Searching for low-mass dark matter particles with a massive Ge bolometer operated above ground

E. Armengaud,[1] C. Augier,[2] A. Benoît,[3] A. Benoit,[2] L. Bergé,[4] J. Billard,[2,*] A. Broniatowski,[4] P. Camus,[3] A. Cazes,[2] M. Chapellier,[4] F. Charlieux,[2] D. Ducimetière,[2] L. Dumoulin,[4] K. Eitel,[5] D. Filosofov,[6] J. Gascon,[2] A. Giuliani,[4] M. Gros,[1] M. De Jésus,[2] Y. Jin,[7] A. Juillard,[2] M. Kleifges,[8] R. Maisonobe,[2] S. Marnieros,[4] D. Misiak,[2] X.-F. Navick,[1] C. Nones,[1] E. Olivieri,[4] C. Oriol,[4] P. Pari,[9] B. Paul,[1] D. Poda,[4] E. Queguiner,[2] S. Rozov,[6] V. Sanglard,[2] B. Siebenborn,[5] L. Vagneron,[2] M. Weber,[8] E. Yakushev,[6] A. Zolotarova,[1,†]

(EDELWEISS Collaboration)

B. J. Kavanagh[10,‡]

[1]*IRFU, CEA, Université Paris-Saclay, F-91191 Gif-sur-Yvette, France*
[2]*Université Lyon, Université Lyon 1, CNRS/IN2P3, IPN-Lyon, F-69622 Villeurbanne, France*
[3]*Institut Néel, CNRS/UJF, 25 rue des Martyrs, B.P. 166, 38042 Grenoble, France*
[4]*CSNSM, Université Paris-Sud, CNRS/IN2P3, Université Paris-Saclay, 91405 Orsay, France*
[5]*Karlsruher Institut für Technologie, Institut für Kernphysik, Postfach 3640, 76021 Karlsruhe, Germany*
[6]*JINR, Laboratory of Nuclear Problems, Joliot-Curie 6, 141980 Dubna,*
Moscow Region, Russian Federation
[7]*Laboratoire de Photonique et de Nanostructures, CNRS, Route de Nozay, 91460 Marcoussis, France*
[8]*Karlsruher Institut für Technologie, Institut für Prozessdatenverarbeitung und Elektronik,*
Postfach 3640, 76021 Karlsruhe, Germany
[9]*IRAMIS, CEA, Université Paris-Saclay, F-91191 Gif-sur-Yvette, France*
[10]*GRAPPA, University of Amsterdam, Science Park 904, 1098 XH Amsterdam, Netherlands*

The EDELWEISS Collaboration has performed a search for dark matter particles with masses below the GeV scale with a 33.4-g germanium cryogenic detector operated in a surface lab. The energy deposits were measured using a neutron-transmutation-doped Ge thermal sensor with a 17.7 eV (rms) baseline heat energy resolution leading to a 60 eV analysis energy threshold. Despite a moderate lead shielding and the high-background environment, the first sub-GeV spin-independent dark matter limit based on a germanium target has been achieved. The experiment provides the most stringent, nuclear-recoil-based, above-ground limit on spin-independent interactions above 600 MeV/c^2. The experiment also provides the most stringent limits on spin-dependent interactions with protons and neutrons below 1.3 GeV/c^2. Furthermore, the dark matter search results were studied in the context of strongly interacting massive particles, taking into account Earth-shielding effects, for which new regions of the available parameter space have been excluded. Finally, the dark matter search has also been extended to interactions via the Migdal effect, resulting for the first time in the exclusion of particles with masses between 45 and 150 MeV/c^2 with spin-independent cross sections ranging from 10^{-29} to 10^{-26} cm^2.

I. INTRODUCTION

Various cosmological observations indicate that 26% of the energy density of the Universe is in the form of cold, nonbaryonic, dark matter (DM) [1]. Weakly interacting massive particles (WIMPs) are suitable cold DM candidates; they arise in extensions of the Standard Model of particle physics, such as supersymmetry, and are naturally produced in the early Universe with the correct abundance (for reviews see e.g., Refs. [2,3]). WIMPs from the Milky Way's dark matter halo can be detected directly on Earth, via the keV-scale recoils produced when they elastically scatter off nuclei [4,5]. In recent decades, significant advances have been made in the search for WIMPs in the GeV/c^2 to TeV/c^2 range that is natural for supersymmetry [6–8]. However, in light of the absence of a signal in that region there is an increasing interest in DM

*[*]j.billard@ipnl.in2p3.fr
[†]Present address: CSNSM, Université Paris-Sud, CNRS/IN2P3, Université Paris-Saclay, 91405 Orsay, France.
[‡]b.j.kavanagh@uva.nl

particles in the GeV/c^2 and sub-GeV/c^2 mass range [9–15]. These searches require experimental thresholds as low as a few tens of eV, a performance that can be attained by cryogenic detectors [16,17]. A particular advantage of such detector technology is that the thermal signal is not affected by the strong quenching effects that tend to severely reduce the amplitude of ionization or scintillation signals at low energy.

This paper describes the results obtained by the EDELWEISS Collaboration with a 33.4-g Ge detector demonstrating that such a device equipped with a neutron-transmutation-doped Ge (Ge-NTD) sensor [18] can reach the sensitivity to probe the sub-GeV domain. As a proof of the relevance of this technology, it is used in a dedicated EDELWEISS-Surf run devoted to an above-ground search for DM particles. Such a search is bound to be limited by the large background induced by cosmic-ray interactions but has the advantage of being able to probe models beyond the simple WIMP paradigm by considering relatively large values for the DM-nucleon cross sections. For sufficiently large values, each DM particle will typically interact many times in the atmosphere, Earth and shielding before reaching an underground detector. A DM particle loses energy with each interaction and may therefore arrive at the detector with insufficient energy to be observed above threshold [19–26]. Above-ground searches minimize this "stopping" effect in Earth and therefore provide good sensitivity to large DM-nucleon cross sections. While most current constraints on strongly interacting massive particles (SIMPs [27]) rely on a reanalysis of public data [30–32], we perform a dedicated search that fully takes into account the detailed detector response, presenting both the smallest and largest cross sections to which the experiment is sensitive.

We also present a search for WIMP-induced nuclear recoils which are accompanied by the ionization of an atomic electron [33–35]. The total energy in the nuclear recoil and ionization is typically larger than what can be deposited by elastic nuclear recoils at a given DM mass [36]. This "Migdal effect" therefore allows one to probe lighter DM particles (for a fixed energy threshold) at the expense of a smaller event rate. Constraints using the Migdal effect have been reported previously for NaI crystal detectors [35] and liquid xenon detectors [37,38]. Here, we present the first Migdal limit using a germanium target.

The paper is organized as follows. In Sec. II, we outline the experimental setup. In Sec. III, we give details of the dark matter search, including data processing, detector calibration and data analysis. In Sec. IV, we present limits on weakly and strongly interacting DM, for both elastic nuclear recoils and the inelastic Migdal effect. The resulting exclusion regions are summarized in Fig. 6. Finally, we conclude in Sec. V.

II. EDELWEISS-SURF

A. Detector

The detector prototype consists of a cylindrical high-purity Ge crystal of 20 mm diameter and 20 mm height,

corresponding to a total mass of 33.4 g. The thermal sensor design has been optimized for enhanced heat energy resolution. It consists of a Ge-NTD of $2 \times 2 \times 0.5$ mm^3, glued on the top surface of the crystal, weakly thermally coupled to the copper housing of the detector thanks to gold wire bonds connecting its electrodes to two gold pads on a Kapton tape. With a total Ge-NTD electrode surface of 2 mm^2 this weak thermal link is about 2.1 nW/K which is subdominant with respect to the electron-phonon coupling of 6.7 nW/K ensuring that the detector properly integrates all of the heat signal [39]. The crystal is held by six polytetrafluoroethylene (PTFE) clamps (three on each side) in order to ensure the mechanical constraints on all three axes of displacement and to minimize the stress due to PTFE elasticity at low temperatures. Unlike the usual EDELWEISS-III FID800 detectors [40], this detector proto-type has only one heat channel and no ionization readout. Therefore, a discrimination between nuclear and electron recoils is not possible. However, as there is no electric field applied across the crystal, the detector acts as a true calorimeter measuring the deposited energy of the recoiling particle independently of its type (nuclear or electronic recoil). Quenching effects on the heat energy scale for nuclear recoils in Ge cryogenic detectors have been shown to be very small [41,42] and are therefore neglected hereafter.

B. Experimental setup

The dark matter search has been performed in the dry dilution cryostat of the Institut de Physique Nucléaire de Lyon installed in a surface building with negligible over-burden; see Sec. IV B. The cryostat is a Hexadry-200 commercially available from Cryoconcept [43], which has been upgraded to reduce the vibration levels of the mixing chamber by mechanically decoupling the cold head of the pulse tube cryocooler from the dilution unit [44]. The vibrations at the detector level were further mitigated with the use of a dedicated suspended tower [45]. The latter consists in a 25-cm-long elastic pendulum, attached to the 1 K stage by a Kevlar string and a stainless steel spring with an elastic constant of 240 N/m, holding the detector tower situated below the mixing chamber at 10 mK. The detector tower is thermally anchored to an intermediate holding structure, via supple copper braids, which also hosts the connectors for the detector readout. This suspended tower design reduces detector vibrations at the sub-μg/$\sqrt{\text{Hz}}$ level, with displacements in the order of a few nanometers (rms) in all three axes, leading to substantial gains in energy resolutions as demonstrated in Ref. [45]. The cold and warm electronics are those described in Ref. [40], with a first Bi-FET preamplifier stage at 100 K and a second stage amplifier at 300 K.

The cryostat is surrounded by a 10-cm-thick cylindrical lead shield covering a solid angle of ∼70% of 4π around the detector. The materials used for the cryostat construction were not selected for low radioactivity, with the exception

of the replacement of the standard glass fiber rods used by Cryoconcept by stainless steel ones, shown to have much less radioactive contamination.

