XENONnT WIMP search: Signal and background modeling and statistical inference

E. Aprile,[1] J. Aalbers,[2] K. Abe,[3] S. Ahmed Maouloud,[4] L. Althueser,[5] B. Andrieu,[4] E. Angelino,[6,7] D. Antón Martin,[8] F. Arneodo,[9] L. Baudis,[10] M. Bazyk,[11] L. Bellagamba,[12] R. Biondi,[13] A. Bismark,[10] K. Boese,[13] A. Brown,[14] G. Bruno,[11] R. Budnik,[15] J. M. R. Cardoso,[16] A. P. Cimental Chávez,[10] A. P. Colijn,[17] J. Conrad,[18] J. J. Cuenca-García,[10] V. D'Andrea,[7,*] L. C. Daniel Garcia,[4] M. P. Decowski,[17] C. Di Donato,[19] P. Di Gangi,[12] S. Diglio,[11] K. Eitel,[20] A. Elykov,[20] A. D. Ferella,[19,7] C. Ferrari,[7] H. Fischer,[14] T. Flehmke,[18] M. Flierman,[17] W. Fulgione,[6,7] C. Fuselli,[17] P. Gaemers,[17] R. Gaior,[4] M. Galloway,[10] F. Gao,[21] S. Ghosh,[22] R. Giacomobono,[23] R. Glade-Beucke,[14] L. Grandi,[8] J. Grigat,[14] H. Guan,[22] M. Guida,[13] P. Gyorgy,[24] R. Hammann,[13,†] A. Higuera,[25] C. Hils,[24] L. Hoetzsch,[13,10,‡] N. F. Hood,[26] M. Iacovacci,[23] Y. Itow,[27] J. Jakob,[5] F. Joerg,[13] Y. Kaminaga,[3] M. Kara,[20] P. Kavrigin,[15] S. Kazama,[27] M. Kobayashi,[27] A. Kopec,[26,§] F. Kuger,[14] H. Landsman,[15] R. F. Lang,[22] L. Levinson,[15] I. Li,[25] S. Li,[28] S. Liang,[25] Y.-T. Lin,[13] S. Lindemann,[14] M. Lindner,[13] K. Liu,[21] J. Loizeau,[11] F. Lombardi,[24] J. Long,[8] J. A. M. Lopes,[16,‖] T. Luce,[14] Y. Ma,[26] C. Macolino,[19,7] J. Mahlstedt,[18] A. Mancuso,[12] L. Manenti,[9] F. Marignetti,[23] T. Marrodán Undagoitia,[13] K. Martens,[3] J. Masbou,[11] E. Masson,[4] S. Mastroianni,[23] A. Melchiorre,[19] M. Messina,[7] A. Michael,[5] K. Miuchi,[29] A. Molinario,[6] S. Moriyama,[3] K. Morå,[1] Y. Mosbacher,[15] M. Murra,[1] J. Müller,[14] K. Ni,[26] U. Oberlack,[24] B. Paetsch,[15] Y. Pan,[4] Q. Pellegrini,[4] R. Peres,[10] C. Peters,[25] J. Pienaar,[8,15] M. Pierre,[17] G. Plante,[1] T. R. Pollmann,[17] L. Principe,[11] J. Qi,[26] J. Qin,[25] D. Ramírez García,[10,¶] M. Rajado,[10] R. Singh,[22] L. Sanchez,[25] J. M. F. dos Santos,[16] I. Sarnoff,[9] G. Sartorelli,[12] J. Schreiner,[13] D. Schulte,[5] P. Schulte,[5] H. Schulze Eißing,[5] M. Schumann,[14] L. Scotto Lavina,[4] M. Selvi,[12] F. Semeria,[12] P. Shagin,[24] S. Shi,[1] J. Shi,[21] M. Silva,[16] H. Simgen,[13] A. Takeda,[3] P.-L. Tan,[18] A. Terliuk,[13,**] D. Thers,[11] F. Toschi,[20] G. Trinchero,[6] C. D. Tunnell,[25] F. Tönnies,[14] K. Valerius,[20] S. Vecchi,[30] S. Vetter,[20] F. I. Villazon Solar,[24] G. Volta,[13] C. Weinheimer,[5] M. Weiss,[15] D. Wenz,[5] C. Wittweg,[10] V. H. S. Wu,[20] Y. Xing,[11] D. Xu,[1] Z. Xu,[1,††] M. Yamashita,[3] L. Yang,[26] J. Ye,[31] L. Yuan,[8] G. Zavattini,[30] and M. Zhong[26]

(XENON Collaboration)[‡‡]

[1]*Physics Department, Columbia University, New York, New York 10027, USA*
[2]*Nikhef and the University of Groningen, Van Swinderen Institute, 9747AG Groningen, Netherlands*
[3]*Kamioka Observatory, Institute for Cosmic Ray Research, and Kavli Institute for the Physics and Mathematics of the Universe (WPI), University of Tokyo, Higashi-Mozumi, Kamioka, Hida, Gifu 506-1205, Japan*
[4]*LPNHE, Sorbonne Université, CNRS/IN2P3, 75005 Paris, France*
[5]*Institut für Kernphysik, Westfälische Wilhelms-Universität Münster, 48149 Münster, Germany*
[6]*INAF-Astrophysical Observatory of Torino, Department of Physics, University of Torino and INFN-Torino, 10125 Torino, Italy*
[7]*INFN-Laboratori Nazionali del Gran Sasso and Gran Sasso Science Institute, 67100 L'Aquila, Italy*
[8]*Department of Physics and Kavli Institute for Cosmological Physics, University of Chicago, Chicago, Illinois 60637, USA*
[9]*New York University Abu Dhabi—Center for Astro, Particle and Planetary Physics, Abu Dhabi, United Arab Emirates*
[10]*Physik-Institut, University of Zürich, 8057 Zürich, Switzerland*
[11]*SUBATECH, IMT Atlantique, CNRS/IN2P3, Université de Nantes, Nantes 44307, France*
[12]*Department of Physics and Astronomy, University of Bologna and INFN-Bologna, 40126 Bologna, Italy*
[13]*Max-Planck-Institut für Kernphysik, 69117 Heidelberg, Germany*
[14]*Physikalisches Institut, Universität Freiburg, 79104 Freiburg, Germany*
[15]*Department of Particle Physics and Astrophysics, Weizmann Institute of Science, Rehovot 7610001, Israel*
[16]*LIBPhys, Department of Physics, University of Coimbra, 3004-516 Coimbra, Portugal*
[17]*Nikhef and the University of Amsterdam, Science Park, 1098XG Amsterdam, Netherlands*
[18]*Oskar Klein Centre, Department of Physics, Stockholm University, AlbaNova, Stockholm SE-10691, Sweden*
[19]*Department of Physics and Chemistry, University of L'Aquila, 67100 L'Aquila, Italy*
[20]*Institute for Astroparticle Physics, Karlsruhe Institute of Technology, 76021 Karlsruhe, Germany*

[21]*Department of Physics and Center for High Energy Physics, Tsinghua University, Beijing 100084, P.R. China*
[22]*Department of Physics and Astronomy, Purdue University, West Lafayette, Indiana 47907, USA*
[23]*Department of Physics "Ettore Pancini," University of Napoli and INFN-Napoli, 80126 Napoli, Italy*
[24]*Institut für Physik and Exzellenzcluster PRISMA$^+$, Johannes Gutenberg-Universität Mainz, 55099 Mainz, Germany*
[25]*Department of Physics and Astronomy, Rice University, Houston, Texas 77005, USA*
[26]*Department of Physics, University of California San Diego, La Jolla, California 92093, USA*
[27]*Kobayashi-Maskawa Institute for the Origin of Particles and the Universe, and Institute for Space-Earth Environmental Research, Nagoya University, Furo-cho, Chikusa-ku, Nagoya, Aichi 464-8602, Japan*
[28]*Department of Physics, School of Science, Westlake University, Hangzhou 310030, P.R. China*
[29]*Department of Physics, Kobe University, Kobe, Hyogo 657-8501, Japan*
[30]*INFN-Ferrara and Dip. di Fisica e Scienze della Terra, Università di Ferrara, 44122 Ferrara, Italy*
[31]*School of Science and Engineering, The Chinese University of Hong Kong, Shenzhen, Guangdong, 518172, People's Republic of China*

(Received 21 June 2024; accepted 25 March 2025; published 19 May 2025)

The XENONnT experiment searches for weakly interacting massive particle (WIMP) dark matter scattering off a xenon nucleus. In particular, XENONnT uses a dual-phase time projection chamber with a 5.9-ton liquid xenon target, detecting both scintillation and ionization signals to reconstruct the energy, position, and type of recoil. A blind search for nuclear recoil WIMPs with an exposure of 1.1 ton-years (4.18 t fiducial mass) yielded no signal excess over background expectations, from which competitive exclusion limits were derived on WIMP-nucleon elastic scatter cross sections, for WIMP masses ranging from 6 GeV$/c^2$ up to the TeV$/c^2$ scale. This work details the modeling and statistical methods employed in this search. By means of calibration data, we model the detector response, which is then used to derive background and signal models. The construction and validation of these models is discussed, alongside additional purely data-driven backgrounds. We also describe the statistical inference framework, including the definition of the likelihood function and the construction of confidence intervals.

DOI: 10.1103/PhysRevD.111.103040

I. INTRODUCTION

Astrophysical and cosmological observations at various scales provide clear evidence for the existence of dark matter (DM) as a fundamental building block of our universe [1,2]. Numerous extensions of the standard model of particle physics predict additional fundamental particles as potential candidates for DM [3,4]. Measuring the direct interaction of particles from the DM halo of the Milky Way with a suitable detector could provide conclusive evidence for the particle DM hypothesis. The XENONnT experiment aims to detect a signal from weakly interacting massive particles (WIMPs) elastically scattering off xenon nuclei with a detector operated underground at the Istituto Nazionale di Fisica Nucleare (INFN) Laboratori Nazionali del Gran Sasso (LNGS) in Italy. The collaboration published the first WIMP search results from a data-taking period between July 6 and November 10 2021 with a total live time of 95.1 days and a fiducial mass of (4.18 ± 0.13) t, referred to as science run 0 (SR0). We performed a blind analysis and observed no significant excess above background expectations [5]. This led to a minimum upper limit on the spin-independent WIMP-nucleon cross section of 2.58×10^{-47} cm^2 for a WIMP mass of 28 GeV$/c^2$ at a 90% confidence level (CL).

The detector used for the XENONnT experiment is a dual-phase xenon time projection chamber (TPC). A particle interaction in the active 5.9 t liquid xenon (LXe) target results in prompt vacuum ultraviolet scintillation light as well as free ionization electrons. The instantaneous light signal, called the S1 signal, is detected by two

*Also at INFN-Roma Tre, 00146 Roma, Italy.
†Contact author: robert.hammann@mpi-hd.mpg.de
‡Contact author: luisa.hoetzsch@mpi-hd.mpg.de
§Present address: Department of Physics and Astronomy, Bucknell University, Lewisburg, Pennsylvania, USA.
‖Also at Coimbra Polytechnic, ISEC, 3030-199 Coimbra, Portugal.
¶Contact author: diego.ramirez@physik.uzh.ch
**Also at Physikalisches Institut, Universität Heidelberg, Heidelberg, Germany.
††Contact author: zihao.xu@columbia.edu
‡‡Contact author: xenon@lngs.infn.it

Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI. Open access publication funded by the Max Planck Society.

arrays of photomultiplier tubes (PMTs) at the top and bottom of the approximately cylindrical TPC. An electric field (23 V/cm average electric drift field during SR0) is applied to the target volume to drift the ionization electrons toward the liquid xenon surface. Here, they are extracted and accelerated into a xenon gas volume above the liquid via a stronger electric field (2.9 kV/cm electric extraction field in liquid during SR0). The kinetic energy of the accelerated electrons in the gas phase is sufficient for the emission of electroluminescent light, which is proportional to the number of extracted electrons. This second light signal is referred to as the S2 signal. The distribution of the S2 signal detected by the PMTs in the top array is used to infer the position of the initial interaction in the horizontal plane (X, Y) parallel to the liquid surface, which defines the radial coordinate $R = \sqrt{X^2 + Y^2}$. The vertical component Z follows from the drift time of electrons, which is determined as the time difference between the S1 and S2 signals. This three-dimensional position reconstruction of the interaction vertex enables the discrimination of multisite events as well as the selection of an inner fiducial volume (FV) within the TPC. This volume benefits from a particularly low background level thanks to the excellent self-shielding properties of LXe, given its high density. Moreover, the relative magnitudes of the S1 and S2 signals can be used to determine if a particle interacted with the xenon nucleus (expected from WIMPs) or xenon shell electrons (typical of the dominant sources of background), i.e., nuclear recoil (NR) or electronic recoil (ER) events. The electric drift and extraction fields are established by three parallel-wire electrodes (cathode, gate, and anode from bottom to top). To reduce wire sagging, two and four horizontal perpendicular wires support the gate and anode electrodes, respectively.

The TPC is nested in an active neutron veto, which in turn is housed in an active muon veto. Both veto systems are water Cherenkov detectors. More details on the TPC, the veto systems, and other subsystems of the detector are provided in [6–8].

The data analysis chain of the WIMP DM search in XENONnT is subdivided into two major parts, like in XENON1T [9,10]. The first part is discussed in [11] and comprises signal and event reconstruction, corrections, event selection, and energy scale calibration. Here, we present the second part of the analysis chain: the detector response to the different interaction types is modeled based on calibration data and discussed in Sec. II. The derived best-fit models are important inputs for the definition of signal and background models, detailed in Sec. III. This section also covers data-driven background models that do not rely on the detector response. Finally, the statistical methods used to compute constraints on DM given the XENONnT data are discussed in Sec. IV.

II. MODELING THE DETECTOR RESPONSE

Modeling the detector response to energy depositions allows for a powerful discrimination between the NR signal and ER background in S1–S2 space. NR events are produced by particles scattering elastically off xenon nuclei, while in ER events, particles scatter off xenon shell electrons. The different energy loss processes involved after the two types of recoil cause ER and NR events to form separate populations in S1–S2 space, which is exploited in the WIMP DM search. The modeling is performed in the region of interest (ROI) designated for the SR0 WIMP search, defined by cS1 in [0, 100] PE (PE = photoelectron) and cS2 in $[10^{2.1}, 10^{4.1}]$ PE. This ROI is selected to contain the majority of our expected signal (cf. Figs. 4 and 5). We build the models using calibration data before unblinding the WIMP search dataset.

