Application of machine learning techniques to search dark matter with ANAIS–112

Iván Coarasa on
behalf of the **ANAIS team**
icoarasa@unizar.es

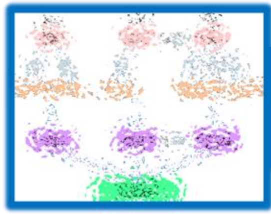
Centro de Astropartículas y Física de Altas Energías
Universidad Zaragoza

MultiDark
Multimessenger Approach for Dark Matter Detection

Outline

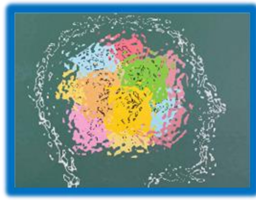
Introduction to dark matter

The ANAIS–112 experiment



Improving filtering protocols with Machine Learning

Annual modulation results with 6 years

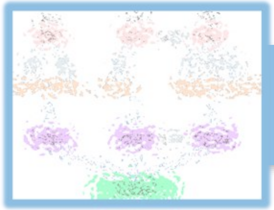


Summary and outlook

Outline

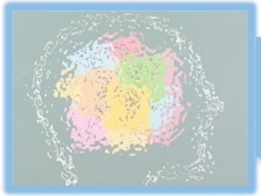
Introduction to dark matter

The ANAIS–112 experiment



Improving filtering protocols with Machine Learning

Annual modulation results with 6 years



Summary and outlook

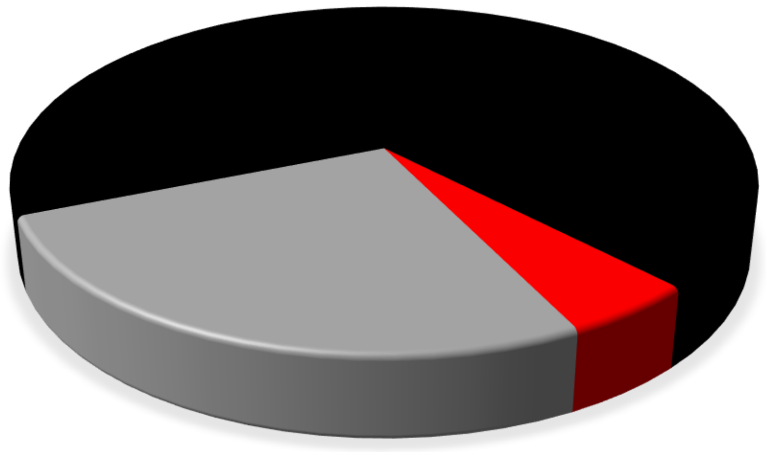
Dark matter in the universe

Planck's satellite (2018)

A large fraction of the Universe budget
is not explained within the Standard Model

Results consistent with
ΛCDM standard model

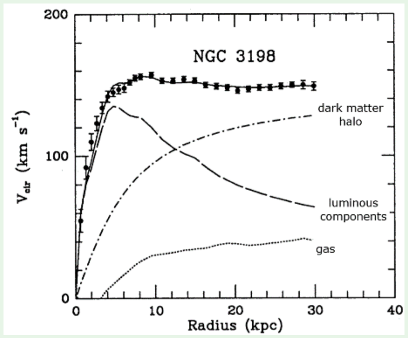
But… dark matter nature
still **unknown**!



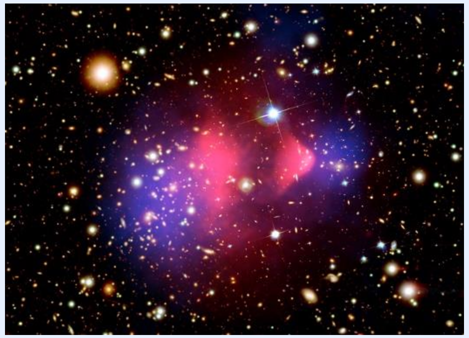
Dark energy
69%

Dark matter
26%

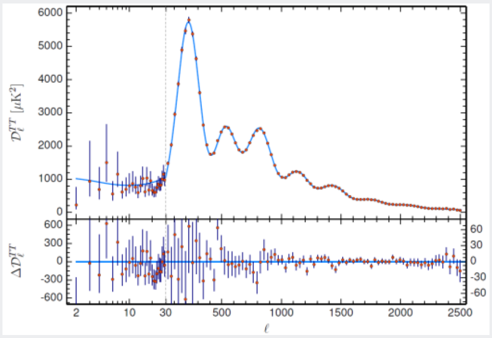
Ordinary matter
5%



Galactic scale



Galactic cluster scale

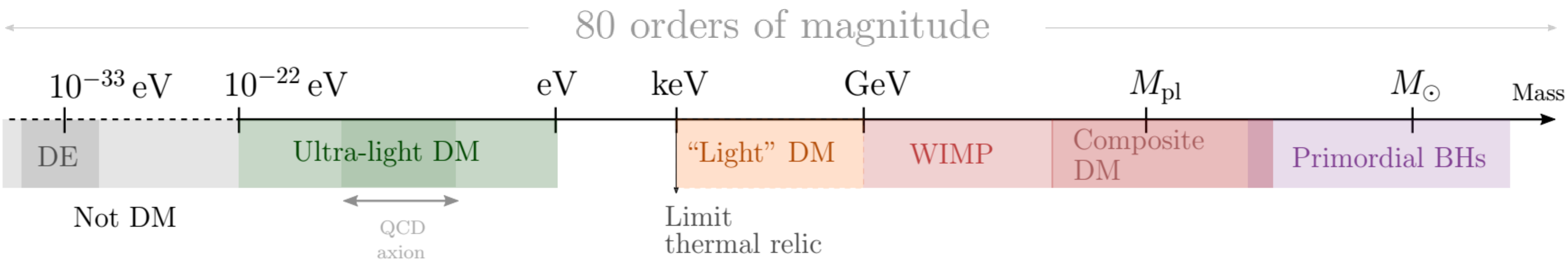


Cosmological scale

Numerous hints
at different scales

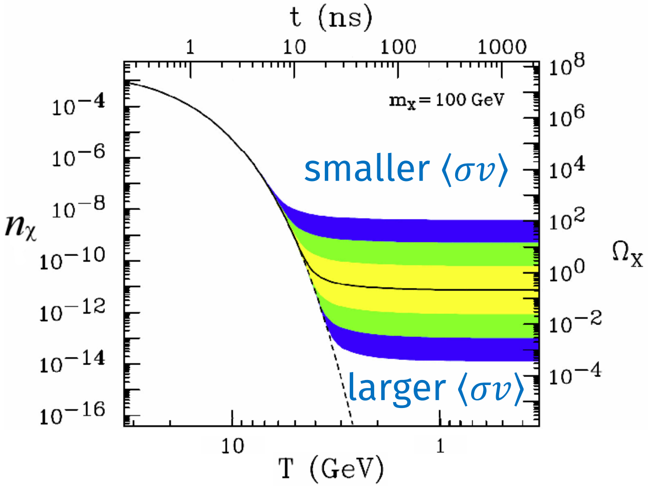
Dark matter candidates

A plethora of **DM candidates** beyond the SM: non-zero-mass, electrically neutral, stable particles having a very low interaction probability with baryonic matter



E. G. M. Ferreira, arXiv:2005.03254 (2021)

Thermal WIMPs: produced at the early Universe via a freeze-out mechanism when SM and DM particles were in thermal equilibrium, producing a constant relic density, reproduced for a wide range of masses from 1 eV to 120 TeV

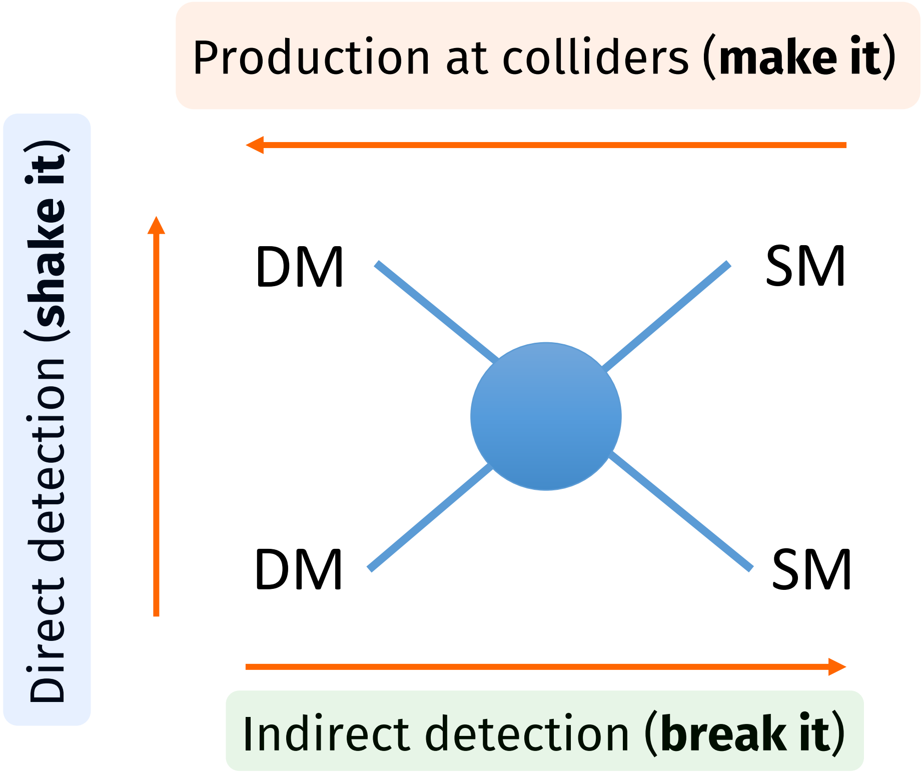


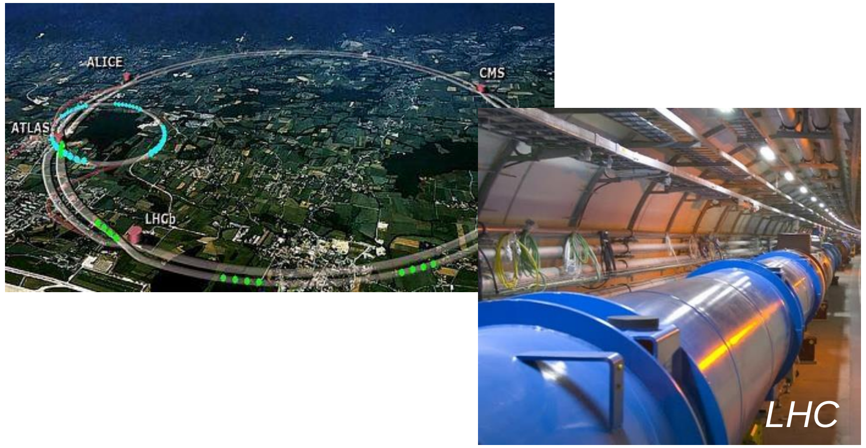
$$\langle\sigma_{ann} v\rangle_f \approx 3 \times 10^{-26}\ \mathrm{cm^3 s^{-1}}$$

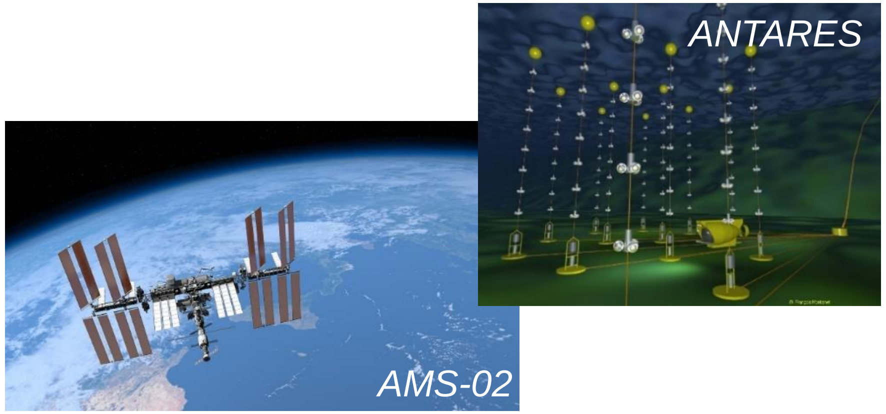
$$\sigma_{ann} \propto \frac{g^4}{m_\chi^2}$$

Dark matter detection

Different **complementary** strategies for detection



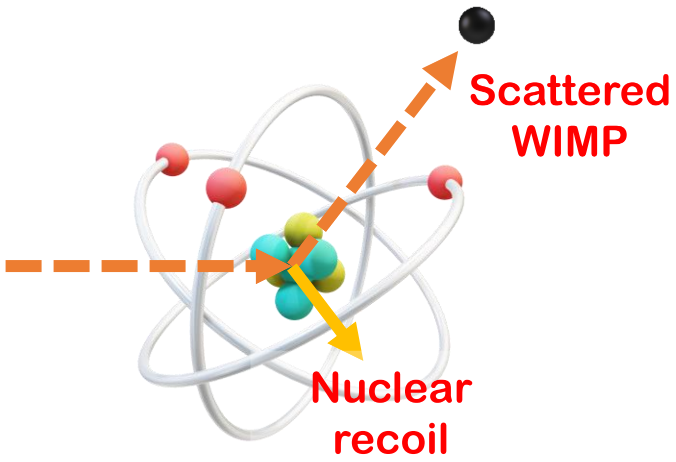




Dark matter direct detection

Nuclear recoils

WIMPs with masses of 10-100 GeV would produce nuclear recoils with energies of 1-100 keV

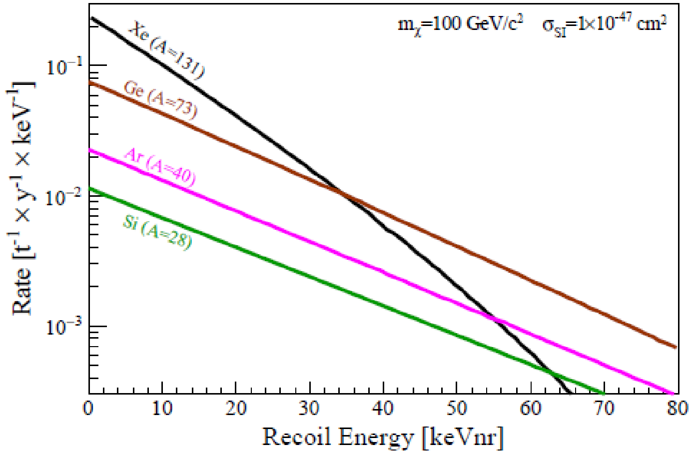


Signals:

- Heat (phonons/bubbles)
- Light (scintillation)
- Charge (ionization)

Detection rate

$$R = \int_{E_{th}}^{E_{máx}} dE_R\, \varepsilon(E_R) \frac{\rho_0}{m_W} \frac{M_{det}}{m_N} \int_{v_{mín}}^{v_{máx}} v f(\vec{v}) \frac{d\sigma_{WN}}{dE_R}(\vec{v}, E_R) d\vec{v}$$



