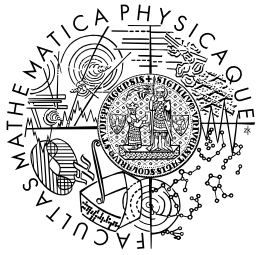

FACULTY
OF MATHEMATICS
AND PHYSICS
Charles University

DOCTORAL THESIS

Tadeáš Dohnal

Precision Measurement of Neutrino Oscillations Parameters

Institute of Particle and Nuclear Physics

Supervisor of the doctoral thesis: prof. RNDr. Rupert Leitner, DrSc.
Study programme: Physics
Study branch: Particle and Nuclear Physics

Prague 2022

I declare that I carried out this doctoral thesis independently, and only with the cited sources, literature and other professional sources. It has not been used to obtain another or the same degree.

I understand that my work relates to the rights and obligations under the Act No. 121/2000 Sb., the Copyright Act, as amended, in particular the fact that the Charles University has the right to conclude a license agreement on the use of this work as a school work pursuant to Section 60 subsection 1 of the Copyright Act.

In ..... date ....

Author's signature

I would like to thank Professor Rupert Leitner for guidance over all the years since my bachelor studies. My deep gratitude goes out to Bedřich Roskovec for a number of insightful discussions and I am also grateful to the rest of our California-Czech-Hong Kong group within the Daya Bay experiment for many valuable comments and suggestions.

Last, but not least, I would like to thank my wife, Hedvika, and the rest of my family for the overall support they provided me with.

Title: Precision Measurement of Neutrino Oscillations Parameters

Author: Tadeáš Dohnal

Institute: Institute of Particle and Nuclear Physics

Supervisor: prof. RNDr. Rupert Leitner, DrSc., Institute of Particle and Nuclear Physics

Abstract: The Daya Bay Reactor Neutrino Experiment, located in Southeast China, was designed to measure short baseline oscillation of electron antineutrinos originating from six 2.9 GW_{th} nuclear reactors. In 2012, it discovered the nonzero value of the θ_{13} mixing angle and currently provides the world's most precise measurement of this neutrino oscillation parameter. The goal of this thesis is to check the performance of the detector nonuniformity correction used by the Daya Bay experiment and then design a new updated version. For that purpose, the Daya Bay experiment is introduced including the procedure of the event reconstruction. Several types of signals are then studied in detail and later they are used to scrutinize the performance of the nonuniformity correction. Several issues are identified, most significantly an absence of radial dependence in the azimuthal part of the correction and inability to deal with dead photomultipliers. A new improved version of the correction that addresses these issues is then designed with the main goal to improve the precision of Daya Bay neutrino oscillation analyses, especially ongoing one which uses neutron capture on hydrogen. The impact of the new improved correction on said analyses is also assessed.

Keywords: Neutrino, Neutrino mixing, Neutrino oscillations, Neutrino experiments

Contents

Introduction

Neutrinos are electrically neutral leptons first proposed in 1930 by Wolfgang Pauli to allow for conservation of energy and angular momentum in beta decays. However, it took over twenty years until they were first observed in 1956 [1]. Since then, three generations of neutrinos have been discovered along with corresponding charged leptons – electron (e^-) and electron neutrino (ν_e), muon (μ^-) and muon neutrino (ν_μ), tau lepton (τ^-) and tau neutrino (ν_τ).

Since neutrinos are the second most abundant particles in the universe after the photons, it is important to study and understand their properties. The Standard Model of particle physics describes neutrinos as massless and interacting only via weak interaction. However, the fact that neutrino flavor can change between its creation and detection in a phenomenon called neutrino oscillation (its experimental discovery was awarded by the Nobel Prize in 2015) showed that neutrinos have nonzero masses. Moreover, it is likely that the oscillations of neutrinos and antineutrinos differ (CP-symmetry[1] is violated), which could shed light on why we live in matter-rich world. Precise measurement of neutrino oscillation and the parameters that drive it is thus crucial for our understanding of the universe.

One of experiments measuring neutrino oscillation is the Daya Bay Reactor Neutrino Experiment, which is very much in the focus of this thesis. It discovered the nonzero value of the θ_{13} neutrino mixing angle in 2012 [2] and currently provides the world's most precise measurement of this parameter [3]. Such results cannot be obtained without precise energy calibration. The main goal of the work presented here is to improve the energy calibration of the Daya Bay detectors.

The thesis is organized as follows. The theoretical framework of neutrino oscillation is presented in Chapter 1 along with introduction of the relevant parameters and overview of experiments that measure these parameters. The Daya Bay experiment is introduced in Chapter 2 with overview of the detector system and the physics program, which covers far more than just neutrino oscillation measurement that we focus on.

In Chapter 3, we will go through the process of event reconstruction in the Daya Bay experiment. One step in the reconstruction chain is the application of the nonuniformity correction, which we need to understand, so that later we can make a better updated version. However, before we get there we first need to identify suitable signals that would allow us to study and quantify the detector response. We will thus dedicate Chapter 4 to a detailed assessment of signals at hand and their further usefulness.

Having identified the appropriate signals, we will move on to scrutinize the original nonuniformity correction and eventually design a new improved version in Chapter 5. Finally, we will use preliminary results from two distinct Daya Bay neutrino oscillation analyses to assess what the impact of the improved correction on the determination of the oscillation parameters is.

[1] Combined charged conjugation and parity symmetry.

1. Neutrino Oscillation

The history of neutrino oscillation as a concept goes back to 1957 when it was proposed by Bruno Pontecorvo with the assumption of neutrinos being composite particles undergoing neutrino-antineutrino transition [4]. In 1962, Maki, Nakagawa and Sakata elaborated on the idea with the difference that they considered mixing between neutrino flavor states instead [5]. This idea was again revisited by Pontecorvo in 1967 with neutrino oscillation described as a transition between neutrinos of different flavors (for example $\nu_e \leftrightarrow \nu_\mu$). In the early 1970s, *solar neutrino problem* arose when only about one-third of expected solar neutrino flux was observed by the Homestake experiment[6] and neutrino mixing offered an explanation. However, it took many years until the first compelling evidence for neutrino oscillation was provided by the Super-Kamiokande experiment in 1998 [7], this time based on the observation of the zenith angle dependence of the atmospheric neutrino flux. Many experiments have followed since then measuring the oscillation of neutrinos from various sources and determining the driving parameters. Almost all of the measurements are consistent with nowadays well-established three active neutrino framework.

Natural units ($c = 1$, $\hbar = 1$) are used throughout this chapter unless stated otherwise.

1.1 Neutrino Mixing

The phenomenon of neutrino oscillation stems from two characteristics of neutrinos. First, their flavor eigenstates $|\nu_e\rangle$, $|\nu_\mu\rangle$, $|\nu_\tau\rangle$ (generally denoted $|\nu_\alpha\rangle$) are not identical to the mass eigenstates $|\nu_1\rangle$, $|\nu_2\rangle$, $|\nu_3\rangle$ (generally denoted $|\nu_i\rangle$). Second, there are differences between masses of the neutrino mass eigenstates (which also means that at least two of the three masses are nonzero).

The neutrino mixing between the flavor and mass eigenstates is given by

$$|\nu_\alpha\rangle = \sum_{i=1}^{3} U^*_{\alpha i} \, |\nu_i\rangle \, , \tag{1.1}$$

where U is a unitary in general complex mixing matrix also known as Pontecorvo-Maki-Nakagawa-Sakata (PMNS) matrix. Its elements are given by

$$U_{\alpha i} \equiv \langle \nu_\alpha | \nu_i \rangle \, . \tag{1.2}$$

Analogous expression for antineutrinos can be derived using the CPT symmetry

$$\bar{U}_{\alpha i} \equiv \langle \bar{\nu}_\alpha | \bar{\nu}_i \rangle = \langle \nu_i | \nu_\alpha \rangle \equiv U^*_{\alpha i}, \tag{1.3}$$

which leads to

$$|\bar{\nu}_\alpha\rangle = \sum_{i=1}^{3} U_{\alpha i} \, |\bar{\nu}_i\rangle \, . \tag{1.4}$$

Let us consider for a moment that there are n neutrino species. The PMNS matrix is then $n \times n$ complex unitary matrix, which can be parameterized by n^2 independent parameters: $\frac{1}{2}n(n-1)$ angles and $\frac{1}{2}n(n+1)$ phases. We can remove n

phases by redefinition of n non-physical phases of charged leptons. In addition to that, $(n-1)$ more phases can be removed if neutrinos are Dirac particles[1]. Even though these phases cannot be removed, if neutrinos are Majorana particles[2], they do not affect neutrino oscillation. Either way, we end up with $(n-1)^2$ parameters relevant to neutrino oscillation, out of which $\frac{1}{2}n(n-1)$ are mixing angles and $\frac{1}{2}(n-1)(n-2)$ are phases.

In particular, in the framework of three active neutrinos, the PMNS matrix is then given by 4 parameters – 3 mixing angles (commonly denoted θ_{12}, θ_{13}, θ_{23}) and one phase δ which if nonzero can lead to CP-symmetry violation in neutrino oscillation. The PMNS matrix takes form

$$U_{\alpha i} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix} \begin{pmatrix} c_{13} & 0 & s_{13}e^{i\delta} \\ 0 & 1 & 0 \\ -s_{13}e^{-i\delta} & 0 & c_{13} \end{pmatrix} \begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{1.5}$$

where s_{ij} is shortened notation for $\sin(\theta_{ij})$ and c_{ij} for $\cos(\theta_{ij})$.

1.2 Neutrino Oscillation in Vacuum

As all neutrinos that we focus on in this thesis have energy of the order of MeV or higher while their masses are on sub-eV scale, we can safely assume that they are ultrarelativistic. That allows us to use simplified formula for energy E_i of the i-th neutrino mass eigenstate $|\nu_i\rangle$

$$E_i = \sqrt{p_i^2 + m_i^2} \cong p_i + \frac{m_i^2}{2E_i}, \tag{1.6}$$

where p_i and m_i denote its momentum and mass respectively. Assuming that neutrino mass eigenstate propagates as a plane wave, its evolution is described as

$$|\nu_i(t,L)\rangle = e^{-iE_i t} e^{+ip_i L} |\nu_i\rangle , \tag{1.7}$$

where L is the distance that the neutrino traversed and t is the time it took.

We will assume now that we have a neutrino created in weak interaction along with a corresponding charged lepton, i.e. neutrino created in a particular flavor eigenstate. As it was shown in Eq. 1.1, such flavor eigenstate can be written as a superposition of mass eigenstates, for its propagation thus follows

$$|\nu_\alpha(t,L)\rangle = \sum_{i=1}^{3} U_{\alpha i}^* |\nu_i(t,L)\rangle . \tag{1.8}$$

This allows us to calculate the probability $P_{\alpha\beta}$ that neutrino created in $|\nu_\alpha\rangle$ flavor eigenstate is later found in $|\nu_\beta\rangle$ flavor eigenstate

$$P_{\alpha\beta} = |\langle \nu_\beta | \nu_\alpha(t,L)\rangle|^2 = |\sum_{i=1}^{3}\sum_{j=1}^{3} U_{\alpha i}^* U_{\beta j} \langle \nu_j | \nu_i(t,L)\rangle|^2. \tag{1.9}$$

[1] If neutrinos are Dirac particles, neutrino and antineutrino are different particles and the total lepton number is in presently known processes conserved.

[2] If neutrinos are Majorana particles, neutrino and antineutrino are in fact the same particle differing only in helicity. If neutrinos also have nonzero masses, there are processes involving them where the lepton number is not conserved.

Combining Eq. (1.7) and Eq. (1.9) we get

$$P_{\alpha\beta} = \sum_{i=1}^{3}\sum_{j=1}^{3} e^{-i\phi_{ij}(L,t)} U^*_{\alpha i} U_{\beta i} U_{\alpha j} U^*_{\beta j}, \tag{1.10}$$

where $\phi_{ij}(t, L)$ is a relative phase factor stemming from propagation of neutrino mass eigenstates $|\nu_i\rangle$ with differing masses. Using ultrarelativistic limit $t = L$ and Eq. 1.6, we get

$$\phi_{ij}(t, L) = (E_i - E_j)t - (p_i - p_j)L \cong \frac{(m_i^2 - m_j^2)L}{2E} \equiv \frac{\Delta m_{ij}^2 L}{2E}. \tag{1.11}$$

Eq. (1.10) can be further altered (along with restoration of $\hbar$ and c)

$$P_{\alpha\beta} = \delta_{\beta\alpha} - 4\sum_{i<j} \mathrm{Re}[U^*_{\alpha i} U_{\beta i} U_{\alpha j} U^*_{\beta j}] \sin^2\left(\frac{\Delta m_{ij}^2 L}{4\hbar c E}\right) + \\ + 2\sum_{i<j} \mathrm{Im}[U^*_{\alpha i} U_{\beta i} U_{\alpha j} U^*_{\beta j}] \sin\left(\frac{\Delta m_{ij}^2 L}{2\hbar c E}\right). \tag{1.12}$$

Looking closer at Eq. (1.12), the first two terms remain the same even if we interchange initial and final neutrino state, i.e. they keep the T-symmetry[3] with $P_{\alpha\beta} = P_{\beta\alpha}$. These two terms are also not affected by interchange of neutrinos for antineutrinos or vice versa, i.e. they keep the CP-symmetry with $P_{\alpha\beta} = P_{\bar{\alpha}\bar{\beta}}$.

On the other hand, the third terms violates these symmetries – if it is nonzero, these symmetries are not conserved in neutrino oscillation. Most notably, CP-symmetry violation introduces difference between oscillation of neutrinos and antineutrinos and thus, generally speaking, difference in behavior of matter and antimatter. One of necessary conditions for that is the PMNS matrix being complex due to $\delta \neq 0, \pi$ (along with nonzero values of all θ_{12}, θ_{23}, θ_{13}). Nevertheless, even if that is satisfied, CPT symmetry still guarantees that $P_{\alpha\beta} = P_{\bar{\beta}\bar{\alpha}}$ even if $P_{\alpha\beta} \neq P_{\beta\alpha}$, $P_{\alpha\beta} \neq P_{\bar{\alpha}\bar{\beta}}$. Note that when survival probability of $|\nu_\alpha\rangle$ neutrino flavor $P_{\alpha\alpha}$ is measured in *neutrino disappearance experiments*[4], this CP- and T-violating term is not relevant; it plays a role in *neutrino appearance experiments* where the presence of $|\nu_\beta\rangle$ neutrino flavor in $|\nu_\alpha\rangle$ neutrino flux is measured.

When we look at the CP- and T-symmetry conserving part of Eq. (1.12), the oscillatory behavior is apparent there – the amplitude is given by the relevant elements of the PMNS matrix and the oscillation length measured in L/E depends on the differences between squares of neutrino masses Δm_{ij}^2. The distance between two maxima (or minima) is then

$$\left(\frac{L}{E}\right)_{\text{period}} = \frac{4\pi\hbar c}{\Delta m_{ij}^2}. \tag{1.13}$$

For practical purposes it is useful to evaluate the constants in phase factor ϕ_{ij} in suitable units

$$\frac{\phi_{ij}}{2} \equiv \frac{\Delta m_{ij}^2 L}{4E\hbar c} \doteq 1.27 \left(\frac{\Delta m_{ij}^2}{\mathrm{eV}^2}\right)\left(\frac{L}{\mathrm{m}}\right)\left(\frac{\mathrm{MeV}}{E}\right) \tag{1.14}$$

[3]Time-reversal symmetry.

[4]Neutrino disappearance experiments measure how neutrino flux of a particular neutrino flavor decreases due to neutrino oscillation. In contrast, neutrino appearance experiments search for a neutrino flavor that was not originally present in the neutrino flux, but appears due to neutrino oscillation.

In neutrino oscillation experiments, detectors are commonly placed in the first disappearance minimum or appearance maximum. Position of such minimum or maximum is given by

$$\left(\frac{L}{\mathrm{m}}\right)\left(\frac{\mathrm{MeV}}{E}\right) \approx \frac{\pi}{2.54}\left(\frac{\mathrm{eV}^2}{\Delta m^2_{ij}}\right). \tag{1.15}$$

More about the physics of neutrino oscillation can be found in Refs. [8], [9] or [10].

1.3 Neutrino Oscillation Parameters

There are six parameters that determine neutrino oscillation in the three active neutrino framework:

- Two independent differences between squares of masses (any 2-element subset of Δm^2_{21}, Δm^2_{31}, Δm^2_{32} is sufficient, the third one is then given by $\Delta m^2_{31} = \Delta m^2_{32} + \Delta m^2_{21}$)
- Three mixing angels (θ_{12}, θ_{13}, θ_{23})
- A CP-violating phase (δ)

These parameters are measured in a number of experiments using variety of neutrino sources, most prominent of which are solar neutrinos (ν_e's created in solar thermonuclear reactions), reactor neutrinos ($\bar{\nu}_e$ emitted from neutron-rich nuclei originating in fission in nuclear reactors), atmospheric neutrinos that are produced when cosmic rays hit atmosphere (ν_μ, $\bar{\nu}_\mu$, ν_e, $\bar{\nu}_e$ coming from decays of pions, muons, kaons etc.) and accelerator neutrinos (ν_μ, $\bar{\nu}_\mu$ from decays of pions produced by accelerators).

The values of neutrino oscillation parameters according to the global fit by the Particle Data Group [11] are listed in Tab. 1.1. There are, however, still several unknowns regarding the neutrino oscillation.

The question of the *neutrino mass ordering* (also called *neutrino mass hierarchy*) is not yet resolved, i.e. whether the ν_3 mass eigenstate is the heaviest (*normal mass ordering*) or the lightest (*inverted mass ordering*) among neutrino mass eigenstates. If the best-fit value of a particular oscillation parameter depends on the mass ordering, values for both mass orderings are included in Tab. 1.1.

Currently, there is experimental preference for the normal mass ordering at $2\,\sigma - 3\,\sigma$ level [11]. JUNO experiment [12], which is under construction in China, has neutrino mass ordering determination as the main goal and hopefully it will help to resolve this question.

Arguably even more important question of whether CP-symmetry is violated in neutrino oscillation or not has also not yet been fully answered. It hinges on the value of the phase δ. Status of its measurement is described in the corresponding section at the end of the chapter.

Table 1.1: Global fit values of neutrino oscillation parameters according to Particle Data Group [11]. When different, values for both mass orderings are listed.

Parameter	Value ± standard deviation
Δm^2_{21}	$(7.53 \pm 0.18) \cdot 10^{-5}\,\mathrm{eV}^2$
Δm^2_{32}	$(2.453 \pm 0.033) \cdot 10^{-3}\,\mathrm{eV}^2$ (normal ordering)
Δm^2_{32}	$(-2.536 \pm 0.034) \cdot 10^{-3}\,\mathrm{eV}^2$ (inverted ordering)
$\sin^2\theta_{12}$	0.307 ± 0.013
$\sin^2\theta_{23}$	0.546 ± 0.021 (normal ordering)
$\sin^2\theta_{23}$	0.539 ± 0.022 (inverted ordering)
$\sin^2\theta_{13}$	0.0220 ± 0.0007
δ	$4.27^{+0.63}_{-0.50}$ rad

Measurement of θ_{12} and Δm^2_{21}

Let us now review how the solar parameters were determined. As the θ_{12} mixing angle is significantly larger than θ_{13}, the oscillation of ν_e and $\bar{\nu}_e$ is mostly determined by θ_{12} and Δm^2_{21}, which means that the survival probability is given by two neutrino approximation up to a certain precision

$$P_{ee}(L/E) = P_{\bar{e}\bar{e}}(L/E) \cong 1 - \sin^2(2\theta_{sol}) \sin^2\left(\frac{\Delta m^2_{sol} L}{4\hbar c E}\right), \tag{1.16}$$

where effective parameters $\Delta m^2_{sol} \cong \Delta m^2_{21}$ and $\theta_{sol} \cong \theta_{21}$ are used. The *sol* in subscript stands for "solar" as these parameters were first measured using solar ν_e. Due to the fact that solar core is significantly larger than oscillation length (and due to decoherence) we cannot observe individual oscillation waves, only their average. Another important factor is the effect that solar matter has on passing neutrinos. Neutrinos with energies $E_\nu \lesssim 0.5$ MeV are rather unaffected by solar matter [10] and their mean survival probability (as ν_e's) is $1 - 0.5\sin^2(2\theta_{sol})$. Meanwhile neutrinos with higher energies tend to leave Sun in mass eigenstate ν_2 (the heavier from ν_1, ν_2) with ν_e content $\sin^2(\theta_{sol}) \simeq 1/3$ which explains the problem of missing solar neutrinos.

Solar neutrinos were observed in experiments like Homestake [6], SAGE [13], GALLEX [14], Super-Kamiokande [15], SNO [16] and Borexino [17], neutrino oscillation guided by Δm^2_{21} and θ_{12} was also measured using reactor antineutrinos in the KamLAND experiment [18] providing the most precise value of Δm^2_{21}. The neutrino oscillation waves observed by the KamLAND experiment are shown in Fig. 1.1.

As listed in Tab. 1.1, the best-fit values of the relevant oscillation parameters were determined to be $\sin^2 2\theta_{12} = 0.85 \pm 0.02$ [11] which corresponds to $\theta_{12} \approx 34°$ and $\Delta m^2_{21} = (7.53 \pm 0.18) \cdot 10^{-5}\,\mathrm{eV}^2$ [11], which means that the first disappearance minimum of ν_e and $\bar{\nu}_e$ lies in the distance of 16.4 km/MeV in vacuum (note that the sign of Δm^2_{21} is known based on the effect of solar matter).

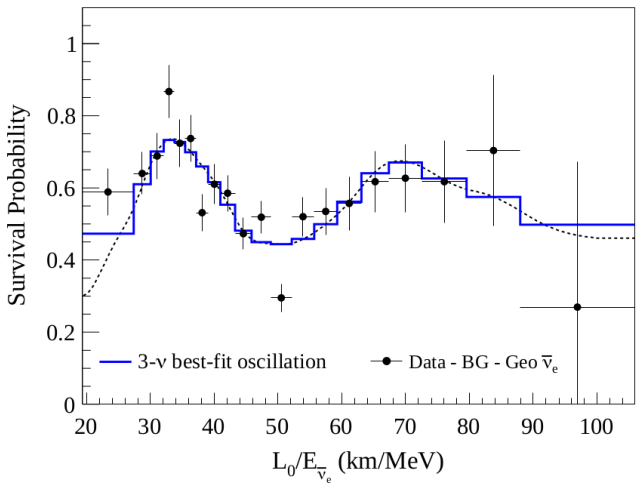

Figure 1.1: KamLAND's measurement of ratio of the observed $\bar{\nu}_e$ spectrum to the spectrum that would be expected without neutrino oscillation as a function of L_0/E, where $L_0 = 180$ km is the flux-weighted average reactor baseline and E is the $\bar{\nu}_e$ energy. Figure taken from Ref. [18].

Measurement of θ_{23} and Δm^2_{32}

The absolute value of Δm^2_{31} or Δm^2_{32} is about 33 times bigger than Δm^2_{21}. This places the first disappearance minimum to about 0.5 km/MeV for both Δm^2_{32} and Δm^2_{31} driven oscillations. However, as mentioned above, it is not yet known whether ν_3 is the heaviest or the lightest mass eigenstate (i.e. the sign of Δm^2_{31} or Δm^2_{32}).

The oscillation driven by Δm^2_{31} and Δm^2_{32} was first observed by the Super-Kamiokande experiment in 1998 [7]. Two-neutrino framework can be used to describe this oscillation with good precision

$$P_{\mu\mu}(L/E) = P_{\bar{\mu}\bar{\mu}}(L/E) \cong 1 - \sin^2(2\theta_{atm}) \sin^2\left(\frac{\Delta m^2_{atm} L}{4\hbar c E}\right), \tag{1.17}$$

where $\theta_{atm} \approx \theta_{23}$ and $\Delta m^2_{atm} \approx \Delta m^2_{31} \approx \Delta m^2_{32}$. The subscript atm stands for "atmospheric" reflecting the fact that oscillation driven by these parameters plays an important role in the observation of atmospheric ν_μ and $\bar{\nu}_\mu$ flux.

Later on, ν_μ and $\bar{\nu}_\mu$ beams produced using particle accelerators became an essential tool to study neutrino oscillation driven by Δm^2_{32} and Δm^2_{31}. These were studied in experiments like K2K [19], succeeded by T2K [20], MINOS, MINOS+ [21] and NOνA [22].

Along with uncertainty regarding the sign of Δm^2_{32} and Δm^2_{31}, there is also ambiguity regarding the value of θ_{23}. As $\sin^2 2\theta_{23} \approx 1$, it follows that $\theta_{23} \approx 45°$, however it is not entirely clear whether $\theta_{23} > 45°$ or $\theta_{23} < 45°$ (i.e. to which octant θ_{23} belongs).

Measurement of θ_{13} and Δm^2_{31}

Being the smallest one out from the three mixing angles, θ_{13} was the last one to be measured, its nonzero value was discovered in 2012 by the Daya Bay experi-

ment [2]. There are currently two strategies employed in θ_{13} determination. One is to use reactor $\bar{\nu}_e$ disappearance in a similar way to θ_{12}, but with shorter baseline as in the experiments Double Chooz [23], RENO [24] and the aforementioned Daya Bay experiment, which provides the world's most precise measurement of $\sin^2 2\theta_{13} = 0.0851 \pm 0.0024$ [3] corresponding to $\theta_{13} \approx 8.5°$.

The other strategy is to use accelerator ν_e ($\bar{\nu}_e$) appearance in ν_μ ($\bar{\nu}_\mu$) beam in the already mentioned experiments T2K, MINOS, MINOS+, NOνA. Apart from $\sin^2 2\theta_{13}$ mixing angle, Δm^2_{31} or Δm^2_{32} mass splitting is also measured in these experiments. The Daya Bay's measurement of the mass splitting is of comparable precision as the ones provided by the leading accelerator neutrino experiments.

Measurement of the CP-violating Phase δ

Our knowledge regarding the δ-phase stems mostly from ν_e ($\bar{\nu}_e$) appearance in accelerator neutrino experiments T2K [25] and NOvA [22], but there is currently tension between results from these two experiments. For CP-symmetry to be violated in neutrino oscillation it is necessary that $\delta \neq 0, \pi$ along with nonzero values of all 3 mixing angles, which is likely satisfied as $\delta = 4.27^{+0.63}_{-0.50}$ rad according to the global fit[11], but no value of δ has been excluded at 5 σ level so far. Next-generation experiments DUNE [26] and Hyper-Kamiokande [27], which are currently under construction, will provide precise measurement of δ.

2. The Daya Bay Experiment

The Daya Bay Reactor Neutrino Experiment was designed to measure the θ_{13} mixing angle by studying reactor antineutrino disappearance caused by neutrino oscillation at about 2 km baseline. It began data taking in December 2011 and finished operation at the end of 2020. The location of the experiment is about 55 km northeast from Hong Kong in Southeast China, in the proximity of the Daya Bay and Ling Ao nuclear power plants as shown in Fig. 2.1. These power plants operate in total six 2.9 GW_{th} pressurized water reactors, which makes them one of the most powerful nuclear complexes in the world. Each reactor isotropically emits $\sim 6 \times 10^{20}$ $\bar{\nu}_e$/s originating in beta-decays of fission fragments and thus serves as an intense and pure source of $\bar{\nu}_e$'s with energies up to ~10 MeV.

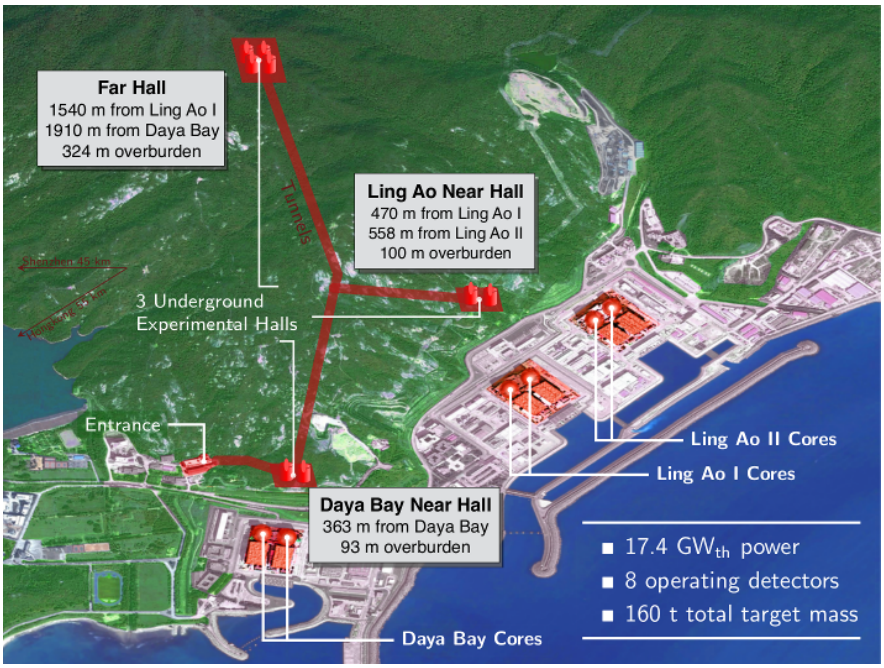

Figure 2.1: Layout of the Daya Bay experiment. Six reactor cores serve as a powerful and pure source of $\bar{\nu}_e$'s. Mostly unoscillated flux and spectrum is sampled by two near experimental halls (EH1 & EH2) while the far hall (EH3) sees the largest effect of the oscillation driven by the θ_{13} mixing angle. All experimental halls are placed underground to suppress cosmic rays with corresponding overburdens indicated.

2.1 Detector Configuration

The Daya Bay experiment nominally utilized 8 antineutrino detectors (ADs) of identical design which were located in 3 underground experimental halls (EHs) as listed[1] in Tab. 2.1 along with overburdens (experimental halls are located

[1]Note that different AD labels based on installation order can be found in the Daya Bay literature.

underground for suppression of cosmic rays and associated backgrounds) and distances from the individual reactor cores. The idea behind the experimental layout is that EH1 (containing 2 ADs) is located close to the Daya Bay nuclear power plant and EH2 (also containing 2 ADs) close to the Ling Ao nuclear power plant (EH1 and EH2 are thus labeled as the "near halls") with the aim to sample the reactor $\bar{\nu}_e$ flux and spectrum only minimally affected by neutrino oscillation. EH3 is located in almost 2 km distance (thus labeled as the "far hall"), where the maximal disappearance of reactor $\bar{\nu}_e$ due to neutrino oscillation driven by the θ_{13} mixing angle can be observed.

Table 2.1: List of experimental halls with corresponding depths, detectors and distances between each detector and each reactor core [28]. D1 and D2 stand for the Daya Bay nuclear power plant reactors, whereas L1, L2, L3 and L4 stand for Ling Ao reactors. List of baselines with full precision of 18 mm can be found in Ref. [29].

Hall	Depth [m]	Detector short (long) label	Baseline [m] with respect to reactor: D1	D2	L1	L2	L3	L4
EH1	93	AD1 (EH1-AD1)	362	372	903	817	1354	1265
		AD2 (EH1-AD2)	358	368	903	817	1354	1266
EH2	100	AD3 (EH2-AD1)	1332	1358	468	490	558	499
		AD4 (EH2-AD2)	1337	1363	473	495	559	501
EH3	324	AD5 (EH3-AD1)	1920	1894	1533	1534	1551	1525
		AD6 (EH3-AD2)	1918	1892	1535	1535	1555	1528
		AD7 (EH3-AD3)	1925	1900	1539	1539	1556	1530
		AD8 (EH3-AD4)	1923	1898	1541	1541	1560	1533

2.2 Data-taking Periods

While the Daya Bay experiment was designed to operate with 8 ADs, not all of them were operational all the time. Only 6 ADs were present from the start of data taking in December 2011 to July 2012 in so called "6-AD period" which was followed by a shutdown during which the other two ADs (EH2-AD2 and EH3-AD4) were installed. The period of full configuration ("8-AD period") began in October 2012 and lasted until December 2016 when EH1-AD1 was repurposed to study liquid scintillator properties for the JUNO experiment currently under construction [12]. The experiment continued in configuration with 7 ADs from January 2017 until the end of data-taking in December 2020 ("7-AD period").

2.3 Detector System

Each hall contains an instrumented water pool in which the corresponding ADs are placed, as shown in Fig. 2.2 for a near hall as an example. The water pool shields the ADs from gammas coming from natural radioactivity in the surrounding rock and muon-induced neutrons. It is optically separated into two

sections – inner and outer one. Both are instrumented by PMTs in order to detect Cherenkov light produced by passing muons and thus serve as a veto. On top of the water pools, 4-layer resistive plate chamber (RPC) system was installed to improve the muon detection, but it was later phased out.

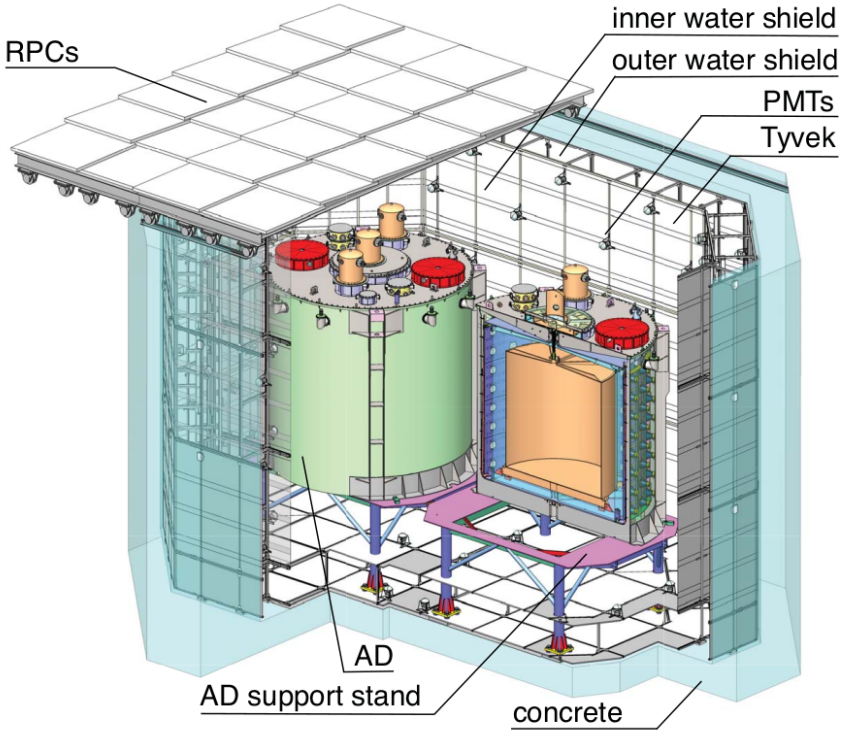

Figure 2.2: Cross section of the detector system of a near hall containing two antineutrino detectors. Instrumented water pool divided into two sections separated by tyvek sheets serves as both passive shielding from radiation and as a Cherenkov detector for muons. System of resistive plate chambers covers the pool for improved muon identification [29].

Each individual AD consists of three nested cylindrical regions as shown in Fig. 2.3. The innermost volume is contained in a vessel made from 11 mm thick acrylic with height and diameter of 3 m and contains 20 t of liquid scintillator doped with gadolinium at 0.1% mass fraction (thus being referred to as "GdLS"). It is surrounded by 22 t of liquid scintillator without gadolinium doping inside 17 mm thick 4 m by 4 m vessel (this volume is referred to as "LS"). These two volumes are nested in a 5 m by 5 m stainless steel vessel (SSV) filled with mineral oil (labeled as "MO"), which does not scintillate and serves as a buffer shielding the LS and GdLS from outside radiactivity. There are 192 inward-facing 8-inch photomultiplier tubes (PMTs) arranged in 8 rows and 24 columns mounted around the perimeter of the mineral oil volume. They collect scintillation light produced by ionizing particles in GdLS and LS volumes. The top and bottom of the AD are not instrumented by PMTs, instead there are specular reflectors that help with collection of the scintillation light.

Three automated calibration units (ACUs) are mounted on the top of SSV. Two are connected to the GdLS volume by calibration tubes – the first one

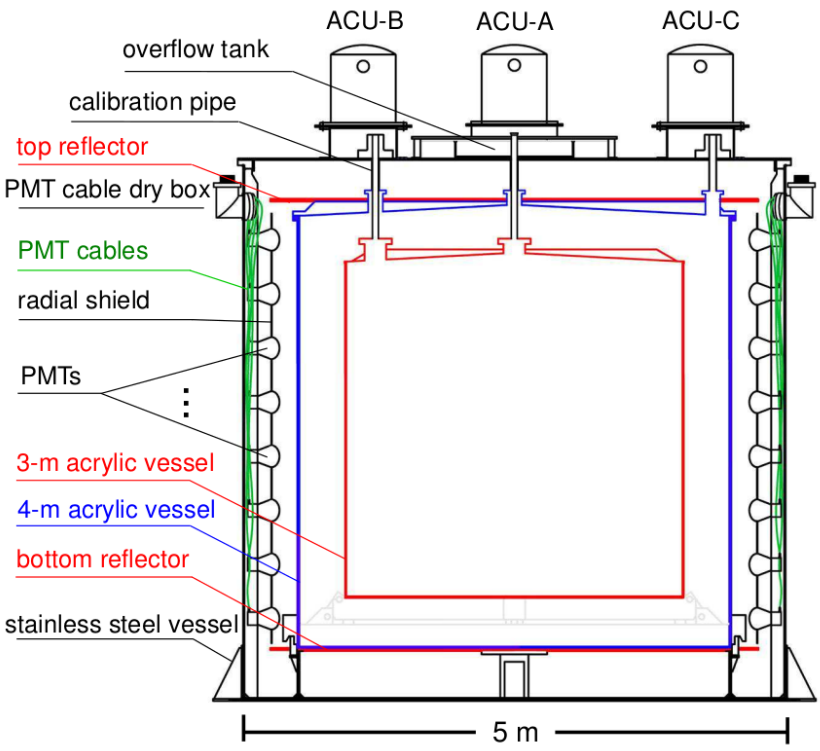

Figure 2.3: Cross section of an antineutrino detector. Gadolinium-doped liquid scintillator (GdLS) in the center is surrounded by scintillator without doping (LS) which in turn is nested in mineral oil (MO). Scintillation light is detected by 192 PMTs mounted in eight rows. On top, three automated calibration units (ACUs) are placed [29].

(ACU-A) is located above the center of the AD and deploys calibration sources or LEDs along the AD axis, the second one (ACU-B) operates at the distance of 1350 mm from the axis. The third one (ACU-C) is connected to the LS volume at 1772.5 mm from the axis. The tube in the center (under ACU-A) also connects the GdLS and LS volumes to concentric overflow tanks located on the SSV lid. The air above the overflow tanks is continuously purged with dry nitrogen, the same applies for the space above scintillator surface in ACU-B and ACU-C calibration tubes.

2.4 Antineutrino Detection

When an ionizing particle passes through the liquid scintillator, photons are emitted and detected by PMTs allowing for determination of position and deposited energy (details about event reconstruction can be found in the Chapter 3). The reactor antineutrinos are detected via the inverse beta decay (IBD) reaction on a free proton

$$\bar{\nu}_e + p \rightarrow e^+ + n, \tag{2.1}$$

which has a threshold of about 1.8 MeV.

As liquid scintillator based on linear alkylbenzene is a hydrocarbon compound,

free protons are abundant there for antineutrinos to interact with. Positron created in IBD quickly looses energy and eventually annihilates emitting in most cases two gammas, which also promptly deposit their energy. Neutron, on the other hand, thermalizes first and then is captured on a nucleus. In GdLS, it is most likely gadolinium (nGd) with mean capture time ∼30 μs. Gamma cascade is emitted with energy ∼8 MeV. Even if a gamma escapes the GdLS, it is likely contained in LS thus minimizing energy leakage. Meanwhile, neutrons in LS are mainly captured on hydrogen (nH) with mean capture time ∼200 μs. A single 2.2 MeV gamma is then emitted. The detection scheme is outlined in Fig. 2.4. The spatial ($\lesssim$ 50 cm) and temporal coincidence of the prompt positron signal and the delayed neutron capture signal allows for a powerful background rejection in general. The energy of the prompt signal E_{prompt} can be directly related to the incoming antineutrino energy $E_{\bar{\nu}_e}$ as

$$E_{\bar{\nu}_e} \simeq E_{\text{prompt}} + 0.78 \text{ MeV}. \tag{2.2}$$

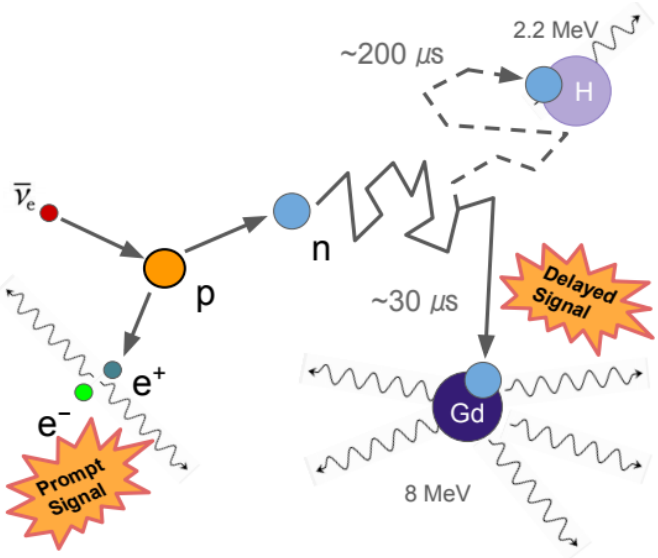

Figure 2.4: Scheme of $\bar{\nu}_e$ interaction and detection in the Daya Bay experiment. As $\bar{\nu}_e$ interacts via IBD, positron and neutron are created. The positron quickly loses energy and annihilates giving us the prompt signal. The neutron first thermalizes and after that gets captured on a nucleus. In GdLS, it is most likely gadolinium, in LS hydrogen with mean capture times ∼30 μs and ∼200 μs respectively. The gamma(s) emitted after the neutron capture give us the delayed signal. Figure taken from Ref. [30].

2.5 Selection of Physics Results

2.5.1 Neutrino Oscillation in the Three Neutrino Framework

The survival probability of $\bar{\nu}_e$ in the three neutrino framework is given by

$$P_{\bar{e}\bar{e}} = 1 - \cos^4(\theta_{13})\sin^2(2\theta_{12})\sin^2\left(\frac{\Delta m^2_{21}L}{4\hbar cE}\right) - \\ - \sin^2(2\theta_{13})\left[\cos^2(\theta_{12})\sin^2\left(\frac{\Delta m^2_{31}L}{4\hbar cE}\right) + \sin^2(\theta_{12})\sin^2\left(\frac{\Delta m^2_{32}L}{4\hbar cE}\right)\right], \quad (2.3)$$

where L is the distance traversed by $\bar{\nu}_e$ and E is its energy. Note that in short baseline experiments such as Daya Bay, the part of the oscillation that is driven by Δm^2_{21} has only minor impact on the oscillation probability.

Eq. 2.3 can be rewritten as

$$P_{\bar{e}\bar{e}} \simeq 1 - \cos^4(\theta_{13})\sin^2(2\theta_{12})\sin^2\left(\frac{\Delta m^2_{21}L}{4\hbar cE}\right) - \sin^2(2\theta_{13})\sin^2\left(\frac{\Delta m^2_{ee}L}{4\hbar cE}\right), \quad (2.4)$$

where we used an effective mass splitting Δm^2_{ee}, which is independent of mass ordering, and it is defined as

$$\sin^2\left(\frac{\Delta m^2_{ee}L}{4E}\right) \simeq \cos^2(\theta_{12})\sin^2\left(\frac{\Delta m^2_{31}L}{4E}\right) + \sin^2(\theta_{12})\sin^2\left(\frac{\Delta m^2_{32}L}{4E}\right). \quad (2.5)$$

Daya Bay was the first experiment to discover the nonzero value of the θ_{13} mixing angle just in 2012 using nGd sample [2]. Since then, several improved measurements have been published ([31, 32, 33] and most notably Ref. [29] where the experiment and measurement are described in great detail) with the latest one corresponding to 3158 days of data taking [3]. During that time, unprecedented sample of over 5.5 million IBD candidates in total with almost 0.7 million in the far hall (EH3) were collected.

By comparing $\bar{\nu}_e$ flux and energy spectral shape across all ADs (the difference between the near halls and the far hall being essential) as shown in the left panel of Fig. 2.5 the values of the parameters driving the oscillation were determined with the results $\sin^2 2\theta_{13} = 0.0852 \pm 0.0024$ and $\Delta m^2_{ee} = (2.519 \pm 0.060) \times 10^{-3}$ eV^2 using the Δm^2_{ee} effective mass splitting. With the full three neutrino framework, the results are $\sin^2 2\theta_{13} = 0.0851 \pm 0.0024$ and $\Delta m^2_{32} = (2.466 \pm 0.060) \times 10^{-3}$ eV^2 (normal mass ordering) or $\Delta m^2_{32} = (-2.571 \pm 0.060) \times 10^{-3}$ eV^2 (inverted mass ordering) [3]. The best fit and the confidence regions are shown in the right panel of Fig. 2.5. This Daya Bay's final measurement of $\sin^2 2\theta_{13}$ using the full data set has the world's best precision of 2.8%. The Δm^2_{32} mass splitting is also measured with a leading precision comparable to that achieved by the accelerator experiments.

The neutrino oscillation measurement can be also performed using IBD sample with neutron capture on hydrogen (nH). So far, only analysis comparing antineutrino rates across all ADs based on 621 days of data taking has been published with the result $\sin^2 2\theta_{13} = 0.071 \pm 0.011$ [34]. Measurement based on nH is largely independent from the nGd measurement, since the statistical samples

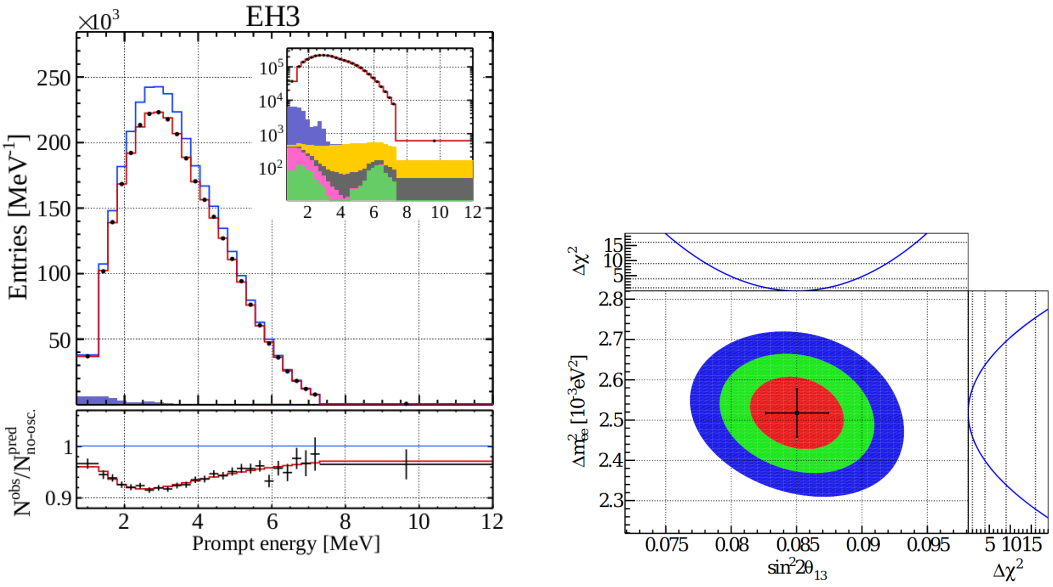

Figure 2.5: (Left) Energy spectrum of reactor antineutrinos detected in the far hall ADs (black points) is compared to the prediction based on the near halls data assuming no oscillations (blue line) and with the best oscillation fit (red line) [3]. (Right) The best fit of $\sin^2 2\theta_{13}$ and Δm^2_{ee} and the 1, 2, 3σ confidence regions [3].

are completely distinct and systematical uncertainties are mostly different too. Analysis using nH suffers from larger systematics, especially larger background dominated by the accidental coincidences of two uncorrelated signals. A spectral analysis is currently under preparation, which means that precise energy calibration is needed especially in LS, where most of the IBD interaction resulting in nH occur. The goal of this thesis is to improve the uniformity of the AD energy reconstruction with focus on this ongoing neutrino oscillation analysis based on nH.