Finally, an ^{55}Fe calibration source was glued on the inner part of the detector's copper housing and facing the crystal surface opposite to the side on which is glued the Ge-NTD.

III. DARK MATTER SEARCH

The dark matter search started two weeks after the mixing chamber reached its base temperature of 10 mK. During these first two weeks, the thermal response of the detector was studied, and its working point was optimized. The best heat energy resolution achieved was 17.7 eV (rms), with the temperature of the suspended tower regulated at 17 mK and the Ge-NTD biased at 1 nA, leading to a steady state resistance of 3.4 MΩ. After the optimization period, it was decided to record six days of data in these conditions from May 22 until May 27, 2018. It was decided beforehand to blind a 24-hr-long data period started at 5 p.m. on May 26. The remaining five-days worth of data were then used to both tune the analysis procedure and selection cuts and build a data-driven background model, in order to derive dark matter constraints from a blind analysis.

A. Data processing

The data acquisition used here is the same as from the EDELWEISS-III experiment located at the Laboratoire Souterrain de Modane [40]. In order to cancel common electronic noise sources and reduce microphonics, the voltage drop across the Ge-NTD is measured differentially and the current across it is modulated from positive to negative values, following a square wave function. A modulation frequency f_s of 400 Hz was chosen as it

resulted in the best achievable signal-to-noise ratio in the experimental conditions considered here. The data are recorded continuously at this effective sampling frequency f_s, in the so-called *stream mode*, such that there is no on-line trigger, unlike the standard EDELWEISS-III data acquisition. Instead, pulse signals are identified off-line thanks to a dedicated signal processing pipeline, described in the following, which optimally filters the data based on the frequency dependence of the observed signal and noise. In order to avoid any noise structures at frequencies below the analysis range, from 1 to 200 Hz, a second-order Butterworth numerical filter with a cutoff frequency of 2 Hz has been applied to the entire data stream before any selection of noise traces and triggers. Unless otherwise stated, the remaining part of the data processing is based on this prefiltered data stream.

1. Noise PSD estimation

The noise power spectral density (PSD) of the heat channel for an hour of recorded data is recursively determined from a random sample of 1-s time traces uniformly selected throughout the entire data stream for that hour. After an initial selection, based on their rms dispersion, the removal of time traces containing pulses proceeds recursively by computing their individual frequency-based χ^2 with respect to the averaged PSD. The procedure stops once the mean χ^2, from both the prefiltered and raw data stream, from all selected time traces is compatible with the expected value of $N_s = 400$, corresponding to the number of time samples per trace. An average noise PSD is thus determined for each of the 137 hours that comprises the entire dataset. These PSDs, corrected for the 2 Hz filter gain, are overlaid in Fig. 1 (left panel). The small dispersion shows that, despite its surface
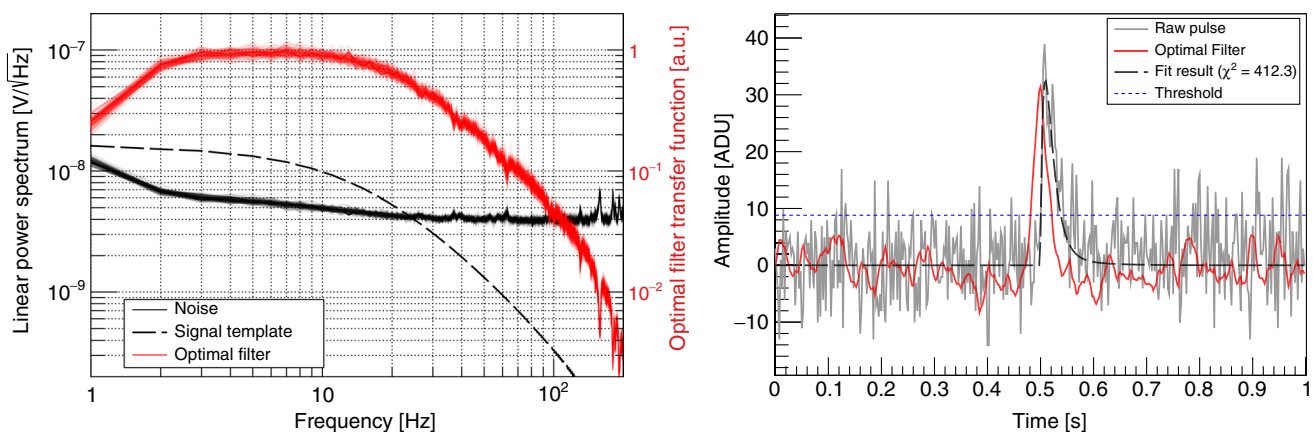

FIG. 1.   Left: Hourly averaged noise power spectral densities (PSDs) (black curves), detector signal bandwidth (black dashed line), and resulting optimal filter transfer functions (red curves) as a function of frequency, for the six days of data acquisition. The 137 separate PSDs and transfer functions are overlayed. Right: Example of a 200 eV pulse: unfiltered raw trace (gray solid line) and output of the optimal filter (red solid line). The trigger level at 3σ is shown as the blue dotted line. The result of the pulse fitting procedure, with a $\chi^2/\text{ndf} = 1.03$, is shown as the black long-dashed line.

operation, the noise is very stable over the entire 6-day-long run.

The PSD plateaus at a value of 4 nV/$\sqrt{\text{Hz}}$ above 20 Hz, with very little electromagnetic pickup contributions. This value is very well explained by the quadratic addition of the Johnson noise of the Ge-NTD (1.8 nV/$\sqrt{\text{Hz}}$) and the current noise of the Bi-FET preamplifier. The slow rise in the noise level below 20 Hz, reaching a value of 6 nV/$\sqrt{\text{Hz}}$ at 2 Hz, is due to internal thermal fluctuation noise from the detector [46].

2. Off-line trigger

The off-line trigger is based on a match filtered approach [47] designed to maximize the signal-to-noise ratio in the estimation of the signal pulse amplitude. The filter $H(f_i)$ considered hereafter is derived from the measured frequency dependence of the signal-to-noise ratio and is therefore defined as

$$H(f_i) = h \frac{s^*(f_i)}{J(f_i)} e^{-j2\pi f_i t_M}, \qquad (1)$$

where $s^*(f_i)$ is the complex conjugate of the signal template (shown as the black dashed line in the left panel of Fig. 1), $J(f_i)$ is the noise PSD in V^2/Hz (black solid lines), t_M corresponds to the time position of the pulse template maximum, and h is a normalization constant that preserves the amplitude of the pulse signal such that

$$h = \left(\sum_i \frac{|s(f_i)|^2}{J(f_i)} \right)^{-1}, \qquad (2)$$

where f_i varies between $-f_s/2$ and $+f_s/2$. An optimally matching filter is determined from each of the recorded 137 hours. Their moduli $|H(f_i)|$ are also shown as the red solid lines in the left panel of Fig. 1. As one can derive from the latter, only the lowest frequencies (below 50 Hz) are relevant to recover the observed pulse amplitude. The filter from Eq. (1) is applied to the data using the numerical procedure described in Ref. [47]. As an example, the right panel of Fig. 1 shows a 200 eV event prior to (gray solid line) and after (red solid line) filtering. Also shown is the best pulse fitting solution following the event processing procedure described in Sec. III A 3. Candidate events are identified when the filtered data exceed a given threshold level which has been defined in terms of a fixed number n of the baseline energy resolution σ_{OF}, where $\sigma_{OF} = \sqrt{h}$. The value of n was chosen such that the rate of noise-induced triggers is significantly smaller than the rate of physical events, which is about ∼1.3 Hz. The dependence on the value of n of the rate of noise induced triggers has been evaluated by simulating a 24-hour-long stream of fake data using the observed noise PSD, without injecting any signal pulses, and applying the same triggering procedure as for real data. We chose $n = 3$ as it resulted in an optimum between a low-energy threshold and a

reasonably low expected noise-induced trigger rate of 0.15 Hz. With the observed value of σ_{OF} of 18 eV, this corresponds to a trigger threshold of 55 eV shown as the blue dotted line in Fig. 1 (right panel).

The triggering procedure searches for candidate events in the filtered data stream, starting with the largest positive deviation from zero. An exclusion interval of ±0.5 s is imposed around each pulse found, and the search is reiterated in the remaining data until no fluctuation greater than $n\sigma_{OF}$ is found. This energy ordering of the pulse finding algorithm affects the energy dependence of the triggering efficiency. For instance, the dead time associated to the search for low-energy events is effectively greater than that associated to large pulses. A dedicated data-driven procedure, described in Sec. III C 2, has been developed to fully take into account these effects in the determination of the efficiency as a function of energy. The resulting trigger rate is almost constant over the six days and approximately equal to 0.58 Hz, as shown in the right panel of Fig. 2.

3. Event processing

Each recorded time trace, corresponding to either triggered events or noise samples, is further processed to estimate its amplitude by minimizing the following χ^2 function defined in the frequency domain as

$$\chi^2(a, t_0) = \sum_i \frac{|\tilde{v}(f_i) - a\tilde{s}(f_i)e^{-j2\pi t_0 f_i}|^2}{J(f_i)}, \qquad (3)$$

where $\tilde{v}(f_i)$ is the Fourier transform of the observed signal, a is the amplitude of the unit-normalized signal template s, and t_0 corresponds to the starting time of the pulse. For triggered events, the value of t_0 was allowed to vary within a 20 ms time window centered around the pulse time found by the trigger algorithm. In addition to delivering a slightly more precise estimation of the pulse amplitude a, this processing step provides a χ^2 value that quantifies the quality of the fit. This is used to reject pulses with shapes that are not consistent with the standard pulse, such as pileups and other categories of spurious events further detailed in Sec. III C 1.

