XENON1T Dark Matter Data Analysis: Signal & Background Models, and Statistical Inference.

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] 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,[15]
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. Molinario,[16] K. Morå,[2, §] Y. Mosbacher,[12] M. Murra,[6] J. Naganoma,[16, 24] K. Ni,[15] U. Oberlack,[5]
K. Odgers,[10] B. Pelssers,[2] F. Piastra,[9] J. Pienaar,[18] V. Pizzella,[11] G. Plante,[1] R. Podviianiuk,[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. Trinchero,[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. Zopounidis[22]

(XENON Collaboration),[¶]

[1]*Physics Department, Columbia University, New York, NY 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, 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, NY 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, CA 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, IL 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, IN 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, TX 77005, USA*
[25]*Physics & Astronomy Department, University of California, Los Angeles, CA 90095, USA*
(Dated: April 30, 2019)

The XENON1T experiment searches for dark matter particles through their scattering off xenon atoms in a 2 tonne liquid xenon target. The detector is a dual-phase time projection chamber, which measures simultaneously the scintillation and ionization signals produced by interactions in target volume, to reconstruct energy and position, as well as the type of the interaction. The background rate in the central volume of XENON1T detector is the lowest achieved so far with a liquid xenon-based direct detection experiment. In this work we describe the response model of the detector, the background and signal models, and the statistical inference procedures used in the dark matter searches with a 1 tonne×year exposure of XENON1T data, that leaded to the best limit to date on WIMP-nucleon spin-independent elastic scatter cross-section for WIMP masses above 6 GeV/c^2.

I. INTRODUCTION

The existence of dark matter (DM), making up about 27% of the mass-energy of the Universe [1] is indicated by a wide range of astronomical and cosmological observations. Direct detection experiments, which search for DM particles interacting with ordinary matter in a terrestrial detector target, have not yet yielded unequivocal evidence for dark matter [2–7]. The XENON1T experiment [8], located in the INFN Laboratori Nazionali del Gran Sasso, Italy, primarily searches for Weakly Interacting Massive Particles (WIMPs), which could scatter elastically off xenon atoms. Using a 1 tonne×year exposure and a nuclear recoil (NR) energy range from 4.9 keV to 40.9 keV, XENON1T has set upper limits on the cross section of spin-independent elastic scattering with a minimum of $4.1×10^{-47}$ cm^2 for a 30 GeV/c^2 WIMP [9]. These are the most stringent constraints set on this interaction for WIMP masses above 6 GeV/c^2. The XENON1T experiment has achieved the lowest background rate among liquid xenon (LXe) detectors to date.

The dual-phase time projection chamber (TPC) used by the XENON1T detector allows for the reconstruction of the deposited energy and the three-dimensional position of interactions in the active liquid xenon target. The observable signals are the scintillation (S1) and ionization (S2) signals produced by energy depositions. Both S1 and S2 signals are detected by arrays of photomultiplier tubes (PMTs) arranged at the top and bottom of the detector. The longitudinal (z) position is reconstructed using the time difference between the prompt S1 signal and the S2 signal. The latter is produced by electroluminescence in gaseous xenon after electrons, drifted upwards by an electric field, get extracted from the liquid into the gas. The position in the (x,y) plane is reconstructed using the S2 signal pattern in the upper PMT array. Background from radioactivity in detector materials can be rejected to a large extent by selecting a three-dimensional fiducial region within the active volume. In addition, the S2-S1 ratio can be used to discriminate between NRs from WIMPs and neutrons and electronic recoils (ERs) from γ and β, which constitute the major backgrounds of the XENON1T experiment. More details on TPC working principles and the XENON1T TPC can be found in [8].

The XENON1T data analysis can be divided into two parts. The first part includes event reconstruction, signal corrections, and event selection, and is reported in [10]. The second part includes the detector response model, the background and WIMP signal models, and the statistical inference, and is presented in detail in this manuscript. These models and techniques are used in the XENON1T DM searches [9, 11]. The detector response model, which will be presented in Section II, describes how an ER or a NR energy deposition is reconstructed in the TPC. The fit of the detector response model to calibration data provides the ER and NR background models and the signal model described in Section III, which also considers background models constructed using data-driven methods. Lastly, the statistical inference is presented in Section IV. A summary is then given in Section V.

II. DETECTOR RESPONSE MODEL

Understanding the conversion from the deposited energy to the observed S1 and S2 signals is critical for interpreting the results of DM searches in XENON1T. The energy region of interest is in the range of a few keV to several tens of keV in searches for elastic scatters between WIMP and xenon nuclei. The conversion of deposited energy to S1 and S2 in this region is non-linear, with fluctuations due to the scintillation and ionization processes in LXe and due to the detector reconstruction.

The model of the XENON1T detector response to ERs and NRs is based on simulations which include a comprehensive description of the signal production process and detailed characterizations of detector detection and reconstruction effects. The model is constrained by a Bayesian simultaneous fit to ER and NR calibration data, which allows to use all available information and treat correlated detector uncertainties coherently.

A. Basic Signal Response in Liquid Xenon

The intrinsic signal response model in LXe follows the approach used in the Noble Element Simulation Technique (NEST) model [12, 13]. There are three forms of energy deposition in LXe: thermalization of the recoiling particle, excitation of xenon atoms, and ionization of xenon atoms. The thermalization energy loss is undetectable in the XENON1T detector. The number of detectable quanta N_q is the sum of the number of excitons N_{ex} and ion-electron pairs N_i, and can be used to reconstruct the deposited energy ε. It follows a Binomial fluctuation due to the potential energy loss to thermalization,

$$N_q \sim \text{Binom}\left(\varepsilon/W, L\right), \qquad (1)$$

where L is the Lindhard factor expressing the fraction of energy loss to heat and W (13.7±0.2 eV from a global fit [12]) is the average energy required to create either an exciton or ion-electron pair in LXe. Negligible energy

———————

*ql2265@columbia.edu

†Also at Coimbra Polytechnic - ISEC, Coimbra, Portugal

‡manfredi@physik.uzh.ch

§knut.mora@fysik.su.se

¶xenon@lngs.infn.it

is lost to thermalization in an ER as the mass of the recoiling electron is much smaller than the xenon nucleus. In NRs, the recoiling xenon atom transfers kinetic energy through elastic scattering off surrounding xenon atoms, resulting in a Lindhard factor [14] of 0.1-0.2 in LXe. The field- and energy-dependence of the Lindhard factor are parametrized following the NEST model [13].

The exciton-to-ion ratio $\langle N_{\mathrm{ex}}/N_i \rangle$ is related to the excitation and ionization cross sections of recoiling particles on xenon atoms. For ERs, it is assumed to be constant, and is given a uniform prior ranging from 0.06 to 0.20 [12]. For NRs, it is parametrized as a function of the deposited energy and the electric field strength F in the active volume, following [13], and is in the range of 0.7 to 1.0 for NR energies from 5 to $40\,\mathrm{keV_{nr}}$ under a field of $81\,\mathrm{kV/cm}$. The binomial fluctuations of N_{ex} and N_i can be written as

$$
\begin{aligned}
N_i &\sim \mathrm{Binom}\left(N_q, \frac{1}{1+\langle N_{\mathrm{ex}}/N_i\rangle}\right), \\
N_{\mathrm{ex}} &= N_q - N_i.
\end{aligned} \tag{2}
$$

Ionization electrons have a probability $1-r$ to escape the cloud of ion-electron pairs, where r is referred to as the recombination fraction,

$$
\begin{aligned}
N_e &\sim \mathrm{Binom}\left(N_i, 1-r\right), \\
N_\gamma &= N_i - N_e + N_{\mathrm{ex}},
\end{aligned} \tag{3}
$$

where N_γ and N_e are the number of photons generated by de-excitation of the initial excitons and ion-electron recombination, and of the escaping electrons, respectively. Due to detector effects, such as field non-uniformity, and intrinsic fluctuations [15], the recombination fraction r fluctuates, and is modeled as a Gaussian distribution,

$$
r \sim \mathrm{Gauss}\left(\langle r \rangle, \Delta r\right). \tag{4}
$$

The mean recombination fraction $\langle r \rangle$ depends on the deposited energy and the electric field, and is described by the Thomas-Imel (TI) box model [16],

$$
\langle r \rangle = 1 - \frac{\ln\left(1 + N_i\varsigma/4\right)}{N_i\varsigma/4}, \tag{5}
$$

where ς is the field-dependent TI model parameter. Δr is the recombination fluctuation, parameterized as

$$
\Delta r = q_2(1 - e^{-\varepsilon/q_3}), \tag{6}
$$

where q_2 and q_3 are free parameters. The parametrization is empirically chosen to take into account both the fact that a constant recombination fluctuation is observed with deposit energy larger than $2\,\mathrm{keV}$ and the assumption that the recombination fluctuation diminishes as deposit energy goes to zero.

For NR, the TI box model, together with the Lindhard factor, has been shown to match data well [13]. In [17], it has been shown that NR response in LXe can be modeled with only statistical fluctuations taken into account, indicating a negligible recombination fluctuation for NR. Thus, the recombination fluctuation for NR is set to 0 in the detector response model. For the parameterization of the NR recombination fraction $\langle r \rangle_{\mathrm{nr}}$, the TI parameter ς_{nr} is expressed, following [13], as a power law function for the field dependence. For low energy ERs (roughly above $3\,\mathrm{keV}$ and below $10\,\mathrm{keV}$), recent measurements [15, 18–20] indicate that the TI box model cannot fully describe the recombination process. Therefore, we use a modified TI box model for the ER recombination fraction $\langle r \rangle_{\mathrm{er}}$,

$$
\langle r \rangle_{er} = \left(1 - \frac{\ln(1 + N_i\varsigma_{\mathrm{er}}/4)}{(N_i\varsigma_{\mathrm{er}}/4)}\right) / \left(1 + e^{-(\varepsilon-q_0)/q_1}\right), \tag{7}
$$

$$
\varsigma_{\mathrm{er}} = \gamma_{\mathrm{er}} e^{-\varepsilon/\omega_{\mathrm{er}}} F^{-\delta_{\mathrm{er}}}, \tag{8}
$$

where the Fermi-Dirac term in Eq. 7 and the exponential term in Eq. 8 were empirically added to the TI box model to account for the deviation of measurements in the $<3\,\mathrm{keV}$ and $>10\,\mathrm{keV}$ energy ranges, respectively. Similarly to NR, the field dependence of the ER TI box parameter ς_{er} follows a power law as introduced in NEST [13]. The free parameters q_0, q_1, γ_{er}, ω_{er}, and δ_{er} are obtained by matching the detector response model to XENON1T data, without any additional constraints. It is worth noting that NEST model has been updated based on a global fit using recent measurements [21], and is compatible with this work in the energy region of interest. Fig. 1 shows the mean photon and charge yields as a function of energy for NR and ER, respectively, together with the measurements from [17, 22–28] for NR and [15, 19, 20, 29] for ER. The mean photon $\langle N_\gamma \rangle/\varepsilon$ and charge yields $\langle N_e \rangle/\varepsilon$ are defined as

$$
\begin{aligned}
\langle N_\gamma \rangle/\varepsilon &= \frac{1}{W} \frac{\langle r \rangle + \langle N_{\mathrm{ex}}/N_i\rangle}{1+\langle N_{\mathrm{ex}}/N_i\rangle}, \\
\langle N_e \rangle/\varepsilon &= \frac{1}{W} \frac{1-\langle r \rangle}{1+\langle N_{\mathrm{ex}}/N_i\rangle}.
\end{aligned} \tag{9}
$$

