

Probing lepton flavor mixing in W_R searches with machine learning at the LHC

Jin-Man Cai^{1,*} and Gang Li^{1,2,†}

¹*School of Physics and Astronomy, Sun Yat-sen University, Zhuhai 519082, China*

²*Guangdong Provincial Key Laboratory of Quantum Metrology and Sensing, Sun Yat-Sen University, Zhuhai 519082, China*



(Received 9 June 2026; accepted 26 August 2026; published 21 September 2026)

Right-handed lepton flavor mixing in the left-right symmetric model directly affects the production and decay of heavy Majorana neutrinos N_R , yet its impact on collider searches remains less explored. Using a deep neural network, we analyze the Keung-Senjanović process $pp \rightarrow W_R \rightarrow \ell_\alpha N_R \rightarrow \ell_\alpha \ell_\beta jj$ with $\ell_{\alpha,\beta} = e, \mu$ at LHC Run 2 and the HL-LHC, considering both same-sign and opposite-sign dilepton channels. We adopt three benchmark mixing scenarios: unmixed, maximal mixing, and Pontecorvo-Maki-Nakagawa-Sakata (PMNS)-like. In the unmixed scenario, the deep neural network improves the expected significance over the cut-based analyses performed by ATLAS, leading to stronger expected sensitivities. For the combined $\ell\ell$ analysis, the HL-LHC is projected to reach m_{W_R} and m_{N_R} values up to 6.7 (6.3) and 4.4 (4.1) TeV, respectively, under maximal (PMNS-like) mixing. For representative mass benchmarks, our analysis indicates that LHC Run 2 can test a significant portion of the $|V_{e1}|$ - $|V_{\mu 1}|$ plane, and the HL-LHC could probe even smaller mixing values, including the corresponding benchmark points for both the maximal and PMNS-like patterns. Finally, we investigate complementarities with low-energy charged lepton flavor-violation processes, where future searches can overlap with or exceed the LHC reach.

DOI: [10.1103/fzws-k9rl](https://doi.org/10.1103/fzws-k9rl)

I. INTRODUCTION

The origin of neutrino masses remains one of the most compelling open questions in particle physics, motivating extensive searches for neutrino-related new physics beyond the Standard Model (SM). Among the proposed frameworks, the left-right symmetric model (LRSM) [1–4] provides an especially attractive explanation for the smallness of neutrino masses. The model extends the SM gauge group to $SU(2)_L \times SU(2)_R \times U(1)_{B-L}$ [5,6], naturally introducing right-handed neutrinos. The spontaneous breaking of $SU(2)_R$ subsequently generates the heavy Majorana masses. Compared to the minimal type-I seesaw mechanism [7–11], the LRSM establishes a more complete framework in which the origin of neutrino masses is directly connected to gauge symmetry breaking.

A key prediction of the LRSM is the existence of heavy gauge bosons $W_R^\pm$ and Z_R that couple with the SM fields and heavy Majorana neutrinos N_R . At the Large Hadron Collider (LHC), the Keung-Senjanović process [12] $pp \rightarrow W_R \rightarrow \ell N_R \rightarrow \ell\ell jj$ via the mediation of W_R serves as a smoking-gun signature for the LRSM. This process not only enables direct searches for W_R boson and the right-handed neutrino N_R , but also offers a direct way to probe the Majorana nature of N_R through the observation of same-sign dilepton events.

Searches at the LHC during Run 2, with a center-of-mass energy $\sqrt{s} = 13$ TeV and an integrated luminosity of 139 fb^{-1} , have placed lower bounds on the masses of W_R and N_R at the TeV scale [13,14]. For instance, the ATLAS Collaboration excludes m_{W_R} up to about 6.4 TeV for $m_{N_R} = 1$ TeV [13]. However, these limits are obtained under the assumption of no lepton flavor mixing in the right-handed sector. By contrast, the structure of right-handed neutrino mixing matrix V_R in the LRSM has drawn increasing theoretical attention [15–24].

Notably, lepton flavor-mixing effects in heavy-neutrino searches have been investigated experimentally in low-scale type-I seesaw models by both ATLAS and CMS at Run 2 [25–30]. Similar flavor-dependent searches in the

*Contact author: caijm7@mail2.sysu.edu.cn

†Contact author: ligang65@mail.sysu.edu.cn

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/). Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI. Funded by SCOAP³.

context of the LRSM are lacking.¹ In Ref. [35], prospects for the 14 TeV LHC with an integrated luminosity of 30 fb⁻¹ were investigated, including the effects of lepton flavor mixing. The determination of V_R at the LHC was also discussed in Ref. [36] through ratios of flavor-dependent final states, with all e , μ , and τ flavors taken into account. Despite these phenomenological studies, a dedicated and stand-alone same-sign dilepton analysis was not performed, as it depends crucially on a proper treatment of SM backgrounds from fake and charge-misidentified leptons [31]. With the substantial improvement in experimental techniques and the much larger Run 2 dataset now available, it is timely and well motivated to derive updated flavor-dependent LHC constraints in the LRSM.

In this work, we conduct a detailed collider study of the Keung-Senjanović process at LHC Run 2 and the high-luminosity LHC (HL-LHC) with $\sqrt{s} = 14$ TeV and $\mathcal{L} = 3$ ab⁻¹, focusing on different e, μ flavor combinations in the final state. Motivated by the rapid development and widespread application of machine learning (ML) in high-energy physics [37,38], we perform an ML analysis of W_R searches using a deep neural network [39–41]. Recent work [42] has also explored ML analyses of this process including lepton flavor-mixing effects. Our work differs in several respects: Beyond using a different ML method, we study the process within the minimal LRSM, while they considered the LRSM with inverse seesaw. Moreover, we include crucial SM backgrounds from fake and charge-misidentified leptons in the same-sign dilepton channel, which are essential for a realistic collider analysis.

The rest of this paper is organized as follows. In Sec. II, we briefly introduce the minimal LRSM. Section III discusses the Keung-Senjanović process with possible flavor effects in the final state. The detailed collider analyses, including background estimation and the ML analysis, are presented in Sec. IV. In Sec. V, we derive sensitivities to the masses of W_R and N_R as well as to the right-handed lepton mixing at LHC Run 2 and the HL-LHC. We also study the complementarity with low-energy charged lepton flavor-violating (CLFV) processes in Sec. VI. Finally, Sec. VII summarizes the results. More details of ML analysis and CLFV searches are provided in the Appendixes.

II. MINIMAL LEFT-RIGHT SYMMETRICAL MODEL

The minimal LRSM [1–6] extends the electroweak gauge symmetry of the SM to $SU(2)_L \times SU(2)_R \times U(1)_{B-L}$.

¹In the type-I seesaw searches cited above, the flavor dependence is controlled by active-heavy neutrino mixing rather than by the right-handed mixing matrix V_R . In the latter LRSM, an early ATLAS analysis at $\sqrt{s} = 7$ TeV with 2.1 fb⁻¹ of data did consider lepton flavor mixing [31], but later searches omitted such possibilities [13,32–34].

Each generation of quarks and leptons is assigned in doublets:

$$Q_{L,R} = \begin{pmatrix} u \\ d \end{pmatrix}_{L,R}, \quad L_{L,R} = \begin{pmatrix} \nu \\ \ell \end{pmatrix}_{L,R}. \quad (1)$$

The scalar sector includes a bidoublet and two triplet fields:

$$\Phi = \begin{pmatrix} \phi_1^0 & \phi_2^+ \\ \phi_1^- & \phi_2^0 \end{pmatrix}, \quad (2)$$

$$\Delta_{L,R} = \begin{pmatrix} \Delta_{L,R}^+/\sqrt{2} & \Delta_{L,R}^{++} \\ \Delta_{L,R}^0 & -\Delta_{L,R}^+/\sqrt{2} \end{pmatrix}. \quad (3)$$

After the spontaneous symmetry breaking, they develop vacuum expectation values as follows:

$$\langle \Phi \rangle = \begin{pmatrix} \kappa/\sqrt{2} & 0 \\ 0 & \kappa' e^{i\alpha}/\sqrt{2} \end{pmatrix}, \quad (4)$$

$$\langle \Delta_L \rangle = \begin{pmatrix} 0 & 0 \\ v_L e^{i\theta_L}/\sqrt{2} & 0 \end{pmatrix}, \quad (5)$$

$$\langle \Delta_R \rangle = \begin{pmatrix} 0 & 0 \\ v_R/\sqrt{2} & 0 \end{pmatrix}. \quad (6)$$

Here, α and θ_L are the spontaneous CP phases, which we assume to vanish for simplicity throughout this work. For a recent review of the model details, see Ref. [43].

The weak eigenstates of charged gauge bosons W_L and W_R mix to form the mass eigenstates:

$$\begin{pmatrix} W_1 \\ W_2 \end{pmatrix} = \begin{pmatrix} \cos \zeta & \sin \zeta \\ -\sin \zeta & \cos \zeta \end{pmatrix} \begin{pmatrix} W_L \\ W_R \end{pmatrix}. \quad (7)$$

The left-right mixing parameter ζ is defined as

$$\sin \zeta = \lambda \sin(2\beta), \quad (8)$$

with $\tan \beta \equiv \kappa'/\kappa$ and $\lambda \simeq m_{W_1}^2/m_{W_2}^2$. The experimental constraints on λ give $\lambda \lesssim 10^{-4}$ [44,45]. Thus, we can safely neglect the left-right mixing² and treat $W_{L,R}$ as $W_{1,2}$, respectively.

The charged-current interactions are described as

$$\mathcal{L}_{CC}^\ell = \frac{g}{\sqrt{2}} \sum_{\ell=e,\mu,\tau} \sum_{i=1}^3 \left[\bar{\ell}_L W_L^- V_{L\ell i} \nu_i + \bar{\ell}_R W_R^- V_{R\ell i} N_i \right] + \text{H.c.}, \quad (9)$$

²For the nontrivial impact of ζ on low-energy observables, see Refs. [46,47].

$$\mathcal{L}_{CC}^q = \frac{g}{\sqrt{2}} \sum_{i,j=1}^3 [\bar{u}_{Li} W_L^+ V_{Lij}^q d_{Lj} + \bar{u}_{Ri} W_R^+ V_{Rij}^q d_{Rj}] + \text{H.c.}, \quad (10)$$

where we have assumed that the gauge couplings are equal, $g_L = g_R \equiv g \simeq 0.65$. In the above, ν_i and N_i ($i = 1, 2, 3$) denote the light and heavy neutrinos, respectively. They are related to the left-handed neutrino $\nu_{L\alpha}$ and right-handed neutrino $\nu_{R\alpha}$ in the flavor basis ($\alpha = e, \mu, \tau$) via lepton flavor mixing³:

$$\nu_{L\alpha} = V_{L\alpha i} \nu_i, \quad \nu_{R\alpha} = V_{R\alpha i} N_i, \quad (11)$$

where $V_L = V_{\text{PMNS}}$ with V_{PMNS} being the Pontecorvo-Maki-Nakagawa-Sakata (PMNS) matrix and V_R is the right-handed counterpart of V_L . V_L^q denotes the Cabibbo-Kobayashi-Maskawa matrix describing flavor mixing in the quark sector, while V_R^q is its right-handed analog. For generalized parity, $V_R^q \simeq V_L^q$ up to a small correction [48], while for generalized charge conjugation the two matrices are equal up to an overall phase [49].

The leptonic Yukawa Lagrangian is also relevant, as it can lead to CLFV processes. The interactions are given by

$$-\mathcal{L}_Y^\ell = \bar{L}_L^\ell \epsilon \Delta_L f_L L_L + \bar{L}_R^\ell \epsilon \Delta_R f_R L_R + \text{H.c.}, \quad (12)$$

where f_L and f_R denote the general Yukawa coupling matrices and $\epsilon \equiv i\tau^2$ with τ^2 the second Pauli matrix.

III. FLAVORED KEUNG-SENJANOVIĆ PROCESS

At the LHC, W_R boson can be produced via the charged-current interactions and decay into a charged lepton and a heavy neutrino. Owing to the Majorana nature of the heavy neutrino, it subsequently decays into a second charged lepton, which can carry either the same or opposite electric charge as the first. This process is known as the Keung-Senjanović process [12], as illustrated in Fig. 1.