B. Detector calibration and stability

1. Calibration

The energy calibration of the reconstructed amplitudes a was ensured by the use of a low-energy x-ray ^{55}Fe source irradiating the bottom side of the Ge crystal, opposite to the Ge-NTD heat sensor, inducing an interaction rate of ∼0.3 Hz. The ^{55}Fe source produces two lines corresponding to the K_α and K_β lines of Mn at 5.90 and 6.49 keV, respectively. They are clearly visible on the left panel of Fig. 2, showing the calibrated energy as a function of time over the 137-hour acquisition period. The energy resolution of these peaks is 34 eV (rms). There is sufficient statistics to calibrate each hour separately and to follow precisely any
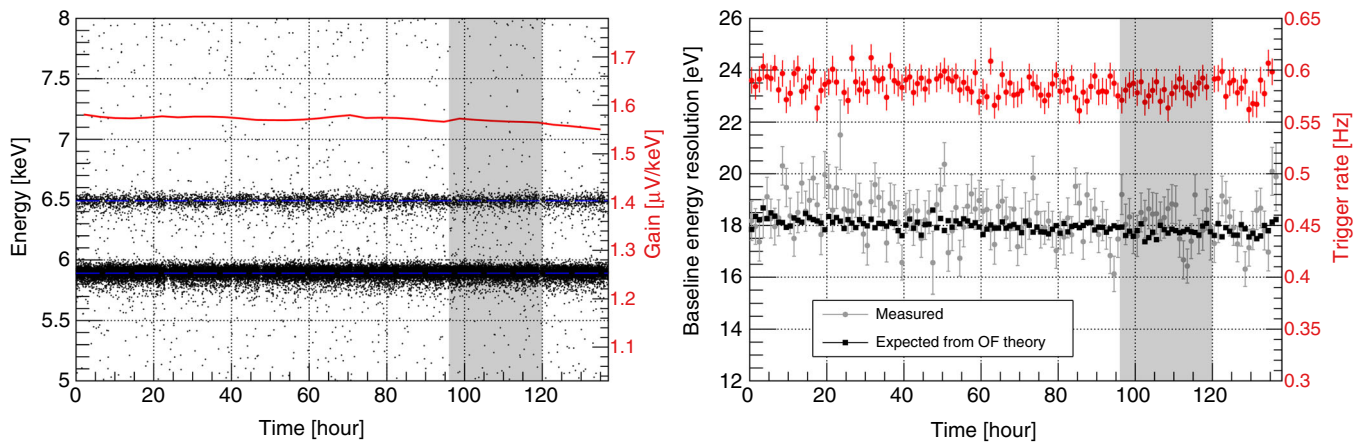
FIG. 2. Left: Event energy distribution between 5 and 8 keV as a function of time. The horizontal bands at 5.90 and 6.49 keV correspond to the K_α and K_β x-ray lines, respectively, of Mn emitted by the ^{55}Fe source. The data have been corrected for the measured time evolution of the detector gain as a function of time, shown as the red line, and corresponding to the right-hand axis. Right: Baseline heat energy resolution (rms) in eV as a function of time. The gray dots are the values derived from a fit to the energy distributions of the noise event selection, and the black squares are those derived from the ratio of the signal and noise PSDs. The corresponding trigger rates in hertz are shown as red dots. Each data point corresponds to one hour. The gray shaded region in both panels corresponds to the interval of the blinded dataset.

time evolution of the detector gain, defined as the voltage sensitivity per unit of deposited energy, and resolution. The red solid line on the same figure shows the variation of the heat sensor gain (in μ V/keV) as a function of time deduced from the peak position. The gain fluctuations are remarkably small during the entire time period, with a mean value of 1.56 μV/keV and a maximal dispersion of ∼2%.

2. Baseline resolution stability

As mentioned is Sec. III A 3, the time traces used for the PSD determination, also called noise events, were also fitted with Eq. (3). However, in this case t_0 was arbitrarily fixed to the center of the 1-s time window, in order to estimate the amplitude distribution associated with noise events. The amplitudes are distributed according to a Gaussian distribution with a standard deviation (rms) σ that corresponds to the baseline energy resolution. This quantity, measured hour by hour, is compared to the prediction of the optimal filter resolution ($\sigma_{\rm OF}$) in the right panel of Fig. 2. The good agreement between these two estimates of the energy resolution validates the assumption that the noise in each frequency bin is independent. The average baseline energy resolution is 18 eV (rms), with a ∼3% overall decrease in $\sigma_{\rm OF}$ over the six days of the search. The average value of σ during the blinded data period is 17.7 eV (rms). The stability of the noise is instrumental at maintaining an almost constant trigger rate, shown as the red dots in Fig. 2 (right panel), over the entire search period.

C. Data analysis

1. Selection cuts

Because of its above-ground operation and the moderate lead shield, the detector was exposed to a rather intense rate of high-energy interactions in the bulk of the crystal as well as in the holding PTFE clamps and the Ge-NTD heat sensor. The pulse shapes of events produced in these cases are very different. In total, the following four types of events were observed.

(i) *Normal events.*—These correspond to the vast majority of observed events for which the incoming particle has interacted in the Ge crystal. They are characterized by a rise time of 6 ms and a decay time of 16 ms, as expected from our thermal model calculations [46].

(ii) *Fast events.*—These are induced by incoming particles impinging the Ge-NTD heat sensor, which has a non-negligible volume of 2 mm^3, and from internal radioactivity of the Ge-NTD. They are characterized by a rise time of 2 ms, faster than normal events as expected, and a decay time constant of 41 ms. This longer decaying time with respect to normal events could be explained by athermal phonons heating up the Au wire bonds from the weak thermal leak. This hypothesis is further strengthened by the fact that the amplitudes of these fast events are much smaller than anticipated from our theoretical calculations [46].

(iii) *Slow events.*—These could be produced by muon interactions in the holding clamps where $\mathcal{O}(1)$ MeV energy is deposited. Despite the very weak thermal coupling of these PTFE clamps to the Ge crystal, such high-energy deposition can still produce some non-negligible rise in the temperature of the crystal. These events are characterized by a mean rise time of 26 ms and a decay time constant of 115 ms, hence much slower than normal events. This population was
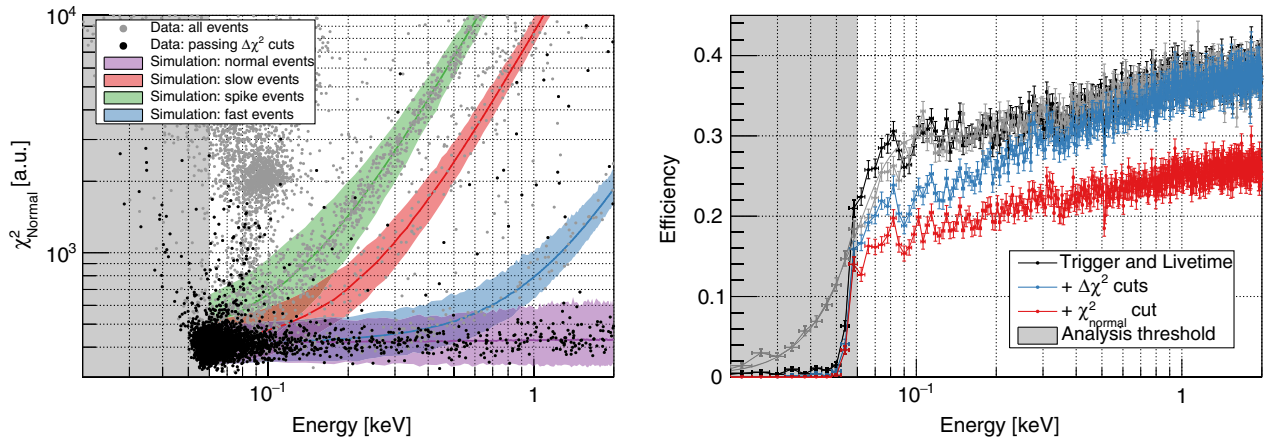
FIG. 3. Left: Distribution of χ^2_{Normal} as a function of energy for the events in the blinded dataset. The events passing all $\Delta\chi^2$ cuts are shown as black dots while those rejected are shown as gray dots. The expected region where different pulse populations are expected is shown in purple (events with normal pulse shape), blue (fast pulses), red (slow pulses) and green (single-sample spikes). Right: Signal efficiency as a function of heat energy for 10^6 simulated events injected in the data stream with energies uniformly selected between 0 and 2.5 keV, shown from 0 to 2 keV corresponding to the region of interest. Gray points: Trigger plus live-time efficiency as a function of the input energy, together with a fit with an error function (gray line). Black points and line: Trigger plus live-time efficiency as a function of the energy resulting from the χ^2 fit of a pulse to the data in the frequency domain. Blue points and line: The same, but after applying all $\Delta\chi^2$ cuts. Red points and line: The final efficiency obtained after applying the additional cut on χ^2_{Normal}. In both panels, the gray area corresponds to energies below the analysis cut.

later observed to be significantly reduced in detectors where the PTFE clamps have been replaced by a combination of sapphire balls with chrysocale sticks.

(iv) *Spike events.*—The data acquisition system used in this experiment (a former version of the one used by EDELWEISS-III [40]) suffers from random synchronization losses that result in occasional octal jumps in the data flow. These are well modeled by delta functions that can easily be discriminated from physical events.

To discriminate normal events from the other populations of events, including also pileups, cuts are performed on the goodness of fit parameter quantified by the χ^2_{Normal} value from Eq. (3), where the subscript "Normal" refers to the use of the standard pulse template s in that equation. Additional pulse-shape-related selection cuts have been designed based on the χ^2 differences $\Delta\chi^2_k = \chi^2_{Normal} - \chi^2_k$, where χ^2_k corresponds to the value calculated by replacing the standard pulse template for normal events s with the templates of the nonstandard event population discussed above; i.e., k stands for *fast*, *slow* and *spike*. All cuts were optimized on the nonblinded dataset prior to unblinding.

