

Improving ANAIS–112 sensitivity to DAMA/LIBRA signal with machine learning techniques

I. Coarasa,^{a,b} J. Apilluelo,^a J. Amaré,^{a,b} S. Cebrián,^{a,b}
D. Cintas,^{a,b} E. García,^{a,b} M. Martínez^{a,b,c} M.A. Oliván,^{a,b,d}
Y. Ortigoza,^{a,b,e} A. Ortiz de Solórzano,^{a,b} T. Pardo,^{a,b}
J. Puimedón,^{a,b} A. Salinas,^{a,b} M.L. Sarsa,^{a,b} and P. Villar^a

^aCentro de Astropartículas y Física de Altas Energías (CAPA), Universidad de Zaragoza, Pedro Cerbuna 12, 50009 Zaragoza, Spain

^bLaboratorio Subterráneo de Canfranc, Paseo de los Ayerbe s.n., 22880 Canfranc Estación, Huesca, Spain

^cFundación ARAID, Av. de Ranillas 1D, 50018 Zaragoza, Spain

^dFundación CIRCE, Av. de Ranillas 3D, 50018 Zaragoza, Spain

^eEscuela Universitaria Politécnica de La Almunia de Doña Godina (EUPLA), Universidad de Zaragoza, Calle Mayor 5, La Almunia de Doña Godina, 50100 Zaragoza, Spain

E-mail: icoarasa@unizar.es, japilluelo@unizar.es, amare@unizar.es,
scebrian@unizar.es, cintas@unizar.es, edgarcia@unizar.es, mariam@unizar.es,
maolivan@unizar.es, ortigoza@unizar.es, alfortiz@unizar.es, tpardo@unizar.es,
puimedon@unizar.es, salinas@unizar.es, mlsarsa@unizar.es, pvillar@unizar.es

Abstract. The DAMA/LIBRA observation of an annual modulation in the detection rate compatible with that expected for dark matter particles from the galactic halo has accumulated evidence for more than twenty years. It is the only hint of a direct detection of the elusive dark matter, but it is in strong tension with the negative results of other very sensitive experiments, requiring *ad-hoc* scenarios to reconcile all the present experimental results. Testing the DAMA/LIBRA result using the same target material, NaI(Tl), removes the dependence on the particle and halo models and is the goal of the ANAIS–112 experiment, taking data at the Canfranc Underground Laboratory in Spain since August 2017 with 112.5 kg of NaI(Tl). At very low energies, the detection rate is dominated by non-bulk scintillation events and careful event selection is mandatory. This article summarizes the efforts devoted to better characterize and filter this contribution in ANAIS–112 data using a boosted decision tree (BDT), trained for this goal with high efficiency. We report on the selection of the training populations, the procedure to determine the optimal cut on the BDT parameter, the estimate of the efficiencies for the selection of bulk scintillation in the region of interest (ROI), and the evaluation of the performance of this analysis with respect to the previous filtering. The improvement achieved in background rejection in the ROI, but moreover, the increase in detection efficiency, push the ANAIS–112 sensitivity to test the DAMA/LIBRA annual modulation result beyond 3σ with three-year exposure, being possible to reach 5σ by extending the data taking for a few more years than the scheduled 5 years which were due in August 2022.

Contents

1	Introduction	1
2	The ANAIS–112 experiment	3
2.1	The apparatus	3
2.2	The analysis	4
2.3	The neutron calibration program	6
2.4	The blank module set-up and performance	7
3	Machine learning technique applied to ANAIS–112 data	10
3.1	Boosted Decision Tree algorithm	12
3.2	Training populations	13
3.3	Input parameters	14
3.4	Training and BDT response	15
4	Bulk scintillation events selection	17
5	BDT output validation	22
6	ANAIS–112 sensitivity improvement with BDT filtering	23
7	Conclusions	23

1 Introduction

For almost a century, evidence on the existence of an unknown matter component in the Universe, from the scale of galaxies and clusters of galaxies up to cosmological scales, has been steadily accumulating. The Λ CDM cosmological model is able to satisfactorily explain all the observations at the cost of introducing 84% of the matter in the form of Dark Matter (DM) [1, 2]. Moreover, 68% of the total matter-energy should be in the form of Dark Energy (DE). Understanding the nature of the DM has shown to be a challenge, remaining one of the most puzzling unresolved questions in Fundamental Physics. Progressing in this quest requires following complementary approaches from the different points of view of cosmology, astronomy, astrophysics, and nuclear and particle physics, and both from theory and experiment. Despite the large efforts invested in the searches for new particles with the right properties to be this DM, by combining searches at accelerators [3], indirect [4] and direct searches [5, 6], the solution to the DM puzzle remains beyond our present sensitivities.

DM direct detection experiments aim to identify the small energy deposited in a convenient target material by the interaction of the DM particles filling the Milky Way’s halo. The expected signal is strongly dependent on the particle DM model and galactic halo model¹, being either unknown or affected by large uncertainties. In the preferred DM scenarios, these interactions should produce nuclear recoils by the elastic scattering off the target nuclei of the detector. In spite of the wide experimental effort, after more than three decades of DM

¹In particular, the distribution of velocities of the DM particles and the particle density at the Solar System position.

direct searches, most of the experimental results are compatible with the estimated backgrounds [6, 7]. Only the DAMA/LIBRA experiment, which uses NaI(Tl) crystal detectors at the Gran Sasso National Laboratory (LNGS) in Italy, has provided a long-standing positive result [8–14], the observation of a highly statistically significant annual modulation in the low energy detection rate (below 6 keV_{ee}²), which is compatible with that expected for DM particles distributed following the standard halo model [15, 16]. The DAMA/LIBRA result has not been reproduced by any other experiment and, for years, the most plausible DM scenarios have been strongly constrained by the results of other experiments using different target materials and detection techniques [17–35]. A model-independent test of the DAMA/LIBRA result requires an experiment using the same target material, a similar (or better) background and threshold energy, large exposure (product of mass and measurement time) and stability, in order to analyze the annual modulation effect. This is the goal of the ANAIS (Annual modulation with NaI Scintillators) experiment [36, 37]: to provide a model independent confirmation or refutation of the annual modulation positive signal reported by DAMA/LIBRA using the same target and technique, but different experimental conditions. ANAIS–112 started taking data at the Canfranc Underground Laboratory (LSC), in Spain, in August 2017 with a mass of 112.5 kg and the last reported annual modulation results, corresponding to 3 years of data taking, are not compatible with DAMA/LIBRA, for a sensitivity of 2.5–2.7 σ , depending on the energy region [38].

Other projects worldwide share the ANAIS’ goal. The COSINE–100 experiment is taking data at YangYang laboratory, in South Korea since September 2016 with an effective mass of ≈ 60 kg. COSINE–100 has released model-dependent results excluding DAMA/LIBRA-compatible parameters regions [39], while the last annual modulation analysis reported point to a compatibility with both DAMA/LIBRA and absence of modulation [40]. Other projects in R&D phase are SABRE [41, 42], PICOLON [43] and COSINE–200, aiming at improving the radiopurity of the NaI(Tl) crystals using an only-scintillation detection approach, whereas the COSINUS project [44] bases the detection on the simultaneous measurement of the light and phonons, which allows identifying the interacting particle by the different sharing between the two energy conversion channels. Previously, DM-Ice17, operating at the South Pole, tried to address this goal, but its sensitivity was too low [45].

ANAIS sensitivity is limited by anomalous scintillation events which dominate the detection rate below 10 keV. Robust protocols for the selection of events associated with the NaI(Tl) scintillation in the crystal bulk, as DM particles should produce, are then mandatory. In this paper, a complete description of a new machine learning procedure developed for the selection of low-energy bulk scintillation events in ANAIS–112 data is presented. The previously used analysis protocol showed a very low acceptance efficiency at energies below 2 keV and, in addition, resulted in a background excess mostly below 2 keV which could be attributed to non-scintillation events leaking the filtering procedure [36]. This motivated the search for more powerful filtering techniques to be applied to ANAIS–112 data, with the consequent improvement in sensitivity.

Machine learning techniques applied in other experiments searching for rare events have demonstrated to be very powerful in discriminating signal from noise [46–51]. Following such a different strategy for the filtering of ANAIS–112 data we expected both, to improve acceptance efficiencies for bulk scintillation events (signal) and better rejection of non-bulk scintillation events (noise) between 1 and 2 keV. For that goal, we have implemented a supervised machine

²Electron-equivalent energy. In the following, we will use just keV for keV_{ee}.

learning technique based on a multivariate analysis, in particular a boosted decision tree (BDT). In a previous work [52, 53], a two-phase training was performed: ^{109}Cd calibration events were considered signal events while noise samples were built from a blank module (i.e., identical to the ANAIS–112 ones, but without scintillating NaI(Tl) crystal) events in the first phase, see Section 2.4, and background events in the second phase. In this paper, we report on the update of this BDT approach, profiting from the development of a dedicated neutron calibration program for the ANAIS–112 experiment, see Section 2.3. We will use as signal events for the algorithm training scintillation events in the bulk of the crystals produced by neutron interactions, which are mostly associated to elastic nuclear recoils in that energy region, and as noise events, those coming from the blank module, without requiring to use background events at all.

The structure of the article is the following: some relevant features of the ANAIS–112 experimental set-up, in particular the complementary blank module set-up and the dedicated neutron calibration program developed in the last year of operation, are described in Section 2; the machine learning algorithm, training populations and parameters selected and the algorithm response are presented in Section 3; Section 4 describes the event selection based on the machine learning algorithm output; in Section 5, the validation of the machine learning response is checked; the sensitivity improvement to the annual modulation observed by DAMA/LIBRA is covered in Section 6. Finally, conclusions are summarized in Section 7.

2 The ANAIS–112 experiment

2.1 The apparatus

ANAIS–112 consists of nine 12.5 kg NaI(Tl) modules in a 3×3 matrix configuration, for a total mass of 112.5 kg. The ANAIS modules were manufactured by Alpha Spectra Inc. in Colorado (US) between 2012 and 2017. All the crystals are cylindrical (4.75" diameter and 11.75" length) and are housed in OFE (Oxygen Free Electronic) copper; the encapsulation has a Mylar window in the lateral face for low energy calibration using external gamma sources. Every crystal was coupled through synthetic quartz windows to two highly efficient Hamamatsu R12669SEL2 photomultipliers (PMTs) at the LSC clean room. The high optical quality of the crystals added to the high-efficiency PMTs result in remarkable light collection at the level of 15 photoelectrons/keV [36, 54] in all the modules.

The ANAIS–112 shielding consists of 10 cm of archaeological lead, 20 cm of low activity lead, an anti-radon box (continuously flushed with radon-free nitrogen gas), an active muon veto system made up of 16 plastic scintillators covering the top and sides of the whole set-up and 40 cm of neutron moderator (a combination of water tanks and polyethylene blocks). The left panel of Figure 1 shows a design of the set-up. The hut housing the experiment is located at the hall B of LSC under a rock overburden equivalent to 2450 m of water. A more detailed description of the ANAIS–112 experimental set-up can be found in [36].

Concerning the ANAIS–112 data acquisition system, each PMT signal is fully processed by the electronic chain and divided into a trigger signal, a low-energy (LE) signal and a high-energy (HE) signal. Triggering of every detector is done by the coincidence (logical AND) of the two PMT signals of any detector at photoelectron (phe) level in a 200 ns window, whereas the main acquisition trigger is the logical OR of individual detectors. The HE signals are fed into CAEN V792 QDC (charge-to-digital-converter) modules, with 1 μs integration window and 12 bits resolution. The LE signals of each PMT are sent to MATAcq-CAEN V1729A digitizers, which sample the waveforms at 2 GS/s in a 1.2 μs window with high

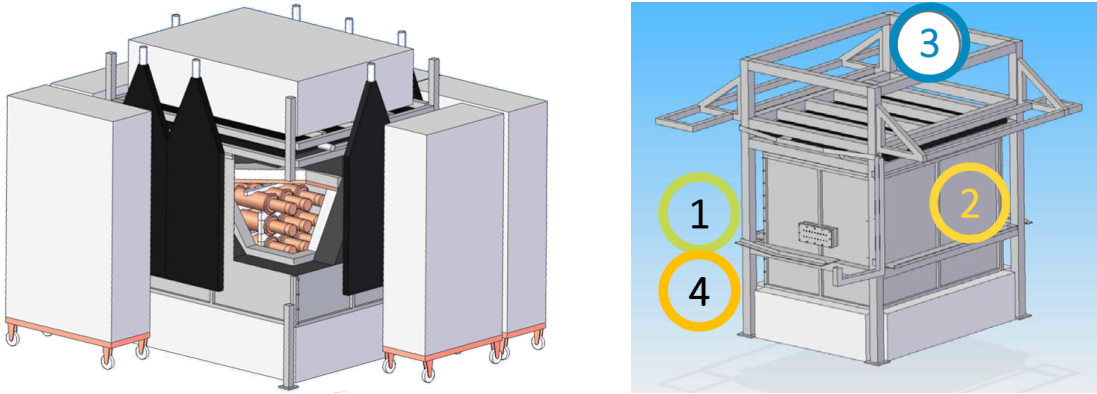


Figure 1: Left: artistic view of the ANAIS–112 set-up at LSC showing the different shielding components (see text for more details). Right: position of the ^{252}Cf neutron source during the different calibration runs carried out: west side (1 and 4), south side (2) and top side (3). The detector axis is aligned parallel to the east-west direction. The source is placed outside the anti-radon box and lead shielding, but inside the muon vetoes and neutron shielding.

resolution (14 bits). In the analysis phase, a software trigger position (t_0) is determined for every waveform by using a threshold level over the baseline, independently from the hardware trigger threshold. t_0 is used later in the calculation of several pulse shape parameters (see Sections 2.2 and 3).

