Signatures of Dirac and Majorana sterile neutrinos in trilepton events at the LHC

Claudio O. Dib,[1,‡] C. S. Kim,[2,†] and Kechen Wang[3,4,*]

[1]*CCTVal and Department of Physics, Universidad Técnica Federico Santa María, Valparaíso, Chile*
[2]*Department of Physics and IPAP, Yonsei University, Seoul 120-749, Korea*
[3]*DESY, Notkestrae 85, D-22607 Hamburg, Germany*
[4]*Center for Future High Energy Physics, Institute of High Energy Physics, Chinese Academy of Sciences, Beijing 100049, China*

Heavy sterile neutrinos with masses below M_W can induce trilepton events at the 14 TeV LHC through purely leptonic W decays of $W^{\pm} \to e^{\pm}e^{\pm}\mu^{\mp}\nu$ and $\mu^{\pm}\mu^{\pm}e^{\mp}\nu$ where the heavy neutrino will be in an intermediate state on its mass shell. Discovery and exclusion limits for the heavy neutrinos are found using both cut-and-count (CC) and multivariate analysis (MVA) methods in this study. We also show that it is possible to discriminate between a Dirac and a Majorana heavy neutrino, even when lepton number conservation cannot be directly tested due to unobservability of the final state neutrino. This discrimination is done by exploiting a combined set of kinematic observables that differ between the Majorana and Dirac cases. We find that the MVA method can greatly enhance the discovery and discrimination limits in comparison with the CC method. For a 14-TeV pp collider with integrated luminosity of 3000 fb^{-1}, sterile neutrinos can be found with 5σ significance if heavy-to-light neutrino mixings $|U_{Ne}|^2 \sim |U_{N\mu}|^2 \sim 10^{-6}$, while the Majorana vs Dirac type can be distinguished if $|U_{Ne}|^2 \sim |U_{N\mu}|^2 \sim 10^{-5}$ or even $|U_{N\ell}|^2 \sim 10^{-6}$ if one of the mixing elements is at least an order of magnitude smaller than the other.

I. INTRODUCTION

Neutrinos are the most esoteric of all particles in the Standard Model (SM): they interact via weak interactions only, which makes them very hard to detect; their interaction is purely of left-handed chirality, so their right-handed component—if there is any—is sterile; their masses, while not all zero, are much smaller than the energies of all detectable processes, so they are only available as extremely relativistic particles; and they exist in three flavors, pairing the charged leptons. However, they exhibit large mixing in their mass eigenstates, thus exhibiting a chameleonic behavior by changing their flavor in flight, which is known as neutrino oscillations. The very observation of neutrino oscillations [1] implies that neutrinos must have mass, contrary to the SM in which the neutrinos are assumed to be massless. Moreover, since neutrinos are electrically neutral, they could be their own antiparticles; i.e., Majorana fermions [2], carrying no charge such as the lepton number, in which case weak interactions involving neutrinos will not conserve the lepton number. Alternatively, if they carry the lepton number, they must be Dirac fermions, neutrinos and antineutrinos will be different particles, and the lepton

number will be conserved. Therefore, one important step toward resolving the origin of the neutrino mass is to ascertain whether they are Dirac or Majorana fermions. The Majorana nature of neutrinos can be revealed in neutrinoless double beta ($0\nu\beta\beta$) decay experiments [3,4], but so far no evidence has been found. Concerning explanations of the smallness of neutrino masses, most are based on seesaw mechanisms [5–14], which imply the existence of additional, heavier neutrinos, which are sterile under the electroweak interactions except for their mixings with the standard neutrinos [15]. The original seesaw models required very large masses for the sterile neutrinos, $M \sim 10^5$ to 10^{15} GeV, beyond detectability in any foreseen experiment, and mixings $U \sim (m_\nu/M)^{1/2}$, which are highly suppressed as well (10^{-8} to 10^{-13}). However, in other versions called *low-scale seesaw, inverse seesaw, etc.,* the smallness of m_ν does not require huge values for M nor tiny values for U. To date, each specific scenario proposes heavy neutrinos with their mass within a given scale, but from one scenario to another this scale can be anywhere from a few eV all the way to grand unification scales. In turn, different experiments put bounds on neutrino masses and mixings, each one in a different and limited mass range within this broad spectrum of possibilities. So far, experimental searches have not found conclusive evidence of their existence [16] either as Dirac or Majorana particles. In particular, the studies at the Large Hadron Collider (LHC) look for same sign dilepton plus dijet events, $\ell^{\pm}\ell^{\pm}jj$, which can be produced and observed if there are heavy

[*]Corresponding author.
kechen.wang@desy.de
[†]cskim@yonsei.ac.kr
[‡]claudio.dib@usm.cl

Majorana fermions with mass above a few tens of GeV and up to a few hundred GeV [17,18]. For neutrino masses below M_W, the jets may not be energetic enough to be separated from the background, and thus trilepton events $\ell^\pm \ell^\pm \ell'^\mp \nu$ would provide clearer signals [19].

In previous works we have studied the potential of these trilepton events to discover heavy neutrinos, especially addressing the discrimination between their Dirac or Majorana nature [20]. We studied the signal $W^\pm \to e^\pm e^\pm \mu^\mp \nu$, which will appear resonantly enhanced provided there exist neutrinos with masses below M_W, through the subprocess $W^\pm \to e^\pm N$ followed by $N \to e^\pm \mu^\mp \nu$ (where N stands for the heavy neutrino). The choice of having no opposite-sign same-flavor (no-OSSF) lepton pairs in the final state helps eliminate a serious SM radiative background $\gamma^*/Z \to \ell^+ \ell^-$ [21]. If N is Majorana, it will induce a lepton number conserving (LNC) process $W^+ \to e^+ e^+ \mu^- \nu_e$ as well as a lepton number violating (LNV) process $W^+ \to e^+ e^+ \mu^- \bar\nu_\mu$, while if it is of Dirac type, it will induce the LNC process only. One could in principle use this feature to discriminate between a Majorana and a Dirac N; however, since the final neutrino escapes detection, the observed final state is just $e^\pm e^\pm \mu^\mp$ or $\mu^\pm \mu^\pm e^\mp$ plus missing energy. It is then not a simple task to distinguish between the LNC and the LNV processes, and hence between the Majorana and Dirac nature of N. In our previous work we found that, in principle, the two cases could be distinguished by constructing the energy spectrum of the opposite-charge lepton.

In our consecutive work [22], we presented a simpler method to distinguish between Majorana and Dirac N's, by examining the full decay rates instead of the spectra, for all the channels $e^\pm e^\pm \mu^\mp$ and $\mu^\pm \mu^\pm e^\mp$, because the discrimination through spectra in rare processes is much more difficult to achieve. However, this discrimination, which is based on full rates, works only if the mixing parameters U_{Ne} and $U_{N\mu}$ are considerably different from each other.

Here we present a comprehensive strategy to detect heavy sterile neutrinos and discriminate between Dirac vs Majorana using all the details of the leptonic decays of $W^\pm \to e^\pm e^\pm \mu^\mp \nu$ or $\mu^\pm \mu^\pm e^\mp \nu$ at the 14 TeV LHC, provided they exist with masses near and below the W boson mass. The strategy is based on both a cut-and-count (CC) method and a multivariate analysis (MVA) method that uses all the features of the leptonic decays that can distinguish between Dirac and Majorana neutrinos.