2.5.2 Search for Sterile Neutrino Mixing

Even though most experiments are in line with the model of three active neutrino mixing, there are some anomalies which do not conform to it and which can be explained by introduction of sterile neutrinos [35], most notably in LSND [36] and MiniBooNE [37] experiments. The possibility of sterile neutrino mixing can be probed using 3 (active) + 1 (sterile) framework. In the Daya Bay experiment, this would introduce additional oscillation driven by new parameters Δm^2_{41} and θ_{14}. Daya Bay's layout with multiple baselines allows it to explore several orders of $|\Delta m^2_{41}|$. Based on 1230 days of data taking, no deviation from the three neutrino model was observed [38] resulting in limits on $\sin^2 2\theta_{14}$ for almost four orders of magnitude in $|\Delta m^2_{41}|$ with the limits for $|\Delta m^2_{41}| < 0.2$ eV^2 being the most stringent ones up to date, as shown in the left panel of Fig. 2.6.

The limits imposed by the Daya Bay experiment can be extended by limits from the Bugey-3 experiment [42]. When combined with the results from the

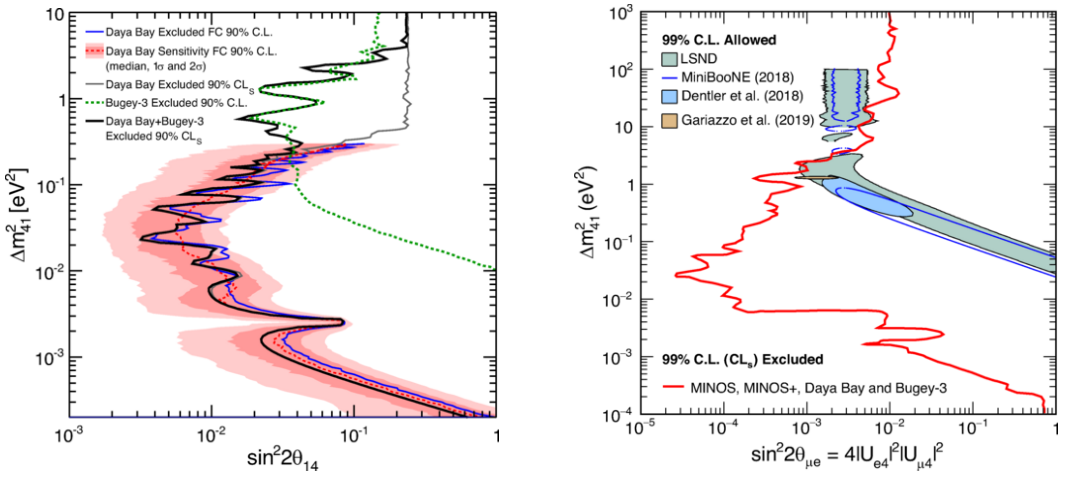

Figure 2.6: (Left) Exclusion region at 90% C.L. from Daya Bay based on Feldman-Cousins (blue line) and CL_s (gray line) statistical methods along with excluded region from Bugey-3 (green dashed line) and combination of both experiments (black line) [38]. (Right) Exclusion region at 99% CL_s based on combination of the Daya Bay, Bugey-3, MINOS and MINOS+ experiments (right of the red line) compared with 99% C.L. allowed region from the LSND and MiniBooNE experiments [38]. A global fit [39, 40] excluding Daya Bay, Bugey-3, MINOS and MINOS+ is ruled at >99% C.L. along with a fit to appearance experiments only [41].

MINOS and MINOS+ experiments [43] they allow to assess the findings of LSND and MiniBooNE which observed an excess of $\nu_e/\bar{\nu}_e$ in $\nu_\mu/\bar{\nu}_\mu$ beam. This excess points to the sterile neutrino mixing with $\Delta m^2_{41} \sim 1$ eV^2 and an effective mixing angle $\sin^2 2\theta_{\mu e} \equiv \sin^2 2\theta_{14} \sin^2 \theta_{24}$. While the value of θ_{14} is constrained by the Daya Bay and Bugey-3 $\bar{\nu}_e$ disappearance experiments, limits on θ_{24} are provided by the MINOS and MINOS+ measurement of $\nu_\mu/\bar{\nu}_\mu$ disappearance. Thus when results from these experiments were combined, strong constrains on the sterile neutrino mixing over seven orders of magnitude in Δm^2_{41} were imposed. As shown in the right panel, the LSND and MiniBooNE 99% C.L. allowed regions were excluded at 99% CL_s for $\Delta m^2_{41} < 1.2$ eV^2. Thus the anomalies observed in these experiments were likely not caused by sterile neutrinos. In addition to that, global fit to the experiments searching for sterile neutrinos [39, 40] as well as fit only to the appearance experiments [41] were fully excluded on more than 99% C.L..

2.5.3 Measurement of the Reactor Antineutrino Flux and Spectrum

Precise measurement of the reactor $\bar{\nu}_e$ flux was performed using 2.2 million IBD candidates collected in the near halls over 1230 days [44]. The average IBD yield was determined to be $(5.91 \pm 0.09) \times 10^{-43}\,\mathrm{cm}^2/\mathrm{fission}$. While it is consistent with previous experimental results as shown in the left panel of Fig. 2.7, when compared to the Huber-Mueller model [45, 46] the ratio of measured to predicted yield is $0.952 \pm 0.014 \pm 0.023$, where the first uncertainty is experimental and the second one comes from the theoretical model. This discrepancy is called "reactor

antineutrino anomaly".

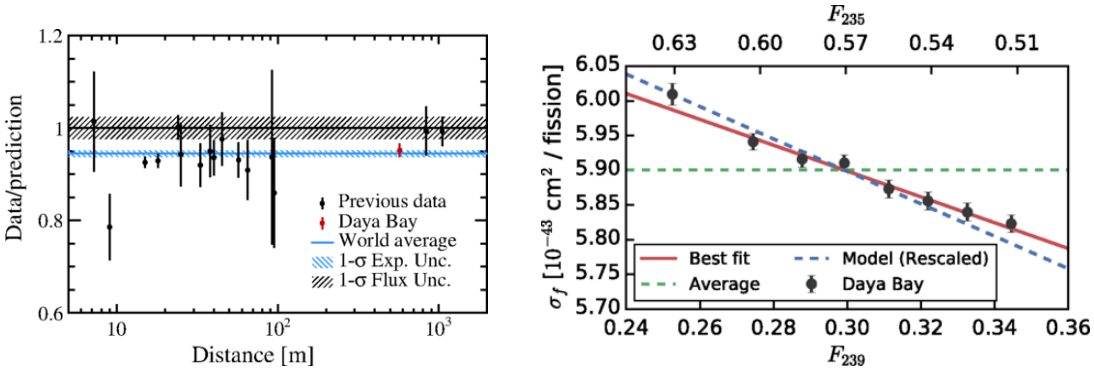

Figure 2.7: (Left) The ratio of measured reactor antineutrino yield to the Huber-Mueller theoretical prediction as a function of the distance from reactor to detector [44]. Each ratio is corrected for the effect of neutrino oscillation. The Daya Bay measurement is shown at the flux weighted baseline (578 m) of the two near halls. (Right) IBD yield per fission, σ_f, versus effective ^{239}Pu (bottom axis) or ^{235}U (top axis) fission fraction [47].

Using the same data set, correlations between reactor core fuel evolution and changes in the reactor antineutrino flux and energy spectrum were also analyzed [47]. Fuel-dependent variation in the IBD yield was observed (as shown in the right panel of Fig. 2.7) rejecting the hypothesis of a constant antineutrino flux as a function of the ^{239}Pu fission fraction at 10 standard deviations and the hypothesis of a constant antineutrino energy spectrum at 5.1 standard deviations. Individual yields of the two most prominent fissile isotopes ^{235}U and ^{239}Pu were determined to be (6.17 ± 0.17) and $(4.27 \pm 0.26) \times 10^{-43}$ cm^2/fission respectively (shown in the left panel of Fig. 2.8) suggesting that ^{235}U is the primary contributor to the reactor antineutrino anomaly.

Along with improved measurement of the prompt energy spectrum of reactor antineutrinos, individual antineutrino spectra from ^{235}U and ^{239}Pu fissions were extracted using evolution of the prompt spectrum as a function of the isotope fission fractions based on 3.5 million IBD candidates collected in the near halls over 1958 days [48]. Total spectrum showed discrepancy of over 5 standard deviations when compared to Huber-Mueller model, especially significant in 4-6 MeV "bump" region ($>$ 6 standard deviations).

Joint analysis with the PROSPECT experiment, which detects $\bar{\nu}_e$'s from highly enriched uranium compact research reactor core, was also performed [49]. It showed that the measured ^{235}U antineutrino spectra are consistent between the two experiments. Combined analysis then reduced correlation between extracted ^{235}U and ^{239}Pu spectra (shown in the right panel of Fig. 2.8) and it also reduced ^{235}U spectrum shape uncertainty.

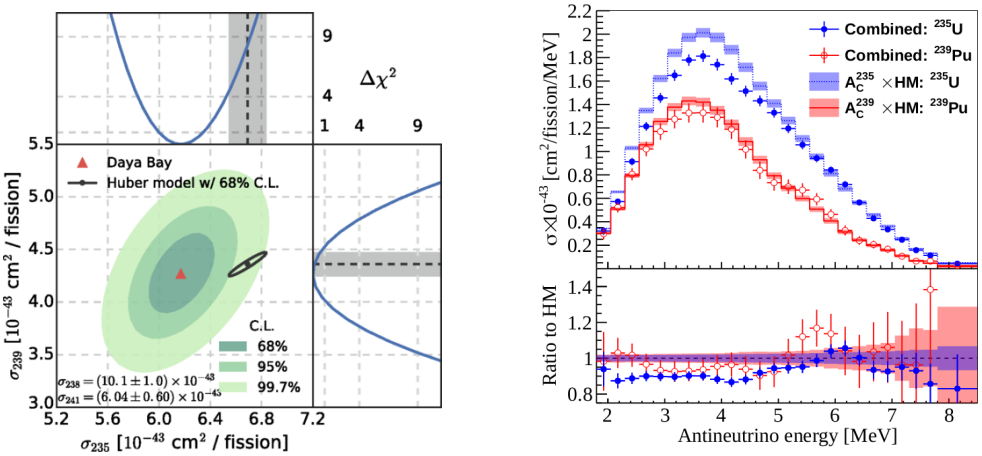

Figure 2.8: (Left) Combined measurement of ^{235}U and ^{239}Pu IBD yields per fission σ_{235} and σ_{239} [47]. The red triangle indicates the best fit with corresponding contours in green while prediction of thr Huber-Mueller model is shown in black. (Right) Top – ^{235}U and ^{239}Pu antineutrino spectra unfolded from the jointly deconvolved Daya Bay and PROSPECT measurements [49]. Bottom – ratio of the measurements to their respective models.

3. Event Reconstruction

When a particle deposits energy in one of the sensitive volumes of the AD, scintillation light is produced, propagates and part of the photons are eventually detected by PMTs. The position and energy of the interaction thus needs to be deduced from the spatial and temporal PMT hit pattern as well as from the information about spatial non-uniformity (resulting mostly from optical properties determining light propagation, geometric acceptance of PMTs etc.), nonlinear energy response of the liquid scintillator and electronics nonlinearity.

First, the analog signal given by a PMT is processed by an analog-to-digital converter (ADC) measuring the collected charge and a time-to-digital converter (TDC) extracting information about the hit time. For the next steps proper calibration is needed in order to extract physical quantities from ADC and TDC values.

3.1 Time and PMT Gain Calibration

LED calibration source was used to synchronize individual PMTs to eliminate the effect of different cable lengths. The absolute timing is then given by GPS-synchronized time stamp with 25 ns precision.

To reconstruct position and deposited energy of an event, we need to determine the conversion between ADC counts and number of photoelectrons registered by a PMT, i.e. to establish the gain of each PMT channel. The gain of a particular PMT depends on many factors such as the input voltage, ambient temperature or orientation with respect to the local magnetic field. It thus follows that the gain is generally not stable in time and needs to be determined for each PMT individually on continuous basis.

PMT dark noise (primarily single photoelectrons originating from thermal emissions in the photocatode) was used on daily basis to establish the mean charge $\overline{Q}^{\mathrm{SPE}}$ induced by a single photoelectron in ADC counts (on average about 19 counts/photoelectron).

3.2 Position Reconstruction

The coordinate system used in event reconstruction has its origin in the center of the GdLS. Both cylindrical (r,z,ϕ) and Cartesian (x, y, z) coordinate systems are used with z-axis points up along the axis of symmetry of the cylindrically-shaped detector and x- and y-axes define the horizontal plane and are connected to the cylindrical coordinates in the standard way ($x = r\cos\phi$, $y = r\sin\phi$).

There are several event reconstruction algorithms used in the Daya Bay experiment, but we will focus on the one referred as "reconstruction B" in Ref. [29] and its extension, which are also called "AdSimple" and "AdSimpleNL" within the Daya Bay collaboration – these labels will be used further on. The AdSimpleNL method will be used in the analysis described in Chapters 4 and 5.

In order to reconstruct event position (it works the same way in both AdSimple and AdSimpleNL), a library of 9600 charge templates generated in Monte Carlo

simulation is employed with each template corresponding to one position on grid with 20 bins in r direction, 20 bins in z direction and 24 bins in ϕ direction. The charge pattern of the event that is being reconstructed is then compared to these templates by calculating χ^2 [29]

$$\chi^2(r,z,\phi) = \sum_{i=1}^{192} \left(-2 \ln \frac{\text{Poisson}(N_i^{\text{obs}} | N_i^{\text{temp}}(r,z,\phi))}{\text{Poisson}(N_i^{\text{obs}} | N_i^{\text{obs}})} \right), \tag{3.1}$$

where $\text{Poisson}(n|\mu)$ stands for Poisson probability of observing n photoelectrons when the mean value is μ, N_i^{obs} is the observed number of photoelectrons by the i-th PMT and $N_i^{\text{temp}}(r,z,\phi)$ is the expected number of photoelectrons in the i-th PMT predicted by the template corresponding to the simulated event position (r, z, ϕ).

The reconstruction precision is further increased by interpolation between few best-matching simulated templates. For each coordinate, the χ^2 of the template that yielded minimal value and the two neighboring ones was fitted by a parabolic function. The final position is then taken from the minimum of the parabolic function.

The performance of the AdSimple postion reconstruction was tested using the calibration system with ^{60}Co source and it was shown that reconstruction bias was < 20 cm and the resolution < 40 cm [29].

3.3 Energy Reconstruction

Let us assume that an event happened in the AD and we have its reconstructed position (r, z, ϕ) and charge collected by each PMT (Q_i for i-th PMT) along with current gain calibration for each PMT ($\overline{Q}_i^{\text{SPE}}$ giving us the number of ADC counts per photoelectron for i-th PMT). The reconstructed energy E_{rec} is then given by [29]

$$E_{\text{rec}} = \left(\sum_{i=1}^{192} f_{\text{SCNL}}^* \frac{Q_i}{\overline{Q}_i^{\text{SPE}}} \right) \frac{f_{\text{act}}(t)}{N^{\text{PE}}(t) f_{\text{AP}}} f_{\text{pos}}(r,z,\phi,t), \tag{3.2}$$

where $f_{\text{act}}(t)$ is the active PMT correction, N^{PE} photoelectron yield, f_{AP} is a factor correcting for PMT afterpulses caused by shower muons and $f_{\text{pos}}(r,z,\phi,t)$ is the non-uniformity correction. The AdSimple reconstruction method does not use single channel nonlinearity correction f_{SCNL}^* (hence the asterix) – it was introduced in its derivative method AdSimpleNL. Let us now look at each of these factors in a bit more detail.

3.3.1 Single Channel Nonlinearity Correction

The single channel nonlinearity correction factor f_{SCNL}^* is only present in AdSimpleNL reconstruction method, which was derived from AdSimple in 2017 after extensive study of Daya Bay energy response nonlinearity [50]. It corrects the nonlinearity of PMTs and front-end electronics, which will be discussed in more detail in Section 3.3.5.

3.3.2 Active PMT Correction

If some PMT channels were not working during a certain period (due to being disconnected, broken or due to the corresponding channel(s) in the high-voltage module being broken), their absence needed to be compensated by factor $f_{\text{act}}(t) = 192/N(t)$, where $N(t)$ stands for the number of PMT channels operating at the given time (out of 192 in total).

3.3.3 Photoelectron Yield

The average number of photoelectrons observed by PMTs as a result of particle interaction of fixed energy varied both from AD to AD and for a single AD in time. The conversion factor between deposited energy and number of photoelectrons thus had to be determined for each AD individually on a regular basis. For that purpose, we can use a peak of known energy and the observed number of photoelectrons to make a conversion between these two quantities. In AdSimple(NL) reconstruction method, peak from neutron capture on Gadolinium (nGd) was used, where the neutrons come from spallation of nuclei (mainly carbon) induced by cosmic-ray muons (SPNs – spallation neutrons). The flux of muons gets lower with increasing overburden which means that the sufficient statistics for a calibration point was accumulated approximately once per day in the shallower near halls and once per week in the deeper far hall.

If the neutron is captured on Gd, it is either ^{155}Gd or ^{157}Gd isotope. The abundance of ^{157}Gd is 15.65%, its thermal neutron capture cross section is 257,000 barn and energy released in gamma cascade is 7.937 MeV. The second isotope ^{155}Gd has similar abundance of 14.80%, but lower neutron capture cross section of 60,700 barn, while the energy released is 8.536 MeV [51]. This means that the neutron captures on ^{155}Gd constitute 22.7% of ^{157}Gd captures and the observed nGd peak in fact consists of two peaks which are mostly merged together due to energy resolution.

The fitting of the nGd peak is performed using double *Crystall Ball* (CB) function [52] which consists of two CB functions. A single (unnormalized) CB function of energy E is defined using parameters μ, σ, n and α

$$f_{\text{CB}}(E;\mu,\sigma,n,\alpha) = \begin{cases} \exp\left(-\frac{(E-\mu)^2}{2\sigma^2}\right) & \text{for } \frac{E-\mu}{\sigma} > -\alpha \quad \text{(peak)} \\ A \cdot \left(B - \frac{E-\mu}{\sigma}\right)^{-n} & \text{for } \frac{E-\mu}{\sigma} \leq -\alpha \quad \text{(tail)} \end{cases} \tag{3.3}$$

$$A = \left(\frac{n}{|\alpha|}\right)^n \cdot e^{-\frac{\alpha}{2}}, \qquad B = \frac{n}{|\alpha|} - |\alpha|.$$

As we can see, CB function has two parts sewn together at the point given by parameter α. Above this point it is a Gaussian function with mean μ and standard deviation σ. If there was no energy leakage, this would be sufficient, however, as the gammas from the cascade may deposit their energy or part of it in non-scintillating volumes, tail towards smaller energies appears. This is reflected by the second part of CB function with parameter n defining its shape.

To fit the nGd double peak we need to introduce double CB function

$$f_{\text{DCB}} = N_1 \cdot f_{\text{CB1}}(E;\mu_1,\sigma_1,n_1,\alpha_1) + N_2 \cdot f_{\text{CB2}}(E;\mu_2,\sigma_2,n_2,\alpha_2), \tag{3.4}$$

where index 1 corresponds to neutron capture on ^{157}Gd and index 2 to capture on ^{155}Gd. Fortunately, the number of free parameters in Eq. 3.4 can be reduced as there are particular relations between these parameters. First, the normalization factor N is directly proportional to the abundance and the neutron cross section of the isotope, which are known quantities. The relation between N_1 and N_2 is thus given as

$$N_2 = 0.227\, N_1. \tag{3.5}$$

Next, we also know the energies released when a neutron is captured on each isotope. These energies correspond to parameters μ_1 and μ_2, the relation between them is then

$$\mu_2 = 1.0755\, \mu_1. \tag{3.6}$$

In a similar fashion with the assumption that $\sigma \propto \sqrt{\mu}$, we get

$$\sigma_2 = \sqrt{1.0755}\, \sigma_1. \tag{3.7}$$

Lastly, we can assume that both ^{157}Gd and ^{155}Gd peaks have the same shape defined by α and n parameters

$$\alpha_2 = \alpha_1,\ n_2 = n_1. \tag{3.8}$$

By fitting the SPN nGd spectrum by the double CB function, we get the value of peak mean μ_1 in units of photoelectrons. We can thus establish the conversion factor between these two quantities, which is on average about 168 photoelectrons/MeV. The exact dependence on AD and time is shown in Fig. 3.1.

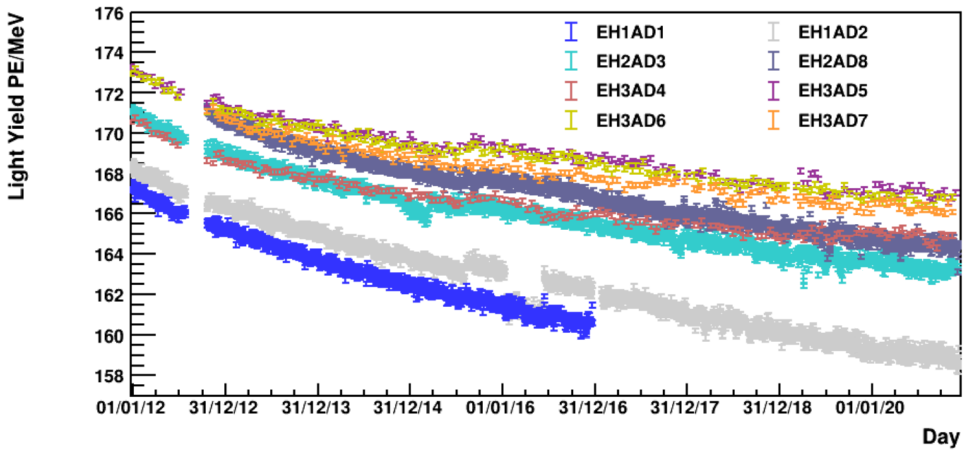

Figure 3.1: Observed photoelectron (PE) yield as a function of time for each AD individually. All ADs showed consistent sub-percent yearly decline of photoelectron yield. Figure taken from [53].

There is, however, one more factor we need to account for. Shower muons cause massive production of scintillation light, which leads to PMT afterpulsing and thus temporarily slightly shifting the energy scale when compared to IBD nGd [54]. The shift increases with muon energy and we thus need f_{AP} correction factor to account for it. The f_{AP} factor was derived using the assumption that when muon energy is extrapolated to zero there should be match between SPN

and IBD nGd energies. The mean muon energy varies between EHs because of different overburdens, which is the source of differences in the values of f_{AP} between EHs. The f_{AP} value of ADs in the same hall tends to be quite similar. Parameters of cosmic muon flux and f_{AP} values are summarized in Tab. 3.1.

Table 3.1: Overburdens of EHs in meters-water-equivalent (mwe), mean muon energy $\langle E_\mu \rangle$ and rate R_μ in each EH. Correction factor f_{AP} is listed for each AD.

Hall	Overburden	$\langle E_\mu \rangle$	R_μ	Detector	f_{AP}
EH1	250 mwe	57 GeV	1.27 Hz/m^2	EH1-AD1	0.9927
				EH1-AD2	0.9934
EH2	265 mwe	58 GeV	0.95 Hz/m^2	EH2-AD1	0.9921
				EH1-AD2	0.9922
EH3	860 mwe	137 GeV	0.056 Hz/m^2	EH3-AD1	0.9901
				EH3-AD2	0.9904
				EH3-AD3	0.9899
				EH3-AD4	0.9895

3.3.4 Nonuniformity correction

The amount of light that reaches the PMTs depends on the position of the light production due to the experiment geometry, geometrical acceptance of PMTs, optical properties of liquid scintillator, acrylic vessels and mineral oil. The quantum efficiency (probability that an incident photon causes electron emission from the photocathode) may vary between PMTs and even the collection efficiency (probability that the photoelectron reaches the first dynode) is affected by orientation of the PMT with respect to the local magnetic field. The reconstructed energy would deviate between about -6% and +15% from the nominal energy[1] depending on the position in the AD (as illustrated in Fig. 3.2) if no correction was applied. That would seriously affect the energy scale and the energy resolution. It is thus clear that precise correction is necessary. Even though all ADs behave in a similar way, there is generally up to 2% difference between the ADs. It thus follows that the correction has to be applied for each AD individually so that these differences can be addressed.

The nonuniformity correction for AdSimple(NL) reconstruction method has a form of a function of reconstructed coordinates r, z, ϕ, t. It is factorized into three parts

$$f_{\mathrm{pos}}(r, z, \phi, t) = f_\phi(\phi) f_{\mathrm{rz}}(r, z) f_{\mathrm{rt}}(r, t), \tag{3.9}$$

where f_ϕ denotes azimuthal correction, f_{rz} the main spatial correction working on r-z coordinate space and f_{rt} time-dependent correction which is also r-dependent. First two correction factors, $f_\phi(\phi)$ and $f_{\mathrm{rz}}(r, z)$, are determined for each AD independently. While the nonuniformity patterns are qualitatively similar, there are quantitative differences that need to be addressed. On the other hand, the last correction factor $f_{\mathrm{rt}}(r, t)$ was found to be even quantitatively similar between ADs enough for it to be common for all ADs.

[1] Energy given by fitting of SPN nGd peak as described in Section 3.3.3. As nGd occurs only in GdLS, nominal photoelectron yield is calculated for GdLS as whole.

Azimuthal Correction

The azimuthal correction aims to eliminate residual effect of local Earth's magnetic field. While the PMTs were wrapped in truncated conical magnetic shields, there is still a residual effect resulting in deviation of reconstructed energy up to 1 % depending on the azimuthal angle ϕ. The correction factor is calculated as

$$f_\phi(\phi) = \frac{1}{1 + \alpha_0^{\text{corr}} \sin(\phi - \phi_0^{\text{corr}})}, \tag{3.10}$$

where α_0^{corr} and ϕ_0^{corr} are data-driven parameters with values listed in Tab. 3.2. The orientation of the correction given by parameter ϕ_0^{corr} is calculated in the local coordinate system of the particular EH. The values of ϕ_0^{corr} are thus rather similar for ADs in the same EH as expected for effect caused by residual geomagnetic field, but differ between EHs as the local coordinate system of each EH is differently oriented with respect to the geomagnetic field. On the other hand, the differences in amplitude α_0^{corr} suggest that the level of shielding (or PMT susceptibility) varies between the ADs, even in the same EH. Apart from geomagnetic field, the azimuthal nonuniformity was also affected by several PMTs which stopped working during the operation of the experiment. The impact of these dead PMTs on nouniformity was not addressed by the aforementioned correction and it seems unlikely that it even could be addressed by a correction in the form of Eq. 3.10 as a dead PMT would be expected to have a rather localized and time-dependent (before vs. after the PMT stopped working) effect. For that and other reasons, the correction given by Eq. 3.10 will be scrutinized and a new improved correction will be presented in Chapter 5.

Table 3.2: Parameters which are used for $f_\phi(\phi)$ azimuthal nouniformity correction factor calculation as described by Eq. 3.10.

Detector	α_0^{corr} $[10^{-2}]$	ϕ_0^{corr} [°]
EH1-AD1	0.348	138.3
EH1-AD2	0.657	141.5
EH2-AD2	0.675	-5.0
EH2-AD2	0.424	7.4
EH3-AD1	0.427	68.0
EH3-AD2	1.115	60.2
EH3-AD3	0.670	35.0
EH3-AD4	0.997	47.5

Spatial Correction

The spatial nonuniformity correction is given by $f_{\text{rz}}(r, z)$ term in Eq. 3.9 and addresses the differences (between -6% and $+15\%$) in observed energy depending on the r-z position in the AD. In contrast to the azimuthal correction described by a continuous function, spatial correction uses discrete 2D map M_{NU} to calculate the corresponding factor

$$f_{\text{rz}}(r, z) = \frac{1}{M_{\text{NU}}(r^2, z)}. \tag{3.11}$$

The M_{NU} nonuniformity map covers the AD sensitive volume with 10 bins on the r^2-axis uniformly spanning interval (0 m^2, 4 m^2) and 10 bins on the z-axis uniformly spanning interval (-2 m, 2 m), in total 100 voxels. The square of radius, r^2, is used instead of r so that each voxel corresponds to the same volume and thus contains approximately the same number of events for the correction determination. The value in each bin indicates relative amount of scintillation light that is detected for an event in that particular spatial voxel compared to the GdLS average. $M_{\mathrm{NU}}(r^2, z)$ in Eq. 3.11 then stands for a value in the voxel corresponding to r^2, z coordinates. If an event is reconstructed outside of the M_{NU} grid, correction factor from the nearest voxel is used instead. An example of the nonuniformity map used for EH1-AD1 is shown in in Fig. 3.2.

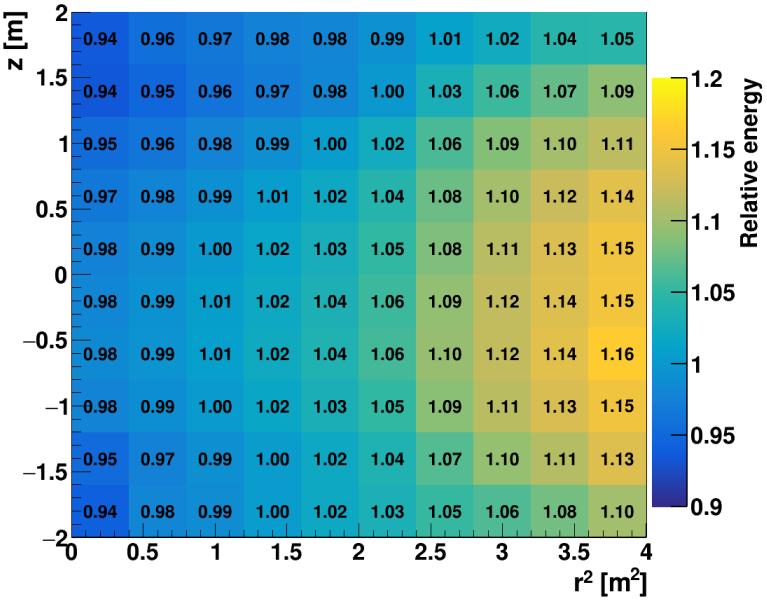

Figure 3.2: An example of the nonuniformity map M_{NU}, which is used in AdSimpleNL reconstruction in EH1-AD1. The value in each bin indicates the relative amount of scintillation light that is detected for an event in that particular voxel compared to the GdLS average. In general more scintillation light is detected with increasing r and decreasing $|z|$. The pattern is very similar for all ADs. During the correction, factor of $1/M_{\mathrm{NU}}$ is applied.

In AdSimple(NL), the map M_{NU} is constructed using SPN nGd peak for voxels that are completely inside GdLS, SPN nH peak for voxels inside LS and arithmetic average of both for bins that cover the boundary of GdLS and LS as illustrated in Fig. 3.3.

The fitting of nGd peak is performed by the double CB function as described by Eqs. 3.3 and 3.4 with additional decreasing exponential function that is aimed to account for tail going to higher energies

$$f_{\mathrm{nGd}} = f_{\mathrm{DCB}} + e^{-\frac{E}{\lambda}},\ \lambda > 0. \tag{3.12}$$

On the other hand, the nH peak is fitted by the so-called Daya Bay (DYB) function also referred to as the Calorimeter function given by Eqs.3.13 and 3.14 and

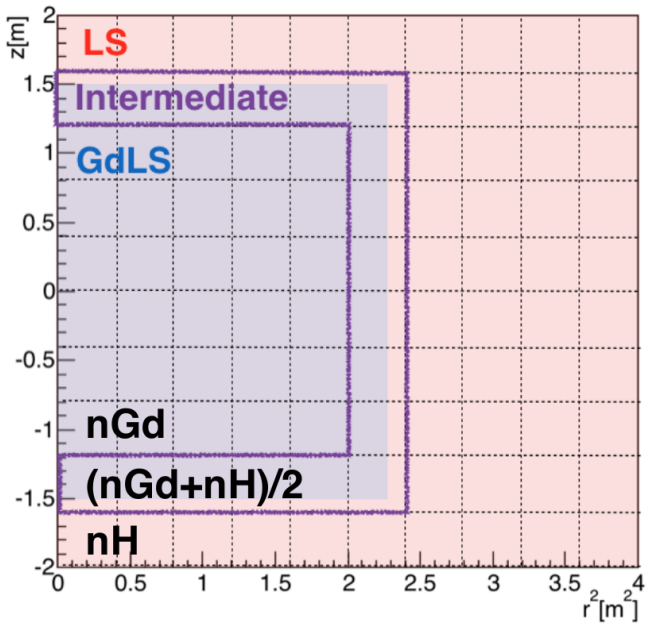

Figure 3.3: Scheme of r^2-z spatial nonuniformity correction. SPN nGd signal is used in GdLS and SPN nH signal in LS. In the intermediate region, arithmetic average of both is used. Figure taken from Ref. [55].

described in detail in Ref. [56]. As suggested in Fig. 3.4, the idea behind the DYB function is that without energy resolution smearing there would be a true energy spectrum. It would consist from a peak represented by delta function and a leakage tail going from the peak energy to lower energies represented by exponential function, constant function or combination of both. The final function is then obtained by applying Gaussian smearing with a common standard deviation for all energies to the true energy spectrum. The crucial difference to the CB function is that there is no sewing point but both smeared peak and tail in principle span the whole energy range.

The building blocks of DYB function are defined as

$$\begin{aligned} f_{\text{peak}}(E;\mu,\sigma) &= \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(E-\mu)^2}{2\sigma^2}} \\ f_{\text{exp-tail}}(E;\mu,\sigma,\lambda) &= \frac{\lambda}{2} e^{\frac{\sigma^2\lambda^2+2\lambda E}{2}} \left[\text{erf}\left(\frac{\mu - E - \sigma^2\lambda}{\sqrt{2}\sigma}\right) - \text{erf}\left(\frac{-E-\sigma^2\lambda}{\sqrt{2}\sigma}\right)\right] \\ f_{\text{const-tail}}(E;\mu,\sigma) &= \frac{1}{\mu}\left[\text{erf}\left(\frac{\mu - E}{\sqrt{2}\sigma}\right) - \text{erf}\left(\frac{-E}{\sqrt{2}\sigma}\right)\right], \end{aligned} \tag{3.13}$$

where μ stands for the true peak energy, σ for the energy resolution and λ describes the steepness of the exponential tail. Each of the DYB function building blocks is normalized to 1 so a scaling factor is needed when fitting data. For the purpose of SPN nH fitting for nonuniformity map determination DYB function with one exponential tail and one constant tail is used

$$\begin{aligned} &f_{\text{nH-map}}(E;N,\alpha,\beta,\mu,\sigma,\lambda) = \\ &N\alpha f_{\text{peak}}(E;\mu,\sigma) + N(1-\alpha)[\beta f_{\text{exp-tail}}(E;\mu,\sigma,\lambda) + (1-\beta) f_{\text{const-tail}}(E;\mu,\sigma)], \end{aligned} \tag{3.14}$$

where N is the normalization, $\alpha \in (0, 1)$ determines the fraction of events that belong to the peak and $\beta \in (0, 1)$ determines how the rest is distributed between the two tails.

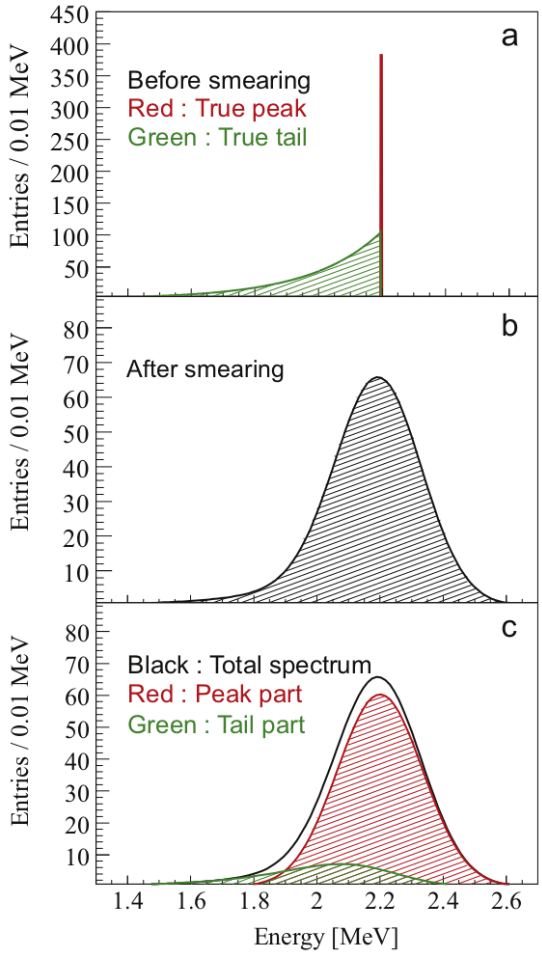

Figure 3.4: The idea behind the Daya Bay function is that (a) before energy smearing the true energy spectrum consists from a peak represented by a delta function and a leakage tail represented by an exponential function (more complicated form is also possible). After Gaussian smearing (b) we get spectrum as it is expected to be measured – we cannot distinguish whether an event belongs to the peak or tail. (c) The decomposition of the contributions of the peak part and the tail part of the measured spectrum is shown – we can see that both parts overlap as there is no sewing like in the CB function. Figure taken from Ref. [56].

Time-dependent Correction

The nonuniormity pattern is not perfectly stable and over time it changed a bit, most likely due to slight degradation of the attenuation length of the liquid scintillator, which means that photons from the center are progressively less likely to reach PMTs and be detected when compared to photons produced near the edge. It was found out that this effect depends only on time t and radius r and

thus the time-dependent correction has a form of

$$f_{rt}(r,t) = 1 - (\beta_0 + \beta_1 r^2)t, \tag{3.15}$$

where the data-driven parameters β_0 and β_1 are common for all ADs and both AdSimple and AdSimpleNL reconstruction methods. In order to construct the correction, each AD was divided into eight slices in r^2 and the data were sampled in half a year periods. The SPN nH peak was fitted in each segment and period so that time-dependent deviation from initial nonuniformity could be quantified as illustrated in Fig. 3.5. Data from all ADs showed that they are rather consistent with each other and fit was made using model described by Eq. 3.15 as shown in Fig. 3.6 determining the values of the parameters $\beta_0 = -1.16 \cdot 10^{-3}\,\mathrm{year}^{-1}$ and $\beta_1 = 7.5 \cdot 10^{-4}\,\mathrm{m}^{-2}\,\mathrm{year}^{-1}$.

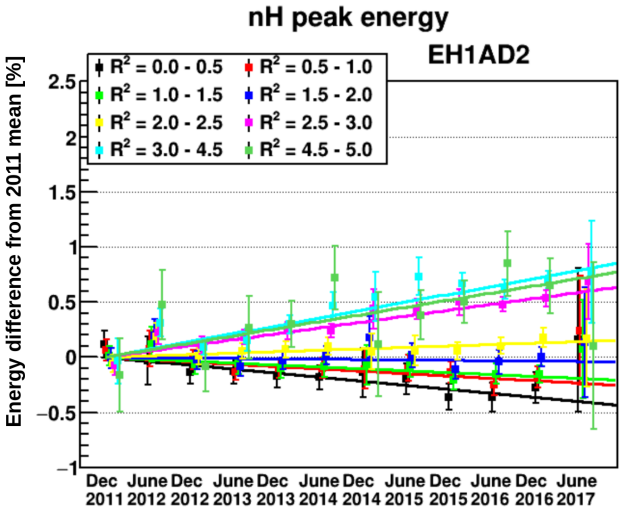

Figure 3.5: An example of the time-dependent non-uniformity obtained by EH1-AD2 (very similar for all ADs) when no time-dependent correction is applied. Each color represents an annular section of the AD covering a range of square of radius R^2 [m^2], most of the sections being 0.5 m^2 wide. We can see that over time (x-axis) the observed energy (y-axis represents relative difference from the initial value) increases for higher radii and decreases for smaller radii [57].

Shortcomings of the Nonuniformity Correction

The nonuniformity correction described above has been successfully employed in a number of analyses (some of them are described in Section 2.5). Most of the analyses measuring neutrinos primarily used nGd and thus were restricted mostly to GdLS, where the nonuniformity correction works very well. However, there is still room for improvement, especially in LS. There are both old phenomena which influence the nonuniformity pattern and which were not fully addressed and new phenomena that require new methods to be dealt with (most notably dead PMTs). Overall, the issues of the nonuniformity correction described above can be summarized as:

- Dead PMTs introduce localized and time-dependent (before vs. after PMT stopped working) decrease of reconstructed energy in their vicinity. The

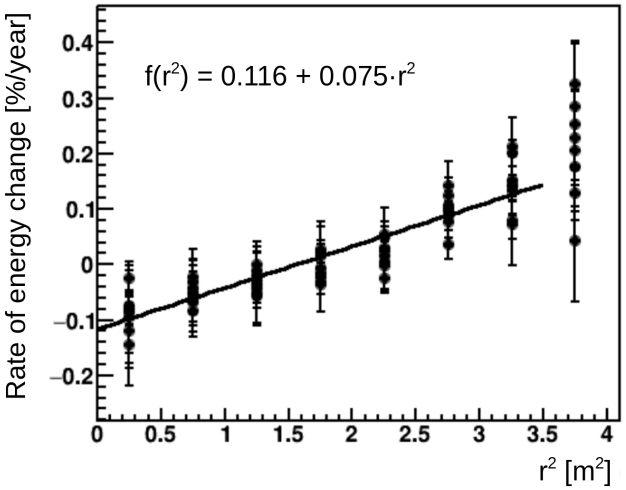

Figure 3.6: Fit of the time-dependent nonuniformity model described by Eq. 3.15 for data from all ADs determining the values of β_0 and β_1 parameters. Each point represents pace of energy change over time (y-axis) for an annular AD section with mean square radius r^2 value indicated on x-axis (essentially slopes of the lines like the ones drawn in Fig. 3.5 are plotted here as a function of r^2) [57].

original correction is not equipped to deal with this phenomenon as the time-dependent part of the correction works only with r-dependence, while strong ϕ- and z-dependence is also present in the effect of dead PMTs. Moreover, the time-dependent correction is common for all ADs, while the dead PMTs are highly AD-specific.

- The azimuthal correction does not fully describe the effect of the residual geomagnetic field as it lacks the r-dependence which is present in the data.

- The fitting of the nGd peak can be improved as it uses double CB function instead of better motivated double DYB function.

- While the r^2-z nonuniformity correction maps used by AdSimpleNL mostly differ by less than 2% AD to AD, there are several outliers which raise suspicion. Most notably the relative energy of EH3-AD4 bottom right voxel shown in the right panel of Fig. 3.7 is between 4.8% – 7.5% higher than in the same voxel in other ADs (which are quite consistent) as well as in neighboring voxels. Analogously, top right voxel of EH3-AD3 in the left panel of Fig. 3.7 shows significantly lower energy. There are more outliers among other ADs, but the two mentioned above are the most significant ones. The cause of these outliers is most likely inaccurate fitting of the nH peak.

Correct and precise energy reconstruction is crucial for neutrino oscillation and other analyses. Generally speaking, most of the issues are less of a problem in GdLS, but grow in significance in LS. It is thus the neutrino oscillation analysis based on nH that is most affected. Our main focus is to improve the correction for this analysis.

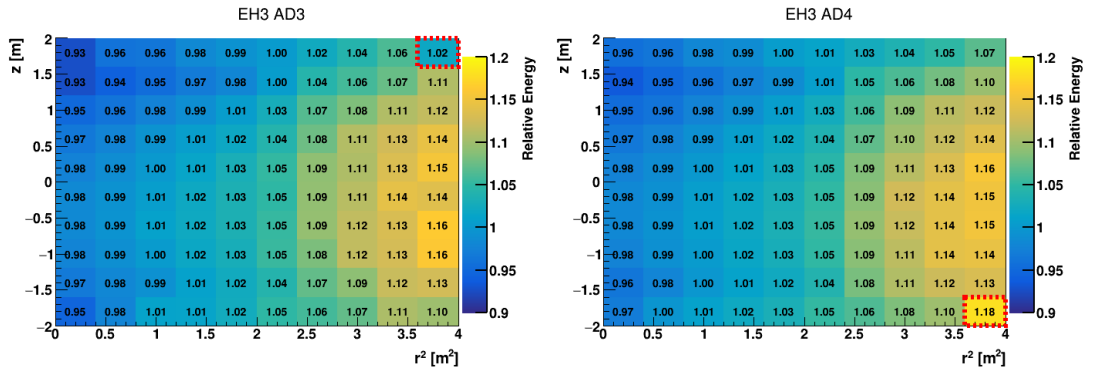

Figure 3.7: The AdSimpleNL nonuniformity map M_{NU} used in EH3-AD3 (left) and EH3-AD4 (right) event reconstruction. Note that the value in the highlighted voxel in EH3-AD3 (EH3-AD4) is suspiciously lower (higher) than in the surrounding voxels (or the same voxel in other ADs – the maximal difference there is about 9% between EH3-AD1 and EH3-AD3).

Especially the dead PMTs have notable impact on the nonuniformity pattern in their vicinity, most of all in EH2-AD2 where two PMTs quite close to each other stopped working around the beginning of 2017. Moreover, the dead PMTs differ from AD to AD, thus potentially introducing a bias between near and far halls if not addressed. Such bias would be most detrimental for the neutrino oscillation measurement as it is based on comparison of measurements from these two groups of ADs. While the problem of the dead PMTs can be up to a certain extend mitigated by not using the data after 2017 as many of the dead PMTs stopped working around that time, it is still very important to actually carefully assess the impact of the dead PMTs and other issues and design new (residual) correction if necessary.

For the reasons stated above, the nonuniformity correction described here needs to be further scrutinized and eventually new updated version needs to be introduced. However, this requires a long procedure starting with selection of suitable signals and devising fitting of their energy spectra as described in Chapter 4 and continues with updated correction construction and implementation described in Chapter 5. But before we get to that, we will first finish the description of the event reconstruction chain.

