Asymptotic analysis of binned likelihoods and the neutrino floor

Jian Tang[†] and Bing-Long Zhang[*]

School of Physics, Sun Yat-Sen University, Guangzhou 510275, China

Observations of suspected coherent elastic neutrino-nucleus scatterings by dark matter direct detection experiments highlight the need for an investigation into the so-called "neutrino floor." We focus on the discovery limit, a statistical concept to identify the neutrino floor, and analyze the asymptotic behavior of the profile binned likelihood ratio test statistic where the likelihood is constructed by variate from events in each bin and pull terms from neutrino fluxes. To achieve the asymptotic result, we propose two novel methods: (i) the asymptotic-analytic method, which furnishes the analytic result for large statistics, is applicable for more extra nuisance parameters, and enables the identification of the most relevant parameters in the statistical analysis; (ii) the quasi-Asimov dataset, which is analogous to the Asimov dataset but with improved speed. Applying our methods to the neutrino floor, we significantly accelerate the computation procedure compared to the previous literature, and successfully address cases where the Asimov dataset fails. Our derivation on the asymptotic behavior of the test statistic not only facilitates research into the impact of neutrinos on the search for dark matter, but may also prove relevant in similar application scenarios.

I. INTRODUCTION

The nature of dark matter (DM) is one of most alluring fields in physics but still remains unknown with several decades of efforts on searching for DM [1–3]. From the perspective of DM direct detection (DD) experiments, however, constraints on DM, particularly a popular candidate—weakly interacting massive particles (WIMPs), have been improved through the dedicated efforts of DD experiments, which inspires us to pursue increasingly sensitive techniques for detecting recoil signals generated by WIMPs. The recent discovery [4] of the coherent elastic neutrino-nucleus scattering (CEνNS) process by COHERENT collaboration represents a recent milestone in this field. DM DD experiments, such as XENON and PandaX [5,6], have announced their observations of suspected solar neutrino signals generated by CEνNS. However, as sensitivities increase, the presence of irreducible neutrino backgrounds poses a challenge to the search for WIMPs, given the similarities between their respective signals. As such, it is essential to quantitatively evaluate the impact of neutrinos on direct searches for WIMPs.

Neutrino floor, proposed to systematically quantify the discovery potential of searching for WIMPs in the presence of neutrino backgrounds [7], is presented by a curve on the parameter space of WIMP-nucleon cross section versus WIMP mass, below which WIMPs cannot be detected significantly. Formally, the neutrino floor can be defined by discovery limits [8]: the minimum cross section required for an experiment to have a 90% probability or 90% confidence level (CL) to detect a WIMP signal with a 3σ significance [9]. Obtaining the neutrino floor traditionally requires Monte Carlo (MC) pseudoexperiments which can be time consuming. Additional studies on the neutrino floor have explored various factors such as astrophysical uncertainties [10], altering the WIMP-nucleon interaction with the nonrelativistic effective field theory formalism [11] and modifying neutrino backgrounds with neutrino related new physics [12–14], etc. There are several studies on how to overcome the neutrino floor, including combining data from experiments using different targets [9], taking advantage of the annual modulation [15] or the diurnal modulation [16], and adopting more realistic directional detection strategy [17–24] with lower statistics. Recently, a new definition on the neutrino floor has been proposed, which does not rely on choices on detector thresholds and exposures [25]. Instead, it uses the gradient of the discovery limit cross section with respect to the exposure to measure the marginal utility on the WIMP searches, resulting in the definition of the "neutrino fog." The neutrino floor is then identified as a boundary of the neutrino fog, representing the transition from statistical to systematical limits [26].

[*]Corresponding author: zhangblong@mail2.sysu.edu.cn
[†]tangjian5@mail.sysu.edu.cn

To reduce computational costs, the confidence level in the previous definition was adjusted to 50% and Asimov dataset [27] was used to obtain the median discovery limit.

There appears to be some unresolved issues for determining the neutrino floor: (i) the statistic methodology for the discovery limit lacks further investigation; (ii) the calculation of the neutrino floor necessitates substantial computational resources, and enhancing the stability of the algorithm is imperative; (iii) Asimov dataset cannot resolve cases that some specific degrees of freedom (DOFs) such as astrophysical uncertainties are considered, and (iv) MC pseudoexperiments increasingly cost time when more parameters are considered. Thus, in this paper, we infer the asymptotic behavior of the profile binned likelihood ratio test statistic (PBLRTS) where the likelihood is constructed by different variables. The similar case, where the likelihood is constructed by variables following the same distribution, has been worked out [28,29]. Subsequently, we propose two new methods to achieve the asymptotic result: (i) asymptotic-analytic method which provides analytic results for large statistics, is general for more extra nuisance parameters and affords a way to determine the most relevant parameters in the statistical analysis; and (ii) quasi-Asimov dataset analogous with but faster than the Asimov dataset method. By applying our methods to the neutrino floor, we significantly accelerate the computation procedure used in previous literature, investigate the neutrino fog taking into account the uncertainty of the weak mixing angle, and handle the case when a specific DOF from the astrophysical uncertainty is considered, which cannot be addressed by Asimov dataset. Moreover, the most relevant parameters from neutrino fluxes can be obtained by the asymptotic-analytic method, so that MC pseudoexperiments can be boosted by neglecting nondominant contributions from neutrino sources. Furthermore, our proposed methods are not only applicable for the neutrino floor and fog [14,19,25], but also might be feasible in searching for the bump of the diffuse flux of high-energy cosmic neutrinos in IceCube [30], studying the impact of DM on DSNB in Hyper-Kamiokande experiment [31], detecting time-varying DM signals with Paleo-Detectors [32] and analyzing the daily modulation on dark photon [33].

This paper is organized as follows. In Sec. II, we provide some key process in obtaining the asymptotic behavior of PBLRTS. In Sec. III, we describe the statistic method behind the neutrino floor and provide the result from our new methods. In addition, the asymptotic-analytic method is used to analytically explain the evolution of the discovery limit cross section with the exposure. In Sec. IV, we provide the neutrino fog and floor with considering the uncertainty of the weak mixing angle, and afford the modified discovery limit curve for experimental configurations with considering nuisance parameter from the velocity of the local standard of rest. Finally, in Sec. V we draw our conclusions and outlooks.

II. ASYMPTOTIC BEHAVIOR OF PROFILE BINNED LIKELIHOOD RATIO TEST STATISTIC

A. Derivation on asymptotic distribution

Our notation adheres to the conventions described in the book on large sample theory [34]. To begin with, we recall a theorem concerning the large sample distribution of the likelihood ratio test statistic (LRTS) [28,29]. In the context of the likelihood ratio test, we consider a parameter set with k DOFs denoted by $\boldsymbol{\theta} \in \Theta \subset \mathbb{R}^k$, where r $(1 \le r \le k)$ is the number of parameters of interest and (k-r) is the number of nuisance parameters. The likelihood function $L(\boldsymbol{\theta})$ is constructed by variables following the same distribution. The likelihood ratio test provides a general way for discriminating the null hypothesis $H_0: \boldsymbol{\theta}^0 \in \Theta_0$ versus the alternative hypothesis $H_1: \boldsymbol{\theta}^1 \in \Theta - \Theta_0$, where $\Theta_0 \subset \Theta$. Defining the log likelihood function $l(\boldsymbol{\theta}) \equiv \log L(\boldsymbol{\theta})$ and $\lambda \equiv \frac{L(\boldsymbol{\theta}^*)}{L(\hat{\boldsymbol{\theta}})}$, the null hypothesis is rejected if the likelihood ratio test statistic $-2 \ln \lambda = -2[\ln l(\boldsymbol{\theta}^*) - \ln l(\hat{\boldsymbol{\theta}})]$ is larger than our expectation, where $\boldsymbol{\theta}^*$ and $\hat{\boldsymbol{\theta}}$ represent the maximum-likelihood estimator (MLE) over Θ_0 and Θ, respectively. In general, Θ_0 is a (k-r)-dimensional subspace of Θ, with r constraint conditions given by: $g_i(\boldsymbol{\theta}) = 0 \, (1 \le i \le r)$. Under the assumption of large samples, if the true model $\boldsymbol{\theta}'$ satisfies $H_0: \boldsymbol{\theta}^0$, the quantity $-2 \ln \lambda$ follows the chi-square distribution with r DOFs, denoted by $-2 \ln \lambda \sim \chi_r^2$. If $\boldsymbol{\theta}'$ deviates from $\boldsymbol{\theta}^0$, then $-2 \ln \lambda \sim \chi_r^2(\phi)$, where ϕ is the noncentral parameter related to $\boldsymbol{\delta} = \boldsymbol{\theta}' - \boldsymbol{\theta}^0$.

However, the aforementioned theorem should be modified to be applied to our scenario, where the binned likelihood comprises variables that follow different distributions and are accompanied by pull terms. Consequently, it becomes imperative to extend the theorem scope. To this end, we invoke the Lyapunov central limit theorem (CLT) [35] to approximate some quantities to its expectation or acquire the asymptotic distribution of some quantities, which are constructed by random variate that might not be from the same distribution, in our proof. Moreover, our derivation in the following is based on the assumption that the likelihood function is smooth and its derivatives are bounded.

Here we simply expound upon the test statistic's asymptotic formula, and please refer to Appendix A for the detailed derivation. The asymptotic formula is constructed by the derivative of the logarithmic likelihood function. One is the first derivative at $\boldsymbol{\theta}'$, i.e., $\dot{l}(\boldsymbol{\theta}')$, which follows a multivariate normal distribution with a mean vector $\boldsymbol{\mu}$ and a covariance matrix $\mathbf{V}$, denoted by $\mathcal{N}(\boldsymbol{\mu}, \mathbf{V})$. The other is the expectation of the second derivative at $\boldsymbol{\theta}'$, i.e., $E(\ddot{l}(\boldsymbol{\theta}'))$. For convenience, we define $\mathcal{F} \equiv -E(\ddot{l}(\boldsymbol{\theta}'))$ and a matrix $\mathbf{H}$ which is also constructed by $E(\ddot{l}(\boldsymbol{\theta}'))$, and details are provided in Appendix A. Utilizing the Taylor's expansion, the Lyapunov CLT and some properties of quantities, the asymptotic formula for the test statistic can be written as:

$$-2\ln\lambda \approx Z^T \mathbf{V}^{\frac{1}{2}}[\mathcal{F}^{-1} - \mathbf{H}]\mathbf{V}^{\frac{1}{2}}Z, \quad (1)$$

where $Z \sim \mathcal{N}(\mathbf{V}^{-\frac{1}{2}}(\boldsymbol{\mu} + \mathcal{F}\boldsymbol{\delta}), I)$ is a vector. To make the result more explicit, we perform a diagonalization: $-2\ln\lambda \approx Z^T U^T \Lambda U Z = Y^T \Lambda Y$, where Λ is a diagonal matrix, U is a special orthogonal matrix and $Y = UZ \sim \mathcal{N}(U\mathbf{V}^{-\frac{1}{2}}(\boldsymbol{\mu} + \mathcal{F}\boldsymbol{\delta}), I)$. As a consequence, $-2\ln\lambda$ asymptotically follows the distribution of a sum over several noncentral chi-square variate with varying weights:

$$-2\ln\lambda \approx \sum_i a_i Y_i^2, \quad (2)$$

where a_i is the ith component of the diagonal term in Λ, and $Y_i^2 \sim \chi_1^2(\phi_i)$ where ϕ_i is the noncentral parameter. For instance, in the case presented in the next section where H$_0$ is true, we have $\boldsymbol{\delta} = 0$ and $Y \sim \mathcal{N}(0, I)$. Consequently, $-2\ln\lambda \sim \chi_1^2$, aligning with Wilk's theorem.[1]

Moreover, it is noteworthy to observe that the entries within the matrix U signify the impact exerted by each parameter. As a result, one can minimize the cost of MC pseudoexperiments by selectively incorporating only the most relevant parameters, and the criteria can be derived from the matrix U in our asymptotic-analytic method. This technique has been applied in our MC realizations, which is provided in Appendix C.

B. Asymptotic-analytic method

The asymptotic-analytic method has been presented previously, but it did not include a discussion on how to obtain the values of $\boldsymbol{\mu}$, $\mathbf{V}$, and $\mathcal{F}$. Here, we provide a concrete application.