Figure 3 shows the event distribution in the χ^2_{Normal} vs energy plane before and after applying the $\Delta\chi^2$ cuts (black and gray points, respectively). The colored bands, estimated from the dedicated pulse simulation discussed in Sec. III C 2, correspond to the 90% C.L. confidence bands for the four different types of events. Normal events are expected to have an averaged χ^2_{Normal} value centered around $N_s = 400$ while the other event categories exhibit a quadratic divergence with increasing energy. As expected,

the population of events that is the most alike to the normal ones, and therefore the hardest to reject efficiently, are the fast events. After the combination of the various $\Delta\chi^2$ cuts and selecting events with $350 < \chi^2_{Normal} < 450$ we found that the survival probabilities of the three types of spurious events reach 10^{-2} above 60 eV for slow and spike events and above 400 eV for fast events. The additional event population observed centered at 100 eV and $\chi^2_{Normal} \sim 2000$ corresponds to noise events where the first half of the prefiltered time trace is affected by the tail of a previous pulse (mostly from ^{55}Fe), occurring a few hundreds of milliseconds before the start of its 1-s time window. Note that this particular population of pileup events, as all the other types of pileup events that appear in the gaps between the colored bands in the left panel of Fig. 3, is fully taken into account in the pulse simulation procedure described in Sec. III C 2 and is very well rejected by the standard χ^2_{Normal} cut.

2. Efficiency

The trigger efficiency was evaluated using a dedicated pulse simulation where pulses are generated at random times throughout the entire real data streams. This procedure samples rigorously all possible baseline fluctuations, including those induced by tails of high-energy events, and other nonstationary behavior, hence avoiding any possible selection-induced biases. It naturally accounts for live-time losses due to the physical event rate and comprehensively takes into account any systematic uncertainties or biases related to the

processing pipeline. The following efficiency estimates, as well as the various simulated DM signals discussed in Sec. IV, were obtained by generating a total of 10^6 simulated events distributed over 1000 Monte Carlo iterations of injecting 1000 simulated pulses in the 24-hour-long blinded data stream. This way, the simulated event rate is about 2 orders of magnitude lower than the real physical event rate, hence inducing negligible additional live-time losses. A full simulation of the DM signal is performed for each mass value in order to evaluate in each case the combined effect of the noise observed in the actual data stream, the trigger selection and the analysis cuts. To better illustrate and understand the detector performance, it is however useful to calculate an efficiency taken as the fraction of simulated events that survive all of these selection criteria, from a population of simulated events with energies uniformly distributed between 0 and 2.5 keV. These resulting efficiencies at different stages of the trigger and analysis are shown in the right panel of Fig. 3. The black and gray points correspond to the variation of the combined trigger and live-time efficiency as a function of the reconstructed and input energy, respectively. The sharp rise around 55 eV of the efficiency curves expressed as a function of the reconstructed energy is due to the fact that the amplitude estimates from the triggering algorithm and the pulse processing are very similar. The rise of the efficiency as a function of the true input energy is much smoother and is well described by an error function centered at 55 eV with

a dispersion of 18 eV, as anticipated from the observed baseline energy resolution and the considered 3σ trigger level. As discussed in Sec. III A 2, the trigger efficiency rises smoothly with energy because of the energy ordering of the algorithm. Hence, between 0.1 and 2.0 keV, the trigger and live-time efficiency rises from 30% to almost 40%. These large dead-time losses can be expected given the size of the ± 0.5 s exclusion time window compared to the observed trigger rate of 0.58 Hz.

It has been tested that the dead time could be reduced by using shorter time traces but at the cost of a slight deterioration of the energy resolution. As it was anticipated that this high-background DM search would not be limited by statistics, the emphasis was put on low thresholds and therefore energy resolution.

The right panel of Fig. 3 also shows the effect of the $\Delta\chi^2$ on the analysis efficiency. Those cuts have a significant effect on the signal below 200 eV, where all three background populations compete with the signal. Finally, the cut on the χ^2_{Normal} value has a more uniform effect as a function of energy, as shown as the red points in the same figure.

3. Observed energy spectrum

The energy spectrum recorded within the blinded day of DM search data is shown in Fig. 4. The spectrum is dominated at high energy by the calibration source lines at 5.90 and 6.49 keV. A continuous and relatively flat
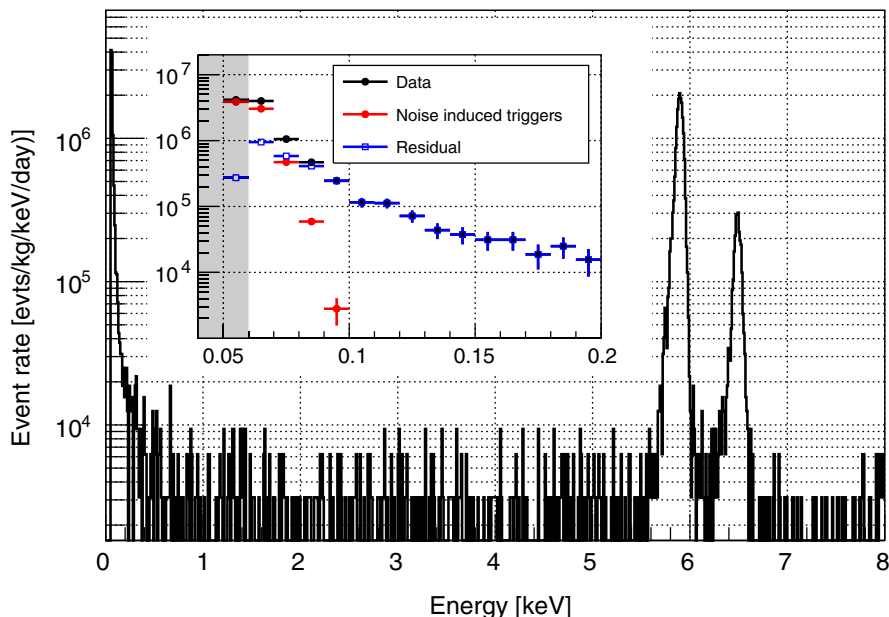

FIG. 4. Energy spectrum recorded in the blinded day of the DM search data, after all cuts. The calibration lines from the ^{55}Fe source at 5.90 and 6.49 keV are clearly visible, and they exhibit an energy dispersion of 34 eV (rms). The data are normalized in events per kilogram per day and per keV and are not corrected for efficiency. The bin size is 10 eV. The inset shows an enlargement of this distribution between 40 and 200 eV (black dots), as well as the energy spectrum observed in the simulated noise streams after all cuts (red dots). The blue squares are the difference between the two spectra.

background of 8×10^3 evt/kg/keV/day is observed between 500 eV and 8 keV [48]. The data below 500 eV, better illustrated in the inset figure, can be described approximately by an exponentially decreasing spectrum with a characteristic slope of 25 eV reaching 10^5 evt/kg/keV/day at 200 eV and culminating at 2×10^7 evt/kg/keV/day at the 60 eV analysis energy threshold. While the flat component of 8×10^3 evt/kg/keV/day corresponds to the expected gamma-ray background for a detector operated in a non-low-radioactivity cryostat with a moderate lead shield, the exponential rise at the lowest energies is not yet fully explained. Dedicated studies are ongoing to better assess their origin, such as adding electrodes to read out the ionization yield associated to these interactions and modifying the detector holders.

Finally, the inset figure also shows the contribution from the noise-induced triggers (red histogram) that has been evaluated using the noise-only simulated data stream discussed in Sec. III A 2. This event population explains very well the observed spectrum below 70 eV and it has a negligible contribution to the total observed event rate above 80 eV. It should be noted that in order to derive conservative constraints on DM interactions, it was decided prior to unblinding that this well-anticipated and -modeled background will not be subtracted.

IV. RESULTS

This section presents the limits on spin-independent interactions of DM particles with nucleons derived from the experimental spectrum of Fig. 4. In Sec. IV A the data are interpreted in the standard WIMP context where the signal is the kinetic energy of the nuclear recoil induced in the collision. In Sec. IV B we extend its interpretation to the case of SIMPs where Earth-shielding effects are taken into account. In Sec. IV C we consider the so-called Migdal effect where the DM-nucleus interaction is inelastic and simultaneously produces a nuclear and an electronic recoil, opening up access to an unexplored domain of cross sections for DM particles with masses well below 100 MeV/c^2. Finally, these results are interpreted in terms of spin-dependent interactions of DM particles in Sec. IV D.

A. Weakly interacting spin-independent dark matter

The expected signal for the standard elastic DM interactions in the detector was simulated using the comprehensive pulse simulation, described in Sec. III C 2, that takes into account all systematic effects related to the detector response and data analysis. A total of 10^6 fake DM-induced events were simulated for each DM mass considered. Their amplitudes were drawn from the theoretical distribution of induced nuclear recoils calculated assuming standard spin-independent interactions and using the standard astrophysical parameters for the DM velocity distribution [49]: a Maxwellian velocity distribution, with

an asymptotic velocity of $v_0 = 220$ km/s, and a galactic escape velocity of $v_{esc} = 544$ km/s. The local DM density at Earth's position is assumed to be $\rho_0 = 0.3$ GeV/c^2/cm^3 and the lab velocity with respect to the DM halo is $V_{lab} = 232$ km/s. Also, the loss of coherence at high momentum is taken into account using the standard Helm form factor.

The objective of the search was to establish upper limits on the DM-nucleon cross section for each considered DM particle mass, using Poisson statistics and assuming that all events observed in a given energy interval are signal candidates. Prior to unblinding the DM search data, the optimal energy intervals for such a purpose were determined by maximizing for each DM particle mass the signal-to-noise ratio between the simulated signals and the background model extracted from the 5 days of non-blinded data [50].

After the unblinding, the numbers of counts in these intervals were extracted from the data shown in Fig. 4. In the left panel of Fig. 5, the derived 90% C.L. upper limits of signals for WIMPs with masses of 0.7, 2 and 10 GeV/c^2 are compared to the experimental data. The background model, i.e., the average spectrum recorded in the five nonblinded days, is also shown.

Finally, the spin-independent WIMP-nucleon cross sections excluded at 90% C.L. as a function of the WIMP mass are shown as the solid red curve in Fig. 6 and compared with the other experimental results from Refs. [6,7,16,17, 37,38,51–55] and the so-called neutrino floor [56]. The EDELWEISS-Surf result is the most stringent, nuclear-recoil-based, above-ground limit on spin-independent interactions above 600 MeV/c^2.

