New studies on cosmogenic induced spallation background for Supernova relic neutrino search in the Super-Kamiokande experiment

Alice Coffani

INSTITUT POLYTECHNIQUE DE PARIS

NNT : 2021IPPAX113

Thèse de doctorat

New studies on cosmogenic induced spallation background for Supernova relic neutrino search in the Super-Kamiokande experiment

Thèse de doctorat de l'Institut Polytechnique de Paris
préparée à l'École polytechnique

École doctorale n°626 École doctorale de l'Institut Polytechnique de Paris (EDIPP)
Spécialité de doctorat : Physique

Thèse présentée et soutenue à Palaiseau, le 16-12-2021, par

ALICE COFFANI

Composition du Jury :

Yves Sirois Directeur de recherche, LLR, École polytechnique	Président
Maria Cristina Volpe Directrice de recherche, Laboratoire Astroparticule et Cosmologie	Rapporteur
Jaques Marteau Maître de conférences, Institut de Physique des deux Infinis de Lyon	Rapporteur
Cécile Jollet-Meregaglia Maître de conférences, Centre d'Etudes Nucléaires de Bordeaux	Examinateur
Joao Pedro Athayde Marcondes De Andre Chargé de recherche, Institut Pluridisciplinaire Hubert Curien	Examinateur
Michel Gonin Directeur de recherche, LLR, École polytechnique	Directeur de thèse

"We have calcium in our bones, iron in our veins,
carbon in our souls, and nitrogen in our brains.
93 percent stardust, with souls made of flames,
we are all just stars that have people names."

Nikita Gill

Acknowledgements

It is a great pleasure to end my thesis thanking the people that accompanied me in this journey over the last three years, their personal and professional support had been crucial for the success of my PhD.

I would like to start by thanking all members of my defense jury, Yves Sirois, Maria Cristina Volpe, Jaques Marteau, Cécile Jollet-Meregaglia and Joao Pedro Athayde Marcondes De Andre. I truly appreciate the time you spent reading my thesis, the valuable comments and enlightening questions you addressed to my work.

I would like to express my gratitude to Michel, my PhD supervisor, for offering me the fascinating opportunity of working in the LLR neutrino group but also for having pushed me to face the challenging experience of spending some months in the remote Japanese village of Mozumi. However arduous it was to run away from wild monkeys, this period was incredibly beneficial for my growth as a researcher.

A big thank you should also go to the members of the neutrino group, all enthusiastic physicists but also delightful persons. I greatly appreciated the opportunity to work closely with them, to take advantage of their advice and the valuable discussions about physics, life, Italian castles and much more. Among them I want to acknowledge, in particular, Sonia and Laura, my true guides in these three years. I could never have successfully concluded my PhD without the fruitful discussions with them, the invaluable support and help in everyday problems but also the pleasant atmosphere they created in our shared office. It is also very appreciated the time Sonia spent to carefully read my thesis, giving me the most useful comments and helping me to shape my work in the best way. I cannot forget to explicitly mention the third office mate, Alberto, a patient colleague and cherished friend.

I also want to thank all the Super-Kamiokande collaborators, especially the members of Low Energy group, it was an enormous pleasure to work in this incredible experiment and have the possibility to meet such passionate researchers.

My sincere gratitude to my friends and family can only truly be expressed in my native language. Prima di tutto devo ringraziare i miei amici parigini: Martina, Matteo, Filippo, Simone, Marianna, Tullio, Marco. Ho avuto la fortuna d'iniziare e finire questa avventura insieme a loro, di condividere momenti stupendi, risate, aperitivi, discussioni sugli argomenti più disparati. Mi hanno accompagnata negli alti e bassi di questi tre anni facendo in modo che non mi sentissi mai sola, una famiglia lontana da casa.

Un ringraziamento speciale va alle mie amiche del cuore, Caterina, Giulia, Marta, Sharon e Silvia. Non ci sono parole per esprime l'importanza che hanno nella mia vita, il supporto che mi hanno sempre saputo dimostrare nonostante la distanza, il valore che un loro semplice messaggio ha avuto nei momenti più difficili. Poche cose mi hanno riempito il cuore come la sorpresa che mi hanno fatto il giorno della mia difesa, un regalo meraviglioso.

Qualche parola va anche alla persona che più mi è stata vicino in questi tre anni, Feli. Non posso che ringraziarlo dal profondo del cuore per avermi costantemente accompagnata, con il suo sostegno, il suo aiuto concreto e quotidiano, la sua perspicacia, il suo senso dell'umorismo, la sua integrità, il suo modo di affrontare la vita con quella dolcezza contagiosa che ti strappa un sorriso anche nei momenti di difficoltà. Una persona rara e preziosa che ho avuto l'incredibile fortuna di avere al mio fianco in questo dottorato. Non vedo l'ora di scoprire insieme cosa la vita ci riserverà.

Infine un grazie enorme alla mia famiglia: mia mamma, mio papà e le mie sorelle, Giorgia e Greta. Per quell'amore incondizionato che non conosce distanze, che ti raggiunge ovunque tu sia e ti dà la spinta per lottare e raggiungere i tuoi obiettivi. La mia

forza è sempre stata in quell'amore, nella consapevolezza che ovunque io sia e qualunque scelta io faccia ci saranno loro a fare il tifo per me. Siete la mia roccia.

Contents

Ai miei genitori, la mia forza.

Introduction

Neutrinos have played a key role in astrophysics, from the characterization of nuclear fusion processes in the Sun to the observation of supernova SN1987A and multiple extragalactic events. The Super-Kamiokande experiment (SK), the world's largest underground neutrino detector, played a major part in these astrophysical studies by investigating low energy neutrinos ($\mathcal{O}(10)$ MeV). It has notably been instrumental in characterizing the ^{8}B solar neutrino spectrum and currently exhibits the best sensitivity for the observation of supernova relic neutrinos (SRN). This signal arises from the cumulative flux of neutrinos originating from all of the supernovae events which have occurred throughout the Universe. The detection of this cosmological background is currently one of the most awaited discoveries in neutrino physics.

Starting from the end of 2020, with the beginning of SK-Gd, the upgraded phase of Super-Kamiokande with the injection of gadolinium salt into the detector's otherwise ultrapure water, a remarkable increase of neutron detection efficiency is expected and thereby an extent of the experiment reach to the currently unobserved supernova relic neutrino flux. A major effort in repairing the detector as well as an implementation of new calibration methods had been instrumental for the transition of the experiment and the first months of data taking have been successful. The common denominator of the work activities presented in this thesis is the preparation for this new exciting phase of Super-Kamiokande.

Low energy searches in SK, like the SRN analysis, face significant challenges due to important backgrounds from cosmic induced spallation: muons reaching SK can break up oxygen nuclei, producing unstable isotopes whose radioactive decays are a major contamination below $\sim$ 20 MeV. The reduction of this major contamination is fundamental in order to reach the sensitivity for the first observation of supernova relic neutrinos. Although current reduction techniques, at the cost of a significant signal loss, remove a large fraction of spallation decays, this background remains dominant at 6-20 MeV kinetic energies. This is the reason why in this work a new method has been developed implementing state-of-the-art neutron tagging algorithms, as well as a thorough characterization of spallation-inducing mechanisms. A new and unique FLUKA-based simulation, incorporating the most advanced nuclear models, played a key role in the investigation of spallation processes and in turn the tuning of new rejection methods. Thanks to the simulation, we now have not only an in-depth study of spallation backgrounds but also a valuable tool for optimizing the analysis strategies of future low energy searches in water-Cherenkov detectors. It is in view of SK-Gd that the new spallation reduction method, carefully optimized thanks to unique insights from simulation, proves its real potential, showing an unprecedented efficiency in spallation identification and opening the door for important new discoveries.

