 50% in average, while it rises over 70% for electrons with $p_T > 2$ GeV.

The collection has been tested in the search presented in this thesis, as described in Appendix A.1. The acceptance increase yielded by the inclusion of electrons with $0.5 < p_T < 2$ GeV and by the higher electron reconstruction efficiency results in approximately doubled signal yields compared to the ones computed using the PF

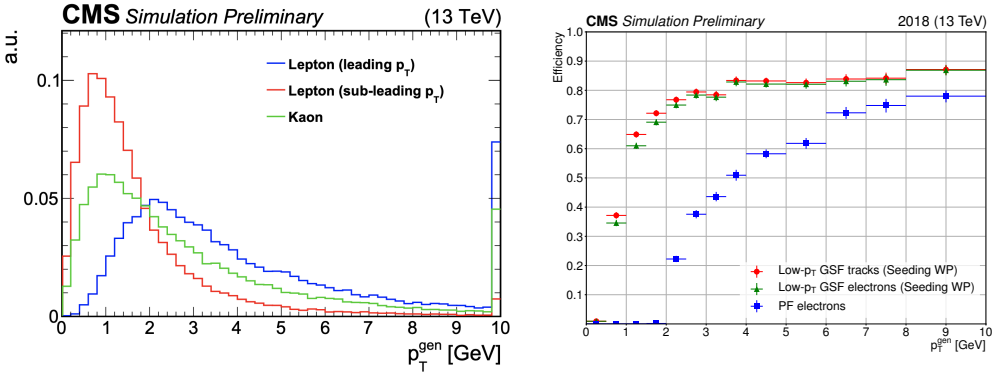


Figure 3.10. On the left, momentum spectra of the decay products in a $B \to K\ell\ell$ process. The spectra are very soft, with maxima in the 1-3 GeV range. The right plot shows the efficiency of different types of electron reconstruction as a function of the electron transverse momentum p_T: the PF reconstruction algorithm is presented in blue, the GSF algorithm for electron tracks reconstruction is reported in red, while the low-p$_T$ reconstruction efficiency is indicated in green.

electrons for all the signal hypotheses. Yet, the background yields for $\mu e\pi$ candidates including *lowPt electrons* are a factor at least four higher than the ones resulting from the usage of the PF electrons only. The increase in the background yields is due the looser requirements on track seeding and lowered thresholds for ECAL supercluster to track matching exploited in the *lowPt* algorithm.

A dedicated event selection has been put in place to control such increased background yields, yet the sensitivity for signal hypotheses of mass 3 GeV and different lifetimes in $\mu e\pi$ candidates exploiting *lowPt electrons* results up to a factor 5 worse than the one for $\mu e\pi$ candidates built with PF electrons. No sensitivity increase would be brought to the analysis by the *lowPt electrons*, which are therefore not included.

A retraining of the standard PF electrons MVA identification criterium has been developed to provide identification tools for electrons with $p_T < 5$ GeV and to boost identification performance at low momenta. The retraining is based on simulated $B \to KJ/\psi(e^+e^-)$ sample. Electron candidates correctly associated to a generated electron from the J/ψ decay are labeled as signal, while all the reconstructed electron candidates are labeled as background ("fak" electrons).

The training is performed using the same set of features as the PF electrons Standard MVA ID, in two separate p_T bins: $p_T \in [2, 5]$ GeV and $p_T \in [5, \infty)$ GeV. It is therefore the first identification tool in CMS for PF electrons in the $2 < p_T < 5$ GeV range. A novel feature with respect to the standard CMS MVA ID lies in the fact that, due to the B meson non-negligible lifetime, signal electrons for the retrained PF MVA ID are slightly displaced with respect to the interaction region. The algorithm is therefore expected to have better performance for displaced electrons with respect to the standard CMS MVA ID.

In Figure 3.11 the ROC curves for the Standard and retrained PF MVA ID computed for the different training bins and three η regions are reported. The top panel of Figure 3.11 shows the comparison of the ROC curves for the standard CMS

MVA ID and several versions of the retrained PF MVA ID in the [2, 5] GeV bin: at fixed background rates, a factor 3 or higher gain can be observed in the retrained PF MVA ID efficiencies with respect to the standard CMS MVA ID.

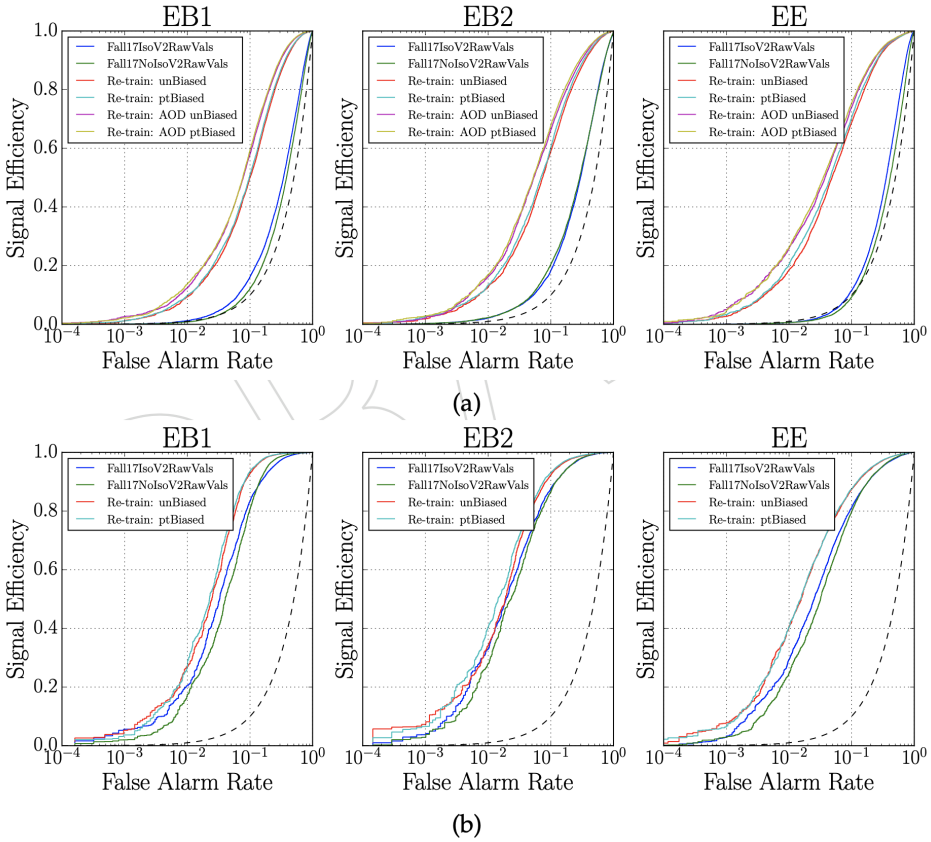


Figure 3.11. Comparison of the ROC curves for different versions of the standard CMS MVA ID (labeled Fall17Iso) and PF MVA ID (labeled Re-train) for the low transverse momentum [2,5] GeV training bin in top plots and for the p_T [5,∞) bin in bottom plots. The comparison is reported in three η regions: from left to right $0 < |\eta| < 1$(EB1), $1 < |\eta| < 1.5$(EB2) and $|\eta| > 1.5$(EE).

This PF MVA ID retraining is based on signal and background electrons extracted from a simulated sample, yet it is a tool for background rejection in data. The identification performance in simulation and data has to be tested to check whether all the signal and background sources present in data are correctly represented by simulated events. Data to simulation agreement is very well established for standard known physics processes, such as the Drell-Yan $Z \to e^+e^-$ decays. This might not be the case when considering very soft objects: in particular, the amount and sources of background might not be correctly accounted for, as CMS simulations are optimized for high-energy processes.

The identification performance on signal electrons can be tested by comparing the PF MVA ID behavior on a known standard model resonance: an example of such procedure is reported in Section 6.4 for the ID requirement chosen for the heavy neutral lepton search presented in this work.

To test whether the background simulated events are representative of all the possible sources of electron misidentification, a control sample is defined in data, enriched in fake electron candidates. The sample is obtained in the B-Parking dataset requiring signatures with 4 muons in the final state. Such signature is considered to be a proxy for events containing two $B \to X\mu^+\mu^-$ decays, which should not contain electrons: to first approximation, every reconstructed electron candidate in these events is a "fake".

Events containing 4 muons are selected and the muons combined in di-muon pairs. The combination of the four muons in two pairs is chosen to be the one for which the two di-muon invariant masses are minimized, within the combinations which have opposite charged muons. The invariant mass of the pairs is further required to be below the mass of the B meson $m_B = 5.27$ GeV. Finally, the di-muon reconstructed vertexes are required to fall within a 0.1 cm distance: this is needed to reject events in which he muons originate from secondary semileptonic decays of the B meson products such as $B \to K(\to \mu + X)$.

A sample of more than 50 000 "fake" electron candidates is obtained by applying this strategy on 10 fb^{-1} of B-Parking data.

A comparison of the p_T, $|\eta|$ and PF MVA ID distributions for the "fake" electrons sample and the signal and background PF MVA ID training samples is reported in Figure 3.13. The PF MVA ID distribution is compatible with the distribution of simulated background events, and this already validates the background training sample. The behavior is also well compatible in η while it has some small deviations in the high p_T, over 10 GeV. This does not represent a concern, as electron p_T distributions for the signals in this search typically lie below this threshold.

A possible source of concern is that the selection applied on the the muon pairs is too loose and therefore contamination from other processes might question the validity of the sample. Therefore a tighter selection is put in place. The two muon pairs selected in the four muons events are further required to be included in reconstructed $B \to K\mu^+\mu^-$ candidates. Such candidates are built by fitting to a common vertex the two muon candidates and combining the built pair with a charged track, which is labeled as kaon (K). If multiple candidates are built with the same muon pair, the one with the best fit to vertex is selected.

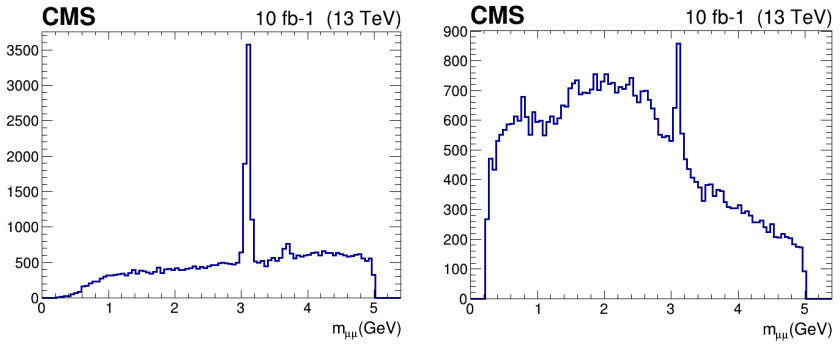


Figure 3.12. Distributions of the invariant mass of the "tag" (left) and "probe" (right) di-muon pairs selected in events holding 4 muons combined in 2 $B \to K\mu^+\mu^-$ candidates.

To ensure the absence of true electron candidates, the selection can be further

tightened, by requiring events in which the di-muon pairs resonate on the J/ψ peak. The two $B \to K\mu^+\mu^-$ candidates can be distinguished by labeling as "tag" $B \to K\mu^+\mu^-$ the one containing the triggering muon for the event, and as "probe" the other one.

The invariant mass distributions for the di muon pairs in the "tag" and "probe" $B \to K\mu^+\mu^-$ candidates is reported in Figure 3.12: the J/ψ resonance is clearly visibile.

Three subsamples can be defined:

- Subsample A : $3 < m_{\mu\mu}^{tag} < 3.15$ GeV
- Subsample B : $3 < m_{\mu\mu}^{probe} < 3.15$ GeV
- Subsample C : $3 < m_{\mu\mu}^{tag} < 3.15$ GeV and $3 < m_{\mu\mu}^{probe} < 3.15$ GeV.

The behavior of the fakes sample in terms of the PF MVA ID is reported in Figure 3.14 for Subsample A (top left), B (top right) and C (bottom).

As expected the sample size is significantly reduced with respect to Figure 3.13, especially for subsample C. Yet, the "fake" electron candidate behavior is consistent through all the subsamples and agrees with the simulated background.

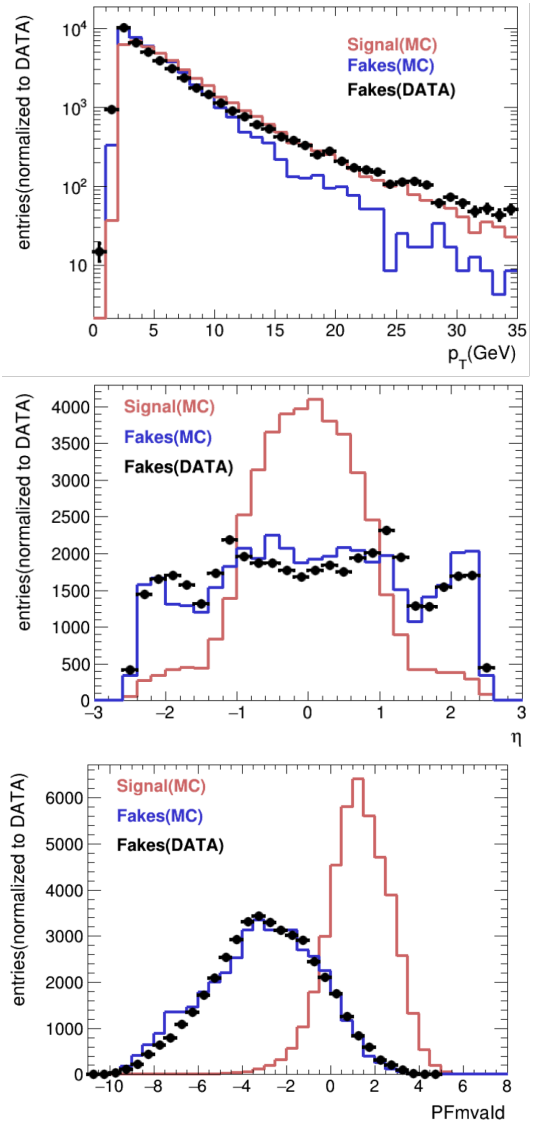


Figure 3.13. Comparison of the p_T,η and soft electrons BDT score distributions for the fake electron candidates sample (black), the background training sample of electron candidates not matched to a generated electron (blue) and the signal training sample (red).The histograms are normalized to the data yields.

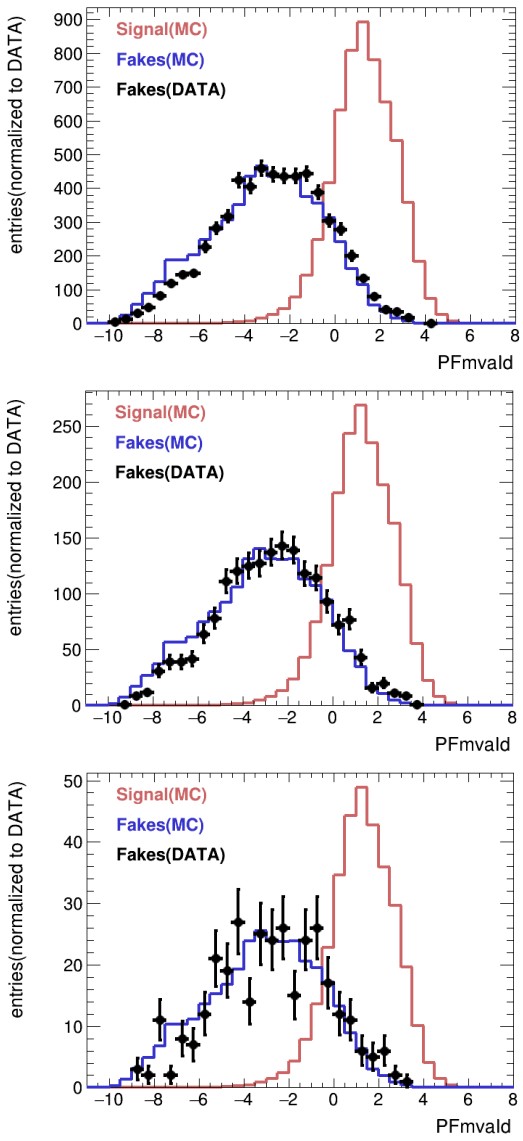


Figure 3.14. Comparison of the soft electrons BDT score distributions for the fake electron candidates sample (black), the background training sample of electron cadidates not matched to a generated electron (blue) and the signal training sample(red). The top left plot is obtained for Subsample A , the top right plot for Subsample B, while the bottom plot represents Subsample C. The histograms entries are normalized to the data yields for each subsample.

Chapter 4

Event Selection

The analysis strategy for a search for Heavy Neutral Leptons in B meson decays is detailed in this chapter. The analysis makes use of the 2018 B-Parking dataset introduced in Section 2.2.4, which has an integrated luminosity of 41.6 fb^{-1}. A detailed description of the B-Parking dataset and of the simulated signal samples is presented in Section 4.1. The signal candidate reconstruction, categorization and selection are detailed in Section 4.2 and Section 4.3. Finally, a summary of the signal efficiencies and background rejection rates is reported in Section 4.4.

4.1 Data Samples

4.1.1 Data

The results presented in this work make use of the full 2018 B-Parking dataset of 41.6 fb^{-1}. Data were collected and processed through the strategies presented in Section 2.2.4. Several trigger paths were employed, with different requirements on the muon transverse momentum p_T^{μ} and its impact parameter (IP) significance, defined as the ratio between the measured impact parameter and its uncertainty $IP/\sigma(IP)$. Different paths were activated and deactivated through each data taking in order to meet the trigger rate requirements. The list of the available B-Parking trigger paths is reported in Table 4.1.

The B-Parking dataset is divided into four data-taking "periods", labeled as A, B, C and D. Each period is further divided in parts of equal integrated luminosity: 6 parts for period A and B, and 5 for period C and D. This is summarized in Table 4.2 together with the integrated luminosities recorded for each period and part, and their sum.

Each part of a given path has identical prescale as the other parts, and the sum of all parts have prescale equal to either 0 or 1.

4.1.2 Simulated Samples

Background yields for this search are extracted from a direct fit to data, therefore no simulated background sample is necessary to the analysis. This section therefore describes the dedicated procedure developed for $B \to \ell N + X$ signal samples simulation.

HLT path	p_T requirement	IP significance requirement
HLT_Mu7_IP4	> 7 GeV	> 4
HLT_Mu8_IP3	> 8 GeV	> 3
HLT_Mu8_IP5	> 8 GeV	> 5
HLT_Mu8_IP6	> 8 GeV	> 6
HLT_Mu8p5_IP3p5	> 8.5 GeV	> 3.5
HLT_Mu9_IP4	> 9 GeV	> 4
HLT_Mu9_IP5	> 9 GeV	> 5
HLT_Mu9_IP6	> 9 GeV	> 6
HLT_Mu10p5_IP3p5	> 10.5 GeV	> 3.5
HLT_Mu12_IP6	> 12 GeV	> 6

Table 4.1. List of the B-parking HLT paths together with the transverse momentum and impact parameter significance requirements.

Generation procedure

Several generators are used for B physics events simulation in CMS. A combination of two generators, PYTHIA 8.2 [60] and EVTGEN-00-11-007 [53], is used to simulate the production and decay of $B \to \ell N(\to \ell\pi) + X$ processes for the $B^{\pm}, B^0$ and B_s species. PYTHIA handles proton-proton collision simulation and collision products hadronization. When a B meson is simulated, it is fed to EVTGEN, which handles its decay.

B_c generation in PYTHIA is inefficient in terms of computing resources, as the rate of events containing a B_c is very small due to the B_c fragmentation function introduced in Section 1.6. Therefore, heavy neutrinos from B_c mesons are generated in a separate sample, thanks to the dedicated generator BCVEGPY [14]. Proton-proton collisions with a B_c meson are generated by BCVEGPY, asking for the B_c meson to have $p_T > 8$ GeV, $|y| < 2.5$, then the events are handled to EVTGEN for the decay simulation.

The EVTGEN generator is used to implement novel particles, such as the heavy neutrino N, by specifying the particle mass, lifetime, spin, charge and width. Once the particle's characteristics are known, decay chains including it or originating from it can be defined.

Event generator filters (ϵ^{gen}) based on the particles allowed within the signal decay chain and their kinematics and characteristics are used to avoid the generation of events which would be eventually discarded at analysis level. A dedicated set of filters has been developed for this generation, based on the following requirements:

- Presence of a B meson;
- Presence of two leptons:
  - a μ_{trg} with $p_T > 6.8$ GeV and $|\eta| < 1.55$, to enforce B-Parking *tag-side* trigger conditions;
  - an electron or a muon with $p_T > 1$ GeV and $|\eta| < 2.45$;

Dataset name	Integrated luminosity (fb^{-1})
/ParkingBPH1/Run2018A-05May2019-v1/MINIAOD	0.774
/ParkingBPH2/Run2018A-05May2019-v1/MINIAOD	0.774
/ParkingBPH3/Run2018A-05May2019-v1/MINIAOD	0.774
/ParkingBPH4/Run2018A-05May2019-v1/MINIAOD	0.774
/ParkingBPH5/Run2018A-05May2019-v1/MINIAOD	0.774
/ParkingBPH6/Run2018A-05May2019-v1/MINIAOD	0.774
B-Parking Run2018A	4.626
/ParkingBPH1/Run2018B-05May2019-v2/MINIAOD	0.911
/ParkingBPH2/Run2018B-05May2019-v2/MINIAOD	0.911
/ParkingBPH3/Run2018B-05May2019-v2/MINIAOD	0.911
/ParkingBPH4/Run2018B-05May2019-v2/MINIAOD	0.911
/ParkingBPH5/Run2018B-05May2019-v2/MINIAOD	0.911
/ParkingBPH6/Run2018B-05May2019-v2/MINIAOD	0.911
B-Parking Run2018B	4.932
/ParkingBPH1/Run2018C-05May2019-v1/MINIAOD	1.103
/ParkingBPH2/Run2018C-05May2019-v1/MINIAOD	1.103
/ParkingBPH3/Run2018C-05May2019-v1/MINIAOD	1.103
/ParkingBPH4/Run2018C-05May2019-v1/MINIAOD	1.103
/ParkingBPH5/Run2018C-05May2019-v1/MINIAOD	1.103
B-Parking Run2018C	5.515
/ParkingBPH1/Run2018D-05May2019promptD-v1/MINIAOD	5.302
/ParkingBPH2/Run2018D-05May2019promptD-v1/MINIAOD	5.302
/ParkingBPH3/Run2018D-05May2019promptD-v1/MINIAOD	5.302
/ParkingBPH4/Run2018D-05May2019promptD-v1/MINIAOD	5.302
/ParkingBPH5/Run2018D-05May2019promptD-v1/MINIAOD	5.302
B-Parking Run2018D	26.510
2018 B-Parking dataset	41.599

Table 4.2. Data samples for the B-Parking dataset with their integrated luminosities.

- The invariant mass of the two leptons is constrained to be less than 10 GeV. This condition is meant to reject events in which the two selected leptons originate from two separate B meson decays;
- Requirement of the μ_{trg} to originate either from a B meson or a N;
- Presence of a π from the N decay , with $p_T > 0.5$ GeV;
- The transverse displacement of the generated heavy neutrino is required to be in the CMS tracker acceptance, $L_{xy} < 1.3$ m. The requirement rejects events that would yield L_{xy}s out of the analysis acceptance.

The number of expected events for a given signal hypothesis $(m_N, c\tau, f_e, f_\mu, f_\tau)$ as generated with the scheme described above results:

$$N^{sig} = \sigma\left(\mathrm{B\bar{B}}\right) \cdot \mathcal{L} \tag{4.1}$$

$$\sum_{\ell_i,\ell_j=\mu\mu,\mu e,e\mu} \sum_{\alpha,X_\alpha} f_\alpha \frac{\Gamma(B_\alpha \to \ell_i N X_\alpha)}{\Gamma(B_\alpha)} f_i |V|^2 \tag{4.2}$$

$$\varepsilon^{sel}_{gen,\ell_i\ell_j} \cdot \epsilon^{gen} \cdot \frac{\Gamma(N \to \ell_j \pi)}{\Gamma_e f_e + \Gamma_\mu f_\mu + \Gamma_\tau f_\tau} f_j \tag{4.3}$$

where:

- $\sigma\left(\mathrm{B\bar{B}}\right)$ is the inclusive $\mathrm{B\bar{B}}$ hadron cross-section at $\sqrt{s} = 13$ TeV, as measured in a phase-space compatible with the B-Parking trigger requirements and then weighted to be fully inclusive (see Section 5.1), and $\mathcal{L}$ is the integrated luminosity of the B-Parking dataset;
- the first sum runs over the i, j lepton pairs in which the N oscillates in the different analysis channels: i represents the ν_{SM} flavor in the $\nu_{SM} \to$ N oscillation while j stands for the ν_{SM} flavor in the oscillation N $\to \nu_{SM}$;
- the second sum runs over $\alpha = u, d, s$ and the respective B_α decay channels;
- f_i and f_j are the flavor coupling fractions introduced in Section 1.6;
- $|V|^2$ is the total coupling of the heavy neutrino N to SM neutrinos, as defined in Section 1.6;
- f_α is the fragmentation fraction for B_α;
- $\Gamma(B_\alpha \to \ell_i N X_\alpha)$ are the partial widths in the BSM theory, where the dependence on $|V|^2$ has been factored out, and $\Gamma(B_\alpha)$ is the SM width;
- $\varepsilon^{sel}_{sim,\ell_i\ell_j} = \frac{N^{sel}_{sim,\ell_i\ell_j}}{N^{sim}_{\ell_i\ell_j}}$ is the analysis selection efficiency for the simulated events in the i, j channel, weighted by the experimental weights;
- ϵ^{gen} is the generator filter efficiency;
- $\Gamma(N \to \ell_j \pi)$ is the partial decay width of the N, with the dependence on the coupling $|V|^2_\ell$ factored out.

The grid of simulated signal samples, parameterized in heavy neutrino mass m_N and coupling $|V|^2$, is summarized in Figure 4.1 for heavy neutrinos in $B^0, B^\pm, B_s$ decays, and in Figure 4.2 for heavy neutrinos produced in B_c decays. Five heavy neutrino mass hypotheses $m_N \in [1.0, 1.5, 2.0, 3.0, 4.5]$ GeV are generated for three or four lifetime hypotheses in the cτ $\in$[0.1,1000] mm range for the $B^0, B^\pm, B^s$ species. As the B_c production channel becomes relevant for heavy neutrino masses $m_N > 3$ GeV, the B_c samples are generated for masses of 3 GeV and 4.5 GeV, for three or four lifetime values in the cτ [0.1,1000] mm range. These samples have been used for the development and optimization of the signal candidate reconstruction and selection respectively described in Section 4.2 and Section 4.3.

An additional set of samples for several m_{HNL} hypotheses is simulated at a fixed lifetime, in order to allow for a fine scan of the neutrino mass hypotheses. The signal samples are simulated in:

- 0.02 GeV steps for $1 < m_{HNL} < 1.5$ GeV for a $c\tau = 10$mm lifetime;
- 0.03 GeV steps for $1.5 < m_{HNL} < 2.0$ GeV for a $c\tau = 10$mm lifetime;
- 0.05 GeV steps for $2.0 < m_{HNL} < 4.0$ GeV for a $c\tau = 10$mm lifetime for signals of mass $m_{HNL} < 3.0$ GeV and $c\tau = 1$mm lifetime for $m_{HNL} < 3.0$ GeV;
- 0.05 GeV steps for $m_{HNL} > 4.0$ GeV for lifetime $c\tau = 0.1$mm.

The mass steps vary according to the different signal resolution, as later defined in Section 5.1, for different mass hypotheses. The signals are generated for a single lifetime and reweighted through the technique described in Section 5.1. For masses $m_{HNL} > 3$ GeV both the B and B_c are simulated.

The $V_e \neq 0, V_\mu \neq 0$ spurious coupling condition is intrinsically realized in the μe *channel*, as pointed out in Section 1.6. No a-priori choice on the coupling scenario is imposed throughout the analysis, and therefore the μe *channel* coupling will be expressed as a function of the coupling fractions to the muon and electron flavors: $f_\mu f_e |V|^2 = \frac{|V_e \cdot V_\mu|^2}{|V|^2}$, where the coupling and coupling fractions are assumed to be real numbers.

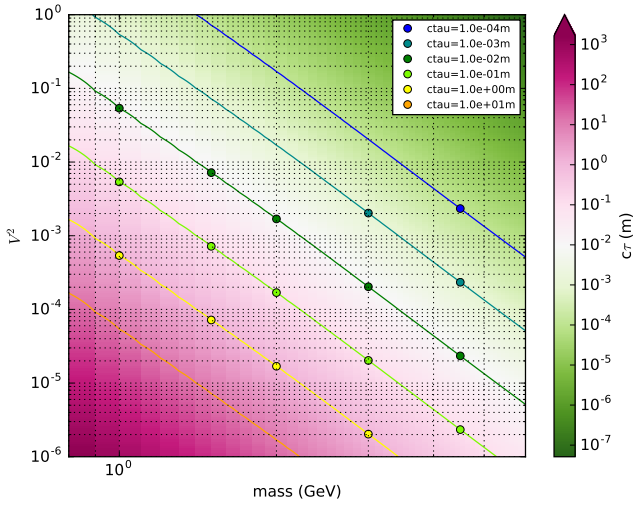


Figure 4.1. Values of the heavy neutrino mass m_N and coupling $|V|^2$ used for signal samples simulation for $B^0, B^\pm, B_s$ decays. The plot show contours at equal lifetimes.

The coupling fractions $f_\mu : f_e : f_\tau$ chosen for the generation can be arbitrary, as the number of events generated for the muon and electron channel can be re-weighted in Equation 4.3 in order to represent any coupling fractions combination.

The different heavy neutral lepton mass hypotheses heavily constrain the signal kinematics. As pointed out in Section 3.3, B mesons and their decay products are

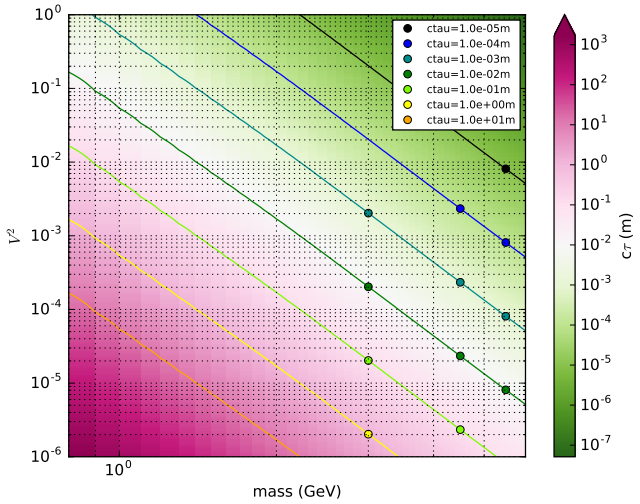


Figure 4.2. Values of the heavy neutrino mass m_N and coupling $|V|^2$ used for signal samples simulation for B_c decays. The plot show contours at equal lifetimes.

low energy objects in LHC. The limited amount of energy available in a B meson decay is split, in the investigated signatures, over the B meson decay vertex and the displaced N vertex. A heavier N mass hypothesis implicitly assigns higher energy content to the displaced vertex, leaving a very constrained phase space available for the $B^{\pm}, B^0, B_s \to \ell X$ decay.

A direct consequence of this energy unbalance is that for heavy N mass hypothesis, the triggering muon has higher probability of being yielded by the displaced vertex rather than the one from the B meson. The rate of generated muon passing B-Parking triggering conditions and coming from a B meson decay or an N decay are reported in Figure 4.3 for the B meson samples.

The rate of events triggered by a muon coming from a B meson decay is higher than 60-70% for masses of 1 or 1.5 GeV, yet it drops to less than 5% for signals of mass 4.5 GeV. The analysis strategy is designed to maximize the sensitivity to both light and heavy N hypotheses in the scan range, as the triggering muon can be originated from either the B or N decay vertexes.

In Figure 4.4, the ratio of B-Parking "tag-side" and "probe-side" events is reported for the simulated samples samples. Probe-side events, which do not fall within the analysis acceptance, represent for most signals less then 10% of the generated yields, with peaks of less than 30% for very displaced samples.

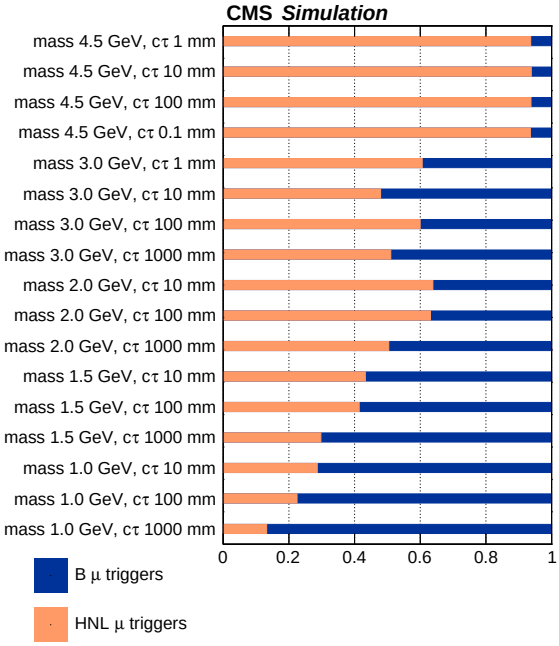


Figure 4.3. Probability of a muon from a B decay (blue) or from an N (labeled HNL) decay (orange) to trigger for the grid of (m_N,cτ) samples.

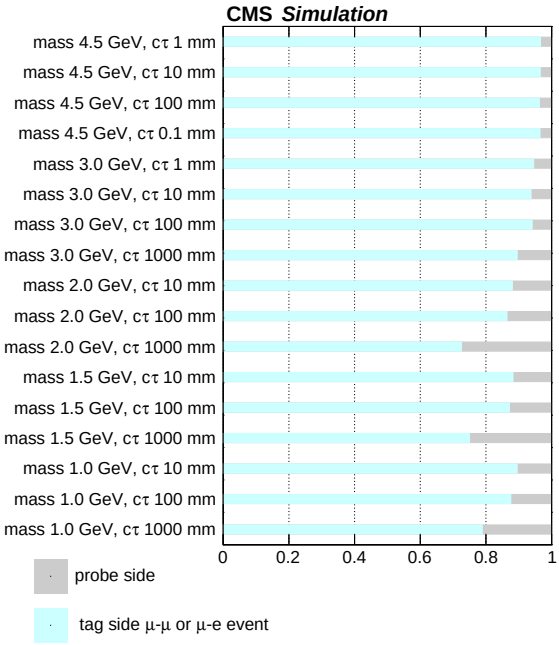


Figure 4.4. Rates of B-Parking tag and probe side events for the grid of (m_N,cτ) samples.

4.2 Analysis Objects and Definitions

This section details the reconstruction of the $B \to \ell_B \ell \pi$ signal candidate. The signal candidates are formed by three objects:

- one lepton ℓ_B from the B meson decay;
- one displaced lepton ℓ from the N decay;
- one displaced pion π.

Either one of the two final state leptons must be a B-Parking triggering muon μ_{trg}, as the analysis is performed on the dataset *tag* side. Standard PF muon, electron and track candidates are used for the signal candidate reconstruction in the analysis.

The signature of the analysis consists of a final state with three low-energy objects: a triggering muon, an electron, and a pion. Two of these particles originate, because of the N non negligible lifetime, from a displaced vertex. Multiple challenges to the reconstruction efficiency and quality of the reconstructed objects are set by the low energy and displaced features of the investigated signature.

The *lowpt electron* collection introduced in Section 3.3 has been extensively tested to be included in the analysis, as reported in the AppendixA.1. The collection resulted not suited for the purposes of this search: the looser requirements on track seeding and lowered thresholds for ECAL SC to track matching cause larger background yields which spoil the sensitivity increase due to the inclusion of electrons with $p_T < 2$ GeV.