2.2 The analysis

The main analysis parameters are drawn from the waveform of LE pulses. In particular, the pulse area (our energy estimator) is calculated by integrating the pulse from the software trigger position, individually for each PMT. Moreover, peaks in the waveform of each PMT are identified using a peak-searching algorithm [55] based on a low pass filter and the detection of a sign change in the derivative waveform. The number of peaks in the waveform of every photomultiplier (n_0 and n_1 for PMT0 and PMT1, respectively) are stored, besides the corresponding positions in the waveform. These variables, as it will be commented below, are very relevant in the ANAIS–112 analysis. Other analysis parameters, in particular those related with the pulse shape, will be introduced in following sections.

The response of the ANAIS–112 modules is monitored through the PMTs gain and the light collection. In particular, the PMT gain is directly monitored by the charge distribution of a single photoelectron for each PMT, the so-called single electron response (SER). We estimate periodically, every two weeks approximately, this SER for each PMT by selecting the last peak identified in the tail of PMT pulses having a very low number of peaks.

At very low energy, the number of peaks in the pulse corresponds to the number of photoelectrons. However, as energy increases, overlapping of photoelectron signals in time produces a reduction of the number of identified peaks. Because of this, we define the number of photoelectrons ($nphe$), independently from the number of peaks, as the pulse area divided by the mean SER area for each PMT ($nphe_0$ and $nphe_1$ for PMT0 and PMT1 signals, respectively), and then we add both to build the total number of photoelectrons for each module, either the blank module or the ANAIS–112 modules. Some examples of pulses with a low number of phe can be found in Section 2.4, Figure 6.

The stability of the ANAIS–112 experiment is monitored by calibrating with external ^{109}Cd sources every two weeks. All the nine modules are simultaneously calibrated using a multi-source system built in a flexible wire, which allows the sources to be introduced from the outside, keeping the system closed and guaranteeing radon-free operation. The energy calibration of the region of interest (ROI, 1–6 keV) is carried out by combining the information from those ^{109}Cd lines and from known lines present in the background at 3.20 and 0.87 keV (following the electron capture decay of ^{40}K and ^{22}Na from bulk crystal contamination, respectively). The events corresponding to those background lines are tagged by coincidences with high energy depositions (1460.8 and 1274.5 keV, respectively) in a second module. The use of $^{22}\text{Na}/^{40}\text{K}$ lines increases the reliability of the energy calibration in the ROI since they are actually either in the ROI or very close to the analysis energy threshold, set at 1 keV [36]. Moreover, $^{22}\text{Na}/^{40}\text{K}$ events provide a clean sample of very low energy depositions converted into scintillation in the crystal bulk unlike ^{109}Cd events, which because of the external origin of the gammas, correspond to scintillation produced near to the surface in the middle position between both PMTs. This makes this $^{22}\text{Na}/^{40}\text{K}$ population very interesting in the ANAIS–112 analysis.

The ANAIS–112 background in the ROI is dominated by non-bulk scintillation events [36], requiring the development of robust filtering protocols (based on the pulse shape and light sharing among the two PMTs) to reduce it. The fine tuning of these event rejection procedures was performed unblinding the $\sim 10\%$ of the data from the first year of measurement of the experiment, chosen randomly distributed days along the data taking period. The efficiency of these selection criteria was calculated with populations of scintillation events (from ^{109}Cd , ^{40}K , and ^{22}Na) in [36], showing to be very close to 100% down to 2 keV, but then decreasing steeply to about 15% at 1 keV. However, the ANAIS–112 modules are effectively triggering below 1 keV, as $^{22}\text{Na}/^{40}\text{K}$ coincident events show [36].

The origin of these non-bulk scintillation events is not fully established, but undoubtedly they are associated to the detection of light in the PMTs and not to spurious coincidences between individual photoelectrons with origin in the PMTs dark current. Some of those events, typically fast and asymmetric in the light sharing, are originated by Cherenkov emission in the glass structure of the PMTs. Nevertheless, other kind of light emissions are observed, which could be produced at the dynodes of the PMTs, quartz windows, optical coupling media, Teflon wrapping, etc.

The filtering protocols initially developed and applied in ANAIS–112 have been exhaustively described in [36] and they are based on standard cuts on a few parameters. To be accepted, an event must fulfill the following criteria:

- (1) pulse shape according to NaI scintillation time constant. A bi-parametric pulse shape cut, combining the fraction of the pulse area in [100, 600] ns after the event trigger (P_1 , see Equation 3.2) and the logarithm of the amplitude-weighted mean of the arrival times corresponding to the identified peaks in the pulse in the digitized window (μ_p , see Equation 3.4) removes fast Cherenkov-like events originated in the PMTs, random coincidences between individual dark current photoelectrons and long phosphorescences in the crystal or pulse tails.
- (2) symmetrical sharing of light between the two PMTs. More than 4 peaks in each PMT signal ($n_0 > 4$ and $n_1 > 4$) are required, removing events having a strongly asymmetric light sharing, with a very low number of peaks in one of the PMTs. These asymmet-

ric events do not have the timing characteristics of afterpulses or pulse tails, and are observed below 2 keV.

- (3) single-hit events. Any event showing coincident energy depositions in more than one ANAIS-112 module is produced by the residual cosmic ray muon flux, radioactive contaminants in the set-up components or environmental radioactivity, but not by DM particles.
- (4) events arriving more than 1 s after the last muon veto trigger. As ANAIS-112 trigger is set at the photoelectron level, photons in the tail of pulses caused by the passage of a very energetic particle are able to trigger the acquisition multiple times, up to hundreds of ms after the original pulse onset. At the same time, other energy depositions associated to delayed processes following the muon interactions in the shielding could be rejected. Removing muon related events is very important for the ANAIS-112 goal, because the muon flux in deep underground laboratories has shown annual modulation [56, 57].

Although this filtering procedure works very well above 2 keV, in the region from 1 to 2 keV it shows some weaknesses. On the one hand, as commented above, the efficiency of acceptance of bulk scintillation events decreases very sharply, reducing the sensitivity, and on the other hand, the measured rate in this region is about 50% higher than expected according to the ANAIS-112 background model [58], pointing at a possible leaking of non-bulk scintillation events as responsible of that excess.

2.3 The neutron calibration program

As stressed in Section 1, a comparison between results obtained with different target nuclei in DM searches is affected by model-dependencies. Using the same target should remove these dependencies completely. For instance, the expected contribution of energy released by Migdal-electrons should be exactly the same for detectors consisting of the same nuclei [59]. There is, however, one caveat: a direct comparison between experiments using the same target is direct only in the case in which the full response function of the detectors is taken into account in such a comparison of results. This is particularly relevant in the testing of the DAMA/LIBRA result because scintillation is strongly quenched for energy deposited by nuclear recoils with respect to the same energy deposited by electrons, and experiments are usually calibrated (i.e., the response function is calculated) using gamma/electron radiation. Because of this, experimental measurements are often presented in terms of electron-equivalent energy (keV_{ee}) and experiments can be directly compared only if energy is deposited by this channel. In the case when DM particles produce nuclear recoils, quenching factors (QF) should be well known to re-calibrate the energy depositions scale.

This is not the case for NaI(Tl) scintillators. Measurements of QF in the energy region relevant for DM searches, and in particular for the comparison with DAMA/LIBRA in case the signal is attributed to nuclear recoils, do not agree, and in fact are affected by strong discrepancies (QF dependent on the energy versus constant values, for instance) [60–70].

The ANAIS-112 experiment’s design included periodical calibration using ^{109}Cd sources to monitor and correct any gain drift that could compromise the experimental goals. Stability in the long term of any experiment searching for an annual modulation in the data at very low energies is mandatory. Such ^{109}Cd calibrations are carried out every two weeks and produce lines at 11.9, 22.6 and 88.0 keV, the first of them attributed to Br-content in the source housing. Besides gain monitoring purposes, these calibrations provided a clean population

of low energy events corresponding to scintillation in the NaI(Tl) crystal which were used to determine the efficiency of the filtering applied for the 1.5, 2 and 3 years analyses [38, 71, 72].

However, in order to fully understand the response of the ANAIS–112 detectors to different particle energy depositions, a specific neutron calibration program was necessary. This program, under development, combines two different approaches: on the one hand, sodium quenching factor measurements were carried out at TUNL (Duke University, North Carolina, US) using small crystals from the same growing batches of the ANAIS–112 crystals (see [60] for preliminary results); on the other hand, neutron calibrations *onsite* have been performed, starting in 2021, using ^{252}Cf sources of low activity³ at LSC. This paper focuses on the latter, which provide a clean population of bulk scintillation events, dominated by nuclear elastic scattering, homogeneously distributed in all the crystals volume and that can be further clean-up by selecting only multiple-hit events. Because the low-energy dependence of the elastic scattering differential spectrum is exponential-like, a large number of events in the ROI is obtained in only a few calibration runs, lasting 3–4 h each. Up to the moment of writing this article, four calibration runs have been carried out, distributed along one year and a half and in different positions of the set-up (see Figure 1, right panel).

GEANT4 [73] simulations of the neutron calibration runs at the ANAIS–112 set-up, including a detailed description of all the shielding and detectors components, allow concluding that events in the ROI correspond to nuclear recoils and were able to reproduce qualitatively the measured energy spectra for single and multiple hits, shown in Figure 2, in spite of the uncertainties in the knowledge of the QF required for the comparison of simulations and measurements. In fact, we are aiming at determining the QF for our crystals by a precise quantitative comparison between simulation and measurement. It can be observed in Figure 2 that background contribution is negligible –except in the ^{210}Pb region, at about 40–50 keV in the single-hits spectrum. Moreover, contributions from electron/gamma events are only relevant at precise energies: 57 keV (inelastic scattering in ^{127}I , also observed in multiple-hits spectrum) and 31 keV (^{128}I decay after neutron capture in ^{127}I). Below 20 keV in the single-hits and 40 keV in the multiple-hits spectra, rates are dominated by multiple elastic scattering on Na and I nuclei.

Figure 3 shows scintillation pulses observed for events corresponding to neutron interactions and electron/gamma events in the ANAIS–112 modules for two energy regions: [1,6] keV (left panel) and [6,10] keV (right panel). As it can be observed, in the ROI, nuclear recoil events (solid red line) show identical temporal behaviour than electron/gamma events (dashed blue line) selected in coincidence with a high energy γ (^{22}Na and ^{40}K), whereas above 6 keV the former are slightly faster, as it was also observed in [74, 75]. Both neutron interactions and $^{22}\text{Na}/^{40}\text{K}$ coincidence events are occurring in the bulk of the crystal, homogeneously distributed, but the latter have a very low number of events. For this reason, we decided to train the BDT with nuclear recoil events in the ^{252}Cf calibration runs. Instead of using for the training events in the full ROI, from 1 to 6 keV, we decided to use only events from 1 to 2 keV, the range where our previous filtering showed to be weak.

2.4 The blank module set-up and performance

A module similar to the ANAIS–112 modules, unless because of the absence of NaI(Tl) crystal, was integrated in the ANAIS–112 set-up in August 2018. We refer to this module as blank module. Two PMTs, identical to those used in the ANAIS–112 modules, were coupled to

³The ^{252}Cf source used for the calibrations referred in the text had an activity of 3 kBq on April 2021.

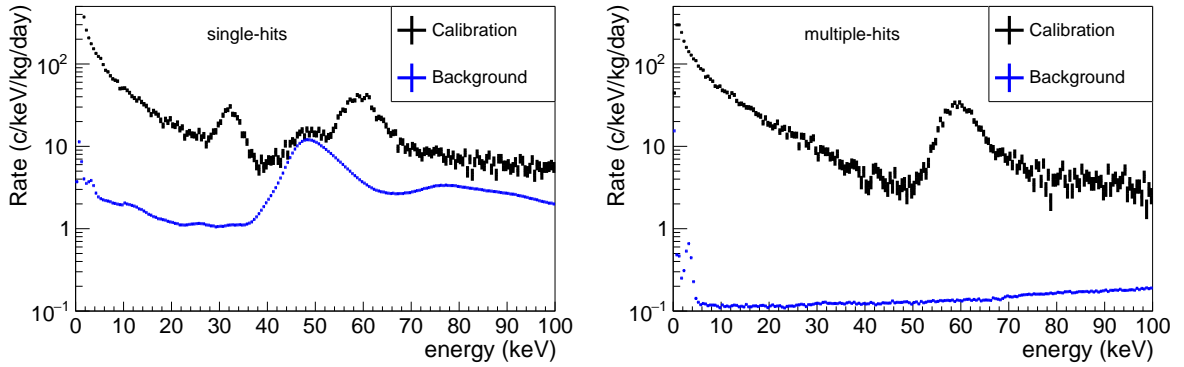


Figure 2: Energy spectrum produced at low energy in one of the ^{252}Cf calibrations (west side) of the ANAIS-112 experiment. Horizontal scale is represented in electron-equivalent energies. The background contribution is shown in blue as reference. Left panel: single-hits; right panel: multiple-hits.

