

Offline tagging of radon-induced backgrounds in XENON1T and applicability to other liquid xenon time projection chambers

E. Aprile¹, J. Aalbers², K. Abe³, S. Ahmed Maouloud⁴, L. Althueser⁵, B. Andrieu⁴, E. Angelino⁶, J. R. Angevaere⁷, D. Antón Martín⁸, F. Arneodo⁹, L. Baudis¹⁰, A. L. Baxter¹¹, M. Bazyk¹², L. Bellagamba¹³, R. Biondi¹⁴, A. Bismark¹⁰, E. J. Brookes⁷, A. Brown¹⁵, G. Bruno¹², R. Budnik¹⁶, T. K. Bui³, J. M. R. Cardoso¹⁷, A. P. Cimental Chavez¹⁰, A. P. Colijn⁷, J. Conrad¹⁸, J. J. Cuenca-García¹⁰, V. D'Andrea^{19,*}, L. C. Daniel Garcia⁴, M. P. Decowski⁷, C. Di Donato²⁰, P. Di Gangi¹³, S. Diglio¹², K. Eitel²¹, A. Elykov²¹, A. D. Ferella^{20,19}, C. Ferrari¹⁹, H. Fischer¹⁵, T. Flehmke¹⁸, M. Flierman⁷, W. Fulgione^{6,19}, C. Fuselli⁷, P. Gaemers⁷, R. Gaior⁴, M. Galloway¹⁰, F. Gao²², S. Ghosh¹¹, R. Glade-Beucke¹⁵, L. Grandi⁸, J. Grigat¹⁵, H. Guan¹¹, M. Guida¹⁴, R. Hammann¹⁴, A. Higuera²³, C. Hils²⁴, L. Hoetsch¹⁴, N. F. Hood²⁵, M. Iacovacci²⁶, Y. Itow²⁷, J. Jakob⁵, F. Joerg¹⁴, A. Joy¹⁸, Y. Kaminaga³, M. Kara²¹, P. Kavragin¹⁶, S. Kazama²⁷, M. Kobayashi²⁷, A. Kopec^{25,†}, F. Kuger¹⁵, H. Landsman¹⁶, R. F. Lang¹¹, L. Levinson¹⁶, I. Li²³, S. Li²⁸, S. Liang²³, Y. T. Lin¹⁴, S. Lindemann¹⁵, M. Lindner¹⁴, K. Liu²², J. Loizeau¹², F. Lombardi²⁴, J. Long⁸, J. A. M. Lopes^{17,‡}, T. Luce¹⁵, Y. Ma²⁵, C. Macolino^{20,19}, J. Mahlstedt¹⁸, A. Mancuso¹³, L. Manenti⁹, F. Marignetti²⁶, T. Marrodán Undagoitia¹⁴, K. Martens³, J. Masbou¹², E. Masson⁴, S. Mastroianni²⁶, A. Melchiorre²⁰, M. Messina¹⁹, A. Michael⁵, K. Miuchi²⁹, A. Molinario⁶, S. Moriyama³, K. Morā¹, Y. Mosbacher¹⁶, M. Murra¹, J. Müller¹⁵, K. Ni²⁵, U. Oberlack²⁴, B. Paetsch¹⁶, J. Palacio¹⁴, Y. Pan⁴, Q. Pellegrini⁴, R. Peres¹⁰, C. Peters²³, J. Pienaar⁸, M. Pierre⁷, G. Plante¹, T. R. Pollmann⁷, L. Principe¹², J. Qi²⁵, J. Qin^{11,23,§}, D. Ramírez García¹⁰, M. Rajado¹⁰, J. Shi²², R. Singh¹¹, L. Sanchez²³, J. M. F. dos Santos¹⁷, I. Sarnoff⁹, G. Sartorelli¹³, J. Schreiner¹⁴, D. Schulte⁵, P. Schulte⁵, H. Schulze Eißing⁵, M. Schumann¹⁵, L. Scotto Lavina⁴, M. Selvi¹³, F. Semeria¹³, P. Shagin²⁴, S. Shi¹, M. Silva¹⁷, H. Simgen¹⁴, A. Takeda³, P.-L. Tan¹⁸, A. Terliuk^{14,||}, D. Thers¹², F. Toschi²¹, G. Trincherro⁶, C. Tunnell²³, F. Tönnies¹⁵, K. Valerius²¹, S. Vecchi³⁰, S. Vetter²¹, G. Volta¹⁰, C. Weinheimer⁵, M. Weiss¹⁶, D. Wenz²⁴, C. Wittweg¹⁰, T. Wolf¹⁴, V. H. S. Wu²¹, Y. Xing¹², D. Xu¹, Z. Xu¹, M. Yamashita³, L. Yang²⁵, J. Ye³¹, L. Yuan⁸, G. Zavattini³⁰, M. Zhong²⁵ and T. Zhu¹

(XENON Collaboration)[¶]

¹Physics Department, Columbia University, New York, New York 10027, USA

²Nikhef and the University of Groningen, Van Swinderen Institute, 9747AG Groningen, Netherlands

³Kamioka 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

⁴LPNHE, Sorbonne Université, CNRS/IN2P3, 75005 Paris, France

⁵Institut für Kernphysik, Westfälische Wilhelms-Universität Münster, 48149 Münster, Germany

⁶INAF-Astrophysical Observatory of Torino, Department of Physics, University of Torino and INFN-Torino, 10125 Torino, Italy

⁷Nikhef and the University of Amsterdam, Science Park, 1098XG Amsterdam, Netherlands

⁸Department of Physics & Kavli Institute for Cosmological Physics, University of Chicago, Chicago, Illinois 60637, USA

⁹New York University Abu Dhabi - Center for Astro, Particle and Planetary Physics, Abu Dhabi, United Arab Emirates

¹⁰Physik-Institut, University of Zürich, 8057 Zürich, Switzerland

¹¹Department of Physics and Astronomy, Purdue University, West Lafayette, Indiana 47907, USA

¹²SUBATECH, IMT Atlantique, CNRS/IN2P3, Université de Nantes, Nantes 44307, France

¹³Department of Physics and Astronomy, University of Bologna and INFN-Bologna, 40126 Bologna, Italy

¹⁴Max-Planck-Institut für Kernphysik, 69117 Heidelberg, Germany

¹⁵Physikalisches Institut, Universität Freiburg, 79104 Freiburg, Germany

¹⁶Department of Particle Physics and Astrophysics, Weizmann Institute of Science, Rehovot 7610001, Israel

¹⁷LIBPhys, Department of Physics, University of Coimbra, 3004-516 Coimbra, Portugal

¹⁸Oskar Klein Centre, Department of Physics, Stockholm University, AlbaNova, Stockholm SE-10691, Sweden

¹⁹INFN-Laboratori Nazionali del Gran Sasso and Gran Sasso Science Institute, 67100 L'Aquila, Italy

²⁰Department of Physics and Chemistry, University of L'Aquila, 67100 L'Aquila, Italy

²¹*Institute for Astroparticle Physics, Karlsruhe Institute of Technology, 76021 Karlsruhe, Germany*

²²*Department of Physics & Center for High Energy Physics, Tsinghua University, Beijing 100084, People's Republic of China*

²³*Department of Physics and Astronomy, Rice University, Houston, Texas 77005, USA*

²⁴*Institut für Physik & Exzellenzcluster PRISMA⁺, Johannes Gutenberg-Universität Mainz, 55099 Mainz, Germany*

²⁵*Department of Physics, University of California San Diego, La Jolla, California 92093, USA*

²⁶*Department of Physics “Ettore Pancini”, University of Napoli and INFN-Napoli, 80126 Napoli, Italy*

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, People's Republic of 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 25 March 2024; accepted 4 June 2024; published 15 July 2024)

This paper details the first application of a software tagging algorithm to reduce radon-induced backgrounds in liquid noble element time projection chambers, such as XENON1T and XENONnT. The convection velocity field in XENON1T was mapped out using ^{222}Rn and ^{218}Po events, and the rms convection speed was measured to be 0.30 ± 0.01 cm/s. Given this velocity field, ^{214}Pb background events can be tagged when they are followed by ^{214}Bi and ^{214}Po decays, or preceded by ^{218}Po decays. This was achieved by evolving a point cloud in the direction of a measured convection velocity field, and searching for ^{214}Bi and ^{214}Po decays or ^{218}Po decays within a volume defined by the point cloud. In XENON1T, this tagging system achieved a ^{214}Pb background reduction of $6.2_{-0.9}^{+0.4}\%$ with an exposure loss of $1.8 \pm 0.2\%$, despite the timescales of convection being smaller than the relevant decay times. We show that the performance can be improved in XENONnT, and that the performance of such a software-tagging approach can be expected to be further improved in a diffusion-limited scenario. Finally, a similar method might be useful to tag the cosmogenic ^{137}Xe background, which is relevant to the search for neutrinoless double-beta decay.

DOI: [10.1103/PhysRevD.110.012011](https://doi.org/10.1103/PhysRevD.110.012011)

I. INTRODUCTION

Liquid xenon time projection chambers (TPCs) such as XENON1T [1], XENONnT [2], and LZ [3] are constructed with the primary goal of searching for dark matter in the form of weakly interacting massive particles (WIMPs) [1,4]. These TPCs as well as the dedicated EXO-200 TPC also search for neutrinoless double-beta decay ($0\nu\beta\beta$) [5–7].

Other physics channels include measurements of double electron capture in ^{124}Xe [8], solar axions, nonstandard neutrino interactions, and bosonic dark matter [9,10].

Achieving low levels of radioactive backgrounds is critical to the aforementioned physics channels because the sensitivity scales as $\text{signal}/\sqrt{\text{background}}$ (see Sec. III C). The decay chain of ^{222}Rn includes ^{214}Pb , an isotope that undergoes beta decay. At low energies, this is a major source of backgrounds in the electronic-recoil (ER) channel of xenon-based dark matter experiments, and is also important to the nuclear-recoil (NR) channel due to imperfect ER/NR discrimination [1–3,11,12]. The reason ^{222}Rn contamination is a major source of backgrounds is because ^{222}Rn is produced from the emanation of ^{226}Ra , which is present at low levels in almost all materials [13]. In addition, ^{222}Rn is miscible with xenon, and the half-life of $t_{1/2} = 3.8$ days [14] allows it to move throughout the detector. The isotope in the ^{222}Rn decay chain that decays to produce the relevant low-energy background is ^{214}Pb (see Fig. 1). Because of this, substantial efforts have been made to reduce the radon level using dedicated hardware solutions [13,15–22]. The

*Also at INFN-Roma Tre, 00146 Rome, Italy.

†Present address: Department of Physics and Astronomy, Bucknell University, Lewisburg, Pennsylvania, USA.

‡Also at Coimbra Polytechnic—ISEC, 3030-199 Coimbra, Portugal.

§Contact author: jq8@rice.edu

¶Also at Physikalisches Institut, Universität Heidelberg, Heidelberg, Germany.

¶¶Contact author: xenon@lngs.infn.it

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

described earlier in [section I A](#).

