

XENON1T dark matter data analysis: Signal reconstruction, calibration, and event selection

E. Aprile,¹ J. Aalbers,^{2,3} F. Agostini,⁴ M. Alfonsi,⁵ L. Althueser,⁶ F. D. Amaro,⁷ V. C. Antochi,² F. Arneodo,⁸ L. Baudis,⁹ B. Bauermeister,² L. Bellagamba,⁴ M. L. Benabderrahmane,⁸ T. Berger,¹⁰ P. A. Breur,³ A. Brown,⁹ E. Brown,¹⁰ S. Bruenner,¹¹ G. Bruno,⁸ R. Budnik,¹² C. Capelli,⁹ J. M. R. Cardoso,⁷ D. Cichon,¹¹ D. Coderre,¹³ A. P. Colijn,^{3,*} J. Conrad,² J. P. Cussonneau,¹⁴ M. P. Decowski,³ P. de Perio,¹ P. Di Gangi,⁴ A. Di Giovanni,⁸ S. Diglio,¹⁴ A. Elykov,¹³ G. Eurin,¹¹ J. Fei,¹⁵ A. D. Ferella,² A. Fieguth,⁶ W. Fulgione,^{16,17} A. Gallo Rosso,¹⁶ M. Galloway,⁹ F. Gao,¹ M. Garbini,⁴ L. Grandi,¹⁸ Z. Greene,¹ C. Hasterok,^{11,‡} E. Hogenbirk,³ J. Howlett,¹ M. Iacovacci,¹⁹ R. Itay,¹² F. Joerg,¹¹ S. Kazama,²⁰ A. Kish,⁹ G. Koltman,¹² A. Kopec,²¹ H. Landsman,¹² R. F. Lang,²¹ L. Levinson,¹² Q. Lin,¹ S. Lindemann,¹³ M. Lindner,¹¹ F. Lombardi,⁷ J. A. M. Lopes,^{7,†} E. López Fune,²² C. Macolino,²³ J. Mahlstedt,² A. Manfredini,^{9,12} F. Marignetti,¹⁹ T. Marrodán Undagoitia,¹¹ J. Masbou,¹⁴ D. Masson,²¹ S. Mastroianni,¹⁹ M. Messina,^{16,8} K. Micheneau,¹⁴ K. Miller,¹⁸ A. Molinaro,¹⁶ K. Morá,² Y. Mosbacher,¹² M. Murra,⁶ J. Naganoma,^{16,24} K. Ni,¹⁵ U. Oberlack,⁵ K. Odgers,¹⁰ B. Pelssers,² R. Peres,^{9,7} F. Piastra,⁹ J. Pienaar,¹⁸ V. Pizzella,¹¹ G. Plante,¹ R. Podviiianuk,¹⁶ H. Qiu,¹² D. Ramírez García,¹³ S. Reichard,⁹ B. Riedel,¹⁸ A. Rizzo,¹ A. Rocchetti,¹³ N. Rupp,¹¹ J. M. F. dos Santos,⁷ G. Sartorelli,⁴ N. Šarčević,¹³ M. Scheibelhut,⁵ S. Schindler,⁵ J. Schreiner,¹¹ D. Schulte,⁶ M. Schumann,¹³ L. Scotto Lavina,²² M. Selvi,⁴ P. Shagin,²⁴ E. Shockley,¹⁸ M. Silva,⁷ H. Simgen,¹¹ C. Therreau,¹⁴ D. Thers,¹⁴ F. Toschi,¹³ G. Trincherro,¹⁷ C. Tunnell,²⁴ N. Upole,¹⁸ M. Vargas,⁶ O. Wack,¹¹ H. Wang,²⁵ Z. Wang,¹⁶ Y. Wei,^{15,§} C. Weinheimer,⁶ D. Wenz,⁵ C. Wittweg,⁶ J. Wulf,⁹ J. Ye,¹⁵ Y. Zhang,¹ T. Zhu,¹ and J. P. Zopounidis²²

(XENON Collaboration)^{||}

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

²*Oskar Klein Centre, Department of Physics, Stockholm University, AlbaNova, Stockholm SE-10691, Sweden*

³*Nikhef and the University of Amsterdam, Science Park, 1098XG Amsterdam, Netherlands*

⁴*Department of Physics and Astronomy, University of Bologna and INFN-Bologna, 40126 Bologna, Italy*

⁵*Institut für Physik & Exzellenzcluster PRISMA, Johannes Gutenberg-Universität Mainz, 55099 Mainz, Germany*

⁶*Institut für Kernphysik, Westfälische Wilhelms-Universität Münster, 48149 Münster, Germany*

⁷*LIBPhys, Department of Physics, University of Coimbra, 3004-516 Coimbra, Portugal*

⁸*New York University Abu Dhabi, 129188 Abu Dhabi, United Arab Emirates*

⁹*Physik-Institut, University of Zurich, 8057 Zurich, Switzerland*

¹⁰*Department of Physics, Applied Physics and Astronomy, Rensselaer Polytechnic Institute, Troy, New York 12180, USA*

¹¹*Max-Planck-Institut für Kernphysik, 69117 Heidelberg, Germany*

¹²*Department of Particle Physics and Astrophysics, Weizmann Institute of Science, Rehovot 7610001, Israel*

¹³*Physikalisches Institut, Universität Freiburg, 79104 Freiburg, Germany*

¹⁴*SUBATECH, IMT Atlantique, CNRS/IN2P3, Université de Nantes, Nantes 44307, France*

¹⁵*Department of Physics, University of California, San Diego, California 92093, USA*

¹⁶*INFN-Laboratori Nazionali del Gran Sasso and Gran Sasso Science Institute, 67100 L'Aquila, Italy*

¹⁷*INFN-Torino and Osservatorio Astrofisico di Torino, 10125 Torino, Italy*

¹⁸*Department of Physics & Kavli Institute for Cosmological Physics, University of Chicago, Chicago, Illinois 60637, USA*

¹⁹*Department of Physics “Ettore Pancini”, University of Napoli and INFN-Napoli, 80126 Napoli, Italy*

²⁰*Kobayashi-Maskawa Institute for the Origin of Particles and the Universe, Nagoya University, Furo-cho, Chikusa-ku, Nagoya, Aichi 464-8602, Japan*

²¹*Department of Physics and Astronomy, Purdue University, West Lafayette, Indiana 47907, USA*

²²*LPNHE, Université Pierre et Marie Curie, Université Paris Diderot, CNRS/IN2P3, Paris 75252, France*

²³*LAL, Université Paris-Sud, CNRS/IN2P3, Université Paris-Saclay, F-91405 Orsay, France*

²⁴*Department of Physics and Astronomy, Rice University, Houston, Texas 77005, USA*²⁵*Physics & Astronomy Department, University of California, Los Angeles, California 90095, USA*

(Received 12 June 2019; published 25 September 2019)

The XENON1T experiment at the Laboratori Nazionali del Gran Sasso is the most sensitive direct detection experiment for dark matter in the form of weakly interacting particles (WIMPs) with masses above $6 \text{ GeV}/c^2$ scattering off nuclei. The detector employs a dual-phase time projection chamber with 2.0 metric tons of liquid xenon in the target. A one metric ton \times year exposure of science data was collected between October 2016 and February 2018. This article reports on the performance of the detector during this period and describes details of the data analysis that led to the most stringent exclusion limits on various WIMP-nucleon interaction models to date. In particular, signal reconstruction, event selection, and calibration of the detector response to nuclear and electronic recoils in XENON1T are discussed.

DOI: [10.1103/PhysRevD.100.052014](https://doi.org/10.1103/PhysRevD.100.052014)

I. INTRODUCTION

The existence of a nonluminous, massive matter component beyond the standard model, called dark matter, is evidenced by numerous astrophysical observations [1]. Among the best-motivated dark matter candidates are weakly interacting massive particles (WIMPs) [2,3]. Ultrasensitive Earth-based detectors provide one possible approach to the direct detection of WIMPs as the particles are expected to scatter off the detector's target nuclei [4,5]. This induces nuclear recoils with mean energies in the order of a few keV.

Experiments that employ the liquid xenon (LXe) time projection chamber (TPC) technology are leading the search for elastic WIMP-nucleon interactions for masses from a few GeV/c^2 up to the TeV/c^2 scale [6–8]. The XENON1T experiment [9] has placed the most stringent upper limit on the spin-independent WIMP-nucleon cross section for masses above $6 \text{ GeV}/c^2$ with a minimum of $4.1 \times 10^{-47} \text{ cm}^2$ at $30 \text{ GeV}/c^2$ and 90% confidence level [6].

Together with Ref. [10], this document reports on the analysis methods employed for the spin-independent dark matter search with XENON1T and subsequent results [11,12]. While the present article describes the techniques of signal reconstruction, event selection, and detector calibration, details on the detector response model, the WIMP signal and background models, and the statistical inference are presented in Ref. [10].

The XENON1T detector was hosted by the Laboratori Nazionali del Gran Sasso and was operated with a total of approximately 3200 kg of ultrapure LXe, with $(2004 \pm 5) \text{ kg}$ contained in the TPC as an active target and the remainder used for shielding. Additional shielding from ambient radioactivity was provided by a 10 m tall and 9.6 m diameter water

tank that served as active Cherenkov muon veto [13] and passive shielding. The TPC was cylindrically shaped with a diameter of 96 cm and height of 97 cm. The top and bottom surfaces were instrumented by arrays of 248 Hamamatsu R11410-21 low-background photomultiplier tubes (PMTs) of 3 in. diameter in total [14,15].

Particles scatter off xenon atoms inducing either nuclear recoils (NRs) or electronic recoils (ERs). The recoil energy is measured by detecting signals from excitation and ionization of xenon atoms, in which the relative contribution of these two channels depends on the recoil type. Excited xenon atoms emit 178 nm scintillation light [16], which is observed by the PMTs and is referred to as the *S1* signal. The ionization electrons are extracted from the interaction site by an electric field and are drifted towards the liquid-gas interface at the top of the TPC where they are extracted into the gas phase by another, stronger electric field and create the proportional scintillation *S2* signal [17]. *S1* and *S2* signals are anticorrelated due to the recombination of electrons with ions creating excited xenon atoms. The two signals are temporally separated by the drift time of the electrons, from which the *z* coordinate of the interaction is reconstructed. The *x* and *y* coordinates are inferred from the light pattern of the *S2* signal on the top PMT array. The *S2/S1* ratio is used to discriminate between ER background (γ and β radiation) and signal-like NR events (WIMPs and neutrons).

XENON1T was operated stably for more than one year from October 2016 to February 2018 (Sec. II). The raw data acquired during this period were converted into physical quantities such as peak amplitude, area, width, etc., by means of a data processor, which is described in Sec. III, together with an event simulation framework used to

* Also at Institute for Subatomic Physics, Utrecht University, Utrecht, Netherlands.

† Also at Coimbra Polytechnic - ISEC, Coimbra, Portugal.

‡ constanze.hasterok@mpi-hd.mpg.de

§ ywei@physics.ucsd.edu

|| xenon@lngs.infn.it

evaluate the processor's performance. The processor also reconstructs the interaction position of each event (Sec. IV) and applies corrections to the measured signals to account for spatial dependencies (Sec. V). The search for WIMP dark matter is based on the selection of a clean sample of single-scatter events (Sec. VI) inside a central fiducial volume which features a reduced background level (Sec. VII). Finally, we present the detector response to NR and ER events in Sec. VIII.

II. DETECTOR OPERATION AND STABILITY

A. Science runs

XENON1T performed its first science run (SR0) between October 2016 and January 18, 2017, when a magnitude 5.7 earthquake temporarily interrupted the detector's operation. In SR0, the drift field yielded 120 V/cm with a rms of 8 V/cm averaged over a 1.3t fiducial volume (FV) (see Sec. VII). The second science run (SR1) was launched February 2, 2017. SR1 featured a lower drift field of 81 V/cm with a rms of 6 V/cm. The field was reduced as a consequence of high-voltage instabilities occurring after the earthquake. Compared to the improvement of the electron lifetime in SR1 (Sec. VA), this field reduction has a negligible impact on the analysis results in view of energy resolution and signal acceptance. The drift field is taken into account in the signal and background models [10], and the electric field values are determined from simulations, using finite element (COMSOL Multiphysics [18]) and boundary element methods (KEMField [19]). SR1 continued until February 24, 2018, marking over one year of stable data taking, as shown in Fig. 1. The blue line presents the accumulated dark matter live time vs calendar time, corrected for the data quality conditions described in Sec. VI A which reduce the

live time by 13.6% and 6.9% in SR0 and SR1, respectively (see Table I).

