Machine Learning for Rare Event Searches

Iván Coarasa

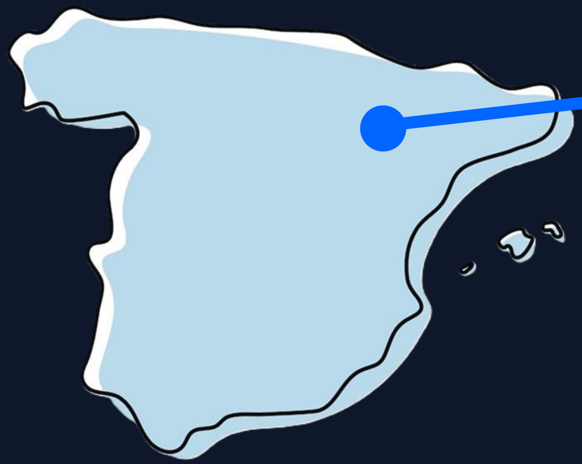
icoarasa@unizar.es

The 4th school on Computing Challenges (COMCHA)
8-15 April 2026, Zaragoza

Centro de Astropartículas y Física de Altas Energías
Universidad Zaragoza

COMCHA
Computing challenges

CAPA: Centro de Astropartículas y Física de Altas Energías



CAPA
Centro de Astropartículas y
Física de Altas Energías
Universidad Zaragoza

Research lines – Experimental area

Dark matter direct detection
- **ANAIS-112**
- TREX-DM
- **DarkSide-20k**/DART

Axion and light particle searches
- IAXO/BabyIAXO
- RADES

Neutrinos physics
- CROSS
- NEXT
- Hyper-Kamiokande

Novel detector development
- LiquidO
- CLOUD

Societal and environmental applications
- LABAC

https://capa.unizar.es/

Outline

1. Rare Event Searches
2. Dark Matter
3. Case Study I: The ANAIS Experiment
4. Case Study II: The DarkSide-20k Experiment

Outline

1. Rare Event Searches
2. Dark Matter
3. Case Study I: The ANAIS Experiment
4. Case Study II: The DarkSide-20k Experiment

Defining Rare Events

We are searching for interactions so rare that ton-scale detectors might only see one event per year

$$\sigma \sim 10^{-45}\ \mathbf{cm^2}$$

The statistical challenge is the "**Needle in a Haystack**" problem, where the haystack is a billion times larger than the needle

Examples:

- Dark matter direct detection
- Neutrinoless double beta decay
- Proton decay
- Charged lepton flavour violation

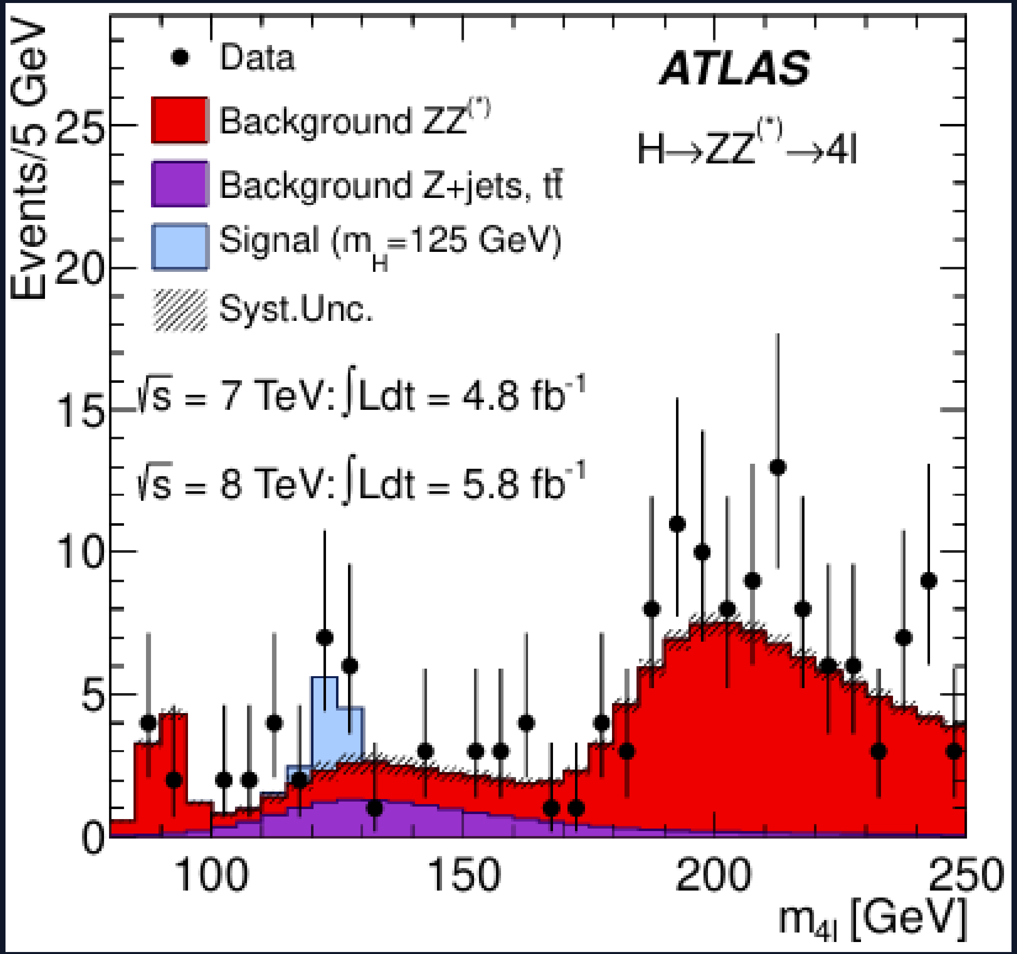
Background: Public Enemy No. 1

Background is the **antagonist**. It is not just noise; it is a mimic that obscures the signal across multiple domains:

- **Natural Radioactivity:** U/Th traces in materials
- **Cosmic Rays:** Muons and high-energy neutrons
- **Instrumental:** PMT thermal noise and electronics

*The **challenge** is not just to find something rare...*

... it's to distinguish it from other events that look almost the same



Phys. Lett. B 716(2012)1

How do we deal with background?

Common features

- The signal is extremely rare
- the background dominates
- and the differences between them can be very subtle

Everything is background!!

So any successful strategy must focus on **reducing or controlling** the background

Background mitigation strategies (I)

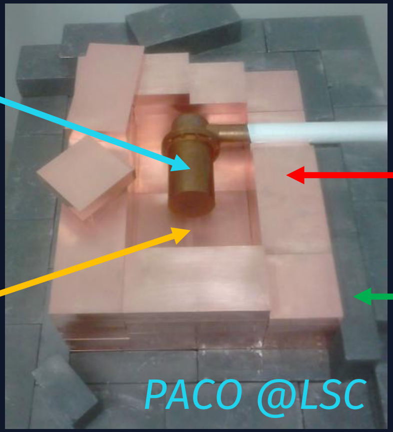
Traditional background suppression

- **Radioactivity** ⟶ Radiopure materials, passive shielding
- **Cosmic Rays** ⟶ Underground laboratories, veto systems
  - **Cosmogenics** ⟶ Moderators, absorbers

Material screening

HPGe detector

Sample

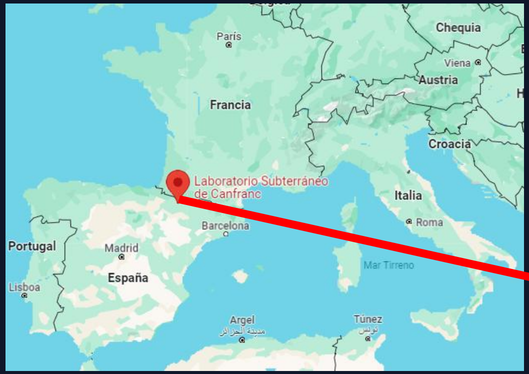


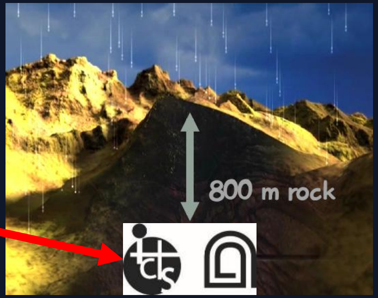
Copper

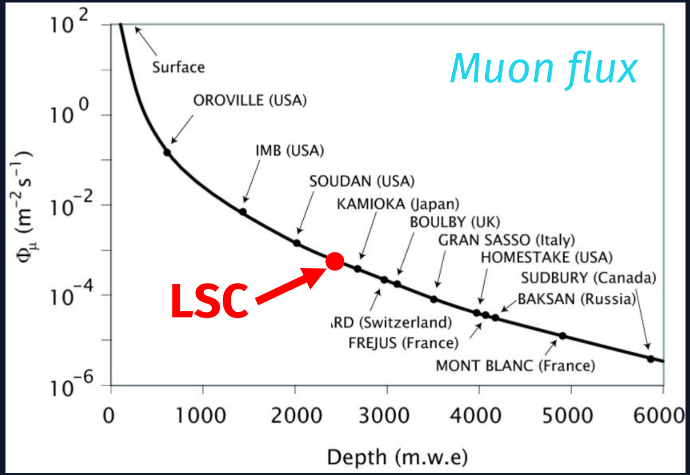
Lead

Passive shielding

Laboratorio Subterráneo de Canfranc (LSC)







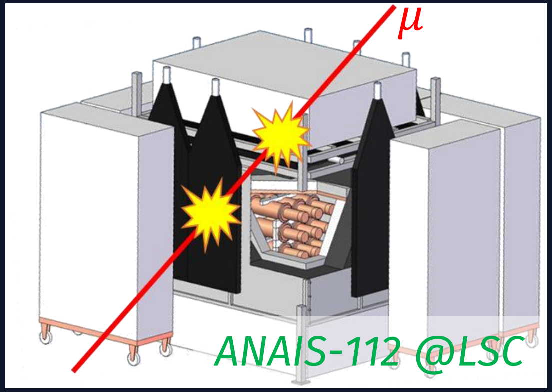
Muon flux about 10^5 lower than at surface

Background mitigation strategies (II)

Event-level discrimination

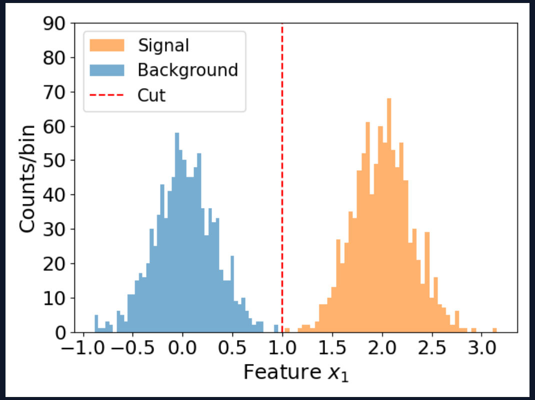
- **Cosmic Rays** ⟶ Active veto systems
- **Instrumental** ⟶ Coincidence requirement, cryogenic cooling, pulse shape discrimination (PSD), fiducialization...

Active veto system



Each muon crossing the shielding is tagged

PSD based on selected features: Traditional cut-based methods



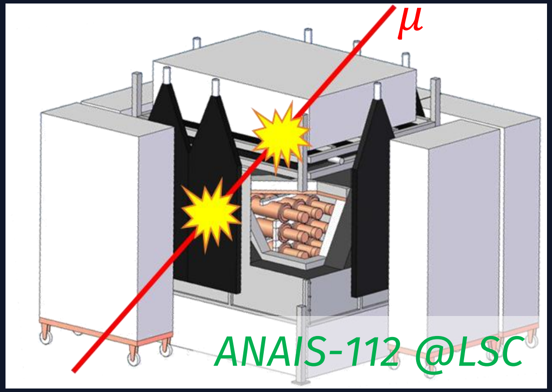
Perfect separation between signal and background based on feature x_1

Background mitigation strategies (II)

Event-level discrimination

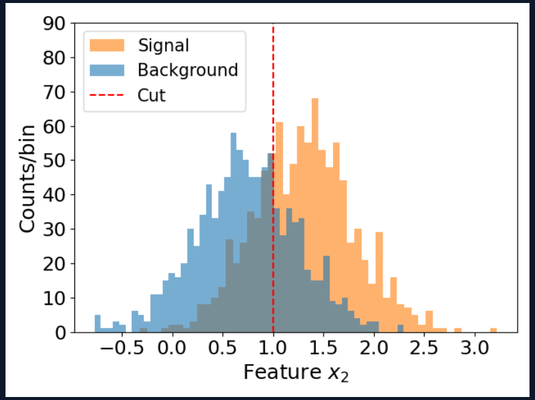
- **Cosmic Rays** ⟶ Active veto systems
- **Instrumental** ⟶ Coincidence requirement, cryogenic cooling, pulse shape discrimination (PSD), fiducialization...

Active veto system



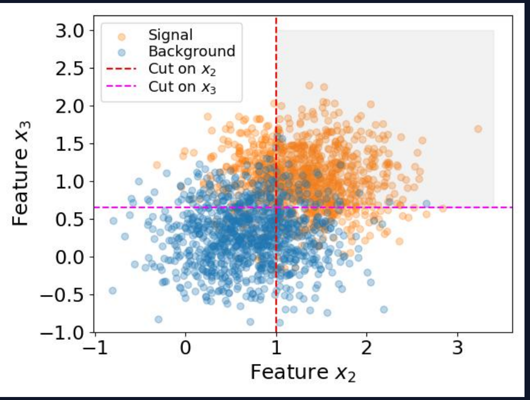
Each muon crossing the shielding is tagged

PSD based on selected features: Traditional cut-based methods



The difficulty is that signal and background often **look very similar**

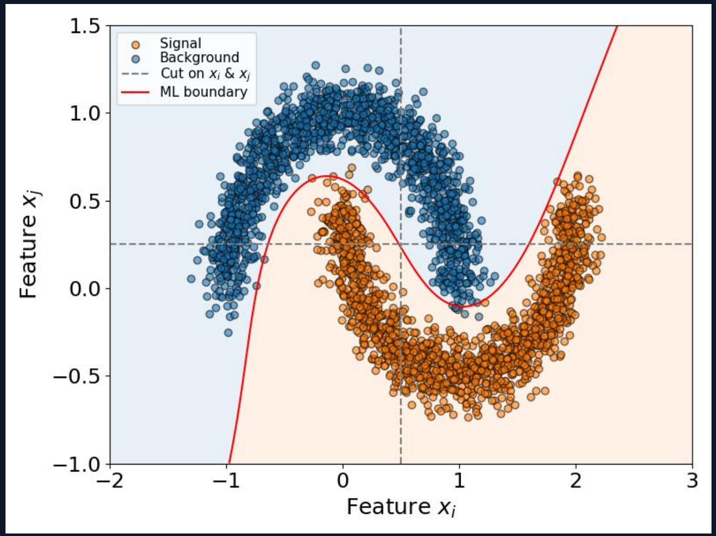
The differences can be slight and spread **across multiple features**



This is exactly where traditional cut-based methods **start to fail**

Background mitigation strategies (III)

PSD based on selected features: Machine Learning methods



Instead of designing selection rules by hand, we can train **machine learning models** to learn optimal decision boundaries directly from the data

- Machine Learning finds the best separation in a multi-dimensional parameter space
- Captures correlations between features (e.g., pulse shape vs. noise characteristics)
- Can improve signal acceptance and background rejection compared to classical cut-based methods
- Trains directly on calibration data, reducing human bias
- Flexible: works with multiple features simultaneously, even complex, non-linear ones

The four pillars of background mitigation

II Shielding

Passive and active barriers to block external radiation flux

IV Machine Learning

Algorithmic refinement to extract signal from noise

I Detector Design

Radiopurity and geometric optimization

III Data selection

Logic-based filtering of events through triggers and fiducial cuts

Final Result

Final high-purity signal sample

Outline

1. Rare Event Searches
2. Dark Matter
3. Case Study I: The ANAIS Experiment
4. Case Study II: The DarkSide-20k Experiment

