Search for $\overline{\nu_e}$ disappearance w/ the SoLi ∂ detector

Hervé Chanal, Stéphane Monteil, Mike Yeresko

BCD ISHEP School 2024, Cargèse (Corsica, France) March 2024

* 曰 > (4 四 > (4 回 > (4 回 > 4 回 >)

In the search for incompleteness



The table of elementary particles of the Standard Model.

CMY

BCD 2024

< E > < E > E|E
27/03/24

< 17 ▶

1/50

In the search for incompleteness



▲ロト ▲母ト ▲ヨト ▲ヨト 三国市 のへで

46

Reactor antineutrino anomaly



The ratios of observed to predicted $\overline{\nu}$ event rates for the different reactor experiments (RAA, left) and the measured to predicted spectral ratios for the commercial reactors (5 MeV bump, right).

- Popped up in 2011 after reevaluation of the reactor flux predictions;
- Might be accounted for with an additional light sterile neutrino state.
- Accompanied by the $\overline{\nu}$ energy spectra distortion (aka "5 MeV bump");
- Very short baseline reactor experiments as a tool to tackle both questions;

(A) = (A)

The SoLi∂ experiment





- Set at the BR2 research reactor (Mol,Belgium)
- Compact core (⊘ 50 cm)
 - $\implies \overline{\nu}$ creation position precisely determined;
- ²³⁵U enriched reactor core (95%)
 Easy energy spectrum simulation;
- Very short baseline experiment [6.3-8.9]m
 Accesses the region, where the most pronounced oscillatory behaviour is expected;
- No nearby experiments, shielded beam ports

 Low reactor-induced bckg environment.
- \approx 140 days of operation/year

Top: An illustration of the BR2 twisted design core.
Bottom: Schematic view of the SoLi∂ detector positioning on the site.

27/03/24

4/50

The SoLi∂ detector





- A sandwich composite of two scintillators;
- 5 cm sided cube are made of polyvinyl-toluene (PVT) and lined with 2 layers of ⁶LiF:ZnS \Rightarrow Highly granular detector.
- Individual cubes are wrapped with Tyveck \Rightarrow Light kept within the cube, where issued.
- Light is taken to the boundaries w/ wavelengthshifting optical fibres and read-out w/ SiPMs;
- Cubes are arranged in planes of 16×16 units;
- Layers are further optically decoupled with two square Tyvek cover sheets:
- 10 planes make a module, 5 modules in total;

← The schematic view of the SoLi∂ basic detection unit cell design (top) and the full-scale geometry (bottom).

27/03/24

5/50

The detection principle



Inverse beta decay (IBD) to detect v:

 $\bar{\nu}_e + p \longrightarrow n + e^+(\gamma\gamma)$

- Neutron scintillation signal [NS]:
 - Generated by the ZnS
 - Energy is issued from n capture on the ⁶Li

$$n + {}^{6}Li \longrightarrow {}^{3}H + \alpha$$

- Electromagnetic scintillation signal [ES]:
 - Generated by the PVT
 - Proton-rich v

 target
 - Measures e⁺ ionisation energy
 - Measures annihilation γ energy
 - High granularity allows to distinguish ionisation and annihilation contributions!

NS and ES correlated in time: $\Delta T = T_{NS} - T_{ES}$

BCD 2024

6/50

- The reconstruction of the ES;
- The relative calibration of the ES;
- The absolute calibration of the ES;
- The ES definition and Topological analysis;
- Verification of the novel methods with an open data set.

Reconstruction basics

- The digitised SiPM readout from the fibers* is the raw detector data;
 - * the waveforms, which passed the triggers and thresholds;
- The waveforms are grouped and clustered;
- Each cluster receives one of the following tags:
 - 1. Muon (track of high amplitude cubes);
 - 2. NS (a single cube, defined by the dedicated trigger);
 - 3. ES (subset of the cubes in the plane, provided by e^+ and γ);
 - \implies the most involved case, requires specific reconstruction strategy;
- ▶ NS and ES clusters within appropriate ΔT are forming a coincidence.

Reconstruction basics

An example of three different types of muon tracks reconstructed in the SoLi ∂ detector.



We will come back to them in few minutes!