The particles' displacement is quantified through transverse (d_{xy}) and longitudinal (d_z) impact parameters. These variables are defined as the distances, in the respective transverse and longitudinal planes, between the first track deposit and the Primary Vertex(PV). The impact parameter significances $d_{xy}/\sigma_{d_{xy}}$ and longitudinal d_z/σ_{d_z} quantify the quality of the impact parameter reconstruction: impact parameter significances of the order of unity or less indicate a displacement measurement for which the uncertainties are comparable or larger than the displacement value itself.

The signal candidate is built requesting two displaced objects, a lepton and a pion, on the B-Parking tag side. The tracks of the two objects are fitted to a common vertex through a kinematic fitter [51]: the fitter is successful only if the two tracks are compatible with the hypothesis of originating from a common vertex.

The fitted vertex is associated with a N candidate, its charge is computed together with the kinematic variables yielded from the fit. Vertex-related quantities are introduced to describe the vertex quality and the N candidate properties:

- The vertex probability prob(NV): the p-value associated to the χ^2 of the kinematic fit of the N candidate;
- The cosine back-pointing angle, $cos(\theta)$, as represented in Figure 4.5. The θ angle formed by the direction of the N candidate, computed as the vector linking its production and decay vertex, and the direction of the sum of the

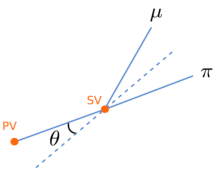


Figure 4.5. Representation of the back-pointing θ angle.

$\vec{p}_T$ of the two fitted tracks. In the sketch, PV stands for the B decay vertex while SV indicates the N decay vertex;

- The transverse displacement of the displaced vertex L_{xy}: the distance from the fitted vertex and the beam spot in the transverse plane;
- The significance of the transverse displacement of the N vertex, defined as the ratio between the transverse displacement and its uncertainty $sL_{xy}/\sigma_{L_{xy}}$

The N candidate is finally associated with a lepton ℓ_B and the kinematic variables for the combination of the $\mu e\pi$ candidate are computed.

The final state objects low energy and possible displacement heavily bias the reconstruction efficiency of the signal candidates. Most of the PF reconstruction algorithms are in fact optimized from high-energy prompt objects coming from the primary vertex.

The reconstruction efficiency is studied on a signal sample of $m_N = 3$ GeV and $c\tau = 1000$ mm.

The reconstruction efficiency is computed as:

$$\varepsilon = \frac{N_{reco}^{gen}}{N_{gen}}, \tag{4.4}$$

where N_{reco}^{gen} is the number of reconstructed events that pass a loose selection and are matched to the full decay chain of a simulated $B \to \mu e\pi$ event, while N_{gen} is the number of generated events falling in the acceptance cuts yielded by the loose selection.

The obtained efficiency is shown in Figure 4.6 as a function of the transverse momentum $p_T^{\ell\pi}$ and of the transverse displacement $l_{xy}^{\ell\pi}$ of the signal candidate. The efficiency decreases with growing displacement and smaller p_T values. In particular, the efficiency for vertexes below 15 cm is in the 10-40% range for the whole transverse momentum spectrum and drops to 2-3% over 15 cm.

4.3 Event Selection and Categorization

The event selection workflow and event categorization are presented in this section. The selection is organized in three steps. The event *preselection*, described in Section 4.3.1 aims at the identification of the best single signal candidate per event, and it is designed to preserve high signal efficiencies. Further selections are applied to veto the presence of standard model resonances in the reconstructed candidates,

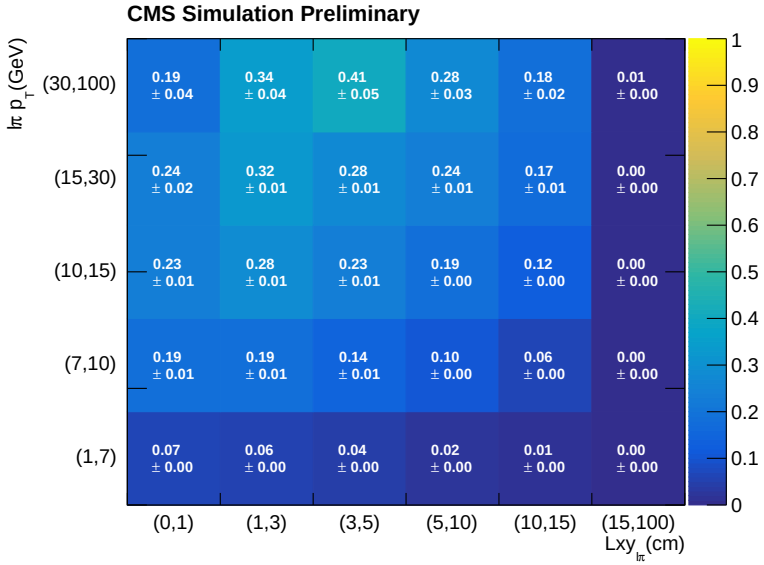


Figure 4.6. Reconstruction efficiency of the $\mu e\pi$ candidates in bin of transverse displacement L_{xy} (x axis) and transverse momentum p_T(y axis) of the N candidate.

in the *baseline selection*, outlined in Section 4.3.2, togehter with the lepton identification criteria. The selected events are then categorized and a final *selection* step, reported in Section 4.3.4 is performed through a parameterized deep learning network in the different analysis categories.

Selection and categorization studies are performed directly on data unblinding 4.91 fb^{-1} of data, corresponding to ∼10% of the full B-Parking integrated luminosity, under the assumption that the signal contamination is negligible. This dataset will be referred to as unblinded 1D dataset, as it is extracted from the part 1D B-Parking data.

4.3.1 Preselection

The preselection has been designed on two criteria: preservation of the signal efficiency and reduction of the number of reconstructed signal candidates per event.

Up to 200 signal candidates are reconstructed for each event without selections applied, therefore the pre-selection criteria are tuned to reduce the number of reconstructed signal candidates per event. A signal efficiency higher than $\sim$ 75% is targeted for all the (m_N,cτ) hypotheses.

The unblinded 1D dataset is used to study the data distribution, while three benchmark signals with different (m$_N$,cτ) are used to provide a representation of the possible kinematics and different lifetimes covered by the search sensitivity. The benchmark signals are:

- $m_N = 1$ GeV, c$\tau = 1000$ mm;
- $m_N = 3$ GeV, c$\tau = 100$ mm;
- $m_N = 4.5$ GeV, c$\tau = 1$ mm;

Data and simulated samples are processed through the candidate reconstruction described in Section 4.2. A set of variables discriminating signal and background is identified through the comparison of the signal samples and data distribution for several variables related to both the single final state particles and the N and $\mu e\pi$ reconstructed candidates.

The comparison of the benchmark signals and data distributions for some of these variables is reported in Figure 4.7 and Figure 4.8.

The final state pion transverse momentum is softer in data with respect to the MC samples, while as expected the vertex quality tends to be higher in the simulated samples rather than in data. The vertex probability and cosine of the back-pointing angle are in fact lower in the background then in simulated samples, indicating an average worse vertex quality for the background candidates with respect to signals. This feature is expected, as one of the main background sources for the analysis lies in combinatorial association on unrelated particles which do not actually originate from a common vertex, as further discussed in 5.2.1.

Displaced related quantities of final state particles can be exploited for discrimination. The comparison between data and the three MC benchmark signals distributions for the transverse impact parameter significances of the three final state particles are reported in Figure 4.7.

The impact parameter significance configures as good discriminator, in both the transverse and longitudinal planes. Background processes final state particles yield smaller impact parameters and smaller impact parameter significances with respect to the long-living signals.

The pre-selection configures as a set of requirements applied with a cut-based approach on discriminating variables. The selection is tuned using a sequential approach, by optimizing each threshold on the candidates that have passed all the previous selections. The preselection requirements consist in:

- A triggering muon with transverse momentum greater than 7 GeV in the barrel region ($|\eta|$ <1.5);
- An electron with transverse momentum greater than 2 GeV in the $|\eta|$ <2.0 region;
- A displaced charged pion from the N candidate decay with transverse momentum greater than 0.7 GeV and in the $|\eta|$ <2 region. The displaced pion is required to have:
  - transverse impact parameter $d_{xy} > 0.005$ cm;
  - longitudinal impact parameter $d_z > 0.005$ cm;
  - transverse displacement significance $d_{xy}/\sigma_{xy} > 3$;
  - longitudinal displacement significance $d_z/\sigma_z > 1.5$.
- A displaced lepton from the N candidate decay (which can be either the electron or the triggering muon) with:
  - transverse impact parameter $d_{xy} > 0.001$ cm;
  - longitudinal impact parameter $d_z > 0.0015$ cm;

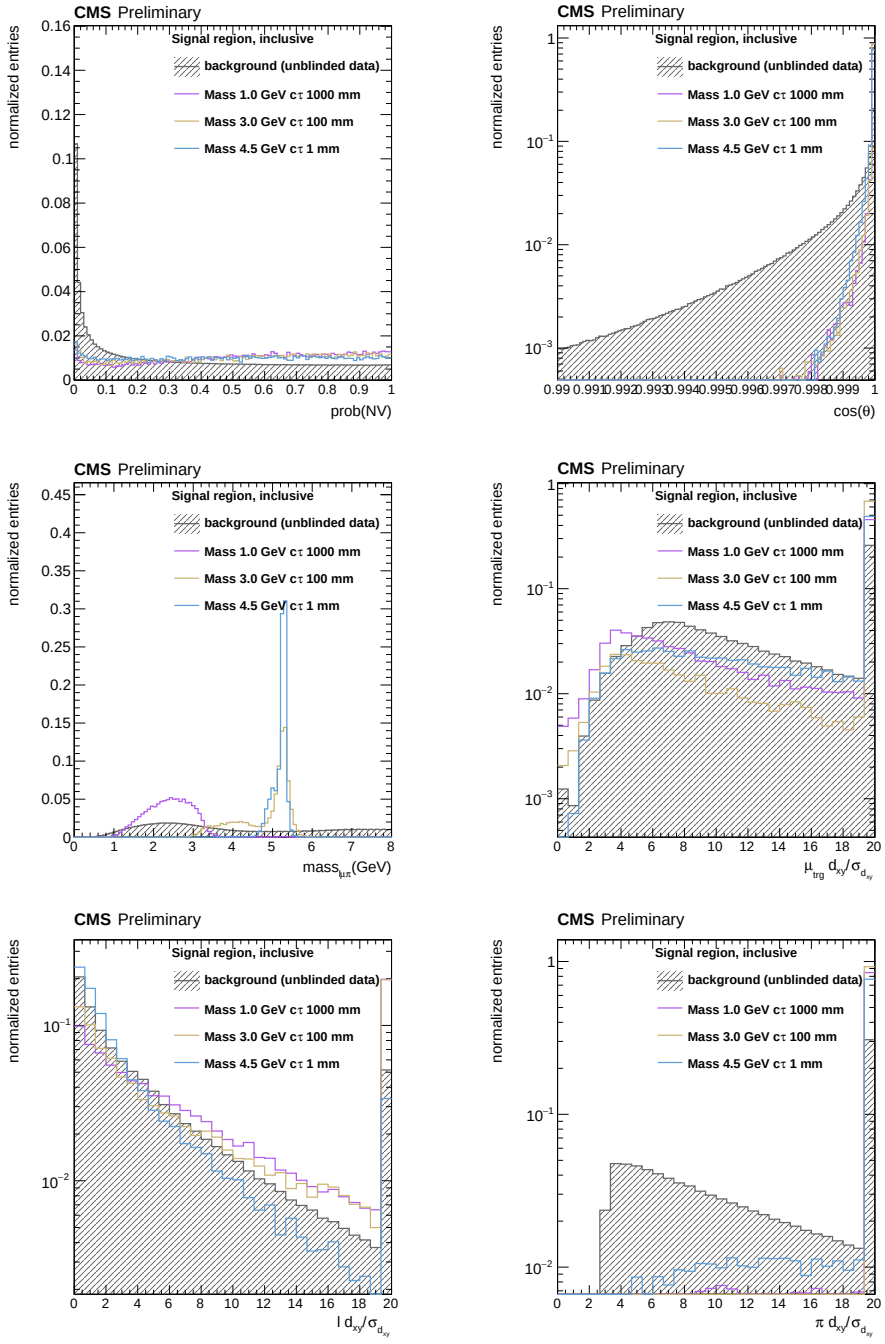


Figure 4.7. Shape comparison between three signal hypotheses and data-driven background distributions of $\mu e\pi$ candidate-related quantities and final state particles displacement-related variables, as obtained with loose preselection criteria. The vertex probability (top left), cosine of the back-pointing angle $cos(\theta)$ (top right), the $\mu e\pi$ invariant mass (center left), distributions and the $d_{xy}/\sigma_{d_{xy}}$ spectra for the triggering muon (center right), electron (bottom left) and pion (bottom right) are shown. The last bin includes the overflow.

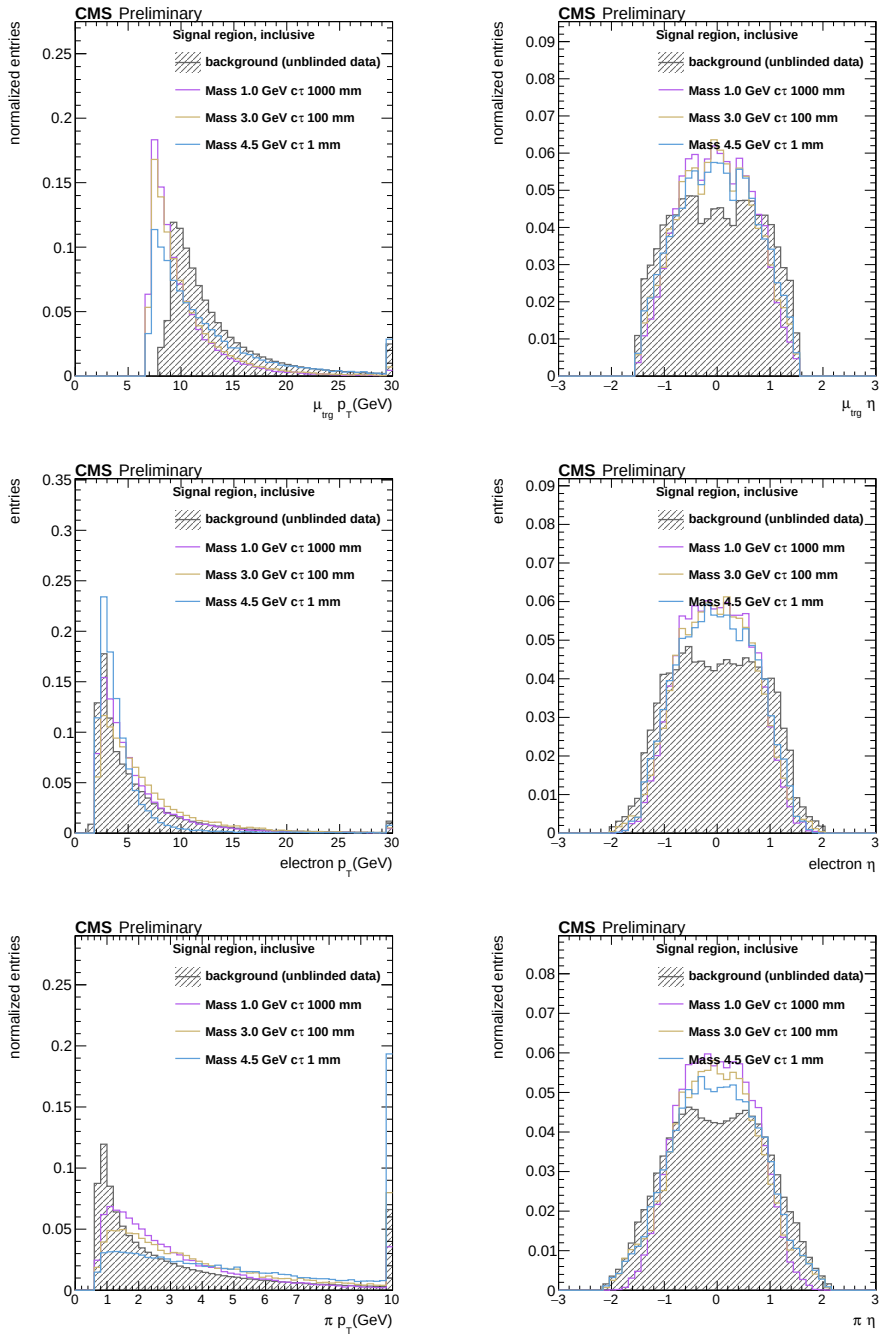


Figure 4.8. Shape comparison between three signal hypotheses and data-driven background distributions of the final state particles kinematics for $\mu e\pi$ candidates, as obtained with loose selection criteria, in the inclusive Signal Region. On the left and right columns are shown the p_T, and $|\eta|$, spectra. The top row shows the distributions for the trigger muon, the middle row the ones for the electron and the bottom row the ones for the displaced pion. The last bin includes the overflow.

  - transverse displacement significance $d_{xy}/\sigma_{xy} > 1.5$;
  - longitudinal displacement significance $d_z/\sigma_z > 1$.

- N vertex probability prob(NV) larger than 0.001;
- cosine of the back-pointing angle $cos(\theta)$ larger than 0.995;
- significance of the transverse displacement of the N candidate vertex $L_{xy}/\sigma_{L_{xy}}$ larger than 15;
- invariant mass of the three particles system $m_{\mu e\pi} < 8$ GeV.

The preselection preserves has efficiency $\varepsilon > 75\%$ for the three benchmark signals, while rejecting 99.72% of the background; a breakdown of the efficiencies at each preselection step is reported in Section 4.4. The average number of reconstructed candidates per event after the preselection is 1.9.

As the analysis strategy relies on the B-Parking *tag-side* only, at most a single $B \to \mu e\pi$ signal candidate is expected for each event. A criterium is therefore designed to identify the best signal candidate. If more than one reconstructed candidate per event passes the preselection, the candidate with the higher $\cos(\theta)$ is selected as signal candidate. The validity of the criterium is tested by evaluating the single candidate selection efficiency defined as:

$$\varepsilon = \frac{N_{gen}^{higher\ cos(\theta)}}{N_{gen}} \tag{4.5}$$

for the three benchmark signals. The efficiency results $\sim 95\%$ for the three signal samples.

4.3.2 Baseline Selection

The baseline selection identifies the analysis Signal Region (SR), implements objects identification requirements, and poses vetoes on the SM resonances appearing in the SR. An extensive discussion on SM resonances vetoes is reported in Section 5.2.1. The selections applied are:

- $\ell\pi$ charge equal to 0: the heavy neutrino N is a neutral particle;
- $L_{xyz} < 100$ cm: this enforces the heavy neutrino to decay within the tracker volume. The requirement $L_{xy} < 130$ cm was already applied when producing the signal samples, as described in Section 4.1.2;
- $|\mathrm{m}_{\ell_B\pi} - 3.097| > 0.05$ GeV in SS events: removal of the J/ψ resonance;
- $|\mathrm{m}_{\ell_B\pi} - 1.76| > 0.15$ GeV in SS events: removal of the D^0 resonance;
- muon identification requirements;
- electron identification requirements;

	1 GeV 1000 mm	3 GeV 100 mm	4.5 GeV 1 mm	background (unblinded data)
PF MVA ID > -3	-13.1%	-4.4%	-1.6%	-39.1%

Table 4.3. Signal and background efficiencies of the electron identification criterion.

- $|\eta_e| < 1.422$. Electron ID efficiencies in data and simulation are evaluated on a $B^{\pm} \to K^{\pm} J/\psi(\to e^+e^-)$ sample with the tag&probe. The number of reconstructed candidates in the endcap regions is limited and yields efficiency uncertainties that would spoil the sensitivity brought to the analysis by including the endcap events. A comprehensive discussion of electron identification efficiencies is reported in Section 6.4.

The baseline selection has an efficiency over 90% on the three benchmark signals, with a $\sim 50\%$ background efficiency.

Muon Identification

Muon candidates are required [42] to be reconstructed in the tracker and in the muon system. Different identification requirements are applied to the displaced muons of the N candidate and muons related to the B decay.

The muon from the B decay is required to have

- a tracker track matched with at least one muon segment in the muon chambers;
- more than 5 hits in the tracker layers;
- at least one hit in the pixel layers.

Displaced muons, on the other hand, are required to be either reconstructed through the combination of the muon chamber hits and the track, or to be tracker-driven muons: the track deposit is matched to hits in the muon system, but the particle kinematics is computed relying only on tracker information.

Electron Identification

The PF MVA ID score introduced in Section 3.3 is used for electron identification in the analysis. The identification criterion is specifically designed on the distribution of the PF MVA ID scores for the signals and background pertaining to the analysis. The PF MVA ID output distributions for the three benchmark signals and a small portion of the unblinded data are reported in Figure 4.9.

The ID criterion is chosen to be PF MVA ID> -3, with signal and background efficiencies reported in Table 4.3.

4.3.3 Event Categorization

The events are categorized based on their transverse displacement significance $L_{xy}/\sigma_{L_{xy}}$ and the relative sign of the leptons in the reconstructed candidate.

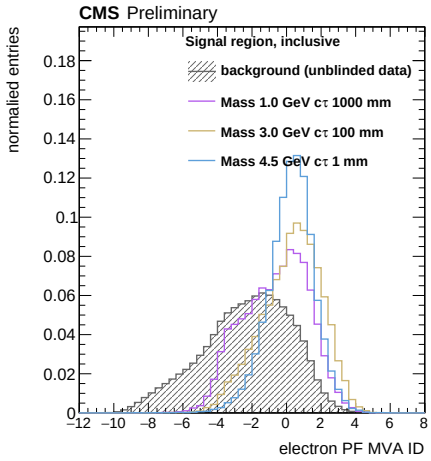


Figure 4.9. Shape comparison between three signal hypotheses and data-driven background distributions of the PF MVA ID, in the inclusive Signal Region defined in Section 4.3.2.

The categorization aims at the enhancement of signals of different lifetimes. The $L_{xy}/\sigma_{L_{xy}}$ distributions for the three benchmark signals and part of the unblinded data is reported in Figure 4.10.

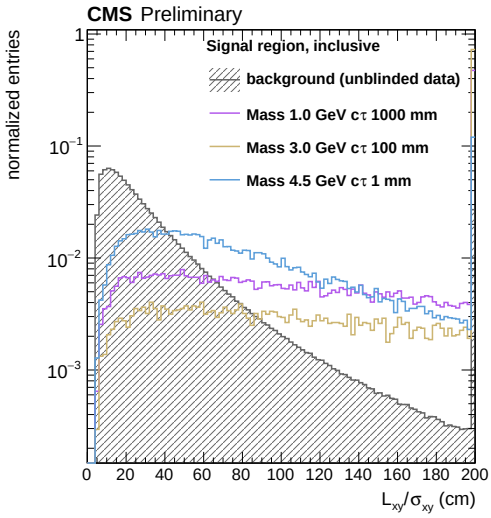


Figure 4.10. Shape comparison between three signal hypotheses and data-driven background distributions of the significance of the transverse displacement of the N candidate in the inclusive Signal Region defined in Section 4.3.2.

The data show a smoothly falling shape in $L_{xy}/\sigma_{L_{xy}}$, as most SM processes typically yield events in which a particle produced in proton-proton collision decays instantly. On the other hand, the three benchmark signals populate the high $L_{xy}/\sigma_{L_{xy}}$ region as well. In particular, the benchmark signals with higher lifetimes, (m_N=3 GeV, cτ=100 mm) and (m_N=1 GeV, cτ=1000 mm), describe long-lived par-

ticles which decay away from the interaction region, yielding high $L_{xy}/\sigma_{L_{xy}}$ values. Short-lived signals as (m_N=4.5 GeV, cτ=1 mm) populate more the low $L_{xy}/\sigma_{L_{xy}}$ range, yet their distribution remains very different from the one of the background.

The $L_{xy}/\sigma_{L_{xy}}$ categories defined in the analysis are:

- low transverse displacement significance category: $L_{xy}/\sigma_{L_{xy}} < 50$;
- medium transverse displacement significance category: $50 < L_{xy}/\sigma_{L_{xy}} < 150$;
- high transverse displacement significance category: $L_{xy}/\sigma_{L_{xy}} > 150$;

A further categorization is built on the relative signs of the charges of the leptons in the $\mu e\pi$ reconstructed candidate. SM events observed up to now, in fact, preserve the total lepton number, while processes involving Heavy Majorana Neutrinos may not. Hence, less background is expected in the SS category, while the categorization allows for a separate study of the different background sources in lepton number violating and non-violating processes.

A total of 6 categories is defined for the analysis. A summary of the categorization is reported in Table 4.4.

A dedicated categorization is introduced to the analysis for heavy neutrino masses in the $m_N > 3$ GeV regime. The B_c heavy neutrino production channel becomes relevant for these masses, and it is treated in the analysis as a signal process orthogonal to the $B^0, B^\pm, B_s \to \ell_B N$ decays. The orthogonality condition is realized by asking for the invariant mass of the three final state particles $m_{\mu e\pi}$ to be smaller than 5.7 GeV for $B^0, B^\pm, B_s \to \ell_B N$ signal candidates and larger than 5.7 GeV for $B_c \to \ell_B N$ decays. The inclusive leptonic and sempileptonic decays of the $B^0, B^\pm, B_s$ species are kinematically limited by the mass of the B meson $m_B = 5.28$ GeV. On the other hand, the B_c decays fully leptonically $B_c \to \ell_B \ell\pi$, and therefore the invariant mass of the three final state particles, $m_{\mu e\pi}$ resonates at the B_c mass value $m_{B_c} = 6.27$ GeV.

The analysis is therefore run on 12 categories for $m_N > 3$ GeV, 6 for events with $m_{\mu e\pi} < 5.7$ GeV ("B categories") and 6 for events with $m_{\mu e\pi} > 5.7$ GeV ("B_c categories").

$L_{xy}/\sigma_{L_{xy}} \leq 50$, OS	$(50 < L_{xy}/\sigma_{L_{xy}} \leq 150)$, OS	$L_{xy}/\sigma_{L_{xy}} > 150$, OS
$L_{xy}/\sigma_{L_{xy}} \leq 50$, SS	$(50 < L_{xy}/\sigma_{L_{xy}} \leq 150)$, SS	$L_{xy}/\sigma_{L_{xy}} > 150$, SS

Table 4.4. Categorisation based on the displacement significance and the relative (OS or SS) leptons charge.

4.3.4 Signal Region Selection

The selection step aims at the optimization of the sensitivity for all the signal hypotheses treated in the analysis. It is based on the development of a Parametric Neural Network (pNN) [11] for signal-to-background discrimination.

Parametric Neural Networks are a relatively new approach to multivariate learning techniques. Neural networks [55] are deep learning algorithms inspired by the

biological neural connections that constitute the human brain, specifically designed to toggle non-linear problems. Deep learning techniques are currently exploited in CMS, for event classification [4], object reconstruction and identification [49] (Section 3.3), or triggering [43]. These algorithms, however, are usually applied to solve a single specific task, such as the discrimination of an interesting physics event against the p-p collisions low energy background.

This search for heavy neutrino deals with different mass and lifetime hypotheses for the N candidate. The signal kinematics and displacement behavior changes with different masses and lifetimes. A standard approach, using cut-based selection or multivariate techniques, would imply the design of a set of isolated selection criteria for each signal hypothesis. Each of those set of selections would be ignorant of the larger context and would therefore lack the ability to interpolate among the different hypotheses.

Parametrized Neural Networks have the ability to toggle a set of multiple related problems. The network takes as input one or several discrete parameters $\vec{\theta}$, together with the standard set of continuously distributed input features $\vec{x}$. The pNN score obtained is a function of all the input parameters:

$$\text{score} = f\left(\vec{\theta}, \vec{x}\right) \tag{4.6}$$

Such networks are able to learn several different signatures, related to different sets of parameters. Moreover, they succeed in interpolating for intermediate θ values which are not included in the training set.

An independent network is trained in each of the six (or twelve, for masses $m_N > 3$ GeV, as discussed in Section 4.3.3) analysis categories introduced in Section 4.3.3. Separate trainings are performed for the B and B_c categories. The signal training sample consists of a mixture of events from the signal samples introduced in Figure 4.1, while part of the unblinded 1D dataset is used as the background sample. The discrete parameter designed for the pNN is the mass of the N candidate, as it allows to achieve a continuous discrimination performance over the wide range of mass hypotheses explored. The mass parameter design strategy is summarized in Figure 4.11 for the mass parameter values, signal and background samples used for the network, for the $L_{xy}/\sigma_{L_{xy}} > 150$, OS category.

The discrete parameter choice for the signal samples is straightforward, as the parameter assigned to be the signal mass hypothesis m_N for each sub-sample of the signal training set. The different mass parameters used in the signal training sample are indicated in different colors in Figure 4.11 left plot.

The background samples are treated by selecting only events in a narrow window of $\sim 10\sigma$ around each mass signal hypothesis, as shown in Figure 4.11 right plot. The σ is the standard deviation of the signal distribution for a specific mass hypothesis, as later defined in Section 5.1. A standard approach for pNN discrete parameter building [55] would assign the m_N parameter value for the background by random sampling of the background distribution: this procedure might be dangerous in this framework, as the $m_{\ell\pi}$ distribution is correlated with most of the discriminating variables that can be included in the training features set, biasing the response of the background sample during training. The mass hypothesis m_N around which the window has been opened is therefore chosen as discrete parameter for each

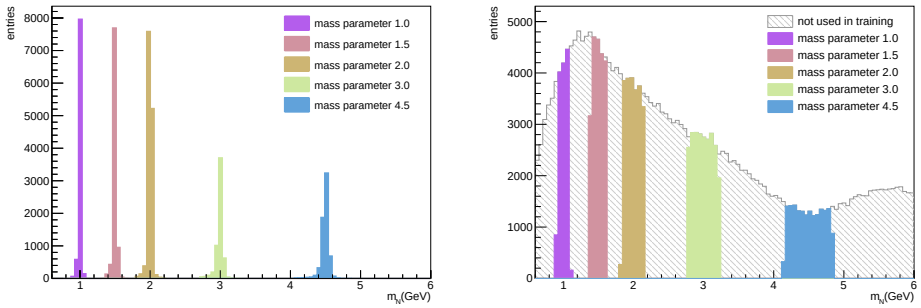


Figure 4.11. mass parameter value and sizes for signal (left) and background (right) pNN training samples for the $L_{xy}/\sigma_{L_{xy}} > 150$, OS analysis category. The different colours correspond to the 5 mass parameters assigned to each sample in the training. The grey shaded area in the background spectrum corresponds to data that is not going to be used in the training.

background subset.

The sample size used for each training ranges from 50 000 events to 120 000 events. As can be seen in Figure 4.11, the training sets are built by requiring for the signal and background samples sizes to be balanced for each mass hypothesis. The total sample size for each mass hypothesis is required to be the same, in order to have equal representation of all the mass hypotheses. No constraint is applied on the number of signal events of different lifetimes for a fixed mass.

The set of features for pNN training is the same in all the categories. It is composed by a mixture of kinematics related and lifetime dependent variables:

- Primary lepton p_T
- Displaced pion p_T
- Displaced lepton p_T
- $\ell_B\ell\pi$ invariant mass
- $\ell_B\ell\pi$ p_T
- N candidate $\cos\theta$
- N candidate $L_{xy}/\sigma_{L_{xy}}$
- prob(NV)
- NV χ^2
- $\ell_B\ell$ invariant mass
- $\ell_B\pi$ invariant mass

A simple network architecture is employed. It consists of three layers:

- input layer: contains as many nodes as the number of features plus one extra node for the mass discrete parameter;
- hidden layer: contains 64 fully connected nodes, and is activated with a Rectified Linear Unit (ReLu) [6] activation function;
- output layer: returns the pNN score normalized to one through the sigmoid activation function. The pNN score can be understood as the probability of an event to be a signal event.

The training is performed through the `KERAS` [41] package of the `TENSORFLOW` framework [61], making use of the Adam optimizer [25]. Since neural networks training is proven to be more efficient when the training features distributions range over values of the same orders of magnitudes, sklearn RobustScaler [26] is used for standardization.

Each network is trained for 50 epochs, over a batch size of 32 units, with dynamic learning rate starting at 0.01 and decreasing of a 0.2 factor if the accuracy does not improve over 5 epochs. A summary of the pNN hyper parameters is reported in Table 4.5.

Hyperparameter	Value
Nodes in hidden layer	64
Learning rate	0.01
Batch size	32
Epochs	50

Table 4.5. List of hyper parameters in the pNN architecture.

The training loss and accuracy curves for training validation are reported in Figure 4.12 and Figure 4.13 respectively. No sign of overtraining is observed.

The ROC curves and score distributions for the different training mass hypotheses and the corresponding lifetimes are respectively reported in Figure 4.14 and Figure 4.15, for the $L_{xy}/\sigma_{L_{xy}} > 150$, OS category. The plots are produced with the full 1D unblinded dataset. The performance plots for the other training categories can be found in Appendix A.

Both the ROC curves and the score distributions show good performances for all the mass hypotheses. The performances are, as expected, slightly worse for signals with very low lifetimes, as the high transverse displacement category is expected to be mostly populated by long-lived signals.

The pNN performance is finally compared with the one yielded by a standard NN trained on a single task. A standard neural network is trained on a mass 3 GeV and lifetime $c\tau$ 100 mm sample and compared in performance over different mass hypotheses with the pNN. The figure of merit chosen for the comparison, the Area Under the ROC Curve (AUC), is reported for the pNN and a NN in Figure 4.16, for the $L_{xy}/\sigma_{L_{xy}} < 50$ OS category.

The pNN and NN performance is compatible for the signal point with mass 3 GeV, as the training on multiple signal hypotheses does not spoil the performance on

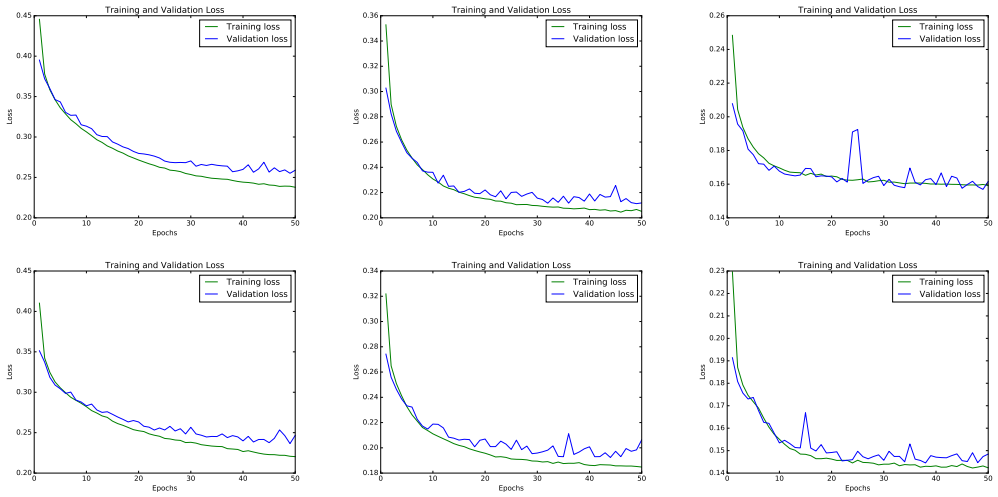


Figure 4.12. Loss curves for the training (green) and validation (blue) sets as a function of the training epochs in the different categories. The top row shows the three OS categories, while the bottom plots the SS ones. The left column shows the $L_{xy}/\sigma_{L_{xy}} < 50$ categories, the center column, the $50 < L_{xy}/\sigma_{L_{xy}} < 150$ ones and finally the right column, the $L_{xy}/\sigma_{L_{xy}} > 150$ ones.