B. Strongly interacting dark matter

Thanks to its above-ground operation, the present DM search can probe SIMPs that would escape detection in underground experiments as the DM particles would be stopped in the rock overburden before reaching the detectors. We therefore extend the data interpretation of Sec. IV A to determine for each mass a range of excluded cross sections that take into account the absorption of the SIMP flux in the material above the detector. Overly conservative SIMP limits can be obtained by including in the analysis only those SIMPs that reach the detector without scattering [59]. Here, more stringent limits are obtained by fully taking into account the effect of scattering on the velocity distribution of the SIMP flux reaching the detector. This velocity distribution is calculated using the publicly available VERNE code [60], introduced in Ref. [32]. It assumes continuous energy losses and straight-line trajectories for SIMPs traveling through the atmosphere, overburden and detector shielding [19]. Reference [61] showed that this simplified formalism leads to constraints similar to more complete but computationally expensive Monte Carlo simulations (e.g., [57,62–64]).
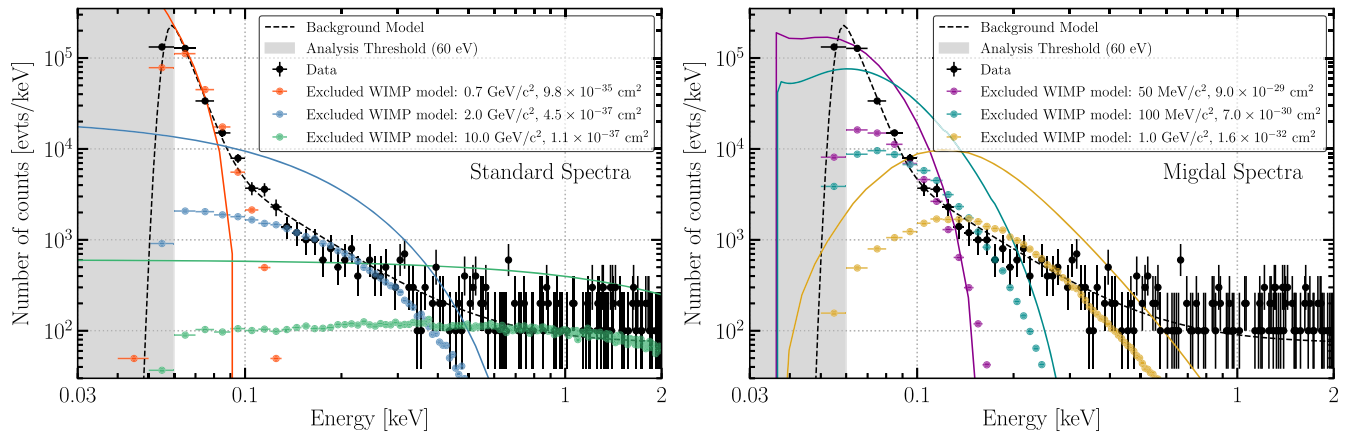
FIG. 5. Comparison of the energy spectrum after all cuts for the blinded data set (black dots) with the background model derived from the five-day nonblinded dataset (black dashed line). Also shown are the spectra for three excluded WIMP signals. The solid colored lines are the unsmeared WIMP spectra, while colored points are the spectra that incorporate resolution effects and all cuts. Left: Standard elastic nuclear recoil spectra that are excluded for WIMPs with masses of 0.7 (red), 2 (blue) and 10 GeV/c^2 (green). Right: Inelastic Migdal spectra that are excluded for WIMPs with masses of 0.05 (magenta), 0.1 (cyan) and 1.0 GeV/c^2 (yellow).

The SIMP flux calculation takes into account the variation of the direction of the mean DM flux [65] over the course of the blinded EDELWEISS-Surf exposure (24 hr, starting 17h00 on 26 May 2018). It also accounts for the atmospheric overburden above the detector as well as shielding provided by the material in the building where the detector is located and the lead, steel and copper components in its vicinity. The dominant sources of
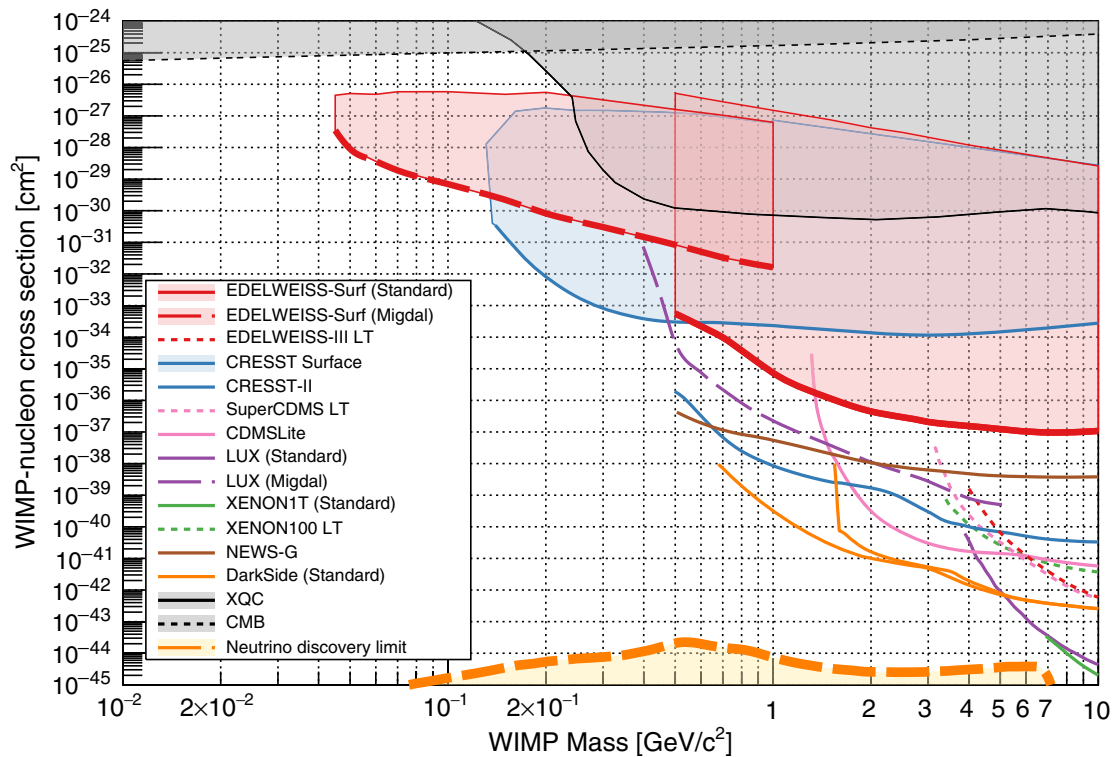

FIG. 6. The 90% C.L. limits on the cross section for spin-independent interaction between a DM particle and a nucleon as a function of the particle mass obtained in the present work. The thick solid red line corresponds to the result from the standard WIMP analysis. The associated red contour is obtained from the SIMP analysis, taking into account the slowing of the DM particle flux through the material above the detector. The thick dashed line and its accompanying red contour are obtained in the Migdal analysis. These results are compared to those of other experiments (see text). Other results using the Migdal effect are shown as dashed lines. The other shaded contours correspond to the SIMP analyses of the CRESST 2017 surface run [31,32,53] (blue contour), the XQC rocket [55,57] (gray contour with full line) and the CMB [58] (gray contour with dashed line).

stopping for downward-traveling particles are 20 cm (40 cm) of concrete from the ceiling (walls) and 10 cm of lead shielding which surrounds the detector in all directions, apart from an opening of around 50° above the detector. Upward-traveling particles are almost entirely stopped by Earth.

The escape velocity from the surface of Earth is around 11 km/s. For DM particles at such low speeds, effects such as gravitational capture [66,67] and gravitational focusing [68] may become important. These effects are not incorporated in the flux calculation. Instead, the DM velocity distribution is conservatively set to zero below $v_{cut} = 20$ km/s when calculating SIMP bounds.

Because of the very large values of cross sections involved and consequently large attenuation of the flux, the simulation of the SIMP signals corresponding to the upper bound of the excluded cross section contour requires samples many orders of magnitude larger than those required in the simple WIMP analysis of Sec. IV A. As scaling up the simulated sample size from 10^6 to $>10^{10}$ was not technically feasible for computational reasons, we developed an analytic model for the detector response based on the simulation of 10^7 events with input energies ranging from 0 to 2.5 keV (see the Appendix). This model describes the probability $P_{OF}(E_{out}|E_{in})$ of reconstructing an energy E_{out} given an initial energy E_{in} when applying the optimal filter algorithm of Sec. III A 3. The observed spectrum of events $\frac{dR}{dE_{out}}$ is thus given by

$$\frac{dR}{dE_{out}} = \eta(E_{out}) \int_0^\infty P_{OF}(E_{out}|E_{in}) \frac{dR}{dE_{in}} dE_{in}. \quad (4)$$

The measured efficiency as a function of output energy is $\eta(E_{out})$, as shown by the red curve in the right panel of Fig. 3. The calculation of P_{OF} and the comparison of the analytic detector response with results of the pulse simulations is discussed in the Appendix.

Using the signal calculated in these simulations, the same statistical procedure described in Sec. IV A is applied to derive the 90% C.L. upper bounds on the excluded cross section interval as a function of SIMP mass, resulting in the red contours shown in Fig. 6. The upper bound reported in this work improves upon the high cross section reach of the CRESST 2017 surface run [53] (thin blue), as reported in Refs. [31,32,61]. This improvement is driven in part by the longer exposure of the EDELWEISS-Surf run, which covers a full day. This includes periods when the mean direction of the DM flux (set by the Sun's velocity) is perpendicular to Earth's surface, reducing the attenuation effect of Earth and the atmosphere.

Sensitivities to 100 MeV/c²-scale SIMPs with spin-independent cross sections of the order of 10^{-30} cm² have been reported in [69]. These were derived based on the dark counts observed in a photomultiplier tube coupled to a liquid scintillator cell operated in a shallow site with 6-m

water equivalent overburden. However these resulting constraints did not take into account Earth-shielding effects, which are particularly relevant for such low-mass dark matter particles with cross section values above 10^{-31} cm², hence preventing them to be compared with the ones presented in this work.