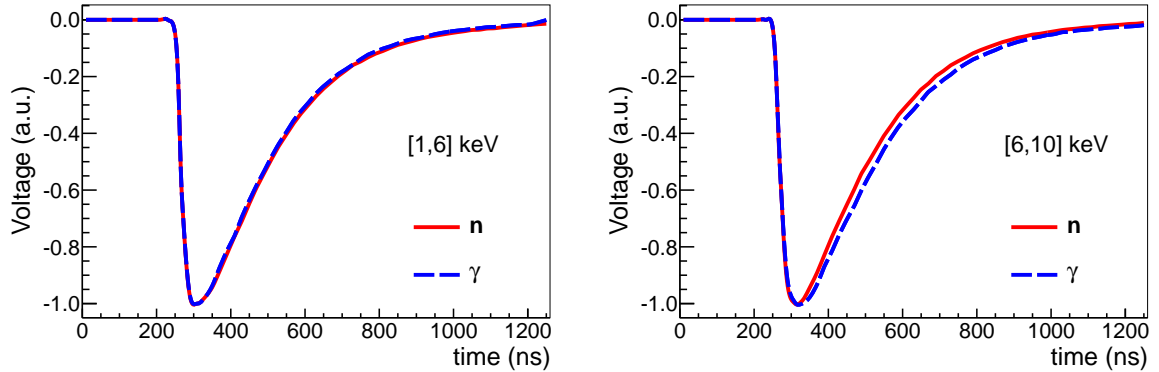


Figure 3: Average pulses corresponding to neutrons (solid red line) and gamma (dashed blue line) events for two energy regions: [1,6] keV (left panel) and [6,10] keV (right panel). Gamma events in the ROI have been selected in coincidence with a high energy γ , while gamma events outside the ROI correspond to single-hit background scintillation events.

quartz optical windows at both sides of the module, consisting of a copper housing with the interior covered by Teflon diffusor film. This module was conceived to monitor specifically non-NaI(Tl) scintillation events that were identified as one of the main drawbacks limiting the sensitivity of ANAIS-112, starting the corresponding data taking with the beginning of the second year of operation of ANAIS-112. It is installed in the ANAIS hut at LSC, inside a specific lead shielding of 10 cm thickness, besides the ANAIS-112 set-up; the readout electronics, DAQ and analysis procedures applied are the same than for the ANAIS-112 modules. Figure 4 shows the set-up of this blank module.

The module has been working uninterruptedly, accumulating a large amount of events of different origin. At the beginning of the blank module operation, the shielding was open to the laboratory air, which suffers from an average radon content of 80 Bq/m^3 but a high variability [76]. It was clearly observed that the detection rate of the blank module was sensitive to this radon content, observing a high correlation between both parameters in the period the blank module was exposed to normal air. In February 2019, the blank module



Figure 4: Setting-up the ANAIS blank module in the ANAIS hut, inside a specific lead shielding, 10 cm thick, next to ANAIS-112.

shielding was tightly closed and flushed with radon-free air from the LSC supply and the detection rate of the blank module stabilised, as it can be seen in Figure 5. The periods affected by shortcut of this radon-free air supply are clearly noticed as sharp peaks in the total trigger rate. The bias of the blank module PMTs was changed along the first year of operation in order to verify the right performance of the PMTs, and the trigger was tuned up. It can be concluded that radon in the air close to the PMTs is responsible of a relevant part of the coincident events, explaining trigger rates at the 0.5 Hz level. On a radon-free atmosphere, the trigger rate fell down to 0.1 Hz. It is worth to remark here that the ANAIS-112 modules are continuously operated in such a radon-free atmosphere. An upper limit of 0.06 Bq/m^3 at 95% C.L. to the radon content was determined by screening the same nitrogen gas used for the flushing into the ANAIS-112 shielding using an HPGe detector at LSC [76].

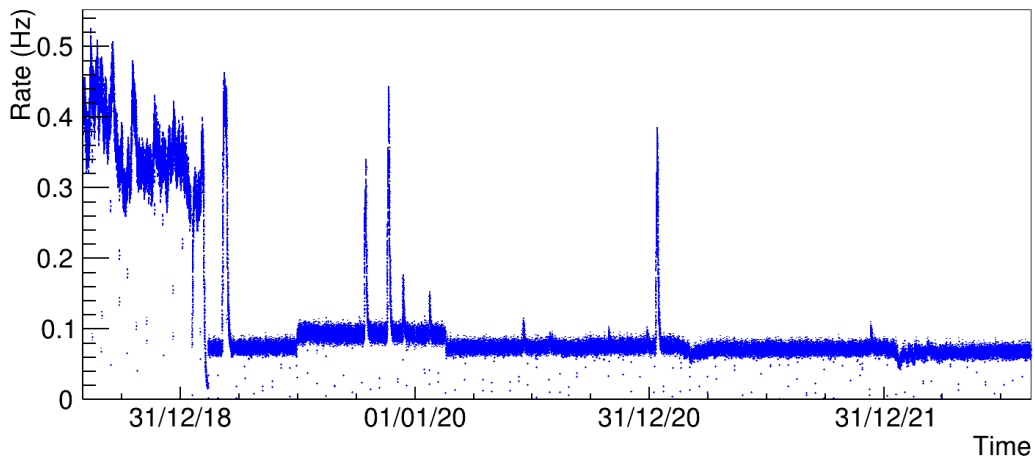


Figure 5: Trigger rate of the blank module corresponding to events with more than two peaks ($n_0 + n_1 > 2$) after filtering for electric noise by removing events with high baseline RMS. Before February 2019, the air near the module had a mean radon content of 80 Bq/m^3 . The strong decrease in rate observed corresponds to the flushing of the blank module shielding with radon-free air. The periods affected by shortcut of this radon-free air supply are clearly noticed as sharp peaks in the trigger rate.

Figure 6 shows some examples of pulses coming from the blank module. Most of the events detected in this module are similar to that shown in the left-upper panel, fast events with large amplitude, non-compatible with the NaI(Tl) scintillation shape but with Cherenkov emission in the PMTs glass. Others, like that shown in the right-upper panel are very asymmetric events; both are easily removed by the filtering designed for ANAIS-112 data and explained in Section 2.2. However, many other types of pulses have been identified that are not easy to discriminate from NaI(Tl) bulk scintillation, as those shown in the two lower panels, which were not removed by that filtering. This observation supported our goal of introducing these events in the design of a new filtering based on machine learning tools, explained in Section 3.

In order to introduce the blank module events in the ANAIS-112 analysis we had to assign them an equivalent energy, although a calibration similar to that of the ANAIS-112 modules was impossible. This is required for selecting those events that would appear in the ROI, for which the filtering should be optimized. For that goal, the number of photoelectrons in the pulse, $nphe$, is the convenient parameter. We convert the $nphe$ for blank module pulses into equivalent energy with the light collection corresponding to each ANAIS-112 module [38, 54].

3 Machine learning technique applied to ANAIS-112 data

Machine learning algorithms have shown to be powerful tools in pattern recognition and decision taking. For our goal (the classification of events as signal or noise), decision trees have shown a high performance. They have been implemented for a similar goal in COSINE-100 [49] and SABRE [77] experiments, in both cases producing higher efficiencies and better rejection than traditional cuts. Other machine learning approaches are being applied in the

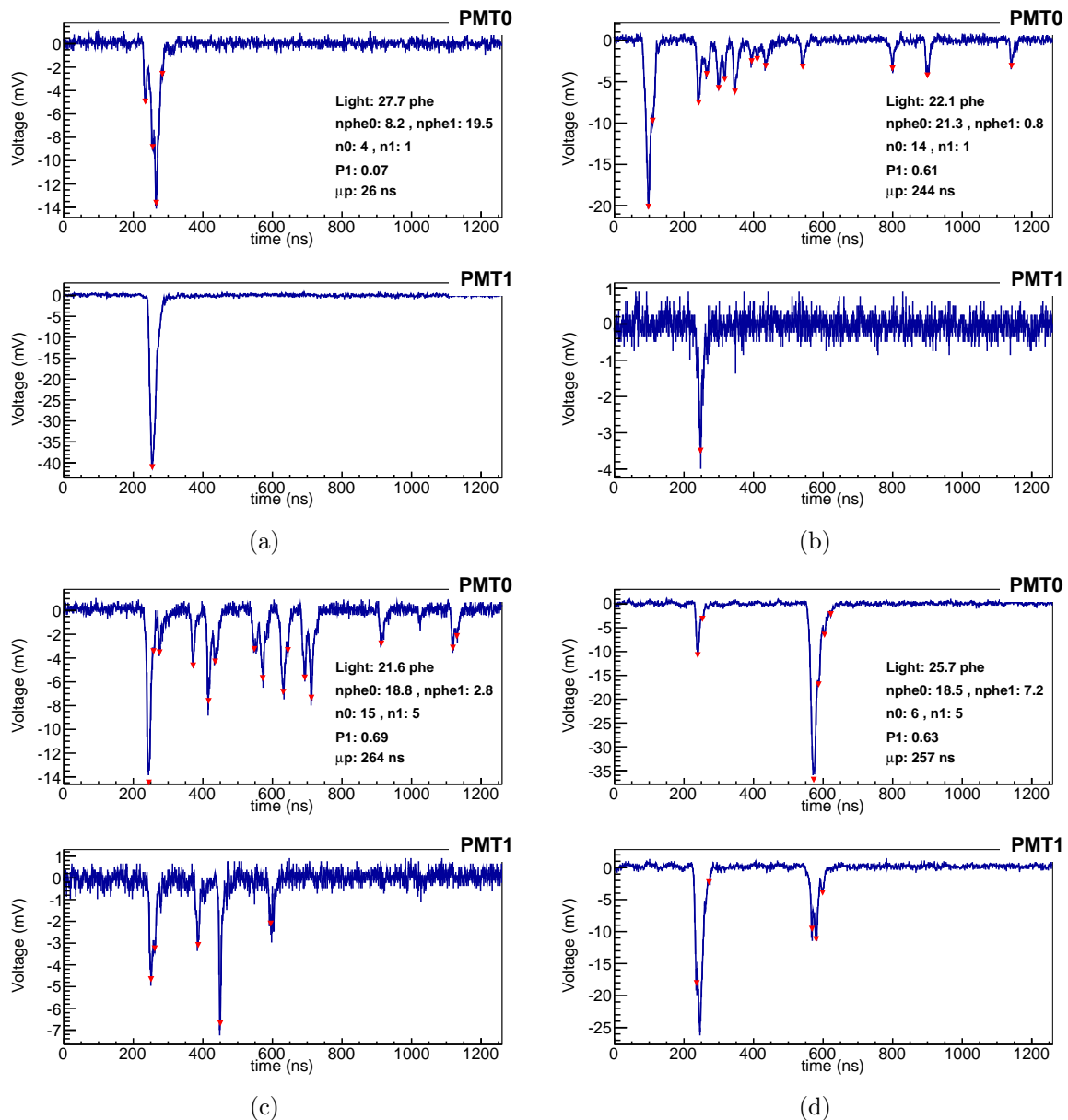


Figure 6: Examples of four events coming from the blank module. The two traces at each panel correspond to the two PMT signals, whereas the legend displays the total number of photoelectrons of the event and at every trace ($nphe_0$ and $nphe_1$), the number of peaks detected by the algorithm at every PMT waveform (n_0 and n_1), and pulse shape parameters P_1 and μ_p . Red triangles: peaks identified by the peak-searching algorithm. Most of the events detected are similar to that shown in the panel (a), fast events with large amplitude, non-compatible with the NaI(Tl) scintillation shape. Other, like that shown in the panel (b) are very asymmetric events, and are also easily removed by the filtering summarized in Section 2.2. However, many other types of pulses have been identified that are not easy to discriminate from NaI(Tl) bulk scintillation, as those shown in panels (c) and (d), which were not removed by that filtering.

direct detection of DM [51]. Unsupervised machine learning methods, in particular, could offer powerful tools to identify anomalous events or unknown systematic errors.

3.1 Boosted Decision Tree algorithm

A boosted decision tree (BDT) is a multivariate analysis technique that, taking into account the correlations between individual input parameters, allows several weak discriminating variables to be combined into a single powerful discriminator [78, 79]. Decision trees are an extension of the commonly used cut-based selection strategy in which each event is analysed by means of pass/no pass decisions taken on a single variable at a time that are iterated until a stopping criterion is achieved. In fact, most events do not have all features of either signal or noise⁴, therefore, decision trees do not reject right away events that fail a criterion, but check whether other criteria may help to classify these events properly. This helps to increase the acceptance efficiency of the filtering with respect to cut-based selection strategy.