This thesis is organized in 7 Chapters: Chapter 1 briefly introduces the main properties of neutrinos, including an overview of neutrino sources and interaction processes. It also contains a concise summary of supernova explosion dynamics and SRN flux characteristics. Details on the SRN search are presented and followed by a description of spallation processes. Chapter 2 is dedicated to the Super-Kamiokande detector: its structure, electronics, water system as well as the customized Monte Carlo simulation and

reconstruction algorithms are presented. Super-Kamiokande was recently upgraded in view of its new phase with gadolinium. The major refurbishment work was completed in Summer 2018, followed by a calibration period where significant efforts were made to introduce new measurements techniques. These studies led to the successful loading of Gd in pure water and the beginning of SK-Gd in August 2020. These topics will be described in Chapter 3. Chapter 4 explains the construction of the simulation to model spallation backgrounds, the models adopted, the main components and their interfaces. The second half of this Chapter is dedicated to the validation of this simulation using SK data, an essential step in the development of an accurate calculation framework. The first insights on spallation processes obtained with the new simulation are presented in Chapter 5: a deeper understanding of muon energy losses, shower development and a better comprehension of radioactive isotope production is crucial to improve spallation removal techniques. Subsequently, Chapter 5 illustrates the new technique developed to identify spallation events in SK-IV. The simulation accurately reproduces data and proves to be a key tool not only to interpret the results of current analysis but also to design future strategies. Finally Chapter 6 enlightens the ability of the spallation simulation to adapt and optimize rejection cuts for SK-Gd, providing performance predictions and guiding new analysis approaches.

Some of the studies presented in this thesis are presented in a publication with the title “New Methods and Simulations for Cosmogenic Induced Spallation Removal in Super-Kamiokande-IV", which can be find in Ref. [1].

Chapter 1

Neutrino Physics

Among the fermions of the Standard Model of Particle Physics, neutrinos are the most elusive species, making our actual knowledge of their properties quite limited. As any other lepton, neutrinos do not participate in strong interactions. Since they have zero electric charge, thus no electromagnetic interactions are possible, they are only subject to weak and gravitational forces. These properties are confirmed by the recent proof of neutrino non-vanishing masses. As a consequence, although neutrinos are one of the most abundant particles in the universe, they interact very weakly with matter. It is due to their peculiar properties that experimental neutrino physics is extremely challenging. However, thanks to their distinctive nature, neutrinos represent the ideal tool to investigate weak interactions and to accelerate the progress of fundamental physics in describing our universe. Moreover, since they travel almost unattenuated and undeflected for enormous distances, even through extremely dense material, they can directly reveal precious information about the source and processes that produced them. As it will be depicted in the following Sections, the history of neutrino physics results is tightly intertwined with the understanding of both weak forces and astrophysical processes.

In this Chapter we will first present a historical overview of neutrino physics, followed by a description of the main properties of neutrinos, such as their interactions, masses and oscillating behavior. We will give a brief overview of the main neutrino sources while offering a deeper description of supernova relic neutrinos, namely neutrinos emitted from all the past core-collapsed supernovae in the universe. Their origin, characteristics and detection channels will be presented in Section 1.6. The experimental search of neutrinos from past supernovae explosions will frequently recur along this thesis, especially because its detection in Super-Kamiokande suffers from serious contamination from spallation background, the core topic of this work. A "spallation" event occurs when a muon traveling through matter loses its energy via disruptive processes that lead to nuclear breakup. Subsequently, the production of unstable daughter nuclei introduces delayed decay signals which can mimic interesting events in SK searches. Details on spallation processes and models will be given in Section 1.7.

1.1 The discovery of neutrinos

Unlike the majority of particles, the neutrino was first postulated theoretically as a consequence of experimental observations that were at the time incomprehensible: in 1923, Chadwick surprisingly discovered that the beta decay energy spectrum is continuous. In contrast with α- and γ-decays, if the beta decay were a two body process, it would contradict the law of conservation of energy, momentum and angular momentum [2, 3]. To address this unsolved problem, in 1930 Pauli speculated in a famous letter that an extremely light particle, with no electric charge and $\frac{1}{2}$ spin, could be emitted together with

the electron in the beta decay process. This particle was named "neutrino" later on by Enrico Fermi, in 1931 [4].

The neutrino was observed for the first time only about 20 years later, when fission reactors became available. In 1946 Pontecorvo proposed a method to detect free neutrinos from reactors, taking advantage of the inverse β-decay process:

$$\bar{\nu}_e + p \rightarrow e^+ + n \tag{1.1}$$

The detection principle was based on the measurement of the 511 keV photons associated with positron annihilation before a neutron capture reaction occurring a few μs later [5]. Based on Pontecorvo's idea, Cowan and Reines used a water tank filled with a ^{108}Cd solution and surrounded by two liquid scintillators, placed in close vicinity to a nuclear reactor. In 1956 they published the first evidence of the existence of neutrinos by detection of inverse beta decays at the Savannah River nuclear power plant [6]. Reines was awarded the Nobel Prize in 1995 for their discovery.

The helicity of neutrinos was measured soon after in 1958: Goldhaber determined the polarization of photons in an electron capture reaction on ^{152}Eu. The experimental results were consistent with the observation of a single chirality state: all neutrinos are left-handed and all anti-neutrinos are right-handed [7].

The muonic neutrino, ν_μ, was detected in 1962 at Brookhaven, using neutrinos produced at an accelerator facility [8]. The experiment used a proton beam hitting a beryllium target that creates secondary pions: their subsequent decays, always accompanied by muons, produced an almost pure ν_μ beam. Moreover, when the τ lepton was indirectly discovered by a series of experiments at SLAC, in the US, between 1974 and 1977, it suggested the existence of tau neutrinos ν_τ. In year 2000 the ν_τ was directly observed in the DONUT experiment at Fermilab [9, 10].

Some years earlier, the LEP experiment at CERN had shown, as a consequence of the measurement of the Z boson decay width, that the number of active neutrinos must necessarily be three [11]. If we want to consider more neutrinos, even without a charged lepton as partner, more neutrino eigenstates could in principle exist, but the coupling with the W boson would be inaccessible. The same arguments would also prohibit the coupling with the Z boson. These hypothetical new neutrino species would hence not be affected by weak interactions within the framework of the Standard Model, and are therefore referred to as "sterile neutrinos". To date, there is no conclusive proof of the existence of sterile neutrinos.

In the same year of the discovery of ν_μ, Maki, Nakagawa and Sakata [12] extended an idea put forward by Pontecorvo in the 50s: neutrino flavors can mix, and their mixing results into one type of neutrino oscillating into another. This effect is called flavor oscillation and, as discussed below, can occur only if neutrinos are massive. A second powerful hint of neutrino flavor changes arises in 1968 when an experiment in the Homestake mine in South Dakota measured the neutrino flux from nuclear fusion in the sun. In landmark papers R. Davis and J. Bachall presented results showing a deficit in the measured rate of solar ν_e compared to that predicted by the standard solar model (SSM)[13, 14]. This deficit was dubbed "the solar neutrino problem". Further hints of oscillation were provided by atmospheric neutrinos in the late 1980s with the measurements of the Kamiokande experiment. This large-scale underground Cherenkov detector found a large deficit in the fraction of ν_μ flux compared to ν_e for neutrinos produced by cosmic rays interacting in Earth's atmosphere. Beyond providing the first evidence for neutrino oscillations, the observation of solar neutrinos represented a major breakthrough as the first detection of astrophysical neutrinos. These pioneering efforts gave birth to neutrino astrophysics and are the reason why Davis and Koshiba shared the 2002 Nobel Prize in Physics. From there

on, a major interest for neutrinos coming from astrophysical sources resulted in the construction of numerous neutrino telescopes that have pushed forward our understanding of the Universe, its components and structures have contributed to the exploration of a range of astrophysical questions.