The article is organized as follows. In Sec. II we present the theoretical aspects and formulation of the problem. In Sec. III we describe the data simulation and background study. In Sec. IV we state the method to detect the existence of heavy neutrinos that induce the trilepton events. In Sec. V we study the capability of the method to discriminate between the Dirac and Majorana characters of N. In Sec. VI we present our summary and conclusions.

II. THEORETICAL FRAMEWORK

Our processes of interest are labeled as $W^\pm \to l_W^\pm l_N^\pm l'^\mp_N \nu$ and depicted in Fig. 1. Here l and l' are charged leptons of different flavors, either e or μ (i.e., $e^\pm e^\pm \mu^\mp \nu$ or $\mu^\pm \mu^\pm e^\mp \nu$), where ν represents a SM neutrino or antineutrino. These are two different processes, but since the final neutrino goes undetected, the observable final state is the same, namely trilepton $l^\pm l^\pm l'^\mp$ plus missing energy, in either process. One process is LNV, while the other is LNC. If N is Majorana, both LNV and LNC processes occur and the events of the two modes add up in the experiment, while if N is Dirac, only the LNC process occurs.

Following the notation of Fig. 1 for the lepton momenta and calling q the momentum of the W boson, the differential rate of the LNV process with flavors, e.g., $e^+ e^+ \mu^-$, is

$$\Gamma(W^+ \to e^+ e^+ \mu^- \bar\nu_\mu)$$
$$= \frac{8\sqrt{2} G_F^3}{3\pi} \frac{m_N (M_W^2 - m_N^2)}{M_W \Gamma_N} |U_{Ne}|^4$$
$$\times \int d\Phi_3 (\ell_N \cdot \ell_\nu) \left\{ (\ell_W \cdot \ell'_N) + \frac{2}{M_W^2} (q \cdot \ell_W)(q \cdot \ell'_N) \right\},$$
$$(1)$$

where Γ_N is the N width, which also depends on m_N and the lepton mixings [20]. We denote by $\int d\Phi_3$ the Lorentz invariant phase space for the three final particles of the N decay in the normalization of the Particle Data Group [23]. Similarly, the differential rate for the LNC process is

$$\Gamma(W^+ \to e^+ e^+ \mu^- \nu_e)$$
$$= \frac{8\sqrt{2} G_F^3}{3\pi} \frac{(M_W^2 - m_N^2)}{m_N M_W \Gamma_N} |U_{Ne} U_{N\mu}|^2$$
$$\times \int d\Phi_3 (\ell'_N \cdot \ell_\nu) \times \left\{ 2(k_N \cdot \ell_N) \left[(k_N \cdot \ell_W) \right.\right.$$
$$+ \frac{2}{M_W^2} (q \cdot k_N)(q \cdot \ell_W) \right]$$
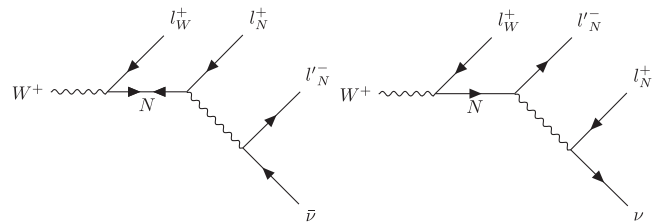$$\left. - m_N^2 \left[(\ell_W \cdot \ell_N) + \frac{2}{M_W^2} (q \cdot \ell_W)(q \cdot \ell_N) \right] \right\}. \quad (2)$$

FIG. 1. Left: LNV process $W^+ \to l_W^+ l_N^+ l'^-_N \bar\nu$, mediated by a heavy sterile neutrino of Majorana type only. Right: the LNC process $W^+ \to l_W^+ l'^-_N l_N^+ \nu$, mediated by a heavy sterile neutrino of either Majorana or Dirac type.

These two expressions seem to be different. Indeed, it is straightforward to show that the two processes have different spectral and angular distributions. Their integrated branching ratios, on the other hand, are equal except for a global lepton mixing factor [20],

$$\mathrm{Br}(W^+ \to e^+ e^+ \mu^- \bar{\nu}_\mu) = f(m_N) \times \frac{|U_{Ne}|^4}{\sum_\ell |U_{N\ell}|^2}, \quad (3)$$

$$\mathrm{Br}(W^+ \to e^+ e^+ \mu^- \nu_e) = f(m_N) \times \frac{|U_{Ne}|^2 |U_{N\mu}|^2}{\sum_\ell |U_{N\ell}|^2}, \quad (4)$$

where

$$f(m_N) \approx 4.8 \times 10^{-3} \left(1 - \frac{m_N^2}{M_W^2}\right)^2 \left(1 + \frac{m_N^2}{2M_W^2}\right).$$

The spectral distributions were studied in a previous work [20]. Concerning the angular distributions, from Eqs. (1) and (2), one can see that the LNV and LNC processes differ in their combinations of scalar products, which translates into different angular distributions among pairs of particles. Previous works have exploited these angular distribution differences for other models [24]. In our case, each one of these differences is not dramatic by itself, but in our analysis we build a combination of several distributions that differ between the LNV and LNC modes, adding up in their capacity to discriminate between a Dirac and a Majorana N.

For convenience, we introduce two parameters: the normalization factor s and the disparity factor r,

$$s \equiv 2 \times 10^6 \frac{|U_{Ne} U_{N\mu}|^2}{|U_{Ne}|^2 + |U_{N\mu}|^2}, \qquad r \equiv \frac{|U_{Ne}|^2}{|U_{N\mu}|^2}. \quad (5)$$

The mixing angles $|U_{Ne}|^2$ and $|U_{N\mu}|^2$ can be expressed in terms of r and s as

$$|U_{Ne}|^2 = \frac{s(1+r)}{2 \times 10^6}, \qquad |U_{N\mu}|^2 = \frac{s(1+\frac{1}{r})}{2 \times 10^6}. \quad (6)$$

In our study we assume for simplicity that only one sterile neutrino N is within the experimental reach, and that it mixes only with the active neutrinos ν_e and ν_μ. The sterile neutrino can be either Dirac or Majorana. According to Eqs. (3) and (4), the branching ratios of W decaying to trilepton final states via the sterile neutrino in the LNV processes go as

$$\mathrm{Br}(W^\pm \to e^\pm e^\pm \mu^\mp \nu) \propto s \times r,$$
$$\mathrm{Br}(W^\pm \to \mu^\pm \mu^\pm e^\mp \nu) \propto \frac{s}{r}, \quad (7)$$

while the branching ratios in the LNC processes go as

$$\mathrm{Br}(W^\pm \to e^\pm e^\pm \mu^\mp \nu) \propto s,$$
$$\mathrm{Br}(W^\pm \to \mu^\pm \mu^\pm e^\mp \nu) \propto s. \quad (8)$$

TABLE I. Scale factors due to lepton mixing, for the production rates of the different trilepton modes.

	Dirac	Majorana
$e^\pm e^\pm \mu^\mp \nu$	s	$s(1 + r)$
$\mu^\pm \mu^\pm e^\mp \nu$	s	$s(1 + 1/r)$

Therefore, the production rates for the Dirac case (LNC process only) and Majorana case (both LNC and LNV processes) corresponding to different trilepton final states are proportional to the scale factors shown in Table I.

In the following section we present our studies of simulated events in pp collisions at 14 TeV at the LHC within this theoretical framework, including SM backgrounds.