For simplicity, we only consider the case of signal discovery where there is only one parameter of interest and all parameters are normalized to unity. For the discovery of a signal, θ_1 represents the signal strength, while $\theta_i (i \geq 2)$ are nuisance parameters. In many literature, the signal strength is always denoted by μ, while we use θ_1 here for convenience. Under the null hypothesis H$_0$: $\boldsymbol{\theta}^0$, we have $\theta_1^0 = 0$, $\theta_i^0 = 1 (i \geq 2)$. Here, the bold symbol represents a vector where the superscript marks the model, and the normal symbol like θ_i^0 represents the ith component of $\boldsymbol{\theta}^0$. Under the alternative hypothesis H$_1$: $\boldsymbol{\theta}^1$, we have $\theta_i^0 = 1 (i \geq 1)$. We adopt the binned likelihood with N bins and M parameters as shown below:

$$L(\theta_1, \theta_2, \ldots, \theta_m) = \prod_{i=1}^{N} \mathcal{P}(n_i | \theta_1 s_i + b_i) \prod_{j=2}^{M} \mathcal{N}_{\theta_j}(1, \sigma_j), \quad (3)$$

where n_i, s_i, and b_i are observed, signal and background events in ith bin, respectively. The variable n_i follows a Poisson distribution with a mean of the expected value $v_i = \theta_1 s_i + b_i$, i.e., $\mathcal{P}(n_i | v_i) = v_i^{n_i} e^{-v_i} / n_i!$. The quantities s_i and b_i are dependent on the nuisance parameters $\theta_i (i \geq 2)$, which follows a normal distribution with a mean of 1 and a standard deviation σ_j: $\mathcal{N}_{\theta_j}(1, \sigma_j) = \frac{1}{\sqrt{2\pi}\sigma_j} \exp\left(-\frac{(\theta_j - 1)^2}{2\sigma_j^2}\right)$.

Furthermore, we obtain the expressions for the expectation vector $\boldsymbol{\mu}$, the variance matrix $\mathbf{V}$ of the first derivative of $l(\boldsymbol{\theta})$ at $\boldsymbol{\theta}'$: $\dot{l}(\boldsymbol{\theta}')$ and the expectation of the second derivative with a minus sign $\mathcal{F}$ as follows:

$$\mu_\alpha = \begin{cases} \sum_{ij} \frac{1}{2} \frac{\partial^2 v_i}{\partial \theta_j^2} \frac{\sigma_j^2}{v_i} \frac{\partial v_i}{\partial \theta_\alpha}, & \alpha = 1, \\[2ex] \sum_{ij} \frac{1}{2} \frac{\partial^2 v_i}{\partial \theta_j^2} \frac{\sigma_j^2}{v_i} \frac{\partial v_i}{\partial \theta_\alpha} - \frac{\theta_\alpha - \theta_\alpha'}{\sigma_\alpha^2}, & 2 \leq \alpha \leq M, \end{cases}$$

$$\mathcal{F}_{\alpha\beta} = \begin{cases} -\sum_{ij} \frac{1}{2} \frac{\partial^2 v_i}{\partial \theta_j^2} \frac{\sigma_j^2}{v_i} \frac{\partial^2 v_i}{\partial \theta_\alpha \partial \theta_\beta} + \sum_{ij} \frac{1}{v_i} \left(1 + \frac{1}{2} \frac{\partial^2 v_i}{\partial \theta_j^2} \frac{\sigma_j^2}{v_i}\right) \frac{\partial v_i}{\partial \theta_\alpha} \frac{\partial v_i}{\partial \theta_\beta}, & \alpha = 1 \text{ or } \beta = 1, \\[2ex] -\sum_{ij} \frac{1}{2} \frac{\partial^2 v_i}{\partial \theta_j^2} \frac{\sigma_j^2}{v_i} \frac{\partial^2 v_i}{\partial \theta_\alpha \partial \theta_\beta} + \sum_{ij} \frac{1}{v_i} \left(1 + \frac{1}{2} \frac{\partial^2 v_i}{\partial \theta_j^2} \frac{\sigma_j^2}{v_i}\right) \frac{\partial v_i}{\partial \theta_\alpha} \frac{\partial v_i}{\partial \theta_\beta} + \delta_\beta^\alpha \frac{1}{\sigma_\alpha^2}, & 2 \leq \alpha, \beta \leq M, \end{cases}$$

$$V_{\alpha\beta} = \sum_{ijkl} \left(\frac{1}{v_i} \delta_j^i + \frac{1}{2} \frac{\partial^2 v_i}{\partial \theta_k^2} \frac{\sigma_k^2}{v_i^2} \delta_j^i + \frac{\partial v_i}{\partial \theta_k} \frac{\partial v_j}{\partial \theta_k} \frac{\sigma_k^2}{v_i v_j} - \frac{1}{4} \frac{\partial^2 v_i}{\partial \theta_k^2} \frac{\partial^2 v_j}{\partial \theta_l^2} \frac{\sigma_k^2 \sigma_l^2}{v_i v_j}\right) \frac{\partial v_i}{\partial \theta_\alpha} \frac{\partial v_j}{\partial \theta_\beta}, \quad 1 \leq \alpha, \quad \beta \leq M, \quad (4)$$

where δ_j^i and δ_β^α are the Kronecker delta symbols. Kindly note that the aforementioned quantities must be assessed at

$\boldsymbol{\theta}$, which can be either $\boldsymbol{\theta}^0$ or $\boldsymbol{\theta}^1$. For more details, please refer to Appendix B.

After completion of the aforementioned step, we can proceed to compute the asymptotic distribution of $-2\ln\lambda$. This procedure is what we call the asymptotic-analytic method. Furthermore, we present a rather simple case here

---

[1]Generally speaking, explicit forms for a_i and ϕ_i are not readily available, and should be computed numerically by evaluating Λ and Y.

for better understanding. In this case, there is only one nuisance parameter considered to modify the background, and the expected value $v_i = \theta_1 s_i + \theta_2 b_i$. According to Eq. (4), we have:

$$\boldsymbol{\mu} = 0, \qquad \mathcal{F} = \begin{pmatrix} \sum_i \frac{s_i^2}{v_i} & \sum_i \frac{s_i b_i}{v_i} \\ \sum_i \frac{s_i b_i}{v_i} & \sum_i \frac{b_i^2}{v_i} + \frac{1}{\sigma_2^2} \end{pmatrix},$$

$$\mathbf{V} = \begin{pmatrix} \sum_{ij} (\frac{1}{v_i} + \frac{b_i b_j}{v_i v_j} \sigma_2^2) s_i^2 & \sum_{ij} (\frac{1}{v_i} + \frac{b_i b_j}{v_i v_j} \sigma_2^2) s_i b_i \\ \sum_{ij} (\frac{1}{v_i} + \frac{b_i b_j}{v_i v_j} \sigma_2^2) s_i b_i & \sum_{ij} (\frac{1}{v_i} + \frac{b_i b_j}{v_i v_j} \sigma_2^2) b_i^2 \end{pmatrix}.$$

However, it is not feasible to determine the asymptotic distribution of the statistic analytically. As demonstrated in Section III C utilizing a numerical method, we ascertain that the test statistic follows a chi-square distribution when H_0 is true. When we ask θ_1 for being positive, the test statistic should follows a distribution of $\frac{1}{2}[\delta(0) + \chi_1^2]$. If H_1 is true, $-2 \ln \lambda \sim \chi_1^2(\phi)$, where ϕ can be written as:

$$\phi = \sum_i \frac{s_i^2}{s_i + b_i} - \frac{(\sum_i \frac{s_i b_i}{s_i + b_i})^2}{\sum_i \frac{b_i^2}{s_i + b_i} + \frac{1}{\sigma_2^2}}, \qquad (5)$$

according to Eq. (B4).

C. Quasi-Asimov dataset method

The Asimov dataset [27] was introduced as a means to readily obtain the median values of the test statistic like $-2 \ln \lambda$ mentioned previously. In this scenario of hypothesis testing, the observations are precisely aligned with their expected values: $n_i = v_i$, where v_i is determined by $\boldsymbol{\theta}$, while the nuisance parameters in the normal distribution are fixed at $\boldsymbol{\theta}'$. The Asimov dataset is considered an effective approximation method in cases where the sample size is sufficiently large, and is commonly employed in some literature [14,19,25,30–33].

The most time-consuming aspect of implementing Asimov dataset is the numerical search for the MLE $\boldsymbol{\theta}^*$ for H_0. However, Eq. (A4) provides a simple yet effective approach to approximate:

$$\boldsymbol{\theta}^* \approx \boldsymbol{\theta}' - [1 - H\mathcal{F}]\boldsymbol{\delta}. \qquad (6)$$

Hence, we can render the acquisition of the test statistic's median feasible without incurring significant time expense. Note that $\boldsymbol{\theta}'$ is consistently set to $\boldsymbol{\theta}^1$, as the requirement for Quasi-Asimov dataset method is to attain the statistic's median assuming the alternative hypothesis is genuine. We call this procedure Quasi-Asimov dataset method.

However, an issue remains with both Asimov dataset and quasi-Asimov dataset method. In the signal discovery scenario, Asimov dataset fails when extra nuisance

parameters only involve s_i. As $\theta_1^0 = 0$, the extra parameters are constrained to their true values, resulting in a result that is unaltered by their presence. For example, suppose a new nuisance parameter, θ_{k+1}, is introduced, with a mean of 1 and standard deviation of σ_{k+1}, and $\theta_{k+1}^0 = \theta_{k+1}^1 = 1$. Then $\theta_1^0 = 0$ and the Gaussian term with σ_{k+1} compels $\theta_{k+1}^* = 1$. Consequently, the median of the test statistic remains unchanged regardless of the inclusion of θ_{k+1}, which is not our expectation. Fortunately, the asymptotic-analytic method can handle this issue.

III. NEUTRINO FLOOR AND FOG

A. Recoil spectrum

In this section, we shall commence by delving into the response of WIMP and neutrinos in the detector as inputs for the neutrino floor. For the sake of simplicity, only the spin-independent WIMP-nucleon interaction is considered here. The differential event rates of WIMP [36] in the detector can be expressed as follows:

$$\frac{dR_{\text{WIMP}}}{dE_r} = \frac{\rho_0 A^2}{2m_\chi \mu_N^2} \sigma_0 F^2(E_r) \int_{v_{\min}(E_r)}^{v_{\text{esc}}} \frac{f(v, v_0)}{v} d^3v, \quad (7)$$

where E_r is the recoil energy, m_χ is the mass of WIMP, σ_0 is the spin-independent WIMP-nucleon cross section, μ_N is the WIMP-nucleon reduced mass, A is the atom number of the target nucleus, and $F(E_r)$ is the nuclear form factor generally presented by the Helm form [37]. The other parameters are taken from the standard halo model (SHM) [38,39]: the WIMP density surrounding the Earth $\rho_\chi = 0.3$ GeV/cm^3, the circular velocity of the local standard of rest (LSR) $v_0 = 220$ km/s, the escape velocity of the Milky Way $v_{\text{esc}} = 544$ km/s, and $f(v, v_0)$ is the velocity distribution of WIMP. The quantity $v_{\min}(E_r)$ represents the minimum speed at which a WIMP may cause the recoil energy E_r, as limited by the kinematics.

Integrating the differential CEνNS cross section multiplied by the neutrino flux, we can derive the recoil spectrum for CEνNS:

$$\frac{dR_\nu}{dE_r} = \frac{1}{m_N} \sum_i \int_{E_\nu^{\min}(E_r)} \frac{d\Phi_i}{dE_\nu} \frac{d\sigma_{\nu N}(E_\nu, E_r)}{dE_r} dE_\nu, \quad (8)$$

where m_N is the nuclear mass, $E_\nu^{\min}(E_r)$ is the minimum neutrino energy to generate E_r, $\frac{d\Phi_i}{dE_\nu}$ is the neutrino flux from the source labeled as i, and $\frac{d\sigma_{\nu N}(E_\nu, E_r)}{dE_r}$ is the differential CEνNS cross section that can be well described by the standard model. There are numerous neutrino sources causing recoil events in DM detectors and we adopt the same neutrino flux model as in Table I of Ref. [19]. Here, we briefly introduce some of the main neutrino sources.