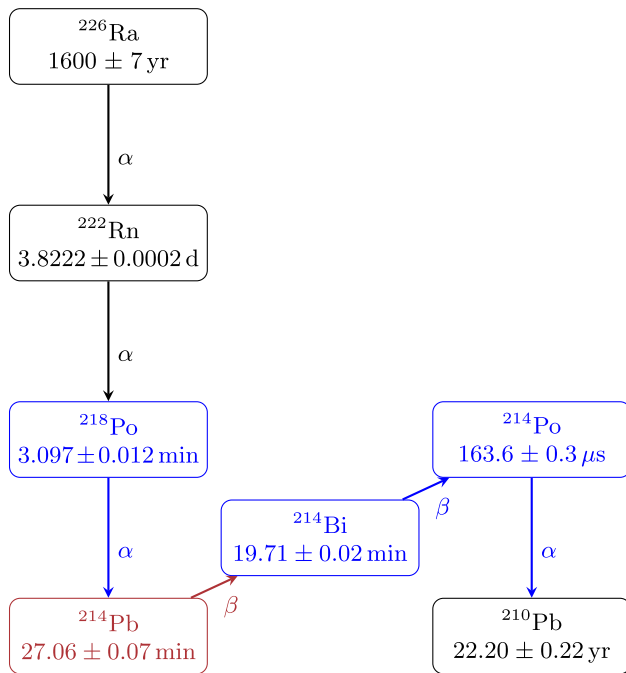


FIG. 1. Decay chain of ^{222}Rn , part of the uranium series. Only branches with branching fraction above 99.5% are shown. Data retrieved using the NNDC ENSDF, with the following nuclear data sheets citations; [14,24–27]. The isotope that decays to produce the background events being tagged in this work, ^{214}Pb , is colored red, whereas the isotopes with decays that are used for the tagging of ^{214}Pb are colored blue.

software-based approach introduced in this paper performs better at lower background levels, thus complementing these hardware-based methods to further reduce radon-induced backgrounds using offline analysis. Such a method can also be used to suppress radon chain backgrounds in liquid argon TPCs, where hardware-based approaches for the mitigation of radon-chain backgrounds are similarly being pursued [23]. The key challenge in a software-based approach is that the convection timescale is $\sim \frac{100 \text{ cm}}{0.3 \text{ cm/s}} = 300 \text{ s}$, which is significantly shorter than the decay times of both ^{214}Pb and ^{214}Bi ; these are 27 min and 20 min, respectively. This implies that over one decay half-life, the radioactive isotopes move substantially in the TPC, ruling out naive approaches that simply veto spherical volumes around ^{218}Po events.

As such, efforts to tag ^{214}Pb events in XENON1T based on other events in the same decay chain require a measurement of the convection velocity field. The measurement of convection in XENON1T and properties of the velocity field are detailed in [Sec. II](#). The algorithm to track isotopes along the measured velocity field and thus veto ^{214}Pb events is detailed in [Sec. III](#). Results, including demonstrations of the technique on XENON1T data,

projections to XENONnT and future liquid xenon detectors, and application to cosmogenic ^{137}Xe which is a background for the search for neutrinoless double-beta decay ($0\nu\beta\beta$), are discussed in [Sec. IV](#). Finally, the conclusion is presented in [Sec. V](#).

A. XENON1T and XENONnT

The XENON1T experiment used a cylindrical dual-phase TPC, with a diameter of 96 cm and a 2.0 tons active liquid xenon target [1]. Two arrays of photomultiplier tubes (PMTs) were installed in the top and bottom of the TPC, respectively. The vertical sides of the TPC are constructed out of UV-reflective polytetrafluoroethylene to increase the light collection efficiency. A cathode at the bottom of the TPC and a gate electrode 97 cm above the cathode produced a drift field of 81 V/cm. The anode was 5 mm above the gate and ~ 2.5 mm above the liquid-gas interface, and with the gate produced an extraction field of 8.1 kV/cm. The cryostat containing the TPC was positioned inside a 740 m³ water Cherenkov muon veto which allowed for the active detection and veto of muons and muon-induced backgrounds.

XENONnT is an in-place upgrade of XENON1T that uses much of the existing infrastructure. It features an enlarged TPC, a novel neutron veto system, and various improvements in the xenon handling system, which allow for an improved xenon purity. The TPC diameter is 1.3 m, and the separation between gate and cathode extended to 1.5 m. This results in an active liquid xenon target mass of 5.9 tons [2]. The muon veto uses the same design as in XENON1T. However, there is now a neutron veto surrounding the main cryostat containing the TPC, which is optically separated from the muon veto. The neutron veto aims to reduce the radiogenic neutron background by detecting neutrons which scatter in the TPC volume and are then captured in the neutron veto [2].

Events in the TPC are detected via two signals. First, a prompt scintillation signal (S1) comprised of 175 nm photons [28] is produced at the site of a particle interaction due to the decay of excited atoms [29]. Additionally, ionization electrons are produced; some of these recombine with xenon ions, contributing to the S1 as well [29]. These photons are detected by the PMTs. Ionization electrons that do not recombine are drifted towards the liquid-gas interface by the drift field, and extracted into the xenon gas by the extraction field, where a secondary scintillation signal (S2) is produced. The size of these S1 and S2 signals are measured in units of photoelectrons (PE).

The 3D position of the interaction can be reconstructed using the hit pattern of the S2 signal on the top PMT array (x, y) and the drift time between the S1 and S2 signals (z). Details regarding position reconstruction methods in XENON1T can be found in [30]. The reconstructed position also allows for the computation of corrected S1

and S2 signal sizes, based on position-dependent signal efficiencies [30].

B. The ^{222}Rn decay chain

The decay chain of ^{222}Rn is shown in Fig. 1. ^{214}Pb is responsible for the low-energy ER background. This is because the beta spectrum of ^{214}Pb extends to low energies and is flat to the percent-level below 50 keV [9]. Alpha decays, on the other hand, are mono-energetic and have a different S1/S2 ratio from ER or NR events [31,32], and hence are easy to select. This differing S1/S2 ratio is related to recombination of electrons and ions, as described earlier in Sec. I A.

^{214}Bi decay does not represent an important background because it is quickly followed by ^{214}Po , an isotope that undergoes alpha decay with a half-life of 164 μs [26]. Because this is much shorter than the drift time in XENON1T and XENONnT, the ^{214}Bi and ^{214}Po events are combined into a single event with two S1s, and two or more S2s. This is a unique event topology that is easy to select. Such events are termed ‘‘BiPo’’ events in this paper.

II. MEASURING CONVECTION IN THE XENON1T DETECTOR

A. Mapping the convection velocity field

Convection has been observed in earlier dual phase liquid xenon TPCs, such as XENON100 and LUX [33,34]. While the exact boundary conditions driving the convection are not known, the convective flow is likely driven by the thermal flux into the TPC, possibly from both recirculation flows and from the cryostat. The relevant temperature gradient might be either horizontal or vertical.

To measure the convective flow in the XENON1T detector, ^{222}Rn and ^{218}Po events were used. ^{222}Rn and ^{218}Po undergo alpha decay. The 3 min half-life of ^{218}Po is short enough that the ^{222}Rn and ^{218}Po events can be paired up, but long enough that there can be significant displacement between pairs of events at ~ 0.1 cm/s speeds. As such, the decays of ^{222}Rn and ^{218}Po are particularly suited to the measurement of convection.

These events were selected using Gaussian mixture clustering [35] using the position-corrected S1, position-corrected S2, width of the S2 peak in nanoseconds, radial coordinate, and z -coordinate of each event [36].

After event selection, ^{222}Rn and ^{218}Po events have to be paired to construct velocity vectors corresponding to the convective flow. However, the rate of ^{222}Rn decays exceeds 10 $\mu\text{Bq/kg}$ [37], corresponding to approximately two ^{222}Rn events every 3 min in a 1 tons fiducial mass. As the half-life of ^{218}Po is ~ 3 min, the pairing of ^{222}Rn and ^{218}Po events cannot be done in a naive manner where every ^{218}Po event is considered to be the daughter of the preceding ^{222}Rn event. Instead, for each pair of ^{222}Rn

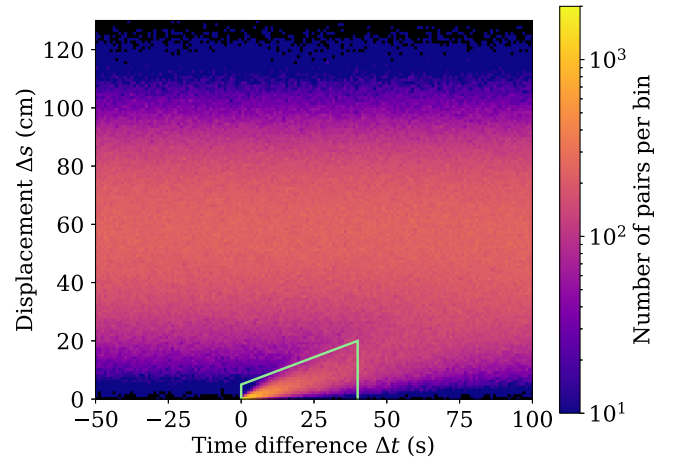


FIG. 2. 2D histogram of the time difference (Δt) and displacement (Δs) of all permutations of ^{222}Rn and ^{218}Po pairs. The excess of pairs where $\Delta s < 20$ cm and $0 \text{ s} < \Delta t < 40$ s, is from correctly-paired events, and is highlighted with the light green box. At negative times, the pairs are unphysical and can be used to profile the distribution of incorrect pairs.

and ^{218}Po events, the time difference (Δt) and displacement (Δs) were plotted on a histogram, see Fig. 2. An excess of pairs where $\Delta s < 20$ cm and $0 \text{ s} < \Delta t < 40$ s becomes apparent. This excess is due to correctly-paired events. In addition, the distribution of ^{222}Rn and ^{218}Po events that are not correctly paired is independent of Δt , and can be determined using pairs where $\Delta t < 0$. The observed distribution of these incorrect pairs is largely due to the TPC geometry, and the maxima at ~ 60 cm is what one would expect from the pairwise distances of two uniform random distributions in the TPC. One can then compute the purity of each histogram bin as $f_{\text{pure}} = 1 - N_{bg}/N_{\text{bin}}$, where N_{bg} is the number of incorrect pairs in a bin at the given Δt estimated using negative-time bins, and N_{bin} is the total number pairs in a bin.

All permutations of ^{222}Rn and ^{218}Po pairs with $0 \text{ s} < \Delta t < 100$ s were used. We iteratively selected the bin with highest f_{pure} , and then removed all pairs which contain one of these ^{222}Rn or ^{218}Po events. This is done because as one iteratively removes events that belong to selected pairs, the total number of events remaining in the pool decreases, thus decreasing the number of incorrectly-matched pairs remaining. This is run for 1000 iterations.

Vectors were constructed from selected pairs of events by computing the velocity from the Δs and Δt values of the pair. The velocity field obtained using the iterative method is shown in Fig. 3. It can be seen that this velocity field is still noisy, and contains outliers that likely correspond to incorrectly paired events.

It is notable that there appears to be a single large convection cell, which is expected for convection cells in a cylinder with an aspect ratio of close to 1 [38].

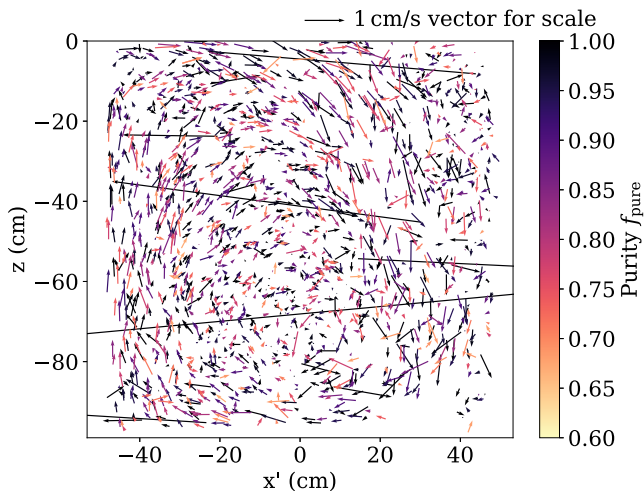


FIG. 3. Slice of velocity field with 107804 vectors showing the convection cell. The x' coordinate is perpendicular to the angular momentum vector. A 1 cm/s velocity vector is shown in the top right for scale. The purity of a vector, f_{pure} , is indicated by the color.