C. Migdal search

As discussed in Sec. II A, the detector acts as a true calorimeter with equal sensitivity to the energy deposited by nuclear and electronic recoils. In this section, we consider the case where the WIMP or SIMP interaction with the target atoms induces simultaneously a nuclear recoil and the ionization of an electron. The final state comprises the two types of recoils. This is true in the case of the so-called Migdal effect, whose rate was calculated numerically in Ref. [36]. The calculations therein were performed for the case of *isolated* atoms and does not fully consider the band structure of the germanium semiconductor which is particularly important for the valence electrons ($n = 4$).

In the absence of detailed calculations, we chose to neglect the contribution from the ionization of electrons in the $n = 4$ valence shell, as well as the much smaller contribution from excitation of electrons *into* the valence shell. The inner electrons $n \leq 2$ are too tightly bound to give an appreciable signal. Therefore, the only contribution considered here is that from the ionization of M-shell ($n = 3$) electrons. Once ionized, electrons are not free (as in Ref. [36]) but instead populate the conduction band. However, the band gap in germanium is typically much smaller than the M-shell ionization energies ($\sim$0.74 eV compared to 35–170 eV, respectively) and here we will also neglect this small correction. Electrons, radiation and nuclear recoils in that energy range have very short absorption lengths in germanium, and it can be considered that the detector will completely collect the energy from all contributions.

The same standard spin-independent DM-nucleus interactions are assumed as in the previous sections, and notably the ($\propto A^2$) dependence of the cross section arising from the coherent coupling to the whole nucleus. The Earth-shielding effects are taken into account as in Sec. IV B. While the observable signal arises from inelastic Migdal events, the predominant stopping power comes from the elastic DM-nucleus collisions, and therefore only these are taken into account in the calculation of Earth-shielding effects.

The right panel of Fig. 5 shows the expected Migdal spectra for WIMPs with masses of 0.05, 0.1 and 1.0 GeV/c² corresponding to the respective interaction cross section values excluded at 90% C.L. The comparison with the standard signal in the left panel for WIMPs with considerably larger masses clearly shows how the Migdal

effect greatly enhances the sensitivity to very light WIMPs, albeit for significantly larger cross section values.

The resulting 90% C.L. exclusion region is shown in Fig. 6, bounded from below by the thick dashed red line. The constraints weaken rapidly for WIMPs with masses below 100 MeV/c^2. For lighter WIMPs, the peak in the expected Migdal signal spectrum overlaps in large part with the observed rapid rise of the data below 100 eV. Nevertheless, limits are obtained for WIMP masses as low as 45 MeV/c^2. Below this, the large cross sections required to give an observable signal lead to substantial stopping effects in the atmosphere and shielding, meaning that no constraints can be obtained anymore.

For comparison, we also show in Fig. 6 the Migdal limit obtained by the LUX Collaboration [38] (purple dashed line) but we do not show the unofficial limits based on XENON1T as presented in Ref. [37].

D. Spin-dependent interactions

In this last section we present our result in the context of a spin-dependent coupling between the dark matter particle and the nucleus. The isotope ^{73}Ge (with a natural abundance of 7.73%) is the only stable germanium isotope with nonzero nuclear spin ($J = 9/2$), hence providing sensitivity to spin-dependent (SD) WIMP-nucleus interactions. For the spin expectation values, we assume $\langle S_p \rangle = 0.031$ and $\langle S_n \rangle = 0.439$ [70], with the larger $\langle S_n \rangle$ value coming from a single unpaired neutron. For the spin-structure functions,

we take the average of the values reported in Table V of Ref. [70] which includes both one- and two-body currents.

We account for Earth-scattering effects, including only the SD scattering of the dark matter particles with nitrogen-14 in the atmosphere, which is the dominant source of stopping for surface-based SD searches [59]. Nitrogen-14 has nuclear spin $J = 1$ and we assume equal proton and neutron spin contents $\langle S_p \rangle = \langle S_n \rangle = 1$. The results, presented as 90% C.L. excluded regions and limits, are shown in Fig. 7 for DM-neutron and DM-proton SD couplings on the left and right panels, respectively. Our results, show as red lines and contours, are compared to those of other direct detection experiments shown as solid lines: LUX [71] (purple), XENON1T [72] (green), PICO-60-II [73] (brown), CDMSLite [74] (pink), and PANDAX-II [75] (blue). The other shaded contours correspond to the SIMP analyses from the X-ray Quantum Calorimeter (XQC) rocket [55,57,59] (black solid line), the Rich, Rocchia, and Spiro (RRS) balloon experiment [59], and the cosmic microwave background (CMB) [58] (gray contour with dashed line). As for the SI case (see Sec. IV B), the above-ground dark matter searches reported in [69,76] do not take into account Earth-shielding effects which are particularly relevant for cross section values above 10^{-31} cm^2 and are therefore not shown in Fig. 7.

The EDELWEISS-Surf results are exploring new parameter space of the SD couplings, on both neutrons and protons, for dark matter masses ranging from 500 MeV/c^2 to
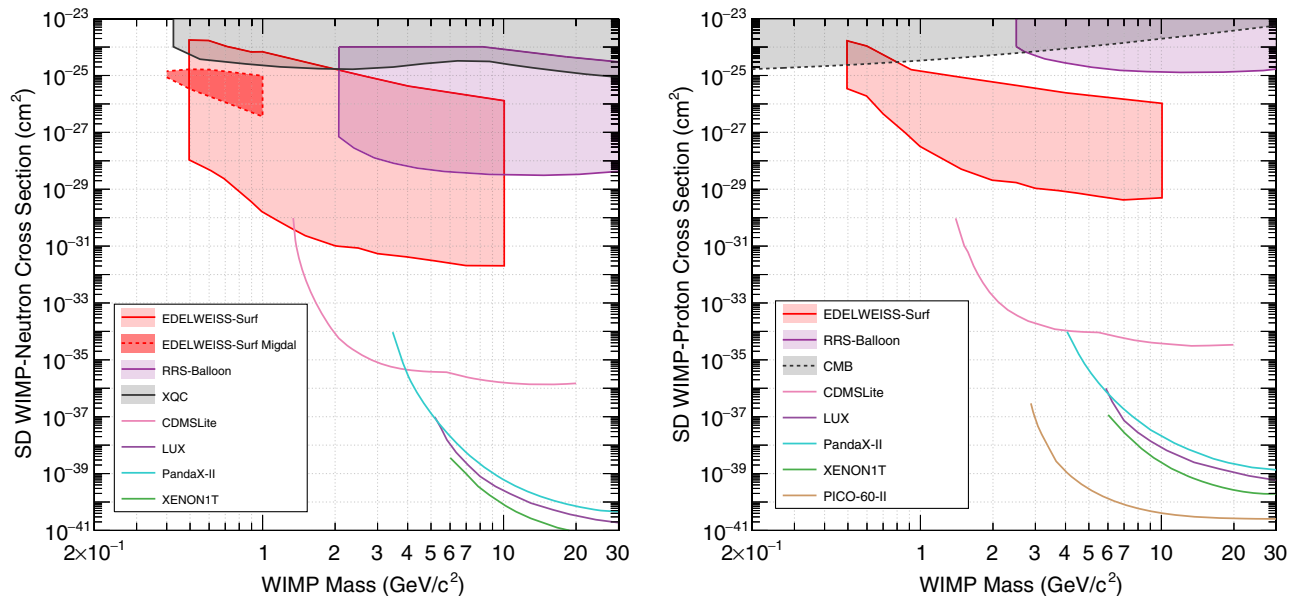

FIG. 7.   The 90% C.L. limits on the cross section for spin-dependent interactions assuming a dark matter coupling only to neutrons (left panel) and to protons (right panel) as a function of the particle mass obtained in the present work. The thick red lines and contours correspond to the presented surface dark matter search taking into account Earth-shielding effects and the so-called Migdal effect (dashed line) which is only relevant for the neutron SD coupling. These results are compared to those of other direct detection experiments shown as solid lines: LUX [71] (purple), XENON1T [72] (green), PICO-60-II [73] (brown), CDMSLite [74] (pink), and PANDAX-II [75] (blue). The other shaded contours correspond to the SIMP analyses from the XQC rocket [55,57,59] (black solid line), the RRS balloon experiment [59], and the CMB [58] (gray contour with dashed line).

1.3 GeV/c^2 considering the standard DM-induced nuclear recoil signal. The Migdal effect, producing both a nuclear recoil and an ionised electron, extends our lower DM mass bound down to 400 MeV/c^2 only in the SD-neutron case because of the dominating Earth-shielding effects happening at such high cross sections. In the case of WIMP-proton SD interactions, we obtain no limit from the Migdal effect due to the strong Earth-shielding effects at such high cross sections. Indeed, the combination of the small proton spin content of ^{73}Ge and the low ionization probability means that the scattering rate in this case is extremely small, hence requiring a large cross section to be detectable (around 10^{-24} cm^2 at 1 GeV/c^2). However, at such large cross sections, atmospheric stopping is too strong for a DM particle to reach our detector. We also note that at low masses [$\mathcal{O}(500$ MeV/c^2] both the Migdal WIMP-neutron and the standard WIMP-proton constraints are affected by Earth-stopping effects at both the upper and lower parts of the excluded regions.

V. CONCLUSION

The EDELWEISS Collaboration has searched for dark matter particles with masses between 45 MeV/c^2 and 10 GeV/c^2 using a 33.4-g germanium detector operated in a surface lab and thus relevant in the search for SIMPs. The energy deposits were measured using a Ge-NTD thermal sensor with a 17.7 eV (rms) heat energy resolution leading to a 60 eV analysis threshold. This performance, combined with the nearly completely stationary behavior of the detector, led to the achievement of the first limit for the spin-independent interaction of sub-GeV WIMPs based on a germanium target. The experiment provides the most stringent, nuclear-recoil-based, above-ground limit on spin-independent interactions above 600 MeV/c^2.