In addition to science data acquisition, various calibration campaigns were performed as shown in Fig. 1 by the vertical colored bands. The ^{220}Rn decay chain includes the ^{212}Pb β decay, which is employed to calibrate the detector response to low-energy ER events for background modeling and to derive event selection criteria [20]. The metastable isotope $^{83\text{m}}\text{Kr}$ decays into its ground state by emitting 32.1 and 9.4 keV conversion electrons with half-lives of 1.83 h and 157 ns [21], respectively. Those decays are used to monitor spatial and time dependencies of detector signals (Sec. V) and reconstruct signal positions (Sec. IV B). The ^{220}Rn and $^{83\text{m}}\text{Kr}$ sources are injected into XENON1T via the gas purification loop. Shortly after injection, the isotopes are distributed homogeneously throughout the TPC [22]. To avoid potential impact on the position and signal correction from the nonuniformity of $^{83\text{m}}\text{Kr}$ events at the beginning of injection, the first hour data were not used for correction map generation. The acquisition of dark matter data was resumed after the calibration campaigns when the trigger rate had fallen to background level. In these science data, we also observed a trace amount of $^{83\text{m}}\text{Kr}$ events, which presumably was caused by a malfunction of the source valve. These events are not in the energy region of interest for standard dark matter searches but can be used to monitor the detector response throughout SR1.

A deuterium-deuterium plasma fusion neutron generator [23] and a vessel containing a $^{241}\text{AmBe}$ [9] source were immersed into the water tank next to the cryostat to calibrate the detector response to NR events for signal modeling.

Additionally, monoenergetic γ lines from radio impurities in the detector materials (^{60}Co and ^{40}K) and xenon isotopes excited after neutron calibrations ($^{129\text{m}}\text{Xe}$ and

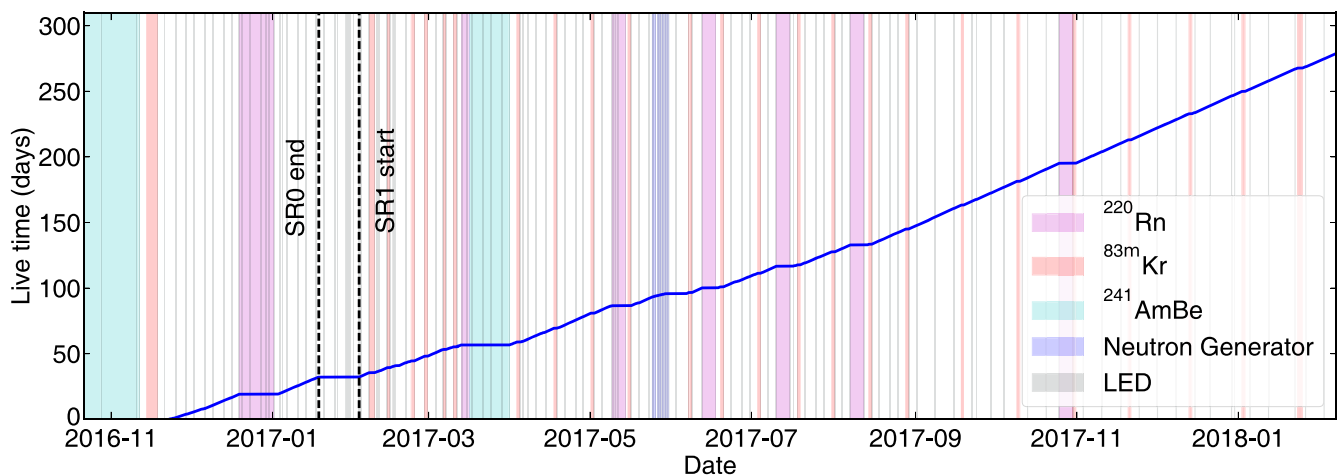


FIG. 1. Accumulated data live time acquired with the XENON1T detector in dark matter search mode and corrected for data quality conditions (Sec. VI A). The black dashed vertical lines mark the end of SR0 and the start of SR1. The various colored bands represent periods of detector calibration with ^{220}Rn (magenta), $^{83\text{m}}\text{Kr}$ (red), $^{241}\text{AmBe}$ (cyan), and neutron generator (blue) sources and with blue LED light (gray).

TABLE I. Comparison of data live time attributes and detector operation parameters among the two science runs of XENON1T. The science data live time reduction refers to the fraction of data removed due to data quality criteria (Sec. VI A). The mean and 1σ values of the detector operation parameters correspond to their temporal distribution, except for the electric fields where the mean and 1σ refer to the spatial homogeneity. Note that PE is the abbreviation of photoelectron.

Data live time attributes	SR0	SR1
Science data live time (days)	32.1	246.7
Science data live time reduction due to data quality (%)	13.6	6.9
Calibration live time (days)	39.2	83.6
Detector operation parameters	SR0	SR1
Drift field: FV-averaged (mean $\pm 1\sigma$) (V/cm)	120 ± 8	81 ± 6
Average electron lifetime (μ s)	290	641
Extraction field: FV-averaged (mean $\pm 1\sigma$) (kV/cm)	8.1 ± 0.1	8.1 ± 0.1
Number [fraction (%)] of excluded PMTs	35 (14.1)	36 (14.5)
Liquid xenon temperature (mean $\pm 1\sigma$) ($^{\circ}$ C)	-96.07 ± 0.04	-96.02 ± 0.02
Xenon gas pressure (mean $\pm 1\sigma$) (bar)	1.934 ± 0.001	1.938 ± 0.001
Charge yield: max. deviation from mean (%)	2	2
Light yield: max. deviation from mean (%)	1	1
Single electron gain (mean $\pm 1\sigma$) (PE/SE)	27.2 ± 0.9	28.2 ± 0.1

^{131m}Xe) as well as the combined 41.5 keV conversion electron signal from ^{83m}Kr were used to monitor the detector stability and calibrate the energy scale over 3 orders of magnitude [9].

B. Detector stability

During both science runs, the detector was operated under stable thermodynamic conditions that guarantee constant signal sizes.

In SR0 (SR1), the 1σ fluctuations of the xenon gas pressure were within 0.05% (0.05%). The pressure was uniformly distributed inside the detector, which also reflected on the stability of the single electron (SE) gain, i.e., the size of S2 signals from SE extracted from the liquid into the gas phase. The SE gain is determined by the gas pressure, the LXe level, and the electric extraction field. SE signals are identified similarly to the procedure described in Ref. [24], and the extracted SE gain has an average of 27.2 photoelectrons (PE)/SE (28.2 PE/SE) and a 1σ time variation of 3.2% (0.5%) in SR0 (SR1) and a SE 1σ width of 7.3 PE/SE (7.4 PE/SE) with a time stability of 11% (5%). The slightly larger variation in SR0 is caused by small fluctuations of the LXe level. The fluctuation of the LXe temperature was within 0.04% (0.02%) for SR0 (SR1). The temperature difference between the top and bottom of the TPC was within 0.25 $^{\circ}$ C. The potential impact of the temperature variation on signal yields was taken into account by the signal corrections in Sec. V. Table I lists the most important detector operation parameters and their 1σ spread around the temporal mean value.

The PMT gains were measured weekly with a pulsed blue light-emitting diode (LED) configured to stimulate the emission of a low number of PEs from the photocathode,

and the gains are extracted using the model-independent approach described in Ref. [25]. Figure 2 shows the gain evolution of three stable PMTs (104, 108, and 133) that are representative for the majority of PMTs in the XENON1T TPC. In SR0 (SR1), 35 (36) PMTs were excluded from analysis [26], with 15 (15) in the top and 20 (21) in the bottom array. Nearly all of these PMTs suffered from vacuum leaks causing decreased performance such as light emission and afterpulses, requiring the bias voltage of the PMT to be lowered, eventually to zero. The criterion for exclusion from analysis is a single photoelectron (SPE) acceptance smaller than 50% [9]. Although vacuum leaks led to the decrease of PMT gains, 19 PMTs with small leaks were operated successfully throughout the science runs, and their gains were monitored and corrected. Two examples, PMTs 61 and 142, are shown in Fig. 2. The gain evolution is modeled empirically by Fermi-Dirac functions that take into account the time when the PMT high-voltage was lowered. The standard deviation of the measured gain with respect to the model is within a few percent for both stable and decreasing PMT gains and is dominated by statistical uncertainties. Compared to the approximately 27% resolution of the PMT response to SPEs [14], the systematic uncertainty of the gain has negligible impact on the energy resolution.

The temporal stability of the S1 and S2 signals is further confirmed by monitoring the light and charge yield (LY and CY, respectively) evolution over time. Using data from monoenergetic sources between 9.4 keV (^{83m}Kr) and 5.6 MeV (^{222}Rn), the measured S1 and S2 signals per incident energy are evaluated following the same procedure as in Ref. [9]. The values are stable throughout both science runs with maximum deviations from the mean of 1% and 2%, respectively.

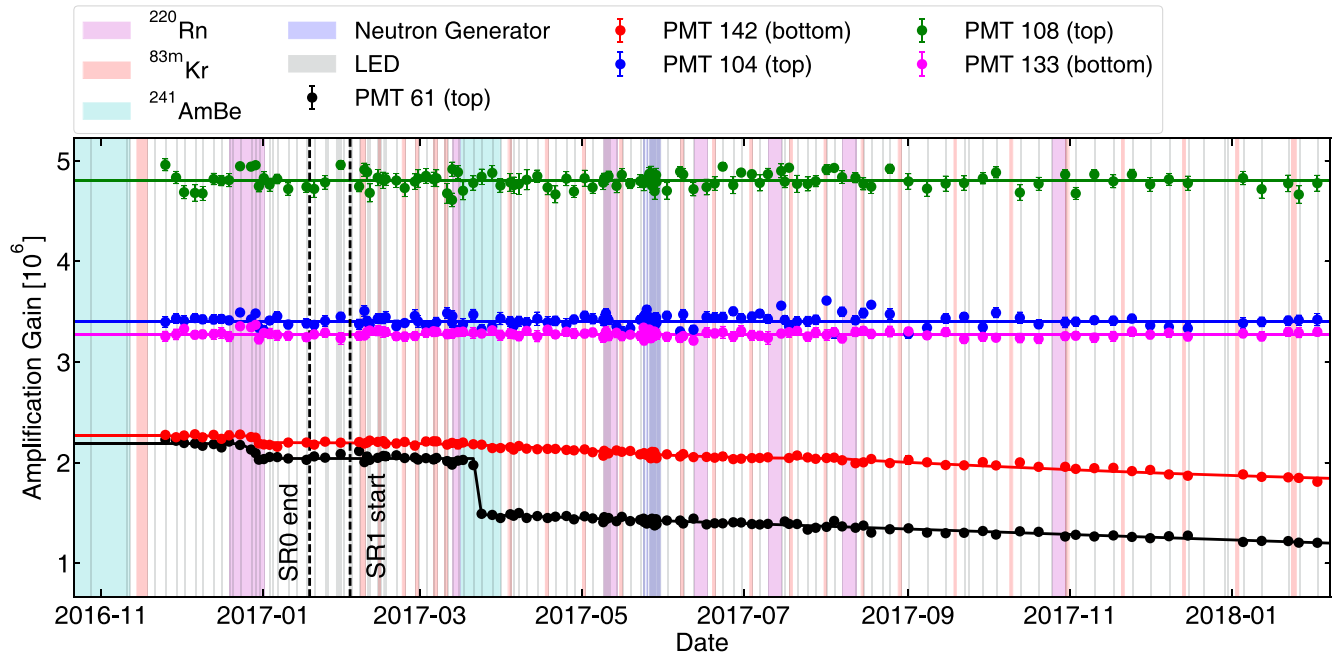


FIG. 2. PMT gains measured by LED calibrations as a function of time for three representative stable PMTs (green, blue, and magenta) and two examples in which the gain decreased due to small vacuum leaks (red and black).