In this section, we describe the TPC response model to energy depositions in LXe via ER or NR interactions, which is obtained from fits to calibration data. This detector response model is separated into two parts. The first one is the empirically parametrized LXe emission model, which describes the production processes of the detectable quanta, i.e., the conversion of deposited energy into scintillation photons and ionization electrons. The second part is the detector reconstruction model, which covers the conversion from the produced photons and electrons into the observed and spacetime-dependence-corrected S1 and S2 signals (cS1 and cS2). Due to its complexity, it is unfeasible in the fit to simulate processes of photon and electron propagation on a per-quantum basis. Instead, toy Monte Carlo (toy-MC) simulations of the detector reconstruction model are used, which sample from effective maps provided either by data-driven analyses described in [11] or by simulation-driven analyses using the XENONnT Monte Carlo (MC) framework [12,13]. All model parameters in the simulations are fit to both ER and NR calibration data. The two parts of the detector response model together with the toy-MC simulation workflow are described in Secs. II A and II B, and the fit to ER and NR calibration data is detailed in Sec. II C.

A. Liquid xenon emission model

The production of quanta from energy depositions in a xenon target is complex and lacks a comprehensive description derived from first principles. Thus, a semi-empirical parametrization of the emission model is commonly used. The parametrization used in XENONnT SR0 is similar to XENON1T [10], which is based on the noble element simulation technique (NEST) models described in [14,15]. The simulation workflow of the emission model is described in detail in Appendixes A and B, and is summarized in the following.

In an ER event, the recoil energy is transmitted into the excitation and ionization of xenon atoms. Recoiling xenon

nuclei from an NR event, on the other hand, lose a significant amount (roughly 80%) of their kinetic energy to atomic motion in collisions with other xenon atoms [14]. This thermalization process is undetectable in LXe TPCs, resulting in an effective signal loss. For both recoil types, the total number of detectable quanta N_q equals the sum of the number of produced excitons N_{ex} and ions N_i. A fraction of ions and electrons recombine, resulting in the production of additional excitons. An exciton can combine with a ground-state xenon atom, forming an excimer, which de-excites, producing a scintillation photon. In an event, N_γ photons are generated, part of which are then detected by PMTs as the S1 signal. The N_e electrons that do not recombine are drifted to the liquid-gas interface and eventually form an S2 signal.

In the SR0 WIMP search ROI, ER interactions have a much smaller ratio of produced excitons to electron-ion pairs than NR interactions (by a factor of up to ∼10 in the ROI), while the fraction of electron-ion pairs that recombine is similar for both interaction types given the same number of detectable quanta. The difference in the exciton-to-ion ratio dominates, leading to larger ratios of S2 to S1 areas for ER interactions, which is the principle behind ER-NR discrimination in LXe TPCs. We parametrize the emission model for ER interactions as in [10], with a constant exciton-to-ion ratio, and a mean recombination fraction following the Thomas-Imel box model [16] with an additional Fermi-Dirac term. For NR interactions, the emission model follows the parametrization used in the NEST v1 model detailed in [15]. While there are newer versions of fits available in the NEST framework that use a different parametrization for NR interactions (which we refer to as NEST v2 in the following), we choose to stay with the previous version (referred to as NEST v1). The model is simpler, fits our data well, and the best-fit values and uncertainties of parameters of NEST v2 were not published at the time of the analysis.

B. Detector reconstruction model

The detector reconstruction model covers all aspects of the signal reconstruction, starting from the produced scintillation photons and ionization electrons up to the measured S1 and S2 signal sizes. The XENONnT detector reconstruction model is almost identical to the one presented in [10] and is briefly outlined here. The detailed simulation workflow of the detector reconstruction model is described in Appendix C.

For S1 signals, the spatial dependence of geometric effects during photon propagation, the photon detection efficiency of the PMTs, and the effect of double photoelectron emission (DPE) from the PMT photocathode [17,18] are all modeled in the simulations. For S2 signals, we model the loss of electrons along their drift path due to attachment to electronegative impurities, as well as due to electric field effects close to the TPC wall, the efficiency of electron extraction from the liquid into the gas phase, and the single-electron gain of extracted electrons to detected PE via the electroluminescence process in gas. In addition to these physical processes, we also account for the influence of software reconstruction effects on the signals. For both signal types, the software reconstruction process can introduce biases and fluctuations in the recorded signal size. We also account for the S1 signal reconstruction efficiency, which vanishes at about 3 PE, as we require that at least 3 PMTs detect a photon within ±50 ns around the maximal amplitude. Finally, we model the uncertainty of the event position reconstruction and the acceptances of data quality selections [11].

The modeling of ER interactions relies on calibration data from sources dissolved in xenon, in particular ^{220}Rn and ^{37}Ar. The ^{220}Rn progeny ^{212}Pb undergoes a β-decay with a Q-value at about 0.6 MeV [19], giving an approximately flat ER energy spectrum in the WIMP ROI, i.e., below about 10 keV. This dataset is used for the fit of the ER model. The ^{37}Ar source provides a monoenergetic ER peak when its K-shell electron capture process leads to a fast cascade of x rays and Auger-Meitner electrons, with a total energy of about 2.8 keV [20]. Both ER sources provide single-scatter (SS) events, producing one S1 and one S2 signal for the total deposited energy, and the above-described detector reconstruction processes are simulated accordingly.

The NR model is calibrated using an external ^{241}AmBe neutron source (referred to as AmBe source in the following). It emits fast neutrons via the $^{9}\mathrm{Be}(\alpha, n)^{12}\mathrm{C}$ capture reaction. Neutrons often scatter multiple times in the LXe target in so-called multiscatter (MS) events, creating distinct energy depositions at each interaction site, thus the topologies of neutron events must be considered in the modeling. Since the neutrons travel at about 10% of the speed of light and the S1 width is $\mathcal{O}(100\ \mathrm{ns})$, the S1 signals from the separate energy depositions in a MS event get reconstructed into one merged S1 signal. The S2 signals might be distinguished based on the separation in space-time between the energy depositions of the individual scatters, but can get (partially) merged if the interaction sites are close together. Due to the low drift field of 23 V/cm in XENONnT SR0, the Z separation resolution between two S2 signals is reduced with respect to the design because of the increased drift time $\mathcal{O}(\mathrm{ms})$ and, consequently, larger S2 spread $\mathcal{O}(\mu\mathrm{s})$ due to the diffusion of electrons. This effect is accounted for in the fit of the NR model to the neutron calibration data: we first produce the photons and electrons for each energy deposition in a MS event separately, based on the emission model. While the primary scintillation photons are summed to produce the merged S1 signal, we only sum S2 signals that will become part of the largest S2 signal after going through the remaining detector reconstruction processes. This information is provided by PMT waveform simulations of neutron

scatters from the AmBe source, using the MC framework of XENONnT [12,13] (see also Sec. III C for details on neutron simulations). The results of these waveform simulations are used as inputs to the toy-MC simulations and fitting framework. Data selections that remove MS events (both resolved and unresolved) are directly applied to these inputs, because their acceptances can only be determined from full simulations, and are correspondingly dropped from the selection acceptance curves for the NR fit.

C. Fit to calibration data

The parameters in the full detector response model are fitted to ^{220}Rn, ^{37}Ar, and AmBe calibration data. The ROI selection is the same as the SR0 WIMP search region defined in Sec. II A. All data quality selections are applied to the calibration data within the ROI. For the AmBe neutron data, an additional selection is applied using coincident gammas detected in the neutron veto. In $^{9}\mathrm{Be}(\alpha, n)^{12}\mathrm{C}$ capture reactions in the AmBe source, ^{12}C is left in the first excited state after neutron emission in about 50% of the cases, estimated with neutron veto data [21]. The ^{12}C excited state promptly decays via the emission of a 4.44 MeV gamma ray, coincident with the neutron emission. The requirement of the time coincidence within 250 μs between the 4.44 MeV gamma ray detected in the neutron veto and the event in the TPC leads to a clean calibration dataset, with only about one expected event from accidental coincidences between the two detectors in the dataset (∼0.05% of the resulting sample). The events recorded in the ROI that pass (fail) the neutron veto coincidence selection are shown in red (gray) before applying any additional data quality selections in Fig. 1.

For the fit to the ER data, an additional background component is added to account for accidental coincidence (AC; see Sec. III E for details) events formed by a wrong pairing of S1 and S2 signals in the calibration dataset.

After all selections, the ^{220}Rn and AmBe datasets consist of ∼2000 events each. We use a downsampled dataset of 10 000 ^{37}Ar events to perform a combined ^{220}Rn-^{37}Ar ER fit. With the downsampling, we avoid overconstraining the fit to the narrow, very-low-energy range of ^{37}Ar.

For the fit, an affine-invariant Markov chain Monte Carlo (MCMC) algorithm [22,23] is used to sample from the high-dimensional posterior distribution of the parameters. In each step of the sampling, a GPU-accelerated toy-MC simulation is performed for every set of the ER and NR model parameters, following the steps described in the previous two sections, producing a model of the detector response in the space of corrected signals, i.e., cS1–cS2. The model parameters are described in detail in the Appendix and are listed in Table II. These toy-MC models are then compared to calibration data by calculating the binned Poisson likelihood in cS1–cS2 space. The likelihoods are multiplied by the prior distributions of the parameters to yield the posteriors. After sufficient iterations

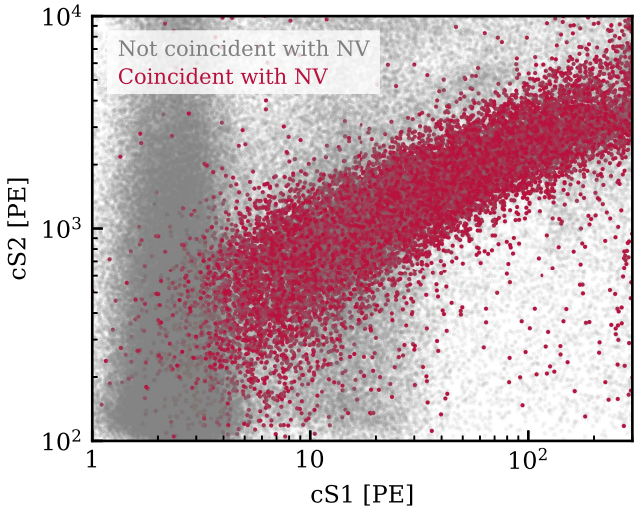

FIG. 1. AmBe neutron calibration events with (red) and without (gray) a coincident signal in the neutron veto. Selecting coincident events ensures a clean nuclear recoil sample for detector response modeling. The accidental coincidence population (cS1 below 5 PE) and misidentified single-electron S1s (cS1 between 10 and 30 PE) are visible in the gray population. After applying additional data quality selections, approximately 2000 events (Fig. 2, right) are used for band fitting.

of the MCMC, the samples in the chain then converge to the posterior distribution of the parameters.

In the ER emission model, the parametrization is the same as in XENON1T [10] with very wide flat priors, referred to as free priors. For the NR emission model, in turn, the parameter priors are taken from [15]. A Gaussian-like distribution with different widths on either side is used as the prior of parameters with asymmetric uncertainties. However, due to the low drift field of 23 V/cm, where no literature data on NR yields exist, the validity of the field dependence in the model is unverified. Therefore, the field dependence in the emission model from [15] was modified. Two scaling parameters were freed [α and γ in Eqs. (B8) and (B11) in the Appendix], and two field dependence parameters (ζ and δ) were fixed to the reported best-fit values. The widths of all other parameter priors were doubled in order to allow more freedom in the fit. Because the dependence of NR signals on the drift field has been shown to be small (see also literature measurements shown in Fig. 3), we consider this a reasonable choice.

The fit results of all emission model parameters are summarized in Table II in the Appendix. No significant tension between the posteriors and the priors of any parameter is found. Suitable goodness-of-fit (GOF) tests are performed to assess whether the best-fit models adequately describe the ^{220}Rn and AmBe calibration data. Specifically, we employ binned Poisson likelihood χ^2 tests in the cS1–cS2 space. We adopt an equiprobable binning scheme using the GOFevaluation package [24], ensuring that

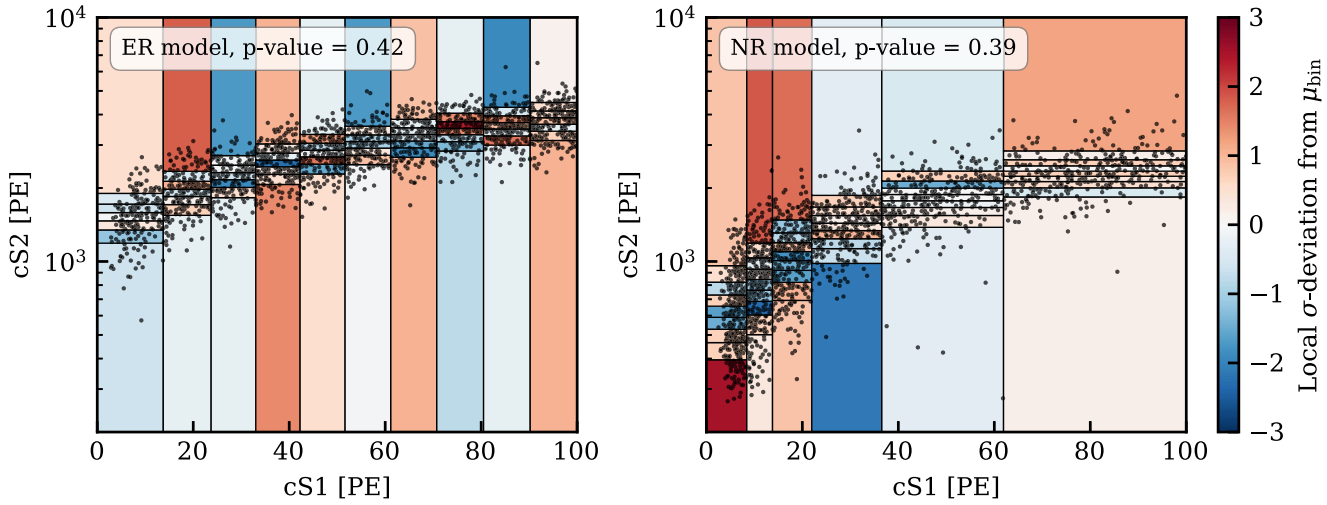

FIG. 2. Comparison between calibration data and the best-fit ER (left) and NR (right) models. The equiprobable binning for the 2D binned Poisson likelihood χ^2 goodness-of-fit tests is shown. The color scale indicates the deviation of the number of data points (overlaid as black dots) in each bin from the best-fit model expectation μ_{bin} in units of sigma, which is 29.3 for the ER model and 36.9 for the NR model.

the number of expected events under the best-fit model is the same in each bin, μ_{bin}. In Fig. 2, the results of the 2D cS1–cS2 Poisson likelihood χ^2 test are shown for both the ER and the NR fit, overlaid with the ^{220}Rn and the AmBe calibration data, respectively. The resulting p-values of the tests indicate no significant discrepancy between the best-fit models and the calibration data. The few outlier events below the NR band in the right panel are modeled by accounting for S2 charge losses near the TPC walls (see also Appendix C).

