

Towards a robust model-independent test of the DAMA/LIBRA dark matter signal: ANAIS-112 results with six years of data

Julio Amaré^{1,2}, Jaime Apilluelo^{1,2}, Susana Cebrián^{1,2}, David Cintas^{1,2}, Iván Coarasa^{1,2,*},
Eduardo García^{1,2}, María Martínez^{1,2}, Ysrael Ortigoza^{1,2,3}, Alfonso Ortiz de Solórzano^{1,2},
Tamara Pardo^{1,2}, Jorge Puimedón^{1,2}, María Luisa Sarsa^{1,2,†} and Carmen Seoane^{1,2}

¹*Centro de Astropartículas y Física de Altas Energías (CAPA),*

Universidad de Zaragoza, Pedro Cerbuna 12, 50009 Zaragoza, Spain

²*Laboratorio Subterráneo de Canfranc, Paseo de los Ayerbe s.n., 22880 Canfranc Estación, Huesca, Spain and*

³*Escuela Universitaria Politécnica de La Almunia de Doña Godina (EUPLA),*

Universidad de Zaragoza, Calle Mayor 5, La Almunia de Doña Godina, 50100 Zaragoza, Spain

(Dated: June 17, 2025)

The nature of dark matter, which constitutes 27% of the Universe’s matter-energy content, remains one of the most challenging open questions in physics. Over the past two decades, the DAMA/LIBRA experiment has reported an annual modulation in the detection rate of ≈ 250 kg of NaI(Tl) detectors operated at the Gran Sasso Laboratory, which the collaboration interprets as evidence of the galactic dark matter detection. However, this claim has not been independently confirmed and is refuted under certain dark matter particle and halo model scenarios. Therefore, it is crucial to perform an experiment with the same target material. The ANAIS experiment uses 112.5 kg of NaI(Tl) detectors at the Canfranc Underground Laboratory and it has been collecting data since August 2017 to model-independently test the DAMA/LIBRA result. This article presents the results of the annual modulation analysis corresponding to six years of ANAIS-112 data. Our results, the most sensitive to date with the same target material, NaI(Tl), are incompatible with the DAMA/LIBRA modulation signal at a 4σ confidence level. Such a discrepancy strongly challenges the DAMA/LIBRA dark matter interpretation and highlights the need to address systematic uncertainties affecting the comparison, particularly those related to the response of detectors to nuclear recoils, which may require further characterization of the DAMA crystals.

While astronomical and cosmological evidence about dark matter (DM) accumulates, its nature remains elusive [1]. Over the past thirty years, various excesses above background have been reported in both indirect and direct dark matter searches, most of which have been resolved. However, a few of them persist and deserve further investigation [2]. One of the most puzzling results is that from the DAMA/LIBRA experiment, taking data at the Laboratori Nazionali del Gran Sasso (Italy), which reports a positive detection for more than twenty years [3, 4]. DAMA/LIBRA detectors observe an annual modulation in the detection rate at low energies, below 6 keV (in electron-equivalent energy¹), consistent with the predictions of weakly interacting massive particles (WIMPs) distributed in the galactic halo [5, 6], and having cross-sections with nucleons in the range of $10^{-40} - 10^{-42}$ cm² [3]. The statistical significance of such a detection is overwhelming, but the origin of such a modulation has not yet been determined. No other experiment has observed a compatible signal [7, 8], but the uncertainties and unknowns in both the dark matter particle and halo models make it very difficult to disprove the interpretation of the result in terms of a DM signal in a model-independent way [2]. This is the objective of ongoing or recently decommissioned experiments,

such as ANAIS-112 [9–13] and COSINE-100 [14, 15], or other upcoming projects like COSINE-100U [16], COSINUS [17], SABRE [18–20], and PICOLON [21].

The DAMA/LIBRA experiment is expected to have completed the data taking by the end of 2024. The final results should then be released. The COSINE-100 experiment was in operation from 2016 to 2023 in the YangYang Underground Laboratory in Korea. After concluding data collection, the COSINE-100 collaboration recently released results from the full dataset [22], corresponding to 6.4 years of operation and a total effective exposure of 358 kg \times yr. The ANAIS-112 experiment, in operation since August 2017 at the Canfranc Underground Laboratory (LSC) in Spain, consists of 112.5 kg of NaI(Tl), distributed in nine detectors, labeled in the following D0 to D8, and produced by the same provider as those used in the COSINE-100 experiment (Alpha Spectra Inc.). A blank module, similar in design to the other modules but without the NaI(Tl) crystal, operates in the same experimental space within an independent shielding, but integrated into the same data acquisition system, to monitor non-NaI(Tl) scintillation events. The trigger threshold is set-up at photoelectron level in each PMT and a coincidence within 200 ns between the two PMT signals of the same module is required for triggering that module. Multiple-hit events are defined as those where two or more module-triggers occur within a coincidence window of 1 μ s, while for single-hit events only one module triggers in that window. In ANAIS, the muon vetoes are readout by an independent DAQ system and they

* icoarasa@unizar.es

† mlsarsa@unizar.es

¹ In this article, we will use keV for referring to electron-equivalent energies while nuclear recoils energies will be expressed in keV_{nr}.

do not contribute neither to the definition of multiple-hit events nor to the dead time. Details on the ANAIS-112 experimental set-up, detector performance and previous results can be found elsewhere [9–13, 23]. ANAIS-112 will complete more than 8 years of data by the end of 2025, achieving a sensitivity exceeding the 5σ confidence level according to our prospects (see Figure 9 within the Supplemental Material [24]). Here, we present the analysis of the annual modulation corresponding to six years of ANAIS-112 data for a total effective exposure of 625.75 kg \times yr, which confirms our sensitivity estimates.

First, we briefly describe the key features of the analysis pipeline applied to the 6 years data, which in some aspects differs from that applied to obtain previous results [10–13].

A more robust calibration procedure for the low energy range has been developed and applied. We use a proportional calibration using the ^{109}Cd line at around 22 keV as reference to correct any possible gain drift. Later, by combining the whole exposure, we apply a recalibration procedure specifically designed for the ROI using information from two internal contaminants in the NaI(Tl) crystals, ^{22}Na and ^{40}K , both homogeneously distributed in the crystal bulk. They provide distinct peaks at 0.9 and 3.2 keV in the module where the decay occurs. These peaks are identified through coincidences with a high-energy gamma ray from the nuclear de-excitation, which can escape and be detected in a second module. The corresponding evolution of the rates for both peaks, shown in Figure 1 within the Supplemental Material [24], is fully consistent with a constant rate in the case of ^{40}K ($T_{1/2}(^{40}\text{K})=1.25\cdot 10^9$ yr) and shows an exponential decay compatible with the isotope lifetime in the case of ^{22}Na (1389 ± 51 d is the lifetime derived from the fit, while 1369 d is the nominal lifetime for ^{22}Na). This evolution of the ^{22}Na low-energy events (just below the ANAIS-112 analysis threshold, set at 1 keV) along the 6 years analyzed, guarantees both the stability of the threshold and the accuracy of the calibration in the ROI.

A new filtering procedure to remove non-bulk scintillation events is applied, following Ref. [23, 25], as done in Ref. [13] for the reanalysis of the 3-year exposure. This filtering protocol is based on Boosted Decision Trees (BDT) and the training does not use background events at all, but uses events from ^{252}Cf calibrations in the [1–2] keV energy region resulting mainly from nuclear elastic scattering from neutrons as signal and non-NaI(Tl) bulk scintillation events from the blank module in an equivalent energy range, as noise. The efficiency of this filtering procedure is derived both from ^{252}Cf and ^{109}Cd calibrations data, and their respective time evolution in the [1–6] keV energy region along the data taking is shown in the Figure 3 of the Supplemental Material [24]. By considering the deviations of the ^{109}Cd efficiency averaged for all the detectors with respect to their mean value in the [1–8] keV energy region, we obtain a standard deviation of 0.13 %, while it increases to 0.3 % by considering the detectors independently. These values are better or of

the same order that those reported for DAMA/LIBRA-phase2 efficiencies in that energy region (0.3 %) [3]. It can be observed [24] that ^{252}Cf and ^{109}Cd efficiencies are compatible. In the current analysis, the average of the efficiencies in the ROI derived from the seven ^{252}Cf calibrations will be used. However, according to the evolution of the ^{109}Cd efficiency, constant efficiencies have been assumed for all modules except D0, D4 and D5. During the first year of data collection, modules D4 and D5 showed instabilities until the operation HV was reduced. Consequently, a different (but still constant) efficiency is used in the analysis for D4 during the first 361 days and for D5 during the first 153 days. In the case of D0, one of the PMTs suffered from a gain loss in the last months of the sixth year, and then, a different (but constant) efficiency is used in the analysis for the last ≈ 200 days of that year.

This new filtering has allowed to significantly increase the efficiency for bulk scintillation event selection, but the remaining background below 3 keV is still showing an excess with respect to our background modeling estimates [23, 26]. Work is in progress in complementary directions towards understanding such an excess.

After selecting bulk scintillation events by the cut on the BDT parameter, the event rate is evaluated in one-day binning using the corresponding live time. High-rate periods which are more than 3 standard deviations over the annually averaged detection rate below 3 keV are removed (see Ref. [13] for more details)².

The time corresponding to the rejected periods is discounted from the effective live time used for calculating the effective exposure. The events arriving within 1 s from a cosmic muon triggering the veto system had been removed before applying the BDT cut and the live time had also been conveniently corrected [9].

The accumulated exposure used for the annual modulation analysis corresponding to the six years of data of the ANAIS-112 experiment is summarized in Table 1 within the Supplemental Material [24]. It also details the dead time, down time, percentage of live time rejected in the analysis, and the corresponding effective exposure after subtracting the latter.

Then, the rate is recalculated for each detector in the energy windows of interest with the corresponding live time and corrected by the average efficiency for each energy window in 0.1-keV bins, before applying the annual modulation analysis. The background after applying the filtering procedure corresponding to the 6 years data for all the nine modules in the ROI is shown in Figure 4

² The effect of this cut on the annual modulation analysis presented later in this manuscript was checked with a toy MC. By injecting a modulation with the DAMA/LIBRA amplitude over a background sampled with ANAIS-112 features and applying a cut on periods having a rate at ± 3 sigma level of the average, the bias observed in the modulation amplitude recovered after the analysis is (0.00 ± 0.02) cpd/ton/keV.

within the Supplemental Material [24], compared to the corresponding background of COSINE-100 and DAMA/LIBRA experiments.