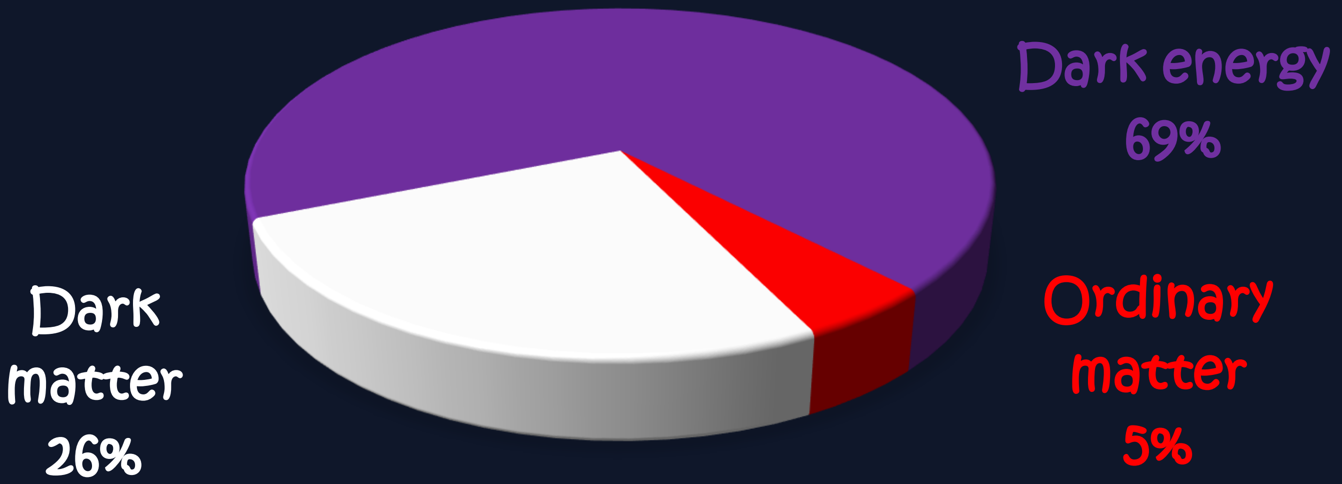
Dark matter in the universe

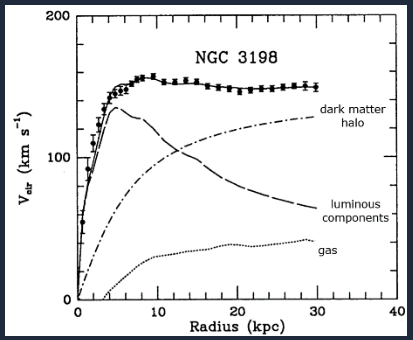
Planck's satellite (2018)

Results consistent with
ΛCDM standard model

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

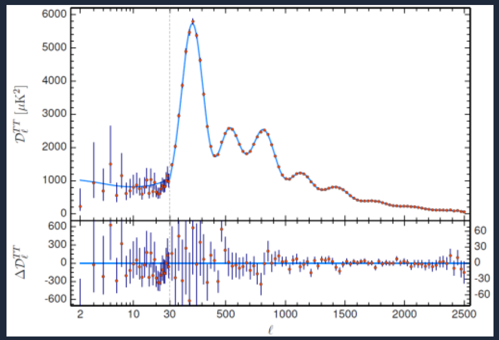
A large fraction of the Universe budget
is not explained within the Standard Model





Galactic scale

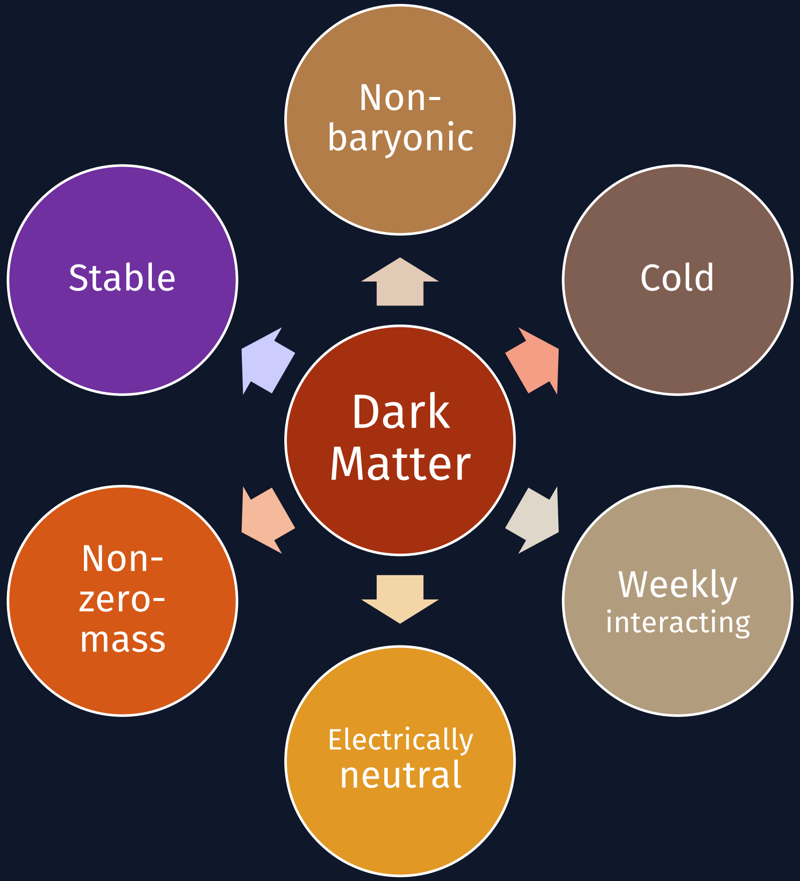
Galactic cluster scale



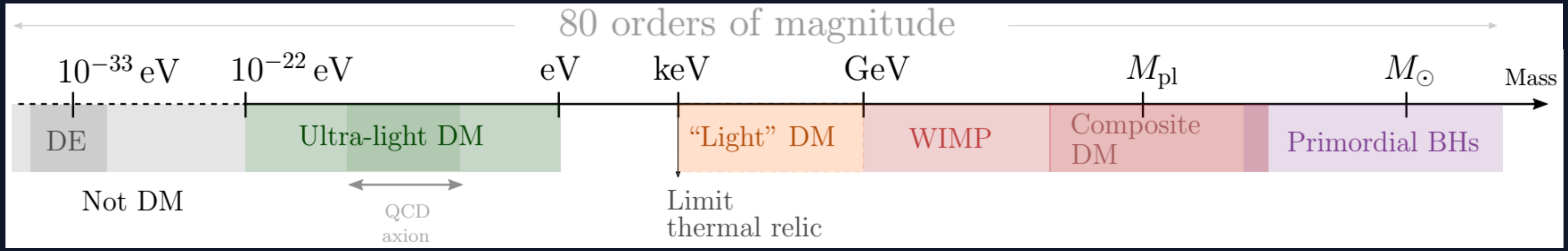
Cosmological scale

Numerous hints
at different scales

Dark matter candidates



A plethora of **DM candidates** beyond the SM having a very low interaction probability with baryonic matter

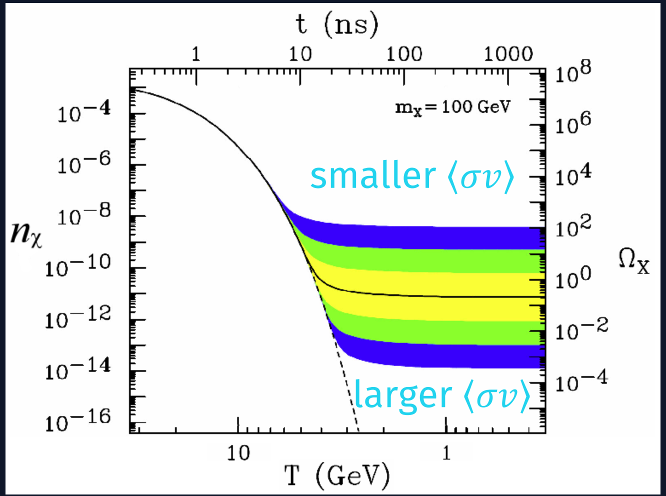


Astron. Astrophys. Rev., 29(2021)7

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

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

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



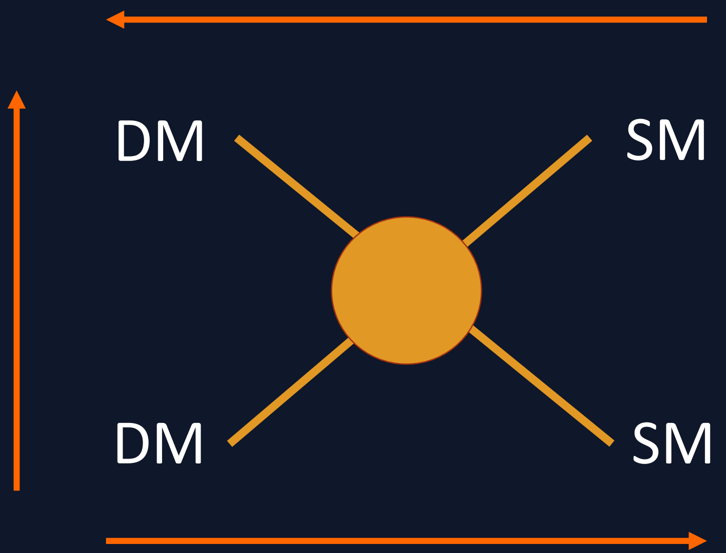
Annu. Rev. Astron. Astrophys. 48(2010)495

Dark matter detection

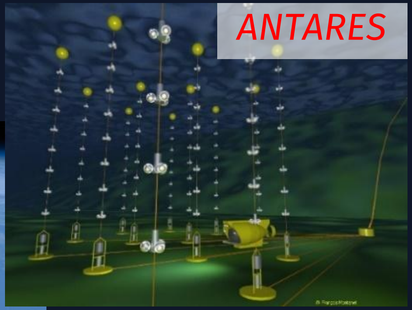
Different **complementary** strategies for detection

Direct detection (**shake it**)

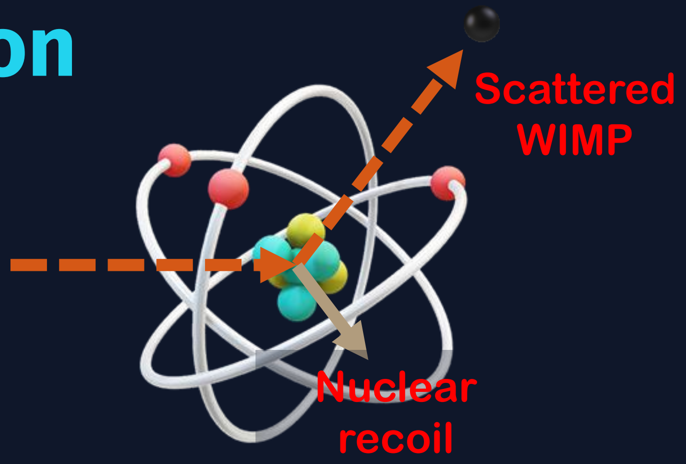
Production at colliders (**make it**)



Indirect detection (**break it**)



Dark matter direct detection



Nuclear recoils

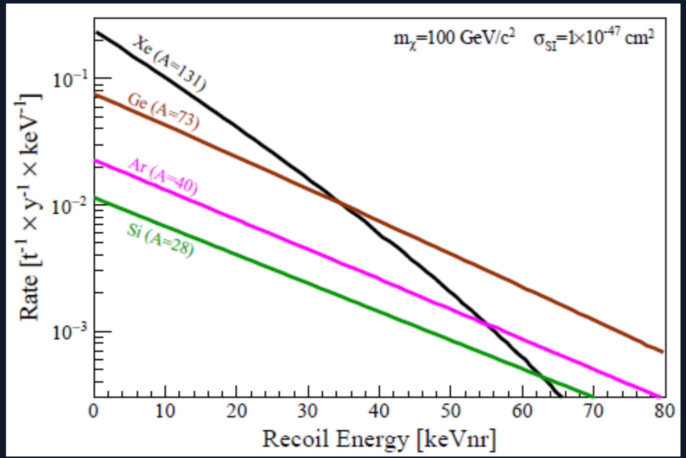
WIMPs with masses of 10-100 GeV would produce nuclear recoils with energies of 1-100 keV

Signals:

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

Detection rate

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



J. Phys. G: Nucl. Part. Phys. **46(2019)**103003

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

Continuum energy spectrum entangled with background

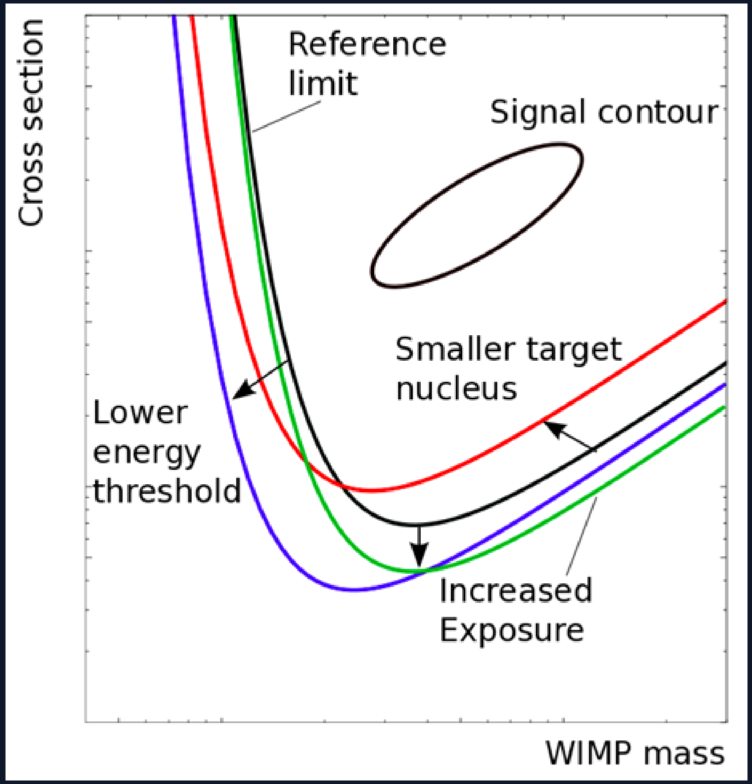
Dark matter direct detection

Direct detection experiments can be focused on different physics cases:

- Looking for an excess of events over background, considering:
  - Spin-Independent or Spin-Dependent interactions (SI /SD)
  - Different ranges of masses
  - Nuclear or Electronic Recoils (NR / ER)

- Searching for distinctive dark matter signatures:
  - **Annual modulation** in the interaction rate
  - **Directionality**