0	κ.	•	0
J	IV	1	Y,

9/50

Reconstruction basics

- The digitised SiPM readout from the fibers* is the raw detector data;
 - * the waveforms, which passed the triggers and thresholds;
- The waveforms are grouped and clustered;
- Each cluster receives one of the following tags:
 - 1. Muon (track of high amplitude cubes)
 - 2. NS (a single cube, defined by the dedicated trigger)
 - 3. ES (subset of the cubes in the plane, provided by e^+ and γ)
 - \implies the most involved case, requires particular reconstruction strategy
- ▶ NS and ES clusters within appropriate ΔT are forming a coincidence.

- The fibres project the deposited energies to the boundaries of the detector
- The digitised SiPM readout from the fibers are the raw detector data
 - \implies Reverse engineering is required to restore the list of involved cubes

- Each layer is a separate problem
- Parametrisation:
 - Unknowns: PVT deposits (E_i)
 - *p_i* are the SiPM measurements
- Challenges:
 - Cube projects through adjacent fibers
 - Fibers can overflow during the run



- The fibres project the deposited energies to the boundaries of the detector
- The digitised SiPM readout from the fibers are the raw detector data
 - \implies Reverse engineering is required to restore the list of involved cubes

- Each layer is a separate problem
- Parametrisation:
 - Unknowns: PVT deposits (E_i)
 - p_i are the SiPM measurements
- Challenges:
 - Cube projects through adjacent fibers
 - Fibers can overflow during the run



- The fibres project the deposited energies to the boundaries of the detector
- The digitised SiPM readout from the fibers are the raw detector data
 - Reverse engineering is required to restore the list of involved cubes \implies
- Each layer is a separate problem
- Parametrisation:
 - Unknowns: PVT deposits (E_i)
 - p_i are the SiPM measurements
- Challenges:
 - Cube projects through adjacent fibers
 - Fibers can overflow during the run



The reconstruction problem can be put down as follows:

$$p = AE$$
,

where A is the so-called system matrix (SM) that embodies the overall response of the detector. This equation has been widely studied in medical imaging and particle physics.

BCD 2024



Similar energy resolution = 13%



27/03/24

12/50

Method	FISTA ¹	FISTA+ML-EM ²	sOMP ³ +ML-EM			
① (%)	15.8	11.4	6.9			
ϵ (%)	75.3	76.3	77.7			
			A = A + A = A A	하 소문에	문기관	200

ES calibration overview

- Calibration w/ horizontal muons:
 - Relative calibration
 - Higher precision
 - Access to the Light Leakages
 - Time evolution of the response
 - Absolute scale calibration(?)
 - dE/dx values

- Calibration w/ radioactive sources:
 - Relative calibration
 - Time evolution of the response
 - Calibration campaigns required
 - Absolute scale calibration:
 - ²²Na: developed, low energy range
 - AmBe: TBD, desired energy range

SoLi∂

Crosschecks:

- Identification of the wellknown sources of the bckg
 - Cosmogenic (¹²B, etc.)

Crosschecks:

- Validation with natural radiation source (²¹⁴Bi)
- Data/MC comparison

Relative calibration with horizontal muons

- Horizontal muons [crossing 1 cube/plane] compose powerful calibration tool;
- Allow to perform the per fibre relative calibration of the whole detector;
- Give an access to the measurement of the light leakages to the neighbouring cubes;
- Provide the ballpark for the absolute energy scale calibration
 - \implies To be compared and checked with the calibration sources;
- Control the time evolution of the detector response.



Relative calibration with horizontal muons



- ▶ 1 crossed cube per plane \Rightarrow clear posed problem;
- >1 impacted cube per plane ⇒ access to the LL!
- Total energy in the plane as a proxy to μ deposit;
- For the hit and adjacent cubes fibers (12 in grand total) define the fractions:

$$f = \frac{E_{\rm fibre}}{E_{\rm plane}}$$

- Describe the light sharing characteristics;
- Constitutes the initial SM elements values;

Top: an illustration of the light sharing for the hit cube in the muon track (1 main + 4 LL = 12 fibres in total). Bottom: An example of the LL distribution for one of the 8 neighboring fibres.