III. SIGNAL RECONSTRUCTION AND SIMULATION

PMT signals exceeding a channel-specific threshold above the baseline, accepting on average 93% of SPE signals, are digitized at a rate of 10^8 samples/sec by the data acquisition (DAQ) system [27]. These signals are referred to as *pulses*. An online event builder groups pulses into events using a simplified algorithm to trigger on S1 and S2 candidates and stores a 1 ms window around each trigger. During offline processing by the custom developed data processor *PAX* [28], pulses are further segmented into smaller intervals, denoted as *hits*, by separating individual signals, which may have been grouped into the same pulse waveform. Hits from different PMT channels are grouped into clusters in time, referred to as *peaks*, and corresponding to individual ionization or scintillation signals. Properties of each peak, such as area, width, and height, are computed by the processor. A peak is classified as S1 (S2) if its waveform rises sufficiently fast (slow) and has at least 3 (4) contributing PMTs. For S1s, only hits with maxima within a 100 ns window centered on the maximum of the sum waveform for all channels are counted for the latter requirement. Finally, each event is searched for a valid S1-S2 pairing, starting with the largest peaks of each type. These pairings are called *interactions*. A further reduction of the processor output is performed by the software package *HAX* [29]. Signal corrections (Sec. V) as well as other higher-level algorithms are included at this level.

The performance of the data processor is studied using emulated PMT signals from a waveform simulator.

The simulation employs data-driven models of XENON1T detector-specific properties like the scintillation light pulse shape, the spatial dependence of the light collection efficiency, the diffusion of electrons during drift, the time profile of SE, PMT afterpulses, the SEs generated by photoionization of impurities, and the electronic noise. The simulated data are validated by comparison to ^{83m}Kr and neutron calibration data and provide the means to optimize the reconstruction algorithms in the data processor and quantify their performance.

The S1 signal reconstruction efficiency is determined from simulated waveforms and is shown in Fig. 3. The efficiency is a function of the number of PMT hits, which at low energies is equivalent to the number of detected photons. The conversion into the S1 peak area (shown in the top axis of Fig. 3) assumes a double electron emission probability at the PMT's photocathode of 21.9% [10]. The efficiency's uncertainty is estimated from the simulation by varying the data-driven model parameters S1 width, PMT afterpulse rates, and rate of photoionization at the gate within their uncertainties. The results are cross-checked with a data-driven method, of which subsets of hits from a large S1 are selected to build an artificial low-energy S1, and the efficiency is calculated based on these low-energy artificial S1s. The result from the data-driven method using ^{220}Rn calibration is in agreement with simulations as shown in Fig. 3. Compared to SR1, the S1 efficiency for SR0 is slightly smaller at the threshold due to a higher PMT noise level in the first third of SR0 dark matter data acquisition that could be reduced by installing low-pass filter boxes at the PMT high-voltage modules [27].

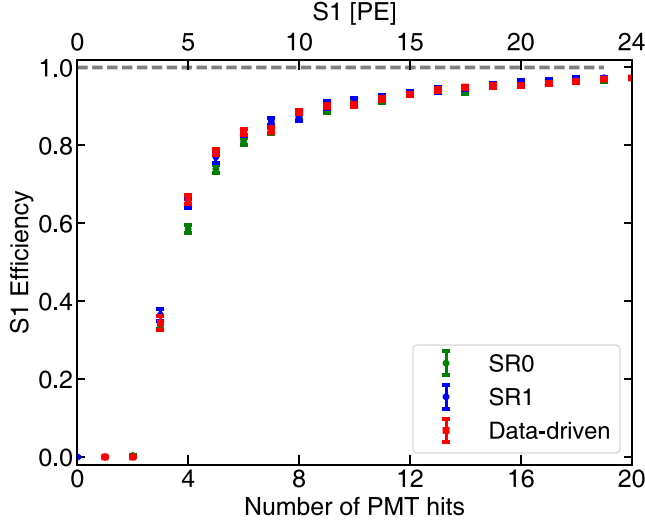
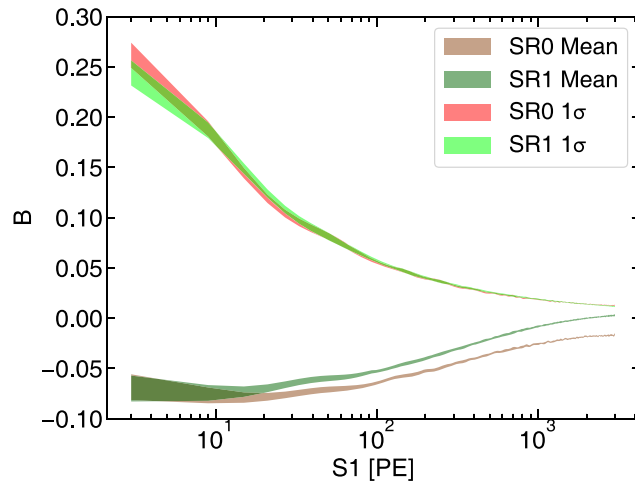


FIG. 3. S1 signal reconstruction efficiency estimated from waveform simulations, as a function of number of PMT hits. Hits are converted into S1 signal size (top axis) using a double PE emission probability at the photocathode of 21.9% [10]. The data-driven efficiency from ^{220}Rn calibration is overlaid for comparison.

The S2 trigger efficiency of the offline event builder is determined by applying the trigger algorithm to simulated S2 waveforms that were generated homogeneously throughout the TPC. The analysis threshold for S2 signal sizes is fixed to 200 PE, where the trigger efficiency yields $(99.8^{+0.2}_{-0.6})\%$ and $(99.4^{+0.4}_{-0.7})\%$ in SR0 and SR1, respectively, and the S2 size is defined using the sum of the signals from the top and bottom PMT arrays. Recently, a data-driven method became available and resulted in about 4% smaller efficiencies at the S2 signal threshold [27]. However, this is expected to have no noticeable impact on the dark matter search since the S1 signal reconstruction efficiency is the dominating parameter determining the detection threshold.



The bias B in reconstructing peak area is estimated by simulating homogeneously distributed S1 and S2 peaks. The expected number A_{true} of detected PEs of a given peak is compared with the number A_{rec} of reconstructed PEs:

$$B = (A_{\text{rec}} - A_{\text{true}})/A_{\text{true}}. \quad (1)$$

The reconstruction bias is predominantly caused by signals from photoionization, particularly coming from the gate electrode, and afterpulse signals in the PMTs. These signals can be merged to or cut off from the primary peak. Figure 4 shows the mean and 1σ width of the Gaussian-shaped distribution of B as functions of S1 and S2 signal sizes. The uncertainty bands are estimated following the same procedure as for the S1 reconstruction efficiency. The mean bias and its width are very similar for the two science runs in the case of S1 signals. However, the two parameters are slightly higher for S2 signals measured in SR1 compared to those in SR0. This difference is most likely caused by a higher PMT afterpulse rate in SR1 as a result of increasing PMT vacuum leaks as explained in Sec. II B.

The S1 reconstruction efficiency and S1 and S2 signal reconstruction biases are input parameters for the signal and background response models [10].

IV. POSITION RECONSTRUCTION AND RELATED CORRECTIONS

Three-dimensional position reconstruction is one of the main advantages of dual-phase TPCs. Most radiogenic background events are located near the boundaries of the TPC and are rejected by selecting a radio-pure inner fiducial volume (Sec. VII). In addition, accurate position reconstruction is required for the development of background models [10] and for position-dependent signal corrections (Sec. V).

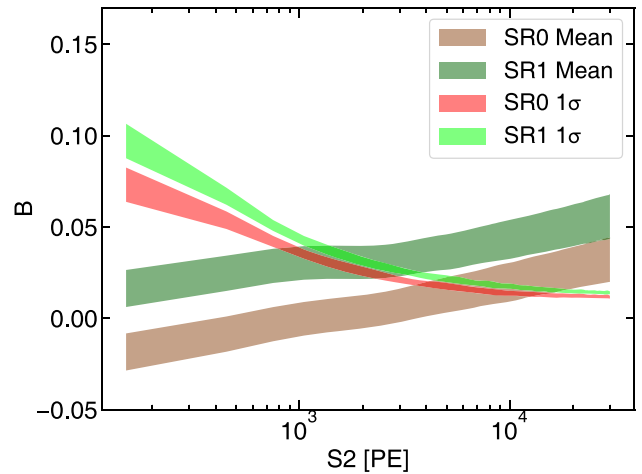


FIG. 4. S1 (left) and S2 (right) signal size reconstruction bias B and its 1σ width as estimated from simulations. The bands are derived by varying the data-driven model parameters within their uncertainties and hence represent the credible region of the shown values.

A. Position reconstruction methods

The vertical coordinate z_{obs} (the subscript obs indicating the position before correction as described in Sec. IV B) of an interaction is determined by the electron drift velocity and the time difference between the prompt S1 and the delayed S2 signal. The origin of the coordinate is at the gate electrode, and the TPC height extends down to -97 cm. Because of the diffusion of the electron cloud during the drift, multiple scatter events with close proximity in z are more difficult to separate at the bottom of the TPC than at the top. By identifying multiple scatter events in NR calibration data, the distance in z for the two interactions is determined. The distribution features a roll-off for small z distances varying with the interaction depth. The minimal value for which two scatters are separable with an acceptance of 50% is found to increase from 2 mm at the gate to 7 mm at the cathode for S2 signals in the region of interest for dark matter searches ($S2 < 25000$ PE).

The horizontal position ($[x_{\text{obs}}, y_{\text{obs}}]$ or $[R_{\text{obs}} = \sqrt{x_{\text{obs}}^2 + y_{\text{obs}}^2}, \phi_{\text{obs}}]$) is obtained from the hit pattern of the S2 signal on the top PMT array. The origin in the x - y plane is set to the center of the TPC. Several position reconstruction algorithms are employed in LXe detectors, such as artificial neural networks (NN) [30], top pattern fit (TPF) [31], support vector machine [32], and statistical light response functions [33]. In XENON1T, a NN is trained using the open-source Fast Artificial Neural Network Library (FANN) [34]. In addition, a TPF algorithm serves as a cross-check for identifying events with poorly reconstructed positions (Sec. VI).

To calibrate the algorithms, data from an optical Monte Carlo (MC) simulation is used. Training data are generated by propagating photons over the full detector geometry [35], which is implemented using the GEANT4 toolkit [36]. Optical parameters such as the refractive index, PTFE (polytetrafluoroethylene, Teflon) reflectivity, xenon absorption length, and Rayleigh scattering length are tuned by matching the simulated light collection to $^{83\text{m}}\text{Kr}$ calibration data.

B. Field distortion correction

Reconstructed spatial distributions exhibit a radial inward bias with increasing depth due to the distortion of the electric drift field and repulsive effects of accumulated negative charges on the lateral PTFE reflector panels that confine the TPC. Figure 5 illustrates this effect integrated over the z_{obs} coordinate. The $x_{\text{obs}}-y_{\text{obs}}$ distribution for $^{83\text{m}}\text{Kr}$ data exhibits a regular geometric distortion from the physical TPC boundary that can be related to the PTFE panel configuration. The distortion is largest at the locations of the 24 panels (black segments) that are not in contact with the ring-shaped copper electrodes surrounding the TPC for drift field shaping [9]. The distortion is smallest at the smaller panels (magenta segments), which are in contact with the electrodes. Figure 6 shows the position of

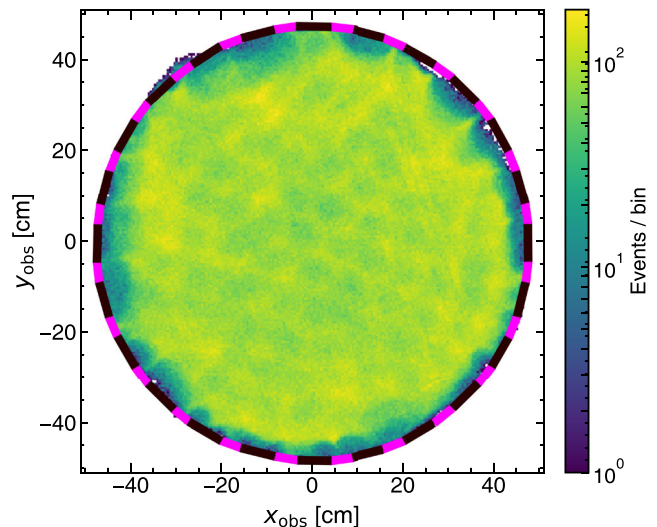


FIG. 5. $x_{\text{obs}}-y_{\text{obs}}$ distribution of $^{83\text{m}}\text{Kr}$ events as reconstructed by the FANN algorithm integrated over z_{obs} . The distortions at high radii coincide with the 24 PTFE reflector panels (black segments) that are not in contact with the ring-shaped electrodes surrounding the TPC for drift field shaping. Magenta segments indicate the panels that are in contact with the electrodes.

the TPC edge in bins of z_{obs} by open markers for several periods throughout the science run, indicating an increasing accumulation of charges during detector operation. The data are derived from the radial distribution of signals from ^{222}Rn progeny on the PTFE surface. Those signals are referred to as surface events [10]. The error bars in the R_{obs} direction indicate the event distribution's radial width, while the error bars in z_{obs} mark the bin width.