Experimental setup

- Target nuclei (sensitive to spin-(in)dependent coupling)
- Detector mass
- Energy threshold

Astrophysical parameters

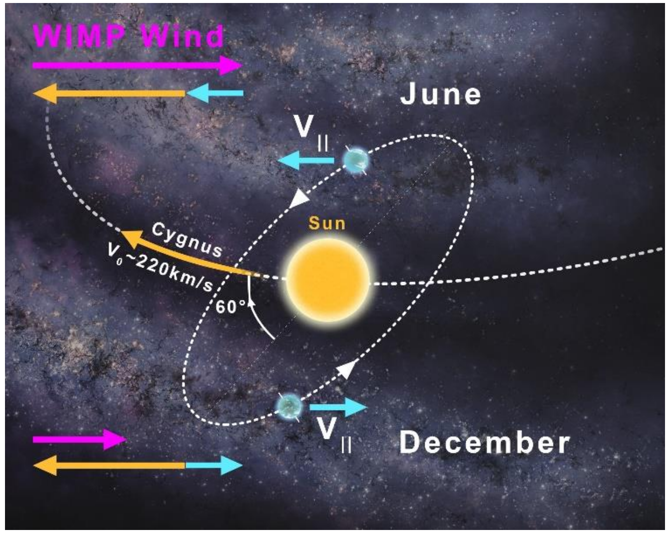
- Local DM density
- Velocity distribution

Theoretical inputs

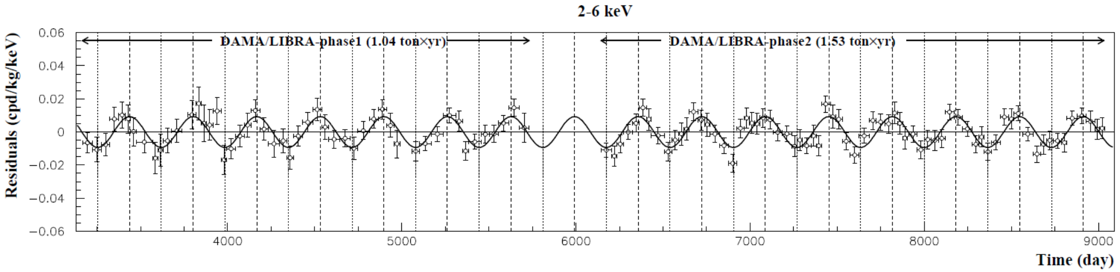
- Differential cross section
- Nuclear uncertainties

A DISTINCTIVE SIGNATURE IS NEEDED

Dark matter annual modulation & DAMA/LIBRA positive signal



DAMA/NaI and DAMA/LIBRA @LNGS (since 1995)



R. Bernabei et al., Nucl. Phys. At. Energy 22 (2021) 329-342

Cosine behaviour: $T = 1$ y, $\phi = 02/\mathrm{Jun}$

Only at low energy

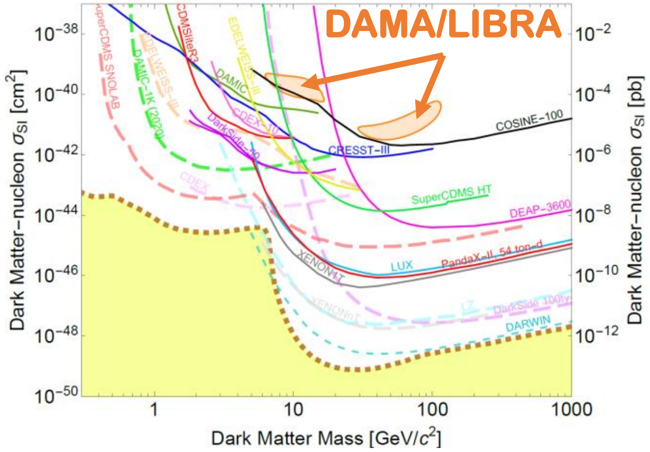
Single-hit events

$S_m/S_0 \lesssim 7\%$

DAMA/NaI: 100 kg NaI(Tl) [1995-2002]
DAMA/LIBRA: 250 kg NaI(Tl) [2003-today]

DAMA clearly observes an **annual modulation** compatible with DM **at more than 13σ**

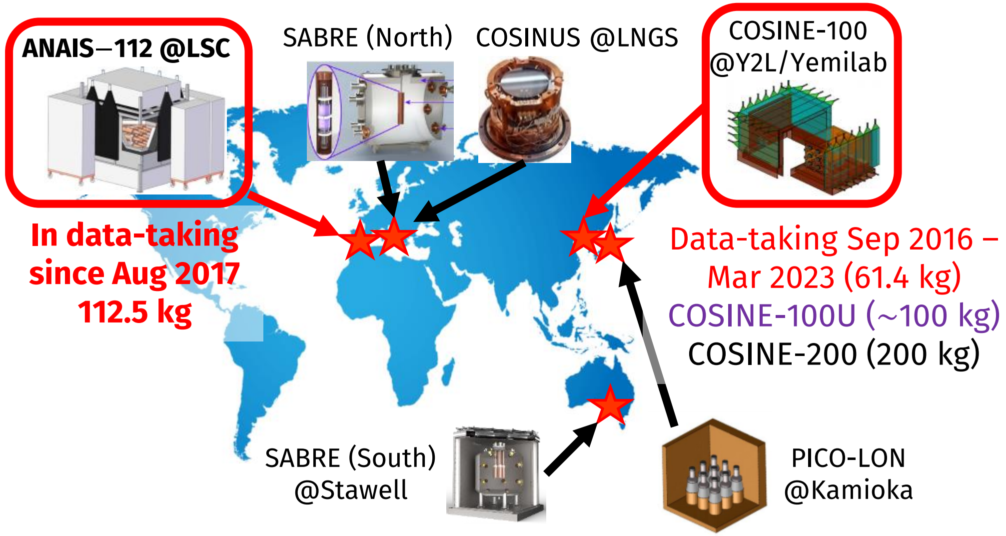
STRONG TENSION



Other very sensitive experiments do not see the signal, but the comparison is **model dependent**

A model independent test is needed using the same target

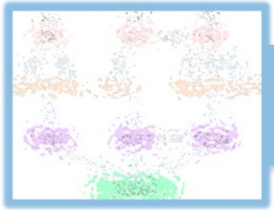
Other NaI experiments around the world



Outline

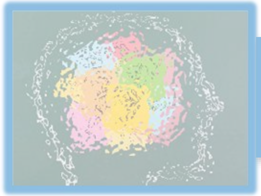
Introduction to dark matter

The ANAIS–112 experiment



Improving filtering protocols with Machine Learning

Annual modulation results with 6 years



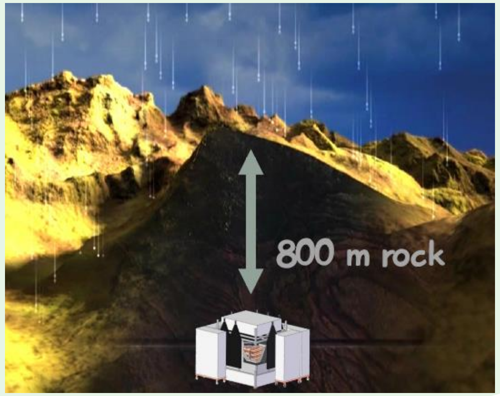
Summary and outlook

The ANAIS experiment

GOAL **ANAIS** (*Annual modulation with NaI(Tl) scintillators*) intends to provide a **model independent** test of the signal reported by DAMA/LIBRA, using the **same target and technique,** but different experimental approach and environmental conditions

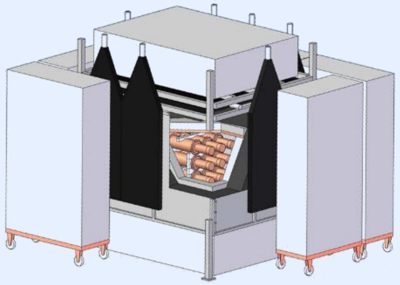
Projected sensitivity: 3σ in 5 years data-taking

WHERE At the **Canfranc Underground Laboratory**, LSC @ SPAIN (under 2450 m.w.e.)



ANAIS–112 SET-UP

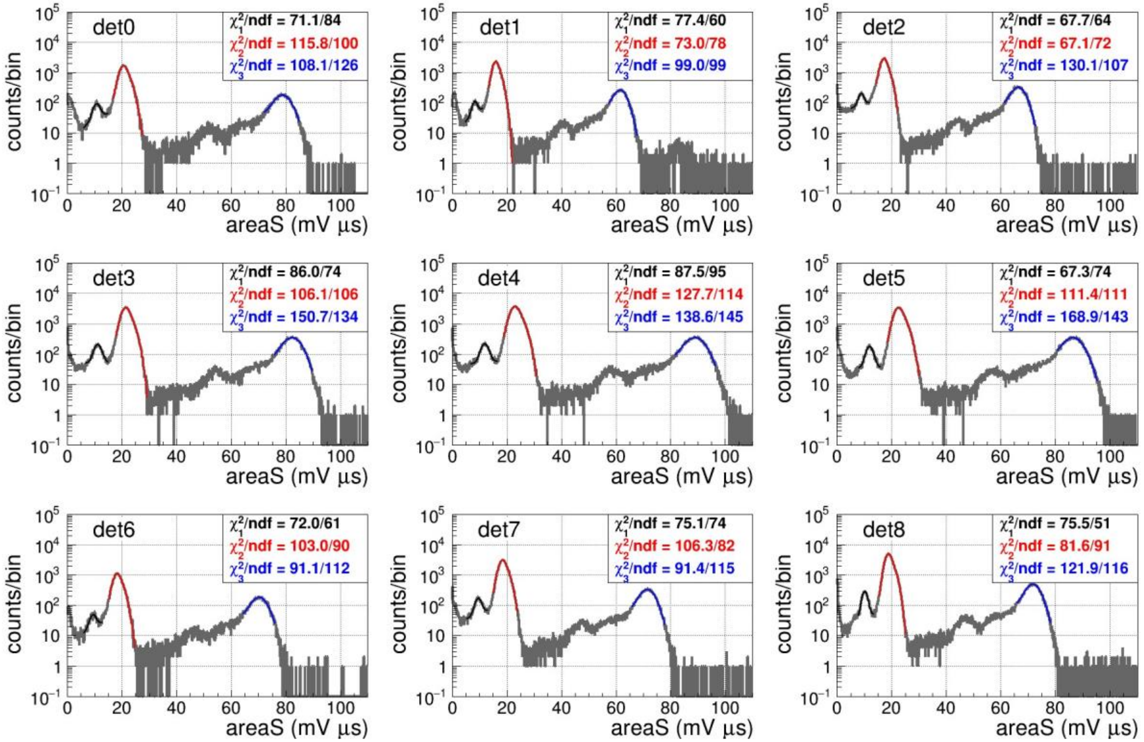
- 9 ultrapure NaI(Tl) crystals 12.5 kg (**112.5 kg**) in 3×3
- Cylindrical modules coupled to 2 high QE PMTs (~40%)



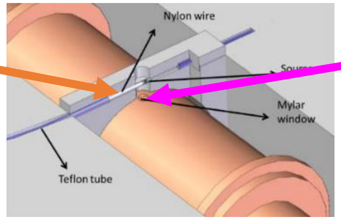
On 3 August 2017, data collection starts

Low energy calibration

Calibration with external **^{109}Cd sources** (11.9, 22.6 and 88.0 keV) every two weeks for gain correction

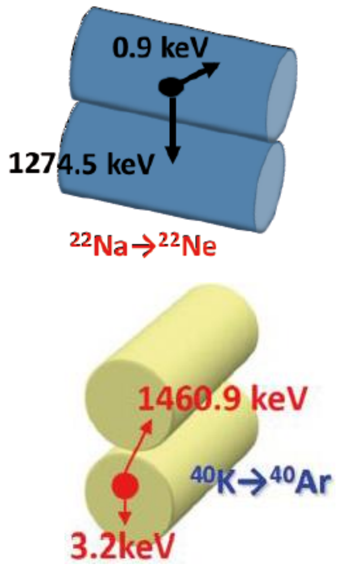


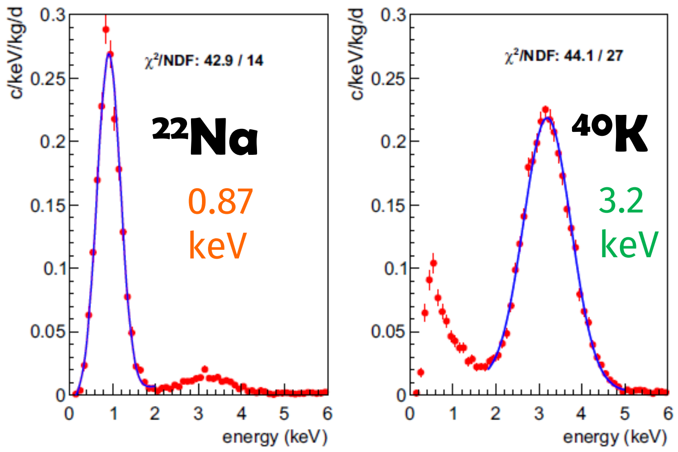
Guides for ^{109}Cd sources



Mylar window

Calibration in the **ROI [1-6] keV** with internal bulk contaminants **^{22}Na (0.9 keV)** and **^{40}K (3.2 keV)** using whole statistics



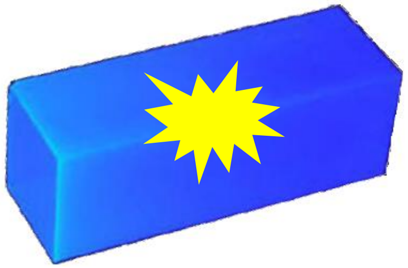


Linear calibration in 2 ranges:

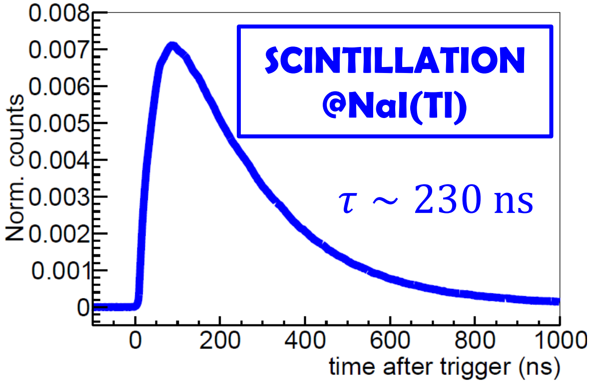
- 1-10 keV [ROI]
- 10-100 keV

What do we expect to see?