We search for the modulation in the overall event rate over time through a least squares fit, by defining the χ^2 function as follows:

$$\chi^2 = \sum_{i,d} \frac{(n_{i,d} - \mu_{i,d})^2}{\sigma_{i,d}^2}, \quad (1)$$

where $n_{i,d}$ represents the number of events in the ROI in the time bin t_i for detector d , obtained by correcting the measured event count using the live time for that specific temporal bin and detector, along with the corresponding acceptance efficiency; $\sigma_{i,d}$ is the Poisson uncertainty associated with the event count, also corrected by the corresponding live time and efficiency, whose uncertainties are combined in quadrature; and $\mu_{i,d}$ denotes the expected number of events in that particular time bin and detector, including a hypothetical dark matter component.

$\mu_{i,d}$ is expected to diminish over time because there are background contributions from radioactive isotopes with half-lives on the order of a few years, primarily ^{210}Pb ($T_{1/2}=22.3\text{ yr}$), ^3H ($T_{1/2}=12.3\text{ yr}$) and ^{22}Na ($T_{1/2}=2.6\text{ yr}$). For detectors D6, D7 and D8, contributions from Te and I, cosmogenically produced isotopes with shorter half-lives, are also relevant. Accurately modeling this background rate decrease is crucial to avoid biasing the fit. Our background modeling is based on the independent determination of the contamination levels of the crystals and other detector components by different techniques³, followed by the Monte Carlo (MC) simulation within the Geant4 package [26]. This background modeling is used to describe the background evolution in time for every detector.

Then, we model $\mu_{i,d}$ as:

$$\mu_{i,d} = [R_{0,d}(f_d\phi_{bkg,d}^{MC}(t_i) + (1 - f_d)\phi_{flat}(t_i)) + S_m \cos(\omega(t_i - t_0))]M_d\Delta E\Delta t, \quad (2)$$

where $\phi_{bkg,d}^{MC}$ is the probability distribution function sampled from the MC model, describing the background evolution at time bin t_i for detector d ; ϕ_{flat} is a constant probability distribution function that accounts for the noise contribution not explained by the background

model (related to the excess below 3 keV previously commented) but found at a constant rate in the data, and the average component of a hypothetical contribution from DM interactions; M_d is the mass of every module; and ΔE and Δt represent energy and time intervals, respectively. $R_{0,d}$ and f_d are nuisance parameters in the fitting procedure: $R_{0,d}$ represents the background index in the considered energy region, while f_d measures the relative weight of the MC estimated time dependence with respect to the total rate. S_m represents the DM annual modulation amplitude, and ω and t_0 are the angular frequency and the phase of the modulation searched for in the data. For the standard halo model, they correspond to a period of 1 year and June, 2, respectively. S_m is set to 0 to test the null hypothesis and allowed to vary freely for the modulation hypothesis.

In the fit, the period and the phase are fixed at one year and to June 2, respectively, in order to directly compare ANAIS-112 with DAMA/LIBRA results, as they appear in Ref. [3]. We perform two independent fits: in the energy region [2–6] keV, which can be compared with the results from the total accumulated exposure of DAMA/NaI and DAMA/LIBRA, and in the [1–6] keV region, which can be compared with those of DAMA/LIBRA-phase2. Results of the fit in the [1–6] keV energy region are shown in Figure 1. Table I shows the results of the fit in the different energy regions analyzed for ANAIS-112 six-years data (this work), together with those of COSINE-100 full dataset [22], and DAMA/LIBRA [3]. It can be observed that the ANAIS-112 results for 6-year exposure are compatible with the absence of modulation within one standard deviation and incompatible with DAMA/LIBRA at 4.0 and 3.5 σ C.L. for [1–6] and [2–6] keV energy regions, respectively⁴. More information is available in the Supplemental Material [24]: fit results corresponding to the [1–6] keV, [2–6] keV energy regions with the individual p-values and χ^2/ndf (Figures 5 and 6), residuals after subtracting the non-modulated component for each detector with the DAMA/LIBRA modulation signal superimposed (Figure 8), and a table with the results obtained for all the nuisance parameters considered in the fits, Table 2).

We assess our sensitivity to the DAMA/LIBRA signal as the ratio $S_m^{\text{DAMA}}/\sigma(S_m)$, which directly gives in σ units the C.L. at which we can test the DAMA/LIBRA result. Then, the standard deviations for the modulation amplitude obtained in the best fit, $\sigma(S_m) = 2.5\text{ cpd/ton/keV}$ for both [1–6] keV and [2–6] keV, correspond to sensitivities of $(4.2\pm0.4)\sigma$ and $(4.1\pm0.3)\sigma$, respectively, where the uncertainty corresponds to the 68% C.L. DAMA/LIBRA result uncertainty. Alternatively, if we define our sensitivity as

³ Crystal and other detector components contamination levels introduced in our background modeling have been determined by using different techniques. For the most relevant crystal contaminations we used coincidence measurements (for ^{40}K , ^{22}Na and other cosmogenically produced isotopes), alpha rate determination (for ^{210}Pb - ^{210}Bi - ^{210}Po), and Bi-Po sequences (for ^{238}U and ^{232}Th chains). For other detector components, HPGe spectroscopy was used, being the PMTs the most relevant background contribution. A thorough description can be found in [26].

⁴ These values take into account the uncertainty in the DAMA/LIBRA result, unlike our previous publications. The improvement in sensitivity of our result makes more relevant the contribution from the DAMA/LIBRA uncertainty.

$S_m^{\text{DAMA}}/\sigma'(S_m)$, where $\sigma'(S_m)$ is the quadrature combination of both uncertainties, we obtain statistical significance of 3.8σ and 3.9σ for the [1–6] and [2–6] keV energy regions, respectively.

This is a highly significant result, showing strong inconsistency with the DAMA/LIBRA signal at an unprecedented level of statistical significance.

We have studied how some systematics could affect such a direct comparison between ANAIS–112 and DAMA/LIBRA results using toy MC simulations of experiments equivalent to 6 years of ANAIS–112 data, with and without adding the modulation observed by the DAMA/LIBRA experiment. Table II shows the bias of the fitting procedure estimated from this analysis for both the modulation and null hypotheses, besides the standard deviation obtained from the fits. No bias is observed and similar standard deviations than found in the ANAIS–112 six-years results, as shown in Table I, are obtained. We have also considered the effect of introducing additional fluctuation in the efficiency, as described in more detail in the End Matter section, finding negligible results in all cases.

Energy region	Exper.	Exposure (kg×y)	p-value mod hyp	S_m (cpd/ton/keV)
[1–6] keV	ANAIS	625.75	0.156	-0.4 ± 2.5
	COSINE	358.00	—	1.7 ± 2.9
	DAMA	1126.40	0.513	10.5 ± 1.1
[2–6] keV	ANAIS	625.75	0.596	1.1 ± 2.5
	COSINE	358.00	—	5.3 ± 3.1
	DAMA	2462.09	0.935	10.2 ± 0.8
(cpd/ton/3.3 keV _{nr})				
[6.7–20] keV _{nr}	ANAIS	625.75	0.566	0.0 ± 2.3
	COSINE	358.00	—	1.3 ± 2.7
	DAMA	2462.09	0.935	10.2 ± 0.8

TABLE I. Summary of the fit results (goodness of the fit and best fit value for the modulation amplitude) searching for an annual modulation with fixed phase in [1–6] and [2–6] keV energy regions for ANAIS–112 six-years data (this work), COSINE–100 full dataset [22], and DAMA/LIBRA [3]. Results in the [6.7–20] keV_{nr} sodium nuclear recoil energy region (corresponding to [2–6] keV for DAMA/LIBRA) are also shown.

Although the inconsistency between the current ANAIS–112 results and DAMA/LIBRA is indisputable in the case of DM particles releasing the energy by electron recoils (ER) in the NaI(Tl) crystal, for DM particles producing nuclear recoils (NR) there is a relevant source of systematic uncertainty affecting the comparison between both experiments. NR in NaI(Tl) result in a much lower light yield than ER releasing the same energy in the material. This is estimated by the relative scintillation efficiency factor or quenching factor (QF), the amount of light produced by a NR with respect to that produced by an ER. Sodium and iodine

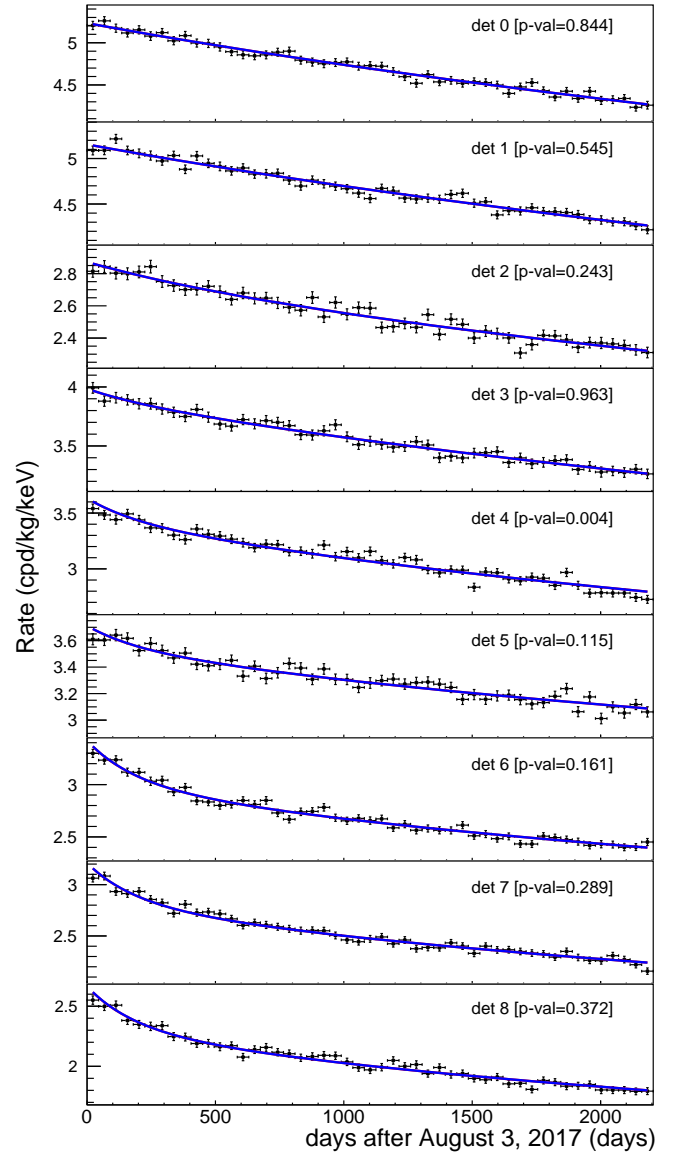


FIG. 1. Results of the fit for the data from the nine modules in the [1–6] keV energy region, under the modulation (blue) and null hypotheses (red). In all the panels, the red line is masked by the blue one, as the fit obtained for the modulated hypothesis is consistent with $S_m = 0$. p-values under the modulation hypothesis are individually displayed for each module. The global results of the fit are: for the null hypothesis, $\chi^2/\text{ndf} = 451.34/423$, (p-value = 0.164), and for the modulation hypothesis, $\chi^2/\text{ndf} = 451.31/422$, (p-value = 0.156). The best-fit modulation amplitude in the latter case is $S_m = (-0.4 \pm 2.5)$ cpd/ton/keV.