The atmospheric neutrino problem was finally solved in 1998 by the high-precision measurements of the Super-Kamiokande experiment, the follow-up of Kamiokande with a 15-times larger mass, which measured a deficit in the atmospheric upward-going ν_μ flux, with a characteristic angular dependence consistent with mass-induced oscillations [15]. The SNO experiment, in Canada, provided the answer to the solar neutrino problem in 2002. Using an ultra-pure heavy water Cherenkov detector they were sensitive to both the rate of solar ν_e, through the charged current reaction $\nu_e + n \rightarrow e^- + p$ but also via the elastic scattering of electrons by neutrinos, and the total flux from all neutrino flavors, through the neutral current reaction $\nu_\alpha + d \rightarrow p + n + \nu_\alpha$, where the neutron was identified by the photon produced by the neutron capture on deuterium. Their results not only confirmed the deficit in the ν_e flux but showed also that the total flux was consistent with that predicted by the SSM, with the conclusion being that the missing ν_e had been converted into ν_μ or ν_τ on their way to the Earth [16].

All of these results are well-explained within the framework of three-flavor neutrino mixing. In the next Sections we give an overview of the main neutrino interactions, the standard theory of neutrino oscillations in vacuum and its extension to oscillations in matter.

1.2 Neutrino interactions

As neutrinos only interact through weak forces the cross sections for all processes are extremely small (order of 10^{-38} cm^2 at the GeV scale) and the mean free path through a solid medium is of the order of tens of light years. In the electroweak theory of the Standard Model, neutrino interactions can occur both via neutral (NC) and charged (CC) current interactions:

$$\begin{aligned} \text{CC} &: \nu_l + N \rightarrow l + X \\ \text{NC} &: \nu_l + N \rightarrow \nu_l + X \end{aligned} \tag{1.2}$$

where l denotes the leptonic flavor, N the nucleon or nucleus, and X stands for the hadronic final state. In the simplest case of neutrino-lepton interaction, N and X in equation 1.2 are substituted with leptonic states. In particular, neutrino-electron scattering, due to its high directional nature, has been extensively exploited for solar neutrino detection. In case ν interacts with a nucleus, the neutral current process can be detected when the nucleus breaks apart as a result of the interaction or, in neutrino coherent scattering experiments, by detecting the low-energy nuclear recoil given to the nucleus kicked by the neutrino. At low momenta, where a nucleon is less likely to be emitted from the nucleus, neutral current events are usually undetected. We underline that NC are flavor independent. On the contrary, in CC interactions the flavor of the lepton provides the flavor of the incoming neutrino [17].

Neutrino cross sections for different interaction processes highly depend on incoming energy. Among the threshold-less processes ($E_\nu \sim 0 - 1$ MeV), figure coherent scattering, namely scattering of neutrino with an entire nucleus, and neutrino capture on radioactive nuclei [17]. For higher neutrino energies, it is possible to penetrate the nucleus at smaller scales and to access nucleons individually. At $E_\nu \sim 1 - 100$ MeV inverse beta decay (IBD) processes start gaining importance. Due to the importance of inverse beta decay for the

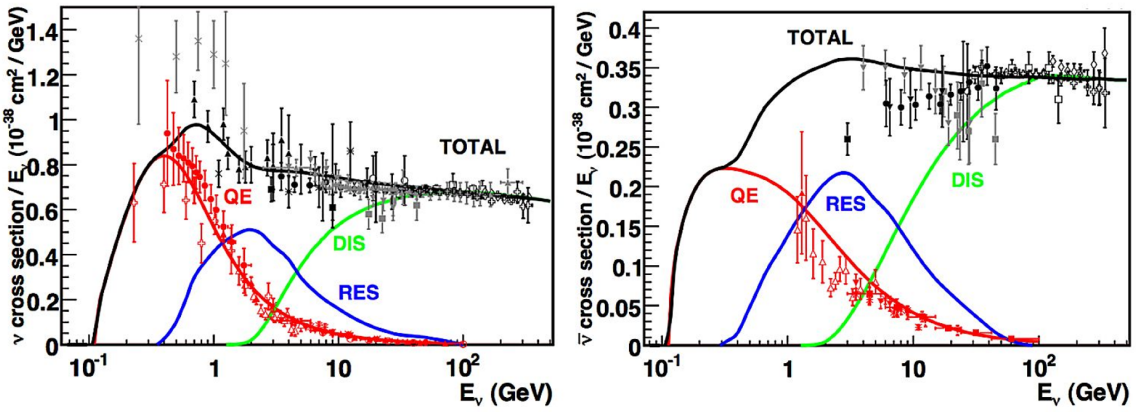


FIGURE 1.1: Total neutrino and antineutrino per nucleon CC cross sections divided by energy as a function of neutrino energy. QE, resonant production (RES), and DIS contribution are shown. Data from various experiments together with the prediction from the NUANCE event generator are also displayed and more details about them can be found in the original Reference [17].

analysis treated in this thesis, it will be discussed in more detail at the end of this Section. For interactions occurring above $\sim$ 2 GeV, due to the high momentum transferred, the nucleon is generally broken up and the final state is composed of a hadronic shower and a lepton (this process is called Deep Inelastic Scattering, DIS). For lower energies ($E_\nu \sim 0.1 - 20$ GeV) the dominant interaction is the quasi-elastic (QE) process, where the hadronic final state consists in a single nucleon. In quasi elastic interactions, having two particles in the final state, the energy of the incoming neutrino can be inferred from the energy and direction of the lepton. The QE interactions apply to both NC and CC modes:

$$\begin{aligned} \text{CCQE} &: \nu_l + n \rightarrow l + p \\ \text{NCQE} &: \nu_l + n(p) \rightarrow \nu_l + n(p) \end{aligned} \tag{1.3}$$

The nucleon hit by the neutrino can also be excited into a resonant state, such as a Δ^{++} or Δ^{+} for CC interactions exchanging a W^+ boson and a Δ for NC interactions with a Z boson. In this case we can have the production of a single pion from the Δ-decay (CCπ^+, NCπ^0). The resulting π^0 is a major source of background in neutrino experiments because it can be misidentified as a lepton of an electron neutrino event and therefore, a good knowledge of resonant NC interactions is of relevance for neutrino oscillation studies. Neutrino and antineutrino cross-sections as a function of the energy are shown in Fig. 1.1.

1.2.1 Inverse beta decay

The inverse beta decay is one of the simplest nuclear interaction channel for neutrinos. It consists in an antineutrino-proton scattering of this form:

$$\bar{\nu_e} + p \rightarrow e^+ + n \tag{1.4}$$

where a positron and a neutron are emitted. Here protons are free, namely hydrogen nuclei, and not bound in heavier nuclei, for which nuclear binding effects suppress interactions at low energies [18]. Figure 1.2 shows the neutrino interaction cross sections in water as a function of energy. It is obtained from the study presented in Reference [19]:

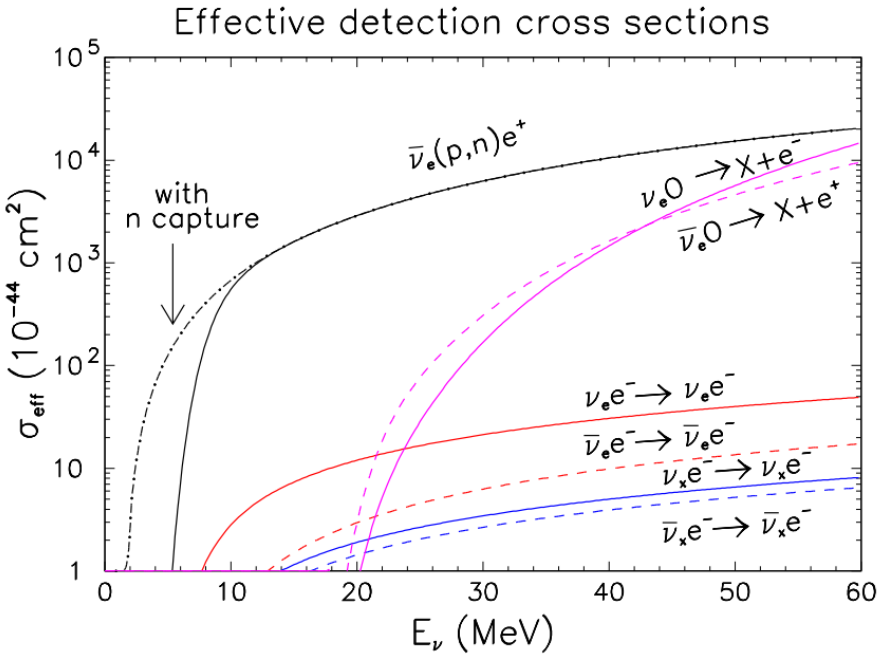


FIGURE 1.2: ν interaction cross sections as a function of neutrino energy for low energy neutrinos ($E_\nu < 60$ MeV) in Cherenkov detectors. Energy resolution and threshold effects are estimated in Reference [19] and included in the model. The possibility of neutron capture on Gd nuclei, enabling the detection below threshold is also shown (black dotted line). From [19].