B. Filtering of the velocity field

The convection field, of which a slice is shown in Fig. 3, was then filtered and discretized onto a grid. The purpose of this was to reduce noise and to speed up computation, as finding the nearest velocity vector to a given position is much faster with data on a regular grid. First, every vector of purity f_{pure} was oversampled $25 \times f_{\text{pure}}$ times, rounded to the nearest integer. The value of 25 was chosen to avoid significant computational cost. The $x - y$ position reconstruction uncertainty of α events was estimated to be $\sigma_x = 0.3$ cm, $\sigma_y = 0.3$ cm, based on the spread observed in ^{210}Pb decays on the TPC surface [39], and the z -position uncertainty was estimated to be $\sigma_z = 0.17$ cm from the displacement of the two decays in BiPo events. Diffusion is not considered as it is much smaller than the position

resolution for the relevant timescales of ^{218}Po decay; this can be seen from Sec. IV C. During oversampling, each vector was perturbed randomly based on the position reconstruction uncertainty. The oversampled population of vectors were then put onto a grid with a grid spacing of 1/3 cm, by computing the geometric median [40] of the nearest 175 vectors at every grid point, as defined by the midpoint of the vectors. The geometric median has been shown to be particularly robust for noisy datasets [41]. The result of this procedure is shown in Fig. 4.

C. Root-mean-square convection speed

The convection vectors obtained in Sec. II A allow a measurement of the bulk convection properties. To avoid biases due to uneven event densities, the detector was divided into 11 bins in $r^2 \in [0, 47.9^2]$ cm², 10 bins in azimuth $\phi \in [-\pi, \pi]$, and 9 bins in $z \in [-96.9, 0]$ cm. Every vector was then assigned to a bin, and given a weight w_i equal to the reciprocal of the number of vectors in that bin. This procedure allows for the computation of a volume-averaged rms speed,

$$v_{\text{rms}} = \sqrt{\frac{\sum_{i=1}^N v_i^2 w_i}{\sum_{i=1}^N w_i}}. \quad (1)$$

The uncertainty on each velocity vector can be estimated using the position reconstruction uncertainty as $\sigma_i = \sqrt{2 \frac{\sigma_x^2 + \sigma_y^2 + \sigma_z^2}{\Delta_i^2}}$. The total uncertainty is then given by

$$\sigma_{\text{rms}} = \sqrt{\frac{\sum_{i=1}^N v_i^2 \sigma_i^2 w_i}{\sum_{i=1}^N w_i}}. \quad (2)$$

The root-mean-square speed was thus found to be 0.30 ± 0.01 cm/s. This is significantly slower than what was observed in XENON100 and LUX [33,34]; however, it

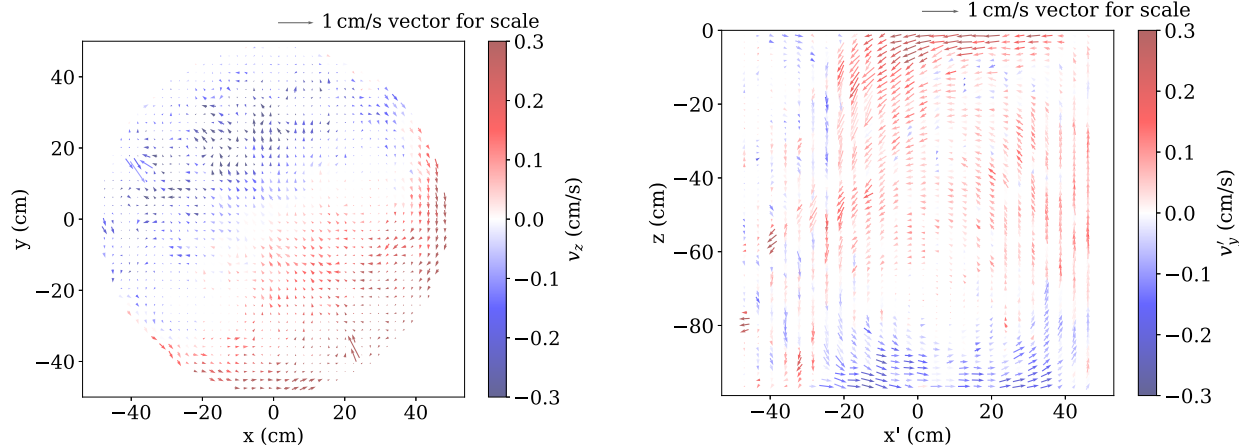


FIG. 4. Slice of the velocity field after it was filtered and put discretized onto a grid, shown from the top view (left) and the side view (right). In the left plot, the x' coordinate is defined as in Fig. 3; in the right plot, a slice at $z = -50$ cm is taken. For clarity, only every sixth vector is displayed. A 1 cm/s velocity vector is shown in the top right for scale.

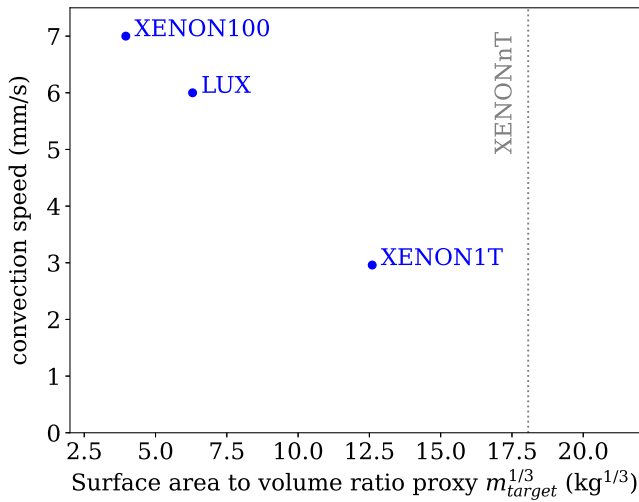


FIG. 5. Convection speed versus $(\text{target mass})^{1/3}$ for XENON100 [33], LUX [34], and XENON1T (this work). $(\text{target mass})^{1/3}$ is used as a proxy for the surface area to volume ratio. It can be seen that convection speed decreases linearly. EXO-200 and LZ are excluded from this plot as convection was not observed in EXO-200 [42], and was found to be subdominant to the drift of charged ions in LZ [43]. The target mass of XENONnT is indicated in gray as convection in XENONnT has not yet been analyzed in detail.

is significantly higher than EXO-200 and LZ where convection is subdominant to the mobility of charged ions [42,43]. This is shown in Fig. 5, plotted against the cube root of the target mass, which is a proxy for linear dimension.

The heat flux into a TPC is likely proportional to the surface area, which is the square of linear dimension; when distributed over the entire target mass, which is related to the cube of linear dimension, one expects convection speed to vary linearly with linear dimension. It can be seen that convection speed decreases linearly with larger detectors.

It is notable that the convection velocity implies a convection timescale of $\sim \frac{100 \text{ cm}}{0.3 \text{ cm/s}} = 300 \text{ s}$, which is significantly smaller than the decay time of both ^{214}Pb and ^{214}Bi , underscoring the difficulty of a software veto.

III. ^{214}Pb VETO ALGORITHM

The veto algorithm first starts from an ER event that is in the energy range to be due to ^{214}Pb decay, termed the ^{214}Pb candidate. A normally distributed set of points centered around the location of this event is generated, and propagated forwards or backwards along the velocity field just described in Sec. II to search for BiPo or ^{218}Pb events respectively. If a BiPo or ^{218}Pb event is found within the point clouds, then the ^{218}Pb candidate event is labeled as the ^{218}Pb background and can be vetoed. In the rest of this section, we detail the algorithm used to select ^{214}Pb events.

A. Generation of noise fields

There is likely uncertainty in the velocity field due to both limited statistics of ^{220}Rn - ^{218}Po pairs and the position reconstruction uncertainty; this needs to be properly accounted for. We address this by adding noise fields to the velocity field to induce fluctuations. This is done because one cannot directly use the uncertainties from the position reconstruction uncertainty when integrating the trajectories of the propagated points; furthermore, there is additional uncertainty introduced due to the purity of the selected vector population (see Sec. II). These are constrained by two conditions; first, the noise fields have to be divergence free to avoid introducing sources and sinks, and second, the fields have to behave smoothly at the TPC boundaries. The constraints are expected to be sufficiently strong to make the arbitrary noise field a realistic proxy for the real conditions. Generation of the noise field started with smoothed Gaussian noise where a $\sigma = 0.8 \text{ cm}$ Gaussian kernel was used for smoothing, chosen to be significantly smaller than the length scale of the convection field as visible in Fig. 4. After this, the curl was taken to ensure the noise is divergence free. 16 noise fields were generated, and then permuted by mirroring and rotating the fields, resulting in a total of 256 noise fields.

TPC surfaces must also be handled smoothly, and the velocity component perpendicular to the surfaces must approach zero at the surfaces. This was ensured by smoothly scaling the perpendicular component of the field to zero, starting 3 cm away from surfaces. This method of

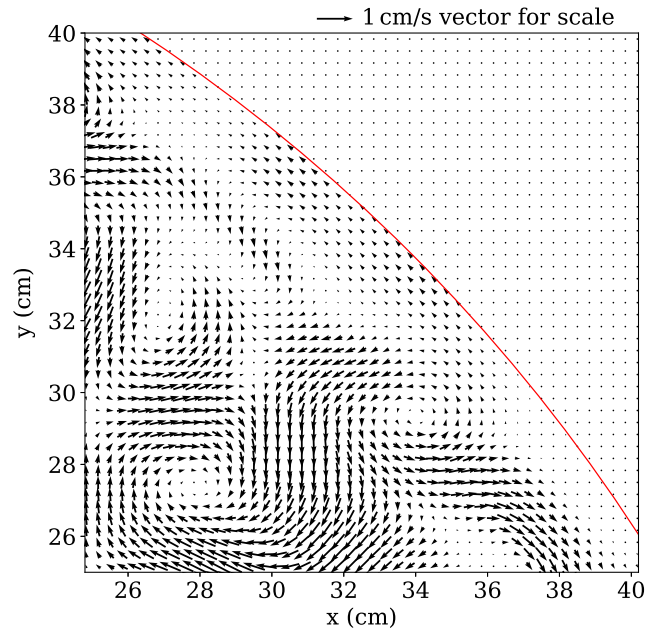


FIG. 6. Zoomed-in sample of the noise field. The edge of the detector is shown in red. It can be seen that the boundaries are handled smoothly, and that the field has no sinks that can trap propagating points.

using the curl to generate divergence-free noise and handling boundaries is described in [44]. A section of the resultant noise field can be seen in Fig. 6.

B. Point-cloud propagation

A veto volume within which one looks for predecessor or daughter events was constructed using a point cloud. The predecessor and daughter events for ^{214}Pb are ^{218}Po and BiPo events (see Fig. 1). BiPo events were selected by choosing events that have two interactions, corresponding to the ^{214}Bi and ^{214}Po decays. These interactions were required to be less than 5 cm apart in each of the x , y , and z directions, and the α -event is further required to have appropriate position-corrected S1 and S2 values. First, a random event was picked from the set of electronic recoils in XENON1T as the ^{214}Pb candidate event. A point cloud was then generated around this ^{214}Pb candidate event and then propagated using the convection and noise velocity fields, with every point in the point cloud exposed to a different randomly-assigned noise field (see Sec. III A). This noise field is introduced to account for the uncertainty of the velocity field; hence, exposing each point to a different noise field can be understood conceptually as exposing each point to a different version of the velocity field, allowing the uncertainty to be sampled. As one only needs to consider ^{214}Pb candidate events within the energy region of interest to a given analysis, propagating a point cloud from every ^{218}Po and BiPo event is more computationally expensive than from the smaller number of low-energy ^{214}Pb events. Point clouds from each ^{214}Pb candidate event are thus propagated in the forward and backward directions to look for BiPo and ^{218}Po events, respectively. These search directions are termed the BiPo and ^{218}Po channels for the rest of this paper. The point clouds are then culled in likelihood-space, based on the log-likelihood of each point. An illustration of a point cloud propagated along the velocity field can be seen in Fig. 7.