3.3.5 Nonlinearity Correction

The reconstructed energy scale is calibrated by the nGd gamma signal at about 8 MeV and is thus equal to the deposited energy for that particular signal. However, when we observe different types of particle or the same type with different energy, the reconstructed energy does not correspond perfectly (or even roughly in some cases) to the deposited energy, as there is also nonlinear energy response which needs to be taken into account. It essentially consists of two parts – liquid scintillator nonlinearity (the relation between deposited energy and the amount of scintillation light produced) and electronics nonlinearity (the relation between the number of photoelectrons collected by a PMT and the number of photoelectrons reconstructed).

Electronics Nonlinearity

Unlike liquid scintillator nonlinearity, electronics nonlinearity is not dependent on the particle type. It is caused by the fact that the PMTs and front-end electronics (FEE) that digitize PMT signal have intrinsic nonlinearities. The main source is the tendency of the summing circuit to underestimate the charge collected by a PMT when there are several pulses that do not perfectly overlap. As a result 10% nonlinearity is introduced in the energy region between 0 MeV and 12 MeV. For that reason, detailed study of electronics and subsequently liquid scintillator nonlinearity was carried out as described in Ref. [50].

In order to study the performance of FEE, 192-channel (corresponding to 192 PMTs) flash analog-to-digital converter (FADC) was installed in EH1-AD1 in 2015 and acquired data simultaneously with the FEE readout system. FADC digitized the raw PMT waveforms at 1 GHz and thus could determine the charge without the aforementioned problems of FEE. By comparing the outputs of FADC and FEE, appropriate correction factor (f^*_{SCNL} in Eq. 3.2) was determined. It is common for all PMTs, but applied for each PMT signal individually. The ratio of charge obtained with FEE and FADC ($1/f^*_{\mathrm{SCNL}}$) is drawn in Fig. 3.8. The corresponding factor f^*_{SCNL} is used in AdSimpleNL energy reconstruction described by Eq. 3.2, but not in the older AdSimple method.

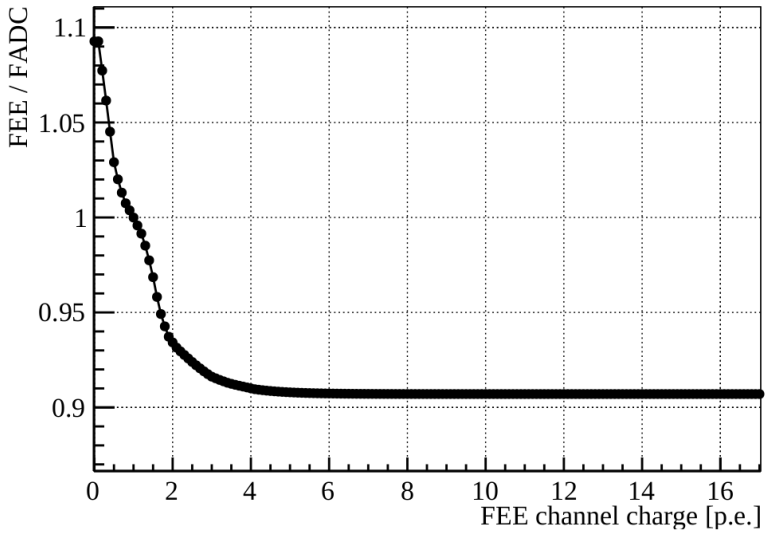

Figure 3.8: The ratio between charge obtained using front-end electronics (FEE) and flash analog-to-digital converter (FADC), inverse of which is used for single channel nonlinearity correction [50]. As we can see, when more photoelectrons (p.e.) are registered by a PMT, FEE tends to underestimate the charge compared to FADC.

Liquid Scintillator Nonlinearity

The amount of light produced in liquid scintillator is in general not directly proportional to the deposited energy. This nonlinearity is mainly caused by two processes – *quenching* and *Cherenkov light* production.

If speed (β) of a charged particle exceeds the speed of light in the medium ($1/n$, where n is the refractive index), Cherenkov light is produced. For an electron or positron the threshold for Cherenkov light production in liquid scintillator is about 0.2 MeV of kinetic energy. As a result, particles with higher energy above the threshold produce Cherenkov light in addition to the scintillation, while the particles below the threshold do not, thus adding to the liquid scintillator nonlinear energy response.

The other and bigger reason for the liquid scintillator nonlinearity are quenching interactions between excited molecules along the path of the incident particle, i.e., interactions that drain the energy which would otherwise go into luminescence. The higher the energy deposition density dE/dx caused by a particle is, the higher the density of excited molecules is and it is thus more likely their excitation energy is drained by quenching instead of production of scintillation light. This phenomenon can be described by a semi-empirical Birks' formula [58] giving us the relation between deposited energy E_{dep} and energy that was actually converted to scintillation photons E_{SC}

$$E_{\text{SC}}(E_{\text{dep}}) = \int_0^{E_{\text{dep}}} \frac{dE}{1 + k_{\text{B}} \frac{dE}{dx}}, \tag{3.16}$$

where k_{B} is the Birks' coefficient.

The effect of quenching is especially impactful when it comes to alpha particles. As mentioned before, the light yield is calibrated using SPN nGd signal. The energy of the gamma cascade is eventually transferred to electrons whose ionizing losses cause emission of scintillation light. It thus follows that for electrons, positrons and gammas the reconstructed energy at least roughly corresponds to the deposited energy. On the other hand, alpha particles of similar energies move much slower, have much higher energy deposition density and thus are much more affected by quenching. As a result, the reconstructed energy of an alpha particle without nonlinearity correction is much lower that the kinetic energy it lost in the liquid scintillator (for example alpha from ^{214}Po decay has kinetic energy of 7.7 MeV [59], but reconstructed energy is only 0.9 MeV).

Once the electronics nonlinearity was measured, liquid scintillator nonlinearity was also decoupled from the overall nonlinearity and determined for gammas, electrons and positron. The nonlinearity curve for gammas is shown in Fig. 3.9, the curve for electrons is qualitatively and to a certain extent quantitatively similar.

Nonlinearity Correction

When we want to deduce the energy deposited by a particle from the reconstructed energy, we need to apply the nonlinearity correction. In AdSimple, there is just one correction that covers both liquid scintillator and electronics nonlinearity. It is common for all ADs, but depends on the particle type.

On the contrary, the electronics nonlinearity in AdSimpleNL is already taken care of by the single channel nonlinearity correction mentioned above. However, liquid scintillator nonlinearity correction still needs to be applied in a similar fashion to AdSimple (common for all ADs, different for different particle types) if we want to deduce the deposited energy.

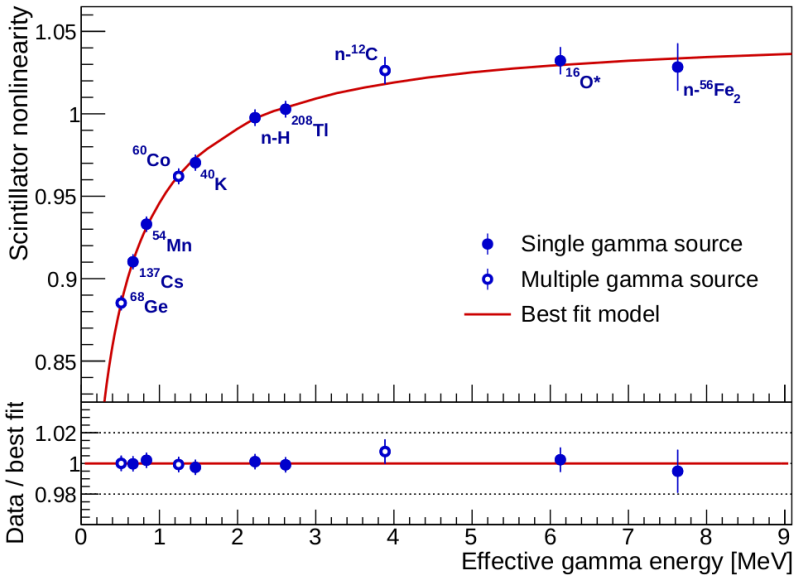

Figure 3.9: Liquid scintillator nonlinearity curve (ratio between energy than went to light production and deposited energy normalized to 1 at 2.2 MeV) for gamma rays [50]. Gamma sources that were used for the measured are marked by blue circles.

While it is crucial when for example determining the reactor antineutrino spectrum, there are other cases where it does not really play a role and reconstructed energy is completely sufficient to work with. For example, in the study of detector nonuniformity, we are interested only in the relative variation of the reconstructed energy as a function of position in the AD and not the absolute energy. For that reason, in the Chapters 4 and 5 we will not apply the nonlinearity correction as it is not needed for our purposes. We will thus work with the reconstructed energy, which is close to the deposited energy for gamma and beta signal, but much lower for alpha signal.

4. Signals for Energy Response Study

There are several signals which can be used for the study of detector nonuniformity and other aspects of the energy response. As mentioned before, SPN nH and nGd provide valuable gamma energy lines. Other promising options are several prominent correlated decays coming from natural radioactivity because the temporal and spatial correlation of prompt and delayed signal allows for great suppression of background as with IBDs. There are three decay chains occurring in nature – Actinium (Ac), Uranium (U) and Thorium (Th) series, however in each of these series there is only one useful pair of correlated decays:

- Th-series: ^{212}Bi $\rightarrow$ ^{212}Po $\rightarrow$ ^{208}Pb correlated beta-alpha decays.
- U-series: ^{214}Bi $\rightarrow$ ^{214}Po $\rightarrow$ ^{210}Pb correlated beta-alpha decays.
- Ac-series: ^{219}Rn $\rightarrow$ ^{215}Po $\rightarrow$ ^{211}Pb correlated alpha-alpha decays.

There are naturally other decays in these decay chains, but corresponding half-lives are too long for efficient coincidence measurement by the Daya Bay experiment. The only suitable correlated pairs (listed above) are centered around short-lived isotopes of polonium – ^{212}Po, ^{214}Po and ^{215}Po. However, in order to reduce confusion, we will refer to each of these correlated decays by the series it belongs to (Ac-series, U-series, Th-series) rather than by the polonium nucleon number.

Ambient natural radioactivity gamma sources like ^{40}K and ^{208}Tl can be also used for certain purposes, but they are less useful for nonuniformity study as they cannot be fully isolated from background.

4.1 Natural Radioactivity Correlated Decays

In the past, correlated decays coming from natural radioactivity were employed to check the performance of the original nonuniformity correction [60, 61]. In contrast to that, our aim is to use these correlated decays not only to scrutinize the original correction, but also to construct a new updated version. We thus need to study their properties in great detail to properly assess which correlated decays are suitable for our purposes and how they can be used to extract the energy scale without introducing any unwanted biases.

In order to study the correlated decays, we must first perform a dedicated data selection. As the amount of unfiltered data is enormous, certain selection cuts must be applied in order to extract a sample of the correlated decays.

4.1.1 Loose Selection

For purpose of this study, full data set covering experiment operation between 2011 and 2020 was used with AdSimpleNL reconstruction method. List of the

Table 4.1: Loose selection cuts for correlated decays.

Coincidence time	(1 μs, 2000 μs)
Prompt energy	(0.5 MeV, 15 MeV)
Delayed energy	(0.5 MeV, 15 MeV)
Flasher cuts	$(\text{Quadrant})^2 + (\frac{Q_{\max}}{0.45})^2 > 1$ $Q_{\text{2-inch}} > 100$ photoelectrons
IWP, OWP muon veto (nHit > 12)	600 μs
AD muon veto (energy > 20 MeV)	1000 μs
Shower muon veto (energy > 2500 MeV)	1 s
Multiplicity cut	No other signal > 0.5 MeV or muon in (-400 μs, 2400 μs) around prompt signal

loose (common for all series) selection criteria is summarized in Tab. 4.1, however, the meaning of individual cuts and the logic behind them also needs to be discussed.

- **Coincidence time (time difference between delayed and prompt signals):** The goal of the selection is to isolate a sample of pairs of correlated decays. Such a pair consists of prompt signal followed by a delayed one with mean time difference (mean lifetime) of about 0.43 μs for Th-series [62], 236 μs for U-series [59] and 2569 μs for Ac-series [63]. As our main focus will be on the U-series, coincidence interval ("window") was set to (1 μs, 2000 μs), which needs to cover enough statistics for both desired signal and background, which would be later subtracted. The coincidence window is not optimal for Ac-series and Th-series, but they suffer from other problems which hinder their usefulness for nonuniformity study anyway as we will see in Sections 4.2 and 4.3. The lower limit of 1 μs is a result of detector dead time needed to read out the prompt event.

- **Prompt & delayed energy:** The lower bound on energy for events that we save in our selection was set to 0.5 MeV as it is close to the detection threshold (the detection efficiency reaches 100% at approximately 0.7 MeV). The reconstructed energy of beta and alpha decays we are interested in goes up to 3.5 MeV[1]. However, since the background rate is very low for energies > 3.5 MeV, expanding the selection up to 15 MeV does not increase the statistics much and also allows for study of IBD nGd.

- **Flasher cuts:** Apart from physical events associated with ionizing particles causing production of scintillation light there are also so-called *flashers* which are a result of a spontaneous light emission by PMTs caused by sparks from electrostatic discharges (the high operating voltage makes PMTs prone to electrostatic breakdown). Appropriate cuts were applied in order to remove the flashers following the procedure described in Ref. [29].

[1]The energy deposited by alpha particles in these decays is higher, but the amount of scintillation light is reduced due to liquid scintillator nonlinearity as described in Section 3.3.5.

- **Muon veto:** Correlated decays that we are interested in come from natural radioactivity in the ADs. However, there are also correlated signals that follow after muon passes through the detector system, such as multiple spallation neutron captures or decays of cosmogenic isotopes. As these would pose contamination in our sample, their presence is minimized by excluding the data which follow in a certain interval after muon was detected – 600 μs if more than 12 PMTs were hit in either section of a water pool, 1000 μs if an event (presumably muon) deposited between 20 MeV and 2500 MeV in the scintillating volume and 1 s if the deposited energy was over 2500 MeV in which case muon-induced shower likely happened.

- **Multiplicity cut:** It is required that there is no signal > 0.5 MeV other than one prompt and one delayed in the coincidence window or 400 μs before or after it in order to avoid multiple counting of individual signals and to make the prompt-delayed selection unambiguous.

The loose selection allows us to overview some properties of the correlated decays and design fine selection cuts tailored for each series separately. However, most of the principles and procedures we will use are common for all series and differ only in values of their parameters. We shall therefore make a general overview first before delving into series-specific details.

4.1.2 Background Subtraction

Even though the loose selection aims to pick out correlated decays, lot of unwanted signals pass the selection too, which then needs to be dealt with. The most prominent group are accidental coincidences of signals from ambient radioactivity. As they are not really correlated and just happen to occur in close temporal (and spatial) proximity, the coincidence time distribution of these accidentals is flat. On the contrary, the distribution of coincidence times for correlated decays is given by an exponential function driven by corresponding polonium mean lifetime. This fact can be utilized for background subtraction.

The method we will use for background subtraction utilizes signal and background windows, where each window corresponds to a certain coincidence time interval. The simplest configuration is one signal window covering interval (t_1, t_2) and one background window (t_3, t_4). Assuming that the lengths of the windows are the same, i.e.

$$t_2 - t_1 = t_4 - t_3, \tag{4.1}$$

we can get a clean sample of correlated signals without uncorrelated background by simply subtracting the data in the background window from the data in the signal window. As the uncorrelated signal has uniform coincidence time distribution, the numbers of events in the signal and background windows are in principle the same and thus they get subtracted completely. On the contrary, correlated signals tend to concentrate at lower coincidence times (presumably in the signal window) and thus part of them remains after the subtraction. This procedure is outlined in Fig. 4.1 along with comparison of coincidence time distributions of all three series.

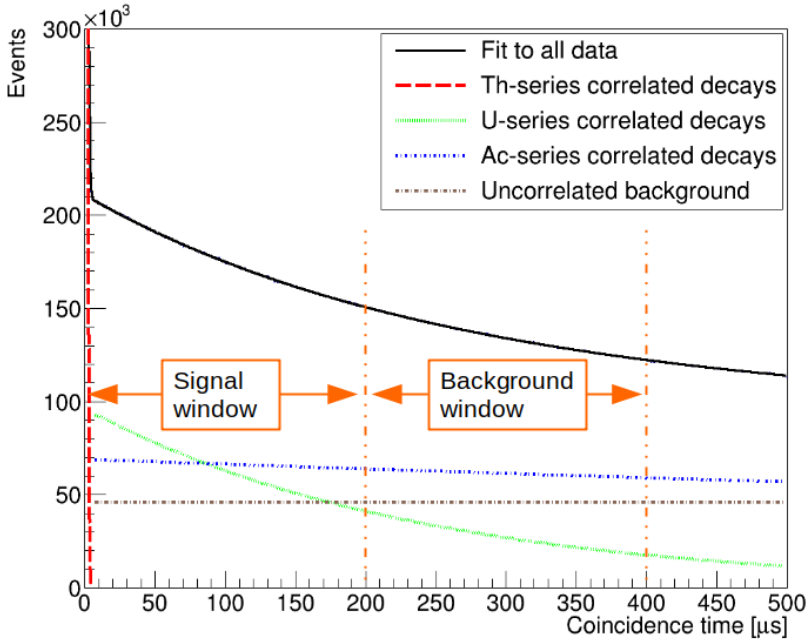

Figure 4.1: Coincidence time distribution obtained with loose selection for correlated decays and additional prompt-delayed distance cut < 750 mm and delayed energy cut (0.7 MeV, 1.4 MeV). Data from all 8 ADs are included. Contribution from individual series based on fit to the data (black line) are shown – Th-series (red), U-series (green), Ac-series (blue) along with contribution of uncorrelated background (brown). One possible configuration of signal and background windows is marked out. Even though it leads to background subtraction, it is not really optimized for any particular one of the correlated decays – signal window is too long for Th-series and too short for U-series and Ac-series.

If the lengths of signal and background windows are different, appropriate weighting factor w_{bkg} for events in background window can be used

$$w_{\mathrm{bkg}} = \frac{t_2 - t_1}{t_4 - t_3}. \tag{4.2}$$

While the length of the signal window is usually chosen in proportion to the corresponding half-life, the background window can be longer in order to reduce statistical fluctuations.

In some cases, we would like to subtract not only uncorrelated background, but another correlated signal with mean lifetime τ_{C} too. For example when we want to study U-series, we want to subtract Ac-series events along with the uncorrelated background. Obviously, this cannot be done with only one signal and one background window – one more needs to be added. Let us first assume that we are interested in correlated signal with mean lifetime τ_{S} which is shorter than the one we would like to subtract ($\tau_{\mathrm{S}} < \tau_{\mathrm{C}}$). Such situation can be solved by adding one more signal window after the background window as illustrated in Fig. 4.2. This way we can balance the number of correlated background events in the signal and background windows as the first signal window has the highest rate of these

events and the second signal window the lowest one while the background window is somewhere in between. Depending on their lengths, the windows are assigned weighting factors. The calculation of these factors is described in Appendix A.

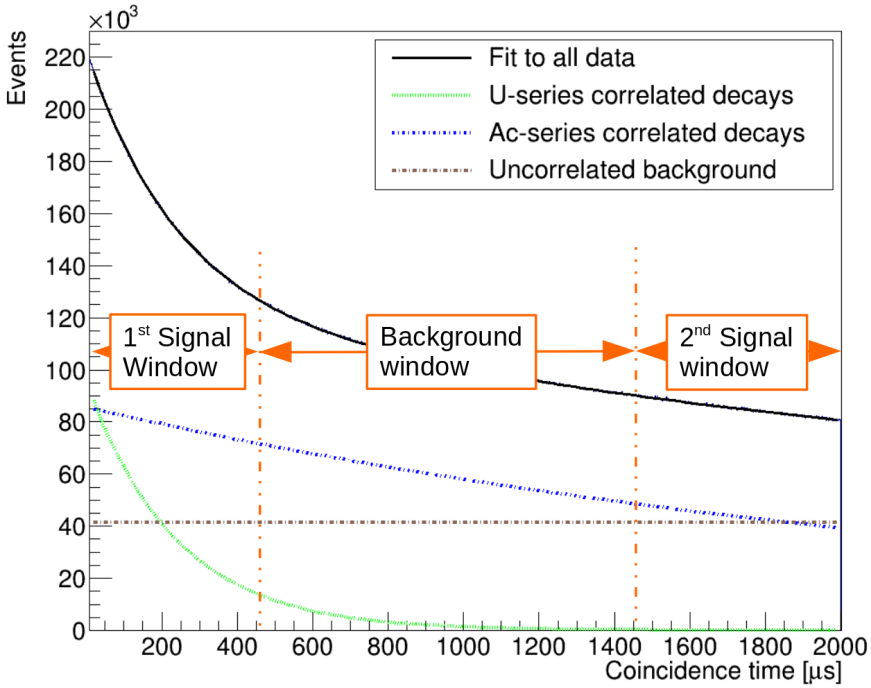

Figure 4.2: Illustration of the background subtraction method using two signal windows and one background window on top of coincidence time distribution with cuts aimed to highlight U-series and Ac-series. The windows are set in the way that weighting factors are all equal (= 1) and their configuration allows us to subtract uncorrelated signal as well one correlated signal (Ac-series in this case). Most of U-series signal is, on the other hand, concentrated in the first signal window and only small fraction in the background window.

Analogous procedure can be applied if we want to subtract a correlated signal with shorter mean lifetime ($\tau_S > \tau_C$), the only difference is that we need to use different configuration of coincidence time windows: background – signal – background. The calculation of corresponding weighting factors is again included in Appendix A.

4.1.3 Quantities of Interest

Before delving into the individual correlated decays, it will be useful to go through the most relevant quantities that we will study as some of them are quite similar for all of the correlated decays and some, on the contrary, can be used to differentiate between them and further reduce the background that needs to be subtracted.

Prompt-Delayed Distance

The distance between prompt and delayed signal is a parameter that allows us to greatly suppress uncorrelated background. This is outlined in Fig. 4.3 showing decomposition of distance distribution obtained with the loose selection into correlated and uncorrelated signals up to 1500 mm of distance. At such distance, there are virtually no correlated signals while the number of uncorrelated signals continues to rise steadily.

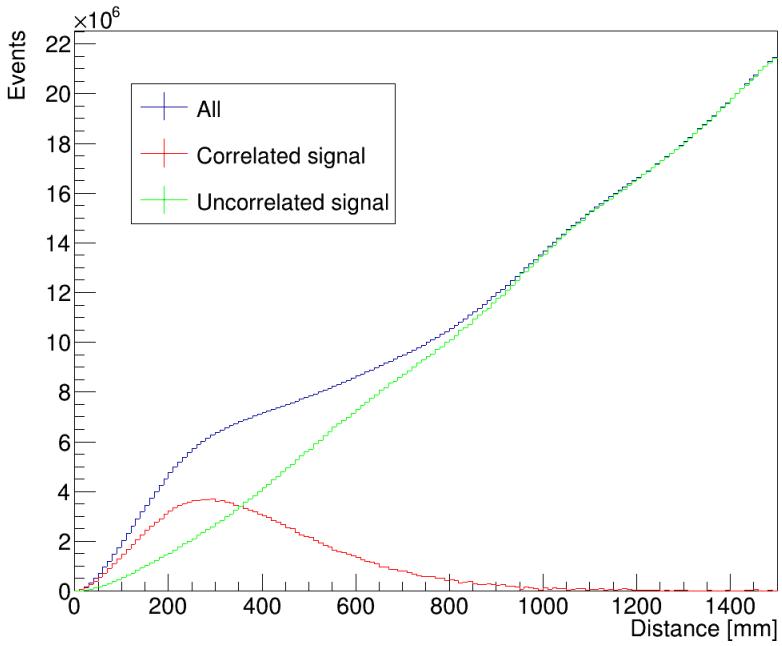

Figure 4.3: Decomposition of the distance distribution obtained with the loose selection into correlated and uncorrelated signals. Coincidence time distribution of each distance bin was analyzed to obtain the correlated and uncorrelated fraction (similar to what is described in the Appendix B). Data from all ADs were aggregated for that purpose.

The distance distributions of the three series have similar shapes as shown in Fig. 4.4. Common value of distance cut was used in the fine selection for all series – events with distance < 750 mm were selected.

The differences between the series are not necessarily caused by actually different physical distances between prompt and delayed events, but rather by the position reconstruction precision, which depends on the energy of the event and the part of AD it took place in.

Prompt and Delayed Energy

Imposing further cuts on prompt and/or delayed energy is another way to reduce both uncorrelated and correlated background. Overview of the situation is depicted in Fig. 4.5, where prompt vs. delayed energy plot is shown with contributions from individual sources marked.

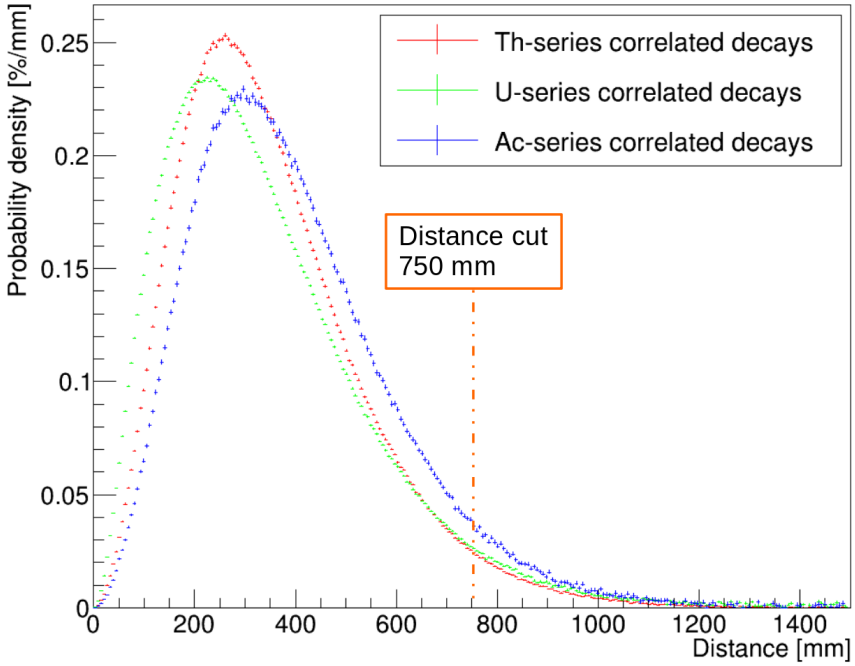

Figure 4.4: Comparison of normalized distance distributions of correlated decays of Th-, U- and Ac-series obtained using fine selection cuts (except distance cut) described in sections dedicated to the particular series. The distance cut at 750 mm common for all series is indicated by orange dash-dottedline.

Out from the various signals (prompt or delayed) at hand, we are mostly interested in the alpha signals for two main reasons. First, the alpha particle deposits energy over a very short track which means there is likely no energy leakage (part of the energy deposited outside of the sensitive volume). Second, if there is only one energy line, we get an easily fitted Gaussian peak with the width corresponding to the energy resolution. This is mostly the case with delayed signals of all series, while the prompt signals are caused by beta decays for U-series and Th-series and alpha for Ac-series. However, all of the prompt signals have a significant admixture of gammas.

The delayed signal of Th-series is quite separated from others at around 1.2 MeV, but U-series and Ac-series delayed energies are very similar around 0.9 MeV. Thus, special care must be taken when disentangling these two. One can use the fact that Ac-series prompt signal does not exceed 1.3 MeV while U-series prompt peak-like structure is positioned around 2.2 MeV.

Pseudorate

While it is possible to calculate the actual rate of signals caused by a particular series, it is quite complicated procedure, which requires a lot of variables to be taken into account (various efficiencies related to the cuts etc.). Moreover, for most purposes we do not need to know the absolute rate, but just whether it is stable or if it changes over time. For that reason, the concept of a pseudorate was introduced.

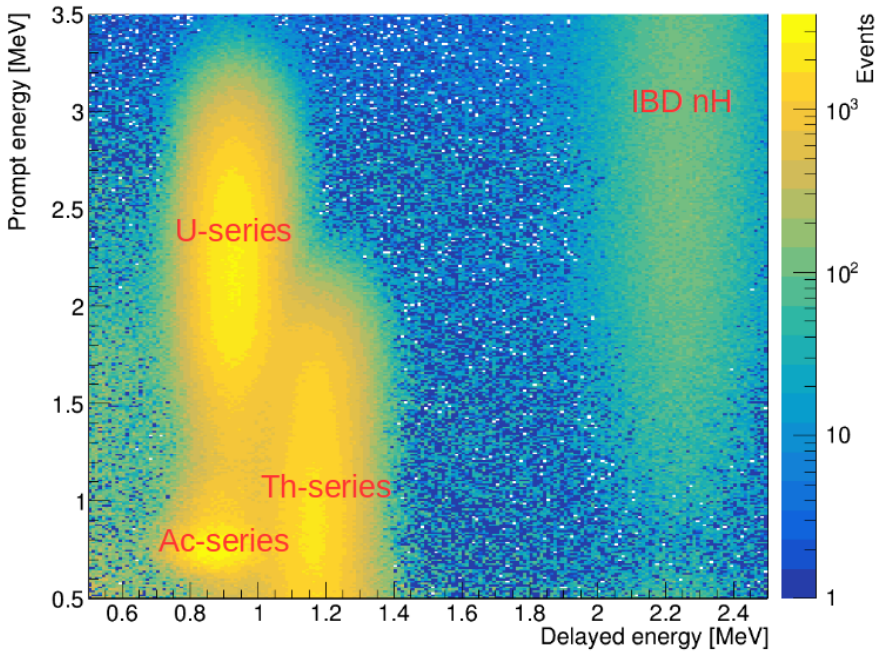

Figure 4.5: The plot of prompt versus delayed energy for pairs of correlated decays obtained with loose selection, signal window (1 μs, 200 μs), background window (201 μ, 400 μs) and distance cut 750 mm for data from all 8 ADs. Contributions from individual series are marked as well as from IBD nH.

Even though the point of pseudorate is to provide a reasonably well obtainable variable directly proportional do the actual rate, the procedure of its calculation is still quite complicated and can be found in Appendix B. The result is indeed a variable which copies the changes of the actual rate over time, which provides us with valuable insights to the properties of the studied correlated decays.

Spatial Distribution

Spatial distribution is an important characteristic of each type of signal. When studying the detector energy response, we need to know whether our observation is related to the whole volume of AD or just part of it. Similarly, we can only study nonuniformity using a particular signal in the part of the AD where that signal is present.

Spatial distribution is usually obtained by dividing the AD to a number of voxels and counting events that belong to the desired series using fine cuts and background subtraction as described above. The obtained values can be used to create a spatial distribution map and assess whether the statistics in each voxel is sufficient for further analysis. The values in the voxels can also be normalized using pseudorate as described in the previous subsection. It is especially useful if we want to study the time stability of spatial distribution.

4.1.4 Fine Selection Cuts

In the following sections, we will look into the individual correlated decays, their features and potential usefulness for nonuniformity and other studies. For that purpose, we will use fine selection cuts tailored for each series individually. These cuts include signal and background windows, distance, prompt and delayed energy cuts. However, it should be noted that not all cuts will be used for all purposes. For example, signal and background windows are irrelevant when studying coincidence time distribution. Calculation of pseudorate uses only one coincidence time window which may or may not coincide with the signal window. Prompt energy cut is not applied when prompt energy spectrum itself is studied and the same applies for delayed energy. On the other hand, study of spatial distribution usually utilizes all the fine selection cuts available.

4.2 Th-Series

When talking about correlated decays from Th-series, we have in mind correlated beta–alpha decays ^{212}Bi $\rightarrow$ ^{212}Po $\rightarrow$ ^{208}Pb. The correlation time is given by ^{212}Po mean lifetime 0.431 μs.

Table 4.2: Fine selection cuts and coefficients for Th-series.

1$^{\text{st}}$ signal window	(1 μs, 3 μs)
Coefficient	1
Background window	(3 μs, 10 μs)
Coefficient (w_{B})	0.5
2$^{\text{nd}}$ signal window	(10 μs, 15 μs)
Coefficient (w_{S})	0.3
Maximal distance	750 mm
Prompt energy	(0.5 MeV, 2.2 MeV)
Delayed energy	(0.95 MeV, 1.4 MeV)

4.2.1 Signal and Background Windows

Even if we start our signal window at the lowest available limit of 1 μs, we still loose over 90% of statistics due to short ^{212}Po mean lifetime. On the other hand, we can make the signal window also very short and thus reduce both correlated and uncorraleted backgrounds. In the fine selection, we use the first signal window (1 μs, 3 μs), which covers 99.9% of Th-series signals that survive over 1 μs, the background window of (3 μs, 10 μs) and the second signal window (10 μs, 15 μs), which allows for complete removal of U-series contamination and lowering the Ac-series contamination to 3×10^{-7}% of the original level. Full list of fine selection parameters is in Tab. 4.2 including the weighting factors for the signal and background windows. The procedure of their calculation is described in Appendix A.

4.2.2 Prompt and Delayed Energy

The prompt signal is caused by beta decay ^{212}Bi $\rightarrow^{212}$ Po with Q-value 2251.5 keV, which occurs in 64% of all ^{212}Bi decays [64], the remaining 36% are alpha decays ^{212}Bi $\rightarrow^{208}$ Tl not interesting for us. Out from the beta decays, 86% lead directly to the ground state of ^{212}Po, the rest lead to excited states followed by gamma(s) emission. The prompt energy spectrum obtained with cuts listed in Tab. 4.2 is shown in the left panel Fig. 4.6. It was not used for energy scale study and therefore no appropriate fitting function was devised.

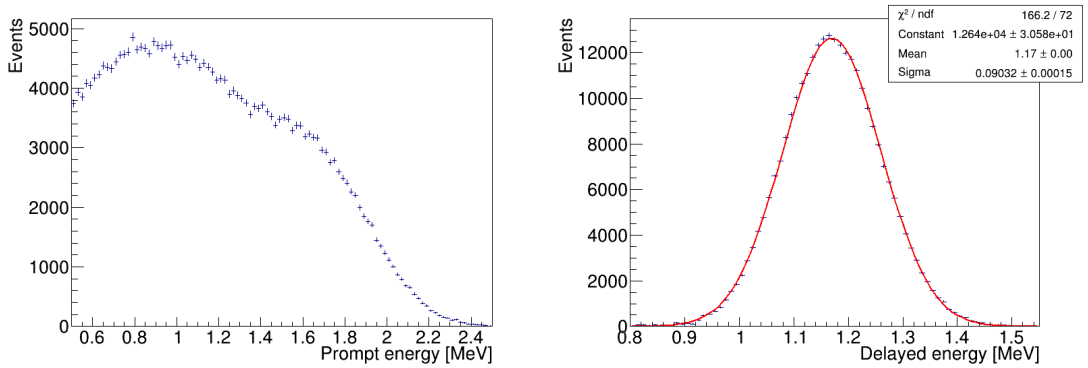

Figure 4.6: Th-series prompt (left) and delayed (right) energy spectra obtained by EH1-AD1 using cuts listed in Tab. 4.2. The prompt spectrum is only minimally affected by energy leakage as majority of the ^{212}Bi beta decays are not accompanied by gamma emission and even if gamma is emitted, it is unlikely to deposit energy outside of the sensitive volume as the Th-series is present only in GdLS. The delayed energy spectrum is fitted by a simple Gaussian function (red line).

The delayed signal of Th-series is caused by ^{212}Po alpha decay with Q-value 8954.1 keV [62], but the reconstructed energy is only about 1.2 MeV due to the liquid scintillator quenching. As there is only one energy line with no accompanying gammas, the signal is very well described by a Gaussian function with width given by energy resolution, as it is illustrated in the right panel of Fig. 4.6. This makes it suitable for energy response studies.

4.2.3 Pseudorate

Pseudorate of Th-series was studied using fine cuts listed in Tab. 4.2, except for the fact that only one coincidence time window is used in pseudorate determination. For Th-series, it is identical to the first signal window (1 μs, 3 μs). As it is shown in Fig. 4.7, the pseudorate is consistent among ADs in the same hall, but differs between the halls due to the multiplicity cut and muon veto efficiencies. The actual rate is expected to be the same for all ADs as they share the same GdLS mixture.

The major finding is that the pseudorate significantly decreased to about a third of the initial value over the 9 years of the Daya Bay operation. Explanation of this phenomenon can be found in the upper part of the Th decay chain assuming

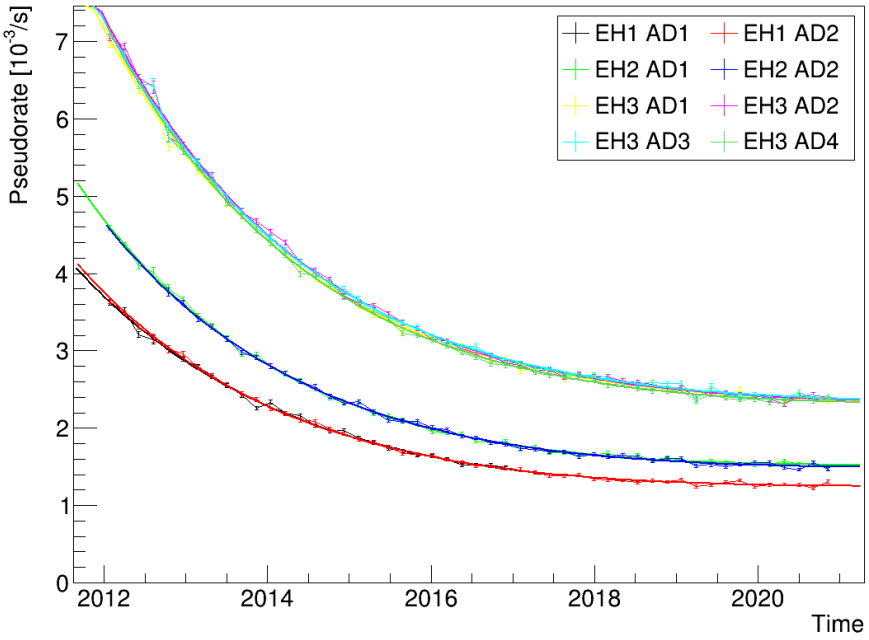

Figure 4.7: Th-series pseudorate of all ADs fitted by the function described by Eq. 4.3. During nine years of Daya Bay operation, the pseudorate decreased to about a third of the initial value.

that it is not necessarily in equilibrium. If we take into the account the fact that we study pseudorate across the span of months and years, there are only 3 decays which cannot be treated as instant:

1. ^{232}Th decay with mean lifetime 20×10^9 years [65], which is long enough for us to consider its activity as constant over the Daya Bay operation.
2. ^{228}Ra decay with mean lifetime 8.3 years [65].
3. ^{228}Th decay with mean lifetime 2.76 years [66]. This is the last "long-lived" isotope in the Th decay chain before the ^{212}Bi $\rightarrow$ ^{212}Po $\rightarrow$ ^{208}Pb correlated decays, its activity thus should correspond to the pseudorate we observed.

If we consider that the initial activities of these isotopes as free parameters, we get following formula for ^{228}Th activity $A_3(t)$ (which should be equal to the ^{212}Po activity and directly proportional to the pseudorate) as a function of time t that passed since the beginning of the Daya Bay data taking

$$A_3(t) = A_{1,0} + (A_{3,0} - A_{1,0})\, e^{-\frac{t}{\tau_3}} + \frac{\tau_2}{\tau_2 - \tau_3}(A_{2,0} - A_{1,0})(e^{-\frac{t}{\tau_2}} - e^{-\frac{t}{\tau_3}}), \qquad (4.3)$$

where $A_{1,0}$, $A_{2,0}$, $A_{3,0}$ stand for initial ($t = 0$) activities of ^{232}Th, ^{228}Ra and ^{228}Th respectively and τ_1, τ_2 and τ_3 are corresponding mean lifetimes. We can use Eq. 4.3 to fit the observed pseudorate and we get a good match to the data when the ^{232}Th and ^{228}Ra initial activities are set to $\sim 36\%$ and $\sim 23\%$ of the initial ^{228}Th activity respectively. The reason why ^{228}Th is initially disproportionately

prevalent compared to the other isotopes is likely linked to the the processes of Gd production and Gd doping preparation as Th-series is present only in GdLS (see Section 4.2.4 for more details).

The fact that Th-series pseudorate changes so much over time does not disqualify it from use for energy response studies, however one needs to bear in mind that the beginning of Daya Bay operation will be overrepresented in the whole data set if no correction is made.

4.2.4 Spatial Distribution

The spatial distribution of Th-series was obtained using fine selection cuts listed in Tab. 4.2 and then normalized to the observed pseudorate. Two sets of spatial distribution maps were created – one shows the AD in r^2-z coordinates and the other one in x-y coordinates as illustrated in Fig. 4.8. It was found out that Th-series contamination is present only in GdLS and is likely linked to the Gd doping as the Gd compound used for that purpose contains traces of other elements, which end up in the scintillator as well. This means that Th-series has only limited use for energy response study despite the fact that its short mean lifetime allows for very efficient background subtraction and the delayed energy peak can be easily fitted.

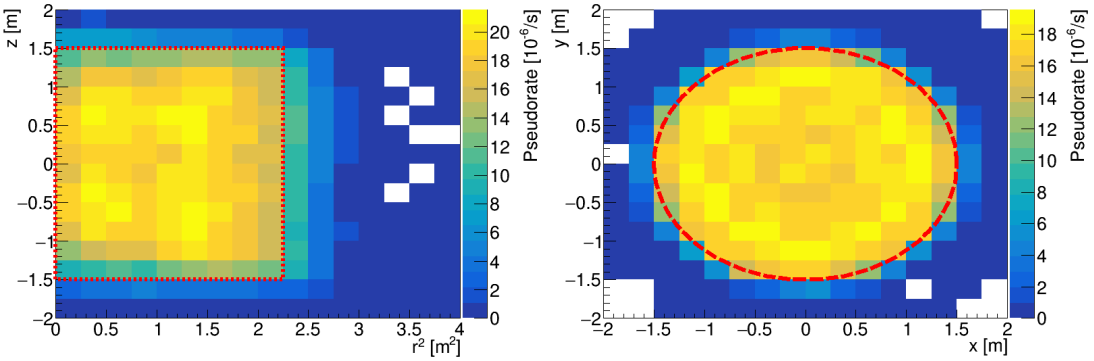

Figure 4.8: Examples of Th-series spatial distribution in r^2-z (left) and x-y (right) coordinates for EH1-AD1. Th-series contamination is present only in GdLS ($|z| < 1.5$ m, $r^2 = x^2 + y^2 < 2.25$ m^2 – the boundary of GdLS is marked by red dashed line), however some events are reconstructed just outside due to position reconstruction resolution. All ADs follow the same pattern.

4.3 Ac-Series

By Ac-series we mean correlated alpha–alpha decays ^{219}Rn $\rightarrow$ ^{215}Po $\rightarrow$ ^{211}Pb with coincidence time given by ^{215}Po mean lifetime 2569 μs.

4.3.1 Signal and Background Windows

As mentioned before, coincidence time window (1 μs, 2000 μs) is not optimal for Ac-series selection as it is even shorter than its mean lifetime. However, it can

Table 4.3: Fine selection cuts and coefficients for Ac-series.

1st background window	(5 μs, 165.2 μs)
Coefficient (w_{B1})	1
Signal window	(165.2 μs, 1162.8 μs)
Coefficient	1
2nd background window	(1162.8 μs, 2000 μs)
Coefficient (w_{B2})	1
Maximal distance	750 mm
Prompt energy	(0.65 MeV, 1.1 MeV)
Delayed energy	(0.7 MeV, 1.05 MeV)

still be used to a certain extent with appropriately chosen signal and background windows.

In order to avoid Th-series contamination, we do not use any pairs with coincidence time < 5 μs (only 10^{-3}% of Th-series decays pass such selection). As both prompt and delayed energy spectra of Ac-series are overlapping with U-series, we cannot use prompt or delayed energy cuts to disentangle Ac-series from U-series (though it is possible the other way round), hence we need to set the signal and background windows to do the job. The procedure described in Appendix A was used to obtain the boundaries of signal and background windows and corresponding weighting factors, which are listed in Tab. 4.3 along with other fine selection parameters. Unfortunately, while removing both Th-series and U-series contamination, the coincidence time selection preserves only 6.4 % of Ac-series correlated decays.

4.3.2 Prompt and Delayed Energy

The prompt signal is caused by ^{219}Rn alpha decay with Q-value 6946.1 keV [67]. Most of the decays (79.4%) go directly to the ground state with no accompanying gamma emission. The peak mean is then only $\sim$0.75 MeV due to liquid scintillator quenching. Rest of the decays lead to the excited states of ^{215}Po. In 12.9% of cases 271.2 keV of gamma energy is released and in 7.5% of cases it is 401.8 keV. There are many more energy lines, but their contributions are negligible.

Since signal caused by a gamma is much less quenched by liquid scintillator than the one caused by an alpha particle, we get a combination of three peaks in the reconstructed spectrum, however the two smaller ones are not distinguishable one from the other as shown in the left panel of Fig. 4.9. Though not optimal, position of the main peak can be in principle used for an energy response study. For that purpose, the prompt energy spectrum can be fitted by double Gaussian function – first Gaussian function for the main peak and the second one for the two minor peaks together. As they are mostly merged, the second Gaussian function covers them reasonably well. If we tried to fit them individually (resulting in triple Gaussian function for the prompt energy spectrum), we would have too many free parameters which could make the fitting unstable. More refined fitting function could be devised, but double Gaussian function was considered sufficient given the limited use of this signal for further studies.

The delayed signal is caused by ^{215}Po alpha decay with Q-value 7526.3 keV [63]

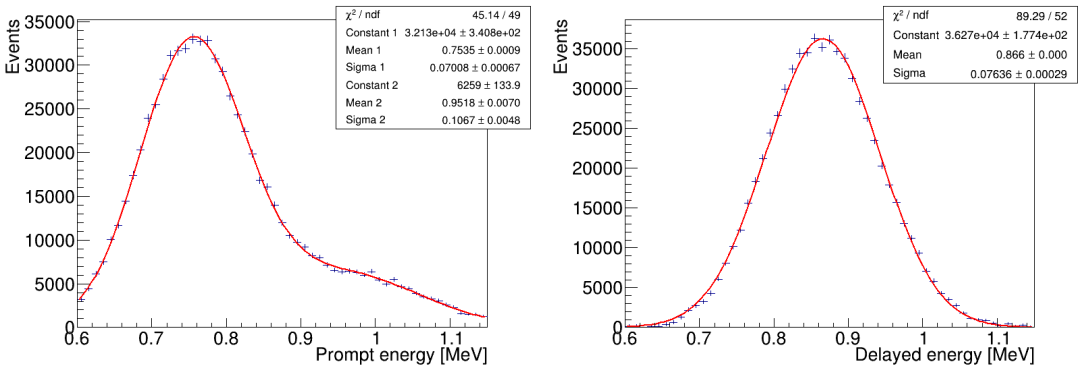

Figure 4.9: Ac-series prompt (left) and delayed (right) energy spectra obtained by EH1-AD1 using cuts listed in Tab. 4.3. The prompt consists of several peaks – the main peak at ∼0.75 MeV corresponds to pure alpha decay, two minor peaks constituting a bump at ∼1.0 MeV are caused by less quenched gammas together with the alpha decay. Double Gaussian function was used to fit the spectrum with index 1 stands for the main peak and index 2 for the minor peaks. The delayed spectrum is fitted by a simple Gaussian function.

leading to the reconstructed energy ∼ 0.9 MeV due to the liquid scintillator quenching. Decays to excited states of ^{211}Pb are possible, but we can neglect them due to their low probability (< 0.1% in total). As a result, the delayed energy spectrum can be very well described by a Gaussian function as shown in the right panel of Fig. 4.9 and is thus in principle suitable for energy response studies.