specific results for water Cherenkov detectors are shown. The inverse beta decay cross section is the largest below 100 MeV; below 30 MeV it is two orders of magnitude larger than electron neutrino elastic scattering, the second-leading channel. This is the reasons why IBD is the preferred detection channel for reactor and supernova anti neutrinos in water Cherenkov detectors. [20]. Inverse beta decay can be initiated only if the incoming anti neutrino kinetic energy is at least 1.806 MeV [21]. This threshold energy can be computed from the difference, in mass, between the products and the scattered particles. Since e^+ have a mass which is more than 3 order of magnitude smaller than the n mass, most of the incoming energy involved in the IBD process is transferred to the positron. For this reason, e^+ plays a key role for detection: it annihilates right after its creation, producing visible energy that can be approximated, at the 0^{th} order, as [22]:

$$T_{e^+} = E_{\bar{\nu}_e} + m_p - m_n - m_{e^+} = E_{\bar{\nu}_e} - 1.806 \tag{1.5}$$

In a Cherenkov or scintillation detector the signal generated from positron interaction is typically called the “prompt event" since the Cherenkov radiation or scintillation light is emitted immediately after the IBD. On the contrary, the neutron signal, referred to as the “delayed event", occurs only after neutron thermalization and capture. In the particular case of IBDs in water, neutron captures on hydrogen nuclei with the emission of a 2.2 MeV γ ray after a thermal propagation of about 200 μs. The reaction is:

$$n + p \rightarrow d + \gamma \tag{1.6}$$

The specific time and space correlation of prompt and delayed events provides a clear signature for IBD detection favoring the discrimination from background signals.

1.3 Neutrino masses

The Standard Model has to be formulated with massless particles in order to guarantee gauge invariance. What gives mass to all the particles is the process of spontaneous symmetry breaking through the so-called Higgs mechanism: interaction between massless particles and a scalar field, the Higgs boson. In particular, fermions acquire masses interacting with the Higgs field with a strength regulated by the coupling to the Higgs boson, called Yukawa coupling [2, 23]. This coupling, expressed in terms of a Lagrangian is:

$$L_{Yukawa} = \sum_{\alpha,\beta=1}^{3} \bar{u}_L^{\alpha} m_{\alpha\beta} u_R^{\beta} + \bar{d}_L^{\alpha} m_{\alpha\beta} d_R^{\beta} + \bar{l}_L^{\alpha} m_{\alpha\beta} l_R^{\beta} \tag{1.7}$$

where l denotes the charged leptons, u and d are the two types of quark fields respectively, and α, β label the three fermion generations (e, μ, τ for leptons and u, c, t and d, s, b for quarks). The L and R subscripts indicate the left and right chirality of the fields. The fields present in Equation 1.7, together with the gauge boson part (here not shown), are those that take part in weak interactions. At this stage, the mass matrices $m_{\alpha\beta}$ are in general non-diagonal, namely the entries outside the main diagonal are in general non-zero, so that they cannot define the physical masses of the particles involved. As the formula structure in Equation 1.7 may suggest, mass terms for fermions require a right handed (RH) and a left handed (LH) component in order to maintain gauge invariance. Neutrinos do not appear in the Lagrangian in Eq. 1.7 because, at the time of the model formulation, they were considered to be massless and a right-handed neutrino was not necessary in the Standard Model formulation (massless fermions do not couple with the Higgs field). Although the absolute value of ν mass remains unknown and only upper limits at eV level have been obtained by β-decay experiments [24], the discovery of neutrino oscillations made clear that at least 2 generations of neutrinos have to be massive [15]. Since a non-zero mass for ν contradicts the original Standard Model formulation, possible extensions of the theory need to be developed.

Dirac case

The most straightforward way to define neutrino mass is to extend the theory by including a right-handed neutrino field, analogously to the Yukawa coupling of other massive fermions. The Yukawa Lagrangian would pick up an extra term:

$$L_{mass} = \bar{\nu}_L^{\alpha} m_{\alpha\beta} \nu_R^{\beta} + h.c. \tag{1.8}$$

where ν denotes the neutrino field and $\alpha, \beta = e, \mu, \tau$ are the three neutrino flavors which couple to the charged lepton flavors in the weak interaction.

As we mentioned above, mass terms are generally not diagonal, so that to define the physical masses of the neutrinos, the mass matrix must undergo a diagonalization process. Thus, as a first step, two unitary matrices $U_{L,R}$ have to be chosen:

$$m' = U_L^{\dagger} m U_R \tag{1.9}$$

which is equivalent to re-defining the fields as:

$$\begin{aligned} \nu_R^{'i} &= U_R^{\beta i} \nu_R^{\beta} \\ \nu_L^{'i} &= U_L^{\beta i} \nu_L^{\beta} \end{aligned} \tag{1.10}$$

The primed fields are called mass eigenstates. The $U_{L,R}$ matrices are unitary and it is therefore equally correct to say that the flavor eigenstates, labeled with α and β, are superposition of the mass eigenstates, labeled with i in eq. 1.10, and vice-versa. The mass eigenstates are those which describe the propagation through space and time while the flavor states obey gauge interactions rules and therefore are the ones coupling with weak bosons. The U matrices that connect the flavor eigenstates to the mass eigenstates of neutrinos are called PMNS (Pontecorvo-Maki-Nakagawa-Sakata) matrices. By convention, all the mixing parameters are attributed to the neutrino sector while the weak eigenstates for charged leptons are the same as their mass ones: in fact there is no oscillation observed for e, μ and τ. Quarks have non trivial mixing matrix but they will not be discussed in this work.

Finally, we remind that weak interaction mediators only couple to left-handed fields, but there is no fundamental reasons why right-handed (also called “sterile") neutrinos should not exist as singlets of the Standard Model gauge group. If they were to exist, sterile neutrinos would be extremely difficult to detect, since neutrinos only interact through the weak and gravitational interaction.

Majorana case

In 1937 Majorana suggested that the neutrino is a self conjugate particle [25]. In fact, since neutrinos are electrically neutral, the charge symmetry does not impose neutrinos to be Dirac: only neutral fermions can be described by a Majorana field. The Majorana condition is:

$$\nu^c = \nu \tag{1.11}$$

Where ν^C represents the charge conjugated neutrino field: charge conjugation is a transformation that switches all particles with their corresponding antiparticles, by changing the sign of all charges. As a consequence, a Majorana spinor is only composed of two independent fields and cannot distinguish particles from antiparticles. Assuming the Majorana condition for neutrinos, it becomes possible to write additional terms in the Standard Model Lagrangian which are Lorentz scalars and just as viable as the Dirac terms ($\bar{\nu}\nu$). More details on the form of Majorana terms can be found in Reference [23]. Analogously to the Dirac case, to transform the flavor eigenstates into the corresponding mass eigenstates the new mass matrix can be diagonalized using unitary operators. The diagonalization lead to a rewriting of the original neutrino fields as a linear combination of two different fields, each of them with mass m_1 and m_2 as eigenvalues of mass matrix, describing the space-time propagation. The Majorana description of a massive neutrino has different phenomenological consequences with respect to the Dirac case: an important one is that the two new fields obey to the so-called See-Saw mechanism [26]. In a very simplistic description, we can say that the two mass values, m_1 and m_2, are strongly correlated in a way that an increase of one of the masses brings as an effect the decrease of the second one. To justify the unexpected small neutrino mass, it has been proposed that the neutrino with the smaller mass contains mostly the left-handed component of the neutrino field: this neutrino interacts with matter and it is called “active" neutrino. On the contrary, the heavy neutrino contains mostly the right-handed component: it is thus sterile with respect to weak interactions. This mechanism is very appealing since it provides a natural explanation for the lighter neutrino masses with respect to the charged fermions but it requires the neutrinos to be Majorana particles. At present, however, we have no evidence that neutrinos obey the Majorana requirement in equation 1.11. In order

to prove that ν is a Majorana spinor several experiments are trying to detect the neutrinoless double β-decay: a process that would not occur if neutrinos were a Dirac particle [27, 28].

