Likelihood approach to the first dark matter results from XENON100

E. Aprile *et al.* (XENON100 Collaboration)

Likelihood Approach to the First Dark Matter Results from XENON100

E. Aprile,[1] K. Arisaka,[2] F. Arneodo,[3] A. Askin,[4] L. Baudis,[4] A. Behrens,[4] K. Bokeloh,[5] E. Brown,[2] T. Bruch,[4]
J. M. R. Cardoso,[6] B. Choi,[1] D. Cline,[2] E. Duchovni,[7] S. Fattori,[3, 8] A. D. Ferella,[4] K.-L. Giboni,[1] E. Gross,[7, *]
A. Kish,[4] C. W. Lam,[2] J. Lamblin,[9] R. F. Lang,[1, †] K. E. Lim,[1] S. Lindemann,[10] M. Lindner,[10] J. A. M. Lopes,[6]
T. Marrodán Undagoitia,[4] Y. Mei,[11] A. J. Melgarejo Fernandez,[1] K. Ni,[12] U. Oberlack,[8, 11] S. E. A. Orrigo,[6]
E. Pantic,[2] G. Plante,[1] A. C. C. Ribeiro,[6] R. Santorelli,[1, 4] J. M. F. dos Santos,[6] M. Schumann,[4, 11] P. Shagin,[11]
A. Teymourian,[2] D. Thers,[9] E. Tziaferi,[4] O. Vitells,[7] H. Wang,[2] M. Weber,[10] and C. Weinheimer[5]

(The XENON100 Collaboration)

[1]*Physics Department, Columbia University, New York, NY 10027, USA*
[2]*Physics & Astronomy Department, University of California, Los Angeles, USA*
[3]*INFN Laboratori Nazionali del Gran Sasso, Assergi, 67100, Italy*
[4]*Physics Institute, University of Zürich, Winterthurerstr. 190, CH-8057, Switzerland*
[5]*Institut für Kernphysik, Wilhelms-Universität Münster, 48149 Münster, Germany*
[6]*Department of Physics, University of Coimbra, R. Larga, 3004-516, Coimbra, Portugal*
[7]*Department of Particle Physics and Astrophysics, Weizmann Institute of Science, 76100 Rehovot, Israel*
[8]*Institut für Physik, Johannes Gutenberg Universität Mainz, 55099 Mainz, Germany*
[9]*SUBATECH, Université de Nantes, 44307 Nantes, France*
[10]*Max-Planck-Institut für Kernphysik, Saupfercheckweg 1, 69117 Heidelberg, Germany*
[11]*Department of Physics, Rice University, Houston, TX 77005 - 1892, USA*
[12]*Department of Physics, Shanghai Jiao Tong University, Shanghai, 200240, China*

Many experiments that aim at the direct detection of Dark Matter are able to distinguish a dominant background from the expected feeble signals, based on some measured discrimination parameter. We develop a statistical model for such experiments using the Profile Likelihood ratio as a test statistic in a frequentist approach. We take data from calibrations as control measurements for signal and background, and the method allows the inclusion of data from Monte Carlo simulations. Systematic detector uncertainties, such as uncertainties in the energy scale, as well as astrophysical uncertainties, are included in the model. The statistical model can be used to either set an exclusion limit or to quantify a discovery claim, and the results are derived with a proper treatment of statistical and systematic uncertainties. We apply the model to the first data release of the XENON100 experiment, which allows to extract additional information from the data, and place stronger limits on the spin-independent elastic WIMP-nucleon scattering cross-section. In particular, we derive a single limit, including all relevant systematic uncertainties, with a minimum of $2.4 \times 10^{-44} \, \mathrm{cm}^2$ for WIMPs with a mass of $50 \, \mathrm{GeV/c}^2$.

I. INTRODUCTION

It is now well established that the dominant mass fraction of our Universe consists of some yet-unknown form of dark matter [1]. Well-motivated models predict Dark Matter in the form of Weakly Interacting Massive Particles (WIMPs) [2] for which searches can be conducted in direct scattering experiments located on Earth. In such experiments, WIMPs are expected to induce nuclear recoils with a roughly exponentially falling spectrum that extends to at most a few tens of keV in energy [3]. In contrast, the dominant background is usually composed of electronic recoils from β- or γ-radiation. Most direct Dark Matter search experiments are able to discriminate nuclear recoils from electronic recoils based on some dis-

crimination parameter, such as the ratio of ionization to phonon signal [4, 5], scintillation to phonon signal [6], ionization to scintillation signal [7, 8], or the pulse shape of the scintillation light [9]. In a plot of this discrimination parameter as function of energy, the signal and background events thus separate. Typically, a signal acceptance area is defined a priori in this parameter space. Only events that fall within this area are considered as WIMP candidate events for the interpretation of the result in terms of a Dark Matter detection, or for the calculation of limits on the WIMP-nucleon scattering cross-section.

Such a hard cut has some obvious disadvantages. Most importantly, the particular location of the signal candidate events within the signal acceptance area is not taken into account. In general, the expected signal and background will follow a different spectrum, which is utilized by the maximum-gap and optimum-interval methods [10], a frequentist approach to calculate limits. The generalization to two dimensions, the maximum-patch

---

method [11], extends this feature to the case where background leaks into the signal acceptance region from high or low discrimination parameters. An intrinsic advantage of these methods is that the resulting limit is robust against small changes of the signal acceptance region, and they are widely used in discriminating direct Dark Matter searches today. However, these methods also have some severe drawbacks: they are designed to always result in an upper limit, and a natural transition to a claim of detection, as for example inherent to the Feldman-Cousins-approach [12], is not possible. In addition, systematic uncertainties are not taken into account. Thus, the resulting limit can be quoted at a given statistical confidence level (CL), typically taken to be 90%, but systematic uncertainties need to be treated separately.