III. DATA SIMULATIONS AND BACKGROUND STUDIES

In this section, we describe in detail our event simulations, the observables that can be used to reject the SM backgrounds, and our strategies to determine the discovery potential of a heavy sterile neutrino N with mass below M_W and to determine its Majorana or Dirac character. For the data simulation, similar to our previous work [22], we build a Universal FeynRules Output [25] model file that extends the SM with additional sterile neutrino interactions using FEYNRULES [26] and implement it into MADGRAPH 5 [27] to generate the signal events. We explore two benchmark points: $m_N = 20$ GeV and 50 GeV, both with $r = s = 1$ (i.e., $|U_{Ne}|^2 = |U_{N\mu}|^2 = 10^{-6}$). The background events are also generated with MADGRAPH 5. The parton showering and hadronization are finished with PYTHIA 6 [28], while the detector simulation is completed with the help of DELPHES 3 [29]. At the parton level, we include up to two extra partons for both signal and background processes, and the jet matching is performed using the MLM-based shower-$k_\perp$ scheme [30]. To maintain consistency through all our study, the production cross sections calculated by MADGRAPH 5 are used to estimate the number of events for both signal and background processes as well.

Although in this trilepton search we demand no lepton pairs with opposite signs and the same flavor in the final state (no-OSSF) in order to reject backgrounds from radiative pairs, there still exist non-negligible backgrounds from various processes. The dominant SM backgrounds can be divided into two categories: (i) from leptonic τ decays and (ii) from fake leptons. In the first category, the dominant process is the pair production of WZ with W decaying leptonically and $Z \to \tau\tau$. The trilepton final states with no-OSSF pairs can arise from the subsequent leptonic decay of τ's. We estimate this background process via Monte Carlo simulations.

The dominant processes of the second category are $\gamma^*/Z + \text{jets}$ and $t\bar{t}$, where two leptons come from

$\gamma^*/Z \to \tau\tau$ or the prompt decay of t and $\bar{t}$, and a third lepton is faked from jets containing heavy-flavor mesons. Although in general fake leptons from such heavy-flavor meson decays are not well isolated, there are still rare occasions when they can pass the lepton isolation criteria [31–33]. Because these background processes ($\gamma^*/Z +$ jets and $t\bar{t}$) have large cross sections and small fake probabilities, it is very challenging to obtain enough statistics for background study in a purely Monte Carlo (MC) simulation. Moreover, simulating such processes requires a detailed modeling of the jet fragmentations, and the current level of the MC simulation may not be accurate enough. For these reasons, data-driven methods are used by the ATLAS and CMS Collaborations to estimate the fake lepton contributions [34–36], a matter beyond the scope of this study.

In this work, similar to our previous study [22], we adopt a phenomenological *fake lepton* (FL) simulation method, originally introduced in Ref. [37] and later also implemented in Ref. [19]. A fake lepton originates from a jet that generates an imprint in the detector that resembles that of a lepton, and therefore it inherits part of the kinematics of the actual jet. For the FL simulation, two modeling functions are introduced: a *mistag efficiency*, $\epsilon_{j\to\ell}$, which is the probability of a particular jet to be faked as a lepton, and a *transfer function*, $\mathcal{T}_{j\to\ell}$, which is a probability distribution that determines how much of the jet momentum is transferred into the fake lepton. These two functions contain some modeling parameters that can be fitted by validating simulated results against those of the actual experiment. We revisited the validation performed in Ref. [19] and found that the modeling parameters they obtained are consistent with the experimental results. Thus, the same set of parameters are used here. We also assume the same fake efficiency for electrons and muons. Details of this FL simulation method and the validation can be found in Refs. [19,37]. Our validation results are shown in Appendix A.

For the analysis, we first select the events with three leptons $l^\pm l^\pm l'^\mp$ that have no lepton pairs with no-OSSF. Then, the following basic cuts for the leptons and jets are applied: $p_{T,l} \geq 10$ GeV and $|\eta_l| \leq 2.5$; $p_{T,j} \geq 20$ GeV and $|\eta_j| \leq 5.0$. We also veto b jets to suppress the $t\bar{t}$ background.

Then, to pick up the correct lepton from the N decay in the same sign same flavor lepton pair ($l^\pm l^\pm$) for the rest of the analysis, we construct a chi-square function

$$\chi^2 = \left(\frac{M_W - m_W}{\sigma_W}\right)^2 + \left(\frac{M_N - m_N}{\sigma_N}\right)^2, \qquad (9)$$

where m_W is the input W mass of 80.5 GeV; m_N is the assumed N mass (20 or 50 GeV in our benchmarks), while M_W and M_N are the reconstructed W and N masses from the invariant mass of the $l^\pm l^\pm l'^\mp \nu$ and $l^\pm l'^\mp \nu$

systems, respectively; σ_W and σ_N denote the widths of the reconstructed mass distributions, which are taken to be 5% of the respective m_W and m_N, for simplicity. When calculating the reconstructed masses M_W and M_N, the final neutrino transverse momentum $\vec{p}_{T,\nu}$ is assumed to be the missing transverse momentum. The neutrino longitudinal momentum $p_{z,\nu}$ and the correct lepton $l^\pm$ from the N decay are determined by minimizing the χ^2 of Eq. (9).

The distributions of the reconstructed W and N masses for Dirac signals using this method, after applying the basic cuts and b-jet vetoes, are shown in Fig. 2. Since usually the N mass cannot be reconstructed correctly with the incorrect neutrino longitudinal momentum $p_{z,\nu}$ or with the wrong lepton from the same-sign lepton pair ($l^\pm l^\pm$), and a good reconstructed W mass also requires a correct $p_{z,\nu}$, the sharp resonances around our benchmarks $m_N = 20$ and 50 GeV for the M_N distributions, and around $m_W = 80.5$ GeV for the M_W distribution, indicate that one can indeed find the correct $p_{z,\nu}$ and identify the lepton from the N decay effectively by minimizing the χ^2 of Eq. (9). Since the leptons from the N decay for the $m_N = 20$ GeV case are softer and affected more by the lepton threshold cuts than for the $m_N = 50$ GeV case, the corresponding peaks in the M_W distribution for the former case are less sharp compared to the $m_N = 50$ GeV case.

Once the correct neutrino longitudinal momentum $p_{z,\nu}$ is found and the right lepton $l^\pm$ from the N decay is identified from the lepton pair $l^\pm l^\pm$ by minimizing the above χ^2 of Eq. (9), a set of different kinematical observables that are sensitive to the presence of a heavy sterile neutrino N, and also sensitive to its Dirac/Majorana character, can be constructed. The full list of these observables will be stated in Sec. IV. In Fig. 3, we present the distributions of some of them for both Dirac N signals with $m_N = 20$ and 50 GeV, and for the SM backgrounds of γ^*/Z, WZ, and $t\bar{t}$ after applying the basic cuts and the b-jet vetoes. One can see that the distributions for signals and backgrounds are quite different; thus these observables can actually be used to reduce the SM backgrounds effectively. Compared with the $m_N = 50$ GeV case, most signal distributions for $m_N = 20$ GeV are more separate from the backgrounds, so these observables can be more efficient to reject the backgrounds
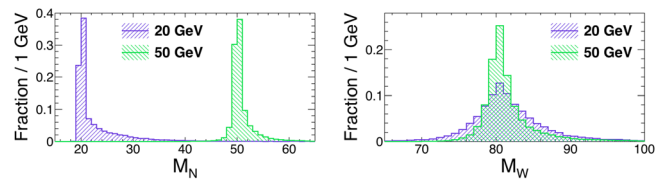