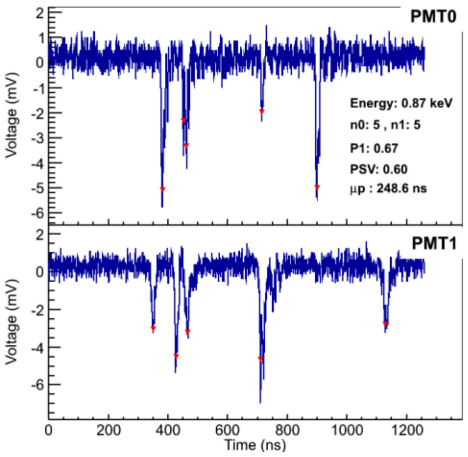
Scintillation light in the NaI(Tl) crystal (bulk)



Expected signal



... but at low energy (15 phe/keV)

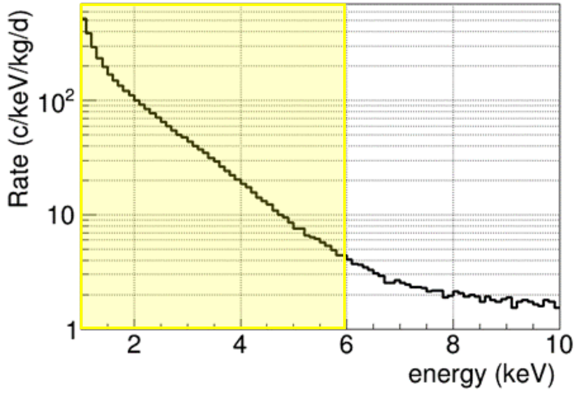


Due to the high light collection, we can see the individual photoelectrons (phe) in each PMT

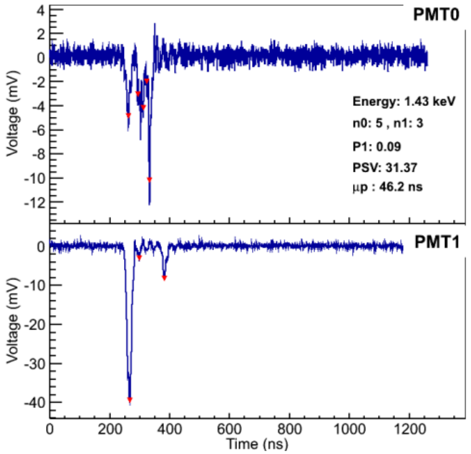
What do we actually see?

The region of interest (1-6 keV) is dominated by **non-bulk scintillation events**

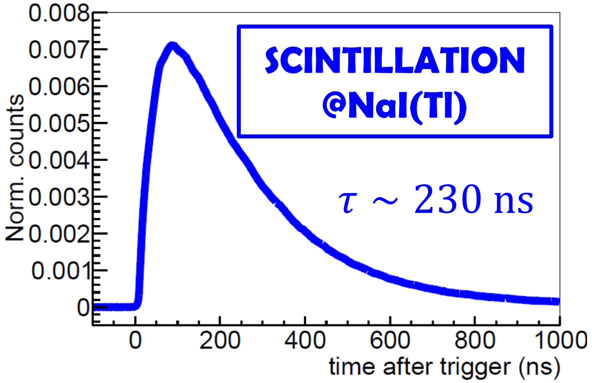
Application of event selection protocols to distinguish scintillation events from noise



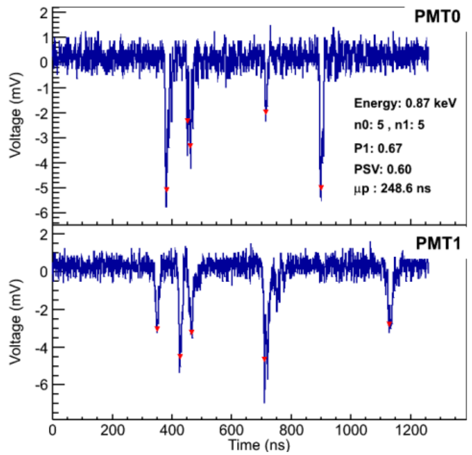
FAST (Cherenkov) @PMTs



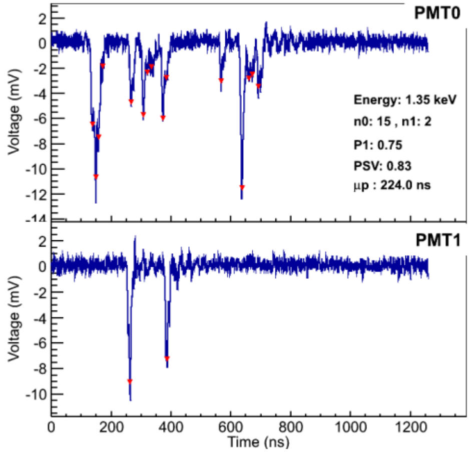
Expected signal



... but at low energy (15 phe/keV)



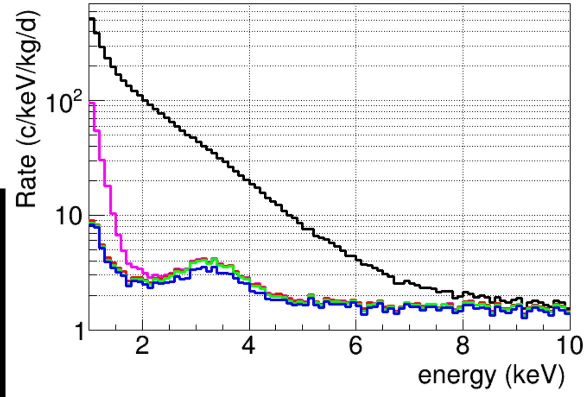
ASYMMETRIC light-sharing @PMTs



ANAIS–112 event selection

The region of interest (1-6 keV) is dominated by **non-bulk scintillation events**

Application of event selection protocols to distinguish scintillation events from noise

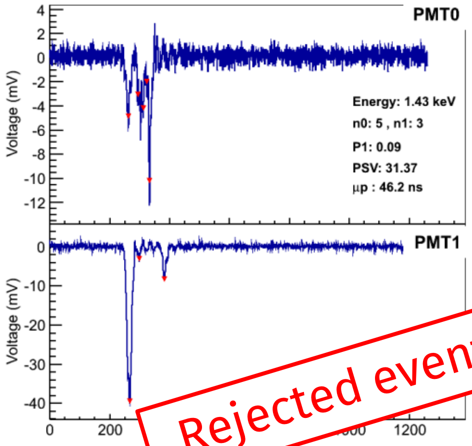


Standard analysis

1) Pulse shape cut to select pulses with NaI(Tl) scintillation constant
2) We remove asymmetric events (<2 keV_{ee}) with origin in the PMT
3) Remove 1 s after a muon passage
4) Multiplicity = 1 (single-hit events)

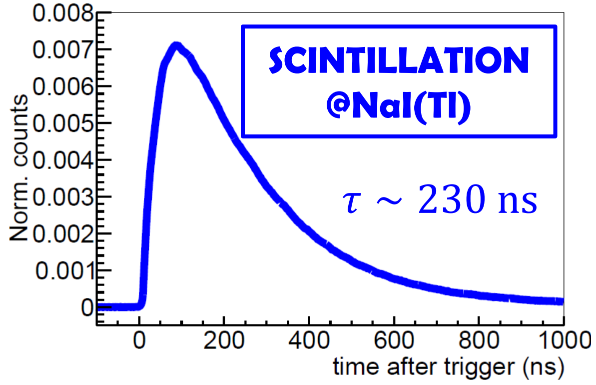
$$P_1 = \frac{\sum_{100\text{ ns}}^{600\text{ ns}} A(t)}{\sum_{0\text{ ns}}^{600\text{ ns}} A(t)} \qquad \mu_p = \frac{\sum_i A_i t_i}{\sum_i A_i} \qquad n_0, n_1$$

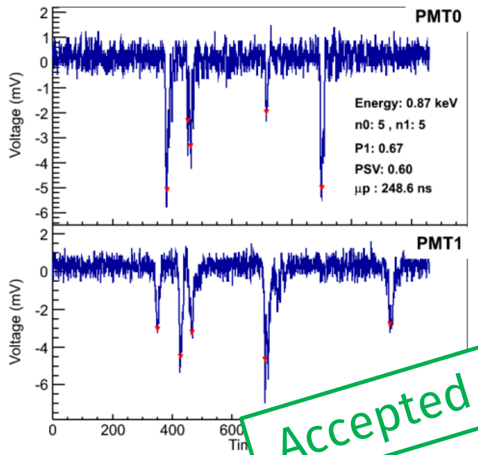
FAST (Cherenkov) @PMTs



Rejected event

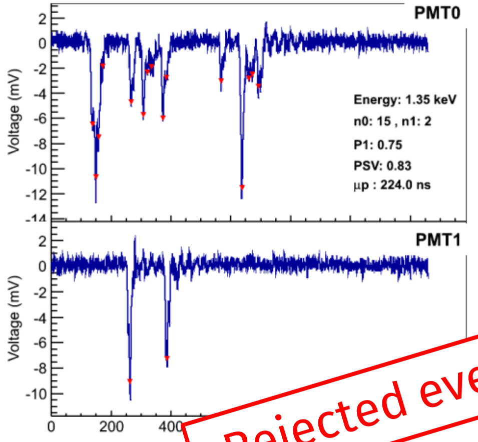
SCINTILLATION @NaI(Tl)





Accepted event

ASYMMETRIC light-sharing @PMTs

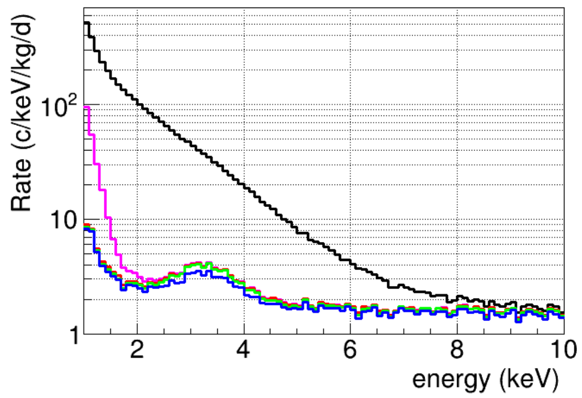


Rejected event

ANAIS—112 event selection

The region of interest (1-6 keV) is dominated by **non-bulk scintillation events**

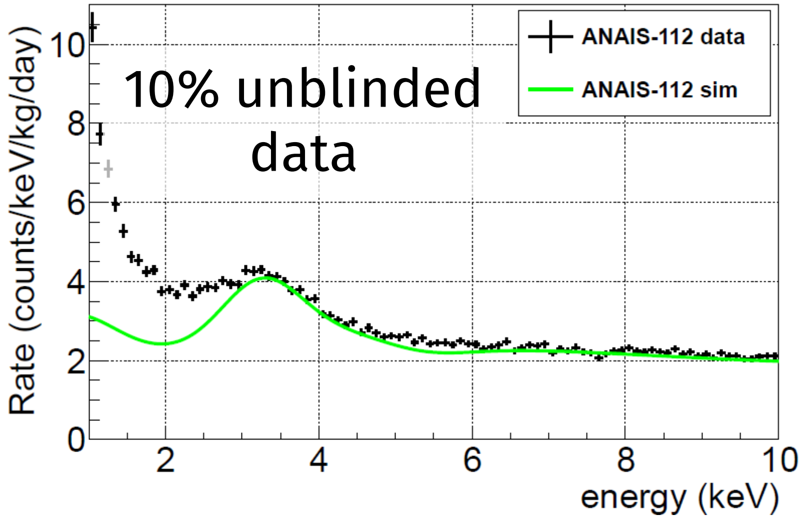
Application of event selection protocols to distinguish scintillation events from noise



Comparing with MC background model...

Strong discrepancy in [1-2] keV

...and very low efficiency at 1 keV (~15%)



The unexplained events <2 keV could be related with:

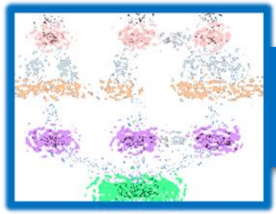
- ➔ background sources not considered in model
- ➔ non-bulk scintillation events not rejected

Machine-learning technique to improve noise rejection in [1-2] keV

Outline

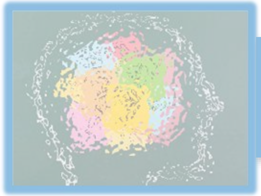
Introduction to dark matter

The ANAIS–112 experiment



Improving filtering protocols with Machine Learning

Annual modulation results with 6 years

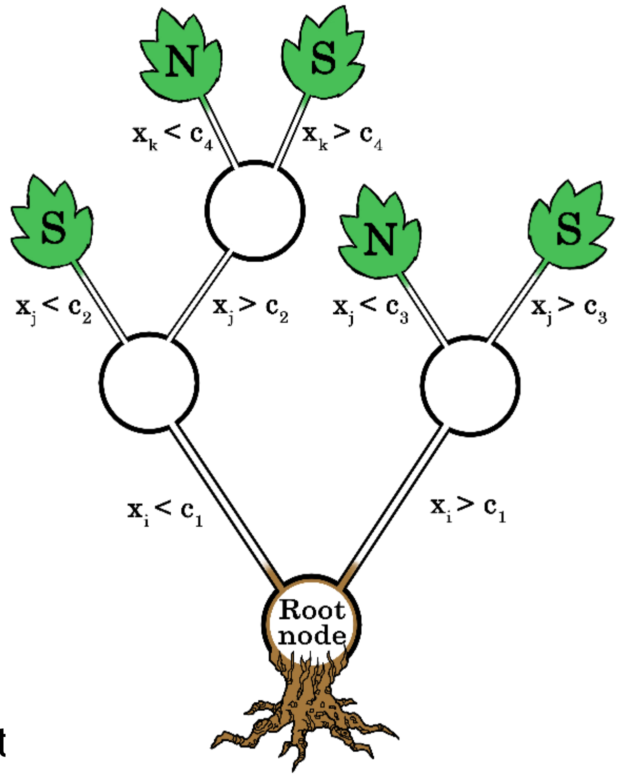


Summary and outlook

Machine-learning techniques for event selection

Boosted Decision Tree (BDT)

- Multivariate analysis
- Combination of several weak discriminating variables into a single powerful discriminator
- Two classes: signal-like and noise-like events
- BDT response: from -1 (noise-like) to +1 (signal-like)



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

AdaBoost

n_{Trees}: number of trees

$\alpha_j = \frac{1-f_j}{f_j}$: boost weight

f_j: fraction of misclassified events of the previous tree

$T_j(\vec{x}_i)$: result of an individual classifier (-1 or +1)

Performance of using ML for event selection in: *JCAP11(2022)048 and JCAP06(2023)E01*

Reanalysis of 3 years data in: *arXiv:2404.17348 (Apr. 2024), Accepted in Comm. Phys.*