Fig. 3 shows the photon and charge yields as a function of deposited energy for the ER and NR emission models using our best-fit parameters (dark red). For comparison, NEST models calculated at a drift field of 23 V/cm and several published measurements at different drift fields are shown. For the NR yields, we also compare our results to the NEST v1 model at a drift field of 23 V/cm (red dashed lines in Fig. 3), which uses the same parametrization as the model in this work (see Sec. II A), showing very good agreement.

III. SIGNAL AND BACKGROUND MODELING

The dominant sources of background in the WIMP search are ER events from intrinsic β-decays from materials and neutrinos (Sec. III B), NR events from radiogenic neutrons from detector materials (Sec. III C) and coherent neutrino scattering (Sec. III D), AC events (Sec. III E), and surface background (Sec. III F). Besides their total expected rates, the distribution of these background events in the analysis space cS1–cS2–R is derived. In Fig. 4, the distributions of background models in the cS1–cS2 space are shown, compared to the signal model for elastic spin-independent scattering of a 200 GeV/c^2 mass WIMP.

A. WIMP signal model

For nonrelativistic, spin-independent elastic WIMP-nucleus coherent scattering, the event rate R_{WIMP} scales with the square of the atomic mass number A^2 of the target and can be written as [40]

$$\frac{\mathrm{d}R_{\text{WIMP}}}{\mathrm{d}E} = \frac{\rho_\chi}{2m_\chi \mu_{\chi N}^2}\left\langle\frac{1}{v}\right\rangle \sigma A^2 F^2(q), \tag{1}$$

where E is the recoil energy, ρ_χ is the local DM density of 0.3 GeV/(c^2 cm^3), m_χ is the WIMP mass, $\mu_{\chi N}$ is the WIMP-nucleon reduced mass, v is the WIMP velocity in the lab-frame, σ is the WIMP-nucleon cross section, and $F(q)$ is the nuclear form factor as a function of the momentum transfer q to the xenon nucleus [41]. The DM velocity distribution is averaged using the parameter values of the standard halo model with values from [42–47], as recommended in [48]. Figure 5 shows the cS1–cS2 distribution of WIMP-nucleus scattering for different WIMP masses. For increasing masses, the 68% contours extend to higher cS1 and cS2 values, up to about 200 GeV/c^2, where the shape does not change significantly anymore due to kinematic effects. For this reason, confidence intervals in σ (see Sec. IV) for $m_\chi \gtrsim$ 200 GeV/c^2 approximately scale proportional to the inverse of the assumed WIMP mass m_χ.

For the spin-dependent WIMP-nucleus interaction, assuming natural abundances of xenon isotopes, only ^{129}Xe and ^{131}Xe can contribute since they are the only two stable isotopes of xenon with nonzero nuclear spin J. The cross section is usually written as

$$\frac{\mathrm{d}\sigma_{\text{SD}}}{\mathrm{d}q^2} = \frac{\sigma}{3\mu_{\chi N}^2 v^2}\frac{\pi}{2J+1} S_N(q). \tag{2}$$

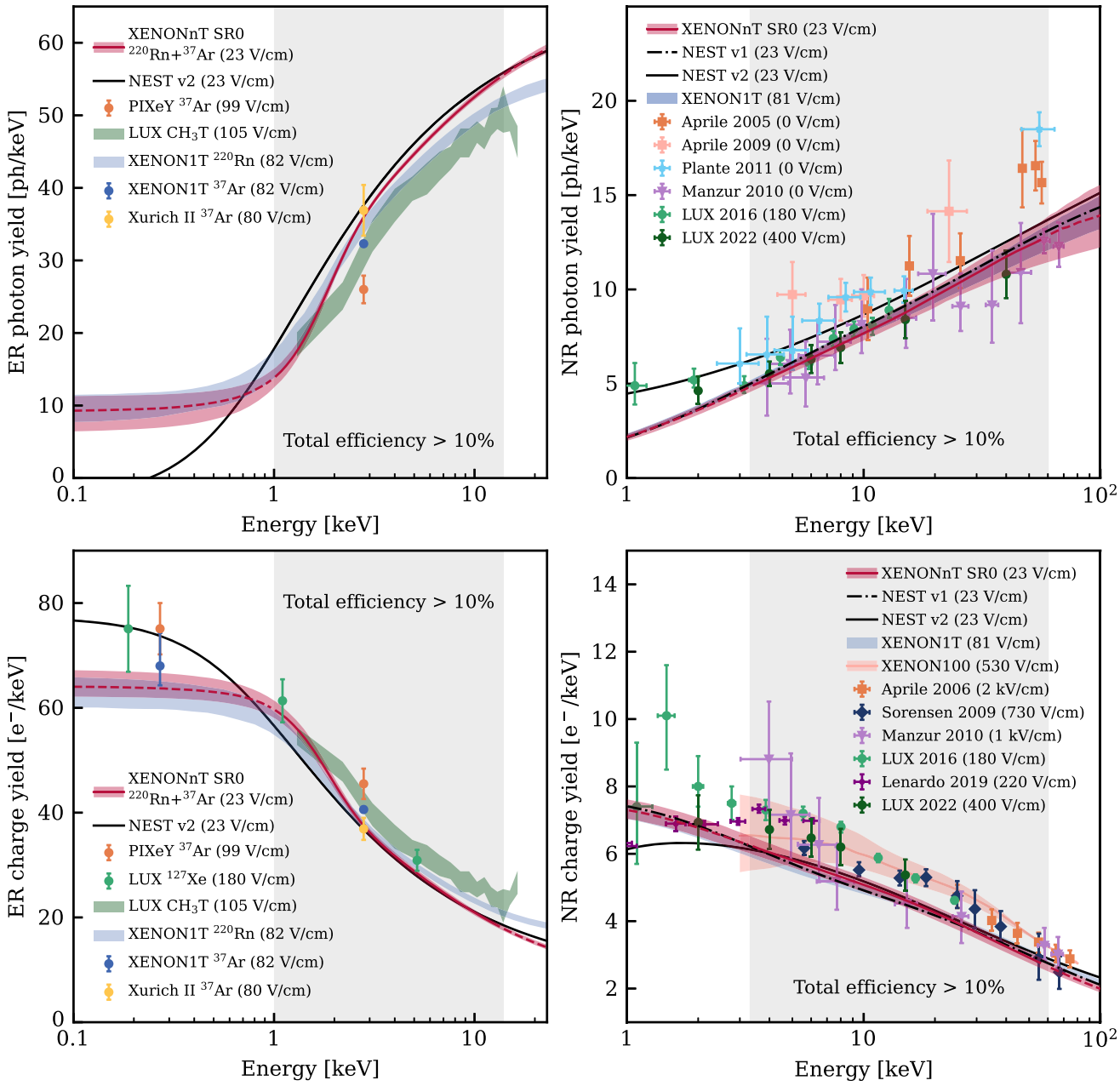

FIG. 3. Photon (top) and charge (bottom) yields as functions of deposited energy for ER events (left) and NR events (right). Dark red lines and shaded bands show the XENONnT liquid xenon emission models with the best-fit parameters and their uncertainties, respectively. The fit results from XENON1T (blue shaded bands) [10,25], and the models from NEST v1 (black dash-dotted line) and NEST v2 (black solid line), both evaluated at 23 V/cm, are shown as well. Several measurements from literature performed at various drift fields as labeled in the legends (ER yields from [10,20,26–29], NR yields from Aprile 2005 [30], Aprile 2006 [31], Aprile 2009 [32], XENON100 [33], Plante 2011 [34], Sorensen 2009 [35], Manzur 2010 [36], LUX 2016 [37], Lenardo 2019 [38], and LUX 2022 [39]) are also shown. Note that for the ER yields, a considerable dependence on the electric drift field is expected, as visible in the literature data points. The gray shaded regions mark the energy ranges in which the total detection + selection efficiency in the WIMP analysis is above 10%.

The axial-vector structure factor of xenon S_N is taken from [49]. Note that the uncertainty on the structure factor dominates, especially for the "neutron-only" case. In the XENONnT SR0 analysis, the mean of the structure factor is used.

Uncertainties from the posterior distribution of the NR model parameters and other efficiencies can be propagated to the NR rate uncertainty. For a 6 GeV/c^2 (50 GeV/c^2) WIMP, the relative rate uncertainty is ∼30% (10%). The

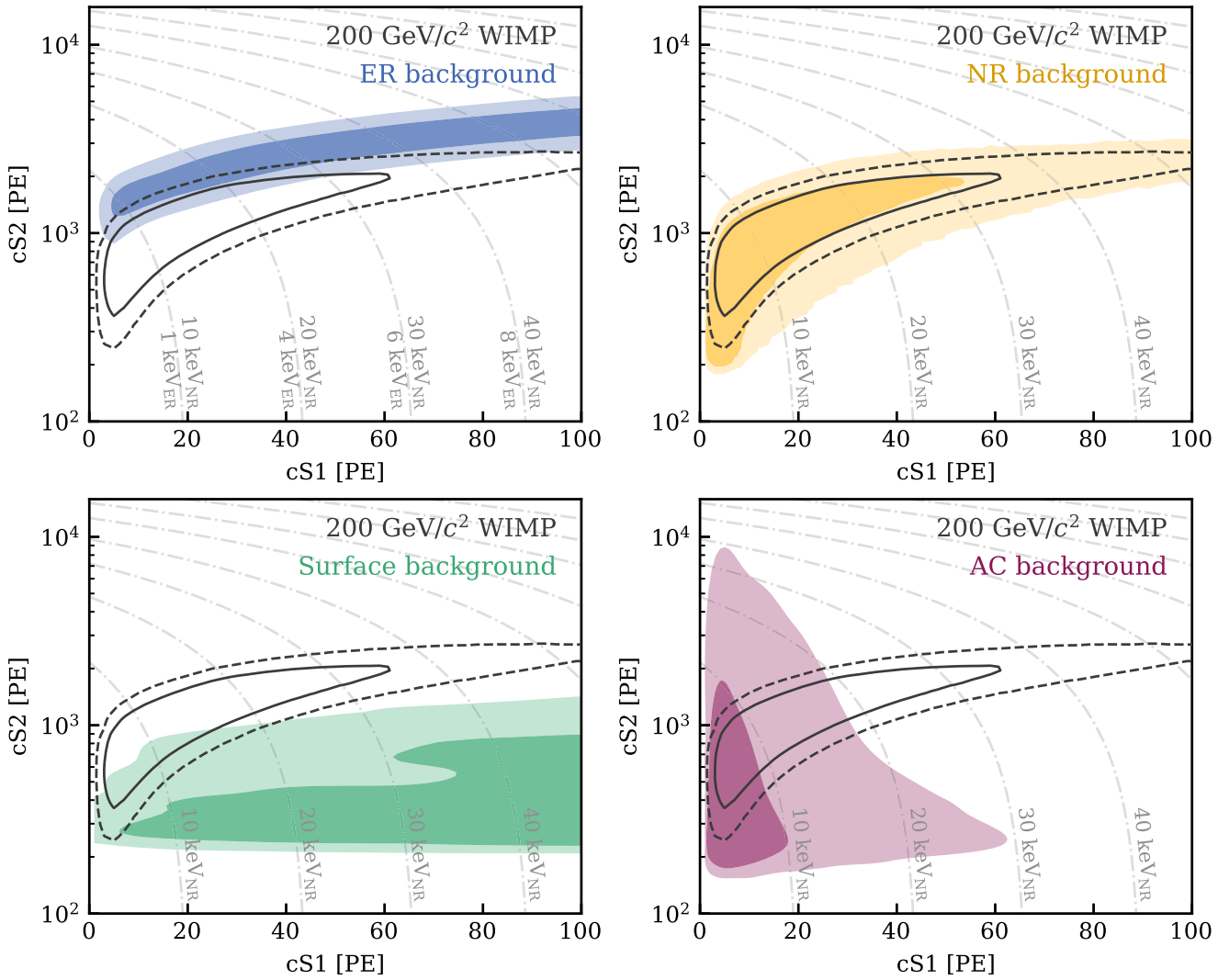

FIG. 4. Distribution of each background component (colored) in cS1–cS2 space. The NR background includes both radiogenic neutron background and NR events by neutrinos. The probability density functions (PDFs) of backgrounds are shown as 1σ (2σ) dark (light) regions, containing 68% (95%) expected events in the ROI. For reference, the contours are also shown for a spin-independent 200 GeV/c^2 WIMP as a solid (dashed) dark gray line. The light gray dash-dotted lines are contours of constant NR-equivalent energy (and ER-equivalent energy in the top left panel) in units of keV.

NR model shape can also be affected by the posterior distribution. However, because of the low statistics of NR events in SR0 (for both background and signal expectation), the impact of the shape uncertainty of the NR model is negligible compared to the uncertainty of its absolute rate. Thus, the NR model shape uncertainty is not propagated to the final likelihood.