We assume that N_1 is the lightest heavy neutrino, which we denote as N_R , and that it is well separated in mass from the other two heavy neutrinos $N_{2,3}$ and lighter than the W_R boson. Thus, N_R is taken to be on shell, with its production and decay factorized in the narrow-width approximation, and interference among different heavy neutrinos is neglected. The parton-level cross section for the production of N_R via $pp \rightarrow \ell_\alpha^\pm N_R$, where ℓ_α denotes the charged lepton of flavor α , is given by⁴ [50]

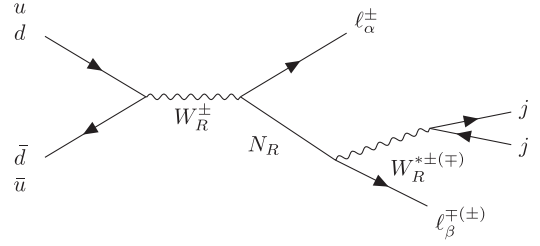


FIG. 1. Feynman diagram of the Keung-Senjanović process at the LHC, illustrating different flavor and electric charge combinations of the charged leptons.

$$\hat{\sigma}_{ij}^{\ell N}(\hat{s}) = \frac{\alpha_W^2 \pi |V_{R\alpha 1}|^2 |V_{Rij}^q|^2}{72 \hat{s}^2} \times \frac{(\hat{s} - m_{N_R}^2)^2 (2\hat{s} + m_{N_R}^2)}{(\hat{s} - m_{W_R}^2)^2 + m_{W_R}^2 \Gamma_{W_R}^2}, \quad (13)$$

where $\alpha_W = g^2/(4\pi)$, $\hat{s}$ denotes the squared partonic center-of-mass energy, and Γ_{W_R} is the decay width of W_R .

The on-shell heavy neutrino N_R decays via the three-body process $N_R \rightarrow \ell_\beta^{\pm} j j'$, mediated by an off-shell W_R boson. The corresponding decay width is [36,51]

$$\Gamma(N_R \rightarrow \ell_\beta^{\pm} j j') = N_c \frac{g^4 |V_{R\beta 1}|^2 m_{N_R}^5}{768 (2\pi)^3 m_{W_R}^4} \sum_{jj'} |V_{Rij'}^q|^2, \quad (14)$$

where N_c is the number of colors and j and j' denote light quarks or antiquarks.⁵ From the expressions in Eqs. (13) and (14), the signal depends on the first column of V_R .

The total decay width sums over all lepton flavors. From the unitarity condition

$$|V_{Re1}|^2 + |V_{R\mu 1}|^2 + |V_{R\tau 1}|^2 = 1, \quad (15)$$

the total width does not depend on the individual mixing entries. Consequently, the event rate for $pp \rightarrow \ell_\alpha \ell_\beta j j'$ is proportional to $|V_{R\alpha 1}|^2 |V_{R\beta 1}|^2$. Final states involving τ leptons are not included in our analysis due to the experimental challenges associated with τ reconstruction, and $V_{R\tau 1}$ enters only indirectly through the unitarity condition in Eq. (15). Thus, for the signal process, the flavor indices are restricted to $\alpha, \beta = e, \mu$.

Recent experimental searches at the LHC have placed stringent bounds on the masses of the W_R boson and heavy neutrino N_R . Using both opposite-sign (OS) and same-sign (SS) dilepton channels, the ATLAS Collaboration [13] excludes $m_{W_R} \lesssim 5\text{--}6$ TeV for $m_{N_R} = 1\text{--}3$ TeV in the

³We work in the charged-lepton mass basis, in which the charged-lepton mass matrices are taken to be diagonal [35]. This choice is without loss of generality in the minimal LRSM for both generalized parity [19] and generalized charge conjugation [15].

⁴As in Ref. [50], we do not apply the narrow-width approximation to the intermediate W_R and retain the full contribution from W_R exchange.

⁵A complementary LHC search targeting the heavy-flavor topology $pp \rightarrow W_R \rightarrow \ell N_R \rightarrow \ell \ell t b$ was proposed in Ref. [52], together with a recast of the existing CMS search at LHC Run 2 for the Keung-Senjanović process in the same-flavor channels [14].

resolved analysis. The OS dilepton channel achieves sensitivity comparable to that of the SS channel in this mass range, while the latter provides a distinctive signature of lepton number violation.

However, most existing studies, including the experimental analyses above, have focused exclusively on same-flavor charged leptons (ee and $\mu\mu$), neglecting the possibility of lepton flavor violation induced by a nontrivial right-handed lepton mixing matrix V_R . In the presence of flavor mixing, charged leptons of different flavors ($e\mu$) also appear in the final state, and the signal rates depend sensitively on the flavor structure of V_R . Moreover, flavor effects significantly affect the experimental bounds on m_{W_R} and m_{N_R} even in the ee and $\mu\mu$ channels, since they alter both the production cross section of N_R and its decay branching ratios.

The expected numbers of SS dilepton events for the Keung-Senjanović process in the final states $e^\pm e^\pm jj$, $\mu^\pm \mu^\pm jj$, and $e^\pm \mu^\pm jj$ are given by

$$N(e^\pm e^\pm) = \mathcal{L} \cdot \epsilon_{ee}^{\text{SS}} \cdot \sigma_{ee}^{\text{SS}} \cdot |V_{e1}|^4, \quad (16a)$$

$$N(\mu^\pm \mu^\pm) = \mathcal{L} \cdot \epsilon_{\mu\mu}^{\text{SS}} \cdot \sigma_{\mu\mu}^{\text{SS}} \cdot |V_{\mu1}|^4, \quad (16b)$$

$$N(e^\pm \mu^\pm) = \mathcal{L} \cdot \epsilon_{e\mu}^{\text{SS}} \cdot \sigma_{ee}^{\text{SS}} \cdot 2 \cdot |V_{e1}|^2 \cdot |V_{\mu1}|^2. \quad (16c)$$

Here, σ_{ee}^{SS} and $\sigma_{\mu\mu}^{\text{SS}}$ are the inclusive cross sections, summed over the two charge configurations, under the assumptions $|V_{e1}| = 1$ and $|V_{\mu1}| = 1$, respectively, which are equal since the masses of e and μ are negligible. The quantities $\epsilon_{ee}^{\text{SS}}$, $\epsilon_{\mu\mu}^{\text{SS}}$, and $\epsilon_{e\mu}^{\text{SS}}$ are the corresponding overall selection efficiencies, and $\mathcal{L}$ is the integrated luminosity. For convenience, we denote the right-handed lepton mixing matrix simply as V (i.e., $V_R \rightarrow V$). In Eq. (16c), the factor of 2 accounts for the two possible flavor sequences in which the electron and muon can appear.

For expected numbers of signal events in the OS dilepton channels e^+e^- , $\mu^+\mu^-$, and $e^\pm\mu^\mp$, analogous expressions hold [35]:

$$N(e^+e^-) = \mathcal{L} \cdot \epsilon_{ee}^{\text{OS}} \cdot \sigma_{ee}^{\text{OS}} \cdot |V_{e1}|^4, \quad (17a)$$

$$N(\mu^+\mu^-) = \mathcal{L} \cdot \epsilon_{\mu\mu}^{\text{OS}} \cdot \sigma_{\mu\mu}^{\text{OS}} \cdot |V_{\mu1}|^4, \quad (17b)$$

$$N(e^\pm\mu^\mp) = \mathcal{L} \cdot \epsilon_{e\mu}^{\text{OS}} \cdot \sigma_{ee}^{\text{OS}} \cdot 2 \cdot |V_{e1}|^2 \cdot |V_{\mu1}|^2, \quad (17c)$$

where the quantities carrying the superscript OS follow definitions similar to their SS counterparts.

IV. COLLIDER ANALYSES

The FeynRules [53] implementation of the LRSM and the corresponding UFO model are publicly available [43,54]. Signal and SM background events for both SS and OS channels are generated with MadGraph5_aMC@NLO [55],

using the NNPDF3.0 parton distribution function set via LHAPDF6 [56]. Within this setup, the decay widths of both N_R and W_R are computed automatically. The following *basic cuts* are applied at the generator level:

$$\begin{aligned} p_{T_j} &> 20 \text{ GeV}, & |\eta_j| &< 2.8, \\ p_{T_\ell} &> 20 \text{ GeV}, & |\eta_\ell| &< 2.5. \end{aligned} \quad (18)$$

Here, p_{T_j} and p_{T_ℓ} denote the transverse momenta of quarks and leptons, respectively, while η_j and η_ℓ are the corresponding pseudorapidities. The parton-level events are passed to Pythia8 [57] for hadronization and Delphes3 [58] for fast detector simulation. In the SS (OS) dilepton channels, the *baseline* selection requires events to contain exactly two charged leptons with the same (opposite) electric charges and at least two jets.

A. Simulation of SM backgrounds

In searches for the Keung-Senjanović process, modeling of SM backgrounds is essential, particularly in the SS dilepton channel, where such backgrounds are rare. The backgrounds contributing to the SS dilepton final state fall into three categories:

- (i) *prompt backgrounds* arising from SM diboson processes ($W^\pm W^\pm$, WZ , ZZ);
- (ii) *fake-lepton backgrounds* originating from jets misreconstructed as leptons or photon conversions ($t\bar{t}$, W + jets); and
- (iii) *charge-flip (CF) backgrounds* arising from SM processes with OS leptons ($t\bar{t}$, Z/γ^* + jets, W^+W^-), where the electric charge of an electron ($e^\pm$) is misidentified.

The dominant processes for each category are indicated in parentheses. These backgrounds have been considered in earlier experimental searches [13,14,25,59,60] and phenomenological studies of TeV-scale lepton number violation at colliders [61,62].

In the OS dilepton channels, prompt backgrounds are dominant. Specifically, we account for backgrounds originating from diboson production (W^+W^- , WZ , ZZ), top pair production ($t\bar{t}$), and single boson production (Z/γ^* + jets) [13,42]. While prompt backgrounds are generated directly via Monte Carlo (MC) simulation, the other backgrounds are modeled as described below.

Instead of the data-driven methods employed in experimental analyses, Ref. [63] proposed a jet-fake approach to estimate fake-lepton backgrounds, which we refer to as jet-fake (JF) backgrounds. This method uses jet-enriched samples to model the fake-lepton backgrounds via two functions: the mistag efficiency and the transfer functions. The mistag efficiency function is defined as

$$\epsilon_{j \rightarrow \ell}(p_{T_j}) = \epsilon_{200} \left[1 - (1 - r_{10}) \frac{200 - p_{T_j}/\text{GeV}}{200 - 10} \right], \quad (19)$$

where $\epsilon_{j \rightarrow \ell}(p_{T_j})$ denotes the probability that a jet with transverse momentum p_{T_j} is mistagged as a prompt lepton. The transfer function, which maps the transverse momentum of the jet p_{T_j} to the reconstructed transverse momentum of the fake lepton p_{T_ℓ} , is parametrized by a truncated Gaussian distribution as follows:

$$T_{j \rightarrow \ell} = \frac{1}{N} \exp \left[-\frac{(\alpha - \mu)^2}{2\sigma^2} \right], \quad (20)$$

where N is the normalization factor and the parameter $\alpha \equiv 1 - p_{T_\ell}/p_{T_j}$ satisfies $0 < \alpha < 1$.

This jet-fake method has been widely used in the literature [62,64–67]. In this work, we implement it as a dedicated module in Delphes3 [62,65] and estimate the dominant backgrounds $t\bar{t}$ and W + jets. The parameters $\{\epsilon_{200}, r_{10}, \mu, \sigma\}$ are then fitted by comparing the simulated differential distributions with those reported by the ATLAS or CMS Collaboration.

Because of bremsstrahlung effects, the electric charge of electrons can be misidentified. Consequently, SM backgrounds arising from charge misidentification of electrons must be taken into account in SS dilepton searches. In contrast, the charge misidentification rate for muons is negligible and is, therefore, ignored. Following the methodology in Refs. [60,68], the charge misidentification probability for electron, $P(p_T, \eta)$, is parametrized as

$$P(p_T, \eta) = \sigma(p_T)f(\eta), \quad (21)$$

where $\sigma(p_T)$ and $f(\eta)$ encode the dependence on the electron transverse momentum and pseudorapidity, respectively. These functions are extracted from a likelihood fit to a dedicated $Z \rightarrow e^+e^-$ data sample [60].⁶

To verify the overall background modeling, we apply the following selection criteria to events from all three background categories [13]:

$$\begin{aligned} p_{T_{\ell_1}} &> 40 \text{ GeV}, & m_{\ell\ell} &> 400 \text{ GeV}, \\ \Delta R(\ell_1, \ell_2) &< 3.9, & m_{jj} &> 110 \text{ GeV}, \\ p_{T_{j_1}} &> 100 \text{ GeV}, & p_{T_{j_2}} &> 100 \text{ GeV}, \\ H_T \equiv p_{T_{j_1}} + p_{T_{j_2}} + p_{T_{\ell_1}} + p_{T_{\ell_2}} &> 400 \text{ GeV}. \end{aligned} \quad (22)$$

⁶As in Ref. [62], we validate this modeling by comparing the ratio of SS to OS event yields from our simulated $Z \rightarrow e^+e^-$ events with that from the ATLAS results and find overall good agreement, with a relative difference of about 17%. This provides an estimate of the associated systematic uncertainty, consistent with the evaluation in Ref. [68], where deviations of up to about 20% in the charge-misidentification rate were found in some pseudorapidity regions for the adopted parametrization.

