An approximate likelihood for nuclear recoil searches with XENON1T data

XENON Collaboration; Aprile, E.; Angevaare, J.R.; Bruenner, S.A.; Colijn, A.P.; Decowski, M. P.; Di Pede, S.; Gaemers, P.W.

DOI
10.1140/epjc/s10052-022-10913-w

Publication date
2022

Document Version
Final published version

Published in
European Physical Journal C

License
CC BY

Citation for published version (APA):
An approximate likelihood for nuclear recoil searches with XENON1T data

XENON Collaboration


1 Physics Department, Columbia University, New York, NY 10027, USA
2 Kamioka Observatory, Institute for Cosmic Ray Research, and Kavli Institute for the Physics and Mathematics of the Universe (WPI), University of Tokyo, Higashi-Mozumi, Kamioka, Hida, Gifu 506-1205, Japan
3 Department of Physics and Astronomy, University of Bologna and INFN-Bologna, 40126 Bologna, Italy
4 LPNHE, Sorbonne Université, CNRS/IN2P3, 75005 Paris, France
5 Institut für Physik and Excellence Cluster PRISMA+, Johannes Gutenberg-Universität Mainz, 55099 Mainz, Germany
6 Institut für Kernphysik, Westfälische Wilhelms-Universität Münster, 48149 Münster, Germany
7 Department of Physics and Astronomy, Weizmann Institute of Science, 76100 Rehovot, Israel
8 INFN-Laboratori Nazionali del Gran Sasso and Gran Sasso Science Institute, 67100 L’Aquila, Italy
9 Physikalisches Institut, Universität Freiburg, 79104 Freiburg, Germany
10 Department of Physics and Kavli Institute for Cosmological Physics, University of Chicago, Chicago, IL 60637, USA
11 New York University Abu Dhabi-Center for Astro, Particle and Planetary Physics, Abu Dhabi, United Arab Emirates
12 Physik-Institut, University of Zürich, 8057 Zurich, Switzerland
13 Department of Physics and Astronomy, Purdue University, West Lafayette, IN 47907, USA
14 INFN-Laboratori Nazionali del Gran Sasso and Gran Sasso Science Institute, 67100 L’Aquila, Italy
15 Physikalisches Institut, Universität Freiburg, 79104 Freiburg, Germany
16 SUBATECH, IMT Atlantique, CNRS/IN2P3, Université de Nantes, 44307 Nantes, France
17 Department of Particle Physics and Astrophysics, Weizmann Institute of Science, 7610001 Rehovot, Israel
18 LIPPhy, Department of Physics, University of Coimbra, 3004-516 Coimbra, Portugal
19 Max-Planck-Institut für Kernphysik, 69117 Heidelberg, Germany
20 Department of Physics “Ettore Pancini”, University of Napoli and INFN-Napoli, 80126 Napoli, Italy
21 Institute for Astrophysics and Astronomy, University of Leipzig, 04109 Leipzig, Germany
22 Department of Physics and Chemistry, University of L’Aquila, 67100 L’Aquila, Italy
23 Department of Physics and Astronomy, Rice University, Houston, TX 77005, USA
24 Department of Physics and Center for High Energy Physics, Tsinghua University, Beijing 100084, China
25 Kobayashi-Maskawa Institute for the Origin of Particles and the Universe, and Institute for Space-Earth Environmental Research, Nagoya University, Furo-cho, Chikusa-ku, Nagoya, Aichi 464-8602, Japan
26 Department of Physics, University of California San Diego, La Jolla, CA 92093, USA
Abstract The XENON collaboration has published stringent limits on specific dark matter – nucleon recoil spectra from dark matter recoiling on the liquid xenon detector target. In this paper, we present an approximate likelihood for the XENON1T 1 t-year nuclear recoil search applicable to any nuclear recoil spectrum. Alongside this paper, we publish data and code to compute upper limits using the method we present. The approximate likelihood is constructed in bins of reconstructed energy, profiled along the signal expectation in each bin. This approach can be used to compute an approximate likelihood and therefore most statistical results for any nuclear recoil spectrum. Computing approximate results with this method is approximately three orders of magnitude faster than the likelihood used in the original publications of XENON1T, where limits were set for specific families of recoil spectra. Using this same method, we include toy Monte Carlo simulation-derived binwise likelihoods for the upcoming XENONnT experiment that can similarly be used to assess the sensitivity to arbitrary nuclear recoil signatures in its eventual 20 t-year exposure.