B. Electronic recoil background model

ER events are one of the dominant background sources in the XENONnT SR0 WIMP search, primarily originating from β-decays of intrinsic radioactive contaminants, such as ^{214}Pb (a product of the ^{222}Rn-decay chain) and ^{85}Kr. Contributions from ^{136}Xe double-β-decays, solar neutrino interactions, and gamma events from the materials are also taken into account. In the ROI for the WIMP search, the total ER energy spectrum is approximately flat. The distribution of ER background events in cS1–cS2 space is generated from the ER model fitted to ^{220}Rn and ^{37}Ar calibration data, as discussed in Sec. II C. The total ER rate is fitted in the final WIMP likelihood, without applying an additional ancillary constraint in order to avoid potential systematic biases, which will be discussed in Sec. IV.

In contrast to the NR model, the uncertainty in the shape of the ER model can affect the sensitivity of the WIMP search. Ideally, the posterior distribution of all parameters should be propagated to the ER model as nuisance parameters in the WIMP search likelihood function. However, an excessive number of correlated nuisance parameters becomes computationally challenging. In XENONnT SR0, a principal component analysis (PCA)

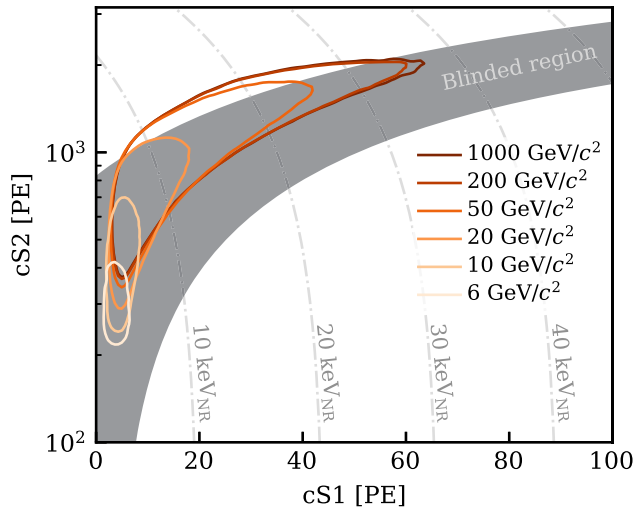

FIG. 5. Expected distribution in cS1–cS2 space from spin-independent WIMP nuclear recoil for WIMP masses from 6 GeV/c^2 to 1000 GeV/c^2. The contours contain 68% of expected events in the ROI. The blinded region, which is exclusively applied to the WIMP search dataset, is shown for reference as a gray band, and dash-dotted lines are constant NR-equivalent energy contours.

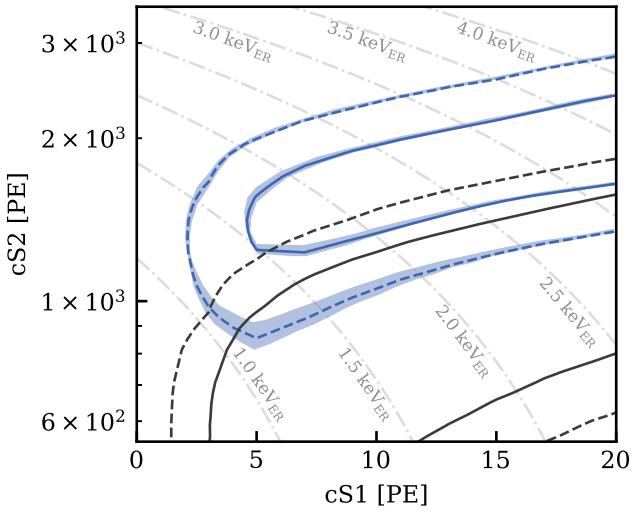

FIG. 6. Shape uncertainty of the ER model constrained by the ^{220}Rn calibration data. The blue (gray) solid and dashed lines represent the contours of the ER background (200 GeV/c^2 WIMP) containing 68% and 95% of expected events in the ROI. The transparent blue regions show the shape uncertainty by varying the two shape nuisance parameters by $\pm 1\sigma$. Dash-dotted lines are contours of constant ER-equivalent energy in units of keV$_{ER}$.

[50] is used to remove correlation among parameters from the MCMC sampler. All ER parameters shown in Table II, together with W, g_1, and g_2 (see details in the Appendix), are included in the PCA. Different from the original PCA, the variance of

$$t \equiv \sum_i \frac{s_i^2}{s_i + b_i} \tag{3}$$

along each principal component is used to quantify its importance. Here, s_i and b_i are the probability of a 50 GeV/c^2 WIMP and ER background in the ith bin of cS1–cS2 space, respectively. A larger variance of t means that the uncertainty of that component can be more impactful for the WIMP sensitivity. In the end, the two most important principal components are kept as nuisance parameters. These two shape parameters are constrained within the ^{220}Rn ancillary term of the WIMP search likelihood function, as shown in Fig. 6. While adding an additional ^{37}Ar ancillary term could further constrain the shape uncertainty at low energies, it was not included to maintain a conservative approach, as ^{220}Rn already provides a sufficiently tight constraint on the shape uncertainty.

C. Neutron background model

Radiogenic neutrons are mainly produced through spontaneous fission and (α, n) reactions in the detector materials due to intrinsic traces of radioactive impurities. The cosmogenic neutron background is subdominant [51], hence it is neglected in XENONnT SR0 analysis. The neutron yield and energy per isotope and material are calculated with the SOURCES-4A software [52], using the radioactive impurity levels of the relevant detector components obtained via the combined XENON1T and XENONnT radioassay campaigns [53,54]. A full-chain simulation pipeline [12] is used to estimate the neutron background rate, the geometrical distribution, and the cS1–cS2 distributions for SR0.

The propagation of the neutrons in the XENONnT detector is simulated with the Geant4 toolkit [55,56], where the recoil type and energy deposition per interaction in the target are recorded. We compute the number of photons and electrons for a given interaction via the custom-developed EPIX package [57], which utilizes the energy-dependent LXe response derived from the SR0 calibrations, shown in Fig. 3. This package also handles the clustering of the individual energy-depositing steps at the LXe microphysics scale before the quanta generation. The photons and electrons produced by EPIX are passed to the waveform simulator (WFSIM) [58], which computes the S1 and S2 signals up to the waveform level, by means of a precise set of simulation- and data-driven corrections that characterize the XENONnT detector response. The event-by-event simulated waveforms share the same data structure as the science data after applying PMT and Data Acquisition (DAQ) effects, which allows us to process them with the same software used for the real data (STRAXEN [59]), as well as to apply the same data selections. The final SS neutron rate arises from weighting the rates obtained via the entire waveform simulation pipeline with the specific activities of the corresponding material and isotopic neutron yield.

When a neutron scatters in the ∼250 kg of LXe between the cathode and the bottom PMT array, only the scintillation light for these events can be detected. The electrons, in turn, are lost due to an electric field pointing in the opposite direction to the active volume. Neutron events that consist of scatters above and below the cathode are referred to as neutron-X events, where "X" means additional S1 contribution from charge-insensitive scatters. They are modeled as a separate background since they have a larger cS1-to-cS2 ratio than normal neutron scatters.

The event parameters having discrimination power on MS interactions, such as the S2 pulse shape and PMT hit patterns of S1, are matched and validated with calibrations, such that the relevant data quality selections can be applied to the simulation outputs. Notably, a validation of the multiple-to-single-scatter ratio and total rate in the TPC of the AmBe calibration data was conducted, from where we obtained the systematic uncertainty associated with the accuracy of the full-chain simulations.

With this agreement between simulations and calibration data, we decided before unblinding the WIMP ROI to proceed with the sideband unblinding of the events tagged by the neutron veto. Initially, this confirmed the simulation-driven neutron background prediction. However, a mistake in the definition of the neutron veto time window was found after the unblinding of the WIMP ROI. After fixing this issue, a mismatch was found, with the neutron background rate being larger than predicted in the ROI [5]. We, therefore, decided to constrain the neutron background rate in a purely data-driven way based on the aforementioned sideband unblinding with the correct neutron veto tagging window. The results of the sideband unblinding are shown in Fig. 7. Based on the multiple-to-single-scatter ratio of 2.2, the 53% neutron veto tagging efficiency, and the three observed MS events and one SS neutron event tagged by the neutron veto in the fiducial volume, a data-driven prediction of $1.1^{+0.6}_{-0.5}$ events was derived for the SR0 exposure. We used the simulation-driven ratio between normal neutron background and neutron + X events to estimate the neutron + X event rate shown in Table I.

No further modification was propagated into the analysis after the data-simulations rate mismatch was identified: the 250 μs time veto window between the TPC and the neutron veto, chosen due to the reduced background rate initially predicted, and the fiducial volume of the WIMP search remained as defined prior to unblinding. An underestimated contamination from some of the surrounding materials is considered as a possible cause of the discrepancy between the expected rate and the observed neutron background. Studies to constrain the material's radiopurity by means of the high-energy gamma rays for specific detector regions are ongoing.

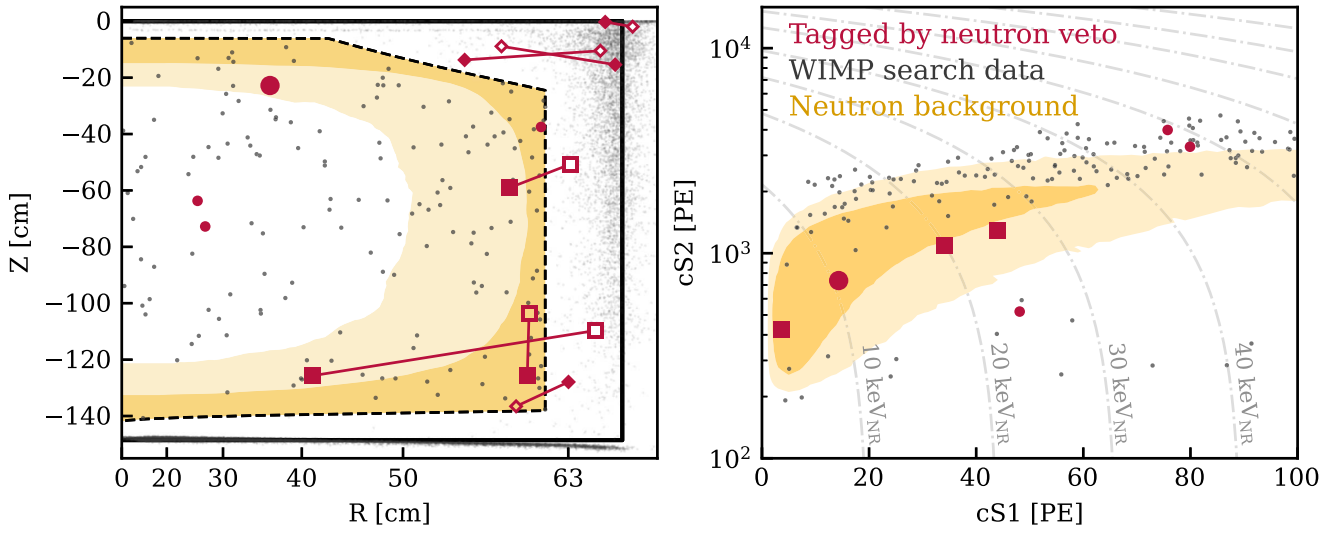

FIG. 7. Spatial (left) and cS1–cS2 (right) distribution of events tagged by the neutron veto (red markers). Black dashed and solid lines in the left plot indicate the FV and physical TPC boundary, respectively. The regions containing 68% (95%) of the total expected neutron background are indicated in dark (light) yellow. The neutron veto tagging yields four SS events (circles) and seven MS events, from which three are reconstructed with their largest S2 signals inside the FV (squares) and the other four outside (diamonds). Solid and hollow markers represent the reconstructed positions of the largest and the second-largest S2 signals, respectively. The right plot only shows events inside the FV and, from these, the four events within the neutron background band (large markers, one SS and three MS) were used to estimate the neutron background rate. The other three SS events are in agreement with the prediction of falsely tagged ER and surface background events (cf. Fig. 4). For reference, the events surviving all data selections, not tagged by the neutron veto, are displayed as gray dots. Also here, the right panel displays only those data points within the FV, which correspond to the unblinded WIMP search events.

TABLE I. Parameters of the XENONnT SR0 WIMP search likelihood function and their constraints. The right column shows best-fit values with an unconstrained 200 GeV/c^2, WIMP signal component (spin-independent coupling). For the rate parameters, all values are given in units of "events in the SR0 exposure." Near- and far-wire refer to the regions in the X–Y plane near and far from the perpendicular wires. In the near-wire region, which constitutes approximately 17% of the total FV, the AC expectation value is comparable to the one in the far-wire region due to the higher AC rate near the wires.

Rate Parameter	Constraint	Nominal	Best Fit
ER	WIMP search data	134.5	135^{+12}_{-11}
Neutron	Ancillary measurement	0.8 ± 0.4	0.8 ± 0.4
Neutron-X	Ancillary measurement	0.31 ± 0.16	0.30 ± 0.15
CEνNS (Solar ν)	Ancillary measurement	0.19 ± 0.06	0.19 ± 0.06
CEνNS (Atm + DSN)	Ancillary measurement	0.05 ± 0.02	0.04 ± 0.02
AC (near-wire)	Ancillary measurement	2.3 ± 0.7	2.3 ± 0.6
AC (far-wire)	Ancillary measurement	2.0 ± 0.6	2.1 ± 0.6
Surface	Ancillary measurement	14 ± 3	12 ± 2
^{220}Rn calibration	^{220}Rn dataset	2062 ± 210	2052 ± 44
Shape Parameter	**Constraint**	**Nominal**	**Best Fit**
ER shape parameter 1	^{220}Rn dataset	0.0	0.4 ± 0.2
ER shape parameter 2	^{220}Rn dataset	0.0	$-1.8^{+0.7}_{-0.6}$
Signal Parameter	**Constraint**	**Nominal**	**Best Fit**
Relative signal efficiency	NR model uncertainty	1.00 ± 0.09	1.0
WIMP cross section [10^{-47} cm^{-2}]		Parameter of interest	3.22

D. CEνNS background model

Neutrinos can also contribute to the NR background via coherent elastic neutrino-nucleus scattering (CEνNS). Signals from solar, atmospheric, and diffuse supernova neutrinos (DSN) will be in the WIMP search ROI. Due to the weak interaction between neutrinos and nuclei, CEνNS events are expected to be single-scatter and spatially uniform.