scintillation QF have been measured by several authors (see Ref. [27] and references therein) and results do not fully agree. This hints either at a dependence on the particular crystal properties (impurities, defects, growth method, etc.) or to systematics in the calculation of the factors (as found in Ref. [27]). In particular, QF values measured by DAMA/LIBRA, both for sodium and

Energy region	Bias [null hyp] (cpd/ton/keV)	$\sigma(S_m)$ [null hyp] (cpd/ton/keV)	Bias [DAMA S_m] (cpd/ton/keV)	$\sigma(S_m)$ [DAMA S_m] (cpd/ton/keV)
[1–6] keV	0.01±0.02	2.34±0.01	0.00±0.02	2.32±0.01
[2–6] keV	0.00±0.02	2.49±0.01	0.01±0.02	2.50±0.01
	(cpd/ton/3.3 keV _{nr})	(cpd/ton/3.3 keV _{nr})	(cpd/ton/3.3 keV _{nr})	(cpd/ton/3.3 keV _{nr})
[6.7–20] keV _{nr}	0.00±0.02	2.23±0.01	0.01±0.02	2.25±0.01

TABLE II. Bias (true value - fitted value) of the fitting procedure derived from 20 000 MC simulations assuming no modulation present (second column) and DAMA/LIBRA observed modulation (fourth column). The third and fifth columns are the standard deviations of the distribution of the best fit modulation amplitudes obtained from the MC for both hypotheses.

iodine NR, are higher than most of the recent measurements, which, in contrast to DAMA/LIBRA, also point to an energy dependence. Further work to improve the understanding and modeling of the scintillation QF is required to solve this issue, but we can analyse the effect of introducing different QF values in the comparison of the results. If we consider the values published by the DAMA/LIBRA collaboration for their crystals ($QF_{Na}=0.3$ and $QF_I=0.09$ [28]) and the values obtained for crystals grown in the same batch than ANAIS-112 crystals assuming constant-with-the-energy QF values ($QF_{Na}=0.210\pm0.003$ and $QF_I=0.066\pm0.022$ [27]), it can be noticed that both fulfill a $\approx 3/2$ proportionality, allowing to convert the [2–6] keV energy region from DAMA/LIBRA into a nuclear recoil energy range of [6.7–20] keV_{nr} for sodium recoils and [22.2–66.7] keV_{nr} for iodine recoils, which can be directly compared with the corresponding region in ANAIS-112 data ([1.3–4.0] keV). We have repeated all of the previous fitting procedures in this energy region. Results are again consistent with the absence of modulation and incompatible with DAMA/LIBRA at 4.2σ C.L., as shown in Table I and in Figure 7 within the Supplemental Material [24].

Although the impact of systematics related to the uncertainty in the QF for NaI(Tl) in the ROI for the DM analysis has not been fully estimated, analyzing the dependence of the modulation amplitude with the energy derived from the different experiments could help to confirm or rule out the presence of a modulation. We have analyzed the dependence of the modulation amplitude in 1-keV energy bins, from 1 to 20 keV, both for single-hit and multiple-hit events, combining the data from the nine modules. Results are shown in Figure 2 besides those of DAMA/LIBRA-phase2 [3], as well as 1σ , 2σ and 3σ bands derived from sensitivity estimates from ANAIS-112 data [29]. The ANAIS results are compatible with the absence of modulation in all of the energy bins.

According to the distribution of the modulation amplitude values, $\chi^2/\text{ndf} = 24.03/5$ (p-value = 2.14×10^{-4}) is obtained for the DAMA/LIBRA hypothesis, taking into account DAMA/LIBRA result uncertainty, and $\chi^2/\text{ndf} = 3.14/5$ (p-value = 0.678) for the null hypothesis, both corresponding to single-hit events in the [1–6] keV energy region. Similar values are obtained in

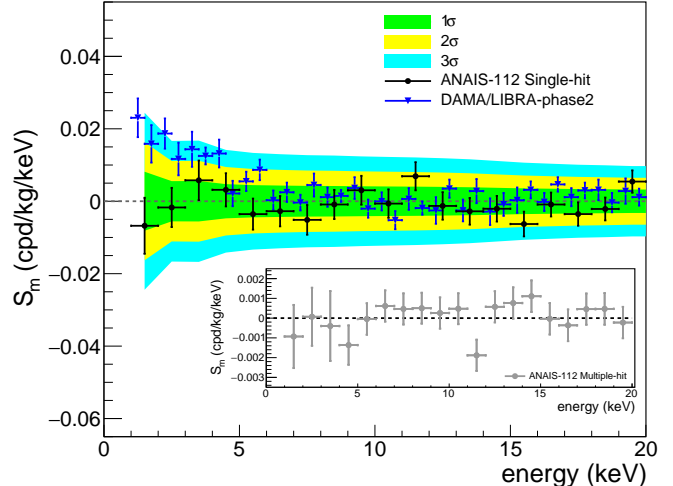


FIG. 2. Modulation amplitude per 1 keV energy bins for single-hit events (black dots) compared with the corresponding DAMA/LIBRA-phase2 result [3] (blue triangles). The 1σ , 2σ and 3σ ANAIS-112 bands derived from sensitivity estimates are also shown. Inset: The same for multiple-hit events.

the [2–6] keV energy region. For events with multiplicity 2, we obtain $\chi^2/\text{ndf} = 2.26/5$ (p-value = 0.812) in the [1–6] keV energy region, in full agreement with the results for single-hits.

Figure 3 shows the modulation amplitude on the scale of sodium nuclear recoils, for 3.3 keV_{nr} energy bins, for single-hit events, combining the data from the nine ANAIS modules (black dots). DAMA/LIBRA-phase2 (blue triangles) and COSINE-100 data (gray diamonds) are also shown for comparison. This figure assumes constant QF for sodium: 0.2 in the case of ANAIS-112 detectors and 0.3 in the case of DAMA/LIBRA. Because the relation between QF for sodium and iodine in DAMA/LIBRA and ANAIS-112 is the same, the energy scale of this figure can be directly converted into iodine nuclear recoils for both datasets, but not for COSINE-100 data. The ANAIS modulation amplitudes derived within this assumption in the [6.7–20] keV_{nr} correspond to $\chi^2/\text{ndf} = 4.22/4$ (p-value = 0.376) for the null hypothesis, and $\chi^2/\text{ndf} = 22.98/4$ (p-value = 1.28×10^{-4}) for the DAMA/LIBRA modulation hypothesis, support-

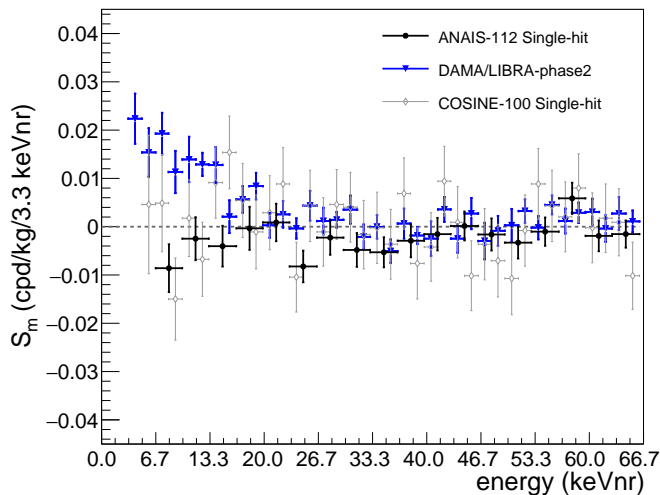


FIG. 3. Comparison of the modulation amplitude for single-hit events in the energy scale of sodium nuclear recoils from ANAIS-112 (black dots) and COSINE-100 data [22] (gray diamonds). The DAMA/LIBRA-phase2 result [3] is also shown as blue triangles.

ing again with high statistical significance the incompatibility between both experimental results. It is worth highlighting that Figure 3 is not a simple re-scaling of the energy axis from Figure 2.

Summarizing, 6-year exposure ANAIS-112 results are incompatible with DAMA/LIBRA at 4.0 and 3.5 σ C.L. for [1–6] and [2–6] keV energy regions and at 4.2 σ C.L. in the [6.7–20] ([22.2–66.7]) keV_{nr} region assuming a constant QF for sodium (iodine), while compatible with full dataset COSINE-100 results, as shown in Figure 6. According to our sensitivity prospects (see Figure 9 within the Supplemental Material), the scheduled ANAIS-112 full dataset (completed by the end of 2025) will be able to provide a robust refutation of the DAMA/LIBRA signal at 5 σ level for ER and for NR, even assuming different scintillation QF for DAMA/LIBRA and ANAIS-112 crystals [27, 28]. However, it is worth to remark that the comparison in terms of NR is affected by relevant systematics related with the uncertainties in the knowl-

edge of the scintillation quenching factors for sodium and iodine recoils. The non-proportionality of the NaI(Tl) light yield in the ROI relevant for the annual modulation analysis may partially contribute to these uncertainties. Therefore, a better modeling of both, the quenching factors and the non-proportionality of NaI(Tl) is needed. We are working to improve the estimates of the QF in ANAIS-112 crystals using onsite ^{252}Cf calibrations, but a better evaluation of the QF for DAMA/LIBRA crystals is also necessary.

ACKNOWLEDGEMENTS

This work has been financially supported by MCIN/AEI/10.13039/501100011033 under grant PID2022-138357NB-C21 and PID2019-104374GB-I00, the Consolider-Ingenio 2010 Programme under grants MultiDark CSD2009-00064 and CPAN CSD2007-00042, the LSC Consortium, the Gobierno de Aragón and the European Social Fund (Group in Nuclear and Astroparticle Physics) and funds from European Union NextGenerationEU/PRTR (Planes complementarios, Programa de Astrofísica y Física de Altas Energías). Authors would like to acknowledge the use of the Servicio General de Apoyo a la Investigación-SAI, Universidad de Zaragoza and technical support from LSC and GIFNA staff.

DATA AND CODE AVAILABILITY

The data and scripts used for preparing the figures and reproducing the results presented in this article will be available at the website of the ORIGINS Excellence Cluster: <https://www.origins-cluster.de/odsl/dark-matter-data-center/available-datasets/anais>.