1.4 Neutrino oscillations

The misalignment of neutrino mass eigenstates and flavor eigenstates, as illustrated in equation 1.10, is the root cause of the neutrino oscillation phenomenon. Eq. 1.10 clearly states that the mixing terms are non zero only if the mass matrix is non trivial, therefore neutrinos oscillate only if at least two of them are massive [2, 23].

As mentioned in the previous section, the three neutrino mass eigenstates $k = 1, 2, 3$ are related to their flavor eigenstates, $\alpha = e, \mu, \tau$ by the PMNS matrix U:

$$|\nu_\alpha\rangle = \sum_k U^*_{\alpha k} |\nu_k\rangle \tag{1.12}$$

The PMNS matrix can be parametrized by three mixing angles and three complex phases. It can be demonstrated that if neutrinos are Dirac particles then two of these phases are non-physical while if they are Majorana particles the neutrino mixing matrix U is multiplied by additional matrix factors and two extra phases arise which do not affect neutrino or anti-neutrino oscillations. This said, in what follows the PMNS matrix will be parameterized by three mixing angles θ_{12}, θ_{13}, and θ_{23} and only one complex phase δ:

$$\begin{aligned} U &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix} \begin{pmatrix} c_{13} & 0 & s_{13}e^{-i\delta} \\ 0 & 1 & 0 \\ s_{13}e^{i\delta} & 0 & c_{13} \end{pmatrix} \begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{13}c_{23} \end{pmatrix} \end{aligned} \tag{1.13}$$

where s and c stand for respectively sine and cosine of the θ angles, indicated with the corresponding numbers. The upper part of equation 1.13 makes explicit a factorization in terms of the three distinct rotations. This representation is very useful in practice because the three angles correspond to oscillations visible in atmospheric/accelerator, reactor/accelerator, and reactor/solar experiments, respectively in the same order as the matrices are presented. A particular experiment can be sensitive to one or more matrix parameters depending on the ν type detected, experimental characteristics of the apparatus, baseline/energy and the mass splitting between the neutrinos.

In order to infer the equation of oscillation probability we can start to consider the massive neutrino states as eigenstates of the Hamiltonian: they satisfy $H|\nu_k\rangle = E_k|\nu_k\rangle$ with energy eigenvalues $E_k = \sqrt{\vec{p}^2 + m_k^2}$. The evolution is determined by the Schrodinger equation with a time dependent term written in such a way that, for plane wave solutions, the flavor composition as a function of time can be calculated as:

$$|\nu_\alpha(t)\rangle = \sum_{\beta=e,\mu,\tau} \left(\sum_k U^*_{\alpha k} e^{-iE_k t} U_{\beta k} \right) |\nu_\beta\rangle \tag{1.14}$$

Therefore, if the mixing matrices are not diagonal, a flavor eigenstate at time $t = 0$ evolves, at $t > 0$, into a superposition of different flavor states. More details can be found in reference [23]. A flavor state α, produced at time $t = 0$, has a probability to be detected in

a different flavor state β at a time t:

$$P_{\nu_\alpha \to \nu_\beta}(t) = |\langle \nu_\beta | \nu_\alpha(t) \rangle|^2 \tag{1.15}$$

Applying the approximations for ultra-relativistic neutrinos, if we define the squared-mass difference $\Delta m^2_{kj} \equiv m^2_k - m^2_j$, the oscillation probability in a practical form for experiments able to measure L (travelled distance) and E (energy) is:

$$P_{\nu_\alpha \to \nu_\beta}(L, E) = \sum_{k,j} U^*_{\alpha k} U_{\beta k} U_{\alpha j} U^*_{\beta j} exp\left(-i \frac{\Delta m^2_{jk} L}{2E}\right) \tag{1.16}$$

The transition probability for which $\alpha \neq \beta$ is called the “appearance probability" while $\alpha = \beta$ case is called the “disappearance probability". Equation 1.16 shows a clear dependence of the oscillations on the distance between the detector and the sources L, the neutrino energy E, and the squared-mass differences. The amplitude of the oscillations are influenced solely by the PMNS mixing matrix U. The oscillation length is inversely proportional to the squared-mass splitting of each state. The squared-mass differences and the components of the PMNS matrix are physical constants of nature and subject to measurement by neutrino oscillation experiments.

For anti-neutrinos, oscillation probabilities can be obtained with a similar procedure but substituting the terms in eq. 1.16 with their complex conjugates: the quadratic product of U gain an opposite sign in the imaginary part of the complex phase ($U^* \neq U$). As a consequence, neutrino and anti-neutrinos have different oscillation probabilities.

A charge parity (CP) transformation links neutrinos with anti neutrinos (interchanges neutrino and anti neutrino while reversing helicity)

$$\nu_\alpha \to \nu_\beta \overset{CP}{\longleftrightarrow} \bar{\nu}_\alpha \to \bar{\nu}_\beta \tag{1.17}$$

In the case of $\delta \neq 0$ but also $\theta_{12}, \theta_{13}, \theta_{23} \neq 0$ or $90°$, by measuring the CP asymmetry

$$A^{CP}_{\alpha\beta} = P_{\nu_\alpha \to \nu_\beta} - P_{\bar{\nu}_\alpha \to \bar{\nu}_\beta} \tag{1.18}$$

neutrino oscillations can probe CP violation in the mixing matrix. Finally, the transitions are invariant under CPT:

$$P_{\nu_\alpha \to \nu_\beta} = P_{\bar{\nu}_\beta \to \bar{\nu}_\alpha} \tag{1.19}$$

1.4.1 Neutrino oscillations in matter

The expression for the transition probability presented in eq. 1.16 holds if neutrinos propagate in vacuum. In 1978 L. Wolfenstein—pioneer of the electroweak theory—suggested that the propagation of neutrinos in matter is perturbed by an effective potential due to coherent forward scattering with the particles that compose the medium [29]. This effect leads to a modification of the flavor states evolution. In 1985 Mikheev and Smirnov generalised Wolfenstein’s work and noticed that neutrinos traveling through a medium with varying density (the most common example is neutrinos coming from a very dense source like stars) tend to incur in a resonant flavor transition characterized by an effective maximal mixing angle of $\pi/4$. This resonant amplification is called the Mikheev-Smirnov-Wolfenstein (MSW) effect [30].In their work a particular enphasis is reseved to describe this effect in the Sun. For neutrinos propagating in matter, an extra term, which depends on the potential exerted by the medium, needs to be added to the Hamiltonian H in the flavor basis and as a result, the probability for changing flavors in matter can