The recoil energy spectrum of solar CEνNS is almost identical to that of a 6 GeV/c^2 spin-independent WIMP [60], and the flux is $(5.25 \pm 0.20) \times 10^6$ cm^{-2} s^{-1} [61,62]. After applying all selections and their corresponding efficiencies, the expected number of events is (0.19 ± 0.06) in SR0, also shown in Table I. Atmospheric neutrinos and DSN mainly affect the search for heavier WIMPs. Their recoil energy spectra are taken from [51], and the SR0 expectation value is (0.05 ± 0.02) events. Due to the low cross section and the similarity to WIMP interactions, CEνNS background will be the major limitation to the WIMP search sensitivity of the next generation of LXe experiments.

E. Accidental coincidence background model

AC events consist of incorrectly paired S1 and S2 signals. These S1 and S2 signals can occur, for example, when either the S1 or the S2 signal of a physical event is not reconstructed due to detector effects, or when a single electron S2 signal is misclassified as an S1 signal. Such signals are referred to as "isolated" S1 and S2 signals. If an isolated S1 and an isolated S2 signal fall within the event-building time interval [9], they form an AC event.

The AC background is modeled with a data-driven approach. Isolated S1 and S2 signals, as well as their surrounding S1 and S2 signals that are close in time, are sampled and paired into events. S1 signals < 150 PE are selected as isolated S1 signals, and the isolated S2 signals are taken from events whose S1 area is < 150 PE, together with all pulses in the event window. The AC event rate is computed as

$$R_{\mathrm{AC}} = R_{\mathrm{isoS1}} \times R_{\mathrm{isoS2}} \times \Delta t, \tag{4}$$

where R_{isoS1} and R_{isoS2} are the isolated S1 and S2 rates, and Δt is the event-building time interval, which is defined according to the maximum drift time of 2.2 ms in XENONnT. Occasionally during SR0, some localized high rates of S2 signals appeared in the TPC. Excluding these periods from the analysis, R_{isoS1} and R_{isoS2} remained stable throughout SR0 at an average rate of 1.5 Hz and 80 mHz, respectively. In XENONnT, the isolated S2 rate is an order of magnitude higher compared to XENON1T [63]. This can be explained by the lower electron extraction efficiency, which causes an increased rate of delayed electrons. The isolated S1 and S2 signals are fed into the data processing pipeline [59,64] to reconstruct the events. This provides the background distribution of all relevant

event properties, especially the distribution in cS1 and cS2, as shown in Fig. 4 (lower right).

To suppress AC events, a selection criterion based on a gradient-boosted decision tree (GBDT) classifier utilizing the S2 pulse shape and the drift time was developed. Due to an insufficient model for the S2 pulse shape near the perpendicular wires, the GBDT classifier is only applied in the region at least 4.45 cm away from the wires (far-wire region), while in the region near the wires (near-wire region), a data-driven S2 shape selection is applied instead. The resulting AC background rate per ton-year in the near-wire region is about 5.6 times higher compared to the rate in the far-wire region. Consequently, the modeling of the TPC response is split into two parts.

Validation performed on calibration data provides the rate and shape uncertainty of the AC model. For this, we used dedicated AC-rich datasets, such as sidebands defined by inverting anti-AC selections. We compared the predicted and observed AC events by performing both 1D and 2D GOF tests with equiprobable bins in all relevant parameters. The AC predictions are provided by the AC modeling method discussed above. Because of the large statistics of ^{37}Ar calibration data, it delivers the most stringent constraint on the AC model. ^{37}Ar events with S2 areas smaller than 400 PE are selected to test the AC model. The predicted number of AC events was 731.6, while 733 events were observed in the data. The statistics in the AC model is very large so the uncertainty of the predicted AC rate is neglected in the following GOF tests. The 1D GOF tests in S1, S2, Z, R, and the 2D test in S1–S2 all yield p-values between 0.05 and 0.95. The result of the 2D test with a p-value of 0.61 is shown in Fig. 8. The model was further validated with events removed by selections targeting the AC background inside the WIMP ROI. Similar tests were also performed on ^{220}Rn calibration data. All these tests show the model and data to be compatible. Due to an overfitting issue found in the GBDT training process, we conservatively estimated the rate uncertainty to be 30% for both the near- and far-wire regions. The expectation value of AC events in the WIMP search dataset in the ROI is (4.3 ± 0.9).

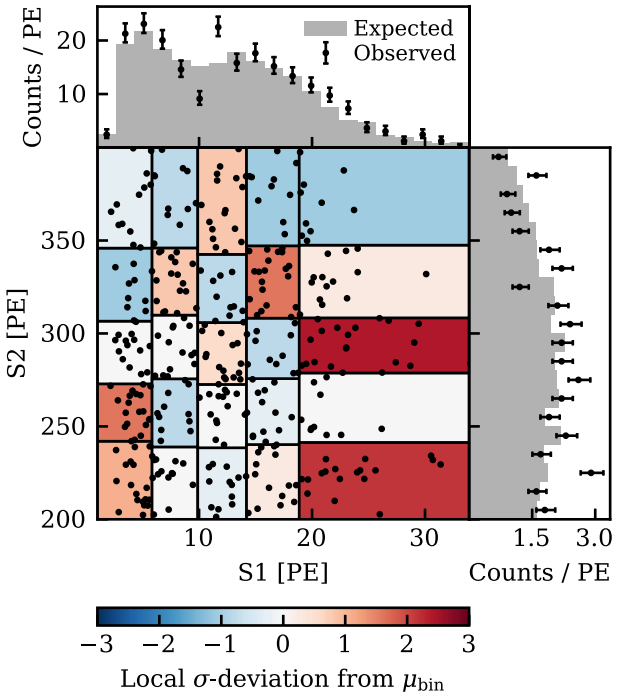

FIG. 8. AC model validation with ^{37}Ar calibration data in S2 versus S1 areas. The AC model was cross-checked against this high-statistics dataset using a 2D GOF test evaluated in an equiprobable binning scheme. The color scale shows the deviation from the predicted number of events in each bin $\mu_{\text{bin}} = 14.0$ in units of sigma. The projections to S1 and S2 are shown above and to the right of the 2D plot. No significant deviations were observed in the bulk of the distribution.

F. Surface background model

In XENONnT, the TPC wall is made of Polytetrafluoroethylene (PTFE), which is known to accumulate isotopes from the decay chain of atmospheric ^{222}Rn, which decays down to ^{210}Pb, an isotope with a half-life of 22.2 years. Radioactive decays on the wall surface can result in events with reduced S2 signals due to charge losses, which gives rise to a background that can leak into the WIMP signal region. Three components contribute to the surface background in the WIMP ROI: the two β-decays of ^{210}Pb and ^{210}Bi, and the recoiling ^{206}Pb following the α-decay of ^{210}Po. Independent screening measurements indicate ^{210}Po α-decay activity on the PTFE surface of (20 ± 3) mBq m^{-2} [54]. Due to the complexity of electric field conditions near the surface and the loss of ionization electrons to the detector walls, we employ a data-driven approach to model this background component in the space of cS1–cS2–R–Z.

The surface background model in the space of R–Z was developed using ^{210}Po α-events. These events are monoenergetic (5.4 MeV) and thus easily identifiable through their characteristic S1 signal, yet they present a wide range of S2 sizes due to variable charge loss to the walls. The R and Z distributions of ^{210}Po α-events are seen to match those of the lower-energy β-events. The radial profiles were modeled using a Student's generalized skew-t distribution. The radial distribution was fitted independently for different S2 sizes, as shown in Fig. 9, to account for the S2 size-dependent position resolution [11]. The Z profile was modeled using a linear function to account for lower rates of surface events observed at the bottom of the TPC, likely due to the increased charge insensitivity near the walls at the bottom of the detector [7].

The cS1–cS2 distribution was modeled using a 2D Gaussian adaptive kernel density estimation (aKDE), built using events reconstructed outside the detector walls. The resulting model was then validated against the events reconstructed between the walls and the fiducial volume,

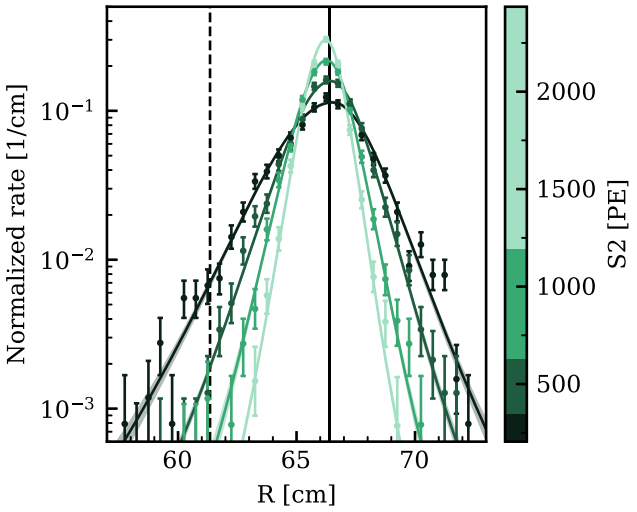

FIG. 9. The radial surface background model is shown for four slices in the S2 area. For each slice, a Student's generalized skew-t distribution is fitted to ^{210}Po α-events. The outer radius of the physical TPC as well as that of the fiducial volume are shown as a solid and a dashed black line, respectively.

in order to rule out any radial bias in the cS1–cS2 model. Fig. 4 (lower left) shows the projection of the four-dimensional model on cS1–cS2 in the ROI. The absolute rate of events in the blinded region was inferred from the radial distributions of adjacent, nonblinded regions in cS1–cS2.

Uncertainties in the model were obtained from the R and Z fit parameter uncertainties, as those are the parameters of interest in the development of the FV. Uncertainties on the overall measured surface event rate were also propagated. Uncertainties in the (cS1 and cS2) aKDE were neglected, as toy-MC tests were performed to show that they had little impact on the overall expectation. For the FV, a maximum radius of 61.35 cm compared to the 63 cm used in [65], was chosen to reduce the number of surface events to (14 ± 3), which improves the robustness against mismodeling.

IV. STATISTICAL INFERENCE

In this section, we describe the statistical methods used to derive the WIMP search results, which generally follow the recommendations formulated in [48]. We first describe the likelihood function used for the search in Sec. IV A and its nuisance parameters in Sec. IV B, followed by an illustration of the procedure for constructing confidence intervals and computing the discovery significance in Sec. IV C. The GOF test to validate the best-fit models and the blinding procedure are described in Sec. IV D. We omit some details already presented in [10], as the methods presented here build upon that previous work.

A. Definition of the likelihood function

The fundamental ingredient for the statistical analysis of our WIMP search data is the likelihood function $\mathcal{L}(\sigma, \boldsymbol{\theta})$. It depends on the WIMP-nucleon cross section $\sigma \geq 0$, which is our parameter of interest, and a set of nuisance parameters $\boldsymbol{\theta}$. The likelihood function is defined as a product of four terms: two for the WIMP search dataset ($\mathcal{L}_{\text{near-wire}}$ and $\mathcal{L}_{\text{far-wire}}$), one for the ^{220}Rn calibration dataset ($\mathcal{L}_{\text{cal}}$), and one for ancillary measurements of nuisance parameters ($\mathcal{L}_{\text{anc}}$),

$$\mathcal{L}(\sigma, \boldsymbol{\theta}) = \mathcal{L}_{\text{near-wire}}(\sigma, \boldsymbol{\theta}) \times \mathcal{L}_{\text{far-wire}}(\sigma, \boldsymbol{\theta}) \times \mathcal{L}_{\text{cal}}(\boldsymbol{\theta}) \times \mathcal{L}_{\text{anc}}(\boldsymbol{\theta}). \quad (5)$$

In the WIMP search data, we categorize events based on their proximity to the perpendicular wires of the gate and anode electrodes, distinguishing between those that are near (≤ 4.45 cm, corresponding to about 17% of the total FV) and far from the wires in the X-Y plane. This approach allows us to account for the higher AC rate observed near the wires, as discussed in Sec. III E, without introducing a full position-dependence in the likelihood. Other backgrounds, in particular radiogenic neutron and surface backgrounds, exhibit a substantial radial dependence. Thus, the likelihood is modeled and evaluated in R, in addition to cS1 and cS2. For the near-wire region, the radial component is not included in the modeling. Each WIMP search term is an extended unbinned likelihood function of the form

$$\mathcal{L}_{\text{region}}(\sigma, \boldsymbol{\theta}) = \text{Pois}(N|\mu_{\text{tot}}(\sigma, \boldsymbol{\theta})) \times \prod_{i=1}^{N} \left[\sum_c \frac{\mu_c(\sigma, \boldsymbol{\theta})}{\mu_{\text{tot}}(\sigma, \boldsymbol{\theta})} \times f_c(\vec{x}_i|\boldsymbol{\theta}) \right], \quad (6)$$

where the index "region" runs over "near-wire" and "far-wire." The index i runs over all N observed events $\vec{x}_i$ in the WIMP search far-wire (near-wire) dataset, where $\vec{x}_i$ is a vector with entries cS1, cS2, and R (cS1 and cS2). The PDFs f_c of each signal and background component c with expectation values μ_c are evaluated for each event. The total expectation value is given by $\mu_{\text{tot}}(\sigma, \boldsymbol{\theta}) = \sum_c \mu_c(\sigma, \boldsymbol{\theta})$.

The term $\mathcal{L}_{\text{cal}}(\boldsymbol{\theta})$ is the extended unbinned likelihood function of the ^{220}Rn calibration dataset. It depends on the two shape parameters introduced in Sec. III B that parametrize the range of models consistent with the ER calibration data, selected using the PCA method. By incorporating this likelihood term, we simultaneously fit the shape parameters to the ^{220}Rn calibration dataset and the ER background model in the WIMP search dataset. Through this procedure, the constraint from the calibration data on the shape of the ER model is propagated to the final inference.

Finally, the ancillary likelihood function $\mathcal{L}_{\text{anc}}(\boldsymbol{\theta})$ is a product of constraint terms for some of the nuisance parameters $\boldsymbol{\theta}$, which are detailed in the following section.