The calibration of low energy ERs in XENON1T is performed using an internal ^{220}Rn source [31]. The energy spectrum of β-decays from ^{212}Pb, one of the progenies of ^{220}Rn, is similar to the dominant ER background, from β-decays of ^{214}Pb originating from ^{222}Rn emanation, in the low energy region ($<10\,\mathrm{keV}$). However, the detector response model built for ERs in XENON1T is, in principle, not applicable to γ-induced ERs that at sufficiently high energies may interact
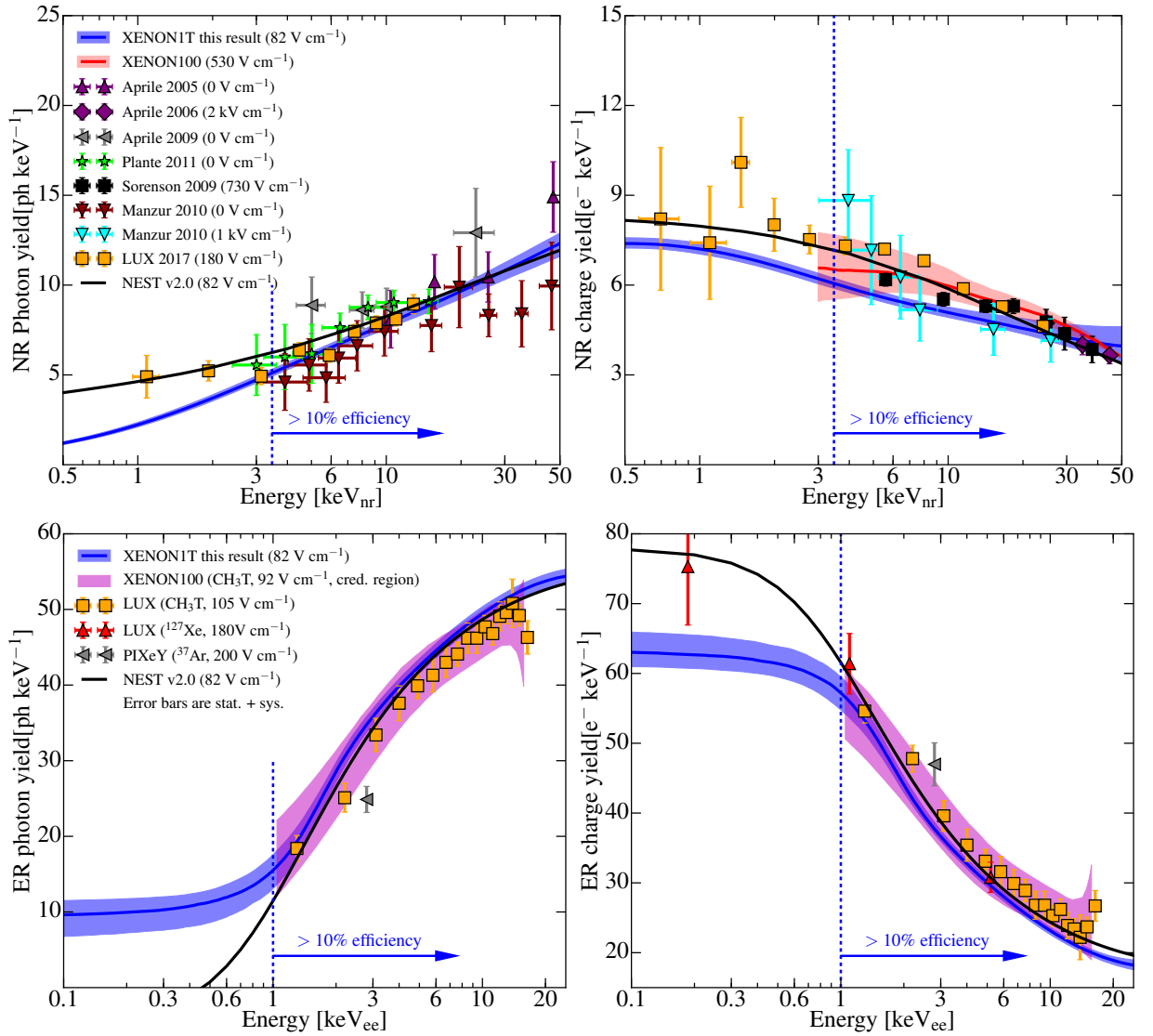
FIG. 1: Mean photon and charge yields of NR (upper panels) and ER (lower panels) in the XENON1T calibration data fit (details of the fit are in Sec. II C). The blue solid line and shaded region represent the point estimation and 15%-85% credible region, respectively, of the posterior. Data points for upper panels are from fixed-angle neutron scattering measurements [22–28]. Results of XENON100 [17] using data-Monte Carlo (MC) matching on the ^{241}AmBe calibration method are shown with the red solid line and shaded region. The best fit from NEST v2.0 [21] is shown with the black solid line. The measurements from [19, 20, 29, 30] are shown along with the best fit of NEST v2.0 [21] in lower panels. The vertical dashed blue lines indicates the energy threshold for XENON1T NR and ER calibrations, below which the detection efficiency drops to less than 10%.

with the inner-shell electrons. When this happens, the vacancy in the inner shell results in either X-ray or Auger electrons emission, both of which further ionize xenon atoms. Consequently, γ-induced ERs can have multiple recoiling electrons instead of one as in β-induced ERs. The binding energy for L-shell electron in xenon is about 4.8-5.5 keV. According to the NIST database [32], the corresponding X-ray has mean free path of about 5 μm. The effect of the separation of electron clouds at this spatial scale on the recombination is not yet understood. In addition, the detection efficiency drops to below 10% when deposited energy is below 1 keV$_{ee}$ for ER or below

3.5 keV$_{nr}$ for NR. In this region, the derived photon and charge yields are extrapolations of the detector response model which is constrained mainly using calibration data with deposited energy above 1 keV$_{ee}$ for ER and above 3.5 keV$_{nr}$ for NR.

B. Detector Reconstruction Effects

Besides the intrinsic response of LXe, detector reconstruction effects on the S1 and S2 signals are modeled. More specifically, the spatial dependence of S1

and S2 signals, the single and double photoelectron (PE) emission of the PMT photocathode [33, 34], the position reconstruction uncertainty, the reconstruction efficiency, bias, and signal fluctuations, and the acceptance of data selections in analysis are taken into account in the model.

Photons from an energy deposition and the subsequent recombination (Eq. 3) are detected by the PMTs as an S1 with an efficiency, which is the product of the light collection efficiency ϵ_L, PMTs' average quantum efficiency ϵ_{QE}, and PMTs' average collection efficiency ϵ_{CE}. Electrons are drifted to the gas-liquid interface under the drift field, and are extracted under the stronger field, amplifying the electron signal (S2) by the gas gain G, which is the number of photoelectrons per electron that is extracted into gaseous xenon. Both ϵ_L and G are spatially dependent, and related to the energy scale parameters g_1' (average number of PEs observed per emitted photon) and g_2' (amplification factor for charge signal), respectively, by

$$
\begin{aligned}
g_1'(x,y,z) &= (1 + p_{\mathrm{dpe}}) \cdot \epsilon_L(x,y,z) \cdot \epsilon_{\mathrm{QE}} \cdot \epsilon_{\mathrm{CE}}, \\
g_2'(x,y) &= \epsilon_{\mathrm{ext}} G(x,y),
\end{aligned}
\tag{10}
$$

where p_{dpe} is the probability for the PMT photocathode to emit two photoelectrons when absorbing one photon [33, 34], and ϵ_{ext} is the extraction efficiency of the drifted electrons which is assumed to be constant in this study. Note that g_1 and g_2 in [10] correspond to the averages of $g_1'(x,y,z)$ and $g_2'(x,y)$, respectively, in Eq. 10 over the active volume. The number of hits detected by PMTs, N_{hit}, and photoelectrons generated from the PMT photocathode, N_{pe}, can be described by a binomial distribution

$$
\begin{aligned}
N_{hit} &\sim \mathrm{Binom}\left(N_\gamma, \epsilon_L(x,y,z) \cdot \epsilon_{\mathrm{QE}} \cdot \epsilon_{\mathrm{CE}}\right), \\
N_{\mathrm{pe}} - N_{\mathrm{hit}} &\sim \mathrm{Binom}\left(N_{\mathrm{hit}}, p_{\mathrm{dpe}}\right).
\end{aligned}
\tag{11}
$$

In addition to the (x, y) dependence caused by the varying charge amplification, S2 signals are a function of the z position, because the electrons attach to electronegative impurities when drifting towards the gaseous phase. The number of electrons N_{ext} that survive the drifting and the extraction into the gas can be modeled as

$$
N_{\mathrm{ext}} \sim \mathrm{Binom}\left(N_e, e^{-z/(\tau_e \cdot \nu_d)} \epsilon_{\mathrm{ext}}\right),
\tag{12}
$$

where τ_e and ν_d are the electron lifetime and electron drift velocity, respectively. The total proportional scintillation light detected, N_{prop}, can be approximated as

$$
N_{\mathrm{prop}} \sim \mathrm{Gauss}\left(N_{\mathrm{ext}} G, \sqrt{N_{\mathrm{ext}}} \Delta G\right),
\tag{13}
$$

where ΔG is the spread of the gas gain. For simplicity, we consider $\Delta G/G$ a constant in the model.

The S1 and S2 signals are constructed from N_{pe} and N_{prop}, respectively, amplified by the PMTs, digitized, and selected by clustering and classification software [35]. To account for biases and fluctuations in this process, the S1 and S2 are written as

$$
\begin{aligned}
\mathrm{S1}/N_{\mathrm{pe}} - 1 &\sim \mathrm{Gauss}\left(\delta_{\mathrm{s1}}, \Delta\delta_{\mathrm{s1}}\right), \\
\mathrm{S2}/N_{\mathrm{prop}} - 1 &\sim \mathrm{Gauss}\left(\delta_{\mathrm{s2}}, \Delta\delta_{\mathrm{s2}}\right),
\end{aligned}
\tag{14}
$$

where δ_{s1} (δ_{s2}) and $\Delta\delta_{\mathrm{s1}}$ ($\Delta\delta_{\mathrm{s2}}$) are the bias and spread, respectively, of the S1 (S2) reconstruction. Reconstruction biases and fluctuations are estimated using a waveform simulation including models of realistic scintillation pulse shape, charge amplification, electronic noise level, PMT single PE spectrum, PMT after-pulses, as well as secondary S2s induced by photoionizations on grids and impurities in the LXe volume [10, 35].

The S1 and S2 signals are corrected for their spatial dependence based on the reconstructed positions $\vec{x}_r$. The z position of an event is reconstructed using the time difference between the S1 and the S2 signals, and has better resolution than the (x, y) position which is reconstructed through the S2 hit pattern on the PMTs in the top array. We assume the reconstruction fluctuations along x and y axes to be identical. The reconstructed position of $\vec{x}_r$ can be written as

$$
\vec{x}_r \sim \mathrm{Gauss}\left(\vec{x}, \sigma_p\right),
\tag{15}
$$

where σ_p is the position reconstruction resolution, and depends on both S2 area and the (x,y,z) position of event. The corrected S1 (cS1) and S2 (cS2) are written as

$$
\begin{aligned}
\mathrm{cS1} &= \mathrm{S1}\frac{g_1}{g_1'(x_r,y_r,z_r)}, \\
\mathrm{cS2} &= \mathrm{S2}\frac{g_2}{g_2'(x_r,y_r)} e^{z/(\tau_e' \cdot \nu_d)},
\end{aligned}
\tag{16}
$$

where τ_e' is the mean electron lifetime measured. The uncertainty of the measured electron lifetime is used to constrain τ_e in the signal response model. The correction is based on the measured electron lifetime τ_e'. In the following analysis, as well as in XENON1T results [9, 11], we use the corrected S2 collected by the bottom PMTs cS2$_b$, which has a more homogeneous spatial dependence. The effect of field distortion is negligible and not implemented in the signal response model because the position correction to account for it is applied to data [9].