4.3.3 Pseudorate

Ac-series pseudorate was obtained using cuts listed in Tab. 4.3 and one coincidence time window set to (5 μs, 2000 μs). The result is shown in Fig. 4.10. As for Th-series, the pseudorate is consistent between ADs that share the same hall and we can again observe its decrease, but this time not so significant (about a quarter in 9 years).

This can be once more explained by the upper part of Ac decay chain not being in equilibrium. The last "long-lived" isotope before the ^{219}Rn is ^{227}Ac with mean lifetime 31.41 years [68], which is preceded by ^{231}Pa with $4.726 \cdot 10^4$ years mean lifetime [68]. If we assume that activity of ^{231}Pa is constant $A_{1,0}$ and in principle different from ^{227}Ac initial activity $A_{2,0}$, we get following expression for ^{227}Ac activity as function of time since the beginning of data taking $A_2(t)$

$$A_2(t) = A_{1,0} + (A_{2,0} - A_{1,0})e^{-\frac{t}{\tau_2}}, \tag{4.4}$$

where τ_2 stands for ^{227}Ac mean lifetime. The observed pseudorate was fitted by this function with $A_{1,0}$ and $A_{2,0}$ as free parameters and it matched the data reasonably well with ^{231}Pa initial activity being on average 7% of the ^{227}Ac initial activity.

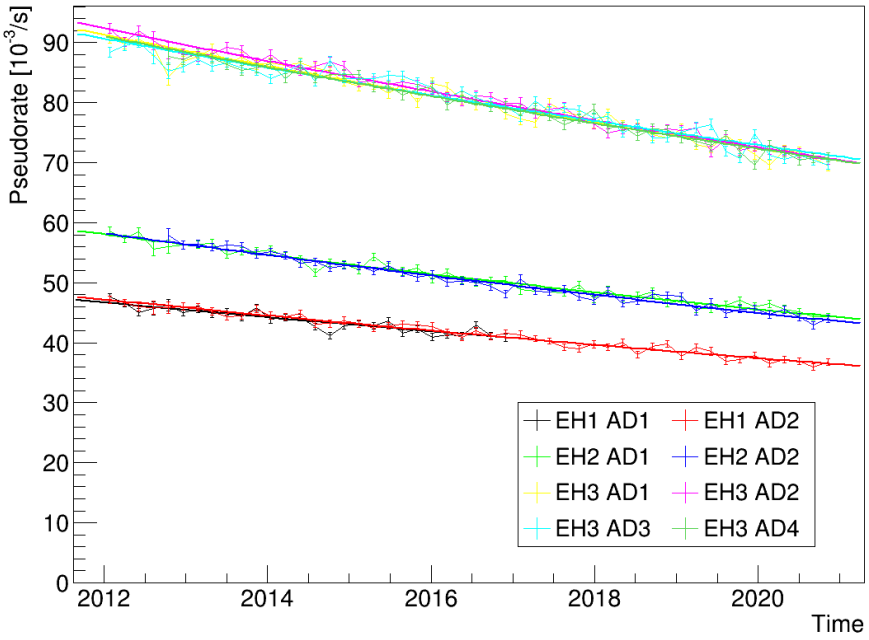

Figure 4.10: Ac-series pseudorate of all ADs fitted by function described by Eq. 4.4. During nine years of Daya Bay operation, the pseudorate decreased by about one quarter of the initial value.

4.3.4 Spatial Distribution

Fine selection cuts listed in Tab. 4.3 were used to obtain spatial distribution of Ac-series contamination and then it was normalized to the observed pseudorate. Both r^2-z and x-y maps were created and an example is shown in Fig. 4.11. Simlarly to Th-series, Ac-series contamination is also present only in GdLS and likely linked to the Gd doping. This along with the fact that Ac-series coincidence time is too long for efficient selection makes Ac-series only marginally useful for energy response studies.

4.4 U-Series

By U-series we mean correlated beta-alpha decays ^{214}Bi $\rightarrow$ ^{214}Po $\rightarrow$ ^{210}Pb. The correlation time is given by ^{214}Po mean lifetime 236.0 μs and it is coincidentally similar to the mean neutron capture time in LS of about 200 μs. There is another rare but observable correlated decay in the U decay chain – ^{234}Pa $\rightarrow$ ^{234m}U $\rightarrow$ ^{234}U. As it is not really useful for nonuniformity study, we will not discuss it this section, instead it is described in Appendix C.

4.4.1 Signal and Background Windows

Similarly to Ac-series fine selection, we can avoid Th-series contamination by using only data with coincidence time > 5 μs. However, as delayed energy of U-

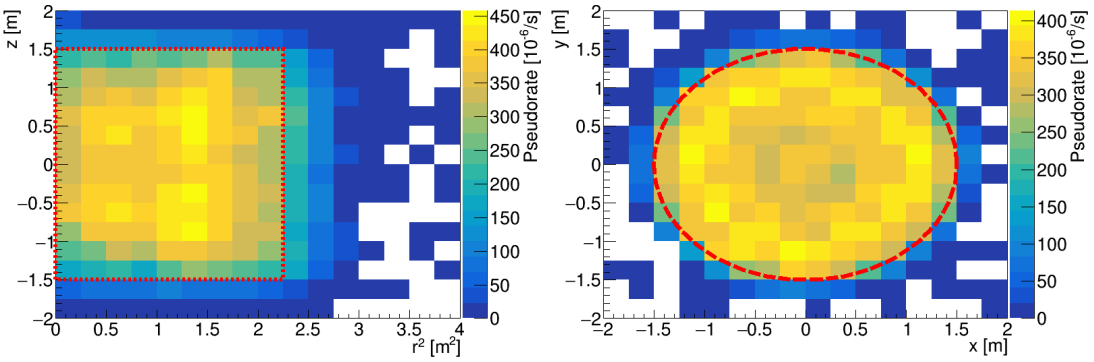

Figure 4.11: Example of Ac-series spatial distribution in r^2-z (left) and x-y (right) coordinates for EH1-AD1. Ac-series contamination is present only in GdLS ($|z| < 1.5$ m, $r^2 = x^2 + y^2 < 2.25$ m^2 – the boundary of GdLS is marked by red dashed line), even though some events are reconstructed just outside due to position reconstruction resolution. All ADs follow the same pattern.

series is almost the same as the one of Ac-series and prompt energy also overlaps, we need to take special care to disentangle these two sets of correlated decays. There are two basic approaches at hand with parameters of both of them listed in Tab. 4.4:

- (Fine selection U1) Using the fact that Ac-series prompt energy goes up to ∼1.25 MeV, while the prompt energy of U-series goes up to 3.5 MeV, we can set the lower bound of prompt energy cut to 1.5 MeV to completely eliminate Ac-series contamination. Signal window (5 μs, 705 μs) and background window (705 μs, 2000 μs) was used in this approach, which yields 79.2% of U-series correlated decays (inefficiency of prompt energy cut not included). However, we cannot study U-series prompt energy below 1.5 MeV this way.

- (Fine selection U2) We can use two signal windows (5 μs, 455.6 μs) and (1453.1 μs, 2000 μs) and one background window (455.6 μs, 1453.1 μs) to eliminate Ac-series contamination, which provides us with 69.3% of U-series correlated decays. This way, we can study even the lower part of U-series prompt energy spectrum (and other properties of the events that have lower prompt energy), but the cost is smaller efficiency and the need to subtract more background which leads to an increase of the statistical uncertainties.

4.4.2 Prompt and Delayed Energy

The prompt signal is caused by ^{214}Bi beta decay with Q-value 3269 keV [69]. The beta decay goes directly to the ground state of ^{214}Po only in 19% of cases, otherwise, deexcitation by a gamma cascade occurs. As a result, the prompt energy spectrum has a peak-like structure around 2.3 MeV if there is no energy leakage. However, as it stems from complicated combination of beta and gamma(s) signals, potential fitting would be difficult. An example of such spectrum is shown

Table 4.4: Fine selection cuts and coefficients for U-series.

	Selection U1	Selection U2
1st Signal window	(5 μs, 705 μs)	(5 μs, 455.6 μs)
Coefficient	1	1
Background window	(705 μs, 2000 μs)	(455.6 μs, 1453.1 μs)
Coefficient w_{B}	0.54	1
2nd Signal window	–	(1453.1 μs, 2000 μs)
Coefficient w_{S}	–	1
Maximal distance	750 mm	750 mm
Prompt energy	(1.5 MeV, 3.2 MeV)	(0.5 MeV, 3.2 MeV)
Delayed energy	(0.7 MeV, 1.2 MeV)	(0.7 MeV, 1.2 MeV)

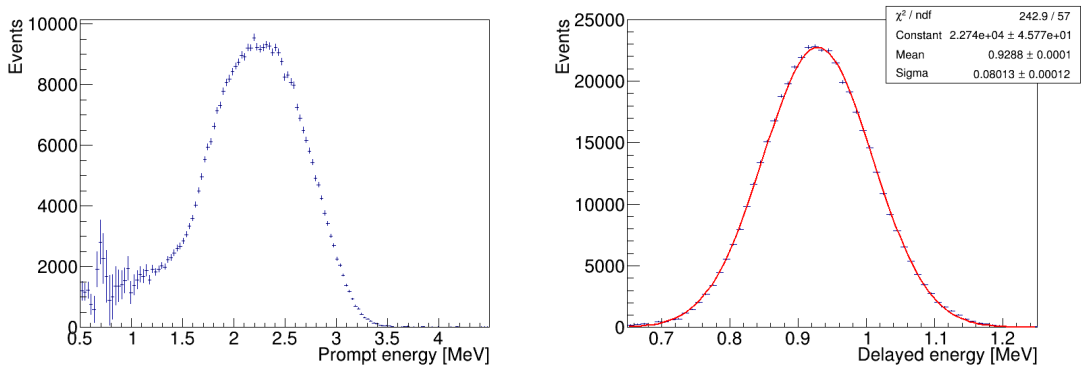

Figure 4.12: U-series prompt (left) and delayed (right) energy spectra obtained by EH1-AD1 using U2 cuts listed in Tab. 4.4 and additional spatial cut to GdLS so that the spectra would be minimally affected by energy leakage. While there is peak-like structure around 2.3 MeV in the prompt spectrum, it is caused by a complicated combination of beta and gamma signals and thus cannot be easily fitted. Large statistical errors between 0.5 MeV and 1.0 MeV there are caused by subtraction of Ac-series spectrum. On the contrary, the delayed spectrum in the GdLS shown in the right panel can be fitted by a simple Gaussian function (red line).

in the left panel of Fig. 4.12, where data only from GdLS were used to avoid distortion caused by energy leakage.

The delayed signal of U-series is caused by ^{214}Po alpha decay with Q-value 7833.5 keV [59], but the reconstructed energy is only about 0.9 MeV due to the liquid scintillator quenching. The decay goes to the ground state of ^{210}Pb in 99.99% of cases. The decay to the excited state with 0.01% probability which is deexcited by 799.7 keV gamma can be thus for most purposes neglected. As a result, we would expect to observe a simple Gaussian peak. That is indeed the case when we restrict the study to only GdLS as shown in the right panel of Fig. 4.12, which makes the situation similar to Th-series and Ac-series where we were naturally restricted to GdLS by their occurrence.

The problems arise once we move closer to the AD edge. Even though alpha particles cannot have energy leakage the same way as gamma particles do (only

decays occurring up to tens of μm from the surface of the acrylic vessel can have parts of the energy deposited in both the sensitive and the insensitive volumes), there are events with significantly lower delayed energy happening at certain places around the ADs' edges. Prompt energy is also affected, however, it is harder to study the impact there as it is also suffers from the regular gamma energy leakage. Since dealing with these so-called "low energy events" is of great importance for energy fitting in further studies, they will be discussed separately in Section 4.4.4.

4.4.3 Pseudorate and Spatial Distribution

For U-series, pseudorate and spatial distribution are closely tied together with several phenomena affecting both. For that reason, we shall first overview the general situation and then look into these phenomena in detail.

In order to study U-series pseudorate, U1 selection cuts listed in Tab. 4.4 were used with only one coincidence time window set to (5 μs, 500 μs). The result is shown in Fig. 4.13. The basic trend of the pseudorate is constant which is to be expected as the last long-lived isotope in the decay chain before ^{214}Bi is ^{226}Ra with mean lifetime 2308 years [70] and thus no fundamental change of ^{214}Bi activity should occur over the span of 9 years. There are, however, other features which make the U-series quite different from Th-series and Ac-series.

- In contrast to Th-series and Ac-series the pseudorate of U-series measured by ADs sharing the same EH is not always consistent. While it is quite similar in EH1, there are noticeable differences in EH3 and EH2-AD2 pseudorate is over 30% higher than that of EH2-AD1 (and slightly decreasing over time).

- There are temporary spikes in pseudorate, most notably in October 2012 in EH1 after the shutdown during which EH2-AD2 and EH3-AD4 were installed and in August 2017 in EH3. There are more for which to be visible better temporal resolution is needed.

These issues will be discussed in more detail later, however we can still get better understanding of what is going on by simply looking at the pseudorate in GdLS and LS separately as shown in Fig. 4.14. The pseudorate baseline is more stable in GdLS, but the spikes in EH1 are actually more pronounced there compared to LS. Moreover, EH2-AD2 pseudorate is consistent with that of EH2-AD1 in GdLS, but almost two times higher in LS, meaning that there is additional U-series contamination there when compared to other ADs. Besides, EH2-AD2 pseudorate clearly decreases over time in LS as well as pseudorate of EH3 ADs.

The spatial distribution was studied using selection U1 as listed in Tab. 4.4 and normalized to pseudorate. Unlike Th-series and Ac-series, there are considerable differences between the ADs. For that reason, spatial distributions are shown in Fig. 4.15 and Fig. 4.16 for each AD individually. The most notable difference to Th-series and Ac-series is that U-series (meaning ^{214}Bi and further) is present not only in GdLS, but also in LS, albeit not uniformly. This leads us to the question whether there is ^{226}Ra (last long-lived isotope before ^{214}Bi) contamination in LS or there is another reason for the ^{214}Bi presence in LS.

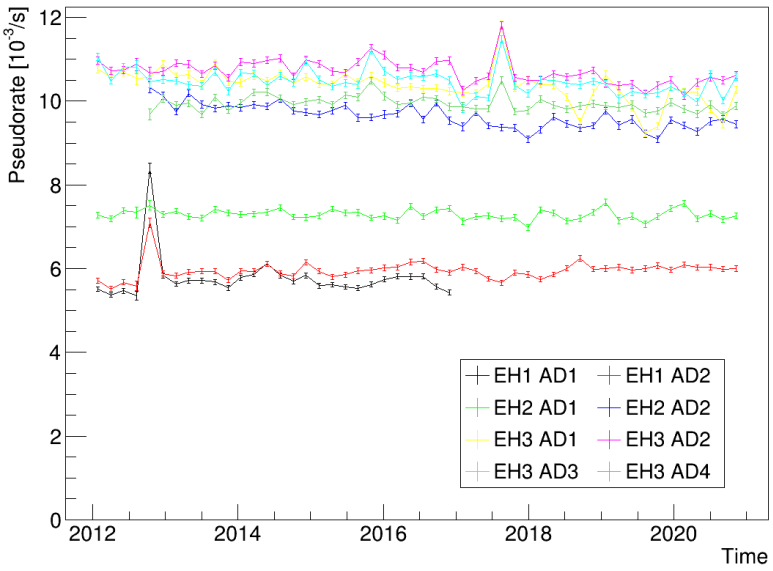

Figure 4.13: U-series pseudorate of all ADs. Whole sensitive volume of each AD was used for the measurement. The pseudorate baseline is rather stable for most ADs, except for occasional spikes. EH2-AD2 which has significantly higher pseudorate than EH2-AD1 and has slightly decreasing trend (observed to a lesser extent also in EH3).

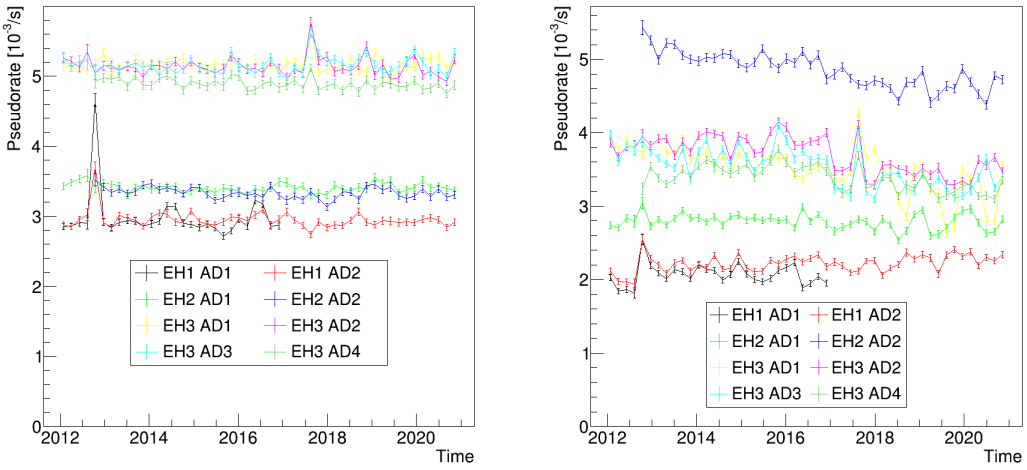

Figure 4.14: Comparison of U-series pseudorate in GdLS (left) and LS (right) of all ADs. Apparently, the excessive EH2-AD2 pseudorate is caused by additional contamination in the LS volume as the EH2-AD2 pseudorate is consistent with EH2-AD1 in GdLS, but much higher in LS. Another thing worth noting is that the pseudorate spikes in EH1 are much more pronounced in GdLS than in LS.

In each decay chain, there is one Rn isotope. As a noble gas, Rn can travel certain distance before decaying, however, ^{220}Rn of Th-series has mean lifetime

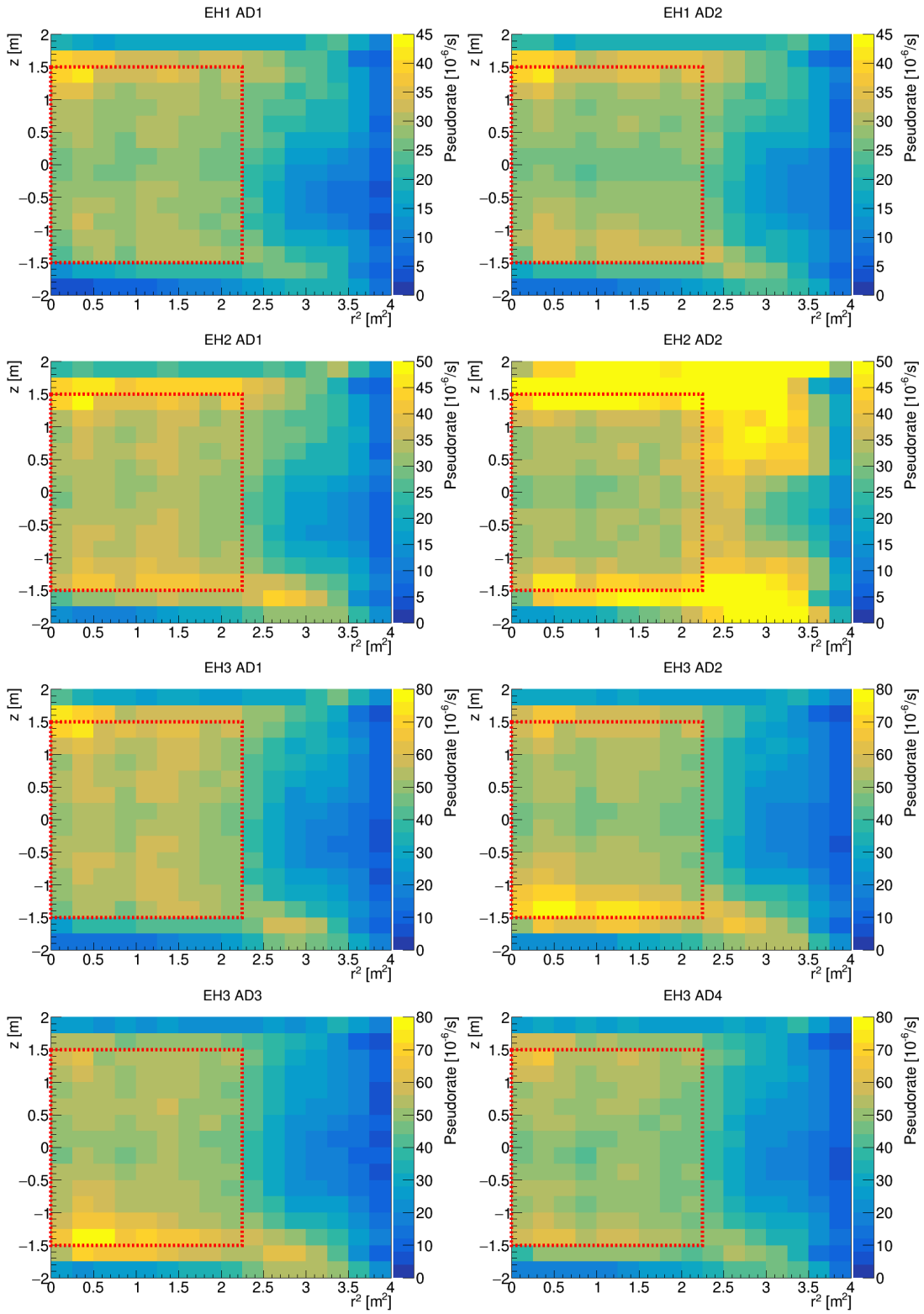

Figure 4.15: U-series spatial distribution in r^2-z coordinates of all ADs. The z-axis scale is the same for ADs sharing the same EH. Apparently, U-series events are present both in GdLS (marked by red dashed line) and LS volumes, but the exact distribution pattern may differ between ADs – espeacially EH2-AD2 stands out.

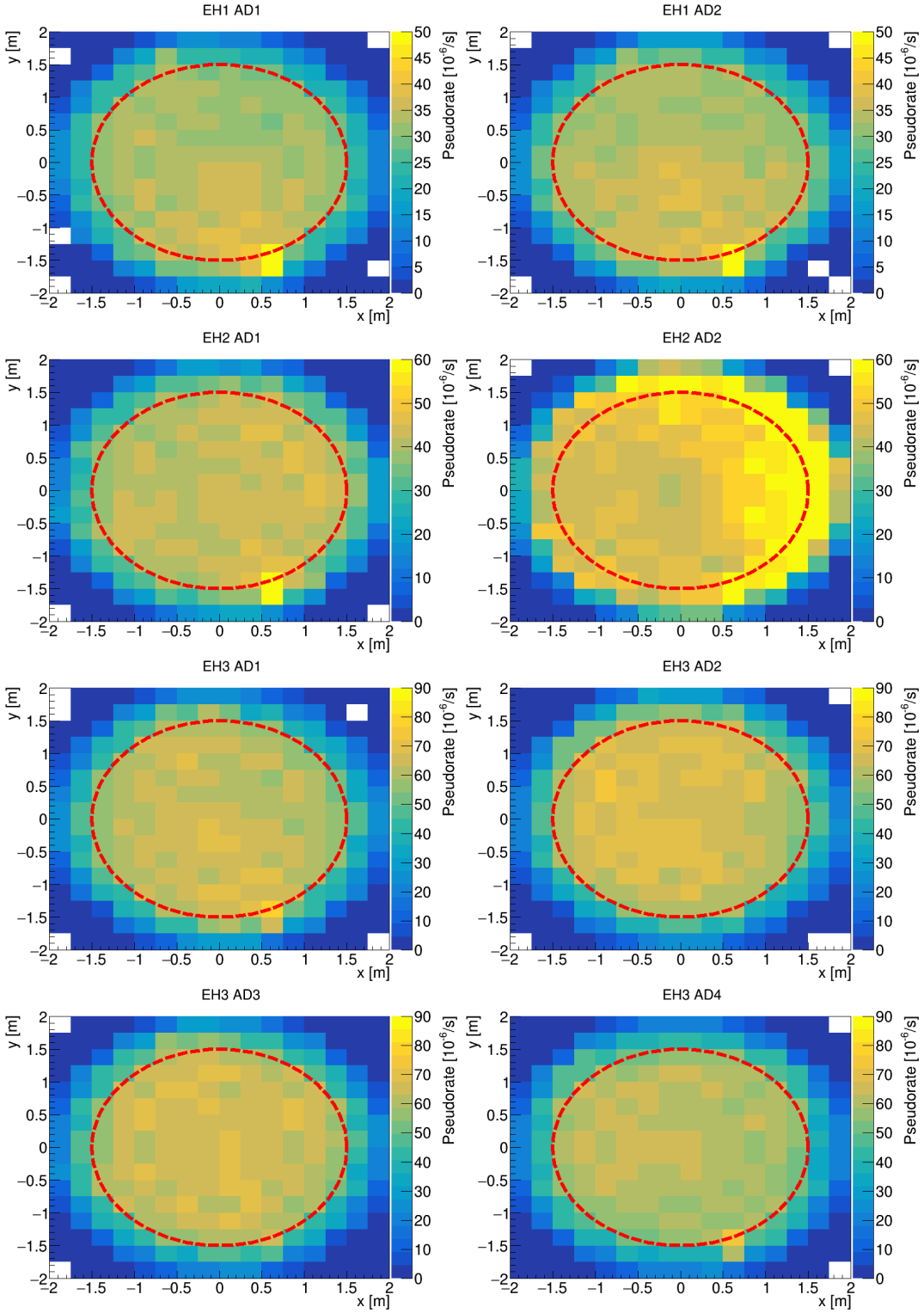

Figure 4.16: U-series spatial distribution in x-y coordinates of all ADs with GdLS volume marked by dashed red line. The z-axis scale is the same for ADs sharing the same EH. The most notable features are the extra contamination in EH2-AD2 and hotspot around $y = -1.5$ m and $x = 0.6$ m which coincides with ACU-B position and is most notable in EH1-AD1, EH1-AD2 and EH2-AD1.

80 s [71] and ^{219}Rn of Ac-series only 5.7 s [67] and thus can travel much smaller distance than ^{222}Rn of U-series with mean lifetime 5.5 days [72]. It is thus conceivable that U-series contamination in the form of ^{226}Ra is present only in GdLS and it is ^{222}Rn that permeates the LS volume. There are several indications for that:

- We have already observed that ^{228}Ra from Th decay chain is present only in GdLS. It would be difficult to explain how ^{226}Ra, different isotope of the same element, ended up in both GdLS and LS.
- The spatial distribution of the rare decay $^{234}\text{Pa} \rightarrow {}^{234m}\text{U} \rightarrow {}^{234}\text{U}$ in the upper part of the U decay chain before ^{222}Rn was studied. Even though there is very limited statistics regarding this decay, there is a strong indication that it is present only in GdLS. The propertion of the $^{234}\text{Pa} \rightarrow {}^{234m}\text{U} \rightarrow {}^{234}\text{U}$ correlated decays are described in detail in Appendix C.

If ^{222}Rn is indeed responsible for U-series contamination in the LS volume, a question of where it comes from arises, which – as we will see – is to a certain extent connected to the phenomenon of pseudorate spikes and the hotspot observed by most ADs around $y = -1.5$ m, $x = 0.6$ m and $z = 2$ m in spatial distribution maps shown in Fig. 4.16.

Pseudorate Spikes

The pseudorate spikes cannot be explained by initial out-of-equilibrium contamination by the isotopes from the U decay chain as it was possible with Th-series and Ac-series pseudorate decreases. The contamination causing the spikes thus has to come from outside, most likely in the form of ^{222}Rn.

There are three calibration pipes leading from the LS and GdLS sensitive volumes to the ACUs as shown in Fig. 4.17, the calibration pipe in the middle under ACU-A also connects the sensitive volumes to the overflow tanks on top of the stainless steel vessel and there is strong evidence suggesting that that is where the extra contamination causing the spikes of pseudorate comes from:

- By dividing the nine years of operation into 500 periods lasting about 6.5 days each, we get time resolution good enough to see how the pseudorate spikes unfold within the AD volume. As illustrated in Fig. 4.18, we can see that the pseudorate increase starts in the upper central part of GdLS below ACU-A, spreads to the rest of GdLS and then starts to disappear as the ^{222}Rn that was sucked into GdLS from the overflow tanks dies out with mean lifetime of 5.5 days.
- Thanks to many sensors, we can check what the AD temperature and liquid scintillator level were at any time. When compared to the measured pseudorate, clear correlation appeared – whenever there was significant drop in temperature, the level of liquid scintillator dropped too and spike of pseudorate appeared as illustrated in Fig. 4.19 . It did not matter whether the temperature first rose above the baseline and then dropped back or first dropped and then rose back to the baseline. The important part is that once the GdLS level drops, scintillator from the overflow tank is sucked

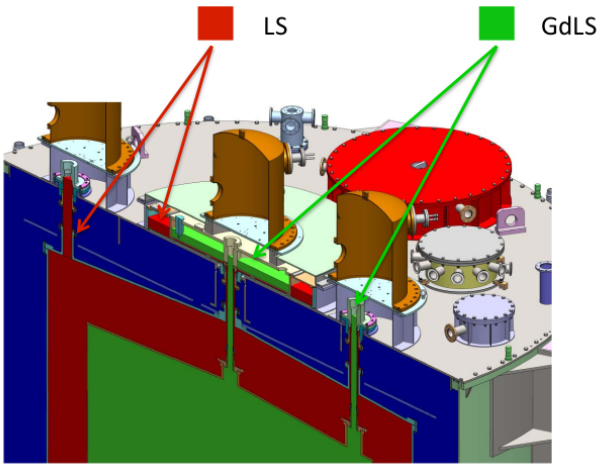

Figure 4.17: Orthogonal slice through the top of an AD showing overflow tanks and calibration tubes used for ACUs to deploy calibration sources to various parts of the ADs – ACU-C on the left-hand side deploys the sources to the LS (dark red), ACU-A in the middle along the central axis of the GdLS (green) and ACU-B on the right-hand along the GdLS boundary. The calibration tube in the middle also connects the LS and GdLS volumes to the overflow tanks on the top of the stainless steel vessel. Figure taken from Ref. [28].

to the main volume. Based on the size of overflow tank, the change of scintillator level and the pseudorate increase, we can asses that the ^{222}Rn contamination in the GdLS overflow tank can be in the order of 1000× higher than in GdLS main volume.

With the origin of the pseudorate spikes understood, we can ask whether this could affect further analyses using U-series. Fortunately, the spikes are quite short-lived so that the statistics they add is on percent level of the overall statistics at maximum. Moreover, once the ^{222}Rn enters the main volume, its and subsequent decays are in principle no different from the decays of intrinsically present ^{222}Rn. The only part of this phenomenon which could have some effect on the nonuniformity study and which we should thus keep in mind are the decays inside the ACU-A calibration pipe, where the light collection efficiency could be slightly different from its surroundings (like LS volume it goes through) and which at the same time can get slightly overrepresented due to the high rate of decays during the pseudorate spikes.

Hotspot under ACU-B

The hotspot located around the position $y = -1.5$ m and $x = 0.6$ m, which coincides with ACU-B position, is most prominent in EH1-AD1, EH1-AD2 and EH2-AD1 as shown in Fig. 4.16. It is observable, but not very distinct in EH3-AD1 and EH3-AD4 and barely observable in EH3-AD2 and EH3-AD3. Regarding EH2-AD2, as there is extra contamination, it is difficult to assess whether higher rate under ACU-B has the same origin as in other ADs or not. However, there is indication that it has the same origin – when we compare the pseudorate within

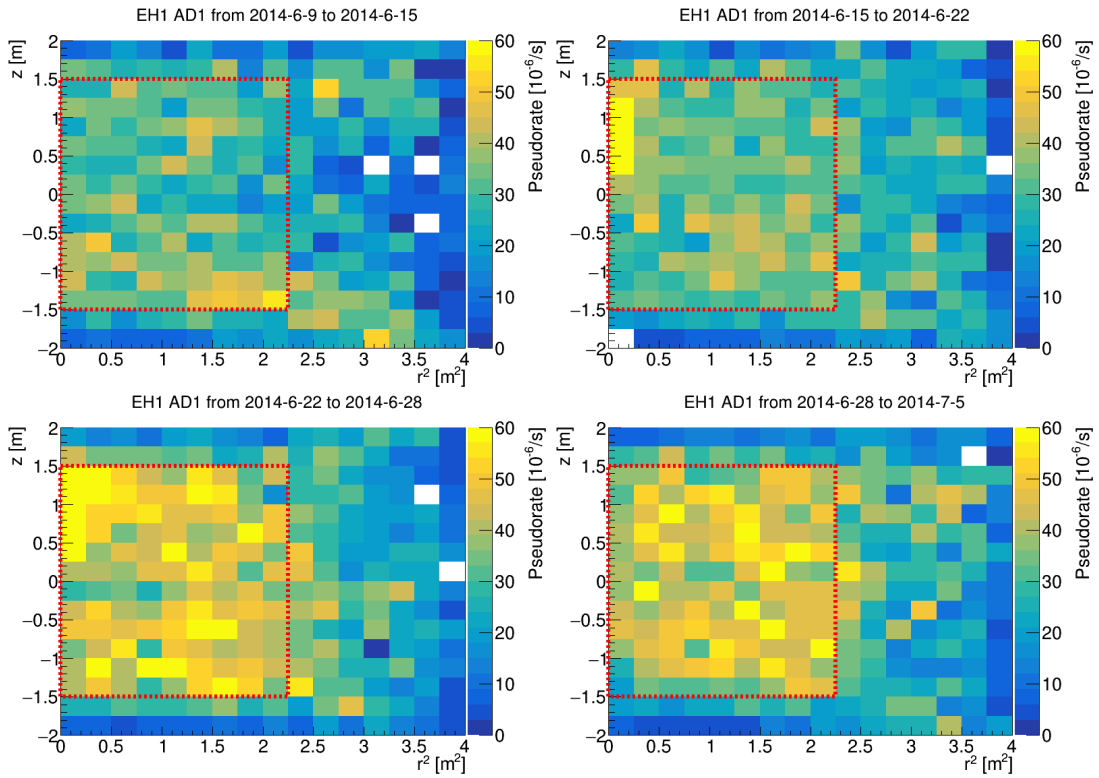

Figure 4.18: The spatial distributions in r^2-z coordinates covering the beginning and peaking of U-series pseudorate spike in EH1-AD1 at the end of June 2014 shown in detail in Fig. 4.19. Each plot covers about 6.5 days. The first one shows the regular situation before the spike begins. In the second plot, we can see beginning of the pseudorate increase in the upper central part of GdLS under ACU-A, which then spreads to the rest of GdLS in the third plot and starts to die out in last plot as the ^{222}Rn contamination has mean lifetime of 5.5 days.

the hotspot (defined as $x \in$ (0.5 m, 0.75 m) and $y \in$ (-1.75 m, -1.25 m), z cut was not used) between ADs in EH2 as shown in Fig. 4.20, we can see that both ADs follow similar pattern with correlation coefficient 0.76. In EH1, the hotspot pseudorate of both ADs is even more consistent with correlation coefficient 0.81. Regarding EH3, the hotspot there is less pronounced, but for EH3-AD1 compared with EH3-AD2, EH3-AD3 and EH3-AD4 we get correlation coefficients 0.46, 0.41 and 0.65 respectively.

The origin of the hotspot is likely similar to the pseudorate spikes – ^{222}Rn contamination, this time entering the main active volume through the ACU-B calibration pipe, but these two phenomena are not correlated in time. The hotspot presence has more continuous character, even though there is quite a big variation in pseudorate as illustrated Fig. 4.20. Whatever the original source of ^{222}Rn is, it is likely common on the level of EH and probably distributed by the cover gas system (more about Daya Bay gas system in Ref. [73]).

Regarding the impact on further analyses, the same remarks as for the pseudorate spikes apply – caution must by taken when this particular part of AD is used for further (nonuniformity) studies as the decays inside the calibration pipe could be overrepresented compared to its immediate surroundings and at the

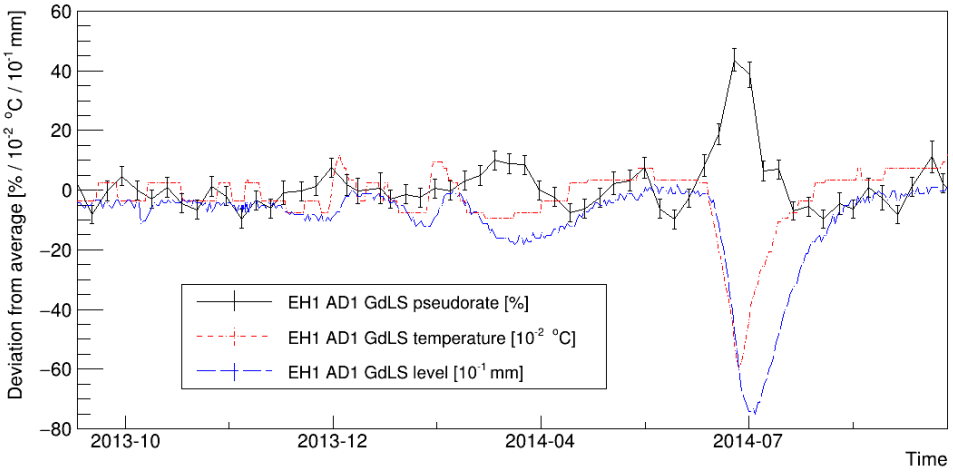

Figure 4.19: Illustration of the connection between EH1-AD1 U-series pseudorate, AD temperature and scintillator level. As the temperature dropped by 0.6°C, the scintillator level dropped by about 8 mm, which led to about 10 l of scintillator being sucked from to overflow tank to the GdLS volume and 40% increase of pseudorate. That corresponds to the ^{222}Rn contamination in overflow tank being of the order of 1000× higher than in GdLS main volume.

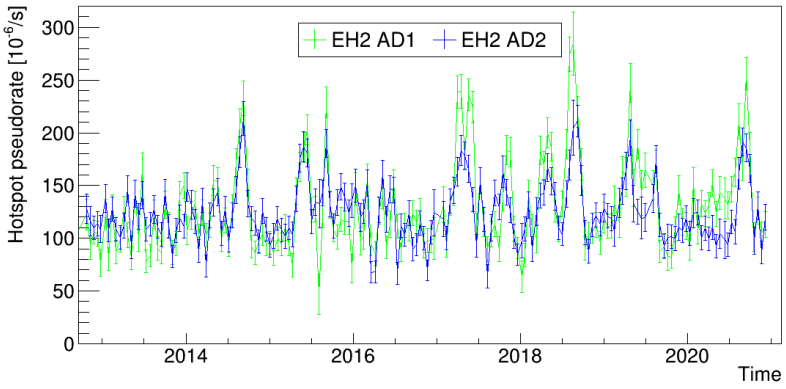

Figure 4.20: Comparison of the hotspot (defined as $x \in$ (0.5 m, 0.75 m) and $y \in$ (-1.75 m, -1.25 m)) pseudorate of EH2-AD1 and EH2-AD2. As we can see, the hotspot pseudorate varies quite a lot, but both ADs follow similar pattern.

same time have slightly different light collection efficiency. As we will see, it will be relevant especially for azimuthal nonuniformity study and the new correction design.

Higher Rate in EH2-AD2

As clearly shown in Figs. 4.15 and 4.16, EH2-AD2 has significantly higher (about 70%) U-series contamination in the LS volume when compared to other ADs. The additional contamination is not distributed uniformly, but tends to create

one or more hotspots which move around the LS as illustrated in Fig. 4.21. The hotspot(s) then tend to stay around the top or the bottom of the AD and more rarely around the middle and their movement along the z-axis can be quantified by mean z-position of the observed U-series signal in a given period. This is shown in Fig. 4.22 with the full 9 years of data divided into 500 periods of 6.5 days each. Interestingly enough, when we plot the mean z-position as function of time of EH2-AD1 over the plot of EH2-AD2, we get almost the same pattern, just with smaller amplitude. The correlation coefficient between these two is 0.96. Whatever the source of the U-series extra contamination in EH2-AD2 is, the movement of the contamination is clearly correlated between the ADs in EH2 and the same actually applies for EH1 with correlation coefficient 0.89 and similar situation is also in EH3. Correlation coefficient between EHs is much lower – ~ 0.3 between EH1 and EH2, ~ 0.1 between EH1 and EH3 and ~ 0.2 between EH2 and EH3. This means that the correlation is not caused by processes that are common for all three halls like stirring of liquid scintillator caused by deployment of the calibration sources from the ACUs on regular weekly basis. Regarding x and y coordinates, no significant correlation between ADs was observed.

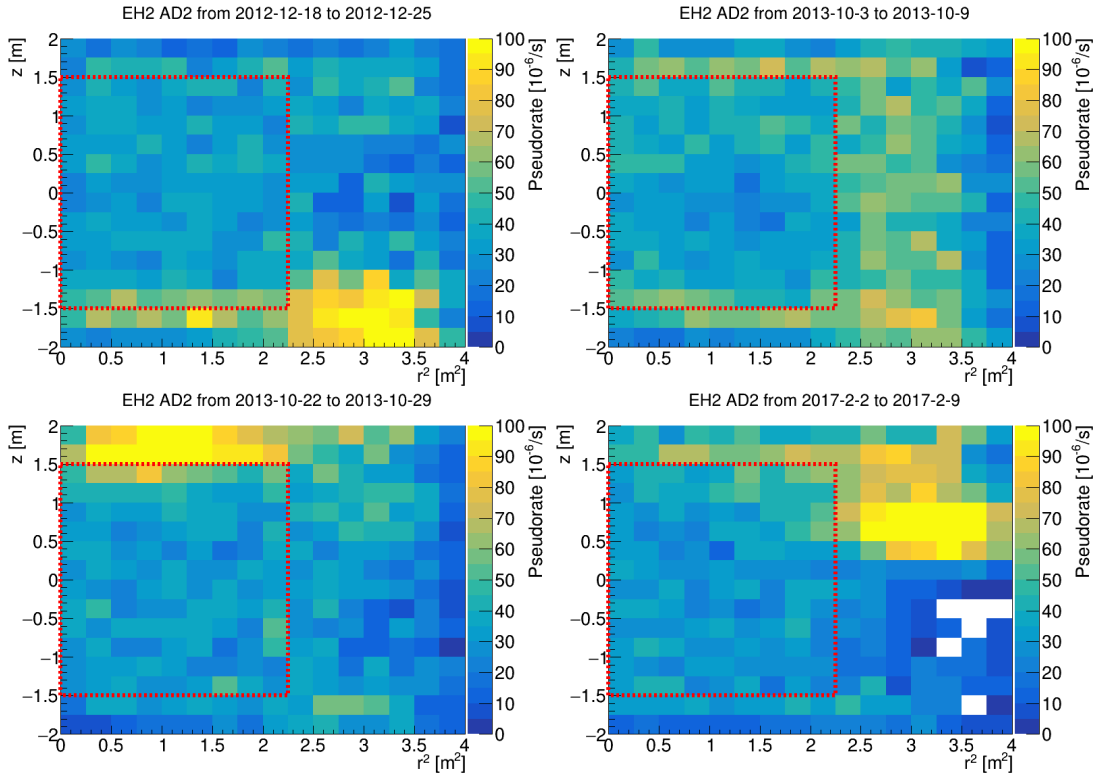

Figure 4.21: Four examples of EH2-AD2 U-series spatial distributions in r^2-z coordinates, where each histogram corresponds to about 6.5 days of data taking. In the top left and bottom plots, we can see the hotspot of higher pseudorate present in various parts of the AD. In the top right plot, there is slightly higher pseudorate in the whole LS (outside of the red dashed line) instead of localized hotspot.

The cause of the change of the U-series mean z-postion remains unclear as no connection to the AD temperature, scintillator level or humidity of the cover

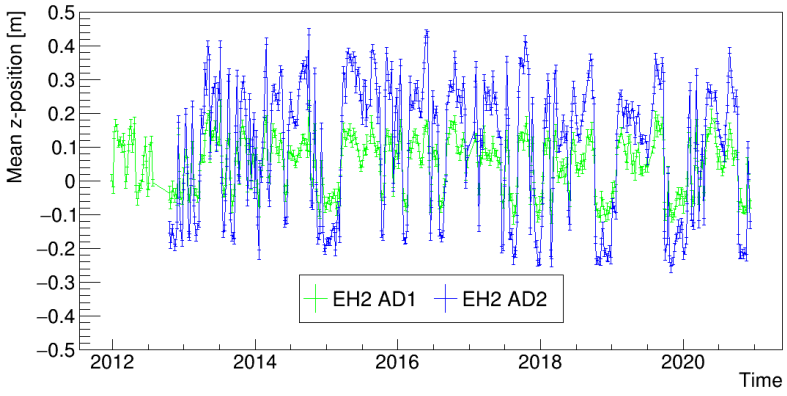

Figure 4.22: Comparison of the evolution of U-series signal mean z-position in EH2-AD1 and EH2-AD2. The nonzero value of the z-mean is caused by uneven U-series contamination distribution in AD. For EH2-AD2, LS is the main contributor to the high z-mean amplitude (in GdLS, mean z-position stays within 0.1 m from zero for both ADs). Even though U-series contamination in LS is significantly higher in EH2-AD2 than EH2-AD1, similar behavior of the z-mean suggest that U-series contamination in EH2-AD2 is not fundamentally different from other ADs in any other aspect than its rate in LS.

gas was found. The U-series extra contamination in EH2-AD2 raises a question of whether EH2-AD2 U-series can be used consistently with other ADs. However, the similarity in mean z-position behavior suggests that EH2-AD2 is not fundamentally different from other ADs in any aspect except the rate in LS.

4.4.4 Low Energy Events

Despite the fact that the U-series delayed energy spectrum has simple Gaussian shape in GdLS as expected and already shown in Fig. 4.12, when we look at the top and bottom of LS, events with lower energy also pass our selection as illustrated in Fig. 4.23. As these low energy events pose a notable hindrance in the fitting of the regular peak, a detailed look into their properties is needed.

As mentioned above, alpha particles deposit energy over a very short distance in liquid scintillator (tens of μm) and thus they cannot be affected by energy leakage in the same way as gammas, which can scatter outside of the sensitive volume depositing only part of their energy inside. We would thus expect a minimum of U-series events with $E_{\text{delayed}} < 0.65$ MeV as it is indeed the case in GdLS, where the delayed peak has a simple Gaussian shape as shown in Fig. 4.12. Since our ultimate goal is to use U-series delayed signal to study the ADs' nonuniformity and propose new correction, we need to be able to extract the peak mean from the delayed energy spectrum at any position of the detector. A simple Gaussian peak can be easily fitted, but if it sits upon a pedestal consisting of low energy events, the fitting might be biased. Therefore, it is necessary to find a way to eliminate the low energy events or at least suppress them. If that is not possible then the fitting function must be altered to accommodate for them.