There are four main steps involved in the generation and propagation of this point cloud:

- (1) A ^{214}Pb candidate event is identified;
- (2) A point cloud is generated around the event, representing the position reconstruction uncertainty. The radial position uncertainty is $\sigma_R = 5.2\text{ cm} - (1.61\text{ cm})\log_{10}(S2/PE) + (0.019\text{ cm})\sqrt{S2/PE}$ [39], whereas the z -position uncertainty is estimated to be $\sigma_z = 0.17\text{ cm}$ from the displacement of the two decays in BiPo events;
- (3) Every 600 time steps (30 s), a probability density function is built out of the point cloud produced in the past 600 time steps. This was done using kernel density estimation (KDE) in four dimensions. To this end, a uniform kernel of radius 3 cm, and 0.3 sec in the time axis is used. Points that fall below a threshold of log-likelihood (ξ) are culled to speed up

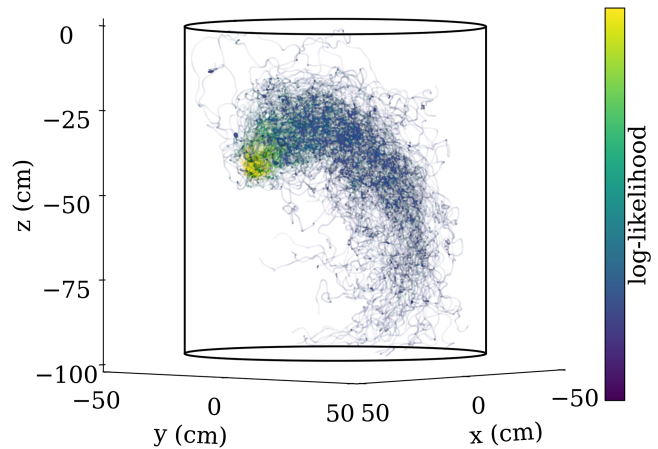


FIG. 7. Plot of a point cloud and the associated log-likelihood at each point. For illustration, the likelihood threshold is relaxed to a value of 8 show the point cloud propagating along the convection velocity field; elsewhere, higher values are used as indicated in the text, varying depending on the use-case. There are 192 points in the initial point cloud, and the time step size is 0.05 s.

computation. This log likelihood threshold is a free parameter;

- (4) DBScan clustering [45] is used to remove outlier points;
- (5) The algorithm repeats from step 3, until all points have been removed.

A time slice of a point cloud at the end of every iteration is shown in Fig. 8, with a log-likelihood constraint of 9.1, which is the same as that used in the final analysis in the BiPo channel.

C. Optimization of veto volume

In this section we will describe the optimization of parameters governing the operation of the ^{214}Pb veto algorithm. As the algorithm looks forward in time for BiPo events, and backwards in time for ^{218}Po events, there are two free parameters representing the log-likelihood thresholds (ξ_{BiPo} , ξ_{Po}) that had to be optimized. To this end, electronic recoil data from science run 0 of XENON1T between 30 and 70 keV was used [9].

To find the optimal values for these two parameters, the software radon veto was run on the entire dataset twice. For one of the two runs, the velocity field and time directions were reversed, so that due to causality, the ^{214}Pb candidate cannot be related to the ^{218}Po or BiPo events. This creates a sample of events that were vetoed purely due to coincidence, allowing for the probability of vetoing an event purely due to coincidence (p_{coinc}) to be profiled. For the aligned-direction run where one searches for BiPo and ^{218}Po events in the correct directions, the fraction of events that was ^{214}Pb (p_{pb}) was determined from a spectral fit from the XENON1T electronic recoil analysis [9]. Whether an event gets vetoed depends on both the probability of

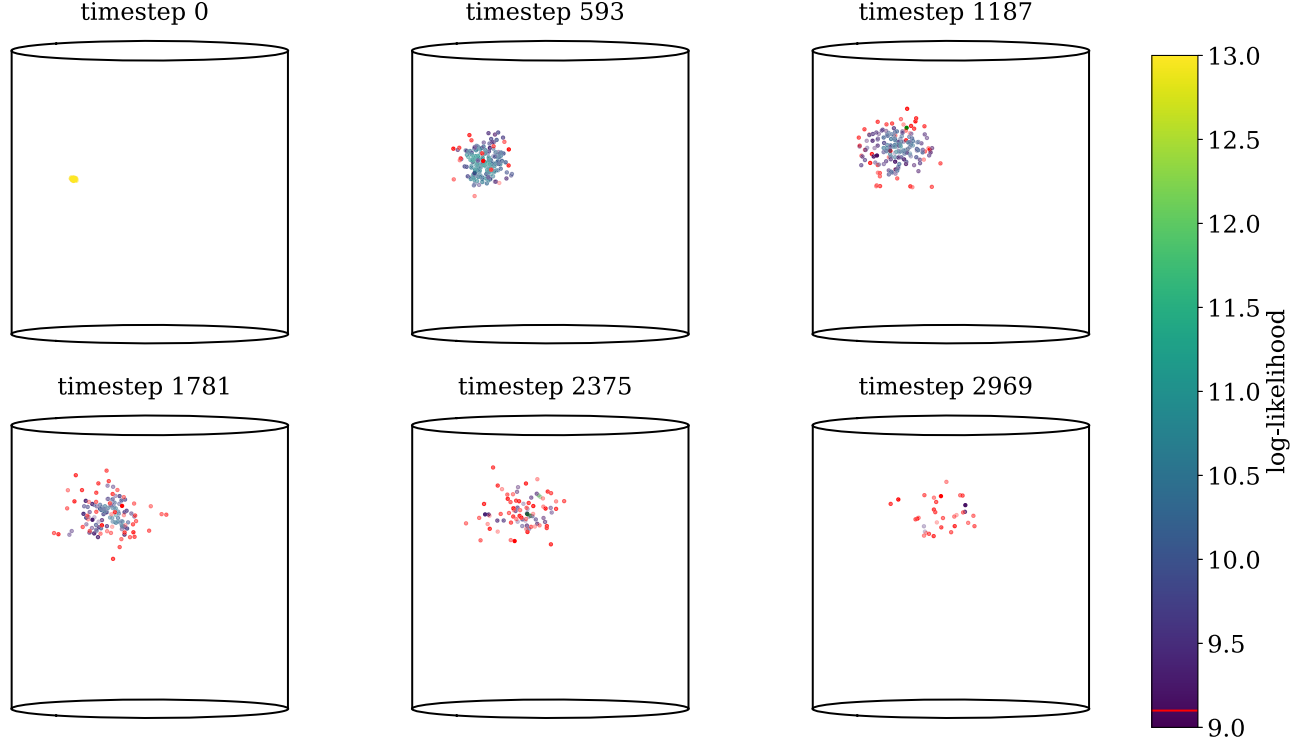


FIG. 8. Plot of point cloud with associated likelihoods at 6 different times steps (0, 29.65, 89.05, 118.75, and 148.45 s). Red points are culled by the log-likelihood limit and not propagated further to speed up computation.

vetoing a ^{214}Pb event (p_{true}) and the probability of vetoing an event purely due to coincidence (p_{coinc}), as illustrated in a tree diagram by Fig. 9. A likelihood function was thus used to fit p_{true} and p_{coinc} . p_{true} and p_{coinc} can also be interpreted as the ^{214}Pb background reduction and the exposure loss, respectively, as the ^{214}Pb background reduction can be given by the probability of vetoing ^{214}Pb events, and the exposure loss can be given by the probability of vetoing events that are not ^{214}Pb , as defined above. The likelihood function for a vetoed event is

$$\begin{aligned} \mathcal{L}_{i,\text{veto}}(p_{\text{true}}, p_{\text{coinc}}) &= p_{\text{Pb}}(E_i) \times p_{\text{true}} \\ &+ (1 - p_{\text{Pb}}(E_i)) \times p_{\text{coinc}}, \end{aligned} \quad (3)$$

where $\mathcal{L}_{i,\text{veto}}(p_{\text{true}}, p_{\text{coinc}})$ is the likelihood for the i th event to be vetoed, E_i is the energy of the event, and $p_{\text{Pb}}(E_i)$ is the fraction of events resulting from the decay of ^{214}Pb , as determined from the XENON1T electronic recoil analysis spectral fit. This can be interpreted as the probability that a given event is ^{214}Pb , multiplied by the probability of vetoing ^{214}Pb events, summed with the probability that a given event is not ^{214}Pb , multiplied by the probability of vetoing events that are not ^{214}Pb .

The likelihood function for a candidate event that is not vetoed for the aligned-direction runs is

$$\begin{aligned} \mathcal{L}_{j,\text{nveto}}(p_{\text{true}}, p_{\text{coinc}}) &= p_{\text{Pb}}(E_j) \times (1 - p_{\text{true}}) \\ &+ (1 - p_{\text{Pb}}(E_j)) \times (1 - p_{\text{coinc}}) \\ &= 1 - p_{\text{Pb}}(E_j) \times p_{\text{true}} \\ &- (1 - p_{\text{Pb}}(E_j)) \times p_{\text{coinc}}, \end{aligned} \quad (4)$$

where $\mathcal{L}_{j,\text{nveto}}(p_{\text{true}}, p_{\text{coinc}})$ is the likelihood for the j th event not vetoed. This can be interpreted as the probability that a given event is ^{214}Pb , multiplied by the probability of not vetoing ^{214}Pb events, summed with the probability that a given event is not ^{214}Pb , multiplied by the probability of not vetoing events that are not ^{214}Pb . This likelihood can similarly be derived from Fig. 9.

For the reversed runs, the likelihoods were changed to

$$\begin{aligned} \mathcal{L}_{i,\text{veto}}(p_{\text{true}}, p_{\text{coinc}}) &= p_{\text{coinc}}, \\ \mathcal{L}_{j,\text{nveto}}(p_{\text{true}}, p_{\text{coinc}}) &= 1 - p_{\text{coinc}}. \end{aligned} \quad (5)$$

This was done because for these runs the candidate event cannot be related to any ^{218}Po or BiPo events found. The likelihood function that was used to fit the probabilities p_{true} and p_{coinc} was then the sum of the log-likelihoods from each individual candidate event. This was done separately for the ^{218}Po and BiPo channels, to obtain four probabilities: the probability of vetoing an event due to coincidence via the BiPo channel ($p_{\text{coinc},\text{BiPo}}$), the probability of vetoing an event due to coincidence via the ^{218}Po

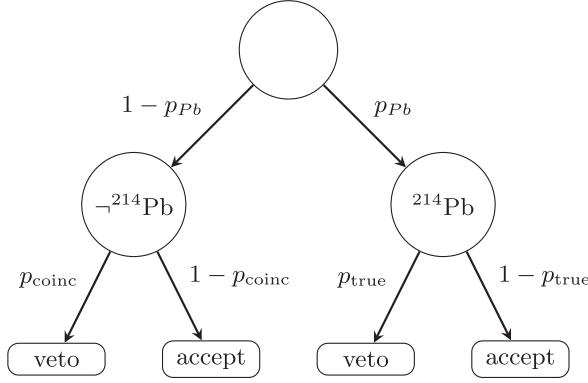


FIG. 9. Tree diagram describing the probabilities involved in whether an event is vetoed for or not for the aligned direction runs. This tree diagram describes the likelihoods shown in Eqs. (3) and (4). Each event has an energy-dependent probability of being a ^{214}Pb event (p_{Pb}); the probability of vetoing an event then depends on whether said event is a ^{214}Pb event or not. A higher probability of vetoing a ^{214}Pb event (p_{true}) and a lower probability of vetoing events that are not ^{214}Pb (p_{coinc}) indicate better performance.

channel ($p_{\text{coinc,Po}}$), the probability of vetoing a ^{214}Pb event via the BiPo channel ($p_{\text{true,BiPo}}$), and the probability of vetoing a ^{214}Pb event via the ^{218}Po channel ($p_{\text{true,Po}}$).