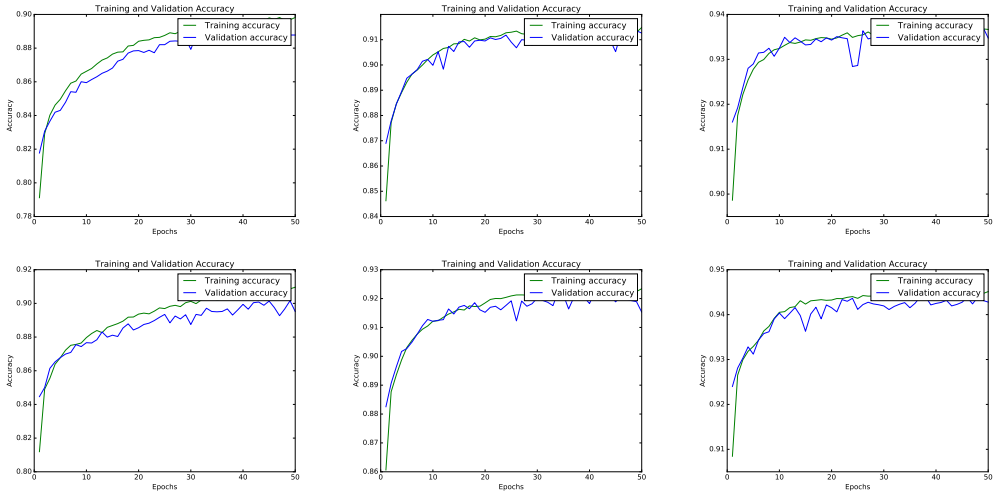


Figure 4.13. Accuracy curves for the training (green) and validation (blue) sets as a function of the training epochs in the different categories.he top row shows the three OS categories, while the bottom plots the SS ones. The left column shows the $L_{xy}/\sigma_{L_{xy}} < 50$ categories, the center column, the $50 < L_{xy}/\sigma_{L_{xy}} < 150$ ones and finally the right column, the $L_{xy}/\sigma_{L_{xy}} > 150$ ones.

a single mass hypothesis. On the other hand, the pNN AUC values are higher with respect to the NN ones for all the other mass points, validating the interpolation and generalization power of the pNN with respect to a single task NN. Analogous considerations apply for all the analysis categories, the performance plots are reported

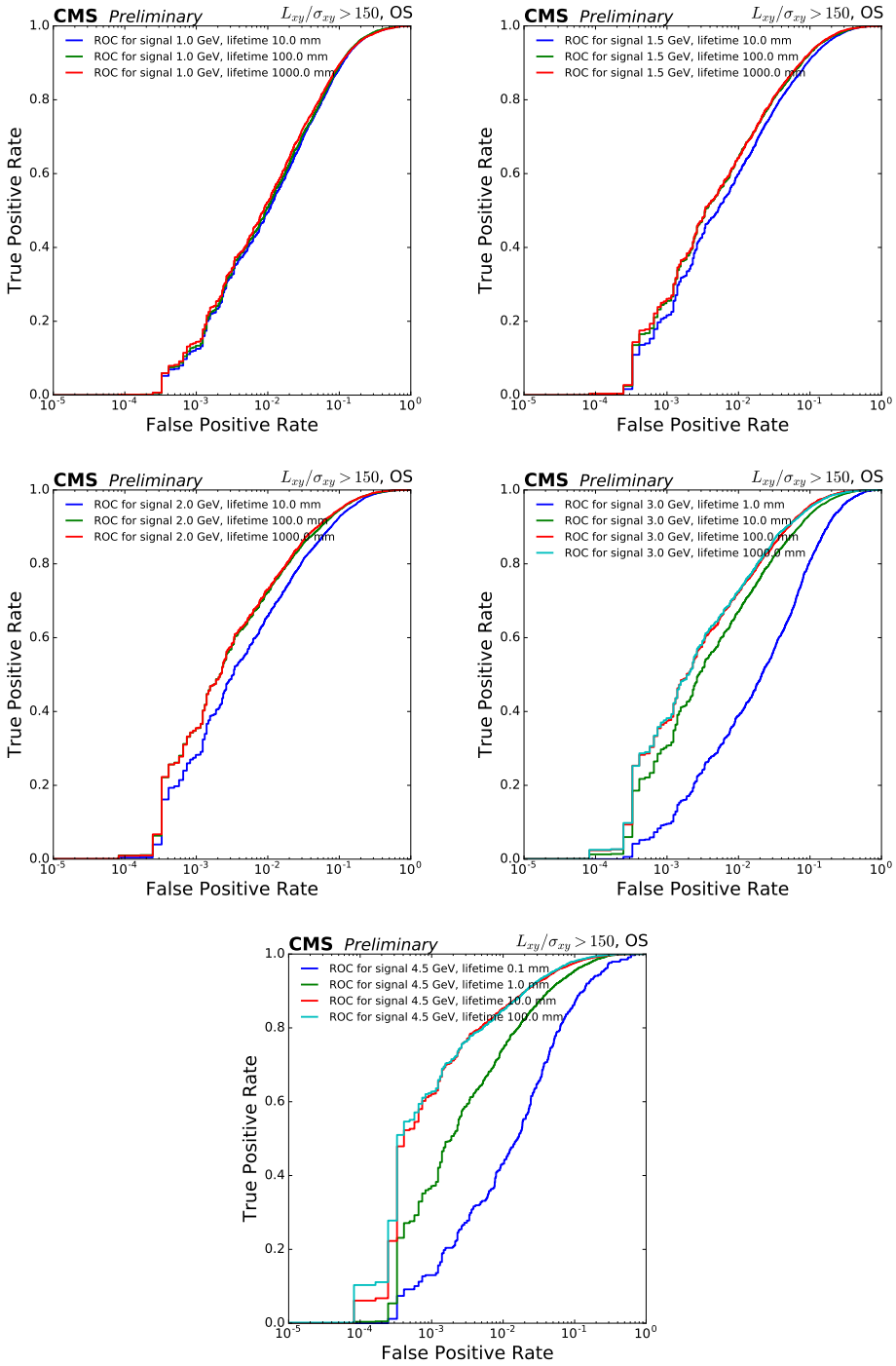


Figure 4.14. ROC curve for the mass hypotheses used in the training in the high L_{xy}/σ_{Lxy} OS category for the $\mu e\pi$ channel.

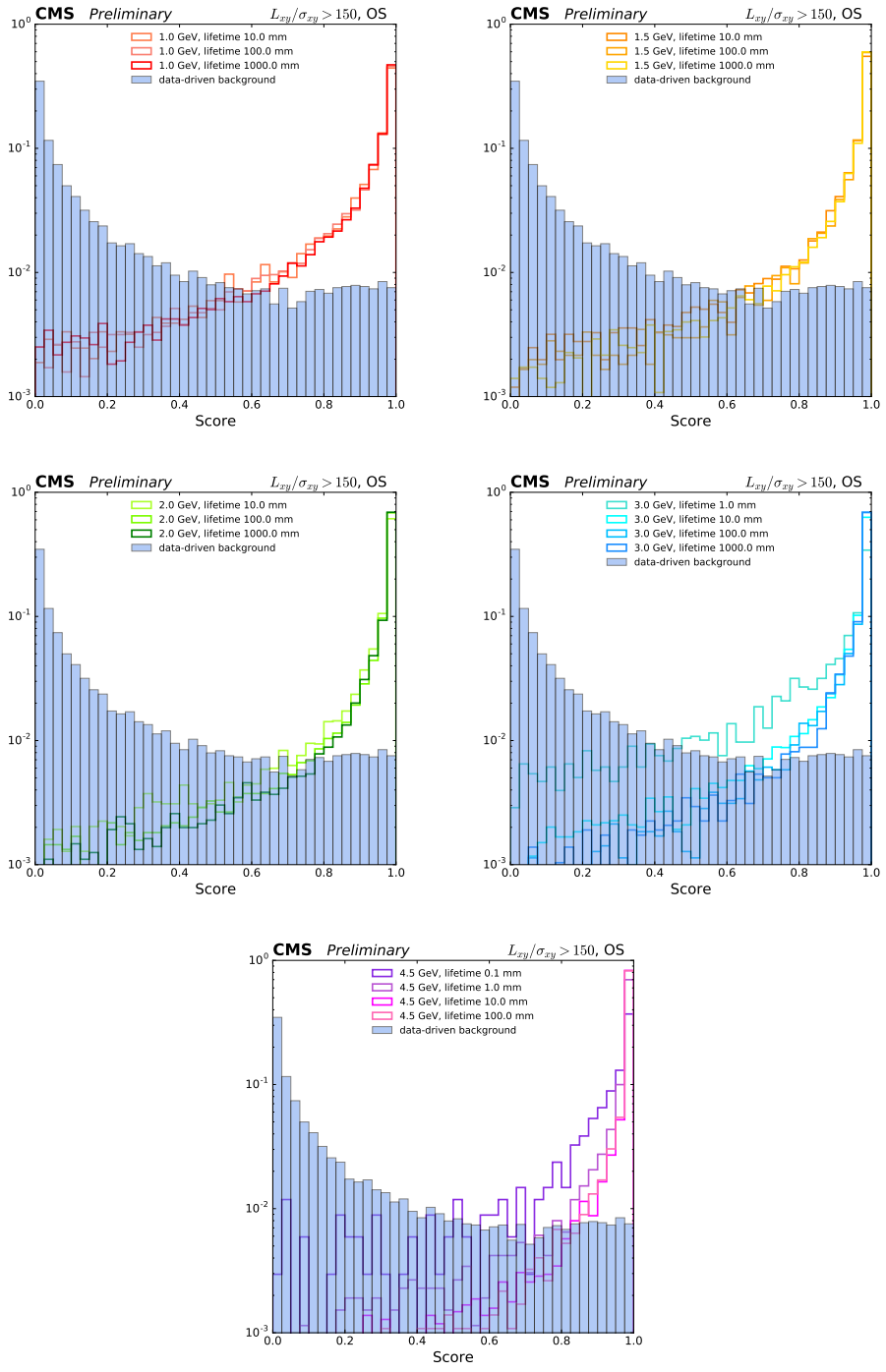


Figure 4.15. Score distribution for the mass hypotheses used in the training in the high $L_{xy}/\sigma_{L_{xy}}$ OS category for the $\mu e\pi$ channel.

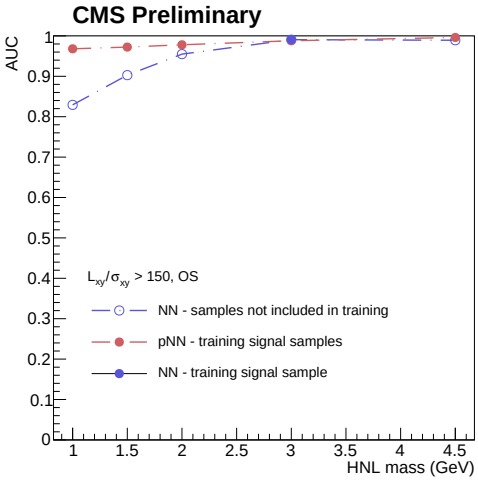


Figure 4.16. Comparison of the Area Under the Curve (AUC) value for ROCs computed for different mass hypotheses using the analysis pNN (red) or a simple NN (blue) trained on a single mass 3 GeV lifetime 100 mm signal sample and for the high $L_{xy}/\sigma_{L_{xy}}$ OS category.

in Appendix A.

The median expected limit on the signal strength μ for each analysis category is chosen as figure of merit for the pNN selection optimization. The limit is computed under the conditions presented in Section 5 and through the statistical tools presented in Section 7.1. The limit value behavior as a function of the pNN selection is presented, as an example, for a signals of mass 1 GeV and lifetime 2000 mm in Figure 4.17.

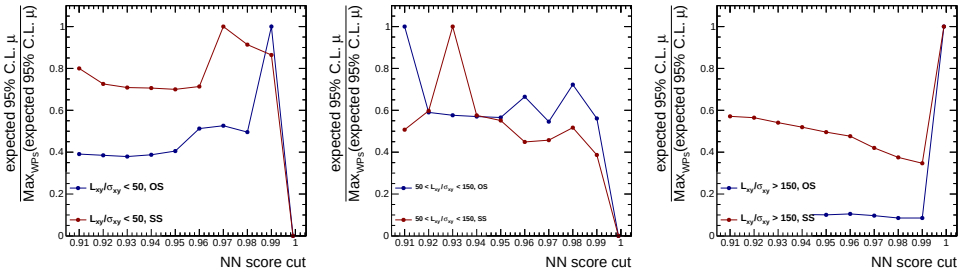


Figure 4.17. Relative median expected limit on the signal strength μ as a function of the cut applied on the pNN score for events in the electron channel for a signal of mass 1 GeV, lifetime 2000 mm. The three plots correspond to the three transverse significance categories: $L_{xy}/\sigma_{L_{xy}} \leq 50$ (left) , $50 \leq L_{xy}/\sigma_{L_{xy}} \leq 150$ (center) and $L_{xy}/\sigma_{L_{xy}} > 150$ (right). The OS and SS categories are reported respectively in blue and red for each plot.

The limit value is normalized, for each WP scan, to the maximum computed limit. The limits shown in Figure 4.17 decrease with larger WPs in the high transverse significance displacement categories, which have the higher sensitivity for this signal. A similar behavior can be observed in the $50 \leq L_{xy}/\sigma_{L_{xy}} \leq 150$ category.

The limit is rather stable over the WP scanned range for the low transverse displacement category; it slightly worsens for high values of the WP, but the fact does not represent a concern as this category holds the lowest sensitivity in this signal scenario.

The optimization procedure is repeated for several signal hypotheses in all the analysis categories. The expected limit shows similar behavior in all the investigated categories and signals in the pNN WP[0.91,0.999] range . The limit is rather stable or decreases with higher pNN cuts. For pNN values of 0.999 and beyond, the signal efficiency drops for all the categories, leading to worse expected limits. A uniform pNN cut throughout all the categories, pNN score$>$ 0.99 is therefore chosen as WP for the analysis.

The pNN learning pattern is tested against the possibility of learning the $m_{\ell\pi}$ mass, as some of the training variables are correlated with this quantity. The pNN would then sculpt the background distribution in favor of the signal hypotheses, classifying background events around the signal hypotheses mass values in the $m_{\ell\pi}$ spectrum with lower scores. Dip-like structures would then appear in the $m_{\ell\pi}$ spectra at each m_N mass hypothesis when selecting events based on their pNN score.

The test is performed by comparing the background distributions around each training m_N hypothesis with and without a pNN selection applied. The comparison of the distributions for the whole 1D unblinded dataset with and without a pNN$>$0.99 selection is reported in Figure 4.18 for the $L_{xy}/\sigma_{L_{xy}} > 150$, OS category. Plots for the other categories can be found in Appendix A. While the shape of the distribution can be modified by the pNN selection, no structure that would bias the signal extraction is observed.

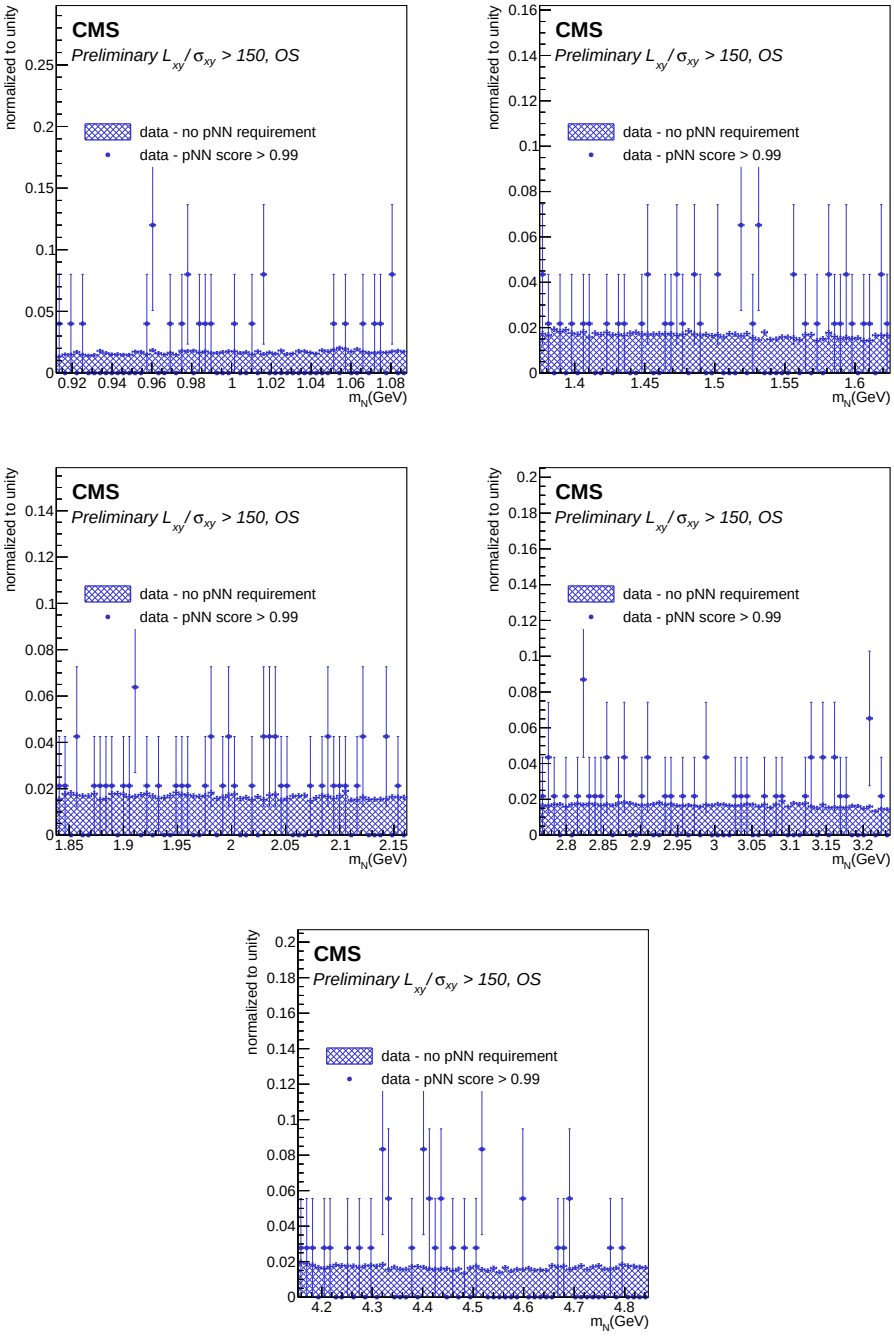


Figure 4.18. Shape comparison of the $\ell\pi$ invariant mass spectrum with and without a cut on the pNN score in mass windows around 1, 1.5, 2, 3 and 4.5 GeV in the high $L_{xy}/\sigma_{L_{xy}}$ OS category for the $\mu e\pi$ channel. The pNN working point is set to 0.99 for all the categories and mass windows. The plots are produced with the full unblinded statistics of the B-Parking dataset (4.91 fb^{-1}).

4.4 Summary of Selection Efficiencies

A summary of the event selection presented in the previous sections is reported below, together with a breakdown of the signal and background efficiencies for each selection step.

The pre-selection and baseline selection efficiencies are reported in Table 4.6 and 4.7 respectively. The efficiencies are computed for the three benchmark signals introduced in Section 4.3.1 and for part of the unblinded data. The selection is cut based, therefore for each selection applied both the sequential efficiency ϵ and the cumulative efficiency ε up to the studied cut are computed.

The pNN selection efficiencies are computed on top of the pre-selection and baseline selection and presented in Table 4.8. The efficiencies are computed for all the signal samples and for the unblided 1D data and for the different analysis categories.

	1 GeV 1000 mm		3 GeV 100 mm		4.5 GeV 1 mm		background (data)	
	ϵ	$\mathcal{E}$	ϵ	$\mathcal{E}$	ϵ	$\mathcal{E}$	ϵ	$\mathcal{E}$
trig muon pt > 7 GeV	98.34%	98.34%	99.54%	99.54%	99.82%	99.82%	100.00%	100.00%
trig muon $\|\eta\|$ < 1.5	100.00%	98.34%	99.53%	99.07%	99.67%	99.49%	95.75%	95.75%
electron pt > 0.7 GeV	99.81%	98.15%	100.00%	99.07%	99.96%	99.45%	95.69%	91.62%
electron $\|\eta\|$ < 2	100.00%	98.15%	100.00%	99.07%	100.00%	99.45%	97.54%	89.36%
pion pt > 0.7 GeV	96.99%	95.19%	97.66%	96.76%	97.10%	96.57%	54.14%	48.38%
pion $\|\eta\|$ < 2	100.00%	95.19%	99.28%	96.06%	98.56%	95.18%	92.56%	44.78%
pion dz > 0.005 cm	98.25%	93.53%	98.80%	94.91%	97.66%	92.96%	96.82%	43.35%
pion dxy > 0.005 cm	98.42%	92.05%	99.27%	94.21%	98.63%	91.68%	61.74%	26.76%
pion dz significance > 1.5	98.19%	90.39%	98.53%	92.82%	99.01%	90.77%	96.71%	25.88%
pion dxy significance > 3	96.93%	87.62%	99.00%	91.90%	99.20%	90.04%	31.41%	8.13%
pion DCA significance > 5	97.47%	85.40%	99.24%	91.20%	97.57%	87.85%	81.91%	6.66%
displaced lepton dz > 0.0015 cm	100.00%	85.40%	100.00%	91.20%	98.26%	86.32%	95.12%	6.33%
displaced lepton dxy > 0.001 cm	98.70%	84.29%	100.00%	91.20%	99.75%	86.10%	95.92%	6.08%
displaced lepton dz significance > 1	97.81%	82.44%	97.97%	89.35%	95.72%	82.42%	89.68%	5.45%
displaced lepton dxy significance > 1.5	96.86%	79.85%	98.45%	87.96%	99.51%	82.01%	77.92%	4.25%
prob(NV) > 0.001	96.99%	77.45%	97.37%	85.65%	97.60%	80.04%	15.29%	0.65%
$\cos(\theta)$ > 0.995	99.52%	77.08%	99.46%	85.19%	99.95%	80.01%	78.53%	0.51%
hnl lxy significance > 20	98.56%	75.97%	99.46%	84.72%	97.13%	77.71%	72.70%	0.37%
$\mu l\pi$ mass < 8 GeV	100.00%	75.97%	100.00%	84.72%	100.00%	77.71%	74.32%	0.28%

Table 4.6. Pre-selection cutflow table for the $\mu e\pi$ candidates. For each sample, the list of sequential cuts is given in the leftmost column ϵ, while the right column $\mathcal{E}$ represents the cumulative efficiency up to the examined cut.

	1 GeV 1000 mm		3 GeV 100 mm		4.5 GeV 1 mm		background (data)	
	ϵ	$\mathcal{E}$	ϵ	$\mathcal{E}$	ϵ	$\mathcal{E}$	ϵ	$\mathcal{E}$
$\ell\pi$ charge = 0	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	54.5%	54.5%
μ_B soft ID or μ_N loose ID	99.0%	99.0%	99.7%	99.7%	99.8%	99.8%	98.8%	53.4%
PF electron MVA ID > -3	87.5%	86.6%	97.8%	97.5%	98.3%	98.1%	58.9%	31.2%
displaced π has high purity track	95.8%	83.0%	96.6%	95.0%	97.4%	93.2%	92.0%	28.7%
Vetos in $\ell_0\pi$ mass spectrum	97.5%	80.9%	97.0%	92.0%	97.8%	90.6%	95.5%	27.3%
$L_{xy}(\sigma) > 0$	100.0%	80.9%	100.0%	92.0%	100.0%	90.6%	100.0%	27.3%

Table 4.7. Baseline selection cutflow table for the $\mu e\pi$ candidates. The list of sequential cuts is given in the leftmost column. For each selection cut, the efficiency ϵ as well as the cumulative efficiency $\mathcal{E}$ up to that cut are reported for the three benchmark signal points and for the background. The efficiencies are computed with the pre-selection of Table 4.6 applied.

Mass (GeV)	Category	Background efficiency	Signal yields			
			0.1 mm	10 mm	1000 mm	10000 mm
1.0			5.4e+00	5.4e-02	5.4e-04	5.4e-05
	L_{xy}/σ <50, OS	0.02 %	-	0.80 %	1.04 %	0.96 %
	50<L_{xy}/σ <150, OS	0.14 %	-	7.24 %	8.10 %	7.96 %
	L_{xy}/σ >150, OS	0.18 %	-	21.22 %	23.10 %	22.19 %
	L_{xy}/σ <50, SS	0.08 %	-	3.70 %	4.77 %	4.27 %
	50<L_{xy}/σ <150, SS	0.10 %	-	8.95 %	9.19 %	9.28 %
	L_{xy}/σ >150, SS	0.46 %	-	35.84 %	38.82 %	38.00 %
1.5			7.1e-01	7.1e-03	7.1e-05	7.1e-06
	L_{xy}/σ <50, OS	0.01 %	-	0.69 %	0.60 %	-
	50<L_{xy}/σ <150, OS	0.06 %	-	5.50 %	6.69 %	-
	L_{xy}/σ >150, OS	0.22 %	-	31.86 %	35.28 %	-
	L_{xy}/σ <50, SS	0.04 %	-	2.75 %	4.35 %	-
	50<L_{xy}/σ <150, SS	0.07 %	-	8.68 %	9.77 %	-
	L_{xy}/σ >150, SS	0.38 %	-	43.09 %	48.91 %	-
2.0			1.7e-01	1.7e-03	1.7e-05	1.7e-06
	L_{xy}/σ <50, OS	0.00 %	-	0.50 %	-	-
	50<L_{xy}/σ <150, OS	0.04 %	-	5.02 %	-	-
	L_{xy}/σ >150, OS	0.21 %	-	38.00 %	-	-
	L_{xy}/σ <50, SS	0.02 %	-	1.81 %	-	-
	50<L_{xy}/σ <150, SS	0.04 %	-	7.22 %	-	-
	L_{xy}/σ >150, SS	0.16 %	-	44.67 %	-	-
3.0			2.2e-02	2.2e-04	2.2e-06	2.2e-07
	L_{xy}/σ <50, OS	-	-	-	-	-
	50<L_{xy}/σ <150, OS	-	-	-	-	-
	L_{xy}/σ >150, OS	0.19 %	-	42.17 %	-	-
	L_{xy}/σ <50, SS	-	-	-	-	-
	50<L_{xy}/σ <150, SS	-	-	-	-	-
	L_{xy}/σ >150, SS	0.23 %	-	48.27 %	-	-
4.5			2.9e-03	2.9e-05	2.9e-07	2.9e-08
	L_{xy}/σ <50, OS	0.43 %	66.65 %	64.61 %	-	-
	50<L_{xy}/σ <150, OS	0.17 %	33.58 %	44.98 %	-	-
	L_{xy}/σ >150, OS	0.21 %	70.49 %	72.75 %	-	-
	L_{xy}/σ <50, SS	0.51 %	63.23 %	63.86 %	-	-
	50<L_{xy}/σ <150, SS	0.28 %	36.81 %	42.29 %	-	-
	L_{xy}/σ >150, SS	0.26 %	75.63%	78.54 %	-	-

Table 4.8. Signal and background efficiency for the mass hypotheses of 1, 1.5, 2, 3 and 4.5 GeV in the different categories for events with pNN score > 0.99. The background yields are computed in the 2 σ window around the mass hypothesis, with σ defined as in Section 5.1. All the efficiencies are computed using 4.91 fb^{-1} statistics of the B-Parking dataset. The efficiencies are computed with the preselection and baseline selection reported in Tables 4.6 and 4.7 applied The dash symbol "-" denotes that the yields have not been computed.

Chapter 5

Signal and Background Modeling

The analysis configures as a search for a peak in the invariant mass $m_{\ell\pi}$ of the displaced decay of an N candidate, therefore two ingredients are needed: signal modeling and background estimation from a fit to the data.

The signal sample parametrization strategies are reported in Section 5.1, together with a description of the signal normalization technique.

Low energy QCD processes are expected to be the main source of background to this analysis: the hypothesis is discussed in Section 5.2.1, together with a focus on SM resonances vetoes applied to several distributions in the signal region. The sliding mass window fit technique used for background yields extraction is presented in Section 5.2.3, together with the discrete profiling strategy for the background shape determination.

A summary of the expected signal and background yields extraction strategies and values are presented in Section 5.3.

5.1 Signal Normalization and Parametrization

5.1.1 Normalization

The number of expected signal yield for each (m_N,$\mathrm{c}\tau$) hypothesis depends, as reported in Equation 4.3, on the $\mathcal{L} \cdot \sigma_{B\bar{B}}$ factor. An independent measurement of the $\mathcal{L} \cdot \sigma_{B^\pm}$ value has been computed in the analysis framework, exploiting the $B^+ \to J/\psi(\to \mu^+\mu^-)K^+$ decays. This channel is chosen as its branching ratios are precisely known, and the final state is fully reconstructed, so that the invariant mass of the decay products peaks at the mass of the B meson.

The $\sigma(B^\pm)\mathcal{L}$ value is estimated as:

$$\sigma(B\pm)\mathcal{L} = \frac{N^{sel}_{data}}{\mathcal{BR}(B^\pm \to K^\pm J/\psi)\mathcal{BR}(J/\psi \to \mu\mu)} \times \frac{N^{gen}_{sim}}{N^{sel}_{sim}\,\epsilon^{gen}}, \tag{5.1}$$

where:

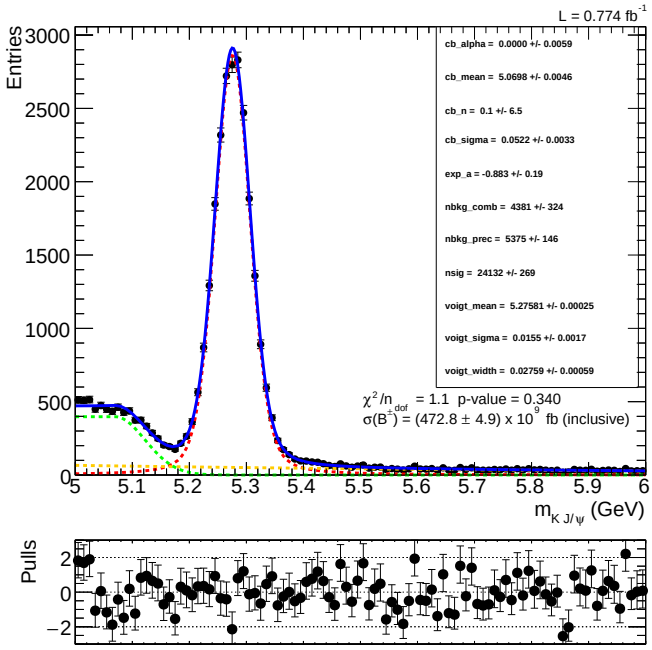


Figure 5.1. Distribution of the $\mu\mu$K invariant mass for a luminosity of 0.774 fb^{-1} of the B-parking dataset. In top pannel, the blue curve corresponds to the full model while the red, green and yellow curves show the models of the signal, composite background and combinatorial background respectively. The yielded parameters of the full model are given as well as the $\chi^2/\mathrm{n_{dof}}$ and the p-value. The extracted value of $\sigma\,(\mathrm{B}^{\pm})$ is also indicated. The bottom panel shows the pull distribution for each point.

- N^{sel}_{data} is the number of selected events in data of $B^{\pm} \to J/\psi(\to \mu\mu)K^{\pm}$ candidates, extracted from a maximum likelihood fit of the invariant mass peak of two muons and track after appropriate event selection;
- N^{sel}_{sim} and N^{gen}_{sim} are the number (sum of the weights) of selected and generated events, respectively, in a simulated $B^{\pm} \to J/\psi(\to \mu\mu)K^{\pm}$ simulated sample;
- ϵ^{gen} is the generator filter efficiency of the simulated sample.

Muons and tracks for the $B^{+} \to J/\psi(\to \mu^{+}\mu^{-})K^{+}$ candidate reconstruction are selected according to the objects definitions described in Section 4.2. Candidate $B^{\pm} \to J/\psi(\to \mu^{+}\mu^{-})K^{\pm}$ events are obtained by reconstructing the vertex formed by a μ_{trig} and another muon μ and associating the di-muon system to a track. A selection is applied on the candidate exploiting the decay kinematics and fit vertex quality variables to enhance the B peak in data over the background.

The invariant mass distribution of the selected data events is shown in Figure 5.1. The results correspond to an integrated luminosity of 0.774 fb^{-1} of the B-parking dataset. The signal peak corresponding to $B^{\pm} \to J/\psi(\to \mu^{+}\mu^{-})K^{\pm}$ events is clearly visible. The invariant mass distribution is fitted with an extended maximum likelihood unbinned fit using the following functional forms:

- Voigtian function for the signal;
- Exponential function + Crystal Ball function for the background.

The measurement yields a value for the cross section of charged B mesons at LHC:

$$\sigma_{B^{\pm}} = 472.8 \pm 4.9\ (stat)\ \mu\text{b}. \tag{5.2}$$

The $B\bar{B}$ cross section can be retrieved from this measurement exploiting the B meson species fragment fractions as follows:

$$\sigma_{B\bar{B}} = \frac{\sigma_{B^{\pm}}}{f_u}. \tag{5.3}$$

The validity of the approach presented above for the measurement of B meson production cross sections in the B-Parking dataset has been tested by repeating the measurement in a fiducial space in which a CMS measurement of the σ_{B^+} had been already performed [45]. The two measurements are in good agreement, thus validating this cross section computation technique.

5.1.2 Lifetime Reweighting

Yields for signals of lifetimes different from the ones reported in Figure 4.1 can be obtained by a reweighting procedure on signal events. The kinematic distributions and yield for a sample of lifetime $c\tau_1$ can be retrieved from an existing sample of lifetime $c\tau_0$ by assigning to each event a weight $w\left(ct, c\tau_0 \to c\tau_1\right)$ computed as:

$$w\left(ct, c\tau_0 \to c\tau_1\right) = \frac{\frac{1}{c\tau_1}\exp\left(-\frac{ct}{c\tau_1}\right)}{\frac{1}{c\tau_0}\exp\left(-\frac{ct}{c\tau_0}\right)} \tag{5.4}$$

where ct is the simulated lifetime of the N in the event, calculated as:

$$ct = \frac{L_{xyz}}{\beta_N \gamma_N}; \tag{5.5}$$

L_{xyz} is the 3D displacement of the N candidate and $\beta_N\gamma_N$ is the boost factor for the generated N. Such event weight w weights the ct distribution from the original $c\tau_0$ sample to match the ct distribution of the new sample with lifetime $c\tau_1$.

The expression in Equation 5.4 can be easily generalized to the case where a set of samples $\{i\}$ generated with $c\tau_i$ rather than a single $c\tau_0$ sample is used as basis for the reweighting:

$$w\left(ct, \{c\tau_i\} \to c\tau_1\right) = \frac{\frac{N^{\text{tot}}_{\text{gen}}}{c\tau_1 \cdot \epsilon^{\text{avg}}_{\text{filter}}}\exp\left(-\frac{ct}{c\tau_1}\right)}{\sum_i \frac{N^{\text{i}}_{\text{gen}}}{c\tau_i \cdot \epsilon^{i}_{\text{filter}}}\exp\left(-\frac{ct}{c\tau_i}\right)}. \tag{5.6}$$

This expression configures as the weighted average over the $\{i\}$ signals of the lifetime weight computed for a single sample in Figure 5.4. The weight involves the total number of generated events N_{gen} and the generator filter efficiency ϵ^{gen} as defined in

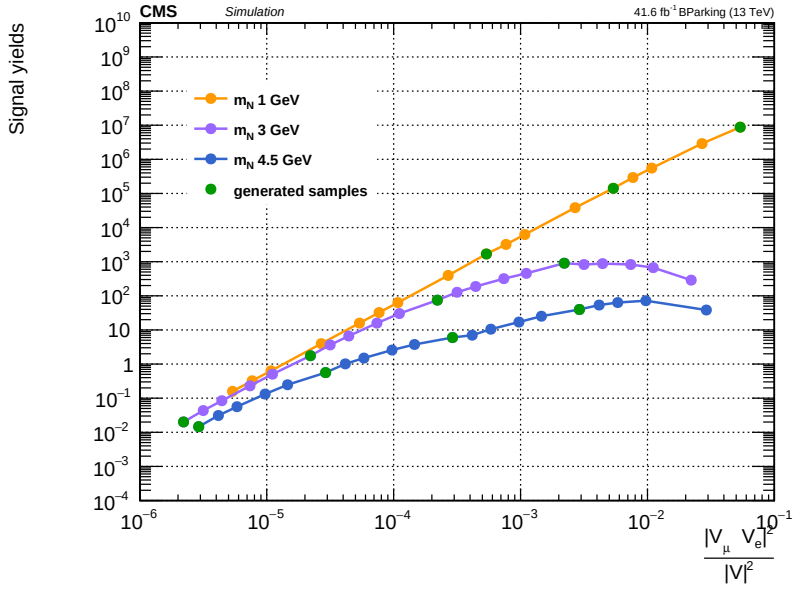


Figure 5.2. Expected $\mu e\pi$ signal yields for an integrated luminosity of 41.6 fb^{-1} as obtained with the pre-selection and baseline selection applied in the inclusive category, are shown as a function of the coupling $|V_\mu \cdot V_e|^2/|V|^2$. Couplings corresponding to generated samples are represented in dark green, while the intermediate points are obtained through the lifetime reweighting procedure.