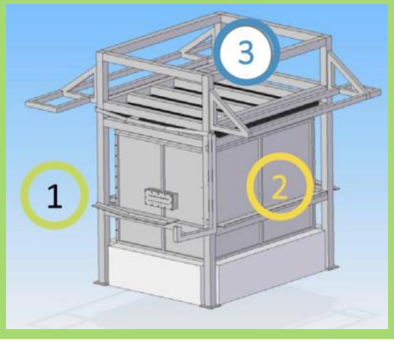
Training populations

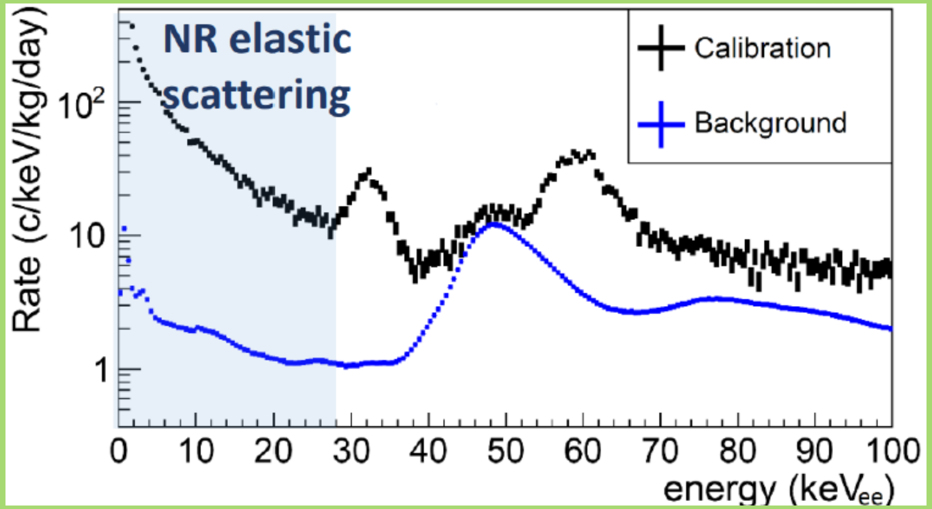
JCAP11(2022)048

SIGNAL EVENTS: Neutron calibrations

Seven calibration runs since April 2021 using ^{252}Cf neutron source at different positions in the ANAIS–112 set-up

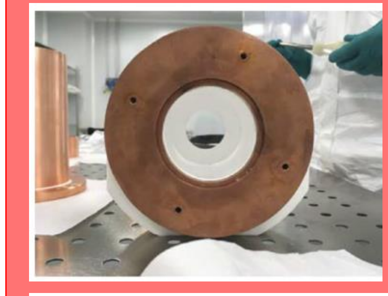
^{252}Cf

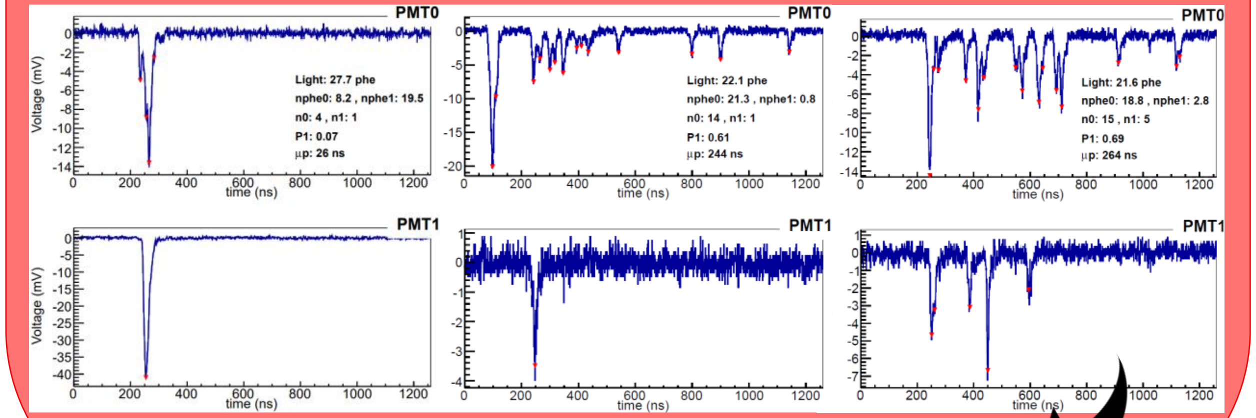




NOISE EVENTS: "Blank" module (No NaI(Tl))

Since 2018 a Blank module (similar to ANAIS–112 modules, but without NaI(Tl) crystal) is taking data with the same DAQ, but in an independent shielding close to ANAIS–112





Not rejected by ANAIS–112 standard cuts

Training parameters

JCAP11(2022)048

15 discrimination parameters combined in a boosted decision tree

instead of the **4** parameters used in the standard analysis

BDT

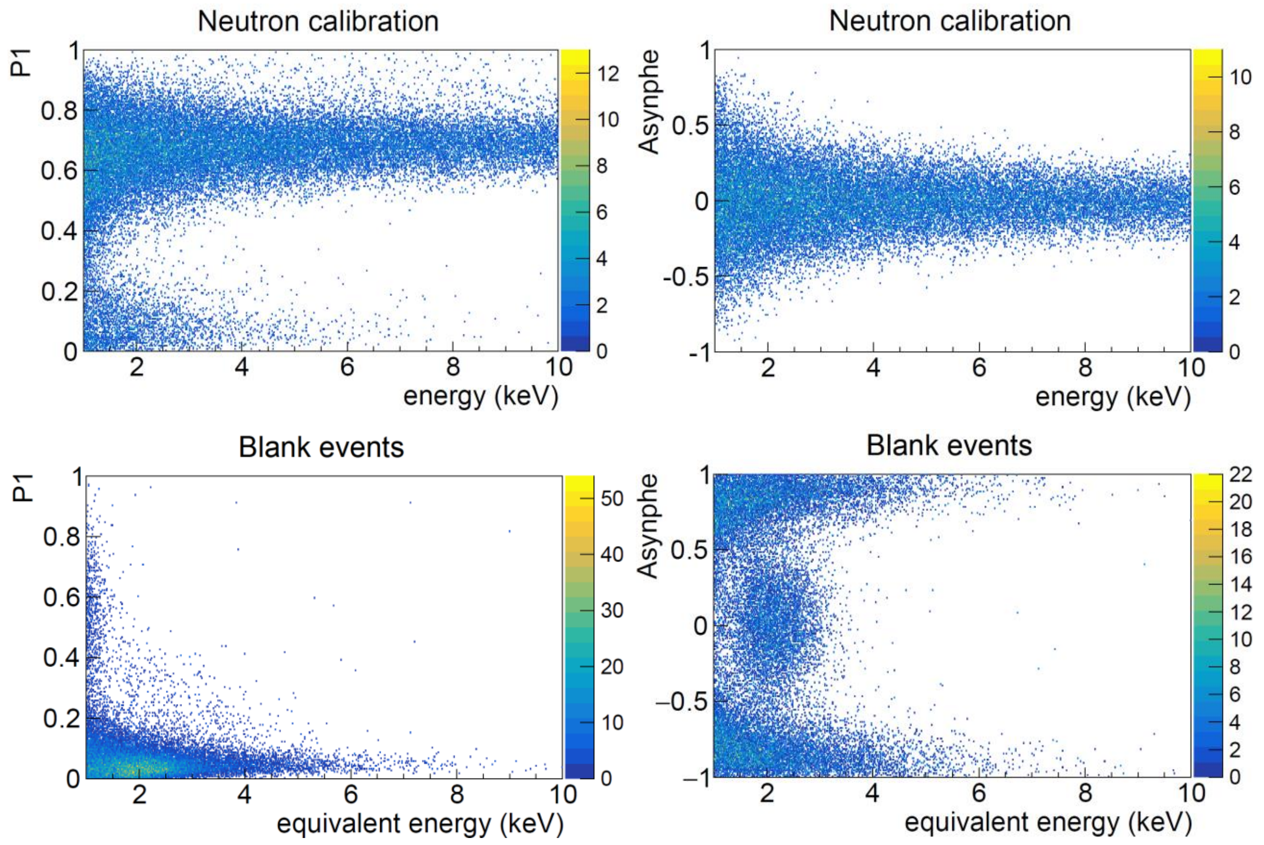
Standard analysis

$$P_1 = \frac{\sum_{100\text{ ns}}^{600\text{ ns}} A(t)}{\sum_{0\text{ ns}}^{600\text{ ns}} A(t)} \qquad \mu_p = \frac{\sum_i A_i t_i}{\sum_i A_i} \qquad n_0, n_1$$

$$P_2 = \frac{\sum_{0\text{ ns}}^{50\text{ ns}} A(t)}{\sum_{0\text{ ns}}^{600\text{ ns}} A(t)} \qquad Asynphe = \frac{nphe_0 - nphe_1}{nphe_0 + nphe_1}$$

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

$$x = 50, 100, 200, 300, 400, 500, 600, 700, 800\text{ ns}$$

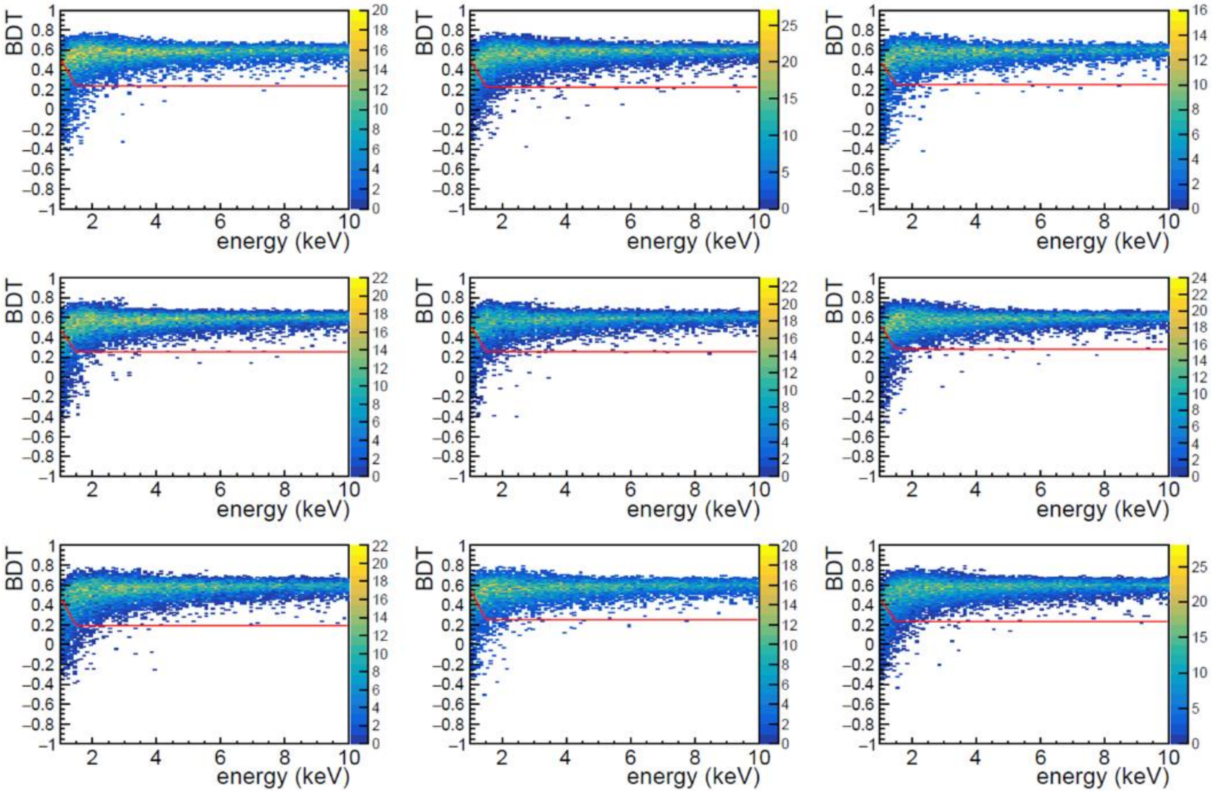


Equivalent energy from $LC = 14.5\text{ phe/keV}$

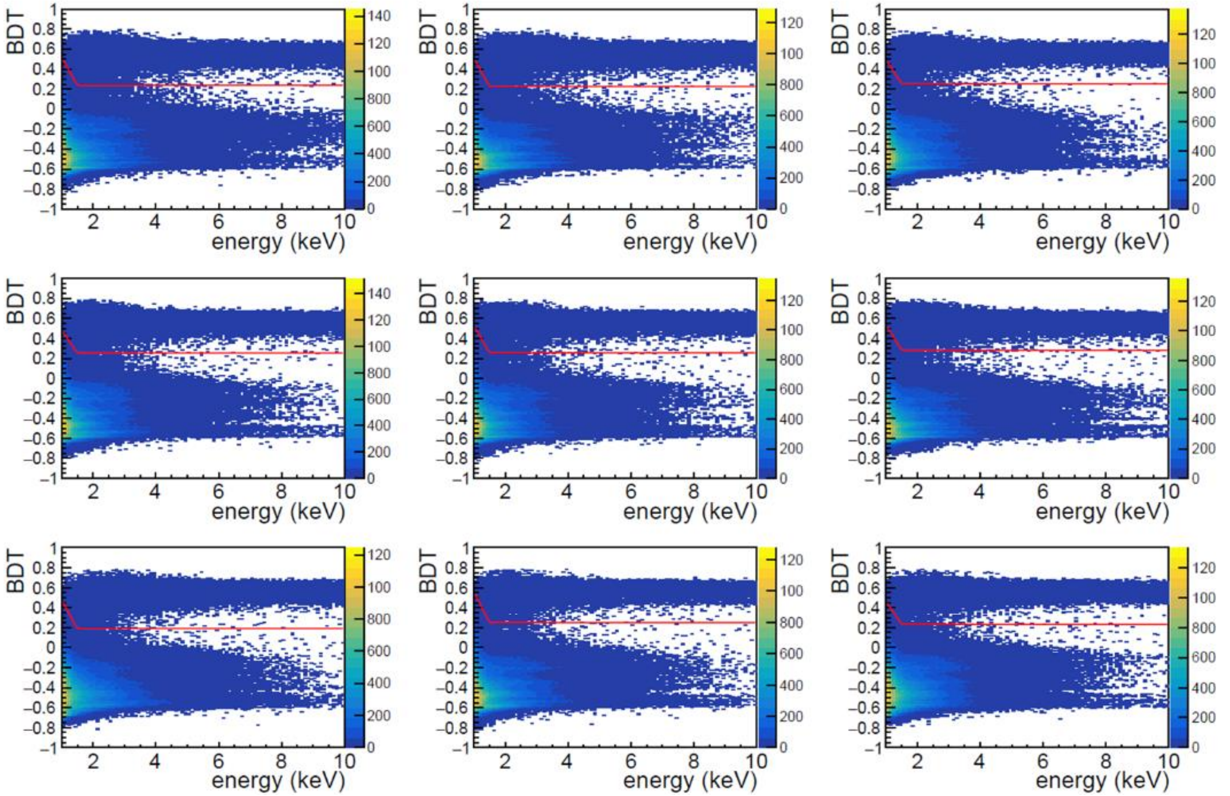
Improved filtering protocols with ML techniques