Here, $p_{T_{\ell_1}}$ is the transverse momentum of the leading lepton, while $p_{T_{j_1}}$ and $p_{T_{j_2}}$ are the transverse momenta of the leading and subleading jets, respectively. Furthermore, $\Delta R(\ell_1, \ell_2)$ denotes the angular separation between the two leptons, $m_{\ell\ell}(m_{jj})$ is the dilepton (dijet) invariant mass, and H_T is the scalar sum of the transverse momenta of the two leading leptons and two leading jets.

Using Eq. (22) and including K factors for SM processes [42], we compare the differential distributions of the SM backgrounds in the $e^\pm e^\pm$ and $\mu^\pm \mu^\pm$ channels with those reported by ATLAS. As shown in Fig. 2 for the $e^\pm e^\pm$ channel, the H_T distributions⁷ for the three dominant background categories are consistent with the ATLAS results in both shape and normalization, using the fitted parameters $\{\epsilon_{200}, r_{10}, \mu, \sigma\} = \{10^{-3}, 1.0, 0.8, 0.2\}$. Similar consistencies are observed for the $\mu^\pm \mu^\pm$ channel, where we obtain the fitted parameters $\{\epsilon_{200}, r_{10}, \mu, \sigma\} = \{4 \times 10^{-5}, 1.0, 0.8, 0.2\}$. The systematic uncertainty associated with the modeling of fake-lepton backgrounds in the JF method is moderate, as indicated by Ref. [63], which found an $\mathcal{O}(10\%)$ variation when these parameters were varied.

The validation of our modeling methods in the same-flavor channels ($ee, \mu\mu$) corroborates the reliability of the simulation across different flavors. Thus, it is straightforward to simulate SM backgrounds in the $e\mu$ channel. Furthermore, instead of applying selection cuts sequentially in the cut-based analysis [13], we will explore LHC sensitivities to the masses of W_R and N_R and the right-handed lepton flavor mixing using machine learning.

We emphasize that the SM backgrounds considered in this work are those for the resolved analysis, where two charged leptons are reconstructed. In the ATLAS analyses [13], the boosted channel is also studied and is particularly effective in the regime $m_{W_R} \gg m_{N_R}$. By contrast, for $m_{N_R} \gtrsim 2.5$ TeV, the resolved analysis yields stronger bounds than the boosted analysis. While lepton flavor effects could, in principle, be incorporated into the boosted channel, such an analysis would require a dedicated treatment and lies beyond the scope of this work. Hence, both the comparison with the ATLAS limits and our flavor-dependent results, including the HL-LHC projected sensitivities, are restricted to the resolved topology.

B. Machine-learning analysis

In high-energy experiments, conventional cut-based analysis relies on sequential cuts applied to a limited set of kinematic variables, which may overlook complex correlations between observables. To more fully exploit the information contained in collision events, we employ a

⁷As additional checks on the background estimation, comparisons with the earlier experimental analyses [25,59,60] are performed and also show consistency.

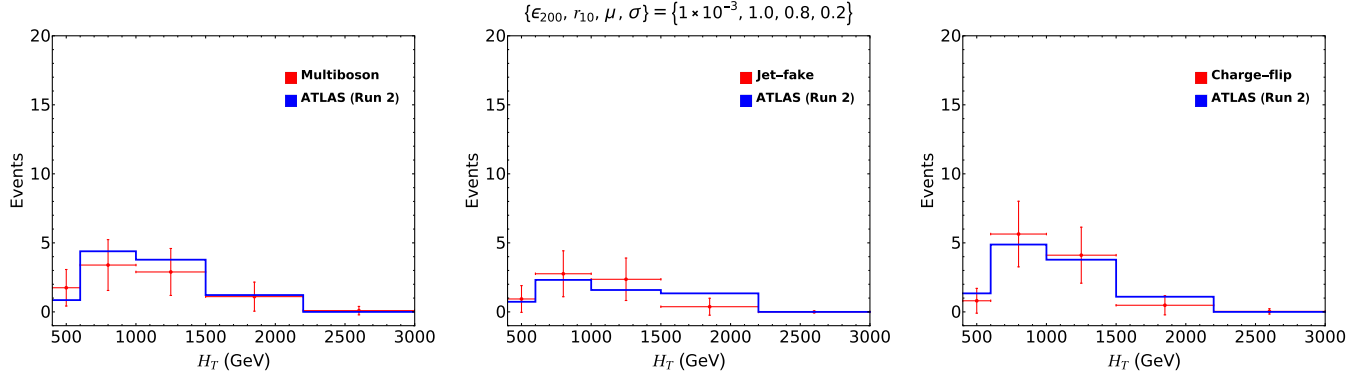


FIG. 2. Comparison of the H_T distributions in the $e^\pm e^\pm$ channel between the simulated backgrounds (red lines) and the ATLAS results (blue lines) from Fig. 7 in Ref. [13]. The left panel shows the multiboson (i.e., prompt) backgrounds; the middle and right panels show the jet-fake and charge-flip backgrounds, respectively. The error bars in each bin denote the statistical uncertainties $\sqrt{N_i}$, where N_i is the number of events in the i th bin. The fitted parameters are $\{\epsilon_{200}, r_{10}, \mu, \sigma\} = \{10^{-3}, 1.0, 0.8, 0.2\}$.

deep neural network (DNN) [39–41] as a classifier to distinguish rare signal events from SM background events.

A DNN is a class of artificial neural networks with multiple hidden layers [39]. In collider analyses, it takes kinematic observables as inputs and learns nonlinear representations that serve as powerful discriminants between signal and background. After training via back-propagation, a selection on the output score yields a signal-enriched sample. Owing to their ability to capture complex correlations in high-dimensional phase space, DNNs outperform conventional cut-based approaches, as is well established in the literature (see Refs. [40,41,69]).

We consider the following variables as training inputs [13,42]:

$$p_{T_{j_1}}, p_{T_{j_2}}, p_{T_{\ell_1}}, m_{jj}, m_{\ell\ell}, H_T, \Delta R(\ell_1, \ell_2), \quad (23)$$

using the same notation as in Eq. (22). Their distributions, shown in Fig. 7, demonstrate that signal events have harder spectra and longer tails than backgrounds—particularly for the heavier mass benchmark—providing a solid basis for discrimination.

Given the wide range in (m_{W_R}, m_{N_R}) , we divide the parameter space into two regions: low-mass region ($m_{W_R} \leq 3$ TeV) and high-mass region ($m_{W_R} > 3$ TeV). We consider two representative benchmark points:

- (i) BP1. $m_{W_R} = 2$ TeV, $m_{N_R} = 0.5$ TeV;
- (ii) BP2. $m_{W_R} = 4$ TeV, $m_{N_R} = 2$ TeV.

Two separate DNN classifiers are trained for the low- and high-mass regimes, using BP1 and BP2, respectively, with an identical set of input variables. To incorporate flavor effects, we adopt the benchmark $V_R = V_{\theta=\pi/4}$, where the mixing matrix in the $e\text{--}\mu$ sector is defined as

$$V_\theta = \begin{pmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix}. \quad (24)$$

Nevertheless, the same selection efficiencies apply to other mixing patterns, as the selection criteria are determined by the masses m_{W_R} and m_{N_R} . Details of the DNN implementation and validation are provided in Appendix B.

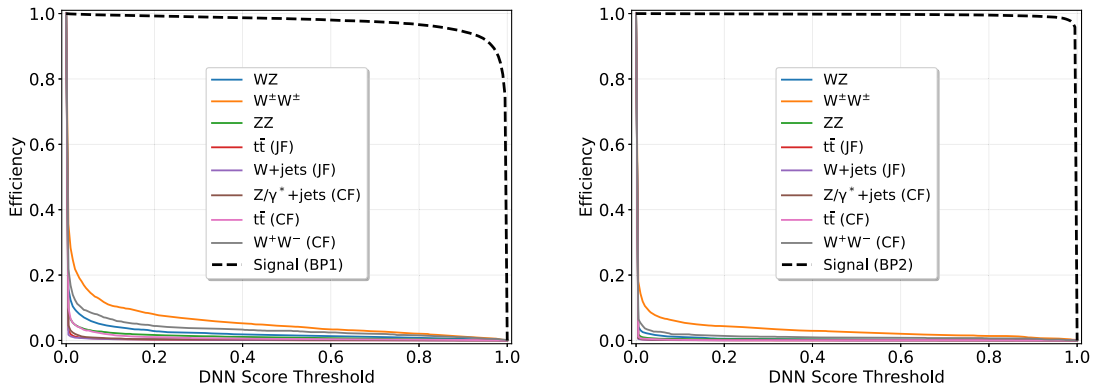


FIG. 3. Signal and background efficiencies for the SS dilepton searches as a function of the DNN score threshold. The left and right panels correspond to the low-mass and high-mass benchmarks, respectively.

TABLE I. Cross sections for the SS dilepton signal and background processes in the case of low-mass model. Values are shown at three stages: after the basic cuts (σ_{basic}), after the baseline selection ($\sigma_{\ell\ell jj}$), and after the optimal DNN score selection (σ_{DNN}). The bottom rows provide the comparison of the total background and signal cross sections.

Process	σ_{basic} (pb)	$\sigma_{\ell\ell jj}$ (pb)	σ_{DNN} (pb)
WZ	4.31×10^{-1}	1.93×10^{-2}	7.74×10^{-5}
$W^\pm W^\pm$	1.17×10^{-2}	2.59×10^{-3}	3.54×10^{-5}
ZZ	4.15×10^{-2}	1.75×10^{-3}	3.15×10^{-6}
$t\bar{t}$ (JF)	1.08×10^{-1}	5.33×10^{-3}	8.03×10^{-6}
W + jets (JF)	9.47×10^{-1}	2.14×10^{-2}	2.96×10^{-5}
$t\bar{t}$ (CF)	1.17×10^{-1}	2.42×10^{-2}	5.47×10^{-6}
$Z/\gamma^* + \text{jets}$ (CF)	2.71×10^{-1}	1.27×10^{-1}	1.39×10^{-5}
$W^+ W^-$ (CF)	1.11×10^{-3}	2.28×10^{-4}	2.53×10^{-6}
Total bkg	1.93	2.02×10^{-1}	1.75×10^{-4}
Signal	5.29×10^{-2}	2.49×10^{-2}	2.36×10^{-2}

Figure 3 shows the selection efficiencies as a function of the DNN score threshold for the low-mass and high-mass models in the SS dilepton searches. The jet-fake and $Z/\gamma^* + \text{jets}$ backgrounds are efficiently suppressed at low thresholds due to their softer kinematics, while the $W^\pm W^\pm$, $W^+ W^-$ (CF), and WZ backgrounds have harder spectra and, therefore, larger surviving fractions. The signal efficiency for the high-mass model decreases slowly up to a threshold that is higher than for the low-mass model, while background efficiencies drop rapidly, resulting in a clean separation.

To establish the optimal selection criteria and derive exclusion limits, the statistical significance Z is evaluated using the asymptotic profile likelihood ratio test for the signal hypothesis [70,71]

$$Z = \sqrt{2 \left(S - B \ln \left(\frac{S+B}{B} \right) \right)}, \quad (25)$$

where S and B denote the number of signal and background events remaining after the DNN selection, respectively. The number of SS signal events in each channel after the optimal DNN selection is given in Eq. (16). For the $\ell^\pm \ell^\pm$ channel, the significance is computed by combining the events from the $e^\pm e^\pm$, $\mu^\pm \mu^\pm$, and $e^\pm \mu^\pm$ channels, with $S = N_{e^\pm e^\pm} + N_{\mu^\pm \mu^\pm} + N_{e^\pm \mu^\pm}$, while B includes all SM background contributions with a pair of SS charged leptons [25].

By maximizing the expected significance Z , we obtain optimal DNN score thresholds of 0.91 for the low-mass model and 0.99 for the high-mass model in the SS dilepton searches. Table I summarizes the cross sections for the signal and SM backgrounds at different selection stages for the low-mass model. The three stages are σ_{basic} after

TABLE II. Event yields after the optimal DNN score selection for the low-mass model. The backgrounds (individual categories and total) and the SS dilepton signal for BP1 are shown for the three SS dilepton channels. The ellipses indicate that the corresponding background contribution is negligible. The last row gives the signal efficiency in each channel.

Process	$e^\pm \mu^\pm$	$e^\pm e^\pm$	$\mu^\pm \mu^\pm$
Multiboson	6	4	6
Jet-fake	3	3	...
Charge-flip	...	3	...
Total bkg	9	10	6
Signal	1615	641	1027
Eff [%]	44	35	56

applying the basic cuts at the generator level [cf., Eq. (18)], $\sigma_{\ell\ell jj}$ after imposing the baseline selection, and σ_{DNN} after requiring the optimal threshold of 0.91. The signal cross section remains robust throughout the selections, while the total background is suppressed by nearly 4 orders of magnitude relative to the generator level, leading to a significantly enhanced signal-to-background ratio.