Selection criteria were applied to data to ensure good quality of the sample and to optimize the signal-to-background ratio for the dark matter search. More details of the data selection can be found in [10]. The detection efficiency loss arises from the software

reconstruction efficiency of S1s, and the S1-related and S2-related event selections. The efficiencies for these are considered as functions of $N_{\rm hit}$, S1 and S2, respectively, in the signal response model. In addition, a realistic selection of single scatters is implemented in the simulation. The rejection of multiple scatters is critical to the search for WIMP signals in XENON1T detector and is based on the areas of the largest and second largest S2s. Not all multiple scatters are rejected by this selection. This is mainly because of PMT after-pulses, photoionization of impurities in the detector, and the spatial resolution of the detector. Each energy deposition that is resolvable by the position reconstruction is taken into account in the simulation of the response model. The same single-scatter selection is applied to the simulated and actual data, in order to accurately address the acceptance of single scatters and the rejection power against multiple scatterings in the response model.

C. Fit to Calibration Data

The detector response model is constrained using the calibration data from ^{220}Rn for ER and ^{241}AmBe and a D-D generator for NR. Events with cS1 ranging from 0 to 100 PE are used to constrain the signal response model, covering the cS1 region of interest (3-70 PE) for WIMP searches [9, 11]. The detector response model obtained from the fit to calibration data is used to construct WIMP signal and background models, which are then input to the statistical inference of dark matter search data [9, 11]. The fit is performed simultaneously using all available XENON1T calibration data taken during the first (SR0 with drift field of 120 V/cm) and second (SR1 with drift field of 81 V/cm) science data taking periods. The fit uses the binned likelihood for distributions in $\log_{10}({\rm cS2}_b/{\rm cS1})$ versus cS1 using the data in the cylindrical fiducial volume defined in SR0 [11]. The likelihood is sampled using affine invariant Markov Chain Monte Carlo (MCMC) [36]. Important nuisance parameters in the detector response model are listed in Table II. There are three parameters for scaling the S1 cut acceptance, S2 cut acceptance, and reconstruction efficiency. These parameters are constrained by the uncertainties estimated for the three efficiencies, which depend on the signal size. The $\delta_{\rm s1}$, $\delta_{\rm s2}$, $\Delta\delta_{\rm s1}$, and $\Delta\delta_{\rm s2}$ in Eq. 14 are signal-size dependent, and thus are not listed in Table II. Their priors can be found in [10]. We choose to use the more conservative uncertainty between the lower and upper uncertainties given in [13] for the NEST parameters that describe the response of LXe to NR, except for η, which parameterizes the Penning quenching of high-energy NRs. Correlations between the NEST parameters are not provided in [13] and are, thus, not considered in the priors of this work, in order to be conservative and avoid potential over-constraint on fit parameters.

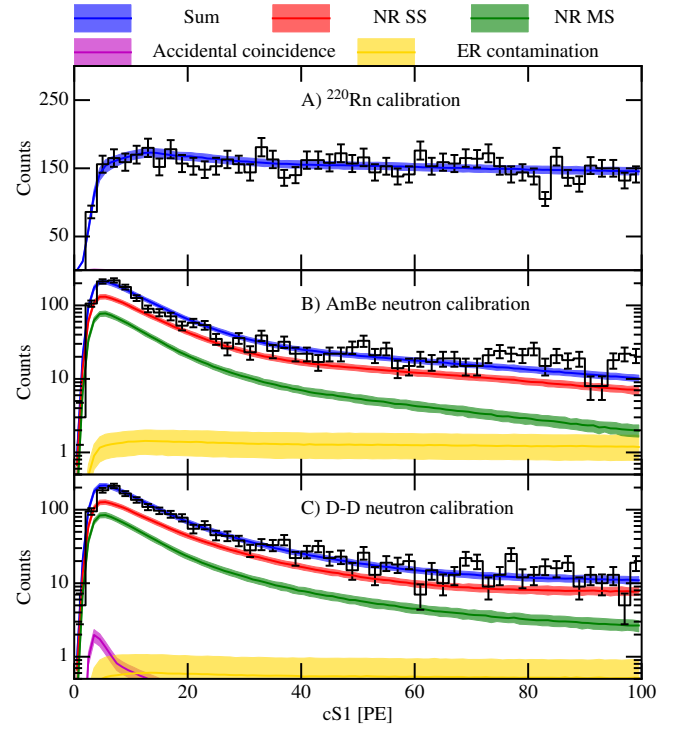Figures 2 and 3 show the comparisons of the

FIG. 2: The cS1 spectra of the SR1 data (black bars) and the signal response models (blue) for ^{220}Rn, ^{241}AmBe, and D-D generator are shown in panel A, B, and C, respectively. Solid lines represent the median of the posterior, and the shaded regions show the 15.4% to 84.6% credible regions of the posterior. The accidental coincidence, ER contamination, single NR scatter (NR SS), and multiple NR scatter (NR MS) components are shown in magenta, gold, red and green, respectively.

cS1 spectra and cS2$_b$ spectra, respectively, between the posterior of tested signal response model and data. For neutron calibrations the ER contamination are considered in the fit. For both ER and NR calibrations, a small fraction of events arise from accidental coincidence (AC) which will be illustrated in section III. The rates of each component are free in the fit, and are not listed in Table II. The matching between the signal response model and the calibration data is good, with the goodness of fit (GoF) p-values (calculated using method in [37]) for the cS2$_b$ distributions comparison in different cS1 ranges shown in Table I. Good agreement between model and data is critical for the WIMP signal and background models, especially for the (cS1, cS2$_b$) region where the WIMP signal is expected. We also use the GoF p-values to check the consistency between the model and data for ^{220}Rn in a reference (cS1, cS2$_b$) region. The upper cS2$_b$ boundary of the reference region for matching ^{220}Rn data is defined by the 97.7% percentiles ($+2\sigma$) of NR events in ^{241}AmBe calibration data. The GoF p-values for ^{241}AmBe and D-D generator data are calculated excluding the lowest S2 region, corresponding to the
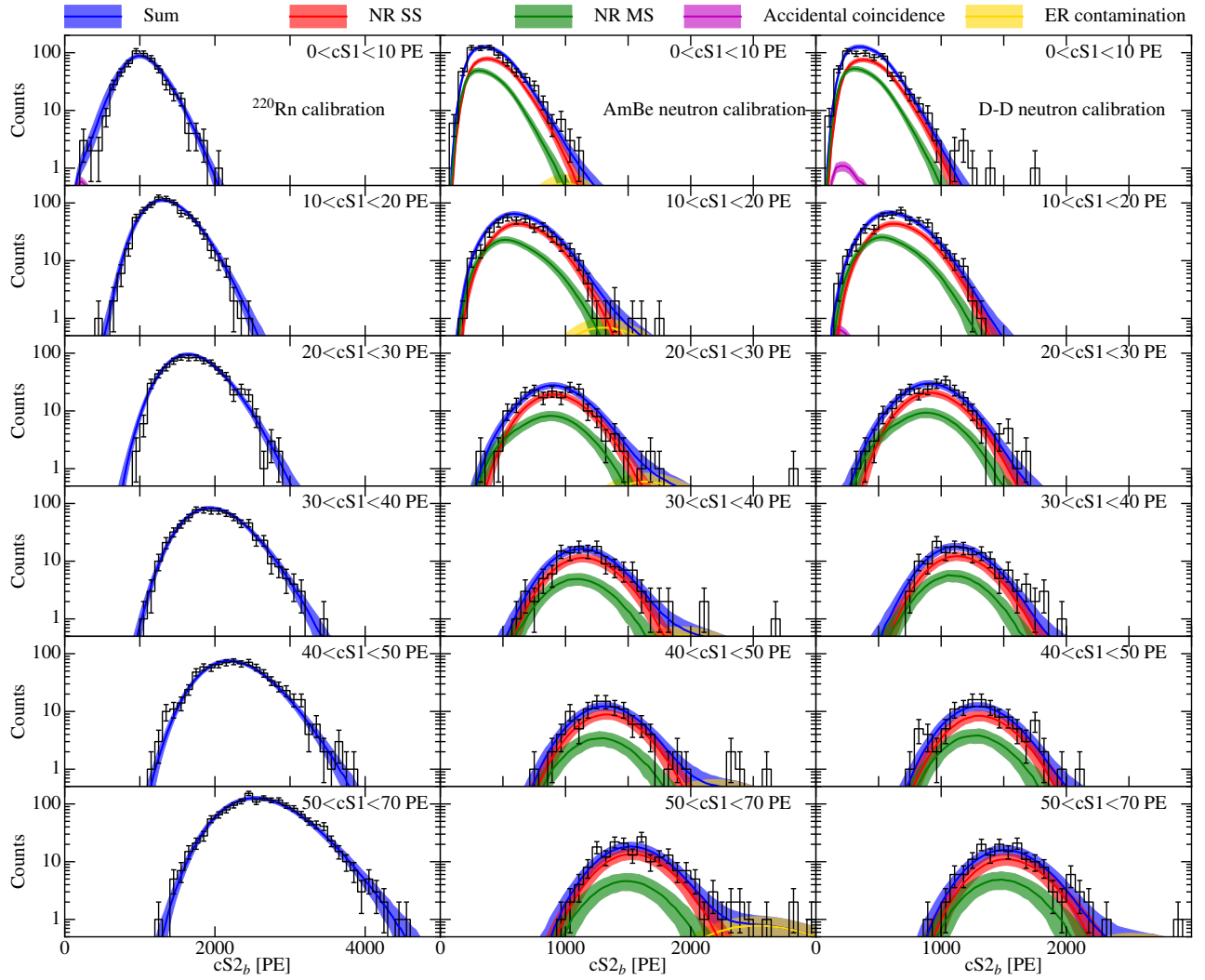
FIG. 3: The $cS2_b$ distributions in different cS1 ranges of the SR1 data (black bars) and the signal response model (blue) for ^{220}Rn, ^{241}AmBe and D-D generator, from left to right, respectively. The figure description is the same as in Fig. 2.

0.13% percentile (-3σ) in S2. The GoF p-values at different cS1 for matching calibration data are all above the 5% threshold set for an acceptable fit.

TABLE I: Goodness of fit p-values of data-model matching on $cS2_b$ spectra in different cS1 ranges for ^{220}Rn, ^{241}AmBe and D-D neutron generator calibration data in SR1.

data	cS1 range (PE)					
	0-10	10-20	20-30	30-40	40-50	50-70
^{220}Rn overall	0.18	0.44	0.51	0.96	0.29	0.49
^{220}Rn reference	0.17	0.28	0.30	0.85	0.21	0.48
^{241}AmBe	0.12	0.50	0.88	0.89	0.62	0.70
D-D	0.10	0.50	0.89	0.73	0.11	0.27

III. BACKGROUND AND DARK MATTER SIGNAL MODELS

Background and DM signal models are crucial in the statistical interpretation of the DM search results in XENON1T. Four background components are modelled in XENON1T: ER, NR, surface, and accidental coincidence (AC). The ER and NR background models, as well as the WIMP signal model, are constructed based on the detector response model illustrated in section II (we call the detector response model "posterior" after fitting to the calibration data). The surface and AC background models are constructed using data-driven methods. The background and WIMP signal models are 3-D distributions in cS1, $cS2_b$, and the spatial coordinate of the detector. The background and DM signal models also include the uncertainties in cS1, $cS2_b$ and spatial distribution, as well as in the absolute rate

TABLE II: Main nuisance parameters in the detector response model of XENON1T. Lower and upper errors of the posterior are calculated as the difference between the median and the 15.4% and 84.6% percentiles of the posterior.