Section 4.1.2, as the lifetime weights are related to the full ct generated distribution and not only to the one yielded by the events passing the event selection.

The expected $\mu e\pi$ signal yields for the total B-Parking dataset integrated luminosity, as obtained with the pre-selection and baseline selection applied in the inclusive category, are shown in Figure 5.2 as a function of the couplings. The yields are normalised to the full B-parking dataset luminosity. The the cτ points for which samples have been generated are represented in dark green, while the intermediate cτ points are obtained through the reweighting procedure.

As can be seen, the reweighting allows for smooth interpolation between the different coupling and lifetime hypotheses.

5.1.3 Resolution Parametrization

The signal resolution for a given signal hypothesis is the parameter defining the width of the fit window for background yield extraction. It is studied as a function of the mass and lifetime in this section. The signal shapes are fitted through a binned maximum likelihood fit with a Double Crystal Ball model:

$$f(x,\alpha_l,n_l,\alpha r,n_r,\mu,\sigma) = N\cdot\begin{cases} G(x,\mu,\sigma) = e^{-\frac{(x-\mu)^2}{2\sigma^2}}, & \frac{(x-\mu)}{\sigma} > -\alpha_l \;\; and \;\; \frac{(x-\mu)}{\sigma} < \alpha_r \\ A_l \cdot \left(B_l - \frac{(x-\mu)}{\sigma}\right)^{-n_l} & \frac{(x-\mu)}{\sigma} < -\alpha_l \\ A_r \cdot \left(B_r - \frac{(x-\mu)}{\sigma}\right)^{-n_r} & \frac{(x-\mu)}{\sigma} > \alpha_r. \end{cases} \tag{5.7}$$

where the x stands in the signal distribution fit for the N candidate invariant mass $m_{\ell\pi}$. The Double Crystal Ball function is formed by a gaussian core $G(x, \mu, \sigma)$ and two power law left and right tails. It is used to describe the signal in order to account for both electron energy losses by bremsstrahlung and possible bias to the invariant mass value computation brought by the electron pion tracks fit to a displaced vertex.

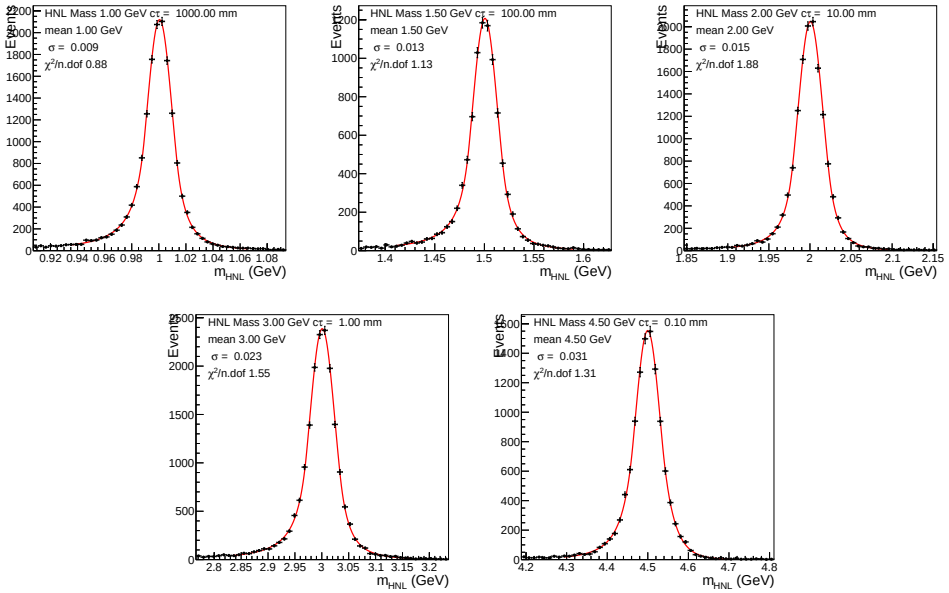


Figure 5.3. Distribution (black) of the invariant mass of the N candidate $m_{\ell\pi}$ for signals of different masses and lifetimes. The Double Crystal Ball fit to the distribution is reported in red.

Examples of the $m_{\ell\pi}$ distribution for signals of different masses and lifetimes are reported in Figures 5.3. The signal resolution is defined as the σ Gaussian parameter, as it quantifies the width of the signal distribution Gaussian core. It is shown, as a function of the heavy neutral lepton mass and for different lifetime hypotheses, in Figure 5.4. The resolution is not influenced by different lifetime hypotheses, as the particle kinematics is completely fixed by the N candidate, B meson and ℓ_B masses and kinematics. The width of the mass peak does not depend on the signal lifetime: the resolution value over different lifetimes varies in 2-10% range for the different mass hypotheses.

The signal resolution shows a linear behavior as a function of the N mass and it can be therefore parametrized in a linear form. The parameters are obtained through a linear fit over the lifetime averaged value of the resolution for each mass point. The fit result is shown in Figure 5.5 and the fit parameters result:

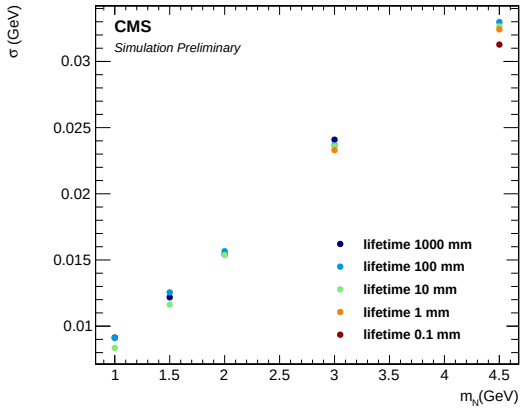


Figure 5.4. Resolution σ as a function of the signal mass in the $\mu e\pi$ channel. For each mass, the different $c\tau$ hypotheses for which a centrally produced simulated sample has been generated are shown.

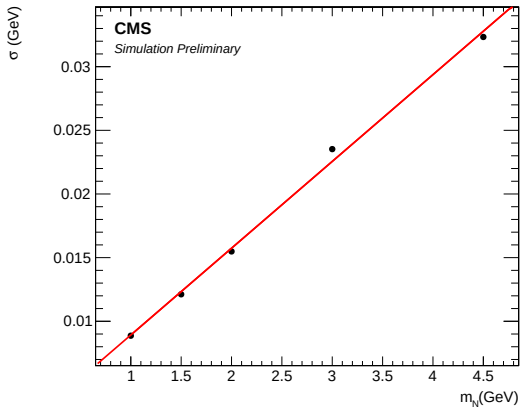


Figure 5.5. Average resolution (in GeV) over the different lifetime hypotheses as a function of the signal mass in the $\mu e\pi$ channel. The points are fitted by a polynomial of first order.

$$\sigma\left(\mathrm{m_N}\right) = p_0 + p_1 \cdot \mathrm{mass}$$
$$p_0 = 1.4 \cdot 10^{-3}\mathrm{GeV} \tag{5.8}$$
$$p_1 = 7.4 \cdot 10^{-3} \tag{5.9}$$
$$\tag{5.10}$$

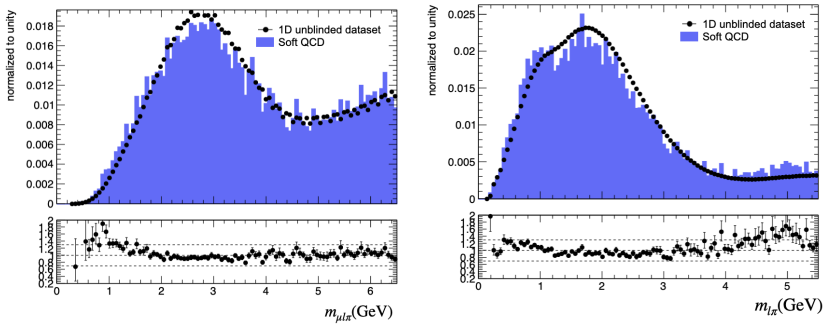


Figure 5.6. Shape comparison of the $\mu\ell\pi$ (left) and $\ell\pi$ (right) invariant mass distributions between data and simulation in the inclusive category of the Control Region (N charge $\neq 0$). In the upper pads, the blue shaded area corresponds to QCD simulated samples while the dots corresponds to the unblinded 1D B-Parking data. The lower pads show the data to simulation ratio.

5.2 Background Estimation

Background shape and yields are computed from a direct fit to the data, as the analysis configures as a search for an invariant mass peak over a continuous background. Background sources are discussed in Section 5.2.1, while SM resonances observed and vetoed in the analysis signal region are discussed in Section 5.2.2. The discrete profiling strategy, which allows for an unbiased choice of the background shape from a direct fit on data, is presented in Section 5.2.3.

5.2.1 Background Composition

The main background component for this search is represented by soft QCD processes. The analysis does not rely on the understanding and modeling of the background components, yet a simple study is here presented to verify the background composition hypothesis and characterize the background sources.

The analysis signature consists of three low-energy objects, two of which stem from a displaced vertex. Most of the SM high-energy processes are ruled out of this signature simply for their prompt nature, which does not allow for a significantly displaced vertex to be reconstructed.

QCD processes entering the analysis background are constituted either by SM B meson decays which are mis-reconstructed in the target analysis signature, or by signal candidates built out of the combinatorial association of unrelated particles.

As the main background component is expected to be combinatorial, and therefore not depending on the signal candidate properties, similar contributions are expected for both the signal region, which is defined in Section 4.3.2 by requiring for the N candidate to be neutral, and the control region for which $Q_N \neq 0$.

A comparison of the $\mu e\pi$ and $\ell\pi$ invariant mass spectra for the unblinded data and simulated soft QCD processes in an analysis CR (N not neutral in charge) is reported in Figure 5.6.

A good agreement can be observed between the data distribution and the QCD

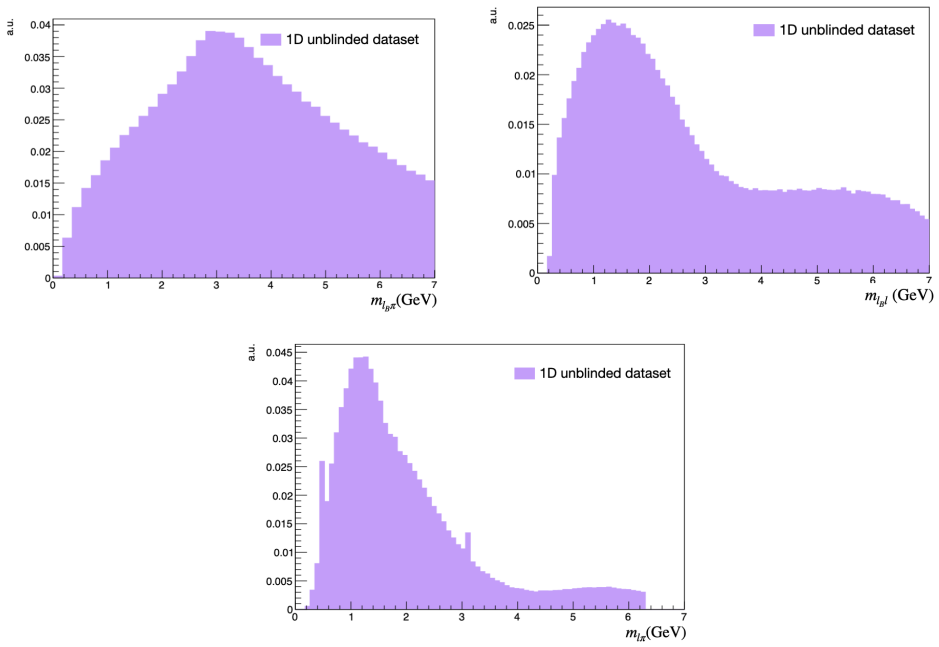


Figure 5.7. Distributions of $m_{\ell\pi}$(left), $m_{\ell_B\pi}$ (center) and $m_{\ell_B\ell}$ (right) for a part of the 1D unblided dataset.

shapes, thus supporting the hypothesis that QCD processes are the main background component for the searched signature.

5.2.2 SM Resonance vetoes

The analysis configures as a search for an invariant mass peak over a smooth background. SM resonances appearing in the invariant mass spectrum of the N candidate $m_{\ell\pi}$ might be misinterpreted as signal, and need to be vetoed. SM resonances appearing in the invariant mass spectra of the other final state particle constitute a known source of background and are therefore vetoed as well.

The distributions of the invariant mass of the N candidate $m_{\ell\pi}$, of the B meson decay lepton and the displaced pion $m_{\ell_B\pi}$ and of the two final state leptons are reported in the signal region and inclusively for the analysis categories for the unblided 1D data sample in Figure 5.7.

The μe invariant mass spectrum does not yield any visible resonance: no known SM resonance decays in a lepton flavor violating pair. No resonance is observed in the $\ell_B\pi$ invariant mass spectrum in Figure 5.7. However, possible resonances yielded by mis-reconstructed $B \to J/\psi + X$ and $B \to \ell D^0$ decays are considered and vetoed in Table 5.1, as the background level at preselection+baseline selection level might mask the peak appearance. These neutral resonances can appear in $m_{\ell_B\pi}$ in the SS region only. Consider a decay chain $B \to \mu^- D^0$ with $D^0 \to K\pi$: the D^0 resonance appears in the $m_{\ell_B\pi}$ invariant spectrum only if the muon from the B decays acts as the displaced ℓ lepton from the heavy neutrino N decay. If a π^+ is misidentified as a lepton ℓ^+, the N candidate has to be $\mu^- K^-$: the N candidate

would not be neutral. On the other hand, if K^- is misidentified with a lepton ℓ^-, $m_{\ell_B\pi}$ resonates to the mass of the D^0, while the candidate will fall in the SS category, as the sign of the other lepton in the signal candidate is negative. Similar considerations apply to the $B \to D^0$ decay conjugates and to the $B \to J/\psi + X$ decay.

Spectrum	resonance	Veto	Relative Lepton Sign
$m_{\ell_B\pi}$	$J/\psi \to \ell\ell$	$\lvert m_{\ell_B\pi} - m_{J/\psi}\rvert > 0.05$ GeV	SS
	$D^0 \to K\pi$	$\lvert m_{\ell_B\pi} - m_{D^0}\rvert > 0.05$ GeV	SS
$m_{\ell\pi}$	$J/\psi \to \ell\ell$	$\lvert m_{\ell_\pi} - m_{J/\psi}\rvert > 0.05$ GeV	OS,SS
	$\Psi(2S) \to \ell\ell$	$\lvert m_{\ell\pi} - m_{\Psi(2S)}\rvert > 0.05$ GeV	OS,SS
	$D^0 \to K\pi$	$\lvert m_{\ell\pi} - m_{D^0}\rvert > 0.05$ GeV	OS,SS

Table 5.1. List of SM resonances and corresponding vetoes in the $m_{\ell_B\pi}$ and $m_{\ell\pi}$ invariant mass spectra. The vetoes are applied in all the transverse displacement significance categories, while the relative lepton sign category in which each veto is applied is indicated. $m_{J/\psi}$,m_{D^0} and $m_{\Psi(2S)}$ are the masses of the three resonances as taken from PDG [36]

The D^0 resonance and J/ψ resonances appear in the signal candidate $m_{\ell\pi}$ spectrum in Figure 5.7 and are vetoed according to Table 5.1. The $D^0 \to K\pi$ resonance appear in this spectrum due to the mis-identification of a track as a lepton. The J/ψ resonance presence, on the other hand, results from the mis-reconstruction of a lepton as a track.

5.2.3 Background Shape and Shape Uncertainty

This heavy neutral lepton search is performed in a continuous range [1, m_B] GeV. To retrieve the expected sensitivity over signals of different mass, the $m_{\ell\pi}$ invariant mass spectrum is scanned with the technique of the sliding mass window.

The background yields are extracted from a direct fit on data in a $\pm 10\sigma$ window around each investigated mass hypothesis m_i; the σ value is extracted through the parametrization presented in Section 5.1.

The background shape changes significantly for different mass hypotheses. To identify the best model for the different background windows, the discrete profiling method [23] is used. A set of functional forms is selected as possible background models, and the background shape choice is treated as a discrete parameter and profiled away when performing the maximum likelihood fit for the background. The method allows for the identification of the best fit function for each background window and for the assignment of a systematic uncertainty for the background shape choice, as the discrete parameter that controls the background shape is treated as a nuisance.

Four families of functions are considered as possible background models:

- Bernstein polynomials:

$$\begin{aligned} \text{bern}_n(x) &= N \textstyle\sum_{\nu=0}^{n} \beta_\nu b_{\nu,n}(x) \\ \text{with } \; b_{\nu,n}(x) &= \tbinom{n}{\nu} x^\nu (1-x)^{n-\nu}, \;\; \nu = 0, 1, .., n \end{aligned} \tag{5.11}$$

  with n free parameters;

- Exponential functions:

$$\exp_n(x) = N \sum_{i=1}^{n} \beta_i e^{\alpha_i x} \tag{5.12}$$

  with $2n$ free parameters (n >1);

- Laurent series:

$$\mathrm{lau}_n(x) = N \sum_{i=1}^{n} \beta_i x^{\left(-4+\sum_{j=1}^{i}(-1)^j(j-1)\right)} \tag{5.13}$$

  with n free parameters;

- Power-law functions:

$$\mathrm{pow}_n(x) = N \sum_{i=1}^{n} \beta_i x^{\alpha_i} \tag{5.14}$$

  with $2n$ free parameters (n >1).

Each family of functions contains multiple polynomials which can be categorized by their order n. To select the polynomial order that, for each family, is in better agreement with the investigated background, the F-test [30] technique is used.

- For each function of order N, define the negative log-likelihood NLL as

$$NLL_{\mathrm{N}} = -ln\mathcal{L} + 0.5N_{\mathrm{par}} \tag{5.15}$$

  , with $\mathcal{L}$ the likelihood and $\mathrm{N_{parameters}}$ the number of parameters describing the function of order N.

- Compute ΔNLL = 2($\mathrm{NLL_N}$ - $\mathrm{NLL_{N+1}}$). According to Wilk's theorem [63], this variable behaves as a χ^2 with one degree of freedom.

- Compute the associated probability as $\mathrm{p^F} = \mathrm{p}(\chi^2 > \Delta\mathrm{NLL})$

- If $\mathrm{p^F} < 0.05$, the function of order N+1 is supported by the data, and the function with order N+2 is going to be tested. Otherwise, the F-test routine will return the order N+1 as the maximum order function supported by the data.

Once the maximum order $\mathrm{N_{max}}$ of the polynomial has been found in each family, a subset of functions, labeled envelope, is built for the background shape modeling. A function of order $\in [1, \mathrm{N_{max}}]$ enters the envelope if it meets either of the following sets of conditions:

1. The function is of order $\mathrm{N_{max}}$ and $\mathrm{p^F}$ <0.1
2. The function is of order less than $\mathrm{N_{max}}$ and $\mathrm{p^F}$ <0.1 and goodness of fit probability > 0.01

The conditions for functions to enter the envelope are kept very loose, as the addition of less performant functions to the envelope does not impact the systematic uncertainty deriving from the envelope.

A closure test on the goodness of the envelope resulting from the F-test can be performed by computing the value of the (double) negative log likelihood, $2 \cdot \text{Log}L + c = 2(\Delta NLL + NLL_0 + NLL)$ over a range of expected signal strengths r, for each function entering the envelope, and the envelope itself. An example of an envelope fit and of the corresponding $2(\Delta NLL + NLL + NLL_0)$ distributions are reported in Figure 5.8 in a background window around a mass hypotheses of 3 GeV, in an arbitrary category with $L_{xy} < 1$, with odd lepton sign charges(OS).

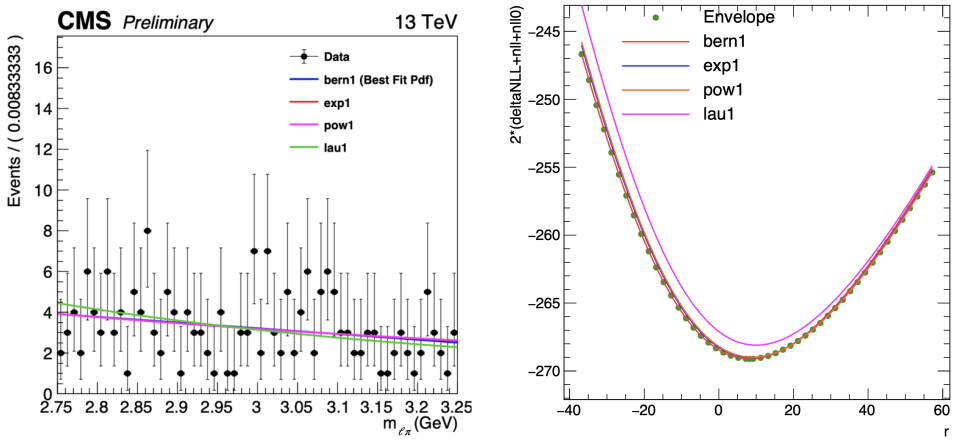


Figure 5.8. Example of an envelope fit (left) and of the corresponding $2(\Delta NLL + NLL + NLL_0)$ distributions (right) in a background window around a mass hypotheses of 3 GeV, for $L_{xy} < 1$ OS category. The left plot shows the different envelope functions over the background they have been fitted to. The data distribution is in an arbitrary category and with a different selection with respect to the one finally used in the analysis, and no veto applied In the right plot, the $2\text{Log}L$ scan of the likelihoods for the envelope and its functions. The function labels indicate the family and order of each background shape: "bern" corresponds to a Bernstein polynomial with $n = 1$,"exp" corresponds to an exponential function of order $n = 1$, "pow" stands for a power-law function of order $n = 1$ and finally "lau1" corresponds to a laurent series of order $n = 1$.

The left plot in Figure 5.8 shows the fits to data for the four functions entering the envelope over the data background they have been fitted to. The envelope functions are a Bernstein polynomial ("bern1"), an exponential function ("exp1"), a power-law function ("pow1") and a laurent series ("lau1"), all of order 1; the best fit function from the F-test results a Bernstein polynomial of first grade. The $2 \cdot \text{Log}L$ behavior for the four functions and the whole envelope is reported in the right plot as a function of the expected signal strength. The envelope picks the function with the best likelihood at each scanned signal strength, while the single different functions maximize the likelihood in different signal strength subranges. The likelihood scan shows that the best fit function correctly minimizes the $2 \cdot \text{Log}L$.

5.3 Expected Yields

A summary of the expected yields per analysis category and for several signal hypotheses is reported.

Mass (GeV)	Category	Background yields	Signal yields			
			0.1 mm	10 mm	1000 mm	10000 mm
1.0			5.4e+00	5.4e-02	5.4e-04	5.4e-05
	L_{xy}/σ <50, OS	2.00E+02	-	3.69E+04	9.87E-01	6.46E-03
	50<L_{xy}/σ <150, OS	3.39E+02	-	1.26E+06	2.79E+01	1.98E-01
	L_{xy}/σ >150, OS	4.24E+01	-	4.42E+06	1.14E+02	7.55E-01
	L_{xy}/σ <50, SS	3.19E+02	-	1.83E+05	4.56E+00	2.96E-02
	50<L_{xy}/σ <150, SS	9.66E+01	-	1.52E+06	3.19E+01	2.32E-01
	L_{xy}/σ >150, SS	5.42E+01	-	7.43E+06	1.91E+02	1.29E+00
1.5			7.1e-01	7.1e-03	7.1e-05	7.1e-06
	L_{xy}/σ <50, OS	9.49E+01	-	1.33E+02	1.93E-02	-
	50<L_{xy}/σ <150, OS	1.85E+02	-	3.76E+03	7.96E-01	-
	L_{xy}/σ >150, OS	7.79E+01	-	4.15E+04	1.06E+01	-
	L_{xy}/σ <50, SS	2.29E+02	-	5.09E+02	1.47E-01	-
	50<L_{xy}/σ <150, SS	1.19E+02	-	5.77E+03	1.18E+00	-
	L_{xy}/σ >150, SS	7.12E+01	-	5.52E+04	1.42E+01	-
2.0			1.7e-01	1.7e-03	1.7e-05	1.7e-06
	L_{xy}/σ <50, OS	2.88E+01	-	1.87E+00	-	-
	50<L_{xy}/σ <150, OS	1.07E+02	-	7.19E+01	-	-
	L_{xy}/σ >150, OS	7.96E+01	-	1.33E+03	-	-
	L_{xy}/σ <50, SS	1.07E+02	-	6.50E+00	-	-
	50<L_{xy}/σ <150, SS	7.63E+01	-	1.04E+02	-	-
	L_{xy}/σ >150, SS	3.56E+01	-	1.51E+03	-	-
3.0			2.2e-02	2.2e-04	2.2e-06	2.2e-07
	L_{xy}/σ <50, OS	4.12E+02	-	-	-	-
	50<L_{xy}/σ <150, OS	2.73E+02	-	1.28E+00	-	-
	L_{xy}/σ >150, OS	7.79E+01	-	1.10E+01	-	-
	L_{xy}/σ <50, SS	6.73E+02	-	-	-	-
	50<L_{xy}/σ <150, SS	1.46E+02	-	1.12E+00	-	-
	L_{xy}/σ >150, SS	5.08E+01	-	1.26E+01	-	-
4.5			2.9e-03	2.9e-05	2.9e-07	2.9e-08
	L_{xy}/σ <50, OS	9.83E+02	7.48E+00	1.08E-02	-	-
	50<L_{xy}/σ <150, OS	2.08E+02	3.49E+00	3.15E-02	-	-
	L_{xy}/σ >150, OS	6.10E+01	1.75E-01	1.54E-01	-	-
	L_{xy}/σ <50, SS	5.66E+02	7.24E+00	1.09E-02	-	-
	50<L_{xy}/σ <150, SS	1.54E+02	3.72E+00	3.03E-02	-	-
	L_{xy}/σ >150, SS	3.39E+01	2.20E-01	1.65E-01	-	-

Table 5.2. Signal and background yields in the electron channel for the mass hypotheses of 1, 1.5, 2, 3 and 4.5 GeV in the different categories for pNN score higher than 0.99. The background yields are computed in the 2σ window around the mass hypothesis, using 4.91 fb^{-1} of the B-Parking dataset and projected to 41.6 fb^{-1}. The signal yields are given for several lifetime points for a luminosity of 41.6 fb^{-1}. The dash symbol "-" denotes that the yields have not been computed.

The signal and background distributions as derived with the strategies described in the previous sections are reported in Figure 5.9 for a signal hypothesis of mass 2 GeV and lifetime 100 mm and for data in the 1D unblinded dataset. The signal yields are normalized to the integrated luminosity of 4.91 fb^{-1} of the 1D unblinded dataset.

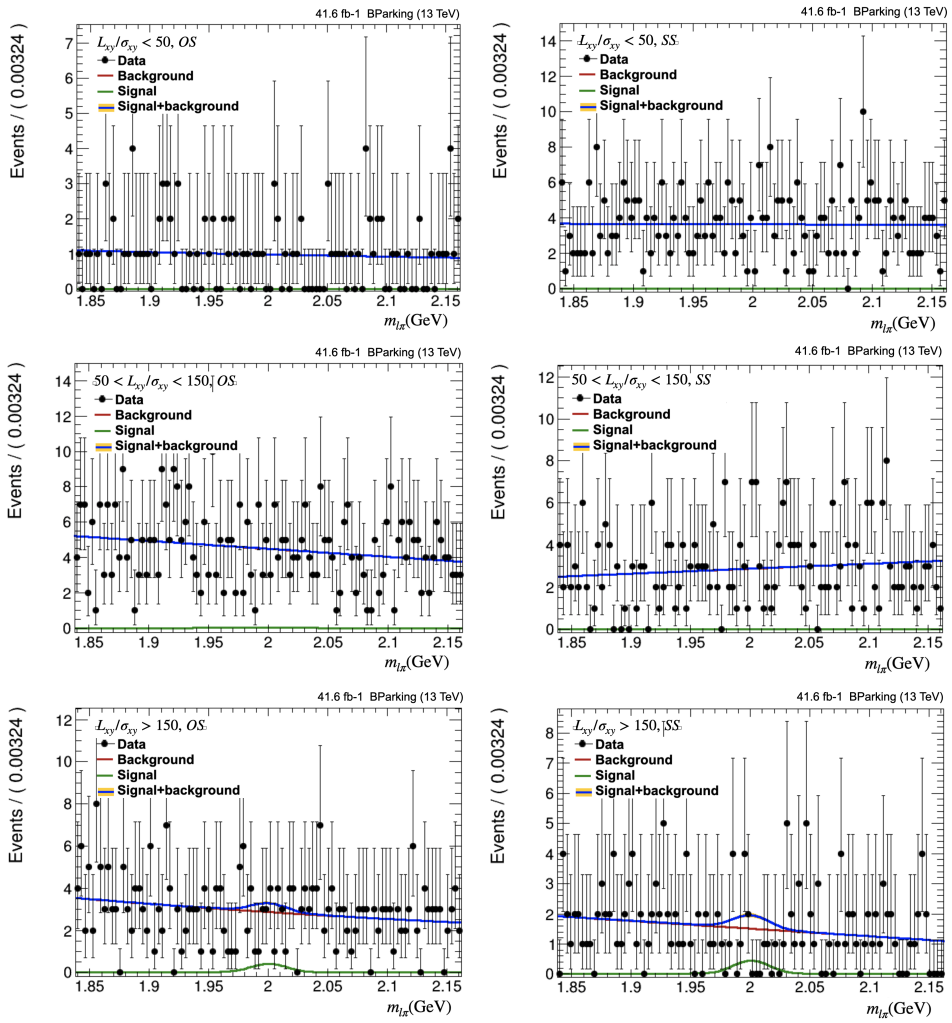


Figure 5.9. Example of data distributions (markers) and signal (green), background (red) and signal+background models (blue) in the $\mu e\pi$ channel for a signal hypothesis of mass 2 GeV and lifetime 300 mm, for the different categories, for pNN score > 0.99 and yields projected to the full B-Parking dataset luminosity. The top plots correspond to the $L_{xy}/\sigma_{L_{xy}} < 50$ OS (left) and SS (right) categories, the central plots represent the $50 < L_{xy}/\sigma_{L_{xy}} < 150$ OS (left) and SS (right) categories and the bottom plots show the signal and background distributions for the $L_{xy}/\sigma_{L_{xy}} < 50$ OS (left) and SS (right) categories. The background models the best fit function as extracted from the discrete profiling procedure described in Section 5.2.3 and the signal shape is the Double Crystal Ball function fit to the signal in the examined category, as reported in Section 5.1.

Chapter 6

Systematic Uncertainties

Possible sources of systematic uncertainties associated with this search are detailed in this Chapter. As the background yields are estimated from a direct fit to data, no additional uncertainties are considered with respect to the ones introduced through the discrete profiling method in Section 5.2.1. The discussion will therefore cover the sources of systematic uncertainty affecting signal: possible sources of systematic uncertainty over the signal efficiency are:

- trigger efficiency;
- track reconstruction;
- muon identification;
- electron identification;
- signal normalization;
- signal selection.

A summary of the sources of systematic uncertainties is reported in Table 6.1, together with their estimated impact on the signal efficiency. For each source of uncertainty, an indication of the underlying distribution and of whether the uncertainty is correlated over the different analysis categories is reported.

Source	Value	Distribution	Correlation
Signal selection	5-10%	log-normal	uncorrelated
Signal normalisation	20%	log-normal	correlated
Tracking scale factors	5%	log-normal	correlated
Trigger scale factors	5%	log-normal	correlated
Muon identification scale factors	1%	log-normal	correlated
Electron identification scale factors	3%	log-normal	correlated

Table 6.1. Summary of the systematic uncertainties applied on the signal. The underlying distribution of the nuisance parameter and an indication of whether the nuisance parameters are treated as correlated across the different categories are reported.

A description of each systematic uncertainty study is reported in the next section. All the sources of uncertainty but the signal normalization are related to the comparison of simulated samples and data under a selection or identification criteria. Corrections to the simulated samples to match the data behavior are discussed and applied whenever possible, and the residual discrepancies are treated as systematic uncertainties.

6.1 Trigger Efficiency

The B-Parking dataset, as highlighted in Section 2.2.4, has been collected thanks to a set of HLT single muon triggers with different thresholds in transverse momentum and impact parameter significance. Simulated signal samples, on the other hand, have been generated as described in Section 4.1.2 with a single minimum transverse momentum requirement for the signal muon ($p_T > 6.8$ GeV) and no constraint on the impact parameter significance. Trigger efficiencies in data and simulated samples are therefore different.

A set of corrections "Scale Factors (SF)"$=\varepsilon_{data}/\varepsilon_{gen}$ are computed in bins of the muon transverse momentum p_T and transverse impact parameter significance $d_{xy}/\sigma_{d_{xy}}$. The simulated samples and data trigger efficiencies are computed using the $B \to K^* J/\psi(\to \mu^+\mu^-)$ SM resonance, using the"tag&probe" technique, which will be described more in detail in Section 6.4. The trigger SFs in bins of muon transverse momentum p_T and transverse impact parameter significance $d_{xy}/\sigma_{d_{xy}}$ are reported in Figure 6.1

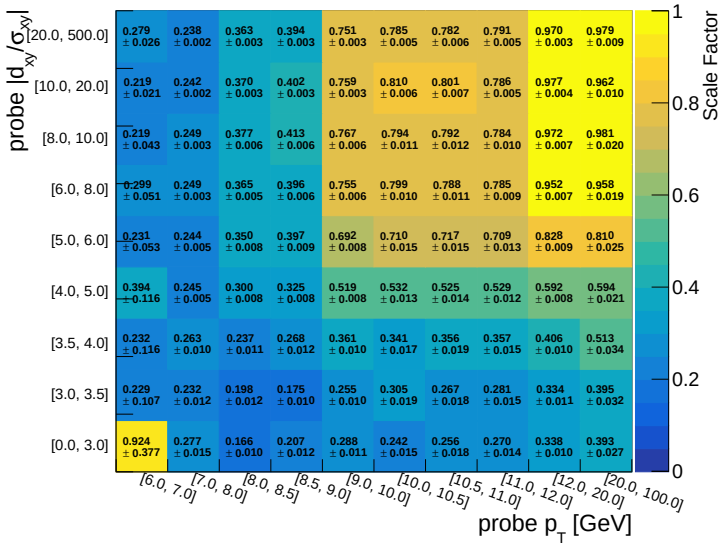


Figure 6.1. Trigger scale factors in bins of the probe muon p_T and $d_{xy}/\sigma_{d_{xy}}$.

Systematic uncertainties over trigger efficiency corrections are evaluated by varying their value by $\pm 1\sigma$. The resulting signal yields are compared with the values obtained with the central value correction applied, for each analysis category and

with the full analysis selection applied. The average of the difference between the nominal signal yield value and the $\pm 1\sigma$ ones is taken as systematic uncertainty. The systematic uncertainty results less than 5% in all the analysis categories. A 5% systematic uncertainty is therefore assigned for trigger efficiency corrections in all the analysis categories.