1 Introduction

Persuasive astrophysical and cosmological evidence for the existence of dark matter has led to numerous direct detection efforts for weakly interacting massive particles (WIMPs) over the last 20 years [2]. Amongst these was the XENON1T [3] experiment, which collected 1 t-year of exposure from 2016 to 2018. It culminated in the then most stringent limits for spin-independent (SI) WIMP nucleon interactions above 6 GeV/c² [4] at the time. Subsequent limits on spin-dependent WIMP interactions with neutrons and protons [5] as well as WIMP-pion couplings [6] have been published. In each of the aforementioned interactions the expected signature in the detector is a nuclear recoil (NR), induced by the single scatter of a WIMP off a xenon atom. All four searches use the same background and detector response models and NR search data set. Other XENON results may be applicable to some NR interactions, such as ionisation-only signatures [7] or Migdal effect searches [8]. For WIMPs below ∼ 10 GeV/c², the best XENON1T limits are provided by a dedicated low-energy NR search searching for solar 8B neutrinos [9]. The SI recoil spectrum and a fixed halo model are the standard for reporting direct-detection WIMP searches [10]. Different interactions or dark matter fluxes, either from alternate dark matter halo models [11] or methods for generating boosted dark matter [12] can yield different spectra. As the exact halo parameters are uncertain, and any candidate dark matter particle may interact through a number of different channels, a robust method to constrain arbitrary nuclear recoil spectra is required.

In the full likelihood used for the XENON1T NR searches there are two data-taking periods, each with an accompanying electronic recoil (ER) calibration set and ancillary measurement terms constraining the detector response and microphysics parameters as well as background models represented in 20 nuisance parameters [13]. Each science data set is modelled in three analysis dimensions (discussed in Sect. 2), with five background components (presented in Sect. 2). This complexity was reflected in the computational expense, requiring about ∼ 30 s for a toy Monte Carlo (toy-MC) simulation of the analysis.

In this paper we present the profiled likelihood of the XENON1T NR search in bins of reconstructed energy, a description of how it may be used to calculate upper limits for a generic NR spectrum and a data release with accompanying code [1] allowing the physics community to use this method to recast the XENON1T result. The computation is fast, taking about ∼ 40 ms to compute an upper limit for a recoil spectrum. We present comparisons to the full toy-MC simulation computation result for several recoil spectra. For heavy WIMPs, the limit computed with the approximate likelihood is typically conservative and within 10% of the full-likelihood computation, while lower-energy recoil spectra see a higher spread around the full-likelihood upper limit of up to ∼ 30%. Finally, we extend this work by including the XENONnT 20 t-year sensitivity projection [14], with 1000 toy-MC simulation binwise likelihoods, so that the sensitivity of this projection can be evaluated for any NR signature.

In Sect. 2 we give an overview of the XENON1T NR search, highlighting the analysis dimensions used in the inference, and in Sect. 3 we discuss the response to NRs. We
present our statistical model in Sect. 4 and the exact methodology in Sect. 5. Section 6 details how to use this approach for approximate limits, and provides estimates of the bias and variance of the method for a selection of NR recoil spectra.

2 XENON1T nuclear recoil search

XENON1T was designed and optimized to detect the low-energy NRs expected from WIMPs recoiling off xenon nuclei [3]. Its primary detector was a dual phase xenon time projection chamber (TPC) containing 2 t of instrumented liquid xenon, observing scintillation and ionization charges from interactions in the target. Prompt scintillation light is observed from the recombination or de-excitation of xenon ions or dimers, respectively, and is referred to as the S1 signal. Ionization electrons are drifted to the liquid-gas interface at the top of the detector by means of a drift field applied between a cathode electrode at the bottom of the chamber and a grounded gate electrode just below the liquid-gas interface. The electrons produce scintillation light proportional to the charge, referred to as the S2 signal, when they are extracted into the gas by a higher extraction field. Xenon scintillation was observed by 248 photomultiplier tubes (PMTs) arranged in two arrays at the top and bottom of the detector. The x–y position of the interaction was inferred from the pattern of S2 photons observed by the top PMT array, while the time separation between the S1 and the S2 signal indicated the z-depth. With access to the full 3D position information we could fiducialize the detector volume, selecting only the innermost 1.3 t xenon volume where contributions from radioactivity in the detector materials are minimized.

Both S1 and S2 signals were corrected to account for the detector’s position dependent light collection efficiency [15], and in the case of the S2 we also corrected for electron attachment to impurities in the liquid xenon volume as the electrons are drifted upwards. These corrected S1 and S2 variables are named cS1 and cS2.

The relative size of the ionisation and scintillation signals, and therefore cS1 and cS2, depends on whether the incident particle scattered off the nucleus (NR) or an electron (ER) of a xenon atom. In Fig. 1 the predicted 1σ contour for interactions of a 50 (6) GeV/c² WIMP, which is expected to interact with the xenon nucleus, producing NRs, is shown in orange (purple). We also illustrate the signal expectation from a mono-energetic 30 keV NR line in red. The 1σ contour of the ER background is shown in blue, demonstrating the separation between nuclear and electronic recoils in XENON1T. Shown in green is the 1σ contour of the “wall” background, which is discussed at the end of this section.