Par.	Prior SR0	Prior SR1	Posterior SR0	Posterior SR1	Reference and note
W	13.7±0.2		13.8±0.2		In unit of eV; Eq. 1
$\langle N_{\mathrm{ex}}/N_i \rangle$	0.06 - 0.20		$0.15^{+0.04}_{-0.06}$		Eq. 2
g_1	0.142±0.002		0.142±0.005		In unit of PE/photon; Eq. 10
g_2	11.4±0.2		11.4±0.2		In unit of PE/e$^-$; Eq. 10, for cS2$_\mathrm{b}$
p_{dpe}	0.18 - 0.24		$0.219^{+0.015}_{-0.023}$		Eq. 10, [33, 34]
$1 - \tau_\mathrm{e}/\tau_\mathrm{e}'$	0±0.04	0±0.02	0.01±0.03	0.01±0.01	Eq. 12 and 16
γ_{er}	free		0.124±0.003		Eq. 8
ω_{er}	free		31±4		In unit of keV; Eq. 8
δ_{er}	free		0.24±0.06		Eq. 8
q_0	free		$1.13^{+0.24}_{-0.32}$		In unit of keV; Eq. 7
q_1	free		$0.47^{+0.18}_{-0.15}$		In unit of keV; Eq. 7
q_2	free	free	0.041±0.006	0.034±0.003	Eq. 6
q_3	free		$1.7^{+1.3}_{-1.1}$		In unit of keV; Eq. 6
α	1.240±0.079		1.280±0.063		NEST parameters [13]
ζ	0.047±0.009		$0.045^{+0.009}_{-0.008}$		NEST parameters [13]
β	239±28		273^{+24}_{-20}		NEST parameters [13]
γ	0.0139±0.007		0.0141±0.006		NEST parameters [13]
δ	0.062±0.006		0.061±0.006		NEST parameters [13]
κ	0.139±0.003		0.138±0.003		NEST parameters [13]
η	3.3±0.7		$3.3^{+0.7}_{-0.6}$		NEST parameters [13]
λ	1.14±0.45		$1.15^{+0.35}_{-0.27}$		NEST parameters [13]
ϵ_{ext}	96%	96%	-	-	Fixed; Eq. 12
$\Delta G/G$	0.24	0.25	-	-	Fixed; Eq. 13
ν_d	0.144	0.134	-	-	In unit of cm/µs; fixed; Eq. 12
F	120	81	-	-	In unit of V/cm;

of the background and DM signal. In this section, the details of the background and DM signal models used in the statistical inference of XENON1T results [9, 11] are given.

The final dark matter search is performed between $3<\mathrm{cS1}<70$ PE, and $50.1<\mathrm{cS2}_b<7940$ PE. Below 3 PE, the S1 acceptance is very small due to the 3-fold PMT coincidence requirement for S1s. The upper cS1 cut is chosen to contain most of spin-independent WIMP recoil spectra, shown in Fig. 8. In previous XENON analyses, fiducial volumes in radius R and z were constructed to provide a low background for the analysis. With the inclusion of a model for the surface background, presented in Section III C, the analysis could be extended to also consider the radius as an analysis variable. The analysis volume is defined by a maximal radius, 42.8 cm, and a R-dependent z-cut, shown in Fig. 5 with a magenta line. The construction of this cut, which was made to include regions of the detector where the total background rate was approximately constant with z, is presented in [10].

The magnitude of the radiogenic background, discussed in Section III B, is attenuated moving towards the center of the detector. In order to optimize the discovery power of the analysis, a partition of the detector, with a clean, "core" volume was proposed. Optimizing for discovery significance yielded a central 0.65 tonne volume, shown with a dashed green line in Fig. 5. The expected radiogenic neutron rate in this volume is 36% of the average rate in the analysis volume.

A. Electronic Recoil Background Model

Although XENON1T achieved an excellent discrimination power between the ER background and NR signal, with an average ER leakage fraction below the NR median of about 0.3% [9], the ER component is the dominant background for the DM search due to its high rate in comparison with the other background sources.

In the energy region of interest for WIMP search

($<100\,\mathrm{keV_{NR}}$), the dominant component contributing to the ER background are β-decays of $^{214}\mathrm{Pb}$. The $^{214}\mathrm{Pb}$ is a progeny of $^{222}\mathrm{Rn}$, which is emanated from $^{238}\mathrm{U}$ daughters in the detector materials, and can convect and diffuse into the inner volume of the detector. Decays of $^{218}\mathrm{Po}$ and $^{214}\mathrm{Bi}$-$^{214}\mathrm{Po}$, which are also $^{222}\mathrm{Rn}$ progenies, can be identified to estimate the rate of $^{214}\mathrm{Pb}$. This selection is based on the unique energy and time profiles of $^{218}\mathrm{Po}$ and $^{214}\mathrm{Bi}$-$^{214}\mathrm{Po}$ decays, and gives activities of 71 ± 5(stat.)±7(sys.) and 29 ± 3(stat.)±3(sys.) events/ton/year/keV (tyu), respectively. The difference between the activities of $^{222}\mathrm{Rn}$ progenies is likely due to their plate-out onto the electrode and polytetrafluoroethylene (PTFE) reflector surfaces. As the $^{214}\mathrm{Pb}$ decay occurs between the $^{218}\mathrm{Po}$ and $^{214}\mathrm{Bi}$ decays, the rate of $^{214}\mathrm{Pb}$ is in range of 29 to 71 tyu which is consistent with the estimate of 56 ± 6(stat.)±6(sys.) tyu from [38], where $10\,\mu\mathrm{Bq/kg}$ of $^{222}\mathrm{Rn}$ were assumed.

The second largest component of the ER background is β-decays of $^{85}\mathrm{Kr}$. The concentration of natural krypton $^{\mathrm{nat}}\mathrm{Kr/Xe}$ was reduced to 0.36 ± 0.06 ppt by the end of SR0 through cryogenic distillation [39]. With regular rare-gas mass spectrometry measurements [40] during SR1, we measured an average natural krypton concentration $^{\mathrm{nat}}\mathrm{Kr/Xe}$ of 0.66 ± 0.11 ppt, resulting in an average decay rate for $^{85}\mathrm{Kr}$ of 7.7 ± 1.3 tyu, using the conversion derived from data with high concentration of krypton at the beginning of XENON1T operation. These high-krypton concentration data also gave a $^{85}\mathrm{Kr}/^{\mathrm{nat}}\mathrm{Kr}$ ratio of $(1.7\pm0.3)\times10^{-11}$ mol/mol. Taking the ER contributions from material radioactivity, solar neutrino scatterings, and β-decays of $^{136}\mathrm{Xe}$ of 8 ± 1, 2.5 ± 0.1, and 0.8 ± 0.1 tyu, respectively, into account, the total ER background rate in the region of interest (ROI) for DM searches is estimated to be between 48 ± 5 and 90 ± 8 tyu. This is consistent with the prediction of 75 ± 6 (sys.) tyu from [38] and with the best-fit of 82^{+5}_{-3}(sys.)±3(stat) tyu of low energy ER background from [9].

The energy distribution of the ER background in the ROI is assumed to be uniform due to the dominance of the flat spectrum from $^{214}\mathrm{Pb}$ beta decay. Uncertainties in the (cS1, cS2$_b$) distribution for the ER background are dominated by the uncertainties in $\langle r \rangle$, and its fluctuations, Δr. The uncertainty in $\langle r \rangle$ is mainly from parameter γ_{er} in Eq. 8. The effects of varying γ_{er} and Δr are shifting the mean and changing the spread, respectively, of the ER distribution in cS2$_b$. Fig. 4 shows the variation of the ER distributions on $\log_{10}(\mathrm{cS2}_b)$ in different cS1 ranges. The distributions are produced based on the detector response model with the rest of the nuisance parameters (shown in Table II) marginalized to the point estimation (median posterior). Due to the computational complexity, the variation of background model in terms of (cS1, cS2$_b$) distribution in the statistical interpretation is practically interpolated using the distributions that are computed at 2.3%, 6.7%, 15.9%, 30.9%, 50.0%, 69.1%, 84.1%, 93.3%, 97.7%

percentiles of the posterior, which correspond to -2, -1.5, -1, -0.5, 0, 0.5, 1.0, 1.5, 2.0 σ, respectively, in standard deviations. Given that the ER background induced by radioactivities in detector materials is subdominant, the ER background is assumed to be spatially uniform inside the analysis volume.
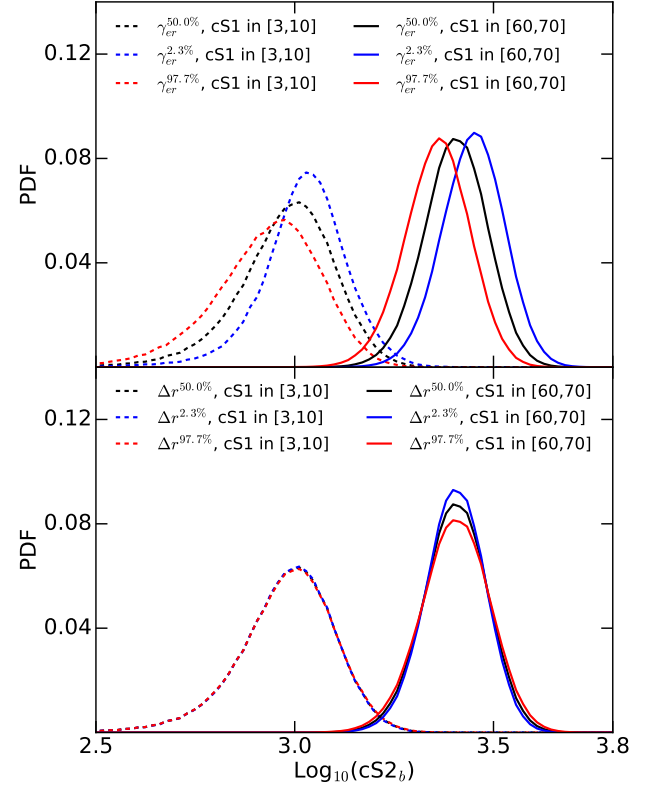

FIG. 4: Variations in $\log_{10}(\mathrm{cS2}_b)$ distributions as γ_{er} (top panel) and Δr (bottom panel) vary, from the 2.3% to 50.0%, and then to 97.7% percentile of the signal model posterior, in different cS1 ranges.

B. Nuclear Recoil Background Model

The NR background, which has a similar (cS1, cS2$_b$) distribution to the WIMP signal, contributes 1.48 events to the $1\,\mathrm{tonne}\times\mathrm{year}$ exposure data [9]. Radiogenic neutrons, muon-induced neutrons and solar neutrinos contribute to this NR background.

Radiogenic neutrons are generated by (α, n) reactions and spontaneous fissions of material radioactive impurities, and are the largest source of NR background. The neutron yields of materials are predicted using SOURCES-4A [41] based on the measured radioactivity of the detector materials [8]. The propagation of generated neutrons is simulated using the GEANT4 toolkit [42].