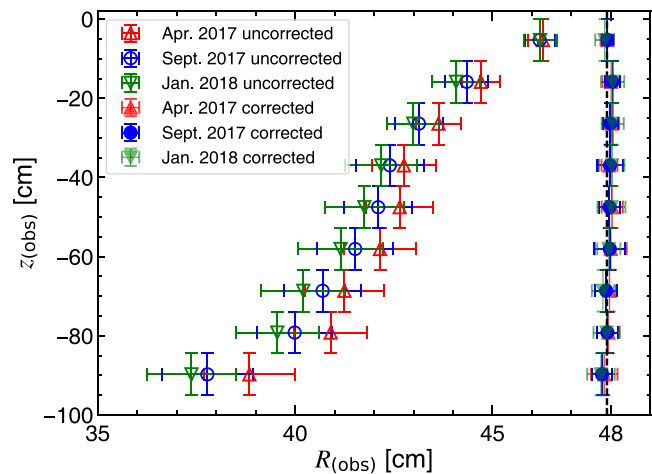


FIG. 6. Maximum TPC radius reconstructed from signals of surface events in three time intervals and in bins of z . Open (filled) markers show radii before (after) position correction. Horizontal error bars indicate the radial width of the event distribution. Vertical error bars mark the z bin width. The black dashed vertical line indicates the geometrical TPC radius.

The first WIMP dark matter results of XENON1T [26] using data recorded in SR0 featured a two-dimensional correction of reconstructed positions ($R_{\text{obs}}, z_{\text{obs}}$) based on a matching of the uniform spatial distribution of $^{83\text{m}}\text{Kr}$ events to the distribution predicted by electric field simulations performed with the COMSOL Multiphysics package [18]. This correction is sufficient for a fiducial mass of up to 1t since the contribution from surface events is negligible in the corresponding volume. During the analysis of SR1 data, an improved understanding of the field distortion and its time evolution was obtained, which led to a $^{83\text{m}}\text{Kr}$ data-driven correction in three dimensions ($R_{\text{obs}}, z_{\text{obs}}, \phi_{\text{obs}}$) for four time intervals throughout SR0 and SR1. To derive the correction, the detector is segmented into 180 bins in ϕ_{obs} and 40 bins in z_{obs} . The event positions in each bin are corrected such that they are evenly spaced in the square of the corrected radial position R^2 . Corrected depths z are subsequently obtained by the geometric relation $z = -\sqrt{z_{\text{obs}}^2 - (R - R_{\text{obs}})^2}$.

The position correction is validated with spatially homogeneous signals from decays of $^{131\text{m}}\text{Xe}$ activated by neutron calibrations as well as nonuniform sources, such as neutron calibration signals, which are compared to MC simulations. A Kolmogorov-Smirnov test [37] yields a goodness of fit p value of 0.55 when comparing $^{131\text{m}}\text{Xe}$ event positions to a uniform distribution.

The impact of the correction is illustrated in Fig. 6. Filled markers indicate the corrected radial position of the TPC edge in bins of z . After correction, the position of surface events coincides with the maximal radial position.

C. Radial position resolution

The position resolution in the radial dimension is dependent on the S2 signal size as well as on the radial event position due to nonfunctional PMTs and light reflection at the TPC boundary.

The two consecutive $^{83\text{m}}\text{Kr}$ decays provide a sample to show the radial position resolution as a function of R for fixed S2 signal sizes. Events in the upper approximately 5 cm of the TPC are selected, as only in that region the two S2 signals can be resolved. The spatial separation of the signals can be neglected with respect to the uncertainty from the reconstruction method. The average path length of the 9.4 keV conversion electrons in LXe is only approximately $10 \mu\text{m}$, and the small half-life of 157 ns does not allow for a reconstructable drift of the atoms by convection. The mean of the distribution of the absolute radial difference ($\Delta R = |R_{32.1} - R_{9.4}|$) between the 32.1 and 9.4 keV signals is shown in bins of R^2 in Fig. 7. The vertical error bars display the distribution's standard deviation. Note that the position reconstruction uncertainties from both decays are convoluted in ΔR . Hence, a direct comparison to the resolution of single events is not possible. While the precision in reconstructed radial positions is in the order

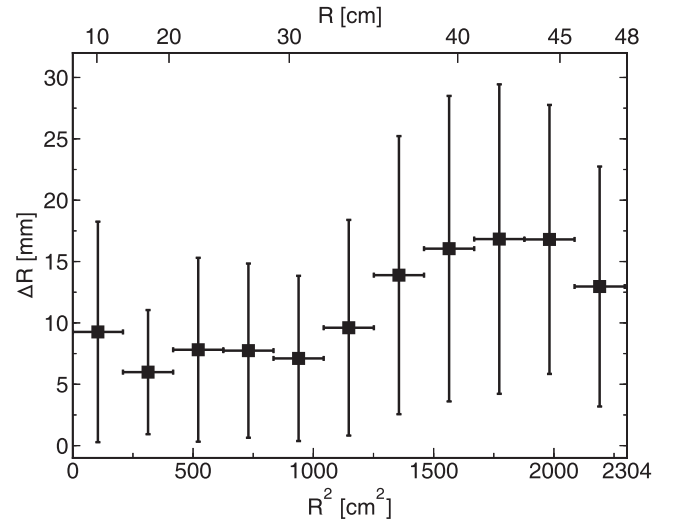


FIG. 7. Absolute radial difference (ΔR) reconstructed between 32.1 and 9.4 keV signals from $^{83\text{m}}\text{Kr}$ in bins of R^2 . The corresponding R scale is shown on the upper horizontal axis. The vertical error bars represent the standard deviation of the ΔR distribution, while horizontal error bars indicate the R^2 bin width.

of approximately 1 cm for $R < 35$ cm, the performance declines by a factor of 1.5 toward larger radii. This is caused by nonfunctional PMTs and light reflection at the TPC boundary as mentioned above.

In addition to the radial dependence of the position resolution for fixed S2 signal sizes, we also investigated its dependence on S2 signal sizes for a fixed position. Surface events provide a sample at the maximum TPC radius and cover a large range of S2 signals down to less than 200 PE due to charge loss. Figure 8 shows the standard deviation σ_R of the radial distribution in bins of S2 signal size. The uncertainties are derived from the Gaussian fit, and

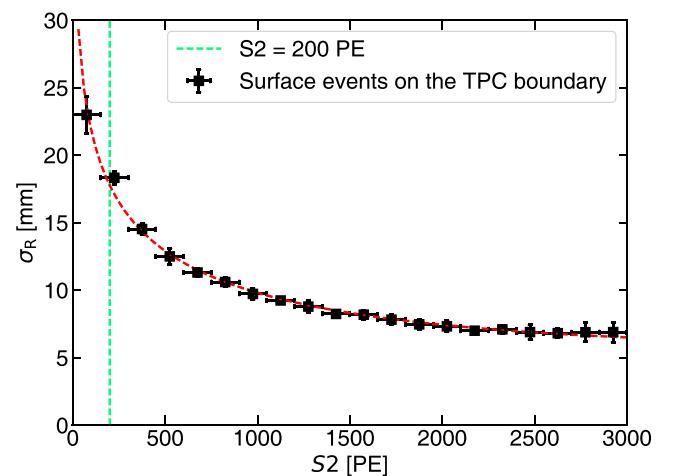


FIG. 8. Radial position resolution σ_R in bins of S2 signal size for surface events. Vertical uncertainties are derived from the Gaussian fit, while horizontal error bars mark the S2 bin width. The red dashed line indicates the best fit of an empirical function.

horizontal error bars mark the S2 bin width. σ_R yields 1.9 cm at the S2 analysis threshold of 200 PE and decreases to values less than 0.8 cm for large S2 signal sizes. The resolution is limited by the accuracy of the optical MC simulation used to train the FANN. Note that the resolution is poorest for surface events due to their location at high radii (Fig. 7) and is expected to improve for interactions in the center of the TPC.

The presented position resolution studies are considered in the background and signal models for final inference of the dark search data [10].

V. SIGNAL CORRECTIONS

The size of the recorded S1 and S2 signals depends on the event location in the detector due to various position-dependent effects, such as electron attachment to impurities in the LXe target, light collection efficiency, field inhomogeneities, variations of the thickness of the proportional scintillation region and nonfunctioning PMTs. In the following, the corrections applied to S1 and S2 signals in order to account for these effects are explained. The corrected signals are denoted as cS1 and cS2.

A. Electron lifetime correction

The loss of ionization electrons due to attachment to electronegative impurities (e.g., O_2) [38] in LXe is a function of the drift time and follows an exponential law with the electron lifetime τ_e as a decay parameter. This effect is the most important correction for S2 signals and is highly dependent on the impurity concentration in the target. Since the xenon is continuously purified, τ_e is a parameter changing over time and has to be continuously monitored [9].

The electron lifetime is evaluated in intervals of two to three weeks by measuring the 41.5 keV signal from the two consecutive ^{83m}Kr decays as a function of electron drift time. Additionally, τ_e is estimated from monoenergetic α decays of ^{222}Rn and ^{218}Po observed in background data that provide sufficient statistics on a daily basis. The τ_e values from these two methods are shown in Fig. 9 and feature an offset of up to 10% of unknown origin that scales with the xenon purity. The best hypothesis of the discrepancy is related to the small inhomogeneity of the drift field. The ionization yield of α decays has stronger field dependence than ERs and NRs; hence, the measured τ_e are different. Because the energy from ^{83m}Kr decays is closer to the region of interest for dark matter searches compared to α decays, furthermore, better energy resolution up to the MeV scale and better discrimination between ER and NR signals (Sec. VIII) can be achieved when applying τ_e derived from ^{83m}Kr decays in the S2 signal correction, we decided to use the τ_e from ^{83m}Kr decays for the final corrections.

The temporal fine structure of the electron lifetime evolution is modeled based on α decays by fitting a model that estimates the evolution of impurity concentrations in the gaseous and LXe phase and takes into account various detector operation parameters like the detector's cooling power and the xenon gas flow in the purification system [39]. The model's best fit (uncertainty) is shown by the gray line (band) in Fig. 9. During the two science runs of XENON1T, several decreases of electron lifetime that coincide with releases of impurities due to changes in the above-mentioned parameters were observed.