Image: A matrix

A = A = A = A = A = A = A

Relative calibration with horizontal muons



- E_{plane} as a proxy to the muon deposit $\Rightarrow dE$;
- Muon track info as a proxy to the path length \Rightarrow dx;
- $\frac{dE}{dx}$ distribution per hit cube is built [1];
- An average detector $\frac{dE}{dx}$ distribution is built [2];
- The SM elements are scaled by $\frac{[2]}{[11]} \implies$ Response homogenisation. Relative calibration is done!

An example of homogenisation procedure impact.

0.30

0.25

0.20

0.15

0.10

0.05

BCD 2024









12

27/03/24

0.30

0.25

0.20

10



ELE SOO

Absolute energy scale calibration

- Split relative and absolute calibrations is a feature of the muon calibration ⇒ various inputs can be used as an absolute energy scale*
 - * also known as Light Yield (LY);
- Available options are the following:
 - From muons: $\frac{dE}{dx}$ value from Geant4 MC simulation or from PDG
 - \implies Disadvantages: high region of the energy spectrum
 - From ²²Na: 1.06 MeV CE. Method developed and crosschecked
 - \implies Disadvantages: low region of the energy spectrum
 - From AmBe: 4.2 MeV CE. At the heart of the desired energy spectrum. Similar to ²²Na techniques can be employed.
- The differences b/w possible choices are currently under investigation

1 = 1 = 1 = 1 = 1

Absolute calibration w/ horizontal muons

- Merge the $\frac{dE}{dx}$ distribution in all cubes from the:
 - Geant4 simulation (all the cubes have the same characteristics)
 - \implies provides an expectation in MeV/cm;
 - ROff data (a single value is enough, since the relative calibration is done)
 - \implies provides the data expectation in ADC/cm;
 - An alignment of the two provides the currently used absolute energy scale;
- Limitations:
 - 1. Precise energy spectrum of μ is required. Otherwise, precision is limited by 10%;
 - 2. Dedicated Geant4 study was performed to support this statement;



18/50

Crosschecks and validations [¹²B]

Cosmogenic bckg searches as a way to check the reconstruction and calibration:

 μ^- + ¹²C $\rightarrow \bar{\nu_{\mu}}$ + ¹²B $\rightarrow \beta^-$

- Executed by S. Gallego from LPC Caen;
- Statistics for 100 ROff days is employed;
- ΔT(Muon-ES) is fit to determine the yield;
- SPlot technique is used to statistically subtract the background (Δ*T* as a discriminative variable) ⇒ access to *E* spectrum;
- It matches the Geant4 prediction!
- Was not the case w/ former machinery.



BCD 2024

19/50

Absolute calibration w/ sources. AmBe

AmBe as an additional ES calibration source (in 60% of the cases)!

²⁴¹ $Am \rightarrow \alpha + {}^{237}Np$; $\alpha + {}^{9}Be \rightarrow n + {}^{12}C^* \rightarrow {}^{12}C + \gamma$ (4.438 MeV)

The *n* has enough energy to create ${}^{12}C^*$ from the interaction with detector materials!

- More energetic sources are providing a promising alternative to CE;
- 3.4 MeV e⁺e⁻ pair as mono-energetic calibration tool!
- The typical energy is right in the middle of the region of interest;
- e^+ is correlated to the *n* signal and provides 2 annihilation $\gamma \implies$ IBD-like signature!



્≡ ્ર્ગેલ્લ 20/50

Absolute calibration w/ sources. AmBe

- Similar to ²²Na strategy, unless the presence of the background (2 templates required);
- Signal definition:
 - 1. The cube w/ the largest energy as the proxy for the e^+e^- pair creation position;
 - 2. $3 \times 3 \times 3$ envelope around it required to be empty;
 - 3. Two ES clusters to be present outside of the envelope (different hemispheres);
 - 4. Low (\leq 5) cube multiplicity;

▶ Resulting distribution for signal and background ⇒ bumpy behaviour in the signal region.



The stacked distributions of the most energetic cubes energy in Topology 20 for the Geant4 (left) and ROSim (right) levels.

CMY

Absolute calibration w/ sources. AmBe

- A systematic scrutiny is ongoing and more sources of errors are going to be checked;
- The best parameter pair preferred by the fit is (0.962, 20.9%).