REFERENCES

- [1] S. Navas et al. (Particle Data Group), Phys. Rev. D **110**, 030001 (2024).
- [2] R. K. Leane et al., *Snowmass2021 Cosmic Frontier White Paper: Puzzling Excesses in Dark Matter Searches and How to Resolve Them* (2022), 2203.06859.
- [3] R. Bernabei et al., Prog. Part. Nucl. Phys. **114**, 103810 (2020).
- [4] R. Bernabei et al., Nucl. Phys. Atom. Energy **22**, 329 (2021).
- [5] A. K. Drukier, K. Freese, and D. N. Spergel, Phys. Rev. D **33**, 3495 (1986).
- [6] K. Freese, J. A. Frieman, and A. Gould, Phys. Rev. D **37**, 3388 (1988).
- [7] J. Billard et al., Rept. Prog. Phys. **85**, 056201 (2022), 2104.07634.
- [8] J. Cooley et al., *Report of the Topical Group on Particle Dark Matter for Snowmass 2021* (2022), 2209.07426.
- [9] J. Amaré et al., Eur. Phys. J. C **79**, 228 (2019), 1812.01472.
- [10] J. Amaré et al., Phys. Rev. Lett. **123**, 031301 (2019), 1903.03973.
- [11] J. Amaré et al., J. Phys. Conf. Ser. **1468**, 012014 (2020), 1910.13365.
- [12] J. Amaré et al., Phys. Rev. D **103**, 102005 (2021), 2103.01175.

- [13] I. Coarasa et al., Commun. Phys. **7**, 345 (2024), 2404.17348.
- [14] G. Adhikari et al. (COSINE-100), Phys. Rev. Lett. **123**, 031302 (2019), 1903.10098.
- [15] G. Adhikari et al. (COSINE-100), Phys. Rev. D **106**, 052005 (2022), 2111.08863.
- [16] D. Lee et al., Commun. Phys. **8**, 135 (2025), 2409.15748.
- [17] G. Angloher et al. (COSINUS), Phys. Rev. D **110**, 043010 (2024), 2307.11139.
- [18] M. Antonello et al. (SABRE), Eur. Phys. J. C **79**, 363 (2019), 1806.09340.
- [19] F. Calaprice et al., Eur. Phys. J. C **82**, 1158 (2022), 2205.13876.
- [20] M. J. Zurewski (SABRE South), Nucl. Instrum. Meth. A **1045**, 167585 (2023).
- [21] K. Fushimi et al., Progress of Theoretical and Experimental Physics **2021** (2021), ISSN 2050-3911, 2101.00759.
- [22] N. Carlin et al., *COSINE-100 Full Dataset Challenges the Annual Modulation Signal of DAMA/LIBRA* (2024), 2409.13226.
- [23] I. Coarasa et al., JCAP **11**, 048 (2022), 2209.14113.
- [24] See Supplemental Material at URL-will-be-inserted-by-publisher for more information on the fits and results.
- [25] I. Coarasa et al., PoS **TAUP2023**, 041 (2024), 2311.03392.
- [26] J. Amaré et al., Eur. Phys. J. C **79**, 412 (2019), 1812.01377.
- [27] D. Cintas et al., Phys. Rev. C **110**, 014613 (2024), 2402.12480.
- [28] R. Bernabei et al., Phys. Lett. B **389**, 757 (1996).
- [29] I. Coarasa et al., Eur. Phys. J. C **79**, 233 (2019), 1812.02000.

END MATTER

Analysis of possible bias associated to variations in the efficiency with toy MC

We have carried out 20 000 toy MC simulations of experiments equivalent to 6 years of ANAIS-112 data (taking into account both the time-dependent background evolution and the measured efficiencies), with and without adding the modulation observed by the DAMA/LIBRA experiment. Table II shows the bias of the fitting procedure estimated from this analysis for both the modulation and null hypotheses, besides the standard deviation obtained from the fits. No bias is observed and similar standard deviations than found in the ANAIS-112 six-years results, as shown in Table I, are obtained.

Then, we have introduced variations in the detection efficiencies of different magnitudes around the mean value and analyzed the modulation hypothesis. The modulation amplitudes obtained by fitting the simulated rate evolution show distributions centered in the DAMA/LIBRA value, with a standard deviation that increases with the fluctuation in the efficiencies introduced, as expected. Figure 4 (left-upper panel) shows the distribution of the modulation amplitudes derived from the fits in the [1–6] keV energy region for fluctuations in the ef-

ficiency of 0, 2, 5 and 10%, while the ANAIS-112 result is shown as a dashed line. The measured efficiency variations in ANAIS-112 are on the order of 0.3%, which implies a negligible probability of obtaining a null result when a modulation as large as that observed by DAMA/LIBRA is present in the data. It is worth highlighting that variations as large as 5% are required to yield a non-negligible probability of hindering a true modulation in the data. On the other hand, the χ^2 value of the fit is highly sensitive to the efficiency fluctuations, as shown in the left-lower panel of Figure 4. The χ^2 values obtained by ANAIS are strongly incompatible with fluctuations larger than 2%.

In addition, taking into account a possible dependence with time of our efficiencies, shown in Figure 3 within the Supplemental Material [24], we have evaluated the effect in the results of our analysis. First, we searched for different time dependences that could be hidden within the uncertainties of the efficiencies, detector by detector, both a linear dependence (decreasing or increasing with time) and annual modulation or antimodulation.

We prepared 20 000 toy MC simulations of experiments equivalent to 6 years of ANAIS-112 data, with and without adding the modulation observed by the DAMA/LIBRA experiment, considering efficiencies with a linear dependence on time in the [1–6] keV energy region. The maximum variation observed in the detectors that show this trend (D2, D6-D8) is 0.6% around the mean value and this is the value that has been considered in the toy MCs for all detectors. The effect is completely negligible both in the modulation amplitude and in the chi-square distribution corresponding to the fits, as can be seen in the middle panels of Figure 4.

On the other hand, no consistent modulation in the efficiency of the ANAIS modules is observed when fitting to a constant or a linear variation over time plus an annual modulation. Seven out of nine modules provide results compatible with the null hypothesis. However, two modules (D4 and D7) result in the fitting with antimodulated efficiencies at 0.1% level. Because a modulation or antimodulation in the efficiency would produce a bias in the determination of the modulation amplitude, we studied the effect with the toy MCs, as done before, including an annual modulation (antimodulation) in the efficiency of 0.1% in 2 modules and constant efficiencies for the other modules. Results are shown in the right panels of Figure 4. The corresponding bias in the modulation amplitude derived from this analysis amounts to about 6.3% of the DAMA modulation amplitude (higher or lower for antimodulation and modulation, respectively). Therefore, an antimodulation of the efficiencies at this level is not sufficient to mask the DAMA/LIBRA signal.

Finally, we have also used equivalent toy MCs for evaluating the ANAIS-112 sensitivity to DAMA/LIBRA result by analyzing the residuals found after subtracting the fitted background. These residuals are shown for our six-year data in Figure 8 within the Supplemental Material [24]. Figure 5 shows the distribution of the mod-

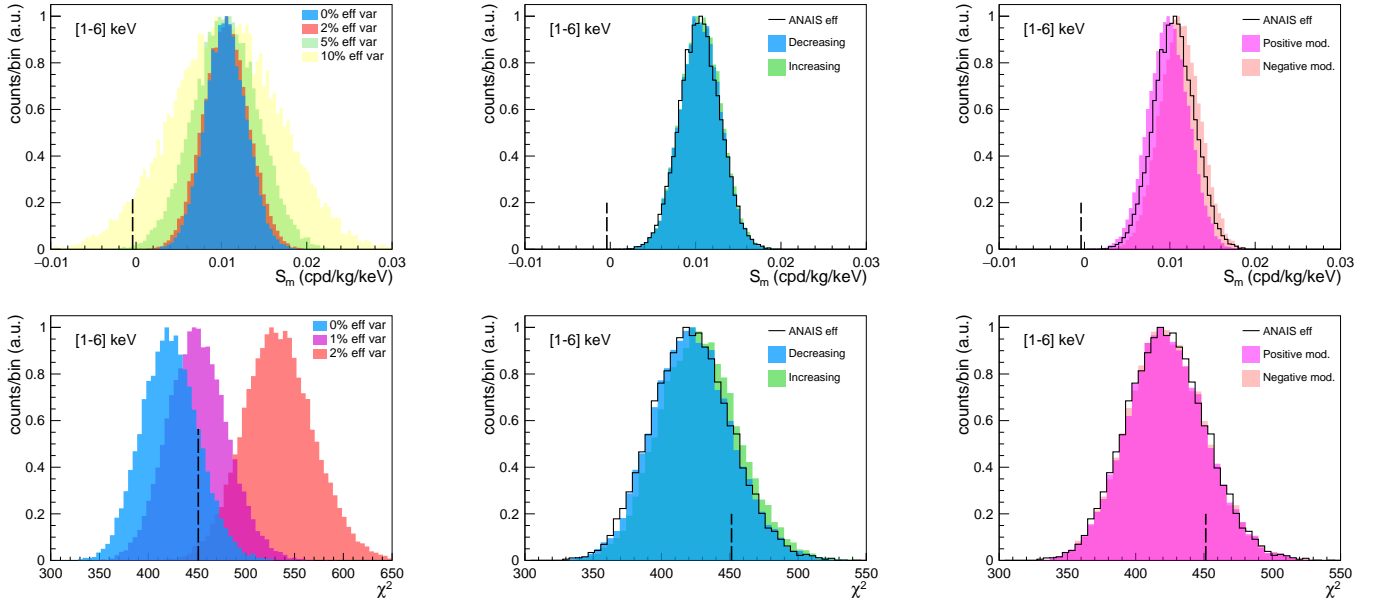


FIG. 4. Results of 20 000 toy MC simulations using the updated ANAIS-112 experimental features for 6 years, adding the modulation observed by DAMA/LIBRA. Upper panels: distribution of modulation amplitudes recovered in the [1–6] keV energy region for fluctuations in the efficiency of 0, 2, 5 and 10% (left panel); for efficiencies with a 0.6% linear variation with time in all detectors, decreasing and increasing (middle panel) and for annually modulated (or antimodulated) efficiencies in 2 modules at 0.1% and constant for the rest (right panel). Lower panels: corresponding χ^2 value distribution of the fits (ndf=422). The ANAIS-112 result is shown as dashed line in all the panels.

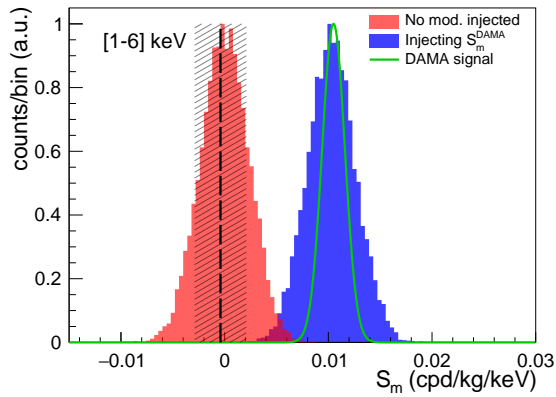


FIG. 5. Distribution of the modulation amplitudes recovered in the [1–6] keV energy region with (blue) and without (red) injecting the DAMA/LIBRA signal in 10 000 toy MC simulations. The ANAIS-112 result is represented by the dashed black line, with the uncertainty shown as a pattern of black lines, and the DAMA/LIBRA signal is displayed in green.