The absolute scale of the electron lifetime model is derived by relating the $1/\tau_e$ data points from the two methods by a linear function which is used to scale from the α measurement to the ^{83m}Kr measurement. The final

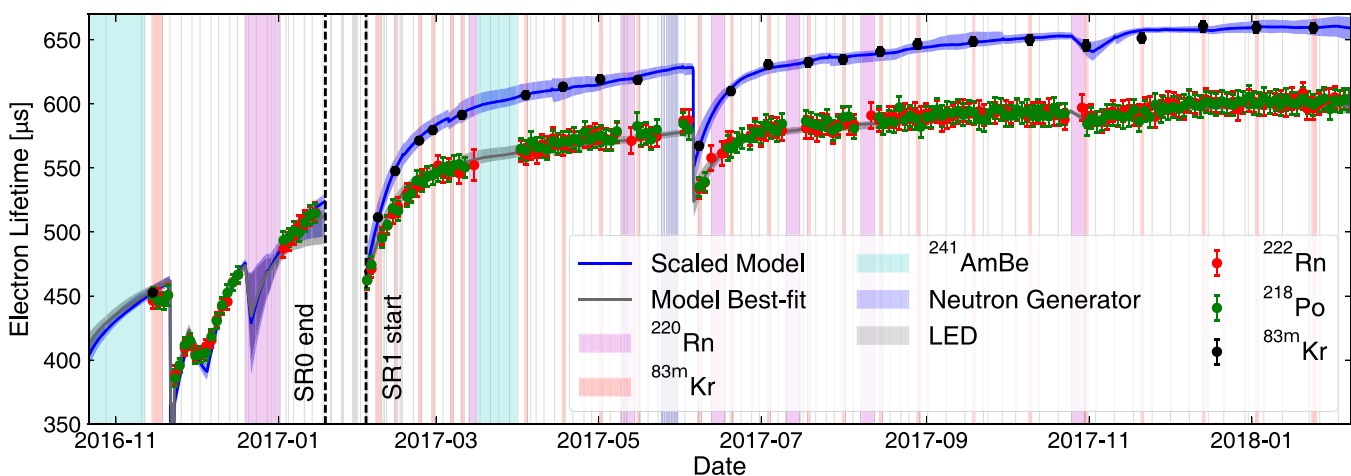


FIG. 9. Electron lifetime evolution during the two science runs measured from ^{83m}Kr (black), ^{222}Rn (red), and ^{218}Po (green) decays. Decreases are caused by releases of impurities due to changes in detector operation parameters like the detector's cooling power and the gas flow in the purification system. The temporal fine structure is modeled based on the α measurements (gray line), while the absolute scale of the electron lifetime is determined from the ^{83m}Kr measurement.

electron lifetime model used for S2 signal correction is shown by the blue solid line together with its uncertainty band.

During SR1, τ_e leveled off at about 650 μs , corresponding to an oxygen equivalent impurity concentration of about 0.5 ppb limited by outgassing materials and the flow in the gas purification circuit.

B. S2 spatial correction

The proportional scintillation signal S2 is produced between the liquid-gas interface and the anode electrode. The fraction of the total S2 signal measured by the top PMT array is $(63 \pm 2)\%$. The signal detected in the top array is highly localized, while the bottom PMT array provides a more uniform distribution that is more resilient to effects from nonfunctional PMTs or variable light collection efficiency. For this reason, only the corrected bottom array signal, cS2_b , is used in the final inference of dark matter search data.

Positional variations in the S2 size arise due to distortion of the electric field at the liquid-gas interface induced by subsidence of the anode caused by its weight, impacting the electron extraction efficiency. These variations are corrected using the 41.5 keV charge signal of $^{83\text{m}}\text{Kr}$. The $x_{\text{obs}}-y_{\text{obs}}$ distribution of S2 signals is fit by a two-dimensional second-order polynomial. The best-fit value of the function's center for S2 signals observed by the bottom array is displaced from the origin by about 1.5 cm to negative x_{obs} and y_{obs} values. This displacement indicates a slight tilt of the TPC. The extraction efficiency is approximately 20%–30% lower at the edge of the detector compared to the center, while the average value yields 96% [9].

C. Light collection efficiency

The LY is impacted by the light collection efficiency ϵ_L , i.e., the number of photons that hit a PMT photocathode per photon emitted at the interaction site; the photon yield (PY), i.e., the number of generated photons per incident energy ϵ ; the PMT quantum efficiency ϵ_{QE} , i.e., the probability that one photon hitting the photocathode induces one PE; and the collection efficiency ϵ_{CE} of photoelectrons within the PMT:

$$\begin{aligned} \frac{S1(R, \phi, z, \epsilon, F)}{\epsilon} &= LY(R, \phi, z, \epsilon, F) \\ &= \epsilon_L(R, \phi, z) \cdot PY(\epsilon, F(R, \phi, z, t)) \cdot \epsilon_{\text{QE}} \cdot \epsilon_{\text{CE}}. \end{aligned} \quad (2)$$

ϵ_L is affected by the number of photon reflections before reaching a PMT photocathode and is, therefore, spatially dependent. The PY depends on the energy and the drift field F , which both impact the electron-ion recombination [40]. The field features variations at the TPC edges, which

vary in time t due to charge accumulation on PTFE surfaces (see Sec. IV).

To correct for the spatial dependence of S1 signals, a three-dimensional correction map is derived from the 32.1 keV signals in $^{83\text{m}}\text{Kr}$ calibration data. The mean of the S1 distribution is evaluated in discrete (R, ϕ) regions and in slices of z and is normalized to its average $\langle S1 \rangle$ across the TPC in order to obtain the relative light collection efficiency L_c that removes the spatial dependencies:

$$\begin{aligned} \frac{LY(R, \phi, z, \epsilon, F)}{L_c(R, \phi, z)} &= \langle LY(\epsilon, F) \rangle \\ &= \langle \epsilon_L \rangle \cdot \langle PY(\epsilon, F(R, \phi, z, t)) \rangle \cdot \epsilon_{\text{QE}} \cdot \epsilon_{\text{CE}}. \end{aligned} \quad (3)$$

The number of bins of the correction map was optimized in each dimension by limiting the maximum variation to be about 2.5% between two adjacent bins. The correction not only averages out the spatial dependence of ϵ_L but also accounts for the spatially dependent PY introduced by field inhomogeneities.

$^{83\text{m}}\text{Kr}$ decay energies lie beyond the region of interest for WIMP searches, and lower-energetic events are less sensitive to changes in the field. Hence, a small bias of approximately 2% [40] is introduced when applying L_c to the WIMP search region, and L_c varies in time by up to 6% due to evolving field inhomogeneities (Sec. IV B). To remove this bias, the spatial distribution of the CY from the 41.5 keV $^{83\text{m}}\text{Kr}$ signal is used to map out local and timely field variations and decouple those from both the L_c and electron lifetime corrections. Since the CY is correlated with the electron lifetime, this procedure is repeated iteratively until convergence is observed resulting in a time stability of L_c within 1.2%.

Figure 10 shows L_c measured as a function of z (left) and ϕ (right) for bins in R . The largest variation is observed along z with a maximum L_c at the bottom center of the detector where the solid angle to the bottom PMT array is largest.

VI. SELECTION CRITERIA AND THEIR ACCEPTANCES

This section describes the criteria applied to the dark matter search data for selecting single scatter events in the region confined by $\text{cS1} \in [3, 70]$ PE and $\text{cS2}_b \in [50, 7940]$ PE corresponding to the energy region of interest, $[1.4, 10.6]$ keV_{ee} ($[4.9, 40.9]$ keV_{nr}) [6]. Note that the cS2_b was used in the analysis due to its uniform distribution as explained in Sec. V B. 50 PE of cS2_b correspond to 100% acceptance for the events with S2 over its 200 PE trigger threshold.

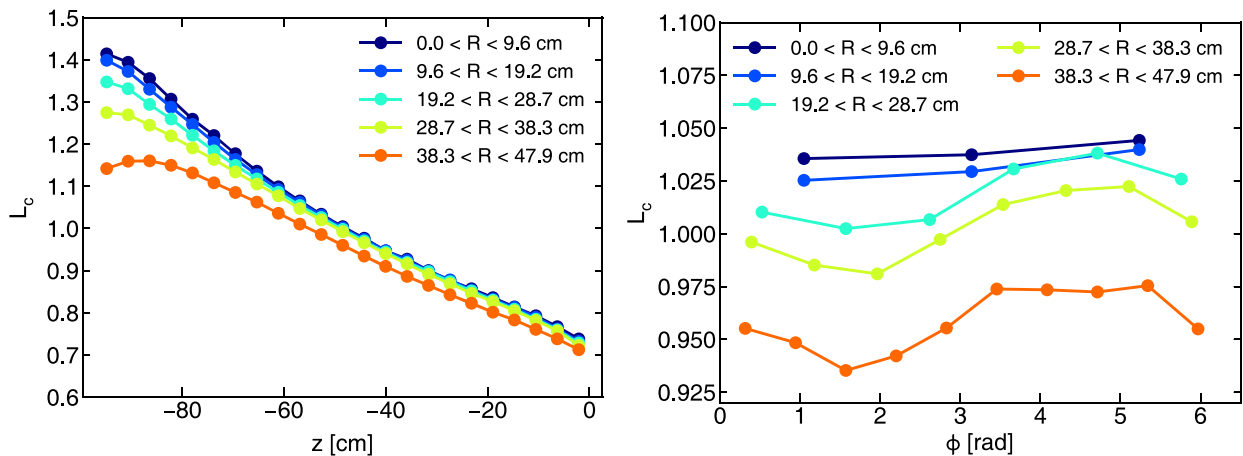


FIG. 10. Spatial dependence of the relative light collection efficiency L_c on z (left) and ϕ (right) for different radial bins (color code). Data points are connected by straight lines to guide the eye.

A. Data quality selection

Operational conditions during data acquisition are necessary for the rejection of certain time periods, regardless of the properties of the events contained within. The corrected live time and respective acceptance after the incremental application of four criteria are summarized in Table II.

The *DAQ veto* ensures that all channels in the DAQ system are able to record data. If this is not the case, a system-wide busy condition is issued. The start and stop times of the busy signal are saved in the data stream, ensuring that those time periods can be removed at the analysis level. The DAQ veto rejects about 6% (1%) of data in SR0 (SR1), with the increased dead time in SR0 caused by nonoptimized DAQ settings during the first XENON1T runs.

The active Cherenkov *Muon veto* triggers if at least eight PMTs in the water tank record signals larger than one PE within a 300 ns time window. Under these conditions, a muon tagging efficiency of 99.5% is achieved, while muon-induced shower events are identified with a probability of 43%. These efficiencies are determined from simulations, and the events are required to exhibit at least one produced neutron that has high enough energy (>10 MeV) to reach the TPC [13]. To remove signals from potential secondary interactions in the TPC, a muon veto trigger is searched in a

window of $[-2$ ms, $+3$ ms] around each TPC event, which is rejected in case of success. The time range is conservatively determined from simulations [13] and data. In addition, all data are removed where the muon veto is inactive. In total, about 2% (1%) of live time after the DAQ veto is removed in SR0 (SR1) due to the muon veto criterion. The muon veto reduces the expected muon-induced neutron background rate by a factor of 2.5.

The XENON1T PMTs can emit bursts of light as previously observed in independent measurements [14]. In SR0 (SR1), a total of 8 (179) light flashes were observed, causing short periods of high pulse rates throughout the TPC from both the primary light and secondary interactions. These incidents are removed from the data by a *Flash veto* that scans the pulse rates for each PMT channel and identifies sudden, drastic increases. A conservative time window of 10 s before and 120 s after each flash is rejected. After the application of the DAQ and muon veto, the flash veto removes 0.04% and 0.12% of live time in SR0 and SR1, respectively. Flashes trigger in most cases the DAQ busy signal. Hence, the criterion is highly correlated with the reduction of live time due to the DAQ veto.

The *S2 tails* criterion addresses delayed S2 signals, e.g., from delayed electron extraction or photoionization on materials and impurities, which are generated especially after large S2s and do not correspond to physical interactions. Those can reduce the detector sensitivity to low-energy interactions for several milliseconds. For each event, the discrimination variable $S2_{\text{pre}}/\Delta t$ is defined as the ratio of the S2 size of a preceding event divided by the time difference to that event, where the preceding 100 events are scanned and the maximum of the parameter is stored. Figure 11 shows the distribution of primary S1 signal sizes vs $S2_{\text{pre}}/\Delta t$ in single scatter events (Sec. VID) of background data, while a subset of noise rejection and reconstruction requirements (Sec. VIB) has been applied as

TABLE II. Summary of data live time and respective acceptance after incrementally applying data quality requirements in the shown order.