B. Nuisance parameters and constraints

In addition to our parameter of interest σ, we parametrize the space of background hypotheses using nuisance parameters $\boldsymbol{\theta}$ in the likelihood function of Eq. (5). These parameters control both background rates and shape. The relative signal efficiency is a special parameter that quantifies the rate uncertainty of the NR emission model, as described in Sec. III A. By construction, it takes the nominal value of 1.0 for an unconstrained fit, but it can depart from this value for fits with constrained cross sections, leading to a broadening of the confidence interval. In total, twelve nuisance parameters are defined, which are listed in Table I. The uncertainties of the nominal expectation values correspond to the width of a Gaussian constraint term. The ancillary likelihood function $\mathcal{L}_{\text{anc}}(\boldsymbol{\theta})$ is the product of all constraint terms. Most parameters are constrained via ancillary measurements, which are obtained from sidebands (e.g., data outside the ROI) or external measurements in combination with simulations. More details on the constraints on neutron, CEνNS, AC, and surface background rates were discussed in Sec. III. The ER rate is a free parameter in the fit and is fully constrained by the WIMP search data. The ER shape parameters obtained with PCA are constrained via the simultaneous fit of the ^{220}Rn calibration dataset via the term $\mathcal{L}_{\text{cal}}(\boldsymbol{\theta})$ in Eq. (5).

The best-fit values from the XENONnT SR0 WIMP search data [5] for a fit including an unconstrained 200 GeV/c^2 signal component are given in the last column of Table I. The corresponding uncertainties represent the two-sided 1-sigma confidence intervals derived from a profile likelihood scan in the respective parameter. More details on the construction of confidence intervals are discussed in the following section.

C. Confidence intervals and discovery significance

The construction of confidence intervals is based on the profile likelihood test statistic

$$q(\sigma) \equiv -2 \ln \frac{\mathcal{L}(\sigma, \hat{\hat{\boldsymbol{\theta}}})}{\mathcal{L}(\hat{\sigma}, \hat{\boldsymbol{\theta}})}. \tag{7}$$

Quantities with a single hat denote the global maximum likelihood estimator of a parameter, while $\hat{\hat{\boldsymbol{\theta}}}$ denotes the set of nuisance parameters that maximize the likelihood if the cross section is fixed to a given σ. The likelihood is defined for a specific signal model, for example, a WIMP with a certain mass and interaction type (e.g., spin-independent or spin-dependent coupling). In our statistical inference, we consider signal models across a range of WIMP masses, from 6 GeV/c^2 to 500 GeV/c^2. For higher WIMP masses, the PDF remains constant, and the flux for a given cross section decreases approximately linearly with the mass, as discussed in Sec. III A. According to this, limits can be extrapolated even beyond 500 GeV/c^2.

Knowing the distribution of $q(\sigma)$ under different hypotheses is essential for calculating discovery significances and confidence intervals. Due to the low rate of background resembling the signal, asymptotic formulae, as discussed in [66], are not necessarily a good approximation. Therefore, we primarily use toy-MCs to estimate the distribution $g(q(\sigma)|\sigma)$ of the test statistic given that σ is the true cross section. We use the custom-developed BLUEICE [67] framework to define the likelihood function. This Python package provides an efficient interpolation ("template morphing") between PDFs evaluated for different discrete values of shape nuisance parameters. This allows shape parameters to be considered for which we have no analytical description.

For each signal model considered, we compute the discovery significance and the confidence interval by comparing the measured test statistic $q(\sigma)$ with the distribution under each hypothesis $g(q(\sigma)|\sigma)$ [68,69]. Testing $\sigma = 0$ yields the (local) discovery significance, while confidence intervals are computed by finding the region where σ is rejected at a 90% CL given the data, which is illustrated for three background-only toy-MCs (three black parabola-like lines) in Fig. 10 (top). The minimum of each curve corresponds to the respective best-fit value $\hat{\sigma}$. Upper limits (ULs) and lower limits are obtained by finding the cross sections at which $q(\sigma)$ reaches the critical region. The threshold of the 90% CL critical region is precomputed as the 90th percentile of the test statistic distribution from toy-MCs with an injected signal corresponding to the respective cross section. For low cross sections, it deviates from the asymptotic two-sided threshold, indicated as a horizontal gray dotted line.

Repeating this procedure $\mathcal{O}(1000)$ times yields the expected distribution of ULs that can be obtained in the absence of any signal, as shown in Fig. 10 (bottom). The experiment's sensitivity is given by the median of these ULs, and the sensitivity band ("Brazil band") marked by the ± 2-sigma and ± 1-sigma quantiles indicates the spread. The distribution of lower limits for background-only toy-MCs is also shown in Fig. 10 (bottom). As recommended in [48], we decide on a discovery significance threshold for reporting two-sided confidence intervals of 3 sigma, following [10].

Figure 11 shows the 3-sigma discovery potential—the expected frequency of achieving a discovery with at least 3-sigma significance—as a function of WIMP mass and cross section. The asymptotic threshold for a 3-sigma discovery was applied. Contours indicating equal signal event counts in the nominal model reveal that the median 3-sigma discovery potential corresponds to approximately 10 events at high WIMP masses, decreasing to about six events at lower masses.

Statistical fluctuations as well as mismodeling, such as overestimated background rates, can yield arbitrarily low ULs, which may result in the spurious exclusion of models

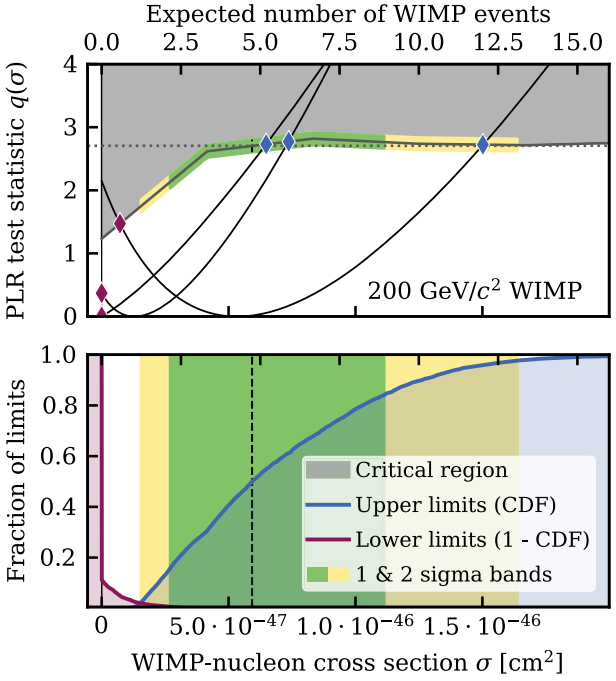

FIG. 10. Illustration of the confidence interval construction and distribution of limits on the WIMP-nucleon cross section σ of a 200 GeV/c^2 WIMP for background-only toy-MCs. Top: the test statistic $q(\sigma)$ as a function of the cross section σ is shown for three toy-MCs. The intersections with the threshold of the critical region (gray line) yield the 90% CL upper (blue diamonds) and lower limits (purple diamonds). Bottom: (inverse) cumulative distribution function of upper (lower) limits for background-only toy-MCs. The bands containing 68 % (green) and 95 % (yellow) of ULs, as well as the median UL (dashed line), are indicated. Note the shared abscissa with the cross section (bottom) and the corresponding expected number of signal events (top).

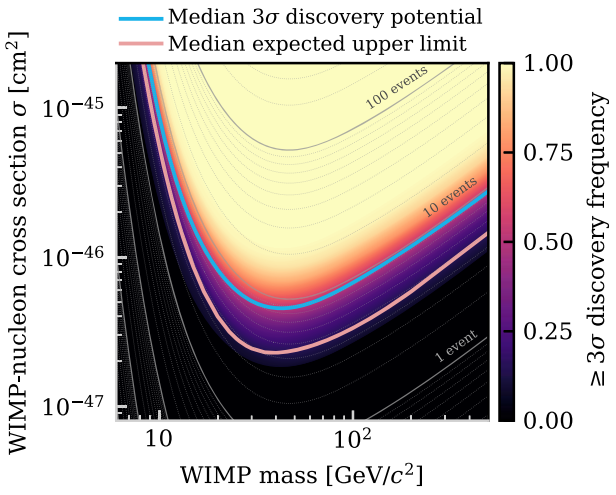

FIG. 11. Three-sigma discovery potential as a function of WIMP mass and cross section. The blue line marks the 50% discovery frequency, while the pink line represents the median expected upper limit from [5]. Thin contour lines indicate regions with an equal number of signal events.

beyond the experiment's sensitivity. To mitigate this issue, various methods have been proposed [70–72]. The XENONnT SR0 WIMP search follows the recommendation of [48] to use power-constrained limits (PCLs) [72]. In this method, a signal size threshold is selected, at which the experiment has a "rejection power" M_{min}—the probability of excluding a given signal under the background-only hypothesis. If an UL falls below this threshold, the threshold value is reported instead. These thresholds correspond to the quantiles of the UL distribution used to compute the sensitivity band illustrated in Fig. 10. For instance, choosing $M_{\text{min}} = 0.16$ sets the threshold to the -1-sigma quantile of the band, while $M_{\text{min}} = 0.5$ truncates ULs at the band's median. In [48], the fiducial choice of $M_{\text{min}} = 0.16$ was suggested. However, the choice was erroneously based on the discovery power, which corresponds to the probability of rejecting the background-only hypothesis given an alternative hypothesis with a specific signal size. In the absence of an updated recommendation, we chose the very conservative choice of $M_{\text{min}} = 0.5$ for the first WIMP search of XENONnT.

D. Goodness of fit and blinding procedure

To verify that the best-fit models describe our WIMP search data well, we defined a GOF test before unblinding and after studies to confirm the power to reject various forms of mismodeling. The test uses a binned Poisson likelihood χ^2 with 15 equiprobable bins in the cS1–cS2 space, defined from the model being tested. To compute a p-value, the distribution of the χ^2 test statistic under the best-fit hypothesis is derived through toy-MCs. Specifying the test and acceptance threshold (90% CL) before unblinding ensured that the results were not influenced by statistical fluctuations expected from the low-statistic WIMP search dataset. For the background-only fit, we found a p-value of 0.67, and for a fit with an additional 200 GeV/c^2 WIMP signal, the p-value is 0.63. Both indicate no strong mismodeling in the predefined parameter space. Using GOF tests with well-studied power to discover signal-like mismodeling was also designed to replace the "safeguard" ER shape parameter defined in [73], due to its computational cost and susceptibility to some kinds of mismodeling.

When searching for new signals in data, it is crucial to avoid the experimenter's bias on the result [74]. In our WIMP search, we adopt a common strategy for bias mitigation: blinding the signal region of the WIMP search dataset until all steps of the data analysis are finalized. Initially, events in the ER and NR bands of the WIMP search dataset with a reconstructed ER energy below 20 keV$_{\text{ER}}$ were blinded. This involved defining the ER and NR bands in cS1 and cS2 based on quantiles in cS2 for slices in cS1 of calibration data. In the first step, all events above 10 keV$_{\text{ER}}$ and those above the -2-sigma quantile of the ER band were unblinded for the analysis presented

in [65]. The signal region for the WIMP search (indicated in Fig. 5) remained blinded until the analysis procedure was finalized. The unblinded SR0 WIMP search data showed no significant excess for any of the tested WIMP masses with local discovery p-values ≥ 0.2. For this reason, we reported new ULs on the spin-independent WIMP-nucleon cross section across WIMP masses ranging from 6 GeV/c^2 to 500 GeV/c^2, with a minimum of 2.58×10^{-47} cm^2 at 28 GeV/c^2.

V. SUMMARY

The XENONnT WIMP dark matter search relies on a detector response model as well as simulation- and data-driven background models. These were combined to construct a statistical model in cS1, cS2, and R, which was used to infer limits on WIMP-nucleus scattering cross sections.

The full detector response model for ER and NR interactions, including both the xenon emission and detector reconstruction models, was successfully fitted to calibration data. Accurate simulations of particle interactions up to the data acquisition waveform level made this possible, in particular, to correctly model the S2 multiplicity of events with several, potentially unresolved energy deposits. Using these models, we derived the distributions for ER and NR backgrounds, as well as the signals, in our analysis space. Except for ER, the background rates were constrained with ancillary measurements. The radiogenic neutron background rate was constrained by first matching the simulated ratio of multiple-to-single-scatter interactions and the neutron veto tagging efficiency with NR calibration data. After unblinding the neutron veto-tagged events, these three inputs were combined to derive a prediction for the remaining nonvetoed single-scatter neutron background. Two shape parameters were propagated to the final inference to account for the uncertainty of the ER model in cS1–cS2 space. Accidental coincidence and the surface background were modeled with data-driven approaches. The validity of the models was confirmed in calibration data as well as science data outside the region of interest of the WIMP search.

A blind analysis was performed for the first science data from XENONnT. The statistical methods largely followed the previous XENON1T approach and community recommendations by using a toy-MC-calibrated profile log-likelihood ratio test statistic. One departure from these recommendations was raising the minimal required power of the PCL threshold from 0.15 to 0.5, corresponding to placing upper limits only at or above the median unconstrained upper limit. Both the calibration fits and the final best-fit model were assessed and found acceptable with goodness-of-fit tests that were chosen based on their mismodeling rejection power, and defined prior to unblinding the data. Analysis of the data with an exposure of 1.1 ton-years revealed no signal excess over backgrounds. Therefore, new upper limits on the spin-independent WIMP-nucleon cross section were derived.