To evaluate flavor-specific sensitivities, Table II provides the event yields after the optimal DNN selection, which correspond to an integrated luminosity of 139 fb^{-1} at LHC Run 2. The last row lists the corresponding selection efficiencies, i.e., the fraction of events that survive the DNN selection: $\epsilon_{\mu\mu} = 56\%$, $\epsilon_{e\mu} = 44\%$, and $\epsilon_{ee} = 35\%$. Thus, the $\mu^\pm \mu^\pm$ channel has the highest efficiency, followed by $e^\pm \mu^\pm$ and $e^\pm e^\pm$. This difference in efficiencies is primarily attributed to the better reconstruction performance of muons compared to electrons at the LHC. We note that the jet-fake backgrounds in the $e^\pm \mu^\pm$ channel arise primarily from $j \rightarrow e$ rather than from $j \rightarrow \mu$.

TABLE III. The same as Table I but for the high-mass model with the optimal DNN score threshold of 0.99. After the DNN selection, only the WZ and $W^\pm W^\pm$ backgrounds are non-negligible; all other backgrounds are suppressed to $\lesssim 10^{-7}$ pb or below.

Process	σ_{basic} (pb)	$\sigma_{\ell\ell jj}$ (pb)	σ_{DNN} (pb)
WZ	4.31×10^{-1}	1.93×10^{-2}	4.18×10^{-6}
$W^\pm W^\pm$	1.17×10^{-2}	2.59×10^{-3}	6.64×10^{-6}
ZZ	4.15×10^{-2}	1.75×10^{-3}	$\lesssim 10^{-7}$
$t\bar{t}$ (JF)	1.08×10^{-1}	5.33×10^{-3}	$\lesssim 10^{-7}$
W + jets (JF)	9.47×10^{-1}	2.14×10^{-2}	$\lesssim 10^{-7}$
$t\bar{t}$ (CF)	1.17×10^{-1}	2.42×10^{-2}	$\lesssim 10^{-8}$
$Z/\gamma^* + \text{jets}$ (CF)	2.71×10^{-1}	1.27×10^{-1}	$\lesssim 10^{-7}$
$W^+ W^-$ (CF)	1.11×10^{-3}	2.28×10^{-4}	$\lesssim 10^{-8}$
Total bkg	1.93	2.02×10^{-1}	1.08×10^{-5}
Signal	4.24×10^{-4}	2.30×10^{-4}	2.23×10^{-4}

TABLE IV. The same as Table II but for the high-mass model with the optimal DNN score threshold of 0.99.

Process	$e^\pm\mu^\pm$	$e^\pm e^\pm$	$\mu^\pm\mu^\pm$
Multiboson	0.67	0.49	0.34
Total bkg	0.67	0.49	0.34
Signal	15.37	6.44	9.17
Eff [%]	52.13	43.58	62.16

Table III shows the cross sections at each stage for the high-mass model in the SS dilepton channel. After the DNN selection, only the WZ and $W^\pm W^\pm$ backgrounds have non-negligible cross sections; all other backgrounds are suppressed to $\lesssim 10^{-7}$ pb or smaller. Table IV presents the corresponding event yields in the three SS dilepton channels. Compared to the low-mass case (cf., Table II), the high-mass model yields significantly fewer background events after the DNN selection.⁸

An analogous procedure is applied to the OS dilepton channels, as detailed in Appendix A. Although the OS channels are subject to significantly larger backgrounds at the generator level, the DNN selection effectively suppresses these contributions, making them comparable to those in the SS dilepton channels. Consequently, the resulting signal significance in the $\ell^+\ell^-$ channel for the low-mass (high-mass) benchmarks is $Z = 77.3(6.3)$, which is close to $Z = 79.5(7.3)$ achieved in the SS dilepton channels.

We perform a parameter scan over the m_{W_R} - m_{N_R} plane for a given right-handed lepton mixing matrix V_R , dividing the analysis into low- and high-mass regimes. For the low-mass regime, we use the low-mass model; for the high-mass regime, we use a model optimized to exploit the distinct kinematics of heavy W_R and N_R decays. For each parameter point, we compute the SS and OS significances using Eq. (25). The combined significance for the SS and OS channels (denoted as SS + OS) is then obtained by taking the square root of the sum of the squares of the individual significances [70]. Exclusion limits at 95% confidence level (CL) for the SS, OS, and SS + OS analyses are derived by requiring the corresponding significance to reach 1.96.

V. LHC SENSITIVITIES

In this section, we evaluate the sensitivity to the Keung-Senjanović process at LHC Run 2, as well as the projected reach at the HL-LHC. Both the sensitivities from SS dilepton searches and their combination with OS dilepton searches are investigated.

⁸For an integrated luminosity of 139 fb^{-1} , some expected event yields are below unity; we show all yields with two decimal places for uniform precision.

To assess the impact of flavor structure of the right-handed lepton mixing matrix, we consider three benchmark mixing scenarios:

- (i) *Unmixed*. $V_{e1} = 1$ or $V_{\mu 1} = 1$, representing the case of no lepton flavor mixing.
- (ii) *Maximal mixing*. $V_R = V_{\theta=\pi/4}$, corresponding to maximal mixing in the e - μ sector.
- (iii) *PMNS-like mixing*. $V_R = V_{\text{PMNS}}^{(*)}$, as commonly adopted in the minimal LRSM with generalized parity and charge conjugation [44].

The signal event yields for the ee and $\mu\mu$ channels scale as $|V_{e1}|^4$ and $|V_{\mu 1}|^4$, respectively. For maximal mixing, $|V_{e1}| = |V_{\mu 1}| = 0.707$. For the PMNS-like pattern, the latest global fits for normal mass ordering give $|V_{e1}| = 0.82$ and $|V_{\mu 1}| = 0.33$ (central values) [72–74]. Consequently, while the scaling factors are symmetric under maximal mixing, the PMNS-like pattern leads to a significantly enhanced ee yield at the expense of a suppressed $\mu\mu$ contribution.

A. Exclusion limits on W_R and N_R masses

The sensitivity of individual channels in both SS and OS dilepton searches depends strongly on the mixing benchmark. In the ee and $\mu\mu$ channels, the signal cross sections scale with $|V_{e1}|^4$ and $|V_{\mu 1}|^4$, respectively, yielding the strongest limits in the unmixed scenario. For the maximal-mixing scenario, the $e\mu$ channel has approximately twice the event yield of the ee or $\mu\mu$ channels [cf., Eqs. (16) and (17)], making it the most sensitive individual channel. In contrast, for the PMNS-like scenario, the sensitivities follow the order $ee > e\mu > \mu\mu$, given the mixing entries.

The combined 95% CL constraints from the SS + OS dilepton searches are shown in Fig. 4 (see Fig. 8 for those from SS dilepton searches). The solid and dashed curves show the LHC Run 2 results and the HL-LHC projections, respectively. The orange curves correspond to the exclusion limits reported by ATLAS in the resolved analysis of the Run 2 data [13].⁹ For the unmixed ee and $\mu\mu$ channels, the expected sensitivities from our DNN analysis (red curves) are stronger than the corresponding ATLAS results, owing to improved signal efficiency and background rejection in the ML approach.

We further combine all flavor contributions in the $\ell\ell$ channel ($\ell = e, \mu$), where the total signal cross section scales as $(|V_{e1}|^2 + |V_{\mu 1}|^2)^2$. Thus, flavor mixing mainly redistributes the signal among the different dilepton final states. In the maximal-mixing scenario ($|V_{e1}| = |V_{\mu 1}| = 0.707$), the total signal cross section

⁹The ATLAS analysis derives upper limits by combining all signal regions without distinguishing between the SS and OS channels. Our limits are obtained using two-sided confidence intervals and without a binned likelihood analysis, both of which lead to a conservative estimate of the sensitivity of the DNN analysis.

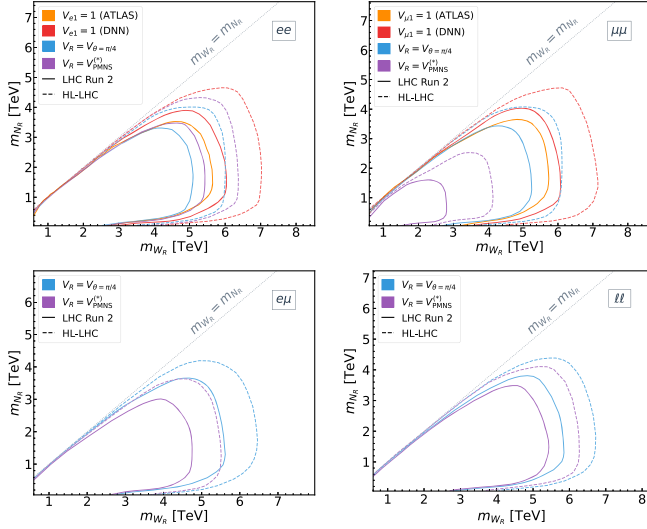


FIG. 4. 95% CL constraints on m_{W_R} and m_{N_R} from combined SS and OS dilepton searches at LHC Run 2 (solid lines) and HL-LHC (dashed lines). The panels correspond to the ee (upper left), $\mu\mu$ (upper right), $e\mu$ (lower left), and combined $\ell\ell$ ($\ell = e, \mu$) (lower right) channels. The blue and purple contours represent the maximal-mixing scenario ($V_R = V_{\theta=\pi/4}$) and the PMNS-like scenario ($V_R = V_{\text{PMNS}}^{(*)}$), respectively. In the ee and $\mu\mu$ channels, the red contours indicate the unmixed cases ($V_{e1} = 1$ or $V_{\mu1} = 1$) obtained with the DNN analysis, while the orange contours indicate the ATLAS results. The gray dotted line in each panel marks the kinematic threshold $m_{W_R} = m_{N_R}$.

coincides with that of either unmixed case ($V_{e1} = 1$ or $V_{\mu1} = 1$) before any selection cuts. However, the combined $\ell\ell$ channel receives background contributions from all three flavor channels, leading to a larger total background yield than in any single-flavor channel. Consequently, the resulting exclusion is weaker than in the individual unmixed cases. In the PMNS-like scenario, $|V_{e1}|^2 + |V_{\mu1}|^2 < 1$; thus, the signal rate is further suppressed, yielding limits that are less stringent than those in the maximal-mixing case.

Using the LHC Run 2 results as a reference, we project the sensitivity at the HL-LHC. The increased luminosity substantially extends the accessible mass range. Under maximal (PMNS-like) mixing, the HL-LHC is expected to probe m_{W_R} up to 6.7 (6.3) TeV for $m_{N_R} = 2$ TeV, and m_{N_R} up to 4.4 (4.1) TeV for $m_{W_R} = 5.6$ (5.4) TeV in the $\ell\ell$ channel. These reaches in m_{W_R} and m_{N_R} are parameter-dependent projections of the sensitivity contours in the (m_{W_R}, m_{N_R}) plane rather than independent end points. Similar improvements are observed in the individual ee , $\mu\mu$, and $e\mu$ channels, as shown in Fig. 4.

B. Bounds on right-handed lepton flavor mixing

We now investigate bounds on the mixing entries $|V_{e1}|$ and $|V_{\mu1}|$ at LHC Run 2 and the HL-LHC. The relations in

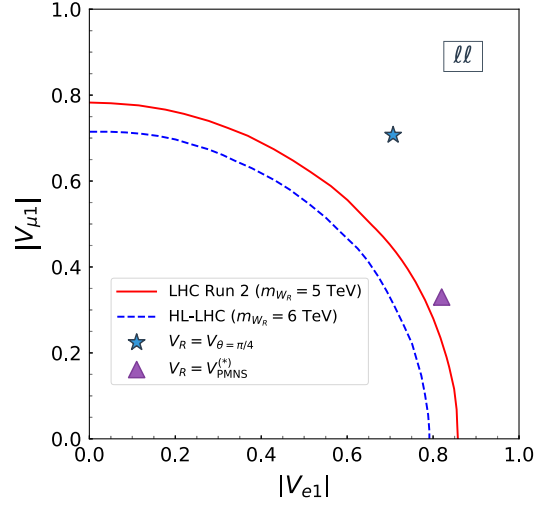


FIG. 5. 95% CL constraints on $|V_{e1}|$ and $|V_{\mu1}|$ for fixed $m_{N_R} = 3$ TeV in the combined SS + OS dilepton channels. The red solid contour shows the limit from LHC Run 2 for $m_{W_R} = 5$ TeV, while the blue dashed contour indicates the HL-LHC sensitivity for $m_{W_R} = 6$ TeV. The star and triangle mark the values of $|V_{e1}|$ and $|V_{\mu1}|$ for the maximal-mixing and PMNS-like scenarios, respectively.

Eqs. (16) and (17) imply that the LHC bounds on $|V_{e1}|$ and $|V_{\mu1}|$ are independent of the assumption of $V_{\tau1}$, and the allowed region in the $|V_{e1}|$ - $|V_{\mu1}|$ plane lies within a quarter circle. Different reconstruction efficiencies of electrons and muons slightly distort this boundary, giving a mild asymmetry.