Data quality criterion	Live time (days)		Incremental acceptance (%)	
	SR0	SR1	SR0	SR1
Without cut	37.2	264.8	100	100
DAQ veto	34.2	261.6	92.1	98.8
Muon veto	33.5	259.1	90.0	97.8
Flash veto	33.5	258.8	89.9	97.7
S2 tails	32.1	246.7	85.8	93.0

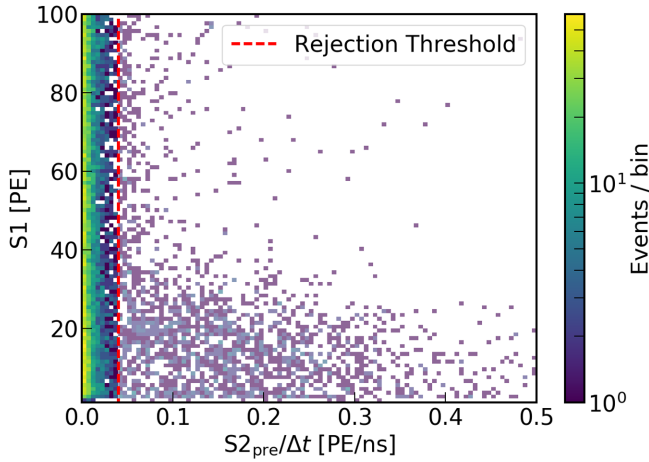


FIG. 11. Distribution of S1 signal sizes vs $S2_{\text{pre}}/\Delta t$ as measured in ^{220}Rn calibration data. The threshold in $S2_{\text{pre}}/\Delta t$ above which events are rejected regardless of their own properties is indicated by the red line.

a preselection. Ionization signals from preceding events can be misidentified as an S1 or S2 of an interaction in subsequent events and therefore appear as a horizontal population in the figure. The $S2_{\text{pre}}/\Delta t$ threshold above which an event is rejected is set to 0.04 PE/ns, which is chosen to remove the most intense region of increased activity while maintaining as much live time as possible. This removes 4% (5%) of live time in SR0 (SR1) in addition to the previously mentioned live time reductions.

B. Noise rejection and reconstruction requirements

A set of conditions is imposed to remove events that are either falsely reconstructed, members of known background populations, or generally of low quality:

- (i) If an event contains a large integral (>300 PE) of signals prior to the primary S2 excepting the primary S1, the event is deemed “noisy” and is removed.
- (ii) The contribution of one channel to an event’s S1 is not allowed to exceed 5% of the S1 plus an offset of four PEs. This criterion prohibits that a single channel exhibiting a PMT malfunction dominates the signal. Typical causes for failing the condition are PMT afterpulses or light emission.
- (iii) S2s originating from single electron signals can be misclassified as S1s at the data-processor level. This can result in events in which two S2 signals, one misidentified and one lone signal, are randomly paired as an interaction. Lone S2 signals originate from delayed electron extraction and pileup and therefore do not feature a corresponding S1. Two machine-learning classification algorithms, a boosted decision tree and a random forest from the SCIKIT-LEARN PYTHON package [41], are employed to

reclassify S1s based on the most important peak properties of width, area, rise time, and signal fraction detected in the top PMT array. Training samples for good S1s are selected from high-quality background events in the ER band and from sampling hits from larger S1 signals, effectively creating smaller signals in the region of interest. Single electron S2 training samples are created by selecting S2 peaks that are isolated from other signals by at least $10\ \mu\text{s}$. The threshold for removing an event, placed on the classifier’s normalized voting, was optimized to achieve a reduction of S2 signals misclassified as S1 signals by a factor of 5.

- (iv) The almost constant fraction of light from S2 signals observed by the top PMT array is used to reject background caused by interactions in the gas phase above the anode electrode or from misreconstructed events. High-quality calibration events are used to model the distribution of true interactions in the liquid xenon depending on the S2 size. Events that exhibit an S2 light fraction in the top array that is smaller or larger than the 99% quantiles of the distribution are rejected.
- (v) Reliable position reconstruction is ensured by demanding the reconstructed $x_{\text{obs}}-y_{\text{obs}}$ coordinates to be consistent with the S2 hit pattern on the top PMT array. This criterion predominantly suppresses pileup of delayed electron signals, double scatters, or events that are misreconstructed at the wrong $x_{\text{obs}}-y_{\text{obs}}$ position, often due to nonfunctional PMT channels. The likelihood of the observed pattern given the position is computed using the same optical MC simulation as employed for the training of the position reconstruction algorithms.
- (vi) To further suppress anomalous $x_{\text{obs}}-y_{\text{obs}}$ reconstruction, events are removed if the difference in the reconstructed positions for the two reconstruction algorithms exceeds the upper 99% quantile of the position difference distribution defined in dependence of S2 size. This distribution is extracted from high-quality calibration data.

The acceptance of each cut described above is evaluated individually based on control samples either from calibration or background data. All data quality criteria together accept $(95 \pm 4)\%$ of true signals that fall into the region of interest.

C. S1-S2 signal correlation requirements

In each time window that is allocated to an event, the processor considers the largest S2 candidate and the S1 with the largest PMT coincidence level before the S2 as the primary interaction. This and the following criteria suppress pileup effects, double scatters, or accidentally pairing lone S1 and S2 signals that arise in charge- and light-insensitive detector regions:

- (i) The width of the S2 peak is required to be correlated with the event's drift time t_d and S2 signal size as expected from physical interactions due to diffusion of the electron cloud. The time interval r_{50} in which 50% of the S2 peak is contained can be modeled by [42]

$$r_{50}^{\text{theory}} = \sqrt{\frac{3.64 \cdot D \cdot t_d}{v_d^2} + w_0^2}, \quad (4)$$

with the free parameters drift velocity v_d , diffusion constant D , and an adjustable constant w_0 . In case of a Gaussian SE response, w_0 would be equal to the width of very shallow S2s. The factor 3.64 adjusts the diffusion model, originally defined for S2 signal widths of one standard deviation, to the 50% width r_{50} . The three parameters are determined from calibration data, and the model is implemented in the waveform generator. Using simulated waveforms, a large data set that extends across the full S2 signal size spectrum, which is not entirely covered by the calibration data, is generated. Figure 12 shows the measured r_{50} distribution normalized to r_{50}^{theory} as a function of S2 signal size for simulated (top) and calibration (bottom) data. The calibration data include both ^{220}Rn and $^{241}\text{AmBe}$ data in order to provide statistics at low as well as at high S2 signal sizes. The broadening of the distribution for small S2 signal sizes is caused by binomial fluctuations of the number of electrons contributing to the signal. The S2 signal width requirement is constructed based on the quantiles (yellow points in Fig. 12) in simulated data such that physical interactions are accepted in more than 99% of the cases. The selection criterion derived from the quantiles is marked by the red lines, and events not contained in between are rejected.

- (ii) The PMT hit pattern of the S1 signal depends on the event position due to geometrical effects. Because the x - y position is extracted from the S2 signal and the z position is computed from the drift time, compatibility of the S1 signal's PMT pattern (S1 pattern) with this position represents an independent confirmation of the interaction pairing. The likelihood for an S1 pattern to originate from the reconstructed event position is computed by means of optical MC simulations. The selection criterion is tuned such that S1-S2 signal pairs from physical interactions defined by calibration control samples are accepted with a probability larger than 99%.
- (iii) Similarly, the fraction of the S1 signal detected by the top PMT array is required to be correlated to the interaction depth due to geometrical effects. The probability of a photon being detected in the top array is evaluated for each position in the target

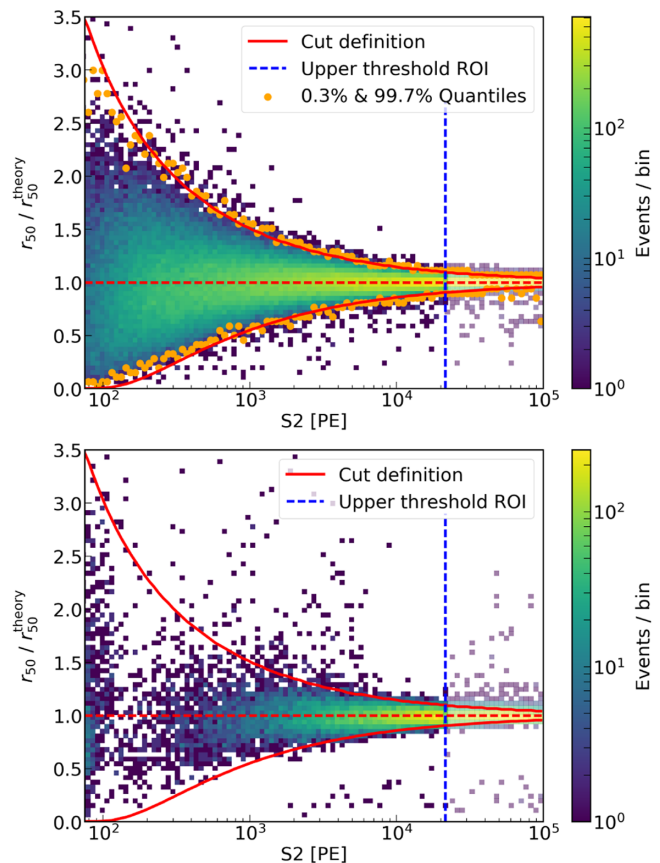


FIG. 12. Distribution of the measured S2 width parameter r_{50} normalized to the width r_{50}^{theory} expected from diffusion as a function of S2 signal size for simulated (top) and combined ^{220}Rn and $^{241}\text{AmBe}$ calibration data (bottom). Yellow points mark the 0.3% and 99.7% quantiles for simulated data, and red solid lines show the chosen cut definition. The red dashed line marks $r_{50} = r_{50}^{\text{theory}}$.

volume by means of optical MC simulation. This number is used to construct a binomial distribution from the total number of photons detected in both PMT arrays for a given event. If the actual fraction of photons detected in the top array falls into the extreme tails of the binomial distribution ($p < 0.001$), the event is removed.

The described S1-S2 signal correlation criteria have a roughly flat acceptance in the region of interest (Fig. 14) with an average value of $(96 \pm 3)\%$ that estimated from control samples in background and calibration data.

D. Single scatter requirements

Given the small expected scattering cross section of dark matter particles, the probability that a WIMP scatters more than once in the TPC target volume is negligible. Hence, the identification of multiple scatter events is a powerful discriminator between signal candidates and background from radiogenic neutrons that induce identifiable multisite events with a probability larger than 80%:

- (i) In the S1 channel, events in which a second S1 signal is accidentally present in the waveform are identified by searching for an additional S1 signal within one maximal drift time interval, i.e., the time interval that corresponds to the height of the TPC, before the primary S2. The drift time of this constructed alternative interaction is tested for a correlation with the primary S2 by using the S2 width selection criterion described in the previous section. If the S2 signal is compatible with an interaction at the depth indicated by this alternate pairing, the event is removed.
- (ii) In the S2 channel, events are removed if a second S2 of sufficient size is identified, with the size threshold defined as a function of the primary S2. The requirement results in a multiple scatter rejection efficiency of about 96% and a single scatter acceptance of 99% as estimated from simulations and NR calibrations.

The described criteria accept more than 99% of true single scatter signals that fall into the region of interest.

E. ^{220}Rn calibration data specific criteria

The first publication of XENON1T results [26] featured a component of the background model that was flat in $cS1$ vs $cS2_b$ space. It was motivated by the observation of events in ^{220}Rn calibration data that featured $cS2_b$ values smaller than expected for regular ERs [26]. Those events are referred to as anomalous leakage. The increased statistics in calibration data acquired in SR1 enabled more detailed studies and a better understanding of anomalous leakage events. Selection criteria were developed to remove this type of events from ^{220}Rn calibration data, obviating the inclusion of an anomalous component to the background models [6]:

- (i) ^{220}Rn calibration data have an increased lone S1 rate in two regions below the cathode where the isotope is injected into the TPC. These S1 signals can get accidentally paired to lone S2 signals. Figure 13 shows the fraction of lone S1 signals between 0 and 70 PE observed by each PMT in the bottom array during SR1 ^{220}Rn calibrations. Events for which the fraction of the S1 signal observed by the PMTs (131, 138, 146, 147, 243, 236, and 237) close to the injection points (magenta star) exceeds an S1 signal size-dependent threshold are rejected. The threshold is defined such that 99% of ^{220}Rn events that pass all previously described selection criteria are accepted.
- (ii) Additionally, lone S1 signals feature larger widths in time if they originate from misidentified single electron signals. The width in which 90% of the signal peak is contained is required to be smaller than an S1 signal size-dependent threshold. The threshold is optimized in the same way as the criterion for the ^{220}Rn injection points.