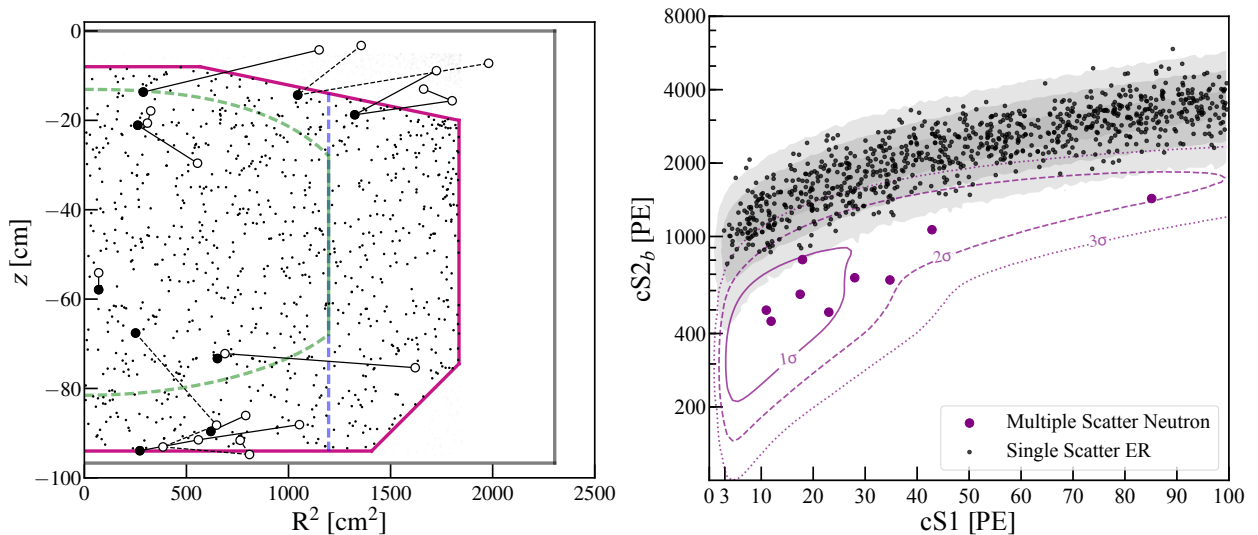Because of the large uncertainty ($\sim$50%) in the estimated neutron rate, multiple neutron scatter events

FIG. 5: Spatial (left panel) and (cS1, cS2$_b$) (right panel) distribution of identified multiple neutron scatters in SR0 and SR1 DM search data, along with blinded DM search data (ER events) shown in black dots. The solid circles in the left panel represent the primary (largest cS2$_b$) scatter positions while the hollow circles show the positions of the secondary scatters. The solid or dashed black line connects the scatters that are in the same event. XENON1T TPC boundaries are shown as solid gray lines. Different volumes defined in [9], 0.65 t (dashed green), 0.9 t (dashed blue) and 1.3 t (solid magenta), are shown. In the right panel, the 1σ, 2σ, and 3σ contours of the expected distribution of neutron multiple scatters are shown as solid, dashed and dotted purple lines, respectively. Solid purple circles show the primary (cS1, cS2$_b$) of the identified neutron multiple scatters in DM data. As a comparison, the shaded black regions display the 1σ (dark) and 2σ (light) probability density percentiles of the ER background component for SR1.

in the DM search data and calibration data are used to further constrain the rate uncertainty. For this purpose, multiple scatter events were unblinded prior to single scatters. Nine neutron multiple scatter events were identified in the DM search data (SR0 and SR1 combined) within the 1.3-tonne fiducial volume (FV), compared to an expectation of 6.4±3.2 events from simulation. The expected rate and uncertainty of reconstructed single scatters from radiogenic neutrons is derived with a likelihood study of the (cS2$_b$, cS1) distribution of these nine identified multiple scatters, and ratios of reconstructed single and multiple neutron scatters from simulation and NR calibration data. The prior on the ratio between reconstructed single and multiple neutron scatters is taken to be uniform between the simulation-predicted ratio and the observed ratio in NR calibration data. The resulting prediction of selected single scatter caused by radiogenic neutrons is 1.43±0.66 in the 1 tonne×year XENON1T DM search data. Misidentified multiple scatters by radiogenic neutrons account for about 35% of these reconstructed single scatters. Fig. 5 shows the spatial and (cS1, cS2$_b$) distributions of the identified neutron multiple scatters in DM search data together with the single scatter events surviving from the blinding cut [9].

Neutrons emitted from bottom PMTs in the bottom array have a probability to scatter in the region between the TPC cathode and PMTs, referred to as the "below-cathode" region. The scintillation light from these scatters is detected, but electrons are lost because the electric field in the below-cathode region drifts them away from the active volume. Such events, with at least one scatter in the below-cathode region and a single scatter in the 1.3 tonne FV, are named neutron-X events, and have a lower cS2$_b$ to cS1 ratio than standard neutron scatters. The relative ratio and difference in (cS1, cS2$_b$) distributions, as well as in spatial distributions, between normal neutron scatters and neutron-X events is taken into account in the simulation for building the NR background model.

Muon-induced neutrons are estimated to be subdominant to radiogenic neutrons in the active volume [38], and the rate of muon-induced neutrons is further suppressed by applying a muon-veto selection. The final rate is approximately 2 orders of magnitude smaller than that of the radiogenic neutrons, and its contribution is neglected in the NR background modelling.

Compared with [11], the NR background model induced by solar neutrino scatters has been updated with the latest results from COHERENT [43], and is 23% smaller. On top of the neutrino flux uncertainty of 14%, the updated NR background model also takes into account the uncertainty of about 21% from [43] and of about 22% from the signal response model at the energy region of interest for neutrino scatters. The resulting rate uncertainty for solar neutrino induced coherent elastic neutrino-nucleus scatters (CEνNS) is 34%.

C. Surface Background Model

Interactions in the bulk region of the detector can be modelled combining knowledge of LXe and detector-related responses, as discussed in the previous section for the ER and NR background models. For background events where the knowledge of the response is incomplete or missing, data-driven methods have to be developed to model the events. In this and the following section, two classes of such background components are analyzed for the XENON1T WIMP searches, events originating from the detector surface, and accidental coincidences of unrelated S1 and S2 signals.

Several experiments [44, 45] have demonstrated that detector surfaces exposed to ambient air during construction are contaminated by a large amount of radon progeny, in particular ^{210}Pb. With a 22 y half-life, ^{210}Pb decays at a constant rate within the lifetime of the XENON1T experiment. For WIMP searches, ion recoils of ^{206}Pb from ^{210}Po α-decays, β-decays and the resulting X-rays and Auger electrons of ^{210}Pb are particularly important. Due to incomplete knowledge of LXe responses and detector physics in presence of PTFE, as well as complicated decay structures, a full model including the relevant physics processes has not yet been achieved in XENON1T. Instead, a data driven approach is adopted to predict the distribution of this background.

Background from radon progeny was modelled in cS1, cS2$_b$, S2, R, and z spaces, with a distribution $f_{\mathrm{Surf}}(\mathrm{cS1, S2, cS2}_b, \mathrm{R}, z)$. Radial position of surface events are reconstructed nearly symmetrically around the TPC boundary, with an uncertainty determined by the S2. Events mis-reconstructed outside the TPC are used to model the background distribution in S2, cS2$_b$, cS1, and z, denoted as $f_{\mathrm{Surf}-1}(\mathrm{cS2}_b, \mathrm{S2, cS1}, z)$, with a kernel-density-estimation method [46]. The distribution in cS2$_b$ and cS1 is shown in Fig. 6. Due to significant charge losses at the PTFE panels, the surface background overlaps significantly with the nuclear recoil region of interest (between nuclear recoil median and -2σ quantile). In contrast, the R distribution provides excellent rejection power. To construct the distribution of surface background in R and S2 space, events originating at the PTFE surface are selected as control sample with an S1 size out of the region of interest, including the ^{210}Po α-decays. In each S2 slice, the radial distribution of the control-sample events is fitted, including an uncertainty estimated by different fitting functions. The 2-dimensional distribution of surface background, denoted as $f_{\mathrm{Surf}-2}(\mathrm{R, S2})$, is combined with $f_{\mathrm{Surf}-1}(\mathrm{cS2}_b, \mathrm{S2, cS1}, z)$ to form a complete model of the surface background as $f_{\mathrm{Surf}}(\mathrm{cS1, S2, cS2}_b, \mathrm{R}, z)$, including the uncertainty in the radial distribution.

The total rate is normalized to the number of events reconstructed outside the TPC boundary. In the WIMP search [9], a radius cut of 42.8 cm is placed to reduce the surface background to $\sim$ 100 events. In the likelihood fit described in section IV, the total surface background
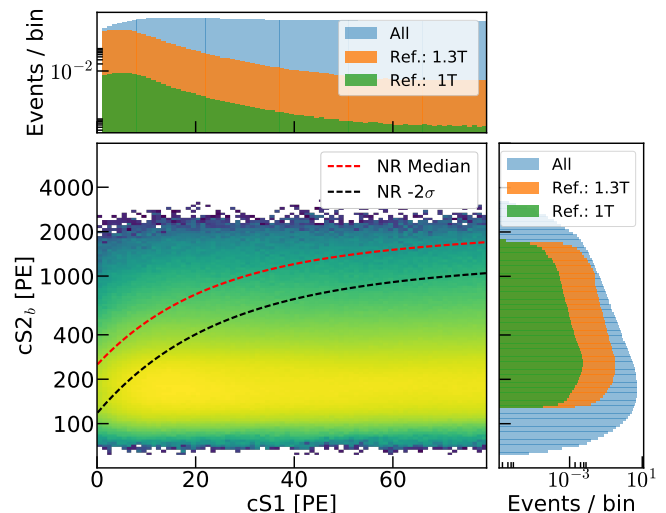

FIG. 6: Illustration of the surface background distributions in cS1 and $\mathrm{Log}_{10}(\mathrm{cS2}_b)$, with projections on each axis showing the expected distribution within the entire analysis space (blue), and in the reference region for the entire 1.3 tonne FV (brown), and the inner 1 tonne (green). The reference region lies between the NR median and -2σ quantile lines, marked by red and black lines, respectively.

expectation value is conservatively treated as a free parameter.

D. Accidental Coincidence Background Model

The accidental coincidence (AC) of uncorrelated S1s and S2s (referred to as lone-S1 and lone-S2, respectively) is the fourth background component considered in the XENON1T model. Lone-S1s and -S2s originate from energy depositions in the non-active regions of the detector, where the scintillation or ionization signal are not detectable. For example, an energy deposition in the below-cathode region does not produce an ionization signal, and a deposited energy very close to the gate may have its scintillation signal blocked by the mesh. Similar to the surface background model described above, AC background is constructed through a data-driven approach by random pile-up of lone-S1 and lone-S2 samples from data.

The lone-S1 sample is obtained by searching for S1s in the time window before the larger primary S1 in each digitized event waveform. The estimated lone-S1 rate ranges from 0.7 to 1.1 Hz, depending on the requirement of noise rate in the search window. The difference in lone-S1 rate between science runs is negligible. The lone-S2 sample is obtained using events with no S1 found in the digitized waveform, or with a reconstructed z position larger than the maximum drift time. The observed lone-S2 rate (with S2 threshold of 100 PE) is determined as 2.6 mHz and stays constant during both science runs. Selection criteria are applied in the lone-S1

and lone-S2 samples directly, while the event selections involving S2 (S1) are excluded. The AC event rate is calculated as $R_{AC} = R_{lS1} \cdot R_{lS2} \cdot \Delta t$, where R_{lS1} and R_{lS2} are the rates of lone-S1s and lone-S2s, respectively. Δt is the coincidence time window, which is fixed to the maximum drift time in the TPC (674 μs for SR0 and 727 μs for SR1). The AC background is nearly uniformly distributed in $(x,\ y,\ z)$ position space. This yields a total AC background of 44.2 events in SR1, which is reduced to 7.0 events with an S2 threshold of 200 PE. Several selection criteria were developed to suppress AC background to a sub-dominant contribution and their rejection efficiency were estimated by simulation.

We simulate the distribution of AC events by sampling and randomly pairing the lone-S1 and -S2 samples. Interaction positions of the event are calculated by applying a field distortion correction to the sampled position. The correction of S1 and S2, depending on the event position, as well as event selections are all applied to the simulated sample. Examples are the drift-time dependent S2 width and S1 fraction in top PMT array cuts. The final AC background prediction is done with the simulation sample where all selection criteria are applied. The AC model was validated before unblinding, using both ^{220}Rn calibration data and WIMP search data outside the ROI (for example the sample with S2 between 100 and 200 PE, or with S1s being identified as single electron S2s). The predicted AC rate after all selection criteria in the ROI for SR0 and SR1 combined is 0.47 - 0.74 events per tonne per year. As shown in Fig. 7, the AC distribution is concentrated in the low cS1 region, making it important in the search for light WIMPs.