Mathematically, decision trees have a binary structure with two classes: signal and noise. In the training process, a sequence of binary splits using the discriminating parameters, \vec{x}_i , is applied to the training events in order to reach the best separation between signal and noise. Depending on the output value of the decision tree, the event is classified as signal or noise (see Figure 7). The response of the decision tree with respect to fluctuations in the training sample can be stabilised using a boosting algorithm. The boosting of a decision tree [80] extends the aforementioned concept from one tree to several trees, forming a forest. The trees are derived from the same training ensemble by reweighting events, and are finally combined into a single classifier which is given by a weighted average of the individual decision trees. In particular, in this work we use AdaBoost (adaptive boosting) [80] to implement the boosting algorithm. Events that were misclassified during the training of a decision tree are given a higher event weight in the training of the following tree. In this way, subsequent iterations focus more on events misclassified by previous trees and a stronger classifier is built. Starting with the original event weights ($=1/N$, being N the number of training events) when training the first decision tree, the subsequent tree is trained using a modified event sample where the weights of previously misclassified events are multiplied by a common boost weight derived from the fraction of misclassified events of the previous tree. The weights of the entire event sample are then renormalised such that the sum of weights remains constant. We adopt the purity, p , of the terminal node (or *leaf*) as a separation criterion. The purity is defined as $p = \frac{s}{s+n}$, where s (n) is the sum of weights of signal (noise) events that ended up in this leaf during training. If $p > 0.5$, $+1$ is attached to the leaf and is classified as signal, otherwise the leaf is set to -1 and labelled as noise. The trees are trained iteratively until a certain number of trees have been grown. If we denote the result of an individual classifier as $T(\vec{x}_i)$, encoded for signal and noise as $T(\vec{x}_i) = +1$ and -1 , respectively, adaptive boosting yields a response defined by:

$$BDT(\vec{x}_i) = \frac{1}{n_{Trees}} \sum_{j=1}^{n_{Trees}} \ln(\alpha_j) \cdot T_j(\vec{x}_i), \quad (3.1)$$

where n_{Trees} is the number of trees in the forest, $\alpha_j = \frac{1-f_j}{f_j} \geq 1$ the j^{th} boost weight and f_j the fraction of misclassified events by the j^{th} decision tree. Small (large) values for $BDT(\vec{x}_i)$ indicate a noise-like (signal-like) event.

⁴In general, *background* is used instead of *noise*. However, in this article *noise* has been preferred to avoid confusion with radioactive background.

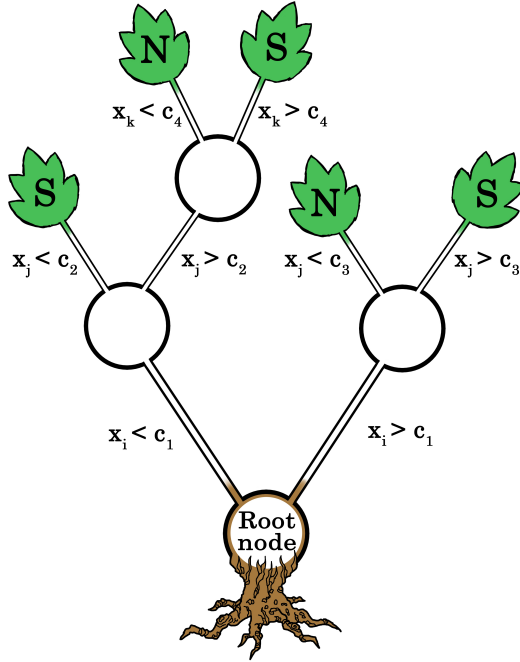


Figure 7: Schematic diagram of a decision tree with the architecture used in our analysis. Starting from the root node, a sequence of binary splits using the discriminating parameters \vec{x}_i is applied to the data. Each split uses the parameter that at this node gives the best separation between signal and noise when being cut on (cut c_p). The same parameter may thus be used at several nodes, while others might not be used at all. The leaf nodes at the top end of the tree are labelled “S” for signal and “N” for noise depending on the majority of events that end up in the respective nodes.

For the implementation of the BDT algorithm in ANAIS–112, we use the Toolkit for Multivariate Data Analysis (TMVA) package [81] integrated in ROOT.

3.2 Training populations

As commented before, training populations for signal and noise are obtained from the neutron calibrations using ^{252}Cf carried out in the last operation years of the ANAIS–112 experiment and blank module events, respectively. The most challenging aspect of a BDT training is to obtain pure event samples that are used to model the signal and noise events. For this reason, a series of preselection cuts on the training populations have been defined to extract representative samples of signal and noise.

The neutron training population consists of events meeting the following requirements:

- Energy in the 1-2 keV range.
- Slow charge (P_1) parameter larger than 0.35 to remove PMT-origin fast pulses that are found in the neutron calibrations and could be, at least partially, related with Cherenkov emissions in the PMTs.
- Trigger time lower than 400 ns from the main acquisition trigger. The software trigger position of the ANAIS–112 triggered events is around 250 ns within the 1260 ns digitization time window. Nevertheless, due to the high multiplicity of neutron interactions,

the moderation of the neutron energy and the propagation time of the neutron between detectors, some multiple events show a delayed software trigger position, resulting in the tail of the pulse not being completely recorded. The estimates for the energy and the other pulse shape related parameters of such events are not correct and, therefore, they are discarded.

For the blank module training events, the criteria are:

- Number of photoelectrons in the 10-28 range, corresponding to an equivalent energy of 1-2 keV.
- Events having $P_1 < 0.35$ are forced to be at 20% of the total training population, although this kind of events are dominant in the detection rate of the blank module. However, they are easy to filter, whereas other events populations identified in the blank module require a training boost.
- Root mean square (RMS) of the baseline level below 0.60 mV in each PMT trace. Some of the blank module events correspond to electrical noise which can be removed based on the RMS of the baseline.

Similar size training populations have been selected for signal and noise (above $3 \cdot 10^4$ events each). Moreover, only 70% of the events of each population has been used for the training. The other 30% has been used for testing in order to check the presence (if any) of overtraining (see Section 5).

Although ideally the training populations should be naturally clean, avoiding preselection cuts, to fully profit from the machine learning analysis potentiality, this is very difficult to achieve from the experimental point of view. The requirements imposed above for the training populations' selection correspond to ROI energy matching, signal quality cuts, and boosting of the noise population more difficult to reject, which do not imply any bias of the training population. Only the cut on the P_1 parameter on the neutron population is required to effectively clean it up from noise events because, even during a neutron calibration, events having PMT origin should be found besides the bulk scintillation events.

3.3 Input parameters

Once the training populations for the scintillation and noise events have been determined, we define the input parameters for the BDT. Some of the discriminating parameters used in the BDT training are based on those best performing in the previous filtering protocol developed for ANAIS-112 and explained in Section 2.2 (P_1 , μ_p , n_0 , n_1), but new parameters have also been added (P_2 , CAP , $Asynphe$). These BDT training parameters are defined as:

- (1) Slow charge (P_1): the fractional charge in [100, 600] ns window with respect to total charge in [0, 600] ns window for the total pulse,

$$P_1 = \frac{\sum_{100 \text{ ns}}^{600 \text{ ns}} A(t)}{\sum_{0 \text{ ns}}^{600 \text{ ns}} A(t)}, \quad (3.2)$$

where $A(t)$ is the pulse amplitude at time t after the software trigger position.

- (2) Fast charge (P_2): the fractional charge in $[0, 50]$ ns window with respect to total charge in $[0, 600]$ ns window for the total pulse,

$$P_2 = \frac{\sum_{0 \text{ ns}}^{50 \text{ ns}} A(t)}{\sum_{0 \text{ ns}}^{600 \text{ ns}} A(t)}. \quad (3.3)$$

- (3) Mean time (μ_p): amplitude-weighted mean time of the arrival times of the peaks identified in the digitized window for the summed pulse,

$$\mu_p = \frac{\sum_i A_i t_i}{\sum_i A_i}, \quad (3.4)$$

being A_i and t_i the amplitude and time of the i^{th} peak identified in the pulse trace, respectively.

- (4) Charge accumulated pulse (CAP): integral of the cumulative distribution of the total pulse charge up to a certain time, x , after the software trigger position,

$$CAP_x = \frac{\sum_{0 \text{ ns}}^x A(t)}{\sum_{0 \text{ ns}}^{t_{max}} A(t)}, \quad (3.5)$$

where t_{max} is the end of the pulse. In particular, we use $x = 50, 100, 200, 300, 400, 500, 600, 700$ and 800 ns.

- (5) Asymmetry in the number of photoelectrons ($Asynphe$): the quotient between the partition in number of photoelectrons among both PMTs and the total number of photoelectrons. That is:

$$Asynphe = \frac{nphe_0 - nphe_1}{nphe_0 + nphe_1}, \quad (3.6)$$

where $nphe_0$ ($nphe_1$) is the number of photoelectrons in the pulse from PMT0 (PMT1).

- (6) Number of peaks identified by the peak-searching algorithm in PMT0 and PMT1 traces (n_0 and n_1 , respectively).

Figure 8 shows the distribution of three representative BDT training parameters for the neutron population (upper panels) and the background data of the $\sim 10\%$ unblinded from the first year (lower panels) in the 1 to 10 keV energy region, and for events coming from the blank module with a number of photoelectrons equivalent to that energy (middle panels). The parameters shown are P_1 (a), $\text{Log}(\mu_p[\mu s])$ (b) and $Asynphe$ (c). For instance, considering the P_1 parameter, we can observe that the neutron population is slightly polluted by fast events with $P_1 < 0.35$ (especially at low energy) and, on the contrary, this kind of events are dominant in the detection rate of the blank module, as commented in Section 3.2.

3.4 Training and BDT response

Using the training populations and input parameters described above, we have carried out a 200-trees⁵ training to develop a BDT score which can discriminate scintillation signal from

⁵The specific number of trees has been previously determined by scanning the value that minimizes the background level between 1 and 2 keV for the same training population. For other configurable parameters such as the maximum depth of the decision tree or the learning rate, the default values (3 and 0.5, respectively) have resulted in good performance.

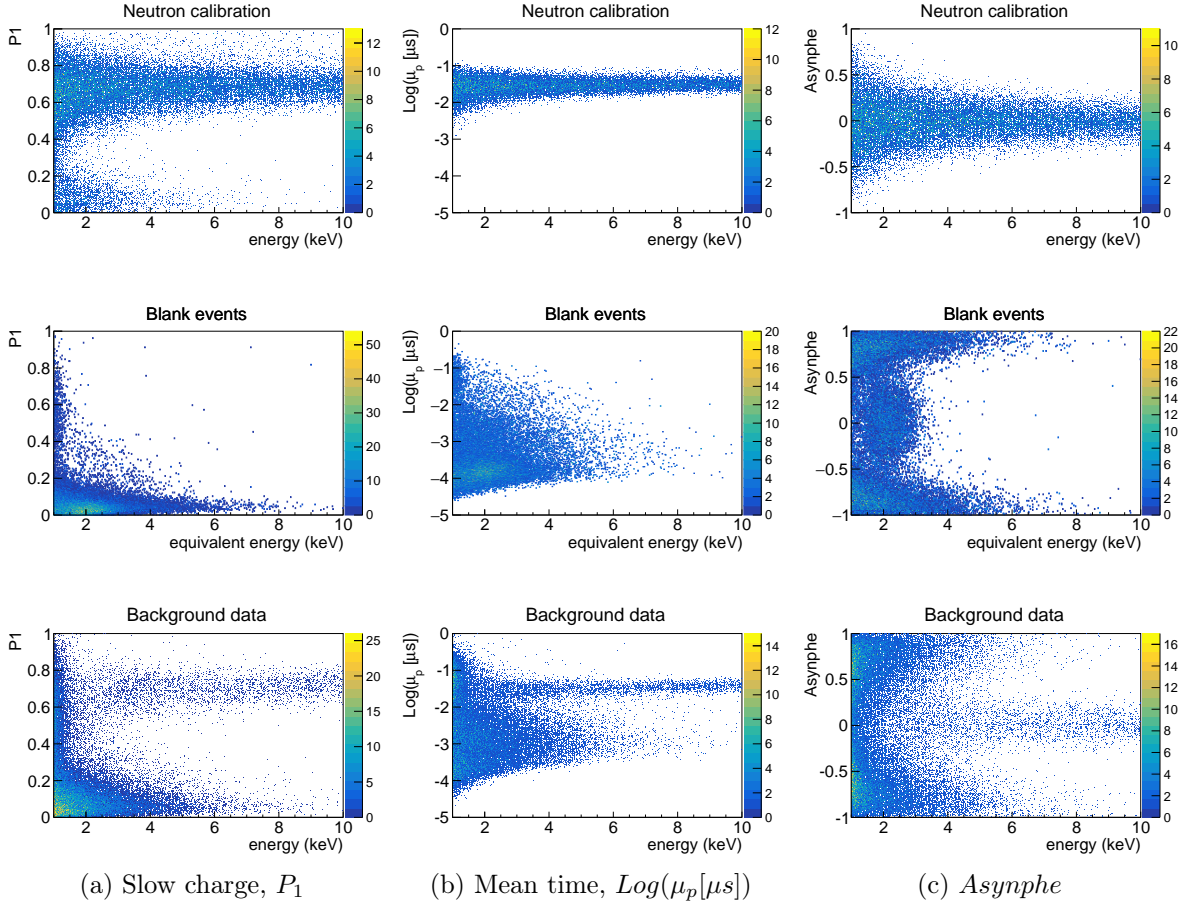


Figure 8: Distribution of three representative input parameters as a function of energy from 1 to 10 keV: (a) Slow charge, P_1 , (b) Mean time, $\text{Log}(\mu_p[\mu s])$ and (c) $Asynphe$. The upper panels show the distribution of events from neutron calibrations, whereas the lower panels display events from the background data. The same for events coming from the blank module with an equivalent energy is shown in the middle panels.