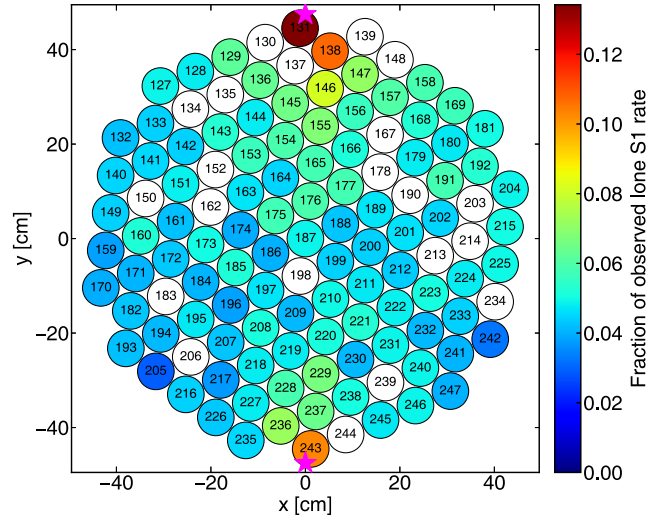


FIG. 13. Fraction of lone S1 signals (color scale) between 0 and 70 PE observed by each PMT (circles) in the bottom array during ^{220}Rn calibrations. Numbers in the circles indicate PMT IDs. Magenta stars mark the points where internal calibration sources are injected into the detector below the cathode. White circles represent PMTs that were nonfunctional in SR1.

The ^{220}Rn -specific selection criteria have been conservatively applied to background data in order to exclude potential artifacts remaining in datasets acquired shortly after ^{220}Rn calibration periods due to the remaining isotopes inside the connection pipes. The combined acceptance of physical interactions is $(98 \pm 1)\%$.

F. Signal acceptances

Figure 14 shows the cumulative signal acceptances of the described categories of selection criteria as functions of uncorrected S1 (left) and S2 (right) signal sizes. The smooth curves are determined by fitting a first-order polynomial function to the data points in the S1 signal space and a function of the form $(a + b \cdot S2)(1 - c \cdot \exp(-S2/d))$ in the S2 signal space. The uncertainty bands account for the statistics present during the acceptance estimation from data as well as for the systematics of the fit which are derived from the deviation of the data points from the best-fit line. Correlations between most of the selection criteria are found to be negligible, and the acceptances are estimated from control samples in background and calibration data. Four requirements on S1 signals (signal size contributed by a single PMT, the likelihood of PMT hit pattern given the reconstructed position, the consistency of signal fraction seen in the top PMT array with the interaction depth, and the single scatter criteria) as well as the ^{220}Rn -specific selection criteria show correlations with other criteria. Hence, their acceptances are conservatively estimated by applying them consecutively to ^{220}Rn calibration data after all other criteria.

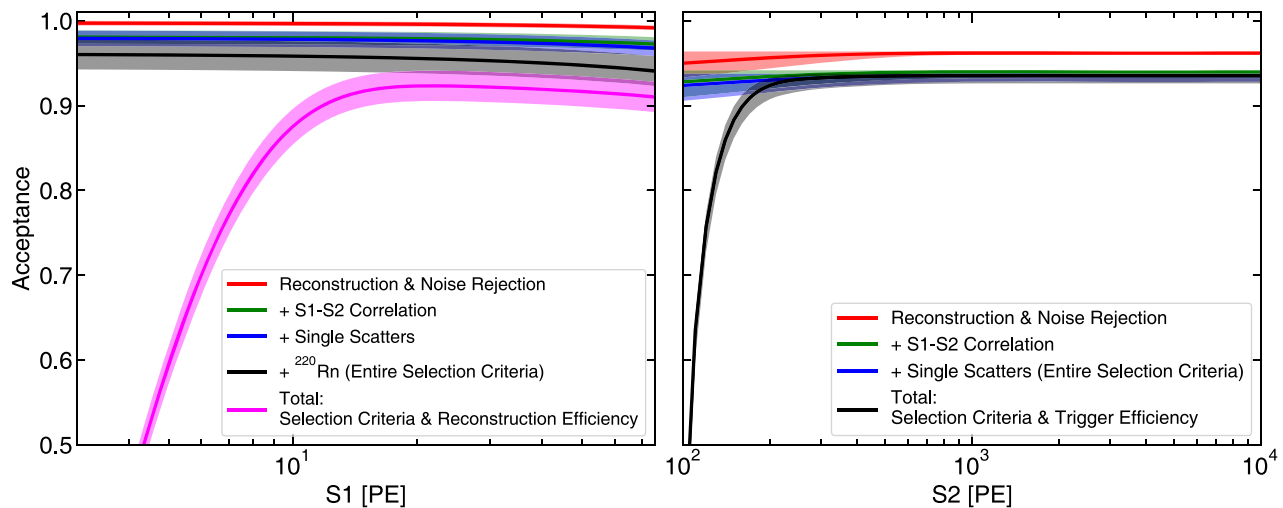


FIG. 14. Evolution of the acceptance of S1 and S2 signals from physical interactions by incremental application of the three categories of selection criteria described in the text as a function of uncorrected S1 and S2 signal sizes. The total acceptance is also shown after considering the reconstruction efficiencies introduced in Sec. III.

In the S1 signal space, the S1-S2 signal correlation conditions and the ^{220}Rn -specific criteria have the largest impact, and the entire selection acceptance is given by the black curve. In the S2 signal space, the acceptance is primarily influenced by the data quality and the S1-S2 signal correlation criteria. The ^{220}Rn -specific criteria have no impact since they do not affect S2 signals. The entire signal selection acceptance averaged over the region of interest yields $(91 \pm 4)\%$.

The total acceptance was overlaid together in Fig. 14 after considering the reconstruction efficiencies that were introduced in Sec. III, which were passed to the statistical inference for dark matter searches [10].

VII. OPTIMIZATION OF THE FIDUCIAL VOLUME

The fiducial volume is optimized using the $[R, z]$ distributions of the background components, shown in Fig. 15. The distributions are evaluated considering only the $cS1-cS2_b$ region in which a $50 \text{ GeV}/c^2$ WIMP signal is expected to have the highest significance compared to the total background. Additionally, the models are limited to $R < 44 \text{ cm}$ in order to exclude biases from signal corrections related to the decreasing position resolution beyond that region (Sec. IV C). ER background (top left) is mostly from homogeneously distributed ^{214}Pb β decays and inhomogeneously distributed emission from detector materials and is modeled with kernel density estimation (TKDE class from the ROOT framework [43]) using blinded WIMP search data. Charge losses exhibited by surface events (bottom left) produce artificially small S2 signals, which can be misreconstructed farther inside the TPC. These events are modeled using sidebands just outside the DM search region. NR events (top right) are only expected from

neutron emission of radio impurities in the detector materials and are modeled with a MC toolkit based on GEANT4 [36]. The simulated distributions are scaled to the expected number of background events predicted with the SOURCES-4A simulation package [44,45] and verified in data with multiple-scatter NR events. The bottom right panel in Fig. 15 shows the total background distribution with all components scaled to their expected relative intensities.

The FV is restricted in depth to $z \in [-94, -8] \text{ cm}$ in order to exclude misreconstructed events from the gas volume and events originating near the cathode, where the electric field exhibits a higher nonuniformity. The maximum radius is set as 42.84 cm in order to exclude biases from field inhomogeneity and to reject the bulk of surface background events. The radial boundary could have been moved slightly inward to reject almost all surface background events, but it was intentionally set such that about 100 surface events are included in order to provide enough statistics to scale the background model [10]. Because the only spatial variable used in statistical inference is R , the outer FV regions are further restricted in z such that each bin in R exhibits a background rate homogeneous within 10%. The result of the optimization is represented by the red solid lines in Fig. 15.

As shown in Fig. 15, top right, the NR background distribution has a strong z dependence at inner radii, which motivated further segmentation of the volume and defining a core volume (dashed line). This region was constructed by maximizing the WIMP signal over the square root of background and has an approximately 80% lower neutron rate than the total FV.

The fiducial mass contained in the total FV is determined by two methods: from the ratio of ^{83m}Kr calibration events in the selected volume with respect to the total sensitive

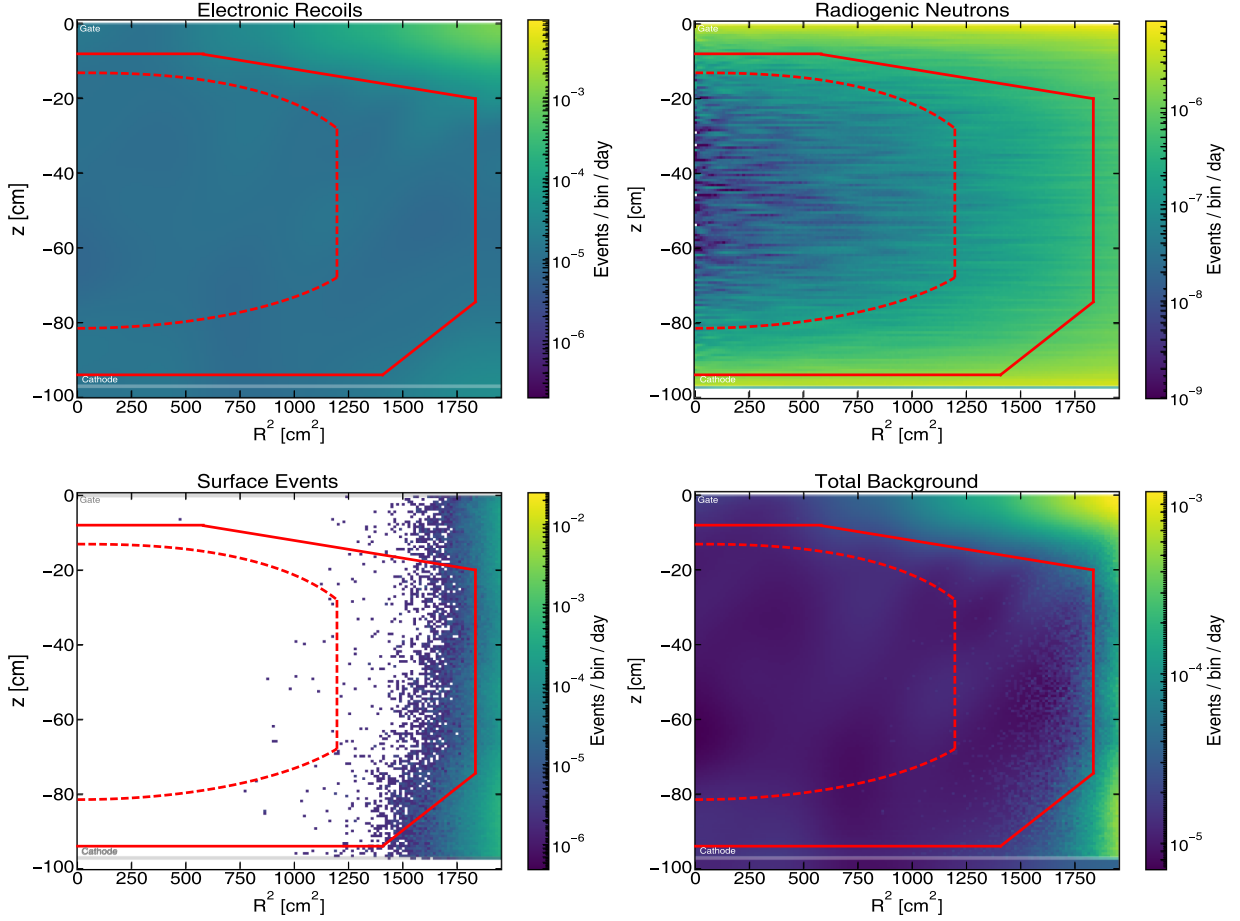


FIG. 15. Spatial distribution of modeled background rates (color scale) from ER (top left), radiogenic neutrons (top right), surface events (bottom left) and total background (bottom right) considering the $cS1$ - $cS2_b$ region in which a 50 GeV WIMP signal is expected to have the highest significance compared to the total background. Note that the rate axis is not the same for all panels. Red solid lines mark the result of the FV optimization (1.3t) and red dashed lines the inner core mass (0.65t).

volume and from geometrical calculation. The relative difference between the two methods increases from 1.6% to 1.9% with increasing radius due to increasing position reconstruction uncertainties. The two values are combined, and their difference is taken into account by a systematic uncertainty. The FV contains (1.30 ± 0.01) t of liquid xenon at -92°C , while the core mass makes up half of the fiducial mass.