E. WIMP Signal Model

The WIMP signal model depends on the dark matter mass, and assumes a uniform distribution of WIMP signals in the FV. The signal energy spectrum is computed in the same way as in [9, 11]. Distributions in (cS1, cS2$_b$), as well as the energy spectra for 10, 50, and 200 GeV/c^2 WIMPs are shown in Fig. 8. The uncertainty in the WIMP signal model in the (cS1, cS2$_b$) distribution, which comes from the uncertainty in the detector response model, is sub-dominant to the uncertainties of background models in the statistical inference. Therefore, we approximate the uncertainty of the WIMP signal model only in the form of its rate uncertainty. The approximated rate uncertainties as a function of WIMP mass are shown in Fig. 8.

IV. INFERENCE WITH XENON1T DATA

In this section we describe the general techniques used in XENON1T for hypothesis testing and construction of confidence intervals. The focus is on the description of the specific likelihood used for the statistical
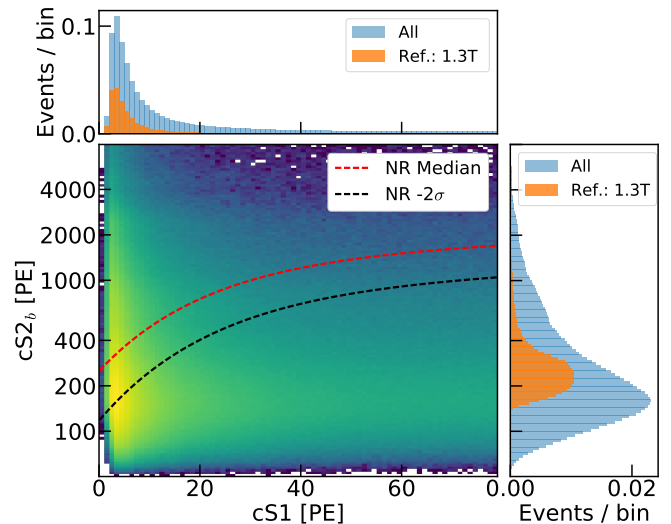

FIG. 7: Illustration of the accidental coincidence background distribution in cS1 and $\log_{10}(cS2_b)$, with projections on each axis showing the expected distribution within the entire analysis space (blue), and in the reference region for 1.3 tonne FV. The reference region lies between the NR median and -2σ quantile lines, marked by red and black lines, respectively.

interpretation of the XENON1T 1 tonne-year WIMPs search data [9, 47].

A. Hypothesis Testing and Confidence Intervals

The profile log-likelihood ratio is used as the test statistic for both confidence intervals and discovery assessment,

$$ q(\sigma) \equiv -2 \cdot \log \frac{\mathcal{L}(\sigma, \hat{\hat{\boldsymbol{\theta}}})}{\mathcal{L}(\hat{\sigma}, \hat{\boldsymbol{\theta}})}, \qquad (17) $$

where $\mathcal{L}$ stands for the XENON1T likelihood defined in Section IV B, $\hat{\sigma}$ and $\hat{\boldsymbol{\theta}}$ are the signal and nuisance parameters that maximize the likelihood overall, and $\hat{\hat{\boldsymbol{\theta}}}$ are the nuisance parameters that maximize the likelihood with the condition that the signal strength is σ. To avoid unphysical regions, the best-fit signal $\hat{\sigma}$ is constrained to be non-negative.

A signal hypothesis H_σ is tested against the data by computing the p-value of the observed test statistic $q(\sigma)$ given H_σ. In case of relatively large expected signal-like background, asymptotic formulae [48, 49] for the distribution of $q(\sigma)$ are convenient. However, given the low background of XENON1T, it was found that these approximations no longer hold. This led to an under-coverage of cross-section and resulted in a overestimated limit of 10% on average in the SR0 results [11]. Therefore, the distribution of $q(\sigma)$ is computed using toy Monte Carlo (toy-MC)
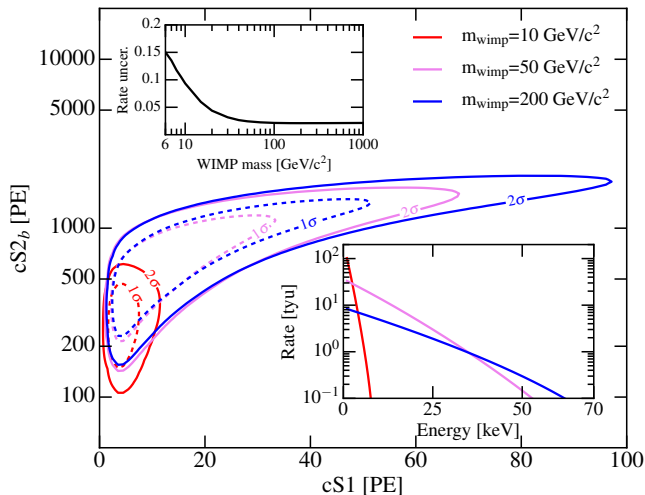
FIG. 8: The main panel shows the 1σ (dashed line) and 2σ contours of the WIMP signal model in (cS1, cS2$_b$) space, for WIMP mass of $10\,\mathrm{GeV/c^2}$ (red), $50\,\mathrm{GeV/c^2}$ (violet), and $200\,\mathrm{GeV/c^2}$ (blue), respectively. The lower-right inset shows the differential energy rate, which is in unit of tyu, for these three WIMP masses, with an assumed WIMP-nucleon cross section of $10^{-45}\mathrm{cm^2}$. The upper-left inset shows the rate uncertainty of WIMP signal model as a function of WIMP mass.
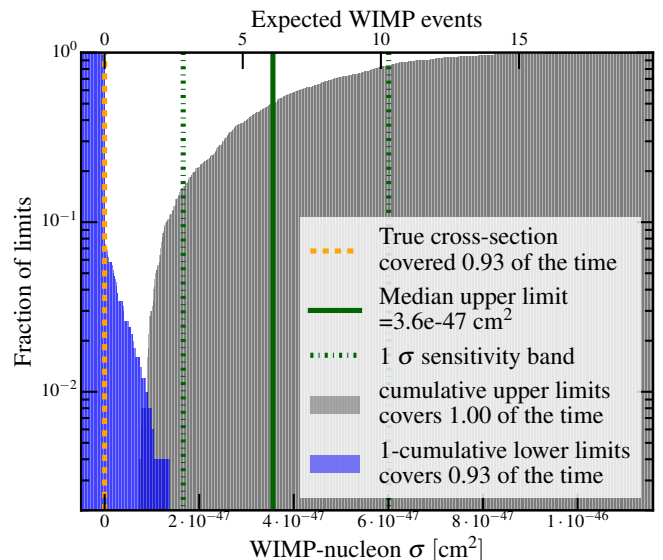
FIG. 9: Distribution of the lower (blue) and upper (gray) toy-MC limits on a 50 GeV/c² WIMP cross-section. The upper (lower) limits are drawn as a cumulative (complementary cumulative) distribution, to show the fraction of limits that include a certain cross-section. The corresponding signal expectation value is shown on the upper horizontal axis. The toy-MC did not include a true signal, indicated with a dashed orange line at 0. All upper limits are above 0, and 0.93 of lower limits are equal to 0, giving a total estimated coverage of 0.93 for the 500 toy-MCs used in this example, part of the data-set used for figure 11. As the cross-section is constrained to be non-negative, the survival fraction below 0 is 1. Green solid and dot-dashed lines show the median upper limit and 1σ sensitivity band.

simulations of the background and signal models. Note that in computing these distributions the auxiliary measurements related to each nuisance parameter are also varied per toy-MC data sample.

A Feldman-Cousins construction in the profile likelihood [50, 51] (also termed profile construction) is used to construct confidence intervals, using $q(\sigma)$. Sensitivities, as well as the coverage properties of the confidence band, detailed in Section IV D, are explored with toy-MC simulations. Fig. 9 shows the distributions of lower and upper limits for a background-only simulation, illustrating the sensitivity computation. Both the science and calibration data-sets are drawn from their model distributions, as are the ancillary measurements according to their uncertainties. For the sensitivity computation, 1000 toy-MCs are run per mass point.

Before unblinding, the XENON1T collaboration resolved to not report the lower edge of confidence intervals unless the discovery significance exceeds 3σ. Since the Feldman-Cousins construction without this threshold yields two-sided confidence intervals when the discovery significance exceeds 1.28σ corresponding to a 90% confidence level, this leads to over-coverage at low WIMP cross section values. The example shown in Fig. 11 shows that this threshold affects limits below the 1σ sensitivity band.

The FC construction may produce upper limits excluding signal strengths to which the experiment has a very small discovery power. In the case of a downwards fluctuation with respect to the background model, a 15% power-constraint would be applied to set a lower threshold for the upper limit, as proposed in [52].

The discovery significance is computed using the test-statistic evaluated at the null-hypothesis, $q_0 \equiv q(\sigma = 0)$. Denoting the distribution of q_0 under the null-hypothesis H_0 as $f_{H_0}(q_0)$, the significance of a given observed test statistic q_0^{obs} can be expressed with a p-value,

$$p_{H_0} = \int_{q_0^{\mathrm{obs}}}^{\infty} [f_{H_0}(q_0)]\,\mathrm{d}q_0. \tag{18}$$

In practice, the p-value is estimated using toy-MC samples to account for non-asymptoticity due to the low signal-background overlap. For the FC construction, the test statistic distribution is estimated for 20 steps in the true signal for each mass. For each signal step, 2500 toy-MC simulations, including ancillary measurements are generated and fitted. A threshold curve, constructed by a smooth interpolation between the 90th percentile of the test statistic for each signal size is compared with the final log-likelihood ratio to construct the FC intervals.

The local discovery significance in Eq. 18 is computed for a single signal hypothesis. In the spin-independent

analysis, the signal hypotheses are the WIMP masses considered between 6 GeV/c^2 and 10000 GeV/c^2. To compute a global discovery significance, the distribution of the most significant p-value for any hypothesis is estimated with 10000 toy-MC simulations. The global significance is found by comparing the local p-value from data with this distribution. For moderate excesses of $\sim 2\sigma$, the global significance is $\sim 0.5\sigma$ lower than the local one.