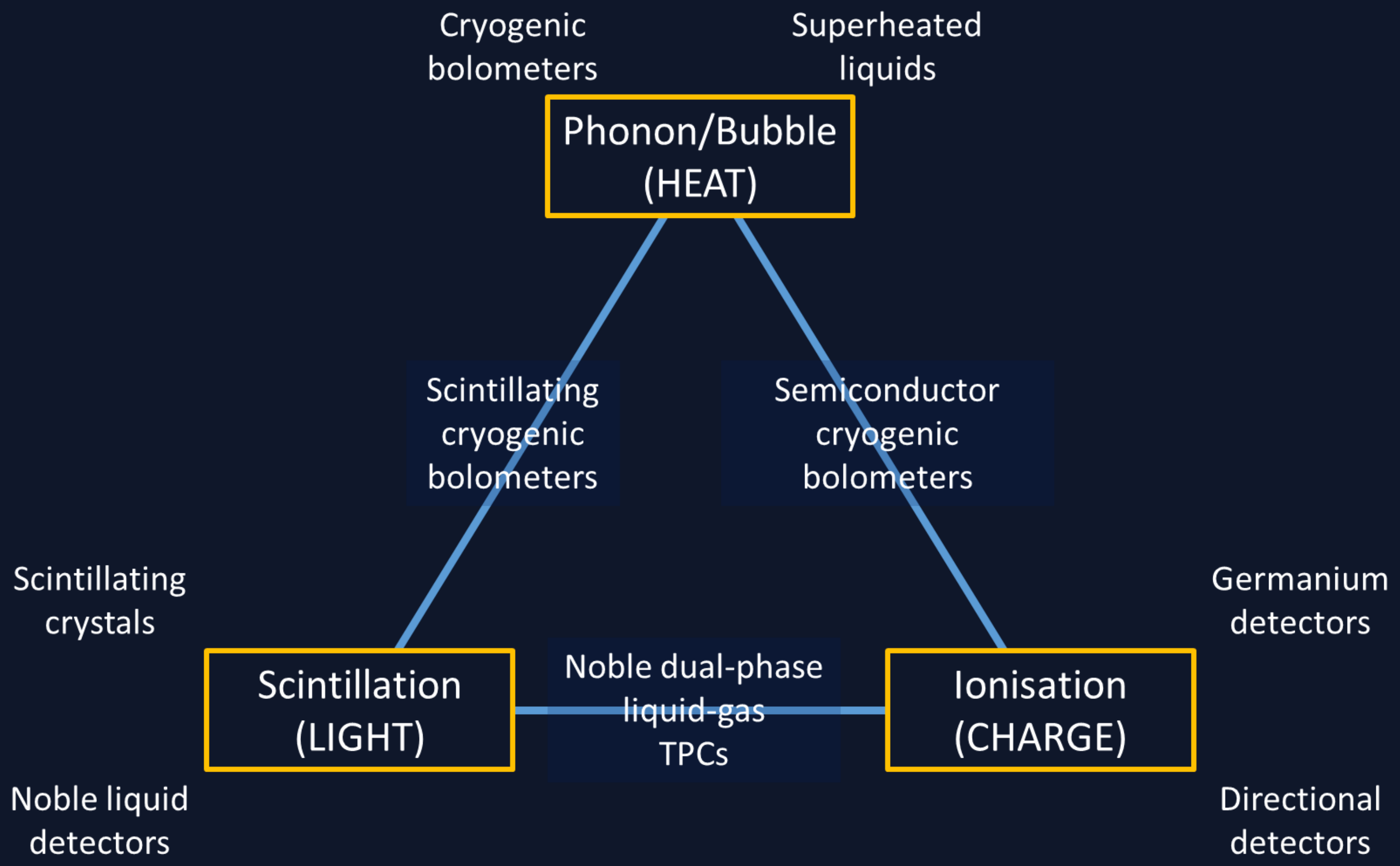
Exclusion plot (m_{WIMP}, σ)



J. Phys. G: Nucl. Part. Phys. 43(2016)1

Dark matter direct detection technologies

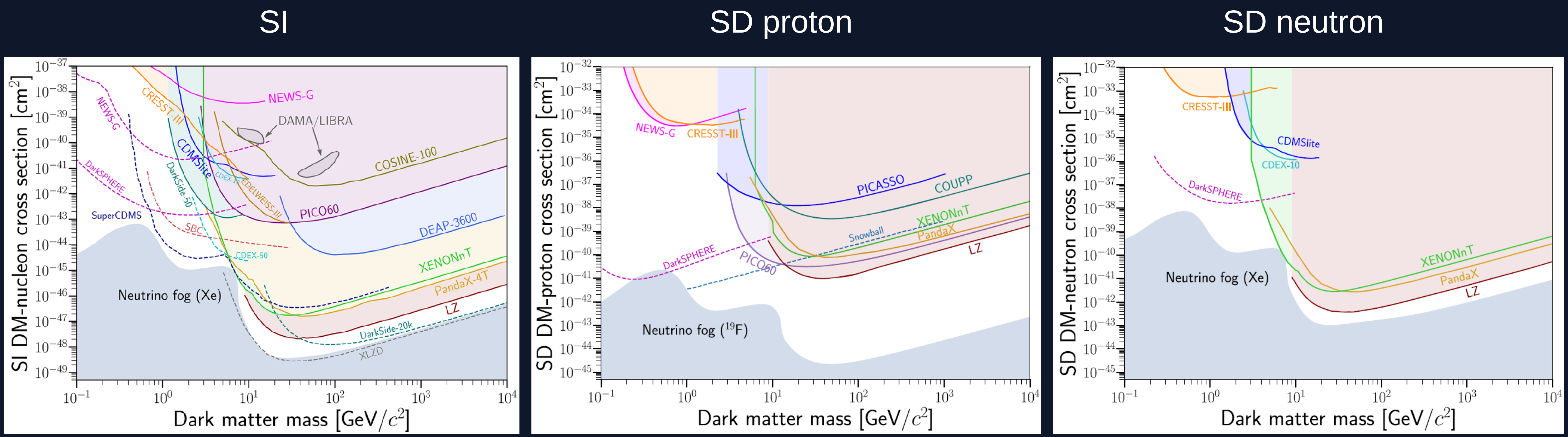
Many different **detection technologies** depending on which observable is exploited



To date, no detector exists that measures all three channels simultaneously

Summary of exclusion results

Present exclusion limits (solid lines) and future sensitivities (dashed lines) at 90% confidence level (C.L.)

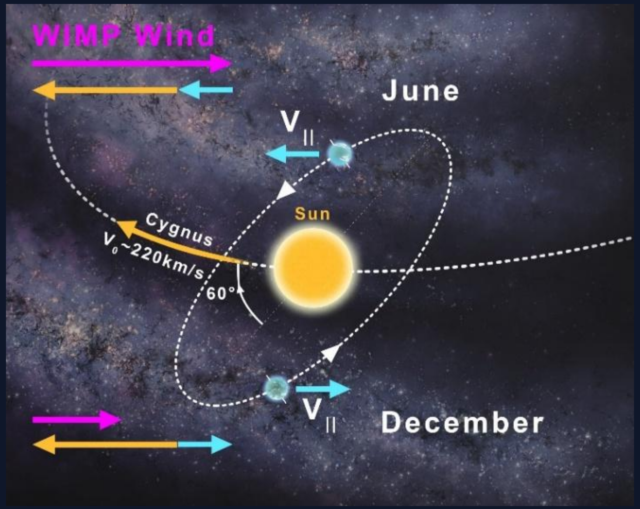


Comm. Phys. 9(2026)105

The comparison is **model-dependent**, based on assumptions about the dark matter particle and its velocity distribution in the galactic halo

Dark matter annual modulation

Due to the revolution of the Earth around the Sun, the speed of dark matter particles in the Milky Way's halo relative to Earth **varies seasonally**, producing an **annual modulation** in the rate of nuclear recoil events in detectors



J. Phys. G: Nucl. Part. Phys., 47(2020)9

$$\boxed{R(t) \approx S_0 + S_m \cdot \cos\omega(t - t_0)}$$

where $\omega = 2\pi/365\ \mathrm{d}^{-1}$, $t_0 = 152.5\ \mathrm{d}$

Detection rate would have a cosine behaviour with a yearly period and maximum around June 2nd

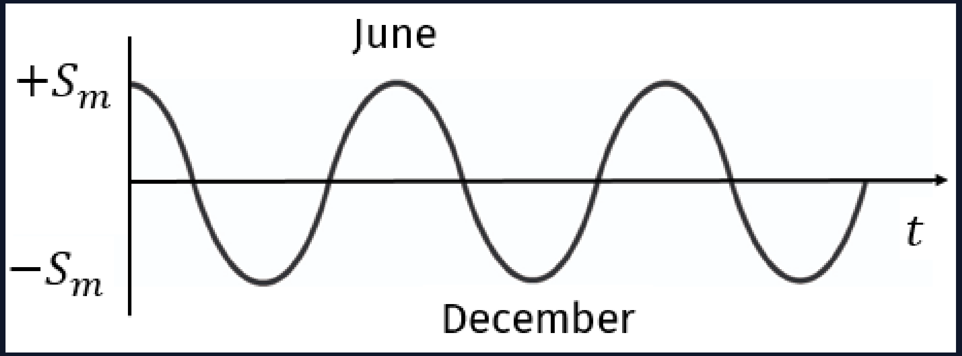
Only at low energy

Single-hit events

$S_m/S_0 \lesssim 7\%$

No background known to mimic the effect

Challenge: years of measurement in very stable conditions

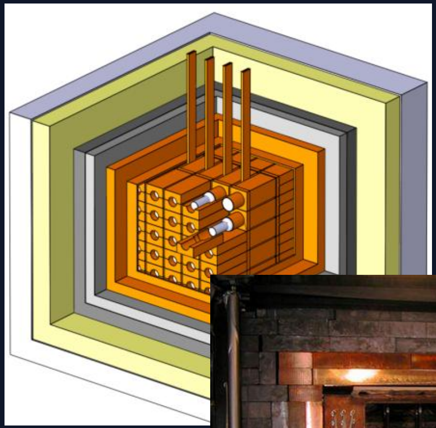


Observed annual modulation signal by the DAMA/LIBRA experiment over 28 years

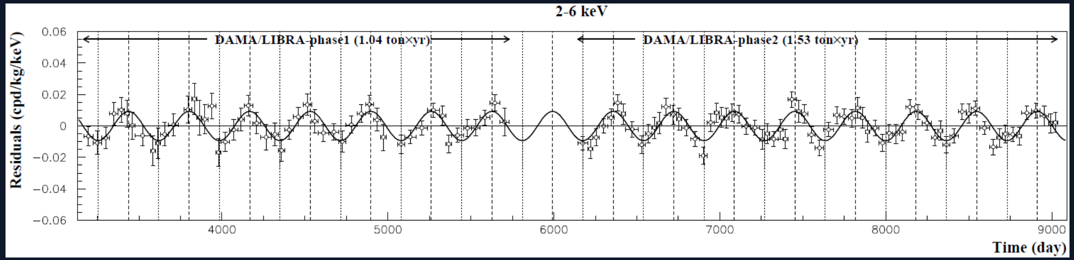
The only positive signal: DAMA/NaI & DAMA/LIBRA

DAMA experiment using NaI(Tl) @LNGS (1995-2024)

Exp. phase	Period	Mass (kg)	Annual cycles	Exposure (t×y)
DAMA/NaI	1995-2002	9×9.7	7	0.29
DAMA/LIBRA-phase1	2003-2010	25×9.7	7	1.04
DAMA/LIBRA-phase2 [Higher *QE* PMTs]	2011-2024 (2019)	25×9.7	14 (8 released)	--- (1.53)



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



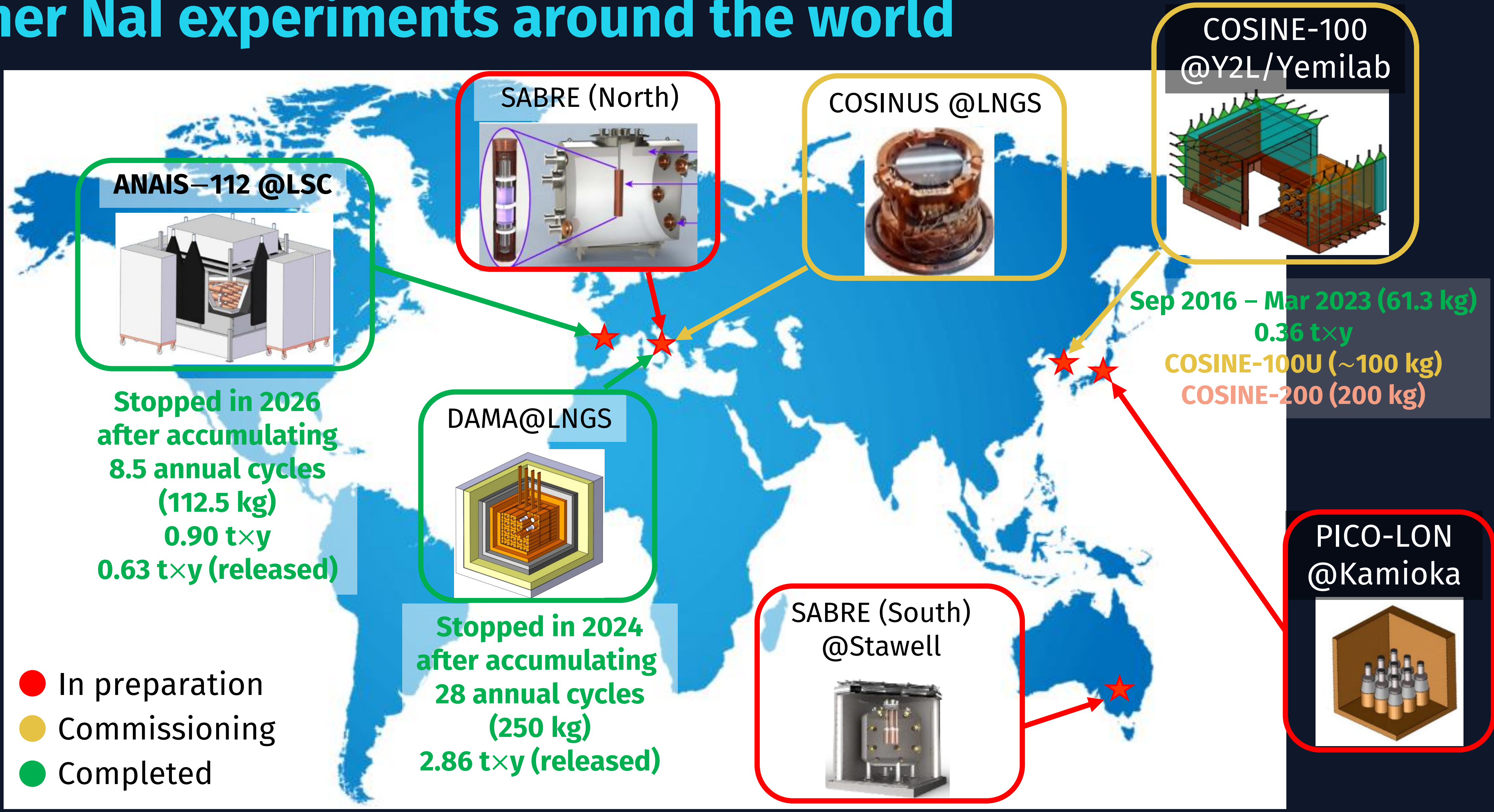
Nucl. Phys. At. Energy 22(2021)329-342

$$[2,6]\ \mathrm{keV}:\ S_m = 0.0102 \pm 0.0008\ \mathrm{cpd/kg/keV}$$

$$[1,6]\ \mathrm{keV}:\ S_m = 0.0105 \pm 0.0011\ \mathrm{cpd/kg/keV}$$

A model independent test is needed using **the same target**

Other NaI experiments around the world



Outline

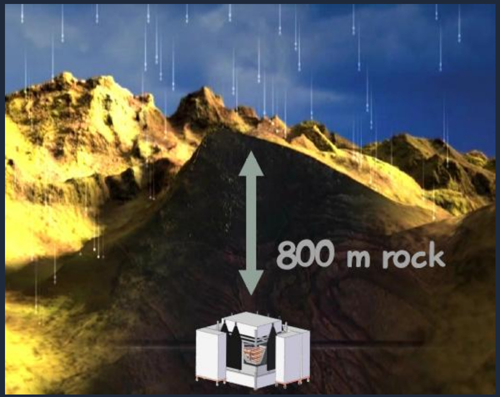
1. Rare Event Searches
2. Dark Matter
3. Case Study I: The ANAIS Experiment
4. Case Study II: The DarkSide-20k Experiment

The ANAIS experiment

GOAL **ANAIS** (*<u>A</u>nnual modulation with <u>NaI</u>(Tl) <u>s</u>cintillators*) 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.)