Figure 5 shows the constraints in the $|V_{e1}|$ - $|V_{\mu1}|$ plane for two benchmark points: $(m_{W_R}, m_{N_R}) = (5 \text{ TeV}, 3 \text{ TeV})$ at LHC Run 2 and $(6 \text{ TeV}, 3 \text{ TeV})$ at the HL-LHC. For the LHC Run 2 benchmark, the sensitivity contour is characterized by $\sqrt{|V_{e1}|^2 + |V_{\mu1}|^2} \sim 0.8$ – 0.85 . For the HL-LHC benchmark, the increased center-of-mass energy and integrated luminosity compensate for the impact of the larger m_{W_R} . Consequently, the sensitivity extends to smaller values of $|V_{e1}|$ and $|V_{\mu1}|$, reaching $\sqrt{|V_{e1}|^2 + |V_{\mu1}|^2} \sim 0.7$ – 0.8 , thereby probing a larger portion of the flavor-mixing plane. In particular, by combining the SS and OS searches, our simulated analysis has the expected sensitivity to exclude the maximal-mixing and PMNS-like benchmark points for the chosen masses.

VI. COMPLEMENTARITIES WITH CLFV SEARCHES

The CLFV processes $\mu \rightarrow e\gamma$, $\mu \rightarrow eee$ and $\mu \rightarrow e$ conversion in nuclei are highly suppressed in the SM. In the LRSM, these processes receive additional contributions from right-handed charged currents as well as from the Yukawa interactions between charged leptons

and the right-handed triplet scalars [cf., Eq. (12)]. These contributions have been extensively studied in Refs. [6,16,35,75–78]. Representative Feynman diagrams are shown in Appendix C.

The current 90% CL upper limits on the branching ratios are given by [79–81]

$$\text{Br}(\mu \rightarrow e\gamma) < 1.5 \times 10^{-13} \quad (\text{MEG II}), \quad (26)$$

$$\text{Br}(\mu \rightarrow eee) < 1.0 \times 10^{-12} \quad (\text{SINDRUM}), \quad (27)$$

$$R_{\text{Au}}(\mu \rightarrow e) < 7.0 \times 10^{-13} \quad (\text{SINDRUM II}). \quad (28)$$

Future experiments are expected to significantly improve these sensitivities. The projected reaches are [79,82,83]

$$\text{Br}(\mu \rightarrow e\gamma) \lesssim 6.0 \times 10^{-14} \quad (\text{MEG II}), \quad (29)$$

$$\text{Br}(\mu \rightarrow eee) \lesssim 2.0 \times 10^{-15} \quad (\text{Mu3e phase I}), \quad (30)$$

$$R_{\text{Al}}(\mu \rightarrow e) \lesssim 6.2 \times 10^{-16} \quad (\text{Mu2e run I}). \quad (31)$$

In addition to W_R and N_i , the charged scalars $\Delta_{L,R}^{\pm\pm}$ and $\Delta_L^\pm$ also contribute to the CLFV processes. Assuming the charged scalar masses are degenerate with W_R [75], the branching ratios take the form given in Eqs. (C1), (C2), and (C7).

If only one heavy neutrino contributes, the resulting constraints are stringent. In contrast, when two or more heavy neutrinos are nearly degenerate, large cancellation can occur, significantly suppressing the CLFV branching ratios compared to the individual contributions [35]. Accordingly, we assume maximal mixing and a relative mass splitting $(m_{N_2} - m_{N_1})/m_{N_1} = 0.01$ to facilitate a direct comparison between the mass reaches of collider searches and CLFV experiments. Under these assumptions, interference effects at the LHC can be neglected [35,84,85], and the signal cross section for the Keung-Senjanović process is approximately twice that for a single N_1 . Such an enhancement is included in the blue LHC sensitivity curves in Fig. 6. This near-degenerate heavy-neutrino spectrum differs from that assumed in Secs. V and III, where N_1 is well separated from $N_{2,3}$.

Using the current limits and projections of CLFV experiments, we derive constraints on m_{W_R} and $m_{N_R} \equiv m_{N_1}$ for the benchmark $m_{\Delta_{L,R}^{\pm\pm}} = m_{\Delta_L^\pm} = m_{W_R}$, as illustrated in Fig. 6. The existing CLFV bounds (red shaded regions) predominantly constrain the domain $m_{N_R} > m_{W_R}$.¹⁰ In combination with the LHC Run 2 exclusion limit obtained from the SS + OS searches in the $e\mu + jj$ channel (solid blue curve), these constraints exclude most of the

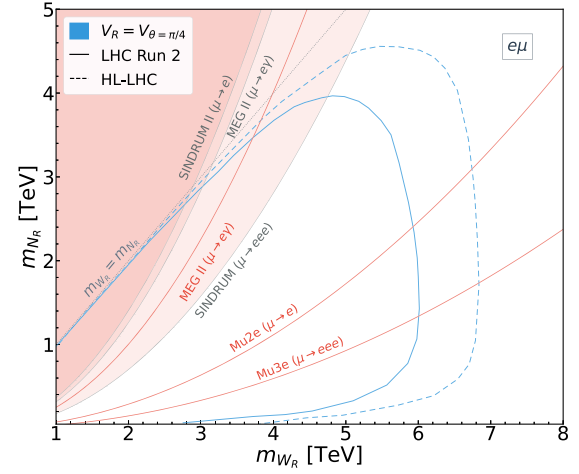


FIG. 6. Comparison of LHC and CLFV constraints on m_{W_R} and $m_{N_R} \equiv m_{N_1}$ in the maximal-mixing scenario, assuming $m_{\Delta_{L,R}^{\pm\pm}} = m_{\Delta_L^\pm} = m_{W_R}$ and $(m_{N_2} - m_{N_1})/m_{N_1} = 0.01$. Blue solid and dashed curves show 95% CL exclusion limits for the $e\mu + jj$ channel at the LHC and HL-LHC, respectively. Current 90% CL bounds from CLFV experiments are shown as red shaded regions with gray boundaries, while red curves indicate future projections. The gray dotted line corresponds to $m_{W_R} = m_{N_R}$.

parameter space with $m_{W_R} \lesssim 6$ TeV, irrespective of the value of m_{N_R} . Projected sensitivities of future CLFV experiments (red curves) are expected to cover a large fraction of the region $m_{W_R} > m_{N_R}$. Notably, they may surpass the HL-LHC reach (dashed blue curve), thereby probing $m_{W_R} \gtrsim 7$ TeV.

Finally, we note that neutrinoless double beta ($0\nu\beta\beta$) decay provides another probe of the LRSM. The short-range contribution to the $0\nu\beta\beta$ decay amplitude scales as $\sim m_p(m_W/m_{W_R})^4 \sum_i (V_{ei})^2/m_{N_i}$ [16], where m_p denotes the nucleon mass. Because of the inverse dependence on m_{N_i} , this contribution is dominated by the lightest of the three right-handed neutrinos. As shown in Refs. [46,88], for TeV-scale W_R and N_R masses, the $0\nu\beta\beta$ -decay constraints from current and future tonne-scale experiments are weaker than those from direct LHC searches for the Keung-Senjanović process [50] and are, therefore, not included in our combined analysis.

VII. CONCLUSION

In this work, we have performed a comprehensive study of the Keung-Senjanović process $pp \rightarrow W_R \rightarrow \ell_\alpha N_R \rightarrow \ell_\alpha \ell_\beta jj$ with $\ell_{\alpha,\beta} = e, \mu$ in the minimal LRSM at LHC Run 2 and the HL-LHC, with particular emphasis on the role of right-handed lepton flavor mixing. To enhance signal sensitivity, we employed a deep neural network optimized for different mass regimes, achieving significantly improved signal efficiency and background rejection over conventional cut-based analyses.

¹⁰The vacuum stability and perturbativity bounds on the right-handed Majorana neutrino masses in the LRSM were investigated in Refs. [86,87].

The expected sensitivity depends strongly on the mixing pattern. Compared to the unmixed case, flavor mixing weakens the exclusion limits in the ee and $\mu\mu$ channels. Under maximal mixing the $e\mu$ channel is the most sensitive individual channel, while under PMNS-like mixing the ee channel provides the strongest sensitivity. Combining all dilepton channels further improves the reach, especially in the maximal-mixing scenario.

At LHC Run 2, the combined $\ell\ell$ channel already provides significant constraints on the parameter space. The HL-LHC, with its increased luminosity, substantially extends the accessible mass range. Under maximal (PMNS-like) mixing, the HL-LHC can exclude m_{W_R} up to 6.7 (6.3) TeV for $m_{N_R} = 2$ TeV and m_{N_R} up to 4.4 (4.1) TeV for $m_{W_R} = 5.6$ (5.4) TeV. These limits are substantially stronger than those obtained at Run 2.

We have also explored generic mixing configurations in the $|V_{e1}|$ - $|V_{\mu1}|$ plane. At the LHC Run 2 benchmark point $(m_{W_R}, m_{N_R}) = (5 \text{ TeV}, 3 \text{ TeV})$, the expected sensitivity of the ee , $e\mu$, and $\mu\mu$ channels decreases as $|V_{\tau1}|$ increases, leading to an exclusion boundary approximately characterized by $\sqrt{|V_{e1}|^2 + |V_{\mu1}|^2} \sim 0.8$ – 0.85 for the SS + OS dilepton analysis. At the HL-LHC, the expected sensitivity is significantly extended, probing down to $\sqrt{|V_{e1}|^2 + |V_{\mu1}|^2} \sim 0.7$ – 0.8 for heavier benchmark masses.

Finally, we investigated complementary constraints from low-energy CLFV observables. We consider a benchmark scenario with degenerate charged scalars ($m_{\Delta_{L,R}^-} = m_{\Delta_{L,R}^+} = m_{W_R}$) and a relative mass splitting $(m_{N_2} - m_{N_1})/m_{N_1} = 0.01$. Under these assumptions, current CLFV bounds, when combined with LHC Run 2 searches, exclude $m_{W_R} \lesssim 6$ TeV for any m_{N_R} . Future CLFV experiments will extend this reach, probing $m_{W_R} \gtrsim 7$ TeV and potentially exceeding the HL-LHC sensitivity. These results demonstrate the strong complementarity between direct collider searches and indirect CLFV probes in testing the LRSM parameter space.

In summary, our results highlight the essential role of right-handed lepton flavor mixing in interpreting collider searches for right-handed gauge bosons and neutrinos. The combination of machine learning techniques with an explicit treatment of e, μ flavor effects provides an effective strategy for probing TeV-scale new physics at current and future experiments.

ACKNOWLEDGMENTS

We thank Frank Deppisch, Bingxuan Liu, Michihisa Takeuchi, and Lailin Xu for helpful discussions and comments. This work is supported by the National Natural Science Foundation of China under Grants No. 12347105 and No. 12505127 and the Guangdong Basic and Applied Basic Research Foundation (2024A1515012668).

DATA AVAILABILITY

There are no publicly available research data or software supporting this manuscript. Requests for further information or data should be sent to the authors.

APPENDIX A: MORE DETAILS OF ML ANALYSIS

The normalized distributions of the input variables for the SS dilepton signal and the corresponding SM backgrounds are shown in Fig. 7. Signal events populate higher-value regions of several observables compared to backgrounds, a pattern that the DNN captures during training. This is especially pronounced for the higher-mass benchmark, providing correlated information that enhances discrimination.

The expected exclusion limits on m_{W_R} and m_{N_R} in the SS dilepton searches at LHC Run 2 are shown in Fig. 8. The sensitivity exhibits the same dependence on mixing scenarios as in the combined SS + OS analysis, as discussed in Sec. VA. Figure 9 shows the constraints in the $|V_{e1}|$ - $|V_{\mu1}|$ plane for the SS dilepton channel. The exclusion boundary is characterized by $\sqrt{|V_{e1}|^2 + |V_{\mu1}|^2} \sim 0.85$ – 0.9 at LHC Run 2 and ~ 0.75 – 0.8 at the HL-LHC. With only SS searches, the maximal-mixing scenario is excluded at both LHC Run 2 and the HL-LHC, whereas the PMNS-like scenario is not excluded by Run 2 but would require HL-LHC sensitivity. This differs from the combined SS + OS analysis (cf., Fig. 5).

The OS dilepton channel shares the same signal topology and kinematic features as the SS channel, so the same set of input variables [cf., Eq. (23)] is used for the DNN classifier. As in the SS analysis, two benchmark models (BP1 and BP2) are employed to ensure stable performance across the parameter space. However, the background composition differs: The $t\bar{t}$ and $Z/\gamma^* + \text{jets}$ processes contribute directly in the OS channel, leading to a substantially larger number of background events after the baseline selection.