Solar neutrinos are the principal source of CEνNS events, which pose an obstacle to the search for the

$\mathcal{O}(10)$ GeV-scale WIMP. These neutrinos stem from nuclear fusion reactions such as the pp Chains and the CNO Cycle with energies less than about 13 MeV. They have been meticulously comprehended with the standard solar mode [40]. As the standard solar model is subject to various observational constraints, the parameters governing the model become increasingly precise. This precision allows for the prediction of uncertainties in solar neutrino fluxes. Experimental measurements of ^{8}B neutrinos have achieved a high level of precision, resulting in a minimum uncertainty of 2% over solar neutrinos. Atmospheric neutrinos produced by cosmic rays interacting with the Earth atmosphere are more energetic but more rare. The atmospheric neutrino flux can be computed through simulations, and the recommended theoretical uncertainty is set at 20%. The diffuse supernova neutrino background (DSNB) originates from the cosmological history of core-collapse supernovae in the visible universe with energies roughly 10–25 MeV at Earth. Due to our incomplete understanding of the DSNB, we set the uncertainty associated with it at 50%.

B. Statistic method and results

The statistic method about the discovery of a positive signal [27] aligns with our discourse on the neutrino floor. As outlined in the Sec. II, the null hypothesis H_0 represents the neutrino background-only model, while the alternative hypothesis H_1 represents the WIMP + neutrino model. The likelihood ratio is given by:

$$\lambda(\theta_1) = \frac{L(\theta_1^0, \hat{\hat{\theta}}_2, \ldots, \hat{\hat{\theta}}_M)}{L(\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_M)}, \tag{9}$$

where $\theta_1^0 = 0$ represents the signal strength in the background-only model, $\theta_i (i \geq 2)$ is the neutrino flux normalization, $\hat{\theta}_i$ and $\hat{\hat{\theta}}_i$ are the component of the MLE θ^* and $\hat{\theta}$, respectively, and M is the number of neutrino sources. The binned likelihood has been shown as Eq. (3), in which $s_i = s_i(\sigma_0)$ and $b_i = \sum_j b_i^j \theta_j (j \geq 2)$ should be the expected WIMP events and neutrino events in the ith bin, respectively. Here, b_i^j represents for the expected neutrino events from the source labeled as i. More information on the uncertainties associated with the neutrino flux can be found in Table I of Ref. [19]. Then the test statistic can be formulated as follows:

$$q_0 = \begin{cases} -2 \ln \lambda(0) & \hat{\theta}_1 \geq 0, \\ 0 & \hat{\theta}_1 < 0. \end{cases} \tag{10}$$

Upon closer examination, it is discernible that the test statistic, denoted by q_0, measures the discrepancy between two hypotheses. This discrepancy is shown to increase when the actual observations deviate from the

background-only model. According to the definition of discovery limits, the 3σ significance of signals corresponds to a p-value $p_0 = 0.0027$. The p-value p_0 is defined by $p_0 = \int_{q_{\rm obs}}^{\infty} f(q_0|H_0) dq_{\rm obs}$, where $f(q_0|H_0)$ is the distribution of the statistic under the assumption of H_0 being true, and is believed to follow a distribution of [27,28,41]. Thus, the 3σ significance corresponds to $q_{\rm obs} = 9$.[2] We also demonstrate it by utilizing the asymptotic-analytic method, and the numerical analysis reveals the existence of only one nonzero diagonal entry within Λ with a value of 1. Therefore, the test statistic $q_0 \approx (\sum_i a_i Y_i)^2$, where $\sum_i a_i^2 = 1$ and $Y_i \sim \mathcal{N}(0, 1)$, indicating that q_0 should follow the chi-square distribution. However, by incorporating the positive condition for the signal strength, the distribution of q_0 should be $\frac{1}{2}[\delta(0) + \chi_1^2]$. The detailed reason of such alteration can be found in Appendix B.

In defining the discovery limits, $P\%$ CL is often requested, with previous literature using $P\% = 90\%$ [7,8] and recent studies opting for $P\% = 50\%$ [14,19]. The $P\%$ CL indicates the percentage of experiments in which the discovery is significant, i.e., $\int_{q_{\rm obs}}^{\infty} f(q_0|H_1) dq_{\rm obs} = P\%$. To obtain the distribution of q_0 when H_1 is real, one typically needs to generate numerous pseudodata from MC simulations and perform time-consuming computations to find the MLE. In high-statistics analyses, however, Asimov dataset [27] provides a time-saving way to compute the median of q_0 when H_1 is real. Thus, a slight modification of the CL value, leading to a minor quantitative difference [19]. Additionally, the asymptotic-analytic method reveals that q_0 follows a noncentral chi-square distribution, denoted as $\chi_1^2(\phi)$, with one DOF and a noncentral parameter ϕ. Note that we can disregard the positive condition for θ_1, because most of θ_1^* and $\hat{\theta}_1$ are positive in our case where $q_{\rm obs} = 9$ and $\int_{q_{\rm obs}}^{\infty} f(q_0|H_1) dq_{\rm obs} = P\%$. Surprisingly, according to Eqs. (B1) and (B4), the asymptotic-analytic method also afford the analytic form of ϕ:

$$\phi = \sum_i \frac{s_i^2}{v_i} - \sum_{jk} \left[\sum_i \frac{s_i b_i^j}{v_i} \right]_j \left[\sum_i \frac{b_i^j b_i^k}{v_i} + \frac{\delta_k^j}{\sigma_k^2} \right]_{jk}^{-1}$$
$$\times \left[\sum_i \frac{s_i b_i^j}{v_i} \right]_k \Bigg|_{\boldsymbol{\theta}=1}, \tag{11}$$

where the quantities enclosed by the square bracket $[\cdots]$ with one and two subscripts represent vectors and matrices, respectively. Besides, $[\cdots]^{-1}$ represents the inverse of a matrix, and δ_k^j is the Kronecker delta symbol.

To demonstrate the effectiveness of our new methods, we reproduce the neutrino floor defined by discovery limits and make a comparison with the result from APPEC report [42] based on MC method. Note that the neutrino floor we

---

[2]The significance $Z = \Phi^{-1}(1 - p_0)$, where Φ^{-1} is the quantile of the standard Gaussian [27].

discuss here relies on the choice on the detect threshold and the exposure. More details can be found in Appendix C. Consequently, the asymptotic-analytic and quasi-Asimov dataset methods are effective, except that sometimes they fails on account of insufficient statistics. Moreover, the asymptotic-analytic technique furnishes a means to assess the influences attributable to individual parameters. As an illustration, the neutrino floor at $m_\chi = 5.5$ GeV, which will be elaborated upon later, is primarily dictated by the ^{8}B neutrino flux.

To avoid the setting of somewhat arbitrary experimental configurations and give the neutrino floor a single consistent interpretation in statistics, a new definition [25] on the neutrino floor has been proposed and widely recognized. Since the neutrino floor relies on the experimental configurations, it is better appreciated that neutrinos should present a "fog": a region of the parameter space where a clear distinction between signal and background is challenging. Define n as the gradient of a discovery limit cross section σ_{DL} with respect to the exposure N: $n = -(d \ln \sigma_{DL}/d \ln N)^{-1}$, which is also called the "opacity" of the neutrino fog. The index n alters as N increases. Briefly, when N is small, i.e., the background-free case, $n = 1$. Then, as N increases, the case becomes the Poissonian background subtraction: $n = 2$. If we have a larger N, WIMP signals are covered by fluctuations of neutrino backgrounds, so that $n > 2$. That means it is difficult to detect the WIMP signal when we only increase the exposure. Therefore, it is more effective to find another way, like the directional detection strategy, to search for WIMPs. Eventually, as N gets large enough, the intrinsic difference between signals and backgrounds help us discover the signal, and $n = 2$ starts to return. In the next section, this evolution of n can be explained analytically with the asymptotic-analytic method. The neutrino floor is defined as the boundary of the neutrino fog, which marks the transition from statistical to systematical limits.

As illustrated in Fig. 1, we employ the quasi-Asimov dataset method to showcase the neutrino floor and fog. The outcomes from the quasi-Asimov dataset method perfectly reproduce the neutrino floor and fog [25,26], albeit with minor discrepancies. In terms of computational efficiency, our methodology outperforms the Asimov dataset, taking only about 10 seconds to calculate the neutrino fog on the same computer, as opposed to the half-hour computation required by the Asimov dataset. Additionally, the quasi-Asimov dataset method exhibits superior computational stability. More details and figures can be found in our public code [43]. Note that we follow the calculation techniques outlined in Ref. [25]. Thus, events are binned for the recoil spectra in the logarithmic scale between 10^{-4} keV and 200 keV. We choose the former detector threshold for the purpose of mapping the neutrino floor down to $m_\chi = 0.1$ GeV, which actually does not impact the height of the limit at other masses.
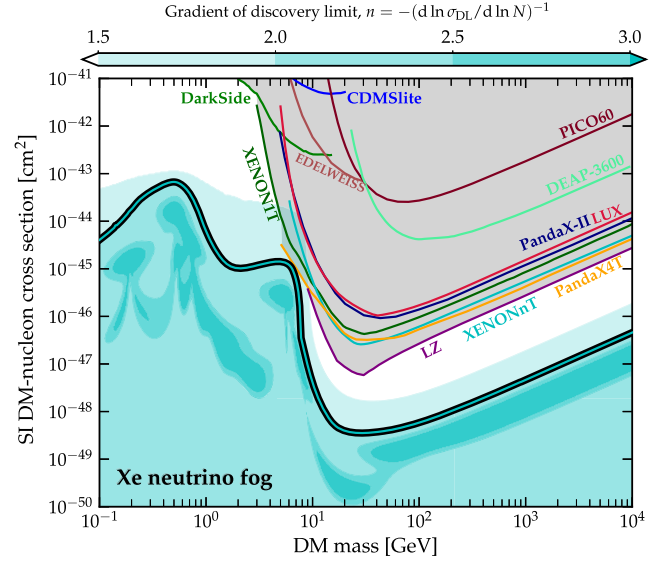

FIG. 1. The neutrino fog for the xenon experiment is presented in the cyan contour map, while the solid line represents the neutrino floor for $n = 2$. Besides, the latest excluded space is shaded gray [44–55] (all experimental limits except the latest exlusion limits from XENONnT and PandaX4T are taken from Ref. [25]).

Moreover, we also present neutrino fogs for six different targets in the DM direct detection experiments. As shown in Fig. 2, neutrino fogs for Xenon, Germanium and Argon targets are similar, while the neutrino fog for the Helium target is quite different and the neutrino floor for the high WIMP mass emerges in the excluded space, since its mass is too low to gain the sensitivity for WIMP with higher mass. Neutrino fogs for composite targets also resemble that derived from Xenon, Germanium and Argon targets, despite of some small differences. For instance, the region around $m_\chi \approx 20$ GeV for the NaI target is more shallow, which indicates that experiments using the NaI target can detect a weaker signal generated by a WIMP of about 20 GeV.