If a signal is much smaller than the background, the median asymptotic discovery significance of a counting experiment scales as $\text{signal}/\sqrt{\text{background}}$ [46]. The reduction in signal can be computed as the probability for events that are not ^{214}Pb to be vetoed. The background reduction can be computed by the reduction in ^{214}Pb background multiplied by the fraction of the background represented by ^{214}Pb , summed with the probability for events that are not ^{214}Pb to be vetoed multiplied by the fraction of background events that are not ^{214}Pb . An event is vetoed is vetoed by either the ^{218}Po or the BiPo channel, thus the probabilities for an event to survive each of the two channels can be

multiplied; a graphical depiction of these probabilities can be found in Fig. 10. One can thus compute a normalized sensitivity for a dark matter search,

$$Z = \frac{\text{signal}}{\sqrt{\text{background}}}, \quad (6)$$

where

$$\begin{aligned} \text{signal} &= \tilde{p}_{\text{coinc,BiPo}} \times \tilde{p}_{\text{coinc,Po}}, \\ \text{background} &= 1 - (1 - \alpha) \\ &\quad \times [1 - \tilde{p}_{\text{coinc,BiPo}} \times \tilde{p}_{\text{coinc,Po}}] \\ &\quad - \alpha [1 - (\tilde{p}_{\text{true,BiPo}} \times \tilde{p}_{\text{true,Po}})], \\ \tilde{p}_{\text{coinc,BiPo}} &= 1 - p_{\text{coinc,BiPo}}(\xi_{\text{BiPo}}), \\ \tilde{p}_{\text{coinc,Po}} &= 1 - p_{\text{coinc,Po}}(\xi_{\text{Po}}), \\ \tilde{p}_{\text{true,BiPo}} &= 1 - p_{\text{true,BiPo}}(\xi_{\text{BiPo}}), \\ \tilde{p}_{\text{true,Po}} &= 1 - p_{\text{true,Po}}(\xi_{\text{Po}}), \end{aligned}$$

where ξ_{Po} and ξ_{BiPo} refer to the likelihood threshold parameters being optimized for the ^{218}Po and BiPo channels, respectively, and α refers to the fraction of the background that can be attributed to ^{214}Pb . This is energy dependent in principle, but is approximated to be a constant $\alpha = 0.8$ for the purposes of this optimization, as given by the average between 0 and 30 keV. A tilde above a parameter, such as \tilde{p} denotes the best-fit value of said parameter. This normalized sensitivity is used as a proxy for optimization of algorithm parameters and evaluation of performance.

Finally, this process was repeated for multiple values of the threshold parameters that govern the veto volume size in the ^{218}Po and BiPo channels. This procedure gives an optimal ^{214}Pb background reduction of 6.3%, and an exposure loss of 1.8%, as defined by the signal and background components in Eq. (6), with optimal thresholds

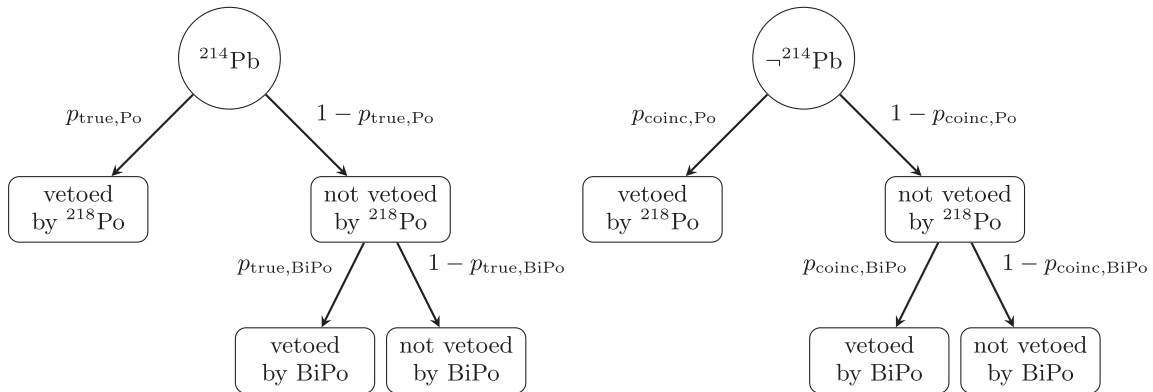


FIG. 10. Tree diagrams describing the probabilities involved in whether ^{214}Pb events (left) and other events (right) are vetoed, given both ^{218}Po and BiPo channels. This tree diagram describes the derivation of Eq. (6). An event is vetoed if the veto is triggered by either the ^{218}Po or BiPo channels.

of 9.7 and 9.0 for the ^{218}Po and BiPo channels respectively. From this, the sensitivity improvement estimated via this procedure is a modest 1.4% in XENON1T. However, as this is an analysis technique, it can still be a cost-effective addition to hardware radon-mitigation efforts, such as the cryogenic distillation system in XENONnT [20]. In addition, as will be shown later in Secs. IV C, much higher performance is possible in systems with lower-background radon levels and which have slower convective flows.

IV. RESULTS AND DISCUSSION

A. Demonstration of software radon veto

The ER analysis dataset from XENON1T [9] was used to demonstrate how this software radon veto would work in practice. For this section, the same data as introduced in Sec. III C was used, but with the energy range [0 keV, 70 keV] instead. Similarly to Sec. III C, both ^{218}Po and BiPo channels were used to tag events as ^{214}Pb . The thresholds used are $\xi_{\text{Po}} = 9.7$ and $\xi_{\text{BiPo}} = 9.0$.

Following this, the likelihoods shown in Eqs. (3) and (4) were used to fit p_{true} and p_{coinc} . This is shown in Fig. 11. The fit corresponds to an exposure loss of $1.8 \pm 0.2\%$, and a $6.2_{-0.9}^{+0.4}\%$ reduction in the ^{214}Pb background. We can see that, as expected, the final fit of the ^{214}Pb background reduction (p_{true}) and the exposure loss (p_{coinc}) agrees with the best fit values obtained Sec. III C.

The expected background spectrum was then computed by multiplying the components of the background fit from [9] that are not from ^{214}Pb with $1 - p_{\text{coinc}}$, multiplying the ^{214}Pb background with $1 - p_{\text{true}}$ and summing the two. This is shown in Fig. 12. It can be seen that the red line is a good fit for the data.

We can also demonstrate that this software radon veto indeed selects ^{214}Pb events by looking at the energy

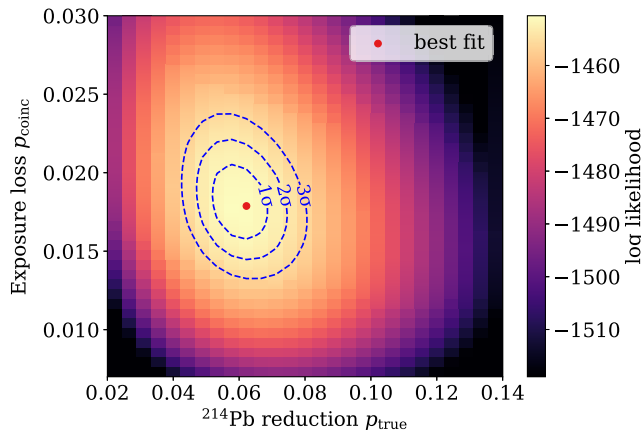


FIG. 11. Likelihood fit using events from the ER analysis dataset and the likelihoods discussed in Sec. III C. In this plot, the best fit values of the ^{214}Pb background reduction (p_{true}) and the exposure loss (p_{coinc}) are marked in red, and the error ellipses are shown in blue.

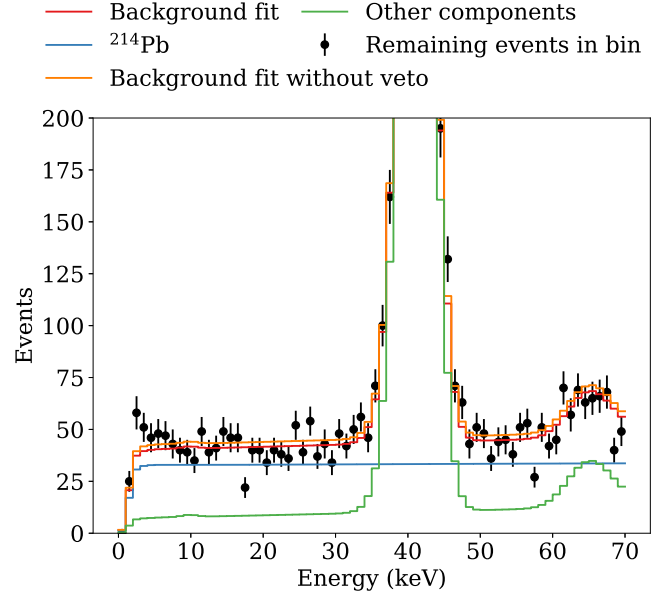


FIG. 12. Spectrum of events remaining after the software radon veto with 1σ Poisson confidence intervals (black), compared with the expected background spectrum based on the signal-free spectral fit from [9] and the inferred values of p_{true} and p_{coinc} (red). The ^{214}Pb component is shown in blue, and other background components are shown in green. The peaks at 42 keV and 64 are due to ^{83m}Kr and ^{124}Xe decay, respectively [9]. The combined fit with all components before applying the software radon veto is also shown in red to indicate the impact of this procedure.

spectrum of vetoed events. To this end, a portion of the fiducialized data from the search for neutrinoless double-beta decays in XENON1T was used [5]. This data corresponds to 22.05 days of exposure with a fiducial mass of 741 ± 9 kg. A spectral fit between 270 and 2000 keV includes both spectral features due to ^{214}Bi excited states at 295 and 352 keV, and the beta decay Q -value of 1018 keV [25], but avoids low-energy features in the spectrum from ^{83m}Kr and ^{131m}Xe . The selected data with a spectral fit is shown in Fig. 13.

The software radon veto was run on the dataset shown in Fig. 13 with thresholds of $\xi_{\text{Po}} = 9.7$ and $\xi_{\text{BiPo}} = 9.4$; these differ from those used in Sec. IV A. These thresholds were picked without an optimization procedure, but did produce a cleaner sample of ^{214}Pb decays. Following that, the same procedure used above for the ER analysis was used to fit p_{true} and p_{coinc} ; however, here events that are tagged as ^{214}Pb are examined instead. Thus, the components of the spectral fit that are not ^{214}Pb were multiplied with p_{coinc} , and the ^{214}Pb component was multiplied by p_{true} .

This result is shown in Fig. 14. It can be seen from the difference in spectral shape between Fig. 13 and Fig. 14 that the tagged population is indeed dominated by the decay of ^{214}Pb . It is also possible to identify relevant spectral features at 295 and 352 keV, as well as the Q -value

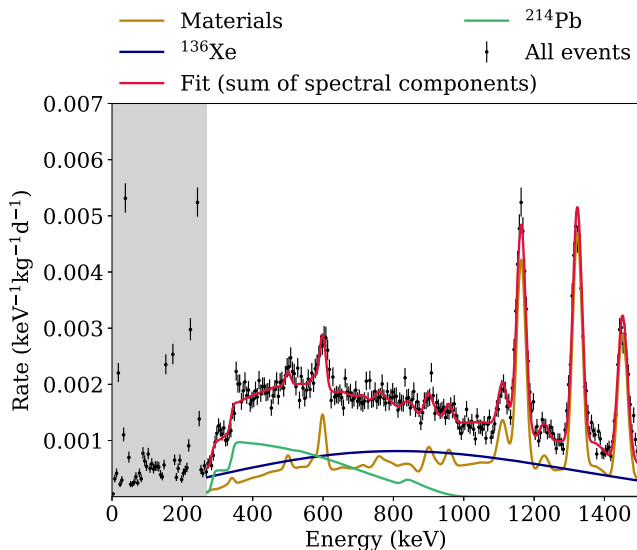


FIG. 13. Spectrum of events in the XENON1T double beta decay dataset [5] corresponding to 22.01 days of exposure, before the application of the radon veto. Data points with 5 keV bins is shown in black. A spectral fit is shown in solid lines, with the summed fit in red. The gray shaded region indicates data that is not used for fitting.

of 1018 keV, though there is insufficient statistics to resolve the two steps at 295 and 352 keV separately [25]. In particular, the ^{214}Pb decay endpoint can be clearly identified in the tagged population in Fig. 14, but not in the full dataset shown in Fig. 13.