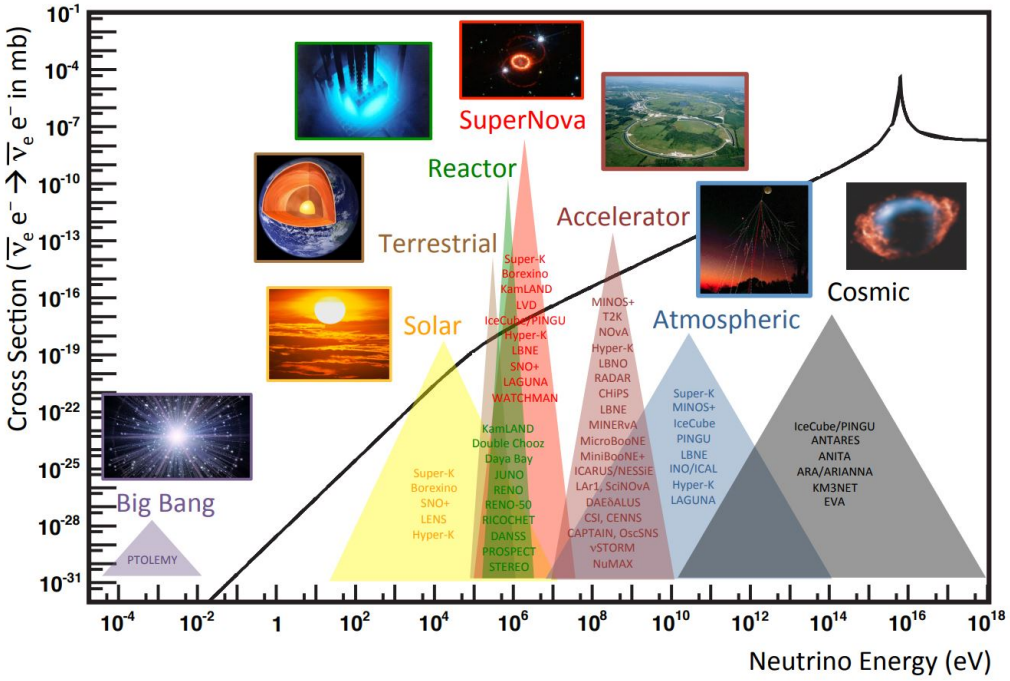


FIGURE 1.3: For illustration: neutrino interaction cross section as a function of energy, showing typical energy regimes accessible by different neutrino sources and experiments. The curve shows the scattering cross section for $\bar{\nu}_e$ elastic scattering on free electrons. From [17].

be rather different than in vacuum and, in particular, strongly enhanced by the resonance condition.

1.5 Overview of neutrino sources

Neutrinos are produced by a wide range of sources through various physics processes that can be detected with different technologies according to the experimental goals. In this section six main categories will be distinguished: solar, atmospheric, astrophysical, geological, reactor and accelerator neutrinos. More generally, neutrinos can also be classified in terms of ν flavor, energy, baseline length and other physical or experimental properties. An illustration of neutrino sources is shown in Figure 1.3, where each neutrino source is described as a function of its energy range and relative cross section. In this Section, a brief overview of the most important neutrino sources is presented while the next section is fully dedicated to a more detailed presentation of supernova relic neutrinos (SRN), due to their importance for this work.

The largest neutrino flux we get on Earth comes from solar and atmospheric neutrinos. The experimental and theoretical study of solar neutrinos is a major topic in neutrino research. Neutrinos from the Sun represent a powerful probe of solar interiors allowing us to study stellar fusion processes from a unique perspective. Moreover, given the specific Sun-Earth distance and the characteristics of the neutrino spectrum, the study of solar neutrinos gives access to oscillation parameters that could not be measured by terrestrial means (specifically the parameters of the third rotation matrix in the factorized part of Eq. 1.13). Solar neutrinos are electron neutrinos originating within the Sun's core at a distance of about 1.5×10^8 km. In the Sun, conversion of protons into neutrons is extremely abundant due to the fact that thermonuclear fusion reactions continuously transform hydrogen

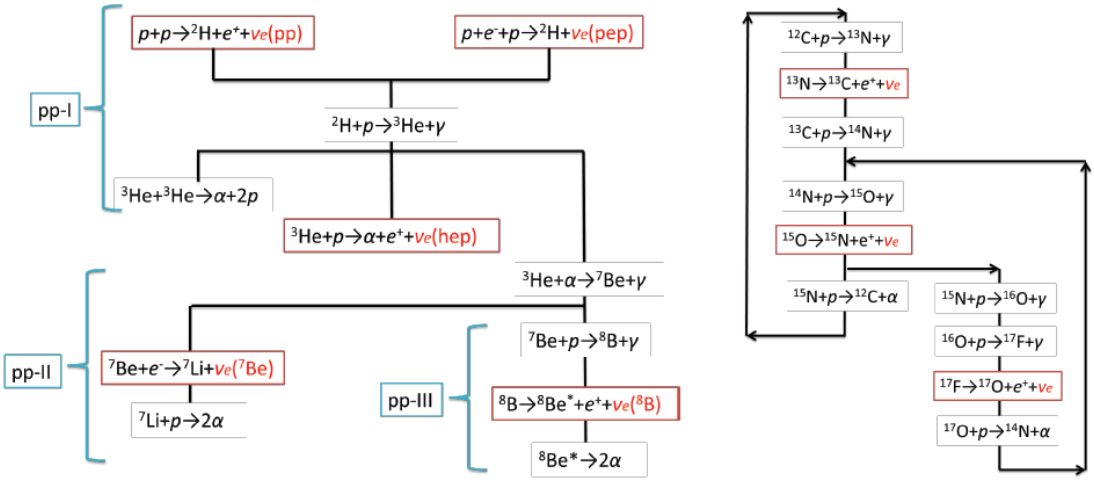


FIGURE 1.4: The schematic illustration of pp-chain (left) and CNO cycle (right). From [33].

in helium within the core, and these processes are accompanied by neutrino production. A well-established theoretical model, the Standard Solar Model, describes solar evolution and predicts the expected solar neutrino fluxes from different fusion chains [31]. Solar energy is generated from thermonuclear reactions within the core where four protons undergo fusion reactions generating ^{4}He, two positrons and two electron neutrinos:

$$4p \rightarrow \alpha + 2e^+ + 2\nu_e + 26.7MeV \tag{1.20}$$

where the energy is released with the emission of photons from the surface [32]. This reaction proceeds via two different branches: proton-proton chain (pp chain) represents 98.4% of the total energy release while Carbon-Nitrogen-Oxygen cycle (CNO cycle) only 1.6% of the total solar luminosity [33]. Figure 1.4 shows a schematic representation of the two fusion chains. ν_e are emitted during the conversion process with mean energy of about 1 MeV [23]. In particular, solar ν_e can be emitted at different steps of the chains, though reactions with fixed or distributed ν energies as shown in Figure 1.5. Due to the weakness of neutrino interactions, solar neutrinos easily escape the Sun's surface and travel freely in space, finally reaching the Earth with a flux of about 6×10^{10} cm^{-2} s^{-1} [34]. Despite this extremely large flux, the vast majority of solar ν_e arrives at Earth's surface with very low energy, a mean value < 0.4 MeV, below the sensitivity threshold of most neutrino experiments. Therefore, most experiments mainly rely on the rarest and most energetic solar neutrinos: ^{8}B solar neutrinos and similarly ν from hep chain with an energy of the order of 10 MeV. In order to investigate such low energies, backgrounds such as the spallation-induced radioactivity studied in this work and treated in Sec. 1.7, need to be drastically reduced. In particular, the Super-Kamiokande experiment measures the solar neutrino flux by detecting elastic scattering interactions [23]:

$$\nu_\alpha + e^- \rightarrow \nu_\alpha + e^- \tag{1.21}$$

This detection channel privileges electron neutrinos, whose interaction cross-section is about six times larger than the one for ν_μ and ν_τ. In order to distinguish solar ν_e from generally isotropic background events, directional correlation of the recoil electron with Sun's position is exploited.