C. Analytical interpretation on sensitivity curves

When the sample size or exposure is sufficiently large, the asymptotic-analytic and quasi-Asimov dataset methods produce consistent results with each other. Therefore, the evolution of the discovery limit cross section σ_{DL} with respect to N can be analytically explained. For the sake of simplicity, we restrict our attention to the ^{8}B neutrino source and $m_\chi = 5.5$ GeV. As discussed before, the noncentral parameter can be obtained from Eq. (5), where the expected WIMP events are proportional to σ_{DL} and N: $s_i \propto \sigma_{DL} N$, the expected background events also have a linear correlation with N: $b_i \propto N$, and the uncertainty $\sigma_2 = \sigma_\nu$, which stands for the uncertainty of the ^{8}B neutrino flux.
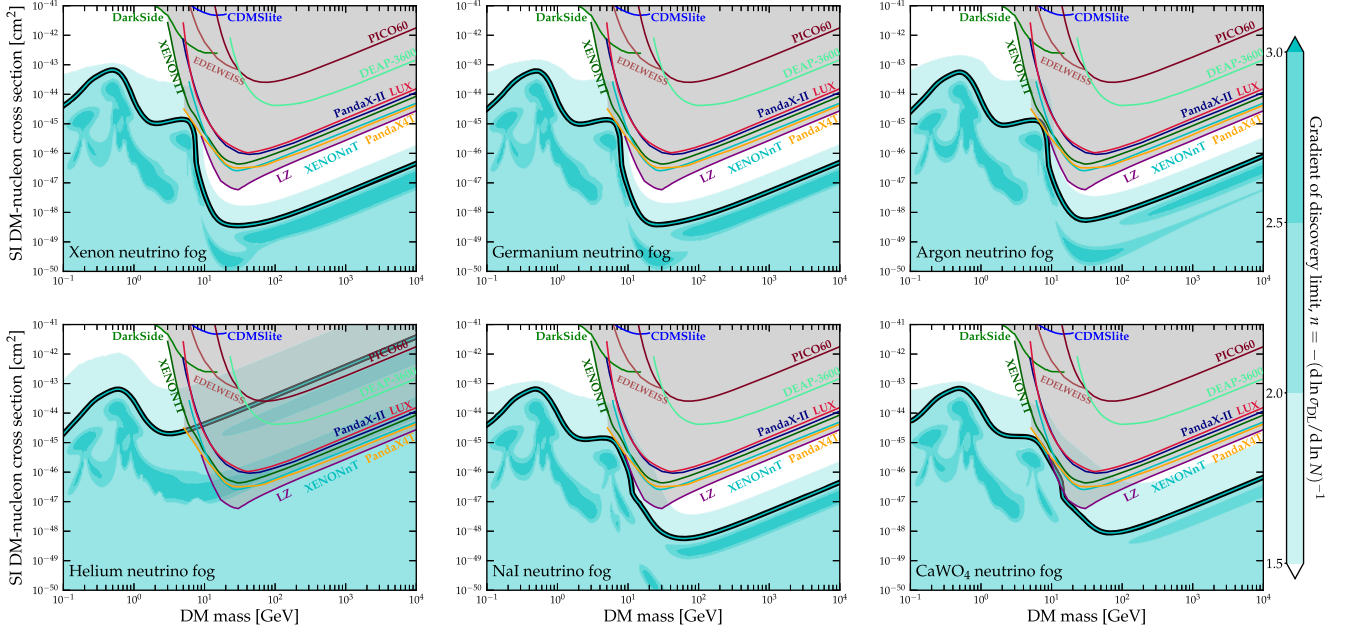
FIG. 2. The neutrino fogs for six different targets in the direct detection experiments are presented in the cyan contour map, while the solid line represents the neutrino floor for $n = 2$. Besides, the latest excluded space is given by the shaded gray region.

Initially, $N < 1$, and WIMP signals dominate. Because $s_i \gg b_i$ and $\frac{1}{\sigma_\nu^2}$ dominates in the denominator of the second term in Eq. (5), the noncentral parameter $\phi \approx \sum_i \frac{s_i^2}{s_i + b_i} \approx \mathcal{O}(s)$. Thus, to maintain the invariance of
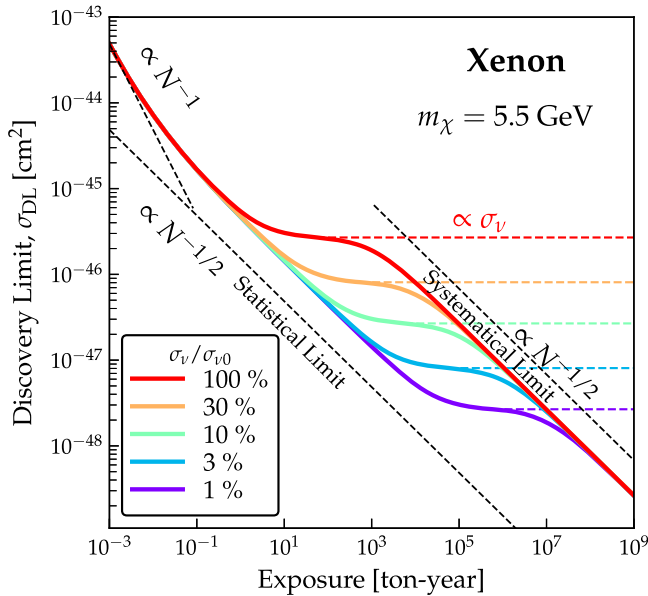

FIG. 3. The sensitivity curves about the evolution of the discovery limit cross section with respect to an exposure for $m_\chi = 5.5$ GeV. Lines in different colors represent the sensitivity curve with different uncertainties of the ^8B neutrino flux. Here we assume the standard uncertainty of the ^8B neutrino flux $\sigma_{\nu 0} = 2\%$.

ϕ, $\sigma_{DL} \propto N^{-1}$. As N increases and σ_{DL} decreases, $s_i < b_i$ while the second term is still controlled by $\frac{1}{\sigma_\nu^2}$. Consequently, we obtain $\phi \approx \sum_i \frac{s_i^2}{s_i + b_i} \approx \mathcal{O}(s^2/b)$, and $\sigma_{DL} \propto N^{-\frac{1}{2}}$. Here we come to the Poissonian background subtraction and the statistical limit as illustrated in Fig. 3.

As N increases further, the signal is lost within the background: $s_i \ll b_i$. Then we have:

$$\phi \approx \sum_i \frac{s_i^2}{b_i} - \frac{(\sum_i s_i)^2}{\sum_i b_i}\left(1 - \frac{1}{\sigma_\nu^2 \sum_i b_i}\right). \quad (12)$$

If there is only one bin for the data or the background mimics the signal: $s_i \propto b_i$, the leading term is the one with σ_ν and $\phi \approx \frac{(\sum_i s_i)^2}{\sigma_\nu^2 (\sum_i b_i)^2} \approx \mathcal{O}(s^2/b^2)$. Hence ϕ remains constant as N increases: $\sigma_{DL} \propto N^0$, which causes the flatness in Fig. 3. Moreover, since $\phi \approx 1/\sigma_\nu^2$ in this case, a smaller value of the uncertainty σ_ν corresponds to a lower σ_{DL}. Thus, $\sigma_{DL} \propto \sigma_\nu$ for the flatness of the curves in Fig. 3. On the other hand, if there are some slight differences between the signal and background even though they are very similar, the term with σ_ν can be neglected and we have $\phi \sim \mathcal{O}(s^2/b)$. Finally, we return to $\sigma_{DL} \propto N^{-\frac{1}{2}}$ and reach the systematical limit as shown in Fig. 3.

IV. ADDITIONAL NUISANCE PARAMETERS

A. Weak mixing angle uncertainty

The weak mixing angle has been well measured at the Z-pole, while it leaves a sizable uncertainties at the low energy (the maximal momentum transfer $q_{max} \lesssim 200$ MeV).

It might be determined by CEνNS process. Conversely, the weak mixing angle also exerts an influence on the response of neutrinos in the DM detector, so we shall explore its impact on the neutrino fog in this section. The weak mixing angle is rather flat over the low energy range, so the neutrino background $b_i \propto Q_W^2$, where $Q_W = N - (1 - 4\sin^2\theta_W)Z$ is the weak charge [14,56,57]. Note that N and Z are the number of neutrons and protons in the target nuclei. As in Ref. [14], we take the central values $\sin^2\theta_W = 0.2387$ with a 10% uncertainty. For simplicity, we assign θ_{M+1} to account for the uncertainty: $Q_W = N - (1 - 4\theta_{M+1}\sin^2\theta_W)Z$, where M represents the number of nuisance parameters from neutrino fluxes. In this case, $v_i(\boldsymbol{\theta}) = \theta_1 s_i + \sum_{2 \leq j \leq M} \theta_j b_i^j(\theta_{M+1})$. As the discussion in Sec. II B, the background-only hypothesis H_0 corresponds to $\boldsymbol{\theta}^0$: $\theta_1^0 = 0$, $\theta_j^0 = 1(2 \leq j \leq M + 1)$, while H_1 corresponds to $\boldsymbol{\theta}^1$: $\theta_j^1 = 1(1 \leq j \leq M + 1)$.

Observed that $\dot{l}(\boldsymbol{\theta}')_{M+1} = \frac{8\sin^2\theta_W Z}{Q_w}\sum_j \dot{l}(\boldsymbol{\theta}')_j (2 \leq j \leq M)$ by making use of $\frac{\partial v_i}{\partial\theta_{M+1}} = \frac{8\sin^2\theta_W Z}{Q_w}\sum_j b_i^j$, there are only M independent components in the normal variate $\mathbf{Z}$. Thus, $\mathbf{V}$

has a dimension of $M \times M$, while $\mathcal{F}$ has a larger dimension of $(M + 1) \times (M + 1)$, as shown in Eq. (14). To solve this problem, we employ a matrix $\mathbf{O}$ to eliminate $\dot{l}(\boldsymbol{\theta}')_{M+1}$, where

$$\mathbf{O} = \left(\begin{array}{cccc} & \mathbf{I}_{M\times M} & & \\ 0, & \underbrace{\dfrac{8\sin^2\theta_W Z}{Q_w}, & ...., & \dfrac{8\sin^2\theta_W Z}{Q_w}}_{M} \end{array} \right),$$

where $\mathbf{I}_{M\times M}$ is a diagonal matrix with a dimension of $M \times M$. Since $\boldsymbol{\delta} = 0$ is H_0 is true and $\boldsymbol{\delta} = \{1, \overbrace{0, ..., 0}^{M}\}$ if H_1 is true, we can use $\mathbf{O}$ to obtain Z with the correct dimension: $Z \sim \mathcal{N}(\mathbf{V}^{-\frac{1}{2}}\mathbf{O}(\boldsymbol{\mu} + \mathcal{F}\boldsymbol{\delta}), I)$. Therefore, Eq. (1) is rewritten as:

$$-2\ln\lambda \approx \mathbf{Z}^T\mathbf{V}^{\frac{1}{2}}\mathbf{O}^T[\mathcal{F}^{-1} - \mathbf{H}]\mathbf{O}\mathbf{V}^{\frac{1}{2}}\mathbf{Z}. \quad (13)$$

According to Eq. (4) we have:

$$\mu_\alpha = \sum_i \frac{1}{2}\frac{\partial^2 v_i}{\partial\theta_{M+1}^2}\frac{\sigma_2^2}{v_i}\frac{\partial v_i}{\partial\theta_\alpha}, \qquad 1 \leq \alpha \leq M + 1,$$

$$\mathcal{F}_{\alpha\beta} = \begin{cases} -\sum_i \frac{1}{2}\frac{\partial^2 v_i}{\partial\theta_{M+1}^2}\frac{\sigma_{M+1}^2}{v_i}\frac{\partial^2 v_i}{\partial\theta_\alpha\partial\theta_\beta} + \sum_i \frac{1}{v_i}\left(1 + \frac{1}{2}\frac{\partial^2 v_i}{\partial\theta_{M+1}^2}\frac{\sigma_{M+1}^2}{v_i}\right)\frac{\partial v_i}{\partial\theta_\alpha}\frac{\partial v_i}{\partial\theta_\beta}, & \alpha = 1 \text{ or } \beta = 1, \\ -\sum_i \frac{1}{2}\frac{\partial^2 v_i}{\partial\theta_{M+1}^2}\frac{\sigma_{M+1}^2}{v_i}\frac{\partial^2 v_i}{\partial\theta_\alpha\partial\theta_\beta} + \sum_i \frac{1}{v_i}\left(1 + \frac{1}{2}\frac{\partial^2 v_i}{\partial\theta_{M+1}^2}\frac{\sigma_{M+1}^2}{v_i}\right)\frac{\partial v_i}{\partial\theta_\alpha}\frac{\partial v_i}{\partial\theta_\beta} + \delta_\beta^\alpha\frac{1}{\sigma_\alpha^2}, & 2 \leq \alpha, \beta \leq M + 1, \end{cases}$$

$$V_{\alpha\beta} = \sum_{ijk}\left(\frac{1}{v_i}\delta_j^i + \frac{1}{2}\frac{\partial^2 v_i}{\partial\theta_{M+1}^2}\frac{\sigma_{M+1}^2}{v_i^2}\delta_j^i + \frac{\partial v_i}{\partial\theta_k}\frac{\partial v_j}{\partial\theta_k}\frac{\sigma_k^2}{v_i v_j} - \frac{1}{4}\frac{\partial^2 v_i}{\partial\theta_{M+1}^2}\frac{\partial^2 v_j}{\partial\theta_{M+1}^2}\frac{\sigma_{M+1}^4}{v_i v_j}\right)\frac{\partial v_i}{\partial\theta_\alpha}\frac{\partial v_j}{\partial\theta_\beta}, \qquad 1 \leq \alpha, \quad \beta \leq M, \quad (14)$$

where the only nonvanishing second derivative $\frac{\partial^2 v_i}{\partial\theta_{M+1}^2}$ is considered, instead of summing over all the nuisance parameters given in Eq. (4). This case is more complicate than the case where only the uncertainties from neutrino fluxes are considered.