FIG. 2. Distributions of the reconstructed N mass M_N (left) and W mass M_W (right) for Dirac signals with $m_N = 20$ and 50 GeV after applying the basic cuts and b-jets veto.
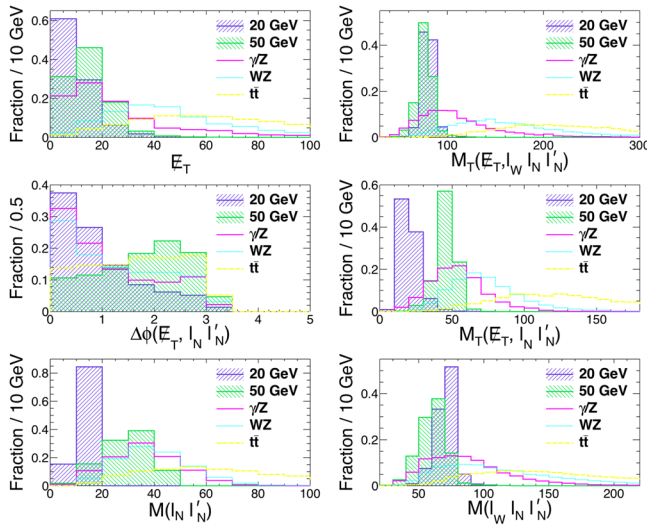
FIG. 3. Kinematial distributions for Dirac signals with $m_N =$ 20 and 50 GeV, and SM backgrounds of $\gamma^*/Z +$ jets, $WZ +$ jets, and $t\bar{t} +$ jets after applying the basic cuts and b-jets veto.

and can lead to larger significances for the $m_N = 20$ GeV benchmark point.

IV. DISCOVERING STERILE NEUTRINOS WITH TRILEPTON MODES

In this section we describe our strategy to search for heavy sterile neutrinos using trileptons at the LHC. After applying the basic cuts and b-jet vetoes, a MVA is performed to exploit useful observables and maximally reduce the SM background. We use the boosted decision trees (BDT) method in the Toolkit for MultiVariate data Analysis with ROOT (TMVA) package [38], and input the following kinematical observables (i)–(viii), which include those presented in Fig. 3, for training and test processes:

(i) the missing energy $\not{E}_T$;
(ii) the scalar sum of p_T of all jets H_T;
(iii) the invariant mass of the system of leptons $M(l_w l_N l'_N)$, $M(l_w l_N)$, $M(l_w l'_N)$, $M(l_N l'_N)$;
(iv) the azimuthal angle difference $\Delta\phi$ between two leptons $\Delta\phi(l_W, l'_N)$, $\Delta\phi(l_N, l'_N)$;
(v) the transverse mass M_T of the system formed by the missing momentum plus lepton(s) $M_T(\not{E}_T, l_W)$, $M_T(\not{E}_T, l_N l'_N)$, $M_T(\not{E}_T, l_w l_N l'_N)$;
(vi) the azimuthal angle difference $\Delta\phi$ between the missing transverse momentum and lepton(s) $\Delta\phi(\not{E}_T, l_N l'_N)$, $\Delta\phi(\not{E}_T, l_W)$;
(vii) the transverse mass M_T of the system formed by the missing momentum plus lepton(s) $M_T(\not{E}_T, l_N)$, $M_T(\not{E}_T, l'_N)$, and $M_T(\not{E}_T, l'_N l_W)$;
(viii) the azimuthal angle difference $\Delta\phi$ between the missing transverse momentum and lepton(s) $\Delta\phi(\not{E}_T, l_N)$, $\Delta\phi(\not{E}_T, l'_N)$, and $\Delta\phi(\not{E}_T, l'_N l_W)$.

The observables (vii) and (viii) in particular are found to differ between the LNC and LNV processes and can be
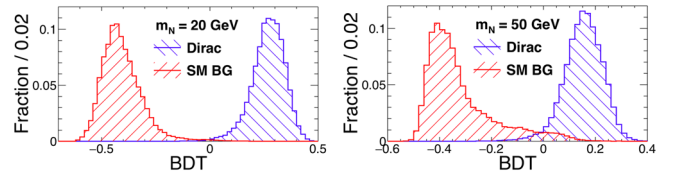

FIG. 4. Distributions of the BDT response for the Dirac signal (blue lines) with $m_N = 20$ (left) and 50 (right) GeV, and total SM backgrounds (red lines) including $\gamma^*/Z +$ jets, $WZ +$ jets, and $t\bar{t} +$ jets.

utilized to determine the Majorana/Dirac nature of N. That part of the study is presented in Sec. V.

For the Dirac (Majorana) sterile neutrinos, the simulation data of the LNC (LNC plus LNV) processes are inputs as the signal sample, while the total SM background data (γ^*/Z, WZ, and $t\bar{t}$ inclusively) are inputs as the background sample for the TMVA training and test processes.

Figure 4 shows the distributions of the BDT response for the Dirac sterile neutrino signals and the total SM background including $\gamma^*/Z +$ jets, $WZ +$ jets, and $t\bar{t} +$ jets, in the two benchmark cases. The kinematical distributions of the signal and of the SM backgrounds differ from each other more for $m_N = 20$ GeV than for $m_N = 50$ GeV.

In Table II we show the number of events for both Dirac and Majorana signals and the SM backgrounds at the 14 TeV LHC with integrated luminosity of 3000 fb^{-1}, for the case with $m_N = 20$ GeV. The first two rows show the number of events after basic cuts and b-jet vetoes. The number of events using the CC method used in Ref. [22] is shown in the third row. The numbers of events for Dirac (Majorana) sterile neutrinos using the BDT method are shown in the fourth (fifth) row. For our benchmark point $m_N = 20$ GeV Dirac (Majorana) sterile neutrino, one can get a statistical significance $SS = N_s/\sqrt{N_s + N_b}$ of about 2.6 (5.8) for the CC method and of about 6.6 (10.7) for the BDT method, where N_s is the number of signal events, while N_b is the corresponding number of total SM background events. Similarly, for the benchmark point $m_N = 50$ GeV the numbers are shown in Table III, where we find significances of about 2.3 (4.8) for the CC method and of

TABLE II. Cut flow for signal and background processes with $m_N = 20$ GeV. Numbers of events correspond to an integrated luminosity of 3000 fb^{-1} at the 14 TeV LHC.

Cuts	Dirac	Majorana	γ^*/Z	WZ	$t\bar{t}$	SS
Basic cuts	54.0	133.2	4220	2658	68588	
$N(b\,\text{jets}) = 0$	53.1	131.1	4063.0	2497.1	31953.5	
CC	44.2	110.9	209.8	25.3	16.9	2.6 (5.8)
BDT > 0.183	46.7	$\cdots$	1.9	1.3	0.0	6.6
BDT > 0.171	$\cdots$	120.7	5.1	1.7	0.8	10.7

TABLE III. Cut flow for signal and background processes with $m_N = 50$ GeV. Numbers of events correspond to an integrated luminosity of 3000 fb^{-1} at the 14 TeV LHC.