The search results were also interpreted in the context of SIMPs, taking into account the screening effect of the atmosphere and material surrounding the detector. The lower part of the excluded region for SIMPs corresponds to the previously mentioned WIMP limit and represent the most stringent constraint for masses above 600 MeV/c^2. The upper part of the excluded region is limited by Earth-shielding effects: it probes the largest SIMP-nucleon cross sections of any direct detection experiment, excluding a value of 10^{-27} cm^2 for a 1 GeV/c^2 WIMP [77]. There are a number of complementary constraints on SIMP DM, including searches for DM annihilation to neutrinos [78]; anomalous heating of Earth [67,79]; heating of Galactic gas clouds [80,81]; and DM-cosmic ray interactions [82]. These typically require additional assumptions about the properties of the DM particle, while the current constraints depend only on its large scattering cross section with nuclei and its interactions near Earth.

The dark matter search has also been extended to interactions via the Migdal effect, resulting in the exclusion for the first time of particles with masses between 45 and 150 MeV/c^2 with cross sections ranging from 10^{-29} to 10^{-26} cm^2. These limits also take fully into account the modeling of Earth-shielding effects essential for obtaining accurate constraints for such large cross section values.

Finally, interpreted in terms of spin-dependent interactions with protons or neutrons, these results exclude new regions of the parameter space below masses of 1.3 GeV/c^2. In this case, atmospheric stopping is significant and has a strong effect on the derived exclusion limits, in particular for WIMP-proton interactions and the Migdal effect.

The next steps for the EDELWEISS Collaboration are twofold. Such high-performance detectors are currently running in the low-background environment of its underground facility [40] at the Laboratoire Souterraine de Modane to better understand the origin of the observed events at low energy. Also, the detectors are equipped with electrodes in order to amplify the signal using the Neganov-Trofimov-Luke amplification [83,84], a particularly enticing prospect for searches using the electron recoils produced by the Migdal effect. The level of performance achieved in this work also opens the possibility of a first experimental measurement of the Migdal effect using a neutron calibration source.

Eventually, this level of detector performance, achieved in an above-ground laboratory with a 30 g-scale massive bolometer, is also very promising in the context of a low-energy and high-precision measurement of the coherent elastic neutrino-nucleus scattering process [85].

APPENDIX: ANALYTIC DETECTOR RESPONSE

This Appendix describes the analytic model for the detector response which is used to set limits in the SIMP and Migdal effect analyses.

As described in Sec. III A 3, the optimal filter (OF) algorithm is used to search for candidate events in the data
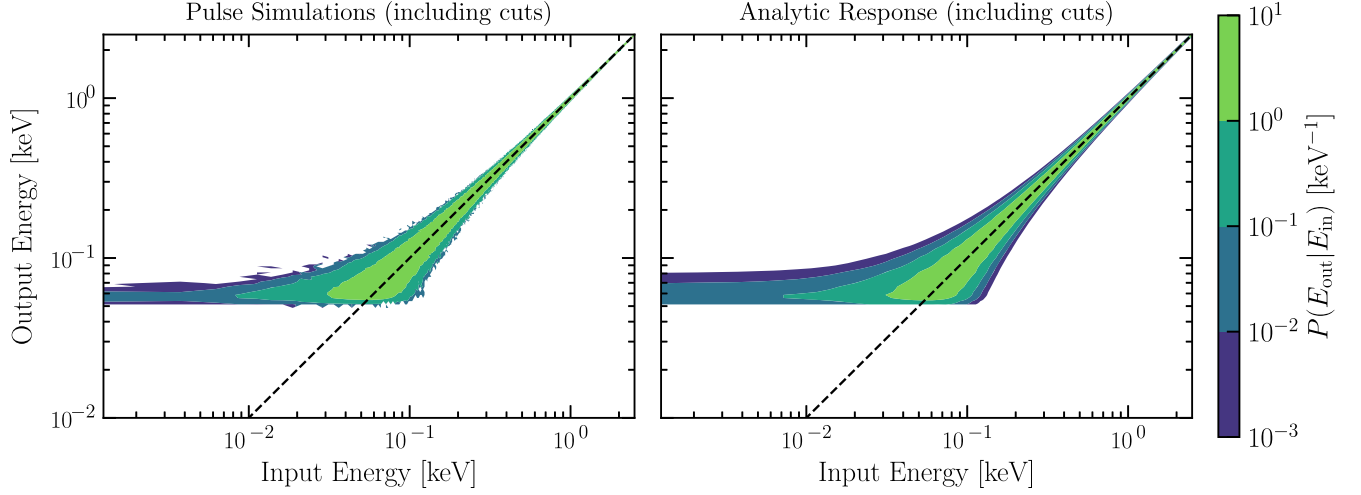
FIG. 8.   Comparison of pulse simulations and analytic detector response. Left: Density function for the output energy E_{out} at a given input energy E_{in} estimated from pulse simulations after applying cuts. Right: Analytic density function described in Eq. (A3), including optimal filter resolution and efficiency cuts. We use bins of 2.5 eV in both E_{in} and E_{out}.

stream and recover the pulse amplitude. The OF algorithm locks on the largest fluctuation away from zero in a given search window. Note that this could be a downward fluctuation, such that the reconstructed pulse amplitude may be negative. The sampling frequency is 400 Hz while the OF search window has a width of 20 ms. If each sample were independent, this would correspond to a total of eight samples. However, correlations in the noise mean that the effective number of independent samples N_s in each search window is smaller.

We now calculate the probability distribution for the largest fluctuation (i.e., deviation from zero) among N_s independent random variables X_i. We denote the probability distribution of each random variable $P_i(X_i)$ and the largest fluctuation X^*. With this, we obtain

$$P(X^*) = \sum_i^{N_s} P_i(X^*) \prod_{k \neq i} P_k(|X_k| \leq X^*). \quad \text{(A1)}$$

This is the probability that one of the random variables lies in the interval $X^* \rightarrow X^* + dX^*$, while the remaining random variables $X_{k \neq i}$ have absolute values smaller than X^*.

We assume Gaussian noise and fix the resolution to $\sigma = 18$ eV (see right panel of Fig. 2). If there are no signal events in the OF search window, then the distribution of the largest fluctuation is given by

$$P(X^*) = \frac{N_s}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{|X^*|^2}{2\sigma^2}\right) \left[\mathrm{erf}\left(\frac{|X^*|}{\sqrt{2}\sigma}\right)\right]^{(N_s-1)}. \quad \text{(A2)}$$

Identifying X^* with E_{out}, we can compare this distribution with the reconstructed energy of noise-only events, for which we find that $N_s = 3$ provides a very good fit.

If a signal event with energy E_{in} appears in one of these $N_s = 3$ independent samples, the resulting random variable

will still be Gaussian distributed with width σ but now centered on E_{in} (the remaining two random variables are Gaussian distributed with mean zero). Again, applying Eq. (A1), we obtain the probability distribution for the largest fluctuation:

$$
\begin{aligned}
&P_{\mathrm{OF}}(X^*|E_{\mathrm{in}}) \\
&= \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(X^* - E_{\mathrm{in}})^2}{2\sigma^2}\right) \left[\mathrm{erf}\left(\frac{|X^*|}{\sqrt{2}\sigma}\right)\right]^{(N_s-1)} \\
&\quad + \frac{(N_s - 1)}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{|X^*|^2}{2\sigma^2}\right) \left[\mathrm{erf}\left(\frac{|X^*|}{\sqrt{2}\sigma}\right)\right]^{(N_s-2)} \\
&\quad \times \frac{1}{2}\left[\mathrm{erf}\left(\frac{|X^*| - E_{\mathrm{in}}}{\sqrt{2}\sigma}\right) - \mathrm{erf}\left(\frac{-|X^*| - E_{\mathrm{in}}}{\sqrt{2}\sigma}\right)\right]. \quad \text{(A3)}
\end{aligned}
$$

At small input energies, this expression accounts for the fact that the largest fluctuation in the OF search window may not correspond to the signal pulse. At large input energies, we recover a standard Gaussian distribution for the output energy.

Finally, we apply the trigger efficiency $\eta(E_{\mathrm{out}})$ as a function of the reconstructed energy, given by the red curve in the right panel of Fig. 3. The final detector response is then

$$P(E_{\mathrm{out}}|E_{\mathrm{in}}) = \eta(E_{\mathrm{out}})P_{\mathrm{OF}}(E_{\mathrm{out}}|E_{\mathrm{in}}). \quad \text{(A4)}$$

The observed spectrum of events is then obtained by a convolution:

$$\frac{dR}{dE_{\mathrm{out}}} = \eta(E_{\mathrm{out}}) \int_0^\infty P_{\mathrm{OF}}(E_{\mathrm{out}}|E_{\mathrm{in}}) \frac{dR}{dE_{\mathrm{in}}} dE_{\mathrm{in}}. \quad \text{(A5)}$$

In Fig. 8, we compare the probability distribution for the output energy (at fixed input energy) estimated using pulse simulations (left panel) and using the analytic prescription

(right panel). We use more than 10^7 pulse simulations, which were performed in order to calculate the standard WIMP limits presented in Sec. IV A. We include only those simulated events passing all cuts and present the results in bins of 2.5 eV in both E_{in} and E_{out}. The analytic response function captures the main features of detector response seen in the pulse simulations.

We have also performed a more quantitative comparison. Binning the events in bins of 1 eV in input energy, we performed a standard Kolmogorov-Smirnov test [86,87] in each bin. We compared the output energy distribution of pulse simulations with the distribution expected from the analytic model. We found that the two distributions are consistent, with p values ranging from 10^{-3} to 1 over the 2000 bins as expected.

Finally, we have verified that the standard WIMP limits calculated using this analytic model are in good agreement with those obtained using the pulse simulations (as described in Sec. IV A). We find that the two methods agree at the 10% level.

[1] N. Aghanim *et al.* (Planck Collaboration), arXiv:1807.06209.

[2] G. Jungman, M. Kamionkowski, and K. Griest, Phys. Rep. **267**, 195 (1996).

[3] G. Bertone and D. Hooper, Rev. Mod. Phys. **90**, 045002 (2018).