Figure 10 shows the selection efficiencies as a function of the DNN score threshold for the low-mass and high-mass models in the OS dilepton searches. For the low-mass benchmark, the DNN achieves strong background rejection while maintaining high signal efficiency. A similar analysis is performed for the high-mass benchmark.

By maximizing the expected significance Z , we obtain optimal DNN score thresholds of 0.99 for both the low-mass and high-mass models in the OS dilepton searches. Table V presents the cross sections after each selection stage for the OS dilepton channel. As in the SS channel, signal efficiency remains high while backgrounds are strongly suppressed. Despite larger contributions from $t\bar{t}$ and $Z/\gamma^* + \text{jets}$, the DNN reduces the total background by nearly 5 orders of magnitude, making the OS background comparable to the SS after the selections.

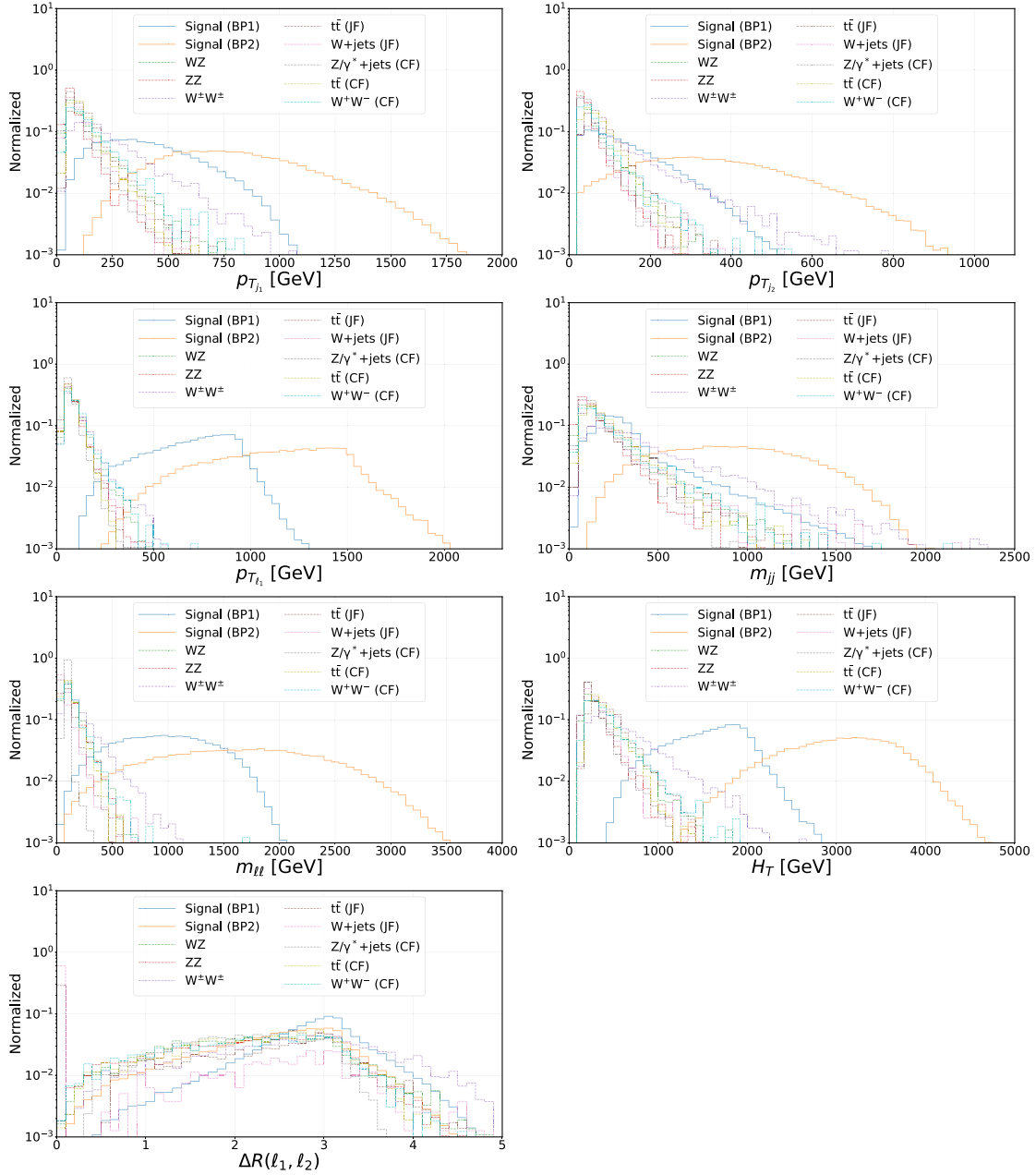


FIG. 7. Normalized distributions for SS dilepton signals and SM backgrounds at the LHC with $\sqrt{s} = 13$ TeV. Signals correspond to $(m_{W_R}, m_{N_R}) = (2 \text{ TeV}, 0.5 \text{ TeV})$ (blue solid line) and $(4 \text{ TeV}, 2 \text{ TeV})$ (orange solid line). Left panels (top to bottom): $p_{T_{l_1}}$, $p_{T_{l_2}}$, m_{jj} , and $\Delta R(\ell_1, \ell_2)$. Right panels (top to bottom): $p_{T_{j_1}}$, $p_{T_{j_2}}$, $m_{\ell\ell}$, and H_T .

The event yields after the DNN selection, expected at LHC Run 2, are summarized in Table VI. The corresponding efficiencies are $\epsilon_{e\mu} = 44\%$, $\epsilon_{ee} = 34\%$, and $\epsilon_{\mu\mu} = 54\%$. This pattern is consistent with the SS case, where better muon reconstruction leads to higher efficiency in the $\mu^+\mu^-$ channel.

From Tables VII and VIII for the high-mass model, background suppression is more efficient than in the low-mass model with the optimal DNN threshold of 0.99. Despite larger backgrounds than in the SS channel, the DNN reduces the OS background to a level comparable to the SS channel after the

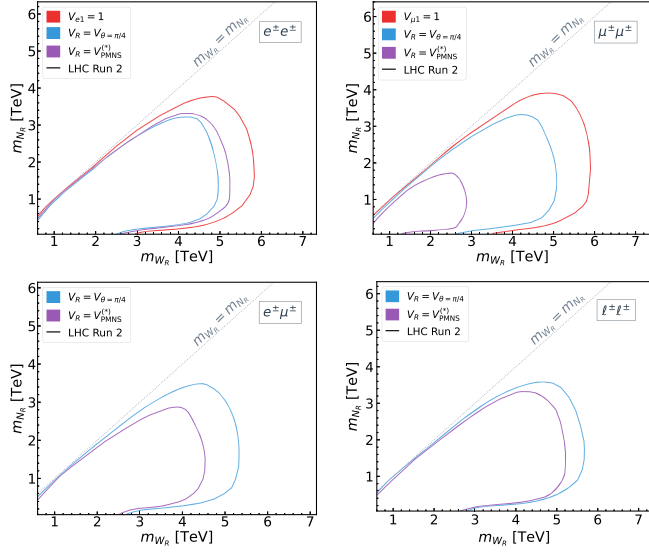


FIG. 8. 95% CL constraints on m_{W_R} and m_{N_R} from SS dilepton searches at LHC Run 2. The panels correspond to the $e^\pm e^\pm$ (upper left), $\mu^\pm \mu^\pm$ (upper right), $e^\pm \mu^\pm$ (lower left), and combined $\ell^\pm \ell^\pm$ ($\ell = e, \mu$) (lower right) channels. The blue and purple contours represent the maximal-mixing scenario ($V_R = V_{\theta=\pi/4}$) and the PMNS-like scenario ($V_R = V_{\text{PMNS}}^{(*)}$), respectively. In the $e^\pm e^\pm$ and $\mu^\pm \mu^\pm$ channels, the red contours denote the unmixed cases ($V_{e1} = 1$ or $V_{\mu1} = 1$). The gray dotted line in each panel marks the kinematic threshold $m_{W_R} = m_{N_R}$.

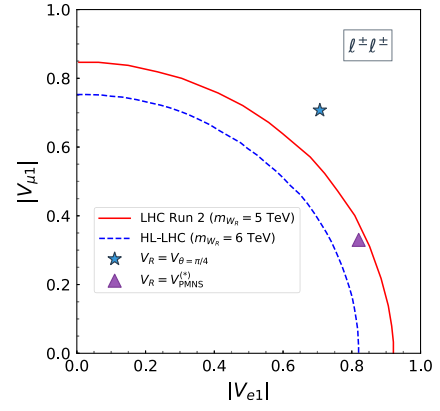


FIG. 9. 95% CL constraints on $|V_{e1}|$ and $|V_{\mu1}|$ with the benchmark $m_{N_R} = 3$ TeV in the SS dilepton channels. The red solid contour shows the limit from LHC Run 2 for $m_{W_R} = 5$ TeV, while the blue dashed contour indicates the HL-LHC sensitivity for $m_{W_R} = 6$ TeV. The star marks the maximal-mixing scenario, and the triangle marks the PMNS-like scenario.

selections. For an integrated luminosity of 139 fb^{-1} , some expected event yields are below unity. Following the convention of Table IV in Sec. IV B, we show all yields with two decimal places for uniform precision.

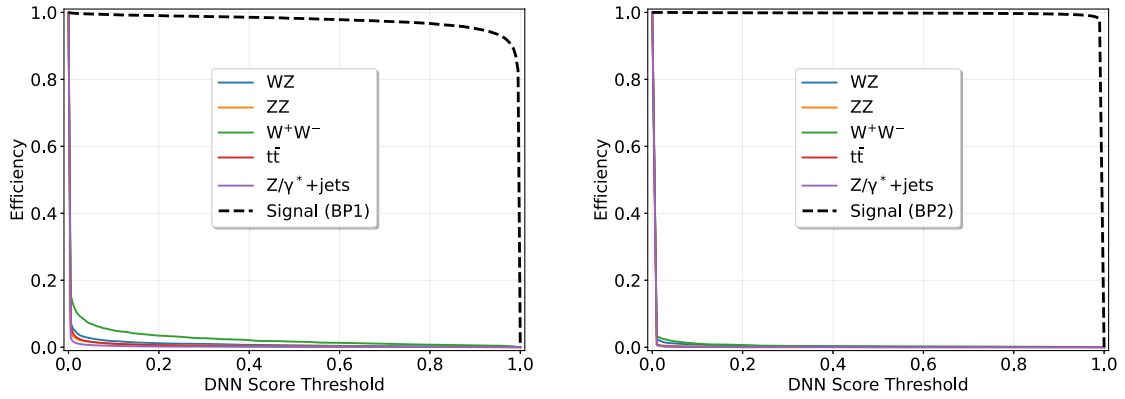


FIG. 10. The same as Fig. 3 but for the OS dilepton searches.

TABLE V. Cross sections for the OS dilepton signal and background processes for the low-mass model. Values are shown at three stages: after the basic cuts (σ_{basic}), after the baseline selection ($\sigma_{\ell\ell jj}$), and after the optimal DNN score selection (σ_{DNN}). The bottom rows provide the comparison of the total background and signal cross sections.

Process	σ_{basic} (pb)	$\sigma_{\ell\ell jj}$ (pb)	σ_{DNN} (pb)
WZ	4.31×10^{-1}	4.49×10^{-2}	3.59×10^{-5}
W^+W^-	3.73×10^{-1}	1.01×10^{-1}	1.92×10^{-4}
ZZ	4.15×10^{-2}	4.34×10^{-3}	8.68×10^{-7}
$t\bar{t}$	1.81×10^1	4.99	1.75×10^{-5}
$Z/\gamma^* + \text{jets}$	5.33×10^1	2.55×10^1	1.39×10^{-4}
Total bkg	7.22×10^1	3.06×10^1	3.85×10^{-4}
Signal	5.21×10^{-2}	2.65×10^{-2}	2.31×10^{-2}

TABLE VI. Event yields after the optimal DNN score selection for the low-mass model. The backgrounds (individual categories and total) and the OS dilepton signal for BP1 are shown for the three OS dilepton channels. The ellipses indicate that the corresponding background contribution is negligible. The last row gives the signal efficiency in each channel.

Process	$e^\pm\mu^\mp$	e^+e^-	$\mu^+\mu^-$
Multiboson	10	12	10
$t\bar{t}$	1	...	1
$Z/\gamma^* + \text{jets}$	...	10	10
Total bkg	11	22	21
Signal	1611	617	983
Eff [%]	44	34	54

TABLE VII. The same as Table V but for the high-mass model.

Process	σ_{basic} (pb)	$\sigma_{\ell\ell jj}$ (pb)	σ_{DNN} (pb)
WZ	4.31×10^{-1}	4.49×10^{-2}	5.23×10^{-6}
W^+W^-	3.73×10^{-1}	1.01×10^{-1}	1.05×10^{-5}
ZZ	4.15×10^{-2}	4.34×10^{-3}	$\lesssim 10^{-7}$
$t\bar{t}$	1.81×10^1	4.99	4.62×10^{-6}
$Z/\gamma^* + \text{jets}$	5.33×10^1	2.55×10^1	2.09×10^{-5}
Total bkg	7.22×10^1	3.06×10^1	4.13×10^{-5}
Signal	4.16×10^{-4}	2.27×10^{-4}	2.20×10^{-4}

TABLE VIII. The same as Table VI but for the high-mass model.