B. Example of a recovered decay chain

A reconstructed example of the portion of the decay chain that is used for the software radon veto is shown in this section (compare Fig. 1). The software radon veto was used to find the ^{218}Po and BiPo events from the ^{214}Pb event.

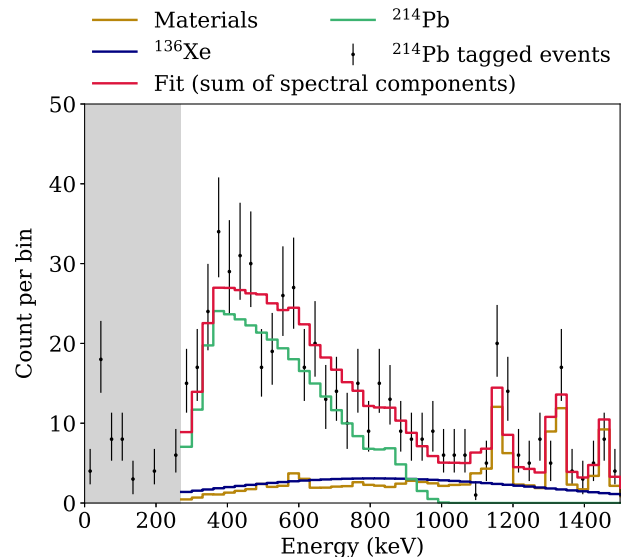
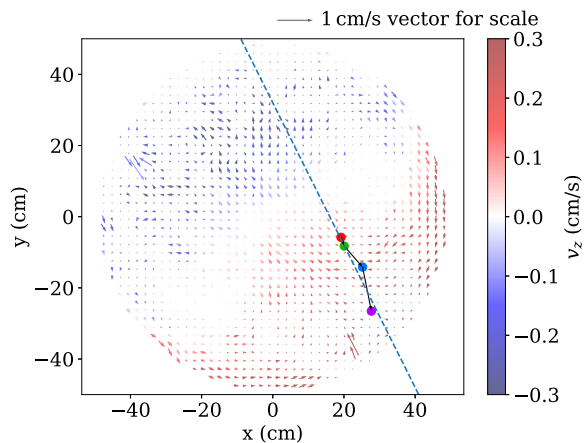


FIG. 14. The spectrum of events in the population of events tagged as ^{214}Pb . It can be seen both from the fit and from the shape of the spectrum that the ^{214}Pb fraction is greatly enhanced in the vetoed sample. The gray shaded region indicates data that is not used for fitting.

Tagging a ^{214}Pb event only requires matching either a ^{218}Po or BiPo event, however, in the chosen example, both ^{218}Po and BiPo events were found. The ^{222}Rn event related to the ^{218}Po was then found via the matching procedure shown in Sec. II A. The four identified events can be seen in Fig. 15, laid over the same velocity field shown in Sec. II B. It can be seen that the events propagate along the velocity field.

C. Projection of performance in XENONnT and future TPCs

To project the performance of this technique to XENONnT, a model of how the software radon veto

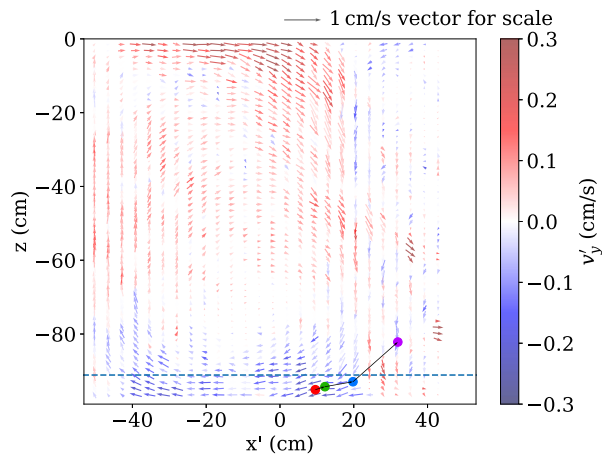


FIG. 15. The positions of a ^{222}Rn event (red), a ^{218}Po event (green), a ^{214}Pb event (blue), and a BiPo event (purple) are shown here, overlaid on top of the velocity field in the detector. The top view is shown on the left, and a side view is shown on the right. The blue dotted line on the top and side views are used to show the slices taken to create the side and top views, respectively.

performs under various conditions must be constructed. Such a model needs to be constructed because the analyses shown in earlier sections is data driven and done using XENON1T data; repeating the full analysis on XENONnT data has not been done yet as of this publication. For each channel, the probability of incorrectly vetoing an event that is not ^{214}Pb is simply given by the size of the point cloud used to construct the veto volume, multiplied by the rate of ^{218}Po or BiPo events. When propagating particles along a three-dimensional flow, chaotic mixing is expected to occur. This makes the point cloud size diverge exponentially with time [47]. As such, the growth of the point cloud volume can be modeled with a Lyapunov exponent; the characteristic exponential divergence of two close trajectories [48]. The probability of incorrectly vetoing an event that is not ^{214}Pb as a function of the time the point cloud is propagated for is given by

$$\begin{aligned} p_{\text{coinc}}(t) &= AC \int_0^t e^{v\lambda\tau} d\tau \\ &= AC \frac{e^{v\lambda t} - 1}{v\lambda}, \end{aligned} \quad (7)$$

where A is the activity of ^{218}Po or BiPo events, depending on the channel being modeled, v is the convection speed, t is the time the point cloud is being propagated, and C and λ are fitting constants.

The probability of correctly vetoing an event that is ^{214}Pb , on the other hand, can be modeled with the exponential decay of the radioactive species, multiplied by the probability of there being a correctly reconstructed ^{218}Po alpha or BiPo event in the detector, p_{branch} . As the efficiency of detecting alphas is high, the probability for the ^{218}Po channel is approximated as $p_{\text{branch,Po}} = 1$. The probability for the BiPo channel has to account for the effect of plate-out onto surfaces in the detector [49] and less efficient selections, and as such is taken to be the ratio of the BiPo rate as measured using fully reconstructed BiPo events in the XENON1T detector and the rate of ^{214}Pb events from the search of dark matter in the electronic recoil channel [9] as $p_{\text{branch,BiPo}} = 0.25$. As this includes selection efficiencies and plate-out, this number might change between detectors, but is kept constant here to estimate the XENONnT performance. The probability of correctly vetoing an event that is ^{214}Pb is thus given by

$$\begin{aligned} p_{\text{true}}(t) &= p_{\text{branch}}\lambda_{\text{decay}} \int_0^t e^{-\lambda_{\text{decay}}\tau} d\tau \\ &= p_{\text{branch}}(1 - e^{-\lambda_{\text{decay}}t}), \end{aligned} \quad (8)$$

where λ_{decay} is the decay constant of the radioactive species relevant to the specific channel, p_{branch} is the multiplicative factor stemming from selection efficiencies and plate out as detailed above, and t is the time the point cloud is being

propagated. Equations (7) and (8) can then be combined to eliminate the time variable and produce,

$$p_{\text{coinc}} = \frac{AC}{v\lambda} \left(\left(1 - \frac{p_{\text{true}}}{p_{\text{branch}}} \right)^{-\frac{v\lambda}{\lambda_{\text{decay}}}} - 1 \right). \quad (9)$$

With Eq. (9), there are only two free parameters, C and λ . These two free parameters can be fit by running the software radon veto on XENON1T data with different veto volumes using both ^{218}Po and BiPo channels. The resultant values from both the channels, $p_{\text{true,Po}}$, $p_{\text{true,BiPo}}$, $p_{\text{coinc,Po}}$, and $p_{\text{coinc,BiPo}}$, are used to fit the values of the fitting constants, as shown in Fig. 16. This is done in a single fit, thus the fit procedure only produces one value each of C and λ . These data points differ from the exposure loss and ^{214}Pb reduction values in Fig. 11 as the values from the ^{218}Po and BiPo channels are presented separately.

The veto volume is parametrized by threshold parameters that attempt to find veto volumes with the highest probability content instead of using a simple time cutoff for how long to propagate the point cloud; that is, outlying points in a point cloud might be propagated for shorter amounts of time than points that are central to the point cloud. The extrapolation to XENONnT and future TPCs considers a constant integration time for each point cloud and is thus approximate. However, as can be seen in Fig. 16, it fits XENON1T data quite well when fit simultaneously on both the ^{218}Po and BiPo channels.

To project the performance of the software radon veto in XENONnT, the fit parameters from above are kept the same, but the activities are scaled down. The ^{218}Po activity

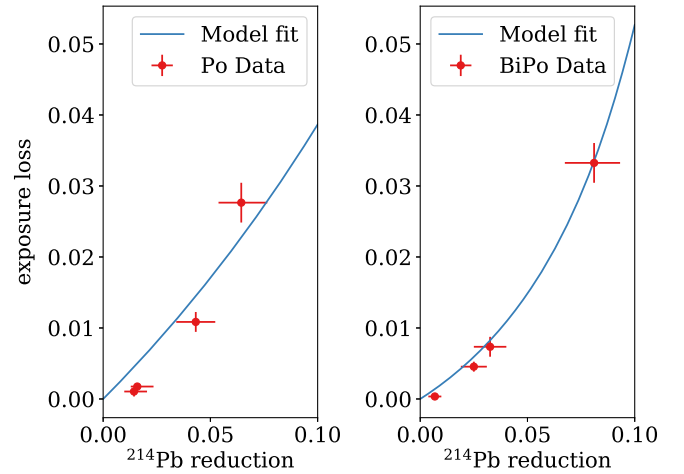


FIG. 16. The exposure loss versus the ^{214}Pb background reduction from the ^{218}Po (left) and BiPo (right) channels. The left and right plots correspond to a simultaneous fit on both datasets; the reason why the curve looks different in the two plots is due to the different half lives, and the different probability of there being a correctly reconstructed ^{218}Po alpha or BiPo event (p_{branch}).

TABLE I. Table showing the estimated optimal improvement in sensitivity ($Z_{\text{optim}} - 1$), at various scenarios of convection speed ($v_{\text{convection}}$) in XENONnT, together with the reduction in ^{214}Pb background ($1 - b_{\text{pb}}$) and the exposure loss ($1 - s$) at the stated optimal sensitivity improvement.

Convection speed (cm/s) $v_{\text{convection}}$	Sensitivity improvement $Z_{\text{optim}} - 1$	^{214}Pb background reduction $1 - b_{\text{pb}}$	Exposure loss $1 - s$
0.8	4.3%	25%	3.5%
0.4	7.6%	41%	5.9%
0.2	12%	59%	8.8%
0.1	17%	75%	11%

in XENONnT is measured to be $1.691 \pm 0.006_{\text{stat}} \pm 0.072_{\text{sys}} \mu\text{Bq/kg}$, and the ^{214}Pb activity is measured to be $1.31 \pm 0.17_{\text{stat}} \mu\text{Bq/kg}$ in XENONnT [50]. It should be noted that this is the XENONnT Science Run 0 radon level, and could be further lowered in future science runs depending on the mode of operation of the radon removal system [20]. The ratio between ^{218}Po and fully reconstructed BiPo activities is kept the same from XENON1T. Due to the lower ^{214}Pb background, the fraction of the background attributed to ^{214}Pb is estimated to be $\alpha = 0.5$ here. The projected performance for various convection speeds, optimized for normalized sensitivity as defined in Eq. (6), is shown in Table I.