In this paper, we present an approach that overcomes both of these problems of the conventional treatment of data from direct Dark Matter detection experiments. As a practical example, we use data from the first Dark Matter results of the XENON100 experiment [8]. This data was previously analyzed in the classical way, by pre-defining a signal acceptance region based on calibration data. In particular, this region was defined using the prompt scintillation signal as a measurement of recoil energy, in the range between 4-20 photoelectrons (PE). In energy, this corresponds to about $9 - 33\,\mathrm{keV_{nr}}$ (keV nuclear recoil), depending on the energy calibration through the assumed relative scintillation efficiency $\mathcal{L}_{\mathrm{eff}}$ [13, 14]. The logarithm of the ratio of proportional scintillation, $S2$, to prompt scintillation, $S1$, was taken as a discrimination parameter [15], and the signal acceptance region was defined below the median of the nuclear recoil band, as measured with a $^{241}\mathrm{AmBe}$ neutron calibration. The electronic recoil discrimination in this signal acceptance region was determined to be larger than 99% based on calibration data from Compton-scattered $^{60}\mathrm{Co}$ gammas. Using 11.17 live days of data, no events were observed in the signal acceptance region. Thus, upper limits on the WIMP-nucleon scattering cross-section were derived, based on zero observed events. Two separate limit curves were calculated for two different parameterizations of $\mathcal{L}_{\mathrm{eff}}$ in the low-energy region where measurements are lacking or are uncertain [16].

Here, we re-analyze this data with the Profile Likelihood Ratio statistical approach [17], that is inherent to the MINOS package commonly used with the MINUIT software [18]. This method relies on the assumption that one can properly model the background via calibration measurements or Monte Carlo simulations, as opposed to the Yellin methods [10], where no information on the background is assumed. This increases the sensitivity of the method at the price of being less robust against systematic uncertainties on the background. However, by incorporating systematic uncertainties into the model, one can adjust the trade-off between sensitivity and robustness to the level desired by the experimentalist. In any case, understanding of the background is of course essential to any method that is used to quantify the sig-nificance of a discovery claim.

The particular Profile Likelihood model used here is presented in Section II. Uncertainties in $\mathcal{L}_{\mathrm{eff}}$, that can serve as an example for systematic uncertainties related to the detector, are treated in Section III, together with systematic uncertainties from the escape velocity v_{esc} as an example for astrophysical uncertainties. Calibration data for both the electronic recoil background as well as the nuclear recoil signal are taken as constraining control measurements. These constrain the full likelihood model, which is constructed in Section IV. The sensitivity of the experiment is calculated, based solely on this model, in Section V. Taking the actual measurement into account, the resulting exclusion curve is derived also in Section V. Discussion and conclusions follow in Section VI.

II. THE STATISTICAL MODEL

Events recorded by the XENON100 detector can be characterized by their prompt and proportional scintillation signals, and are therefore treated in this analysis as a set of points in the $(S1, S2)$ plane. The distribution of signal events in $S1$ can be predicted from theoretical models taking into account the detector response, as will be described below. The joint distribution in $S1$ and $S2$, for signal and background events, is estimated from the calibration measurements. The current measurements of $\mathcal{L}_{\mathrm{eff}}$ and v_{esc} are similarly treated as another set of control measurements. The statistical model describing the outcome of the measurements thus includes a number of unknown quantities, i.e. *nuisance parameters*. Those are the expected number of background events N_b given the exposure and the $(S1, S2)$ parameter range in question, a set of probabilities $\epsilon_s = \{\epsilon_s^j\}$ and $\epsilon_b = \{\epsilon_b^j\}$ describing the distribution of signal and background events in the $(S1, S2)$ plane, the relative scintillation efficiency $\mathcal{L}_{\mathrm{eff}}$, as well as the escape velocity v_{esc}. The full likelihood function $\mathcal{L}$, for a given WIMP mass m_χ and cross-section σ, is written as a product of five terms:

$$\begin{aligned}\mathcal{L} =\ & \mathcal{L}_1(\sigma, N_b, \epsilon_s, \epsilon_b, \mathcal{L}_{\mathrm{eff}}, v_{\mathrm{esc}}; m_\chi) \\ & \times \mathcal{L}_2(\epsilon_s) \times \mathcal{L}_3(\epsilon_b) \\ & \times \mathcal{L}_4(\mathcal{L}_{\mathrm{eff}}) \times \mathcal{L}_5(v_{\mathrm{esc}}).\end{aligned} \quad (1)$$

The first term $\mathcal{L}_1$ describes the main measurement of the XENON100 detector, while the following terms describe the subsidiary measurements that are used to constrain the nuisance parameters in the main likelihood term $\mathcal{L}_1$. The precise definition of these terms will be given in the following sections. Additional uncertainties can easily be incorporated by adding additional Likelihood terms, but are not relevant for the data set analyzed here.

The signal cross-section σ is the one parameter of interest. All other parameters are nuisance parameters which are *profiled out* with a profile likelihood ratio, as explained below (Equation 3). We follow the procedure

of a hypothesis test based on the profile likelihood ratio [19, 20]. This technique can be used both to exclude a WIMP with a specific mass and cross-section, or to establish the significance of a discovery.