Through the numerical computations discussed in Sec. II B, the test statistic still follows the asymptotic $\frac{1}{2}[\delta(0) + \chi_1^2]$ distribution if H_0 is real, while the distribution of q_0 is distorted by the variation of $\sin^2\theta_W$ if H_1 is real. When H_1 is real, the asymptotic-analytic method reveals that there is only one nonzero diagonal element in Λ with its value at 1, which means that there is only one χ^2 variate, and the coefficient is not unity. Thus, the distribution of q_0 can be expressed as $q_0 \sim a\chi_1^2(\phi)$. In Appendix C, Monte Carlo realizations and the asymptotic-analytic method for some benchmark points are shown. It should be noted that although there may be slight deviations between our results and the Monte Carlo realizations, our method remains effective.

Similar to the scenario we present in Sec. III, quasi-Asimov dataset method is utilized to obtain the discovery limit cross section. As shown in the left panel of Fig. 4, the uncertainty of $\sin^2\theta_W$ significantly affect the discovery limit. It is evident that a greater uncertainty leads to a larger σ_{DL} at the same exposure. However, as the exposure increases, these curves converge to the same systematic limit. Note that we only consider the ^{8}B neutrino as depicted in Fig. 3, for the purpose of comparison. Moving to the right panel of Fig. 4, the neutrino fog considering the uncertainty of $\sin^2\theta_W$ is presented. One can see that the uncertainty of $\sin^2\theta_W$ remarkably modifies the neutrino fog shown in Fig. 1 over the low mass range ($m_\chi \lesssim 1$ GeV), while the region over the higher mass range remains unchanged. This fact confirms the result in Ref. [14]. The neutrino floor is elevated when compared to the neutrino floor without considering the uncertainty of $\sin^2\theta_W$, thereby demonstrating the consistency of the new definition of the neutrino floor.
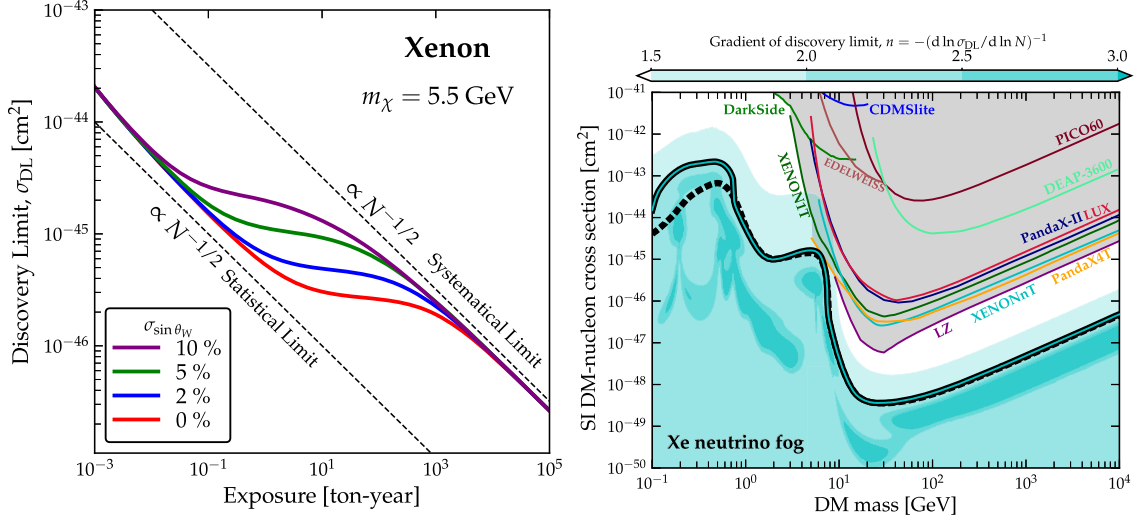
FIG. 4. The sensitivity curves about the evolution of $\sigma_{\rm DL}$ with respect to an exposure for $m_\chi = 5.5$ GeV, involving the uncertainty of $\sin^2 \theta_W$ (left). Lines in different colors represent the sensitivity curve with different uncertainties of $\sin^2 \theta_W$. The uncertainty of the ^{8}B neutrino flux is set to 2%. The neutrino fog for the xenon experiment is presented in the cyan contour map, involving the uncertainty of $\sin^2 \theta_W$ (right). The solid line represents the neutrino floor for $n = 2$, while the dashed black line represents the neutrino floor without considering the uncertainty of $\sin^2 \theta_W$.

B. Astrophysical uncertainty

As elaborated in Sec. II, a quandary regarding the Asimov dataset still persists. In the context of the signal discovery, when some extra nuisance parameters are only implicated in s_i, the zero signal strength $\theta_1^0 = 0$ forces the MLEs for $L(\theta_1^0, \hat{\hat{\theta}}_2, ..., \hat{\hat{\theta}}_M, \hat{\hat{\theta}}_{M+1})$ to be indistinguishable from the scenario where the extraneous parameters are absent. Here, an extra nuisance parameter θ_{M+1} is only involved in the signal, i.e., $s_i = \theta_1 s_i(\theta_{M+1})$. This outcome is contrary to our expectations, and fortunately, our asymptotic-analytic method is adept at resolving this predicament.

For the sake of convenience, we shall confine our attention to the velocity of the local standard of rest (LSR) v_0 and its uncertainty as the extra nuisance parameter. According to the previous investigation [10], we simply surmise that v_0 is subject to a normal distribution, with a mean of 220 km/s and a standard deviation of 50 km/s. Analogous to our handling of neutrino fluxes, the velocity is scaled to unity, and its standard deviation is 21.2%. Denoted by θ_{M+1} as in Sec. IV A, this nuisance parameter is distinguished from θ_j $(2 \leq j \leq M)$, which pertains to the neutrino flux. In this case, $v_i(\boldsymbol{\theta}) = \theta_1 s_i(\theta_{M+1}) + \sum_{2 \leq j \leq M} \theta_j b_i^j$ no longer suffices the simple linear form as θ_{M+1} modifies the shape of WIMP spectrum. Consequently, we must reexamine Eqs. (1) and (4) to elicit the asymptotic distribution of the test statistic. In this case, $\mathbf{V}$ and $\mathcal{F}$ has the same dimension of $(M + 1) \times (M + 1)$, which can be directly computed by Eq. (14) where the dimension of $\mathbf{V}$ should be changed into $(M+1) \times (M+1)$. As the discussion in Sec. II B, the background-only

hypothesis H_0 corresponds to $\boldsymbol{\theta}^0$: $\theta_1^0 = 0$, $\theta_j^0 = 1 (2 \leq j \leq M + 1)$, while H_1 corresponds to $\boldsymbol{\theta}^1$: $\theta_j^1 = 1 (1 \leq j \leq M + 1)$. Comparatively, utilization of the Asimov dataset reveals that the discovery limits, with or without consideration of v_0, coincide, a result which is at odds with the findings in the existing literature [10].

Instead of utilizing the laborious Monte Carlo realizations, our asymptotic-analytic method is available to address this case. Note that we adhere to the 90% CL for the discovery limit to illustrate the feature of our method. In this instance, as $\hat{\hat{\theta}}_{M+1}$ has been firmly established to 1, we are in fact in pursuit of the MLE for $L(0, \hat{\hat{\theta}}_2, ..., \hat{\hat{\theta}}_M, 1)$ which corresponds to the case of the two parameters of interest. When H_0 is real, the computational outcome manifests that the test statistic still follows the asymptotic $\frac{1}{2}[\delta(0) + \chi_1^2]$ distribution. While for H_1 is real, the test statistic is asymptotic distributed to the $a_1 \chi_1^2(\phi_1) + a_2 \chi_1^2(\phi_2)$ distribution, where a_i and ϕ_i $(i = 1, 2)$ can be obtained from the numerical solution as presented in Sec. II B. Numerically, the asymptotic distribution of q_0 can be obtained by performing the inverse Fourier transformation on its characteristic function. To show our method's effectiveness, we compare the test statistic's distribution from Monte Carlo realization and the asymptotic-analytic method for some benchmark points, which can be found in Appendix C.

Nevertheless, it is worth noting that for extremely large exposures, the numerical solution becomes unreliable, since the matrices $\mathbf{V}$ and $\mathcal{F}$ are no longer positive definite. Consequently, it is not possible to accurately present the neutrino fog considering the uncertainty of θ_2, while
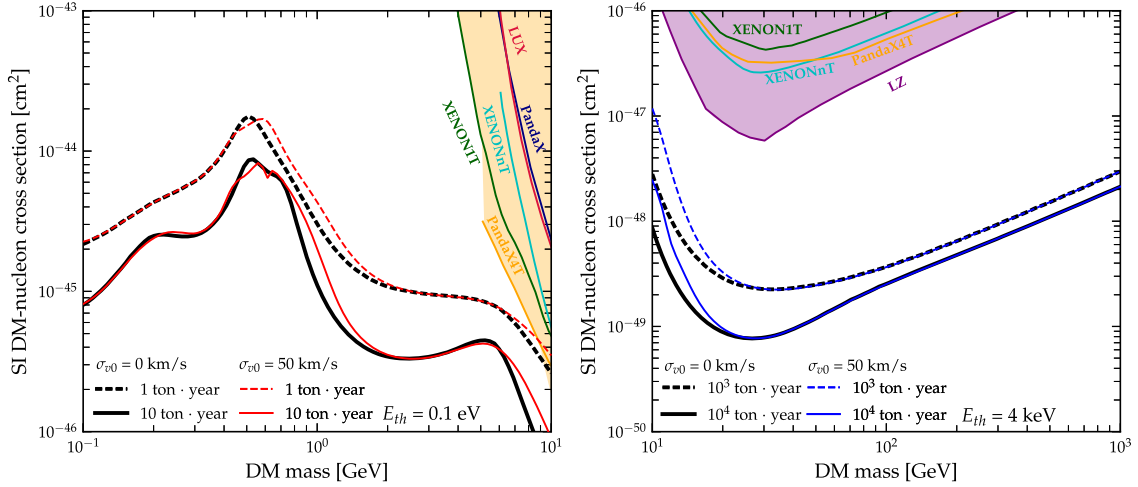
FIG. 5. Discovery limits with and without considering the extra parameter from the velocity of LSR, which are displayed in colored and black lines. The left panel corresponds to the case where the threshold equals to 0.1 eV with two different exposures: 1 ton × years and 10 ton × years. The right panel corresponds to the case where the threshold is 4 keV with two different exposures: 10^3 ton × years and 10^4 ton × years. The solid and dashed lines stand for lower and higher exposures, respectively. The latest excluded limits are also shown.

discovery limits for two benchmark scenarios are presented instead. Note that the covariance matrix **V** should be always positive definite, and a positive definite $\mathcal{F}$ guarantees the stationary point is the minimum. As illustrated in Fig. 5, we implement the asymptotic-analytic method to calculate the discovery limit that pertains to the nuisance parameter from v_0 on the WIMP parameter space. Additionally, we have provided the outcome that involves only the nuisance parameters from neutrino fluxes for the purpose of comparison. In the low mass region, $m_\chi \lesssim 10$ GeV, we adopt the optimal threshold of 0.1 eV as previously utilized, since it facilitates the mapping of the limit down to $m_\chi = 0.1$ GeV. Besides, we choose the exposure of 1 ton × years and 10 ton × years. Generally speaking, introducing an extra nuisance parameter should raise the limit, as the region around $m_\chi \approx 1$ GeV in Fig. 5. However, certain interesting things have been observed around the 0.5 GeV and 6 GeV, where the limits involving an extra parameter become lower on the contrary. We demonstrate the feature with MC realizations which can be found in Appendix C. The similar phenomenon has been discussed in Ref. [10], while it still needs further investigation. For other regions, no significant changes on the limit are observed, while the distribution of the test statistic has been distorted. For further information, please refer to Appendix C.

For the larger mass range $m_\chi \gtrsim 10$ GeV, we adopt a more realistic threshold at 4 keV, while it needs larger exposures to gain enough statistics. So we choose the exposure of 10^3 ton × years and 10^4 ton × years so that events from atmosphere and DSNB are sufficient for CEvNS in the same nuclear recoil energy region stand out as clear signals. It can be seen from Fig. 5 that the

introduction of v_0 significantly raises up the limit around $m_\chi \approx 10$ GeV, while the limit for $m_\chi \gtrsim 30$ GeV. This fact indicates that the effect of the astrophysical uncertainties cannot be ignored especially when we analyze the neutrino floor for the WIMP mass around 10 GeV.