An example of the unbinned event-by-event ML fit to the calibration data. An obtained parameter pair is (0.962, 20.9%).

IBD analysis

Ratio of signal and background events as a function of the ΔT in the raw data sets.



Comparison of the raw signal to background rates reveals overwhelming level of latter;

- Extensive use of the electromagnetic part of the signal physics properties is in order;
- E.g. Selecting geometrical properties of the signal events according to their EM features;
- Complex* patterns implies more discriminative power; *reconstruction of both ann γ.

IBD analysis

Ratio of signal and background events as a function of the ΔT in the raw data sets.



Comparison of the raw signal to background rates reveals overwhelming level of latter;

- Extensive use of the electromagnetic part of the signal physics properties is in order;
- E.g. Selecting geometrical properties of the signal events according to their EM features;
- Complex* patterns implies more discriminative power; *reconstruction of both ann γ.

Signal definition



- The signal definition of the current analysis is based on the topological properties of the ES
- The detector output is converted into a list of cubes
- The possible cubes of interest
 - 1. The cube where e^+ annihilated [AC]
 - 2. The cube where n was captured [NC]
 - [If applicable] The clusters corresponding to the e⁺ annihilation gammas [EM1, EM2]
- The positron and γ deposits can be distinguished to remove the backgrounds (granularity of the detector is used maximally)

Envelope condition	Ο γ	1γ	$1\gamma^*$	2γ	$2\gamma^{\star}$	2 γ**
Only AC AC & 1 cube AC & 1+ cubes	Topo00 Topo01 Topo02	Topo10 Topo11 Topo12	Topo13 Topo14	Topo20 Topo21 Topo22	Торо23 Торо24	Торо25

The definition of the topologies, including the case where the γ deposit is found inside the envelope (tagged with *)

Identification of the background

With the fit of the ΔT distribution we can

estimate the contamination of the backaround sample by the different sources

The specific selection for each source can

Enriched samples are devised

help in the further background rejection: The $\alpha - n$ shape discrimination ΔT [BiPo] and ΔR values [Cosm]

The FPNT sample [Acc]

Natural radiation [BiPo decay chain]

- delayed signal: α decay
- prompt signal: β decay
 ΔT ~ 236μs [²¹⁴Po decay time]



- Cosmogenic background
 - delayed signal: fast or spallation neutrons
 - prompt signal: proton recoils or β decays
 - $\Delta T \sim 60 \mu s$ [the same as for the signal]
- Accidental background
 - No time or space correlation



The ΔT fit of ROff Topo10 sample performed with [1]

イロト イポト イヨト イヨト

 $\Delta T = N_{\text{acc}} + N_{\text{fast}} \cdot e^{-\frac{\Delta T}{8.5\,\mu\text{s}}} + N_{\text{atm}} \cdot e^{-\frac{\Delta T}{62\,\mu\text{s}}} + N_{\text{BiPo}} \cdot e^{-\frac{\Delta T}{235.8\,\mu\text{s}}}$ [1]

Discriminative variables

- Additional features to further suppress background;
- "Back-to-backness" of annihila-► tion γ as the main signature;
- Some variables are discrimina-tive per se;
- Another bring additional info through correlations;
- Hence we need a tool to use it
 - \implies multivariate analysis!





Discriminative variables

- Additional features to further suppress background;
- "Back-to-backness" of annihila-► tion γ as the main signature;
- Some variables are discriminative per se;
- Another bring additional info through correlations;
- Hence we need a tool to use it
 - \implies multivariate analysis!



Enriched samples



 ΔT fits for the BiPo (left) and Atm (right) enriched samples.

- Selected w/ PSD variable and T;
- Equivalent cuts for all topologies;
- Enormous statistics + 90% purity;
- Validated with MC;

- Selected w/ PSD variable and E_{AC};
- Cuts tuned per individual topology;
- Very poor statistics + 90% purity;
- No MC available;

Both are employed in the multivariate analysis training.