Regarding the prompt energy spectrum, it is much more difficult to assess the

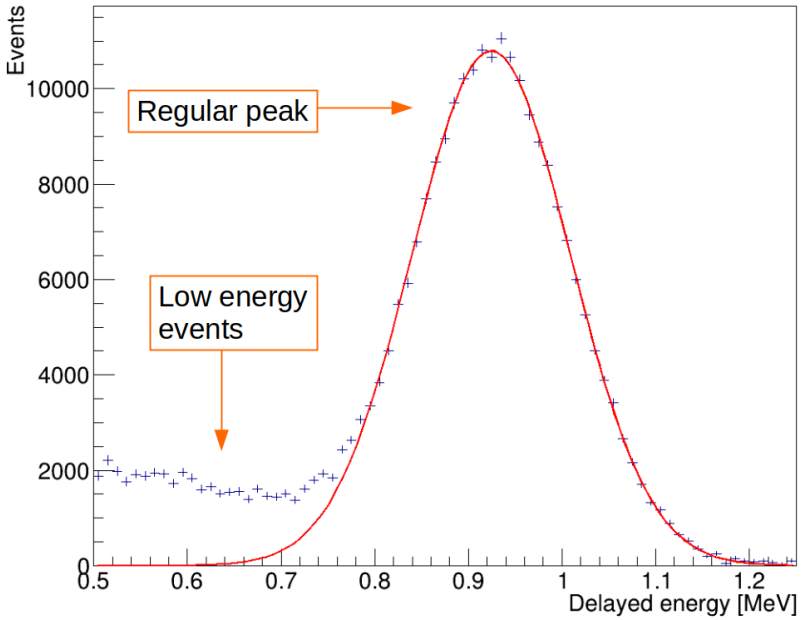

Figure 4.23: Illustration of the phenomenon of low energy events in the U-series delayed spectrum obtained by EH3-AD2 with U1 fine selection cuts and additional spatial cut $|z|$ >1.75 m. The Gaussian shape of the regular Gaussian peak is highlighted by the red line.

impact of the low energy events for two reasons. First, there is no lower bound in the spectrum below which we could clearly identify the low energy events as with the delayed spectrum. Second, due to the gamma contribution to the prompt spectrum, it is affected by regular energy leakage around the edges of the sensitive volume, which obscures whatever effect the cause of low energy events has.

In order to study spatial distribution, pseudorate and coincidence time of the low energy events, delayed energy cut $E_{\text{delayed}} \in$ (0.5 MeV, 0.65 MeV) was used instead the one listed in U1 and U2 selections parameters. Modified prompt energy cut $E_{\text{prompt}} \in$ (0.5 MeV, 1.5 MeV) can be also used to highlight the low energy events, but its effect varies quite a bit depending on the position inside the AD.

Generally speaking, the number of the low energy events compared to the regular events varies widely between ADs, but they tend to appear at the same spots for all ADs. An example of the spatial distribution obtained with both modified prompt and delayed energy cuts is shown in Fig. 4.24. EH3-AD2 was chosen as an example as it suffers from the low energy events relatively heavily. Large part of such events there (and in other ADs too) are concentrated around the upper outer edge of LS (labeled as “Spot A”). Then there are two less prominent spots – Spot B in the bottom outer part of LS and Spot C in the upper central part of the AD. Each of these spots have a slightly different properties and thus they need to be studied individually.

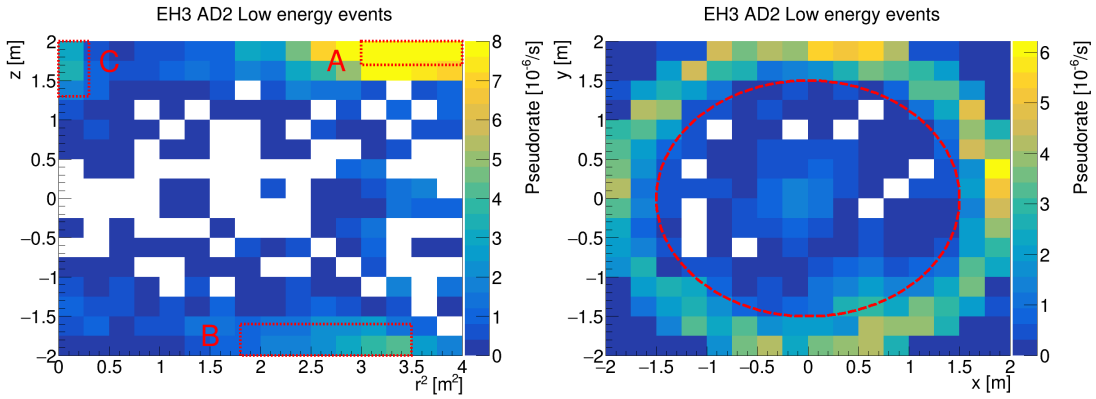

Figure 4.24: Spatial distribution of low energy events in EH3-AD2 obtained using U2 selection with modified cuts $E_{\text{prompt}} \in$ (0.5 MeV, 1.5 MeV) and $E_{\text{delayed}} \in$ (0.5 MeV, 0.65 MeV). While most of the low energy events are observed around the outer top part of the AD, they are not restricted to a particular azimuth thus ruling out ACUs as a direct source. Even though the (pseudo)rate of low energy events differs quite a lot between ADs, the spatial distribution is quite consistent.

Spot A

For the study of the low energy events in Spot A, spatial cuts $r^2 \in$ (3 m^2, 4 m^2) and $z \in$ (1.7 m, 2 m) were used. Unlike the other spots, there is sufficient statistics for us to make a check of how the pseudorate of these low energy events evolved over time. The result is shown in Fig. 4.25. The low energy events are not restricted to a certain time interval (like a U-series pseudorate spikes), but occur during the whole operation more or less uniformly. It is thus clear that we cannot get a cleaner sample of these low energy events by restricting the selection to a certain period of time. It is also apparent that the pseudorate differs quite a bit between ADs, even the ones in the same EH. There is especially high spread in EH3 with EH3-AD2 having triple the low energy events pseudorate of EH3-AD4.

The coincidence time distribution is an important parameter that can help determine whether the low energy events are really caused by sequential ^{214}Bi and ^{214}Po decays or by a different signal that passes our selection. The coincidence time distribution of each AD was fitted as illustrated in Fig. 4.26 for EH3-AD2. The fitted mean lifetime of each AD was within statistical uncertainty from the ^{214}Po mean lifetime of 236.0 μs. There is no other known source besides U-series correlated decay that would show this exact coincidence time distribution.

As mentioned before, the prompt energy spectrum is a bit more difficult to analyze. Nevertheless, we can still compare the prompt spectrum that corresponds to regular delayed energy $E_{\text{delayed}} \in$ (0.7 MeV, 1.2 MeV) and the one corresponding to low delayed energy $E_{\text{delayed}} \in$ (0.5 MeV, 0.65 MeV). An example is shown in Fig. 4.27. Generally speaking, lower delayed energy is correlated with lower prompt energy, hence we can increase the fraction of low energy events in our selection for Spot A by modifying the prompt energy cut to $E_{\text{prompt}} \in$ (0.5 MeV, 1.5 MeV). We will use this fact to study the delayed spectrum

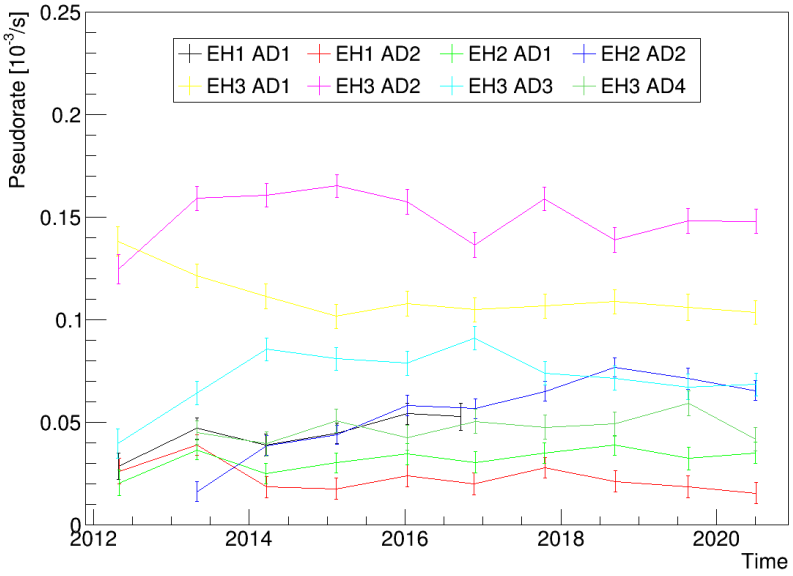

Figure 4.25: Pseudorate of U-series low energy events in Spot A. Even though, it is not perfectly stable, there are no significant changes that would hint at the origin of these events.

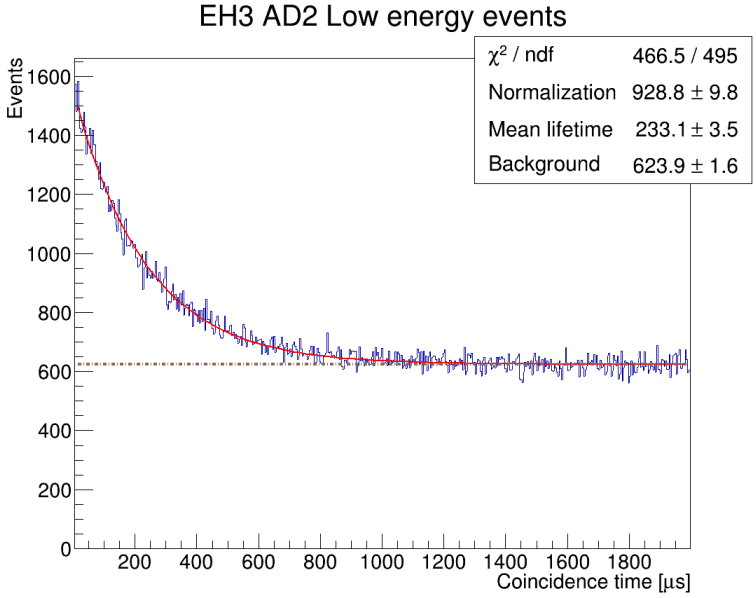

Figure 4.26: Coincidence time distribution of Spot A low energy events in EH3-AD2. As it consistent with ^{214}Po mean lifetime (236.0 μs), we can conclude that the low energy events in Spot A are caused by U-series correlated decays.

of Spot A.

In the delayed spectrum, the low energy events manifest as a peak-like structure centered about 0.5 MeV – 0.6 MeV (it should be noted that the detection efficiency under 0.7 MeV starts to fall below 100%, which influences the shape)

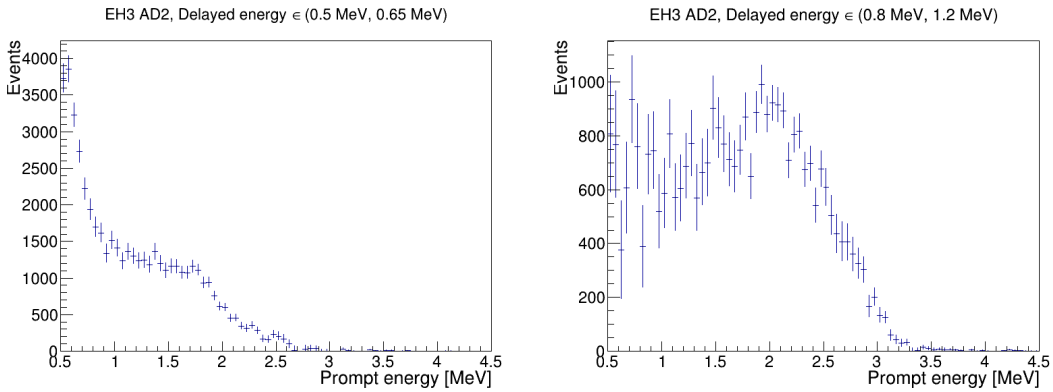

Figure 4.27: Comparison of two prompt energy spectra obtained by EH3-AD2 with Spot A spatial cuts $r^2 > 3$ m and $z > 1.7$ m and most of U1 fine selection cuts except for modified delayed energy cut. (Left) $E_{\text{delayed}} \in (0.5\text{ MeV}, 0.65\text{ MeV})$ cut used to obtain the prompt spectrum corresponding to the low energy events. (Right) $E_{\text{delayed}} \in (0.8\text{ MeV}, 1.2\text{ MeV})$ cut used to obtain a regular spectrum for reference.

separated or semi-separated from the main peak with mean around 0.9 MeV (this applies for Spot A, situation is different in Spot B and C). Four examples are shown in Fig. 4.28 along with fitting function which will be discussed in the next paragraphs. Even though it is conceivable that there are U-series decays happening outside of the sensitive volume with only gammas penetrating inside, it does not really match the observation in Spot A (but might explain Spot B as we will see). Unlike ^{214}Bi, there is only one relevant gamma line occurring in ^{214}Po decay – 799.7 keV with 0.01% intensity. If fully absorbed in sensitive volume, peak around 0.75 MeV could be expected to be observed when liquid scintillator nonlinearity is taken into account. However, that is where the dip between the regular peak and the low energy peak is located. The position of the low energy peak has a better correspondence to the Compton edge expected at 0.55 MeV, but there is no explanation at hand as to why only Compton edge and no full energy peak would be visible in a detector of the size and shape used in Daya Bay.

Whatever the origin of the low energy events is, it is clear that there is no feasible way to completely remove them from the delayed energy spectrum. The prompt energy cut used by U1 selection, $E_{\text{prompt}} \in (1.5\text{ MeV}, 3.2\text{ MeV})$, fortunately removes bigger part of the low energy events. In theory, we could set the lower bound even higher to remove almost all of them, but as a consequence we would lose much of the statistics needed for nonuniformity study and new correction design. It is thus necessary to adjust the fitting function to include the low energy events. As their origin is unclear we have to search for the proper function empirically. The simplest choice is to add second Gaussian function, so that the first one corresponds to the regular peak and the second one to the low energy events peak-like structure. This proved to proved match the low energy events in Spot A very well. Few examples are shown in Fig. 4.28. The fitting function can be further tested when we look at low energy events which happen to be reconstructed outside of the sensitive volume with $r^2 > 3$ m^2 and $z > 2$ m as

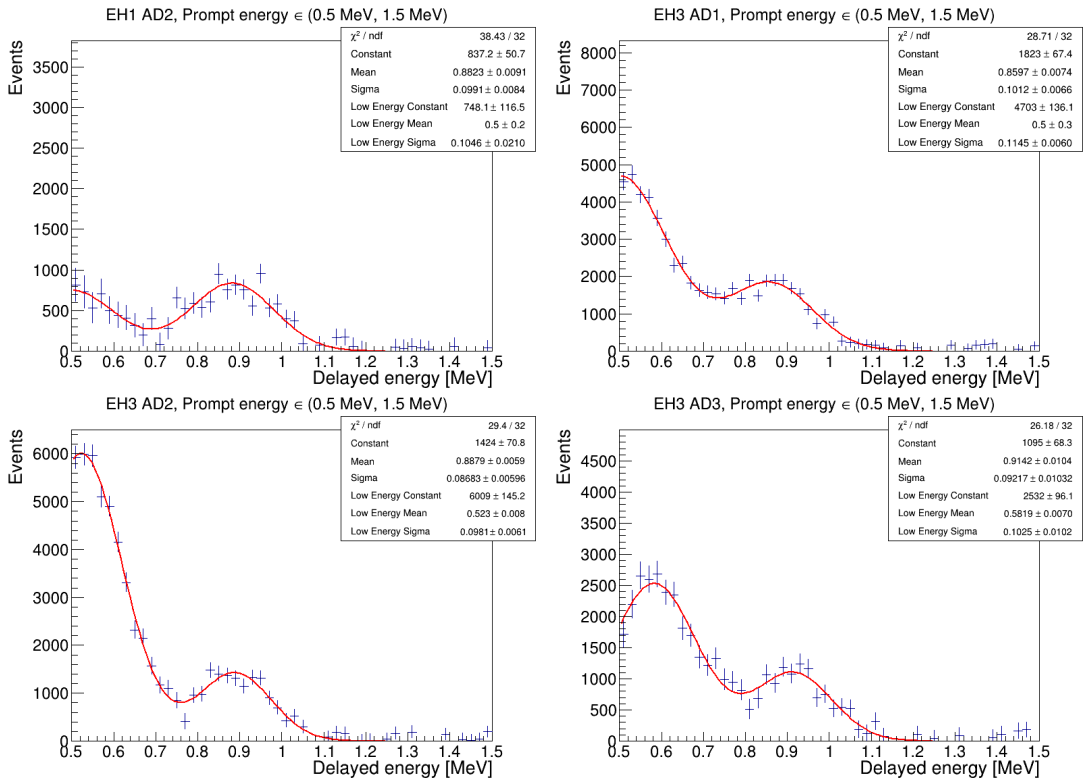

Figure 4.28: Four example of delayed spectra obtained with U2 fine selection cuts with modification $E_{\text{prompt}} \in$ (0.5 MeV, 1.5 MeV) and Spot A spatial cuts. The ranges of y-axes were set in the way that the main peak has the same height in all plots so that the low energy events contamination can be easily compared. EH1-AD2 and a group of EH3 ADs were chosen for that purpose as they have the lowest and highest admixture of low energy events respectively. The spectra are fitted by a double Gaussian function.

the regular peak is even smaller there. While we cannot infer much from clearly not perfectly reconstructed data, the fact that the fitting function still matches the data there (two examples shown in Fig. 4.29) is a good sign. In conclusion, double Gaussian function can be used to fit the delayed spectrum of Spot B and – as we will see – in other parts of the ADs as well.

Spot B

In order to study low energy events in Spot B, spatial cuts $r^2 \in$ (1.8 m^2, 3.4 m^2) and $z \in$ (-2 m, -1.6 m) were used. However, low energy event properties like pseudorate evolution over time cannot be studied in the same detail as we did with Spot A because there is significantly lower statistics in Spot B. The highest Spot B contamination is in EH3-AD3 followed by EH3-AD2, EH3-AD1 and EH2-AD1, meanwhile there is almost none in EH2-AD2.

To make sure U-series correlated decays are really responsible for the observed low energy events, the coincidence time distribution was studied and it was found out that it is consistent with ^{214}Po mean lifetime.

In contrast to Spot A, the low energy events in the delayed spectrum of Spot

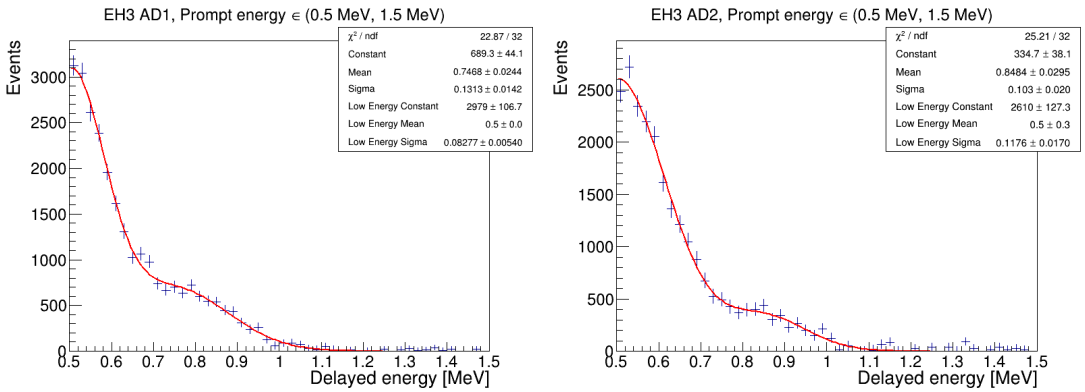

Figure 4.29: Two examples of delayed energy spectra reconstructed outside the sensitive volume ($r^2 > 3$ m^2, $z > 2$ m) above the Spot A, where there is very low fraction of regular events (otherwise U2 fine selection cuts with modification $E_{\text{prompt}} \in$ (0.5 MeV, 1.5 MeV) were used). Delayed energy spectra are fitted by a double Gaussian function, which matches the data reasonably well.

B do not manifest as a separate peak-like structure around 0.5 MeV, but instead they take a form of what looks like an increased tail from the regular peak towards lower energies as illustrated in Fig. 4.30 for EH3-AD3 which is the most affected AD and EH2-AD2 which is the least affected one. We can also see that the double Gaussian function introduced before matches the data very well.

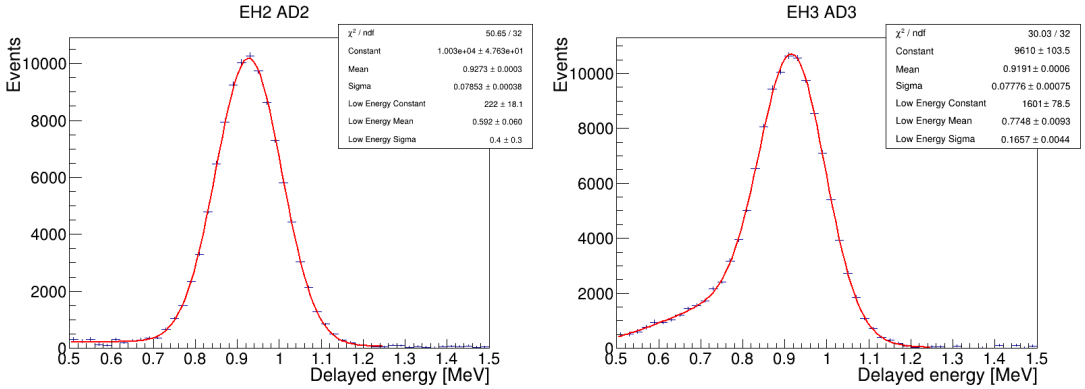

Figure 4.30: Two examples of Spot B delayed energy spectra that were obtained using U1 fine selection along with the Spot B spatial cut $r^2 \in$ (1.8 m^2, 3.4 m^2) and $z \in$ (-2 m, -1.6 m). The ADs we chosen in order to highlight the differences in Spot B low energy events contamination levels – EH2-AD2 is the least affected one and EH3-AD3 is the most affected one. Double Gaussian function was used for the fitting.

The fitted low energy mean tends to be around 0.75 MeV which is where we would expect the 799.7 keV gamma present in 0.01% of ^{214}Po decays to be reconstructed if the alpha particle did not reach the sensitive volume. If that is indeed the case, DYB function as described by Eq. 3.13 and Eq. 3.14 (with one exponential tail for simplicity) would be in theory more appropriate to model the low energy events spectrum than a simple Gaussian function. There is quite

an important distinction between the two as illustrated in Fig. 4.31 – while the DYB function has a tail towards lower energies incorporated in the formula, the Gaussian function does not have such tail. Instead, the fitting converges on a large value of sigma. This makes the fitted Gaussian function wider towards both lower and *higher* energies. Thus, when there is a tail towards lower energies in the data, both functions can somehow match it, but as a result they end up differing greatly in the shape towards higher energies. If we view the function describing the low energy events as a pedestal upon which the regular peak sits, it is obvious that the shape of the pedestal directly influences the fitted mean of the regular peak. To highlight this fact, the fitting functions of the regular peak and that of the low energy events are drawn separately in Fig. 4.31 so that the impact of the latter on the former can be seen.

There is, however, an important caveat regardind the DYB function. If we use it to model the low energy events, we end up with a lot of free parameters (3 from the regular Gaussian peak and 5 from the DYB function) which makes the fitting rather unstable and prone to converging to non-physical values. This problem can be solved by fixing the low energy mean to the values expected for 799.7 keV gamma – 0.75 MeV mean and 0.07 MeV sigma. After that the fitting tends to converge successfully and the influence of the low energy events model on regular peak mean fitting can be assessed. As shown for EH3-AD3 in Fig. 4.31, there is about 0.4% difference in fitted regular peak mean between the Gaussian and the DYB function low energy events model. As it is the most affected AD, the difference in other ADs is similar or smaller.

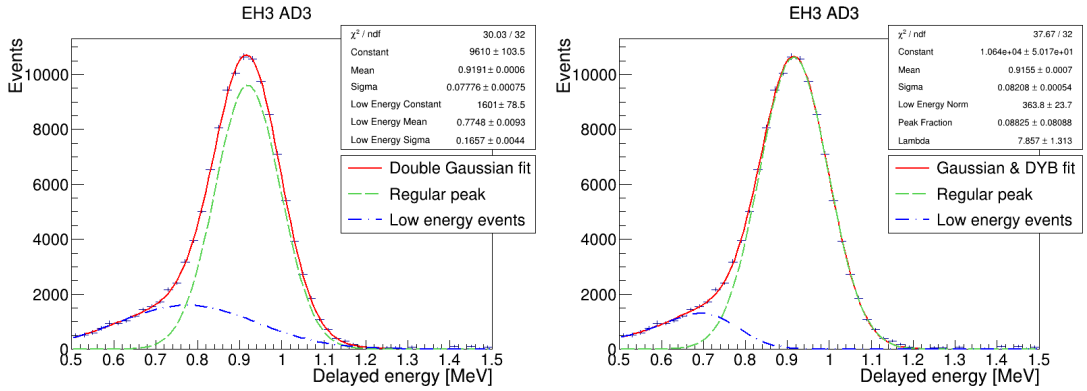

Figure 4.31: Comparison of how two different functions fit the Spot B delayed spectrum in EH3-AD3. We can see the decomposition of the fitting function into the part corresponding to the regular peak (Gaussian function in both panels, drawn in green) and the part responsible for the low energy events (Gaussian function in the left panel, DYB function with one exponential tail in the right panel, both drawn in blue).

The question of which low energy model in Spot B is correct cannot be answered with absolute certainty, but we can assess that a wrong model might introduce up to 0.4% error in the regular peak energy determination.

Spot C

The spatial cuts used to investigate Spot C were set to $r^2 \in (0\ \mathrm{m}^2, 0.3\ \mathrm{m}^2)$ and $z \in (1.3\ \mathrm{m}, 2.0\ \mathrm{m})$, which makes this spot smaller in volume (and corresponding statistics) than the other spots. Another difference is that the low energy events properties in Spot C are very consistent across all ADs. It was also found out that the prompt energy cut does not have a big impact on the shape of the delayed spectrum and thus regular cut $E_{\mathrm{prompt}} \in (1.5\ \mathrm{MeV}, 3.2\ \mathrm{MeV})$ could be used to further study Spot C. Using this cut along with $E_{\mathrm{delayed}} \in (0.5\ \mathrm{MeV}, 0.65\ \mathrm{MeV})$ a check of the coincidence time distributions in all ADs was made confirming that they are consistent with ^{214}Po mean lifetime.

Regarding the delayed energy spectrum, the low energy events in Spot C take a form of a flat tail with height of about 1⁄10 of the regular peak as illustrated in Fig.. 4.32. Even though it is not completely clear what their origin is, it is likely related to the calibration tube connecting the sensitive volumes with the overflow tanks. As we have already seen when discussing pseudorate spikes, there is much higher U-series contamination in the overflow tanks which in turn might lead to excessive decays inside the calibration tube, where the light collection efficiency could be lower than in its surroundings resulting in lower reconstructed energy. Unfortunately, even if that is the case, we cannot predict the shape of the low energy events delayed spectrum without a detail Monte Carlo study. The most straightforward option of how to include Spot C low energy events in the fitting is again to use a Gaussian function as we did with the other spots. Once again, it matches the data well as illustrated in the left panel of Fig. 4.32.

Another low energy model which matches the data reasonably well is the constant tail of DYB function described in Eq. 3.13 with the endpoint (μ in Eq. 3.13) set to the regular peak energy and with the same sigma. Fitting using this model is shown in the right panel of Fig. 4.32 and it leads to the regular peak mean being reconstructed at up to 0.4% higher energy than with Gaussian function model of low energy events.

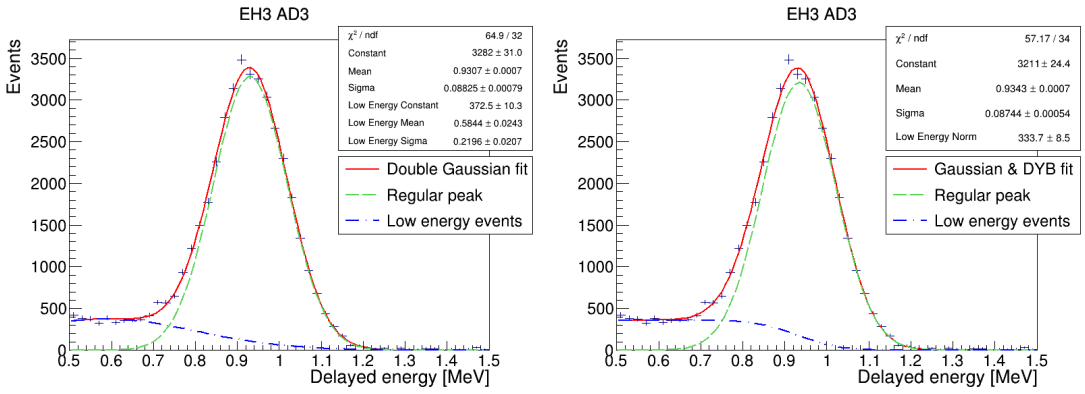

Figure 4.32: The delayed spectrum of EH3-AD3 obtained with U1 fine selection cuts and Spot C spatial cut $r^2 \in (0\ \mathrm{m}^2, 0.3\ \mathrm{m}^2)$ and $z \in (1.3\ \mathrm{m}, 2.0\ \mathrm{m})$. The spectra in other ADs are very similar. Two different models for low energy events were used – (left) Gaussian function and (right) constant tail from DYB function.

As mentioned before, we cannot decide with certainty which low energy model

is correct, but if we use simple Gaussian model, our bias in regular peak mean determination is likely within 0.4%. Moreover, even if we base a new nonuniformity correction on the wrong assumption regarding Spot C, it will be consistent across all ADs and will not create a near – far bias, something that we especially would like to avoid.

4.4.5 Delayed energy peak fitting

As we have shown, the regular delayed peak can be fitted by a Gaussian function. In some parts of the ADs, there are also low energy events which cannot be removed and thus need to be included in the fit. The simplest option which matches the data in all parts of the ADs is to use a second Gaussian function. As that brings the total number of parameters to 6, it was necessary to put some constrains on their values to ensure good convergence. The final fitting function for U-series thus has a form of

$$\begin{aligned} F_{\mathrm{U}}(E; N_{\mathrm{R}}, \mu_{\mathrm{R}}, \sigma_{\mathrm{R}}, N_{\mathrm{L}}, \mu_{\mathrm{L}}, \sigma_{\mathrm{L}}) &= \\ = N_{\mathrm{R}} \exp\left(-\frac{(E-\mu_{\mathrm{R}})^2}{2\sigma_{\mathrm{R}}^2}\right) &+ N_{\mathrm{L}} \exp\left(-\frac{(E-\mu_{\mathrm{L}})^2}{2\sigma_{\mathrm{L}}^2}\right) \\ \mu_{\mathrm{L}} &\in (0.5\,\mathrm{MeV}, 0.8\,\mathrm{MeV}) \\ \sigma_{\mathrm{L}} &\in (0.06\,\mathrm{MeV}, 0.2\,\mathrm{MeV}) \end{aligned} \tag{4.5}$$

where index "R" denotes the regular peak and "L" the low energy events. The corresponding fitting range is (0.5 MeV, 1.25 MeV), the lower bound is set so that the low energy events shape can be properly parameterized. This fitting function and range will be used for U-series throughout the rest of this thesis, unless stated otherwise.

4.4.6 Summary

The U-series correlated decays are present in the whole AD (unlike Th-series and Ac-series), which makes its delayed alpha signal a promising candidate for detector nonuniformity study and new updated correction construction. The pseudorate of U-series is mostly stable with occasional spikes, which are unlikely to affect the measurement of nonuniformity. The same applies for higher rate in EH2-AD2. Low energy events can in principle affect the nonuniformity measurement. Their origin is not fully understood and while the double Gaussian function matches the data well, it cannot be ruled out that the low energy events cause a certain bias in the fitted mean of the regular alpha peak in some parts of the ADs. Nevertheless, out of the all the correlated decays available, U-series delayed alpha is still the signal that is the most suitable for nonuniformity study and new correction construction.

4.5 Spallation Neutrons Captures

Spallation neutrons (SPNs) originate in interactions of cosmic-ray muons with ADs and surrounding matter. Similarly to neutrons from IBDs, SPNs also lose energy and eventually get captured on a nucleus. The capture time is about 30 μs

in GdLS and the neutron is most likely captured on gadolinium (nGd) resulting in about 8 MeV released in a gamma cascade (details in Section 3.3.3) and less likely (about 15%) on hydrogen (nH) with 2.2 MeV gamma released. In LS, nH is predominant with capture time of about 200 μs. While these signals were already used in the original nonuniformity correction, their selection and fitting can be improved so that we could obtain more precise results.

The basic idea of the selection we are going to use is to identify suitable showering muons and save the events that follow in a certain time interval while discarding non-physical flasher events. From the selected interval one signal and one background window are carved out. In order to remove events uncorrelated to the incident showering muon, spectrum from the background window is then subtracted from the spectrum corresponding to the signal window in a similar way to what we did with correlated decays in Section 4.1.2

We will employ only one selection common for both nH and nGd with values of the cuts listed in Tab. 4.5. The reasoning behind them is following:

- **Showering muon energy:** Generally speaking, the number of spallation neutrons per muon increases with increasing observed muon energy. The lower bound is thus set in the way that there are not too many muons which produce no spallation neutrons. On the contrary, the upper bound is aimed to reduce the pile-up of spallation neutron captures.

- **Flasher cuts:** Group of cuts commonly used in the Daya Bay experiment aimed to remove events caused by PMTs sparking were applied. More details can be found in Ref. [29].

- **SPN event energy:** Starts at 0.7 MeV where the detection efficiency reaches 100% and ends at 20 MeV thus covering both nH and nGd energy ranges.

- **Signal and background windows:** The signal windows start at 20 μs after the showering muon in order to avoid re-triggering which can happen after large amount of energy is deposited in the AD. Unfortunately, it also means loss of part of the statistics, more in GdLS with shorter neutron capture time and less in LS. On the other hand, the signal window ending at 500 μs covers over 80% of nH captures in the LS.

Table 4.5: Selection cuts for SPN nGd and nH

Shower muon energy	(1.8 GeV, 2.5 GeV)
Flasher cuts	$(\text{Quadrant})^2 + (\frac{Q_{\max}}{0.45})^2 > 1$ $4 \times (1 - \text{time\_PSD})^2 + 1.8 \times (1 - \text{time\_PSD1})^2 > 1$ $Q_{\text{2inch}} > 100$ p.e. time_psd_local_rms>13 & kurtosis>1 & flasher_ring=8 flasher_column ≥ 4 & flasher_column ≤ 9
SPN event energy	(0.7 MeV, 12.0 MeV)
Signal window	(20 μs, 500 μs)
Background window	(520 μs, 1000 μs)

As a result of the selection and background subtraction, we obtain an energy spectrum which contains both nGd and nH peaks. However, the number of nH and nGd events we obtain varies greatly between the halls – there is far lower muon rate in the deepest EH3 than in the shallower near halls as summarized in Tab. 3.1.

4.5.1 SPN nH

SPN nH signal is caused by a single 2.2 MeV gamma. While the nH capture itself is dominant in LS and occurs only for about 15% of neutrons in GdLS, the spatial distribution illustrated in Fig. 4.33 is also affected by the fact that the gamma can travel a certain distance before depositing its energy, which mostly manifests in some gammas from nH captures in LS depositing energy in GdLS and being reconstructed there. Therefore we do not see a sharp GdLS-LS boundary in the spatial distribution, but just continuous decrease from LS to GdLS. Either way, there is statistics available for nonuniformity study and new correction creation in the whole AD.

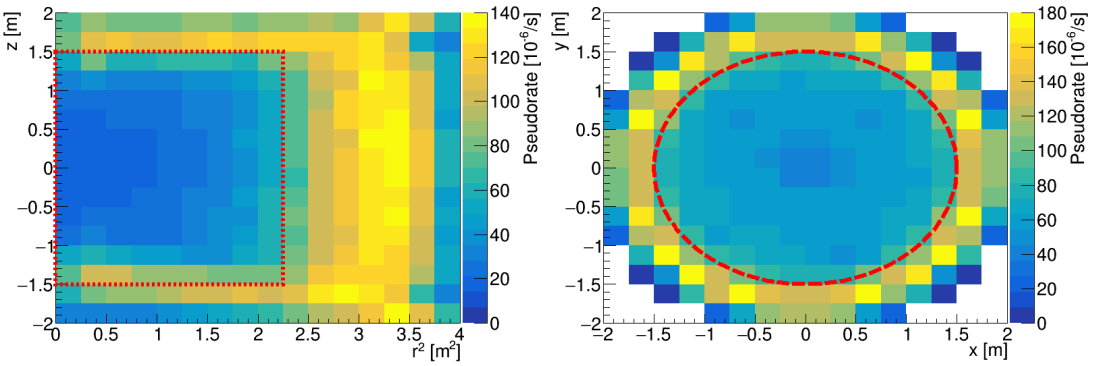

Figure 4.33: Example of a spatial distribution of SPN nH events obtained by EH1-AD1 with cuts listed in Tab. 4.5 and additional energy cut (1.1 MeV, 2.7 MeV). Note that while most events are reconstructed in LS, there is also significant number in GdLS.

The fitting function for the nH peak is the same as the one used for the original correction – DYB function with one exponential and one constant tail described by Eqs. 3.13 and 3.14. The difference is in the fitting range – while in the original correction (1.9 MeV, 3.0 MeV) was used, now we will change it to (1.1 MeV, 2.7 MeV). The reason is that given the number of free parameters the fitting may not converge properly on too narrow range resulting in the issues illustrated in Fig. 3.7. Extended fitting range should mitigate this issue. Two examples of nH DYB fits are shown in Fig. 4.34, one from the middle of the AD, where there is almost no leakage tail, and one from the outer part, where the leakage tail is quite prominent. There is, however, a question whether the DYB function really can fit this wide range of spectral shapes (from no leakage tail to large leakage tail) without introducing a bias into the peak mean value, when we consider the fact that it is still a heuristic function that does not match the data absolutely perfectly. Moreover, there are also SPN nH gammas from MO that can

scatter inside LS depositing part of their energy there further altering the shape of the energy spectrum. Whether or not these issues are a significant problem for the nonuniformity determination is unfortunately impossible to decide without a very precise and detailed Monte Carlo study. While IBD nH toy Monte Carlo is available, its applicability to SPN nH is limited as shown in Ref. [74].

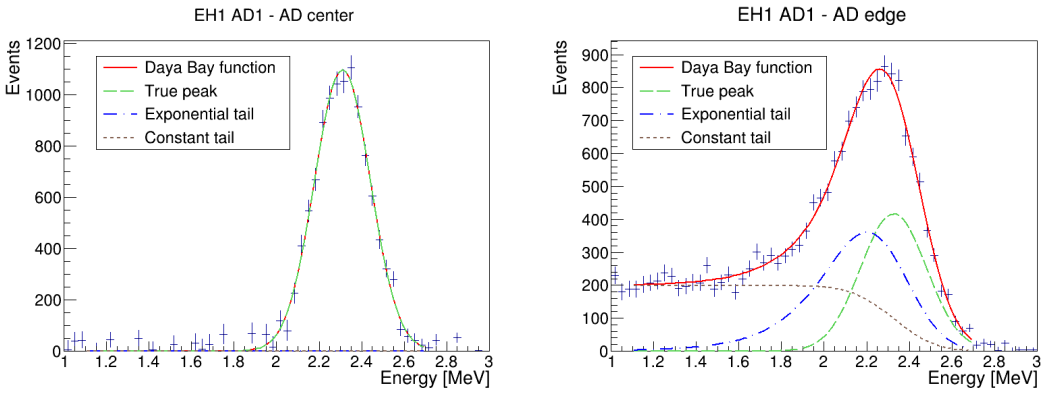

Figure 4.34: Two examples of SPN nH peak obtained by EH1-AD1 selected to highlight the variety of spectral shapes. In the left panel, data from AD center defined as $r^2 < 0.4$ m^2, $|z| < 0.4$ m were used resulting in essentially no leakage tail. The opposite situation is shown in the right panel – there is huge leakage tail around the edge of the AD defined as $r^2 > 3.6$ m^2 and $z > 1.6$ m. The energy spectrum is fitted by the DYB function described by Eqs. 3.13 and 3.14 and the contributions of the true peak and individual leakage tails are drawn in different colors.

4.5.2 SPN nGd

The nGd neutron capture can occur only in GdLS and the following gamma cascade is less likely to have the centroid of the deposited energy far from the nucleus when compared to nH single gamma. The spatial distribution of nGd events is thus restricted to GdLS and its close surroundings as shown in Fig. 4.35. Therefore if we want to use it for creating residual nonuniformity correction, it needs to be combined with another signal that occurs in LS.

The fitting of the SPN nGd we will use is notably different from the one described by Eq. 3.12 in Section 3.3.4. First, double DYB function is used instead of the double CB function as it is better motivated [56]. We will use the version of DYB function that has single exponential tail as a building block and combine two of them

$$\begin{aligned} f_{\text{nGd-map}}(E; N, \alpha, \beta, \mu, \sigma, \lambda) = \\ N_1[\alpha_1 f_{\text{peak}}(E; \mu_1, \sigma_1) + (1-\alpha_1) f_{\text{exp\_tail}}(E; \mu, \sigma_1, \lambda_1)] + \\ N_2[\alpha_2 f_{\text{peak}}(E; \mu_2, \sigma_2) + (1-\alpha_2) f_{\text{exp\_tail}}(E; \mu, \sigma_2, \lambda_2)]. \end{aligned} \quad (4.6)$$

The next step is analogous to what we did with the double CB function – we will keep the same relations between the parameters described by Eqs. 3.5, 3.6,

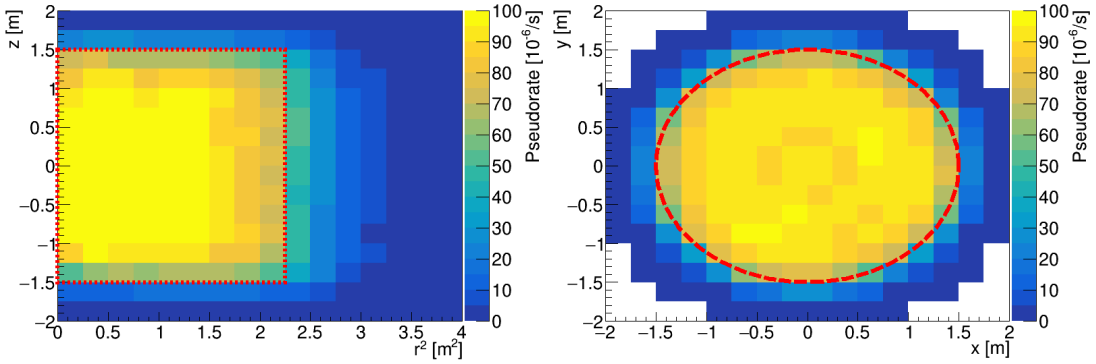

Figure 4.35: Example of a spatial distribution of SPN nGd events obtained by EH1-AD1 with cuts listed in Tab. 4.5 and additional energy cut (6.0 MeV, 10.4 MeV). As expected, the vast majority is reconstructed in GdLS.

3.7. Only instead of double CB-specific Eq. 3.8 we will assume that

$$\alpha_1 = \alpha_2, \ \lambda_1 = \lambda_2, \tag{4.7}$$

but the meaning (ensuring that both peaks have the same shape) is the same. Another difference is that we will drop the decreasing exponential pedestal introduced in Eq. 3.12 as the tail towards higher enegies is greatly reduced by introducing the 2.5 GeV upper bound on the showering muon energy [55]. The change of the fitting function leads to up to 0.7% different fitted peak mean around the edge of GdLS [55]. The fitting range is (6.0 MeV, 10.4 MeV) and two examples of fits are shown in Fig. 4.36.

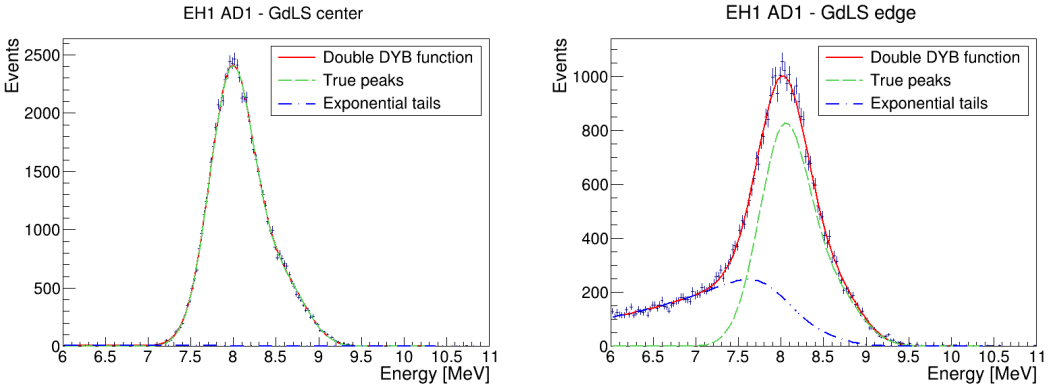

Figure 4.36: Two examples of SPN nGd peak obtained by EH1-AD1. The plot in the left panel corresponds to the AD center ($r^2 < 0.4$ m^2, $|z| < 0.4$ m) where there is virtually no energy leakage. The plot in the right panel comes from the edge of GdLS (2 m^2 $< r^2 <$ 2.4 m^2, -1.6 m $< z <$ −0.8 m) and shows notable leakage tail. The spectra are fitted by the double DYB function (red line) with contributions from true peaks and leakage tails drawn separately in green and blue respectively.

5. Improved Nonuniformity Correction

Now that we have described which signals are available to us and how the energy scale can be determined using them, we can move on to further scrutinize the original nonuniformity correction described in Section 3.3.4 and eventually design a new improved one.

5.1 Azimuthal Nonuniformity

The residual Earth magnetic field (EMF) and the dead PMTs are the main contributors to the azimuthal nonuniformity, both highly AD-dependent. As we use U-series to study azimuthal nonuniformity, there is also the issue of the hotspot under ACU-B described in Section 4.4.3 which needs to be addressed. Each of these phenomena has slightly different properties, hence we will discuss them in the following sections before we design new residual correction.

5.1.1 EMF Effect

In the fully reconstructed data, the EMF effect is partially addressed by the correction described in Section 3.3.4. This actually hinders further analysis as we would like to measure the EMF effect without any distortion added later on. There are two ways to get around this problem:

- The azimuthal correction can be simply retracted by applying the inverse correction factor, i.e. by multiplying the reconstructed energy E by factor from Eq. 3.10 we get energy with retracted azimuthal correction E_{RAC}

$$E_{\mathrm{RAC}} = E \times f_{\phi}^{-1}(\phi) = E \times [1 + \alpha_0^{\mathrm{corr}} \sin(\phi - \phi_0^{\mathrm{corr}})]. \tag{5.1}$$

  The parameters α_0^{corr} and ϕ_0^{corr} are listed in Tab. 3.2. This way, the full EMF effect should be restored so that it can be properly parameterized. The r^2-z and time-dependent corrections are kept resulting in what we will call *semi-corrected energy* E_{RAC}.