V. CONCLUSIONS AND OUTLOOK

In this paper, we have investigated the asymptotic behavior of the profile binned likelihood ratio test statistic, in which the likelihood is constructed from different variables and pull terms, drawing inspiration from the seminal works of Wilk and Ward [28,29]. Based on our findings, we have proposed two new methods: asymptotic-analytic method, which can provide analytical results for large statistics, handle situations involving some specific nuisance parameters and affords a way to determine the most relevant parameters in the statistical analysis; and the quasi-Asimov dataset method, which is similar to but faster than Asimov dataset. We make a comparison on the computational speed, accuracy of results and extensibility for four methods in Table. I. Our proposed methods are not only applicable for the neutrino floor and fog, but also feasible for other studies utilizing the Asimov dataset in experimental analysis and phenomenology. Moreover, the current methodology in statistics will pave the way to scrutinize the origin of tiny discrepancy in a comparison of theoretical predictions and experimental data from the DM experiments where new physics might be hidden [58].

We have employed our newly proposed methods on the neutrino floor and fog. By utilizing Quasi-Asimov dataset, we have achieved near-perfect reproduction of the neutrino

TABLE I. Comparison on the computational speed, accuracy of results and extensibility for four methods.

	Speed	Accuracy	Extensible	References
Monte Carlo simulation	~1 day	✓	✓	[7,9]
Asimov dataset	~30 minutes	✓	sometimes ×	[14,19,25]
Asymptotic-analytic method	~10 seconds	✓ for big statistics	✓	This work
Quasi-Asimov dataset	~10 seconds	✓	sometimes ×	This work

floor and fog [25,26] with a computational speed that is two orders of magnitude faster and improved stability. Likewise, we also consider the uncertainty of the weak mixing angle in the context of the neutrino fog, and verify the consistency of the new definition of the neutrino floor. On the other hand, the asymptotic-analytic method provides an analytical formula to quantitatively explain the evolution of the discovery limit cross section with exposure, and it offers a solution for cases involving astrophysical uncertainties that cannot be dealt with by Asimov dataset. Additionally, our methods are capable and effective when considering more degrees of freedom in the context of the neutrino floor, such as the detector efficiency, the resolution and new physics beyond the standard model. Furthermore, MC pseudoexperiments can be boosted by only considering the most relevant parameters, which can be obtained by the asymptotic-analytic method, from neutrino fluxes.

Our method with the decent derivation with an approximation to speed up the computation has worked very well in the context of the neutrino floor and fog, as demonstrated by the aforementioned numerical calculations. Nevertheless, there are a few drawbacks to our proposed methods that we should give warnings and address carefully. First, the asymptotic-analytic method may be out of service when the sample size is too small. However, quasi-Asimov dataset method remains effective in obtaining the median of the test statistic, as the Asimov dataset does. Second, the asymptotic-analytic method requires small uncertainties; otherwise, higher-order corrections are necessary, which can complicate matters.

APPENDIX A: THE ASYMPTOTIC FORMULA OF THE TEST STATISTIC

Assuming the log likelihood function $l(\boldsymbol{\theta})$ can be approximated as the quadratic function around $\hat{\boldsymbol{\theta}}$ or $\boldsymbol{\theta}'$ in the limit of large samples, $l(\boldsymbol{\theta}^*)$ is expanded about $\hat{\boldsymbol{\theta}}$:

$$l(\boldsymbol{\theta}^*) \approx l(\hat{\boldsymbol{\theta}}) + \dot{l}(\hat{\boldsymbol{\theta}})(\boldsymbol{\theta}^* - \hat{\boldsymbol{\theta}}) + \frac{1}{2}(\boldsymbol{\theta}_n^* - \hat{\boldsymbol{\theta}})^T \ddot{l}(\hat{\boldsymbol{\theta}})(\boldsymbol{\theta}^* - \hat{\boldsymbol{\theta}})$$
$$+ \mathcal{O}\left(\frac{1}{\sqrt{N}}\right),$$

where $E(\ddot{l}(\hat{\boldsymbol{\theta}}))$ is the expectation of the second derivative of the log likelihood function at $\hat{\boldsymbol{\theta}}$, and N represents the sample size and the last term can be safely neglected when N is large enough. As a rule of thumb, the total event number N should be $\mathcal{O}(100)$. Otherwise, the approximation here might be out of service. Assuming that $\hat{\boldsymbol{\theta}}$ is close enough to $\boldsymbol{\theta}'$ for the large sample case, we have $\ddot{l}(\hat{\boldsymbol{\theta}}) \approx \ddot{l}(\boldsymbol{\theta}')$. According to the Lyapunov CLT, $\ddot{l}(\boldsymbol{\theta}')$ comprising numerous variate can asymptotically approximate to its expectation $E(\ddot{l}(\boldsymbol{\theta}'))$, i.e., $\ddot{l}(\boldsymbol{\theta}') \approx E(\ddot{l}(\boldsymbol{\theta}'))$. Thus, with $\dot{l}(\hat{\boldsymbol{\theta}}) = 0$ from the MLE condition and defining $\mathcal{F} = -E(\ddot{l}(\boldsymbol{\theta}'))$, we have:

$$-2\ln\lambda = 2(l(\boldsymbol{\theta}^*) - l(\hat{\boldsymbol{\theta}})) \approx (\boldsymbol{\theta}^* - \hat{\boldsymbol{\theta}})^T \mathcal{F}(\boldsymbol{\theta}^* - \hat{\boldsymbol{\theta}}).$$

Thanks to the Lyapunov CLT, we can infer that the first derivative of the log likelihood function at $\boldsymbol{\theta}'$, i.e., $\dot{l}(\boldsymbol{\theta}')$ follows a multivariate normal distribution, denoted by $\mathcal{N}(\boldsymbol{\mu}, \mathbf{V})$. Here, $\boldsymbol{\mu} = E(\dot{l}(\boldsymbol{\theta}'))$ represents the expectation vector, and $\mathbf{V} = \text{var}(\dot{l}(\boldsymbol{\theta}'))$ is the variance matrix. Since $\dot{l}(\boldsymbol{\theta}') \sim \mathcal{N}(\boldsymbol{\mu}, \mathbf{V})$, we better relate $(\boldsymbol{\theta}^* - \hat{\boldsymbol{\theta}})$ with $\dot{l}(\boldsymbol{\theta}')$ to obtain the asymptotic distribution. Expand $\dot{l}(\boldsymbol{\theta}^*)$ about $\hat{\boldsymbol{\theta}}$:

$$\dot{l}(\boldsymbol{\theta}^*) \approx \dot{l}(\hat{\boldsymbol{\theta}}) + E(\ddot{l}(\boldsymbol{\theta}'))(\boldsymbol{\theta}^* - \hat{\boldsymbol{\theta}}) \approx -\mathcal{F}(\boldsymbol{\theta}^* - \hat{\boldsymbol{\theta}}). \quad \text{(A1)}$$

Besides, denote the deviation from $\boldsymbol{\theta}'$ to $\boldsymbol{\theta}^0$ as $\boldsymbol{\delta} = \boldsymbol{\theta}' - \boldsymbol{\theta}^0$, and expand $\dot{l}(\boldsymbol{\theta}^*)$ about $\boldsymbol{\theta}'$:

$$\dot{l}(\boldsymbol{\theta}^*) \approx \dot{l}(\boldsymbol{\theta}') - \mathcal{F}(\boldsymbol{\theta}^* - \boldsymbol{\theta}^0) + \mathcal{F}\boldsymbol{\delta}. \quad \text{(A2)}$$

Let

$$\mathcal{F} \equiv \begin{bmatrix} \mathbf{G}_1(r \times r) & \mathbf{G}_2(r \times (k-r)) \\ \mathbf{G}_2^T((k-r) \times r) & \mathbf{G}_3((k-r) \times (k-r)) \end{bmatrix},$$

$$\mathbf{H} \equiv \begin{bmatrix} 0 & 0 \\ 0 & \mathbf{G}_3^{-1} \end{bmatrix}.$$

where $\mathbf{G}_1, \mathbf{G}_2, \mathbf{G}_3$ are the block matrices inside $\mathcal{F}$, and r, k represent their dimensions. Because the last k-r components of $\dot{l}(\boldsymbol{\theta}^*)$ and the first r components of $(\boldsymbol{\theta}^* - \boldsymbol{\theta}^0)$ are zero, multiplying $\mathbf{H}$ on the left side of Eq. (A2) we have $\mathbf{H}\dot{l}(\boldsymbol{\theta}^*) = 0$ and then:

$$0 \approx \mathbf{H}\dot{l}(\boldsymbol{\theta}') - (\boldsymbol{\theta}^* - \boldsymbol{\theta}^0) + \mathbf{H}\mathcal{F}\boldsymbol{\delta}. \tag{A3}$$

Substitute it into Eq. (A2) and return to Eq. (A1), we have:

$$\boldsymbol{\theta}^* - \hat{\boldsymbol{\theta}} \approx -\mathcal{F}^{-1}\dot{l}(\boldsymbol{\theta}^*) \approx -(\mathcal{F}^{-1} - \mathbf{H})(\dot{l}(\boldsymbol{\theta}') + \mathcal{F}\boldsymbol{\delta}). \tag{A4}$$

Thus, with $\mathbf{H}\mathcal{F}^{-1}\mathbf{H} = \mathbf{H}$, we finally obtain:

$$-2\ln\lambda \approx (\dot{l}(\boldsymbol{\theta}') + \mathcal{F}\boldsymbol{\delta})^T[\mathcal{F}^{-1} - \mathbf{H}](\dot{l}(\boldsymbol{\theta}') + \mathcal{F}\boldsymbol{\delta}). \tag{A5}$$

For convenience, let $Z = \mathbf{V}^{-\frac{1}{2}}(\dot{l}(\boldsymbol{\theta}') + \mathcal{F}\boldsymbol{\delta}) \sim \mathcal{N}(\mathbf{V}^{-\frac{1}{2}}(\boldsymbol{\mu} + \mathcal{F}\boldsymbol{\delta}), I)$, we obtain Eq. (1):

$$-2\ln\lambda \approx Z^T\mathbf{V}^{\frac{1}{2}}[\mathcal{F}^{-1} - \mathbf{H}]\mathbf{V}^{\frac{1}{2}}Z.$$

APPENDIX B: EVALUATING QUANTITIES IN THE ASYMPTOTIC-ANALYTIC METHOD WITH AN EXAMPLE

From Eq. (3), we obtain the log likelihood and its derivations:

$$l(\boldsymbol{\theta}) = -\sum_{i=1}^{N}[v_i - n_i \ln v_i] - \sum_{j=1}^{M}\frac{(\theta_j - \theta_j')^2}{2\sigma_j^2} + \text{const.},$$

$$\dot{l}(\boldsymbol{\theta})_\alpha = \begin{cases} \sum_{i=1}^{N}[\frac{n_i}{v_i} - 1]\frac{\partial v_i}{\partial \theta_\alpha}, & \alpha = 1, \\ \sum_{i=1}^{N}[\frac{n_i}{v_i} - 1]\frac{\partial v_i}{\partial \theta_\alpha} - \frac{\theta_\alpha - \theta_\alpha'}{\sigma_\alpha^2}, & 2 \leq \alpha \leq M, \end{cases}$$

$$\ddot{l}(\boldsymbol{\theta})_{\alpha\beta} = \begin{cases} \sum_{i=1}^{N}[\frac{n_i}{v_i} - 1]\frac{\partial^2 v_i}{\partial \theta_\alpha \partial \theta_\beta} - \sum_{i=1}^{N}\frac{n_i}{v_i^2}\frac{\partial v_i}{\partial \theta_\alpha}\frac{\partial v_i}{\partial \theta_\beta}, & \alpha = 1 \text{ or } \beta = 1, \\ \sum_{i=1}^{N}[\frac{n_i}{v_i} - 1]\frac{\partial^2 v_i}{\partial \theta_\alpha \partial \theta_\beta} - \sum_{i=1}^{N}\frac{n_i}{v_i^2}\frac{\partial v_i}{\partial \theta_\alpha}\frac{\partial v_i}{\partial \theta_\beta} - \delta_\beta^\alpha\frac{1}{\sigma_\alpha^2}, & 2 \leq \alpha, \beta \leq M. \end{cases}$$

Prior to calculating the values of $\boldsymbol{\mu}$, $\mathbf{V}$, and $\mathcal{F}$, one can easily compute the $E(n_i)$ as follows:

$$E(n_i) = \prod_\alpha \int_\alpha d\theta_\alpha \mathcal{N}(\theta_\alpha, \sigma_\alpha)\sum_j \mathcal{P}(n_i|v_i(\boldsymbol{\theta})) = \prod_\alpha \int_\alpha d\theta_\alpha \mathcal{N}(\theta_\alpha, \sigma_\alpha)v_i(\boldsymbol{\theta}) = v_i(\boldsymbol{\theta}) + \sum_\alpha \frac{1}{2}\frac{\partial^2 v_i}{\partial \theta_\alpha^2}\sigma_\alpha^2,$$

where only the term with $\mathcal{O}(\sigma_j^2)$ is taken into consideration, assuming that σ_j is sufficiently small. However, one can opt for more perturbation orders to achieve a more precise outcome. Analogously, we possess: $E(n_i n_j) = v_i v_j + \sum_\alpha \frac{1}{2}\frac{\partial^2 v_i v_j + \delta_j^i v_j}{\partial \theta_\alpha^2}\sigma_\alpha^2 + \delta_j^i v_j$.