Atmospheric neutrinos are mainly muon neutrinos in the GeV range arising from

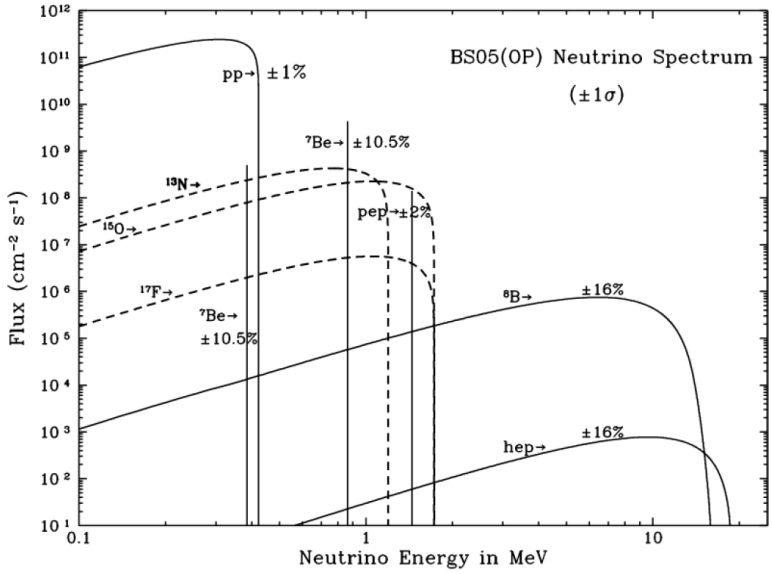


FIGURE 1.5: The solar neutrino energy spectrum at Earth's surface predicted by [31]. The values in percentage correspond to total flux uncertainties.

hadronic showers triggered by collisions of cosmic rays with the Earth's atmosphere (distance of about 10^4 km). Indeed, Earth's atmosphere is constantly bombarded by cosmic-ray particles, mainly high-energy protons and, to a smaller extent, heavier nuclei. The collision with air molecules in the atmosphere lead to the production of these hadronic showers. Pions, which are the most abundant particle created as a consequence of cosmic-ray interactions, decay into muon neutrinos and muons, that can, in turn, decay generating ν_e. Kaons contribute to the flux of atmospheric neutrinos at higher energies. Atmospheric neutrinos are produced with a very wide range of energies, from a few MeV to several hundreds of GeV, reflecting the fact that the primary cosmic-ray particles also have a huge energy range.

Among astrophysical neutrinos we can list supernova neutrinos from core-collapse supernovae (discussed in the next Section), neutrinos from extragalactic and galactic sources of various nature (black holes, pulsars, energetic explosions called gamma-ray bursts, and active galactic nuclei) and relic neutrinos originating from the Big Bang. Due to their ability to travel through dense matter, neutrinos can provide crucial information about the dynamics of these astrophysical sources. SN 1987A, a supernova that exploded on February 1987 in the Large Magellanic Cloud, a satellite galaxy smaller than the Milky Way (about 1.6×10^5 light years of distance), is nowadays the only extragalactic astrophysical object detected through its neutrino burst. 11 neutrinos were detected by Kamiokande II [35], 8 antineutrinos by IMB experiments [36] and 5 by Baksan [37].

In addition, among the natural sources of neutrinos geo neutrinos come from the decay-chain of ^{238}U and ^{232}Th within the Earth.

There also exist two man-made sources neutrinos which have significantly contributed to the measurements of ν properties: those coming from nuclear power plants (reactor neutrinos: $\bar{\nu}_e$ from β-decay with an energy of few MeV) and specifically designed neutrino beams created in accelerator complexes (accelerator neutrinos). The energy of accelerator neutrinos can be tuned by changing the primary proton energy and secondary hadrons momentum selection. It is usually of the order of hundreds of MeV to a few GeV.

1.6 Supernova relic neutrinos

The term Supernova Relic Neutrinos (SRN) refers to a cosmological neutrino background, often called Diffuse Supernova Neutrino Background (DSNB), arising from the totality of neutrinos emitted from all past supernova explosions everywhere in the observable universe. Being a diffuse background, the SRN signal is isotropic and time-independent, permeating the space with a cosmic energy density comparable to that of photons from stars, about 0.01 eV cm^{-3} (10 times smaller than cosmic microwave background) [18]. Observing the SRN spectrum would further our understanding of supernova dynamics, which is crucial to interpret cosmic history and stellar evolution. Moreover, ν from supernova explosions arise from extreme conditions, impossible to replicate in terrestrial laboratories, which can shed light on neutrino properties which may be still unknown or poorly-known. In this section we will illustrate supernova explosion dynamics, supernova relic neutrino properties and the status and characteristics of the SRN search.

1.6.1 Core-collapse supernova

Supernovae are some of the most energetic phenomena in the universe: they mark the catastrophic death of stars and they can lead to massive explosions or the formation of black holes. The kinetic energy emitted with the burst can reach 10^{51} erg [23]. Due to their extreme brightness, many extragalactic supernovae have been observed by optical instruments over the centuries. It is however quite a rare event within our galaxy: their rate is estimated to be of few events per century [23, 38, 39].

Supernova classification

Supernovae are classified into different types: as illustrated in Figure 1.6 two main categories, type I and II, are distinguishable by their spectroscopic characteristics (with or without hydrogen lines, respectively). Subcategories exist for further classification. The trigger mechanism of the explosion is another feature used for supernovae identification: bursts of type Ia supernovae are provoked by thermonuclear reactions, namely drastic fusion processes in stars supported by pressures of the Fermi gas degenerate electrons; these processes happen in relatively small stars, such as white dwarfs. Ia supernovae are the brightest of all categories. On the contrary, supernovae triggered by gravitational forces are categorized as Ib, Ic and II. In this case the star's core collapses under the huge gravitational pressure of the star. This occurs only if the mass of the progenitor star satisfies $M > 6M_{sun}$, where $M_{sun} = 1.989 \times 10^{30}$ kg [39]. The vast majority (99%) of the energy released by core-collapse supernovae is carried away by neutrinos emitted in large numbers and in all directions. Due to their importance in neutrino physics, from now on we will focus only on core-collapse supernovae.

Supernova burst dynamics

Although the dynamics of core-collapse supernova bursts can depend on the progenitor chemical composition and mass, several common processes involved in the explosion can be distinguished and can be summarized in different phases, as described below [23, 40, 41]. Figure 1.7 shows a schematic illustration with the different steps of the explosion, whose numbers are reported for reference along the text (in parenthesis).

- *Photon-disintegration (1-4)*: stars initially support themselves against gravity with pressure produced by nuclear fusion. This process ends when silicon burning produces iron nuclei, whose high binding energy makes it very stable. Once the core is

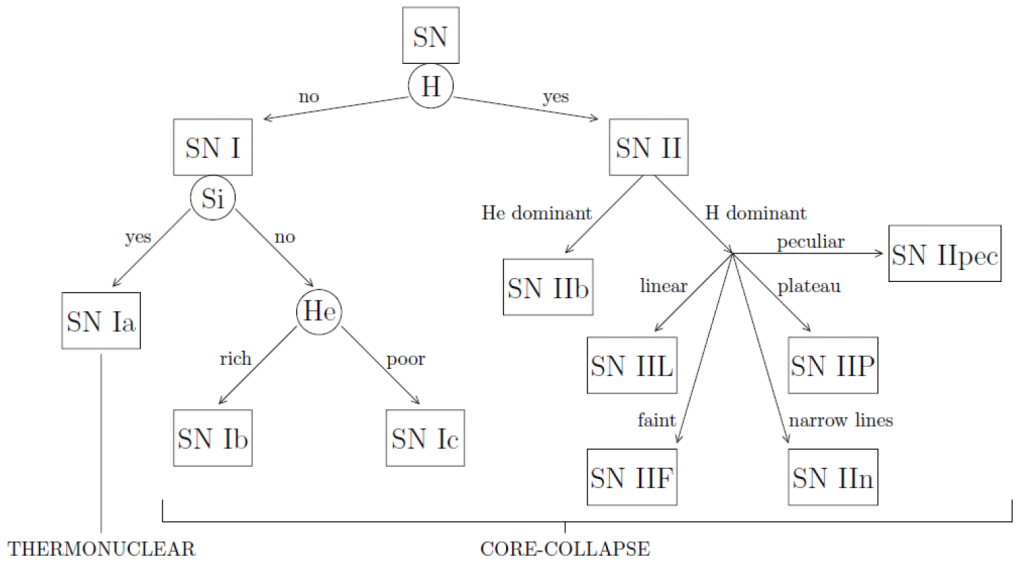


FIGURE 1.6: Scheme of supernovae classification, from [23].