6.2 Track Reconstruction Efficiency

The presence of a displaced vertex in this search signature represents a challenge in terms of reconstruction efficiency, as pointed out in Section 4.2. In particular, displaced related quantities are prone to mismodeling in simulated samples, as they represent fairly new signatures for CMS physics and some material and detector effects might lack proper representation. Moreover, vertex reconstruction fails more for displaced tracks, as less precise information on the track direction can be retrieved by the tracker layer deposits.

To evaluate the track reconstruction efficiency on displaced signature the procedure used in a search for heavy neutral leptons in W boson decays [18] in CMS is used, and a 5% systematic uncertainty for track reconstruction efficiency is applied on all the analysis categories and on the inclusive displacement range.

6.3 Muon Identification

Muon identification efficiencies for the criteria defined in Section 4.3.2 are compared in simulated samples and data and corrections are derived to match the simulated samples identification efficiency to the one measured in data. Systematic uncertainty over the efficiency corrections is evaluated analogously to the trigger efficiency systematic uncertainty in Section 6.1, by computing the impact of $\pm 1\sigma$ variations on the identification efficiency correction. The resulting systematic uncertainty on the signal efficiency for muon identification is 1% for all the analysis categories.

6.4 Electron Identification

In this section the efficiency of the ID criterium presented in Section 4.3.2 for electron candidates is computed and compared in data and simulated samples. This si the first study holding a comparison of the PF MVA ID behavior in simulated samples and data.

The ID efficiencies are computed for both data and simulated samples using the “tag&probe” method. The “tag&probe” approach consists in the construction of di-lepton candidates in a SM resonance invariant mass range, using one lepton with very tight ID and quality requirements, the “tag” lepton, and a "probe" lepton which may or may not pass the ID criterion that has to be checked. The di-lepton candidates for which the “probe” lepton passes the ID criterion define the “pass” region, while the ones not fulfilling the ID requirement fall in the fail region. The ID efficiency is then computed as:

$$\varepsilon = \frac{N_{pass}}{N_{pass} + N_{fail}} \tag{6.1}$$

Standard studies for electron identification efficiencies are carried out on the Z resonance, yet the B-Parking dataset holds low energy objects, and therefore the statistics available to compute the ID efficiencies around the Z resonance would be extremely low. For this reason, the tag&probe method in this study is applied on to the J/ψ resonance. This is the first electron ID tag&probe study in CMS using the J/ψ resonance as SM candle for efficiency computation.

A $B \to KJ/\psi \to Ke^+e^-$ sample is used for the simulated samples efficiency computation, while the full B-Parking dataset, as reported in Table 4.2 in section 4.1, is used as input for data.

The di-electron candidate is extracted from a fully reconstructed $B \to KJ/\psi \to Ke^+e^-$ candidate. The J/ψ candidate is reconstructed by fitting the electron tracks to a common vertex and then coupled to a track for full B meson candidate construction. The B meson candidate is therefore built through the combination of the Ke^+e^- final state particles. The $B \to KJ/\psi \to Ke^+e^-$ does not carry a muon and therefore lies on the "probe-side" of a $B\bar{B}$ B-Parking event.

All the reconstructed candidates per event are considered for this study. A selection applied on the B candidates for background rejection in the J/ψ kinematic region:

- Triggering muon in the event satisfying a B-Parking HLT path with $p_T > 9$ GeV and $IP_{sig} > 6$;
- B meson $p_T > 5$ GeV;
- vertex probability for the J/Psi fitted vertex > 0.1;
- cosine for the J/ψ fitted vertex $\cos(\theta) > 0.99$;
- B mass $m_B \in [5, 5.4]$ GeV.

The tag electron requirements are:

- PF electron with PF MVA ID > 3 **or** LowPt electron with MVA ID > 5;
- $p_T > 7$ GeV only for barrel electrons.

The efficiencies are computed in p_T and η bins:

- $p_T \in [2, 5, \infty]$ GeV - same binning as the one used for PF MVA ID training;
- $|\eta| \in [0, 1.442, 1.56, 2.5]$.

The function chosen to fit the resonance shape are:

- Double Crystal Ball for the J/ψ resonance;
- Exponential function for the background.

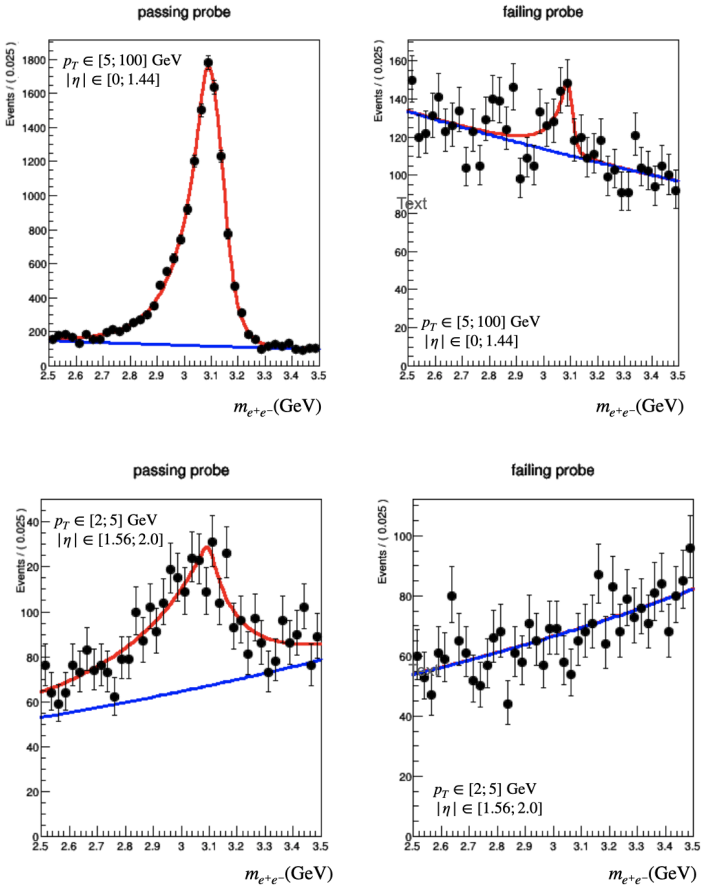


Figure 6.2. Example of fits on the J/ψ resonance. Top: $p_T \in [5; 100]$ GeV and $|\eta| \in [0; 1.442]$. Bottom: $p_T \in [2; 5]$ GeV and $|\eta| \in [1.56; 2.5]$.

Systematics uncertainties for the background fit function choice are assigned by using a first grade polynomial in place of the exponential function.

In Figure 6.2 examples of the pass and fail regions for the designed ID cut are reported for different p_T and $|\eta|$ bins.

The electron identification efficiencies in data and simulated samples, and their ratios are reported in Figure 6.3, as a function of the electron transverse momentum and in pseudorapidity $|\eta|$ bins.

The data and simulation efficiencies are all above 95% for the barrel region, while the data/simulated samples ratios are all within 2% from unity.

No J/ψ peak is visible in the endcap bins fail region, therefore the efficiency estimates suffer from larger uncertainties. A possible explanation might lie in the fact that the chosen ID criterium rejects only very badly reconstructed electrons, which do not allow for a proper invariant mass reconstruction in the fail region even when selecting a very clean electron on the tag side.

Because of this, events with $|\eta| > 1.442$ are dropped from the selection for this search. This results in a minimal loss in signal efficiency , as can be seen in Table 6.2,

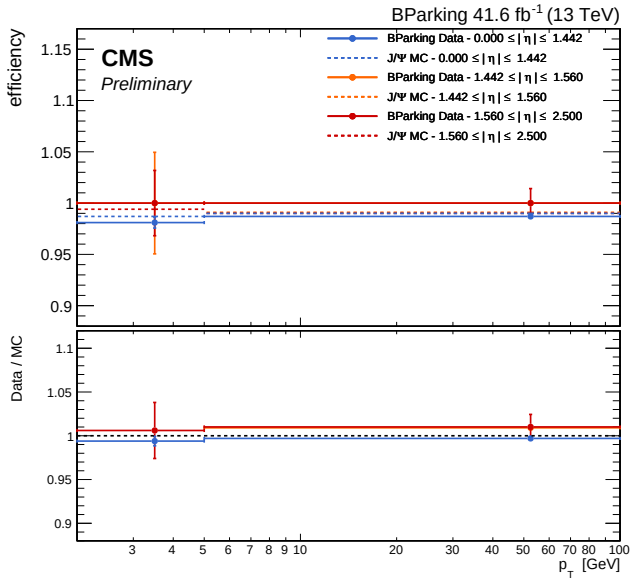


Figure 6.3. Simulated samples and data efficiencies and scale factors for the electron identification criterium PF MVA ID > -3 for two p_T bins: $p_T \in [2;5]$ and $p_T \in [5;100]$.

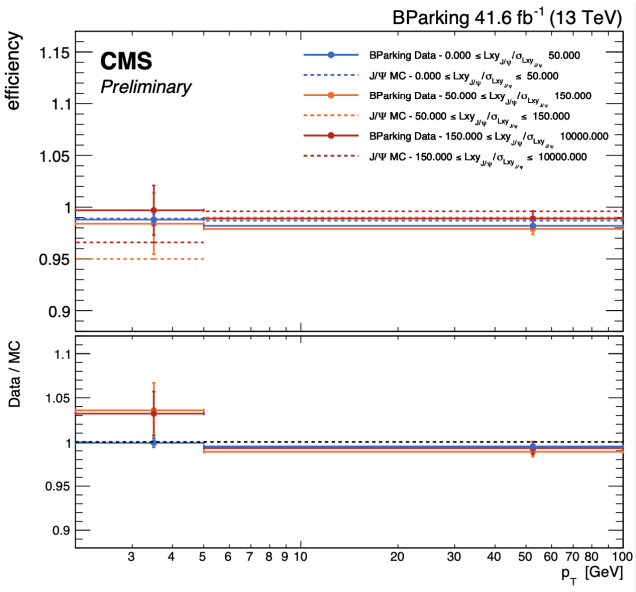


Figure 6.4. Simulation and data efficiencies and scale factors for the electron identification criterium PF MVA ID > -3 for two p_T bins: $p_T \in [2;5]$ and $p_T \in [5;100]$. The different color represent different $L_{xy}/\sigma_{L_{xy}}$ categories.

where the fraction of events falling in the endcap region is shown for some signal hypotheses

The barrel measurement is repeated in categories of transverse displacement

signal	fraction(%) of events in $\|\eta_e\| > 1.442$
$m_N = 1$ GeV, $c\tau = 1000$ mm	1.0
$m_N = 1$ GeV, $c\tau = 100$ mm	1.1
$m_N = 1$ GeV, $c\tau = 10$ mm	1.1
$m_N = 1.5$ GeV, $c\tau = 1000$ mm	1.1
$m_N = 1.5$ GeV, $c\tau = 100$ mm	1.0
$m_N = 1.5$ GeV, $c\tau = 10$ mm	1.0
$m_N = 2$ GeV, $c\tau = 1000$ mm	1.0
$m_N = 2$ GeV, $c\tau = 100$ mm	0.7
$m_N = 2$ GeV, $c\tau = 10$ mm	1.0
$m_N = 3$ GeV, $c\tau = 1000$ mm	2.0
$m_N = 3$ GeV, $c\tau = 100$ mm	1.9
$m_N = 3$ GeV, $c\tau = 10$ mm	2.0
$m_N = 3$ GeV, $c\tau = 1$ mm	2.6
$m_N = 4.5$ GeV, $c\tau = 100$ mm	0.8
$m_N = 4.5$ GeV, $c\tau = 10$ mm	1.0
$m_N = 4.5$ GeV, $c\tau = 1$ mm	1.0
$m_N = 4.5$ GeV, $c\tau = 0.1$ mm	1.0

Table 6.2. Fraction of signal events in the endcaps for centrally generated samples.

significance to ensure that all the analysis categories are covered by the computed identification efficiencies. The efficiencies and data/simulation efficiency ratios for the barrel region only and in transverse displacement significance categories are reported in Figure 6.4.

The derived efficiencies and corrections show no significant deviation with respect to the inclusive transverse displacement significance measurements reported in Figure 6.3, therefore the conclusion reported for the inclusive measurement apply to the single analysis categories as well. A flat systematic of 3% is therefore applied in all the analysis categories; no bin-per-bin correction on the simulated samples is applied.

6.5 Signal Normalization

The number of expected signal events in the μe *channel*, as introduced in Equation 4.3 can be expressed in a simplified expression:

$$N_{\text{sig}}^{\mu e}(m, c\tau) = \sigma_{B\bar{B}} \times \sigma_{\text{HNL}}(\mu e, m, c\tau) \times \mathcal{L} \times \epsilon_{\text{sim}(\mu e, m, c\tau)} \quad (6.2)$$

where $\sigma_{B\bar{B}}$ is the cross section of the $B\bar{B}$ process, σ_{HNL} is the signal cross section, $\mathcal{L}$ is the luminosity and ϵ_{sim} is the simulated samples selection efficiency. The measurement of $\sigma_{B\bar{B}}$ is presented in Section 5.1 and makes use of the luminosity $\mathcal{L}$ in the denominator. The expected signal yields do not depend on the luminosity uncertainty, as $\mathcal{L}$ cancels out. On another note, the other uncertainties on simulated samples efficiency corrections do not cancel out in the expression, as the $\sigma_{B\bar{B}}$ was computed on part of the full B-Parking dataset and in a specific fiducial phase

space. The systematic uncertainty is assigned to the $\sigma_{B\bar{B}}$ by repeating it in the same fiducial region as the one in which a similar measurement [45] has been performed in CMS.

The repetition of this measurement yields $\sigma(\mathrm{B}^{\pm}) = (9.9 \pm 0.4)\,\mu\mathrm{b[stat]}$, while the result quoted in [45] is $\sigma(\mathrm{B}^{\pm}) = (12.5 \pm 2.7)\,\mu\mathrm{b[stat + sys]}$. The two measurements agree within 1σ. The systematic uncertainty of 20% quoted in [45] is applied to the $\sigma_{B\bar{B}}$ measurement.

6.6 Signal Selection

The signal selection output is compared in simulation and data on $B \to KJ/\psi(\to \mu\mu)$ to assess the selection systematic uncertainties. The study compares the signal efficiencies in data and simulation in bins of the pNN score introduced in Section 4.3.4, and computed over the events passing the preselection and baseline selection criteria.

The pNN model used in the analysis does not make use of variables exclusive to the *$\mu\mu$-channel* or *$e\mu$-channel* and relies on the kinematics and displacement properties of the final state particles rather than on their flavor. Since the aim of the study is to compute and compare the selection efficiency in data and simulation, pNN models trained on either the *$\mu\mu$-channel* or *$e\mu$-channel* can be evaluated over the chosen SM resonance regardless of the flavor of the final state leptons. This strategy is therefore applied to evaluate the selection efficiency on both channels, and solves the conundrum of finding a SM standard candle yielding a μe lepton flavor violating final signature to test the *$e\mu$-channel* pNN model.

The study is performed on part 1D unblinded dataset and a $B \to KJ/\psi(\to \mu\mu)$ simulated sample. The events in this study resonate in J/ψ in the $\mu\mu$ invariant spectrum and are vetoed from the signal region in the *$\mu\mu$-channel*. Events are selected in the simulated sample by requiring a match to the generated events with a $B \to KJ/\psi(\to \mu\mu)$ decay, and applying further selection on both data and simulation to isolate the B and J/ψ peaks from the underlying background in data:

- $2.9 < m_{\mu\mu} < 3.3$ GeV;
- OS event;
- $5.15 < m_{\mu\mu K} < 5.4$ GeV.

The reconstructed $m_{\mu\mu K}$ invariant mass is corrected to remove experimental resolution broadening, by building the variable

$$m_{B_corr} = m_{\mu\mu K} - m_{\mu\mu} + m_{J/\psi} \tag{6.3}$$

The s-plot technique [50] is then applied to enhance the signal component over the background in data. The $\mu\mu K$ invariant mass distribution is fitted in categories of transverse displacement significance. The fit is performed using a superposition of two Gaussian distributions for the signal and an exponential function for the background. Each event entering the selection reported above is then assigned a weight based on the ratio of signal and background events present for each bin in

the $m_{\mu\mu K}$ invariant mass spectrum. Thanks to this weight, the data distribution is reweighted in a signal-like distribution. The $m_{\mu\mu K}$ invariant mass spectra together with the fits are reported in Figure 6.5 for the three transverse displacement significance categories of the analysis.

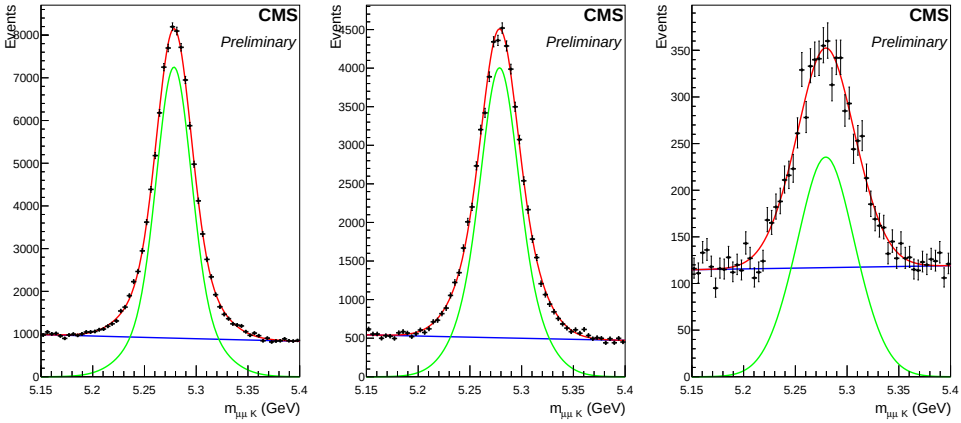


Figure 6.5. Distributions of the $m_{\mu\mu K}$ invariant mass and signal+background fit for s-weights derivation in the $L_{xy}/\sigma_{Lxy} < 50$ (left), $50 < L_{xy}/\sigma_{Lxy} < 150$ (center) and $L_{xy}/\sigma_{Lxy} > 150$ (right) categories. The signal fit is reported in green, the background shape in blue, while the signal+background model is represented in red.

A closure test on the s-plot technique is performed over the $J/\psi \to \mu\mu$ mass peak. The comparison of the distributions of the $J/\psi \to \mu\mu$ events in data and in simulation is reported in Figure 6.6 with the s-weights applied. Good closure between simulated and s-weighted data is observed in all the transverse displacement significance categories.

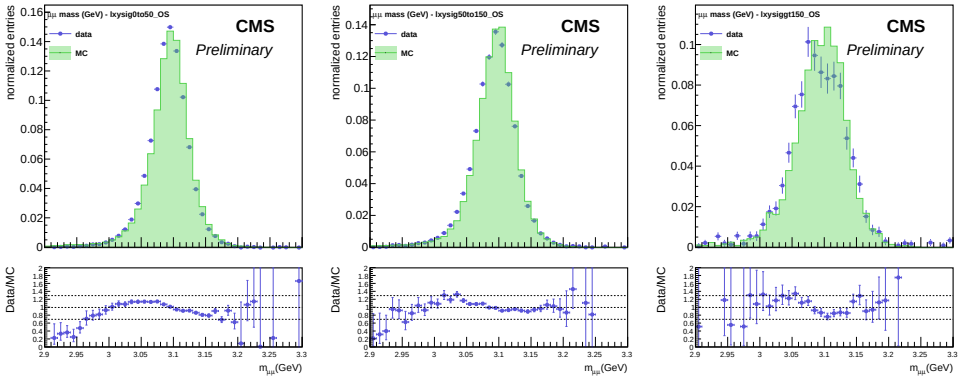


Figure 6.6. Distributions of the $m_{\mu\mu}$ invariant mass in data (markers) and simulation (green) with s-weights applied in the $L_{xy}/\sigma_{Lxy} < 50$ (left), $50 < L_{xy}/\sigma_{Lxy} < 150$ (center) and $L_{xy}/\sigma_{Lxy} > 150$ (right) categories. The lower plot represents the data/simulation ratio.

The comparison of the pNN score distributions for simulation and s-weighted data is reported for the three transverse displacement categories of the analysis in Figure 6.7 for the μe *channel* and Figure 6.8 for the $\mu\mu$ *channel* . The ratio between data and simulation is within ∼30% in all the bins for all the transverse displacement categories.

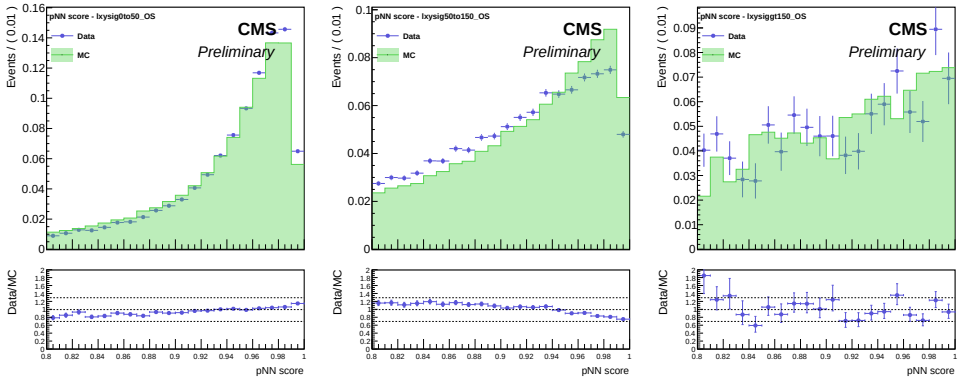


Figure 6.7. Distributions of the pNN score in s-weighted data (markers) and simulation (green) for the μe *channel* trained model in the $L_{xy}/\sigma_{Lxy} < 50$ (left), $50 < L_{xy}/\sigma_{Lxy} < 150$ (center) and $L_{xy}/\sigma_{Lxy} > 150$ (right) categories. The lower plot represents the data/simulation ratio.

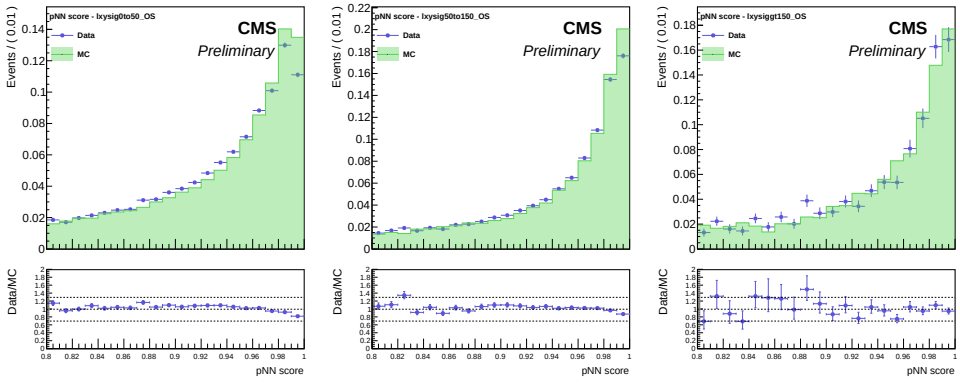


Figure 6.8. Distributions of the pNN score in s-weighted data (markers) and simulation (green) for the $\mu\mu$ *channel* trained model in the $L_{xy}/\sigma_{Lxy} < 50$ (left), $50 < L_{xy}/\sigma_{Lxy} < 150$ (center) and $L_{xy}/\sigma_{Lxy} > 150$ (right) categories. The lower plot represents the data/simulation ratio.

The computation of the systematic uncertainty for the specific pNN working point is derived by integrating the distribution in Figure 6.7 over the pNN values accepted within the selection. The efficiencies for the analysis WP at 0.99 for the two analysis channels are reported in Figure 6.9 in the analysis transverse displacement categories. The ratio between data and simulation efficiencies for different selections is less than 30% in all the analysis transverse displacement significance categories.

Since efficiencies discrepancies between data and simulation vary from category to category, a correction is applied where needed on simulation efficiencies and a systematic uncertainty of the order of the errors of the computed SF is assigned for the signal selection in each category.

The breakdown of the simulation efficiencies correction and assigned systematic uncertainties for thw two analysis channels is reported below:

μe **channel:**

- $L_{xy}/\sigma_{Lxy} < 50$: no correction on the simulation efficiency, 15% selection systematic uncertainty on top of the correction;
- $50 < L_{xy}/\sigma_{Lxy} < 150$: 24% correction on simulation efficiency, 5% selection systematic uncertainty on top of the correction;
- $L_{xy}/\sigma_{Lxy} > 150$: no correction on simulation efficiency, 10% selection systematic uncertainty.

$\mu\mu$ **channel:**

- $L_{xy}/\sigma_{Lxy} < 50$: 18% correction on the simulation efficiency, 5% selection systematic uncertainty on top of the correction;
- $50 < L_{xy}/\sigma_{Lxy} < 150$: 12% correction on simulation efficiency, 5% selection systematic uncertainty on top of the correction;
- $L_{xy}/\sigma_{Lxy} > 150$: no correction on simulation efficiency, 10% selection systematic uncertainty.

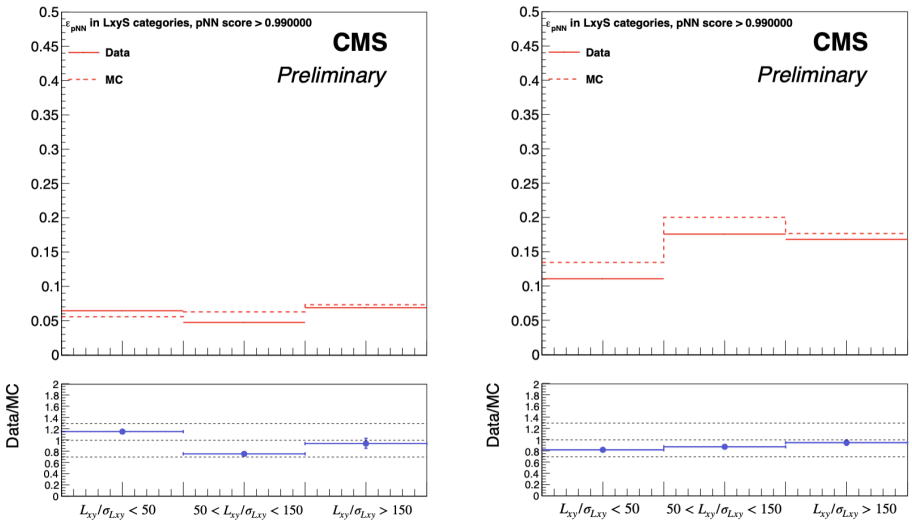


Figure 6.9. Efficiencies for 0.99 pNN working point in Data (full line) and simulation (dashed line) for the μe *channel* (left) and the $\mu\mu$ *channel* (right). The three bins represent the three analysis transverse significance displacement categories. The lower plot represents the data/simulation ratio.

Chapter 7

Results and Interpretation

The statistical methods used to convert the signal and background yields estimation in exclusion limits over the heavy neutrino mass and lifetime hypotheses are presented in this chapter.

Expected and observed limit for the μe *channel* are derived for the full 41.6 fb^{-1} B-Parking integrated luminosity through the techniques discussed in the previous chapters. The limits over the μe *channel* coupling coefficient $|V_\mu \cdot V_e|^2/|V|^2$ are derived with no explicit choice over the flavor coupling scenario.

Expected exclusion limits over the total coupling value $|V|^2$ to SM neutrinos are derived for four different mass hypotheses and 66 sets of different coupling scenarios through the combination of the results derived from the μe *channel* and the parallel $\mu\mu$ *channel* introduced in Section 1.6.

7.1 Statistical Analysis

The limits are computed using the CLs technique [39] with a binned profile likelihood ratio in the asymptotic approximation as the test statistic. Limits are set for each signal hypothesis on the signal strength modifier $\mu = N^{sig}_{@95\text{C.L.}}/N^{sig}_{exp}$, where N^{sig}_{exp} is computed as in Equation 4.3, while the discrete parameter for the background functional shape introduced in Section 5.2.3 and the continuous systematic uncertainties described in Chapter 6 act as nuisance parameters. The signal strength modifier μ is computed for each (m_N,cτ) signal hypothesis through the combination of the fits over the 6 orthogonal analysis categories defined in Section 4.3.3.

The limits are computed using the CLs criterion with a binned profile likelihood ratio [39]. A confidence level of 95% is required for the exclusions in this search. The expected and observed signal strength modifier value is therefore derived by fixing $CLs = 0.05$ and exploiting the relation

$$CL_s(\mu) = \frac{p_\mu}{1 - p_b} \tag{7.1}$$

where p_μ and p_b are the p-values associated respectively with the signal+background and background only hypothesis. The p-values are computed with a binned profile likelihood ratio in the asymptotic approximation [22] as the test statistic.

The expected CL_s limit on μ is evaluated in its central value, assuming $1 - p_b = 0.5$ and equipped with the $\pm 1\sigma$ (68%) and $\pm 2\sigma$ (95%) bands.

7.1.1 Signal Injection Test

A search for a peak over a background distribution heavily relies on the fit procedure to correctly extract the value of a possible observed signal. Tests on the signal strength extraction from the signal+background fit and on the impact of the nuisance parameters over the extracted μ are performed to ensure that no bias is applied by the fitting routine on the extracted signal strength value.

A signal injection test is implemented to verify that the fitting procedure extracts the correct value for arbitrary signal strengths. The test is performed by generating signal+background distributions where the signal yield is scaled by the injected signal strength, and the background is modeled through the F-test best-fit function defined in Section 5.2.3 for each analysis category. A sample of 1000 pseudo-experiments is generated through this method and for each experiment the expected signal strength is extracted as the mean over the 1000 pseudo-experiments and compared to the injected one.

The pull distributions, defined as $(\mu_{ext} - \mu_{inj})/\sigma_{ext}$ are reported for a set of injected signal strengths in Figure 7.1, for a signal hypothesis of mass 2 GeV and lifetime 100 mm, and on the combined analysis categories.

The pull distributions result symmetric for small injected signal strength values, and become slightly asymmetric when injecting higher μ values.

The mean of the extracted expected signal strength of the 1000 pseudo-experiment is reported in Figure 7.2 as a function of the injected signal strength: the fit correctly returns the injected signal strength value.

7.1.2 Nuisance Parameters Impact

The impact of a nuisance parameter θ on a parameter of interest μ is defined as the shift $\Delta\mu$ that is induced as θ shifts by $\pm 1\sigma_\theta$, where σ_θ is the uncertainty over the nuisance parameter after the fit, with all other nuisance parameters profiled as normal. Impacts are an effective measurement of the correlation of the parameter of interest to the different nuisance sources. This correlation is useful to determine which nuisance has the larger effect on the signal strength uncertainty.

The impact of the different nuisance parameters θ on the extracted signal strength is reported in the right column of Figure 7.3, together with the relative difference of the nuisance value before and after the fit (left column) for a signal of mass 2 GeV and lifetime 100 mm and for the combination of the analysis categories.

The nuisance parameter impacts reported in Figure 7.3 show that the selection systematic uncertainty for the high transverse significance SS category has the highest correlation to the expected signal strength modifier. This behavior is expected, as this category is the one yielding the highest sensitivity for the tested signal hypothesis.

The relative difference of the nuisance value before and after the signal strength modifier extraction reported in Figure 7.3 show no significant variation of the fitted nuisance with respect to the pre-fit values: the fit is not relying on large oscillations on the nuisance parameter to converge. The absolute value of the extracted signal strength modifier is therefore not biased by the nuisance parameters, which only contribute to its uncertainty.

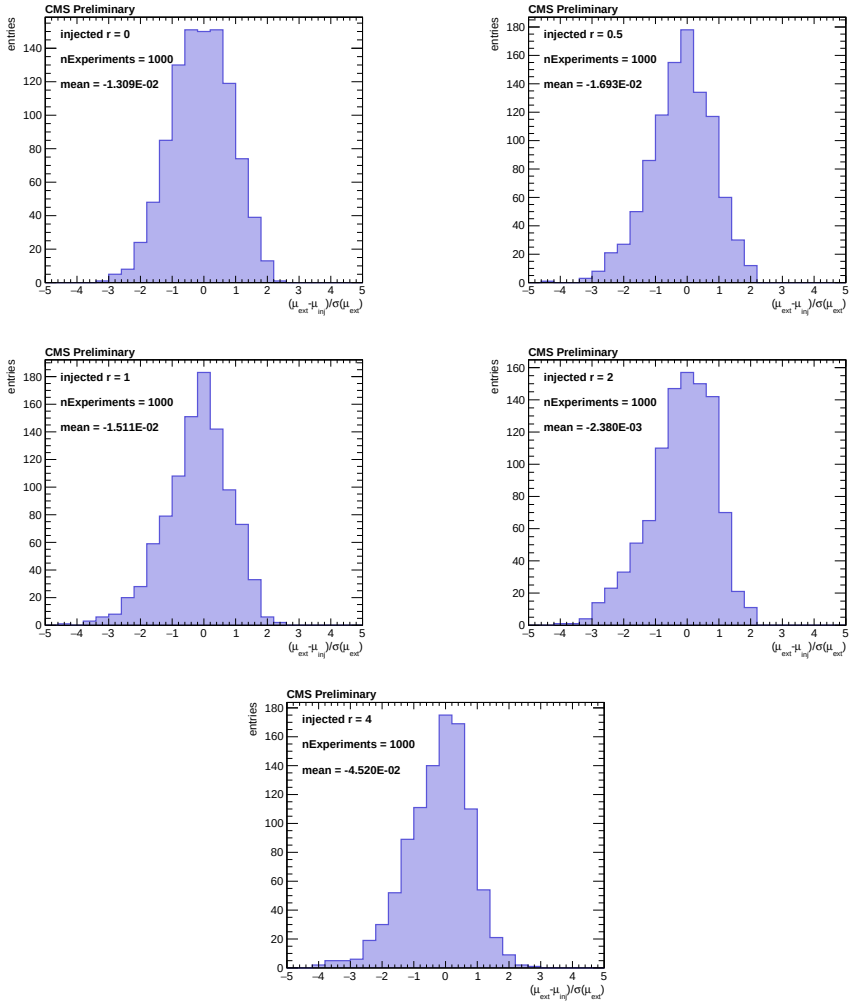


Figure 7.1. Pull distributions in the electron channel extracted from 1000 toys for signals of injected strength 0, 0.5, 1, 2 and 4 (from left to right, top to bottom) for a signal hypothesis of mass 2 GeV and lifetime 100 mm.

7.2 Results on the μe Channel

The expected (dashed line) and observed (full line) limits at 95% confidence level for the signal strength parameter μ are reported, as a function of the coupling $|V_\mu \cdot V_e|^2/|V|^2$, for each simulated mass hypothesis in Figure 7.4. This means that the yields for this channel are computed as in Equation 4.3 without multiplying by the coupling fractions f_μ and f_e. In other words, these limits are "agnostic" with respect to coupling fractions.

For each mass, the limits on μ are computed for all the lifetimes (and couplings) hypotheses obtained through the signal reweighting procedure presented in Section 5.1. The limits are derived using the full 41.6 fb^{-1} B-Parking luminosity.

The plots represent the 4 different tested mass hypotheses of 1 GeV (top right), 1.5 GeV (top left), 2 GeV (bottom right) 3 GeV (bottom left). The exclusion in set

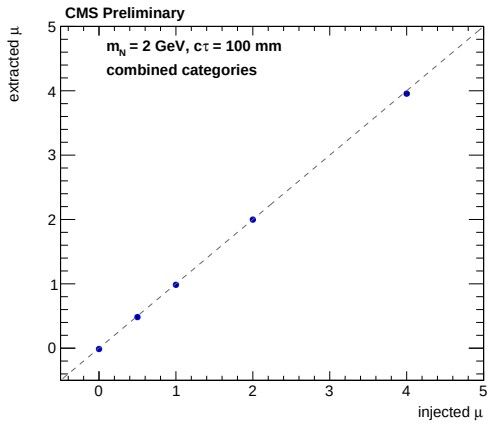


Figure 7.2. Mean value of the extracted signal strength as a function of the injected signal strength for the electron channel and a signal hypothesis of mass 2 GeV and $c\tau$ 100 mm.