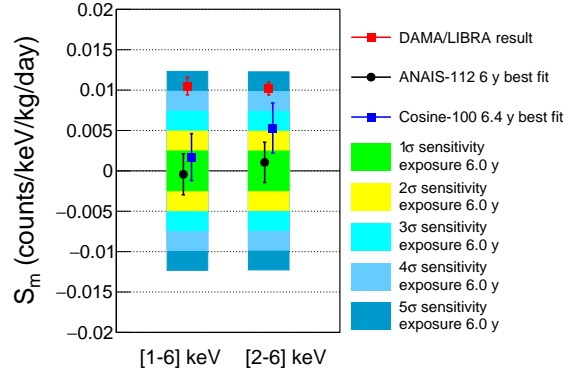


FIG. 6. Comparison of the results from ANAIS-112, COSINE-100 [22], and DAMA/LIBRA [3].

Model-independent comparison of results

Figure 6 compares the results of ANAIS-112 (this work), COSINE-100 [22] and DAMA/LIBRA [3]. The colored bands correspond to the ANAIS-112 sensitivity estimates for the six years exposure. It is worth to highlight that this comparison is independent from the DM particle and halo models and the main systematics remaining is that related with the scintillation QF for NR (already commented in the Letter).

modulation amplitudes recovered with and without injecting the DAMA/LIBRA modulation, which confirms the statistical significance of our result.

Author contributions statement

The main analysis presented in this article was conducted by IC supervised by MM and MLS. The manuscript was written by IC, MM and MLS. All authors have read and agreed to the published version, and are listed alphabetically by their last names. EG, AOdS and MLS contributed to the experiment design; JA, SC, AOdS and MLS contributed to the experiment setting-

up and commissioning; JA, JAp, SC, DC, IC, MM, YO, AOdS, TP, CS and MLS performed calibration and maintenance tasks; MM developed DAQ software tools; IC, MM and MLS developed analysis tools; IC, EG, MM, JP and MLS analyzed the data; SC and TP developed background simulations; EG, AOdS and JP performed radiopurity measurements; IC, EG, MM and JP contributed to sensitivity estimates.

Towards a robust model-independent test of the DAMA/LIBRA dark matter signal: ANAIS-112 results with six years of data.

SUPPLEMENTAL MATERIAL

DETECTOR STABILITY

Many checks to guarantee a stable energy threshold along the 6 years of data taking have been carried out. Figure 1 shows the time evolution of the rate of events corresponding to ^{22}Na and ^{40}K at 0.9 and 3.2 keV, respectively, selected by the coincidence with a high-energy gamma in a second module. ^{22}Na events, in particular, fell below the 1 keV analysis threshold, and ^{40}K events are fully contained in the ROI of the ANAIS-112 experiment. Both rates (integrated for the nine modules) show the expected time behavior, a constant rate in the case of ^{40}K and an exponential decay compatible with the isotope lifetime in the case of ^{22}Na (1389 \pm 51 d is the lifetime derived from the fit, while 1369 d is the nominal lifetime for ^{22}Na). Instabilities, either in the energy calibration or in the energy threshold, would have been identified by an anomalous time evolution of these rates. Figure 2 shows the ^{22}Na and ^{40}K events at low energy, normalized to cpd/kg/keV, selected by the coincidence with the corresponding high-energy gamma in a second module, and accumulated along the 6-year exposure of the experiment. These peaks are obtained by adding all the background runs conveniently corrected for gain drifts using the procedure explained in the introduction section and then, recalibrated. In case the gain correction was introducing additional uncertainties, these peaks would appear deformed and with larger resolution than expected. None anomalous effect has been observed. While ^{40}K events are evenly distributed along the 6-years of data taking, ^{22}Na events are scarce in the last year of the analysed exposure. It can be observed well below the 1 keV analysis threshold for ^{40}K events the peak corresponding to the L-shell-EC, confirming a sound trigger for events in the sub-keV energy range.

Figure 3 shows the time evolution of the total detection efficiency in the ROI of ANAIS-112 ([1–6] keV), estimated using both, ^{109}Cd calibration events (in blue) and neutron calibration events (in orange). Neutron calibrations are not evenly distributed along the data taking, but are concentrated in the second half. The solid blue and orange lines correspond to the mean values for each detector, and the shaded regions represent the standard deviations. Both are compatible within uncertainties. ^{109}Cd data allow to consider constant efficiencies in all the modules, but D0, D4 and D5, as explained in the introduction section. The efficiency used in the analysis is also represented in Figure 3 by a solid magenta line.

LOW ENERGY BACKGROUND

ANAIS-112 background in the ROI after applying the event selection procedures based on BDT and explained in the introduction section is shown in Figure 4. COSINE-100 and DAMA/LIBRA backgrounds are also shown for comparison. It can be clearly noticed the ^{40}K peak at 3.2 keV present in ANAIS-112 data and the excess of events below 2 keV which is still under study and the background model is not able to account for.

EXPOSURE

The accumulated exposure used for the annual modulation analysis corresponding to the six years of data of the ANAIS-112 experiment is summarized in Table I. It also details the dead time (measured using latched counters during the data taking), down time (primarily due to bi-weekly ^{109}Cd calibrations and seven ^{252}Cf calibrations carried out in the referred period), percentage of live time rejected in the analysis, and the corresponding effective exposure after subtracting the latter.

ANNUAL MODULATION FIT RESULTS

The results of the fits for the annual modulation analysis in the three energy regions considered for the nine modules are shown in Figures 5, 6 and 7. χ^2/ndf and p-values under the modulation hypothesis are also individually displayed for each module, besides the global analysis χ^2/ndf and p-values for the null and modulation hypothesis and the S_m corresponding to the best fit for the latter. In all the energy regions and detectors, we obtain good fits to the PDF built with our MC background model time evolution. We can conclude that our data are fully consistent with no modulation and incompatible with DAMA/LIBRA. Fit results for the modulation amplitude and corresponding p-values were shown in Table-I in the Letter. Table II reports on the nuisance parameters derived from the fits, both background index and f for each module, which measures the deviation of the fit from the MC background model detector by detector. Only detector 5 shows systematically values of f below 85%. Figure 8 shows the results of the fit after subtracting the non-modulated component. The modulation observed by DAMA/LIBRA is shown in the same plot for comparison.

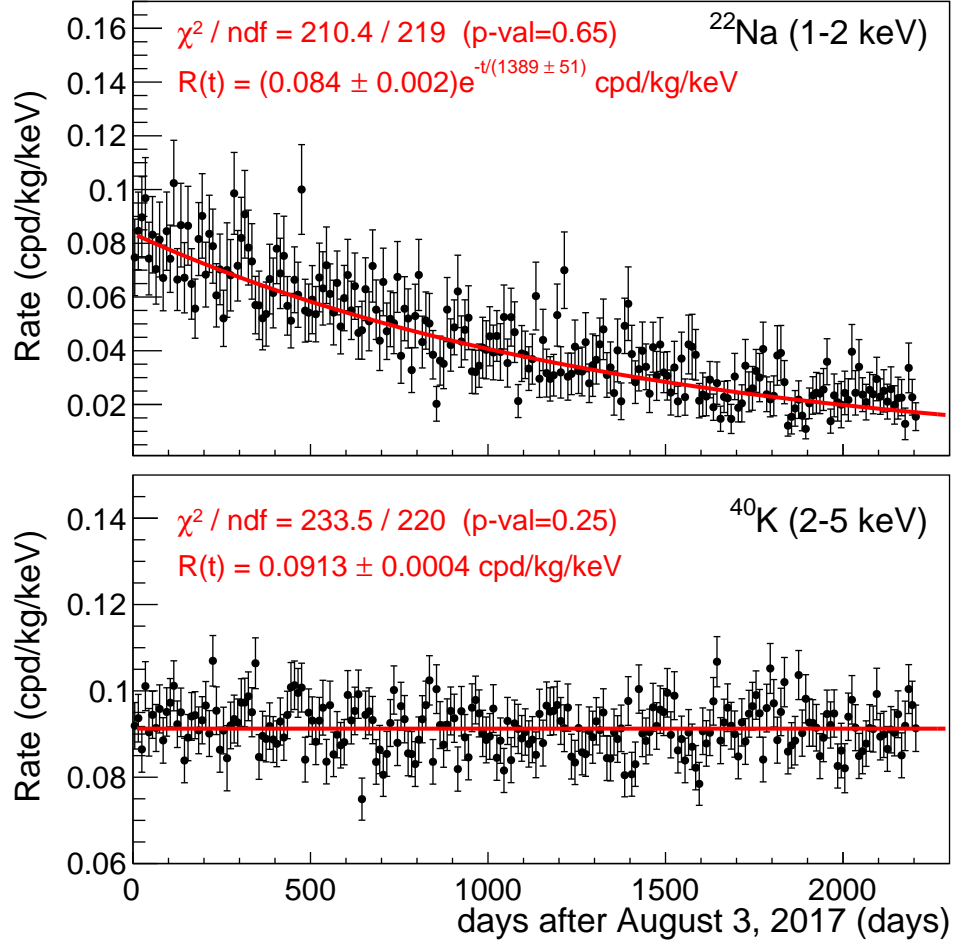


FIG. 1. Time evolution of the rate of events corresponding to ^{22}Na (upper panel) and ^{40}K (lower panel) at low energy identified by the coincidence with the corresponding high-energy gamma in a second module.

ANAIS SENSITIVITY PROSPECTS

Figure 9 displays in dark blue line the ANAIS-112 sensitivity projection in the [1–6] keV energy region following Ref. [3], conveniently updated to the effective exposure, background level and detection efficiency presented in this work. Similar sensitivities are obtained for the [2–6] keV energy region. Cyan band takes into account

the 68% uncertainty in S_m^{DAMA} . The black dot is the sensitivity derived from the 6-year result presented here, in good concordance with our estimate. This result supports our expectation of achieving a 5σ sensitivity to the DAMA/LIBRA result by the end of 2025.

REFERENCES

-
- [1] G. Adhikari *et al.* (COSINE-100), Strong constraints from COSINE-100 on the DAMA dark matter results using the same sodium iodide target, *Sci. Adv.* **7**, abk2699 (2021), arXiv:2104.03537 [hep-ex].
 - [2] R. Bernabei *et al.*, The DAMA project: Achievements, implications and perspectives, *Prog. Part. Nucl. Phys.* **114**, 103810 (2020).
 - [3] I. Coarasa *et al.*, ANAIS-112 sensitivity in the search for dark matter annual modulation, *Eur. Phys. J. C* **79**, 233 (2019), arXiv:1812.02000 [astro-ph.IM].