B. The XENON1T Likelihood

The log-likelihood used in the spin-independent analysis is a sum of extended un-binned log-likelihoods for the two science runs. Additional terms are the extended un-binned likelihoods for ER calibration data, and terms expressing ancillary measurements of nuisance parameters θ_m,

$$\log \mathcal{L}_{\text{total}}(\sigma, \boldsymbol{\theta}) = \sum_{\text{SR}} \log \mathcal{L}_{\text{SR sci}}(\sigma, \boldsymbol{\theta})$$
$$+ \sum_{\text{SR}} \log \mathcal{L}_{\text{SR cal}}(\boldsymbol{\theta})$$
$$+ \sum_{m} \log \mathcal{L}_m(\theta_m), \qquad (19)$$

where SR runs over data-taking periods, SR0 and SR1, σ is the WIMP-nucleon cross-section, and $\mathcal{L}_{\text{sci}}$, $\mathcal{L}_{\text{cal}}$ are the likelihood terms for the WIMP search data and ^{220}Rn calibration data, respectively. Ancillary measurements of nuisance parameters θ_m are included in $\log \mathcal{L}_m(\theta_m)$. The un-binned science likelihood is defined in three dimensions: cS1, cS2$_b$ and R, the radius of the reconstructed event. Each background and signal distributions are defined and normalized in this three-dimensional space. The un-binned likelihoods take the form

$$\mathcal{L}_{\text{SR sci}}(\sigma, \boldsymbol{\theta}) = \text{Pois}(N_{\text{SR}}|\mu_{\text{tot}}(\sigma, \boldsymbol{\theta}))$$
$$\times \prod_{i=1}^{N_{\text{SR}}} \left[\sum_c \frac{\mu_c(\sigma, \boldsymbol{\theta})}{\mu_{\text{tot}}(\sigma, \boldsymbol{\theta})} \cdot f_c(\vec{x}_i|\boldsymbol{\theta}) \right],$$
$$(20)$$
$$\mu_{\text{tot}}(\sigma, \boldsymbol{\theta}) \equiv \sum_c \mu_c(\sigma, \boldsymbol{\theta}). \qquad (21)$$

Here, the index i runs over events in the relevant science data-set, and c runs over each signal or background component, with expectation value $\mu_c(\sigma, \boldsymbol{\theta})$, which may be a nuisance parameter or a function of nuisance parameters. The probability density functions (PDFs), $f_c(\vec{x}_i|\boldsymbol{\theta})$, for each component are functions of the analysis coordinates $\vec{x}_i = (\text{cS1}_i, \text{cS2}_{b\,i}, \text{R}_i)$ and are evaluated for each event in the likelihood. The models for the science

data measurements, both the WIMP signal model, and the ER, neutron, CEνNS, surface and AC backgrounds are described in detail in the previous sections. In the case of the calibration likelihood, which utilizes a smaller volume, 0< R <36.94 cm, the likelihood uses only the cS1 and cS2 dimensions, but the structure of $\mathcal{L}_{\text{SR cal}}$ is otherwise identical to Eq. 21. At high radius, the radiogenic neutron background described in section III B increases towards the wall, and the surface background, described in section III C dominates the total rate. Extending the detector modelling to include R rather than applying an optimized cut improves the sensitivity to a 50 GeV/c^2 WIMP by 10%. At the radial cut of 42.8 cm, no significant improvement in sensitivity can be obtained by increasing the volume. In addition to the three analysis dimensions in which the distributions are modelled, events in the science data-sets are also classified as inside or outside a core mass of 0.65 t (shown in Fig. 5). Practically, this can be considered as a combination of two separate un-binned likelihoods in the shape of Eq. 21, where the relative expectations inside and outside the core mass are determined using the R, z distribution of the different components.

C. Nuisance Parameters

The signal and background models, consisting of expectation values and distributions in analysis space, depend on several nuisance parameters. Table III lists all the parameters of the combined likelihood, the data sets that mainly constrain them and their best-fit value. Uncertainties on the nuisance parameters are computed using the profiled likelihood in each nuisance parameter, but unlike the signal confidence intervals, the asymptotic construction is applied. Nuisance parameters are grouped in rate parameters, the expectation value for all background components, signal efficiency parameters and shape parameters that affect the model distributions. Expectation values for the 5 modelled backgrounds in the science data, as well as AC and ER rates in the ^{220}Rn calibration likelihood, are all nuisance parameters in the likelihood. In the following, a separate nuisance parameter and term is applied for each science run, with the exception of the radiogenic rate.

The radiogenic rate, as well as the expected CEνNS rates for each science run, is constrained by ancillary measurements expressed as Gaussian likelihoods,

$$\mathcal{L}_m(\mu_c) = \text{Gaus}(\hat{\mu}_c|\mu_c, \sigma_c). \qquad (22)$$

These likelihoods are defined by the PDF for the ancillary measurement of the component $\hat{\mu}_c$, given a true expectation value μ_c and measurement uncertainty σ_c. The signal expectation is

$$\mu_{\text{sig}} = \sigma \cdot \epsilon \cdot \mu_{\text{ref}}(M)/\sigma_{\text{ref}} \qquad (23)$$

TABLE III: Table of parameters of the full XENON1T likelihood, best-fit values, and $1\,\sigma$ confidence interval for a 200 GeV/c^2 WIMP. In addition, the data-set that constrains the parameter is noted. Parameters are arranged according to whether they affect background rates, the shape of the background distributions, or the signal distribution. For the shape parameters, the models affected by the parameter are also noted. Due to the mismodelling parameter, the signal distribution, which changes with the WIMP mass, may also affect the ER model best-fit. The rates of the models in each data-set are expressed as the expected numbers of events in each of the data-sets; SR0 (SR1) expectation values refer to a 0.12 t × y (0.88 t × y) exposure. The radiogenic neutron event rate, shared between the science runs, refers to the 1.0 t × y exposure.

Rate Parameter		Constraint	Expectation Value
SR0 AC events		Ancillary measurement	$0.051^{+0.035}_{-0}$
SR1 AC events		Ancillary measurement	$0.42^{+0.29}_{-0}$
SR0 ^{220}Rn AC events		Ancillary measurement	$0.73^{+0.51}_{-0}$
SR1 ^{220}Rn AC events		Ancillary measurement	$2.8^{+1.9}_{-0}$
SR0 CEνNS events		Ancillary measurement	$0.0040^{+0.0014}_{-0.0014}$
SR1 CEνNS events		Ancillary measurement	$0.050^{+0.017}_{-0.017}$
SR0 science data ER		SR0 data-set	$73.4^{+8.8}_{-8.4}$
SR1 science data ER		SR1 data-set	554^{+24}_{-24}
SR0 calibration data ER		SR0 ^{220}Rn data-set	689^{+27}_{-26}
SR1 calibration data ER		SR1 ^{220}Rn data-set	5264^{+73}_{-72}
Radiogenic events in SR0+SR1		Ancillary measurement	$1.44^{+0.65}_{-0.66}$
SR0 surface events		SR0 science data	$12.5^{+3.9}_{-3.2}$
SR1 surface events		SR1 science data	$93.7^{+10}_{-9.5}$
Shape Parameter	Model Affected	Constraint	Value
ER photon yield	All ER models	^{220}Rn calibration	$-0.043^{+0.053}_{-0.052}$
SR0 ER recombination fluctuation	SR0 ^{220}Rn, science models	SR0 ^{220}Rn data-set	$0.64^{+0.58}_{-0.70}$
SR1 ER recombination fluctuation	SR1 ^{220}Rn, science models	SR1 ^{220}Rn data-set	$0.32^{+0.41}_{-0.39}$
SR0 safeguard as fraction of ER	SR0 ^{220}Rn, science models	SR0 ^{220}Rn data-set	$-0.0060^{+0.0049}_{-0.0036}$
SR1 safeguard as fraction of ER	SR1 ^{220}Rn, science models	SR1 ^{220}Rn data-set	$-0.0049^{+0.0014}_{-0.0045}$
SR0 surface shape parameter	SR0 surface model	SR0 science data	$1.00^{+0.0}_{-0.67}$
SR1 surface shape parameter	SR0 surface model	SR1 science data	$-0.39^{+0.24}_{-0.22}$
Signal Parameter		Constraint	Value
SR0 signal efficiency		NR model uncertainty	$1.000^{+0.036}_{-0.036}$
SR1 signal efficiency		NR model uncertainty	$1.000^{+0.021}_{-0.021}$
WIMP cross-section [10^{-45}cm^2]			$0.0421^{+0.1165}_{-0}$
WIMP mass [GeV/c^2]	Signal, ER mismodelling term	Fixed in analysis	200

where $\mu_{\rm ref}$ is the reference expectation for a WIMP of mass M and cross-section $\sigma_{\rm ref}$, given by the signal acceptance, and ϵ is a multiplicative factor expressing the uncertainty on the signal expectation for a fixed cross-section. This expectation uncertainty is constrained by the NR model posterior taking all model variations into account,

$$\mathcal{L}_{\rm m}(\epsilon) = {\rm Gaus}(1|\epsilon, \sigma_\epsilon(M)), \qquad (24)$$

where the uncertainty $\sigma_\epsilon(M)$ depends on the WIMP mass, ranging from 0.15 at 6 GeV/c^2 to 0.03 at 200 GeV/c^2, as shown in the inset of Fig. 8. The nominal value of 1 reflects that the best-fit expectation is expressed by the reference expectation.

The AC rate $\mu_{\rm AC}$ is constrained between two extreme estimates of its rate and is assigned a uniform PDF between the lower and upper reference, written as $0.6 \cdot \hat{\mu}_{\rm AC}$ and $\hat{\mu}_{\rm AC}$ for convenience,

$$\mathcal{L}_{\rm m}(\mu_{\rm AC}) = {\rm Uniform}(\mu_{\rm AC}|0.6 \cdot \hat{\mu}_{\rm AC}, \hat{\mu}_{\rm AC}). \quad (25)$$

The ER and surface background rates are not assigned auxiliary measurements, as their high statistics in the science data sample constrain them. In the case of the surface background, the region of highest signal overlap (at low cS1 and R) was blinded. This motivated the conservative procedure of not placing an auxiliary constraint on the surface shape.

In addition to uncertainties on the rate, the ER and surface background PDFs are also assigned shape uncertainties. The surface background sideband fit, described in Section III C, is also used to construct an uncertainty for the radial slope of the background. To avoid over-constraining this distribution, which could lead to spurious excesses or too-tight confidence intervals, no constraint is placed on the radial slope from the sideband measurement. The science data fit finally

constrains the radial slope with approximately 2.5 times smaller uncertainty than the sideband fit.

The ER model is described in detail in Section III A. The nuisance parameters γ_{er} and Δr described there are propagated to the likelihood, with the former shared between science runs. The value of these nuisance parameters and their associated errors are determined in the combined fit by including the ER calibration likelihood term for ^{220}Rn data. This ER model includes the nominal ER model, as well as variations due to changing the photon yield or recombination fluctuation parameter. It is slightly different from the one illustrated in Section II, as the uncertainties in the model from nuisance parameters other than γ_{er} and Δr are not included in the nominal ER model.

A mismodelling term, or "safeguard", proposed in [53], is a shape-uncertainty added to the ER model, consisting of a signal-like component added to or subtracted from the nominal ER model. This ensures that regardless of other nuisance parameters, the ER model will have the freedom to fit the calibration data in the signal-like region. A spurious signal-like over- or under-estimation in the ER background will have the greatest impact on the inference, giving spuriously constraining limits or spuriously significant excesses, respectively. The safeguard mainly affects the ER model tail which overlaps with the signal region, as shown in Fig. 10. The ER model, including the mismodelling term is constrained by the calibration likelihood terms included in the total likelihood, this allows the safeguard component to be constrained by the much higher (approximately 10 times for SR0 and SR1) statistics of the ^{220}Rn calibration data compared to the science data. The ER background PDF including the safeguard term, PDF $f'_{ER}(\vec{x}|\boldsymbol{\theta}_{ER})$ can be written as

$$ f'_{ER} \equiv b \cdot [(1 - a_{mm}) \cdot f_{ER}(\vec{x}|\boldsymbol{\theta}_{ER}) + a_{mm} \cdot f_{sig}(\vec{x})], \tag{26} $$

where a_{mm} is the safeguard nuisance parameter and f_{sig} the signal PDF. If the safeguard causes the PDF to be negative in a region, it is truncated to 0. The pre-factor b, a function of a_{mm} and nuisance parameters that affect the ER distribution $\boldsymbol{\theta}_{ER}$, ensures that the PDF is normalized in the analysis space.