ACKNOWLEDGMENTS

We gratefully acknowledge support from the National Science Foundation, Swiss National Science Foundation, German Ministry for Education and Research, Max Planck Gesellschaft, Deutsche Forschungsgemeinschaft, Helmholtz Association, Dutch Research Council (NWO), Fundacao para a Ciencia e Tecnologia, Weizmann Institute of Science, Binational Science Foundation, Région des Pays de la Loire, Knut and Alice Wallenberg Foundation, Kavli Foundation, JSPS Kakenhi, JST FOREST Program, and ERAN in Japan, Tsinghua University Initiative Scientific Research Program, Domaine d'intérêt majeur en astrophysique et conditions d'apparition de la vie+ (DIM-ACAV+) Région Ile-de-France, and Istituto Nazionale di Fisica Nucleare. This project has received funding/support from the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie Grant Agreement No. 860881-HIDDeN. We gratefully acknowledge support for providing computing and data-processing resources of the Open Science Pool and the European Grid Initiative, in the following computing centers: the CNRS/IN2P3 (Lyon, France), the Dutch national e-infrastructure with the support of Sanford Underground Research Facility (SURF) Cooperative, the Nikhef Data-Processing Facility (Amsterdam, Netherlands), the INFN-Centro Nazionale Analisi Fotogrammi (CNAF) (Bologna, Italy), the San Diego Supercomputer Center (San Diego, USA), and the Enrico Fermi Institute (Chicago, USA). We acknowledge the support of the Research Computing Center (RCC) at The University of Chicago for providing computing resources for data analysis. We are grateful to Laboratori Nazionali del Gran Sasso for hosting and supporting the XENON project.

APPENDIX A: DETAILS ON THE ER EMISSION MODEL

Let E be the recoil energy of an ER event. The total number of detectable quanta N_q is sampled from a normal distribution

$$N_q \sim \mathrm{Norm}(E/W, \sqrt{fE/W}), \tag{A1}$$

where W is the mean energy needed to generate a quantum, and f is the Fano factor of 0.059 [75,76]. The ER energy deposit will produce both excitons and electron-ion pairs. Using the mean ratio between the number of excitons and ions $\langle N_{ex}/N_i \rangle$, we simulate their numbers from a binomial distribution,

$$N_i \sim \mathrm{Binom}\left(N_q, \frac{1}{1+\langle N_{ex}/N_i \rangle}\right), \tag{A2}$$

$$N_{\mathrm{ex}} = N_{\mathrm{q}} - N_{\mathrm{i}}. \tag{A3}$$

A fraction r of ions recombine with electrons, depending on the electric drift field F. We parametrize the mean value of r in the same way as in XENON1T [10],

$$\langle r\rangle = \frac{1}{\mathrm{e}^{-(E-q_0)/q_1}+1}\left(1-\frac{\log(1+\langle N_{\mathrm{i}}\rangle\varsigma)}{\langle N_{\mathrm{i}}\rangle\varsigma}\right), \tag{A4}$$

where

$$\varsigma = \frac{1}{4}\gamma \mathrm{e}^{-E/\omega}F^{-\delta}. \tag{A5}$$

The fluctuation of r is parametrized via

$$\Delta r = q_2(1-\mathrm{e}^{-E/q_3}), \tag{A6}$$

and the true recombination fraction r is then sampled from

$$r \sim \mathrm{Norm}(\langle r\rangle, \Delta r). \tag{A7}$$

All fitted ER parameters are listed in Table II. Finally, the numbers of produced electrons and photons are given by

$$N_{\mathrm{e}} \sim \mathrm{Binom}(N_{\mathrm{i}}, 1-r), \tag{A8}$$

$$N_{\gamma} = N_{\mathrm{q}} - N_{\mathrm{e}}. \tag{A9}$$

APPENDIX B: DETAILS ON THE NR EMISSION MODEL

Let E be the recoil energy of an NR event. The total number of produced quanta N is

$$N \sim \mathrm{Norm}(E/W, \sqrt{fE/W}). \tag{B1}$$

In contrast to electron recoils, recoiling xenon nuclei lose a significant amount of their recoil energy via atomic motion and collisions with other xenon atoms, which are processes that are undetectable in a LXe TPC. This quanta loss is modeled following the Lindhard theory [77], with the so-called Lindhard quenching factor L, such that the number of detectable quanta becomes

$$N_{\mathrm{q}} \sim \mathrm{Binom}(N, L). \tag{B2}$$

Following the parametrization in [15], the Lindhard factor is given by

$$L = \frac{\kappa g(\epsilon)}{1+\kappa g(\epsilon)}, \tag{B3}$$

where $g(\epsilon)$ is a function of deposited energy via

$$g(\epsilon) = 3\epsilon^{0.15} + 0.7\epsilon^{0.6} + \epsilon, \tag{B4}$$

$$\epsilon = 11.5Z^{-7/3}\ (E/\mathrm{keV}), \tag{B5}$$

with the nuclear charge number $Z = 54$ for xenon. The numbers of produced ions and excitons are then simulated by

$$N_{\mathrm{i}} \sim \mathrm{Binom}\left(N_{\mathrm{q}}, \frac{1}{1+\langle N_{\mathrm{ex}}/N_{\mathrm{i}}\rangle}\right), \tag{B6}$$

$$N_{\mathrm{ex}} = N_{\mathrm{q}} - N_{\mathrm{i}}, \tag{B7}$$

with the exciton-to-ion ratio parametrized as

$$\langle N_{\mathrm{ex}}/N_{\mathrm{i}}\rangle = \alpha F^{-\zeta}(1-\mathrm{e}^{-\beta\epsilon}). \tag{B8}$$

Unlike ERs, the recombination fluctuation in NRs is usually small, thus the number of photons produced from the recombination of electron-ion pairs is

$$N_{\gamma}^{\mathrm{re}} \sim \mathrm{Binom}(N_{\mathrm{i}}, r), \tag{B9}$$

where r follows the Thomas-Imel box model [78]

$$r = 1 - \frac{\log(1+N_i\varsigma)}{N_i\varsigma}, \tag{B10}$$

$$\varsigma = \gamma F^{-\delta}. \tag{B11}$$

The number of excitons is further reduced by bi-excitonic and Penning quenching effects [79], such that the number of photons produced from de-excitation becomes

$$N_{\gamma}^{\mathrm{de}} \sim \mathrm{Binom}(N_{\mathrm{ex}}, f_l), \tag{B12}$$

with the scintillation quenching factor

$$f_l = \frac{1}{1+\eta\epsilon^{\lambda}}. \tag{B13}$$

Then, the final numbers of photons and electrons are

$$N_{\gamma} = N_{\gamma}^{\mathrm{de}} + N_{\gamma}^{\mathrm{re}}, \tag{B14}$$

$$N_{\mathrm{e}} = N_{\mathrm{i}} - N_{\gamma}^{\mathrm{re}}. \tag{B15}$$

APPENDIX C: DETAILS ON THE DETECTOR RESPONSE MODEL

The simulations of S1 and S2 signals from N_{γ} and N_{e} are almost independent of each other.

To simulate S1 signals, we first introduce the spatially dependent scintillation gain $\tilde{g}_1$ with DPE excluded,

$$\tilde{g}_1(\mathrm{X},\mathrm{Y},\mathrm{Z}) = \frac{g_1}{1+p_{\mathrm{DPE}}}\cdot \mathrm{S1corr}^{-1}(\mathrm{X},\mathrm{Y},\mathrm{Z}), \tag{C1}$$

where g_1 is the averaged scintillation gain, S1corr is the relative spatial correction factor, and p_{DPE} is the probability

of the double photoelectron emission effect [18]. Then, for a given number of photons N_γ and the position of the recoil X, Y, Z, the number of photons detected by the PMTs is

$$N_{\mathrm{PhD}} \sim \mathrm{Binom}(N_\gamma, \tilde{g}_1(\mathrm{X},\mathrm{Y},\mathrm{Z})). \tag{C2}$$

Now, accounting for the DPE effect, the expected number of PEs in the S1 signal is simulated via

$$N_{\mathrm{DPE}} \sim \mathrm{Binom}(N_{\mathrm{PhD}}, p_{\mathrm{DPE}}), \tag{C3}$$

$$N_{\mathrm{S1,PE}} \sim N_{\mathrm{PhD}} + N_{\mathrm{DPE}}. \tag{C4}$$

However, the S1 is not always equal to the number of PEs due to reconstruction bias. This is modeled as

$$\mathrm{S1}/N_{\mathrm{S1,PE}} - 1 \sim \mathrm{Norm}(\delta_{\mathrm{S1}}, \Delta_{\mathrm{S1}}). \tag{C5}$$

Here, both δ_{S1} and Δ_{S1} are functions of N_{PhD} obtained from the XENONnT MC. In the end, the relative spatial correction is applied back to the S1 signal, yielding

$$\mathrm{cS1} = \mathrm{S1} \cdot \mathrm{S1corr}(\mathrm{X}', \mathrm{Y}', \mathrm{Z}), \tag{C6}$$

where X′ and Y′ are the event positions, as reconstructed in the data, which are smeared by the S2 size-dependent resolution of the position reconstruction.

To get the S2 signal from N_{e}, the first step is to simulate the electron loss during the drift. The survival probability due to attachment to electronegative impurities in LXe is given by

$$p_{\mathrm{loss}}(\mathrm{Z}) = \mathrm{e}^{-\mathrm{Z}/(\tau v)}, \tag{C7}$$

where τ is the electron drift survival time ("electron lifetime") and v is the drift velocity of electrons in LXe. Inside the FV, the electric field is uniform enough to approximate v as a constant. Electron losses due to drift field effects close to the wall are accounted for via a spatially dependent charge-insensitive-volume (CIV) probability function $p_{\mathrm{CIV}}(\mathrm{R},\mathrm{Z})$ from field simulations of XENONnT [7]. The number of surviving electrons is then given by

$$N_{\mathrm{surv}} \sim \mathrm{Binom}(N_{\mathrm{e}}, p_{\mathrm{loss}}(\mathrm{Z}) \cdot p_{\mathrm{CIV}}(\mathrm{R},\mathrm{Z})). \tag{C8}$$

At the liquid-gas interface, a fraction of electrons are extracted,

$$N_{\mathrm{extr}} \sim \mathrm{Binom}(N_{\mathrm{surv}}, \epsilon(\mathrm{X},\mathrm{Y})). \tag{C9}$$

The extraction efficiency is X–Y-dependent and can be calculated by

$$\epsilon(\mathrm{X},\mathrm{Y}) = g_2 \cdot \mathrm{S2corr}^{-1}(\mathrm{X},\mathrm{Y})/G, \tag{C10}$$

where g_2 is the averaged ionization gain, S2corr is the relative spatial correction factor, and $G \sim 31$ PE/e$^-$ is the single electron gain. Assuming the fluctuation of the secondary scintillation process is Poisson-like, the expected number of S2 PEs is

$$N_{\mathrm{S2,PE}} \sim \mathrm{Norm}(N_{\mathrm{extr}}G, \sqrt{N_{\mathrm{extr}}G}). \tag{C11}$$

Similar to S1, the S2 reconstruction bias is modeled by

$$\mathrm{S2}/N_{\mathrm{S2,PE}} - 1 \sim \mathrm{Norm}(\delta_{\mathrm{S2}}, \Delta_{\mathrm{S2}}). \tag{C12}$$

Finally, the correction is applied to the S2,

$$\mathrm{cS2} = \mathrm{S2} \cdot \mathrm{S2corr}(\mathrm{X}', \mathrm{Y}') \cdot \mathrm{e}^{\mathrm{Z}/\tau' v}. \tag{C13}$$

Here, τ' is the electron lifetime extracted from the data, which is used to define the correction. Note that τ' can be slightly different from the true value τ due to small variations over time or nonuniformities along the drift path of electrons. The ratio τ'/τ is one of the parameters to fit in the band fitting. In principle, this effect can be absorbed into an additional smearing of S2, however in the fitting framework, it is still considered separately from others.

For both S1 and S2 signals, data quality selection efficiencies are applied, which depend on the respective signal sizes. For S1 signals, we additionally apply the S1 reconstruction efficiency, which is a function of the number of detected photons N_{PhD}.

For NR multiscatter events, the simulation is slightly altered. Here, the total number of detected photons is defined as the sum of the detected photons for each single energy deposition i,

$$N_{\mathrm{PhD}} = \sum_i N_{\mathrm{PhD}}^{(i)}. \tag{C14}$$

Similarly, we sum over the number of S2 PEs of only those scatters that contribute to the S2,

$$N_{\mathrm{S2,PE}} = \sum_{i_{\mathrm{S2}}} N_{\mathrm{S2,PE}}^{(i_{\mathrm{S2}})}. \tag{C15}$$

Reconstruction biases, signal corrections, and selection efficiencies are correspondingly applied to the merged signals.

APPENDIX D: DETAILS ON THE CALIBRATION FIT RESULTS

The parameters in the ER and NR band fits are summarized in Table II, showing both the applied prior and the marginal posterior values. Parameters where prior values are given without uncertainties are fixed in the fit. The prior for $\langle N_{\mathrm{ex}}/N_{\mathrm{i}}\rangle$ is a uniform distribution in the stated range. For the posteriors, the stated values and uncertainties correspond to the median and the central 68% of the marginal posterior distributions.

TABLE II. Prior and marginal posterior distributions of each parameter in the ER and NR emission models.

Parameter	Prior	Marginal posterior	Unit
W	13.7 ± 0.2	$13.7^{+0.2}_{-0.2}$	eV
f	0.059	0.059	
ER parameters			
$\langle N_{ex}/N_i \rangle$	0.06–0.20	$0.13^{+0.04}_{-0.04}$	
γ	Free	$0.13^{+0.03}_{-0.02}$	
δ	Free	$0.34^{+0.07}_{-0.07}$	
ω	Free	57^{+15}_{-12}	keV
q_0	Free	$1.32^{+0.17}_{-0.20}$	keV
q_1	Free	$0.47^{+0.07}_{-0.05}$	keV
q_2	Free	$0.030^{+0.002}_{-0.002}$	
q_3	Free	$0.47^{+0.40}_{-0.31}$	keV
NR parameters			
ζ	0.047	0.047	
δ	0.062	0.062	
α	Free	$0.92^{+0.07}_{-0.06}$	
γ	Free	$0.016^{+0.001}_{-0.001}$	
β	239^{+56}_{-18}	334^{+40}_{-43}	
κ	$0.139^{+0.006}_{-0.005}$	$0.138^{+0.005}_{-0.006}$	
η	$3.3^{+10.6}_{-1.4}$	$10.0^{+6.8}_{-5.9}$	
λ	$1.14^{+0.90}_{-0.18}$	$1.40^{+0.61}_{-0.38}$	

[1] G. Bertone and D. Hooper, History of dark matter, Rev. Mod. Phys. **90**, 045002 (2018).