XENONnT is a larger detector than XENON1T; hence, due to considerations discussed in Sec. II A should be expected to have much lower convection speeds. However, we consider higher convection velocity conditions as well due to the introduction of liquid xenon recirculation, which may affect the convective flow in the TPC. As can be seen, due to the reduced radon level in XENONnT, the background reduction is improved greatly over XENON1T in all of the considered convection speed scenarios. These results can also be seen in Fig. 17.

Future large dual-phase TPCs might also not have a convective flow; in such a situation, the movement of daughter nuclides after a radioactive decay becomes dominated by ion drift [42]. Here, we consider the performance of software tagging of radon-chain backgrounds in the limiting case of this regime, where any stochastic motion is entirely due to diffusion. For simplicity, and because of the unknown effects of plate-out and BiPo reconstruction in future detectors, only the ^{218}Po channel is considered here, resulting in a conservative estimate of the algorithm's performance.

A simple analytic model can be used to estimate the performance of software tagging in the true diffusion-limited regime. The probability density function of the displacement of a particle diffusing in one dimension is given by the 1D diffusion equation [51],

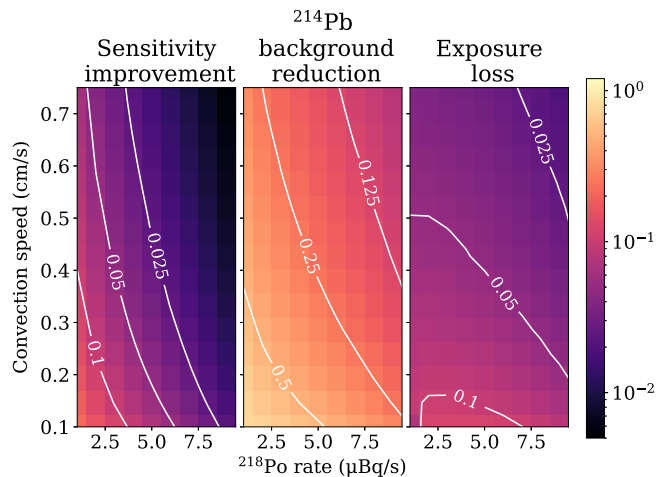


FIG. 17. Estimated optimal improvement in sensitivity ($Z_{\text{optim}} - 1$) (left), reduction in ^{214}Pb background ($1 - b_{\text{pb}}$) (center), and exposure loss ($1 - s$), as a function of ^{218}Po activity and convection speed. The ratio between ^{218}Po and ^{214}Pb activities is kept at a constant 1.691/1.31, based on [50]. Contours for specific values of each panel are shown in white. It can be seen that we can expect significantly improved background reduction at lower activities and convection speeds.

$$\frac{\partial \rho_x}{\partial t} = D \frac{\partial^2 \rho_x}{\partial x^2}. \quad (10)$$

Using the 1D diffusion equation leads to no loss of generality because the distribution of displacement of a diffusing point is independent in different orthogonal axes.

The solution to Eq. (10) with an initial Dirac delta function, $\delta(x)$, corresponding to the known position of the original particle, is a normal distribution with $\mu = 0$ and $\sigma^2 = 2Dt$; in 3D, this corresponds to a spherical normal distribution with $\sigma_x = \sigma_y = \sigma_z = \sqrt{2Dt}$. It can be noted here that these are also the Green's function of the 1D and 3D heat equations, respectively, as the isotropic diffusion equation is the heat equation [52].

The true diffusion constant can be estimated using Einstein's relation [53] and the mobility of $\mu = 0.219 \pm 0.004 \text{ cm}^2/(\text{kV/s})$ as measured by EXO-200 [42],

$$\begin{aligned} D &= \frac{\mu k_b T}{q} \\ &= \frac{k_b 0.219 \text{ cm}^2/(\text{kV s}) 170 \text{ K}}{q_e} \\ &\approx 3.2 \times 10^{-6} \text{ cm}^2/\text{s}. \end{aligned} \quad (11)$$

With the diffusion constant shown in Eq. (11), the daughter of a ^{218}Po decay would diffuse approximately $\sqrt{3 \times 2 \times D \times (5 \times 27.06 \text{ min})} \approx 0.4 \text{ cm}$ in five half-lives. An activity of $1.7 \mu\text{Bq/kg}$ as achieved in XENONnT SR0 [50], and a liquid xenon density of

$\sim 3 \text{ g/cm}^3$ [54] corresponds to an activity per unit volume of $4.6 \times 10^{-9} \text{ Bq/cm}^3$, or 0.4 decays per 10 L per day. As $1 \text{ L} = (10 \text{ cm})^3$, this implies that as long as the position reconstruction uncertainty remains significantly below 10 cm, decays would be essentially spatially isolated without fluid flows, and hence one can reject radon-chain backgrounds with a tagging efficiency of near unity.

D. Application to ^{137}Xe

The decay of cosmogenic ^{137}Xe is expected to be a major background in the search for $0\nu\beta\beta$ decay in ^{136}Xe in XENONnT [5], and next-generation liquid xenon TPCs [11,55,56]. ^{137}Xe is produced due to the capture of muon-induced neutrons or radiogenic neutrons by ^{136}Xe [5,55,57], and subsequently undergoes beta-decay to ^{137}Cs , as shown in Fig. 18. In this section, we focus on the ^{137}Xe background arising from muon-induced neutrons.

A similar methodology to Sec. IV C can be used to estimate the performance of a ^{137}Xe veto; however, this estimate is more speculative. This is because the performance of such a veto would rely on the reconstruction of neutron-capture gammas and a detailed analysis to search of these neutron-capture events has not been done in this study. These neutron-captures gammas represent the progenitor events. Point clouds generated at the position of ^{137}Xe decay candidates are thus used to look for these neutron-capture events, which should appear as ER events that are coincident with muon veto triggers.

The relationship between p_{true} and p_{coinc} can be derived from Eq. (9). However, the fit parameters from Sec. IV C

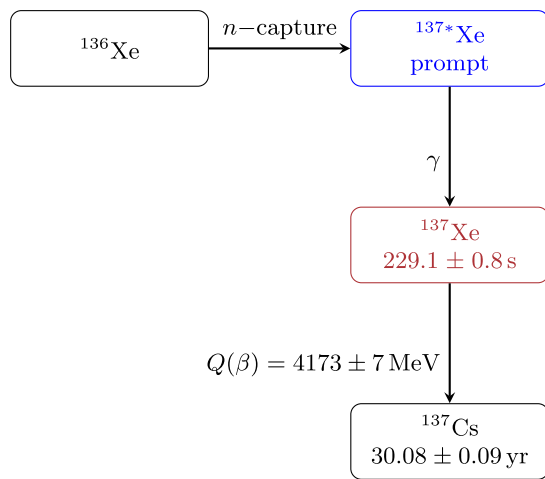


FIG. 18. Neutron capture of ^{136}Xe and subsequent decay of ^{137}Xe . Data regarding the decay of ^{137}Xe and ^{137}Cs retrieved using the NNDC ENSDF, with original data from nuclear data sheets [58]. The isotope the decays to produce the relevant background, ^{137}Xe , is colored red, whereas the excited state which produces the gamma events that are used for the tagging of the ^{137}Xe background are colored blue.

have to be adapted for this study. The initial point cloud has to be much bigger, because the uncertainty on the true location of the neutron capture is not dominated by position reconstruction uncertainties, but by the mean-free path of gammas. In the absence of a detailed analysis, the minimum attenuation between 10^{-2} MeV and 10^1 MeV is conservatively applied. This is $0.036 \text{ cm}^2/\text{g}$ according to the XCOM database [59], corresponding to a maximum mean-free path of 9.8 cm. Thus, to account for this, the fit parameter C in Eq. (7) which should scale with the initial point cloud size, is divided by the position reconstruction uncertainty volume, and multiplied by the volume of a sphere with a radius of 9.8 cm in liquid xenon. Further, the half-life of ^{137}Xe , which is $229.1 \pm 0.8 \text{ s}$ [58], is applied.

The activity rate A is also different in this scenario. In XENONnT, the rate in the muon veto is observed to be $\approx 0.035 \text{ Hz}$. As the neutron capture time in liquid xenon is $\sim 100 \mu\text{s}$ [60], a 1 ms window after each muon trigger to search for neutron captures can be considered, leading to a live time fraction of 3.5×10^{-5} within which neutron captures would be searched for. The emitted gammas are expected to be of energies $\sim 1 \text{ MeV}$ [60] where ^{136}Xe decay is the dominant background. Thus, the background rate can be approximated using the fraction found above, multiplied by the rate of ^{136}Xe decays in natural xenon, $\approx 4.2 \mu\text{Bq/kg}$ [5,61], resulting in $A = 1.5 \times 10^{-4} \mu\text{Bq/kg}$. Using these values, which represent adaptations of the fit values used in Sec. IV C, the performance for different convection velocities is shown in Fig. 19.

It can be seen that for all of the velocity scenarios, almost all of the cosmogenic ^{137}Xe background can be rejected. In particular, for convection velocities around or below 0.2 cm/s , the reduction of the cosmogenic ^{137}Xe background approaches unity for a 10% reduction in exposure. However, it should be noted that the reconstruction efficiency of neutron-capture gammas has not been

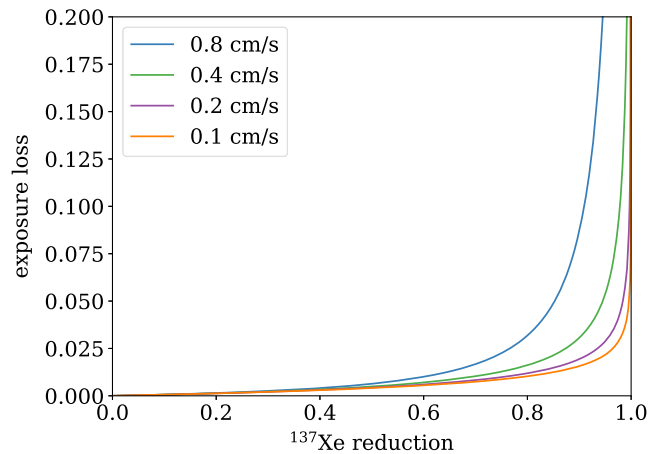


FIG. 19. Projected exposure loss versus ^{137}Xe background reduction when tagging ^{137}Xe backgrounds for different scenarios of convection velocity.

measured, and will proportionally reduce p_{true} . In addition, the projections presented here use the worst case mean-free path of 9.8 cm; in reality, the initial point cloud size could be potentially much smaller, depending on the spectrum of the gammas emitted after a neutron capture.

V. CONCLUSIONS

In this paper, the design and performance of an algorithm for tagging radon-chain backgrounds in liquid noble element TPCs were presented. The presented algorithm performs tagging of the ^{214}Pb background, which is part of the ^{222}Rn decay chain. This was demonstrated on XENON1T datasets used for the ER analysis and the search for neutrinoless double beta decay; the original analyses can be found in [5,9]. It was shown that for the ER analysis, an exposure loss of $1.8 \pm 0.2\%$ and a $6.2_{-0.9}^{+0.4}\%$ reduction in the ^{214}Pb background can be expected. The neutrinoless double beta-decay dataset is used to produce a high-purity sample of ^{214}Pb decay events, as can be seen from a spectral fit. This sample also displays relevant features in the spectrum such as the peak at 352 keV and falling off at the Q -value of approximately 1 MeV.

While the demonstrated background reduction is small, the cost of such a software-based background-reduction technique can be minimal, making deployment cost effective. In addition, much higher performance can be expected in larger detectors with lower intrinsic radon levels, due to individual radon-chain events being further apart in the detector in both space and time. In XENONnT due to the lower radon level, performance is expected to be significantly higher than in XENON1T, with an optimal ^{214}Pb background reduction of between 25% and 75%, depending on the convection speed in the detector, with a corresponding exposure loss of between 3.5% and 11%. If the motion is dominated by diffusion, near-perfect tagging of radon chain backgrounds can be expected.