PMT-noise more efficiently. All crystals are trained at the same time on a single BDT. Table 1 lists the separation signal/noise of the input parameters before training (2^{nd} column) and the parameter ranking after training (3^{rd} column). The ranking of the BDT input parameters is derived by counting how often the parameters are used to split decision tree nodes, weighting each split occurrence by the separation gain-squared it has achieved, and by the number of events in the node [81]. The top parameter in the table is the best ranked. In particular, the parameters describing the asymmetry in the light sharing among both PMTs have the largest discriminating power (n_0 , n_1 , and $Asynphe$), followed by the new parameters defined on the integral of the total pulse. The other parameters used in the ANAIS-112 standard analysis (P_1 and μ_p) have a discriminating power similar to the integrals of the pulse, but it should be noted that the discriminating capability of P_1 has been reduced by forcing the population of noise events coming from the blank module to be only lightly polluted by fast events (20%, as explained in Section 3.2).

As a result of the training, we obtain the new BDT parameter that allows to better discriminate noise from scintillation events. Figure 9 shows the BDT distribution for ^{252}Cf

Parameter	Parameter separation before training (%)	Parameter ranking after training (%)
n_0	40.10	11.40
Asynphe	48.40	10.00
n_1	33.40	9.40
CAP ₅₀	21.30	7.90
CAP ₈₀₀	24.50	7.80
μ_p	33.00	7.40
CAP ₇₀₀	25.20	6.50
P ₁	21.50	6.20
CAP ₄₀₀	23.90	6.00
CAP ₆₀₀	25.10	5.80
CAP ₂₀₀	16.30	5.60
CAP ₁₀₀	12.90	5.50
CAP ₅₀₀	24.80	5.30
CAP ₃₀₀	22.00	4.90
P ₂	1.70	0.30

Table 1: Signal/noise separation of input parameters before training and parameter ranking after training. The top parameter is the best ranked.

neutron calibration events (in orange) and background events of the $\sim 10\%$ unblinded data (in blue) in the [1,2] keV energy region, and for events coming from the blank module with a number of photoelectrons equivalent to that energy (in black). We can observe that most of the events from the blank module are satisfactorily distributed around negative values of the BDT parameter and are negligible above 0.5. On the contrary, the events corresponding to neutron calibrations are distributed around positive values of BDT ~ 0.60 , although they present a tail of events that extends down to -0.5 . The distribution of background events combines features of both populations, being dominated by noise-like events between 1 and 2 keV.

4 Bulk scintillation events selection

Figure 10 displays the BDT response as a function of energy for the nine ANAIS-112 modules both for the neutron calibration sample (a) and the $\sim 10\%$ unblinded background (b). We can observe a clear separation between scintillation and fast PMT-noise events for energies greater than 2 keV, but below this energy the populations overlap. For the event selection, we define an energy-dependent BDT cut (solid red lines in Figure 10) and keep only events above it. This selection has been optimized for each detector and energy bin to ensure the lowest efficiency-corrected background from the $\sim 10\%$ unblinded data of the first year of data taking, guaranteeing acceptance efficiencies greater than those obtained with the current ANAIS-112 filtering procedures. The corresponding efficiency is estimated for each detector independently by using ^{252}Cf neutron calibration events. The ratio of the events which pass

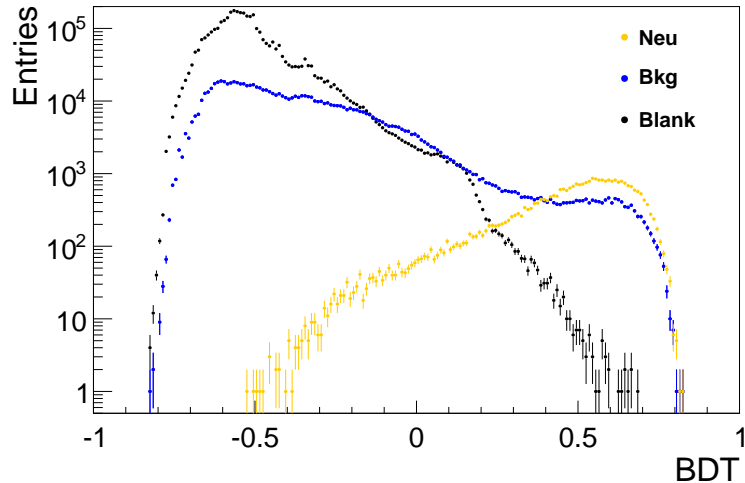


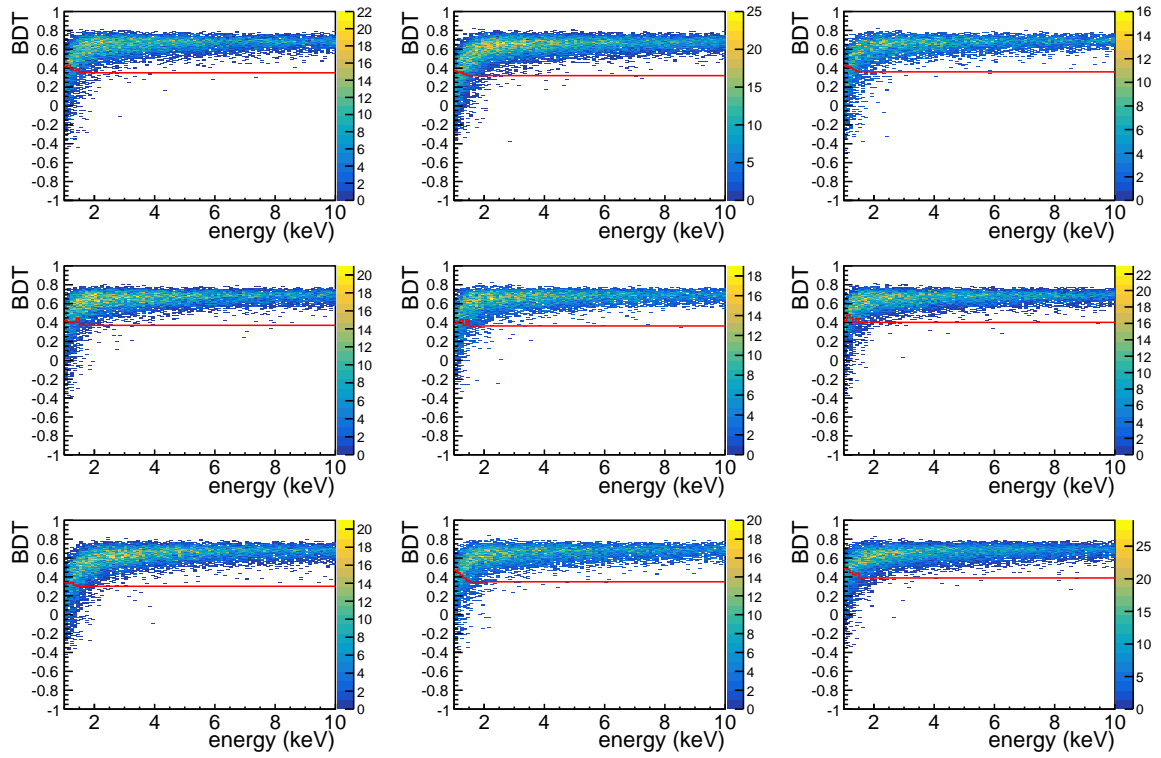
Figure 9: BDT distribution for neutron calibration events (in orange) and background events of the $\sim 10\%$ unblinded data (in blue) in the $[1,2]$ keV energy region. The same for events coming from the blank module with an equivalent energy is shown in black.

the signal selection to the total events is the acceptance efficiency⁶. The average efficiency of the BDT cut is shown in orange in Figure 11, while that obtained with the ANAIS–112 filtering protocols is displayed in black for comparison. We can observe that the acceptance efficiency derived from the BDT cut is significantly higher (around 30% in $[1,2]$ keV, see Table 2) with respect to the previous ANAIS–112 filtering procedure. Furthermore, we use two additional populations to cross-check the consistency of the selection criteria and efficiency estimation: ^{109}Cd calibration single-hit events (in red), and the $^{22}\text{Na}/^{40}\text{K}$ low energy events selected in coincidence with a high energy γ (in blue). It has to be remarked that the latter efficiency has only been calculated around the 0.87 and 3.2 keV peaks from ^{22}Na and ^{40}K , respectively, to avoid spurious coincident events, and that this efficiency estimate has a higher uncertainty due to the low number of ^{22}Na and ^{40}K coincidence events. It can be observed that all of them are well compatible.

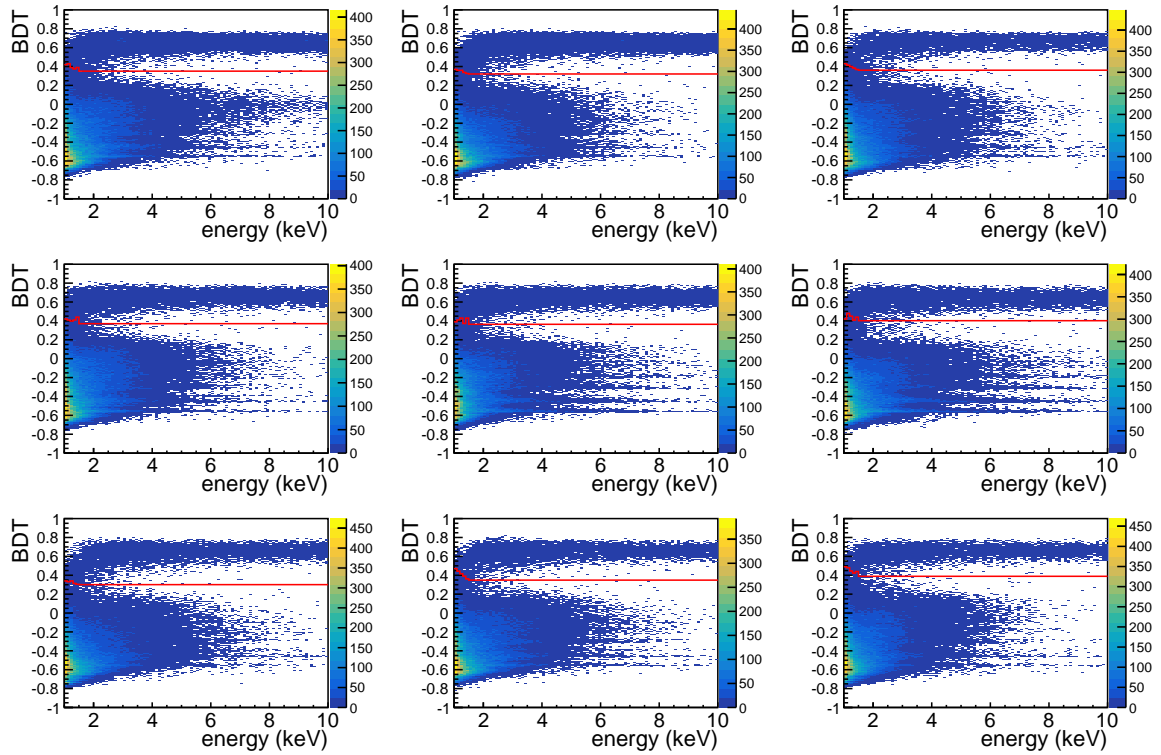
In Figure 12, the efficiency derived from the BDT cut using ^{252}Cf neutron calibration events is shown in orange for each ANAIS–112 module. These efficiencies have clearly improved below 3 keV in all the modules with respect to those obtained with the previous ANAIS–112 filtering protocols (in black).

The anticoincidence energy spectra in the ROI for the nine ANAIS–112 detectors after BDT event selection and efficiency correction are shown in orange in Figure 13. The background level at 2 keV ranges from 2 to 5 c/keV/kg/d, depending on the detector, and then increases up to 2–9 c/keV/kg/d at 1 keV. For comparison, the anticoincidence spectra using the previous ANAIS–112 filtering procedure are depicted in black in the figure. We can observe that the BDT method developed significantly reduces the background level below 2 keV for all detectors with respect to that obtained by the current ANAIS–112 procedure. In particular, the integral rate from 1 to 2 keV is 5.78 ± 0.06 and 4.69 ± 0.05 c/keV/kg/d for the ANAIS–112 filtering procedure and the BDT method, respectively, which represents a reduction of the background of 19% (see Table 2).

⁶Following [36], the efficiency estimated with the BDT cut has been multiplied by the trigger efficiency derived from the size of the coincidence window between both PMTs (200 ns).