WIMPs are expected to scatter at most once off a target nucleus due to their small interaction cross sections, therefore XENON1T optimized its search strategy to look for single scatter NR events. The analysis space spans from 3 to 70 photoelectrons (PE) in the cS1 space, where the lower boundary is driven by the detection efficiency, and the upper boundary is chosen to include the bulk of the expected WIMP signal. We use the light observed in the bottom PMT array to determine magnitude of the position corrected S2s, referred to as cS2b, due to the more uniform response of this array in the x–y plane. The cS2b space is chosen to fully contain the expected background and signal models in our chosen cS1 region and spans from 50 to 7940 PE, corresponding to approximately 1.5–250 electrons.

The full XENON1T exposure was collected in two science campaigns, SR0 and SR1, between November 2016 and February 2018, with drift fields of 120 V/cm and 81 V/cm, respectively. Continual purification of the xenon improved the electron lifetime from 380 µs at the start of SR0 to ∼650 µs at the end of SR1. The final data, after quality selections detailed in [15] and fiducialization consisted of 739 events in a 1 t-year exposure, shown in Fig. 1 as grey circles.

The response of the detector to low-energy ER and NR interactions was calibrated with 220Rn, the decay products of which produce low-energy beta-decays, and 241AmBe and deuterium-deuterium fusion generator neutron sources. We
used a detector response model based on a fast detector simulation to fit the calibration data and model ER and NR sources in XENON1T [13].

Background models for five sources of interactions within XENON1T were considered, detailed in [13]. The largest background is ERs induced by the $^{214}$Pb decay product of $^{222}$Rn or decays of $^{85}$Kr. The second largest background expectation is referred to as the “wall” background. These are events which occur close to the polytetrafluoroethylene walls of the detector, and consequently lose a portion of the ionization electrons to the wall as they drift upwards. The lower S2 signal, observed close to the detector edge will result in larger position reconstruction errors, and this population will therefore bleed into the fiducial volume. For this reason, we include the radius, denoted by R, as an analysis dimension along with cS1 and cS2b for the background and signal models. The 1σ contours of these two dominant backgrounds is shown in Fig. 1 in blue (ERs) and green (wall). The other backgrounds considered are radiogenic neutrons from detector materials, coherent elastic neutrino-nucleus scattering (CE$
u$NS) of solar $^8$B neutrinos, and accidental pairing of lone S1 and S2 signals.

3 Analysis variables and detector response

Previous XENON1T searches for WIMP interactions [13, 15] directly used the observed cS1 and cS2b variables as described in Sect. 2. Since the total number of quanta produced is dependent on the original energy deposition, the number of prompt scintillation photons and ionization electrons, observed as the cS1 and cS2b, respectively, are intrinsically anti-correlated. Additionally, the fraction of quanta observed as ionization electrons or scintillation photons is energy dependent. Thus for a given cS1 selection, different NR energies yield different distributions in cS2b space. To reduce the dependence on the recoil energy, we transform our analysis space to explicitly feature reconstructed energy as one dimension.

3.1 Reconstructed energy

The reconstructed ER energy $E_{recER}$ of the original interaction can be obtained from cS1 and cS2b quantities as:

$$E_{recER}(cS1, cS2b) \equiv W \cdot \frac{cS1}{g_1} - \frac{cS2b}{g_2}.$$  

where $W = 13.7 \text{ eV}$ is the average amount of energy required to produce one electron or photon in xenon [16]. The detector dependent quantities $g_1$ and $g_2$ represent the number of photoelectrons observed in the PMT arrays per emitted scintillation photon and the number of photoelectrons observed per extracted electron respectively.

Since the approximate likelihood will be presented in bins of $E_{rec}$, it is necessary that other analysis dimensions are as independent of recoil energy as possible. Therefore, we also introduce $E_{rec}^\perp$,

$$E_{rec}^\perp(cS1, cS2b) \equiv W \cdot \frac{cS1}{g_2} - \frac{cS2b}{g_1},$$  

which is constructed so that $E_{recER}$ and $E_{rec}^\perp$ contours are perpendicular.

Performing the analysis in $E_{rec}$, $E_{rec}^\perp$ coordinates rather than in cS1, cS2b is only a coordinate transformation, and does not affect the XENON1T inference results.

In order to obtain the reconstructed recoil energy for NR events ($E_{rec}$), one must also account for the energy dependent quenching effect, where NR energy is lost to unobserved heat. We estimate the quenching magnitude at a given energy from an empirical comparison between the true NR energy and the reconstructed ER energy using the detector response model described in [15]. The constant NR energy lines obtained from the above procedure are shown in Fig. 1 as gray shaded bands.