[4] M. W. Goodman and E. Witten, Phys. Rev. D **31**, 3059 (1985).

[5] A. K. Drukier, K. Freese, and D. N. Spergel, Phys. Rev. D **33**, 3495 (1986).

[6] E. Aprile *et al.* (XENON Collaboration), Phys. Rev. Lett. **121**, 111302 (2018).

[7] D. S. Akerib *et al.* (LUX Collaboration), Phys. Rev. Lett. **118**, 021303 (2017).

[8] A. Tan *et al.* (PandaX-II Collaboration), Phys. Rev. Lett. **117**, 121303 (2016).

[9] R. Essig, J. Kaplan, P. Schuster, and N. Toro, arXiv:1004.0691.

[10] C. Cheung, J. T. Ruderman, L.-T. Wang, and I. Yavin, Phys. Rev. D **80**, 035008 (2009).

[11] D. Hooper and W. Xue, Phys. Rev. Lett. **110**, 041302 (2013).

[12] A. Falkowski, J. T. Ruderman, and T. Volansky, J. High Energy Phys. 05 (2011) 106.

[13] K. Petraki and R. R. Volkas, Int. J. Mod. Phys. A **28**, 1330028 (2013).

[14] K. M. Zurek, Phys. Rep. **537**, 91 (2014).

[15] G. Bertone and T. M. P. Tait, Nature (London) **562**, 51 (2018).

[16] G. Angloher *et al.* (CRESST Collaboration), Eur. Phys. J. C **76**, 25 (2016).

[17] R. Agnese *et al.* (SuperCDMS Collaboration), Phys. Rev. D **97**, 022002 (2018).

[18] E. E. Haller, K. M. Itoh, J. W. Beeman, W. L. Hansen, and V. I. Ozhogin, Instrum. Astron. **VIII**, 630 (1994).

[19] G. D. Starkman, A. Gould, R. Esmailzadeh, and S. Dimopoulos, Phys. Rev. D **41**, 3594 (1990).

[20] J. I. Collar and F. T. Avignone III, Phys. Lett. B **275**, 181 (1992).

[21] J. I. Collar and F. T. Avignone III, Phys. Rev. D **47**, 5238 (1993).

[22] F. Hasenbalg, D. Abriola, J. I. Collar, D. E. Di Gregorio, A. O. Gattone, H. Huck, D. Tomasi, and I. Urteaga, Phys. Rev. D **55**, 7350 (1997).

[23] C. Kouvaris and I. M. Shoemaker, Phys. Rev. D **90**, 095011 (2014).

[24] C. Kouvaris, Phys. Rev. D **93**, 035023 (2016).

[25] R. Bernabei *et al.*, Eur. Phys. J. C **75**, 239 (2015).

[26] B. J. Kavanagh, R. Catena, and C. Kouvaris, J. Cosmol. Astropart. Phys. 01 (2017) 012.

[27] We use the term "SIMP: here to refer to DM with large scattering cross sections with ordinary matter. However, we note that the term SIMP may also be used for hidden sector DM with strong self-interactions [28,29].

[28] Y. Hochberg, E. Kuflik, T. Volansky, and J. G. Wacker, Phys. Rev. Lett. **113**, 171301 (2014).

[29] Y. Hochberg, E. Kuflik, H. Murayama, T. Volansky, and J. G. Wacker, Phys. Rev. Lett. **115**, 021301 (2015).

[30] I. F. M. Albuquerque and L. Baudis, Phys. Rev. Lett. **90**, 221301 (2003); **91**, 229903(E) (2003).

[31] J. H. Davis, Phys. Rev. Lett. **119**, 211302 (2017).

[32] B. J. Kavanagh, Phys. Rev. D **97**, 123013 (2018).

[33] J. D. Vergados and H. Ejiri, Phys. Lett. B **606**, 313 (2005).

[34] C. C. Moustakidis, J. D. Vergados, and H. Ejiri, Nucl. Phys. **B727**, 406 (2005).

[35] R. Bernabei *et al.*, Int. J. Mod. Phys. A **22**, 3155 (2007).

[36] M. Ibe, W. Nakano, Y. Shoji, and K. Suzuki, J. High Energy Phys. 03 (2018) 194.

[37] M. J. Dolan, F. Kahlhoefer, and C. McCabe, Phys. Rev. Lett. **121**, 101801 (2018).

[38] D. S. Akerib *et al.* (LUX Collaboration), arXiv:1811.11241.

[39] M. Pyle, E. Feliciano-Figueroa, and B. Sadoulet, arXiv:1503.01200.

[40] E. Armengaud *et al.* (EDELWEISS Collaboration), J. Instrum. **12**, P08010 (2017).

[41] A. Benoit *et al.* (EDELWEISS Collaboration), Nucl. Instrum. Methods Phys. Res., Sect. A **577**, 558 (2007).

[42] R. Agnese *et al.* (SuperCDMS Collaboration), Appl. Phys. Lett. **113**, 092101 (2018).

[43] http://cryoconcept.com/ (accessed on 2019).

[44] E. Olivieri, J. Billard, M. De Jesus, A. Juillard, and A. Leder, Nucl. Instrum. Methods Phys. Res., Sect. A **858**, 73 (2017).

[45] R. Maisonobe, J. Billard, M. De Jesus, A. Juillard, D. Misiak, E. Olivieri, S. Sayah, and L. Vagneron, J. Instrum. **13**, T08009 (2018).

[46] D. Misiak *et al.* (to be published).

[47] S. Di Domizio, F. Orio, and M. Vignati, J. Instrum. **6,** P02007 (2011).

[48] The quoted background rates are corrected for the final efficiency as a function of energy (see red curve in Fig. 3).

[49] J. D. Lewin and P. F. Smith, Astropart. Phys. **6,** 87 (1996); C. Savage, K. Freese, and P. Gondolo, Phys. Rev. D **74,** 043531 (2006).

[50] In the presence of a signal, this procedure overestimates the background and yields conservative bounds.

[51] L. Hehn *et al.* (EDELWEISS Collaboration), Eur. Phys. J. C **76,** 548 (2016).

[52] P. Agnese *et al.* (DarkSide Collaboration), Phys. Rev. Lett. **121,** 081307 (2018).

[53] G. Angloher *et al.* (CRESST Collaboration), Eur. Phys. J. C **77,** 637 (2017).

[54] Q. Arnaud *et al.* (NEWS-G Collaboration), Astropart. Phys. **97,** 54 (2018).

[55] A. L. Erickcek, P. J. Steinhardt, D. McCammon, and P. C. McGuire, Phys. Rev. D **76,** 042007 (2007).

[56] J. Billard, L. Strigari, and E. Figueroa-Feliciano, Phys. Rev. D **89,** 023524 (2014).

[57] M. S. Mahdawi and G. R. Farrar, J. Cosmol. Astropart. Phys. 10 (2018) 007.

[58] V. Gluscevic and K. K. Boddy, Phys. Rev. Lett. **121,** 081301 (2018).

[59] D. Hooper and S. D. McDermott, Phys. Rev. D **97,** 115006 (2018).

[60] B. J. Kavanagh, verne v1.0 [computer software], Astrophysics Source Code Library, record ascl:1802.005 (archived on Zenodo).

[61] T. Emken and C. Kouvaris, Phys. Rev. D **97,** 115047 (2018).

[62] T. Emken and C. Kouvaris, J. Cosmol. Astropart. Phys. 10 (2017) 031.

[63] M. S. Mahdawi and G. R. Farrar, J. Cosmol. Astropart. Phys. 12 (2017) 004.

[64] M. S. Mahdawi and G. R. Farrar, arXiv:1712.01170.

[65] F. Mayet *et al.*, Phys. Rep. **627,** 1 (2016).

[66] A. Gould, Astrophys. J. **328,** 919 (1988).

[67] G. D. Mack, J. F. Beacom, and G. Bertone, Phys. Rev. D **76,** 043523 (2007).

[68] C. Kouvaris and N. G. Nielsen, Phys. Rev. D **92,** 075016 (2015).

[69] J. I. Collar, Phys. Rev. D **98,** 023005 (2018).

[70] P. Klos, J. Menéndez, D. Gazit, and A. Schwenk, Phys. Rev. D **88,** 083516 (2013); **89,** 029901(E) (2014).

[71] D. S. Akerib *et al.* (LUX Collaboration), Phys. Rev. Lett. **118,** 251302 (2017).

[72] E. Aprile *et al.* (XENON Collaboration), arXiv:1902.03234.

[73] C. Amole *et al.* (PICO Collaboration), arXiv:1902.04031.

[74] R. Agnese *et al.* (SuperCDMS Collaboration), Phys. Rev. D **97,** 022002 (2018).

[75] C. Fu *et al.* (PandaX-II Collaboration), Phys. Rev. Lett. **118,** 071301 (2017); **120,** 049902(E) (2018).

[76] A. H. Abdelhameed *et al.* (CRESST Collaboration), arXiv:1902.07587.

[77] Note that CDMS-II provides the strongest SIMP constraints for DM masses above around 10^8 GeV/c^2 [30,32].

[78] I. F. M. Albuquerque and C. Perez de los Heros, Phys. Rev. D **81,** 063510 (2010).

[79] G. D. Mack and A. Manohar, J. Phys. G **40,** 115202 (2013).

[80] A. Bhoonah, J. Bramante, F. Elahi, and S. Schon, Phys. Rev. Lett. **121,** 131101 (2018).

[81] A. Bhoonah, J. Bramante, F. Elahi, and S. Schon, arXiv:1812.10919.

[82] R. H. Cyburt, B. D. Fields, V. Pavlidou, and B. D. Wandelt, Phys. Rev. D **65,** 123503 (2002).

[83] B. Neganov and V. Trofimov, Otkryt. Izobret. **146,** 215 (1985).

[84] P. N. Luke, J. Appl. Phys. **64,** 6858 (1988).

[85] J. Billard *et al.*, J. Phys. G **44,** 105101 (2017).

[86] A. Kolmogorov, G. Ist. Ital. Attuari **4,** 83 (1933).

[87] N. Smirnov, Ann. Math. Stat. **19,** 279 (1948).