Process	$e^\pm\mu^\pm$	$e^\pm e^\pm$	$\mu^\pm\mu^\pm$
Multiboson	0.89	0.67	0.62
$t\bar{t}$	0.23	0.16	0.25
$Z/\gamma^* + \text{jets}$	...	1.36	1.55
Total bkg	1.12	2.19	2.42
Signal	14.84	6.52	9.24
Eff [%]	50.31	44.15	62.69

APPENDIX B: DNN IMPLEMENTATION AND VALIDATION

The DNN classifiers are implemented in PyTorch Lightning [89] and take the kinematic variables defined in Eq. (23) as inputs. The input variables are standardized using the StandardScaler routine from SCIKIT-LEARN [90] before being passed to the network. The network consists of two fully connected hidden layers with 64 nodes each and ReLU activation functions. The output layer contains two nodes, corresponding to the background and signal classes. The signal score is obtained from the softmax probability of the signal class. Training is performed with the Adam optimizer [91], with an initial learning rate of 10^{-3} , a batch size of 1024, and a cross-entropy loss. The network is trained for 50 epochs. The simulated samples are split into training, validation, and test samples with fractions of 80%, 10%, and 10%, respectively. No event weights are applied in the DNN training loss; all simulated events entering the training samples are treated with equal weight. The process cross sections are not used during training but are applied after the DNN selection to convert the signal and background efficiencies into expected yields. The training and validation loss and accuracy are monitored throughout the training. Both the training and validation losses decrease steadily at the beginning of the training and subsequently reach stable plateaus, while the corresponding accuracies show consistent convergence. No increase of the validation loss after convergence or divergence between the training and validation performance is observed, indicating no evident sign of overtraining.

To validate the performance of the DNN classifiers used in the analysis, we show the receiver operating characteristic (ROC) curves for the SS dilepton channel in Fig. 11. Two independent DNN classifiers are trained using BP1 and BP2, corresponding to the low-mass and high-mass models, respectively.

The ROC curves display the background efficiency versus the signal efficiency, with the background efficiency plotted on a logarithmic scale to highlight the rejection power in the high-purity (tight DNN-score) regime targeted by the final event selection. For both the low-mass and high-mass classifiers, the resulting curves are clearly separated

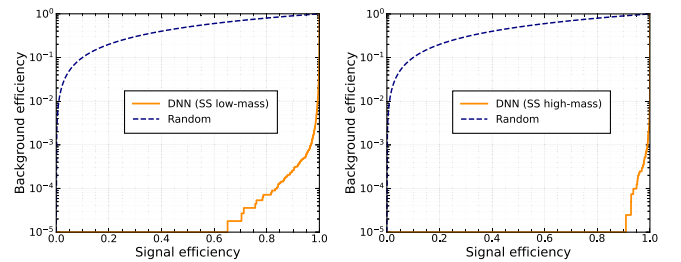


FIG. 11. ROC curves for the low-mass and high-mass DNN classifiers used in the SS dilepton analysis, shown in the left and right panels, respectively.

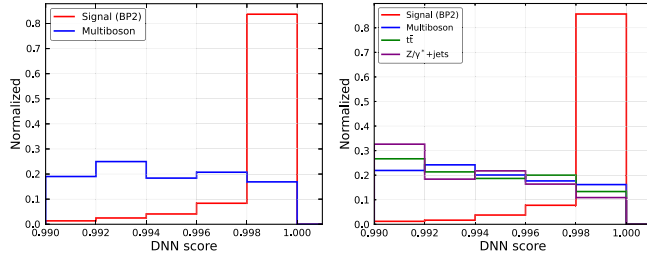


FIG. 12. Normalized DNN-score distributions in the high-score region after applying $D > 0.99$ for the high-mass model. The left and right panels show the SS and OS dilepton analyses, respectively, including the signal and the corresponding SM backgrounds.

from the random-classifier baseline, demonstrating efficient discrimination between signal and background.

The OS dilepton channel is treated analogously. As an additional check of the DNN behavior in the high-score region, Fig. 12 shows the normalized DNN-score distributions in the tail region $D > 0.99$ for the high-mass classifiers used in the SS and OS dilepton analyses. In the SS channel, $\mathcal{O}(10^6)$ unweighted simulated multiboson events pass the baseline selection, with $\mathcal{O}(10^2)$ – $\mathcal{O}(10^3)$ events remaining after the DNN selection. The JF and CF samples contain $\mathcal{O}(10^6)$ and $\mathcal{O}(10^7)$ unweighted events after the baseline selection, respectively, but are reduced to only $\mathcal{O}(10)$ – $\mathcal{O}(10^3)$ events or fewer after the tight DNN requirement, depending on the specific background.¹¹

After rescaling to the expected cross sections, the JF and CF contributions are at least one order of magnitude smaller than the multiboson backgrounds (cf., Table III), even after accounting for their modeling uncertainties, taken as 10% and 17% [63,68], respectively. Therefore, the JF and CF backgrounds are not shown in the DNN-score tail distributions in Fig. 12 and are not included in Table IV. In contrast, for the OS channel, the DNN-score tail distributions of the three dominant backgrounds are all shown in Fig. 12. After the DNN selection, up to $\mathcal{O}(10^2)$ unweighted events for each background component in our simulation are retained, which keeps the MC statistical uncertainties in the tail region under control. These validation checks support the robustness of the DNN-based event selection.

For each signal region, the fixed trained DNN classifier is applied to all scanned mass points using the same preprocessing and DNN-score requirement. The signal efficiency is evaluated separately at each mass point and is then used to compute the expected yield and significance and derive the expected exclusion limits.

To assess the applicability of the trained classifiers in different regions, we examine the benchmark points

¹¹For the low-mass classifier, the multiboson, JF, and CF backgrounds are comparable after the DNN selection (cf., Table I). At least $\mathcal{O}(10^2)$ unweighted events remain for each background component in our simulation, indicating that the MC statistical uncertainty in the tail region is under control.

TABLE IX. Validation of the trained DNN classifiers at representative mass points. The asterisks indicate the training benchmark points BP1 and BP2.

Region	m_{W_R} (TeV)	m_{N_R} (TeV)	m_{N_R}/m_{W_R}	ϵ_{DNN} (%)
Low*	2	0.5	0.250	94.8
Low	2	1.0	0.500	92.1
Low	2	1.5	0.750	78.7
High	4	0.5	0.125	92.3
High	4	1.0	0.250	95.4
High*	4	2.0	0.500	96.9
High	4	3.0	0.750	94.8
High	4	3.5	0.875	82.3

together with representative nontraining mass points covering different values of m_{N_R}/m_{W_R} . Table IX shows the corresponding DNN cut efficiencies, $\epsilon_{\text{DNN}} \equiv \sigma_{\text{DNN}}/\sigma_{\ell\ell jj}$. Here, “low” and “high” indicate that the signal point is evaluated using the low-mass and high-mass classifiers trained using BP1 and BP2, respectively, as defined in Sec. IV B. The asterisks indicate the training benchmark points.

As shown in Table IX, the DNN cut efficiencies remain high for both the benchmark points and the additional nontraining mass points. The efficiencies remain stable over the tested regions, with only an expected moderate decrease at larger values of m_{N_R}/m_{W_R} , where the N_R and W_R masses become closer. Nevertheless, the signal still retains characteristic high-energy kinematic features, allowing it to be effectively distinguished from the background.

This validation shows that the trained classifiers provide stable signal selection away from the benchmark points. While the exclusion contours could be further improved by training classifiers at multiple values of m_{N_R}/m_{W_R} , the stability check presented here suggests that such a finer training strategy would not qualitatively change our conclusions.

APPENDIX C: CLFV PROCESSES: ANALYTIC EXPRESSIONS

The CLFV processes $\mu \rightarrow e\gamma$, $\mu \rightarrow eee$, and $\mu \rightarrow e$ conversion are depicted in Figs. 13–15, respectively. The

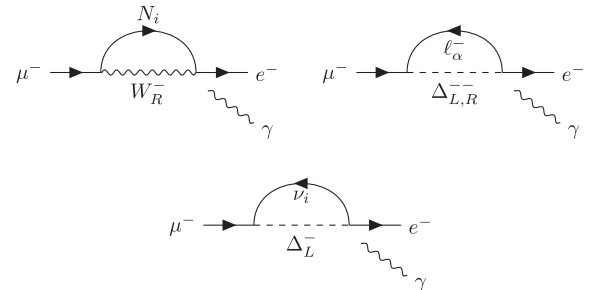
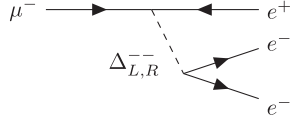


FIG. 13. Representative Feynman diagrams for $\mu \rightarrow e\gamma$. The photon can be radiated from any charged external leg or internal propagator.

FIG. 14. Representative Feynman diagram for $\mu \rightarrow eee$.

decay branching ratios of the former two are given by [16,75]

$$\text{BR}(\mu \rightarrow e\gamma) = \frac{3\alpha_{\text{em}}}{2\pi} (|G_L^\gamma|^2 + |G_R^\gamma|^2), \quad (\text{C1})$$

$$\text{Br}(\mu \rightarrow eee) = \frac{1}{2} |h_{e\mu} h_{ee}^*|^2 \left(\frac{m_{W_L}^4}{m_{\Delta_R^{--}}^4} + \frac{m_{W_L}^4}{m_{\Delta_L^{--}}^4} \right). \quad (\text{C2})$$

Here, $\alpha_{\text{em}} = e^2/(4\pi)$ is the electromagnetic coupling constant, the form factors G_L^γ and G_R^γ are expressed as

$$G_L^\gamma = \sum_i V_{\mu i} V_{ei}^* \left(\frac{m_{W_L}^2}{m_{W_R}^2} G_1^\gamma(y_i) + \frac{2y_i}{3} \frac{m_{W_L}^2}{m_{\Delta_R^{--}}^2} \right), \quad (\text{C3})$$

$$G_R^\gamma = \sum_i V_{\mu i} V_{ei}^* x_i \left(\frac{2}{3} \frac{m_{W_L}^2}{m_{\Delta_L^{--}}^2} + \frac{1}{12} \frac{m_{W_L}^2}{m_{\Delta_L^{--}}^2} \right), \quad (\text{C4})$$

with $x_i \equiv (m_{N_i}/m_{W_L})^2$, $y_i \equiv (m_{N_i}/m_{W_R})^2$. The loop function $G_1^\gamma(y)$ is given as

$$G_1^\gamma(y) = -\frac{2y^3 + 5y^2 - y}{4(1-y)^3} - \frac{3y^3}{2(1-y)^4} \ln y. \quad (\text{C5})$$

The effective coupling is defined as

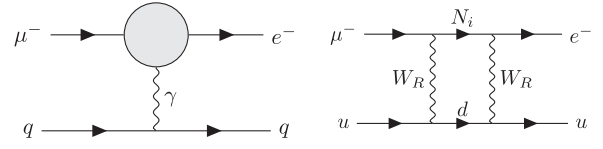


FIG. 15. Representative Feynman diagrams for $\mu \rightarrow e$ conversion in nuclei. The effective $\mu e \gamma^*$ vertex (shaded circle) incorporates contributions from the diagram shown in Fig. 13.

$$h_{\alpha\beta} = \sum_i V_{ai} V_{\beta i} \left(\frac{m_{N_i}}{m_{W_R}} \right), \quad \alpha, \beta = e, \mu. \quad (\text{C6})$$

If the heavy particles W_R , $\Delta_{L,R}^{--}$, and Δ_L^- are degenerate in mass, the $\mu N \rightarrow eN$ conversion branching ratio is approximated by [75]

$$R_N(\mu \rightarrow e) \simeq X_N \times 10^{-7} |g_{e\mu}|^2 \left(\frac{1 \text{ TeV}}{m_{\Delta_R^{--}}} \right)^4 \times \alpha_{\text{em}} \left(\ln \frac{m_{\Delta_R^{--}}^2}{m_\mu^2} \right)^2, \quad (\text{C7})$$

where the factor X_N depends on the nucleus, with $X_{(\text{Al,Ti,Au})} \simeq (0.8, 1.3, 1.6)$, and m_μ denotes the mass of muon. The effective coupling is defined as

$$g_{e\mu} = \sum_i V_{ei}^* V_{\mu i} \left(\frac{m_{N_i}}{m_{W_R}} \right)^2. \quad (\text{C8})$$

Notice that, in addition to the tree-level diagram shown in Fig. 14, there exist additional one-loop contribution arising from the effective $\mu e \gamma^*$ vertex, which is much suppressed [75].