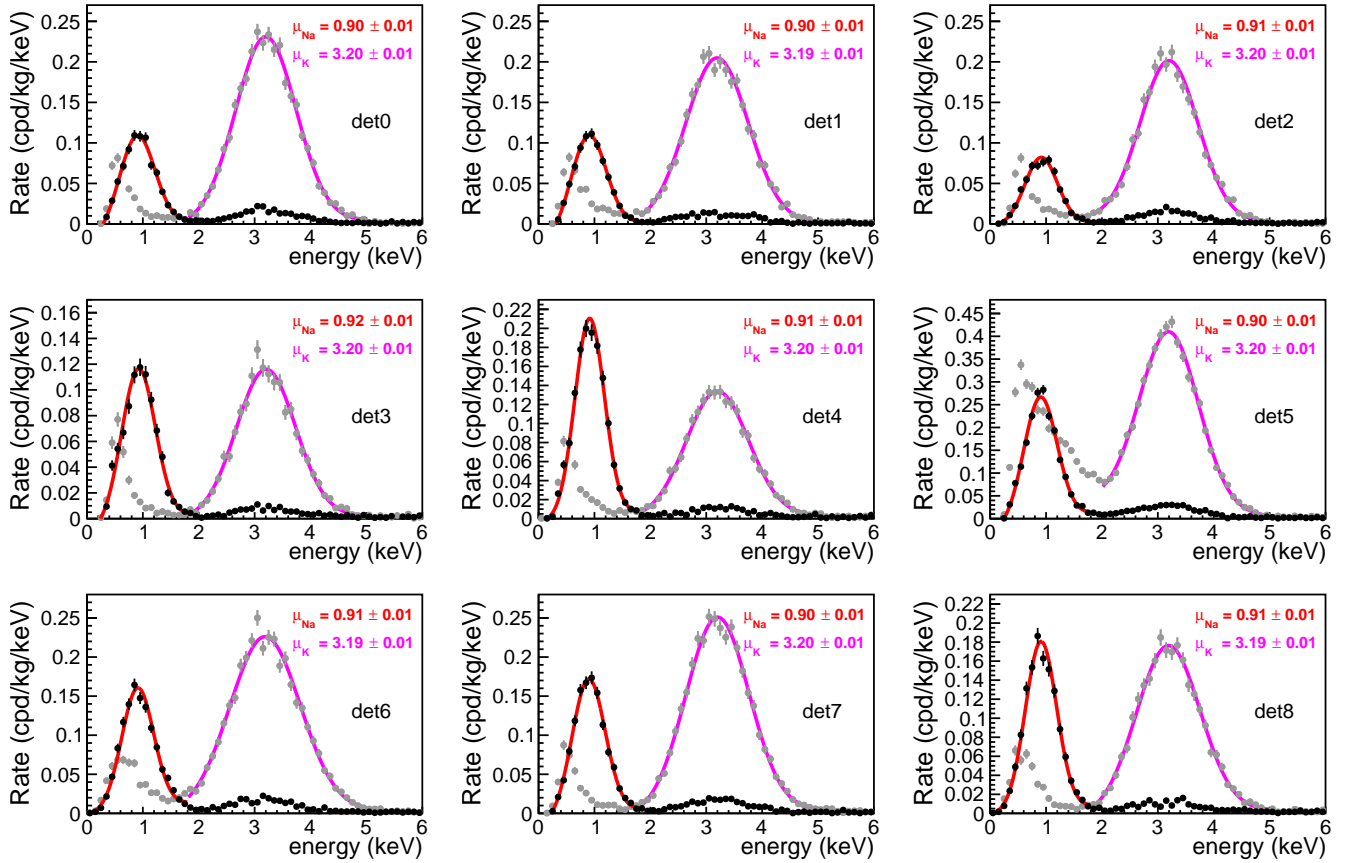


FIG. 2. Low energy spectra of the nine modules in coincidence with a high-energy gamma in the range [1200–1340] keV (black points) and [1340–1560] keV (gray points) in a second module (after recalibration described in the introduction section). Data correspond to the 6-year exposure of ANAIS–112. Peaks at 0.9 and 3.2 keV, attributed to ^{22}Na and ^{40}K decay in the NaI bulk, respectively, can be clearly observed. Solid lines: fit to a Gaussian plus linear background lineshape. Mean value of the energy corresponding to the main peaks is also shown in each panel with the corresponding uncertainty, in keV.

Time period	Exposure (kg×yr)	Dead time (%)	Down time (%)	Rejected periods (%)		Effective exposure (kg×yr)
				muon cut	rate cut	
Aug 3, 2017 – July 31, 2018	104.80	2.88	3.19	2.64	0.60	101.19
Aug 1, 2018 – Aug 28, 2019	115.39	2.07	2.42	2.64	0.38	111.75
Aug 29, 2019 – Aug 13, 2020	102.86	2.38	2.54	2.53	0.38	99.72
Aug 14, 2020 – Aug 3, 2021	104.40	2.42	2.44	2.59	0.34	101.19
Aug 4, 2021 – Aug 30, 2022	116.86	1.89	1.39	2.83	0.28	113.12
Aug 31, 2022 – Aug 17, 2023	102.24	1.80	3.96	2.74	0.46	98.78
TOTAL	646.55					625.75

TABLE I. For each of the six years of data collection, first column: start and end dates; second column: exposure calculated by multiplying live time by mass (112.5 kg); third column: percentage of dead time; fourth column: percentage of down time; fifth and sixth columns: percentages of time with respect to the live time corresponding to the two types of rejected periods (one second after a muon triggering the veto and one day if the daily averaged rate is above three standard deviations from the annually averaged rate, respectively); last column: effective exposure after subtracting the rejected periods.

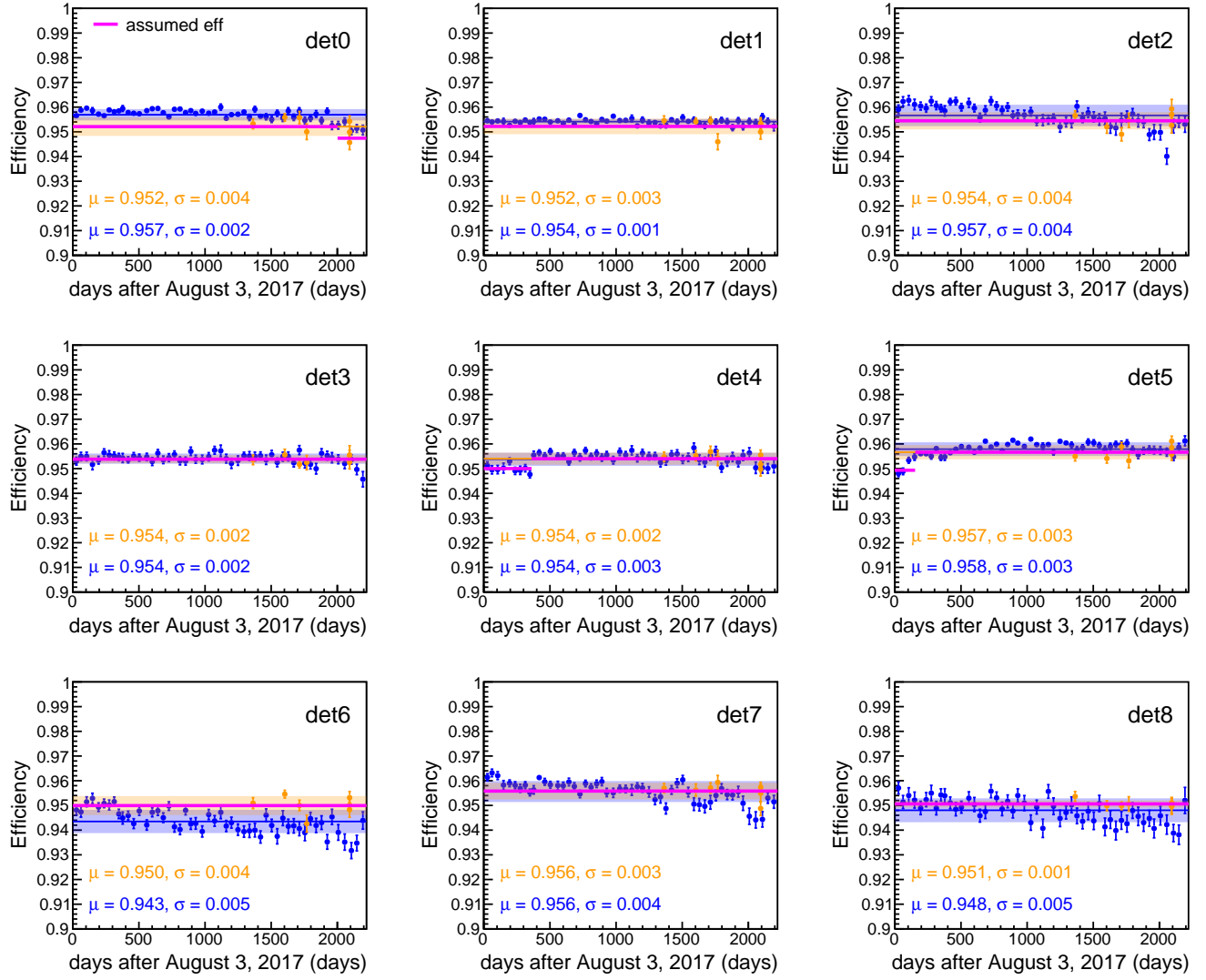


FIG. 3. Evolution of the total detection efficiency in [1–6] keV estimated from ^{109}Cd calibration (in blue) and neutron calibration (in orange) along the six years of data taking for the nine ANAIS-112 modules. The solid blue and orange lines correspond to the mean values for each detector, and the shaded regions represent the standard deviations. The mean value and the standard deviation for each module are also shown in the panels. The efficiency used in the analysis is also represented by a solid magenta line.