Cuts	Dirac	Majorana	γ^*/Z	WZ	$t\bar{t}$	SS	
Basic cuts	108.4	228.8	4220	2658	68588		
$N(b\,\mathrm{jets}) = 0$	106.7	225.2	4063.0	2497.1	31953.5		
CC	91.9	193.9	1283.1	120.7		48.9	2.3 (4.8)
BDT > 0.138	64.4	$\cdots$	25.7	47.5		21.1	5.1
BDT > 0.138	$\cdots$	143.2	31.0	52.8		27.0	9.0

FIG. 6. Discovery and exclusion limits for Majorana sterile neutrinos with $m_N = 20$ (left) and 50 (right) GeV. The blue curves marked with squares correspond to the 3-σ limit, while the red curves correspond to the 5-σ limit; solid lines are for the BDT method, and dashed lines are for the CC method.

about 5.1 (9.0) for the BDT method. From Fig. 4, one can see that the BDT cut is more efficient to reject the SM backgrounds for the $m_N = 20$ GeV than for 50 GeV; thus higher significances can be expected for the $m_N = 20$ GeV benchmark point. This is indeed what is found by comparing Tables II and III.

Figure 5 shows the discovery and exclusion curves for Dirac sterile neutrinos, for both the BDT and the CC methods. According to Table I, for a given m_N the production rates for Dirac sterile neutrinos (LNC processes) depend on the factor s only, so the observables at the LHC in the Dirac scenario can just constrain the parameter s. By exploiting more useful kinematical observables and better optimization compared with the CC method, the BDT method can greatly enhance the discovery and exclusion limits. Nevertheless, the performance of the BDT method becomes close to that of the CC method for small s values, due to the small number of signal events. Using the BDT method, for the benchmark $m_N = 20$ GeV one can get significances above $5.0\sigma(3.0\sigma)$ when $s \geq 0.55(0.25)$, and for the benchmark $m_N = 50$ GeV similar significances are reached when $s \geq 1.02(0.55)$.

Figure 6 shows the discovery and exclusion contour curves for Majorana sterile neutrinos, for both the BDT and the CC methods. Since the production rate for Majorana N involves both LNC and LNV processes, it depends on both the normalization s and the ratio r (see Table I). Thus the observables at the LHC in the Majorana scenario can be used to constrain both s and r. Using the BDT method, when $r = 1$ one can get significances above $5.0\sigma(3.0\sigma)$ with $s \geq 0.24(0.11)$ in the case $m_N = 20$ GeV, and with $s \geq 0.46(0.25)$ in the case $m_N = 50$ GeV. For a given value of s, the significance becomes larger when either $r \gg 1$ or $r \ll 1$, due to the larger number of signal events.
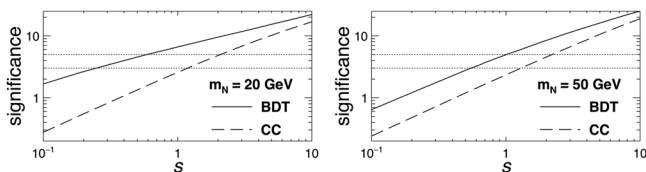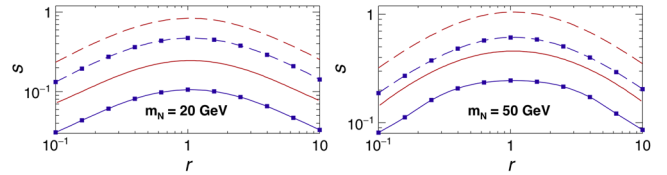
For example, when $r \approx 10$, one can get significances above $5.0\sigma(3.0\sigma)$ with $s \geq 0.08(0.03)$ in the case $m_N = 20$ GeV, and with $s \geq 0.16(0.09)$ in the case $m_N = 50$ GeV.

V. DISTINGUISHING BETWEEN DIRAC AND MAJORANA

In this section, we show how to distinguish between Dirac and Majorana sterile neutrinos using the trilepton events. Recalling that a Majorana N induces both LNV and LNC processes, while a Dirac N induces LNC processes only, a discrimination between Dirac and Majorana N can be achieved based on the differences between the LNC and LNV processes. As mentioned in Sec. IV, among all input observables, the distributions of the observables (vii) and (viii) are found to differ between the LNC and LNV processes. The corresponding theoretical expressions can be deduced from Sec. II. In Fig. 7 we show, as an example for $m_N = 20$ GeV, the distribution of $\Delta\phi(E_T, l'_N)$ in the N rest frame, where we check that the simulations (right plot) indeed reproduce the distinguishing features of the theoretical behavior (left plot) and are not purely an effect of statistical fluctuations. We use these observables to try to distinguish the Majorana nature of N from a Dirac scenario, the latter taken as the null hypothesis.

In order to exploit the differences between the Dirac and Majorana processes, we must first reduce as much SM background as possible; otherwise, the distributions will be dominated by the SM backgrounds and the differences will become imperceptible.
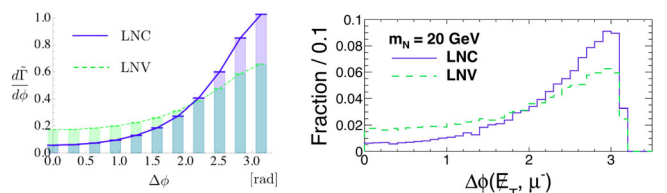

FIG. 5. Discovery and exclusion limits for Dirac sterile neutrinos with $m_N = 20$ (left) and 50 (right) GeV.

FIG. 7. The $\Delta\phi(E_T, \mu^-)$ distributions in the N rest frame from the $W^+ \to e^+e^+\mu^-\nu$ process for the benchmark point $m_N = 20$ GeV by theoretical calculation (left) and data simulation (right). Solid blue and dashed green lines correspond to LNC and LNV processes, respectively.

TABLE IV.   Cut flow for benchmark point with $m_N = 20$ GeV and SM backgrounds. From SM backgrounds, l denotes either e or μ. Numbers of events correspond to an integrated luminosity of 3000 fb^{-1} at the 14 TeV LHC.

Cuts	$e^+e^+\mu^-$		$\mu^+\mu^+e^-$		$e^-e^-\mu^+$		$\mu^-\mu^-e^+$		$l^\pm l^\pm l'^\mp$	$l^+l^+l'^-$	$l^-l^-l'^+$	$l^\pm l^\pm l'^\mp$
	LNC	LNV	LNC	LNV	LNC	LNV	LNC	LNV	γ^*/Z	W^+Z	W^-Z	$t\bar{t}$
Basic cuts	13.6	19.5	15.0	22.0	12.1	18.2	13.3	19.5	1055.0	779.0	550.0	17147.0
$N(b\,\text{jets}) = 0$	13.4	19.2	14.7	21.7	11.9	17.9	13.1	19.2	1015.8	731.8	516.7	7988.4
BDT1 > 0.171	12.2	17.7	13.5	20.0	10.9	16.5	12.0	17.7	1.2	0.5	0.4	0.2

Therefore, as a first step, after applying the basic cuts and the b-jet vetoes, we perform the first BDT analysis and input the kinematical observables (i)–(vi) listed in the first paragraph of Sec. IV to suppress the SM backgrounds. Simulated Majorana data are input as the signal sample, while the total SM background data are input as the background sample for the TMVA training and testing processes.