- *Raw energy* with no nonuniformity correction applied at all can be also used. However, as we would like to eventually construct new residual correction on top of the regular energy, going back to raw energy is a bit overkill, but it is still very useful to check the findings obtained with the semi-corrected energy.

We thus have three data-sets available using different energy variables – regular, semi-corrected and raw energy – to study the EMF effect and particularly what its dependence on r is[1].

[1]It should be noted that the EMF effect has minimal impact on the incident particle and scintillation light production, but it affects the collection efficiency of PMTs depending on their orientation with respect to the EMF. The reconstructed energy is thus dependent on how

In order to study the r-dependence of the EMF effect, ADs were divided into 4 concentric rings (defined as intervals in r^2 coordinate) of similar volumes. Each concentric ring was then divided further into 12 bins along the ϕ-axis, thus dividing each AD into 48 voxels in total. U-series delayed energy spectrum for each voxel was obtained, fitted according to Section 4.4.5 and relative energy deviation E_{rel} of each voxel from the ring's average was calculated. This way we obtain $E_{\text{rel}}(\phi)$ dependence for each ring, examples of which are shown in the left panels of Fig. 5.1. EH3-AD2 before March 2017 was picked as the example because the EMF effect there is strong and at the same time it was minimally affected by dead PMTs. The $E_{\text{rel}}(\phi)$ dependence is shown for raw energy, semi-corrected energy and regular energy respectively so that clear comparison can be made. As expected, revoking the azimuthal correction (semi-corrected energy in the middle left panel) very well restores the full EMF effect seen with raw energy in the top left panel. Meanwhile, it is apparent that the EMF effect is still present in regular energy after azimuthal correction (bottom left panel), only the amplitude is shifted. The shift is generally such that the effect is on average canceled in GdLS (as was the aim of the original correction), i.e. the black and red points tend to have similar absolute values but opposite sign.

The $E_{\text{rel}}(\phi)$ dependence was fitted by a function analogous to Eq. 3.10

$$E_{\text{rel}}(\phi) = \alpha \cos(\phi - \phi_0), \tag{5.2}$$

which provides us with the amplitude of the EMF effect α and its orientation ϕ_0 with respect to the AD's coordinate system. This allows us to determine the $\alpha(r)$ dependence, where the average value of r was determined as the average radius of events in each particular ring. As illustrated in the right panels of Fig. 5.1 the dependence can be reasonably well modeled as linear function

$$\alpha(r) = \kappa r + \alpha_0 \tag{5.3}$$

with slope κ and offset α_0. All types of energy variable used have similar slope of $\alpha(r)$ dependence, but the offset differs. For semi-corrected and raw energies, we get the fitted value $\alpha_0 \sim 0$ as expected – there should be no azimuthal nonuniformity right in the center of the detector[2] unless it is artificially introduced by a flawed correction, which is exactly what happened to the regular energy variable. The shape of $\alpha(r)$ dependence remained the same, but all the values shifted by $\alpha_0 \simeq \alpha_0^{\text{corr}}$, where α_0^{corr} is the value used in the original correction. All ADs follow the same principles, but the slope differs quite a lot even between ADs in the same EH. Besides that, in some ADs the effect of dead PMTs interferes with the EMF effect. More examples of the EMF effect and the dependence of its amplitude on radius can be found in Appendix D.

As we have clearly shown, the original azimuthal nonuniformity correction is insufficient and it is thus more practical to revoke it and base the new correction on the semi-corrected energy for which the EMF effect amplitude dependence on radius can be modeled as

$$\alpha(r) = \kappa r \tag{5.4}$$

the scintillation light is distributed between the PMTs. Generally speaking, as r increases, increasing proportion of the scintillation light gets detected by lower number of PMTs. If these PMTs are affected by EMF one way or the other, the impact on the reconstructed energy also gets bigger. Similar argument also applies for dead PMTs.

[2]When $r = 0$ m, changing ϕ does not lead to any actual change of position.

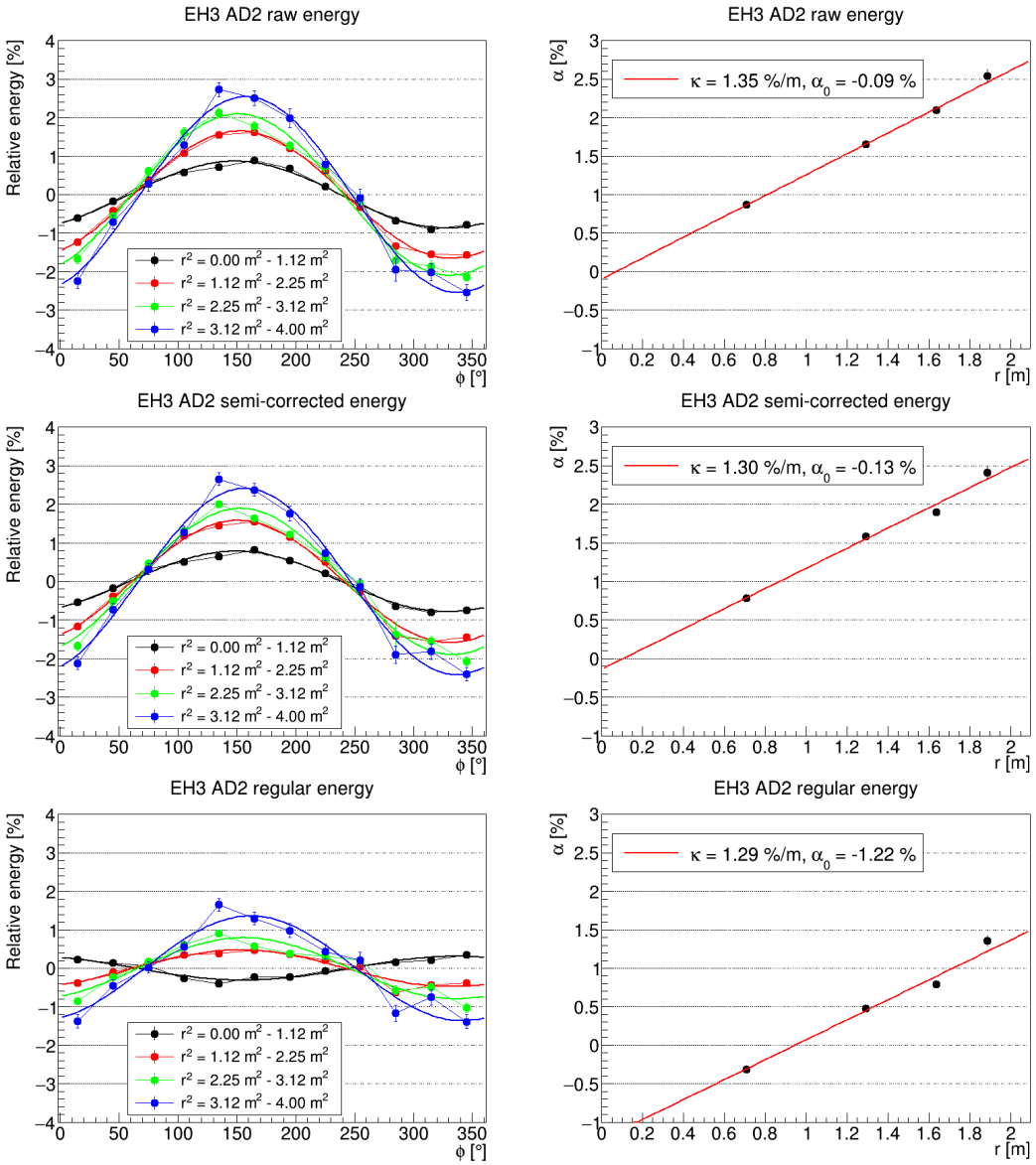

Figure 5.1: The plots to the left show the EMF effect in the dependence of relative energy on the azimuthal angle for 4 concentric rings in EH3-AD2, each ring shown in a different color. The dependence is fitted by $\alpha \cos(\phi - \phi_0)$, where α is the amplitude of the EMF effect in the particular ring. The dependence of the α on radius r is drawn in the plots to the right, showing that it can be modeled by a linear function $\alpha(r) = \kappa r + \alpha_0$. Three rows of plots correspond to different energy variables – raw energy, semi-corrected energy and regular energy respectively. Data from before March 2017 were used in order to avoid the effect of dead PMTs.

without any offset. We will use this variable for the following study of the effect of the dead PMTs and the calibration tube below ACU-B.

5.1.2 Dead PMTs

Next to the EMF effect, dead PMTs have the biggest impact on the azimuthal nonuniformity. The difference is that their effect is more localized, not only in ϕ, but also in z. In order to study and quantify it, ADs were divided into 96 voxels – 8 rows along the z-axis (we will refer to them as "z-rows") with 12 columns along the ϕ-axis each ("ϕ-columns"). As with the EMF effect, U-series delayed energy spectrum in each voxel was fitted and relative energy deviation from the average was calculated. This provides us with ϕ-z nonuniformity map, example of which is shown in Fig. 5.2. Such map reflects both the EMF effect and the dead PMTs effect averaged over r. The calibration tube under ACU-B also plays a role and it will be discussed in the next section. EH2-AD2 was chosen as an example because the EMF effect there is rather weak so that we can focus on the two dead PMTs close to the top. The impact they have on the nonuniformity map is apparent – the reconstructed energy is in general several percent lower in the direction of the dead PMTs. The effect gets bigger with increasing r (i.e. closer to the dead PMT), but the exact form of its dependence on r is difficult to determine due to limited statistics and it seems to vary case by case (probably due to the dead PMT position, interference with other effects or proximity of other dead PMTs if there are any). As we have already established linear dependence on r for the EMF effect, it is reasonable to use the same model for the effect of dead PMTs as that will allow us to create a relatively simple and straightforward correction in ϕ-z coordinates with r-dependence. It may be a bit imprecise, but it will be still a great improvement when compared to the original correction with no r-dependence.

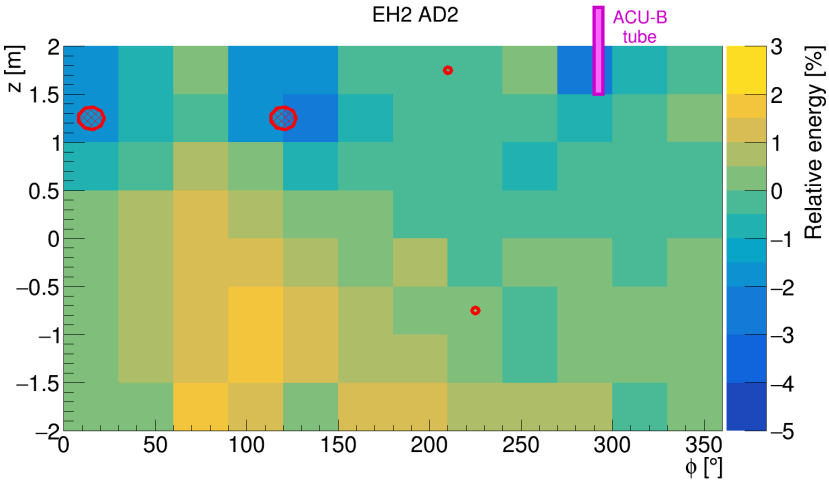

Figure 5.2: Example of a ϕ-z nonuniformity map obtained by EH2-AD2 using U-series semi-corrected energy. It is based on data after December 2016, around which time the two PMTs near the top stopped working. The dead PMTs are marked by red circles with diameter proportional to the fraction of time they were dead in the particular period. Apart from the dead PMTs, we can also see lower energy observed around the ACU-B calibration tube and the EMF effect is also visible in the lower half with maximal energy about 100°, though it is rather weak in EH2-AD2 compared to other ADs.

Table 5.1: AD-specific time division points and number of dead PMTs (not working $> 30\%$ of data-taking).

AD	Time division point	Dead PMTs
EH1-AD1	None	0
EH1-AD2	2018-02-28	1
EH2-AD1	2015-03-31	1
EH2-AD2	2016-12-31	2
EH3-AD1	2016-03-31	0
EH3-AD2	2018-02-28	1
EH3-AD3	2018-03-31	1
EH3-AD4	2015-09-30	2

Another issue with the effect of dead PMTs is that there is a discontinuity in the nonuniformity pattern at the moment the particular PMT stopped working. In theory, when we construct new azimuthal correction, it can be done for the period of PMT working and not working separately, but in practice it is not always possible as minimal length of such period is limited by available statistics. Still, for most ADs the whole duration of the operation can be divided into two parts that reasonably well reflect the PMT dying moments and retain sufficient statistics. The time divisions for each AD are listed in Tab. 5.1. There is an exception of EH1-AD1 which has only one time part as it was decommissioned from Daya Bay experiment already in 2016 and had not suffered from any significant dead PMTs during its operation. The number of PMTs that were not working for more than 30% of Daya Bay operation can be also found in Tab. 5.1.

5.1.3 ACU-B Tube Effect

As illustrated in Fig. 5.2, in EH2-AD2 the reconstructed energy under the ACU-B is lower than in the surrounding areas, region with $\phi \in (270°, 330°)$ and $z \in (1\text{ m},\ 2\text{ m})$ will be considered as affected. A similar effect can be clearly observed in EH1-AD1, EH3-AD1 and EH3-AD2 too, other ADs unfortunately have dead PMTs in the same area or nearby, which makes it difficult to assess what the exact impact of the ACU-B effect is. The question then arises whether this effect is universal for all signal types or unique to the U-series. The evidence suggests that the latter is the case. The connection between U-series spatial distribution, ACU-B and its possible implications have been already discussed in Section 4.4.3. Moreover, there is no indication of ACU-B effect in analogous ϕ-z nonuniformity maps obtained using SPN nH (other signals such as Ac-series, Th-series see no effect either, but their informative value is limited as they occur only in GdLS).

With ACU-B effect being unique to the U-series, it is necessary to address it before creating a new correction intended for use in a neutrino oscillation analysis. Using data from EH2-AD2, EH3-AD1 and EH3-AD2 (EH1-AD1 was omitted due to its short operation) it was found out that the effect is mostly limited to the events with $r \in (0.75\text{ m}, 1.5\text{ m})$, thus we can work with a concentric ring analogous to the ones we used in the study of EMF effect[3]. Within this ring,

[3]ACU-B effect is likely spatially limited even more, but too narrow ring does not have

the effect can be approximately quantified by calculating the difference to the average of neighboring bins[4] as illustrated in Fig. 5.3. Four voxels marked in

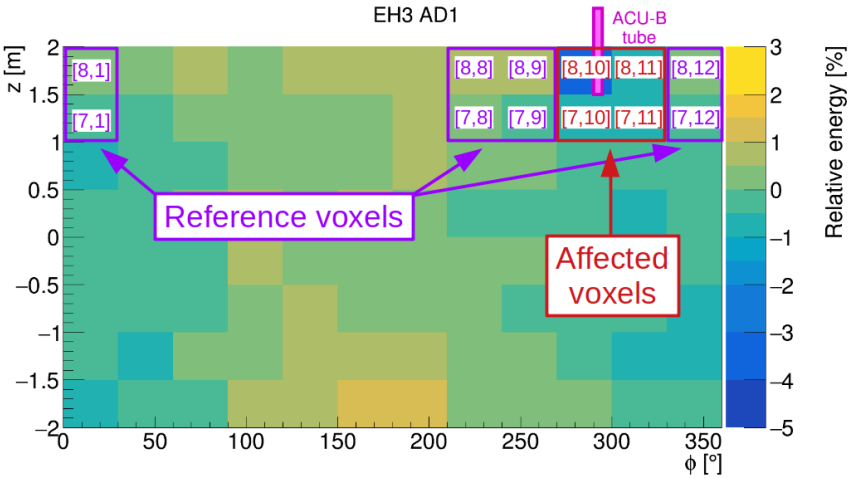

Figure 5.3: The scheme of the calculation of the localized ACU-B correction. The voxels labeled in red are considered as affected (the numbers in square brackets denote to which z-row and ϕ-column the voxel belongs) and average energy deviation from reference voxels in purple is calculated and used for the localized correction. The ϕ-z nonuniformity map was obtained by EH3-AD1 with radial cut $r \in (0.75\text{ m}, 1.5\text{ m})$.

red are considered to be affected by the ACU-B effect – two in the top (8^{th}) z-row and two in the z-row below (7^{th}) and eight voxels marked in purple are considered to be reference ones. Then, for each affected voxel, average difference to the reference voxels in the same row is calculated

$$\begin{aligned} f^k_{\text{ACU}}[i,j] &= E^k_{\text{rel}}[i,j] - \frac{1}{4}(E^k_{\text{rel}}[i,8] + E^k_{\text{rel}}[i,9] + E^k_{\text{rel}}[i,12] + E^k_{\text{rel}}[i,1]), \\ i &\in \{7,8\}, \\ j &\in \{10,11\} \end{aligned} \tag{5.5}$$

where $E^{\text{k}}_{\text{rel}}[i,j]$ stands for relative energy deviation of the voxel in the i-th z-row and j-th ϕ-column in the k-th AD. Next we calculate the localized ACU-B correction factors from the average value from the three ADs in which ACU-B effect is unobscured by dead PMTs

$$\begin{aligned} \tilde{f}_{\text{ACU}}[i,j] &= \frac{1}{3}\sum_k f^k_{\text{ACU}}[i,j], \\ k &\in \{\text{EH2-AD2, EH3-AD1, EH3-AD2}\}, \end{aligned} \tag{5.6}$$

their values are listed in Tab. 5.2. The localized correction factor is then used at the stage of U-series fine data selection for events with $r \in (0.75\,\text{m}, 1.5\,\text{m})$ and

statistics sufficient to study the effect.

[4]This in principle also brings a certain bias coming from the EMF effect, but it is expected to be about one order of magnitude smaller then the ACU-B effect itself, main reason being that the EMF effect is rather weak in EH2-AD2 and EH3-AD1.

the values of z-, ϕ-coordinates corresponding to the voxels affected by ACU-B effect. Reconstructed energy of these events is multiplied by factor

$$1 - \tilde{f}_{\mathrm{ACU}}[i, j]. \tag{5.7}$$

An example of ϕ-z nonuniformity map after the localized ACU-B effect correction is shown in Fig. 5.4.

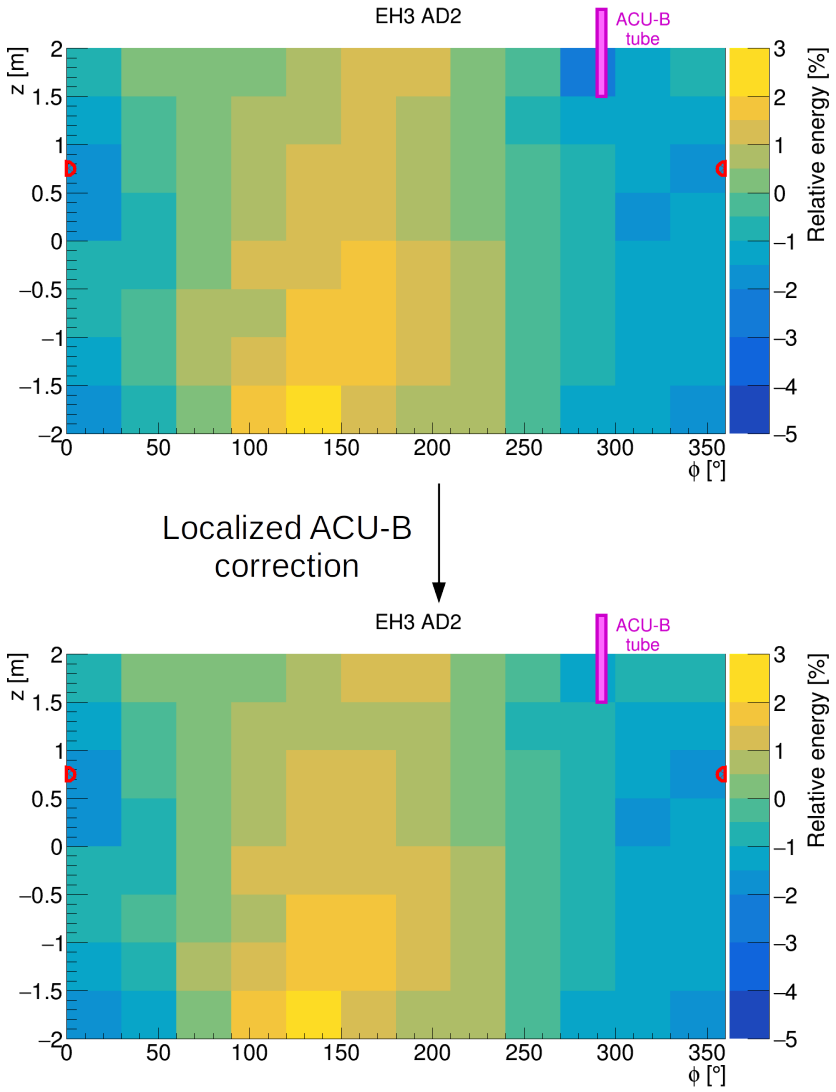

Figure 5.4: Example of the localized ACU-B correction impact on the ϕ-z nonuniformity map. This correction aims to remove the area of low relative energy around the ACU-B tube. Note that the relative energy values in other voxels change too as the AD average is also different after application of the localized ACU-B correction.

5.1.4 Construction of the New Azimuthal Correction

Having discussed all the necessary building blocks for the new azimuthal correction (or ϕ-z(-r-t) correction to be more precise), which will be applied on top

Table 5.2: Values of local ACU-B effect correction factor $\tilde{f}_{\mathrm{ACU}}$

z-row \ ϕ-column	10	11
8	-3.46%	-1.20%
7	-1.02%	-0.42%

of the regular energy already containing the original nonuniformity correction described in Section 3.3.4, we can now proceed to its construction. It has been shown that the ϕ-z nonuniformity maps reflect both EMF and dead PMTs effects. However, there is a question of its normalization. If the values in the voxels were defined as relative U-series delayed peak mean deviation from the whole AD average, then such map would also partially correct the nonuniformity along the z axis, which is something the following residual r^2-z correction is supposed to do. Moreover, the original r^2-z correction has issues, which may lead to shifting of the reconstructed energy of the whole z-row. In such shifted z-row, the simple r-dependence of the EMF and dead PMTs effects described by Eq. 5.4 might not work properly. To avoid this and to keep the new azimuthal and residual r^2-z corrections decoupled, the values in the voxels of the ϕ-z nonuniformity maps are calculated as relative deviation from the z-row average.

The fact that the nonuniformity pattern suddenly changes when a PMT stops working can be mostly addressed by dividing the whole duration of operation into two AD-specific periods – before and after the PMT stopped working. The time division points for each AD are listed in Tab. 5.1. However, they do not always perfectly match the date when the PMT(s) in the particular AD stopped working for one or two reasons. First, we had to always ensure that each time period contains sufficient statistics, i.e. it is not too short. Second, in certain ADs more PMTs died at different times resulting in the division point selected sometime in between if possible.

Regarding the r-dependence of EMF and dead PMTs effects, both can be modeled as directly proportional to r if the original azimuthal correction is revoked. Therefore we start our new correction by applying Eq. 5.1 and creating ϕ-z nonuniformity maps based on E_{RAC} with ACU-B local correction also applied. Such maps have 12 bins along the ϕ-axis, 8 bins along the z-axis (96 voxels in total) and 2 AD-specific time bins. Moreover, when a map is created, each voxel corresponds to a certain average radius r_0, which is determined by the distribution of the U-series events in the AD. This radius is generally z-dependent[5] and it can be easily obtained from data for the 8 z-rows of each map. The ϕ-z(-r-t) correction thus takes a form

$$E_{\mathrm{AZ}}(\phi,z,r,t) = E_{\mathrm{RAC}}(\phi) \times \left(1 - M_{\mathrm{RAC}}(\phi,z,t)\frac{r}{r_0(z)}\right) = \\ = E \times [1 + \alpha_0^{\mathrm{corr}} \sin(\phi - \phi_0^{\mathrm{corr}})] \times \left(1 - M_{\mathrm{RAC}}(\phi,z,t)\frac{r}{r_0(z)}\right), \quad (5.8)$$

where E_{AZ} is the new ϕ-z(-r-t)-corrected energy. To obtain it, the original az-

[5]As there is higher rate of U-series in the GdLS, it pulls the average radius r_0 to smaller values for $|z| < 1.5$ m.

imuthal correction is revoked from the reconstructed energy E, which leads to the semi-corrected energy E_{RAC} defined in Eq 5.1. Using it together with the localized ACU-B correction, we can construct the ϕ-z nonuniformity map M_{RAC} and the associated mean radius $r_0(z)$ for each AD and corresponding time bin individually. M_{RAC} and $r_0(z)$ are then used to correct the azimuthal nonuniformity caused mainly by the EMF and dead PMTs effects.

In practice, radius-normalized maps defined as

$$\tilde{M}_{\mathrm{RAC}}(\phi, z, t) = \frac{M_{\mathrm{RAC}}(\phi, z, t)}{r_0(z)}. \tag{5.9}$$

are used in the correction implementation. All the individual $\tilde{M}_{\mathrm{RAC}}$ maps can be found in Appendix E.

5.2 Residual r^2-z Nonuniformity

After the application of the new azimuthal correction, r^2-z correction also needs to be updated using the new E_{AZ} energy variable (it will be used as the default energy variable throughout this section). Although the new azimuthal correction is constructed in the way that it has only minimal impact on r^2-z nonuniformity, any differences it causes still need to be addressed. Moreover, it is also an opportunity to address the remaining effect of the dead PMTs and the issues of SPN nH and nGd fitting, which likely cause the outlier voxels in the original r^2-z correction.

We can either use a single type of signal in the whole AD or sew the correction map using two types of signals in a similar fashion to the original correction. Regarding the first option, there are only two signals present in the whole AD – U-series delayed alpha and SPN nH. Both signals have issues around the edge of the LS – the desirable simple Gaussian spectrum of U-series alpha signal is polluted by low energy events, while SPN nH leakage tail causes the spectrum to have significantly different shape than in the center, possibly biasing the fitted mean value – but there are no better options. We will thus create the residual correction maps for both signals. In addition to that, we will also sew a third set of maps using SPN nGd in GdLS and U-series alpha in LS.

Regarding the U-series alpha signal, there is also the issue with events under ACU-B, which have lower reconstructed energy as we have seen in Section 5.1.3. There we have concluded that the ACU-B effect is specific for U-series and not present with other types of signal. In order to avoid any bias this might cause, events with $r^2 \in (1.6\ \mathrm{m}^2, 2.4\ \mathrm{m}^2)$, $z \in (1.2\ \mathrm{m}, 2.0\ \mathrm{m})$, $\phi \in (270°, 300°)$ will not be used for U-series alpha r^2-z residual correction construction.

The new residual r^2-z correction will follow the template described in Section 3.3.4. The same division of AD into 10 uniform bins along r^2-axis and 10 uniform bins along z-axis creating in total 100 voxles will be used. The energy spectrum in each voxel is fitted by functions described in Sections 4.4.5, 4.5.1 and 4.5.2. This way we obtain absolute energy maps. As a next step we need to transform them to relative energy maps, which we will use in the correction. However, there is question of which benchmark should be used for the relative energy calculation. As the overall energy scale is calibrated by nGd, we would

like the SPN nGd peak mean not to be shifted by the correction. Since SPN nGd occurs only in GdLS, we will use the average of the voxels covering GdLS (i.e. voxels with $r^2 < 2.4$ m^2 and $|z| < 1.6$ m) and calculate the relative energy with respect to this average. This way we obtain relative energy maps, examples of which are shown in Fig. 5.5. The first one is map based on U-series delayed alpha, the second one on SPN nH and the third is sewn from SPN nGd in GdLS (and little part of LS) inside the red dashed line and U-series alpha in LS outside of it. The red dashed line also marks the area in respect to which the relative energy deviation is calulated. When we look at the map based on SPN nH, it is apparent that there are several outliers, which are not present in the U-series alpha map. This is likely caused by the SPN nH fitting and it only confirms the fact that the convergence of the SPN nH fit is quite hard and unstable. It is especially notable in EH3, where the SPN nH (and nGd) statistics is much lower than in the near halls. All the r^2-z residual correction maps can be found in Appendix E.

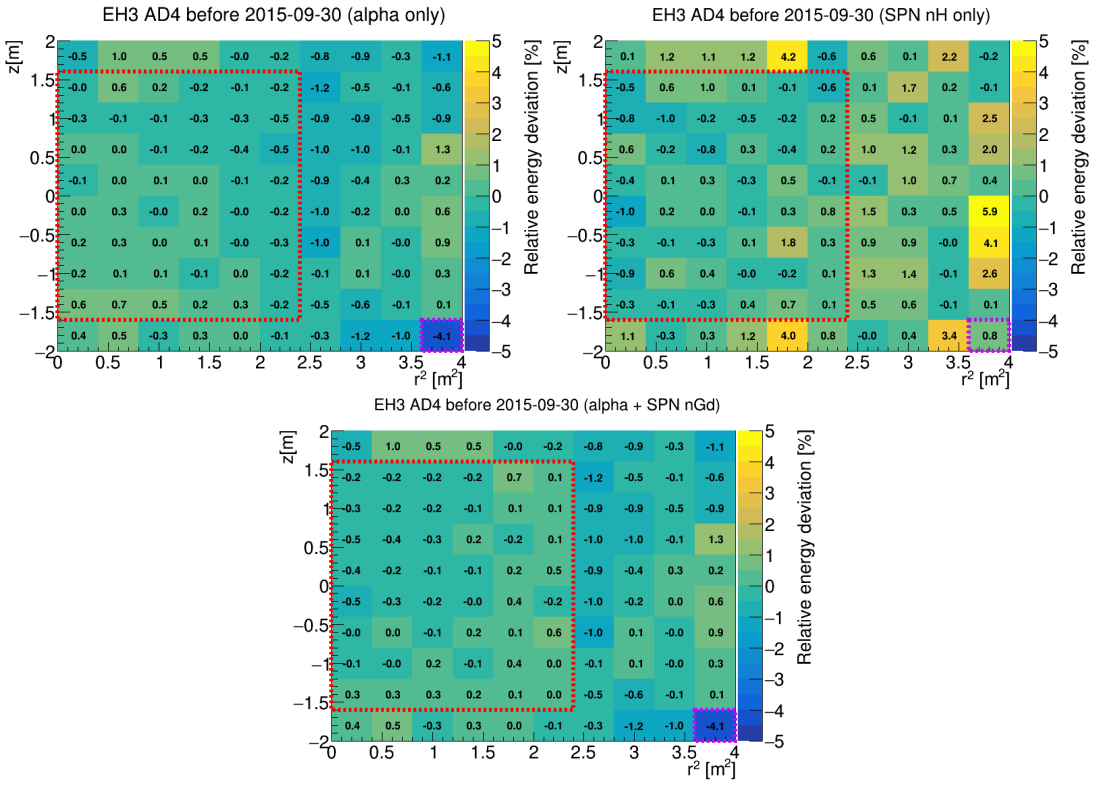

Figure 5.5: Comparison of EH3-AD4 r^2-z residual nonuniformity correction maps based on U-series alpha (top left), SPN nH (top right) and combination of SPN nGd and U-series alpha (bottom), where alpha signal is used in the volume outside the dashed red rectangle (and the values are thus the same as in the top left U-series alpha map) and SPN nGd in the volume inside (mostly GdLS). In the bottom right of each map, the outlier voxel from the original correction shown in Fig. 3.7 is highlighted by violet dashed line.

5.2.1 Outlier Voxels of the Original Correction

EH3-AD4 (first time period) was selected as an example in Fig. 5.5 because there is the most significant outlier of the original correction – the voxel in the bottom right part of the AD map in the right panel of Fig. 3.7, where the correction factor differs on average by 6.6% from the other ADs. Our suspicion was that this outlier in the original map was caused by the fitting unable to find a correct value. The maps shown in Fig. 5.5 show residual nonuniformity after the application of the original correction (and the new azimuthal one, which does not have much effect here) and if we look at the maps based on U-series (top left and bottom panel), we see 4% lower energy in this particular voxel. This fact only supports the suspicion that EH3-AD4 does not really differ that much from other ADs regarding this voxel, but instead fitting used in original correction construction is responsible for the outlier. On the other hand, the new SPN-nH map is quite consistent with the original correction in this particular voxel. This either means that our new SPN nH fitting suffered from the same issues as the original one in this particular voxel or the difference is caused by the different properties of gamma and alpha signal in general. However, the fact that the U-series alpha maps are much more consistent with the original correction in all other ADs when it comes to this particular voxel suggests that the former is the case.

Regarding other outliers in the original maps, we can generally observe corresponding relative energy deviations to a certain extent in all the residual nonuniformity maps, confirming they are a result of wrong fitting (or other issues) of the original correction.

5.2.2 Effect of the Dead PMTs

The effect of the dead PMTs can be also observed in the residual r^2-z nonuniformity maps[6], especially in EH2-AD2, where two PMTs stopped working just before or soon after the end of 2016. The EH2-AD2 maps are shown in Fig. 5.6, where we can compare the situation before and after the end of 2016 as observed using U-series alpha and SPN nH signals. The effect of dead PMTs in other ADs is not as noticeable, but still observable in EH1-AD2, EH2-AD1, EH3-AD2, EH3-AD3 and EH3-AD4 when using the U-series alpha signal. It manifests as lower reconstructed energy in the voxels with z coordinate close to the dead PMT and $r^2 > 2.5$ m^2. When using SPN nH, the effect is less pronounced and in certain ADs not even clearly observable.

The presence of the dead PMTs effect (especially when using U-series alpha signal) makes it clear that the original correction is not sufficient as it lacks the ability to address this localized, strongly time-dependent effect.

[6]Note that the previously applied new azimuthal correction has not completely removed the effect of the dead PMTs, it only leveled the effect along the AD perimeter. If before the new azimuthal correction there was significantly lower energy around the ϕ and z coordinates corresponding to a dead PMT, after this correction, there is slightly lower energy uniformly distributed along the whole perimeter. To get it to the nominal level, subsequent (residual) r^2-z correction is necessary.

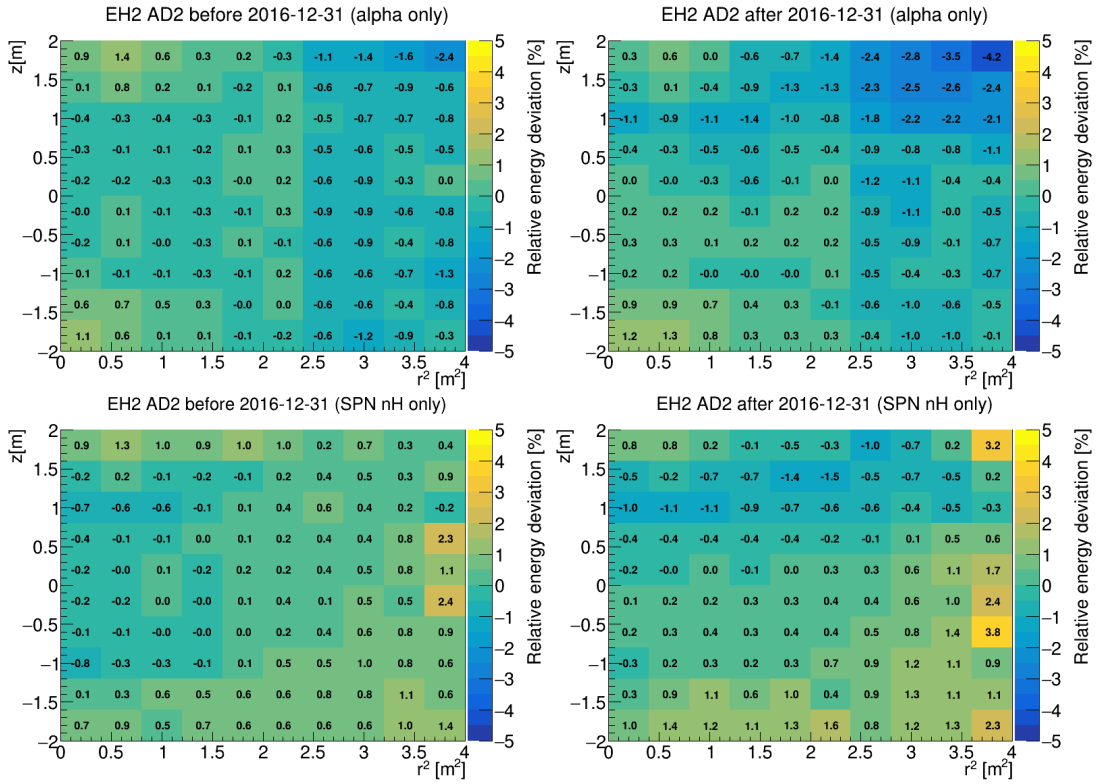

Figure 5.6: EH2-AD2 r^2-z residual nonuniformity maps based on U-series alpha signal (top row) and SPN nH (bottom row) for period before the end of 2016 (left plots) and after (right plots). The effect of two dead PMTs, which stopped working around the end of 2016, can be clearly seen as lower reconstructed energy in the upper outer part of the AD with alpha signal. It is also visible with SPN nH, but much less distinctively. This is likely caused by the fact that while the alpha particle energy deposition is essentially point-like, nH gamma tends to scatter and deposit its energy over a certain volume.

5.2.3 Comparison of Maps Based on Different Signals

The r^2-z residual nonuniformity maps based on various signals do not differ only in the areas affected by dead PMTs or in outlier voxels but also in general features. It is clearly visible when we make average maps of all ADs, which are shown in Fig. 5.7. First time period was used as we would like to see and compare features unrelated to the dead PMTs.

Starting our comparison in GdLS, the differences between the signals are not that big there. The average maps based on U-series alpha signal and SPN nGd (from the map sewn from SPN nGd & U-series alpha) are mostly within 0.5% from each other as well as from the original correction. The average SPN nH map is slightly more different, but still within 1% from every other one.

The situation changes once we move to LS. First, there is no SPN nGd available there. Second, the difference between U-series alpha and SPN nH maps (as well as between them and the original) is notably higher than in GdLS. The values in the SPN nH map are on average ∼1.5% higher than in the U-series alpha map and generally tend to continuously increase as we get closer to the edge. In

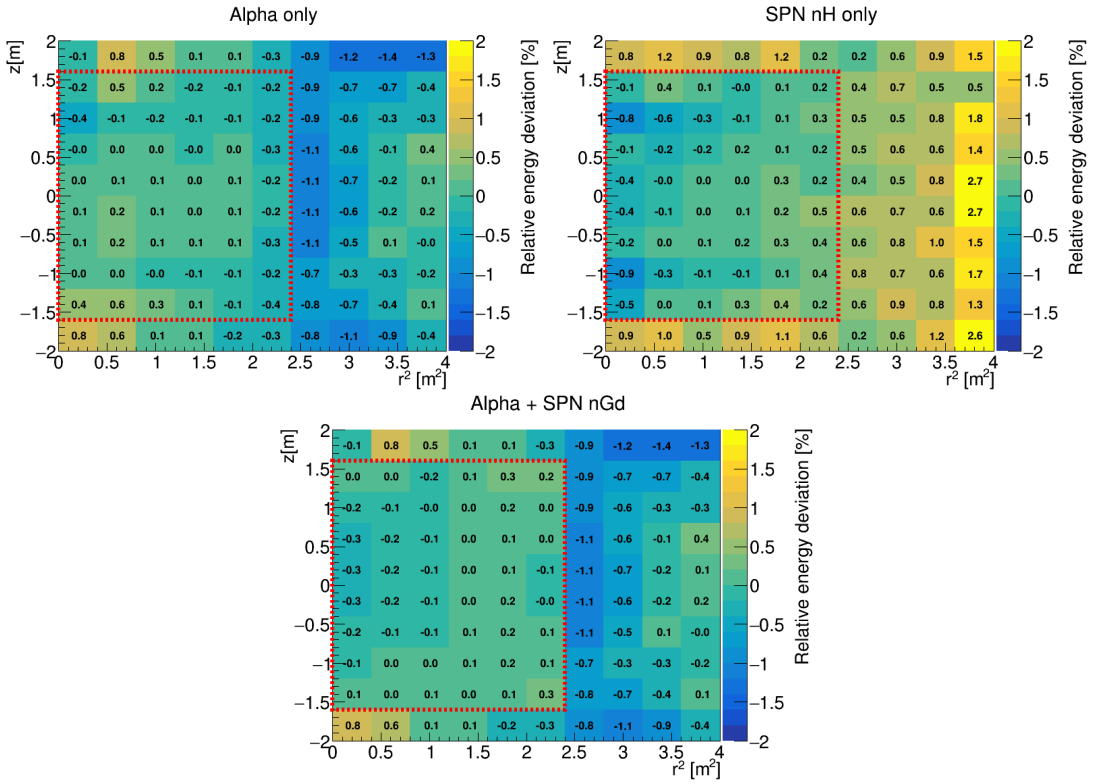

Figure 5.7: Comparison of r^2-z residual nonuniformity maps based on various signals. The maps are averaged over all ADs. The top left map is based on U-series alpha, the top right one on SPN nH and the bottom one is sewn from SPN nGd and U-series alpha. The red dashed rectangle approximately corresponds to GdLS. The top right and bottom maps are identical outside the red dashed line as they both use U-series alpha signal there. Note that the z-axis range is narrower than in previous plots.

contrast to that, with U-series alpha the area of lowest observed energy is located just outside the GdLS ($2.4\ \mathrm{m}^2 < r^2 < 2.8\ \mathrm{m}^2$) and at the top (and to lesser extent bottom) in the outer part of the AD ($|z| > 1.6$ m, $r^2 > 2.4\ \mathrm{m}^2$). The outer top of the AD is also the place where most low energy events occur as described in Section 4.4.4. However, there is no clear correlation between the rate of low energy events in an AD and relative energy deviation in the outer top area of ther^2-z residual nonuniformity maps. The low energy events are thus unlikely to be a major cause of this energy deviation. Moreover, the U-series alpha spectrum is purely Gaussian for $r^2 \sim 2.5\ \mathrm{m}^2$ and $z \sim 0$ m and we observe lower reconstructed energy there nonetheless. Hence, we can conclude that the differences between U-series alpha and SPN nH maps are for the most part not caused by U-series low energy events, but by different energy deposition (point-like alpha vs. scattering gamma) or by bias of SPN nH fitting stemming from the interplay of the main peak and the two tails.

5.2.4 Construction of the Residual r^2-z Correction

With the residual nonuniformity maps at hand, we can proceed to the construction of the r^2-z part of our improved correction, building on top of the new azimuthal correction that was presented in the previous section

$$E_{\mathrm{NC}}(\phi, z, r, t) = E_{\mathrm{AZ}}(\phi, z, r, t) \times \frac{1}{1 + M_{\mathrm{RZ}}(r, z, t)}, \tag{5.10}$$

where E_{NC} is our new fully corrected energy, which takes the new azimuthally corrected energy E_{AZ} and applies a correction factor based on the residual r^2-z nonuniformity map $M_{\mathrm{RZ}}(r, z, t)$. Each voxel of such map contains the relative energy deviation with respect to the GdLS average as we have described above. Following the azimuthal correction, M_{RZ} maps are created for each AD and two AD-specific time periods individually with the time division points listed in Tab. 5.1. All the maps can be found in Appendix E.

5.3 Improved Correction – Summary

Even though the whole procedure described above might seem complicated, lot of what we went through are technicalities of correction maps creation and properties. The application of the improved correction is rather straightforward and it can be summarized in three steps:

1. Revoking of the original azimuthal correction.
2. Application of the ϕ-z correction map with built-in r-dependence and time dependence reflected by two time periods. The map is based on U-series alpha signal.
3. Application of the residual r^2-z correction map, which is also time period dependent. Three sets of maps are available based on various signals – U-series alpha, SPN nH and combination of SPN nGd and U-series alpha.

All the correction factors, maps and time periods are AD-dependent.

As there are three options to choose from regarding the residual r^2-z correction, the question arises as to which one should be primarily used. Since the differences between the U-series alpha, SPN nH and SPN nGd & U-series alpha correction maps can be directly or indirectly attributed to the nature of used signals and their energy deposition, we can assume that the preferable correction map is the one that is based on a signal behaving in a similar way to the one we would like to study. If we want to use gamma signals in our study, then SPN nH or SPN nGd map in GdLS might be appropriate. If we are interested in signal with more or less point-like energy deposition like alpha or beta, then U-series alpha map is more suitable. However, in most analyses we would like to precisely determine the energy of an incident neutrino, which is related to the energy of positron created in IBD interaction by Eq. 2.2. The positron does not match neatly to any of our categories. It first deposits its kinetic energy in a very localized manner (across units of mm) and then annihilates creating two gammas of 511 keV each. This is all detected simultaneously as IBD prompt signal. When measuring reactor antineutrinos, the maximum of the corresponding

prompt spectrum is at about 3 MeV, which means that ∼2 MeV is deposited locally and ∼1 MeV via 511 keV gammas. Moreover, the mean free path of a 511 keV gamma is about two times shorter than that of 2.2 MeV gamma used in SPN nH map construction, which also leads to more localized energy deposition. Overall, we would expect the U-series alpha correction map to be better suited for the measurement of reactor antineutrino energy.

5.4 Impact of the Improved Nonuniformity Correction

The importance of the improved nonuniformity correction varies greatly depending on which part of the AD is used. We can thus expect that the impact of the correction choice will differ greatly between the neutrino oscillation analysis based on nGd and the one based on nH. The reason is that the IBD interactions resulting in nGd occur predominantly in GdLS, where our improved correction introduces only small changes and the original one works very well. In contrast to that, IBD interaction resulting in nH occur mostly in LS, where the improved correction has much more significant impact.

Regarding the choice of the correction version, three options have been employed in the aforementioned analyses – the original correction, the improved correction with r^2-z map sewed using SPN nGd signal in GdLS and U-series alpha in LS and and the one with r^2-z map based on U-series alpha only. For simplicity, we will refer to them as "original correction", "SPN nGd & alpha correction" and "alpha only correction" respectively. However, it should be noted that only simplified versions of the improved correction were employed in the oscillation analysis based on nGd as we will discuss in the following section.

5.4.1 Neutrino Oscillation Analysis Based on nGd

Since the impact of the improved correction (compared to the original one) on the oscillation analysis based on nGd was expected to be minimal, only a simplified version of the improved correction was used. This version did not include the azimuthal part and thus had only one step – application of the r^2-z residual correction map. These maps were based on regular reconstructed energy and used fixed time division point (2017-03-31) for all ADs, otherwise their construction followed the exact same procedure as in the full version of the improved correction described in this thesis.

There were several groups working on the oscillation analysis. We overview preliminary results from two of them to assess the impact of the simplified version of the improved correction. These preliminary results are listed in Tab. 5.3 based on Ref. [75] (Analysis 1) and Ref. [76] (Analysis 2). As we can see there, the impact of the correction choice on this analysis is minimal – the values of $\sin^2 2\theta_{13}$ and Δm^2_{32} parameters remain almost unchanged. While there is an improvement of χ^2_{min}/NDF in certain configurations, there is no strong preference for either correction.

It was eventually decided that the simplified SPN nGd & alpha correction would be applied in the final neutrino oscillation analysis based nGd described

Table 5.3: Preliminary results from the neutrino oscillation analysis based on nGd for several nonuniformity correction options. Values taken from Ref. [75] (Analysis 1) and Ref. [76] (Analysis 2). Normal mass ordering was assumed.