3.2 Migration matrix

In order to convert an arbitrary NR spectrum into the reconstructed energy spectrum expected to be observed in XENON1T we account for detector effects. The complete detector response model, derived from fits to calibration data and accounting for detection efficiency, resolution and correction effects is described in [15]. Using this model, we calculate the spread in reconstructed energy space of a fine grid of true NR recoil energies. The migration matrix is shown in Fig. 2, where the components of the migration matrix

$$\mathcal{P}_{r,t} = P(E_{rec} \text{ in bin } r \mid E_{true} \text{ in bin } t),$$

represent the probability for a NR recoil in some true recoil bin to be reconstructed in a given reconstructed energy bin. The transformation of the true recoil energy spectra for a 6 and 50 GeV/c$^2$ WIMP into bins in reconstructed energy space is shown in purple and orange respectively. Also shown is the transformation of a mono-energetic 30 keV line, illustrating the broadening of the signal spectrum from detector effects.

4 Statistical model

We use a profiled log-likelihood ratio test statistic and toyMC simulations of the test statistic distribution to compute discovery significances and confidence intervals. The likelihood $\mathcal{L}_{total}$ used for NR searches with XENON1T is presented in [13]. It is a product of:

- $\mathcal{L}_{SR}^\text{unbinned} (s, \theta \mid x)$: unbinned, extended likelihood terms in three analysis dimensions: cS1, cS2b, and R, for the
two science data-taking periods, labeled SR0 and SR1 (indexed with SR). The likelihood is a function of the signal strength parameter $s$, and the set of nuisance parameters $\theta$, and is evaluated for the data $x$.

- $L_{\text{SR}}(\theta | x)$: unbinned, extended likelihood terms in two analysis dimensions; cS1 and cS2 for the $^{220}$Rn calibration data taken for each science data-taking period. Since this calibration source is uniformly distributed in the detector, R is not included.

- $L_{\text{anc}}(\theta | x_{\text{anc}})$: terms representing ancillary measurements of background rates and the signal detection efficiency, with $x_{\text{anc}}$ being the ancillary measurements.

The aim of this paper is to present an approximate likelihood applicable to any NR signal in an easily publishable format. To that end, we first reparameterise the signal and background models to be in $E_{\text{rec}}$, $E_{\perp}^{\text{rec}}$ and R, and write separate likelihood terms, primed to mark the reparameterisation, $L_{\text{r,SR}}^{\text{prim}}$ for bins $r$ in reconstructed energy. These two changes leave the likelihood unaltered (up to a constant factor)

$$L^{\text{tot}}(s, \theta) = \prod_r \prod_{\text{SR}} L_{\text{r,SR}}^{\text{prim}}(s, \theta) \times L_{\text{cal}}^{\text{cal}}(\theta) \times L_{\text{anc}}^{\text{anc}}(\theta).$$

The per-bin science data likelihood for bin $r$ with observed events $N_r$ and lower and upper edges $E_{\text{rec},d}$ and $E_{\text{rec},u}$ is

$$L_{\text{r,SR}}^{\text{prim}}(s, \theta) = \text{Poisson}(N_r | \mu_r^{\text{tot}}(s, \theta)) \times \prod_i f^{\text{tot}}(E_{\text{rec},i}, E_{\perp,\text{rec},i}^{\perp}, R_i | s, \theta)$$

where $f^{\text{tot}}(E_{\text{rec}}, E_{\perp}^{\text{rec}}, R | s, \theta)$ is the total probability density function (PDF) in the transformed analysis variables, and $S_r = \{i | E_{\text{rec},d} < E_{\text{rec},i} < E_{\text{rec},u}\}$ is the set of events in the science run with $E_{\text{rec}}$ in bin $r$. The total expected number of events in each bin $r$, and the expectation from each source $j$ in that bin are defined as

$$\mu_r^{\text{tot}}(s, \theta) \equiv \sum_j \mu_{j,r}(s, \theta)$$

$$\mu_{j,r}(s, \theta) \equiv \mu_j(s, \theta) \times \int_{E_{\text{rec},d}}^{E_{\text{rec},u}} \left( \int_{E_{\text{rec},d}}^{E_{\text{rec}}^{\perp}} f_j(E_{\text{rec}}, E_{\perp}^{\text{rec}}, R | s, \theta) \, dE_{\text{rec}} \right) \, dE_{\text{rec}}.$$

and the science likelihood for the bin to one using this averaged PDF,

$$L_{\text{r,SR}}^{\text{prim}}(s, \theta) = \text{Poisson}(N_r | \mu_r^{\text{tot}}(s, \theta)) \times \prod_i f^{\text{tot}}(E_{\text{rec},i}, R_i | s, \theta)$$

$$\prod_{\text{SR}} L_{\text{r,SR}}^{\text{prim}}(s, \theta) \equiv \prod_{\text{SR}} L_{\text{r,SR}}^{\text{cal}}(s, \theta).$$