CMY

BCD 2024

27/03/24

27/50

The multivariate tools

- Multivariate analyses can tackle the high level of background
- There are three tools. Trained against:
 - 1. BiPo [based on the EM features]
 - 2. Atm [based on the same EM features]
 - 3. Atm [based on geometrical features]
- 10-fold cross-validation is applied (to use the full ROff statistics). Overtraining checked.
- 5D optimisation:
 - 3 BDT scores
 - ΔT and BiPonator cut
 - Fixed cut E_{AC} > 1.5 MeV
 - Aimed at the signal significance maximization $(S/\sqrt{S+B})$
- Four cuts are applied at this stage w/o geometrical BDT score and the ΔT. The ΔT and the ΔR variables are used for the further signal and background discrimination
- Partly optimized 2γ B/S = 2 (\sim 30 $\bar{\nu}$ /d)



Simultaneous fit technique

- The tool is required to distinguish the remaining background. Solely ΔT can not do it;
- The signal events are sharing the same moderation time with the atmospheric background; Additional discrimination is required. The Δ*R* variable can fit this demand
- > The p.d.f. for the simultaneous unbinned ML fit are derived from dedicated samples:
 - MC IBD [Signal]
 - Enriched samples: optimized [Acc, BiPo], preselected [Atm (lack of statistics)]
- The yields for each source directly measured from the simultaneous fit



The simultaneous $\Delta T \cdot \Delta R$ fit for the optimised 2γ ROff sample

	4	ロマ・4回マ・4回マ 山田	900
CMY	BCD 2024	27/03/24	29/50

Predictions. Control plots



◆□▶◆圖▶◆言▶◆言▶ 割片 のへの

30/50

27/03/24

BCD 2024

Predictions. Performance

- Optimized 2γ S/B = 1.5 (19.1 ± 1.3 ± 1.5 ν/d), where 1 number is 5D optimised MC predict, 2 number - statistical error and 3 - systematic;
- The blind topological analysis is concluded with the staged scrutiny of the 21 days open data set

15

S/B

SoLid

10



Quick reminder

2) The unblinding steps

- For each ROn cycle, we have the two ROff periods [before b and after a]. They are used to provide predictions of the neutrino yield.
- Step 1: compare accidentals (ROn ROffb) and (ROn ROffa). Pass if the two results are agreeing within 2.5 standard deviations.
- Step 2: compare BiPo enriched yields (ROn ROffb) and (ROn ROffa). Pass if the two results are agreeing within 2.5 standard deviations.
- Step 3: Fit results on ROffa-b. Pass if 2.5 stat. agreement.
- Step 4: BLIND fit. Check the consistency with the background predictions while keeping the plots blinded. Pass if agreement within 2.5 standard deviations.
- Step 5: BLIND fit: GoF: binned χ^2 test . Pass if χ^2 / d.o.f. < 1.6.

Quick reminder

- Over time and acquired understanding the steps have slightly changed;
- ► The ROff data are addressed in the more general way:
 - BiPo and FPNT components have to agree within 2.5σ over Phasel;
 - Atm part can only be addressed in terms of pressure anti-correlation;
- ROn background composition requirements remained almost the same:
 - BiPo component has to agree within 2.5σ wrt ROff over Phasel;
 - FPNT component is compared for [ROn ROffb] and [ROn ROffa], due to increase for the reactor on cycles. It has to agree within 2.5σ;
- Once the background composition consistency is verified the control plots are considered. The comparison has to be kept blind with solely goodness of fit being verified. Test is passed if \(\chi^2/d.o.f < 1.64\);</p>

ROff composition scrutiny. BiPo rate.



- The BiPo yields in the BiPo enriched samples for Topology 30 (gathered 2γ, only statistical uncertainty) representation. Each point is obtained from the per-cycle ΔT fit;
- Each yield is normalised per day according to the data-taking time;

ROff composition scrutiny. BiPo rate.



- ▶ 2 operating modes \implies chronologically consistent with the chiller issue;
- Are there consequences for the discriminative variables inputs?
- ► 2 departing periods (12, 15) \implies differences for the ΔT and ΔR shapes?

ROff composition scrutiny. FPNT rate.



- Normalised FPNT rates from the preselected Topology 30 △T fits;
- Same 2 operation modes are present, no statistically significant outliers;
- Is there an additional information to extract from accidentals?

CMY

27/03/24

ROn background composition. BiPo rates.