Table IV shows the number of events after these cuts for the benchmark case $m_N = 20$ GeV. After the first BDT cut, the total number of events including all four different final states ($e^\pm e^\pm \mu^\mp$ and $\mu^\pm \mu^\pm e^\mp$) for the Dirac signals (the LNC rate only), for the Majorana signals (LNC plus LNV rates), and for the SM backgrounds (γ^*/Z, $W^\pm Z$, and $t\bar{t}$ inclusively), are 48.5, 120.4, and 7.3, respectively. The SM backgrounds are reduced to a negligible level, so the sample will be dominated by the N-induced signal events.

Since the parameter s is an unknown global scale, the Dirac and Majorana cases cannot be experimentally discriminated purely by the number of events. Therefore, as a second step we adjust s for the Dirac hypothesis to a value s_D that matches the number of events for the Majorana scenario, so that our simulation does not distinguish the two scenarios simply by the rates.

Just as in Ref. [22], the best matched value of s_D is found by minimizing the chi-square function

$$\chi_H^2 = -2\min_s \left\{ \ln\left(\prod_i \text{Poiss}[N_i^{\text{expc}}, N_i^{\text{obs}}(s)] \right) \right\}, \quad (10)$$

where i indicates a particular trilepton final state, and $\text{Poiss}(N^{\text{expc}}, N^{\text{obs}})$ denotes the probability of observing N^{obs} events in Poisson statistics when the number of expected events is N^{expc}. Here N^{expc} is the expected number of events for the Majorana scenario (LNC + LNV + SM background), while N^{obs} is the observed number of events for the Dirac hypothesis (LNC + SM background). The best matched parameter s_D found in this way gives for the Dirac hypothesis a number of events closest to those of the Majorana case for all four different final states ($e^\pm e^\pm \mu^\mp$ and $\mu^\pm \mu^\pm e^\mp$). For our $m_N = 20$ GeV benchmark, the matched parameter s_D is found to be around 2.44. After

matching, the Dirac hypothesis will have 125.6 events, which is close to the 127.6 events of the Majorana scenario.

As a third step, we perform a second BDT analysis to distinguish a Majorana scenario from the Dirac hypothesis, by exploiting the differences in the kinematical distributions between the LNC and LNV processes. The input observables are those (vii) and (viii) listed in the first paragraph of Sec. IV. The distributions of these observables for the LNC and LNV processes, after applying the basic cuts, b-jet vetoes, and the first BDT cut, are presented in Figs. 8 and 9 for the $m_N = 20$ and 50 GeV benchmarks, respectively.

In the second BDT analysis, the simulated data for the Majorana scenario (LNV + LNC + SM background after the first BDT cut) are input as the signal sample, while the simulated data for the Dirac hypothesis (LNC + SM background, with matching s_D) are input as the background sample in the TMVA training and testing processes. The BDT distribution will then indicate the differences between the Majorana scenario and the Dirac hypothesis. Figure 10 shows the BDT response for the Majorana case and the Dirac hypothesis for $m_N = 20$ GeV (left) and 50 GeV (right) benchmark points. Comparing the left and right plots, we see that the histograms are more separated in
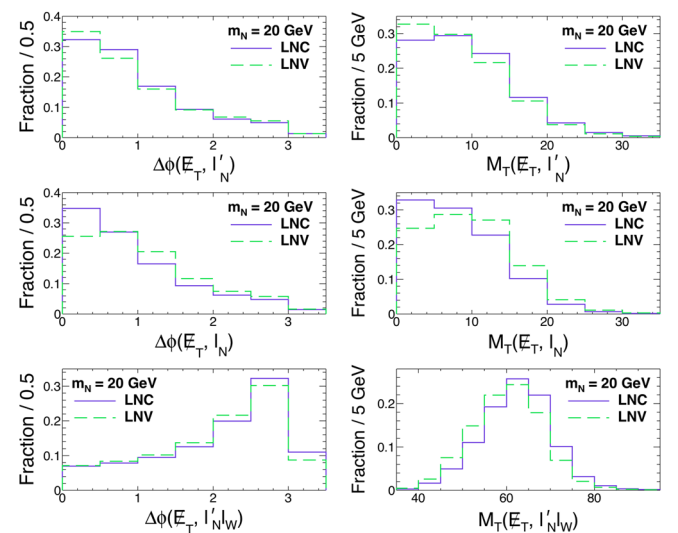

FIG. 8.   Distributions for the benchmark point $m_N = 20$ GeV after applying the basic cuts, b-jet vetoes, and the first BDT cut.
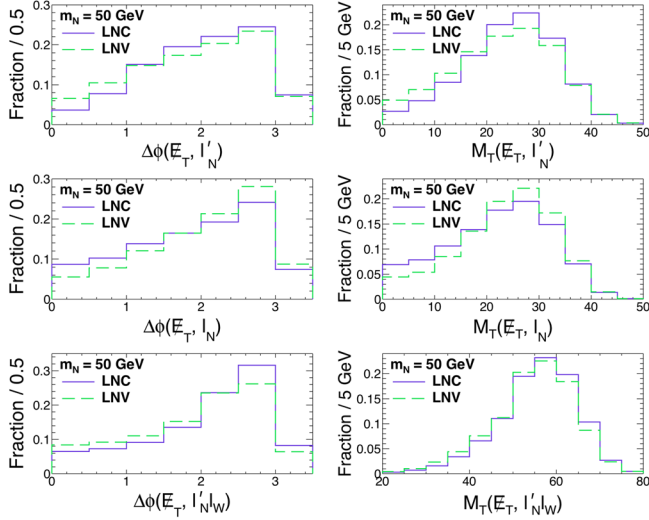
FIG. 9. Distributions for the benchmark point $m_N = 50$ GeV after applying the basic cuts, b-jet vetoes, and the first BDT cut.

the left plot, leading to a better BDT cut efficiency and thus a higher significance for the $m_N = 20$ GeV benchmark point.

With an optimized second BDT cut of about 0.020, the Majorana case ends up with 46.1 events, while the Dirac hypothesis has 34.1 events. After defining the number of events corresponding to the excess in the Majorana case from the Dirac hypothesis as the "signal" events N_s, and the number of events corresponding to the Dirac hypothesis as the "background" events N_b, the statistical significance for distinguishing the Majorana scenario from the Dirac hypothesis can be calculated as $SS = N_s/\sqrt{N_s + N_b} = (46.1 - 34.1)/\sqrt{46.1} \approx 1.8$.