Then we acquire the analytical results of $\boldsymbol{\mu}$, $\mathbf{V}$ and $\mathcal{F}$ in Eq. (4). Especially, in our neutrino floor case $v_i(\boldsymbol{\theta}) = \theta_1 s_i + \sum_{j\geq 2}\theta_j b_i^j$, and we have:

$$\mu_\alpha = 0, \qquad 1 \leq \alpha \leq M,$$

$$\mathcal{F}_{\alpha\beta} = \begin{cases} \sum_i \frac{1}{v_i}\frac{\partial v_i}{\partial \theta_\alpha}\frac{\partial v_i}{\partial \theta_\beta}, & \alpha = 1 \text{ or } \beta = 1, \\ \sum_i \frac{1}{v_i}\frac{\partial v_i}{\partial \theta_\alpha}\frac{\partial v_i}{\partial \theta_\beta} + \delta_\beta^\alpha\frac{1}{\sigma_\alpha^2}, & 2 \leq \alpha, \beta \leq M, \end{cases}$$

$$V_{\alpha\beta} = \sum_i \frac{1}{v_i}\frac{\partial v_i}{\partial \theta_\alpha}\frac{\partial v_i}{\partial \theta_\beta} + \sum_{ijk}\frac{\partial v_i}{\partial \theta_k}\frac{\partial v_j}{\partial \theta_k}\frac{\sigma_k^2}{v_i v_j}\frac{\partial v_i}{\partial \theta_\alpha}\frac{\partial v_j}{\partial \theta_\beta}, \qquad 1 \leq \alpha, \qquad \beta \leq M. \tag{B1}$$

When H_0 is real: $\boldsymbol{\theta}' = \boldsymbol{\theta}^0$, the numerical solution tell us that there is only one non-zero diagonal element in Λ with its value at 1. Therefore, the statistic $q_0 = -2\ln\lambda \approx a_1 Y_1^2$, $a_1 = 1$, $Y_1 \sim \mathcal{N}(0, 1)$. Furthermore, the signal strength θ_1 should be positive in this situation, which leads to the fact that q_0 asymptotically follows the $\frac{1}{2}[\delta(0) + \chi_1^2]$ distribution instead of χ_1^2. It can be explained by the Eq. (A4) for $\boldsymbol{\delta} = 0$:

$$\boldsymbol{\theta}^* - \hat{\boldsymbol{\theta}} \approx -(\mathcal{F}^{-1} - \mathbf{H})\dot{l}(\boldsymbol{\theta}'), \tag{B2}$$

where every component of $\boldsymbol{\theta}^* - \hat{\boldsymbol{\theta}}$ is a normal variate. With the positive condition for θ_1, all negative values of $\theta_1^* - \hat{\theta}_1$ are forced to be zero. Consequently, the distribution of the statistic is altered to $\frac{1}{2}[\delta(0) + \chi_1^2]$.

When H_1 is real: $\boldsymbol{\theta}' = \boldsymbol{\theta}^1$, similarly, the numerical solution reveals that there is only one non-zero diagonal element in Λ with value of 1. Thus, the statistic $q_0 \approx a_1 Y_1^2$, $a_1 = 1$, $Y_1 \sim \mathcal{N}(\sqrt{\phi}, 1)$, where ϕ is:

$$\phi = \boldsymbol{\delta}^T\mathcal{F}[\mathcal{F}^{-1} - \mathbf{H}]\mathcal{F}\boldsymbol{\delta}. \tag{B3}$$

It can be directly obtained from Eq. (A5), where $\dot{l}(\boldsymbol{\theta}')$ is neglected since its mean value μ is zero in this case. Making use of the only nonvanishing component in $\boldsymbol{\delta}$ is $\delta_1 = 1$, we perform a more useful form of ϕ:

$$\phi = \delta_1(\mathbf{G}_1 - \mathbf{G}_2\mathbf{G}_3^{-1}\mathbf{G}_2^T)\delta_1, \qquad (B4)$$

where the dimension of $\mathbf{G}_1$ is (1×1).

APPENDIX C: COMPARED WITH RESULTS FROM MC METHOD

As demonstrated in Fig. 6, two new methods in this study are used to reproduce the neutrino floor, and the results closely match the neutrino floor from APPEC report for $m_\chi \gtrsim 6$ GeV. However, there are some discrepancies between our results and APPEC results, which could be caused by the differences in the threshold settings, neutrino fluxes, and uncertainties. To check this, we select some benchmark points along the neutrino floor and obtain the distribution of the test statistic using MC pseudoexperiments, as shown in Fig. 7. Furthermore, as depicted in Fig. 6, the neutrino floor from the asymptotic-analytic method is slightly higher than that from the quasi-Asimov dataset method. This is because the sample size is not large enough to invalidate the approximation utilized in our deduction in Sec. II.

In order to validate our methods' effectiveness, we reproduce the neutrino floor from APPEC report [42] based on MC method. Since the neutrino floor defined by discovery limits depends on the detector configuration, we need to combine several neutrino spectra for different setup to reach the final neutrino floor. The detector threshold can be chosen to be realistic or ideal, while its multiplication with the exposure is assumed to contribute to about 500 neutrino events [7]. As shown in Fig. 6, two new methods are utilized to reproduced the neutrino floor and the results closely match the neutrino floor from APPEC report for $m_\chi \gtrsim 6$ GeV. However, there are some tiny discrepancies between our results and APPEC's result, which might be caused by the differences in the settings on the detector thresholds, neutrino fluxes and their uncertainties. For $m_\chi \lesssim 6$ GeV, since the exposure corresponds to the threshold we choose is not large enough (only 0.018 ton × years), the distribution of the test statistic somehow deviates from the non-central chi-square distribution. To check this, we have chosen some benchmark points along the neutrino floor and obtained the distribution of the test statistic using MC pseudoexperiments, as shown in Fig. 7. Besides, it can be seen from Fig. 6 that the neutrino floor from the asymptotic-analytic method is slightly higher than that from the quasi-Asimov dataset method. That is because the sample size is not large enough
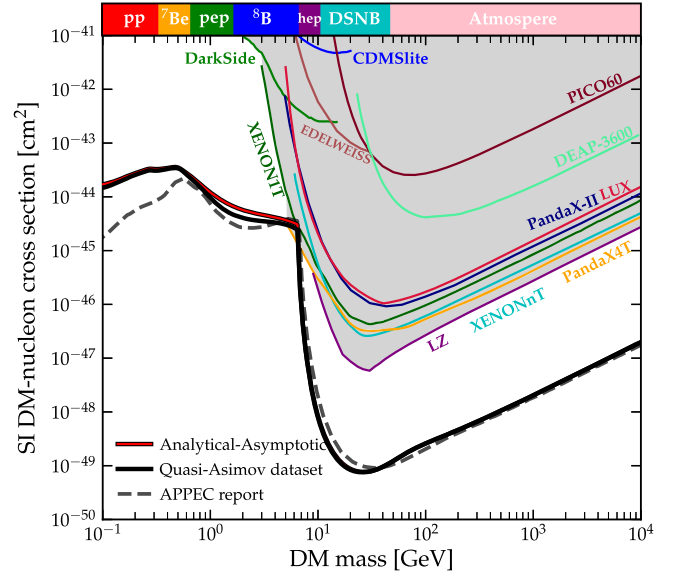

FIG. 6. Neutrino floors defined by discovery limits in the DM parameter space. The red and black solid lines stand for the results from the asymptotic-analytic and quasi-Asimov dataset methods, respectively. The black dashed line is taken from APPEC report. The color-coded neutrino sources, causing primary effects on the neutrino floor over the WIMP mass range, is shown above the figure.

so that the approximation utilized in our deduction in Sec. II becomes invalid.

In order to validate the effectiveness of our method as discussed in Sec. IV A, MC pseudoexperiments are presented here. Note that we fix the exposure at 10 ton × years, and choose the appropriate cross section to demonstrate our results. As shown in the left panel of Fig. 8, we take three benchmark points of interest and run 10000 MC pseudoexperiments for each point. It can be seen from the left panel of Fig. 8 that there are some sizable discrepancies between the result of $m_\chi = 0.51$ GeV and the corresponding prediction. For $m_\chi = 0.1$ GeV and 5.5 GeV, our predictions are in agreement of MC pseudoexperiments.

We also use MC realizations to confirm our statement in Sec. IV B. Note that we use the technique described in Sec. II B to save computational expense by only taking more relevant parameters as inputs. As shown in Fig. 8, we take four benchmark points of interest and run 10000 MC pseudoexperiments for each point. More details can be found in our public codes [43]. One can see that the results from the asymptotic-analytic method are consistent with those from MC realizations, despite of tiny discrepancies. For $m_\chi = 10$ GeV, we might need higher-order corrections since the standard deviation of 21.2% is relatively large in this case.
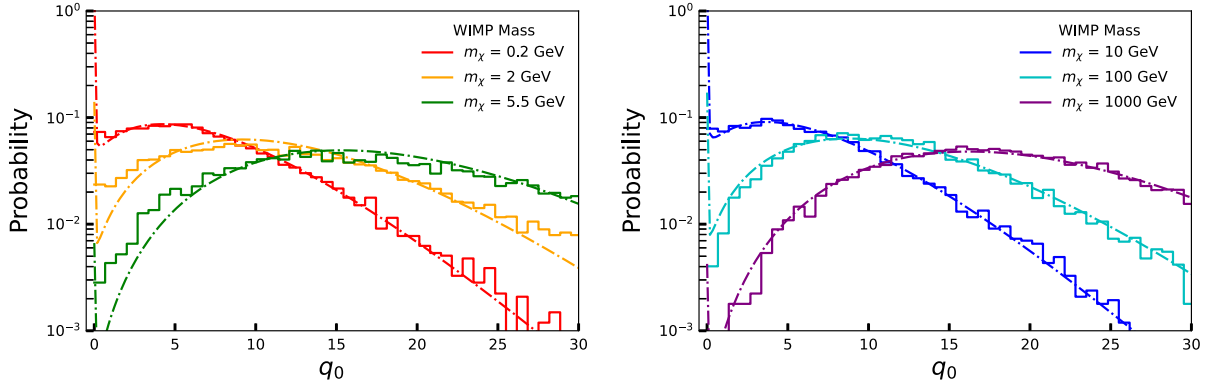
FIG. 7. Comparison between the test statistic's distribution from the theoretical asymptotic prediction and the 10000 MC pseudoexperiments. The non-central parameter ϕ is gained by the quasi-Asimov dataset method. More details can be found in our public codes [43].
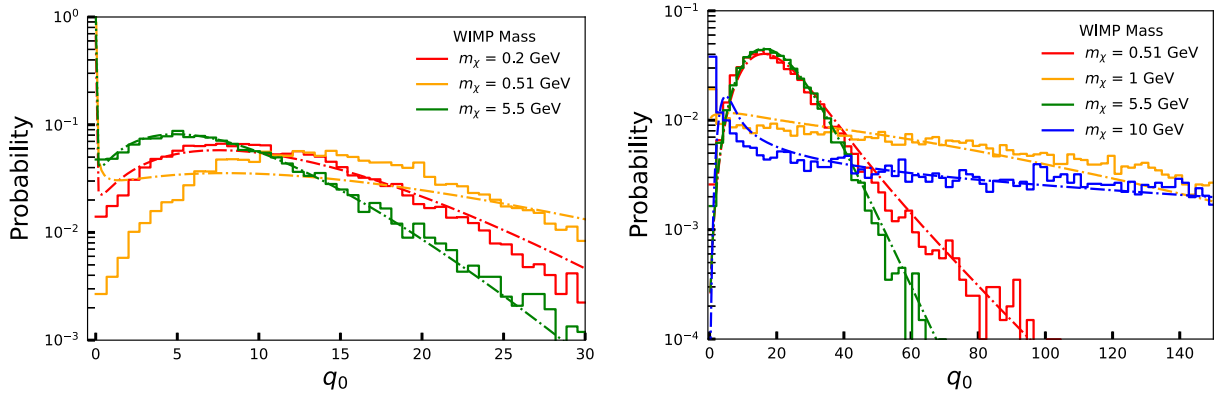

FIG. 8. Comparison between the test statistic's distribution from the theoretical asymptotic prediction and 10000 MC pseudoexperiments, where the test statistic involves the weak mixing angle (left) and the nuisance parameter from the velocity of LSR (right).