The fact that the performance of a software veto for ^{214}Pb backgrounds improves with larger detectors and lower intrinsic radon levels makes it complementary to hardware-based approaches such as the cryogenic distillation system used by XENONnT [20], or a charcoal trap [21]. This is because as detector size increases, these hardware-based approaches require increasing mass flow rates to retain the same performance, whereas algorithmic approaches do not suffer from this scaling. In addition, software-based approaches perform better if the radon level

is already low due to radiopurity controls or hardware-based radon removal methods; in the limiting case where there is on average much less than one ^{218}Po in the TPC at any given time, there can simply be a veto on all data within a few half-lives of a ^{218}Po alpha decay to remove almost all of the ^{214}Pb background. It should be noted that there are also hardware approaches that do require increasing mass flow rates to retain performance, such as material selection and screening [13,19,22], detector design [62], and material coating [63].

The methods outlined in this paper can also be used to suppress radon chain backgrounds in liquid argon TPCs, where hardware-based approaches for the mitigation of radon-chain backgrounds are similarly being pursued [23].

In addition, the performance of a similar approach applied to reduce the cosmogenic ^{137}Xe background was also estimated. This background is expected to be a major background in the search for $0\nu\beta\beta$ decay in ^{136}Xe in XENONnT [5], LZ [57], and next-generation liquid xenon TPCs [55]. It was found that if neutron capture gammas can be selected with high efficiency, then the ^{137}Xe background can be tagged in XENONnT with an efficiency of $> 90\%$, resulting in a $< 9\%$ background reduction, depending on the convection speed.

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), Weizmann Institute of Science, Israeli Science Foundation, Binational Science Foundation, Région des Pays de la Loire, Knut and Alice Wallenberg Foundation, Kavli Foundation, JSPS Kakenhi and JST FOREST Program in Japan, Tsinghua University Initiative Scientific Research Program and Istituto Nazionale di Fisica Nucleare. This project has received funding/support from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie Grant Agreement No. 860881-HIDDeN. Data processing is performed using infrastructures from the Open Science Grid, the European Grid Initiative and the Dutch national e-infrastructure with the support of SURF Cooperative. We are grateful to Laboratori Nazionali del Gran Sasso for hosting and supporting the XENON project.

- [1] E. Aprile *et al.* (XENON Collaboration), *Eur. Phys. J. C* **77**, 881 (2017).
- [2] E. Aprile *et al.* (XENON Collaboration), *J. Cosmol. Astropart. Phys.* **11** (2020) 031.
- [3] D. S. Akerib *et al.* (LZ Collaboration), *Nucl. Instrum. Methods Phys. Res., Sect. A* **953**, 163047 (2020).
- [4] E. Aprile *et al.* (XENON Collaboration), *Phys. Rev. Lett.* **121**, 111302 (2018).
- [5] E. Aprile *et al.* (XENON Collaboration), *Phys. Rev. C* **106**, 024328 (2022).
- [6] G. Anton *et al.* (EXO-200 Collaboration), *Phys. Rev. Lett.* **123**, 161802 (2019).
- [7] D. S. Akerib *et al.* (LZ Collaboration), *Phys. Rev. C* **104**, 065501 (2021).
- [8] E. Aprile *et al.* (XENON Collaboration), *Nature (London)* **568**, 532 (2019).
- [9] E. Aprile *et al.* (XENON Collaboration), *Phys. Rev. D* **102**, 072004 (2020).
- [10] E. Aprile *et al.* (XENON Collaboration), *Phys. Rev. Lett.* **129**, 161805 (2022).
- [11] J. Aalbers *et al.* (DARWIN Collaboration), *J. Cosmol. Astropart. Phys.* **11** (2016) 017.
- [12] H. Zhang *et al.* (PandaX Collaboration), *Sci. China Phys. Mech. Astron.* **62**, 31011 (2019).
- [13] E. Aprile *et al.* (XENON Collaboration), *Eur. Phys. J. C* **82**, 599 (2022).
- [14] B. Singh *et al.*, *Nucl. Data Sheets* **160**, 405 (2019).
- [15] K. Abe *et al.*, *Nucl. Instrum. Methods Phys. Res., Sect. A* **661**, 50 (2012).
- [16] M. Arthurs, D. Q. Huang, C. Amarasinghe, E. Miller, and W. Lorenzon, [arXiv:2009.06069](https://arxiv.org/abs/2009.06069).
- [17] H. Chen, R. Gibbons, S. J. Haselschwardt, S. Kravitz, Q. Xia, and P. Sorensen, *Phys. Rev. D* **109**, L071102 (2024).
- [18] D. S. Akerib *et al.* (LUX Collaboration), *Astropart. Phys.* **97**, 80 (2018).
- [19] D. S. Akerib *et al.* (LZ Collaboration), *Eur. Phys. J. C* **80**, 1044 (2020).
- [20] M. Murra, D. Schulte, C. Huhmann, and C. Weinheimer, *Eur. Phys. J. C* **82**, 1104 (2022).
- [21] K. Pushkin *et al.*, *Nucl. Instrum. Methods Phys. Res., Sect. A* **903**, 267 (2018).
- [22] E. Aprile *et al.* (XENON Collaboration), *Eur. Phys. J. C* **77**, 890 (2017).
- [23] A. Avasthi *et al.*, in *Snowmass 2021* (2022), [arXiv:2203.08821](https://arxiv.org/abs/2203.08821).
- [24] S. Singh, A. K. Jain, and J. K. Tuli, *Nucl. Data Sheets* **112**, 2851 (2011).
- [25] S. Zhu and E. A. McCutchan, *Nucl. Data Sheets* **175**, 1 (2021).
- [26] M. Shamsuzzoha Basunia, *Nucl. Data Sheets* **121**, 561 (2014).
- [27] F. G. Kondev, *Nucl. Data Sheets* **109**, 1527 (2008).
- [28] K. Fujii, Y. Endo, Y. Torigoe, S. Nakamura, T. Haruyama, K. Kasami, S. Mihara, K. Saito, S. Sasaki, and H. Tawara, *Nucl. Instrum. Methods Phys. Res., Sect. A* **795**, 293 (2015).
- [29] M. Szydagis *et al.*, [arXiv:2211.10726](https://arxiv.org/abs/2211.10726).
- [30] E. Aprile *et al.* (XENON Collaboration), *Phys. Rev. D* **100**, 052014 (2019).
- [31] E. Aprile, C. E. Dahl, L. DeViveiros, R. Gaitskell, K. L. Giboni, J. Kwong, P. Majewski, K. Ni, T. Shutt, and M. Yamashita, *Phys. Rev. Lett.* **97**, 081302 (2006).
- [32] F. Jörg, D. Cichon, G. Eurin, L. Höttsch, T. Undagoitia Marrodán, and N. Rupp, *Eur. Phys. J. C* **82**, 361 (2022).
- [33] E. Aprile *et al.* (XENON Collaboration), *Phys. Rev. D* **95**, 072008 (2017).
- [34] D. C. Malling, Measurement and analysis of WIMP detection backgrounds, and characterization and performance of the large underground Xenon dark matter search experiment, Ph.D. thesis, Brown U., 2014, [10.7301/Z0057D9F](https://doi.org/10.7301/Z0057D9F).
- [35] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and Édouard Duchesnay, *J. Mach. Learn. Res.* **12**, 2825 (2011).
- [36] D. Masson, Novel ideas and techniques for large dark matter detectors, Ph.D. thesis, Purdue U., 2018, <https://docs.lib.purdue.edu/dissertations/AAI10792046/>.
- [37] E. Aprile *et al.* (XENON Collaboration), *Eur. Phys. J. C* **81**, 337 (2021).
- [38] G. Ahlers, E. Bodenschatz, R. Hartmann, X. He, D. Lohse, P. Reiter, R. J. A. M. Stevens, R. Verzicco, M. Wedi, S. Weiss, X. Zhang, L. Zwirner, and O. Shishkina, *Phys. Rev. Lett.* **128**, 084501 (2022).
- [39] J. Ye, Searches for WIMPs and axions with the XENON1T experiment, Ph.D. thesis, University of California, San Diego, 2020.
- [40] J. B. S. Haldane, *Biometrika* **35**, 414 (1948).
- [41] H. P. Lopuhaa and P. J. Rousseeuw, *Ann. Stat.* **19**, 229 (1991).
- [42] J. B. Albert *et al.* (EXO-200 Collaboration), *Phys. Rev. C* **92**, 045504 (2015).
- [43] J. Aalbers *et al.* (LZ Collaboration), *Phys. Rev. D* **108**, 012010 (2023).
- [44] R. Bridson, J. Houriham, and M. Nordenstam, *ACM Trans. Graph.* **26**, 46–es (2007).
- [45] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu, in *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining*, KDD'96 (AAAI Press, Washington DC, 1996), p. 226–231, <https://dl.acm.org/doi/10.5555/3001460.3001507>.
- [46] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, *Eur. Phys. J. C* **71**, 1554 (2011); **73**, 2501(E) (2013).
- [47] M. Speetjens, G. Metcalfe, and M. Rudman, *Appl. Mech. Rev.* **73**, 030801 (2021).
- [48] J. Taylor, *Classical Mechanics* (University Science Books, 2005), Chap. 12.
- [49] S. Bruenner, D. Cichon, G. Eurin, P. Herrero Gómez, F. Jörg, T. Marrodán Undagoitia, H. Simgen, and N. Rupp, *Eur. Phys. J. C* **81**, 343 (2021).
- [50] E. Aprile *et al.* (XENON Collaboration), *Phys. Rev. Lett.* **129**, 161805 (2022).
- [51] A. Einstein, *Investigations on the Theory of the Brownian Movement*, Dover Books on Physics (Dover Publications, Mineola, NY, 1956).
- [52] A. D. A. D. Polianin, *Handbook of Linear Partial Differential Equations for Engineers and Scientists*, 2nd ed. (Chapman & Hall/CRC, Boca Raton, 2002).

- [53] R. Richert, L. Pautmeier, and H. Bässler, *Phys. Rev. Lett.* **63**, 547 (1989).
- [54] E. Aprile and T. Doke, *Rev. Mod. Phys.* **82**, 2053 (2010).
- [55] J. Aalbers *et al.*, *J. Phys. G* **50**, 013001 (2023).
- [56] L. Baudis, *Nucl. Phys.* **B1003**, 116473 (2024).
- [57] D. S. Akerib *et al.* (LZ Collaboration), *Phys. Rev. C* **102**, 014602 (2020).
- [58] E. Browne and J. K. Tuli, *Nucl. Data Sheets* **108**, 2173 (2007).
- [59] M. Berger, J. Hubbell, S. Seltzer, J. Chang, J. Coursey, R. Sukumar, D. Zucker, and K. Olsen, Xcom: Photon cross section database (version 1.5) (2010), <http://physics.nist.gov/xcom>.
- [60] C. S. Amarasinghe, R. Coronel, D. Q. Huang, Y. Liu, M. Arthurs, S. Steinfeld, R. Gaitskell, and W. Lorenzon, *Phys. Rev. D* **106**, 032007 (2022).
- [61] J. B. Albert *et al.* (EXO-200 Collaboration), *Phys. Rev. C* **89**, 015502 (2014).
- [62] J. Dierle, A. Brown, H. Fischer, R. Glade-Beucke, J. Grigat, F. Kuger, S. Lindemann, M. R. Silva, and M. Schumann, *Eur. Phys. J. C* **83**, 9 (2023).
- [63] F. Jörg, From ^{222}Rn measurements in XENONnT and HeXe to radon mitigation in future liquid xenon experiments, Ph.D. thesis, Heidelberg U., 2022, [10.11588/heidok.00031915](https://doi.org/10.11588/heidok.00031915).