This three-step method can be extended to the cases where $r \neq 1$. For a given value of the parameter s, when $r \gg 1$ or $r \ll 1$, from Table I one can see that the relative number of events for different trilepton states will be quite different in the Majorana scenario, but not so in the Dirac scenario. This feature helps in the Majorana/Dirac discrimination and results in higher significances. Figure 11 shows the confidence levels for distinguishing between Majorana and Dirac scenarios, obtained with the above three-step method. When $r \approx 1$, one can have significances near $5.0\sigma(3.0\sigma)$ when $s \geq 7.93(3.10)$ in the $m_N = 20$ GeV
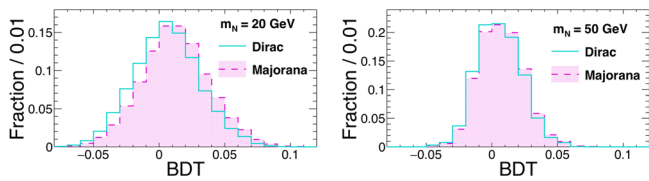

FIG. 10. Distributions of the BDT response in the second BDT analysis, for the Dirac hypothesis (dashed line with filled area) and the Majorana scenario (solid line), for the benchmarks $m_N = 20$ GeV (left) and 50 GeV (right).
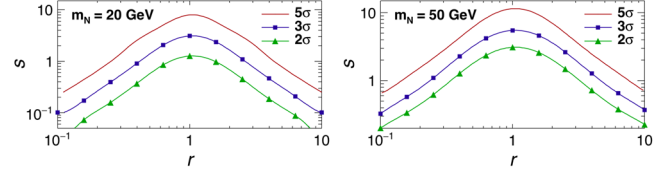

FIG. 11. Confidence levels of distinguishing between Dirac and Majorana neutrinos for $m_N = 20$ (left) and 50 (right) GeV.

benchmark, and when $s \geq 11.44(5.47)$ in the $m_N = 50$ GeV benchmark. As $r \approx 10$, to reach the same significance, the parameter s can be as low as 0.25 (0.10) for $m_N = 20$ GeV, and as low as 0.72 (0.38) for $m_N = 50$ GeV.

VI. CONCLUSIONS

We present a method to detect and distinguish Dirac and Majorana heavy sterile neutrinos with masses near or below the W boson mass, based on the experimental search of the purely leptonic decays $W^\pm \to e^\pm e^\pm \mu^\mp \nu$ and $W^\pm \to \mu^\pm \mu^\pm e^\mp \nu$ at the 14 TeV LHC, which are induced by a heavy neutrino in the intermediate state. The method is based on both a CC as well as a MVA. Our analysis sets discovery limits on the heavy-to-light lepton mixings $|U_{Ne}|^2$ and $|U_{N\mu}|^2$, which we express here in terms of the parameters $s = 2 \times 10^6 \times |U_{Ne}U_{N\mu}|^2/(|U_{Ne}|^2 + |U_{N\mu}|^2)$ and $r = |U_{Ne}/U_{N\mu}|^2$. The discovery potential of heavy Dirac neutrinos depends on s only, while in the case of Majorana neutrinos it depends on both s and r. The best results are found with the MVA method; nevertheless, the performance of the MVA method becomes close to that of the CC method for small s values, due to the smaller number of signal events. We use two benchmark points for the heavy neutrino mass: $m_N = 20$ and 50 GeV, and assume an LHC integrated luminosity of 3000 fb^{-1}.

Using the MVA method, we find that Dirac sterile neutrinos can be discovered with a significance of at least $5.0\sigma(3.0\sigma)$ when $s \geq 0.55(0.25)$ in the case $m_N = 20$ GeV, or when $s \geq 1.02(0.55)$ in the case $m_N = 50$ GeV. Let us recall that, for $r = 1$, the mixings are $|U_{Ne}|^2 = |U_{N\mu}|^2 = s \times 10^{-6}$.

For Majorana sterile neutrinos, if $r = 1$, the same level of significance can be reached for lower values of s because now the events come from both the LNC and LNV processes. Indeed, a significance of $5.0\sigma(3.0\sigma)$ is reached when $s \geq 0.24(0.11)$ in the case $m_N = 20$ GeV, or when $s \geq 0.46(0.25)$ in the case $m_N = 50$ GeV. For the same s but $r \neq 1$ the significances also become larger for a Majorana neutrino, due to the larger number of events: when, e.g., $r = 10$, a significance of $5.0\sigma(3.0\sigma)$ is reached when $s \geq 0.08(0.03)$ for the benchmark $m_N = 20$ GeV, and $s \geq 0.16(0.09)$ for the benchmark $m_N = 50$ GeV. Let us now recall that, for $r = 10$, the mixings are $|U_{Ne}|^2 = 10|U_{N\mu}|^2 = 5.5s \times 10^{-6}$.

Finally, the Dirac and Majorana hypotheses can be distinguished from each other at the $5.0\sigma(3.0\sigma)$ level of significance when $r \approx 1$ provided that $s \geq 7.93(3.10)$ for the benchmark $m_N = 20$ GeV, and $s \geq 11.44(5.47)$ for the benchmark $m_N = 50$ GeV. For $r \ll 1$ or $r \gg 1$, lower values of s are required: as $r \approx 10$, to reach the same significance the parameter s can be as low as 0.25 (0.10) for $m_N = 20$ GeV, or as low as 0.72 (0.38) for $m_N = 50$ GeV.

ACKNOWLEDGMENTS

APPENDIX: VALIDATION FOR FAKE LEPTON SIMULATION

In this appendix, we present our validation results for the fake lepton simulation used in this work. We follow closely the same validation done in Ref. [19], and find out that using their modeling parameters, the simulation results can indeed be consistent with the experimental results given in Ref. [36]. Specifically, we take $r_{10} = 1$, $\mu = 0.5$, $\sigma = 0.3$, and $\epsilon_{200} = 4.6 \times 10^{-3}$. In fact, the suggested mistag rate of $\epsilon_{200} = 4.6 \times 10^{-3}$ coincides with the "rule-of-thumb" introduced in Ref. [33]; i.e., isolated electrons and muons from heavy-flavor decay are about $1/200$ times the rates of b and c quark production. For the other input parameters of r_{10}, μ, and σ, the authors of Ref. [19] find that varying them does not substantially change the fitting to the data, provided the overall fake efficiency of ϵ_{200} remains fixed.

Our validation results are shown in Fig. 12. Each bin represents an event category according to the CMS trilepton search given in Ref. [19], namely, (1) 0-b jet, 1-OSSF,
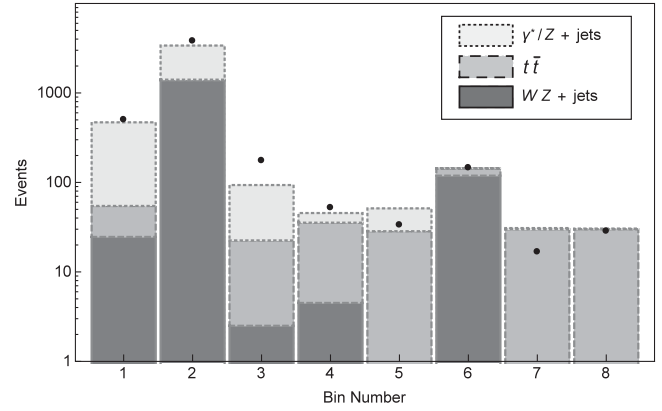

FIG. 12. Validation results for fake lepton simulation. Black dots indicate experimental results in Ref. [36]. Our simulated results for γ^*/Z + jets, $t\bar{t}$, and WZ + jets are given by light gray bars, middle brown bars, and bottom pink bars, respectively. Eight bin categories are as follows: (1) 0-b jet, 1-OSSF, $M_{\ell_+,\ell_-} < 75$ GeV; (2) 0-b jet, 1-OSSF, $|M_{\ell_+,\ell_-} - M_Z| < 15$ GeV; (3) 0-b jet, 1-OSSF, $M_{\ell_+,\ell_-} > 105$ GeV; (4) 0-b jet, 0-OSSF; (5)–(8) are the same as the first four bins, but with at least one b jet.