[1] Marc Schumann, Direct detection of WIMP dark matter: Concepts and status, J. Phys. G **46**, 103003 (2019).

[2] Andrea Gaspert, Pietro Giampa, and David E. Morrissey, Neutrino backgrounds in future liquid noble element dark matter direct detection experiments, Phys. Rev. D **105**, 035020 (2022).

[3] Marco Cirelli, Gennaro Corcella, Andi Hektor, Gert Hutsi, Mario Kadastik, Paolo Panci, Martti Raidal, Filippo Sala, and Alessandro Strumia, PPPC 4 DM ID: A poor particle physicist cookbook for dark matter indirect detection, J. Cosmol. Astropart. Phys. 03 (2011) 051; 10 (2012) E01.

[4] D. Akimov *et al.*, Observation of coherent elastic neutrino-nucleus scattering, Science **357**, 1123 (2017).

[5] E. Aprile *et al.*, Search for Coherent Elastic Scattering of Solar ^{8}B Neutrinos in the XENON1T Dark Matter Experiment, Phys. Rev. Lett. **126**, 091301 (2021).

[6] Wenbo Ma *et al.*, Search for Solar B8 Neutrinos in the PandaX-4T Experiment Using Neutrino-Nucleus Coherent Scattering, Phys. Rev. Lett. **130**, 021802 (2023).

[7] J. Billard, L. Strigari, and E. Figueroa-Feliciano, Implication of neutrino backgrounds on the reach of next generation dark matter direct detection experiments, Phys. Rev. D **89**, 023524 (2014).

[8] J. Billard, F. Mayet, and D. Santos, Assessing the discovery potential of directional detection of dark matter, Phys. Rev. D **85**, 035006 (2012).

[9] F. Ruppin, J. Billard, E. Figueroa-Feliciano, and L. Strigari, Complementarity of dark matter detectors in light of the neutrino background, Phys. Rev. D **90**, 083510 (2014).

[10] Ciaran A. J. O'Hare, Dark matter astrophysical uncertainties and the neutrino floor, Phys. Rev. D **94**, 063527 (2016).

[11] James B. Dent, Bhaskar Dutta, Jayden L. Newstead, and Louis E. Strigari, Effective field theory treatment of the neutrino background in direct dark matter detection experiments, Phys. Rev. D **93**, 075018 (2016).

[12] M. C. Gonzalez-Garcia, Michele Maltoni, Yuber F. Perez-Gonzalez, and Renata Zukanovich Funchal, Neutrino discovery limit of dark matter direct detection experiments in the presence of non-standard interactions, J. High Energy Phys. 07 (2018) 019.

[13] Enrico Bertuzzo, Frank F. Deppisch, Suchita Kulkarni, Yuber F. Perez Gonzalez, and Renata Zukanovich Funchal, Dark matter and exotic neutrino interactions in direct detection searches, J. High Energy Phys. 04 (2017) 073.

[14] D. Aristizabal Sierra, V. De Romeri, L. J. Flores, and D. K. Papoulias, Impact of COHERENT measurements, cross section uncertainties and new interactions on the neutrino floor, J. Cosmol. Astropart. Phys. 01 (2022) 055.

[15] Jonathan H. Davis, Dark matter vs. neutrinos: The effect of astrophysical uncertainties and timing information on the neutrino floor, J. Cosmol. Astropart. Phys. 03 (2015) 012.

[16] Sebastian Sassi, Abolfazl Dinmohammadi, Matti Heikinheimo, Nader Mirabolfathi, Kai Nordlund, Hossein Safari, and Kimmo Tuominen, Solar neutrinos and dark matter detection with diurnal modulation, Phys. Rev. D **104**, 063037 (2021).

[17] Ciaran A. J. O'Hare, Anne M. Green, Julien Billard, Enectali Figueroa-Feliciano, and Louis E. Strigari, Readout strategies for directional dark matter detection beyond the neutrino background, Phys. Rev. D **92**, 063518 (2015).

[18] Ciaran A. J. O'Hare, Bradley J. Kavanagh, and Anne M. Green, Time-integrated directional detection of dark matter, Phys. Rev. D **96**, 083011 (2017).

[19] Ciaran A. J. O'Hare, Can we overcome the neutrino floor at high masses?, Phys. Rev. D **102**, 063024 (2020).

[20] Philipp Grothaus, Malcolm Fairbairn, and Jocelyn Monroe, Directional dark matter detection beyond the neutrino bound, Phys. Rev. D **90**, 055018 (2014).

[21] F. Mayet *et al.*, A review of the discovery reach of directional dark matter detection, Phys. Rep. **627**, 1 (2016).

[22] Tarso Franarin and Malcolm Fairbairn, Reducing the solar neutrino background in dark matter searches using polarized helium-3, Phys. Rev. D **94**, 053004 (2016).

[23] S. E. Vahsen *et al.*, CYGNUS: Feasibility of a nuclear recoil observatory with directional sensitivity to dark matter and neutrinos, arXiv:2008.12587.

[24] Sven E. Vahsen, Ciaran A. J. O'Hare, and Dinesh Loomba, Directional recoil detection, Annu. Rev. Nucl. Part. Sci. **71**, 189 (2021).

[25] Ciaran A. J. O'Hare, New Definition of the Neutrino Floor for Direct Dark Matter Searches, Phys. Rev. Lett. **127**, 251802 (2021).

[26] D. S. Akerib *et al.*, Snowmass2021 cosmic frontier dark matter direct detection to the neutrino fog, arXiv:2203.08084.

[27] Glen Cowan, Kyle Cranmer, Eilam Gross, and Ofer Vitells, Asymptotic formulae for likelihood-based tests of new physics, Eur. Phys. J. C **71**, 1554 (2011); **73**, 2501(E) (2013).

[28] S. S. Wilks, The large-sample distribution of the likelihood ratio for testing composite hypotheses, Ann. Math. Stat. **9**, 60 (1938).

[29] Abraham Wald, Tests of statistical hypotheses concerning several parameters when the number of observations is large, Trans. Am. Math. Soc. **54**, 426 (1943).

[30] Damiano F. G. Fiorillo and Mauricio Bustamante, Bump hunting in the diffuse flux of high-energy cosmic neutrinos, Phys. Rev. D **107**, 083008 (2023).

[31] Nicole F. Bell, Matthew J. Dolan, and Sandra Robles, Dark matter pollution in the diffuse supernova neutrino background, J. Cosmol. Astropart. Phys. 11 (2022) 060.

[32] Sebastian Baum, William DeRocco, Thomas D. P. Edwards, and Saarik Kalia, Galactic geology: Probing time-varying dark matter signals with paleodetectors, Phys. Rev. D **104**, 123015 (2021).

[33] Andrea Caputo, Alexander J. Millar, Ciaran A. J. O'Hare, and Edoardo Vitagliano, Dark photon limits: A handbook, Phys. Rev. D **104**, 095029 (2021).

[34] Thomas S. Ferguson, *A Course in Large Sample Theory* (Routledge, Boca Raton, 2017), https://doi.org/10.1201/9781315136288.

[35] Soumendra N. Lahiri and Krishna B. Athreya, *Measure Theory and Probability Theory* (Springer, New York, NY, 2006), https://doi.org/10.1007/978-0-387-35434-7.

[36] J. D. Lewin and P. F. Smith, Review of mathematics, numerical factors, and corrections for dark matter experiments based on elastic nuclear recoil, Astropart. Phys. **6**, 87 (1996).

[37] Richard H. Helm, Inelastic and elastic scattering of 187-Mev electrons from selected even-even nuclei, Phys. Rev. **104**, 1466 (1956).

[38] A. K. Drukier, Katherine Freese, and D. N. Spergel, Detecting cold dark matter candidates, Phys. Rev. D **33**, 3495 (1986).

[39] N. Wyn Evans, Ciaran A. J. O'Hare, and Christopher McCabe, Refinement of the standard halo model for dark matter searches in light of the Gaia Sausage, Phys. Rev. D **99**, 023012 (2019).

[40] Edoardo Vitagliano, Irene Tamborra, and Georg Raffelt, Grand unified neutrino spectrum at earth: Sources and spectral components, Rev. Mod. Phys. **92**, 045006 (2020).

[41] Herman Chernoff, On the distribution of the likelihood ratio, Ann. Math. Stat. **25**, 573 (1954).

[42] Julien Billard *et al.*, Direct detection of dark matter—APPEC committee report, Rep. Prog. Phys. **85**, 056201 (2022).

[43] https://github.com/zhangblong/AsymptoticAnalysisAndNeutrinoFog.

[44] Xiangyi Cui *et al.*, Dark Matter Results From 54-Ton-Day Exposure of PandaX-II Experiment, Phys. Rev. Lett. **119**, 181302 (2017).

[45] E. Aprile *et al.*, Dark Matter Search Results from a One Ton-Year Exposure of XENON1T, Phys. Rev. Lett. **121**, 111302 (2018).

[46] C. Amole *et al.*, Dark Matter Search Results from the PICO-60 C_3F_8 Bubble Chamber, Phys. Rev. Lett. **118**, 251301 (2017).

[47] P. Agnes *et al.*, DarkSide-50 532-day dark matter search with low-radioactivity argon, Phys. Rev. D **98**, 102006 (2018).

[48] L. Hehn *et al.*, Improved EDELWEISS-III sensitivity for low-mass WIMPs using a profile likelihood approach, Eur. Phys. J. C **76**, 548 (2016).

[49] P. Agnes *et al.*, Low-Mass Dark Matter Search with the DarkSide-50 Experiment, Phys. Rev. Lett. **121**, 081307 (2018).

[50] D. S. Akerib *et al.*, Results from a Search for Dark Matter in the Complete LUX Exposure, Phys. Rev. Lett. **118**, 021303 (2017).

[51] R. Ajaj *et al.*, Search for dark matter with a 231-day exposure of liquid argon using DEAP-3600 at SNOLAB, Phys. Rev. D **100**, 022004 (2019).

[52] R. Agnese *et al.*, Search for low-mass dark matter with CDMSlite using a profile likelihood fit, Phys. Rev. D **99**, 062001 (2019).

[53] E. Aprile *et al.*, First Dark Matter Search with Nuclear Recoils from the XENONnT Experiment, Phys. Rev. Lett. **131**, 041003 (2023).

[54] Jianglai Liu, The first results of PandaX-4T, Int. J. Mod. Phys. D **31**, 2230007 (2022).

[55] J. Aalbers *et al.*, First Dark Matter Search Results from the LUX-ZEPLIN (LZ) Experiment, 2022.

[56] K. S. Kumar, Sonny Mantry, W. J. Marciano, and P. A. Souder, Low energy measurements of the weak mixing angle, Annu. Rev. Nucl. Part. Sci. **63**, 237 (2013).

[57] Jens Erler and Michael J. Ramsey-Musolf, The weak mixing angle at low energies, Phys. Rev. D **72**, 073003 (2005).

[58] Jian Tang and Bing-Long Zhang (to be published).