Following
JCAP11(2022)048

Neutron calibration



10% unblinded 3 years background events

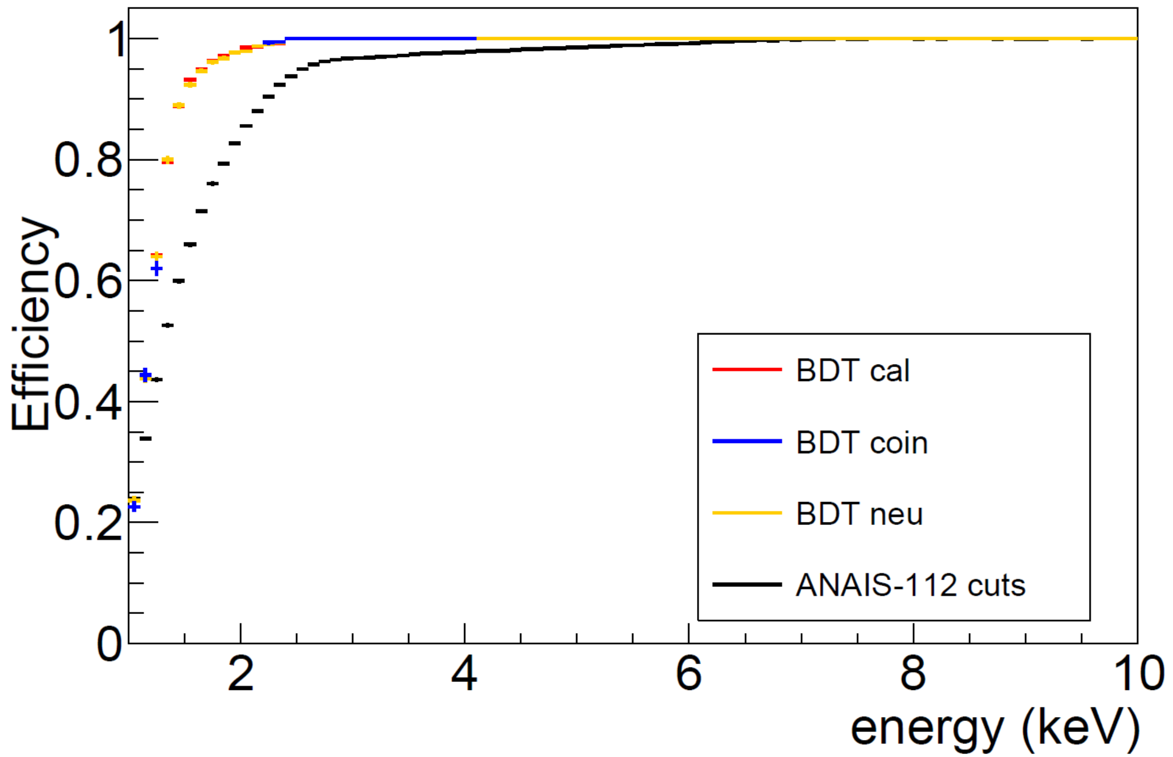


CUT on BDT parameter applied to background

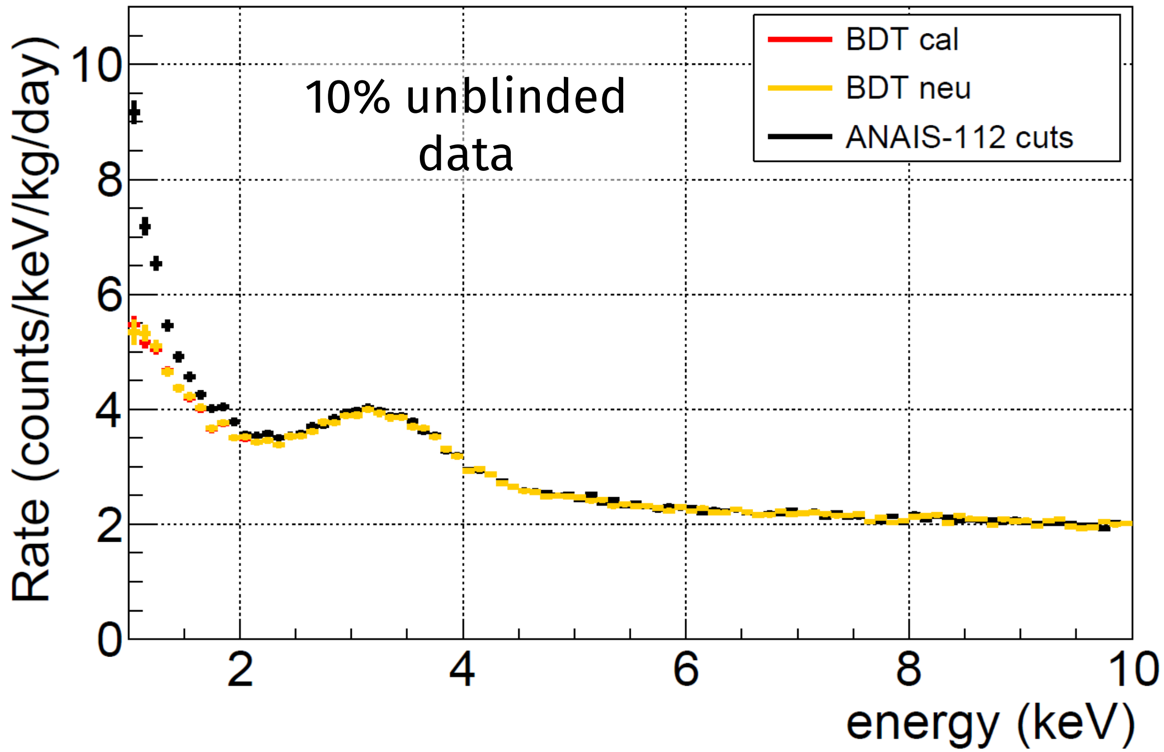
Improved filtering protocols with ML techniques

Following JCAP11(2022)048

Acceptance efficiency



Efficiency-corrected background

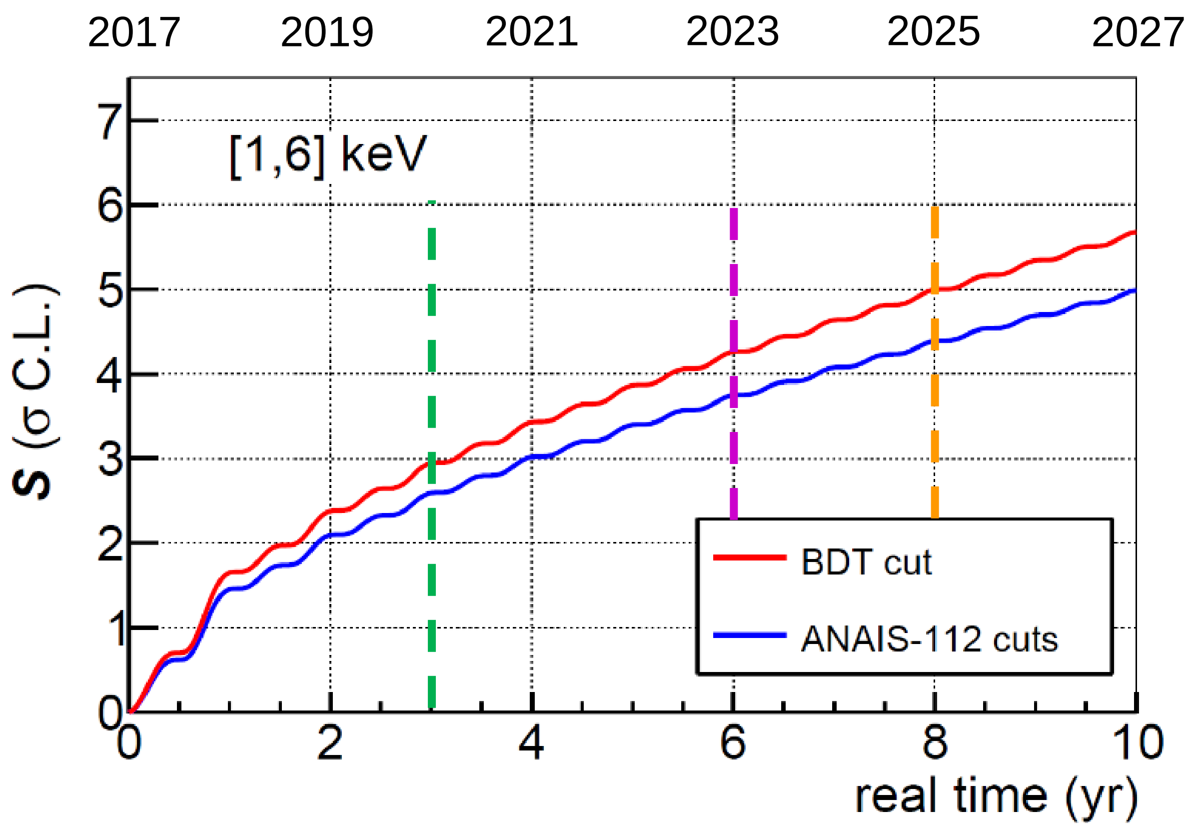


~30% improvement in efficiency in [1-2] keV

~18% background reduction in [1-2] keV

Improving ANAIS–112 sensitivity prospects

Sensitivity to DAMA/LIBRA result as $S = \frac{S_m^{DAMA}}{\sigma(\hat{S}_m)} \propto \sqrt{\frac{M\, T\, \varepsilon}{B}}$



The experimental sensitivity is given by the standard deviation of the modulation amplitude $\sigma(S_m)$, that can be estimated from:

- Updated background
- Efficiency estimate and its error
- Live time distribution

3σ sensitivity with 3 y

$>4\sigma$ sensitivity with 6 y (NOW)

5σ sensitivity in late 2025

Annual modulation analysis strategy

Focus on **model independent** analysis searching for modulation

- ➔ In order to better compare with DAMA/LIBRA results
  - ➔ use the same energy regions ([1-6] keV, [2-6] keV)
  - ➔ Fix period 1 year and phase to June 2nd
- ➔ Simultaneous fit of the 9 detectors in 10-day bins. Chi-square minimization: $\chi^2 = \sum_i (n_i - \mu_i)^2/\sigma_i^2$, where the expected number of events μ_i for detector d in time bin i is given by:

$$\mu_{i,d} = \left[R_{0,d}\left(1 + f_d \phi^{MC}_{bkg,d}(t_i)\right) + \boldsymbol{S_m} \cos\left(\omega(t_i - t_0)\right)\right] M_d \Delta E \Delta t$$

Annual modulation analysis strategy

Focus on **model independent** analysis searching for modulation

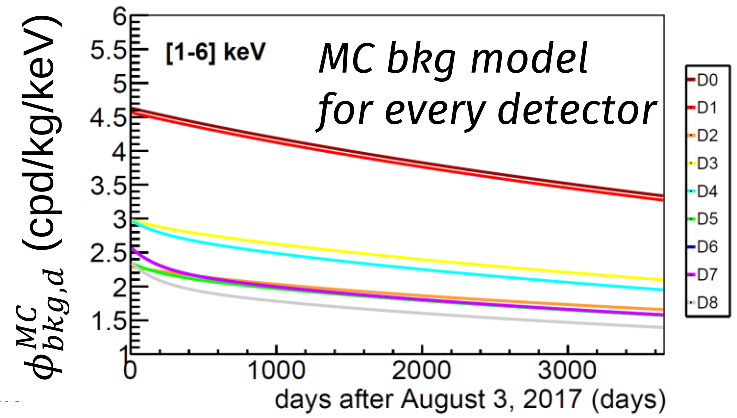
- In order to better compare with DAMA/LIBRA results
  - use the same energy regions ([1-6] keV, [2-6] keV)
  - Fix period 1 year and phase to June 2nd
- Simultaneous fit of the 9 detectors in 10-day bins. Chi-square minimization: $\chi^2 = \sum_i (n_i - \mu_i)^2/\sigma_i^2$, where the expected number of events μ_i for detector d in time bin i is given by:

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

Constant background (long-lived isotopes and residual noise)

Decaying background, modeled by MC

Modulation signal (fixed period and phase)