The compatibility between the best-fit to data and the safeguard-equal-zero hypothesis is performed using the profiled log-likelihood ratio to compute the 1-sigma error on the best-fit safeguard. The calibration data indicates that the signal-like tail of the ER is less pronounced than in the nominal model, reflected in a negative safeguard fit. A zero safeguard is $2.9\,\sigma$ from the SR1 +SR0 combined fit using this method.
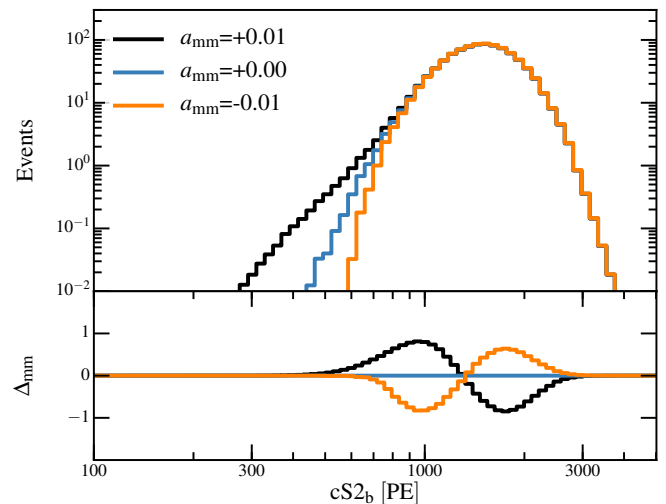

FIG. 10: Illustration of the effect of a mismodelling term on the ER model for a slice of parameter space $20 < cS1 < 30$ PE, projected onto cS2$_b$. The upper curve shows the ER model for safeguard fractions -0.01, 0 and 0.01, showing how a positive (negative) mismodelling term raises (lowers) the signal-like tail of the ER model with respect to the nominal model in blue. The lower panel shows Δ_{mm}, the difference between each model and the nominal, demonstrating both the effect on the tail at low cS2$_b$ and the opposite at higher cS2$_b$, due to the normalization of the PDF.
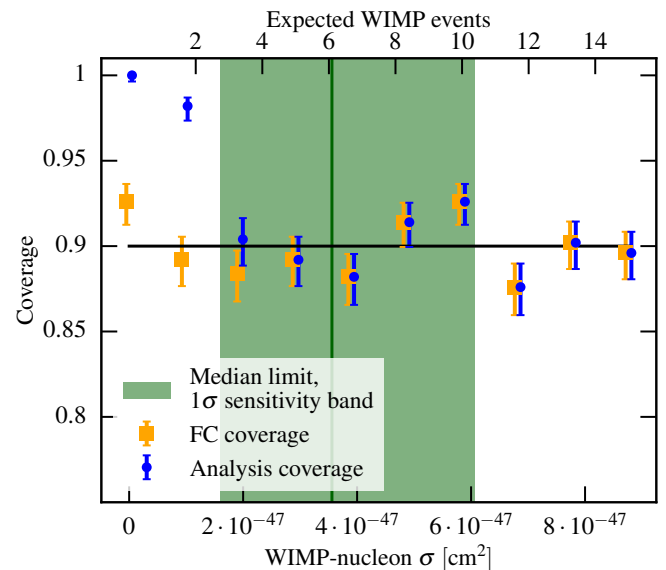

FIG. 11: Figure displaying the coverage of confidence intervals for the spin-independent 1 tonne-year analysis and a 50 GeV/c^2 WIMP. Error-bars indicate $1\,\sigma$ confidence intervals around the best estimate. Orange squares show the result using the profile construction, while the blue circles show the coverage of the XENON1T analysis including the $3\,\sigma$ threshold for reporting upper limits. The upper and lower x-axes show the WIMP cross-section and expectation (setting $\epsilon = 1$) respectively, and the green band and line highlight the sensitivity and $1\,sigma$ band of upper limits for the analysis.

D. Coverage

The fraction of repeated experiments where the confidence interval contains the true parameter is called the coverage. Perfect coverage is equal to the confidence interval, 0.9 in the case of XENON1T. While the nominal FC construction, introduced in section IV A, provides coverage by construction, the coverage of the profile construction must be investigated for the likelihood in question. As we decided to report only the upper edge of the confidence interval for discovery significances $< 3\,\sigma$, there will be over-coverage at very low signal sizes. This has a similar effect to the power constraint. Fig. 11 shows the coverage for a 50 GeV/c^2 WIMP, both for the profile construction (orange points), and the analysis including the $3\,\sigma$ threshold (blue points). The green band shows the $-1\,\sigma$ to $1\,\sigma$ sensitivity band. The result is consistent with perfect coverage (black line), with overcoverage for the $3\,\sigma$ threshold only under the $-1\,\sigma$ edge of the sensitivity band. The effect on coverage of mismeasuring the nuisance parameters was also studied. In the case of the mismodelling term, a shift in the parameter by three times the observed value was required for a one percentage point shift in the coverage.

V. SUMMARY

In this manuscript we have reported details of the detector response model, the background and WIMP signal models, and the statistical inference of XENON1T analysis chain. These have been used in the interpretation of results in the search for spin-independent elastic WIMP-nucleon interactions in XENON1T [9, 11], and is also used in searches for alternative dark matter candidates or interactions using XENON1T data [47, 54]. The background model, using a simulation-based detector response model, has been improved with respect to previous analyses by reduction of the systematic bias and a better treatment of parameter uncertainties and correlations. The statistical inference was developed to use a combined, unbinned likelihood that is adaptable to multiple analysis spaces and additional data sets. A signal-like mismodeling term is introduced for the first time in an analysis as a background model shape uncertainty. In addition, the inference now employs toy Monte Carlo simulations extensively to construct and validate the confidence bands.

Acknowledgement

[1] N. Aghanim et al. (Planck Collaboration) (2018), arXiv:1807.06209.
[2] D. Akerib et al., Phys. Rev. Lett. **118**, 021303 (2017).
[3] E. Aprile et al., Phys. Rev. D **94**, 122001 (2016).
[4] X. Cui et al., Phys. Rev. Lett. **119**, 181302 (2017).
[5] K. Abea et al. (2018), arXiv:1804.02180.
[6] F. Petricca et al. (2017), arXiv:1711.07692.
[7] R. Agnese et al. (2015), arXiv:1509.02448.
[8] E. Aprile et al., Eur. Phys. J. C **77**, 881 (2017).
[9] E. Aprile et al., Phys. Rev. Lett. **121**, 111302 (2018).
[10] E. Aprile et al. (XENON Collaboration) (2019), In preparation.
[11] E. Aprile et al., Phys. Rev. Lett. **119**, 181301 (2017).
[12] M. Szydagis, N. Barry, K. Kazkaz, J. Mock, D. Stolp, M. Sweany, M. Tripathi, S. Uvarov, N. Walsh, and M. Woods, JINST **6**, P10002 (2011).
[13] B. Lenardo, K. Kazkaz, A. Manalaysay, J. Mock, M. Szydagis, and M. Tripathi, IEEE Trans. Nucl. Sci **62**, 1412 (2015).
[14] J. Lindhard, V. Nielsen, M. Scharff, and P. Thomsen, Mat. Fys. Medd. Dan. Vid. Selsk **33**, 1 (1963).
[15] D. Akerib et al., Phys. Rev. D **93**, 072009 (2016).
[16] J. Thomas and D. Imel, Phys. Rev. A **36**, 614 (1987).
[17] E. Aprile et al., Phys. Rev. D **88**, 012006 (2013).
[18] L. Goetzke, E. Aprile, M. Anthony, G. Plante, and M. Weber, Phys. Rev. D **96**, 103007 (2017).
[19] D. Huang, Bulletin of the American Physical Society **60** (2015).
[20] E. Boulton et al., JINST **12**, P08004 (2017).
[21] M. Szydagis et al., *Noble element simulation technique v2.0* (2018), URL https://doi.org/10.5281/zenodo.1314669.
[22] E. Aprile, K. Giboni, P. Majewski, K. Ni, M. Yamashita, R. Hasty, A. Manzur, and D. McKinsey, Phys. Rev. D **72**, 072006 (2005).
[23] E. Aprile, C. Dahl, L. De Viveiros, R. Gaitskell, K.-L. Giboni, J. Kwong, P. Majewski, K. Ni, T. Shutt, and M. Yamashita, Phys. Rev. Lett. **97**, 081302 (2006).
[24] E. Aprile, L. Baudis, B. Choi, K. Giboni, K. Lim, A. Manalaysay, M. Monzani, G. Plante, R. Santorelli, and M. Yamashita, Phys. Rev. C **79**, 045807 (2009).
[25] G. Plante, E. Aprile, R. Budnik, B. Choi, K.-L. Giboni,

L. Goetzke, R. Lang, K. Lim, and A. M. Fernandez, Phys. Rev. C **84**, 045805 (2011).

[26] P. Sorensen et al., Nucl. Inst. and Meth. A **601**, 339 (2009).

[27] A. Manzur, A. Curioni, L. Kastens, D. McKinsey, K. Ni, and T. Wongjirad, Phys. Rev. C **81**, 025808 (2010).

[28] D. Akerib et al., arXiv:1608.05381 (2016).

[29] E. Aprile et al., Phys. Rev. D **97**, 092007 (2018).

[30] D. Akerib et al., Phys. Rev. D **93**, 072009 (2016).

[31] E. Aprile et al., Phys. Rev. Lett. **119**, 181301 (2017).

[32] C. Wagner, A. Naumkin, A. Kraut-Vass, J. Allison, C. Powell, and J. Rumble Jr, U. S. Department of Commerce. (2003).

[33] C. Faham, V. Gehman, A. Currie, A. Dobi, P. Sorensen, and R. Gaitskell, JINST **10**, P09010 (2015).

[34] B. L. Paredes, H. Araújo, F. Froborg, N. Marangou, I. Olcina, T. Sumner, R. Taylor, A. Tomás, and A. Vacheret, Astropart. Phys. **102**, 56 (2018).

[35] XENON Collaboration, *The pax data processor v6.8.0* (2018), URL https://doi.org/10.5281/zenodo.1195785.

[36] J. Goodman and J. Weare, Comm. App. Math. Comp. Sci. **5**, 65 (2010).

[37] A. Gelman, X.-L. Meng, and H. Stern, Statistica sinica pp. 733–760 (1996).

[38] E. Aprile et al., J. Cosmol. Astropart. Phys. **2016**, 027 (2016).

[39] E. Aprile et al., Eur. Phys. J. C **77**, 275 (2017).

[40] S. Lindemann and H. Simgen, Eur. Phys. J. C **74**, 2746 (2014).

[41] W. Wilson et al., Tech. Rep., Technical Report LA-13639-MS, Los Alamos (1999).

[42] S. Agostinelli, Nucl. Inst. and Meth. A **506**, 250 (2003).

[43] D. Akimov et al., Science **357**, 1123 (2017).

[44] R. Agnese et al., Phys. Rev. D **92**, 072003 (2015).

[45] P.-A. Amaudruz et al., Phys. Rev. Lett. **121**, 071801 (2018).

[46] F. Pedregosa et al., J. Mach. Learn. Res. **12**, 2825 (2011).

[47] E. Aprile et al. (XENON Collaboration) (2019), arXiv:1902.03234.

[48] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, Eur. Phys. J. **C71**, 1554 (2011), [Erratum: Eur. Phys. J.C73,2501(2013)].

[49] S. S. Wilks, Ann. Math. Statist. **9**, 60 (1938), URL https://doi.org/10.1214/aoms/1177732360.

[50] G. J. Feldman and R. D. Cousins, Phys. Rev. **D57**, 3873 (1998).

[51] C. Patrignani et al. (Particle Data Group), Chin. Phys. **C40**, 100001 (2016).

[52] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, pre-print (2011), [physics.data-an/1105.3166], 1105.3166.

[53] N. Priel, L. Rauch, H. Landsman, A. Manfredini, and R. Budnik, JCAP **1705**, 013 (2017).

[54] E. Aprile et al. (XENON Collaboration) (2019), arXiv:1811.12482.