The total approximate likelihood is the product of each binwise contribution times the calibration and ancillary constraint terms,

$$L^{\text{tot}}(s, \theta) = \prod_r \left( L_{\text{r}}^{\text{prim}}(s, \theta) \right) \times L_{\text{cal}}^{\text{cal}}(\theta) \times L_{\text{anc}}^{\text{anc}}(\theta).$$

Fig. 2 Illustration of the migration matrix included in the data release [1], as defined in Eq. 3, showing the conversion between true NR recoil energy and the reconstructed energy. The bottom panel shows the true NR spectrum of a 30 keV line in red, and spin-independent (SI) WIMP recoil spectra for a 6 GeV/c² and 50 GeV/c² WIMP in purple and orange, respectively, all with arbitrary normalisation. The left panel shows the same spectra in reconstructed energy after multiplication with the migration matrix. The matrix is normalized such that selections in $E_{\text{rec}}$ account for our overall detection efficiency.
5 Binwise profiling

For the binwise-averaged likelihoods to be a good approximation to the unbinned likelihood, the bins must be small with respect to the XENON1T resolution. In Sect. 6.1, we choose the bin number \( n \) to minimise bias and maximise accuracy. To produce a likelihood for any signal shape, we wish to compute profiled likelihood ratios for each bin separately. However, the chosen binning is so narrow that many nuisance parameters in \( \theta \), for instance the normalisation of the wall background, cannot be constrained in each bin separately. In practice, no nuisance parameter is strongly pulled from its best-fit value in the original XENON1T upper limit computation. Therefore, our second approximation is to first compute \( \hat{\theta}_0 \), the value of the nuisance parameters that optimises \( \mathcal{L}^{\text{cal}}(0, \theta) \), and fix the nuisance parameters to this value. The exception is the ER mismodelling term (and therefore also the ER normalisation) that requires special attention:

Over- or under-estimating a signal-like tail of the background model would bias results towards too-strict limits or spurious discoveries, respectively. Therefore, the XENON1T WIMP search likelihood [13] includes an ER mismodelling term [17] that takes the form of a signal-like component added to the ER model,

\[
f_{\text{ER}}(x) \rightarrow \gamma(\alpha) \times \max[(1 - \alpha)f_{\text{ER}}(x) + \alpha f_{\text{SIG}}(x), 0]
\]

(12)

where \( f_{\text{ER}}(x), f_{\text{SIG}}(x) \) are the PDFs in \( x \) of the ER background and (WIMP) signal, respectively, \( \alpha \) is the size of the ER mismodelling term and \( \gamma(\alpha) \) is a normalisation term to ensure that the total PDF is normalised even for negative \( \alpha \). Since this term depends on the signal model considered, it cannot be determined by the background-only fit, and must be profiled per bin. The total ER distribution used in the likelihood becomes

\[
f_{\text{ER}, r}(x | \alpha_r) = \begin{cases} \gamma(\alpha_r) \times \max[0, (1 - \alpha_r)f_{\text{ER}}(x)] + \alpha_r \times f_{\text{SIG}}(x), & \text{if } E_{\text{rec}, d} < E_{\text{rec}, u} \\ \gamma(\alpha_r) \times f_{\text{ER}}(x | \hat{\theta}_0), & \text{otherwise}, \end{cases}
\]

(13)

which is used both in the calibration and science data likelihoods. Each bin has its own mismodelling component, parameterized with \( \alpha_r \), which therefore can more freely fit the calibration data shape, resulting in an improved fit. The total calibration PDF is the sum of \( f_{\text{ER}}(x | \hat{\theta}_0) \) and the accidental background component making \( \mathcal{L}^{\text{cal}}(\alpha_r) \) have the same form as Eq. 5. Since the mis-modelling term only affects the shape of the background, the normalisation in the calibration term is fixed to the best-fit value.

Using the ER model of Eq. 13 and the best-fit nuisance parameters for the no-signal fit \( \theta_0 \), thereby fixing \( \mathcal{L}^{\text{cal}} \), we construct the likelihood in each bin of reconstructed energy,

\[
\mathcal{L}^{\text{tot}}(s, \alpha_r, \mu_r^\text{ER})_r = \mathcal{L}^{\text{sci}}(s_r, \alpha_r, \mu_r^\text{ER}, \hat{\theta}_0) \times \mathcal{L}^{\text{cal}}(\alpha_r, \hat{\theta}_0).
\]

(14)

Here, \( s_r \) is the signal expectation in each reconstructed energy bin \( r \), which relates to the expectation in each bin of true energy \( t \) via the migration matrix

\[
s_r = \sum_t \mathcal{P}_{r,t} \cdot s_t,
\]

(15)

and the signal expectation in each true energy bin \( t \) in turn is given by

\[
s_t = s \int_{E_{\text{true}, u}}^{E_{\text{true}, d}} g(E) dE_{\text{true}},
\]

(16)

where \( g(E) \) is the signal PDF in true recoil energy \( E_{\text{true}} \), and \( s \) the expected number of true signal events.