A. Exclusion

A test statistic q_σ reduces the observed data to only one value and is constructed in order to test the signal hypothesis H_σ. It is given by

$$q_\sigma = \begin{cases} -2\ln\lambda(\sigma) & \hat{\sigma} < \sigma \\ 0 & \hat{\sigma} > \sigma \end{cases} \qquad (2)$$

where $\hat{\sigma}$ is the maximum likelihood estimator (MLE) of σ, i.e. the value of σ that maximizes the likelihood Equation 1. $\lambda(\sigma)$ is the Profile Likelihood ratio and is given by

$$\begin{aligned} \lambda(\sigma) &= \frac{\max\limits_{\sigma \text{ fixed}} \mathscr{L}\left(\sigma; \mathcal{L}_{\text{eff}}, v_{\text{esc}}, N_b, \epsilon_s, \epsilon_b\right)}{\max \mathscr{L}\left(\sigma, \mathcal{L}_{\text{eff}}, v_{\text{esc}}, N_b, \epsilon_s, \epsilon_b\right)} \\ &\equiv \frac{\mathscr{L}\left(\sigma, \hat{\hat{\mathcal{L}}}_{\text{eff}}, \hat{\hat{v}}_{\text{esc}}, \hat{\hat{N}}_b, \hat{\hat{\epsilon}}_s, \hat{\hat{\epsilon}}_b\right)}{\mathscr{L}\left(\hat{\sigma}, \hat{\mathcal{L}}_{\text{eff}}, \hat{v}_{\text{esc}}, \hat{N}_b, \hat{\epsilon}_s, \hat{\epsilon}_b\right)}. \end{aligned} \qquad (3)$$

The double-hat parameters in the numerator are the conditional MLEs of the nuisance parameters when the signal cross-section is fixed to a given value σ. The 'single-hat' parameters in the denominator are the MLEs of all parameters allowing also σ to vary. By construction, $0 \leq \lambda(\sigma) \leq 1$, hence $q_\sigma \geq 0$. q_σ equals zero when the best-fit value of the cross-section ($\hat{\sigma}$) equals the hypothesized value (σ), which corresponds to the most signal-like outcome. Larger values of the test statistic q_σ indicate that the data are less compatible with the signal hypothesis H_σ. Since we are concerned with calculating a one-sided upper bound, we only consider outcomes with $\hat{\sigma} < \sigma$ as an evidence against the signal hypothesis and set q_σ to zero otherwise.

Let $f(q_\sigma|H_\sigma)$ be the probability distribution function of the test statistic q_σ under the signal hypothesis H_σ, and let q_σ^{obs} be the value of the test statistic obtained with the observed data. The signal p-value p_s, is the probability that the outcome of a hypothetical, random XENON100 experiment results in a test statistic larger (less signal-like) than the observed one, when the signal hypothesis H_σ is *true*. Therefore, p_s given by

$$p_s = \int_{q_\sigma^{\text{obs}}}^{\infty} f(q_\sigma|H_\sigma)\, \mathrm{d}q_\sigma. \qquad (4)$$

The signal hypothesis H_σ is rejected at 90% CL if $p_s \leq 10\%$.

Downward fluctuations of the background might lead to exclusions of very small cross-sections to which the experiment is not sensitive. To protect against such an effect, p_s is modified [21, 22] to

$$p_s' = \frac{p_s}{1 - p_b} \qquad (5)$$

where

$$1 - p_b = \int_{q_\sigma^{\text{obs}}}^{\infty} f(q_\sigma|H_0)\, \mathrm{d}q_\sigma \qquad (6)$$

is the probability of the test statistic q_σ to be larger than the observed one under the background-only hypothesis $H_{\sigma=0} \equiv H_0$. Other protection procedures would also be conceivable, but this particular extension has a conservative over-coverage nature. It has been verified by Monte Carlo simulations that the coverage of our claimed 90% CL upper bound on the cross-section is in the range (92-95)% and thus larger than 90%, in accordance with the conservative nature of the method.

The upper limit $\sigma^{\text{up}}(m_\chi)$ on the cross-section σ for a given WIMP mass m_χ is found by solving

$$p_s'(\sigma = \sigma^{\text{up}}(m_\chi)) = 10\%. \qquad (7)$$

Wilks' theorem [23] states that q_σ follows a chi-square distribution in the limit of a large number of observations, which in this context not only includes the observed WIMP candidate events, but also all of the control measurements. It has been verified by Monte Carlo simulations that this asymptotic behavior indeed describes the distribution of the test statistic to a good approximation. We therefore use the chi-square approximation in order to estimate the signal p-value p_s. p_b is estimated from Monte Carlo simulations of background only events.

B. Discovery

We can also use the above statistical method in a natural way to quantify the significance of a possible additional event population or signal discovery. To this end, we test the background-only hypothesis ($\sigma = 0$), and try to reject it. Similar to Equation 2, the discovery test statistic q_0 is defined as

$$q_0 = \begin{cases} -2\ln\lambda(0) & \hat{\sigma} > 0 \\ 0 & \hat{\sigma} < 0. \end{cases} \qquad (8)$$

Setting $\sigma = 0$ in Equation 3 yields

$$\lambda(0) = \frac{\mathscr{L}\left(\sigma = 0, \hat{\hat{\mathcal{L}}}_{\text{eff}}, \hat{\hat{v}}_{\text{esc}}, \hat{\hat{N}}_b, \hat{\hat{\epsilon}}_s, \hat{\hat{\epsilon}}_b\right)}{\mathscr{L}\left(\hat{\sigma}, \hat{\mathcal{L}}_{\text{eff}}, \hat{v}_{\text{esc}}, \hat{N}_b, \hat{\epsilon}_s, \hat{\epsilon}_b\right)}. \qquad (9)$$

The p-value is defined as

$$p_0 = \int_{q^{\text{obs}}}^{\infty} f(q_0|H_0)\, \mathrm{d}q_0, \qquad (10)$$

where $f(q_0|H_0)$ is the probability density function of q_0 under the background-only hypothesis H_0. A p-value of $p_0 = 0.00135$ corresponds to a 3σ signal evidence. Following Wilks' theorem, q_0 asymptotically distributes as a chi-square distribution under the background-only hypothesis. Under Wilks' theorem, the discovery significance is given by $Z\sigma$ with $Z = \sqrt{q^{obs}}$. As a general remark, the Profile Likelihood can result in the discovery of a process that looks like the signal hypothesis. However, the interpretation of the physical origin of such a signal is of course left to scientific discourse.