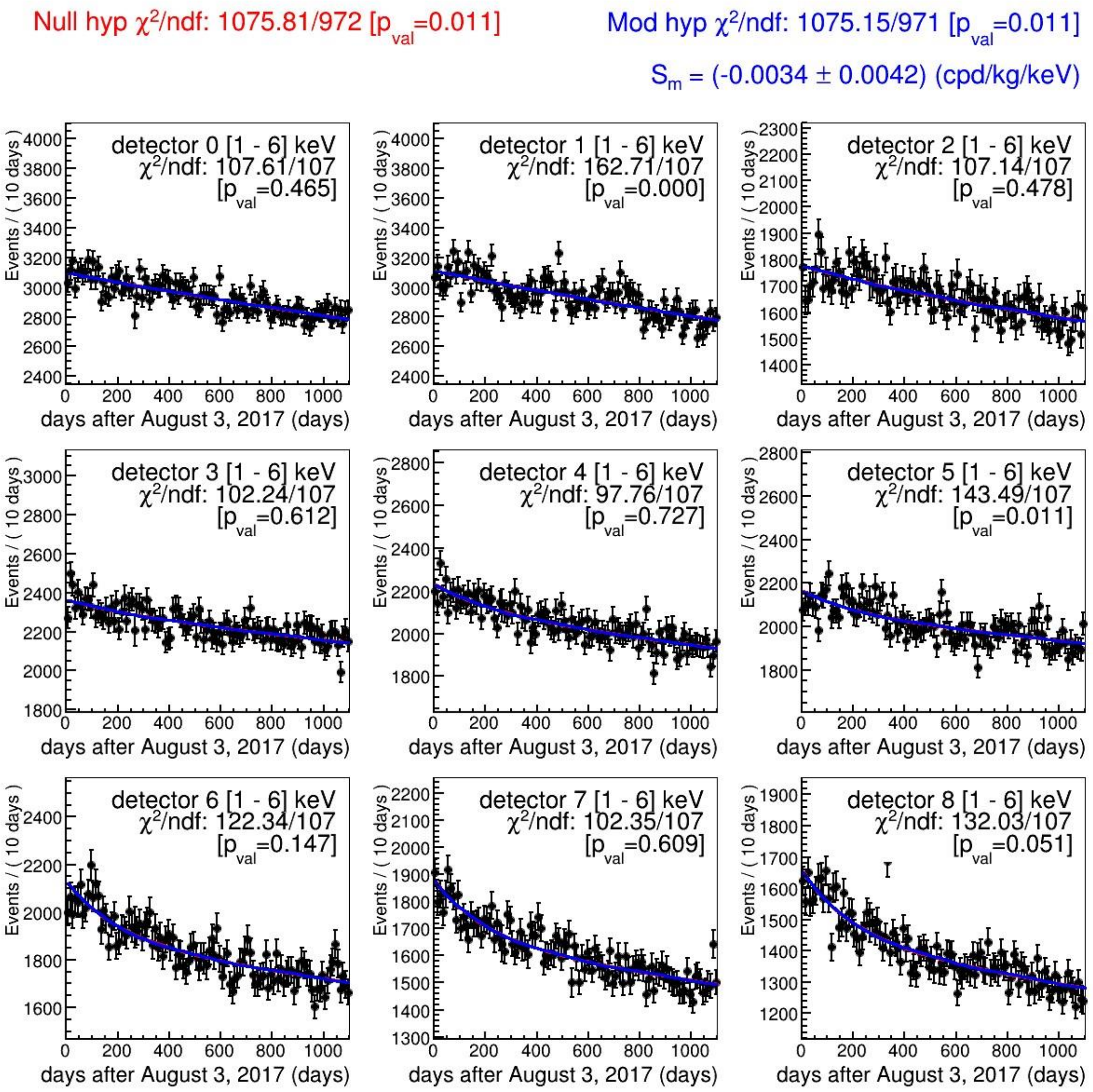


19 free parameters: $R_{0,d}, f_d, S_m$

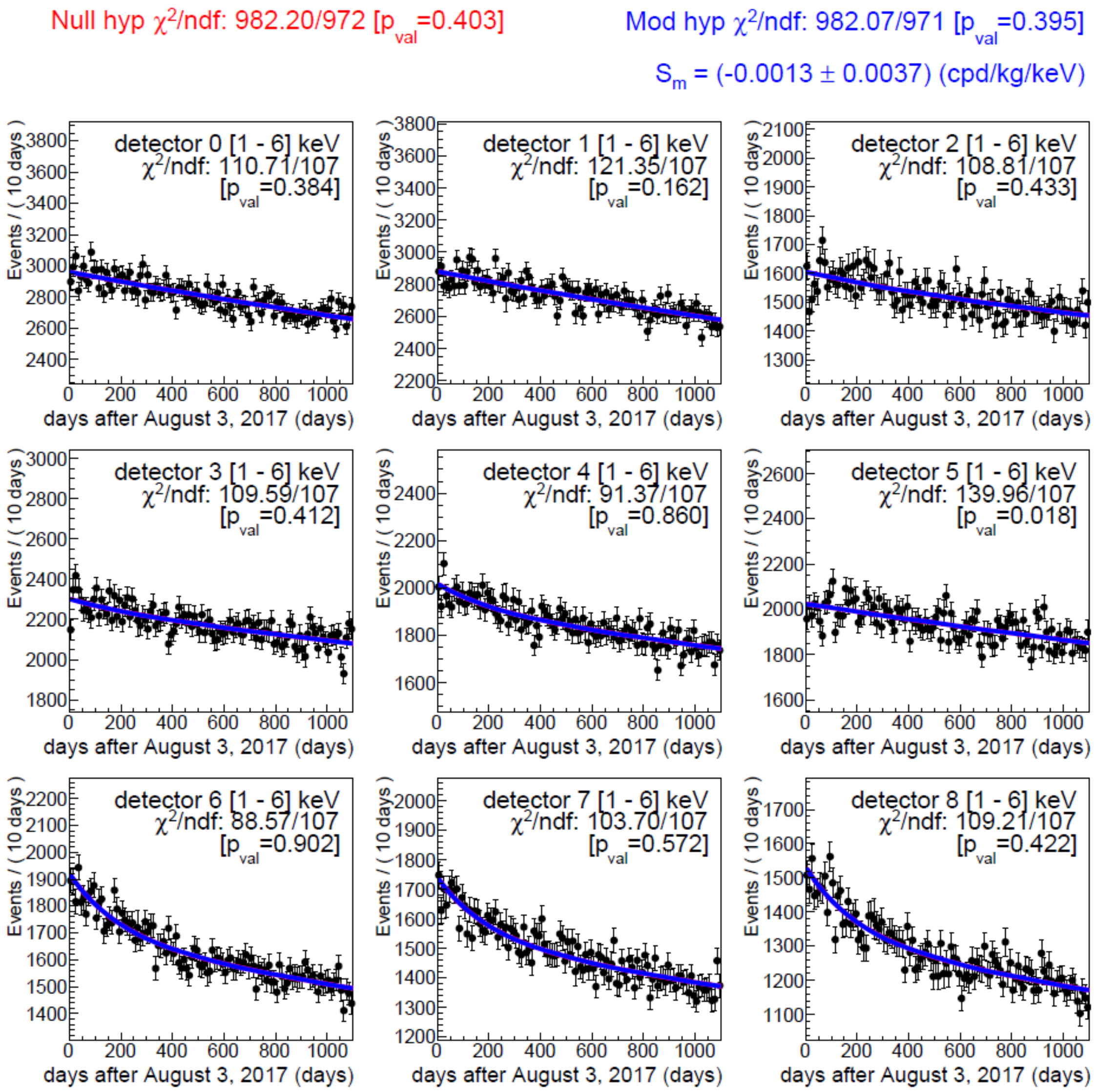
Improved 3-year results [1-6] keV

$2.5\sigma \rightarrow 2.8\sigma$

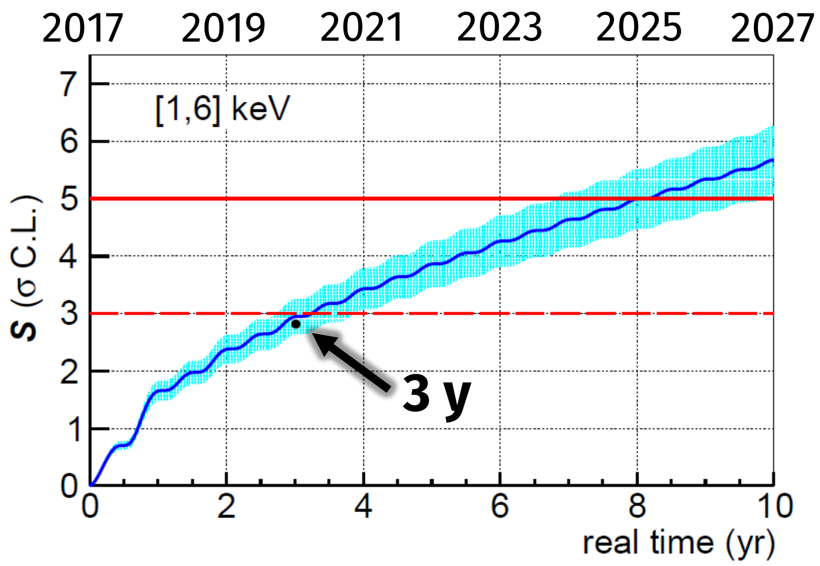
PRD103(2021)102005

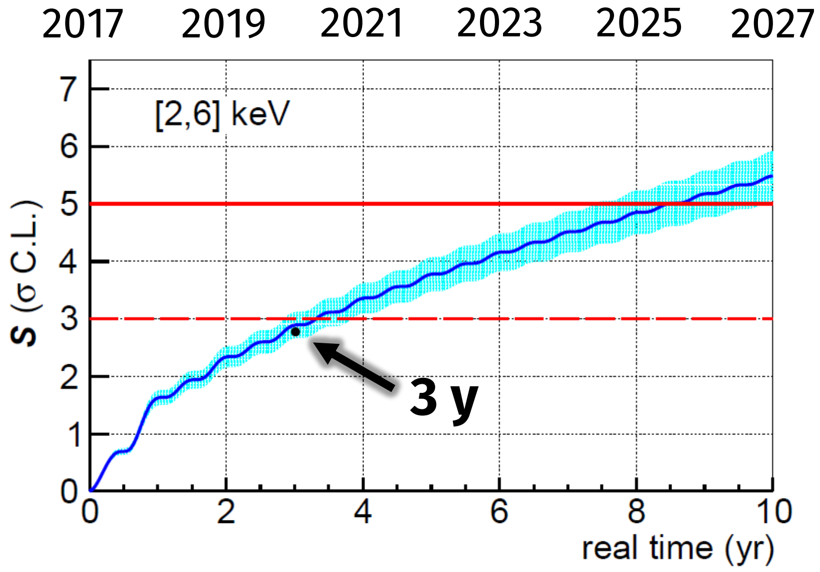


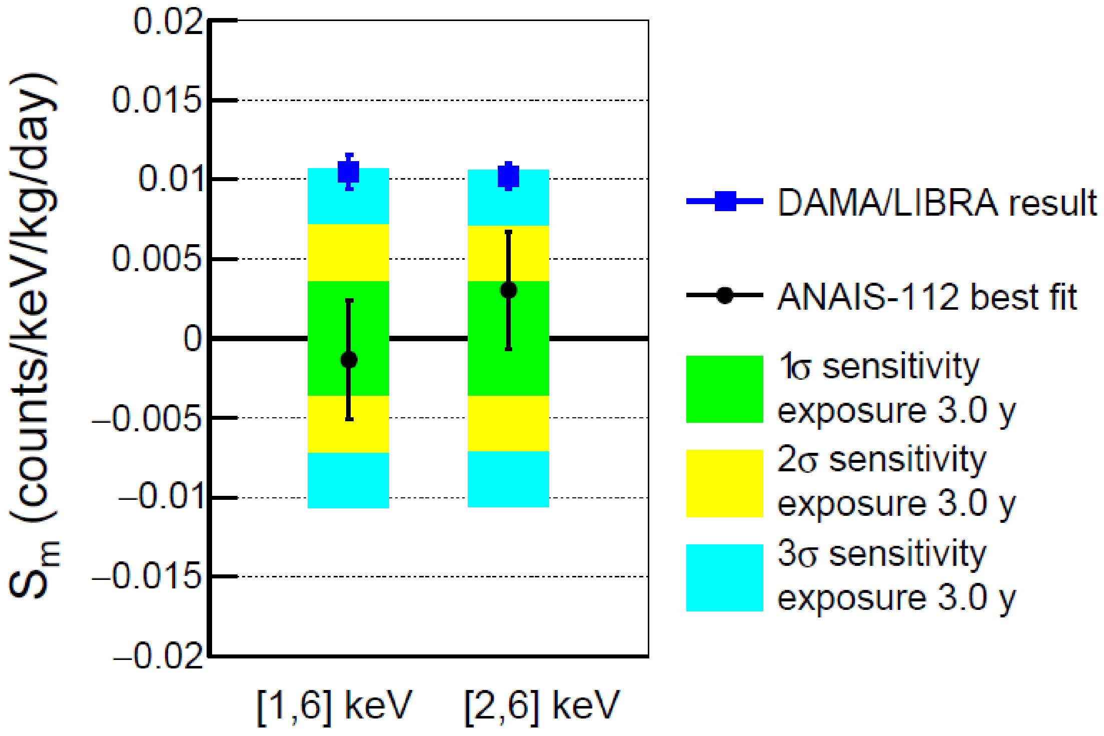
arXiv:2404.17348



3-year annual modulation with BDT cut







Best fit modulation amplitudes **compatible with zero** at $\sim 1\sigma$

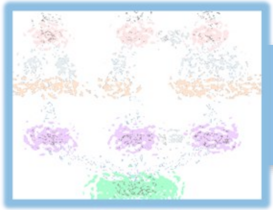
Best fit **incompatible with DAMA/LIBRA** at 3.2 (1.9) σ for [1-6] ([2-6]) keV

Sensitivity with 3 years data: 2.8σ for [1-6] and [2-6] keV

Outline

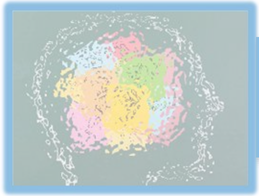
Introduction to dark matter

The ANAIS–112 experiment



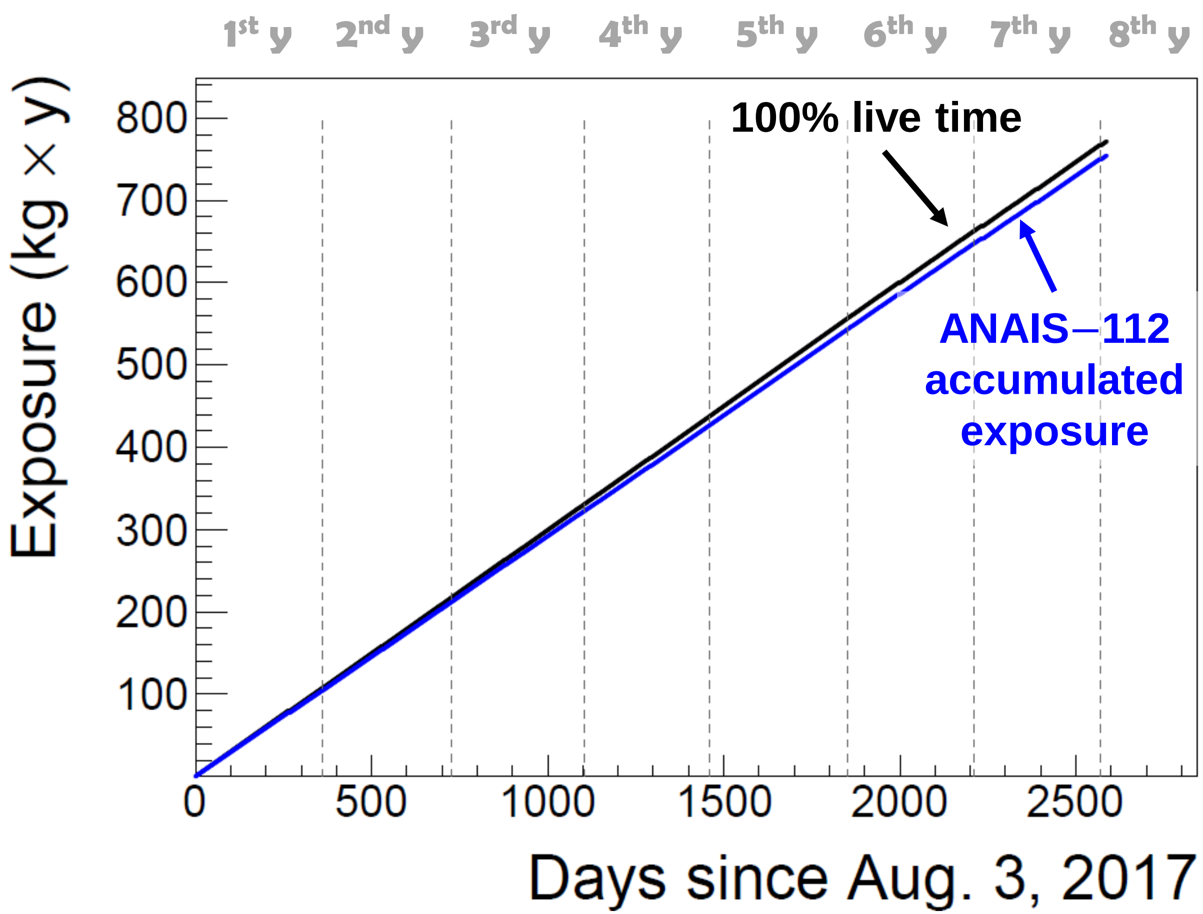
Improving filtering protocols with Machine Learning