The binwise profiling follows the approach in [18, 19], where the likelihood is profiled separately in sections of the analysis variable space. The profiled likelihood in each bin is

\[
\lambda_r(s) = -2 \times \log \left( \frac{\mathcal{L}^{\text{tot}}(s, \hat{s}, \hat{\alpha}_r, \hat{\mu}_r^\text{ER})}{\mathcal{L}^{\text{tot}}(\hat{s}, \hat{\alpha}_r, \hat{\mu}_r^\text{ER})} \right)
\]

(17)

which where \( \hat{s}, \hat{\alpha}_r, \hat{\mu}_r^\text{ER} \) is the signal expectation value, mismodelling fraction and ER rate that maximises the likelihood, and \( \hat{\alpha}_r, \hat{\mu}_r^\text{ER} \) maximise the conditional likelihood. Figure 3 shows the profiled binwise likelihood for each bin as function of \( s \). The per-bin likelihoods show the expected fluctuation from a lower-statistics sample – some prefer a positive signal, others no signal. To compute a full result, they must be combined into one likelihood.

6 Inference using the binwise likelihood

Using the energy migration matrix defined in Sect. 3 to compute bin-wise signal expectations \( s_r \), together with the likelihood ratio for each bin defined in Eq. 17, we can write our approximation of the log-likelihood written in Eq. 4

\[
\Lambda_{\text{tot}}(s) = \sum_r \lambda_r(s_r)
\]

(18)

and the corresponding log-likelihood ratio

\[
\lambda_{\text{tot}}(s) = \Lambda_{\text{tot}}(s) - \Lambda_{\text{tot}}(\hat{s}).
\]

(19)
Using this approximate likelihood induces only a moderate systematic and random error in confidence intervals with respect to the ones computed with the full, computationally much slower XENON1T likelihood. Best-fit and upper limits are then computed using the standard asymptotic formulae [20,21].

6.1 Fidelity of the approximate likelihood method

Differences between the unbinned and approximate binwise results are the binning in $E_{\text{rec}}$, the per-bin ER mismodelling and profiling, and the slight change in the signal distribution in $E_{\text{rec}}$ in individual bins for different signal shapes.

We validated the performance of the binwise likelihood approach by computing upper limits for a range of signal spectra and different numbers of bins in $E_{\text{rec}}$. Table 1 shows the median ratio between the limits computed with the approximate and full likelihood, and errors corresponding to the 15th and 85th percentiles of the ratio between the two. Increasing the number of bins beyond 80 bins between 0 and 60 keV did not markedly improve either the bias or spread of the upper limits for the binwise likelihood. Therefore we choose to report the result of this work using 80 bins in reconstructed energy space. For heavy WIMPs, the bias and errors are both on the order of 10%. The more peaked low-mass WIMP signals or lower-energy monoenergetic lines, both

### Table 1

<table>
<thead>
<tr>
<th>Signal Type</th>
<th>n = 60 Bins</th>
<th>n = 80 Bins</th>
<th>n = 120 Bins</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flat spectrum</td>
<td>$1.01^{+0.08}_{-0.06}$</td>
<td>$1.01^{+0.07}_{-0.05}$</td>
<td>$1.01^{+0.07}_{-0.04}$</td>
</tr>
<tr>
<td>NR lines</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 keV</td>
<td>$1.06^{+0.33}_{-0.16}$</td>
<td>$1.04^{+0.29}_{-0.15}$</td>
<td>$1.03^{+0.23}_{-0.15}$</td>
</tr>
<tr>
<td>5 keV</td>
<td>$1.13^{+0.24}_{-0.14}$</td>
<td>$1.11^{+0.19}_{-0.13}$</td>
<td>$1.11^{+0.17}_{-0.12}$</td>
</tr>
<tr>
<td>7 keV</td>
<td>$1.14^{+0.14}_{-0.10}$</td>
<td>$1.13^{+0.16}_{-0.12}$</td>
<td>$1.13^{+0.15}_{-0.12}$</td>
</tr>
<tr>
<td>10 keV</td>
<td>$1.11^{+0.15}_{-0.12}$</td>
<td>$1.11^{+0.14}_{-0.11}$</td>
<td>$1.11^{+0.12}_{-0.10}$</td>
</tr>
<tr>
<td>20 keV</td>
<td>$1.05^{+0.14}_{-0.09}$</td>
<td>$1.04^{+0.12}_{-0.08}$</td>
<td>$1.05^{+0.12}_{-0.08}$</td>
</tr>
<tr>
<td>30 keV</td>
<td>$1.04^{+0.11}_{-0.09}$</td>
<td>$1.03^{+0.10}_{-0.08}$</td>
<td>$1.04^{+0.10}_{-0.08}$</td>
</tr>
<tr>
<td>SI WIMP signals</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6 GeV/c²</td>
<td>$1.07^{+0.21}_{-0.21}$</td>
<td>$1.08^{+0.21}_{-0.19}$</td>
<td>$1.01^{+0.25}_{-0.17}$</td>
</tr>
<tr>
<td>10 GeV/c²</td>
<td>$1.08^{+0.14}_{-0.14}$</td>
<td>$1.06^{+0.19}_{-0.14}$</td>
<td>$1.05^{+0.17}_{-0.13}$</td>
</tr>
<tr>
<td>50 GeV/c²</td>
<td>$1.07^{+0.10}_{-0.08}$</td>
<td>$1.06^{+0.09}_{-0.07}$</td>
<td>$1.07^{+0.07}_{-0.07}$</td>
</tr>
<tr>
<td>100 GeV/c²</td>
<td>$1.06^{+0.10}_{-0.08}$</td>
<td>$1.06^{+0.09}_{-0.06}$</td>
<td>$1.06^{+0.07}_{-0.06}$</td>
</tr>
</tbody>
</table>