(a)



(b)

Figure 10: BDT response as a function of energy for the nine modules for the neutron calibration sample (a) and for the $\sim 10\%$ unblinded dark matter search data (b). Solid red line: energy-dependent event selection cut.

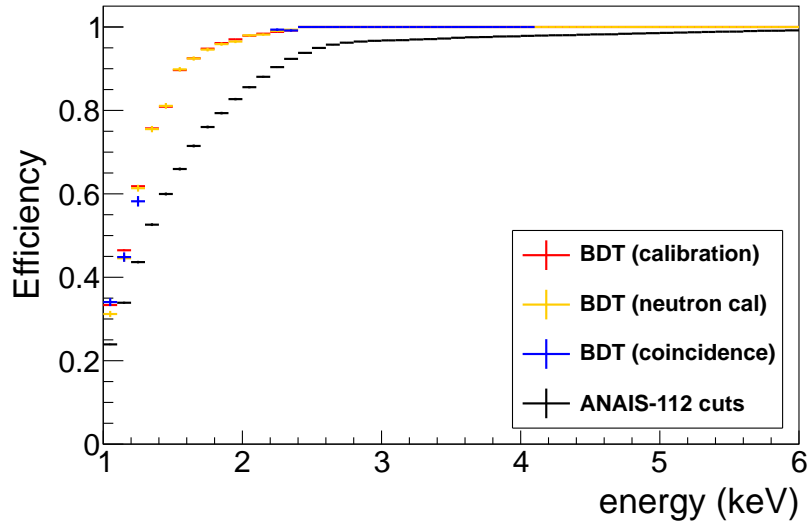


Figure 11: Average total efficiency obtained using the BDT cut estimated from neutron calibration (in orange), ^{109}Cd calibration (in red) and $^{22}\text{Na}/^{40}\text{K}$ coincidences (in blue). Efficiency with the previous ANAIS-112 filtering protocols is shown in black for comparison.

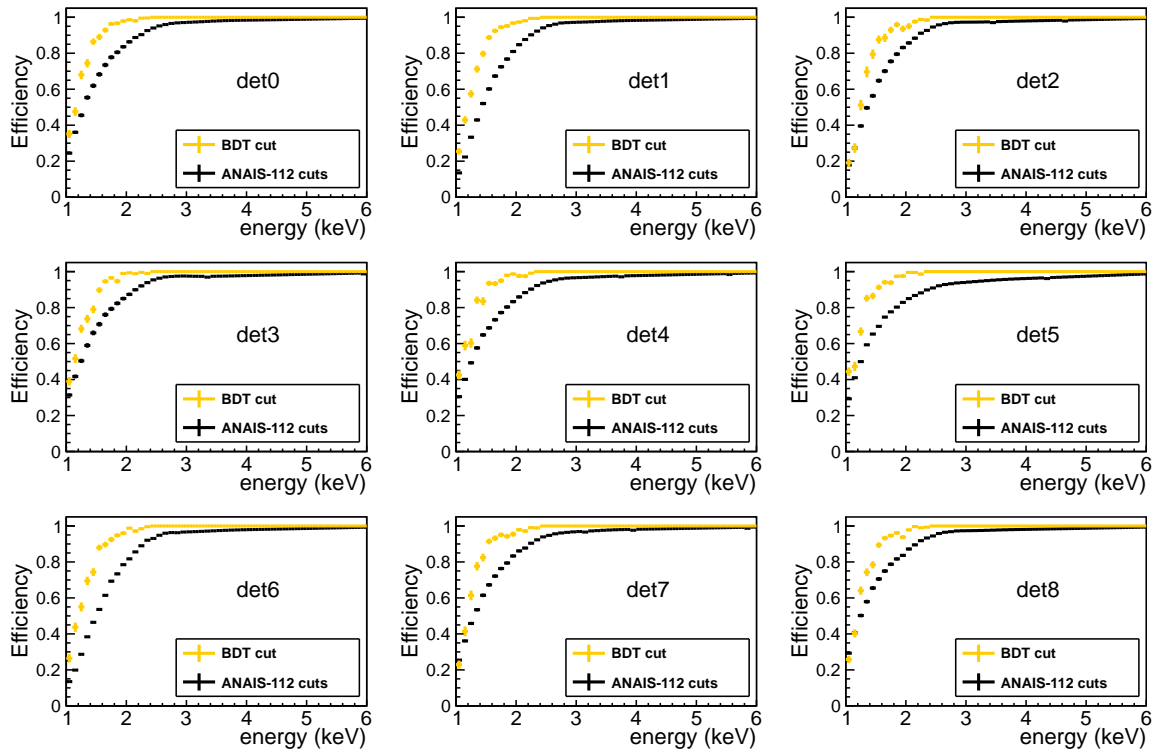


Figure 12: Total efficiency for each ANAIS-112 detector obtained using the BDT cut estimated from neutron calibrations (in orange). The efficiency with the previous ANAIS-112 protocols is shown in black for comparison.

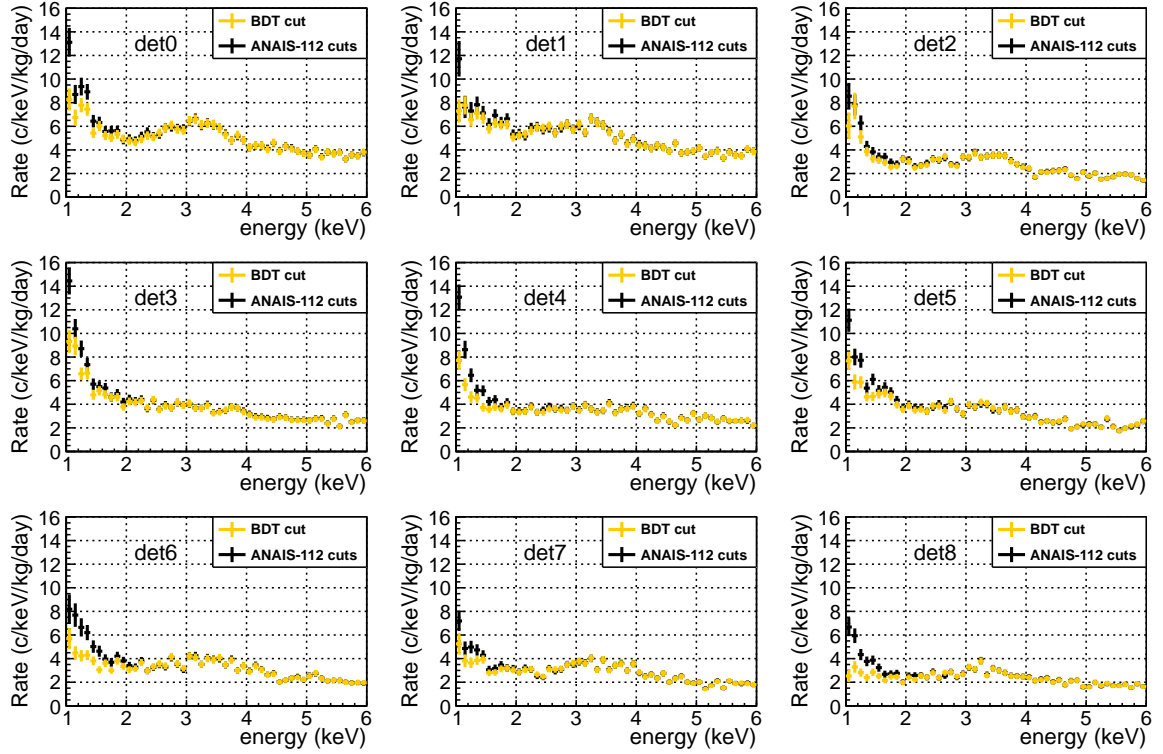


Figure 13: Anticoincidence energy spectrum measured in the ROI for each detector after BDT event selection and efficiency correction estimated from neutron calibration (in orange). In black, the same but using the previous ANAIS–112 filtering procedure. Data correspond to the $\sim 10\%$ data unblinded for the first year of operation.

Detector	Measured rate			Detection efficiency		
	ANAIS–112 cuts (c/keV/kg/d)	BDT cut (c/keV/kg/d)	Variation (%)	ANAIS–112 cuts (%)	BDT cut (%)	Variation (%)
D0	7.45 ± 0.21	6.22 ± 0.17	-17	60.7 ± 0.3	78.4 ± 0.6	29
D1	7.30 ± 0.24	6.48 ± 0.18	-11	52.2 ± 0.3	74.4 ± 0.5	43
D2	4.66 ± 0.18	4.04 ± 0.19	-13	56.4 ± 0.3	70.5 ± 0.7	25
D3	7.12 ± 0.20	5.92 ± 0.17	-17	64.2 ± 0.3	78.6 ± 0.6	22
D4	5.84 ± 0.18	4.45 ± 0.14	-24	62.5 ± 0.3	80.8 ± 0.6	29
D5	6.20 ± 0.18	5.07 ± 0.15	-18	63.0 ± 0.2	80.4 ± 0.6	28
D6	5.40 ± 0.21	3.93 ± 0.15	-27	48.3 ± 0.2	73.0 ± 0.6	51
D7	4.21 ± 0.15	3.55 ± 0.14	-16	60.1 ± 0.3	75.5 ± 0.6	26
D8	3.80 ± 0.14	2.55 ± 0.11	-33	63.3 ± 0.3	75.0 ± 0.6	18
ANAIS–112	5.78 ± 0.06	4.69 ± 0.05	-19	59.0 ± 0.1	76.3 ± 0.2	29

Table 2: Measured rates after BDT filtering and efficiency correction in the [1,2] keV energy region for each ANAIS–112 detector and on average, from the $\sim 10\%$ of unblinded data of the first year of data taking. The corresponding measured rates derived from the previous ANAIS–112 cuts are also presented for comparison in addition to the estimated percentage of variation. The last columns show the same for detection efficiency.

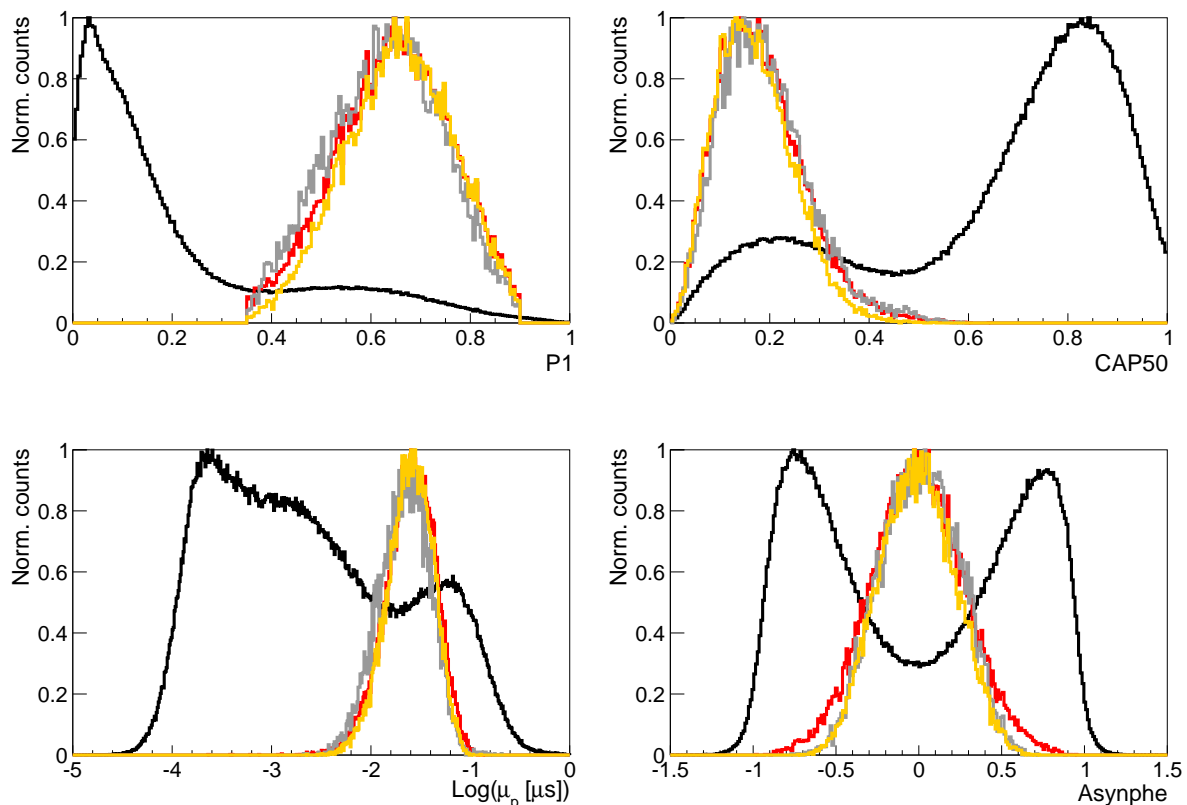


Figure 14: Distribution of some representative input parameters used to validate the BDT output response. The black and red lines correspond to physics-run and ^{252}Cf neutron calibration data before the BDT selection, respectively. After applying the BDT criterion, scintillating-like events from the physics-run data are shown as gray line, whereas selected ^{252}Cf neutron calibration events are shown in orange.