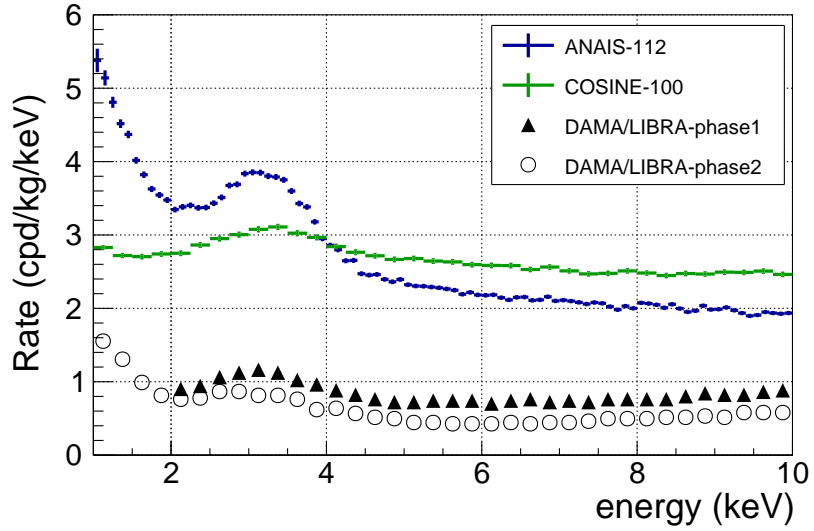


FIG. 4. Comparison of the total single-hit energy spectrum measured after six years of data in ANAIS-112 (blue points), COSINE-100 [1] (green points) and DAMA/LIBRA [2] (phase1, black triangles, and phase2, open circles).

Detector	[1–6] keV		[2–6] keV		[6.7–20] keV _{nr}	
	Bkg index	f	Bkg index	f	Bkg index	f
	(cpd/kg/keV)		(cpd/kg/keV)		(cpd/kg/3.3 keV _{nr})	
0	4.711 ± 0.007	0.98 ± 0.02	4.396 ± 0.007	0.97 ± 0.03	3.512 ± 0.006	0.98 ± 0.03
1	4.672 ± 0.007	0.90 ± 0.02	4.438 ± 0.007	0.98 ± 0.03	3.550 ± 0.006	0.91 ± 0.03
2	2.550 ± 0.005	0.96 ± 0.03	2.334 ± 0.005	0.97 ± 0.04	1.959 ± 0.004	0.90 ± 0.04
3	3.563 ± 0.006	0.85 ± 0.03	3.064 ± 0.006	0.94 ± 0.03	2.637 ± 0.005	0.85 ± 0.03
4	3.100 ± 0.005	0.88 ± 0.02	2.871 ± 0.005	0.95 ± 0.03	2.244 ± 0.005	0.86 ± 0.03
5	3.308 ± 0.06	0.63 ± 0.02	2.805 ± 0.005	0.80 ± 0.03	2.507 ± 0.005	0.64 ± 0.03
6	2.698 ± 0.005	0.95 ± 0.02	2.572 ± 0.005	0.99 ± 0.02	2.097 ± 0.005	0.99 ± 0.03
7	2.525 ± 0.005	0.96 ± 0.02	2.337 ± 0.005	1.05 ± 0.03	1.934 ± 0.004	0.96 ± 0.03
8	2.049 ± 0.004	0.97 ± 0.02	1.956 ± 0.004	0.99 ± 0.03	1.546 ± 0.004	0.98 ± 0.03

TABLE II. Summary of the nuisance parameters obtained in the fits searching for an annual modulation in the six years of ANAIS-112 data taking for different energy regions.

Null hyp χ^2/ndf : 451.34/423 [$p_{\text{val}}=0.164$]

Mod hyp χ^2/ndf : 451.31/422 [$p_{\text{val}}=0.156$]

$S_m = (-0.0004 \pm 0.0025)$ (cpd/kg/keV)

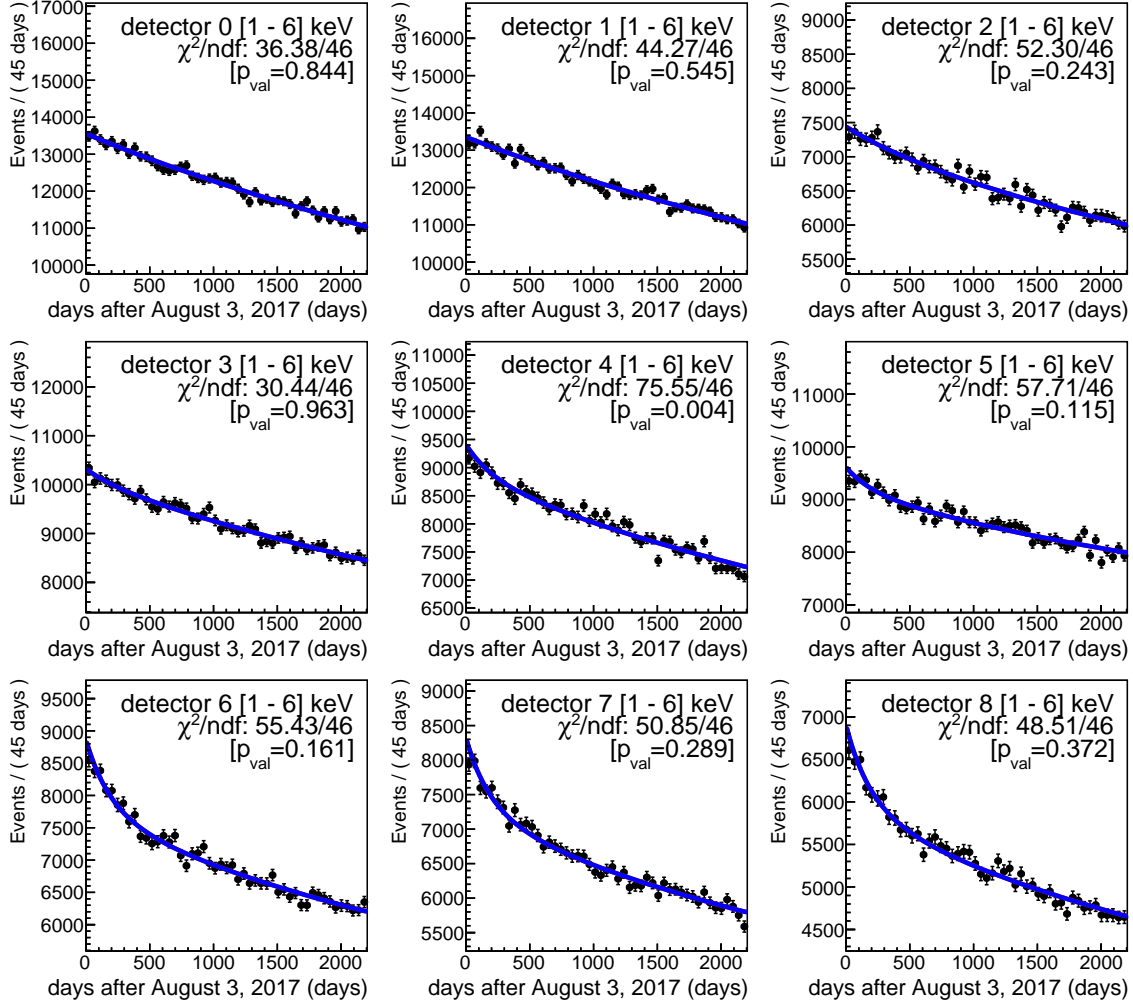


FIG. 5. Results of the fit for the data from the nine modules in the [1-6] keV energy region, under the modulation (blue) and null hypotheses (red). In all the panels, the red line is masked by the blue one, as the fit obtained for the modulated hypothesis is consistent with $S_m = 0$. χ^2/ndf and p-values under the modulation hypothesis are also individually displayed for each module.

Null hyp χ^2/ndf : 414.46/423 [$p_{\text{val}}=0.607$]

Mod hyp χ^2/ndf : 414.28/422 [$p_{\text{val}}=0.596$]

$S_m = (0.0011 \pm 0.0025)$ (cpd/kg/keV)

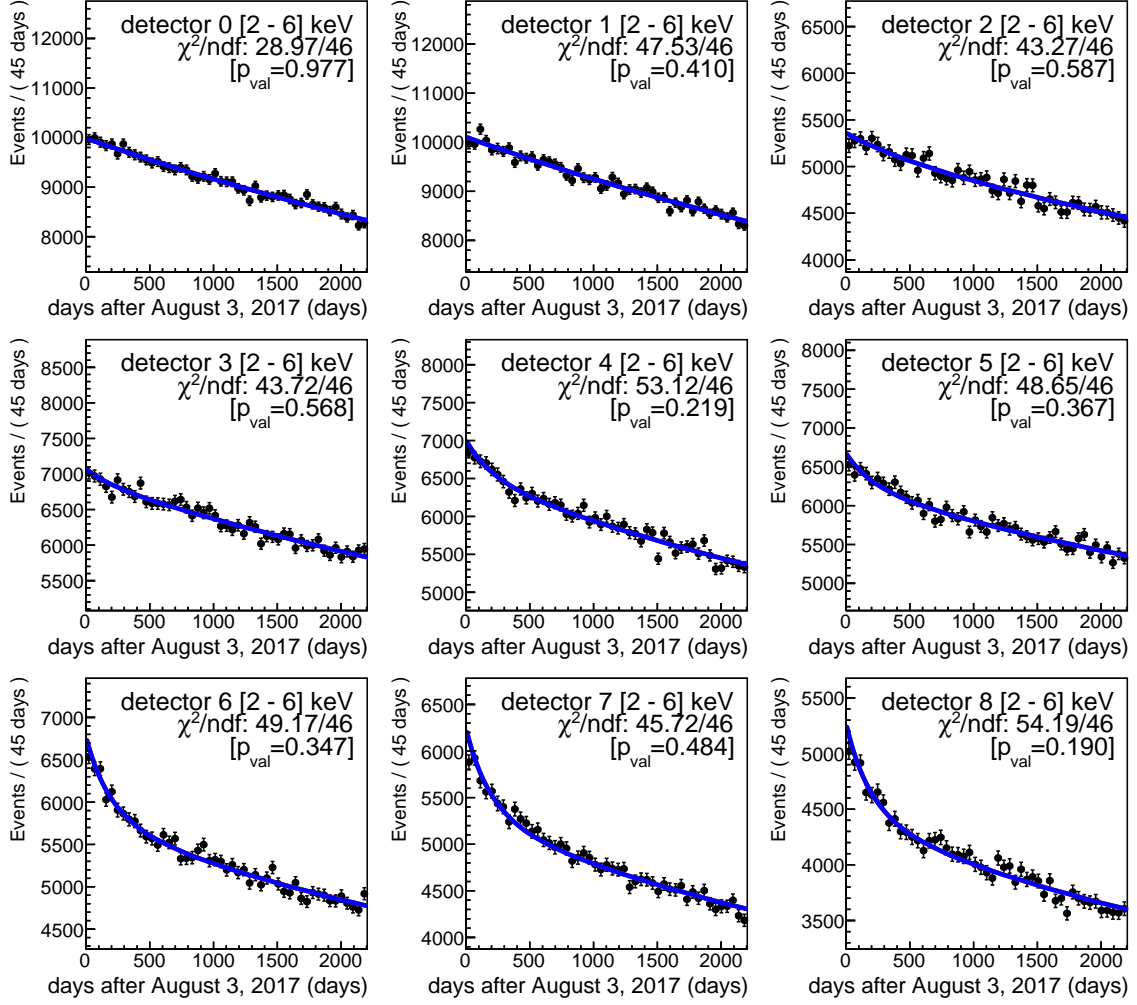


FIG. 6. Results of the fit for the data from the nine modules in the [2–6] keV energy region, under the modulation (blue) and null hypotheses (red). In all the panels, the red line is masked by the blue one, as the fit obtained for the modulated hypothesis is consistent with $S_m = 0$. χ^2/ndf and p-values under the modulation hypothesis are also individually displayed for each module.