-
- [1] J. C. Pati and A. Salam, *Phys. Rev. D* **10**, 275 (1974); **11**, 703(E) (1975).
 - [2] R. N. Mohapatra and J. C. Pati, *Phys. Rev. D* **11**, 2558 (1975).
 - [3] G. Senjanovic and R. N. Mohapatra, *Phys. Rev. D* **12**, 1502 (1975).
 - [4] G. Senjanovic, *Nucl. Phys.* **B153**, 334 (1979).
 - [5] R. N. Mohapatra and G. Senjanovic, *Phys. Rev. Lett.* **44**, 912 (1980).
 - [6] R. N. Mohapatra and G. Senjanovic, *Phys. Rev. D* **23**, 165 (1981).
 - [7] P. Minkowski, *Phys. Lett. B* **67**, 421 (1977).
 - [8] M. Gell-Mann, P. Ramond, and R. Slansky, in *International Symposium on Fundamentals of Quantum Theory and Quantum Field Theory* (1979), [arXiv:hep-ph/9809459](https://arxiv.org/abs/hep-ph/9809459).
 - [9] M. Gell-Mann, P. Ramond, and R. Slansky, *Conf. Proc. C* **790927**, 315 (1979), <https://inspirehep.net/literature/9686>.
 - [10] T. Yanagida, *Conf. Proc. C* **7902131**, 95 (1979), <https://inspirehep.net/literature/143150>.
 - [11] S. L. Glashow, *NATO Sci. Ser. B* **61**, 687 (1980).
 - [12] W.-Y. Keung and G. Senjanovic, *Phys. Rev. Lett.* **50**, 1427 (1983).
 - [13] G. Aad *et al.* (ATLAS Collaboration), *Eur. Phys. J. C* **83**, 1164 (2023).
 - [14] A. Tumasyan *et al.* (CMS Collaboration), *J. High Energy Phys.* **04** (2022) 047.
 - [15] M. Nemevsek, G. Senjanovic, and V. Tello, *Phys. Rev. Lett.* **110**, 151802 (2013).
 - [16] J. Barry and W. Rodejohann, *J. High Energy Phys.* **09** (2013) 153.

- [17] P. S. Bhupal Dev, S. Goswami, and M. Mitra, *Phys. Rev. D* **91**, 113004 (2015).
- [18] G. Senjanović and V. Tello, *Phys. Rev. Lett.* **119**, 201803 (2017).
- [19] G. Senjanovic and V. Tello, *Phys. Rev. D* **100**, 115031 (2019).
- [20] G. Senjanovic and V. Tello, *Int. J. Mod. Phys. A* **38**, 2350067 (2023).
- [21] X. Zhang, J.-H. Yu, and B.-Q. Ma, *Nucl. Phys.* **B976**, 115670 (2022).
- [22] G. Li, D.-Y. Luo, and X. Zhao, *Phys. Rev. D* **110**, 035030 (2024).
- [23] X. Chen and X. Zhang, [arXiv:2603.03874](https://arxiv.org/abs/2603.03874).
- [24] V. Tello, *Phys. Rev. D* **113**, 115060 (2026).
- [25] A. M. Sirunyan *et al.* (CMS Collaboration), *J. High Energy Phys.* **01** (2019) 122.
- [26] A. Hayrapetyan *et al.* (CMS Collaboration), *J. High Energy Phys.* **03** (2024) 105.
- [27] G. Aad *et al.* (ATLAS Collaboration), *Phys. Lett. B* **856**, 138865 (2024).
- [28] A. Hayrapetyan *et al.* (CMS Collaboration), *J. High Energy Phys.* **06** (2024) 183.
- [29] A. Hayrapetyan *et al.* (CMS Collaboration), *J. High Energy Phys.* **02** (2025) 036.
- [30] G. Aad *et al.* (ATLAS Collaboration), *J. High Energy Phys.* **07** (2025) 196.
- [31] G. Aad *et al.* (ATLAS Collaboration), *Eur. Phys. J. C* **72**, 2056 (2012).
- [32] G. Aad *et al.* (ATLAS Collaboration), *J. High Energy Phys.* **07** (2015) 162.
- [33] M. Aaboud *et al.* (ATLAS Collaboration), *J. High Energy Phys.* **01** (2019) 016.
- [34] M. Aaboud *et al.* (ATLAS Collaboration), *Phys. Lett. B* **798**, 134942 (2019).
- [35] S. P. Das, F. F. Deppisch, O. Kittel, and J. W. F. Valle, *Phys. Rev. D* **86**, 055006 (2012).
- [36] J. C. Vasquez, *J. High Energy Phys.* **05** (2016) 176.
- [37] G. Karagiorgi, G. Kasieczka, S. Kravitz, B. Nachman, and D. Shih, [arXiv:2112.03769](https://arxiv.org/abs/2112.03769).
- [38] M. Abdughani, J. Ren, L. Wu, J. M. Yang, and J. Zhao, *Commun. Theor. Phys.* **71**, 955 (2019).
- [39] Y. LeCun, Y. Bengio, and G. Hinton, *Nature (London)* **521**, 436 (2015).
- [40] P. Baldi, P. Sadowski, and D. Whiteson, *Nat. Commun.* **5**, 4308 (2014).
- [41] L. de Oliveira, M. Kagan, L. Mackey, B. Nachman, and A. Schwartzman, *J. High Energy Phys.* **07** (2016) 069.
- [42] S.-Y. Chen, Y.-P. Jiao, S.-Y. Wang, Q.-S. Yan, H.-H. Zhang, and Y. Zhang, [arXiv:2504.12141](https://arxiv.org/abs/2504.12141).
- [43] J. Kriewald, M. Nemevšek, and F. Nesti, *Eur. Phys. J. C* **84**, 1306 (2024).
- [44] J. de Vries, G. Li, M. J. Ramsey-Musolf, and J. C. Vasquez, *J. High Energy Phys.* **11** (2022) 056.
- [45] G. Li, Y.-Y. Li, S. Lu, and Y.-L. Zhou, *Phys. Lett. B* **880**, 140829 (2026).
- [46] G. Li, M. Ramsey-Musolf, and J. C. Vasquez, *Phys. Rev. Lett.* **126**, 151801 (2021).
- [47] M. J. Ramsey-Musolf and J. C. Vasquez, *Phys. Lett. B* **815**, 136136 (2021).
- [48] G. Senjanović and V. Tello, *Phys. Rev. Lett.* **114**, 071801 (2015).
- [49] A. Maiezza, M. Nemevsek, F. Nesti, and G. Senjanovic, *Phys. Rev. D* **82**, 055022 (2010).
- [50] M. Nemevšek, F. Nesti, and G. Popara, *Phys. Rev. D* **97**, 115018 (2018).
- [51] J. C. Helo, M. Hirsch, and S. Kovalenko, *Phys. Rev. D* **89**, 073005 (2014); **93**, 099902(E) (2016).
- [52] M. Frank, B. Fuks, A. Jueid, S. Moretti, and O. Ozdal, *J. High Energy Phys.* **02** (2024) 150.
- [53] A. Alloul, N. D. Christensen, C. Degrande, C. Duhr, and B. Fuks, *Comput. Phys. Commun.* **185**, 2250 (2014).
- [54] J. Kriewald, M. Nemevšek, and F. Nesti, Complete Left-Right symmetric model file, https://feynrules.irmp.ucl.ac.be/wiki/LRSM_NLO.
- [55] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H. S. Shao, T. Stelzer, P. Torrielli, and M. Zaro, *J. High Energy Phys.* **07** (2014) 079.
- [56] A. Buckley, J. Ferrando, S. Lloyd, K. Nordström, B. Page, M. Rüfenacht, M. Schönherr, and G. Watt, *Eur. Phys. J. C* **75**, 132 (2015).
- [57] T. Sjöstrand, S. Ask, J. R. Christiansen, R. Corke, N. Desai, P. Ilten, S. Mrenna, S. Prestel, C. O. Rasmussen, and P. Z. Skands, *Comput. Phys. Commun.* **191**, 159 (2015).
- [58] J. de Favereau, C. Delaere, P. Demin, A. Giammanco, V. Lemaître, A. Mertens, and M. Selvaggi (DELPHES 3 Collaboration), *J. High Energy Phys.* **02** (2014) 057.
- [59] A. M. Sirunyan *et al.* (CMS Collaboration), *Eur. Phys. J. C* **77**, 578 (2017).
- [60] M. Aaboud *et al.* (ATLAS Collaboration), *Eur. Phys. J. C* **78**, 199 (2018).
- [61] T. Peng, M. J. Ramsey-Musolf, and P. Winslow, *Phys. Rev. D* **93**, 093002 (2016).
- [62] J. Harz, M. J. Ramsey-Musolf, T. Shen, and S. U. Quiroga, *Phys. Rev. D* **110**, 035024 (2024).
- [63] D. Curtin, J. Galloway, and J. G. Wacker, *Phys. Rev. D* **88**, 093006 (2013).
- [64] E. Izaguirre and B. Shuve, *Phys. Rev. D* **91**, 093010 (2015).
- [65] M. Nemevšek, F. Nesti, and J. C. Vasquez, *J. High Energy Phys.* **04** (2017) 114.
- [66] C. O. Dib, C. S. Kim, K. Wang, and J. Zhang, *Phys. Rev. D* **94**, 013005 (2016).
- [67] C. O. Dib, C. S. Kim, and K. Wang, *Phys. Rev. D* **95**, 115020 (2017).
- [68] M. Muškinja, Search for new physics processes with same sign leptons in the final state with the ATLAS detector, Ph.D. thesis, Ljubljana U., 2019.
- [69] D. Guest, K. Cranmer, and D. Whiteson, *Annu. Rev. Nucl. Part. Sci.* **68**, 161 (2018).
- [70] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, *Eur. Phys. J. C* **71**, 1554 (2011); **73**, 2501(E) (2013).
- [71] P. N. Bhattiprolu, S. P. Martin, and J. D. Wells, *Eur. Phys. J. C* **81**, 123 (2021).
- [72] I. Esteban, M. C. Gonzalez-Garcia, M. Maltoni, I. Martinez-Soler, J. P. Pinheiro, and T. Schwetz, *J. High Energy Phys.* **12** (2024) 216.
- [73] A. Abusleme *et al.* (JUNO Collaboration), *Nature (London)* **654**, 343 (2026).
- [74] I. Esteban, M. C. Gonzalez-Garcia, M. Maltoni, I. Martinez-Soler, J. P. Pinheiro, and T. Schwetz, *J. High Energy Phys.* **04** (2026) 089.

- [75] V. Cirigliano, A. Kurylov, M. J. Ramsey-Musolf, and P. Vogel, *Phys. Rev. D* **70**, 075007 (2004).
- [76] F. F. Deppisch, T. E. Gonzalo, S. Patra, N. Sahu, and U. Sarkar, *Phys. Rev. D* **91**, 015018 (2015).
- [77] G. Bambhaniya, P. S. B. Dev, S. Goswami, and M. Mitra, *J. High Energy Phys.* **04** (2016) 046.
- [78] D. Borah and A. Dasgupta, *J. High Energy Phys.* **07** (2016) 022.
- [79] K. Afanaciev *et al.* (MEG II Collaboration), *Eur. Phys. J. C* **85**, 1177 (2025); **85**, 1317(E) (2025).
- [80] U. Bellgardt *et al.* (SINDRUM Collaboration), *Nucl. Phys.* **B299**, 1 (1988).
- [81] W. H. Bertl *et al.* (SINDRUM II Collaboration), *Eur. Phys. J. C* **47**, 337 (2006).
- [82] K. Arndt *et al.* (Mu3e Collaboration), *Nucl. Instrum. Methods Phys. Res., Sect. A* **1014**, 165679 (2021).
- [83] F. Abdi *et al.* (Mu2e Collaboration), *Universe* **9**, 54 (2023).
- [84] J. Gluza and T. Jeliński, *Phys. Lett. B* **748**, 125 (2015).
- [85] A. Das, P. S. B. Dev, and R. N. Mohapatra, *Phys. Rev. D* **97**, 015018 (2018).
- [86] R. N. Mohapatra, *Phys. Rev. D* **34**, 909 (1986).
- [87] A. Maiezza, M. Nemevšek, and F. Nesti, *Phys. Rev. D* **94**, 035008 (2016).
- [88] V. Tello, M. Nemevsek, F. Nesti, G. Senjanovic, and F. Vissani, *Phys. Rev. Lett.* **106**, 151801 (2011).
- [89] W. Falcon *et al.*, [10.5281/zenodo.3828935](https://zenodo.org/record/3828935).
- [90] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, *J. Mach. Learn. Res.* **12**, 2825 (2011), <https://jmlr.org/papers/v12/pedregosa11a.html>.
- [91] D. P. Kingma and J. Ba, *International Conference on Learning Representations (ICLR)* (2014), [arXiv: 1412.6980](https://arxiv.org/abs/1412.6980).