III. EXPECTED SIGNAL WITH $\mathcal{L}_{\text{eff}}$ AND v_{esc} AS CONTROL MEASUREMENTS

The relative scintillation efficiency $\mathcal{L}_{\text{eff}}$ is the largest systematic uncertainty for the data analyzed here. It relates the expected number of $S1$ photoelectrons to the recoil energy E_{nr}. The recoil energy dependence of $\mathcal{L}_{\text{eff}}$ together with its uncertainty are taken from a Gaussian fit of all available direct measurements [24–30] as shown in figure 1. To extrapolate below $5\,\text{keV}_{\text{nr}}$, where there are no available measurements of $\mathcal{L}_{\text{eff}}$ yet, the 68% confidence interval for $\mathcal{L}_{\text{eff}}$ is constructed such that it follows the two different extrapolations used in [8] to facilitate the comparison of the resulting limit. We emphasize that this extrapolation follows both the new data published in [31], the theoretical expectations spelled out in [32], as well as the trend expected from simulations [33].
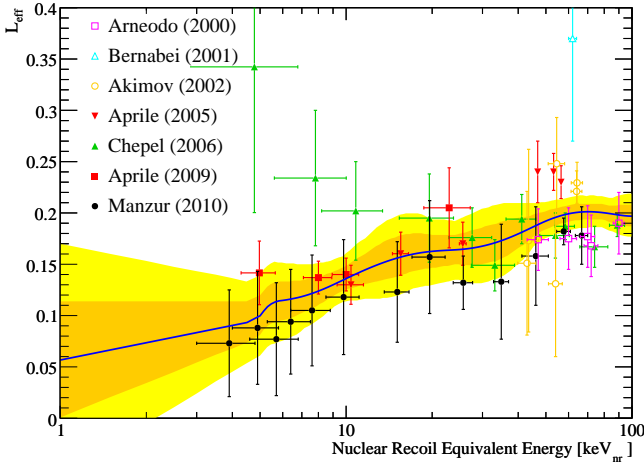

FIG. 1: All direct measurements of $\mathcal{L}_{\text{eff}}$ [24–30] together with a Gaussian fit (solid blue line) and the uncertainty band indicating the ± 1 and $\pm 2\sigma$ regions (shaded dark and light blue, respectively).

We parametrize the uncertainty of $\mathcal{L}_{\text{eff}}$ with a single nuisance parameter t, which is defined to be normally distributed with zero mean and unit variance. Thus, the $\mathcal{L}_{\text{eff}}$-term in Equation 1 corresponds simply to

$$\mathscr{L}_4(\mathcal{L}_{\text{eff}}(t)) = \exp(-(t - t_{\text{obs}})^2/2) \quad (11)$$

with $t_{\text{obs}} = 0$ representing the $\mathcal{L}_{\text{eff}}$-median in Figure 1. $t = -1$ corresponds to the lower boundary of the 68% confidence region of $\mathcal{L}_{\text{eff}}$ (dark shaded band in Figure 1), etc..

The number N_s of expected signal events is given by the integral over the predicted WIMP energy spectrum dR/dE_{nr} [34] as seen in the detector. dR/dE_{nr} is calculated for a given WIMP mass m_χ, cross-section σ, Galactic escape velocity v_{esc} as well as a given target mass and exposure time.

For the uncertainty on the escape velocity v_{esc}, we use the asymmetric probability distribution f_v as shown in Figure 7 of [35], yielding $498\,\text{km/s} < v_{\text{esc}} < 608\,\text{km/s}$ at 90% CL, with the median being $v_{\text{obs}} = 544\,\text{km/s}$. The likelihood term is defined accordingly to be

$$\mathscr{L}_5(v_{esc}) = f_v(v_{\text{obs}}|v_{\text{esc}}). \quad (12)$$

Let $\nu(E_{\text{nr}})$ be the expected number of photoelectrons (PE) for a given recoil energy E_{nr}, given by

$$\nu(E_{\text{nr}}) = E_{\text{nr}} \times \mathcal{L}_{\text{eff}}(E_{\text{nr}}) \times \frac{S_{\text{nr}}}{S_{\text{ee}}} \times L_y. \quad (13)$$

$S_{\text{ee}} = 0.58$ and $S_{\text{nr}} = 0.95$ [15] are the scintillation quenching factors due to the electric field for electronic and nuclear recoils, respectively. The normalization light yield L_y for this data sample was measured to be $L_y(122\,\text{keV}_{\text{ee}}) = (2.20 \pm 0.09)\,\text{PE/keV}_{\text{ee}}$ [8]. Errors on these parameters are small and have been neglected for the current analysis.

The signal rate in number of photoelectrons n is then given by

$$\frac{dR}{dn} = \int_0^\infty dE\, \frac{dR}{dE}(E; m_\chi, v_{\text{esc}}) \times \text{Poiss}(n|\nu(E)). \quad (14)$$

Taking into account also the finite average single-photoelectron resolution $\sigma_{\text{PMT}} = 0.5\,\text{PE}$ of the XENON100 photomultipliers, the resulting $S1$-spectrum is given by

$$\frac{dR}{dS1} = \sum_{n=1}^\infty \text{Gauss}(S1|n, \sqrt{n}\sigma_{\text{PMT}})$$
$$\times \frac{dR}{dn} \times \zeta_{\text{cuts}}(S1) \quad (15)$$

where $\zeta_{\text{cuts}}(S1)$ is the acceptance of the applied cuts [8].