Experimental requirements for testing DAMA

- **Target: NaI / NaI(Tl)**
- Large exposure ✓
- Very stable operation conditions ✓
- Energy threshold at 1 keV_{ee} or below ✓
- Background level as low as possible (a few cpd/kg/keV) ✓
- Good knowledge of the detector response, in particular to nuclear recoils ✗

Eur. Phys. J. C 79(2019)228

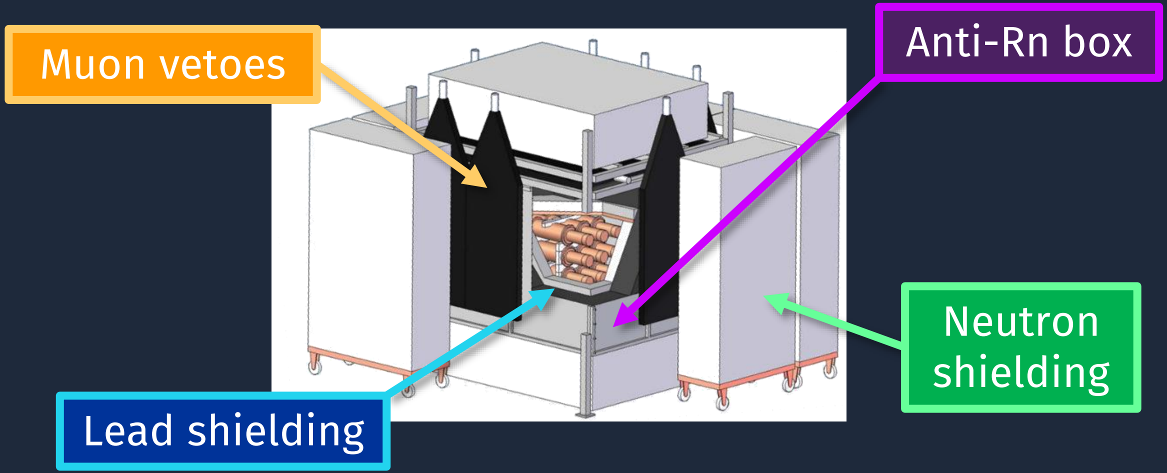
The ANAIS—112 experimental set-up

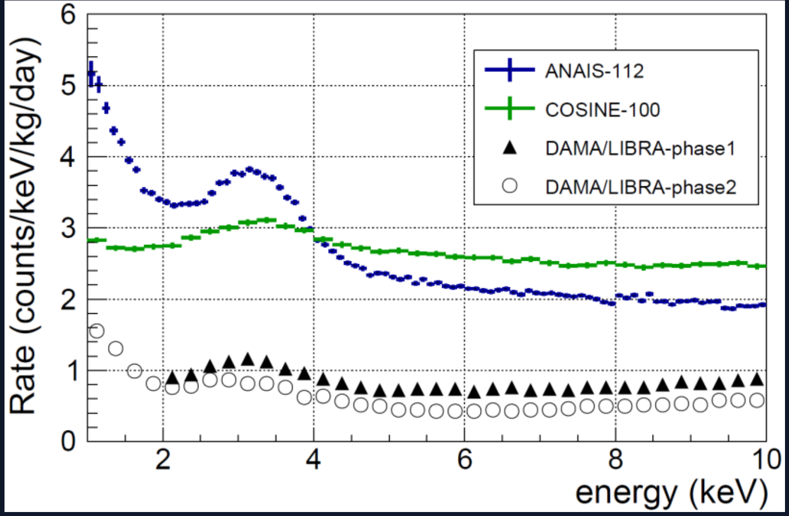
NaI(Tl) scintillating detectors

- 9 cylindrical crystals, 12 cm ϕ ×30 cm, **12.5 kg** each
- Alpha Spectra (same as COSINE)
- PMTs: Hamamatsu R12669SEL2 (QE~40%)
- Quartz windows
- Encapsulated in OFE copper
- Outstanding light collection of ~15 phe/keV

Shielding

- Gamma shielding: 10 cm of ancient Pb + 20 cm of Pb
- Anti-Rn: metallic box fluxed with N_2 gas
- **Active muon vetoes:** 16 scintillator plastics
- Neutron shielding: 40 cm Polyethylene/water tanks

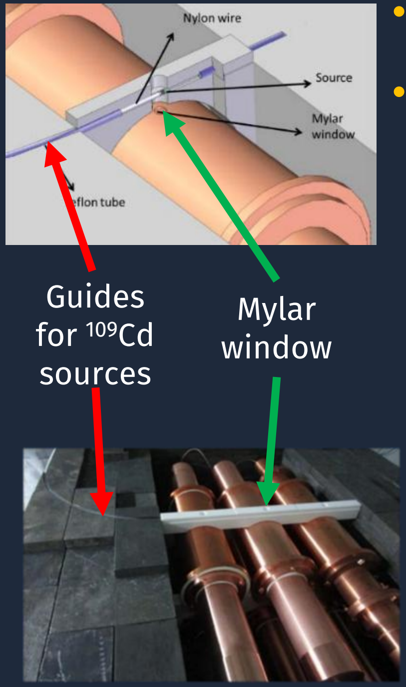




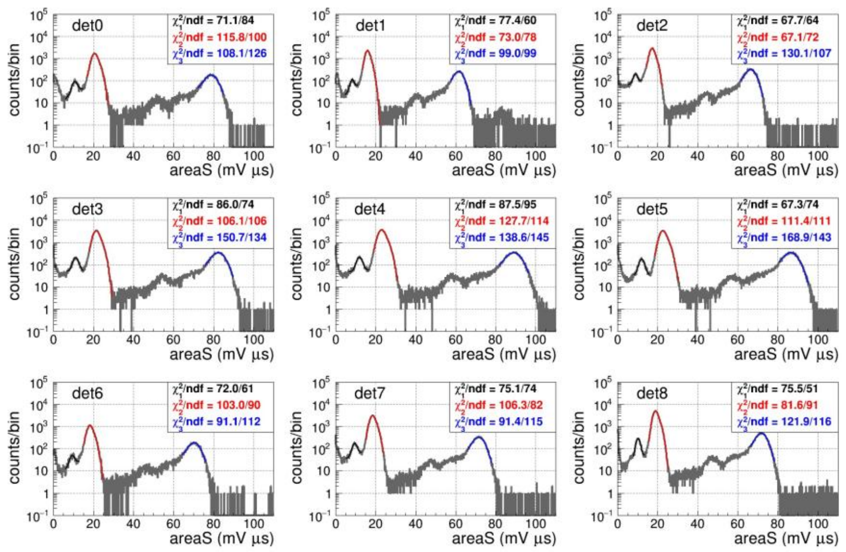
ANAIS-112: 10% unblinded (6 years)
COSINE-100: *Sci. Adv. 7 (2021) eabk2699*
D/L-ph1&2: *Prog. Part. Nucl. Phys. 114 (2020) 103810*

Low energy calibration

Linear calibration in 10-100 keV



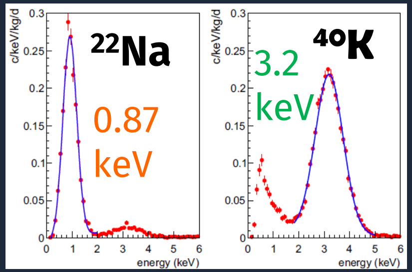
- Detectors equipped with a **Mylar window**
- Calibration with **^{109}Cd sources** (11.9, 22.6 and 88.0 keV) **every two weeks** for gain correction

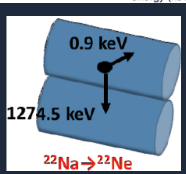


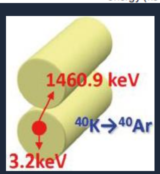
Linear calibration in 1-10 keV [ROI]

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

Coincidence events





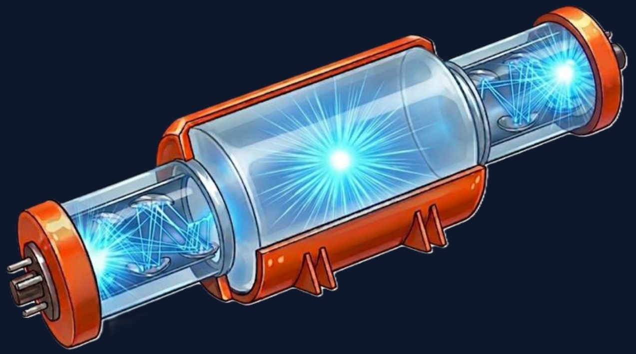


Energy threshold at 1 keV_{ee} or below

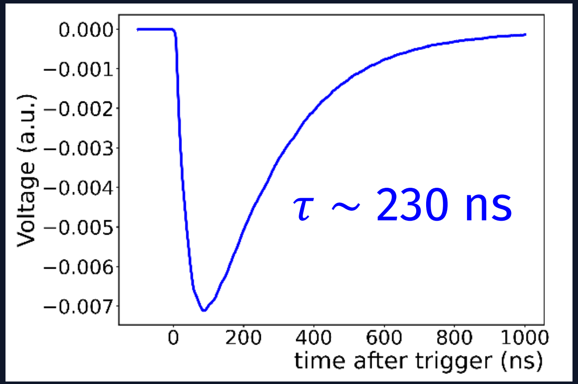
Comm. Phys. 7(2024)345

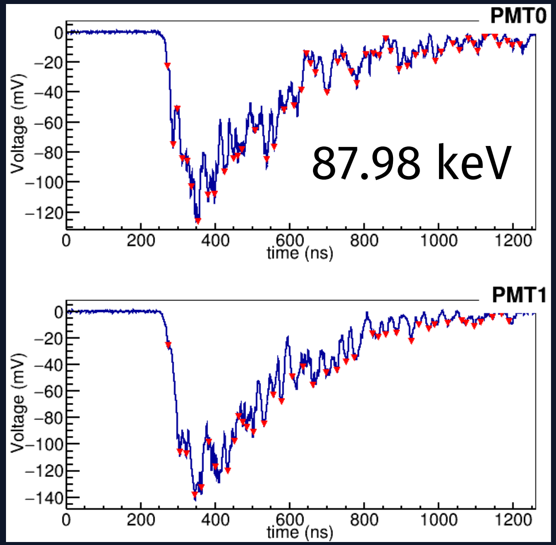
What do we expect to see?

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

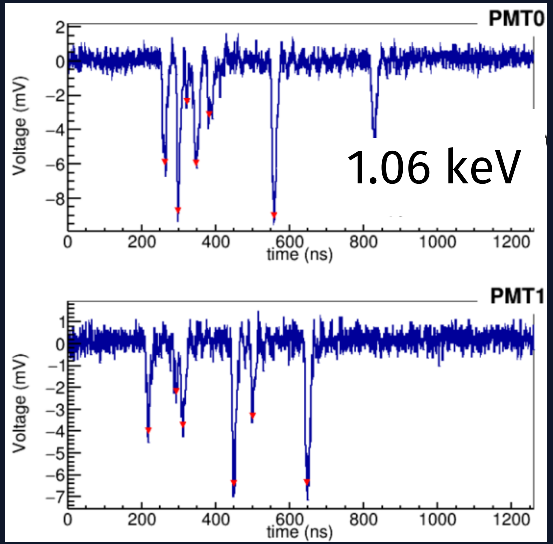


Expected signal





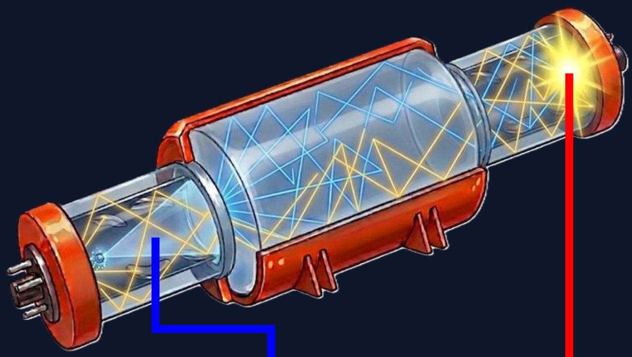
... but at low energy (15 phe/keV)

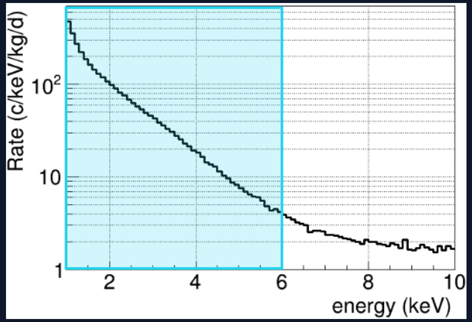


What do we actually see?

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

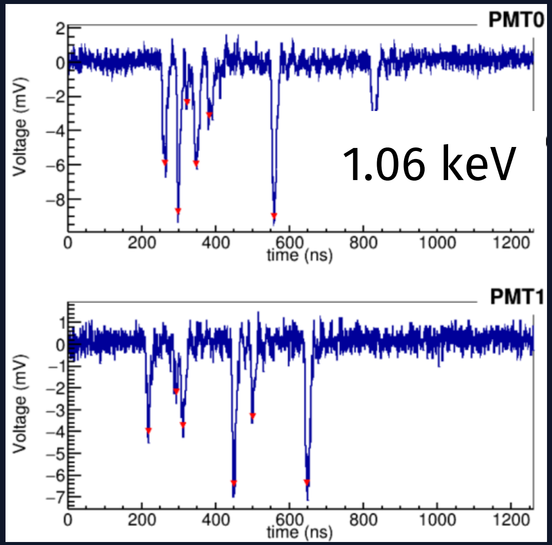
Scintillation light + light emitted in the materials surrounding the NaI(Tl) crystal



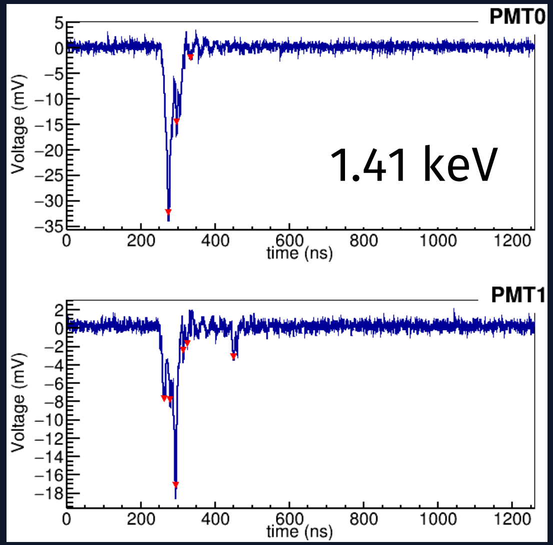


Filtering protocols are mandatory

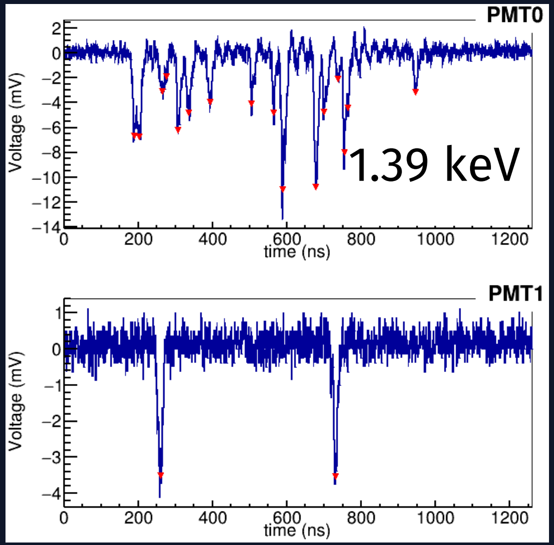
Scintillation light in NaI(Tl)



Cherenkov light @PMTs



Asymmetric light sharing @PMTs



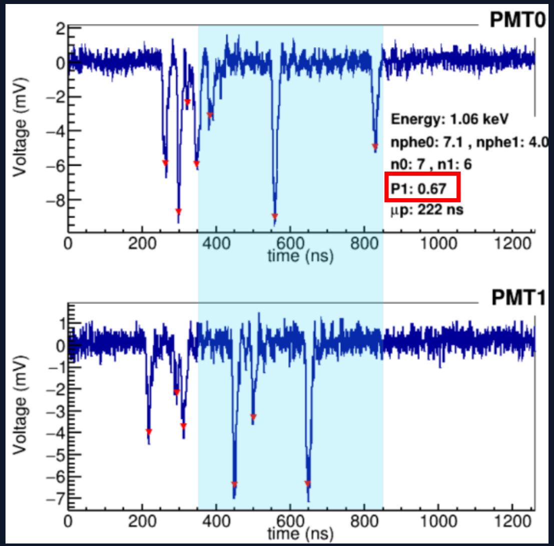
Event selection (I)