5 BDT output validation

The BDT method provides a good event separation between scintillation and noise events, however, the validation of the algorithm becomes mandatory to ensure that neutron calibration events used in the training have the same behaviour as events corresponding to the background data. In order to do this, we compare the distribution of the input parameters for the neutron calibration training sample and for the events selected from the independent dark matter search data. Figure 14 displays the validation of some representative input parameters used to build the ANAIS-112 BDT. The black line corresponds to the raw background, whereas the red line is the preselected ^{252}Cf neutron calibration data. After the BDT event selection, scintillation-like events from the physics-run (^{252}Cf neutron calibration) data are shown as gray (orange) lines. We can observe that there exists a good agreement in the shape of distributions for parameters between ^{252}Cf neutron calibration events and the background selected scintillation data. Similar results are obtained for the other parameters. The consistency between the two independent samples provides an indirect validation of the procedure.

We have also checked the possibility of overtraining in our BDT by comparing the

efficiency curve obtained (for each detector), as shown in Figure 12, with those obtained by applying the training to the test population (see Section 3.2). Since the efficiency curves are compatible with each other, we can conclude that the BDT is not overtrained.

6 ANAIS–112 sensitivity improvement with BDT filtering

The ANAIS–112 sensitivity is determined by the standard deviation of the modulation amplitude estimator, $\sigma(\hat{S}_m)$ [82]. Then, we quote our sensitivity to the annual modulation signal measured by the DAMA/LIBRA collaboration at a given exposure as the ratio between the modulation amplitude measured by DAMA/LIBRA in the corresponding energy region (0.0105 ± 0.0011 and 0.0102 ± 0.0008 c/keV/kg/d in [1,6] and [2,6] keV, respectively [12]) and $\sigma(\hat{S}_m)$, $\mathcal{S}(\hat{S}_m, t) = S_m^{\text{DAMA}} / \sigma(\hat{S}_m)$, because it gives directly (in σ units) the compatibility of the DAMA/LIBRA signal with the null hypothesis ($S_m = 0$) in ANAIS–112. To estimate the ANAIS–112 sensitivity, we exploit the Monte Carlo technique described in [52]. The expected number of events for every time bin t_i and detector d can be written (following [38]) as

$$\mu_{i,d} = [R_{0,d}(1 + f_d \phi_{bkg,d}^{MC}(t_i)) + S_m \cos(\omega(t_i - t_0))] M_d \Delta E \Delta t, \quad (6.1)$$

where M_d is the mass of every module, $\phi_{bkg,d}^{MC}$ is the PDF sampled from the MC background evolution in time calculated independently for every module, and $R_{0,d}$ and f_d are considered free parameters. Then, assuming $S_m=0$, for each ten-day time bin i and detector d , a Gaussian distribution of mean $\mu_{i,d}$ and standard deviation $\sigma_{i,d}$ is sampled for the $R_{0,d}$ and f_d values derived from the ANAIS–112 three-year exposure [38]. In this calculation, we correct the constant term (nonvarying rate), $R_{0,d}$, by the percentage of reduction attained with the BDT method (see Table 2). The mean value $\mu_{i,d}$ in the i^{th} time bin and detector d is computed by evaluating Equation 6.1 at the centre of the time bin t_i , whereas the standard deviation $\sigma_{i,d}$ is obtained by multiplying the mean value $\mu_{i,d}$ by the expected relative error for the rate of events at t_i . The relative error in the rate of events has been estimated assuming that the live time distribution of ANAIS–112 is homogeneous over time, as well as the detection efficiency in [1,6] and [2,6] keV, and considering the background evolution modelled by $\phi_{bkg,d}^{MC}$. In this way, the standard deviation of the modulation amplitude in the i^{th} time bin is estimated by fitting the simulated rates of events up to such time bin to Equation 6.1.

Figure 15 (dark blue lines) shows our sensitivity projection for the two studied energy ranges, whereas the cyan bands represent the 68% uncertainty in S_m^{DAMA} as reported in [12]. Figure 16 shows the improvement of the new sensitivity estimates using the BDT method with respect to the previous ones [38]. According to our new sensitivity estimates, the ANAIS–112 experiment would test the DAMA/LIBRA annual modulation result beyond 3σ sensitivity by applying this new BDT filtering to the three-year exposure, being possible to achieve 5σ sensitivity by extending the data taking for 2–3 more years than the scheduled 5 years which were due in August 2022.

7 Conclusions

A new low-energy noise event filtering protocol based on the boosted decision tree technique has been developed for the ANAIS–112 experiment. For the BDT training, we have used genuine nuclear recoils from dedicated onsite neutron calibrations as signal events and events coming from a blank module (i.e., identical to the ANAIS–112 ones, but without scintillating NaI(Tl) crystal) as noise sample, without requiring to use background events at all.

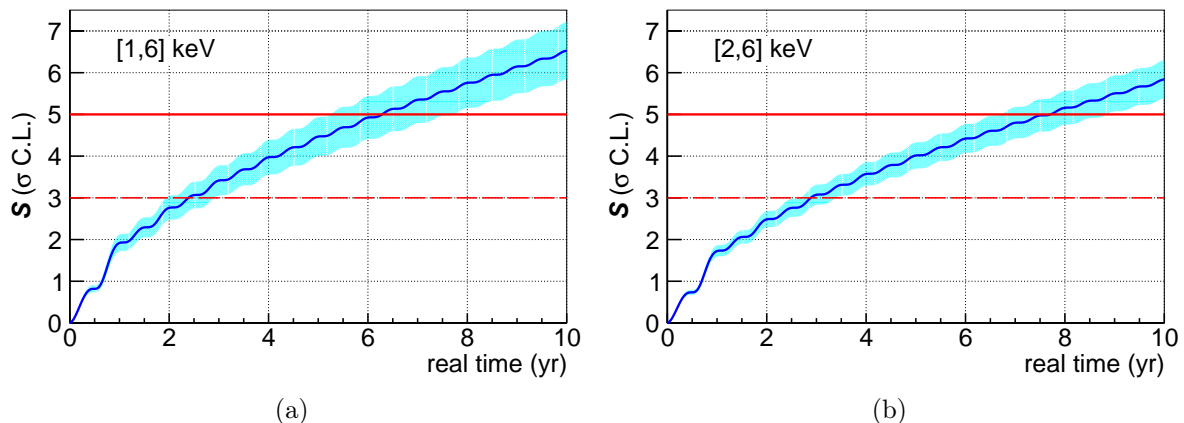


Figure 15: ANAIS-112 sensitivity to the DAMA/LIBRA signal in σ C.L. units as a function of real time in the [1,6] keV (a) and [2,6] keV (b) energy regions considering decreasing background after applying the BDT event selection. The cyan bands represent the 68% C.L. DAMA/LIBRA uncertainty.

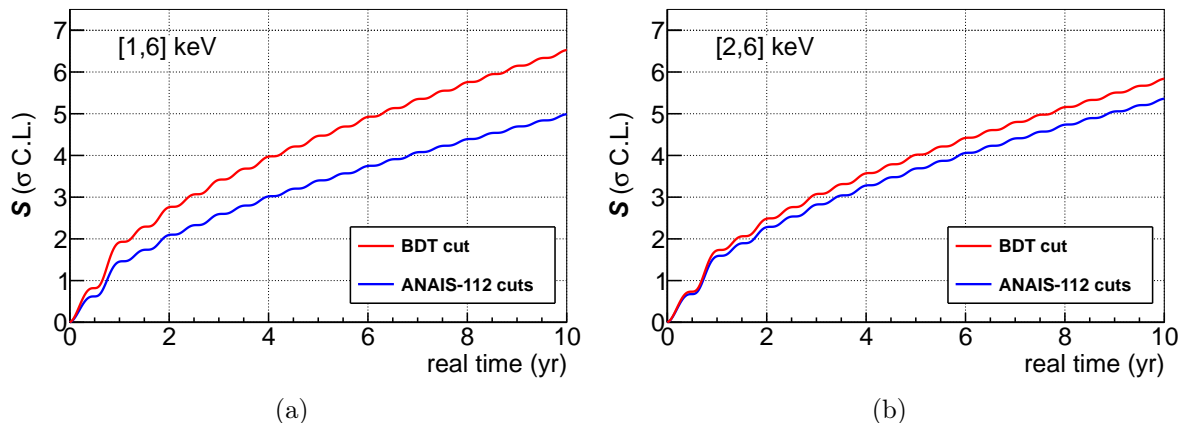


Figure 16: ANAIS-112 sensitivity to the DAMA/LIBRA signal in σ C.L. units as a function of real time in the [1,6] keV (a) and [2,6] keV (b) energy regions considering decreasing background. The blue lines show our sensitivity projection derived from applying the previous ANAIS-112 filtering procedure, whereas the red lines display the expected sensitivity from applying the BDT method, as shown in Figure 15.

With this filtering procedure, a background level reduction of around 20% has been achieved between 1 and 2 keV. However, the excess of events below 2 keV with respect to our background model is still present, pointing at a background contribution in this energy range unaccounted for in the model and supporting the convenience of the revision of the ANAIS-112 background model.

Moreover, this new filtering increases significantly the efficiency of the event selection, pushing the ANAIS-112 sensitivity to test the DAMA/LIBRA annual modulation result beyond 3σ with three-year exposure. A reanalysis of annual modulation in the first 3 years of ANAIS-112 [38] with the new filtering technique is ongoing and will appear soon. According to our sensitivity projections, ANAIS-112 is able to reach 5σ sensitivity to DAMA/LIBRA

signal by extending the data taking for a few more years than the scheduled 5 years, which were due in August 2022.

Acknowledgments

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

References

- [1] PARTICLE DATA GROUP collaboration, *Review of Particle Physics*, *PTEP* **2022** (2022) 083C01.
- [2] G. Bertone and D. Hooper, *History of dark matter*, *Rev. Mod. Phys.* **90** (2018) 045002 [[1605.04909](#)].
- [3] O. Buchmueller, C. Doglioni and L.T. Wang, *Search for dark matter at colliders*, *Nature Phys.* **13** (2017) 217 [[1912.12739](#)].
- [4] J. Conrad and O. Reimer, *Indirect dark matter searches in gamma and cosmic rays*, *Nature Phys.* **13** (2017) 224 [[1705.11165](#)].
- [5] J. Liu, X. Chen and X. Ji, *Current status of direct dark matter detection experiments*, *Nature Phys.* **13** (2017) 212 [[1709.00688](#)].
- [6] J. Billard et al., *Direct detection of dark matter—APPEC committee report**, *Rept. Prog. Phys.* **85** (2022) 056201 [[2104.07634](#)].
- [7] M. Schumann, *Direct detection of WIMP dark matter: concepts and status*, *Journal of Physics G: Nuclear and Particle Physics* **46** (2019) 103003 [[1903.03026](#)].
- [8] DAMA collaboration, *Search for WIMP annual modulation signature: Results from DAMA/NaI-3 and DAMA/NaI-4 and the global combined analysis*, *Phys. Lett. B* **480** (2000) 23.
- [9] DAMA collaboration, *First results from DAMA/LIBRA and the combined results with DAMA/NaI*, *Eur. Phys. J. C* **56** (2008) 333 [[0804.2741](#)].
- [10] DAMA/LIBRA collaboration, *New results from DAMA/LIBRA*, *Eur. Phys. J. C* **67** (2010) 39 [[1002.1028](#)].
- [11] R. Bernabei et al., *Final model independent result of DAMA/LIBRA-phase1*, *Eur. Phys. J. C* **73** (2013) 2648 [[1308.5109](#)].
- [12] R. Bernabei et al., *First model independent results from DAMA/LIBRA-phase2*, *Nucl. Phys. Atom. Energy* **19** (2018) 307 [[1805.10486](#)].
- [13] R. Bernabei et al., *The DAMA project: Achievements, implications and perspectives*, *Prog. Part. Nucl. Phys.* **114** (2020) 103810.
- [14] R. Bernabei et al., *Further results from DAMA/LIBRA-phase2 and perspectives*, *Nucl. Phys. Atom. Energy* **22** (2021) 329.
- [15] A.K. Drukier, K. Freese and D.N. Spergel, *Detecting Cold Dark Matter Candidates*, *Phys. Rev. D* **33** (1986) 3495.
- [16] K. Freese, J.A. Frieman and A. Gould, *Signal Modulation in Cold Dark Matter Detection*, *Phys. Rev. D* **37** (1988) 3388.