Annual modulation results with 6 years



Summary and outlook

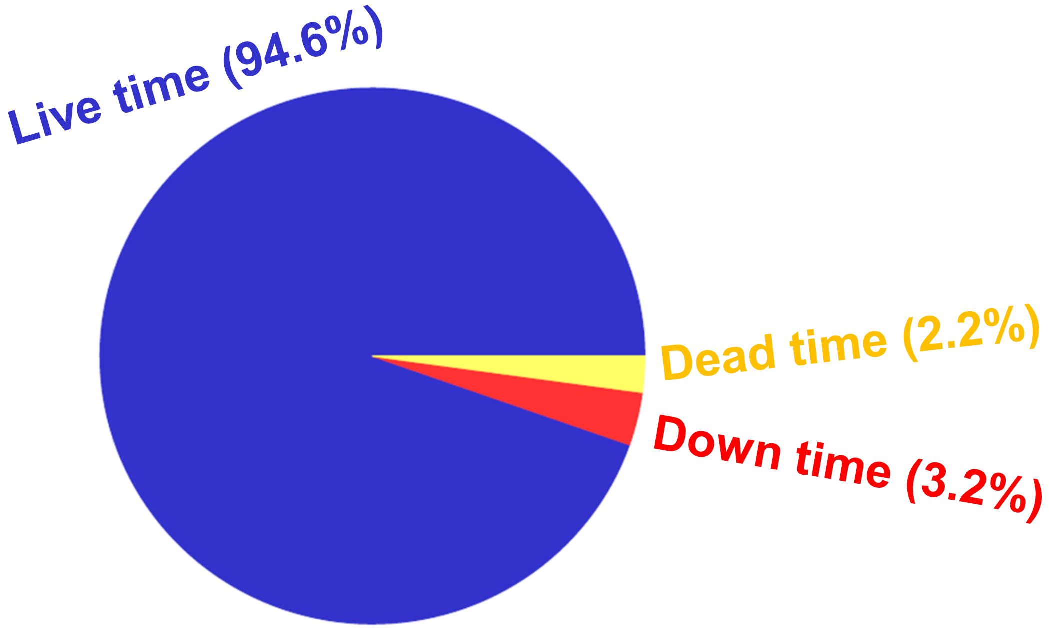
Data-taking overview



About 95% of live time

ANAIS–112 accumulated exposure

754.89 kg×y
@ September 2, 2024



Seven-year exposure has already been completed this August

Annual modulation results with 6 years

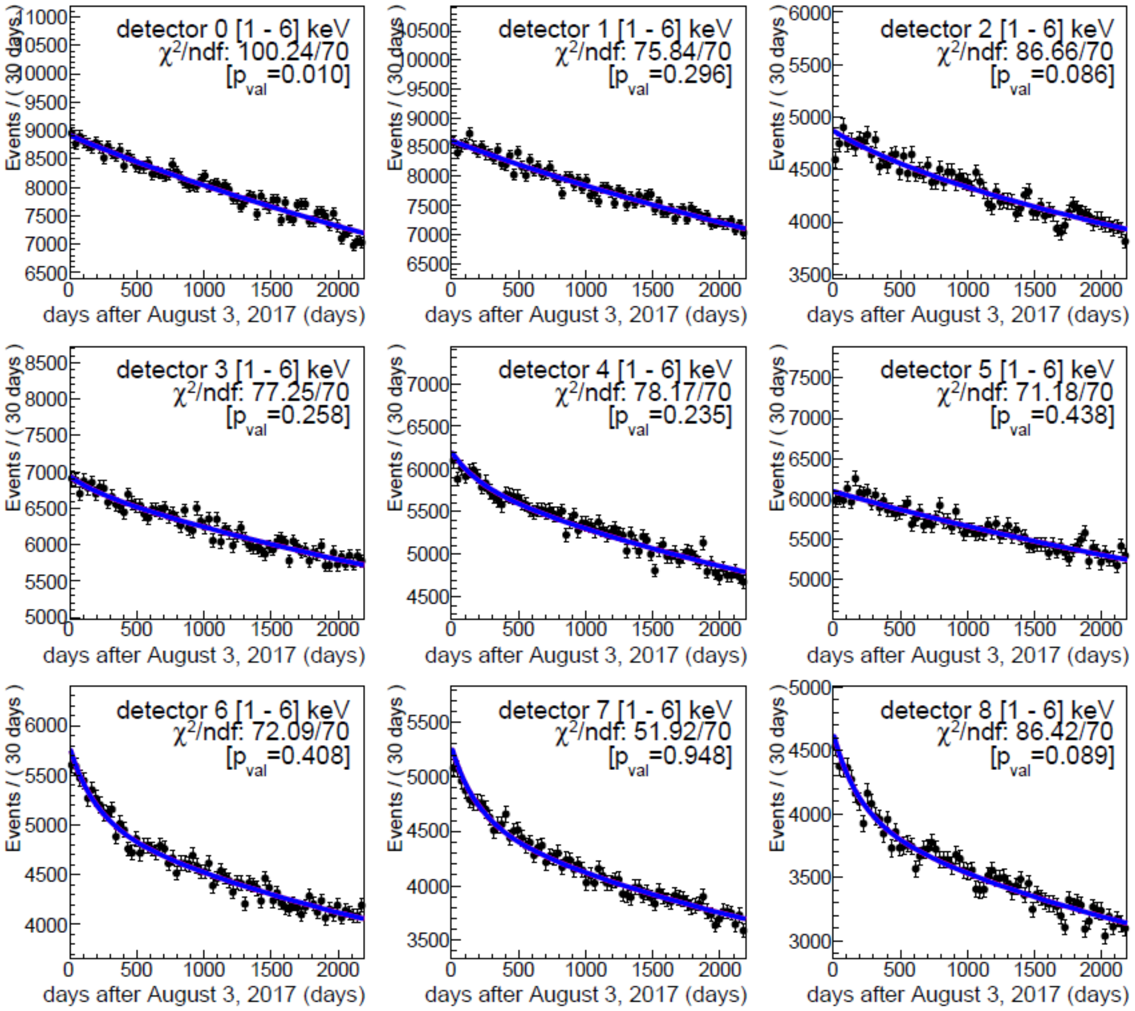
642.05 kg × y

[1-6] keV: **4.2σ**

Null hyp χ^2/ndf: 699.60/639 [p_{val}=0.048]

Mod hyp χ^2/ndf: 699.53/638 [p_{val}=0.046]

S_m = (0.0007 ± 0.0025) (cpd/kg/keV)

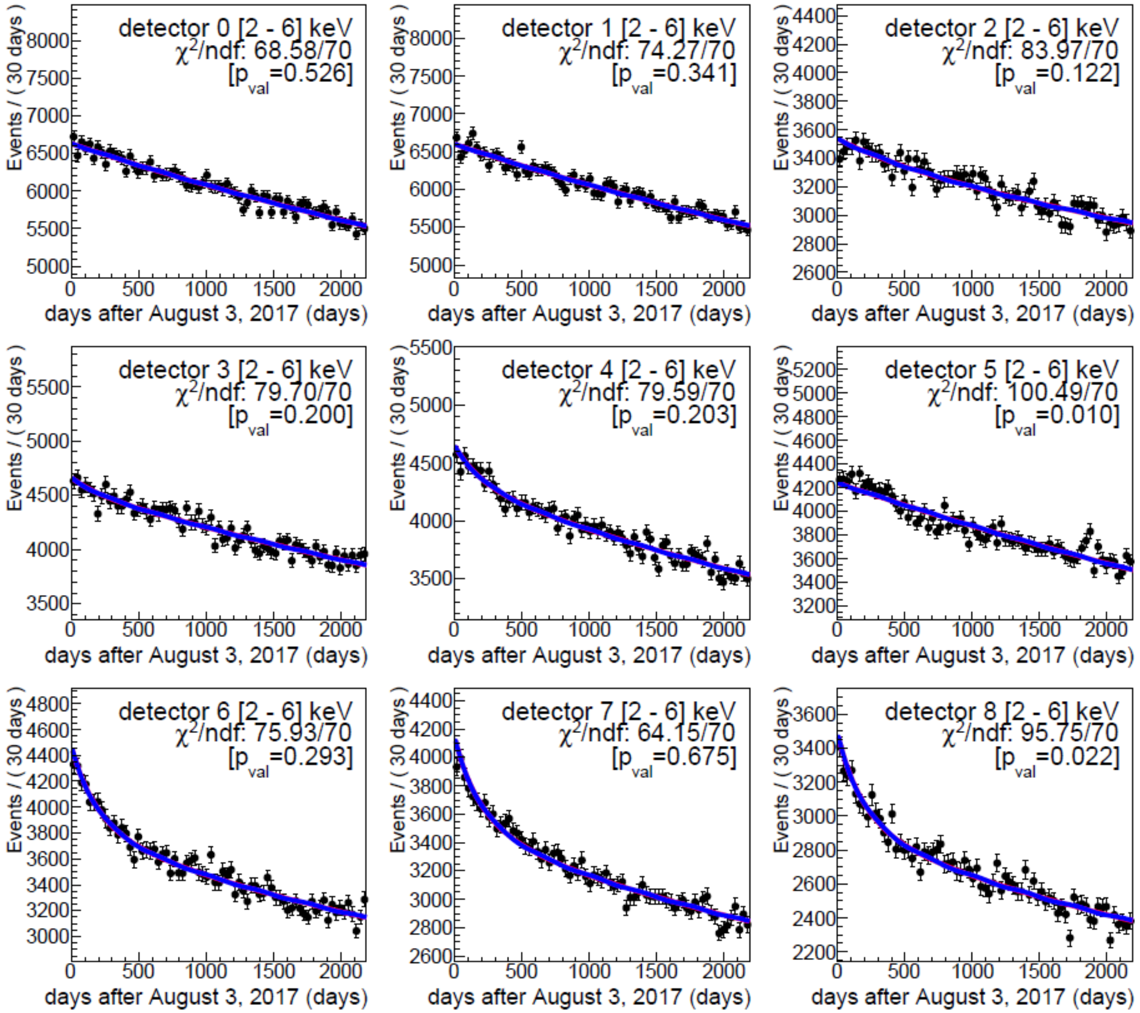


[2-6] keV: **4.1σ**

Null hyp χ^2/ndf: 723.68/639 [p_{val}=0.011]

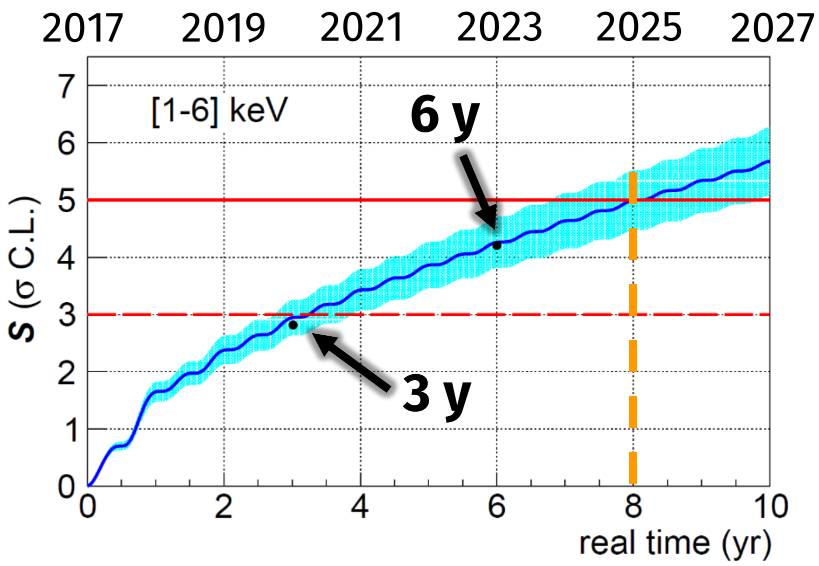
Mod hyp χ^2/ndf: 722.17/638 [p_{val}=0.011]

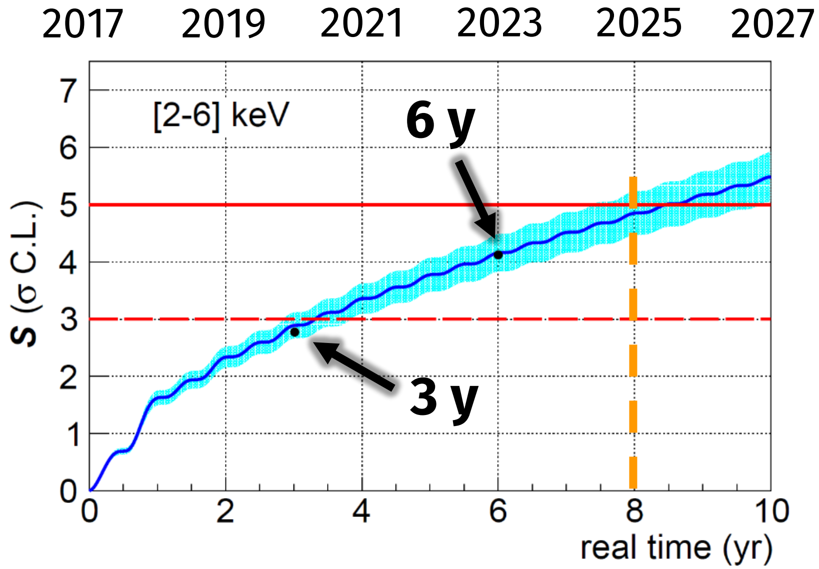
S_m = (0.0030 ± 0.0025) (cpd/kg/keV)