The total number of expected signal events $N_s(t, v_{\text{esc}}, \sigma)$ is calculated from the integral

$$N_s(t, v_{\text{esc}}, \sigma) = \int_{S1_{\text{lower}}}^{S1_{\text{upper}}} \frac{dR}{dS1}(m_\chi)\, dS1 \quad (16)$$

where $S1_{\text{lower}} = 4\,\text{PE}$ and $S1_{\text{upper}} = 20\,\text{PE}$ is the energy interval considered in the analysis. The effect of varying t (Equation 11) for different WIMP masses is shown in Figure 2. It is calculated for an exposure of 11.17 days with 40 kg of xenon and a cross-section of $\sigma = 10^{-39}\,\text{cm}^2$.
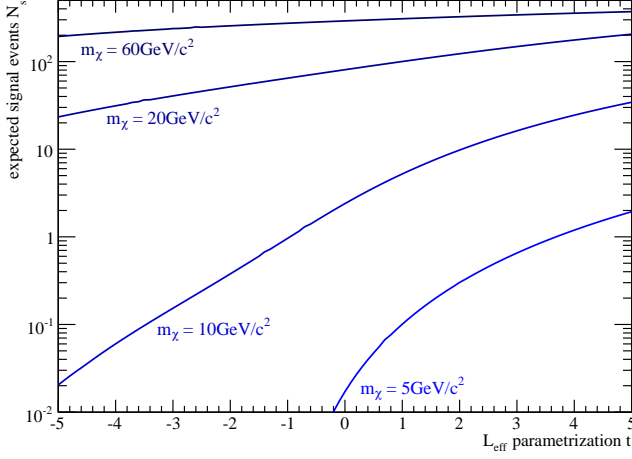
FIG. 2: Expected number of signal events N_s as a function of the $\mathcal{L}_{\text{eff}}$ parametrization t, for various WIMP masses m_χ as indicated in the figure, and for a cross-section $\sigma = 10^{-39}\,\text{cm}^2$. The lower the WIMP mass m_χ, the stronger the systematic uncertainty from $\mathcal{L}_{\text{eff}}$.

Finally, the predicted normalized WIMP spectrum $f_s(S1)$ is given by

$$f_s(S1; m_\chi) = \frac{\dfrac{\mathrm{d}R}{\mathrm{d}S1}(m_\chi)}{\displaystyle\int_{S1_{\text{lower}}}^{S1_{\text{upper}}} \dfrac{\mathrm{d}R}{\mathrm{d}S1}(m_\chi)\,\mathrm{d}S1}. \quad (17)$$

Variation of t modifies the expected number of total signal events at a given cross-section, but the shape of the $S1$ distribution is found to be hardly affected, see Figure 3. We therefore assume that f_s is independent of t, which greatly simplifies the analysis.
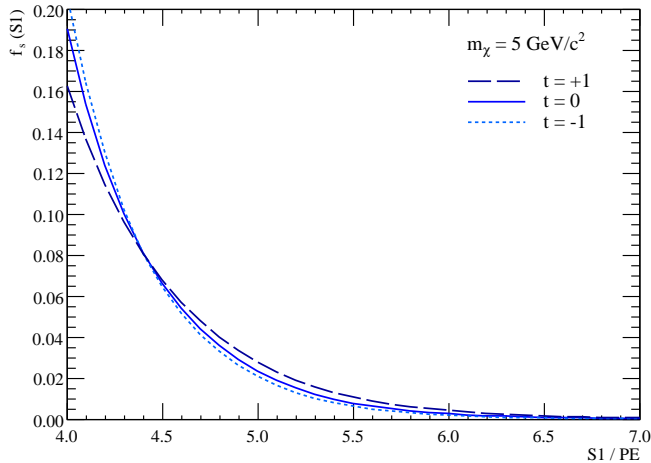

FIG. 3: The normalized WIMP recoil spectrum $f_s(S1)$ for $m_\chi = 5\,\text{GeV/c}^2$ and three possible parameterizations of $\mathcal{L}_{\text{eff}}$ through the parameter t (Equation 11). Even for low WIMP masses, the spectral shape hardly changes.

IV. LIKELIHOOD CONSTRUCTION

The total number of calibration events used for this analysis in the 4 PE-20 PE $S1$-region was $M_s = 18907$ nuclear recoils from ^{241}AmBe and $M_b = 5053$ electronic recoils from Compton-scattered ^{60}Co gammas. Those were divided into $j = 1,\ldots,(K = 23)$ bands in the $(S1, S2)$ plane, containing approximately equal fractions (about 4%) of nuclear recoils in each band, see Figure 4. ϵ_s^j and ϵ_b^j are the corresponding probabilities for a signal or background event to fall in a given band j. These bands were constructed such that the fraction of nuclear recoil events in each band is independent of $S1$, so that the signal $S1$ spectrum f_s is the same in all bands. The binning resolution of the $(S1, S2)$-plane is limited by the amount of available neutron calibration data. It was verified that the resulting limit is not sensitive to the number of bands.
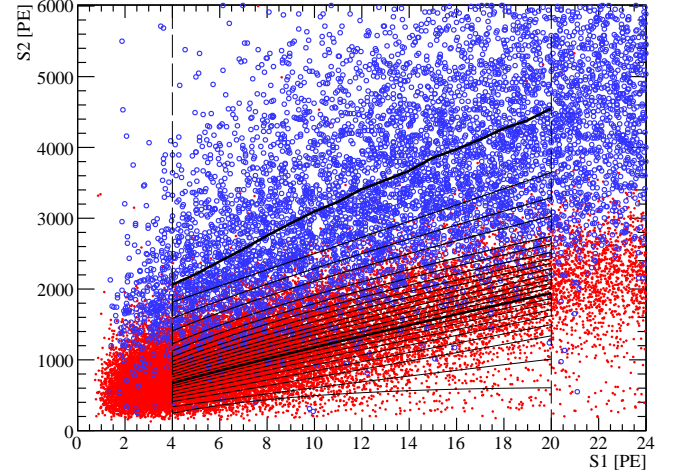