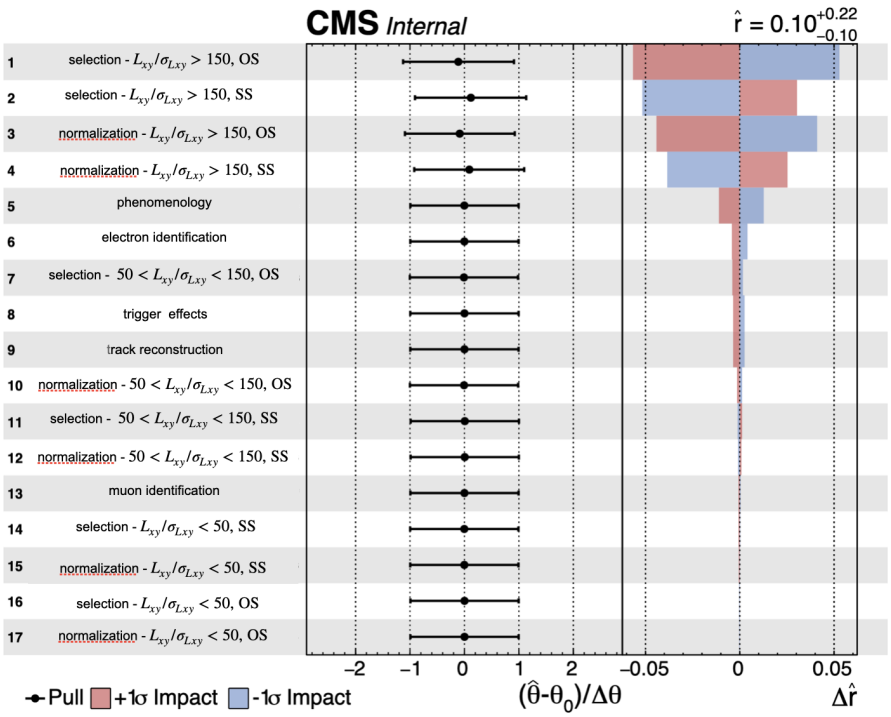


Figure 7.3. Example of an impact plot for a signal hypothesis of mass 2 GeV and lifetime 100 mm in the μe *channel*.

on the $|V_\mu \cdot V_e|^2/|V|^2$ that yields an expected μ value of 1.

The minimum expected and observed excluded coupling as a function of the mass are reported in Figure 7.5 for all the mass hypotheses introduced in Section 4.1.2. No exclusion is reached for masses $m_{HNL} > 3.2GeV$ and no significant variation

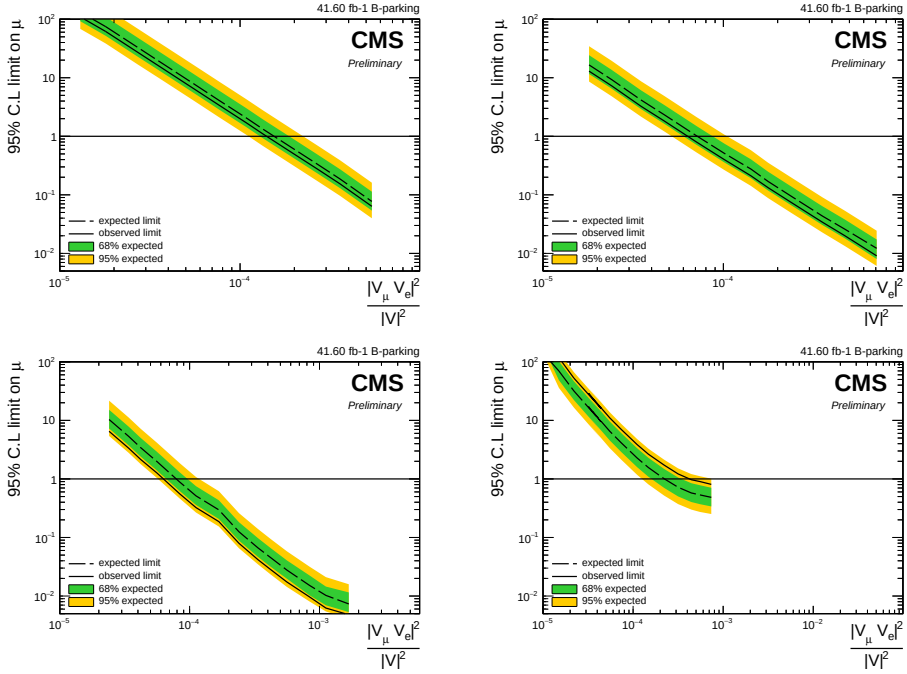


Figure 7.4. Expected (dashed line) and observed (full line) exclusion limits as a function of the $|V_\mu \cdot V_e|^2/|V|^2$ for mass 1 GeV (top left), 1.5 GeV (top right), 2 GeV (center left), 3 GeV (center right).

in the observed limit with respect to the expected limits is visible. The observed sensitivity of the μe *channel* ranges from $\sim 5 \cdot 10^{-5}$ for mass 2.25 GeV to $4 \cdot 10^{-4}$ for mass 3 GeV.

7.3 Channel Combination

The μe *channel* expected results are combined with results from the $\mu\mu$ *channel*. The analysis strategy in the $\mu\mu$ *channel* mirrors the one presented for the μe *channel* in the previous chapters, in terms of event selection and categorization, signal and background yields extraction and statistical analysis. The observed sensitivity for the full B-Parking luminosity on the muon channel ranges between $\sim 3 \cdot 10^{-5}$ for mass 2 GeV to $2 \cdot 10^{-2}$ for mass 4.5 GeV, for a muon dominant coupling scenario ($f_\mu : f_e : f_\tau = 1 : 0 : 0$).

The results on the two channels are combined and interpreted in terms of mixed couplings scenarios. Since the two channels are orthogonal in signature, parallel in analysis strategy and no significant (less than 1%) overlap in the selected events space is measured, the channel combination is carried out considering the two channels as independent signal processes. Expected limits for the combination are derived on the full B-Parking luminosity.

The signal yields are a function of the coupling fractions, as reported in Equation 4.3 and can be therefore expressed, for an arbitrary coupling scenario $f_\mu : f_e : f_\tau$,

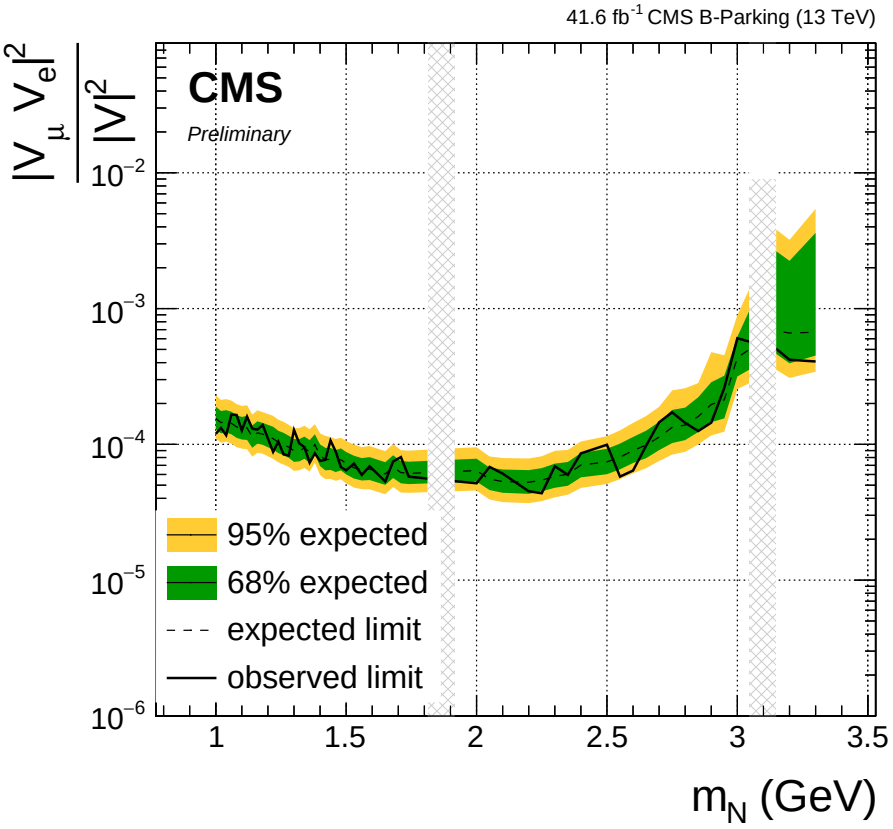


Figure 7.5. Expected and observed exclusion limits on $|V_\mu \cdot V_e|^2/|V|^2$ as a function of the heavy neutrino mass $m_{HNL} \leq 3$ GeV. No coupling scenario is explicitly chosen for these limits. The gray areas represent the vetoes applied for the SM resonance discussed in Section 5.2.

as:

$$\begin{aligned} N_{\mu\mu}^{(\mathrm{f}_\mu:\mathrm{f}_\mathrm{e}:\mathrm{f}_\tau)} &= N_{\mu\mu}^{(1,0,0)} \cdot f_\mu^2 \\ N_{e\mu}^{(\mathrm{f}_\mu:\mathrm{f}_\mathrm{e}:\mathrm{f}_\tau)} &= N_{e\mu}^{\mathrm{agnostic}} \cdot f_\mu f_e, \end{aligned}$$

where $N_{\mu\mu}^{(\mathrm{f}_\mu:\mathrm{f}_\mathrm{e}:\mathrm{f}_\tau)}$ and $N_{e\mu}^{(\mathrm{f}_\mu:\mathrm{f}_\mathrm{e}:\mathrm{f}_\tau)}$ are the expected yields for the probed coupling scenario, f_μ and f_e are the coupling fractions to the muon and electron flavor respectively. $N_{\mu\mu}^{(1,0,0)}$ is the expected yield on the $\mu\mu$ *channel* for a muon dominant coupling scenario, and finally $N_{e\mu}^{\mathrm{agnostic}}$ is the expected yield on the μe *channel* as treated in the previous section, with no explicit choice on the coupling fractions. In this approach, the coupling fractions are assumed to be real numbers [27]. By construction, the sum of the coupling fractions must be equal to 1. When scanning over the fractions with a 0.1 step, this gives 66 coupling scenarios to probe. For each of the different coupling scenarios, the extraction of the expected exclusion limits is repeated. For each mass, the coupling $|V|^2$ at which an exclusion is derived is

reported in the ternary plots shown in Figure 7.6.

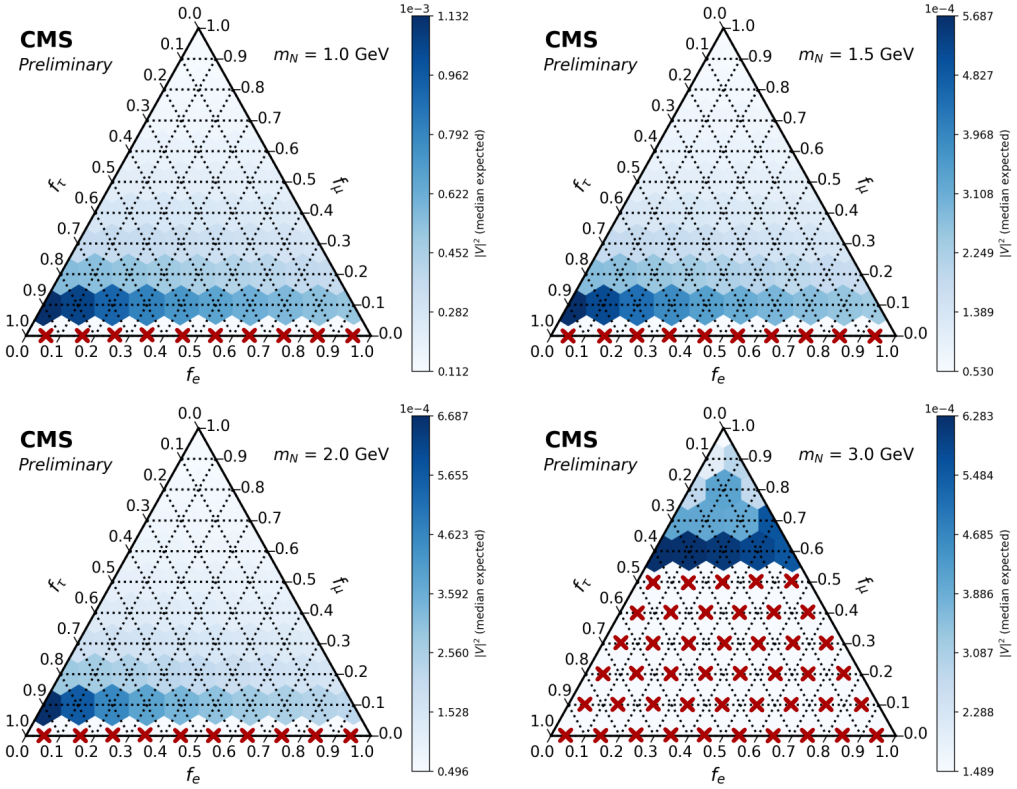


Figure 7.6. Exclusion limit on $|V|^2$ as a function of the coupling scenarios based on 66 sets of couplings (f_μ, f_e, f_τ). The different plots show the performance for mass hypotheses fixed at 1, 1.5, 2 and 3 GeV. The results are extracted with the combined flavor channels. The points towards the basis of the triangle filled in white means that there is not exclusion found for that point. The red cross represents coupling scenarios for which the exclusion is not reached.

The analysis sensitivity is driven by the $\mu\mu$ channel, as the excluded $|V|^2$ becomes smaller the larger the muon coupling fraction for all the probed mass points. Yet, a decrease of the exclusion limit can be observed for fixed f_μ value and growing f_e values. For mass 2 GeV, as an example, and $f_\mu = 0.1$ the exclusion limit value decreases from $\sim 6.3 \cdot 10^{-4}$ when $f_e : f_\tau = 0 : 0.9$ to $\sim 3 \cdot 10^{-4}$ when $f_e : f_\tau = 0.9 : 0$.

Chapter 8

Conclusions

This thesis presents the first search for Majorana Heavy Neutral Leptons in B meson decays with the CMS experiment at CERN. The search exploits data from the proton-proton collisions produced at the CERN LHC at a center of mass energy of 13 TeV. The analysis is made possible by the CMS 2018 B-Parking dataset, a sample of 10^{10} $B\bar{B}$ events with high purity corresponding to 41.6 fb^{-1} of integrated luminosity, recorded with novel trigger and data-taking strategy.

The analysis targets Heavy Majorana Neutral Leptons in the [1;m_B] GeV mass range, with lifetimes ranging between $c\tau = 0.0001$ m and $c\tau = 1$ m. The final state signature is formed by three low-energy objects (a muon, an electron and a pion) two of which stem from the heavy neutrino displaced vertex. The reconstruction and identification of low-energy displaced objects represent one of the main challenges to the analysis, as the standard CMS reconstruction algorithms are optimized for high-energy objects originating from the proton-proton interaction region. On the other hand, as most SM processes yield decay vertexes in the interaction region, the displacement proves as a powerful tool for SM background rejection.

This thesis focuses on final states with an electron. Electrons yield complex signatures in CMS, as they require deposits from two sub-detectors, ECAL and the tracker, to be correctly reconstructed, and lose energy through bremsstrahlung radiation. ECAL energy measurement is crucial for electron reconstruction and undergoes a thorough correction and calibration procedure. The validity of such procedure is tested, for Run2 Data, through the ECAL energy linearity measurement presented in this work. The reconstruction and identification of low-energy electrons represent a challenge for the analysis, which has been overcome through the development of dedicated identification tools.

The search is performed for several Heavy Neutral Lepton mass and lifetime hypotheses, yielding different kinematic and displacement-related properties. The analysis sensitivity is therefore optimized through criteria that maximize the signal efficiency preservation while exploiting the signal displacement features. The analysis is performed in categories of displacement-related quantities; for each category, a parametric deep learning discriminator allows to maintain high performance across a wide range of heavy neutrino masses.

Exclusion limits on the coupling of the Majorana Heavy Neutral Leptons are derived through a sliding window fit over the invariant mass spectrum of the two

displaced final state particles. Observed results are presented for the μe *channel* and for no specific choice of the coupling scenario, on the full B-Parking integrated luminosity. No significant excess is observed in the probed range. The observed sensitivity of the μe *channel* ranges from $5 \cdot 10^{-5}$ for mass 2.25 GeV to $4 \cdot 10^{-4}$ for mass 3 GeV.

Expected results are presented for several mass hypotheses and multiple flavor coupling scenarios through the combination of the $\mu\mu$ and μe *channels*. The analysis reaches an expected sensitivity, expressed as the best 95% exclusion limit on the total coupling $|V|^2$, which ranges between $5 \cdot 10^{-5}$ for a mass hypothesis of 2 GeV to $1.5 \cdot 10^{-4}$ for mass 3 GeV.

Appendix A

LowPt Electrons

A.1 LowPt electrons studies

The CMS 2018 B-Parking dataset, as discussed in Section**??** is equipped with a low energy electron collection. Such collection is included in the BParking dataset as B meson decay products have very soft spectra, with transverse momentum p_T lying in the < 5 GeV region, and the standard PF electron algorithm has null efficiency for electrons with $p_T < 2$ GeV. The LowPt collection extends the transverse momentum range for the electron candidates down to 0.5 GeV in CMS 2018 B-Parking Reco dataset. In this section, an extensive study on the sensitivity performance of signal candidates built with lowPt electrons is presented. The study was carried out on signatures with triggering muon as primary lepton of the decay.

The lowPt collection is built by exploiting the GSFTrack collection, with looser tracks seed requirements, and hit path identified through a Gaussian Smear Filter. The tracks' trajectories are projected to ECAL and the electron candidate is built for tracks that match with an ECAL SuperCluster. The momentum estimates are based on the track info only. The colection is equipped with its own energy regression algorithm and a MVA ID.

A.1.1 Selection

Signal candidates are built using the same strategy reported in Section 4.2. A preselection + baseline selection + signal candidate selection analogous to the one performed on the PF electron channel is implemented. The applied preselection and the baseline selection are identical to the one presented for the electron channel in Section 4.3.1 and Section 4.3.2. The overall preselection efficiencies for signal of mass 3 GeV, $c\tau$ 184 mm is 68.4%, while the background rejection evaluated on a portion of unblinded data is 0.16%. A dedicated multivariate identification algorithm is used for the lowPt electron. It has analogous features to the retraining descrived for the PF MVA ID in Section 3.3. The identification criterium is defined as LowPt MVA ID $>$ -1.

The expected signal for all the available central MC samples and background yields in a 2σ window around 3 GeV for the PF electron and the LowPt channels are reported in FigureA.1.

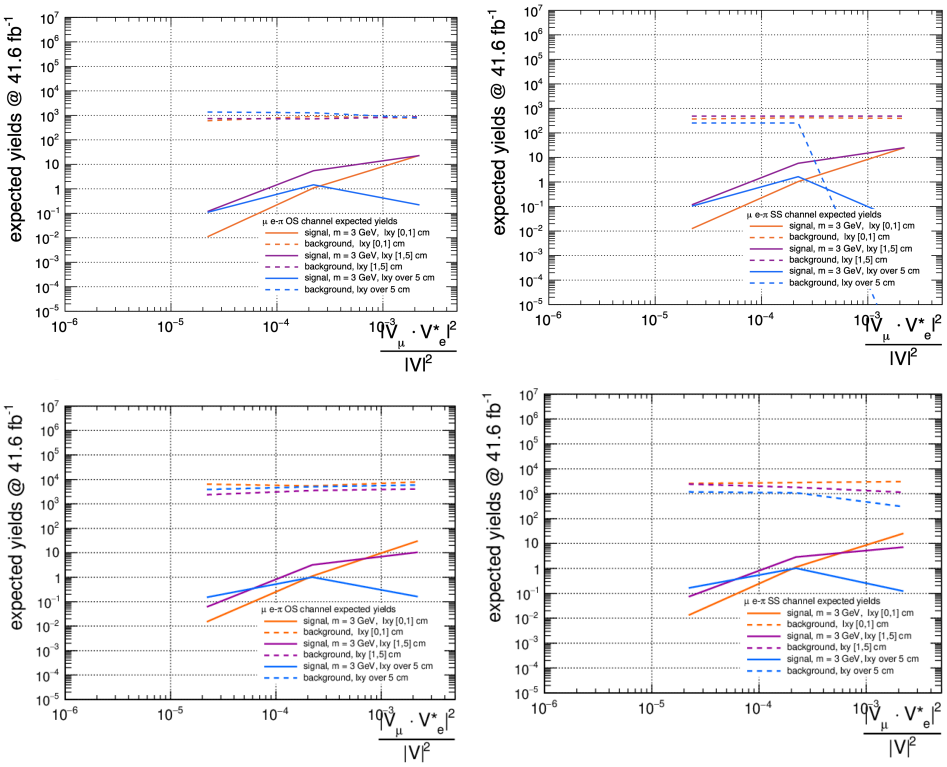


Figure A.1. Expected signal (full line) and background (dashed line) yields for the PF (top plots) and lowPt (bottom plots) channel after preselection+baseline selection, for signals of mass 3 GeV and $c\tau = 100$ mm and the unblinded dataset. The yields are computed in bins of L_{xy} displacement and in the OS (left plots) and SS (right plots) categories. All the yields are projected to the full 2018 B-Parking luminosity.

The yields are computed in three N displacement categories, $L_{xy} < 1$, $1 < L_{xy} < 5$ and $L_{xy} > 5$, in OS and SS categories. The signal yields in the two channels are comparable at preselection level, while the background yields are at least a factor 3 higher in the LowPt channel.

A likelihood discriminator is built for the PF and lowPt channels to further reject the backgrounds. The likelihood is computed for both channels on kinematic and displaced related variables. The selected set of features is reported below:

- the minimum transverse momentum for the HNL daughters
- the value of the LowPt MVA ID or PF MVA ID for the electron used in the signal candidate
- the electron transverse impact parameter significance
- the pion transverse impact parameter significance
- the $cos(\theta)$ for the HNL reconstructed vertex
- the N vertex probability prob(NV)

The likelihood is computed using PDFs extracted, from a signal of mass 3 GeV, $c\tau$ 10 mm for both the PF and lowPt channels. A WP with signal efficiency 75% for likelihood > 0.9, is chosen for both the PF and lowPt channels. The background rate for this WP results ∼2% for the PF channels and ∼3% for the lowPt channel.

A.1.2 Analysis sensitivity with lowPt electrons

A sensitivity study is finally performed on likelihood selected data for both the PF and LowPt channels. The study is performed on central simulated samples with mass 3 GeV, computing limits in the counting experiment approach. The performance for the PF and LowPt channel are respectively reported in the left and right plots of FigureA.2.

The LowPt channel performance is worst with respect to the PF channel by at least a factor 5 all over the probed coupling spectrum. Combination of such sensitivity performance over the two channel would not yield a significant gain. The expected signal yields at preselection level and likelihood performances are similar in the two channels, while the background yields at preselection are a factor >3 higher in the LowPt channel with respect to the PF one. A tighter likelihood WP in the lowPT channel would spoil the signal efficiency while decreasing the background rates. The resulting sensitivity would again under-perform with respect to the PF channel, leading to negligible gains in the combination. It was therefore decided to not include this collection in the analysis.

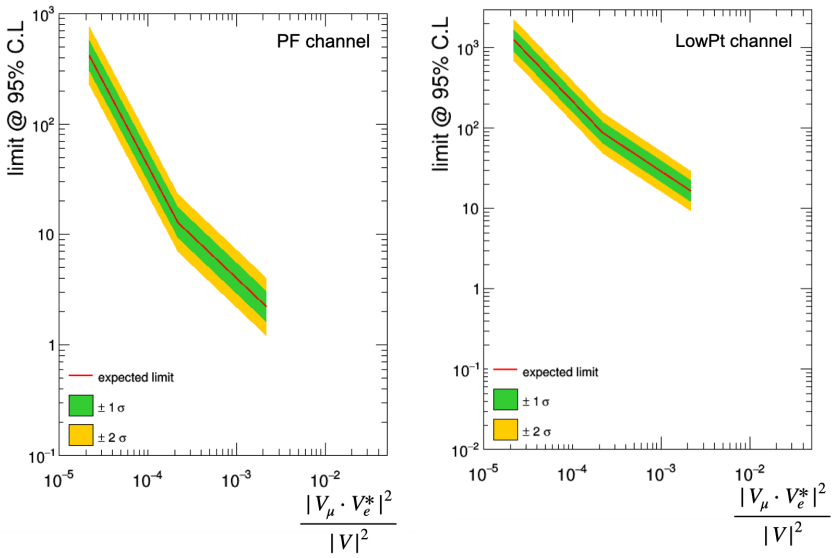


Figure A.2. Expected limits obtained with the counting experiment approach on signals of mass 3 GeV and $c\tau$ =[10,100,1000] mm. Limits obtained unsing $\sim$ 2.91 fb^{-1} of BParking statistics and projected to full luminosity.

Appendix A

pNN performance plots

In this appendix performance plots for the pNN trainings in the different analysis categories are reported.

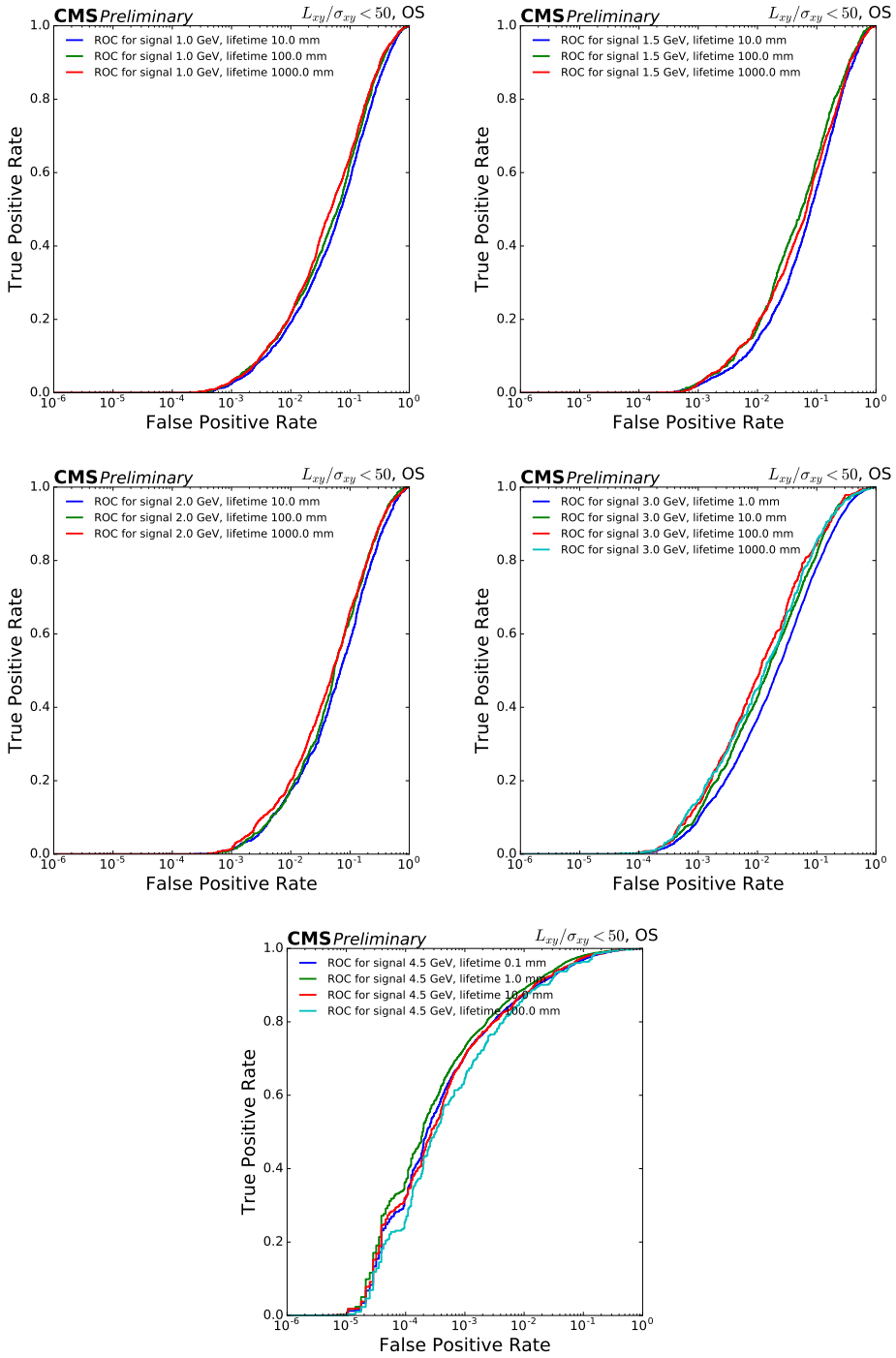


Figure A.1. ROC curve for the mass hypotheses used in the training in the low L_{xy}/σ_{Lxy} OS category for the $\mu e\pi$ channel.

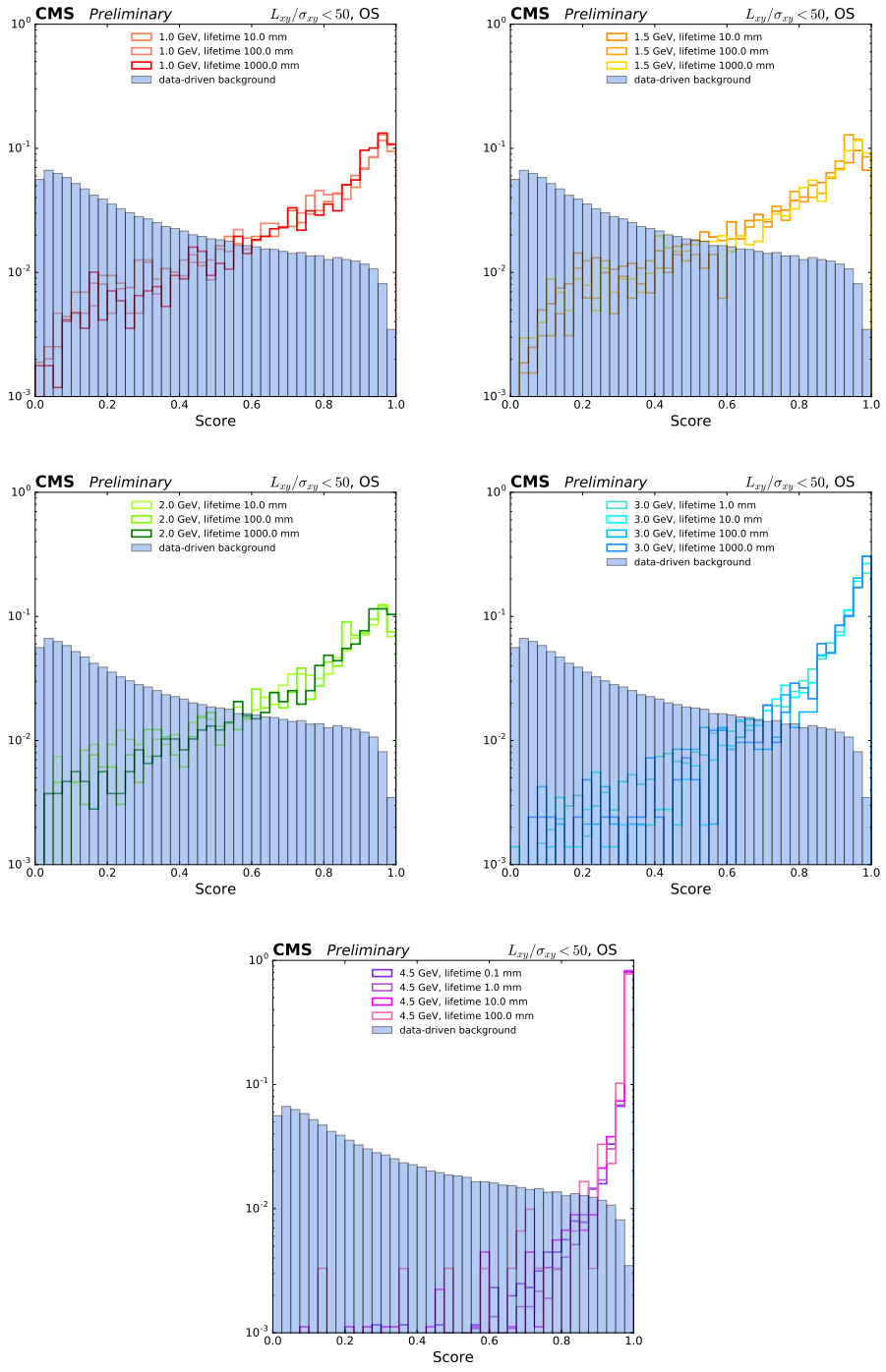


Figure A.2. Score distribution for the mass hypotheses used in the training in the low $L_{xy}/\sigma_{L_{xy}}$ OS category for the $\mu e\pi$ channel.

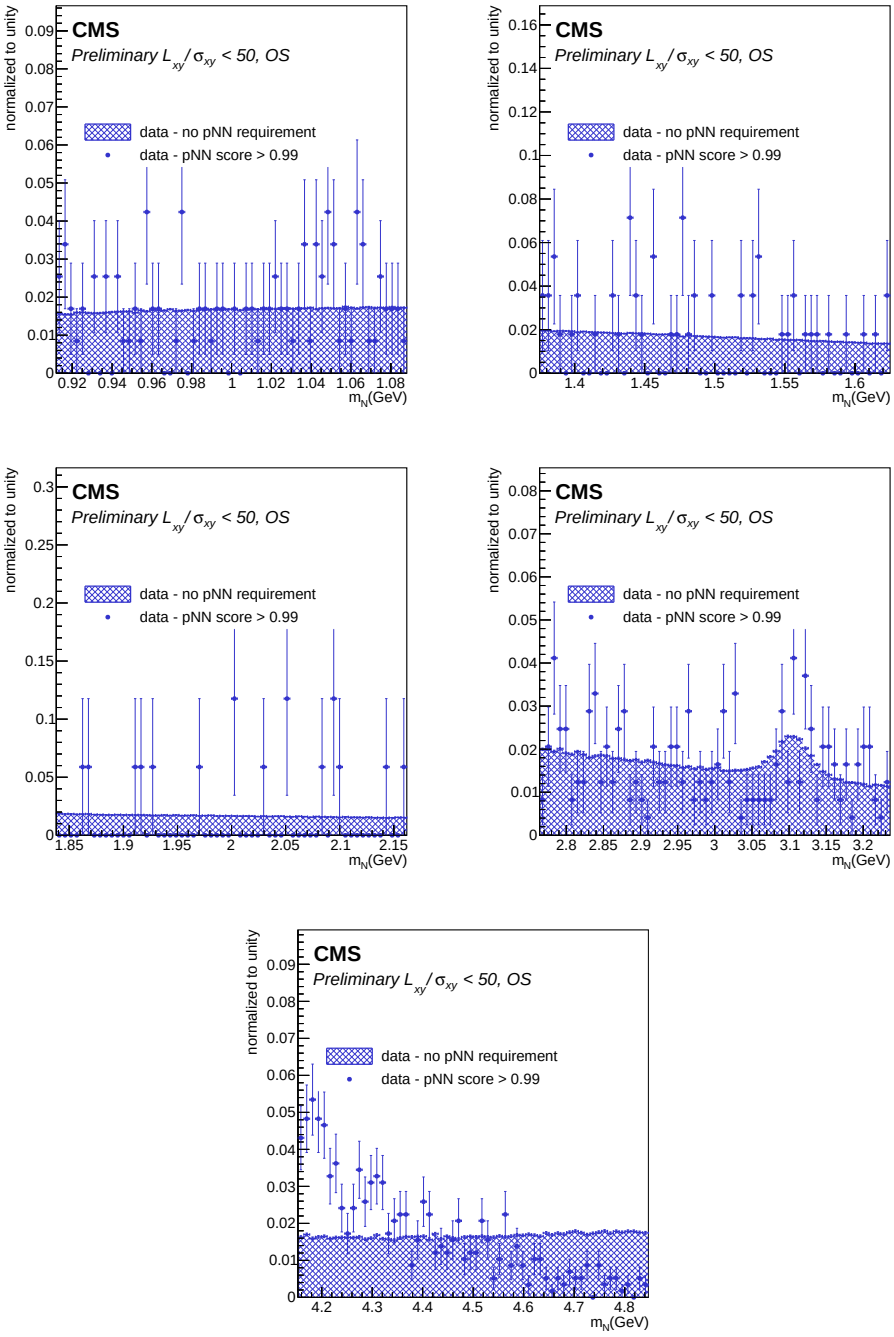


Figure A.3. Shape comparison of the $\ell\pi$ invariant mass spectrum with and without a cut on the pNN score in mass windows around 1, 1.5, 2, 3 and 4.5 GeV in the low $L_{xy}/\sigma_{L_{xy}}$ OS category for the $\mu e\pi$ channel. The pNN working point is set to 0.99 for all the categories and mass windows. The plots are produced with the full unblinded of the B-parking dataset (4.91 fb^{-1}).