![Fig. 3](Image 82x549 to 248x713)

**Fig. 3** Illustration of the binwise, profiled log-likelihood $\lambda_r(s)$ for bins in reconstructed NR energy. The total approximate likelihood is obtained by summing over the entry in each reconstructed energy bin at the expected signal, as in Eq. 18. The purple, orange and red lines indicate the expectation values in each bin for a 6 GeV/c² and a 50 GeV/c² spin-independent WIMP signals and a 30 keV NR line signal respectively, at their respective upper limits derived from the XENON1T dataset. White bins at the highest and lowest reconstruction energies reflect bins for which the migration matrix $0$.

![Fig. 4](Image 259x550 to 267x713)

**Fig. 4** Top: The 90-percentile threshold of the approximate log-likelihood ratio test statistic as function of the true signal expectation. Thresholds estimated with toy-MC simulations for a range of monoenergetic signals are shown with black dots, and the magenta line shows the smoothed maximum. The threshold converges to the asymptotic value for around ~4 expected signal events. Bottom: The coverage of 95, 90 and 68-percent confidence level upper limits are shown with diamonds, squares and circles, respectively, for five NR recoil spectra: Flat (blue), a 3 keV monoenergetic line (red), a 6 GeV/c² SI WIMP (purple) and a 50 GeV/c² SI WIMP (orange).
Concentrated in only a few bins, have a larger range of deviation from the full result, up to 30% scatter with respect to upper limits with the full likelihood. The bias and errors in Table 1 give an indication of how well the approximate likelihood should be expected to perform for different signal shapes and energy ranges.

6.2 Correcting for non-asymptoticity

The XENON1T results were computed from test statistic distributions estimated using toy-MC simulations of datasets.

This was necessary due to the non-asymptotic nature of the distributions for the low signal-numbers considered [13]. Since generating datasets depends on the signal model, this approach must be amended if a similar correction should be applied to the likelihood ratio of Eq. 18.

Our approach is motivated by the observation that the non-asymptotic behaviour of the XENON1T likelihood is driven by the signal-to-background discrimination that leaves the signal region almost background free. Computing the test statistic distribution for a range of monoenergetic NRs will include the best signal-to-background discrimination. Other NR signals will be broader and therefore feature less extreme
ER-NR discrimination than these monoenergetic signals. Therefore, we compute the 90th percentile thresholds of the test statistic for a fine grid of monoenergetic NR signals, and choose the 90th percentile of all thresholds, to avoid statistical fluctuations, before smoothing this threshold using a Gaussian filter. Figure 4 (top) shows the thresholds computed as a function of the signal, while Fig. 4 (bottom) shows the coverage for several signal spectra using the smoothed upper envelope of these thresholds together with Eq. 19 to compute upper limits for several recoil spectra. All show either the nominal coverage, or conservatively over-cover at low signal expectations. We therefore recommend using this threshold rather than the asymptotic $\chi^2$ threshold to compute frequentist confidence intervals. In the data release, we include 68, 90 and 95-percentile thresholds.

Upper limits computed with the binwise approximation and with the binwise approximation plus the non-asymptotic threshold are compared with all XENON1T high-mass NR searches in Fig. 5. Close agreement is seen except at low masses, where the binwise approximation yields a higher, and thus more conservative, upper limit.

7 Summary

This paper and the accompanying code and data release provide a fast and flexible method to compute approximate results of the XENON1T NR search [4] for any NR spectrum. As many spectra can be tested, care should be taken when interpreting the likelihood to compute discovery significances. On the other hand, we have validated with toy-MC simulations and comparisons with the XENON1T full likelihood that good agreement is found for confidence intervals. We also provide a method to ensure that these confidence intervals have, on average, correct or over-coverage only. In the appendix, we also show how this method can be employed to provide recasts of sensitivity projections, in this case of the 20 t-year XENONnT projection presented in [14]. Together with the XENON1T ER spectral search [22,23] and ionisation-only [7,24] publications, the approximate NR likelihood provides a range of recastable legacy results of the XENON1T experiment.