FIG. 4: Scatter plot of electronic recoils from a ^{60}Co source (blue circles) and nuclear recoils from an ^{241}AmBe source (red dots). Superimposed are the border lines of the bands j (thin lines) as well as the median of the electronic (upper thick line) and nuclear (lower thick line) recoil bands in the relevant $S1$ parameter range (dashed vertical lines).

We take the calibration measurements as control measurements of m_s^j and m_b^j events in band j, in order to constrain ϵ_s^j and ϵ_b^j. m_s^j and m_b^j are assumed to be Poisson distributed with expectations $\epsilon_s^j M_s$ and $\epsilon_b^j M_b$. The corresponding likelihood terms in Equation 1 are thus

$$\mathcal{L}_2(\epsilon_s^j) = \prod_j^K \text{Poiss}(m_s^j | \epsilon_s^j M_s), \quad (18)$$

$$\mathcal{L}_3(\epsilon_b^j) = \prod_j^K \text{Poiss}(m_b^j | \epsilon_b^j M_b). \quad (19)$$

Finally, given a set of n^j data points $(S1_i, S2_i)$ in a band j, the likelihood that they emerge from a given

WIMP spectrum is given by

$$\mathscr{L}_1 = \prod_j^K \mathrm{Poiss}(n^j | \epsilon_s^j N_s + \epsilon_b^j N_b)$$

$$\times \prod_{i=1}^{n^j} \frac{\epsilon_s^j N_s f_s(S1_i) + \epsilon_b^j N_b f_b(S1_i)}{\epsilon_s^j N_s + \epsilon_b^j N_b} \quad (20)$$

where f_s is given by Equation 17. In accordance with both observation and Monte Carlo simulations, the electronic recoil background spectrum f_b is assumed to be flat in energy [36].

Collecting all the likelihood terms together, the full likelihood function reads

$$\mathscr{L} = \prod_{j=1}^K \mathrm{Poiss}(n^j | \epsilon_s^j N_s + \epsilon_b^j N_b)$$

$$\times \prod_{i=1}^{n^j} \frac{\epsilon_s^j N_s f_s(S1) + \epsilon_b^j N_b f_b(S1)}{\epsilon_s^j N_s + \epsilon_b^j N_b}$$

$$\times \mathrm{Poiss}(m_b^j | \epsilon_b^j M_b) \times \mathrm{Poiss}(m_s^j | \epsilon_s^j M_s)$$

$$\times e^{-(t - t^{\mathrm{obs}})^2/2} \times f_v(v_{\mathrm{obs}} | v_{\mathrm{esc}}). \quad (21)$$

Differentiating $\log \mathscr{L}$ with respect to the expected number of signal events N_s and with respect to $\epsilon_s^j, \epsilon_b^j$, we find the following relations between the MLEs:

$$N \equiv \hat{N}_s + \hat{N}_b = \sum_{j=1}^K n^j \quad (22)$$

$$(N_s + M_s)\hat{\epsilon}_s^j + (N_b + M_b)\hat{\epsilon}_b^j = n^j + m_s^j + m_b^j. \quad (23)$$

These relations prove useful in the minimization of the likelihood.

V. WIMP EXCLUSION

To assess the statistical power of the experiment, the exclusion sensitivity can be calculated even before analyzing the actual data. The sensitivity of the experiment is the limit that can be expected for given exposure and experimental conditions. The test statistic q_σ is obtained by plugging the likelihood (Equation 21) into equations 2 and 3. To verify the validity of Wilks' theorem, we generate a signal for a given WIMP mass and add it to a background simulation. The resulting test statistic distribution $f(q_\sigma | H_\sigma)$ for a $5\,\mathrm{GeV/c^2}$ WIMP is shown in Figure 5 (thick red histogram). One can clearly see that the distribution is well approximated by a chi-square distribution. This allows to estimate p_s analytically.

To estimate the exclusion sensitivity, background-only experiments are simulated, based on a Poisson distribution with the expectation of 22 events as observed in [8]. The test statistic distribution $f(q_\sigma | H_0)$ is shown as the thin (blue) curve in Figure 5. The sensitivity is defined
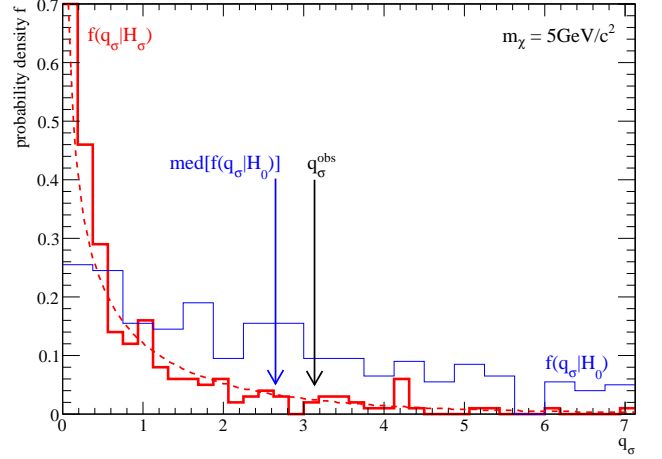