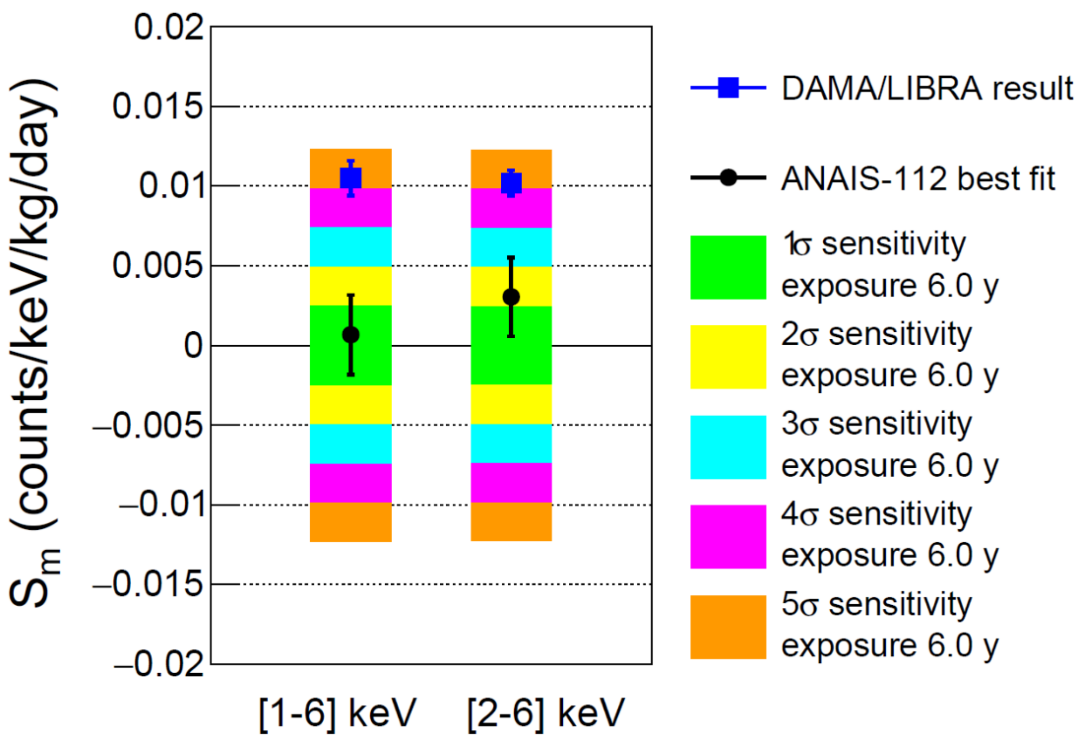


PRELIMINARY

Annual modulation results with 6 years







PRELIMINARY

Best fit modulation amplitudes **compatible with zero** at $\sim 1\sigma$

Best fit **incompatible with DAMA/LIBRA** at 3.9 (2.9) σ for [1-6] ([2-6]) keV

Sensitivity with 6 years data: 4.2 (4.1) σ for [1-6] ([2-6]) keV

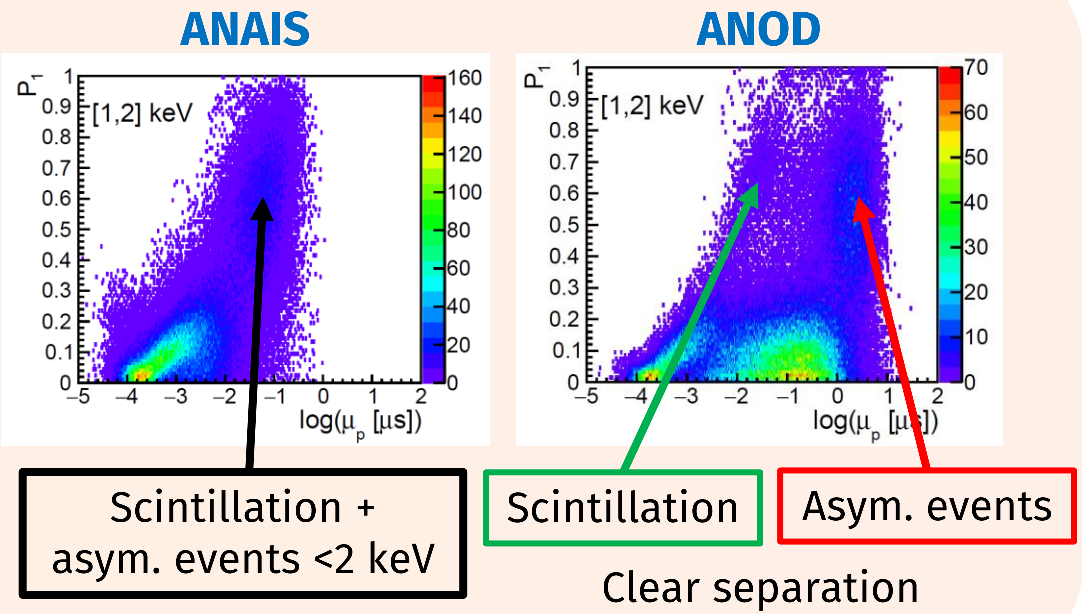
5σ sensitivity in late 2025

Next steps

New parallel DAQ system in ANAIS–112

To better understand (and eventually remove) anomalous events appearing at low energy with asymmetric light-sharing

- Extending the digitization window from 1.25 to **8 μs** and free of dead time (**ANOD**, Anais NO Dead time)
- ANOD is working smoothly since winter 2023 (CAEN DT5730, 8 channels)
- By now, only 4 crystals (8 PMTs) are readout, but **very promising results**! We have acquired a VX2730 CAEN card (32 channels, 14 bit, 500 MS/s, memory 83 MS/ch) that will allow to digitize the 9 detectors + blank module



Clear separation in the mean time!

Improving the background model

Understanding the background evolution is essential for the modulation fit

- Using the full non-blinded information [9 detectors, >7 years]
- Adding full PMT description + surface components
- **Multiparametric fit** to the different components present in the bkg model

Improving ML training populations

Simulating pulses through the response function of ANAIS–112 detectors

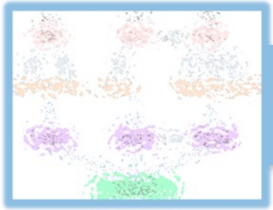
ANAIS+

Replacing the PMTs by SiPMs (at low T)

Outline

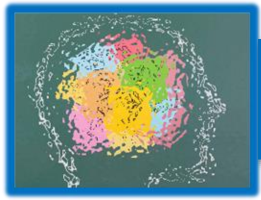
Introduction to dark matter

The ANAIS–112 experiment



Improving filtering protocols with Machine Learning

Annual modulation results with 6 years



Summary and outlook

Summary and outlook

- ANAIS–112 is leading the international efforts in the **independent test** of the DAMA/LIBRA signal, working properly after 7 years of data-taking
- Low-energy event selection and sensitivity have been improved with **machine-learning techniques**
- Preliminary results for **6 years**: ANAIS–112 is compatible with the absence of modulation and incompatible with the DAMA/LIBRA signal at $\mathbf{4\sigma}$ ($\mathbf{3\sigma}$) in [1-6] keV ([2-6] keV), for a sensitivity of $\mathbf{4.2\sigma}$ ($\mathbf{4.1\sigma}$) at [1-6] keV ([2-6] keV)
- **5σ sensitivity in late 2025**

- **New parallel DAQ** in ANAIS working since winter 2023 for 4 crystals. Promising results for improving PSD event selection. 9 crystals + blank at the end of the year
- Plan to improve our **background model** with the accumulated exposure
- ANAIS–112/COSINE–100 working to combine results. Preliminary results in presented this summer in IDM 2024

- **Open Data Policy**: ANAIS–112 3-year annual modulation analysis and the reanalysis can be downloaded at https://www.origins-cluster.de/odsl/dark-matter-data-center/available-datasets/anais. 6 years in the near future

Acknowledgements

Thank you for your attention!

ANAIS research team

J. Amaré, J. Apilluelo, S. Cebrián, D. Cintas, **I. Coarasa**, E. García, M. Martínez, Y. Ortigoza, A. Ortiz de Solórzano, T. Pardo, J. Puimedón, M. L. Sarsa

Centro de Astropartículas y Física de Altas Energías
Universidad Zaragoza

GOBIERNO DE ARAGON

Plan de Recuperación, Transformación y Resiliencia

Financiado por la Unión Europea
NextGenerationEU

MultiDark
Multimessenger Approach for Dark Matter Detection

ANAIS experiment operation is presently financially supported by MICIU/AEI/10.13039/501100011033 (Grants No. PID2022-138357NB-C21 and PID2019-104374GB-I00), and Unión Europea NextGenerationEU/PRTR (AstroHEP) and the Gobierno de Aragón. Funding from Grant FPA2017-83133-P, Consolider-Ingenio 2010 Programme under grants MULTIDARK CSD2009-00064 and CPAN CSD2007-00042, the Gobierno de Aragón and the LSC Consortium made possible the setting-up of the detectors. The technical support from LSC and GIFNA staff as well as from Servicios de Apoyo a la Investigación de la Universidad de Zaragoza (SAIs) is warmly acknowledged.

Backup

Annual modulation results with 5 years

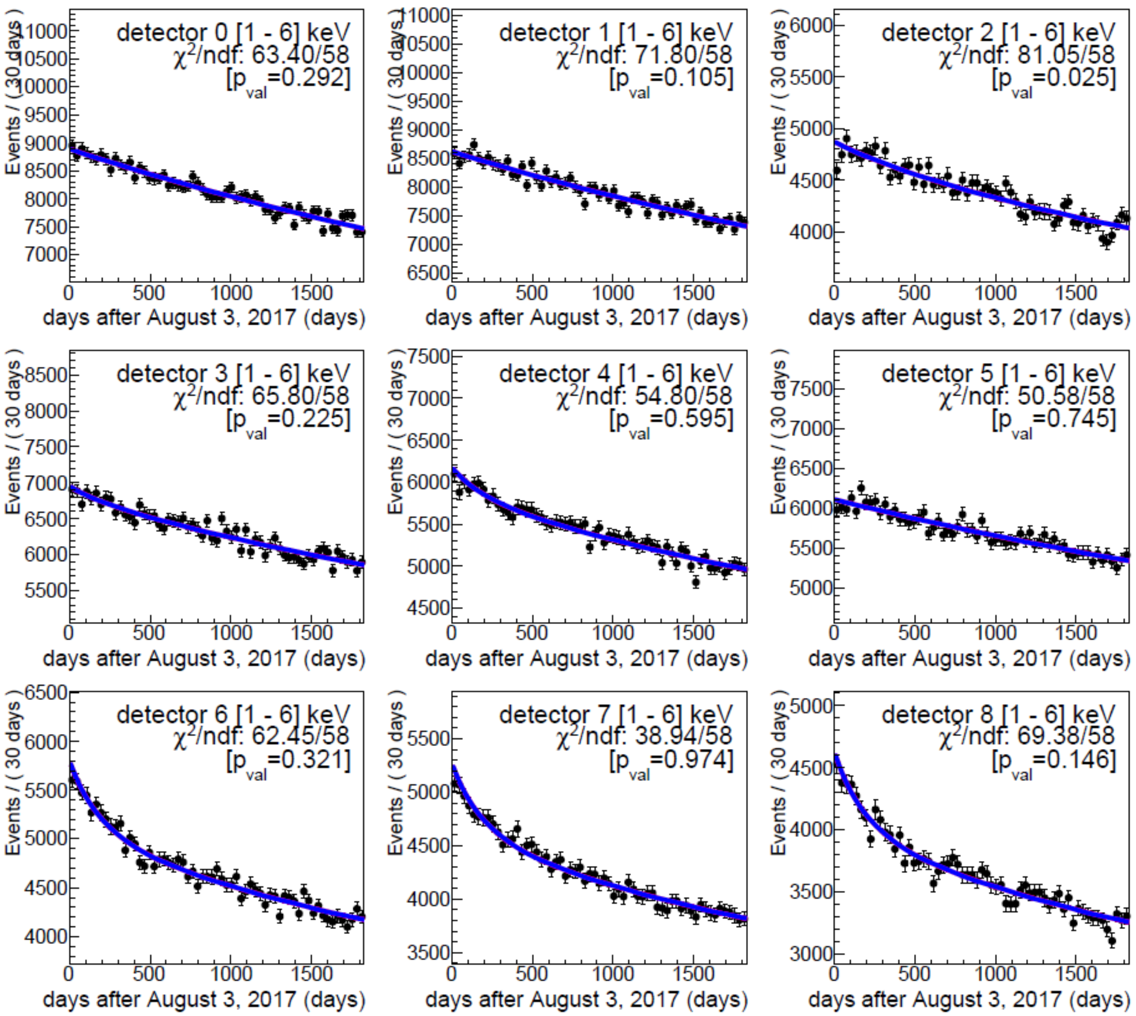
537.44 kg × y

[1-6] keV: **3.8σ**

Null hyp χ^2/ndf: 558.21/531 [p_{val}=0.200]

Mod hyp χ^2/ndf: 557.96/530 [p_{val}=0.194]

S_m = (0.0014 ± 0.0028) (cpd/kg/keV)

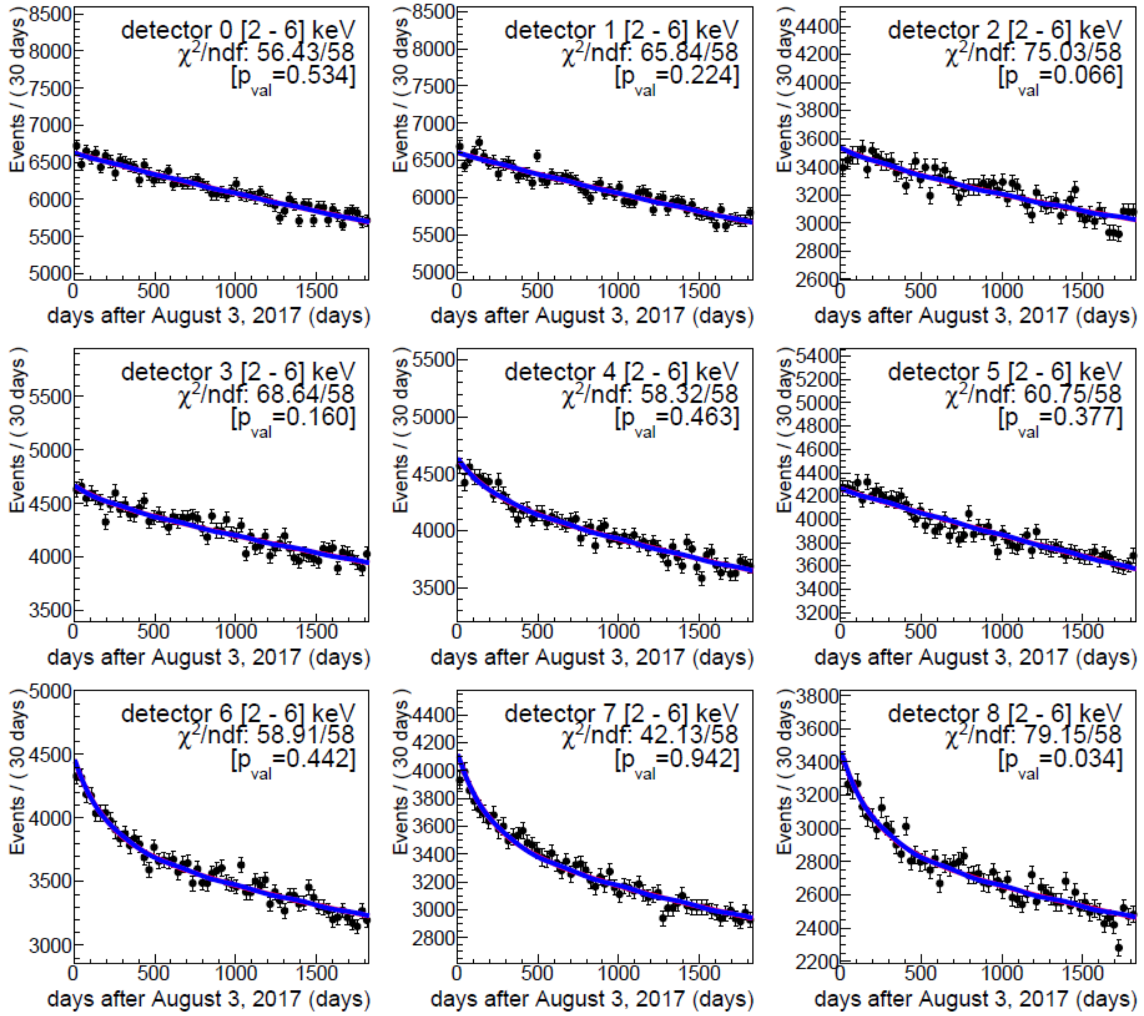


[2-6] keV: **3.7σ**

Null hyp χ^2/ndf: 566.38/531 [p_{val}=0.139]

Mod hyp χ^2/ndf: 565.00/530 [p_{val}=0.142]

S_m = (0.0032 ± 0.0027) (cpd/kg/keV)



PRELIMINARY