- Normalised BiPo rate from the BiPo enriched Topology 30 ΔT fits;
- **•** ROff P12, 15 excluded; All periods are in agreement \implies BiPo is stable!

ROn background composition. BiPo shapes.



BiPo enriched, Topology 30, all Phasel cycles, FPNT subtracted;

315

All shapes are in 5% agreement.

ROn background composition. FPNT rates.



- Normalised FPNT rates from the preselected Topology 30 △T fits;
- As expected, ROn yield is higher \implies reactor power correlation?

CMY

39/50

ROn background composition. FPNT shapes.



FPNT sample, Topology 30, all Phasel cycles, FPNT subtracted;

イロト イヨト イヨト イヨト

315

All shapes are in 10% agreement.

ROn background composition. FPNT stability.



- The relative FPNT increase is consistent along the Phase I;
- An amount of extra FPNTs is correlated with the reactor power.

CMY

27/03/24 41/50

Control plots. Quick reminder.

- The control variables were originally selected from the variables that do not participate in the MVA with an intention to test the combined prediction of MC IBD + ROff to the ROn shape-wise, which complemets the yield comparison provided by the simultaneous ΔT-ΔR fit;
- The control variables include: ΔT , ΔR , BiPonator and BDT_{ATM2} scores;
- ► E_{AC} , $\bar{\nu}$ energy estimator and AC_Z are also interesting, but contain physics;
- They can be kept for control plots and χ² tests for the individual ROn periods in which physics is hidden by statistics. However, they will be removed as soon as we will move to the samples with larger statistics.

42/50

Control plots. χ^2 test. Result.



The total number of plots above the threshold is statistically affordable;

They are evenly split along topologies, variables and periods.

CMY

27/03/24 43/50

Initial conditions. Reminder.



Ratio of signal and background events as a function of the ΔT in the raw data sets.

	<	三郎 スポッスポッス セン	. nac
CMY	BCD 2024	27/03/24	44/50

Final conditions.



Ratio of signal and background events as a function of the ΔT in the 5D optimised data sets.

	4	미 에 소 🖓 에 소 필 에 소 필 어	三日 りょつ
CMY	BCD 2024	27/03/24	45/50

Predictions. Summary

S/B

SoLid

F

CMY

10

- Optimized 2γ S/B = 1.5 (19.1 ± 1.3 ± 1.5 ν/d), where 1st number is 5D optimised MC predict, 2nd number - statistical unc. and 3rd - systematic unc.;
- The blind topological analysis is concluded with the staged scrutiny of the 21 days open data set

15

20



Analysis of the open data set

- Optimized 2γ S/B = 1.5 (19.1 ± 1.3 ± 1.5 $\bar{\nu}$ /d)
- Optimized 2γ S/B = 1.8 (21.9 \pm 2.1 \pm 1.5 $\bar{\nu}$ /d)
- The blind topological analysis is concluded with the staged scrutiny of the 21 days open data set

S/B



180È

160

140

120

OTAL

 $r^2 = 0.88$

preliminar

Analysis of the open data set

- Optimized 2γ S/B = 1.5 (19.1 ± 1.3 ± 1.5 $\bar{\nu}$ /d)
- Optimized 2γ S/B = 1.8 (21.9 \pm 2.1 \pm 1.5 $\bar{\nu}$ /d)
- The blind topological analysis is concluded with the staged scrutiny of the 21 days open data set



180

120F 100

> 80 60

> 40

20

preliminary

 $^{2} = 1.36$

Events/Period0 160 140F



- Analysis developed to maximally use the granularity of the detector;
- Focusing on the most complicated topologies, a decent S/B is obtained;
- Novel reconstruction method introduced to project optimally the measured fiber energy into cubes (CCubes);
- PVT response calibrated by horizontal muons.

Work in progress and Outlook

- Absolute energy scale to be finalised;
- Full Phase I data to be analysed following the unblinding strategy;
- ► The CCube and calibration notes are in progress to be published shortly;

- Around 6k $\bar{\nu}$ are expected for SoLid Phase I (2 gamma) w/ a decent S/B;
- Sensitivity study to oscillation parameters has to be conducted;
- Antineutrino spectrum to be analysed.

BACKUP