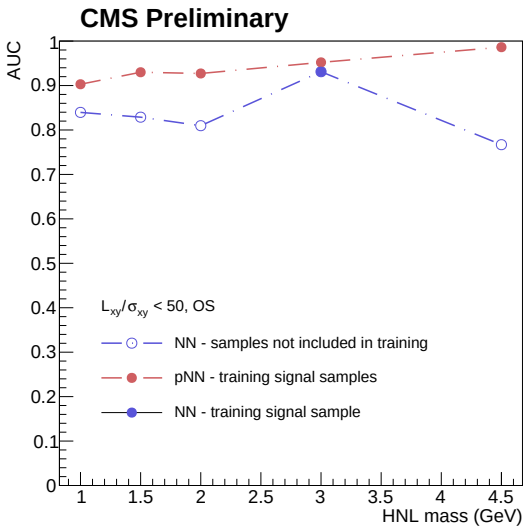


Figure A.4. Comparison of the Area Under the (ROC) Curve value computed for different mass hypothesis using the analysis pNN (red) or a simple NN (blue) trained on a single mass 3 GeV lifetime 100 mm signal sample and for the low $L_{xy}/\sigma_{L_{xy}}$ OS category

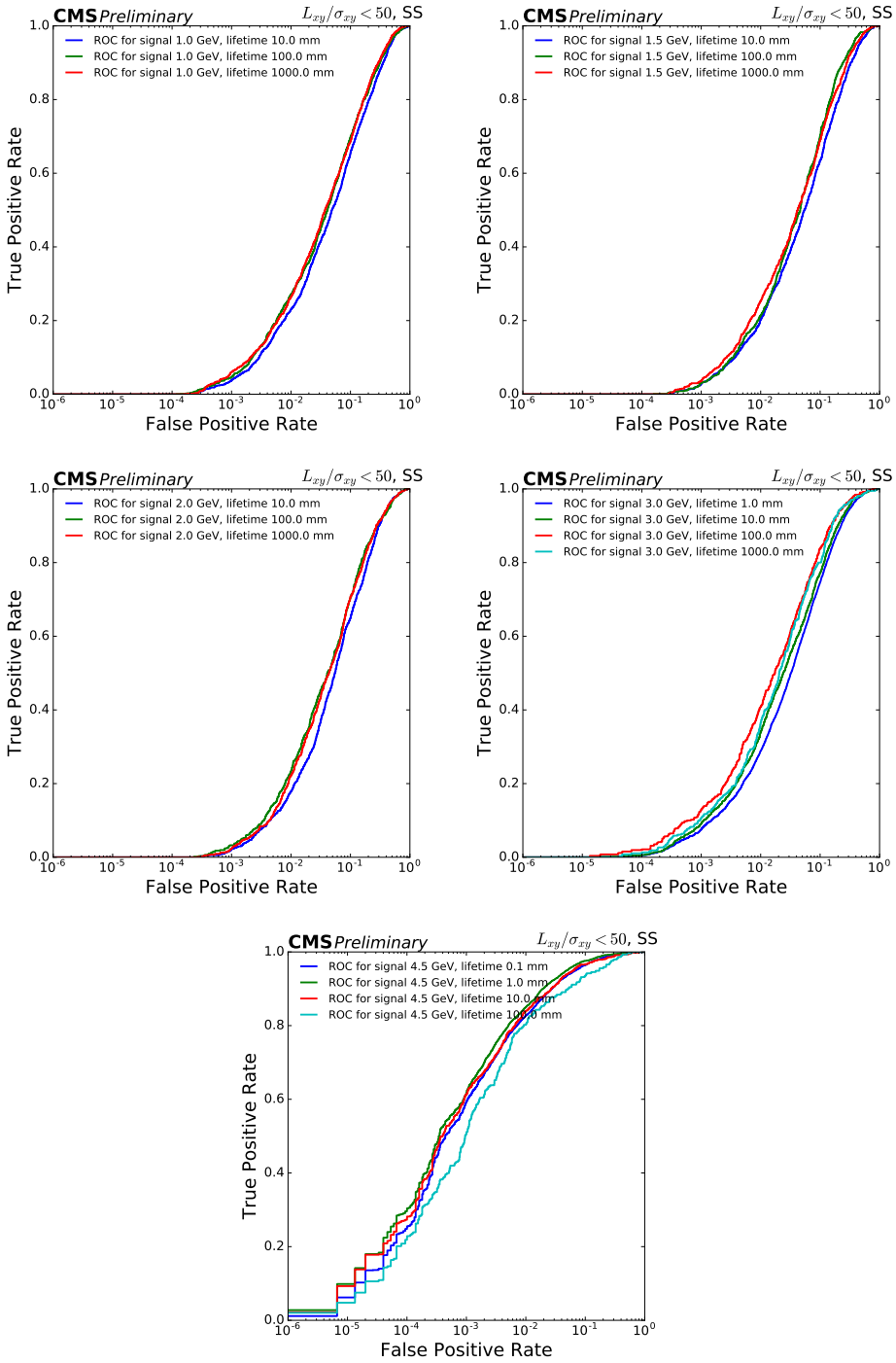


Figure A.5. ROC curve for the mass hypotheses used in the training in the low L_{xy}/σ_{Lxy} SS category for the $\mu e\pi$ channel.

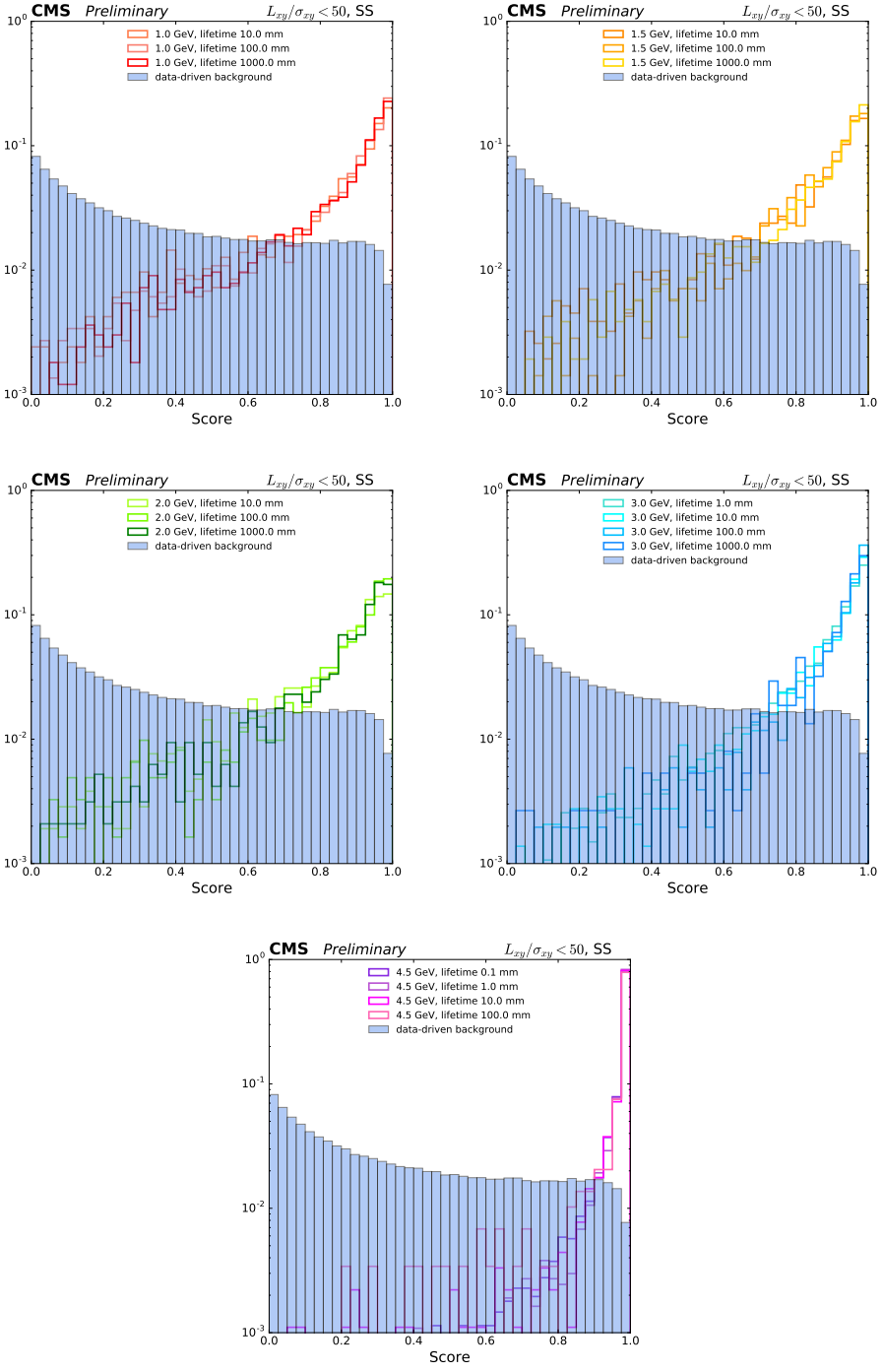


Figure A.6. Score distribution for the mass hypotheses used in the training in the low $L_{xy}/\sigma_{L_{xy}}$ SS category for the $\mu e\pi$ channel.

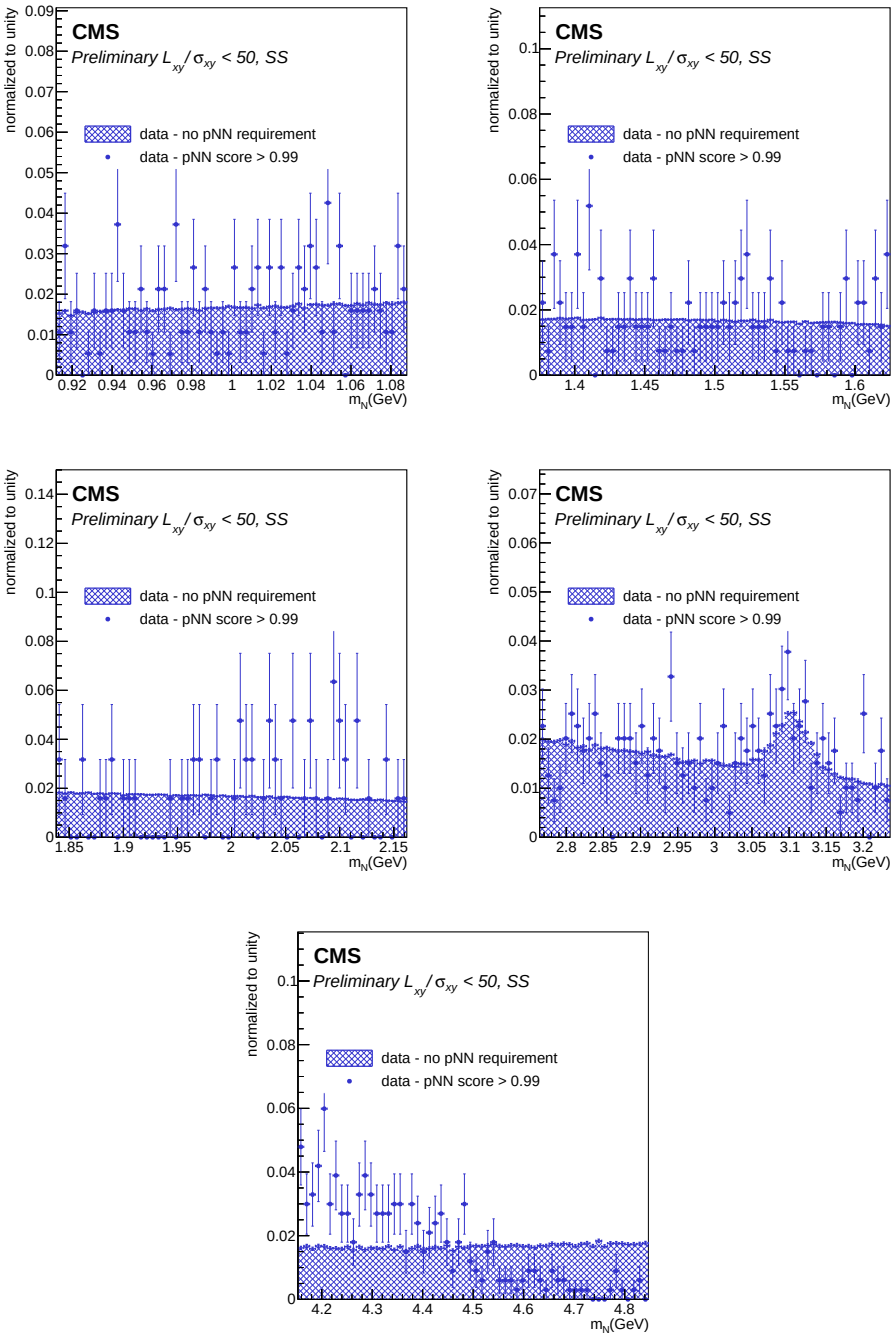


Figure A.7. Shape comparison of the $\ell\pi$ invariant mass spectrum with and without a cut on the pNN score in mass windows around 1, 1.5, 2, 3 and 4.5 GeV in the low $L_{xy}/\sigma_{L_{xy}}$ SS category for the $\mu e\pi$ channel. The pNN working point is set to 0.99 for all the categories and mass windows. The plots are produced with the full of the B-parking dataset (4.91 fb^{-1}).

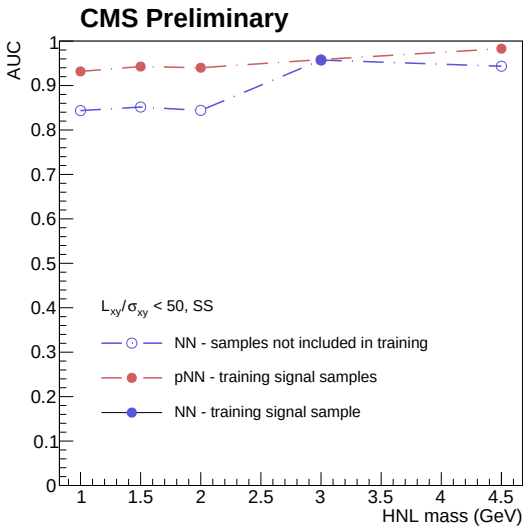


Figure A.8. Comparison of the Area Under the (ROC) Curve value computed for different mass hypothesis using the analysis pNN (red) or a simple NN (blue) trained on a single mass 3 GeV lifetime 100 mm signal sample and for the low $L_{xy}/\sigma_{L_{xy}}$ SS category.

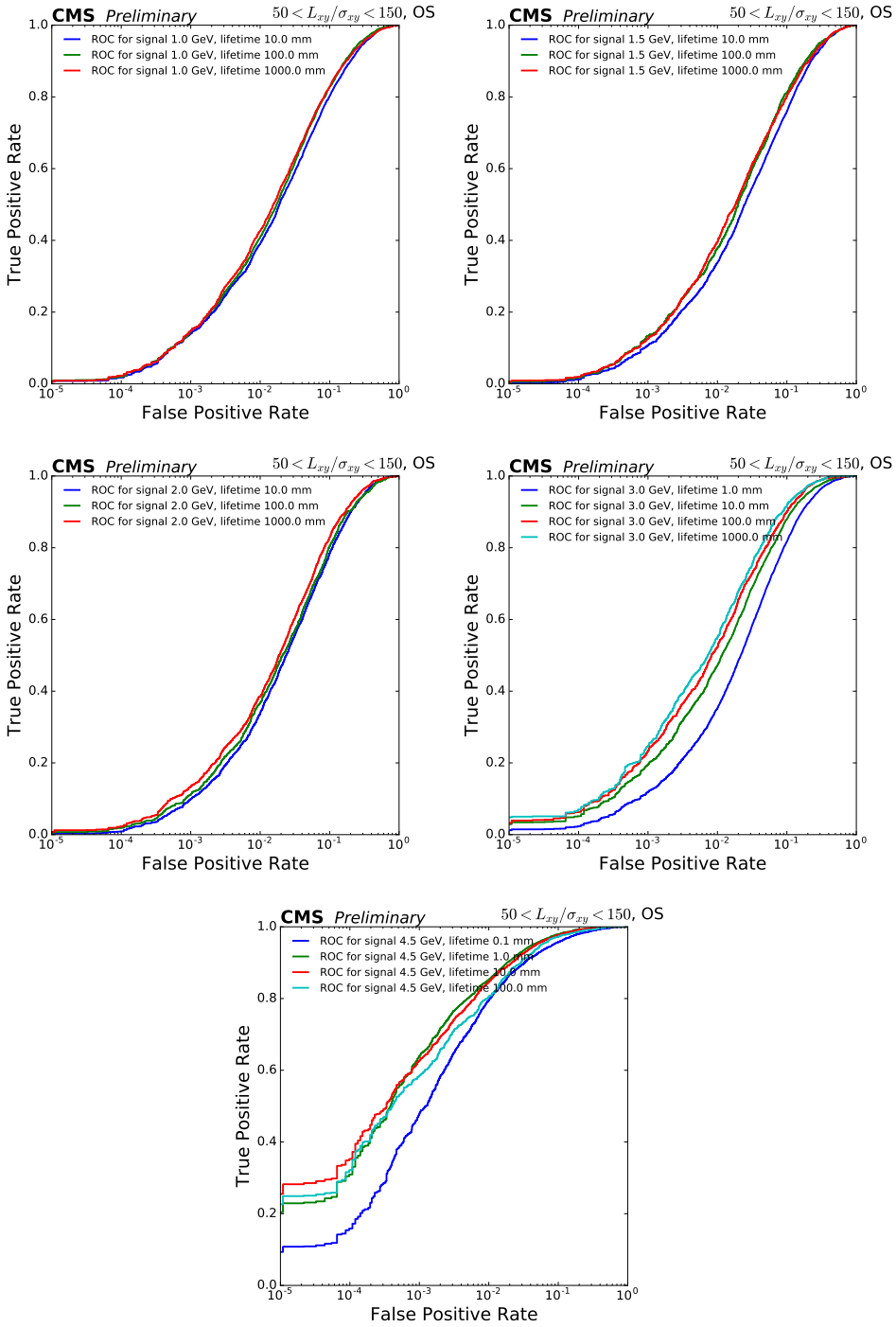


Figure A.9. ROC curve for the mass hypotheses used in the training in the medium L_{xy}/σ_{Lxy} OS category for the $\mu e\pi$ channel.

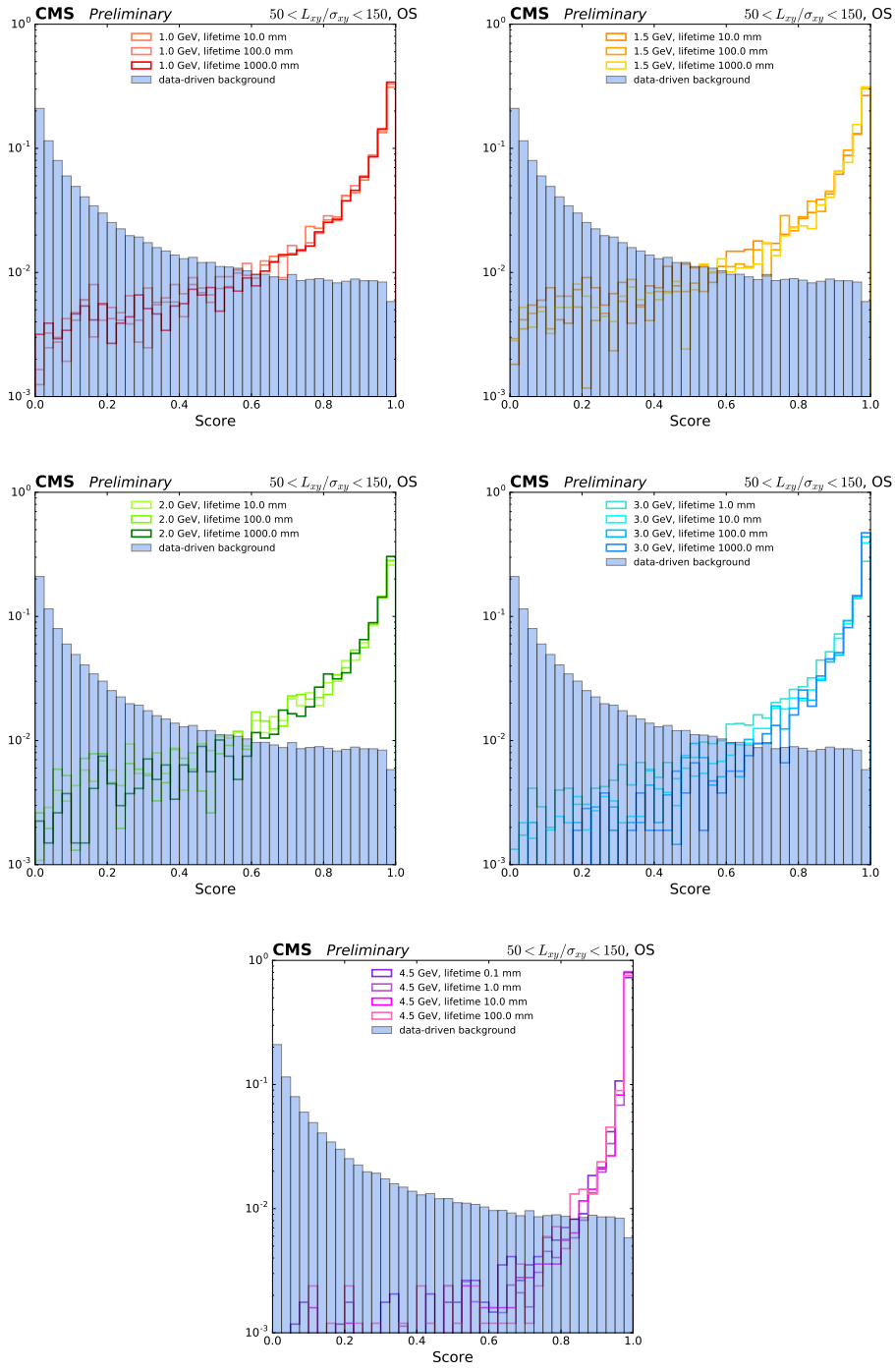


Figure A.10. Score distribution for the mass hypotheses used in the training in the medium $L_{xy}/\sigma_{L_{xy}}$ OS category for the $\mu e\pi$ channel.

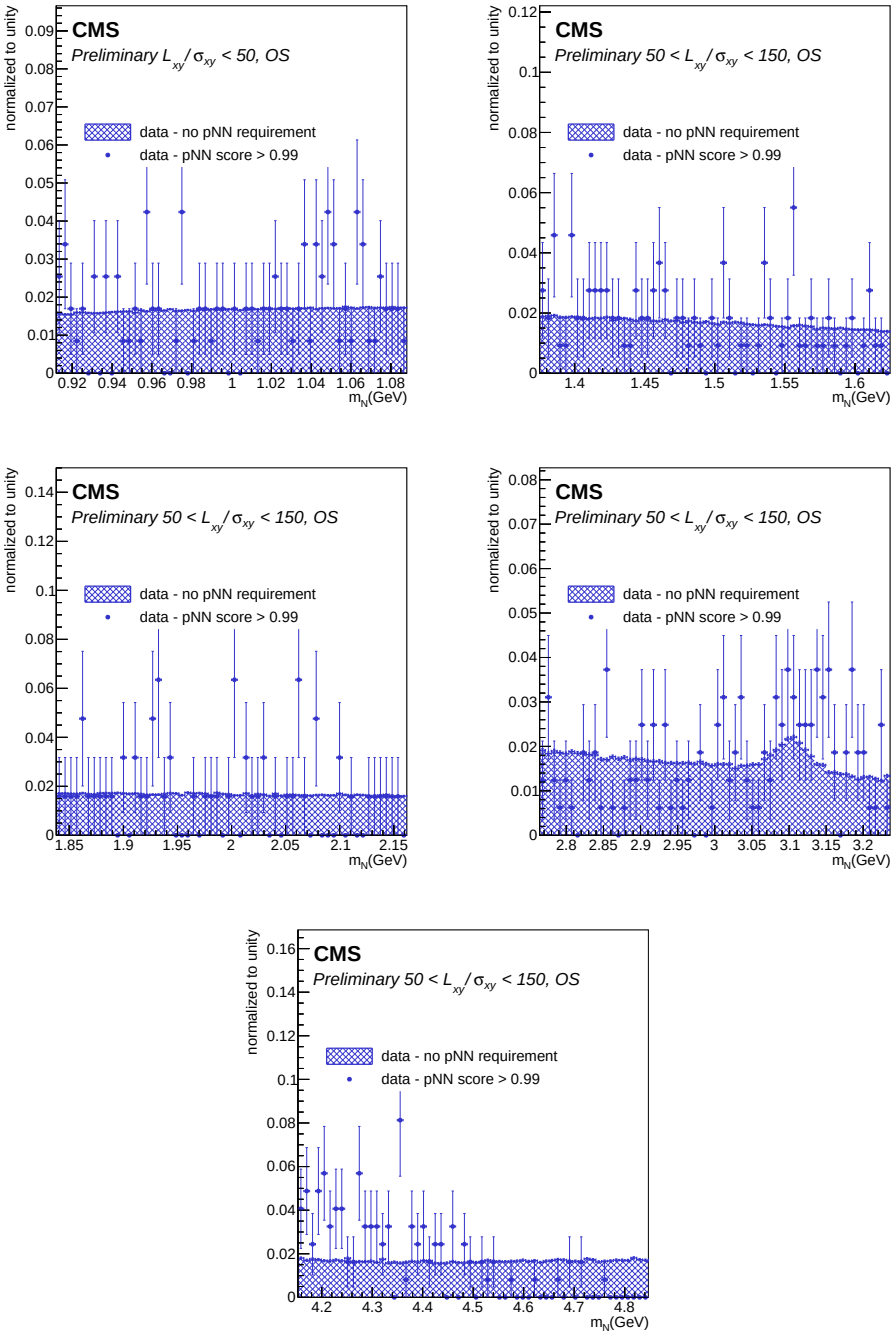


Figure A.11. Shape comparison of the $\ell\pi$ invariant mass spectrum with and without a cut on the pNN score in mass windows around 1, 1.5, 2, 3 and 4.5 GeV in the medium $L_{xy}/\sigma_{L_{xy}}$ OS category for the $\mu e\pi$ channel. The pNN working point is set to 0.99 for all the categories and mass windows. The plots are produced with the full 1D unblinded dataset (4.91 fb^{-1}).

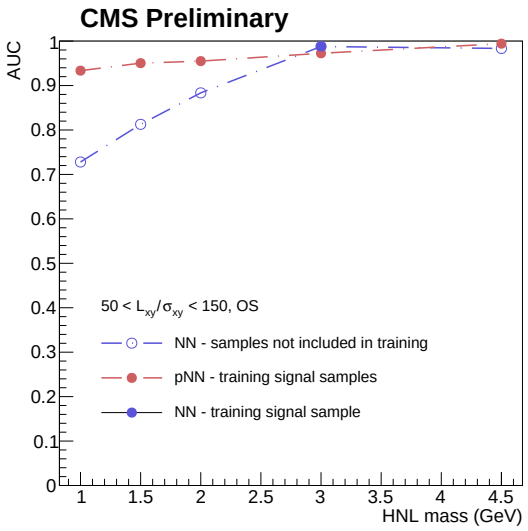


Figure A.12. Comparison of the Area Under the (ROC) Curve value computed for different mass hypothesis using the analysis pNN (red) or a simple NN (blue) trained on a single mass 3 GeV lifetime 100 mm signal sample the medium $L_{xy}/\sigma_{L_{xy}}$ OS category.

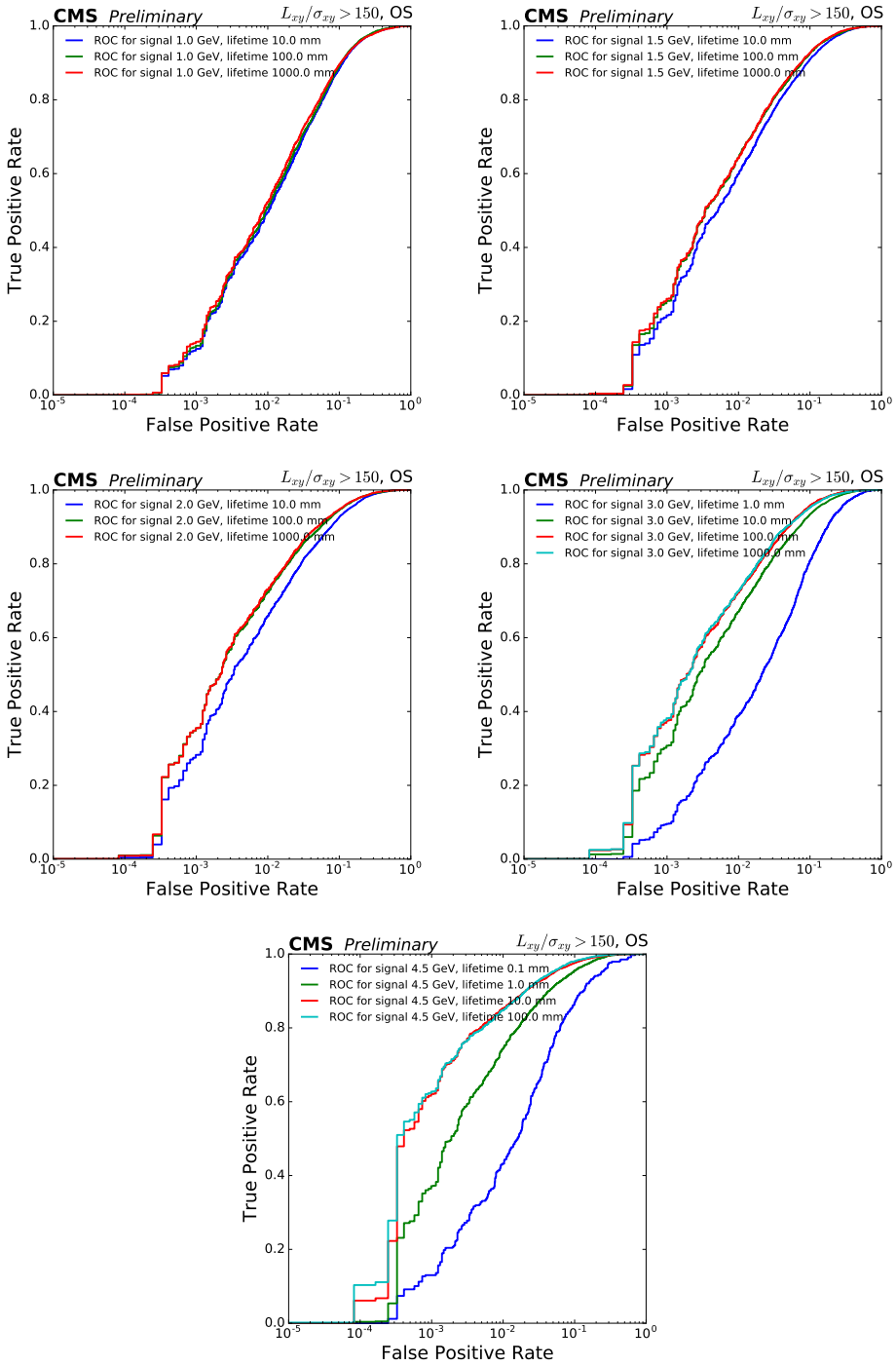


Figure A.13. ROC curve for the mass hypotheses used in the training in the high L_{xy}/σ_{Lxy} OS category for the $\mu e\pi$ channel.

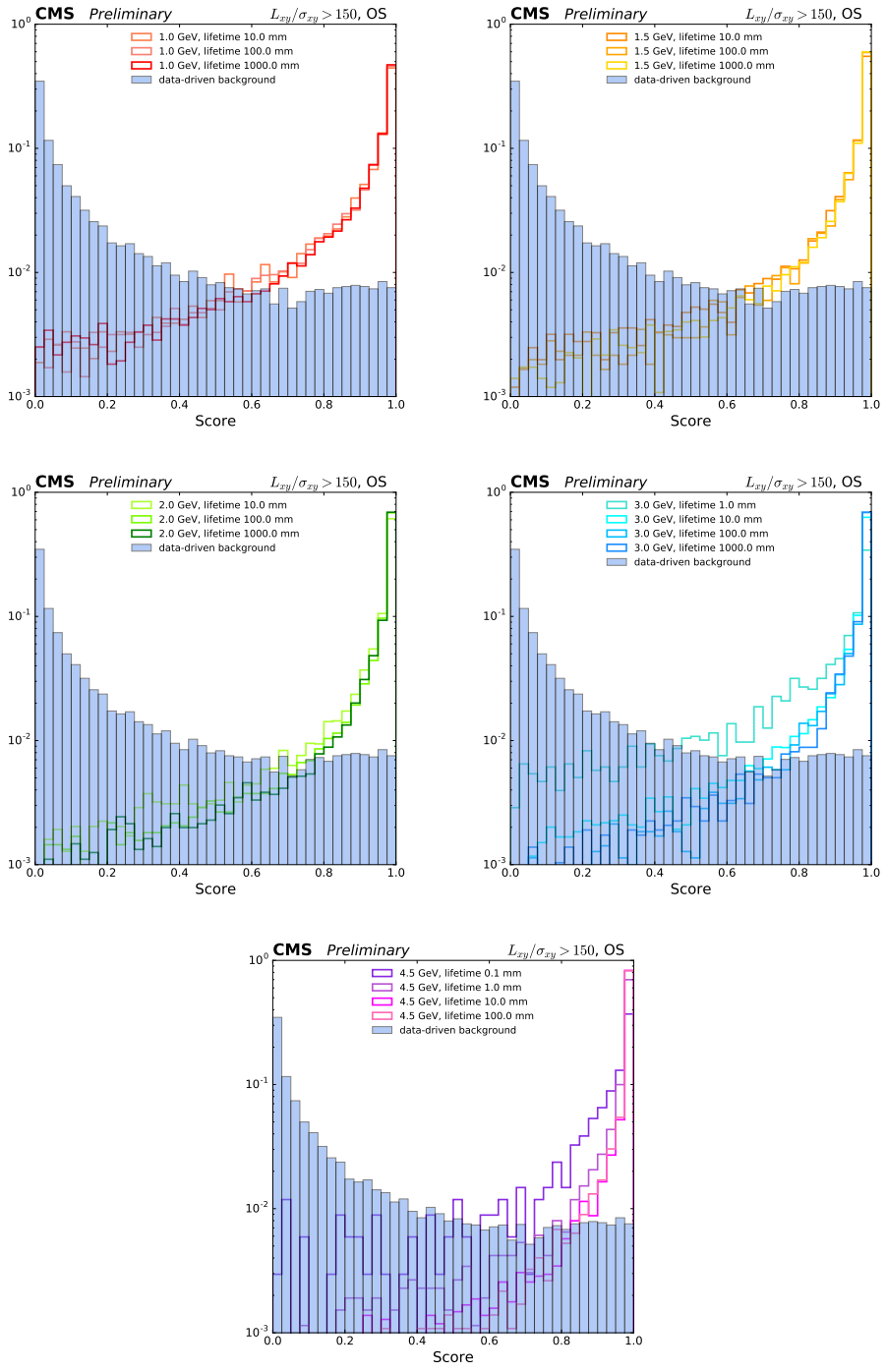


Figure A.14. Score distribution for the mass hypotheses used in the training in the high $L_{xy}/\sigma_{L_{xy}}$ OS category for the $\mu e\pi$ channel.