FIG. 5: The probability density functions $f(q_\sigma | H_\sigma)$ (thick red histogram) and $f(q_\sigma | H_0)$ (thin blue histogram) for a WIMP with $m_\chi = 5\,\mathrm{GeV/c^2}$ and $\sigma = \sigma^{up}$. One can clearly see the distribution is well approximated by a chi-square distribution (dashed red line). The median value of $f(q_\sigma | H_0)$ is indicated together with the test statistic q_σ^{obs} observed in data.

as the median $\mathrm{med}[q_\sigma]$ of $f(q_\sigma | H_0)$, also indicated in the figure. One can see that a large fraction of experiments under the background-only hypothesis (thin blue histogram) result in a test statistic q_σ that is similar to or even more signal-like (the area to the left of q_σ^{obs}) than that of experiments with a signal (thick red histogram). Figure 6 shows the sensitivity as its 1σ (dark shaded) and 2σ (light shaded) bands. For completeness of the statistical analysis we note that the p-value of the background-only hypothesis is 50% as there are less events than expected, $\hat{\sigma} < 0$ and $q_0 = 0$.

To derive the upper bound on the WIMP cross-section as function of WIMP mass, we solve Equation 7 for the cross-section $\sigma = \sigma^{\mathrm{up}}(m_\chi)$ that satisfies $p'_s = 10\%$ (see Section II). The test statistic q_σ^{obs} actually observed in data is shown in Figure 5, and the actual limit is shown in Figure 6. This limit has a minimum of $\sigma^{\mathrm{up}} = 2.4 \times 10^{-44}\,\mathrm{cm^2}$ at $m_\chi \sim 50\,\mathrm{GeV/c^2}$.

Two major advantages of the method presented here manifest themselves when comparing this new limit to the conventionally calculated one previously published in [8]. In the high WIMP mass region, the limit obtained here is stronger by about 20%-30%. This simply reflects the fact that in [8], only half of the available parameter space (50% nuclear recoil acceptance) was defined a priori as the signal acceptance region. In contrast, the current analysis considers the full available discrimination parameter space, also taking the background in each band j into account. Since there are no events observed in the data even in the 1σ region above the nuclear recoil median (see Figure 3 in [8]), the limit improves accordingly.

The second major difference between these two analyses appears in the low WIMP mass region, which is most
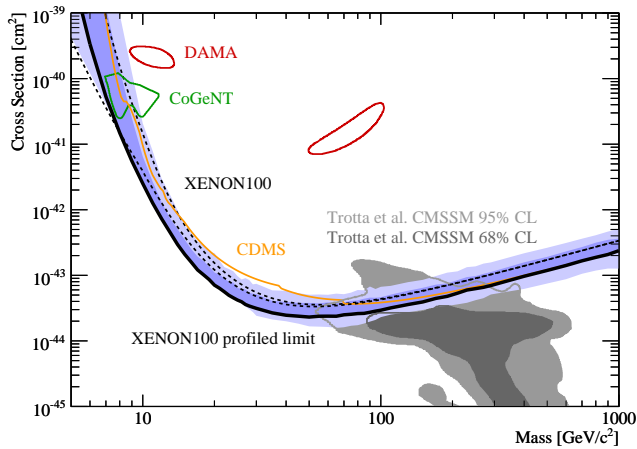
FIG. 6: Parameter space of spin-independent elastic WIMP-nucleon cross-section σ as function of WIMP mass m_χ. The sensitivity for the data set analyzed here is shown as light and dark (blue) shaded areas at 1σ and 2σ CL, respectively. The actual limit at 90% CL, taking into account all relevant systematic uncertainties as derived with the Profile Likelihood method, is shown as the thick (black) line. Two limits from the same data set, derived for two assumptions of the behavior of $\mathcal{L}_{\rm eff}$, are shown as dotted lines [8]. A limit from CDMS [4] is shown as thin (orange) line, re-calculated assuming an escape velocity of $544\,{\rm km/s}$ and $v_0 = 220\,{\rm km/s}$. Expectations from a theoretical model [37] are shown, as well as the areas (at 90% CL) favored by CoGeNT (green) [38] and DAMA (red, without channeling) [39].

sensitive to uncertainties in $\mathcal{L}_{\rm eff}$. The limit obtained here is more stringent than the limit shown in [8], constraining the areas favored by CoGeNT [38] and DAMA [39, 40]. This comes from our definition of the uncertainty bands of $\mathcal{L}_{\rm eff}$ as shown in Figure 1, where we set the 1σ interval to equal the two extrapolations used in the previous analysis. We emphasize that here we do not assume any specific behavior of $\mathcal{L}_{\rm eff}$, but rather consider it as an unknown quantity with the appropriate measurement uncertainty, as described in Section III. Such systematic uncertainty is incorporated in a natural way in the calcu-

lation of the limit and results in a single limit at 90% CL, including all considered uncertainties.

VI. CONCLUSIONS

We have introduced and applied the frequentist statistical method based on the profile likelihood test statistic to re-analyze the first data release from the XENON100 direct Dark Matter search experiment. This method avoids the need to a priori define a signal acceptance region, but instead takes all measured data into account. The background was estimated using calibration data as control measurements. Uncertainties in the relative scintillation efficiency $\mathcal{L}_{\rm eff}$ and the Galactic escape velocity $v_{\rm esc}$ were taken into account in the construction of the likelihood model. Using the profile likelihood test statistic allows to calculate the sensitivity of the experiment with its uncertainty bands, and to set a well-defined single limit, taking all systematic uncertainties into account. Applying the method to the previously published XENON100 data results in an improvement of the limit over a wide WIMP mass range, with a minimum $\sigma^{\rm up} < 2.4 \times 10^{-44}\,{\rm cm}^2$ for WIMPs with mass $m_\chi = 50\,{\rm GeV/c}^2$. In addition, this method can easily be applied for the discovery of a WIMP signal.