	Correction	$\sin^2 2\theta_{13}$	$\Delta m^2_{32}[10^{-3}\,\mathrm{eV}^2]$	χ^2_{min}/NDF
Analysis 1	Original	0.0858 ± 0.0023	2.430 ± 0.056	613.1/578
	nGd & alpha	0.0857 ± 0.0023	2.433 ± 0.056	606.6/578
	Alpha only	0.0856 ± 0.0023	2.433 ± 0.056	620.0/578
Analysis 2	Original	0.0854 ± 0.0025	2.465 ± 0.053	559.5/541
	nGd & alpha	0.0855 ± 0.0025	2.465 ± 0.053	565.4/541
	Alpha only	0.0854 ± 0.0025	2.465 ± 0.053	543.2/541

in Ref. [3].

5.4.2 Neutrino Oscillation Analysis Based on nH

The application of the improved nonuniformity correction grows significantly in importance when events occurring predominantly in LS are analyzed as in the neutrino oscillation analysis based on nH.

While performing the whole oscillation analysis would be far beyond the scope of this thesis, we can still take a look at what is being observed in the currently ongoing oscillation analysis, which is aimed towards the upcoming publication of neutrino oscillation measurement based on nH. Unlike the previous oscillation analysis published in Ref. [34], which fitted $\sin^2 2\theta_{13}$ using only the measured reactor antineutrino rates across the ADs, this ongoing analysis fits both antineutrino rates and energy spectral shapes to obtain both $\sin^2 2\theta_{13}$ and Δm^2_{ee}. To assess the impact of the improved nonuniformity correction, we will show preliminary results from one of the working groups – collaboration of University of California Berkeley, Charles University, University of California Irvine and Shandong University [77]. This analysis uses data only up to August 2017 and thus avoids the effects of some of the dead PMTs that our improved nonuniformity correction aims to address. Nevertheless, the impact of the improved correction is still considerable.

While the impact of the correction choice on the θ_{13} and Δm^2_{ee} oscillation parameters determination is of the highest interest to us, we will first take a look at some intermediate steps. In Fig. 5.8, we can see examples of the prompt and delayed IBD energy spectra with the three nonuniformity correction options mentioned above. The differences in the prompt spectra, which are related to the incident antineutrino energy by Eq. 2.2, are barely observable (both the SPN nGd & alpha and alpha only corrections cause slight shift towards higher energies), but it is enough to influence θ_{13} and Δm^2_{ee} determination. The right panel shows delayed spectra with the IBD nH peak fitted by the DYB function with one exponential tail, which can be described by Eqs. 3.13 and 3.14.The fitting range is (1.6 MeV, 2.8 MeV). As the nH peak is narrower than the prompt spectrum, the impact of the correction choice there is more notable. Moreover, using parameters obtained from the fit, we can compare how IBD nH peak mean and energy resolution are affected across the ADs.

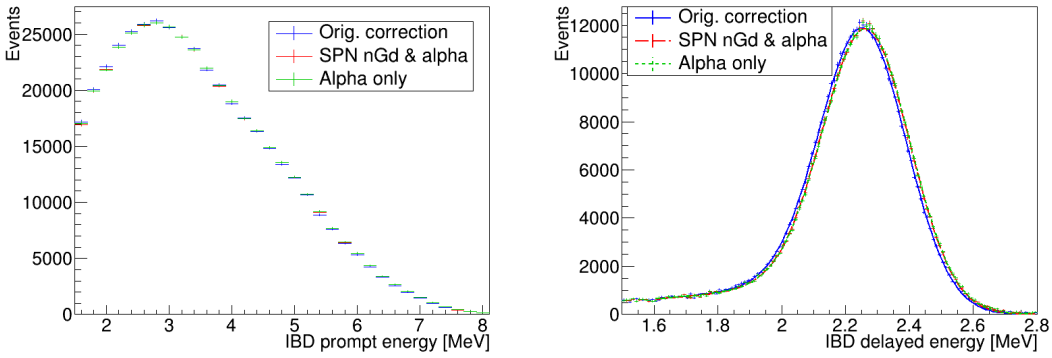

Figure 5.8: Comparison of prompt (left) and delayed (right) IBD spectra obtained using three nonuniformity correction options. The delayed spectrum is fitted by the DYB function with one exponential tail. EH2-AD1 was chosen for this example as it has the highest statistics of all ADs.

In Fig. 5.9 we can observe how the IBD nH peak mean and energy resolution change depending on the nonuniformity correction choice. Generally speaking, both the SPN nGd & alpha and alpha only corrections have similar effect of shifting the IBD nH peak to higher energies when compared to the original correction. Depending on the AD, the shift is 0.15% – 0.45% with the SPN nGd & alpha correction and 0.2% – 0.5% with the alpha only correction. Both of these correction versions also lead to a slightly better alignment between the ADs meaning that the means of the IBD peaks are within 0.45%, 0.40% and 0.35% bands around the average using the original, alpha only and SPN nGd & alpha corrections respectively. Ideally, we would observe the same IBD nH peak mean for all ADs.

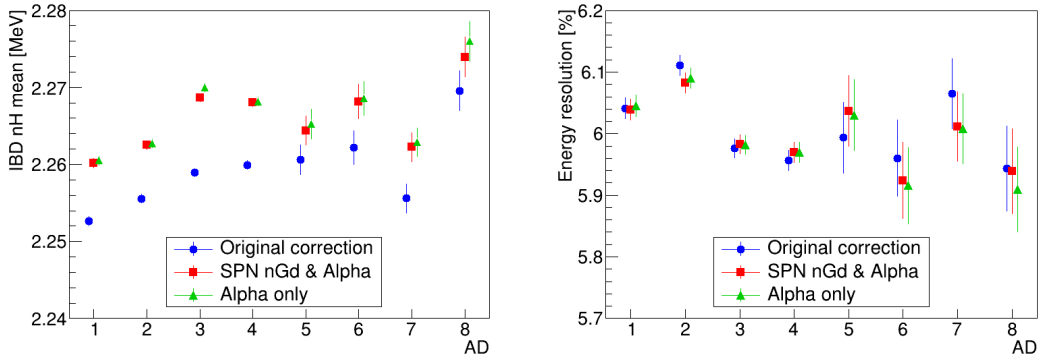

Figure 5.9: IBD nH peak mean (left) and energy resolution (right) for three nonuniformity correction options. The values were obtained by fitting of the IBD nH peak, example of which is shown in Fig. 5.8.

Regarding the energy resolution of IBD nH peak shown in the right panel of Fig. 5.9, there is no unambiguous trend. Both the SPN nGd & alpha and alpha only corrections tend to improve the energy resolution more often than worsen it, the alpha only correction performs mostly better than the SPN nGd & alpha correction in this regard. However, in several ADs, the original correction per-

Table 5.4: Preliminary results from neutrino oscillation analysis based on nH for three nonuniformity correction options.

Correction	$\sin^2 2\theta_{13}$	Δm^2_{ee} $[10^{-3}\,\mathrm{eV}^2]$	χ^2_{min}/NDF
Original	$0.0807^{+0.0073}_{-0.0075}$	$2.65^{+0.13}_{-0.14}$	147.5/134
SPN nGd & alpha	$0.0785^{+0.0075}_{-0.0073}$	$2.66^{+0.14}_{-0.14}$	154.6/134
Alpha only	$0.0824^{+0.0072}_{-0.0072}$	$2.66^{+0.13}_{-0.14}$	149.4/134

forms the best. While we would like to see the SPN nGd & alpha and alpha only corrections to improve the energy resolution in all ADs, we need to keep in mind that both of these corrections use the r^2-z map based on U-series alpha in LS, while the original correction uses SPN nH, i.e. signal of the same nature as IBD nH. The alpha signal was chosen for r^2-z map construction to provide the best correction not for the delayed IBD nH signal, but for the prompt positron signal, which is more similar in energy deposition to alphas than to gammas. While the delayed signal serves to identify the IBD events, the prompt spectrum actually goes to the fitter and is responsible for $\sin^2 2\theta_{13}$ and Δm^2_{ee} determination. Unfortunately, as the energy spectrum of observed reactor antineutrinos is continuous with complicated shape, we cannot easily fit the prompt spectrum and determine how the choice of the correction affects the energy resolution there.

The impact of the correction choice on $\sin^2 2\theta_{13}$ and Δm^2_{ee} determination is summarized in Tab. 5.4. The relative uncertainty of $\sin^2 2\theta_{13}$ measurement is 9%. While the SPN nGd & alpha correction causes almost 3% decrease in the measured value (about a third of the uncertainty) when compared to the original correction, the alpha only correction causes the value of $\sin^2 2\theta_{13}$ to increase by little over 2% (about a quarter of the uncertainty). This is rather surprising as we would expect both these corrections to influence the $\sin^2 2\theta_{13}$ fitting in the same direction, because they have similar impact on the prompt and delayed spectra. However, we need to keep in mind that the results are still preliminary. Either way, while the choice of the correction version does not completely change the $\sin^2 2\theta_{13}$ value, its impact is certainly not negligible.

The relative uncertainty of Δm^2_{ee} determination is about 5% and the impact of the correction choice there is smaller than for $\sin^2 2\theta_{13}$. The SPN nGd & alpha and alpha only corrections cause shift higher of 0.6% and 0.4% respectively when compared to the original correction, which is about a tenth of the uncertainty.

To sum up, the impact of the choice of the nonuniformity correction version on the neutrino oscillation analysis based on nH is considerable unlike the oscillation analysis based on nGd where it was negligible. While the original correction has the lowest χ^2_{min}, we have more confidence in the results obtained with the SPN nGd & alpha and alpha only corrections, because they address several issues of the original correction as we have discussed in this chapter. Moreover, we have seen that the energy resolution of IBD nH peak on average improved despite the fact that both the SPN nGd & alpha and alpha only corrections are tuned for particles with localized energy deposition in LS and not gammas as is the original correction. However, we need to bear in mind that the work on the neutrino oscillation analysis based on nH is still in progress and the results we

have shown are preliminary.

Currently, there is no clear preference between the SPN nGd & alpha and alpha only corrections as to which one will be used for the final result.

Conclusion

The main goal of this work was to scrutinize the detector nonuniformity correction used in the Daya Bay experiment and design a new improved version especially with regards to the neutrino oscillation analysis based on nH which is currently being worked on.

To provide context for this task, we started the thesis by description of the physics behind neutrino oscillation, parameters that drive it and ways these parameters are measured. After that the Daya Bay experiment was introduced along with overview of its detector system and physics program. Next we focused on the procedure of event reconstruction with emphasis on the original nonuniformity correction and especially its shortcomings, most notably the absence of radial dependence in the azimuthal part of the correction and general lack of means to deal with dead PMTs. It was clearly shown that a revision of the original correction is necessary. However, such revision would not be possible without a suitable set of signals that allow us to extract information about the energy scale in various parts of the detectors.

We thus proceeded to a very detailed overview of various types of signals, which I carried out in order to assess their usefulness for a nonuniformity study. While there are several naturally occurring correlated decays, it was found out that their presence is linked to gadolinium doping and thus limited to GdLS, except for U-series with $^{214}\mathrm{Bi} \rightarrow {}^{214}\mathrm{Po} \rightarrow {}^{210}\mathrm{Pb}$ correlated decays, which are also present in LS due to $^{222}\mathrm{Rn}$ contamination. Moreover, its delayed alpha signal has only one significant energy line, so it can be reasonably well fitted. This makes it a promising signal for a nonuniformity study despite some issues which were also investigated and addressed. Besides these correlated decays, gamma signals from captures of spallation neutrons on hydrogen and gadolinium were also checked. Eventually, these three signals (U-series alpha, SPN nH and SPN nGd) were selected to be used for the construction of an improved version of the nonuniformity correction.

Having the signals, their selection and fitting prepared, we moved on to my scrutiny of the original nonuniformity correction and finally creation of an improved one. Starting with the azimuthal part of the correction, it was shown that the original version is inadequate, so much so that the easiest way to deal with it was to completely revoke it and then design a new azimuthal correction essentially from scratch using U-series alpha signal. This new azimuthal correction reflects both the effect of dead PMTs and Earth's magnetic field with data-driven radial dependence included.

The issues of the original r^2-z correction were not that serious, hence it was decided to make a residual correction on top of the current one. Three residual correction maps were constructed based on the aforementioned signals. They reflect the effect of the dead PMTs and also address other issues of the original correction.

At the end of the thesis, I briefly assessed the impact of the improved correction on the neutrino oscillation analyses using data provided by several working groups. The impact was minimal in the analysis based on nGd, because its statistics comes mostly from GdLS, where the differences between the original and the

improved correction are rather small. On the contrary, in the oscillation analysis based on nH, the impact of the improved correction is quite considerable, especially for determination of $\sin^2 2\theta_{13}$. Besides that, the improved correction on average improves the energy resolution of the IBD nH peak.

Eventually, the simplified version of the improved correction based on SPN nGd in GdLS and U-series alpha in LS was employed in the final neutrino oscillation analysis based on nGd. The full versions of the improved correction based on SPN nGd in GdLS and U-series alpha in LS and the one based on U-series alpha in the whole AD are currently employed in the ongoing oscillation analysis based on nH, which fits both antineutrino rates and energy spectral shapes. It has not been yet decided which one will be selected for the final result.

Bibliography

[1] C. L. Cowan et al. Detection of the free neutrino: a confirmation. *Science*, 124(3212):103–104, 1956.

[2] F. P. An et al. Observation of electron-antineutrino disappearance at Daya Bay. *Phys. Rev. Lett.*, 108:171803, Apr 2012.

[3] Daya Bay collaboration et al. Precision measurement of reactor antineutrino oscillation at kilometer-scale baselines by Daya Bay. *arXiv e-prints*, page arXiv:2211.14988, November 2022.

[4] B. Pontecorvo. Mesonium and anti-mesonium. *Sov. Phys. JETP*, 6:429, 1957.

[5] Z. Maki et al. Remarks on the Unified Model of Elementary Particles. *Progress of Theoretical Physics*, 28(5):870–880, 11 1962.

[6] J. N. Bahcall and R. Davis. Solar neutrinos: A scientific puzzle. *Science*, 191(4224):264–267, 1976.

[7] Y. Fukuda et al. Evidence for oscillation of atmospheric neutrinos. *Phys. Rev. Lett.*, 81:1562–1567, Aug 1998.

[8] K. Zuber. *Neutrino Physics*. Second edition. Taylor and Francis Group, Boca Raton, 2012.

[9] C. Giunti and C. W. Kim. *Fundamentals of Neutrino Physics and Astrophysics*. Oxford University Press, 03 2007.

[10] R. Leitner and D. T. *Elementární částice od prvních objevů po současné experimenty*. První vydání. Matfyzpress, Praha, 2012.

[11] P. Zyla et al. Review of Particle Physics. *PTEP*, 2020(8):083C01, 2020. and 2021 update.

[12] JUNO physics and detector. *Progress in Particle and Nuclear Physics*, 123:103927, 2022.

[13] V. Gavrin et al. Solar neutrino results from SAGE. *Nuclear Physics B - Proceedings Supplements*, 91(1):36–43, 2001. Neutrino 2000.

[14] D. Vignaud. The GALLEX solar neutrino experiment. *Nuclear Physics B - Proceedings Supplements*, 60(3):20–29, 1998.

[15] K. Abe et al. Solar Neutrino Measurements in Super-Kamiokande-IV. *Phys. Rev. D*, 94(5):052010, 2016.

[16] Q. R. Ahmad et al. Direct evidence for neutrino flavor transformation from neutral-current interactions in the Sudbury Neutrino Observatory. *Phys. Rev. Lett.*, 89:011301, Jun 2002.

[17] D. D'Angelo et al. Recent Borexino results and prospects for the near future. *EPJ Web Conf.*, 126:02008, 2016.

[18] A. Gando et al. Reactor On-Off Antineutrino Measurement with KamLAND. *Phys. Rev. D*, 88(3):033001, 2013.

[19] M. H. Ahn et al. Measurement of neutrino oscillation by the K2K experiment. *Phys. Rev. D*, 74:072003, Oct 2006.

[20] K. Abe et al. Improved constraints on neutrino mixing from the T2K experiment with 3.13×10^{21} protons on target. *Phys. Rev. D*, 103:112008, Jun 2021.

[21] P. Adamson et al. Precision constraints for three-flavor neutrino oscillations from the full MINOS+ and MINOS dataset. *Phys. Rev. Lett.*, 125:131802, Sep 2020.

[22] M. A. Acero et al. Improved measurement of neutrino oscillation parameters by the NOvA experiment. *Phys. Rev. D*, 106(3):032004, 2022.

[23] H. de Kerret et al. Double Chooz θ_{13} measurement via total neutron capture detection. *Nature Phys.*, 16(5):558–564, 2020.

[24] J. H. Choi et al. Observation of Energy and Baseline Dependent Reactor Antineutrino Disappearance in the RENO Experiment. *Phys. Rev. Lett.*, 116(21):211801, 2016.

[25] K. Abe et al. Constraint on the matter–antimatter symmetry-violating phase in neutrino oscillations. *Nature*, 580(7803):339–344, 2020. [Erratum: Nature 583, E16 (2020)].

[26] B. Abi et al. Long-baseline neutrino oscillation physics potential of the DUNE experiment. *Eur. Phys. J. C*, 80(10):978, 2020.

[27] H.-K. Proto-Collaboration et al. Physics potential of a long-baseline neutrino oscillation experiment using a J-PARC neutrino beam and Hyper-Kamiokande. *Progress of Theoretical and Experimental Physics*, 2015(5), 05 2015. 053C02.

[28] F. An et al. The detector system of the Daya Bay reactor neutrino experiment. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 811:133–161, 2016.

[29] F. P. An et al. Measurement of electron antineutrino oscillation based on 1230 days of operation of the Daya Bay experiment. *Phys. Rev. D*, 95:072006, Apr 2017.

[30] O. Dalager. APS april meeting 2021 - recent results from Daya Bay. Internal Daya Bay document DocDB-12351-v2†, 2021.

[31] F. P. An et al. Spectral measurement of electron antineutrino oscillation amplitude and frequency at Daya Bay. *Phys. Rev. Lett.*, 112:061801, Feb 2014.

[32] F. P. An et al. New measurement of antineutrino oscillation with the full detector configuration at Daya Bay. *Phys. Rev. Lett.*, 115:111802, Sep 2015.

[33] D. Adey et al. Measurement of the electron antineutrino oscillation with 1958 days of operation at Daya Bay. *Phys. Rev. Lett.*, 121:241805, Dec 2018.

[34] F. P. An et al. New measurement of θ_{13} via neutron capture on hydrogen at Daya Bay. *Phys. Rev. D*, 93:072011, Apr 2016.

[35] S. Böser et al. Status of light sterile neutrino searches. *Progress in Particle and Nuclear Physics*, 111:103736, 2020.

[36] A. Aguilar et al. Evidence for neutrino oscillations from the observation of $\overline{\nu}_e$ appearance in a $\overline{\nu}_\mu$ beam. *Phys. Rev. D*, 64:112007, Nov 2001.

[37] A. A. Aguilar-Arevalo et al. Significant excess of electronlike events in the MiniBooNE short-baseline neutrino experiment. *Phys. Rev. Lett.*, 121:221801, Nov 2018.

[38] P. Adamson et al. Improved constraints on sterile neutrino mixing from disappearance searches in the MINOS, MINOS+, Daya Bay, and Bugey-3 experiments. *Phys. Rev. Lett.*, 125:071801, Aug 2020.

[39] S. Gariazzo et al. Updated global 3+1 analysis of short-baseline neutrino oscillations. *Journal of High Energy Physics*, 6:135, 2017.

[40] S. Gariazzo et al. Model-independent $\bar{\nu}_e$ short-baseline oscillations from reactor spectral ratios. *Physics Letters B*, 782:13–21, 2018.

[41] M. Dentler et al. Updated global analysis of neutrino oscillations in the presence of eV-scale sterile neutrinos. *Journal of High Energy Physics*, 8:10, 2018.

[42] B. Achkar et al. Search for neutrino oscillations at 15, 40 and 95 meters from a nuclear power reactor at Bugey. *Nuclear Physics B*, 434(3):503–532, 1995.

[43] P. Adamson et al. Search for sterile neutrinos in MINOS and MINOS+ using a two-detector fit. *Phys. Rev. Lett.*, 122:091803, Mar 2019.

[44] D. Adey et al. Improved measurement of the reactor antineutrino flux at Daya Bay. *Phys. Rev. D*, 100:052004, Sep 2019.

[45] P. Huber. Determination of antineutrino spectra from nuclear reactors. *Phys. Rev. C*, 84:024617, Aug 2011.

[46] T. A. Mueller et al. Improved predictions of reactor antineutrino spectra. *Phys. Rev. C*, 83:054615, May 2011.

[47] F. P. An et al. Evolution of the reactor antineutrino flux and spectrum at Daya Bay. *Phys. Rev. Lett.*, 118:251801, Jun 2017.

[48] D. Adey et al. Extraction of the ^{235}U and ^{239}Pu antineutrino spectra at Daya Bay. *Phys. Rev. Lett.*, 123:111801, Sep 2019.

[49] F. An et al. Joint determination of reactor antineutrino spectra from ^{235}U and ^{239}Pu fission by Daya Bay and PROSPECT. *Physical Review Letters*, 128, 02 2022.

[50] D. Adey et al. A high precision calibration of the nonlinear energy response at Daya Bay. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 940:230–242, 2019.

[51] G. Leinweber et al. Neutron capture and total cross-section measurements and resonance parameters of gadolinium. *Nuclear Science and Engineering*, 154(3):261–279, 2006.

[52] T. Skwarnicki. *A study of the radiative CASCADE transitions between the Upsilon-Prime and Upsilon resonances*. PhD thesis, Cracow, INP, 1986.

[53] C. Marshall. SPN energy scale P21A. Internal Daya Bay document DocDB-12347†, 2021.

[54] Y. Nakajima. Updated estimation of spallation energy scale correction with P13A data. Internal Daya Bay document DocDB-9418†, 2013.

[55] B. Roskovec. Revision of the non-uniformity correction for AdSimple(NL). Internal Daya Bay document DocDB-12272†, 2020.

[56] J.-H. Cheng et al. Determination of the total absorption peak in an electromagnetic calorimeter. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 827:165–170, 2016.

[57] C. Marshall. AdSimple TDNU update. Internal Daya Bay document DocDB-11527†, 2017.

[58] J. B. Birks. Scintillations from organic crystals: Specific fluorescence and relative response to different radiations. *Proceedings of the Physical Society. Section A*, 64(10):874–877, oct 1951.

[59] M. Shamsuzzoha Basunia. Nuclear data sheets for A = 210. *Nuclear Data Sheets*, 121:561–694, 2014.

[60] B. Roskovec and Z. Yu. Residual non-uniformity. Internal Daya Bay document DocDB-8619-v1†, 2013.

[61] D. Adey. Non-uniformity from alpha decays. Internal Daya Bay document DocDB-11495-v1†, 2017.

[62] M. Martin. Nuclear data sheets for A = 208. *Nuclear Data Sheets*, 108(8):1583–1806, 2007.

[63] B. Singh et al. Nuclear data sheets for A = 211. *Nuclear Data Sheets*, 114(6):661–749, 2013.

[64] K. Auranen and E. McCutchan. Nuclear data sheets for A = 212. *Nuclear Data Sheets*, 168:117–267, 2020.

[65] K. Abusaleem. Nuclear data sheets for A = 228. *Nuclear Data Sheets*, 116:163–262, 2014.

[66] S. Singh and B. Singh. Nuclear data sheets for A = 224. *Nuclear Data Sheets*, 130:127–182, 2015.

[67] B. Singh et al. Nuclear data sheets for A = 215. *Nuclear Data Sheets*, 114(12):2023–2078, 2013.

[68] F. Kondev et al. Nuclear data sheets for A = 227. *Nuclear Data Sheets*, 132:257–354, 2016.

[69] S. Zhu and E. McCutchan. Nuclear data sheets for A = 214. *Nuclear Data Sheets*, 175:1–149, 2021.

[70] S. Singh et al. Nuclear data sheets for A = 222. *Nuclear Data Sheets*, 112(11):2851–2886, 2011.

[71] S.-C. Wu. Nuclear data sheets for A = 216. *Nuclear Data Sheets*, 108(5):1057–1092, 2007.

[72] B. Singh et al. Nuclear data sheets for A = 218. *Nuclear Data Sheets*, 160:405–471, 2019.

[73] H. R. Band et al. Daya Bay antineutrino detector gas system. *Journal of Instrumentation*, 7(11):P11029–P11029, nov 2012.

[74] B. Roskovec. Look at Monte Carlo nH leakage tail shape. Internal Daya Bay document DocDB-12349†, 2021.

[75] Z. Chen et al. The most precise measurement of $\sin^2 2\theta_{13}$ using the full Daya Bay data set (IHEP group). Internal Daya Bay document DocDB-12668-v3†, 2022.

[76] Z. Hu et al. SYSU nGd oscillation analysis on the full data set. Internal Daya Bay document DocDB-12667-v4†, 2022.

[77] O. Dalager et al. Berkeley-Czech-Irvine-Shandong nH analysis note. Internal Daya Bay document DocDB-12593-v2†, 2022.

[78] F. An et al. Seasonal variation of the underground cosmic muon flux observed at Daya Bay. *Journal of Cosmology and Astroparticle Physics*, 2018(01):001–001, jan 2018.

[79] S. Bjørnholm et al. Energy levels in 234U populated in the beta-decay of 234Pa(UZ). *Nuclear Physics A*, 118(2):261–301, 1968.

†Internal Daya Bay documents are not publicly displayed, but they are available upon request on my email dohnal@ipnp.mff.cuni.cz.

A. Calculation of Background Subtraction Factors

When performing the fine selection of a particular correlated decay with mean lifetime τ_{S}, it can happen that we would like to subtract not only uncorrelated background, but another correlated signal with mean lifetime τ_{C} too. Such task can be done using three coincidence time windows. If $\tau_{\mathrm{S}} < \tau_{\mathrm{C}}$, then the appropriate configuration is signal – background – signal window. Assuming that the first signal window covers coincidence time interval (t_1, t_2), background window (t_3, t_4) and second signal window (t_5, t_6), with corresponding sets of correlated pairs denoted as S_1, B and S_2 respectively, we get the final data-set F by background subtraction

$$F = S_1 - w_{\mathrm{B}} B + w_{\mathrm{S}} S_2, \tag{A.1}$$

where w_{B} and w_{S} are weighting factors for the background and second signal window respectively. This procedure could be also viewed as background addition for S_2 as there is minimum of the desired correlated signal, but it allows to cancel out the unwanted correlated signal.

The weighting factors in Eq. A.1 are set in the way that both uncorrelated and one unwanted type of correlated signal are subtracted. The uncorrelated signal has flat coincidence time distribution

$$P_{\mathrm{U}}(t) = N_{\mathrm{U}}. \tag{A.2}$$

When we plug it in the Eq.A.1, we get following condition

$$\int_{t_1}^{t_2} P_{\mathrm{U}}(t)\mathrm{d}t - w_{\mathrm{B}} \int_{t_3}^{t_4} P_{\mathrm{U}}(t)\mathrm{d}t + w_{\mathrm{S}} \int_{t_5}^{t_6} P_{\mathrm{U}}(t)\mathrm{d}t = 0. \tag{A.3}$$

For the unwanted correlated signal with exponential coincidence time distribution

$$P_{\mathrm{C}}(t) = N_{\mathrm{C}} \cdot e^{-\frac{t}{\tau_{\mathrm{C}}}} \tag{A.4}$$

we get analogous condition

$$\int_{t_1}^{t_2} P_{\mathrm{C}}(t)\mathrm{d}t - w_{\mathrm{B}} \int_{t_3}^{t_4} P_{\mathrm{C}}(t)\mathrm{d}t + w_{\mathrm{S}} \int_{t_5}^{t_6} P_{\mathrm{C}}(t)\mathrm{d}t = 0. \tag{A.5}$$

Thus, we have two linear Eqs. A.3 and A.5 with two variables w_{B} and w_{S} which can be easily solved giving us values of w_{B} and w_{S} for given signal and background windows (of non-zero length)

$$w_{\mathrm{B}} = \frac{(t_2 - t_1)(e^{-\frac{t_5}{\tau_{\mathrm{C}}}} - e^{-\frac{t_6}{\tau_{\mathrm{C}}}}) - (t_6 - t_5)(e^{-\frac{t_1}{\tau_{\mathrm{C}}}} - e^{-\frac{t_2}{\tau_{\mathrm{C}}}})}{(t_4 - t_3)(e^{-\frac{t_5}{\tau_{\mathrm{C}}}} - e^{-\frac{t_6}{\tau_{\mathrm{C}}}}) - (t_6 - t_5)(e^{-\frac{t_3}{\tau_{\mathrm{C}}}} - e^{-\frac{t_4}{\tau_{\mathrm{C}}}})} \tag{A.6a}$$

$$w_{\mathrm{S}} = \frac{(t_2 - t_1)(e^{-\frac{t_3}{\tau_{\mathrm{C}}}} - e^{-\frac{t_4}{\tau_{\mathrm{C}}}}) - (t_4 - t_3)(e^{-\frac{t_1}{\tau_{\mathrm{C}}}} - e^{-\frac{t_2}{\tau_{\mathrm{C}}}})}{(t_4 - t_3)(e^{-\frac{t_5}{\tau_{\mathrm{C}}}} - e^{-\frac{t_6}{\tau_{\mathrm{C}}}}) - (t_6 - t_5)(e^{-\frac{t_3}{\tau_{\mathrm{C}}}} - e^{-\frac{t_4}{\tau_{\mathrm{C}}}})}. \tag{A.6b}$$

If we want to subtract correlated signal with shorter mean lifetime than that of correlated signal we are interested in ($\tau_{\mathrm{S}} > \tau_{\mathrm{C}}$), we can use an analogous

procedure employing three coincidence time windows with following ordering: background – signal – background. The subtraction then takes form

$$F = -w_{\mathrm{B1}} B_1 + S - w_{\mathrm{B2}} B_2, \tag{A.7}$$

which leads to following formulae for the weighting factors

$$w_{\mathrm{B1}} = \frac{(t_4 - t_3)(e^{-\frac{t_5}{\tau_{\mathrm{C}}}} - e^{-\frac{t_6}{\tau_{\mathrm{C}}}}) - (t_6 - t_5)(e^{-\frac{t_3}{\tau_{\mathrm{C}}}} - e^{-\frac{t_4}{\tau_{\mathrm{C}}}})}{(t_2 - t_1)(e^{-\frac{t_5}{\tau_{\mathrm{C}}}} - e^{-\frac{t_6}{\tau_{\mathrm{C}}}}) - (t_6 - t_5)(e^{-\frac{t_1}{\tau_{\mathrm{C}}}} - e^{-\frac{t_2}{\tau_{\mathrm{C}}}})} \tag{A.8a}$$

$$w_{\mathrm{B2}} = \frac{(t_2 - t_1)(e^{-\frac{t_3}{\tau_{\mathrm{C}}}} - e^{-\frac{t_4}{\tau_{\mathrm{C}}}}) - (t_4 - t_3)(e^{-\frac{t_1}{\tau_{\mathrm{C}}}} - e^{-\frac{t_2}{\tau_{\mathrm{C}}}})}{(t_4 - t_3)(e^{-\frac{t_5}{\tau_{\mathrm{C}}}} - e^{-\frac{t_6}{\tau_{\mathrm{C}}}}) - (t_6 - t_5)(e^{-\frac{t_3}{\tau_{\mathrm{C}}}} - e^{-\frac{t_4}{\tau_{\mathrm{C}}}})}. \tag{A.8b}$$

If we then want to keep the coefficients close to 1, we can set the length of the B_1 window to roughly one half-life of the unwanted correlated signal so that it accumulates the same statistics as the following longer signal window.

B. Pseudorate Calculation

The concept of pseudorate was introduced in order to provide a reasonably well obtainable variable for correlated decays which is directly proportional to the actual rate and thus copies its changes over time given that some basic assumptions are met. Pseudorate for a particular type of correlated decay is calculated in two steps:

1. Applying fine cuts on the loose selection[1], histogram of time differences u between subsequent pairs of correlated decay candidates (including uncorrelated background as we cannot distinguish these two on pair to pair basis) is made for the part of AD and period of interest. Assuming that the decay is a random process of constant rate over the selected period (as well as any background), the u distribution should follow an exponential function as shown in Fig. B.1

$$P(u) = N_{\mathrm{u}} e^{-\frac{u}{\langle u \rangle}}, \tag{B.1}$$

where N_{u} is normalization factor and $\langle u \rangle$ is the mean time difference between subsequent candidate pairs inversely proportional to the rough pseudorate $\tilde{\rho}$ which encompasses both correlated and uncorrelated signals

$$\tilde{\rho} = \frac{1}{\langle u \rangle}. \tag{B.2}$$

There are in principle two ways of extracting the $\langle u \rangle$ – either by performing the exponential fit or by just taking the mean value of the distribution. Even though the former approach can in principle have some problems[2], most of the time both can be used with similar results as illustrated in Fig. B.3. As it is simpler, the latter approach is utilized from now on.

2. While the rough pseudorate could be a useful variable, it contains by construction contributions from both correlated decays and uncorrelated signal. The way to deal with that is to make a histogram of coincidence times (t) for each period and AD part as well. Such histogram then can be fitted as shown in Fig. B.2, presumably with one exponential function corresponding to the chosen correlated decay of mean lifetime τ and constant function corresponding to the uncorrelated background contribution

$$P(t) = N_{\mathrm{S}} e^{-\frac{t}{\tau}} + N_{\mathrm{B}}, \tag{B.3}$$

where "S" in the subscript stands for signal and "B" for background. Using the values N_{S}, N_{B} from the fit and τ either from the fit or as a tabulated value, we can calculate the contribution from the correlated decay. For a

[1] Note that the pseudorate calculation does not use signal and background windows, instead only one coincidence time window is employed, which may or may not correspond to the signal window.

[2] It can happen, albeit rarely, that the pseudorate quickly changes in the period of interest and as a result the u distribution becomes distorted essentially into a combination of exponential functions with various slopes, which cannot be easily fitted by a simple exponential function.

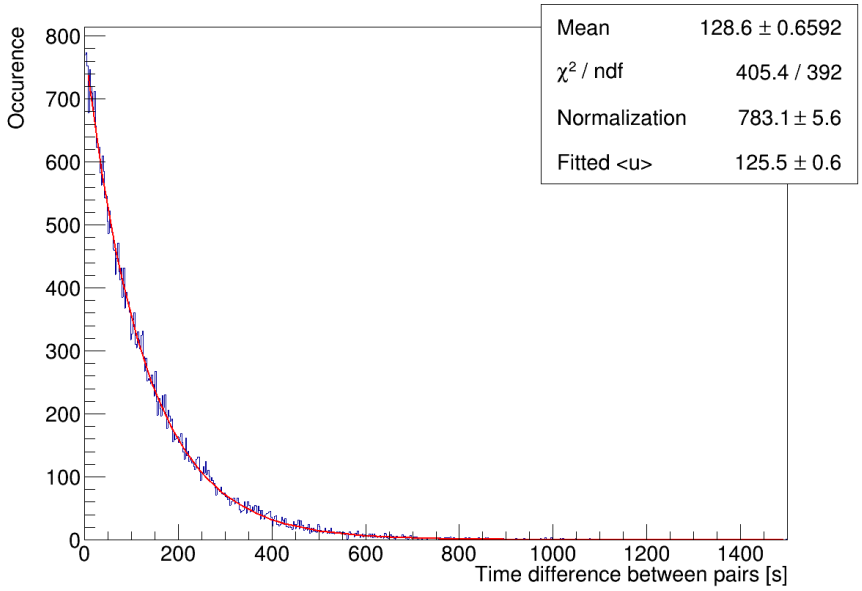

Figure B.1: An example of the distribution of time differences between subsequent pairs of correlated decays obtained using U-series fine cuts as listed in Tab. 4.4 on first two months of EH1-AD1 data. Mean value of the histogram does not match the fitted value $\langle u \rangle$ perfectly, but for most purposes either can be used.

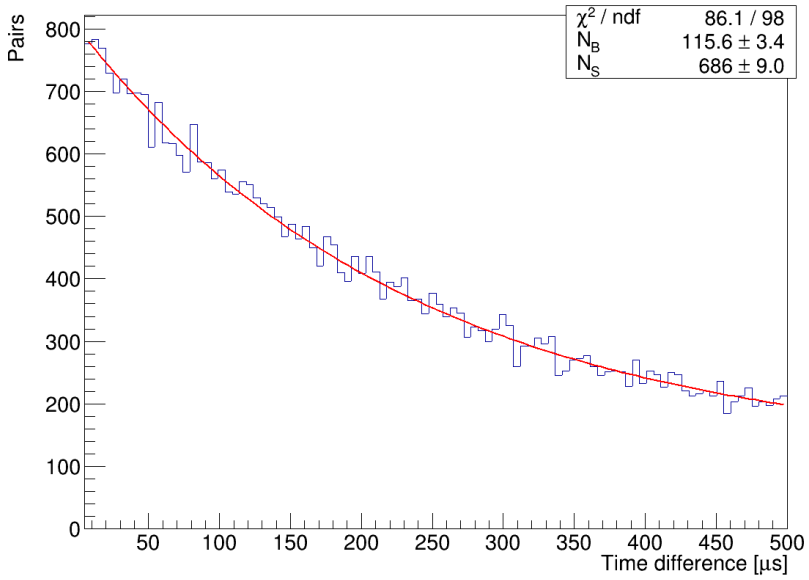

Figure B.2: An example of correlation time histogram used for U-series pseudorate correction. The histogram is fitted by formula described by Eq. B.3, values of parameters N_{S} and N_{B} are determined and used to calculate pseudorate correction using Eq. B.4.

window covering (t_1, t_2) interval of coincidence time, the correlated decay relative contribution is

$$\eta = \frac{\int_{t_1}^{t_2} N_{\mathrm{S}} e^{-\frac{t}{\tau}} \mathrm{d}t}{\int_{t_1}^{t_2} N_{\mathrm{S}} e^{-\frac{t}{\tau}} \mathrm{d}t + \int_{t_1}^{t_2} N_{\mathrm{B}} \mathrm{d}t} = \frac{N_{\mathrm{S}}\tau \left(e^{-\frac{t_1}{\tau}} - e^{-\frac{t_2}{\tau}}\right)}{N_{\mathrm{S}}\tau \left(e^{-\frac{t_1}{\tau}} - e^{-\frac{t_2}{\tau}}\right) + N_{\mathrm{B}}(t_2 - t_1)}. \quad \text{(B.4)}$$

Combining Eq. B.2 and Eq. B.4 we get the final formula for pseudorate ρ that corresponds to correlated decays of interest only

$$\rho = \eta\tilde{\rho} = \frac{1}{\langle u \rangle} \frac{N_{\mathrm{S}}\tau \left(e^{-\frac{t_1}{\tau}} - e^{-\frac{t_2}{\tau}}\right)}{N_{\mathrm{S}}\tau \left(e^{-\frac{t_1}{\tau}} - e^{-\frac{t_2}{\tau}}\right) + N_{\mathrm{B}}(t_2 - t_1)}, \quad \text{(B.5)}$$

given the assumption that only one type of correlated signal has made it pass our cuts. Nevertheless, this approach can be easily extended to account for more types of correlated decays passing the cuts simultaneously by appropriately adding them to the coincidence time fitting.

The downside of using coincidence time distribution to correct the rough pseudorate is that it also brings additional uncertainties and statistical fluctuations as illustrated in Fig. B.3.

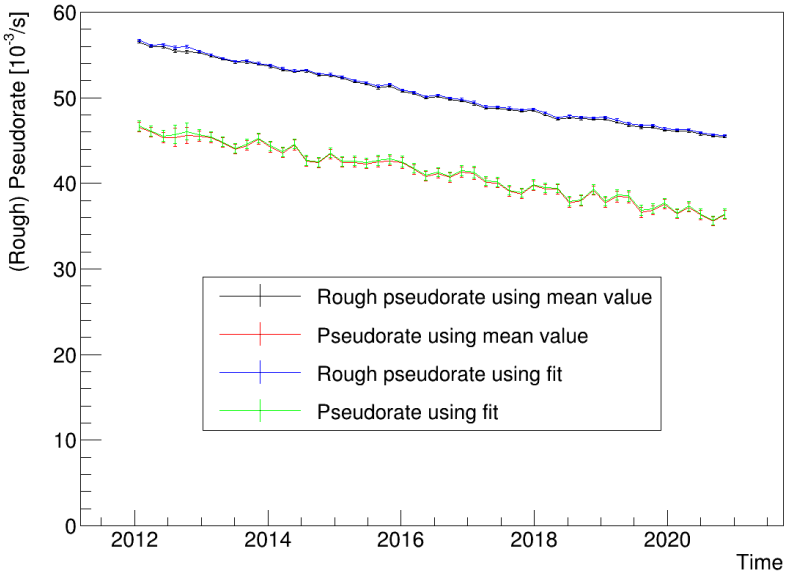

Figure B.3: Comparison of the rough pseudorate $\tilde{\rho}$ corresponding to both correlated and uncorrelated signals and the final pseudorate ρ for Ac-series EH1-AD2 for the whole period of data-taking. Moreover, each of these quantities are determined both by directly taking the mean time difference between correlated pairs $\langle u \rangle$ and by performing an exponential function fit described by Eq. B.1 and shown in Fig. B.1. While the choice of $\langle u \rangle$ determination (mean value or fit) has little impact, the determination of the correlated decays contribution (as described by Eq. B.4) introduces uncertainties which are propagated to the final pseudorate.

One of the assumptions we have made is that the rate of uncorrelated signals causing accidental coincidences is constant. If it changes over time, the correction described by Eq. B.4 prevents direct influence on pseudorate measurement, however, changing singles rate can still influence it indirectly by changing the multiplicity cut efficiency. Fortunately, the singles rate is quite stable except for a slight temporary excess at the beginning of data taking in EH3 which is illustrated in Fig. B.4 by the means of accidental coincidences pseudorate which is roughly proportional to square of singles rate. As shown in Figs. 4.7, 4.10, 4.13, this excess is not significant enough to cause the pseudorate measurement of individual series to be distorted.

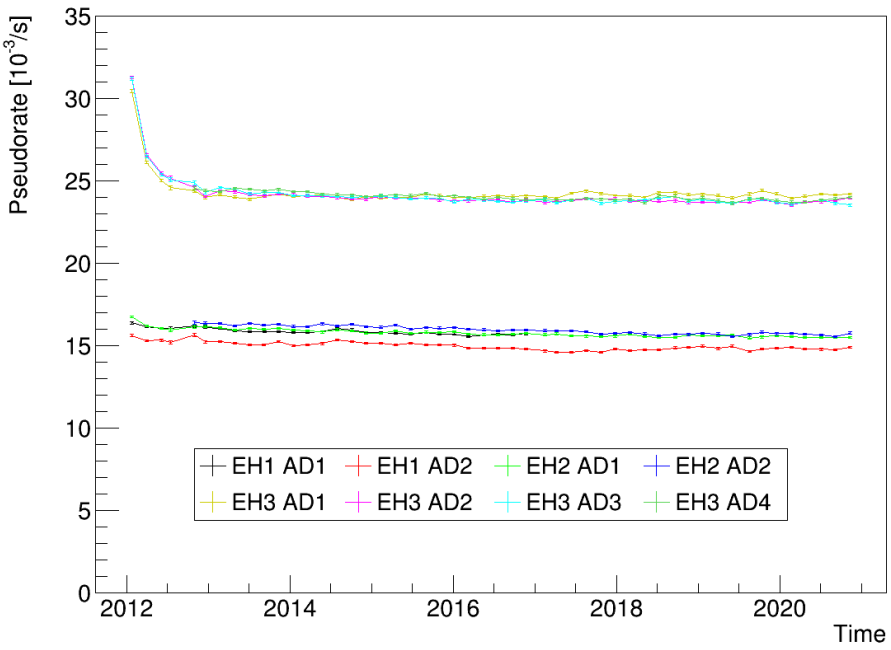

Figure B.4: Pseudorate of accidental coincidences caused by singles with $E_{\text{prompt}} \in$ (1.5 MeV, 3.2 MeV), $E_{\text{delayed}} \in$ (0.5 MeV, 3.2 MeV) and coincidence time in (1000 μs, 2000 μs) which is roughly proportional to the square of corresponding singles rate. There is only slight excess of the accidentals pseudorate in the beginning of data taking in EH3, otherwise the pseudorate of all ADs is stable.

Apart from the singles rate, muon rate also has an effect on detection efficiency and if it changed substantially, the pseudorate measurement could be distorted. Fortunately, muon rate is stable except for up to 1% seasonal variation [78], which has negligible impact on pseudorate measurement.

C. Rare ^{234}Pa – ^{234m}U – ^{234}U Correlated Decay

While the correlated decays centered around ^{212}Po, ^{214}Po, ^{215}Po are most prominent and most useful for detector response study, there is one more decay present in the ^{238}U decay chain with mean coincidence time suitable for measurement: ^{234}Pa $\rightarrow$ ^{234m}U $\rightarrow$ ^{234}U. The ^{234m}U isomer mean lifetime is quite convenient 48 μs, but this state is populated only in 0.03% of all ^{234}Pa decays [79]. This means we have very limited statistics to work with, but it is still insightful to look into this correlated decay in more detail.

Table C.1: Fine selection cuts for ^{234}Pa $\rightarrow$ ^{234m}U $\rightarrow$ ^{234}U correlated decay.

Signal window	(5 μs, 100 μs)
Coefficient	1
Background window	(100 μs, 2000 μs)
Coefficient w_{B}	0.05
Maximal distance	250 mm
Prompt energy	(0.5 MeV, 0.7 MeV)
Delayed energy	(1.25 MeV, 1.5 MeV)

C.1 Fine Selection

The fine selection parameters are listed in Tab. C.1. The signal coincidence time window starts at 5 μs to avoid Th-series contamination and ends at 100 μ at about two mean lifetimes of 48 μs yielding about 78% of ^{234}Pa $\rightarrow$ ^{234m}U $\rightarrow$ ^{234}U decays.

Meanwhile, Ac-series and U-series contamination is minimized by setting lower bound on the delayed energy to 1.25 MeV. In addition to that, the maximal distance was set to 250 mm in order to improve the signal to uncorrelated background ratio. This leads to further loss of already limited statistics. In order to offset it, data from all ADs were merged together. As a result, we can take a look into the prompt and delayed energy spectra and most importantly the spatial distribution. However, the statistics is still not sufficient for a pseudorate evolution study.

C.2 Coincidence Time

The coincidence time distribution obtained with relevant cuts from Tab. C.1 is shown in Fig. C.1. The mean lifetime was fitted in order to verify that ^{234m}U is responsible for the observed signal. The result is indeed consistent with the expectation of 48 μs.