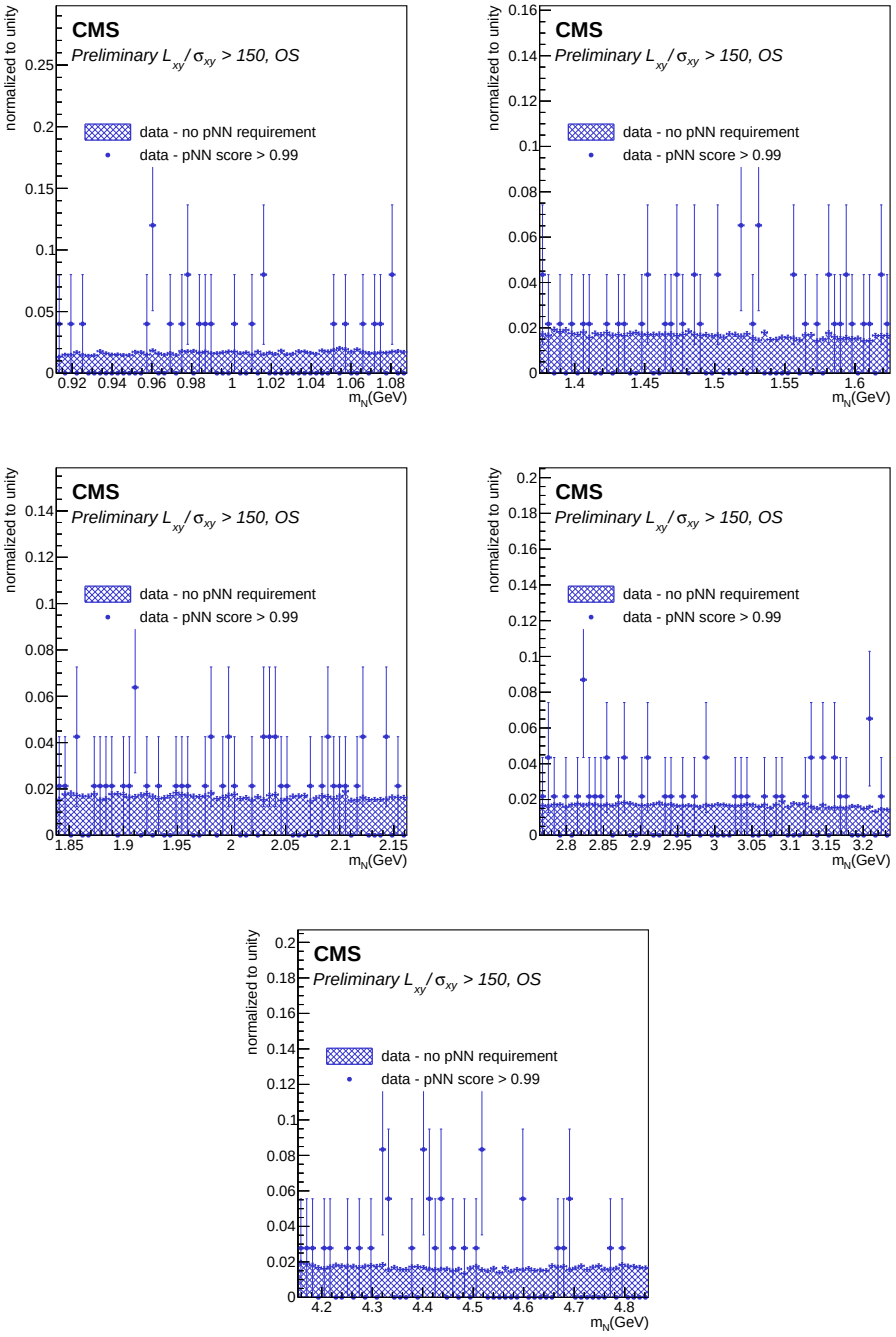


Figure A.15. Shape comparison of the $\ell\pi$ invariant mass spectrum with and without a cut on the pNN score in mass windows around 1, 1.5, 2, 3 and 4.5 GeV in the high $L_{xy}/\sigma_{L_{xy}}$ OS category for the $\mu e\pi$ channel. The pNN working point is set to 0.99 for all the categories and mass windows. The plots are produced with the full 1D unblinded data.

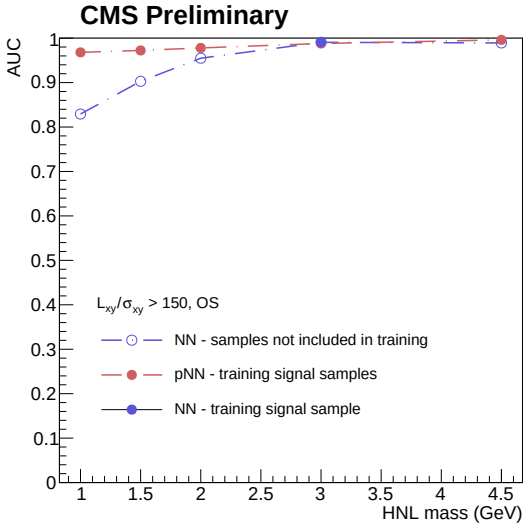


Figure A.16. Comparison of the Area Under the (ROC) Curve value computed for different mass hypothesis using the analysis pNN (red) or a simple NN (blue) trained on a single mass 3 GeV lifetime 100 mm signal sample and for the high $L_{xy}/\sigma_{L_{xy}}$ OS category.

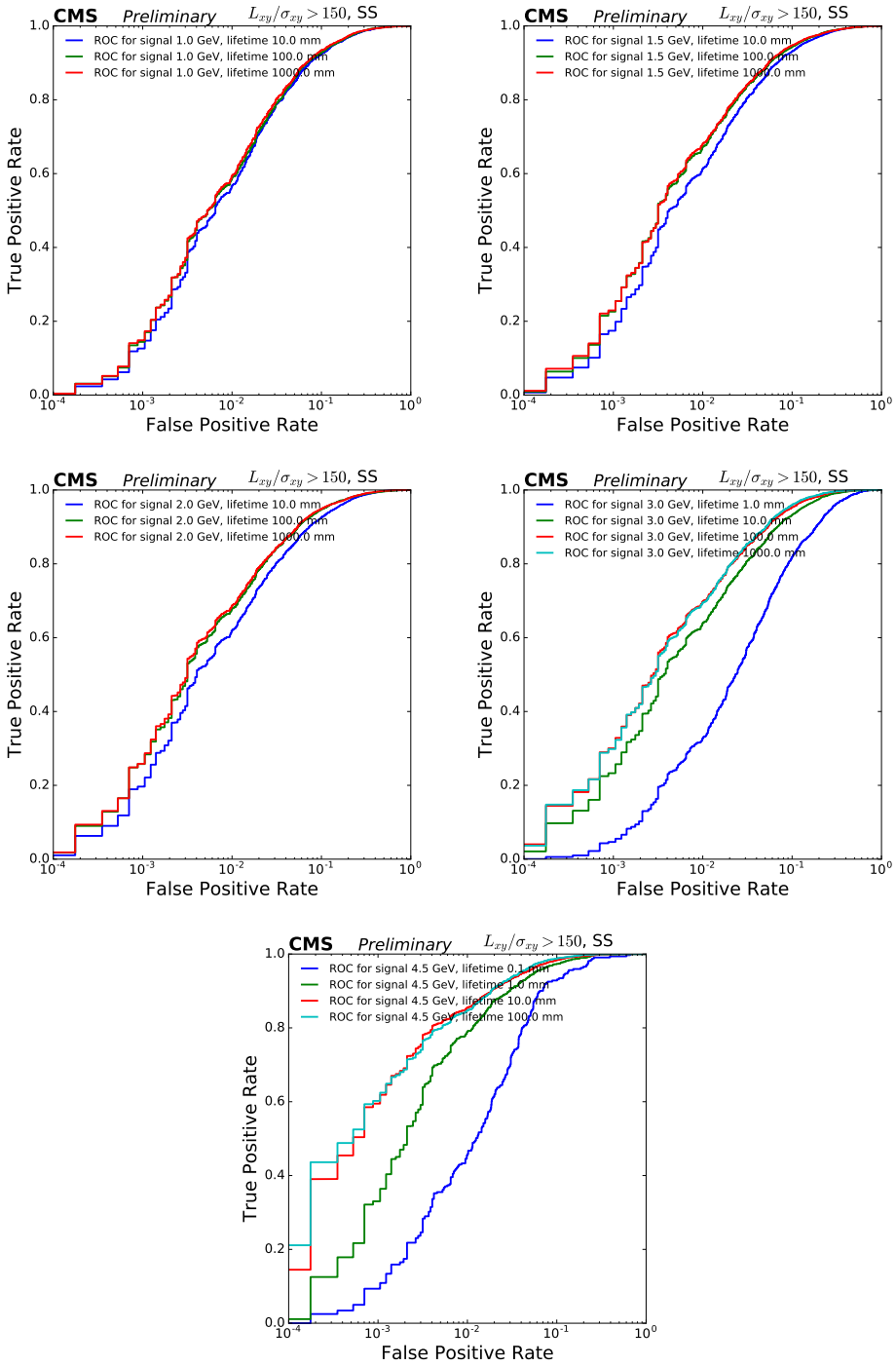


Figure A.17. ROC curve for the mass hypotheses used in the training in the high L_{xy}/σ_{Lxy} SS category for the $\mu e\pi$ channel.

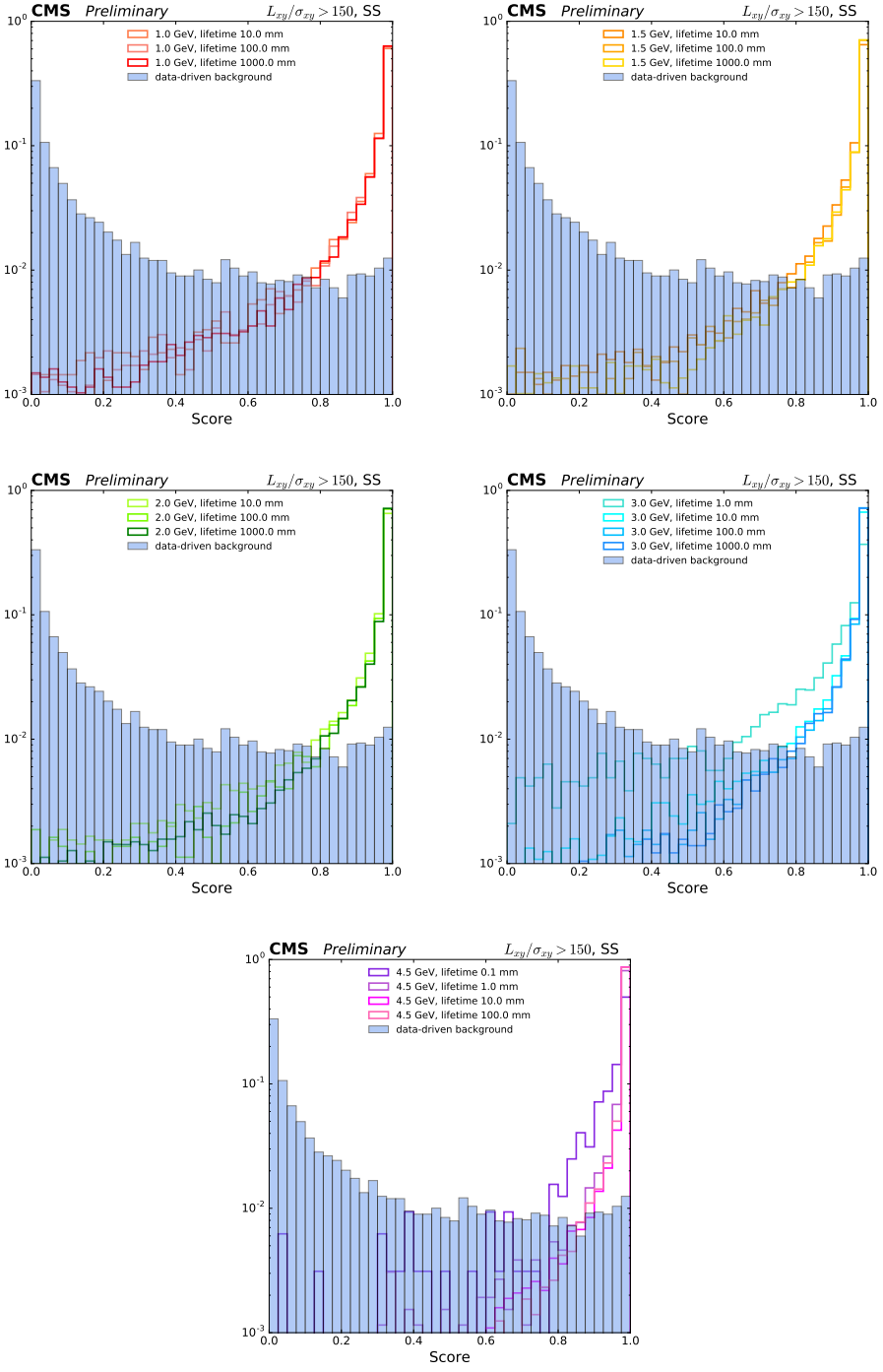


Figure A.18. Score distribution for the mass hypotheses used in the training in the high $L_{xy}/\sigma_{L_{xy}}$ SS category for the $\mu e\pi$ channel.

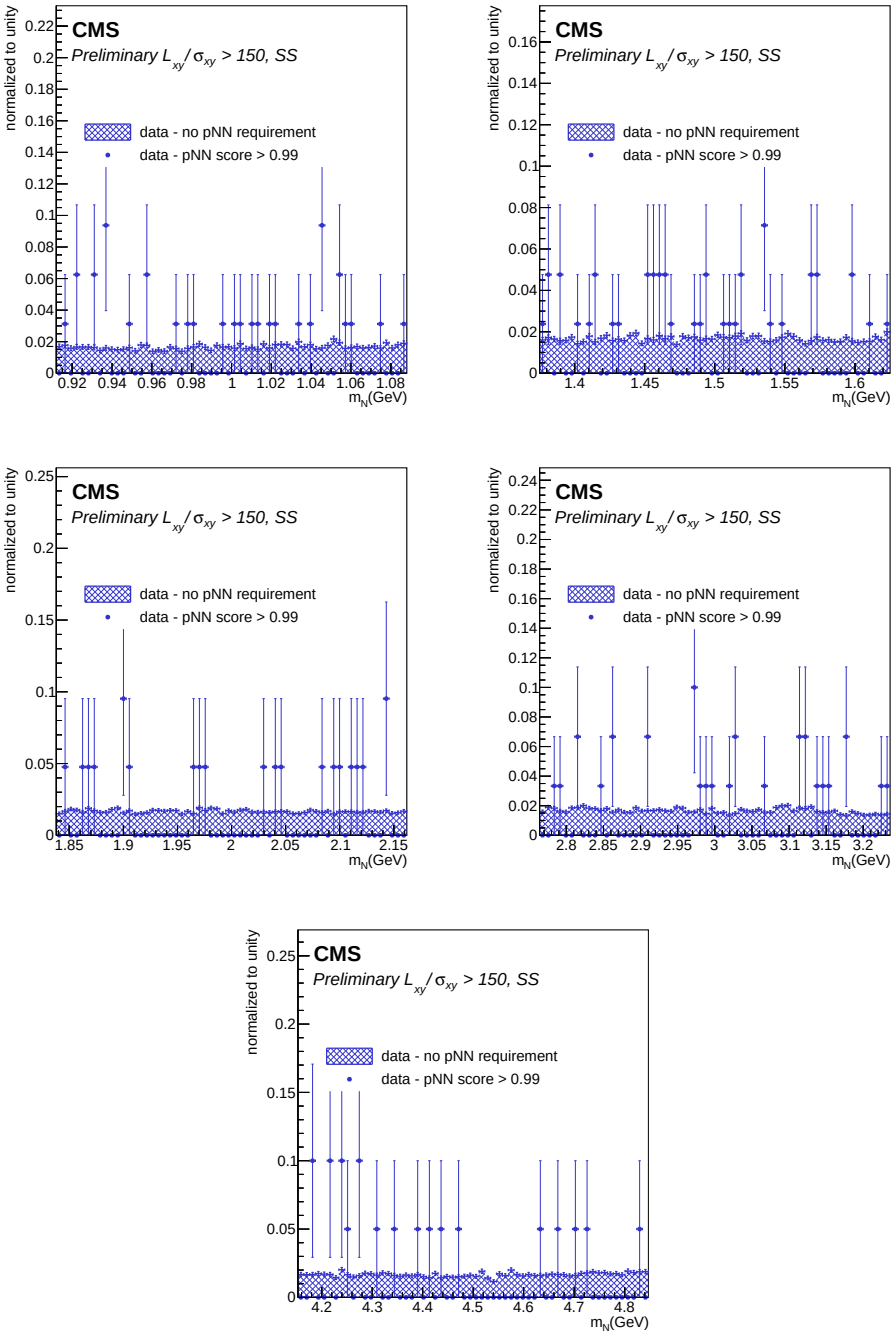


Figure A.19. Shape comparison of the $\ell\pi$ invariant mass spectrum with and without a cut on the pNN score in mass windows around 1, 1.5, 2, 3 and 4.5 GeV in the high $L_{xy}/\sigma_{L_{xy}}$ SS category for the $\mu e\pi$ channel. The pNN working point is set to 0.99 for all the categories and mass windows. The plots are produced with the full unblinded of the B-parking dataset (4.91 fb^{-1}).

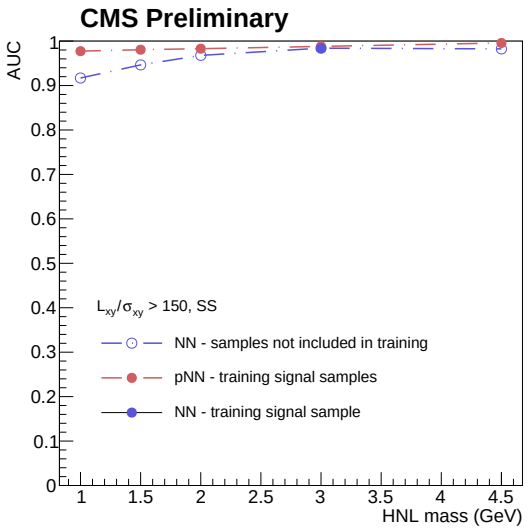


Figure A.20. Comparison of the Area Under the (ROC) Curve value computed for different mass hypothesis using the analysis pNN (red) or a simple NN (blue) trained on a single mass 3 GeV lifetime 100 mm signal sample and for the high $L_{xy}/\sigma_{L_{xy}}$ SS category.

Bibliography

[1] G. Aad et al. "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC". In: *Physics Letters B* 716.1 (Sept. 2012), pp. 1–29. DOI: 10.1016/j.physletb.2012.08.020. URL: https://doi.org/10.1016%5C%2Fj.physletb.2012.08.020.

[2] R. Aaij et al. "Meson production fraction and asymmetry in 7 and 13 TeV collisions". In: *Physical Review D* 100.11 (). DOI: 10.1103/physrevd.100.112006. URL: https://doi.org/10.1103%5C%2Fphysrevd.100.112006.

[3] K. Abe et al. "Indication of Electron Neutrino Appearance from an Accelerator-Produced Off-Axis Muon Neutrino Beam". In: *Physical Review Letters* 107.4 (July 2011). DOI: 10.1103/physrevlett.107.041801. URL: https://doi.org/10.1103%5C%2Fphysrevlett.107.041801.

[4] P. Abreu et al. "Classification of the hadronic decays of the Z0 into b and c quark pairs using a neural network". In: *Physics Letters B* 295.3 (1992), pp. 383–395. ISSN: 0370-2693. DOI: https://doi.org/10.1016/0370-2693(92)91580-3. URL: https://www.sciencedirect.com/science/article/pii/0370269392915803.

[5] N. Agafonova et al. "Final Results of the OPERA Experiment on ν_τ appearance in the CNGS Neutrino Beam". In: *Physical Review Letters* 120.21 (May 2018). DOI: 10.1103/physrevlett.120.211801. URL: https://doi.org/10.1103%5C%2Fphysrevlett.120.211801.

[6] Abien Fred Agarap. *Deep Learning using Rectified Linear Units (ReLU)*. 2018. DOI: 10.48550/ARXIV.1803.08375. URL: https://arxiv.org/abs/1803.08375.

[7] B. Aharmim et al. "Determination of the ν_e and total ^{8}B solar neutrino fluxes using the Sudbury Neutrino Observatory Phase I data set". In: *Phys. Rev. C* 75 (4 Apr. 2007), p. 045502. DOI: 10.1103/PhysRevC.75.045502. URL: https://link.aps.org/doi/10.1103/PhysRevC.75.045502.

[8] M. Anfreville et al. "Laser monitoring system for the CMS lead tungstate crystal calorimeter". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 594.2 (2008), pp. 292–320. ISSN: 0168-9002. DOI: https://doi.org/10.1016/j.nima.2008.01.104. URL: https://www.sciencedirect.com/science/article/pii/S0168900208001599.

[9] G. Apollinari et al. "High Luminosity Large Hadron Collider HL-LHC". In: *CERN Yellow Rep.* 5 (2015), pp. 1–19. DOI: 10.5170/CERN-2015-005.1. arXiv: 1705.08830 [physics.acc-ph].

[10] G. Bak et al. "Fuel-Composition Dependent Reactor Antineutrino Yield at RENO". In: *Physical Review Letters* 122.23 (June 2019). DOI: 10.1103/physrevlett.122.232501. URL: https://doi.org/10.1103%5C%2Fphysrevlett.122.232501.

[11] Pierre Baldi et al. "Parameterized neural networks for high-energy physics". In: *The European Physical Journal C* 76.5 (Apr. 2016). DOI: 10.1140/epjc/s10052-016-4099-4. URL: https://doi.org/10.1140%5C%2Fepjc%5C%2Fs10052-016-4099-4.

[12] Florian Beaudette. "The CMS Particle Flow Algorithm". In: *Proceedings, International Conference on Calorimetry for the High Energy Frontier (CHEF 2013): Paris, France, April 22-25, 2013*. 2013, pp. 295–304. arXiv: 1401.8155 [hep-ex].

[13] Kyrylo Bondarenko et al. "Phenomenology of GeV-scale heavy neutral leptons". In: *Journal of High Energy Physics* 2018.11 (2018), p. 32. DOI: 10.1007/JHEP11(2018)032. URL: https://doi.org/10.1007/JHEP11(2018)032.

[14] Chao-Hsi Chang, Jian-Xiong Wang, and Xing-Gang Wu. "BCVEGPY2.0: A Upgrade version of the generator BCVEGPY with an addendum about hadroproduction of the P-wave B(c) states". In: *Comput. Phys. Commun.* 174 (2006), pp. 241–251. DOI: 10.1016/j.cpc.2005.09.008. arXiv: hep-ph/0504017.

[15] S. Chatrchyan et al. "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC". In: *Physics Letters B* 716.1 (2012), pp. 30–61. ISSN: 0370-2693. DOI: https://doi.org/10.1016/j.physletb.2012.08.021. URL: https://www.sciencedirect.com/science/article/pii/S0370269312008581.

[16] Bruce T. Cleveland et al. "Measurement of the Solar Electron Neutrino Flux with the Homestake Chlorine Detector". In: *The Astrophysical Journal* 496.1 (Mar. 1998), p. 505. DOI: 10.1086/305343. URL: https://dx.doi.org/10.1086/305343.

[17] CMS Collaboration. *CMS Physics: Technical Design Report Volume 1: Detector Performance and Software*. Technical Design Report CMS. There is an error on cover due to a technical problem for some items. Geneva: CERN, 2006. URL: https://cds.cern.ch/record/922757.

[18] CMS Collaboration. "Search for long-lived heavy neutral leptons with displaced vertices in proton-proton collisions at $\sqrt{s}$=13 TeV". In: *Journal of High Energy Physics* 2022.7 (July 2022). DOI: 10.1007/jhep07(2022)081. URL: https://doi.org/10.1007%5C%2Fjhep07%5C%282022%5C%29081.

[19] Yann Coadou. "Boosted Decision Trees". In: (Feb. 2022), pp. 9–58. DOI: 10.1142/9789811234033_0002. URL: https://doi.org/10.1142%5C%2F9789811234033_0002.

[20] CMS collaboration et al. "Determination of jet energy calibration and transverse momentum resolution in CMS". In: *Journal of Instrumentation* 6.11 (2011), P11002.

[21] The CMS Collaboration. "Description and performance of track and primary-vertex reconstruction with the CMS tracker". In: *Journal of Instrumentation* 9.10 (Oct. 2014), P10009. DOI: 10.1088/1748-0221/9/10/P10009. URL: https://dx.doi.org/10.1088/1748-0221/9/10/P10009.

[22] Glen Cowan et al. "Asymptotic formulae for likelihood-based tests of new physics". In: *The European Physical Journal C* 71.2 (Feb. 2011). DOI: 10.1140/epjc/s10052-011-1554-0. URL: https://doi.org/10.1140%5C%2Fepjc%5C%2Fs10052-011-1554-0.

[23] P.D. Dauncey et al. "Handling uncertainties in background shapes: the discrete profiling method". In: *Journal of Instrumentation* 10.04 (Apr. 2015), P04015–P04015. DOI: 10.1088/1748-0221/10/04/p04015. URL: https://doi.org/10.1088%5C%2F1748-0221%5C%2F10%5C%2F04%5C%2Fp04015.

[24] Frank F Deppisch, P S Bhupal Dev, and Apostolos Pilaftsis. "Neutrinos and collider physics". In: *New Journal of Physics* 17.7 (Aug. 2015), p. 075019. DOI: 10.1088/1367-2630/17/7/075019. URL: https://dx.doi.org/10.1088/1367-2630/17/7/075019.

[25] *Documentation on Keras Adam optimiser*. https://keras.io/api/optimizers/adam/.

[26] *Documentation on the RobustScaler*. https://scikit-learn.org/stable/modules/generated/sklearn.preprocessing.RobustScaler.html.

[27] Marco Drewes et al. "NA62 sensitivity to heavy neutral leptons in the low scale seesaw model". In: *Journal of High Energy Physics* 2018 (Jan. 2018). DOI: 10.1007/JHEP07(2018)105.

[28] F. Englert and R. Brout. "Broken Symmetry and the Mass of Gauge Vector Mesons". In: *Phys. Rev. Lett.* 13 (9 Aug. 1964), pp. 321–323. DOI: 10.1103/PhysRevLett.13.321. URL: https://link.aps.org/doi/10.1103/PhysRevLett.13.321.

[29] Ivan Esteban et al. "Updated fit to three neutrino mixing: exploring the accelerator-reactor complementarity". In: *Journal of High Energy Physics* 2017.1 (Jan. 2017). DOI: 10.1007/jhep01(2017)087. URL: https://doi.org/10.1007%5C%2Fjhep01%5C%282017%5C%29087.

[30] Ronald A Fisher. "On the interpretation of χ 2 from contingency tables, and the calculation of P". In: *Journal of the royal statistical society* 85.1 (1922), pp. 87–94.

[31] S. L. Glashow, J. Iliopoulos, and L. Maiani. "Weak Interactions with Lepton-Hadron Symmetry". In: 2.7 (Oct. 1970), pp. 1285–1292. DOI: 10.1103/PhysRevD.2.1285.

[32] Sheldon L. Glashow. "The renormalizability of vector meson interactions". In: *Nuclear Physics* 10 (1959), pp. 107–117. ISSN: 0029-5582. DOI: https://doi.org/10.1016/0029-5582(59)90196-8. URL: https://www.sciencedirect.com/science/article/pii/0029558259901968.

[33] J Goldstone. "Field theories with "superconductor" solutions". In: *Nuovo Cimento* 19 (1961), pp. 154–164. DOI: 10.1007/BF02812722. URL: http://cds.cern.ch/record/343400.

[34] M. C. Gonzalez-Garcia and Yosef Nir. "Neutrino masses and mixing: evidence and implications". In: *Reviews of Modern Physics* 75.2 (Mar. 2003), pp. 345–402. DOI: 10.1103/revmodphys.75.345. URL: https://doi.org/10.1103%5C%2Frevmodphys.75.345.

[35] M.C. Gonzalez-Garcia and Michele Maltoni. "Phenomenology with massive neutrinos". In: *Physics Reports* 460.1-3 (Apr. 2008), pp. 1–129. DOI: 10.1016/j.physrep.2007.12.004. URL: https://doi.org/10.1016%5C%2Fj.physrep.2007.12.004.

[36] Particle Data Group et al. "Review of Particle Physics". In: *Progress of Theoretical and Experimental Physics* 2020.8 (Aug. 2020). 083C01. ISSN: 2050-3911. DOI: 10.1093/ptep/ptaa104. eprint: https://academic.oup.com/ptep/article-pdf/2020/8/083C01/34673740/rpp2020-vol2-2015-2092_18.pdf. URL: https://doi.org/10.1093/ptep/ptaa104.

[37] M. J. Herrero. *The Standard Model*. 1998. DOI: 10.48550/ARXIV.HEP-PH/9812242. URL: https://arxiv.org/abs/hep-ph/9812242.

[38] Peter W. Higgs. "Broken Symmetries and the Masses of Gauge Bosons". In: *Phys. Rev. Lett.* 13 (16 Oct. 1964), pp. 508–509. DOI: 10.1103/PhysRevLett.13.508. URL: https://link.aps.org/doi/10.1103/PhysRevLett.13.508.

[39] Thomas Junk. "Confidence level computation for combining searches with small statistics". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 434.2-3 (Sept. 1999), pp. 435–443. DOI: 10.1016/s0168-9002(99)00498-2. URL: https://doi.org/10.1016%5C%2Fs0168-9002%5C%2899%5C%2900498-2.

[40] Kamiokande Collaboration et al. *Solar Neutrino Measurements in Super-Kamiokande-IV*. 2016. DOI: 10.48550/ARXIV.1606.07538. URL: https://arxiv.org/abs/1606.07538.

[41] *Keras documentation*. https://keras.io/.

[42] Ilse Krätschmer. "Muon reconstruction and identification in CMS Run I and towards Run II". In: *Nuclear and Particle Physics Proceedings* 273-275 (Apr. 2016), pp. 2500–2502. DOI: 10.1016/j.nuclphysbps.2015.09.438.

[43] Leif Lönnblad, Carsten Peterson, and Thorsteinn Rögnvaldsson. "Finding gluon jets with a neural trigger". In: *Phys. Rev. Lett.* 65 (11 Sept. 1990), pp. 1321–1324. DOI: 10.1103/PhysRevLett.65.1321. URL: https://link.aps.org/doi/10.1103/PhysRevLett.65.1321.

[44] Z. Maki, M. Nakagawa, and S. Sakata. "Remarks on the Unified Model of Elementary Particles". In: *Progress of Theoretical Physics* 28.5 (Nov. 1962), pp. 870–880. DOI: 10.1143/PTP.28.870.

[45] "Measurement of the differential inclusive B+ hadron cross sections in pp collisions at $\sqrt{s}$=13 TeV". In: *Physics Letters B* 771 (2017), pp. 435–456. ISSN: 0370-2693. DOI: https://doi.org/10.1016/j.physletb.2017.05.074. URL: https://www.sciencedirect.com/science/article/pii/S0370269317304379.

[46] P. Meridiani and C Rovelli. *Position resolution at the 2006 ECAL testbeam.* Tech. rep. 2007.

[47] Emmy Noether. "Invariant variation problems". In: *Transport Theory and Statistical Physics* 1.3 (Jan. 1971), pp. 186–207. DOI: 10.1080/00411457108231446. arXiv: physics/0503066 [physics.hist-ph].

[48] "Performance of electron reconstruction and selection with the CMS detector in proton-proton collisions at $\sqrt{s}$ = 8 TeV. Performance of electron reconstruction and selection with the CMS detector in proton-proton collisions at sqrt(s)=8 TeV". In: *JINST* 10 (2015). Replaced with published version. Added journal reference and DOI, P06005. DOI: 10.1088/1748-0221/10/06/P06005. arXiv: 1502.02701. URL: https://cds.cern.ch/record/1988091.

[49] Carsten Peterson. "Track finding with neural networks". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 279.3 (1989), pp. 537–545. ISSN: 0168-9002. DOI: https://doi.org/10.1016/0168-9002(89)91300-4. URL: https://www.sciencedirect.com/science/article/pii/0168900289913004.

[50] M. Pivk and F.R. Le Diberder. ": A statistical tool to unfold data distributions". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 555.1-2 (Dec. 2005), pp. 356–369. DOI: 10.1016/j.nima.2005.08.106. URL: https://doi.org/10.1016%5C%2Fj.nima.2005.08.106.

[51] Speer Th Prokofiev K. *A Kinematic fit and a decay chain reconstruction library.* Tech. rep. Sept. 2004. URL: https://twiki.cern.ch/twiki/pub/CMSPublic/SWGuideKinematicVertexFit/CMS_IN_2004_20.pdf.

[52] "Recording and reconstructing 10 billion unbiased b hadron decays in CMS". In: (2019). URL: https://cds.cern.ch/record/2704495.

[53] Anders Ryd et al. "EvtGen: A Monte Carlo Generator for B-Physics". In: (May 2005).

[54] Abdus Salam and J. C. Ward. "Weak and electromagnetic interactions". In: *Il Nuovo Cimento* 11.4 (Feb. 1959), pp. 568–577. DOI: 10.1007/BF02726525.

[55] Jürgen Schmidhuber. "Deep learning in neural networks: An overview". In: *Neural Networks* 61 (2015), pp. 85–117. ISSN: 0893-6080. DOI: https://doi.org/10.1016/j.neunet.2014.09.003. URL: https://www.sciencedirect.com/science/article/pii/S0893608014002135.

[56] Hale Sert. "CMS Run 2 High Level Trigger Performance". In: *PoS* EPS-HEP2019 (2020), p. 165. DOI: 10.22323/1.364.0165.

[57] A.M. Sirunyan et al. "Electron and photon reconstruction and identification with the CMS experiment at the CERN LHC". In: *Journal of Instrumentation* 16.05 (May 2021), P05014. DOI: 10.1088/1748-0221/16/05/p05014. URL: https://doi.org/10.1088%5C%2F1748-0221%5C%2F16%5C%2F05%5C%2Fp05014.

[58] A.M. Sirunyan et al. "Performance of the CMS Level-1 trigger in proton-proton collisions at s = 13 TeV". In: *Journal of Instrumentation* 15.10 (Oct. 2020), P10017. DOI: 10.1088/1748-0221/15/10/P10017. URL: https://dx.doi.org/10.1088/1748-0221/15/10/P10017.

[59] A.M. Sirunyan et al. "Reconstruction of signal amplitudes in the CMS electromagnetic calorimeter in the presence of overlapping proton-proton interactions". In: *Journal of Instrumentation* 15.10 (Oct. 2020), P10002. DOI: 10.1088/1748-0221/15/10/P10002. URL: https://dx.doi.org/10.1088/1748-0221/15/10/P10002.

[60] Torbjörn Sjöstrand et al. "An introduction to PYTHIA 8.2". In: *Comput. Phys. Commun.* 191 (2015), p. 159. DOI: 10.1016/j.cpc.2015.01.024. arXiv: 1410.3012 [hep-ph].

[61] *Tensorflow documentation.* https://www.tensorflow.org/?hl=it.

[62] Steven Weinberg. "A Model of Leptons". In: 19.21 (Nov. 1967), pp. 1264–1266. DOI: 10.1103/PhysRevLett.19.1264.

[63] S. S. Wilks. "The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses". In: *The Annals of Mathematical Statistics* 9.1 (1938), pp. 60–62. DOI: 10.1214/aoms/1177732360. URL: https://doi.org/10.1214/aoms/1177732360.