VII. ACKNOWLEDGMENTS

[1] N. Jarosik et al., Astrophys. J. Suppl. **192**, 14 (2011), `arXiv:1001.4744`.
[2] K. Nakamura et al. (Particle Data Group), J. Phys. **G37**, 075021 (2010).
[3] D. G. Cerdeno and A. M. Green, *Direct detection of WIMPs* (Cambridge University Press, 2010), Chapter 17 in: *Particle Dark Matter: Observations, Models and Searches*, ed. G. Bertone, `arXiv:1002.1912`.
[4] Z. Ahmed et al., Science **327**, 1619 (2010), `arXiv:0912.3592`.
[5] E. Armengaud et al., Phys. Lett. **B687**, 294 (2010), `arXiv:0912.0805`.
[6] G. Angloher et al., Astropart. Phys. **31**, 270 (2009),
`arXiv:0809.1829`.
[7] J. Angle et al., Phys. Rev. Lett. **100**, 021303 (2008), `arXiv:0706.0039`.
[8] E. Aprile et al., Phys. Rev. Lett. **105**, 131302 (2010), `arXiv:1005.0380`.
[9] W. H. Lippincott et al., Phys. Rev. **C78**, 035801 (2008), `arXiv:0801.1531`.
[10] S. Yellin, Phys. Rev. **D66**, 032005 (2002), `arXiv:physics/0203002`, See also S. Yellin (2007), *Extending the optimum interval method*, `arXiv:0709.2701`.
[11] S. Henderson, J. Monroe, and P. Fisher, Phys. Rev. **D78**, 015020 (2008), `arXiv:0801.1624`.

[12] G. J. Feldman and R. D. Cousins, Phys. Rev. **D57**, 3873 (1998), \texttt{arXiv:physics/9711021}.
[13] A. Hitachi, Astropart. Phys. **24**, 247 (2005).
[14] J. Lindhard and M. Scharff, Phys. Rev. **124**, 128 (1961).
[15] E. Aprile et al., Phys. Rev. Lett. **97**, 081302 (2006), \texttt{arXiv:astro-ph/0601552}.
[16] A. Manalaysay (2010), \texttt{arXiv:1007.3746}.
[17] W. Eadie, D. Drijard, F. James, M. Roos, and B. Sadoulet, *Statistical Methods in Experimental Physics* (North-Holland Publ. Co., Amsterdam, 1971), see in particular pp. 203–205.
[18] F. James and M. Roos, Comput. Phys. Commun. **10**, 343 (1975).
[19] W. A. Rolke, A. M. Lopez, and J. Conrad, Nucl. Instrum. Meth. **A551**, 493 (2005), \texttt{arXiv:physics/0403059}.
[20] G. Cowan, K. Cranmer, E. Gross, and O. Vitells (2010), \texttt{arXiv:1007.1727}.
[21] T. Junk, Nucl. Instrum. Meth. **A434**, 435 (1999), \texttt{arXiv:hep-ex/9902006}.
[22] A. L. Read, J. Phys. **G28**, 2693 (2002).
[23] S. S. Wilks, Ann. Math. Statist. **9**, 60 (1938).
[24] F. Arneodo et al., Nucl. Instrum. Meth. **A449**, 147 (2000).
[25] D. Akimov et al., Phys. Lett. **B524**, 245 (2002), \texttt{arXiv:hep-ex/0106042}.
[26] R. Bernabei et al., Eur. Phys. J. direct **C3**, 11 (2001).
[27] E. Aprile et al., Phys. Rev. **D72**, 072006 (2005), \texttt{arXiv:astro-ph/0503621}.
[28] V. Chepel et al., Astropart. Phys. **26**, 58 (2006).
[29] E. Aprile et al., Phys. Rev. **C79**, 045807 (2009), \texttt{arXiv:0810.0274}.
[30] A. Manzur et al., Phys. Rev. **C81**, 025808 (2010), \texttt{arXiv:0909.1063}.
[31] G. Plante et al. (2011), \texttt{arXiv:1104.2587}.
[32] F. Bezrukov, F. Kahlhoefer, and M. Lindner (2010), \texttt{arXiv:1011.3990}.
[33] M. Szydagis et al. (2011), \texttt{arXiv:1106.1613}.
[34] J. D. Lewin and P. F. Smith, Astropart. Phys. **6**, 87 (1996), see also *Detector response corrections: correction* and *Spin factors - revised tables*, available from \texttt{http://hepwww.rl.ac.uk/UKDMC/pub/publications.html}.
[35] M. C. Smith et al., Mon. Not. Roy. Astron. Soc. **379**, 755 (2007), \texttt{arXiv:astro-ph/0611671}.
[36] E. Aprile et al., Phys. Rev. **D83**, 082001 (2011), \texttt{arXiv:1101.3866}.
[37] R. Trotta, F. Feroz, M. P. Hobson, L. Roszkowski, and R. Ruiz de Austri, JHEP **12**, 024 (2008), \texttt{arXiv:0809.3792}.
[38] C. E. Aalseth et al. (2010), \texttt{arXiv:1002.4703}.
[39] C. Savage, G. Gelmini, P. Gondolo, and K. Freese, JCAP **0904**, 010 (2009), \texttt{arXiv:0808.3607}.
[40] R. Bernabei et al., Eur. Phys. J. **C67**, 39 (2010), \texttt{arXiv:1002.1028}.