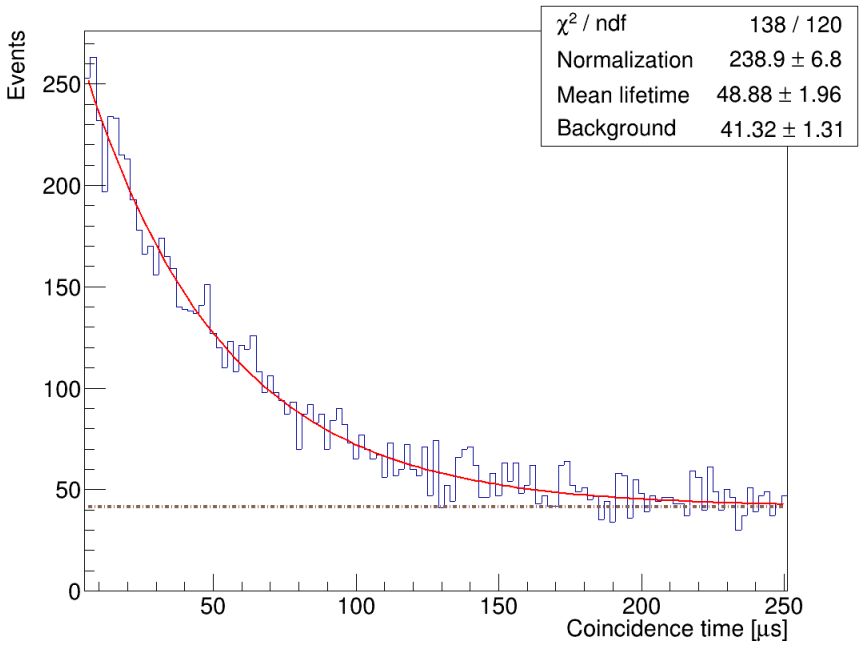

Figure C.1: Coincidence time distribution of ^{234}Pa $\rightarrow$ ^{234m}U $\rightarrow$ ^{234}U correlated decays obtained with relevant cuts from Tab. C.1. The measured mean lifetime is consistent with expectation of 48 μs thus confirming that the ^{234}Pa $\rightarrow$ ^{234m}U $\rightarrow$ ^{234}U is indeed responsible for the observed correlated signals.

C.3 Prompt Energy

The prompt signal which corresponds to populating of the $^{234\mathrm{m}}$U metastable state is caused by a combination of beta and gamma decay of a particular ^{234}Pa isomer with overall probability of 0.03% out from all ^{234}Pa decays. The maximal energy of the beta decay is 668 keV and it is immediately followed by a 131 keV gamma emission [79], the observed prompt energy thus should go up to $\sim$0.8 MeV. This is consistent with the data shown in Fig. C.2..

C.4 Delayed Energy

The delayed signal is caused by deexcitation of $^{234\mathrm{m}}$U by a gamma cascade with total energy of 1421 keV [79]. As we can see in Fig. C.3, the reconstructed energy mean is only $\sim$1.29 MeV, likely due to the liquid scintillator nonlinearity. While the peak can be fitted and theoretically used for detector energy response study, the statistics is far too limited for that purpose.

C.5 Spatial Distribution

As shown in Fig. C.4, the ^{234}Pa $\rightarrow$ ^{234m}U $\rightarrow$ ^{234}U correlated decays are present only in GdLS with just a small part reconstructed in LS. A rather surprising feature is that in GdLS the number of observed events decreases with radius.

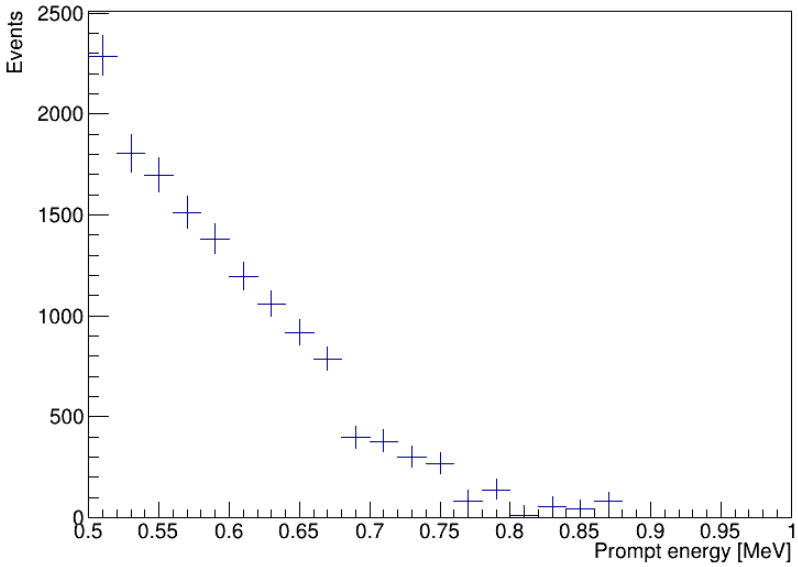

Figure C.2: Prompt energy spectrum of ^{234}Pa $\rightarrow$ ^{234m}U $\rightarrow$ ^{234}U correlated decay obtained with relevant cuts from Tab. C.1 and data from all ADs. Note that the shape of the spectrum below 0.7 MeV is affected by imperfect detection efficiency.

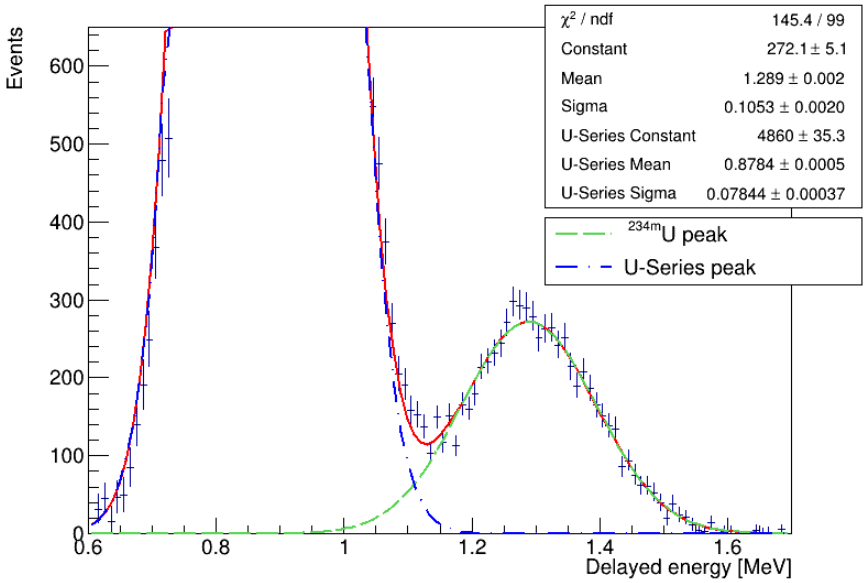

Figure C.3: Delayed energy spectrum obtained with relevant cuts from Tab. C.1 and data from all ADs. The ^{234m}U peak is centered around 1.3 MeV. The much bigger peak to the left is caused by a combination of Ac-series and U-series signal – for that reason, lower bound of 1.25 MeV on the delayed energy cut was adopted. The spectrum is fitted by double Gaussian function (red line) with both contributing parts also drawn separately in green and blue.

However, this is most likely caused by detection efficiency for events with energy < 0.7 MeV decreasing towards the center of the AD and not by the actual physical spatial distribution of ^{234}Pa. The fact that the presence of ^{234}Pa from the upper part of the U decay chain is limited to the GdLS suggests that the observed contamination of ^{214}Bi $\rightarrow$ ^{214}Po $\rightarrow$ ^{210}Pb decays (labeled as U-series) in LS is caused by ^{222}Rn coming from outside of LS.

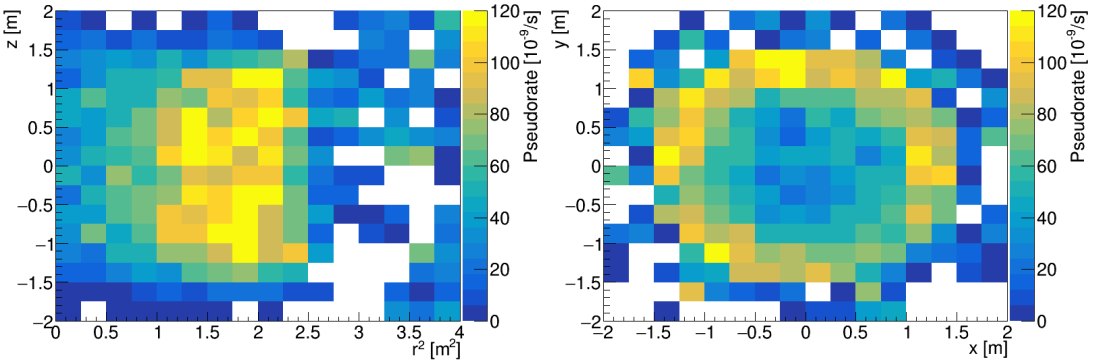

Figure C.4: Spatial distribution of ^{234}Pa $\rightarrow$ ^{234m}U $\rightarrow$ ^{234}U correlated decays obtained with cuts from Tab. C.1. Most of the events are observed in the outer part of GdLS as the detection efficiency for the prompt energy < 0.7 MeV required by the selection gets lower with decreasing radius. Meanwhile, the ^{234}Pa contamination in the LS seems to be negligible (there are quite large statistical fluctuations there though).

D. More Examples of EMF Effect on Azimuthal Nonuniformity

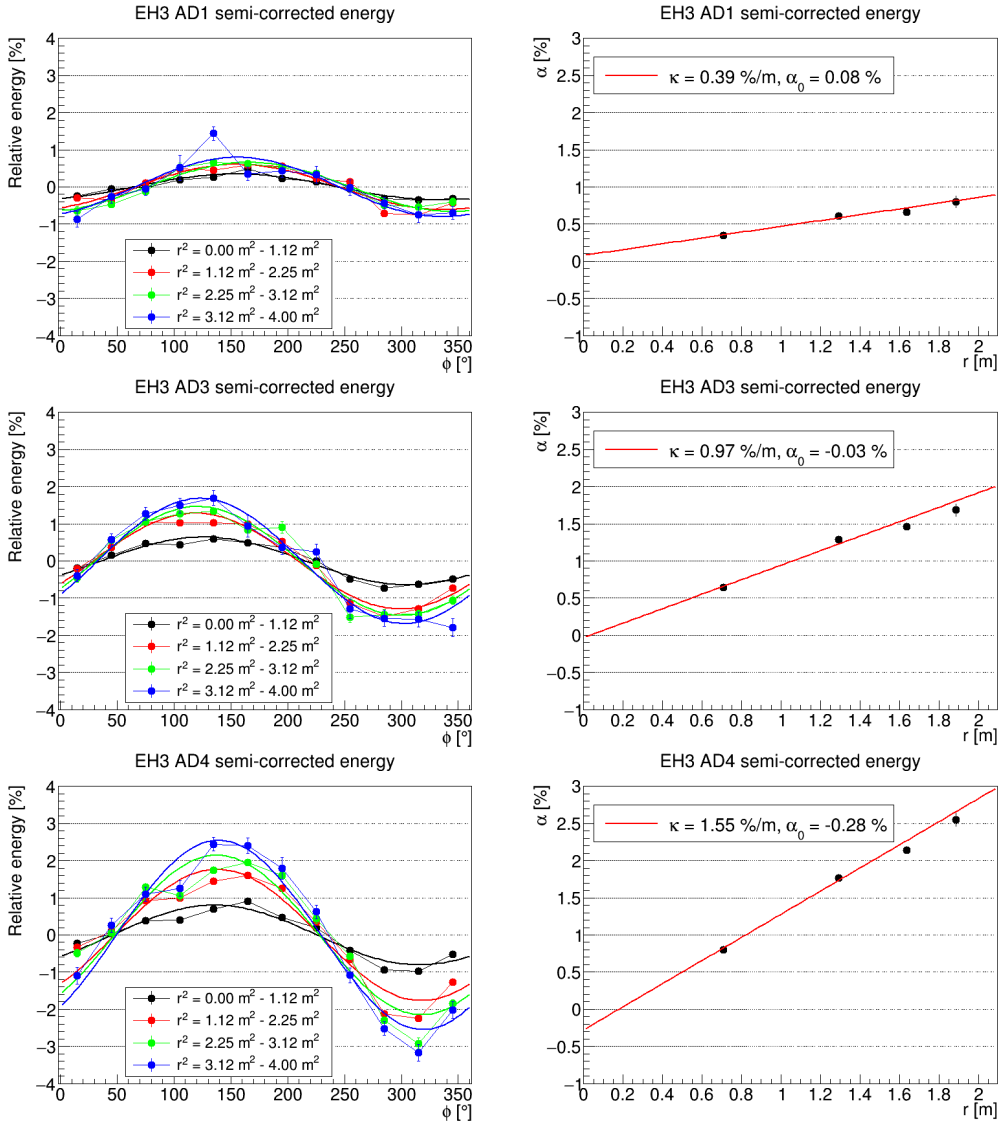

Figure D.1: Three more examples of the EMF effect on azimuthal nonuniformity using semi-corrected energy. (Left) Dependence of the relative energy on the azimuthal angle for 4 concentric rings in each AD, each ring shown in a different color. The dependence is fitted by $\alpha\cos(\phi-\phi_0)$, where α is the amplitude of the EMF effect in the particular ring. (Right) The dependence of α on radius r fitted by a linear function $\alpha(r)=\kappa r+\alpha_0$. ADs with pronounced EMF effect were selected and data from before March 2017 were used in order to reduce the impact of the dead PMTs.

E. Nonuniformity Maps Used in the Improved Correction

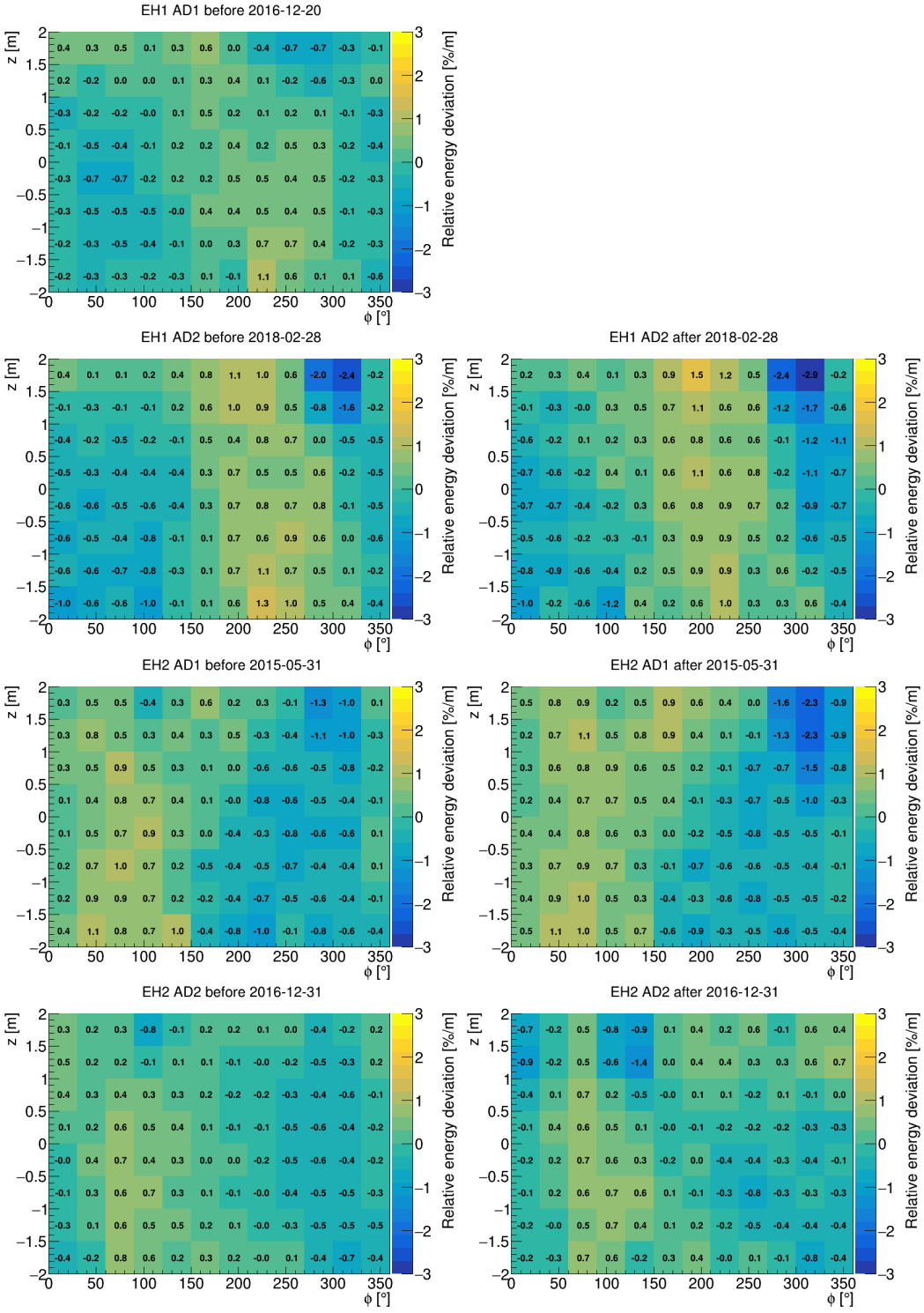

Figure E.1: Radius-normalized $\tilde{M}_{\mathrm{RAC}}$ maps used by the new azimuthal correction in the near halls. Note that EH1-AD1 has only one time period.

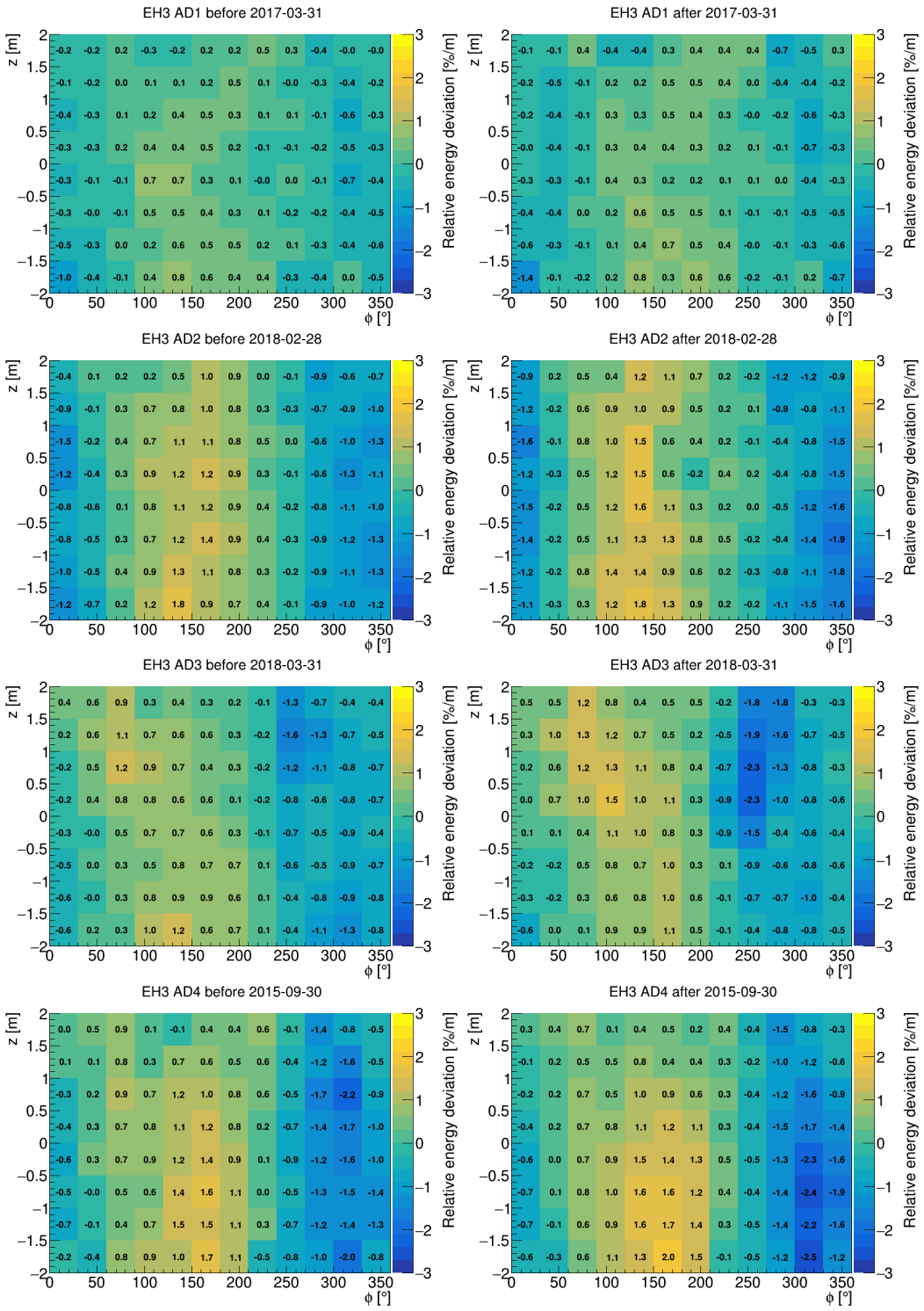

Figure E.2: Radius-normalized $\tilde{M}_{\mathrm{RAC}}$ maps used by the new azimuthal correction in the far hall.

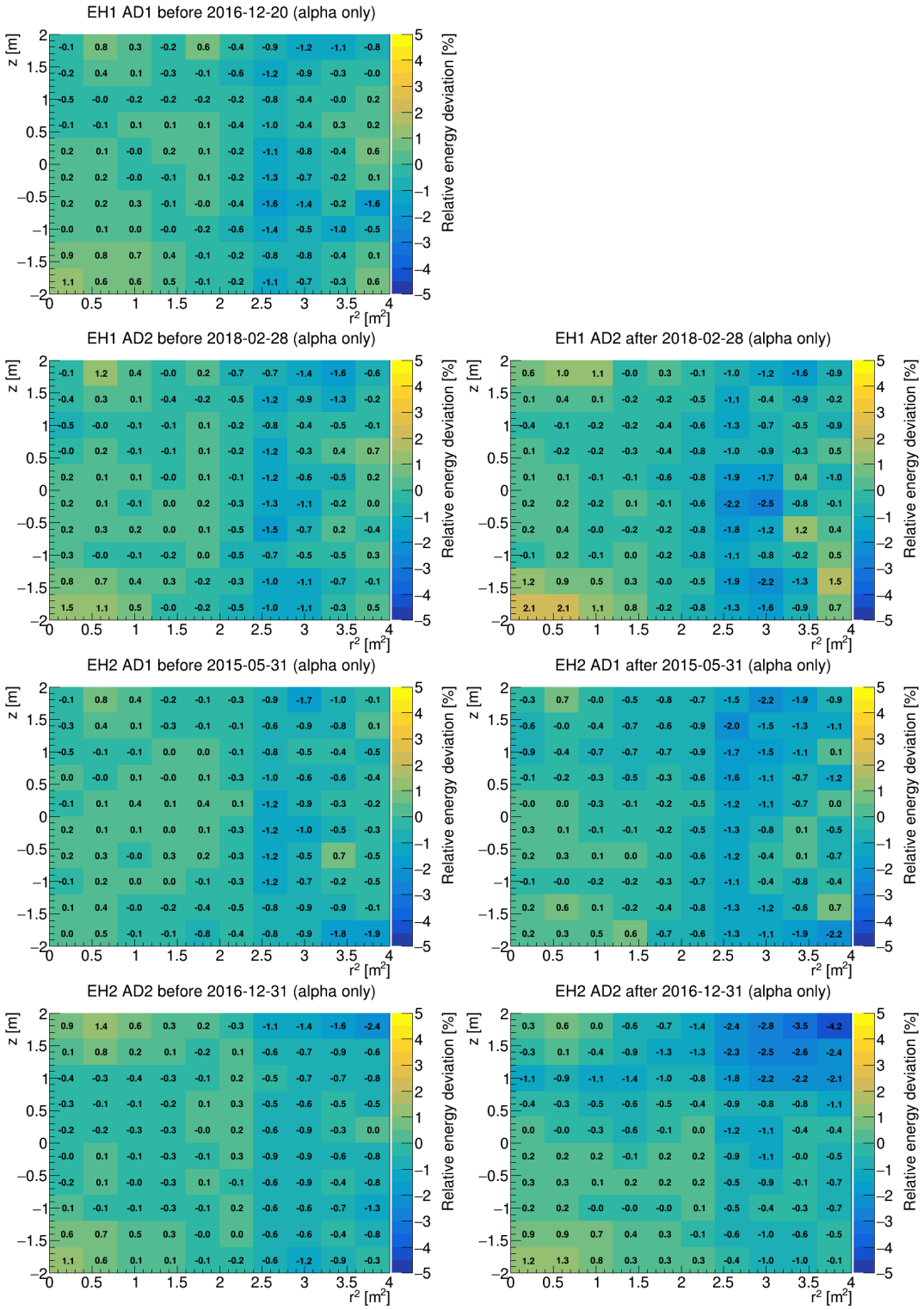

Figure E.3: Residual r^2-z nonuniformity maps M_{RZ} based on U-series alpha signal for ADs in the near halls. Note that EH1-AD1 has only one time period.

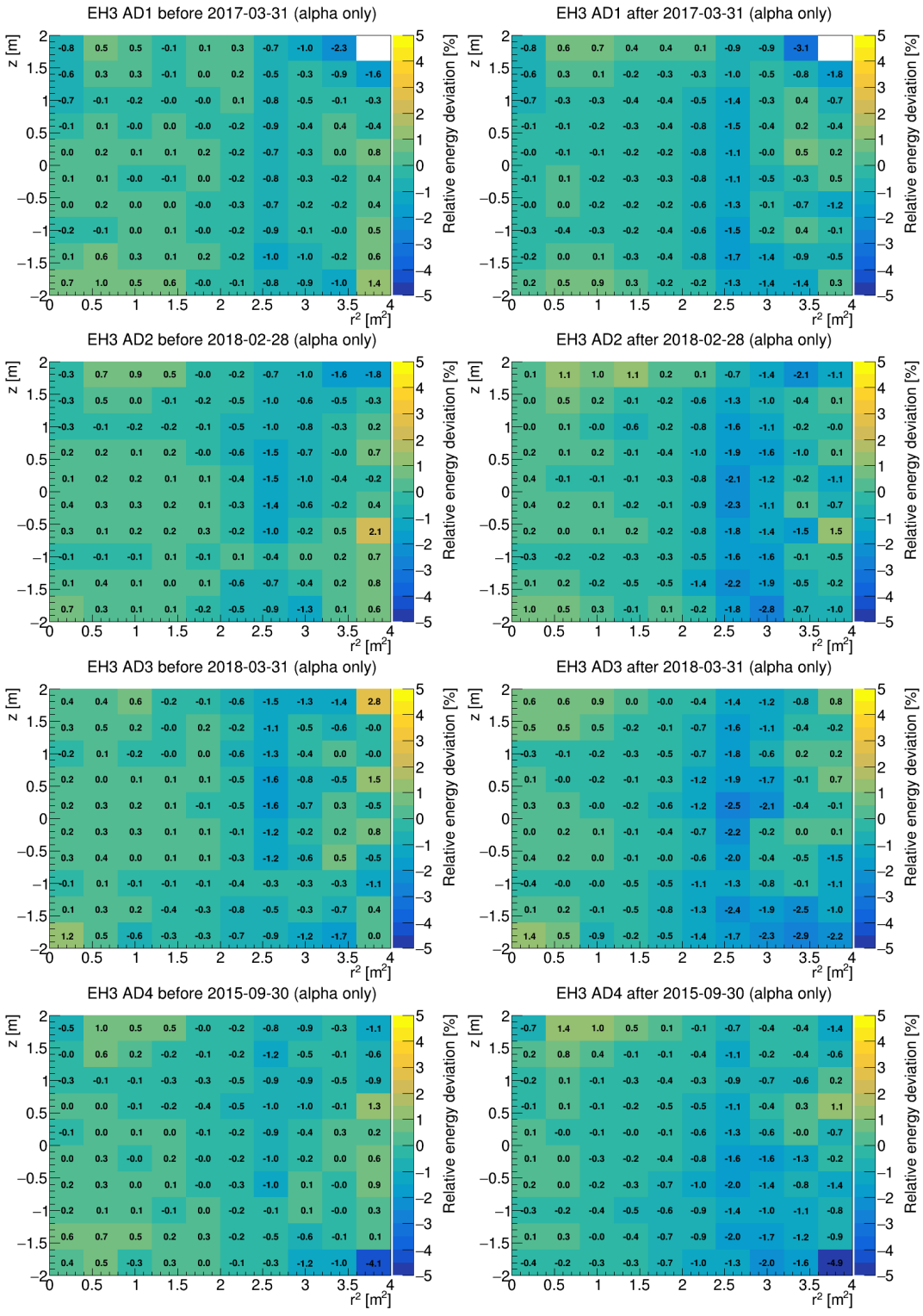

Figure E.4: Residual r^2-z nonuniformity maps M_{RZ} based on U-series alpha signal for ADs in the far hall.

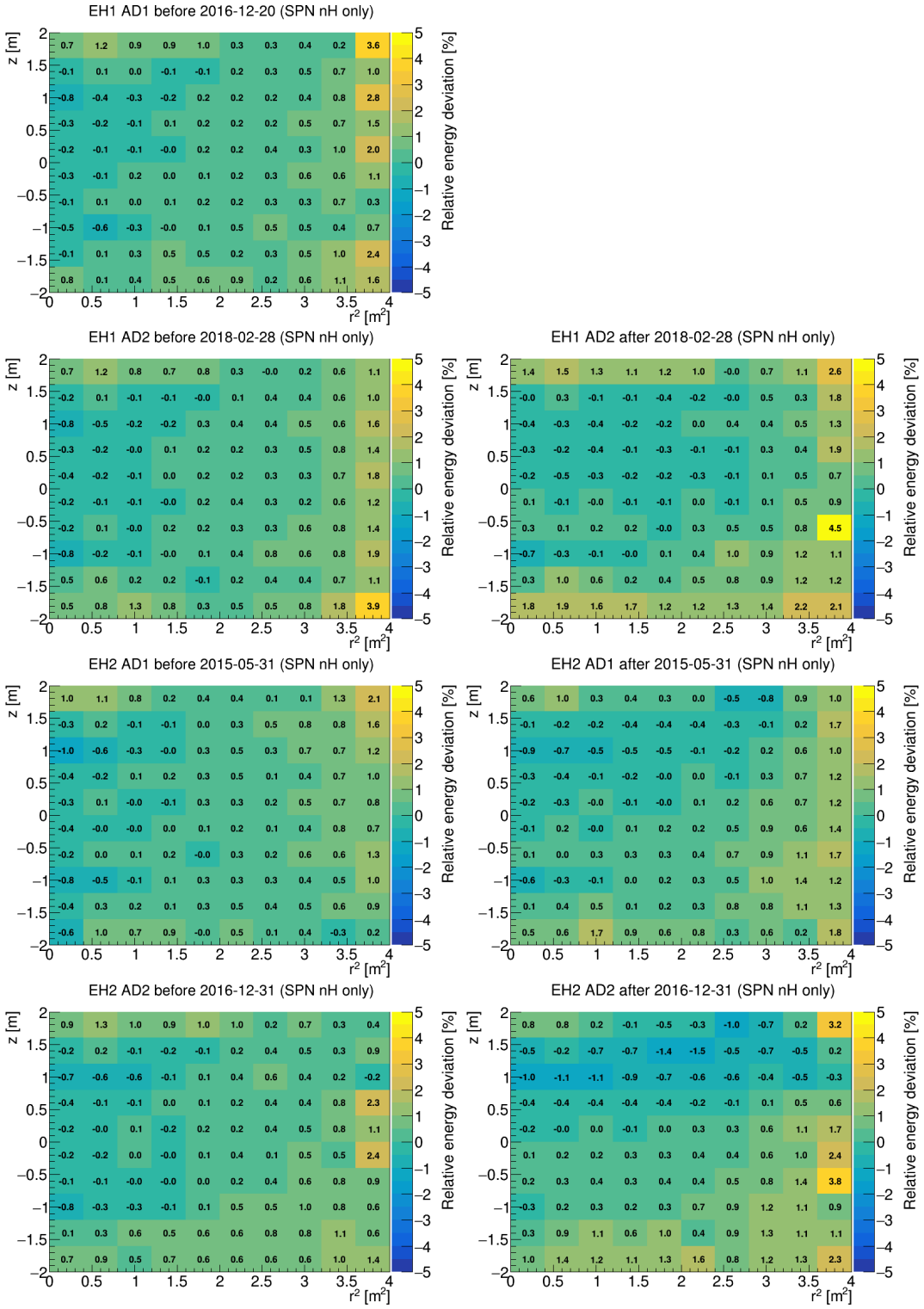

Figure E.5: Residual r^2-z nonuniformity maps M_{RZ} based on SPN nH signal for ADs in the near halls. Note that EH1-AD1 has only one time period.

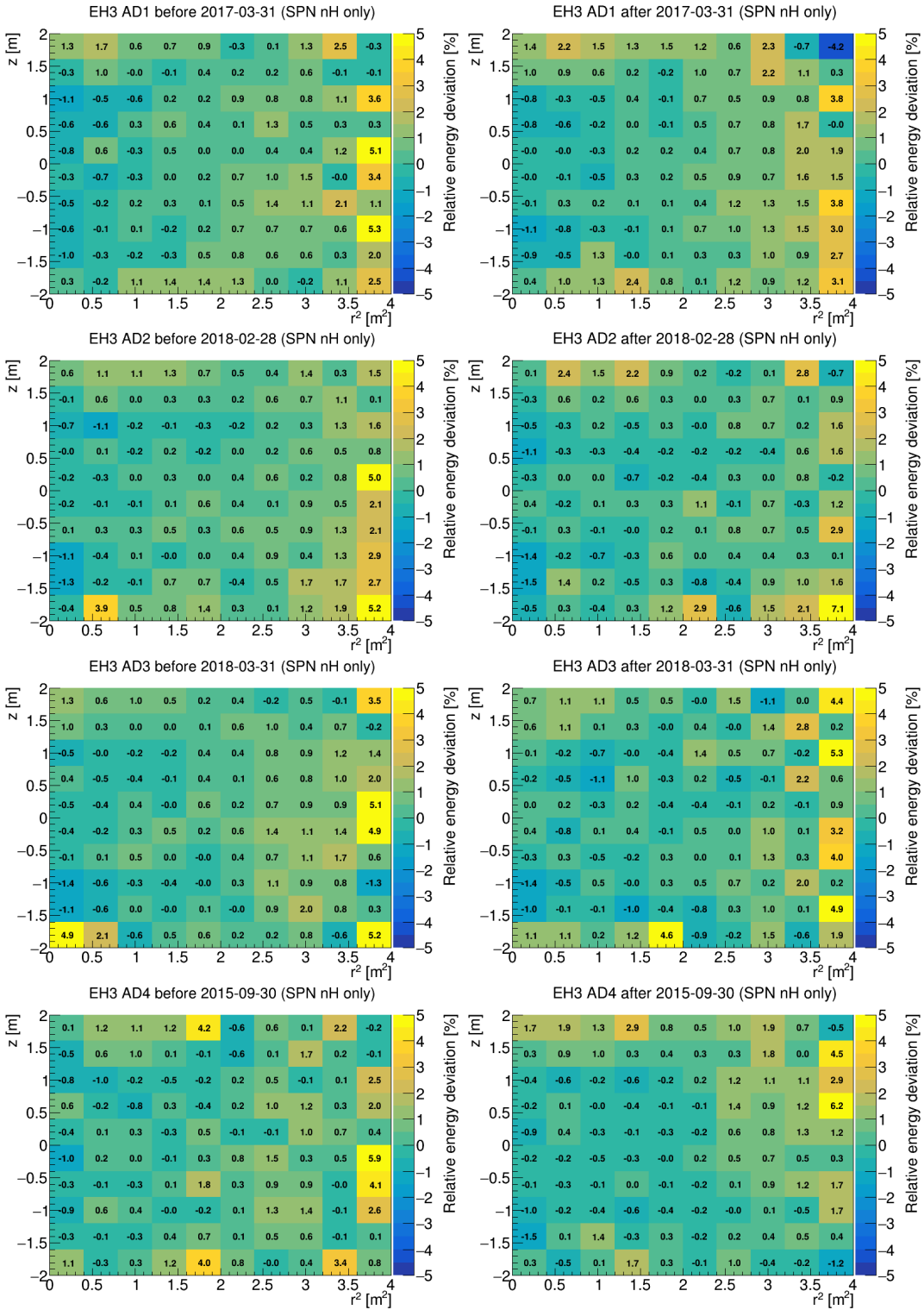

Figure E.6: Residual r^2-z nonuniformity maps M_{RZ} based on SPN nH signal for ADs in the far hall.

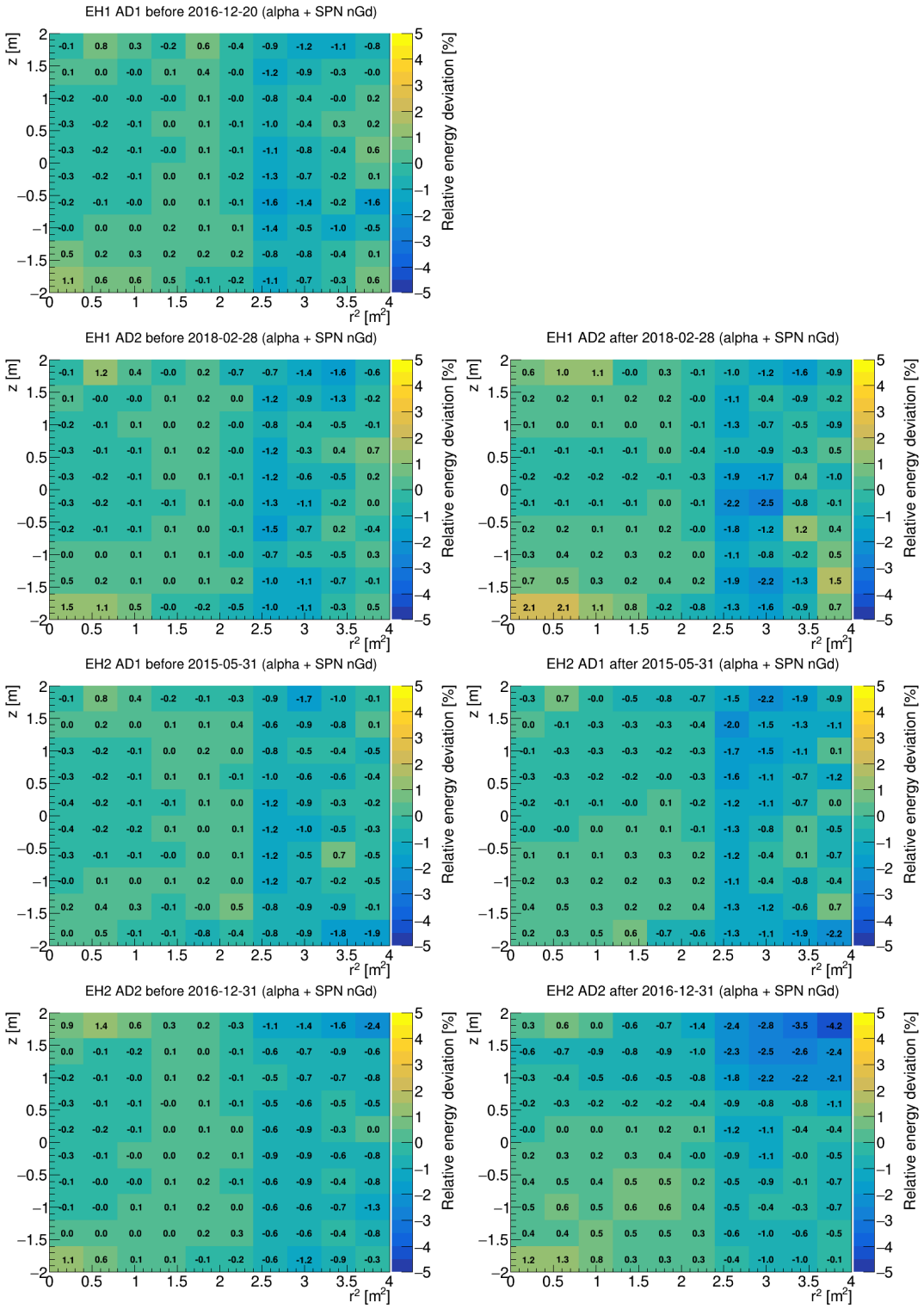

Figure E.7: Residual r^2-z nonuniformity maps M_{RZ} based on SPN nGd signal in GdLS and U-series alpha in LS for ADs in the near halls. Note that EH1-AD1 has only one time period.

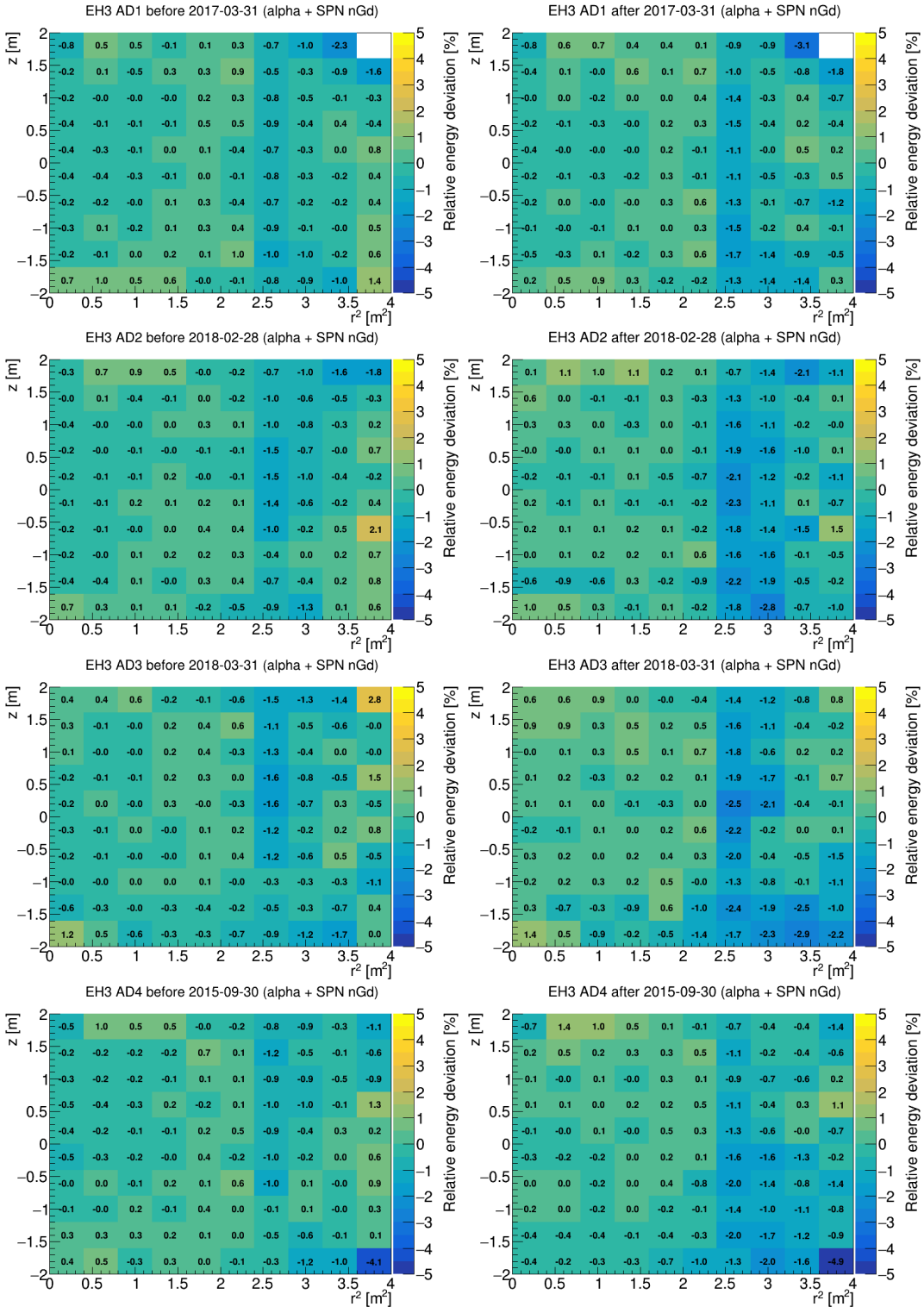

Figure E.8: Residual r^2-z nonuniformity maps M_{RZ} based on SPN nGd signal in GdLS and U-series alpha in LS for ADs in the far hall.

List of Publications

Daya Bay articles

D. Adey *et al.* Search for a time-varying electron antineutrino signal at Daya Bay. *Physical Review D* **98**, 092013 (2018)

D. Adey *et al.* A high precision calibration of the nonlinear energy response at Daya Bay. *Nuclear Instruments and Methods in Physics Research Section A* **940**, 230-242 (2019)

D. Adey *et al.* Extraction of the ^{235}U and ^{239}Pu antineutrino spectra at Daya Bay. *Physical Review Letters* **123**, 111801 (2019)

P. Adamson *et al.* Improved constraints on sterile neutrino mixing from disappearance searches in the MINOS, MINOS+, Daya Bay, and Bugey-3 experiments. *Physical Review Letters* **125**, 071801 (2020)

F. P. Ann *et al.* Search for electron-antineutrinos associated with gravitational-wave events GW150914, GW151012, GW151226, GW170104, GW170608, GW-170814, and GW170817 at Daya Bay. *Chinese Physics C,* **45(5)**, 055001 (2021)

F. P. Ann *et al.* Antineutrino energy spectrum unfolding based on the daya bay measurement and its application. *Chinese Physics C* **45(7)**, 073001 (2021)

F. P. Ann *et al.* Precision measurement of reactor antineutrino oscillation at kilometer-scale baselines by Daya Bay. Preprint: *arXiv:2211.14988* (2022)

JUNO articles

A. Abusleme *et al.* Calibration Strategy of the JUNO Experiment. *Journal of High Energy Physics* **03**, 004 (2021)

A. Abusleme *et al.* JUNO sensitivity to low energy atmospheric neutrino spectra. *European Physical Journal C* **81**, 10 (2021)

A. Abusleme *et al.* The design and sensitivity of JUNO's scintillator radiopurity pre-detector OSIRIS. *European Physical Journal C* **81(11)**, 973 (2021)

A. Abusleme *et al.* Radioactivity control strategy for the JUNO detector. *Journal of High Energy Physics* **11**, 102 (2021)

A. Abusleme *et al.* Feasibility and physics potential of detecting ^{8}B solar neutrinos at JUNO. *Chinese Physics C* **45(2)**, 023004 (2021)

A. Abusleme *et al.* Optimization of the JUNO liquid scintillator composition using a Daya Bay antineutrino detector. *Nuclear Instruments and Methods in Physics Research Section A* **988**, 164823 (2021)

Proceedings

T. Dohnal, V. Vorobel, and T. Tměj. Measurement of liquid scintillator nonlinearity. *PoS ICHEP2020* **390**, 887 (2020)

T. Dohnal (for Daya Bay). Latest Results from the Daya Bay Experiment. *PoS ICRC2021*, 1175 (2021)

Daya Bay technical notes

O. Dalager *et al.* Berkeley-Czech-Irvine-Shandong nH analysis note. *Internal Daya Bay document DocDB-12593-v2*[†] (2022)

[†]Internal Daya Bay documents are not publicly displayed, but they are available upon request on my email dohnal@ipnp.mff.cuni.cz.