$M_{\ell_+,\ell_-} < 75$ GeV; (2) $0 - b$ jet, 1-OSSF, $|M_{\ell_+,\ell_-} - M_Z| < 15$ GeV; (3) 0-b jet, 1-OSSF, $M_{\ell_+,\ell_-} > 105$ GeV; (4) 0-b jet, 0-OSSF; (5)–(8) are the same as the first four bins, but with at least one b jet. The actual experiment results are indicated by black dots, while our simulated results are given by upper light bars, middle dark bars, and bottom light bars for the processes γ^*/Z + jets, $t\bar{t}$, and WZ + jets, respectively. As one can see, our results agree with the experimental results reasonably well within the statistical uncertainties, especially in bin-4, whose selection criteria mostly resemble the ones stated in our main text. Moreover, a good agreement with the results given in Fig. 10 of Ref. [19] is also found, although in some bins we differ in the individual fractions of events from different processes.

[1] Y. Fukuda *et al.* (Super-Kamiokande Collaboration), Phys. Rev. Lett. **81**, 1562 (1998); Q. R. Ahmad *et al.* (SNO Collaboration), Phys. Rev. Lett. **89**, 011301 (2002).

[2] E. Majorana, Nuovo Cimento **14**, 171 (1937); G. Racah, Nuovo Cimento **14**, 322 (1937).

[3] H. V. Klapdor-Kleingrothaus *et al.*, Eur. Phys. J. A **12**, 147 (2001); A. M. Bakalyarov *et al.* (C03-06-23.1 Collaboration), Pis'ma Fiz. Elem. Chast. Atom. Yadra **2005**, 21 (2005) [Phys. Part. Nucl. Lett. **2**, 77 (2005)]; H. V. Klapdor-Kleingrothaus and I. V. Krivosheina, Mod. Phys. Lett. A **21**, 1547 (2006); M. Agostini *et al.* (GERDA Collaboration),

Phys. Rev. Lett. **111**, 122503 (2013); A. Pocar (EXO-200 and nEXO Collaborations), Nucl. Part. Phys. Proc. **265**–**266**, 42 (2015); Y. Gando (KamLAND-Zen Collaboration), Nucl. Part. Phys. Proc. **273**–**275**, 1842 (2016); A. Gando *et al.* (KamLAND-Zen Collaboration), Phys. Rev. Lett. **117**, 08 2016; **117**, 109903(E) (2016).

[4] J. Engel and J. Menndez, Rep. Prog. Phys. **80**, 046301 (2017).

[5] P. Minkowski, Phys. Lett. **67B**, 421 (1977).

[6] M. Gell-Mann, P. Ramond, and R. Slansky, arXiv:1306.4669; in *Supergravity*, edited by D. Freedman *et al.* (North-Holland, Amsterdam, 1979).

[7] T. Yanagida, Conference : Proceedings / The Australasian Corrosion Association **7902131**, 95 (1979).

[8] R. N. Mohapatra and G. Senjanović, Phys. Rev. Lett. **44**, 912 (1980); Phys. Rev. D **23**, 165 (1981).

[9] S. L. Glashow, in *Quarks and Leptons*, edited by M. Levy *et al.* (Plenum, New York, 1980), p. 707.

[10] J. Schechter and J. W. F. Valle, Phys. Rev. D **22**, 2227 (1980); **25**, 774 (1982).

[11] M. Magg and C. Wetterich, Phys. Lett. **94B**, 61 (1980).

[12] T. P. Cheng and L. F. Li, Phys. Rev. D **22**, 2860 (1980).

[13] G. Lazarides, Q. Shaf,i, and C. Wetterich, Nucl. Phys. **B181**, 287 (1981).

[14] R. Foot, H. Lew, X. G. He, and G. C. Joshi, Z. Phys. C **44**, 441 (1989).

[15] W. Y. Keung and G. Senjanovic, Phys. Rev. Lett. **50**, 1427 (1983); A. Atre, T. Han, S. Pascoli, and B. Zhang, J. High Energy Phys. 05 (2009) 030; G. Cvetic, C. Dib, S. K. Kang, and C. S. Kim, Phys. Rev. D **82**, 053010 (2010); C. O. Dib and C. S. Kim, Phys. Rev. D **92**, 093009 (2015).

[16] O. Seon *et al.* (BELLE Collaboration), Phys. Rev. D **84**, 071106 (2011); J. P. Lees *et al.* (*BABAR* Collaboration), Phys. Rev. D **85**, 071103 (2012); J. P. Lees *et al.* (*BABAR* Collaboration), Phys. Rev. D **89**, 011102 (2014); B. Adeva (LHCb Collaboration), J. Phys. Conf. Ser. **447**, 012062 (2013); B. Khanji (LHCb Collaboration), Nucl. Phys. B, Proc. Suppl. **248**–**250**, 91 (2014).

[17] S. Chatrchyan *et al.* (CMS Collaboration), Phys. Lett. B **717**, 109 (2012); V. Khachatryan *et al.* (CMS Collaboration), Phys. Lett. B **748**, 144 (2015); J. High Energy Phys. 04 (2016) 169.

[18] G. Aad *et al.* (ATLAS Collaboration), Eur. Phys. J. C **72**, 2056 (2012); J. High Energy Phys. 07 (2015) 162.

[19] E. Izaguirre and B. Shuve, Phys. Rev. D **91**, 093010 (2015).

[20] C. O. Dib and C. S. Kim, Phys. Rev. D **92**, 093009 (2015).

[21] G. Cvetic, C. O. Dib, and C. S. Kim, J. High Energy Phys. 06 (2012) 149.

[22] C. O. Dib, C. S. Kim, K. Wang, and J. Zhang, Phys. Rev. D **94**, 013005 (2016).

[23] C. Patrignani *et al.* (Particle Data Group), Chin. Phys. C **40**, 100001 (2016).

[24] T. Han, I. Lewis, R. Ruiz, and Z. g. Si, Phys. Rev. D **87**, 035011 (2013); **87**, 039906(E) (2013).

[25] C. Degrande, C. Duhr, B. Fuks, D. Grellscheid, O. Mattelaer, and T. Reiter, Comput. Phys. Commun. **183**, 1201 (2012).

[26] N. D. Christensen and C. Duhr, Comput. Phys. Commun. **180**, 1614 (2009).

[27] 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.

[28] T. Sjostrand, S. Mrenna, and P. Z. Skands, J. High Energy Phys. 05 (2006) 026.

[29] 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.

[30] J. Alwall, S. de Visscher, and F. Maltoni, J. High Energy Phys. 02 (2009) 017.

[31] Z. Sullivan and E. L. Berger, Phys. Rev. D **74**, 033008 (2006).

[32] Z. Sullivan and E. L. Berger, Phys. Rev. D **78**, 034030 (2008).

[33] Z. Sullivan and E. L. Berger, Phys. Rev. D **82**, 014001 (2010).

[34] V. Khachatryan *et al.* (CMS Collaboration), Phys. Lett. B **748**, 144 (2015).

[35] G. Aad *et al.* (ATLAS Collaboration), J. High Energy Phys. 03 (2015) 041.

[36] S. Chatrchyan *et al.* (CMS Collaboration), Phys. Rev. D **90**, 032006 (2014).

[37] D. Curtin, J. Galloway, and J. G. Wacker, Phys. Rev. D **88**, 093006 (2013).

[38] A. Hoecker, P. Speckmayer, J. Stelzer, J. Therhaag, E. von Toerne, and H. Voss, *Proc. Sci.*, A CAT2007 (2007) 040 [arXiv:physics/0703039].