Application of event selection protocols to distinguish scintillation events from noise

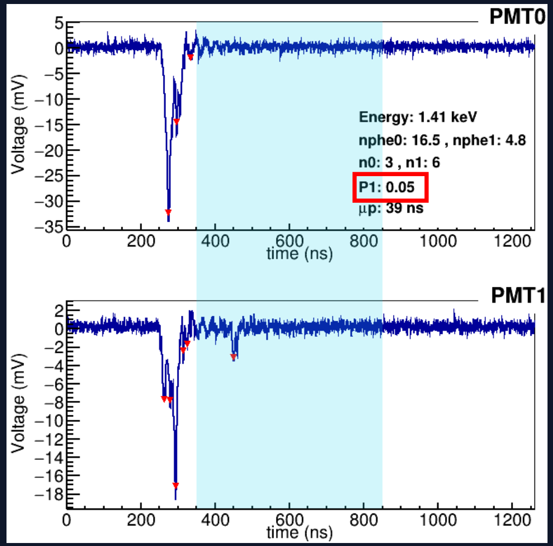
1st approach: Traditional cut-based method

1) Pulse shape cut to select pulses with NaI(Tl) scintillation constant $P_1 = \frac{\sum_{100\text{ ns}}^{600\text{ ns}} A(t)}{\sum_{0\text{ ns}}^{600\text{ ns}} A(t)}$ $\mu_p = \frac{\sum_i A_i t_i}{\sum_i A_i}$

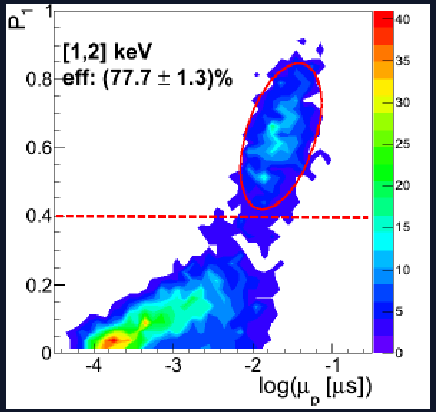
Scintillation light in NaI(Tl)



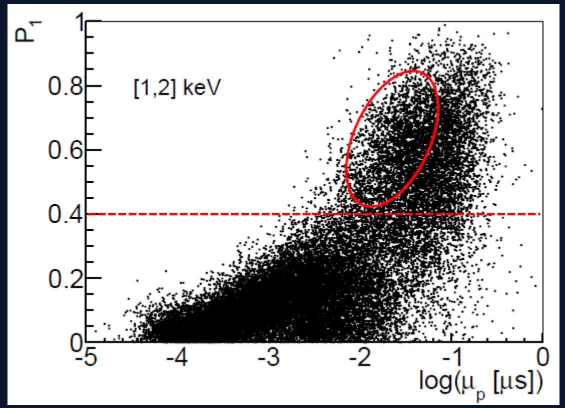
Cherenkov light @PMTs



Bulk scint. sample



Background



Cherenkov-like events and long phosphorescence or pulse tails are removed

The mean efficiency decreases to 78% in the 1–2 keV range

Event selection (I)

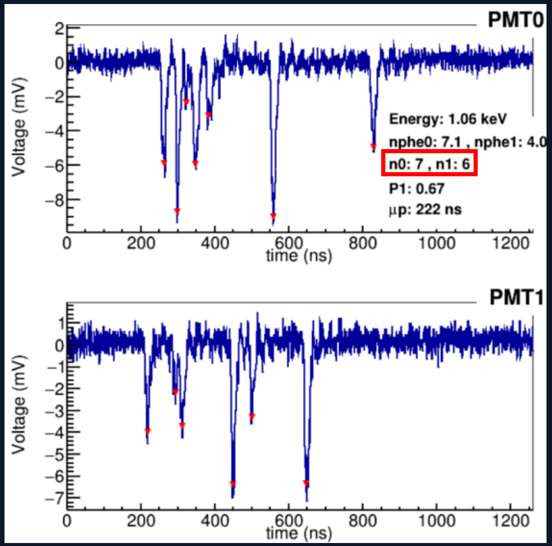
Application of event selection protocols to distinguish scintillation events from noise

1st approach: Traditional cut-based method

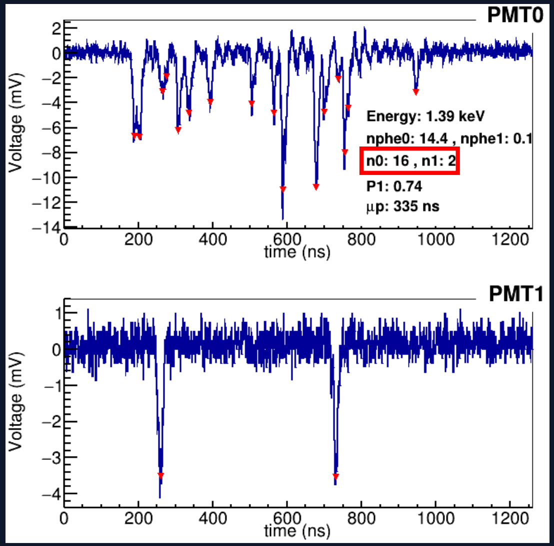
1) Pulse shape cut to select pulses with NaI(Tl) scintillation constant $\quad P_1 = \frac{\sum_{100\text{ ns}}^{600\text{ ns}} A(t)}{\sum_{0\text{ ns}}^{600\text{ ns}} A(t)} \qquad \mu_p = \frac{\sum_i A_i t_i}{\sum_i A_i}$

2) Number of peaks cut to remove asymmetric events (<2 keV_{ee}) with origin in the PMT $\quad n_0, n_1$

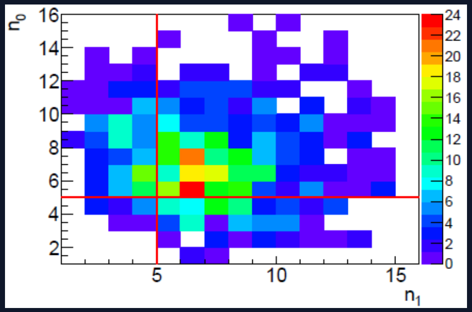
Scintillation light in NaI(Tl)



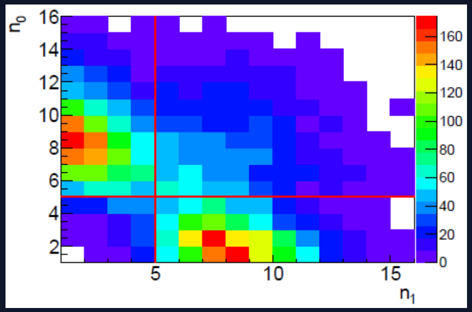
Asymmetric light sharing @PMTs



Bulk scint. sample



Background



Asymmetric events are removed

The selection efficiency decreases to 15% at 1 keV

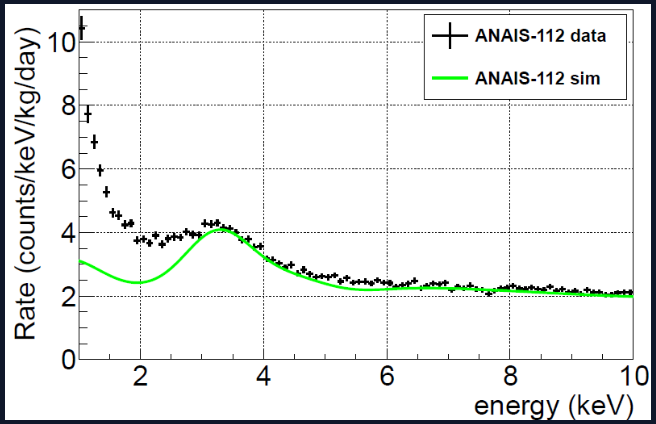
Event selection (I)

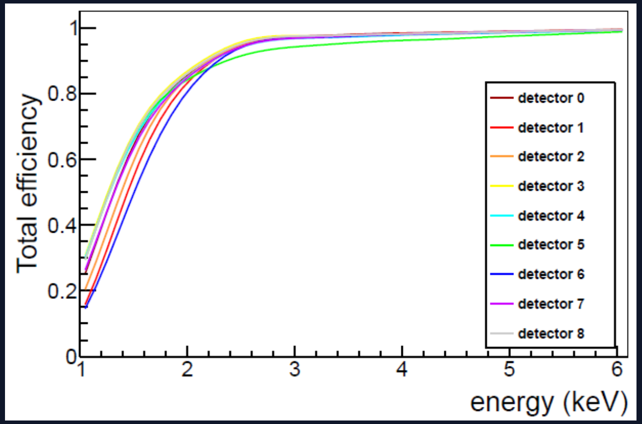
Application of event selection protocols to distinguish scintillation events from noise

1st approach: Traditional cut-based method

1) Pulse shape cut to select pulses with NaI(Tl) scintillation constant $P_1 = \frac{\sum_{100\text{ ns}}^{600\text{ ns}} A(t)}{\sum_{0\text{ ns}}^{600\text{ ns}} A(t)}$ $\mu_p = \frac{\sum_i A_i t_i}{\sum_i A_i}$

2) Number of peaks cut to remove asymmetric events (<2 keV_{ee}) with origin in the PMT n_0, n_1
3) Removal 1 s after a muon passage
4) Single-hit events (multiplicity=1)





Comparing with MC background model:
- Strong discrepancy in [1-2] keV
- 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

Event selection (II)

Application of **event selection protocols** to distinguish scintillation events from noise

2nd approach: Machine learning method

Boosted Decision Trees (BDT) implemented in Toolkit for Multivariate Data Analysis (ROOT)

Training parameters

$P_1, \mu_p, n_0 \ldots$

Training populations

Hyperparameters

BDT

Supervised learning algorithm

✓

^{109}Cd calibration

^{22}Na and ^{40}K coincidences

Neutron calibration

Synthetic events

✗

Background

Blank module

Number of trees

Max. depth

Min. node size

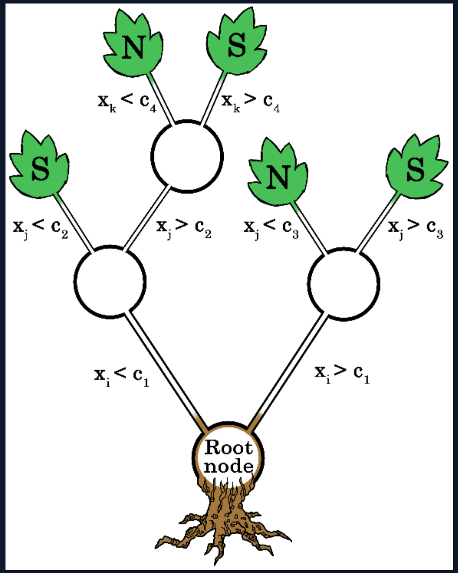
Event selection (II)

Application of **event selection protocols** to distinguish scintillation events from noise

2nd approach: Machine learning method

Boosted Decision Trees (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)



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

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)

Event selection (II)

Application of **event selection protocols** to distinguish scintillation events from noise

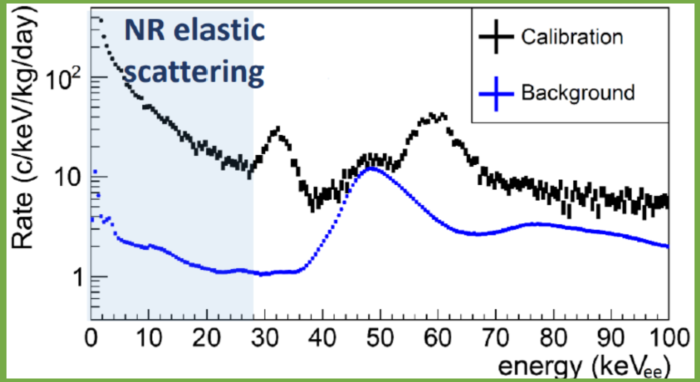
2nd approach: Machine learning method

Training populations Balanced training populations (>30 000 evts), split 70% (30%) for training (test)

Limited size of the training sample

SIGNAL EVENTS: Neutron calibrations

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

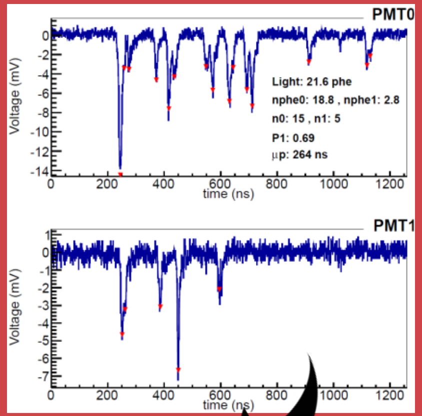


JCAP11(2022)048 and JCAP06(2023)E01

Comm. Phys. 7(2024)345 *Phys. Rev. Lett. 135(2025)051001*

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

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



Not rejected by ANAIS–112 standard cuts

Event selection (II)

Application of **event selection protocols** to distinguish scintillation events from noise

2nd approach: Machine learning method

Training parameters

15 discrimination parameters combined in a boosted decision tree

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

BDT

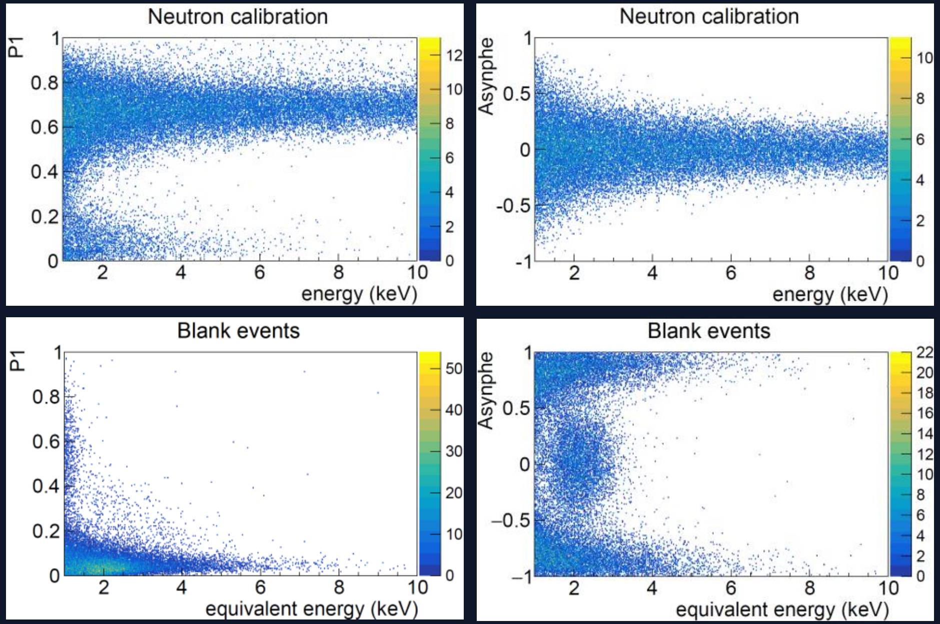
Standard analysis

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

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

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

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



Eq. energy estimated from $LC = 14.5\text{ phe/keV}$

Event selection (II)

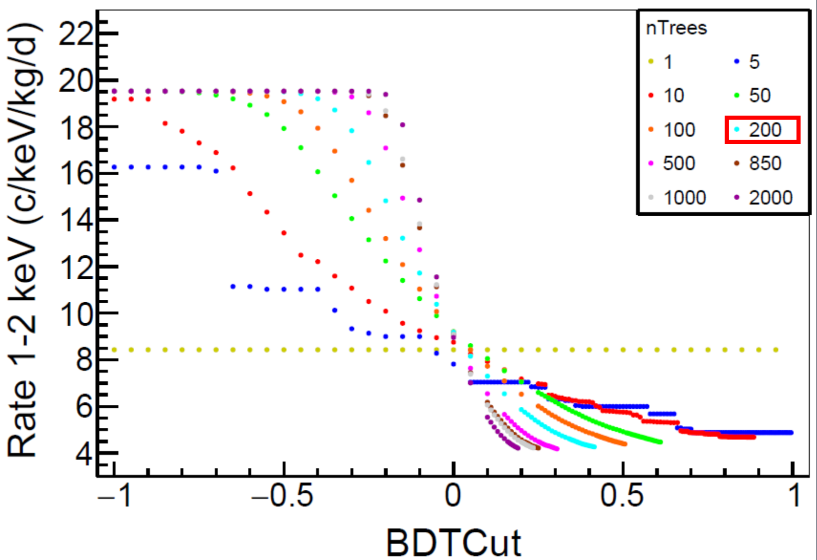
Application of **event selection protocols** to distinguish scintillation events from noise