VIII. SIGNAL AND BACKGROUND DISCRIMINATION

The detector response to ERs is modeled with a Markov chain Monte Carlo fit to β decays of ^{212}Pb selected from ^{220}Rn calibration data [10]. The fit uses 1867 (14138) events in SR0 (SR1) reconstructed at radii smaller than 34.6 cm in order to reject all external and surface backgrounds from the training sample that feature a loss in the charge signal. This smaller volume contains a fiducial mass of approximately 1t. The same method is applied to model the response to NRs with 3247 (4889)

events in SR0 (SR1) from $^{241}\text{AmBe}$ and neutron generator calibration data.

Figure 16 shows ^{220}Rn (top) and neutron generator (bottom) calibration data that form band-shaped distributions in the $cS2_b$ vs $cS1$ WIMP search region of interest. The extracted models are indicated by shaded colored regions together with the respective 10%-50%-90% (dotted-solid-dotted) contour lines. Colored points differentiate events found in the 1t FV (black) or outside the 1t FV but inside the 1.3t FV (green). Surface events visible below the ER band are primarily located in the outer detector region. Because the neutron generator is located outside the TPC, the NR calibration dataset features a larger fraction of events at larger radii.

The ability to discriminate between ER and NR processes is crucial for background reduction and therefore for the experimental sensitivity to WIMP dark matter. The separation of the ER and NR bands is quantified by the expected fraction of ER events reconstructed below the median of the NR distribution (50% NR acceptance, solid red line in Fig. 16). According to the extracted

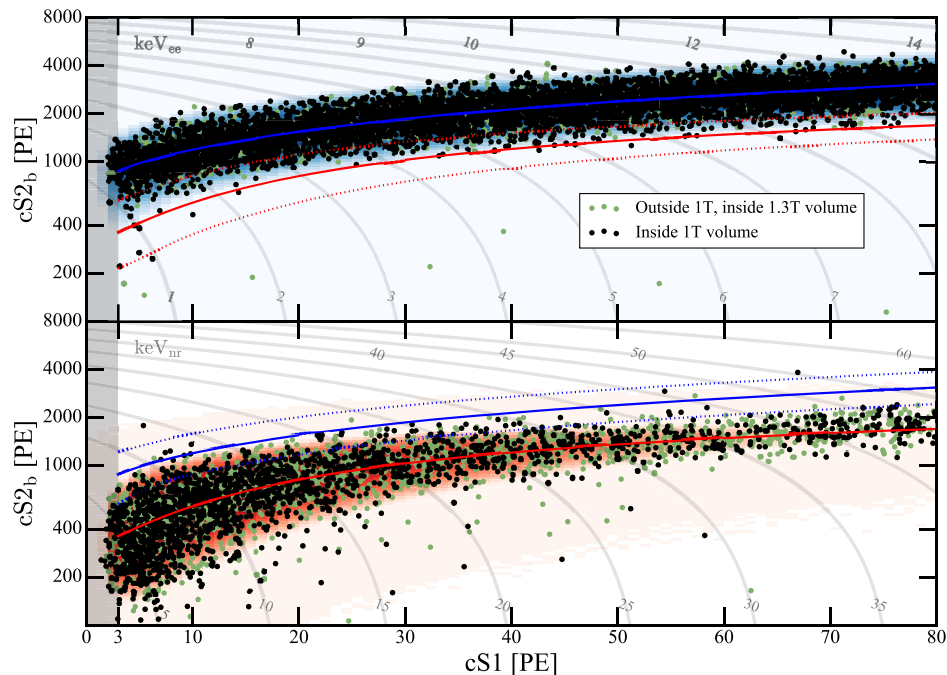


FIG. 16. ^{220}Rn (upper panel) and neutron generator (lower panel) calibration data in ER/NR discrimination space for events within the inner 1t cylindrical volume (black) and outside the 1t volume but within the 1.3t volume (green). The models extracted from the calibration data for ER (blue) and a $200 \text{ GeV}/c^2$ WIMP (red) are shaded and additionally drawn as 10%-50%-90% (dotted-solid-dotted) contour lines. The energy axes for ER (top) and NR (bottom) are drawn as gray contours. More details on background and signal modeling are given in Ref. [10].

models, the leakage fraction yields $(0.3 \pm 0.1)\%$ between 3 and 70 PE in cS1 and within the 1.3t FV. In addition to ER background originating from ^{214}Pb β decays homogeneously distributed within the detector, radon progeny accumulated on the PTFE surface and radiogenic backgrounds from materials can also contaminate the signal region, making the radial position another strong discriminator between signal and background (Sec. VII) [10].

IX. SUMMARY AND OUTLOOK

In this article, we described details on signal reconstruction, event selection, and calibrations used to search for elastic WIMP-nucleon interactions in XENON1T [6,11,12]. Most methods shown here are also being used in current and future results for WIMP and alternative dark matter models as well as other low background searches. The experiment was operated under stable conditions for more than one year. Periodic calibrations with internal sources such as $^{83\text{m}}\text{Kr}$ and ^{220}Rn were used to model time-dependent parameters such as electron live time and progressing electric field distortion due to charge accumulation on PTFE surfaces.

Compared to the predecessor experiment XENON100 [30], the MC simulation efforts were significantly strengthened in the analysis chain of XENON1T. Simulated waveforms were used to compute the peak finding efficiencies,

peak reconstruction biases, and event reconstruction performance of the data processor. MC simulations for light propagation within the TPC were employed to tune position reconstruction algorithms and to determine a goodness of fit between expected and measured patterns on the PMT arrays.

XENON1T was decommissioned in December 2018, and the imminent detector upgrade XENONnT is being constructed. XENONnT will feature an increased target mass of 5.9t and a sensitivity enhanced by more than an order of magnitude [35]. First steps towards this goal were realized by upgrades of the XENON1T purification system after SR1. The improvements resulted in the highest electron lifetime achieved in the experiment of approximately 1 ms and a reduction of the dominating background from ^{222}Rn of about 50%.

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, Netherlands Organisation for Scientific Research (NWO), Netherlands eScience Center (NLeSC) with the support of the SURF Cooperative, Weizmann Institute of Science, Israeli Centers Of Research Excellence (I-CORE),

Pazy-Vatat, Initial Training Network Invisibles (Marie Curie Actions, Grant No. PITNGA-2011-289442), Fundacao para a Ciencia e a Tecnologia, Region des Pays de la Loire, Knut and Alice Wallenberg Foundation, Kavli Foundation, and Istituto Nazionale di

Fisica Nucleare. Data processing was performed using infrastructures from the Open Science Grid and European Grid Initiative. We are grateful to Laboratori Nazionali del Gran Sasso for hosting and supporting the XENON project.

-
- [1] G. Bertone and D. Hooper, *Rev. Mod. Phys.* **90**, 045002 (2018).
- [2] L. Roszkowski, E. M. Sessolo, and S. Trojanowski, *Rep. Prog. Phys.* **81**, 066201 (2018).
- [3] G. Jungman, M. Kamionkowski, and K. Griest, *Phys. Rep.* **267**, 195 (1996).
- [4] T. Marrodñ Undagoitia and L. Rauch, *J. Phys. G* **43**, 013001 (2016).
- [5] L. Baudis, *J. Phys. G* **43**, 044001 (2016).
- [6] E. Aprile *et al.* (XENON Collaboration), *Phys. Rev. Lett.* **121**, 111302 (2018).
- [7] D. S. Akerib *et al.* (LUX Collaboration), *Phys. Rev. Lett.* **118**, 021303 (2017).
- [8] X. Cui *et al.* (PandaX Collaboration), *Phys. Rev. Lett.* **119**, 181302 (2017).
- [9] E. Aprile *et al.* (XENON Collaboration), *Eur. Phys. J. C* **77**, 881 (2017).
- [10] E. Aprile *et al.* (XENON Collaboration), *Phys. Rev. D* **99**, 112009 (2019).
- [11] E. Aprile *et al.* (XENON Collaboration), *Phys. Rev. Lett.* **122**, 071301 (2019).
- [12] E. Aprile *et al.* (XENON Collaboration), *Phys. Rev. Lett.* **122**, 141301 (2019).
- [13] E. Aprile *et al.* (XENON Collaboration), *J. Instrum.* **9**, P11006 (2014).
- [14] P. Barrow *et al.*, *J. Instrum.* **12**, P01024 (2017).
- [15] E. Aprile *et al.* (XENON Collaboration), *Eur. Phys. J. C* **75**, 546 (2015).
- [16] N. Basov, E. M. Balashov, O. V. Bogdankevitch, V. A. Danilychev, G. N. Kashnikov, N. P. Lantzov, and D. D. Khodkevitch, *J. Lumin.* **1–2**, 834 (1970).
- [17] A. Lansart *et al.*, *Nucl. Instrum. Methods Phys. Res., Sect. A* **135**, 47 (1976).
- [18] COMSOL Multiphysics, www.comsol.com.
- [19] D. Furse *et al.*, *New J. Phys.* **19**, 053012 (2017).
- [20] E. Aprile *et al.* (XENON Collaboration), *Phys. Rev. D* **95**, 072008 (2017).
- [21] B. N. Laboratory, National nuclear data center, <http://www.nndc.bnl.gov>.
- [22] D. S. Akerib *et al.* (LUX Collaboration), *Phys. Rev. D* **96**, 112009 (2017).
- [23] R. F. Lang, J. Pienaar, E. Hogenbirk, D. Masson, R. Nolte, A. Zimbal, S. Röttger, M. L. Benabderrahmane, and G. Bruno, *Nucl. Instrum. Methods Phys. Res., Sect. A* **879**, 31 (2018).
- [24] E. Aprile *et al.* (XENON Collaboration), *J. Phys. G* **41**, 035201 (2014).
- [25] R. Saldanha, L. Grandi, Y. Guardincerri, and T. Wester, *Nucl. Instrum. Methods Phys. Res., Sect. A* **863**, 35 (2017).
- [26] E. Aprile *et al.* (XENON Collaboration), *Phys. Rev. Lett.* **119**, 181301 (2017).
- [27] E. Aprile *et al.* (XENON Collaboration), *J. Instrum.* **14**, P07016 (2019).
- [28] E. Aprile *et al.* (XENON Collaboration), The processor for analyzing xenon, 2018, <https://doi.org/10.5281/zenodo.1195785>.
- [29] E. Aprile *et al.* (XENON Collaboration), Handy analysis for xenon, 2018, <https://doi.org/10.5281/zenodo.1464018>.
- [30] E. Aprile *et al.* (XENON Collaboration), *Astropart. Phys.* **54**, 11 (2014).
- [31] B. Pelssers, Position reconstruction and data quality in XENON, Master's thesis, Utrecht University, 2015.
- [32] E. Aprile *et al.* (XENON Collaboration), *Astropart. Phys.* **35**, 573 (2012).
- [33] D. S. Akerib *et al.* (LUX Collaboration), *J. Instrum.* **13**, P02001 (2018).
- [34] Fast Artificial Neural Network Library (FANN), <http://leenissen.dk/fann/wp/>.
- [35] E. Aprile *et al.* (XENON Collaboration), *J. Cosmol. Astropart. Phys.* **04** (2016) 027.
- [36] S. Agostinelli *et al.* (GEANT4 Collaboration), *Nucl. Instrum. Methods Phys. Res., Sect. A* **506**, 250 (2003).
- [37] M. J. Slakter, *J. Am. Stat. Assoc.* **60**, 854 (1965).
- [38] G. Bakale, U. Sowada, and W. F. Schmidt, *J. Phys. Chem.* **80**, 2556 (1976).
- [39] E. Aprile *et al.* (XENON Collaboration) (in preparation).
- [40] M. Szydagis *et al.*, Noble element simulation technique v2.0 (2018), <https://doi.org/10.5281/zenodo.1314669>.
- [41] Scikit-learn: Machine Learning in Python, <https://scikit-learn.org>.
- [42] P. Sorensen, *Nucl. Instrum. Methods Phys. Res., Sect. A* **635**, 41 (2011).
- [43] R. Brun and F. Rademakers, *Nucl. Instrum. Methods Phys. Res., Sect. A* **389**, 81 (1997).
- [44] W. Wilson *et al.*, Los Alamos National Laboratory, U.S.A. Technical Report (Source-4a) LA-13639-MS, 1999.
- [45] V. Tomasello, V. A. Kudryavtseva, and M. Robinsona, *Nucl. Instrum. Methods Phys. Res., Sect. A* **595**, 431 (2008).