Null hyp χ^2/nfd : 416.51/423 [$p_{\text{val}}=0.580$]

Mod hyp χ^2/nfd : 416.51/422 [$p_{\text{val}}=0.566$]

$S_m = (0.0000 \pm 0.0034)$ (cpd/kg/keV)

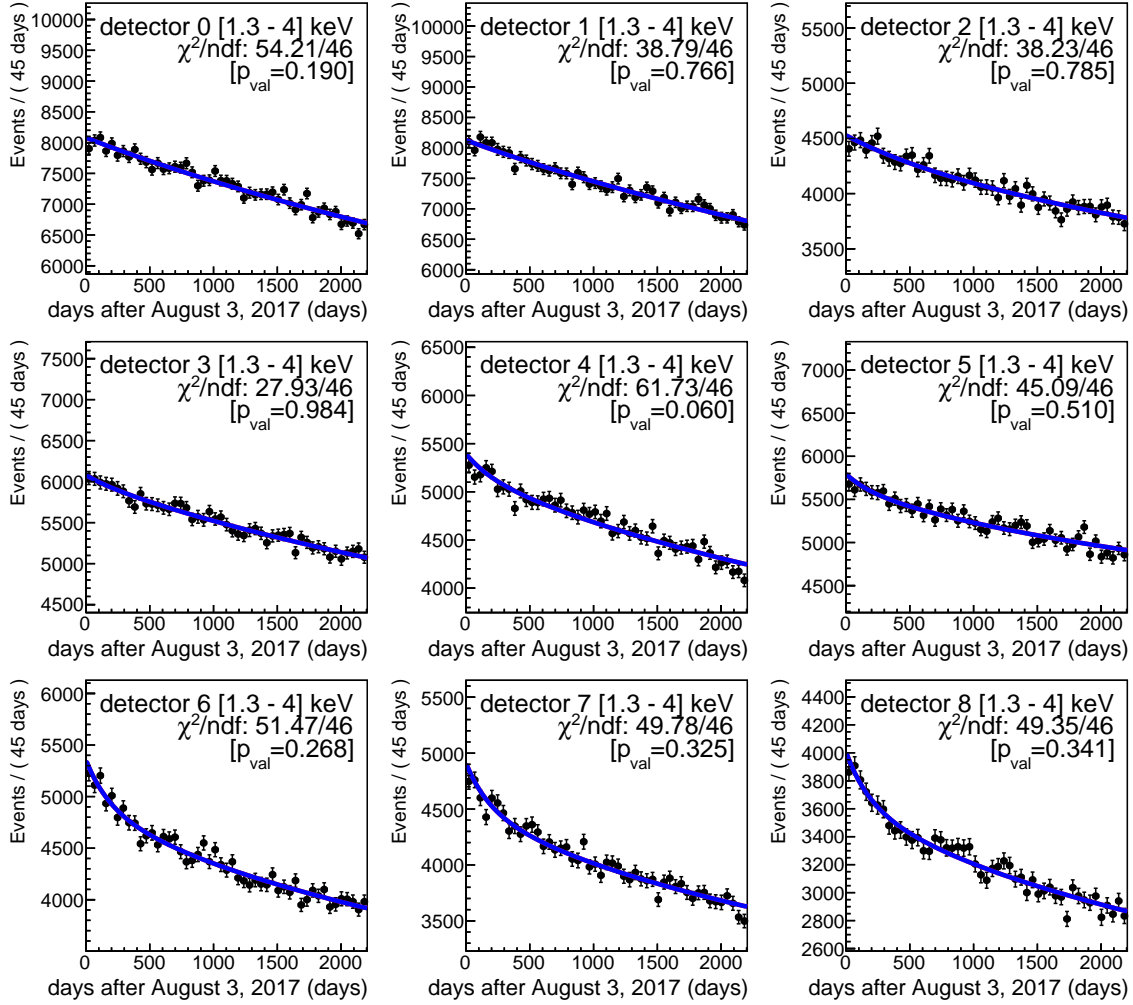


FIG. 7. Results of the fit for the data from the nine modules in the [1.3–4] keV energy region, corresponding to the sodium nuclear recoil energy range of [6.7–20] keV_{nr}, under the modulation (blue) and null hypotheses (red). In all the panels, the red line is masked by the blue one, as the fit obtained for the modulated hypothesis is consistent with $S_m = 0$. χ^2/nfd and p -values under the modulation hypothesis are also individually displayed for each module.

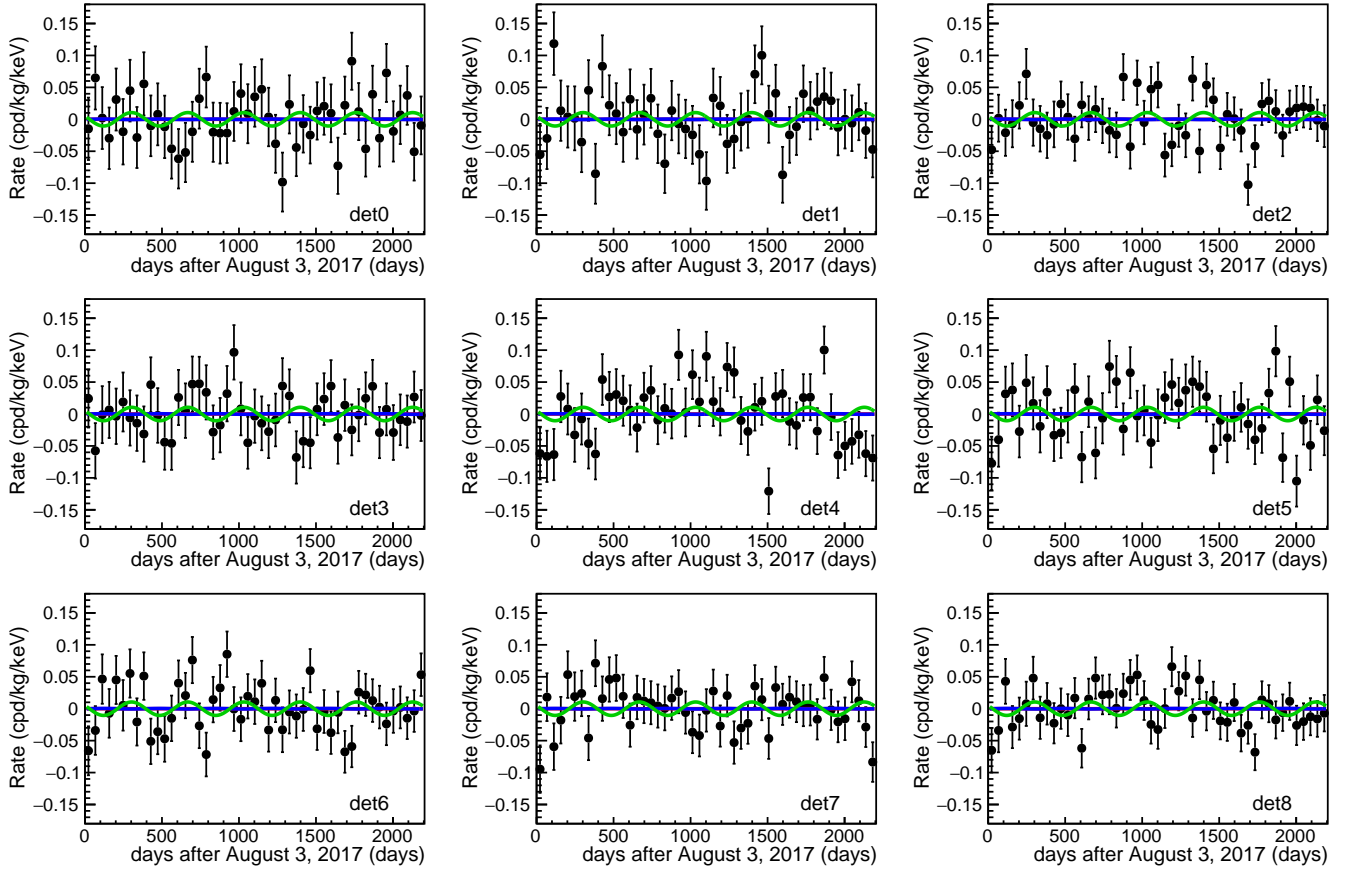


FIG. 8. Fit results for the data from the nine modules in the [1–6] keV energy region after subtracting the non-modulated term from Eq. 2. Blue and red lines are the result of the fit for the modulation and null hypothesis, respectively, after subtracting the non-modulated term from Eq. 2. The modulation observed by DAMA/LIBRA is shown in green.

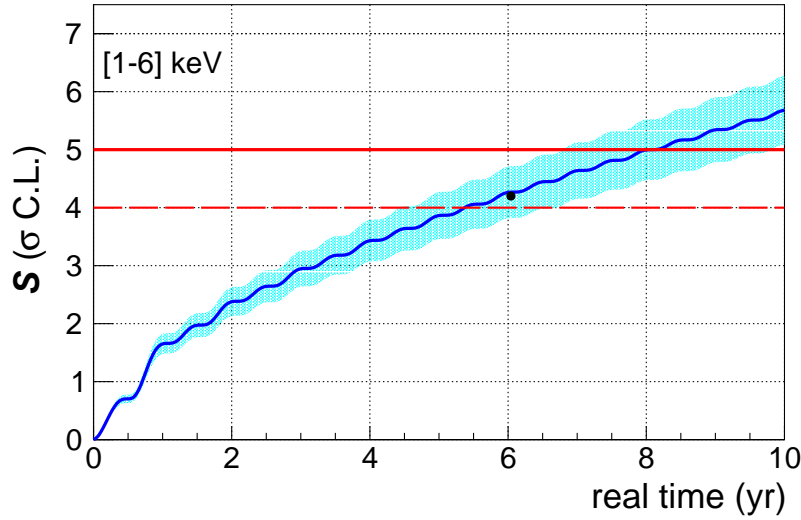


FIG. 9. ANAIS-112 sensitivity to the DAMA/LIBRA signal in σ C.L. units as a function of real time in the [1–6] keV energy region. The black dot is the sensitivity measured experimentally for 6-year exposure. The cyan band represents the 68% C.L. DAMA/LIBRA uncertainty.