2nd approach: Machine learning method

BDT training

- Multiple BDT models trained varying key **hyperparameters**
- Each trained model produces a **ranking of variables** contributing to the predictions

Example: *nTrees*



Ranking

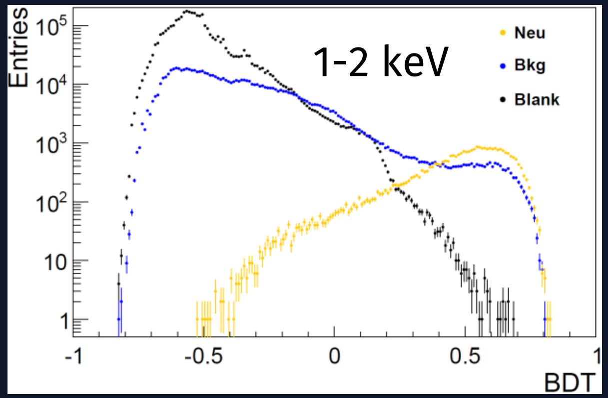
```
: Ranking result (top variable is best ranked)
: Rank : Variable  : Variable Importance
:    1 : Asynphe   : 1.155e-01
:    2 : n0        : 9.636e-02
:    3 : P2        : 8.342e-02
:    4 : n1        : 7.885e-02
:    5 : P1        : 6.713e-02
:    6 : CAP800    : 6.324e-02
:    7 : FM        : 6.250e-02
:    8 : CAP600    : 6.240e-02
:    9 : CAP400    : 5.990e-02
:   10 : CAP500    : 5.562e-02
:   11 : CAP300    : 5.510e-02
:   12 : CAP100    : 5.251e-02
:   13 : CAP700    : 5.233e-02
:   14 : CAP200    : 4.935e-02
:   15 : CAP50     : 4.576e-02
```

Event selection (II)

Application of **event selection protocols** to distinguish scintillation events from noise

2nd approach: Machine learning method

BDT score

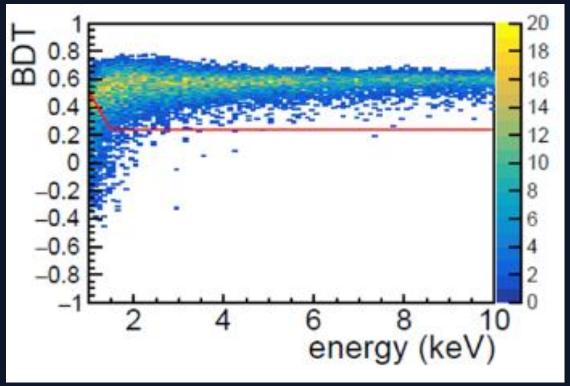


↑30% in efficiency [1-2] keV

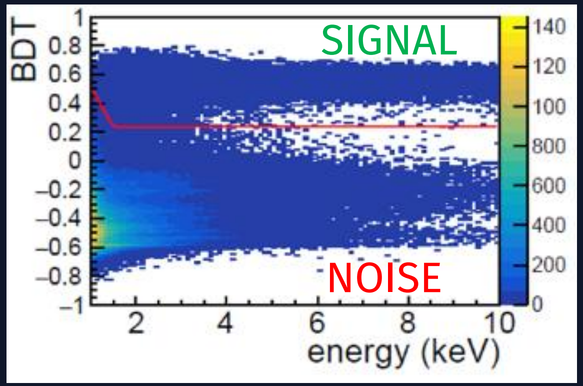
↓20% in background [1-2] keV

Efficiency better than DAMA/LIBRA above 1.5 keV, but sharply decreasing below, **limiting the analysis energy threshold to 1 keV**

Neutron calibration

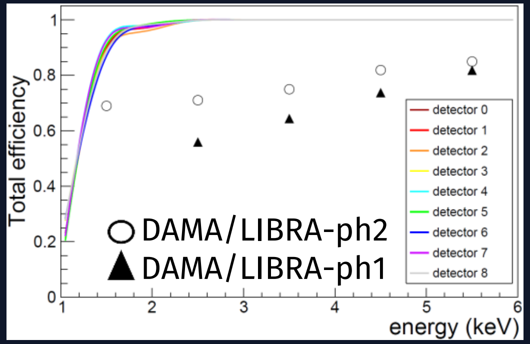


Background

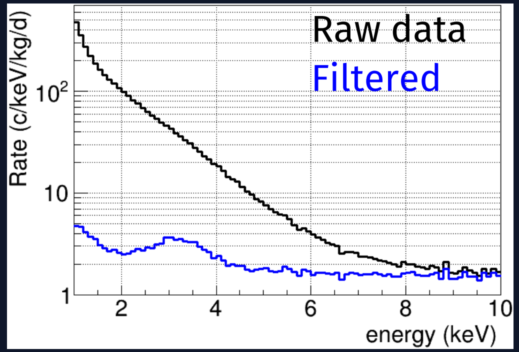


Cut on BDT parameter
(defined for every detector and energy bin)

Efficiency



Energy spectrum



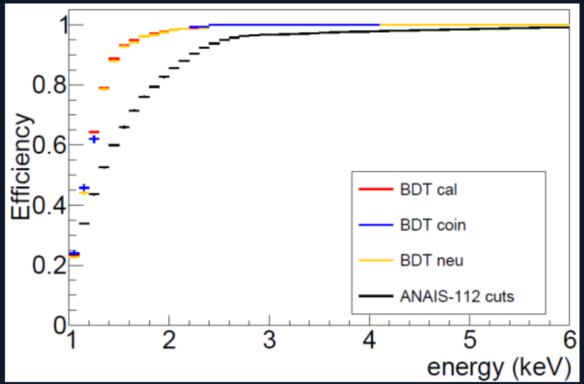
Event selection (II)

Application of **event selection protocols** to distinguish scintillation events from noise

2nd approach: Machine learning method

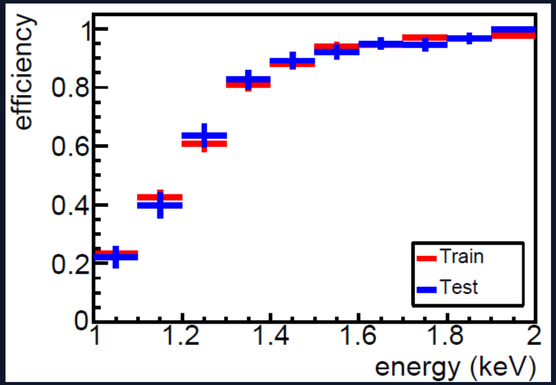
Additional mandatory validation checks

Cross-check of consistency across different bulk scintillation populations



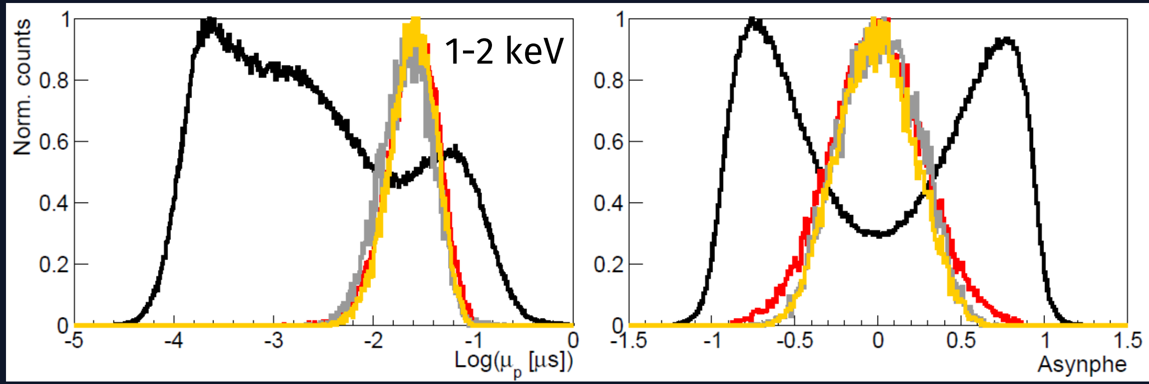
No bias observed

Check for overtraining



Confirmed no overfitting

Training parameter distributions (pre-/post-selection)



Background before | Background after
Neutron cal. before | Neutron cal. after

Event selection (II)

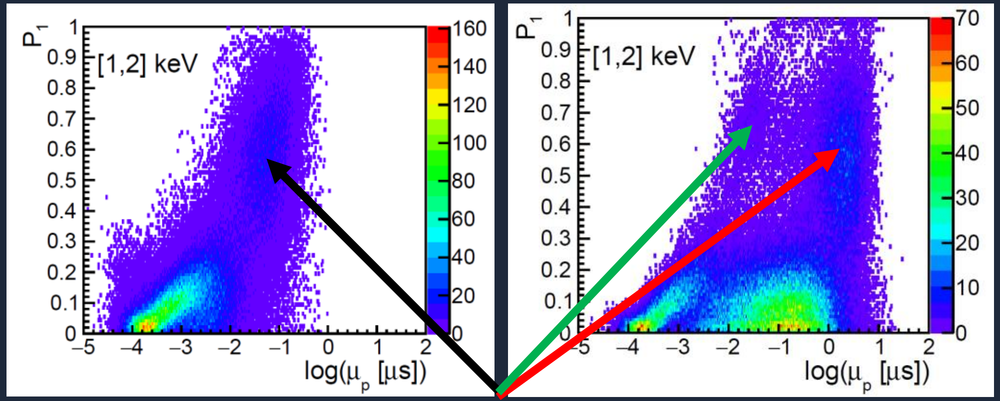
Application of **event selection protocols** to distinguish scintillation events from noise

2nd approach: Machine learning method

Ongoing improvements

New parallel DAQ system in ANAIS–112 since winter 2024, extending the digitization window from 1.25 to 8 μs and free of dead time (**ANOD**)

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



ANAIS Scint+Asy. events **ANOD**

Clear separation in the mean time!

Improving the background model

Using the full dataset [9 dets, >8 y]

Improving ML training populations

Simulating pulses through the response function of ANAIS–112 detectors

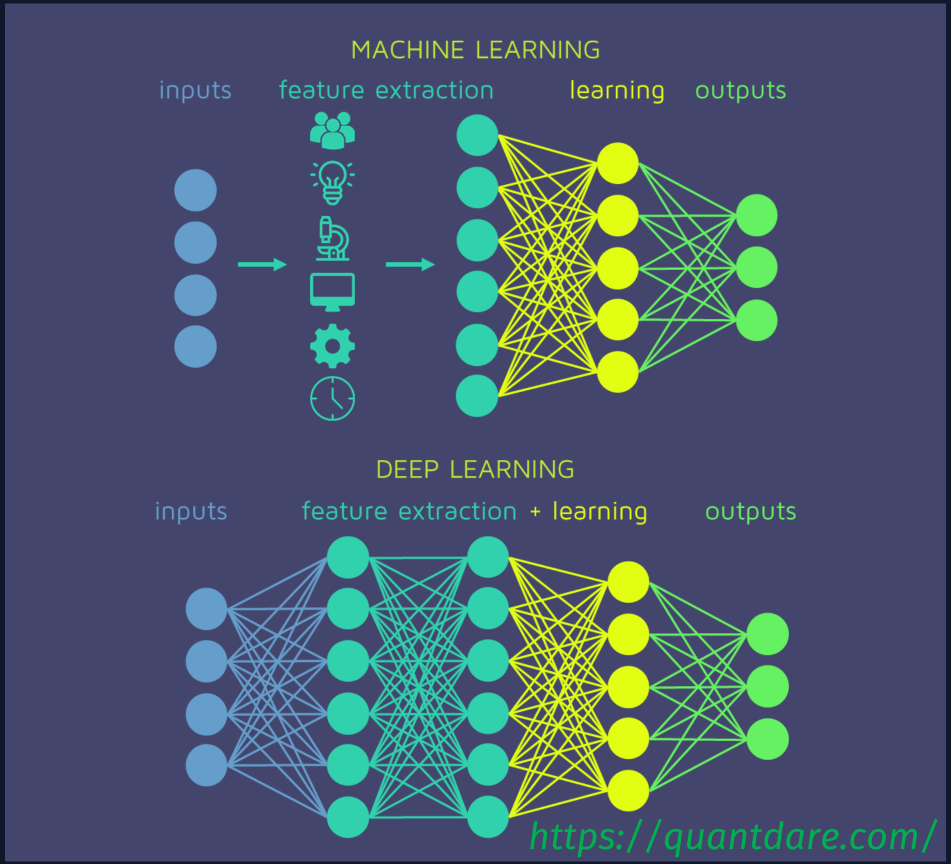
Improving ML algorithm

A more advanced technique is under development to improve event classification

Event selection (III)

Application of event selection protocols to distinguish scintillation events from noise

3^{rd} approach: Deep learning method



Traditional machine learning relies on hand-crafted features, which are tedious and costly to develop

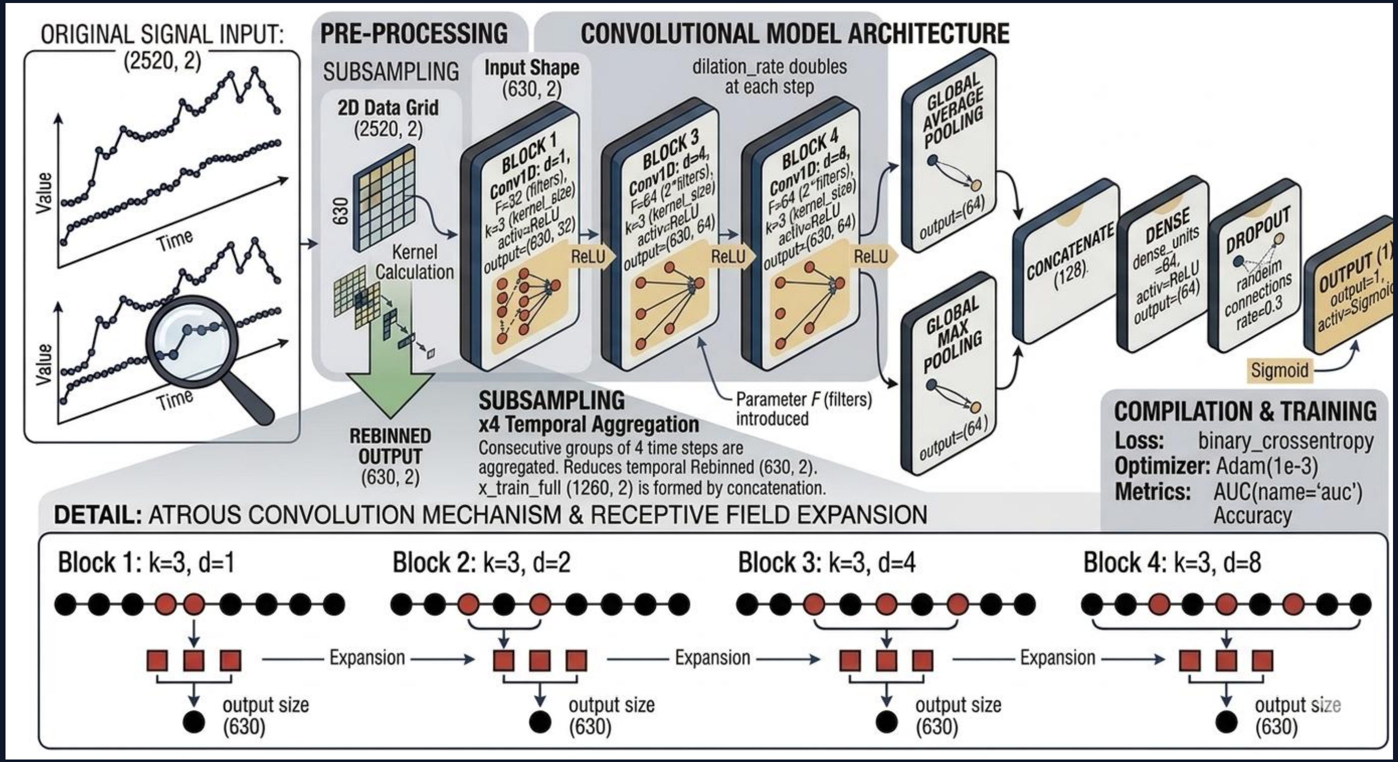
Deep learning learns hierarchical representations directly from data and scales effectively with larger datasets