[2] N. Aghanim *et al.* (Planck Collaboration), Planck 2018 results: VI. Cosmological parameters, Astron. Astrophys. **641**, A6 (2020).

[3] G. Bertone, D. Hooper, and J. Silk, Particle dark matter: Evidence, candidates and constraints, Phys. Rep. **405**, 279 (2005).

[4] L. Roszkowski, E. M. Sessolo, and S. Trojanowski, WIMP dark matter candidates and searches—current status and future prospects, Rep. Prog. Phys. **81**, 066201 (2018).

[5] E. Aprile *et al.* (XENON Collaboration), First dark matter search with nuclear recoils from the XENONnT experiment, Phys. Rev. Lett. **131**, 041003 (2023).

[6] E. Aprile *et al.* (XENON Collaboration), The XENONnT dark matter experiment, Eur. Phys. J. C **84**, 784 (2024).

[7] E. Aprile *et al.* (XENON Collaboration), Design and performance of the field cage for the XENONnT experiment, Eur. Phys. J. C **84**, 138 (2024).

[8] E. Aprile *et al.* (XENON Collaboration), The neutron veto of the XENONnT experiment: Results with demineralized water, arXiv:2412.05264.

[9] E. Aprile *et al.* (XENON Collaboration), XENON1T dark matter data analysis: Signal reconstruction, calibration, and event selection, Phys. Rev. D **100**, 052014 (2019).

[10] E. Aprile *et al.* (XENON Collaboration), XENON1T dark matter data analysis: Signal and background models and statistical inference, Phys. Rev. D **99**, 112009 (2019).

[11] E. Aprile *et al.* (XENON Collaboration), XENONnT analysis: Signal reconstruction, calibration and event selection, Phys. Rev. D **111**, 062006 (2025).

[12] D. Ramírez García, Ph.D. thesis, Albert-Ludwigs-Universität Freiburg, 2022, 10.6094/UNIFR/228338.

[13] T. Zhu, Ph.D. thesis, Columbia University, 2022, 10.7916/7y46-k318.

[14] M. Szydagis, N. Barry, K. Kazkaz, J. Mock, D. Stolp, M. Sweany, M. Tripathi, S. Uvarov, N. Walsh, and M. Woods, NEST: A comprehensive model for scintillation yield in liquid xenon, J. Instrum. **6**, P10002 (2011).

[15] B. Lenardo, K. Kazkaz, A. Manalaysay, J. Mock, M. Szydagis, and M. Tripathi, A global analysis of light and charge yields in liquid xenon, IEEE TNS **62**, 3387 (2015).

[16] J. Thomas and D. A. Imel, Recombination of electron-ion pairs in liquid argon and liquid xenon, Phys. Rev. A **36**, 614 (1987).

[17] B. López Paredes, H. Araújo, F. Froborg, N. Marangou, I. Olcina, T. Sumner, R. Taylor, A. Tomás, and A. Vacheret, Response of photomultiplier tubes to xenon scintillation light, Astropart. Phys. **102**, 56 (2018).

[18] C. Faham, V. Gehman, A. Currie, A. Dobi, P. Sorensen, and R. Gaitskell, Measurements of wavelength-dependent double photoelectron emission from single photons in VUV-sensitive photomultiplier tubes, J. Instrum. **10**, P09010 (2015).

[19] F. Jörg, S. Li, J. Schreiner, H. Simgen, and R. F. Lang, Characterization of a ^{220}Rn source for low-energy electronic recoil calibration of the XENONnT detector, J. Instrum. **18**, P11009 (2023).

[20] Y. Song, D. Jia, W. Zhang, and A. Hosaka (XENON Collaboration), Low-energy calibration of XENON1T with an internal ^{37}Ar source, Eur. Phys. J. C **83**, 1 (2023).

[21] D. Wenz, Ph.D. thesis, Johannes Gutenberg-Universität Mainz, 2023.

[22] D. Foreman-Mackey, D. W. Hogg, D. Lang, and J. Goodman, EMCEE: The MCMC hammer, Publ. Astron. Soc. Pac. **125**, 306 (2013).

[23] D. Foreman-Mackey *et al.*, EMCEE v3: A Python ensemble sampling toolkit for affine-invariant MCMC, J. Open Source Softwaare **4**, 1864 (2019).

[24] XENON Collaboration, XENONnT/GOFevaluation: v0.1.1, 10.5281/zenodo.6619120 (2022).

[25] E. Aprile (XENON Collaboration), Erratum: XENON1T dark matter data analysis: Signal and background models and statistical inference [Phys. Rev. D 99, 112009 (2019)], Phys. Rev. D **111**, 059904 (2025).

[26] L. Baudis, P. Sanchez-Lucas, and K. Thieme, A measurement of the mean electronic excitation energy of liquid xenon, Eur. Phys. J. C **81**, 1060 (2021).

[27] E. M. Boulton *et al.*, Calibration of a two-phase xenon time projection chamber with a ^{37}Ar source, J. Instrum. **12**, P08004 (2017).

[28] D. S. Akerib *et al.* (LUX Collaboration), Ultralow energy calibration of LUX detector using ^{127}Xe electron capture, Phys. Rev. D **96**, 112011 (2017).

[29] D. S. Akerib *et al.* (LUX Collaboration), Tritium calibration of the LUX dark matter experiment, Phys. Rev. D **93**, 072009 (2016).

[30] E. Aprile, K. L. Giboni, P. Majewski, K. Ni, M. Yamashita, R. Hasty, A. Manzur, and D. N. McKinsey, Scintillation response of liquid xenon to low energy nuclear recoils, Phys. Rev. D **72**, 072006 (2005).

[31] E. Aprile, C. E. Dahl, L. de Viveiros, R. J. Gaitskell, K. L. Giboni, J. Kwong, P. Majewski, K. Ni, T. Shutt, and M. Yamashita (XENON Collaboration), Simultaneous measurement of ionization and scintillation from nuclear recoils in liquid xenon as target for a dark matter experiment, Phys. Rev. Lett. **97**, 081302 (2006).

[32] E. Aprile, L. Baudis, B. Choi, K. L. Giboni, K. Lim, A. Manalaysay, M. E. Monzani, G. Plante, R. Santorelli, and M. Yamashita (XENON Collaboration), New measurement of the relative scintillation efficiency of xenon nuclear recoils below 10 keV, Phys. Rev. C **79**, 045807 (2009).

[33] E. Aprile *et al.* (XENON Collaboration), Response of the XENON100 dark matter detector to nuclear recoils, Phys. Rev. D **88**, 012006 (2013).

[34] G. Plante, E. Aprile, R. Budnik, B. Choi, K.-L. Giboni, L. W. Goetzke, R. F. Lang, K. E. Lim, and A. J. Melgarejo Fernandez, New measurement of the scintillation efficiency of low-energy nuclear recoils in liquid xenon, Phys. Rev. C **84**, 045805 (2011).

[35] P. Sorensen *et al.* (XENON Collaboration), The scintillation and ionization yield of liquid xenon for nuclear recoils, Nucl. Instrum. Methods Phys. Res., Sect. A **601**, 339 (2009).

[36] A. Manzur, A. Curioni, L. Kastens, D. N. McKinsey, K. Ni, and T. Wongjirad, Scintillation efficiency and ionization yield of liquid xenon for mono-energetic nuclear recoils down to 4 keV, Phys. Rev. C **81**, 025808 (2010).

[37] D. S. Akerib *et al.* (LUX Collaboration), Low-energy (0.7–74 keV) nuclear recoil calibration of the LUX dark matter experiment using D-D neutron scattering kinematics, arXiv:1608.05381.

[38] B. G. Lenardo *et al.*, Low-energy physics reach of xenon detectors for nuclear-recoil-based dark matter and neutrino experiments, Phys. Rev. Lett. **123**, 231106 (2019).

[39] D. S. Akerib *et al.* (LUX Collaboration), Nuclear recoil calibration at sub-keV energies in LUX and its impact on dark matter search sensitivity, Phys. Rev. Lett. **134**, 061002 (2025).

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

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

[42] 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).

[43] M. C. Smith *et al.*, The RAVE survey: Constraining the local galactic escape speed, Mon. Not. R. Astron. Soc. **379**, 755 (2007).

[44] C. McCabe, The Earth's velocity for direct detection experiments, J. Cosmol. Astropart. Phys. 02 (2014) 027.

[45] R. Schoenrich, J. Binney, and W. Dehnen, Local kinematics and the local standard of rest, Mon. Not. R. Astron. Soc. **403**, 1829 (2010).

[46] J. Bland-Hawthorn and O. Gerhard, The galaxy in context: Structural, kinematic, and integrated properties, Annu. Rev. Astron. Astrophys. **54**, 529 (2016).

[47] R. Abuter *et al.* (GRAVITY Collaboration), Improved gravity astrometric accuracy from modeling optical aberrations, Astron. Astrophys. **647**, A59 (2021).

[48] D. Baxter *et al.*, Recommended conventions for reporting results from direct dark matter searches, Eur. Phys. J. C **81**, 907 (2021).

[49] P. Klos, J. Menéndez, D. Gazit, and A. Schwenk, Large-scale nuclear structure calculations for spin-dependent wimp scattering with chiral effective field theory currents, Phys. Rev. D **88**, 083516 (2013).

[50] K. Pearson, LIII. On lines and planes of closest fit to systems of points in space, London Edinburgh Dublin Philos. Mag. J. Sci. **2**, 559 (1901).

[51] E. Aprile *et al.* (XENON Collaboration), Projected WIMP sensitivity of the XENONnT dark matter experiment, J. Cosmol. Astropart. Phys. 11 (**2020**) P031.

[52] D. G. Madland *et al.*, SOURCES4A: A code for calculating (alpha,n), spontaneous fission, and delayed neutron sources and spectra, 10.2172/15215 (1999).
[53] E. Aprile *et al.* (XENON Collaboration), Material radioassay and selection for the XENON1T dark matter experiment, Eur. Phys. J. C **77**, 890 (2017).
[54] E. Aprile *et al.* (XENON Collaboration), Material radiopurity control in the XENONnT experiment, Eur. Phys. J. C **82**, 599 (2022).
[55] S. Agostinelli *et al.* (Geant4 Collaboration), Geant4: A simulation toolkit, Nucl. Instrum. Methods Phys. Res., Sect. A **506**, 250 (2003).
[56] J. Allison *et al.* (Geant4 Collaboration), Recent developments in Geant4, Nucl. Instrum. Methods Phys. Res., Sect. A **835**, 186 (2016).
[57] XENON Collaboration, XENONnT/EPIX: v0.3.4, 10.5281/zenodo.7516942 (2023).
[58] XENON Collaboration, XENONnT/WFSIM: v1.0.2, 10.5281/zenodo.7216324 (2022).
[59] XENON Collaboration, XENONnT/STRAXEN: v1.8.3, 10.5281/zenodo.6854329 (2022).
[60] E. Aprile *et al.* (XENON Collaboration), Search for coherent elastic scattering of solar ^{8}B neutrinos in the XENON1T dark matter experiment, Phys. Rev. Lett. **126**, 091301 (2021).
[61] B. Aharmim *et al.* (SNO Collaboration), Combined analysis of all three phases of solar neutrino data from the Sudbury Neutrino Observatory, Phys. Rev. C **88**, 025501 (2013).
[62] M. Agostini *et al.* (BOREXINO Collaboration), Comprehensive measurement of pp-chain solar neutrinos, Nature (London) **562**, 505 (2018).
[63] E. Aprile *et al.* (XENON Collaboration), Dark matter search results from a one Tonne × Year exposure of XENON1T, Phys. Rev. Lett. **121**, 111302 (2018).
[64] J. Aalbers *et al.*, AxFoundation/STRAX: v1.2.3 (2022).
[65] E. Aprile *et al.* (XENON Collaboration), Search for new physics in electronic recoil data from XENONnT, Phys. Rev. Lett. **129**, 161805 (2022).
[66] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, Asymptotic formulae for likelihood-based tests of new physics, Eur. Phys. J. C **71**, 1554 (2011).
[67] J. Aalbers, K. D. Morå, and B. Pelssers, JelleAalbers/BLUEICE: v1.1.0 (2021).
[68] G. J. Feldman and R. D. Cousins, Unified approach to the classical statistical analysis of small signals, Phys. Rev. D **57**, 3873 (1998).
[69] C. Patrignani (Particle Data Group Collaboration), Review of particle physics, Chin. Phys. C **40**, 100001 (2016).
[70] A. L. Read, Presentation of search results: The CLs technique, J. Phys. G **28**, 2693 (2002).
[71] V. L. Kashyap, D. A. van Dyk, A. Connors, P. E. Freeman, A. Siemiginowska, J. Xu, and A. Zezas, On computing upper limits to source intensities, Astrophys. J. **719**, 900 (2010).
[72] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, Power-constrained limits, arXiv:1105.3166.
[73] N. Priel, L. Rauch, H. Landsman, A. Manfredini, and R. Budnik, A model independent safeguard against background mismodeling for statistical inference, J. Cosmol. Astropart. Phys. 05 (2017) 013.
[74] J. R. Klein and A. Roodman, Blind analysis in nuclear and particle physics, Annu. Rev. Nucl. Part. Sci. **55**, 141 (2005).
[75] E. Aprile and T. Doke, Liquid xenon detectors for particle physics and astrophysics, Rev. Mod. Phys. **82**, 2053 (2010).
[76] T. Doke, A. Hitachi, S. Kubota, A. Nakamoto, and T. Takahashi, Estimation of fano factors in liquid argon, krypton, xenon and xenon-doped liquid argon, Nucl. Instrum. Methods **134**, 353 (1976).
[77] J. Lindhard, M. Scharff, and H. E. Schioett, Range concepts and heavy ion ranges (Notes on atomic collisions, II.), Mat. Fys. Medd. Dan. Vid. Selsk. **33**, 1 (1963).
[78] J. Thomas and D. A. Imel, Recombination of electron-ion pairs in liquid argon and liquid xenon, Phys. Rev. A **36**, 614 (1987).
[79] D. M. Mei, Z. B. Yin, L. C. Stonehill, and A. Hime, A model of nuclear recoil scintillation efficiency in noble liquids, Astropart. Phys. **30**, 12 (2008).