- [17] XENON collaboration, *Search for Electronic Recoil Event Rate Modulation with 4 Years of XENON100 Data*, *Phys. Rev. Lett.* **118** (2017) 101101 [[1701.00769](#)].
- [18] XENON collaboration, *Dark Matter Search Results from a One Ton-Year Exposure of XENON1T*, *Phys. Rev. Lett.* **121** (2018) 111302 [[1805.12562](#)].
- [19] XENON collaboration, *Light Dark Matter Search with Ionization Signals in XENON1T*, *Phys. Rev. Lett.* **123** (2019) 251801 [[1907.11485](#)].
- [20] LUX collaboration, *Results from a search for dark matter in the complete LUX exposure*, *Phys. Rev. Lett.* **118** (2017) 021303 [[1608.07648](#)].
- [21] LUX-ZEPLIN collaboration, *First Dark Matter Search Results from the LUX-ZEPLIN (LZ) Experiment*, [2207.03764](#).
- [22] PANDAX-II collaboration, *Dark Matter Results From 54-Ton-Day Exposure of PandaX-II Experiment*, *Phys. Rev. Lett.* **119** (2017) 181302 [[1708.06917](#)].
- [23] PANDAX-4T collaboration, *Dark Matter Search Results from the PandaX-4T Commissioning Run*, *Phys. Rev. Lett.* **127** (2021) 261802 [[2107.13438](#)].
- [24] XMASS collaboration, *Search for sub-GeV dark matter by annual modulation using XMASS-I detector*, *Phys. Lett. B* **795** (2019) 308 [[1808.06177](#)].
- [25] DEAP collaboration, *Search for dark matter with a 231-day exposure of liquid argon using DEAP-3600 at SNOLAB*, *Phys. Rev. D* **100** (2019) 022004 [[1902.04048](#)].
- [26] DARKSIDE collaboration, *Low-Mass Dark Matter Search with the DarkSide-50 Experiment*, *Phys. Rev. Lett.* **121** (2018) 081307 [[1802.06994](#)].
- [27] DARKSIDE collaboration, *DarkSide-50 532-day Dark Matter Search with Low-Radioactivity Argon*, *Phys. Rev. D* **98** (2018) 102006 [[1802.07198](#)].
- [28] SUPERCDMS collaboration, *Search for Low-Mass Weakly Interacting Massive Particles with SuperCDMS*, *Phys. Rev. Lett.* **112** (2014) 241302 [[1402.7137](#)].
- [29] SUPERCDMS collaboration, *Search for Low-Mass Dark Matter with CDMSlite Using a Profile Likelihood Fit*, *Phys. Rev. D* **99** (2019) 062001 [[1808.09098](#)].
- [30] EDELWEISS collaboration, *Searching for low-mass dark matter particles with a massive Ge bolometer operated above-ground*, *Phys. Rev. D* **99** (2019) 082003 [[1901.03588](#)].
- [31] CRESST collaboration, *First results from the CRESST-III low-mass dark matter program*, *Phys. Rev. D* **100** (2019) 102002 [[1904.00498](#)].
- [32] CRESST collaboration, *Probing spin-dependent dark matter interactions with ${}^6\text{Li}$: CRESST Collaboration*, *Eur. Phys. J. C* **82** (2022) 207 [[2201.03863](#)].
- [33] KIMS collaboration, *Search for low-mass dark matter with CsI(Tl) crystal detectors*, *Physical Review D* **90** (2014) 052006 [[1404.3443](#)].
- [34] PICO collaboration, *Dark Matter Search Results from the Complete Exposure of the PICO-60 C_3F_8 Bubble Chamber*, *Phys. Rev. D* **100** (2019) 022001 [[1902.04031](#)].
- [35] CDEX collaboration, *Search for Light Weakly-Interacting-Massive-Particle Dark Matter by Annual Modulation Analysis with a Point-Contact Germanium Detector at the China Jinping Underground Laboratory*, *Phys. Rev. Lett.* **123** (2019) 221301 [[1904.12889](#)].
- [36] J. Amaré, S. Cebrián, I. Coarasa, C. Cuesta, E. García, M. Martínez et al., *Performance of ANAIS-112 experiment after the first year of data taking*, *The European Physical Journal C* **79** (2019) 228 [[1812.01472](#)].
- [37] J. Amaré et al., *Dark Matter Searches Using NaI(Tl) at the Canfranc Underground Laboratory: Past, Present and Future*, *Universe* **8** (2022) 75.

- [38] J. Amare et al., *Annual Modulation Results from Three Years Exposure of ANAIS-112*, *Phys. Rev. D* **103** (2021) 102005 [2103.01175].
- [39] COSINE-100 collaboration, *Strong constraints from COSINE-100 on the DAMA dark matter results using the same sodium iodide target*, *Sci. Adv.* **7** (2021) abk2699 [2104.03537].
- [40] COSINE-100 collaboration, *Three-year annual modulation search with COSINE-100*, *Phys. Rev. D* **106** (2022) 052005 [2111.08863].
- [41] SABRE collaboration, *The SABRE project and the SABRE Proof-of-Principle*, *Eur. Phys. J. C* **79** (2019) 363 [1806.09340].
- [42] F. Calaprice et al., *High sensitivity characterization of an ultrahigh purity NaI(Tl) crystal scintillator with the SABRE proof-of-principle detector*, *Phys. Rev. D* **104** (2021) L021302 [2105.09225].
- [43] K. Fushimi, Y. Kanemitsu, S. Hirata, D. Chernyak, R. Hazama, H. Ikeda et al., *Development of highly radiopure NaI(Tl) scintillator for PICOLON dark matter search project*, *Progress of Theoretical and Experimental Physics* **2021** (2021) [2101.00759].
- [44] G. Angloher, Carniti, P., Cassina, L., Gironi, L., Gotti, C., Gütlein, A. et al., *The COSINUS project: perspectives of a NaI scintillating calorimeter for dark matter search*, *The European Physical Journal C* **76** (2016) 441.
- [45] DM-ICE collaboration, *First search for a dark matter annual modulation signal with NaI(Tl) in the Southern Hemisphere by DM-Ice17*, *Phys. Rev. D* **95** (2017) 032006 [1602.05939].
- [46] EDELWEISS collaboration, *Optimizing EDELWEISS detectors for low-mass WIMP searches*, *Phys. Rev. D* **97** (2018) 022003 [1707.04308].
- [47] C.K. Khosa, L. Mars, J. Richards and V. Sanz, *Convolutional neural networks for direct detection of dark matter*, *Journal of Physics G: Nuclear and Particle Physics* **47** (2020) 095201 [1911.09210].
- [48] LUX collaboration, *Improving sensitivity to low-mass dark matter in LUX using a novel electrode background mitigation technique*, *Phys. Rev. D* **104** (2021) 012011 [2011.09602].
- [49] COSINE-100 collaboration, *Lowering the energy threshold in COSINE-100 dark matter searches*, *Astropart. Phys.* **130** (2021) 102581 [2005.13784].
- [50] DRIFT collaboration, *Improved sensitivity of the DRIFT-II-d directional dark matter experiment using machine learning*, *JCAP* **07** (2021) 014 [2103.06702].
- [51] J. Herrero-Garcia, R. Patrick and A. Scaffidi, *A semi-supervised approach to dark matter searches in direct detection data with machine learning*, *JCAP* **02** (2022) 039 [2110.12248].
- [52] I. Coarasa, *ANAIS-112. Searching for the annual modulation of dark matter with a 112.5 kg NaI(Tl) detector at the Canfranc Underground Laboratory*, Ph.D. thesis, Universidad de Zaragoza, 2021.
- [53] I. Coarasa, J. Apilluelo, J. Amaré, S. Cebrián, D. Cintas, E. García et al., *Machine-learning techniques applied to three-year exposure of ANAIS-112*, *Journal of Physics: Conference Series* **2156** (2021) 012036 [2110.10649].
- [54] M. Oliván, J. Amaré, S. Cebrián, C. Cuesta, E. García, M. Martínez et al., *Light yield determination in large sodium iodide detectors applied in the search for dark matter*, *Astroparticle Physics* **93** (2017) 86 [1703.01262].
- [55] M.Á. Oliván, *Design, scale-up and characterization of the data acquisition system for the ANAIS dark matter experiment*, Ph.D. thesis, Universidad de Zaragoza, 2015.
- [56] BOREXINO collaboration, *Modulations of the Cosmic Muon Signal in Ten Years of Borexino Data*, *JCAP* **02** (2019) 046 [1808.04207].

- [57] OPERA collaboration, *Measurement of the cosmic ray muon flux seasonal variation with the OPERA detector*, *JCAP* **10** (2019) 003 [[1810.10783](#)].
- [58] J. Amaré, S. Cebrián, I. Coarasa, C. Cuesta, E. García, M. Martínez et al., *Analysis of backgrounds for the ANAIS-112 dark matter experiment*, *The European Physical Journal C* **79** (2019) 412 [[1812.01377](#)].
- [59] M. Ibe, W. Nakano, Y. Shoji and K. Suzuki, *Migdal Effect in Dark Matter Direct Detection Experiments*, *JHEP* **03** (2018) 194 [[1707.07258](#)].
- [60] D. Cintas et al., *Quenching Factor consistency across several NaI(Tl) crystals*, *J. Phys. Conf. Ser.* **2156** (2021) 012065 [[2111.09590](#)].
- [61] L.J. Bignell et al., *Quenching factor measurements of sodium nuclear recoils in NaI:Tl determined by spectrum fitting*, *JINST* **16** (2021) P07034 [[2102.02833](#)].
- [62] KIMS-NAI collaboration, *Measurement of nuclear recoil responses of NaI(Tl) crystal for dark matter search*, *PoS ICRC2017* (2018) 882.
- [63] T. Stiegler, C. Sofka, R.C. Webb and J.T. White, *A study of the NaI(Tl) detector response to low energy nuclear recoils and a measurement of the quenching factor in NaI(Tl)*, [1706.07494](#).
- [64] J. Xu et al., *Scintillation Efficiency Measurement of Na Recoils in NaI(Tl) Below the DAMA/LIBRA Energy Threshold*, *Phys. Rev. C* **92** (2015) 015807 [[1503.07212](#)].
- [65] J.I. Collar, *Quenching and channeling of nuclear recoils in NaI(Tl): Implications for dark-matter searches*, *Phys. Rev. C* **88** (2013) 035806 [[1302.0796](#)].
- [66] H. Chagani, P. Majewski, E.J. Daw, V.A. Kudryavtsev and N.J.C. Spooner, *Measurement of the quenching factor of Na recoils in NaI(Tl)*, *JINST* **3** (2008) P06003 [[0806.1916](#)].
- [67] G. Gerbier, J. Mallet, L. Mosca, C. Tao, B. Chambon, V. Chazal et al., *Pulse shape discrimination with NaI(Tl) and results from a WIMP search at the Laboratoire Souterrain de Modane*, *Astropart. Phys.* **11** (1999) 287.
- [68] D.R. Tovey, V. Kudryavtsev, M. Lehner, J.E. McMillan, C.D. Peak, J.W. Roberts et al., *Measurement of scintillation efficiencies and pulse-shapes for nuclear recoils in NaI(Tl) and CaF-2(Eu) at low-energies for dark matter experiments*, *Phys. Lett. B* **433** (1998) 150.
- [69] R. Bernabei et al., *New limits on WIMP search with large-mass low-radioactivity NaI(Tl) set-up at Gran Sasso*, *Phys. Lett. B* **389** (1996) 757.
- [70] N.J.C. Spooner, G.J. Davies, J.D. Davies, G.J. Pyle, T.D. Bucknell, G.T.A. Squier et al., *The Scintillation efficiency of sodium and iodine recoils in a NaI(Tl) detector for dark matter searches*, *Phys. Lett. B* **321** (1994) 156.
- [71] J. Amaré et al., *First Results on Dark Matter Annual Modulation from the ANAIS-112 Experiment*, *Phys. Rev. Lett.* **123** (2019) 031301 [[1903.03973](#)].
- [72] J. Amaré et al., *ANAIS-112 status: two years results on annual modulation*, *J. Phys. Conf. Ser.* **1468** (2020) 012014 [[1910.13365](#)].
- [73] GEANT4 collaboration, *GEANT4—a simulation toolkit*, *Nucl. Instrum. Meth. A* **506** (2003) 250.
- [74] H.S. Lee et al., *Pulse-shape discrimination between electron and nuclear recoils in a NaI(Tl) crystal*, *JHEP* **08** (2015) 093 [[1503.05253](#)].
- [75] KIMS collaboration, *Limits on Interactions between Weakly Interacting Massive Particles and Nucleons Obtained with NaI(Tl) crystal Detectors*, *JHEP* **03** (2019) 194 [[1806.06499](#)].
- [76] J. Amaré et al., *Long term measurement of the ^{222}Rn concentration in the Canfranc Underground Laboratory*, *Eur. Phys. J. C* **82** (2022) 891 [[2203.13978](#)].

- [77] F. Calaprice et al., *Performance of a SABRE detector module without an external veto*, [2205.13876](#).
- [78] Y. Coadou, *Decision trees*, *The European Physical Journal Conferences* **4** (2010) 02003.
- [79] Y. Coadou, *Boosted decision trees*, in *Artificial Intelligence for High Energy Physics*, pp. 9–58 (2022), DOI [[2206.09645](#)].
- [80] Y. Freund and R.E. Schapire, *Experiments with a new boosting algorithm*, in *Proceedings of the Thirteenth International Conference on International Conference on Machine Learning*, ICML'96, (San Francisco, CA, USA), pp. 148–156, Morgan Kaufmann Publishers Inc., 1996.
- [81] A. Hoecker, P. Speckmayer, J. Stelzer, J. Therhaag, E. von Toerne, H. Voss et al., *TMVA – Toolkit for Multivariate Data Analysis*, [physics/0703039](#).
- [82] I. Coarasa et al., *ANAIS-112 sensitivity in the search for dark matter annual modulation*, *Eur. Phys. J. C* **79** (2019) 233 [[1812.02000](#)].