Event selection (III)

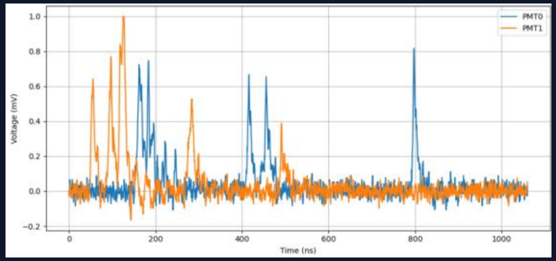
Application of event selection protocols to distinguish scintillation events from noise

3rd approach: Deep learning method

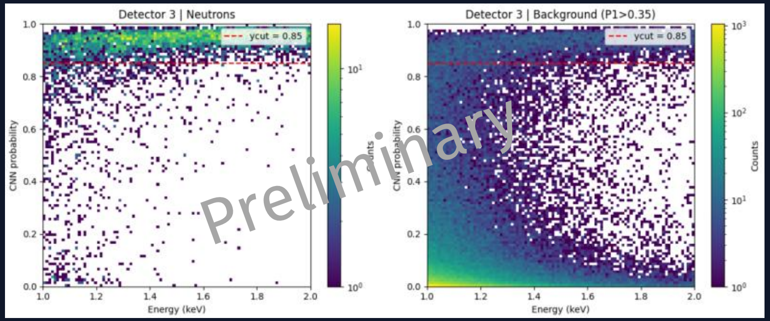
Convolutional Neural Network (CNN) Implemented in Keras (Tensorflow)



Input (630, 2)



Output example: CNN score



Neutron cal. Background

Event selection (III)

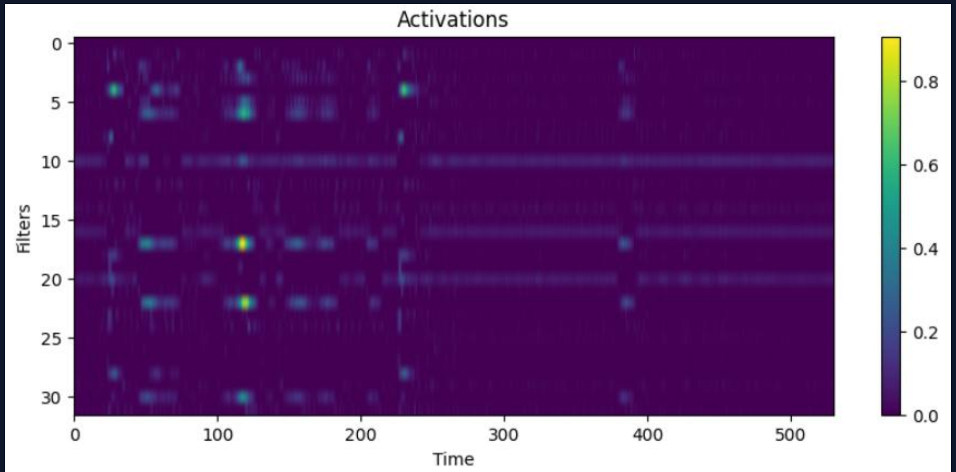
Application of event selection protocols to distinguish scintillation events from noise

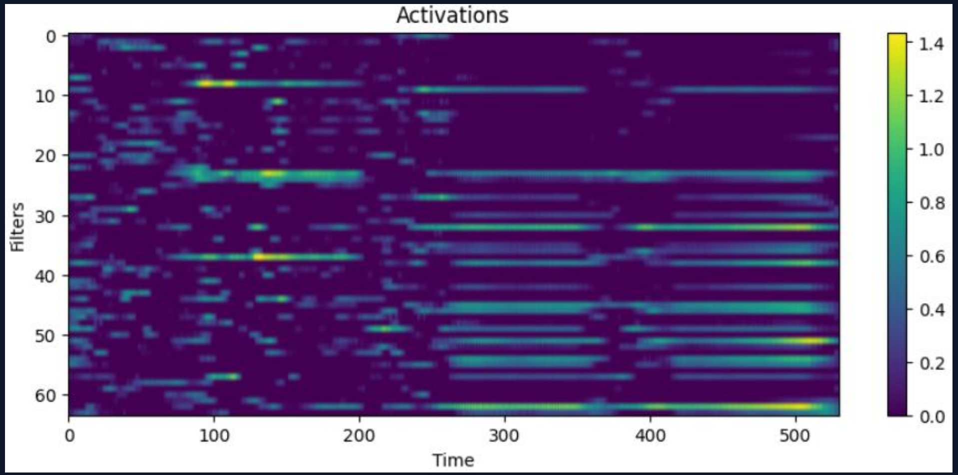
3rd approach: Deep learning method

Convolutional Neural Network (CNN)

- **Less interpretable** compared to rule-based models (e.g., BDTs)
- **Interpretation methods** include:
  - **Grad-CAM / saliency maps** → highlight input regions driving predictions
  - **SHAP / LIME** → estimate contribution of individual features

Layer activations
show what patterns each layer detects in the input



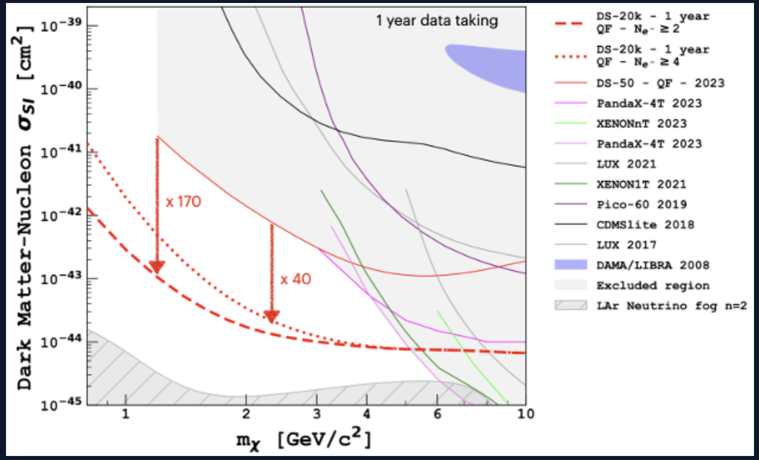


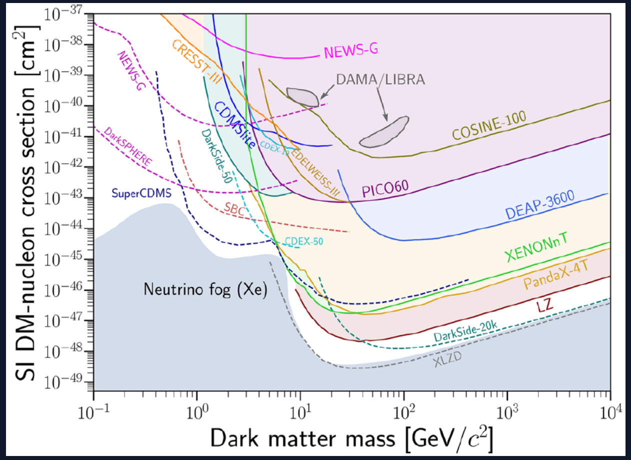
Outline

1. Rare Event Searches
2. Dark Matter
3. Case Study I: The ANAIS Experiment
4. Case Study II: The DarkSide-20k Experiment

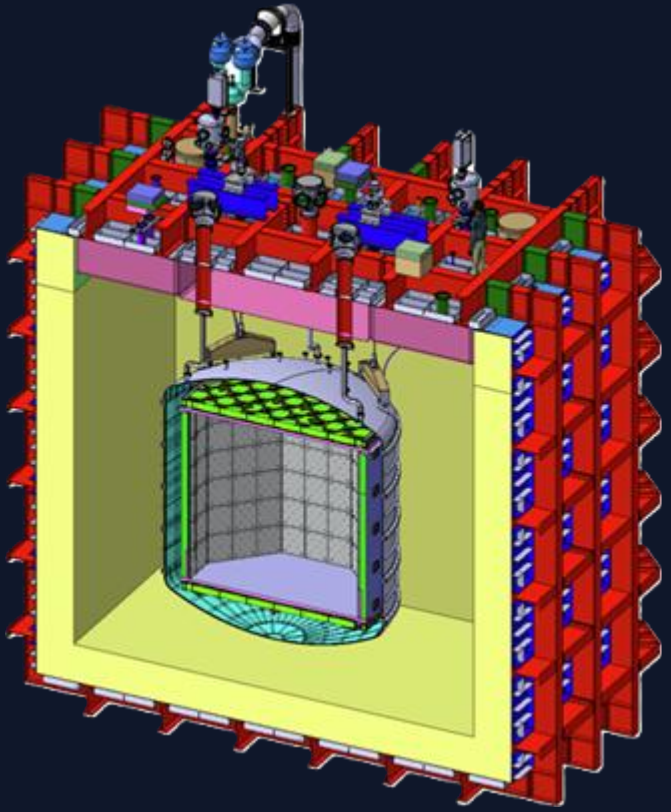
The DarkSide-20k experiment

- Next generation LAr detector for direct detection of dark matter
- Underground in LNGS (Italy)
- Dual-phase TPC containing 50 t of underground Ar (~20 t fiducial volume)
- Instrumented with 21 m² of SiPM-based light detectors
- Global Argon Dark Matter Collaboration (400 scientists from 14 countries)





Nat. Comm. Phys. 7(2024)422

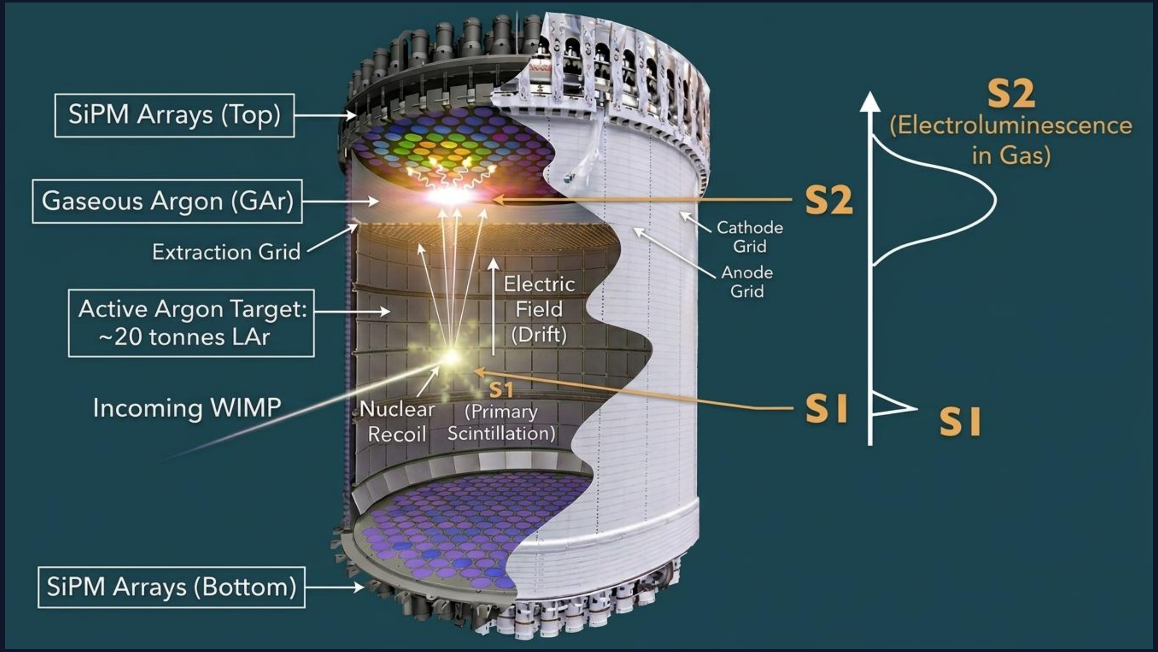


Sensitive to both low-mass (~1 GeV) and high-mass WIMPs

DarkSide-20k mini-clip

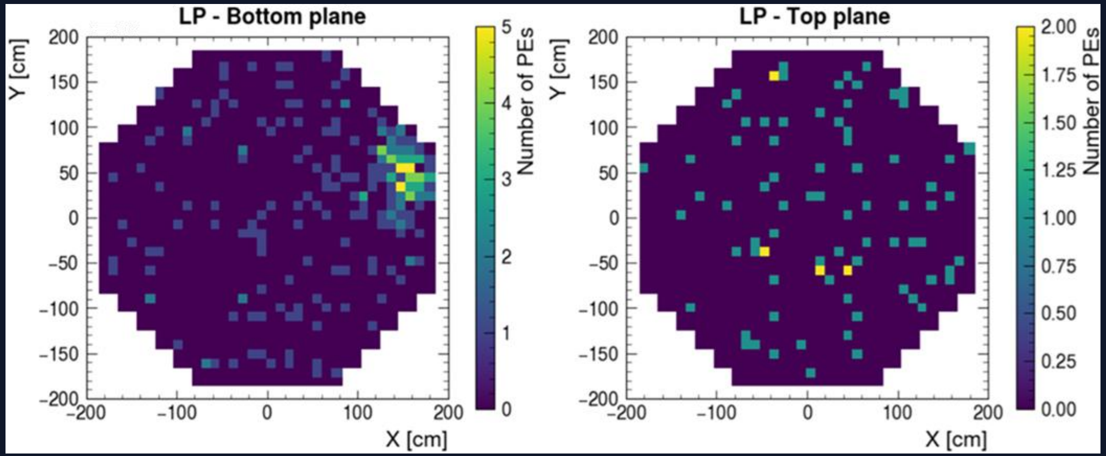
DARKSIDE

Working principle of the dual-phase TPC

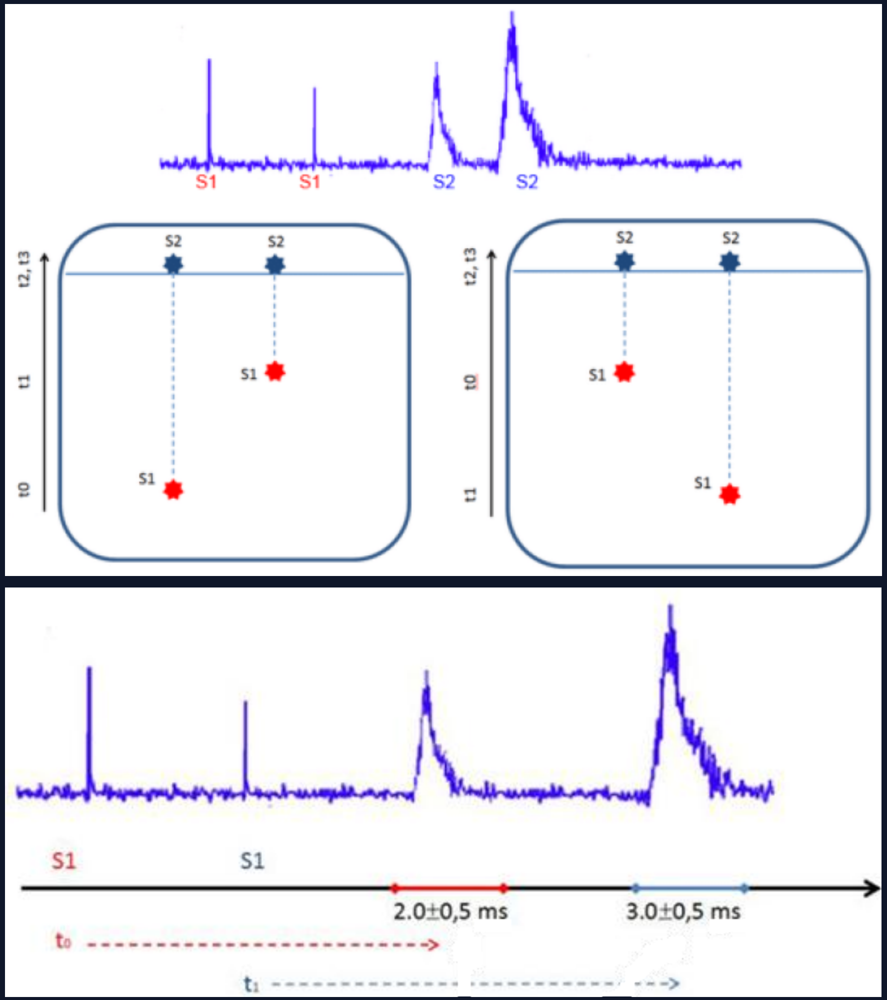


- Two readout channels: S1 (primary scintillation) +
  S2 (charge signal)
- S2 light pattern gives x-y position
- Drift time gives z position
- S1-S2 relative size gives particle information