Acknowledgements We gratefully acknowledge support from the National Science Foundation, Swiss National Science Foundation, German Ministry for Education and Research, Max Planck Gesellschaft, Deutsche Forschungsgemeinschaft, Helmholtz Association, Dutch Research Council (NWO), Weizmann Institute of Science, Israeli Science Foundation, Fundacao para a Ciencia e a Tecnologia, Region des Pays de la Loire, Knut and Alice Wallenberg Foundation, Kavli Foundation, JSPS Kakenhi in Japan and Istituto Nazionale di Fisica Nucleare. This project has received funding/support from the European Union’s Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No 860881-HIDDeN. Data processing is performed using infrastructures from the Open Science Grid, the European Grid Initiative and the Dutch national e-infrastructure with the support of SURF Cooperative. We are grateful to Laboratori Nazionali del Gran Sasso for hosting and supporting the XENON project.

Data Availability Statement This manuscript has associated data in a data repository. [Authors’ comment: The code and the data representing the approximate likelihood presented in this paper are available in a python package at [1].]

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

Funded by SCOAP³. SCOAP³ supports the goals of the International Year of Basic Sciences for Sustainable Development.

Appendix A: XENONnT projection

The successor to XENON1T, XENONnT, operates under the same principle but is designed with three times the active volume of XENON1T and a lower background [14]. As an example of how the approximate likelihood approach can be applied to projections as well, we computed toy-MC binwise likelihoods for 1000 no-signal simulations of the experiment, and compare to the published projections using the full likelihood. Using these, the limit-setting potential of the assumed detector model and exposure can be estimated for any NR signal.

In the projections of its sensitivity [14] in a 20 t-year exposure, we assumed an electron lifetime of 1 ms at a drift field of 200 V/cm. The overall ER background is assumed to be reduced by a factor of 6 from that reported in [4] through selective choice of detector materials and the introduction of a radon distillation column to further reduce the $^{214}$Pb background. A neutron veto is added around the cryostat which contains the time projection chamber of XENONnT in order to suppress the NR background by rejecting 87% of single-scatter neutron interactions in the active volume.

We do not include a wall model in the sensitivity projections, but select a 4 t fiducial volume further from the detector walls than in XENON1T, to minimise this contribution. Without the addition of the wall model, we choose to model the remaining backgrounds in only the cS1 and cS2b parameter spaces, treating their radial dependence as uniform. No background model is considered for accidental coincidence of lone S1s and S2s either. Additionally, rather than implementing a single model for CEνNS, we implement
two models, one for solar neutrinos and one for the diffuse supernova neutrino background and atmospheric neutrinos.

We assume the same low-energy detection efficiency as in XENON1T, therefore the lower boundary of our analysis space in cS1 remains at 3 PE. The higher light collection efficiency expected in XENONnT implies that more light should be detected for each scintillation photon produced, and we therefore adjust our upper boundary to 100 PE to contain the full spin-independent WIMP recoil spectrum.

Table 2 validates the recasting of the XENONnT 20 t-year projection, again showing good performance, and Fig. 6 compares the spin-independent WIMP-nucleon limits using this method and the previously published projection.

References


Table 2 Table of bias and spread of the binwise upper limit with respect to the full likelihood result

<table>
<thead>
<tr>
<th>Model</th>
<th>Bias</th>
<th>Spread</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flat spectrum</td>
<td>0.98</td>
<td>0.05</td>
</tr>
<tr>
<td>NR lines</td>
<td>1.00</td>
<td>0.21</td>
</tr>
<tr>
<td>3 keV</td>
<td>1.00</td>
<td>0.31</td>
</tr>
<tr>
<td>5 keV</td>
<td>1.00</td>
<td>0.13</td>
</tr>
<tr>
<td>7 keV</td>
<td>1.00</td>
<td>0.11</td>
</tr>
<tr>
<td>10 keV</td>
<td>1.00</td>
<td>0.10</td>
</tr>
<tr>
<td>20 keV</td>
<td>1.01</td>
<td>0.06</td>
</tr>
<tr>
<td>30 keV</td>
<td>1.01</td>
<td>0.08</td>
</tr>
</tbody>
</table>

Fig. 6 Comparison between 90% confidence level sensitivity bands for a projected 20 t-year XENONnT search for spin-independent WIMP-nucleon interactions show good agreement between the approximate likelihood and the full likelihood. The published result using the full likelihood [14] is in black. Blue line and band indicate the upper limit using the binwise approximate likelihood presented in this paper.