Example of simulated event near the bottom plane



Limitation: position reconstruction and pile-up



t_0 and t_1 calculated from the reconstructed z coordinate using the drift velocity

- The expected event rate of around 120 Hz and the maximum drift time of 3.5 ms would result in approximately 42% pile-up (a second event can occur while electrons of a previous event drift) ⟶ loss of live time
- A possible solution to avoid such loss is to develop an algorithm to reconstruct the position of S1 pulses
- If the z coordinate can be reconstructed from the S1 light pattern, S1 and S2 pulses can be unambiguously associated when the events are spatially separated
- In this case, a specific time window can be defined in which the S2 corresponding to a given S1 is expected

Proposal:

Geant4 simulation +

Training a 2D CNN using the light patterns from the top and bot planes

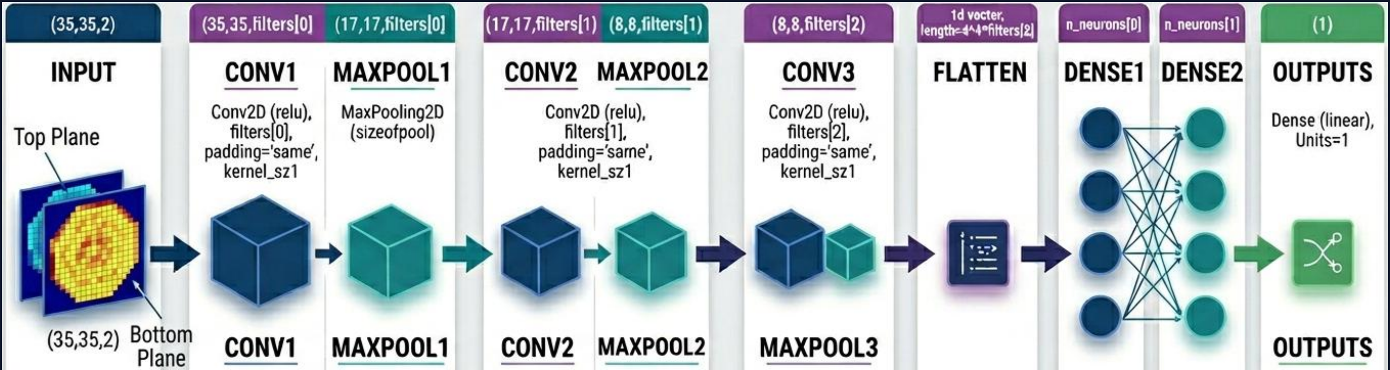
S1-based Z reconstruction

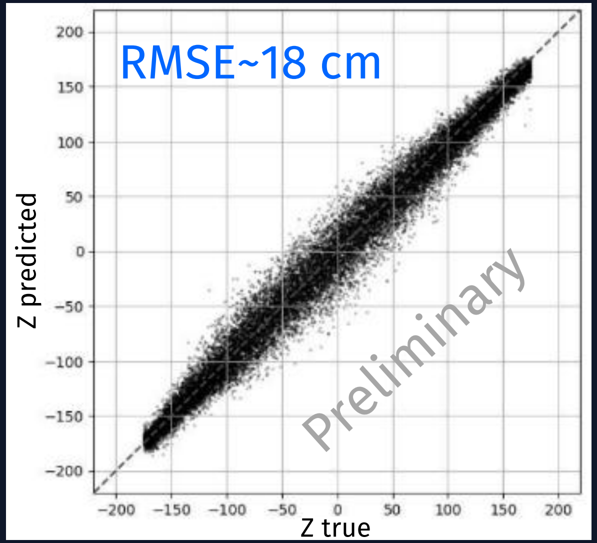
Geant4 simulation

- e^- 30 keV uniformly distributed in TPC
- Light Patterns from top and bottom planes
- TPC interaction position (x, y, z) for validation

Supervised regression with a 2D CNN to estimate the Z coordinate from light patterns

- Training sample size: 10^5 events
- Randomly split into 75% training and 25% test





*Previous estimates using KDE had an RMSE of approximately 60 cm, whereas the **CNN-based approach** reduces it to **18 cm***

Summary and conclusions

Rare Event Searches: Dark Matter

- Signal is extremely rare and hidden under dominant background
- Mitigation is multi-layered: detector design, shielding, radiopure materials, event-level discrimination

Case Study I: The ANAIS Experiment

- Traditional cut-based methods are simple but limited
- Machine learning (BDT) improves signal-background separation
- Deep learning method (CNN-1D) exploits waveform correlations to increase sensitivity

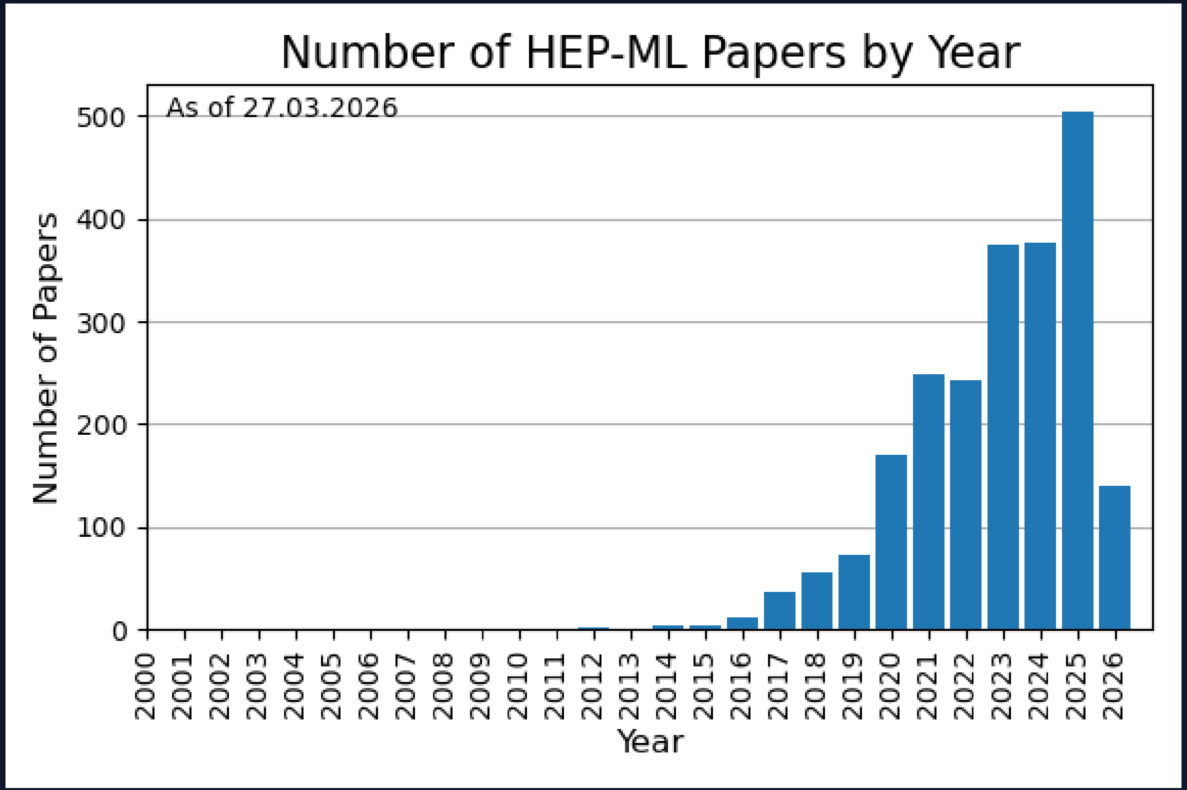
Case Study II: The DarkSide-20k Experiment

- CNN-2D applied to S1 signals allows Z-position reconstruction
- ML enables precise spatial reconstruction and background rejection

Final Remarks

- Rare event searches require combining hardware, traditional analysis, and ML for maximal sensitivity
- Machine learning is not a replacement, but a powerful extension of classical methods
- Understanding the full workflow from detector design to analysis is crucial for future discoveries
- These methods are broadly applicable to other rare event physics beyond dark matter

Yearly HEP-ML Publications



https://github.com/iml-wg/HEPML-LivingReview?tab=readme-ov-file

Thank you for your attention!

Centro de Astropartículas y Física de Altas Energías
Universidad Zaragoza

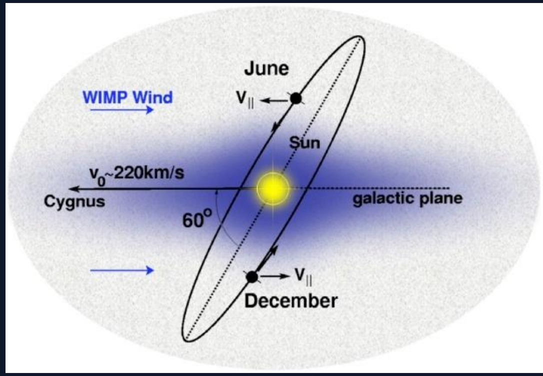
Back-up

Dark matter directionality

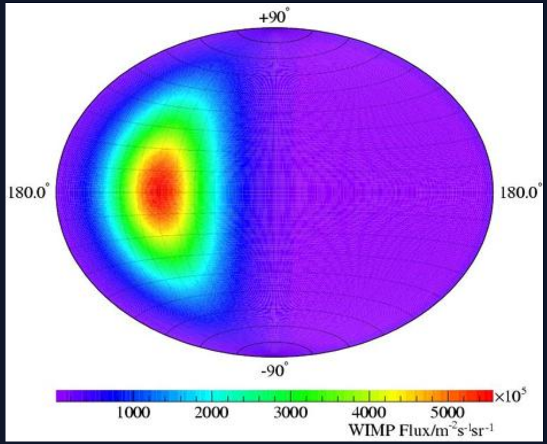
The average direction of the "WIMP wind" through the solar system comes from the constellation of Cygnus, as the Sun is moving around the Galactic center

The **measured track direction** of NR could be used to distinguish a DM signal from background events (expected to be uniformly distributed) and to prove the galactic origin of a possible signal

Challenge: to reconstruct very short tracks (~1 mm in gas, ~0.1 mm in solids) for keV scale NR

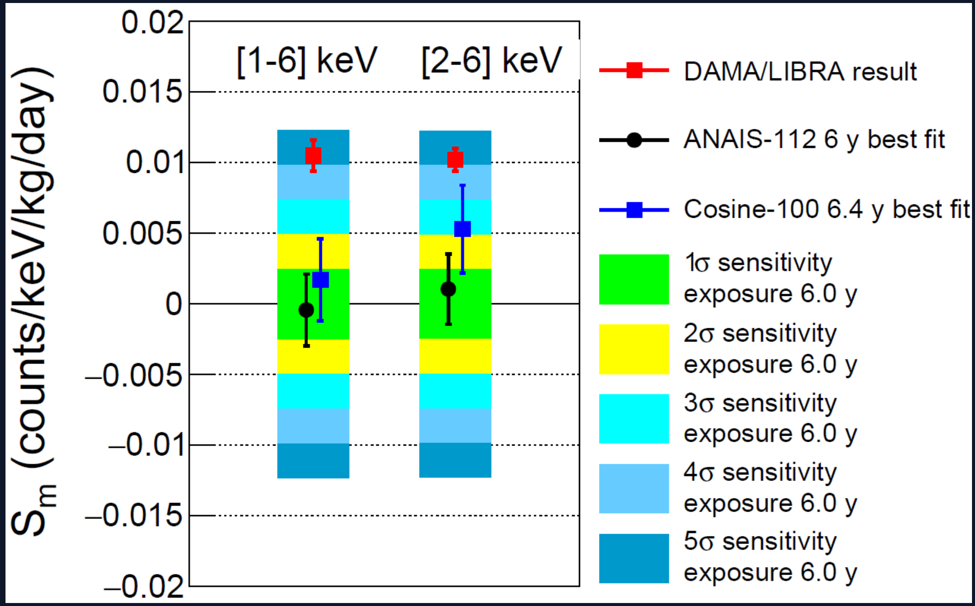


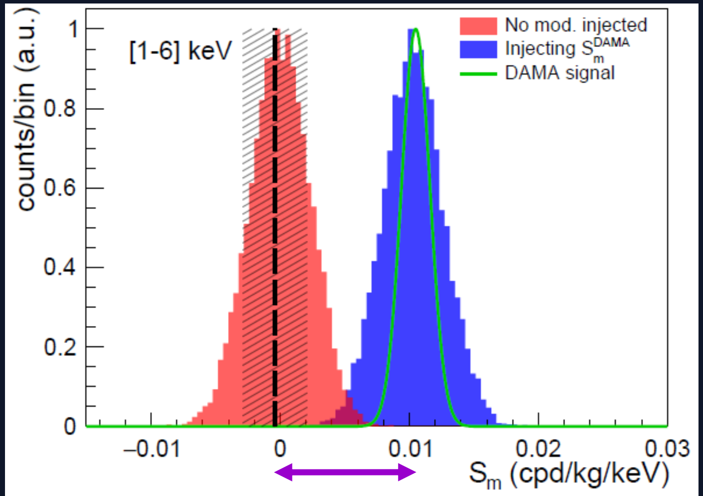
www.hep.shef.ca.uk



Phys. Rev. D 71(2005)103507

Annual modulation results with 6 years





ANAIS−112 *Phys. Rev. Lett. 135(2025)051001*

Best fit modulation amplitudes **compatible with zero** at 1σ

Incompatible with DAMA/LIBRA at 4.0 (3.5) σ for [1-6] ([2-6]) keV

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

COSINE−100 full dataset *Sci. Adv. 11(2025)eadv6503*

Incompatible with DAMA/LIBRA at 2.8 (1.5) σ for [1-6] ([2-6]) keV

Sensitivity: 3.6 (3.3) σ for [1-6] ([2-6]) keV

Sensitivity to DAMA/LIBRA result as: $\mathcal{S} = \frac{S_m^{DAMA}}{\sigma(\hat{S}_m)}$

Incompatibility to DAMA/LIBRA result as: $\frac{|\hat{S}_m - S_m^{DAMA}|}{\sqrt{\sigma^2(\hat{S}_m) + \sigma^2_{DAMA}}}$

DS20k background mitigation

Background source	Mitigation strategy
^{39}Ar β decay	Underground Ar + PSD
γ from rock and γ, e^- from materials	PSD, material screening and selection
Radiogenic neutron, mostly (α,n) reaction in detector material	Material screening and selection Definition of TPC fiducial volume Rejection from neutron veto
Surface contamination from Rn progeny	Reduction of surfaces Surface cleaning Radon abatement system
Muon induced background	Cosmogenic veto
Neutrino coherent scattering	*Irreducible*