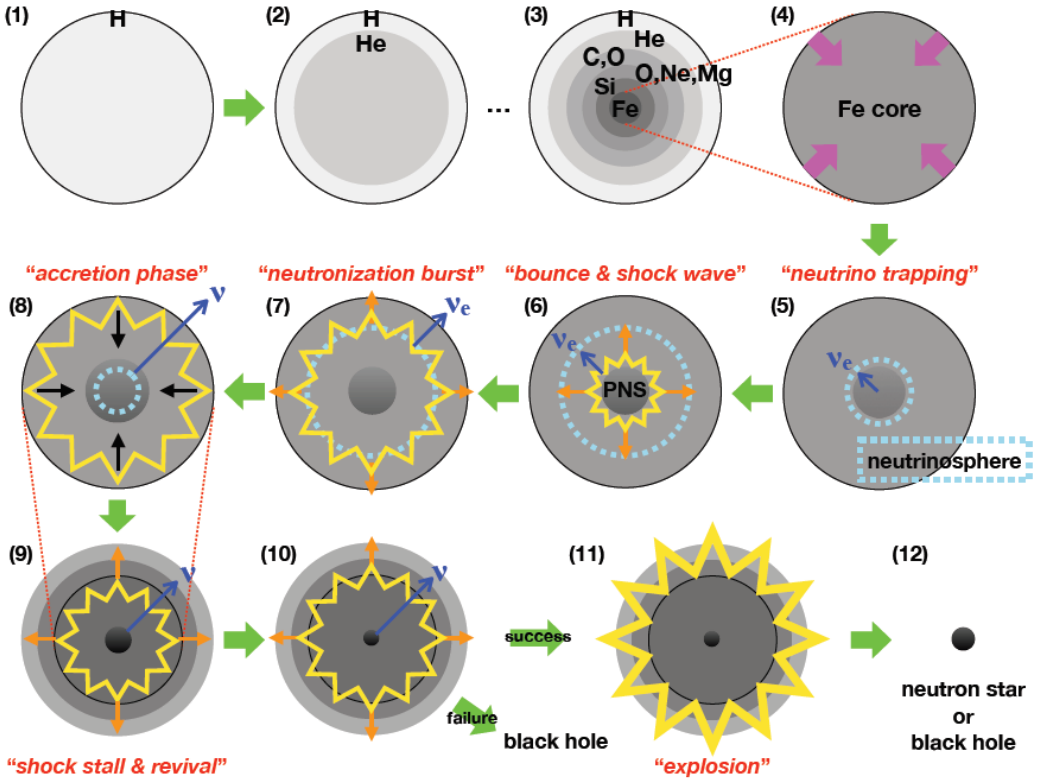


FIGURE 1.7: Illustration of supernova explosion dynamics in its different phases. Reference to the numbers of this Figure (from 1 to 12) can be found in the text with further descriptions. Taken from [39].

mostly composed of Fe, and therefore no more fusion combustible is available, the star's development is marked by a subtle equilibrium between the pressure of the degenerate gas of relativistic electrons in the core and that from gravitational forces. However, while core temperature and density keep rising, electron captures on Fe

$$^{56}Fe + e^- \rightarrow ^{56} Mn + \nu_e \tag{1.22}$$

become more and more favored. As a consequence, the electron degeneracy pressure starts decreasing and core collapse is accelerated. Moreover, endothermic photodisintegration of Fe nuclei enhances core instability. Electronic neutrinos produced at this stage, although they represent a negligible fraction of the total ν emission, escape freely from what is called the "neutrinosphere", namely the surface where neutrinos are scattered last and from which they can be emitted. It divides the stellar material transparent to neutrinos to the region where they are trapped.

- *Neutrino trapping (5)*: Once the core density reaches $\rho = 10^{11}$ g cm^{-3}, the neutrinosphere is contained into the star core. This means that ν are trapped, undergoing very frequent coherent scattering off nuclei

$$\nu_e + N \rightarrow \nu_e + N \tag{1.23}$$

which increase the opacity of the core medium.

- *Core bounce (6-7)*: the collapse of the star core starts decelerating once density approaches nuclear values ($\rho = 3 \times 10^{14}$ g cm^{-3}). The repulsive nuclear pressure becomes so intense that, when the collapse stops, an outward shock wave is triggered. Dissociation of nuclei into free nucleons reinforce the production of electron neutrinos through electron capture processes

$$e^- + p \rightarrow n + \nu_e \tag{1.24}$$

just behind the shock. ν_e produced within the neutrinosphere remain trapped due to the medium opacity, while those generated after the shock wave passes the neutrinosphere are decoupled from matter and freely propagate ahead. This results in a sudden emission of free electron neutrinos which goes under the name "neutronization burst". The luminosity peak reaches 10^{53} erg s^{-1} for a duration of about 10 ms (see Figure 1.8). The total energy released is of the order of 10^{51} erg. For galactic supernovae, the signal caused by this electron neutrino burst is expected to be detectable by Super-Kamionande. Simultaneously, in the innermost part of the core a proto-neutron star is formed.

- *Accretion phase (8)*: After the passage of the shock wave, matter (nucleons, electrons, and positrons) starts falling onto the star core: gravitational energy transforms into thermal energy leading to the production of positrons whose captures induce the generation of many $\bar{\nu}_e$ via $e^+ + n \rightarrow p + \bar{\nu}_e$. This phase is called "accretion phase" since matter start to accumulate on the proto-neutron star. In addition to antineutrinos, high-energy neutrinos of all species are thermally produced in the hot and dense star core throughout different reactions, such as electron-positron annihilation, nucleon-nucleon or electron-nucleon bremsstrahlung and ν-ν annihilation:

$$\begin{aligned} e^+ + e^- &\rightarrow \nu + \bar{\nu} \\ N + N &\rightarrow N + N + \nu + \bar{\nu} \\ e^\pm + N &\rightarrow e^\pm + N + \nu + \bar{\nu} \\ \nu_e + \bar{\nu}_e &\rightarrow \nu + \bar{\nu} \end{aligned} \tag{1.25}$$

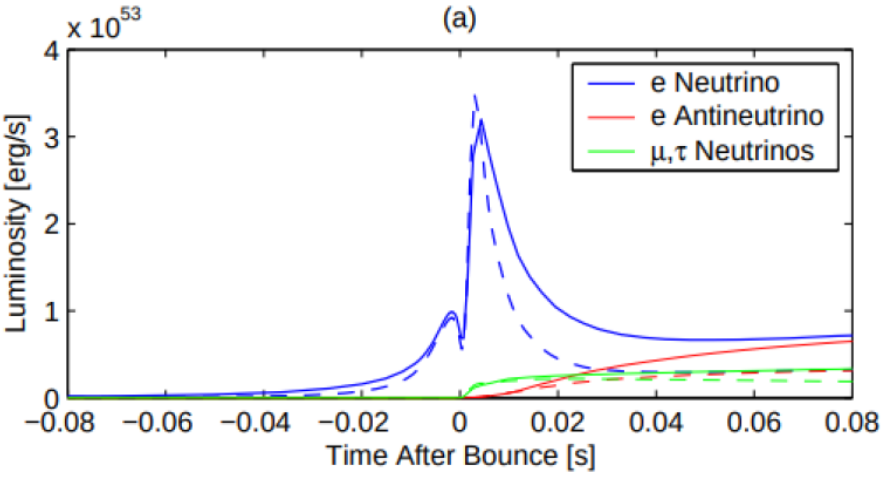


FIGURE 1.8: Luminosities of neutrino emission as a function of time, where $t = 0$ is at the core bounce. Dashed lines represent simulation results for a start with 13 solar masses while solid lines refer to a start with 40 solar masses. The first peak before zero represents the first neutrino emission due to electron captures on Fe. The second peak, few ms after the bounce, is the neutronization burst. Taken from [40].

.

Further electron neutrinos are generated via CC interactions:

$$e^- + p \rightarrow n + \nu_e \tag{1.26}$$

The production of electron antineutrinos remains dominant and the majority of them is generated during this phase.

- *Delayed explosion (9-12)*: the shock wave loses energy and stalls due to accretion material pressure, nuclear dissociation into free nucleons and neutrino emission. At this point there is an equilibrium between the stalled accretion and the proto-neutron star. A new mechanism plays a key role to revive the shock wave and lead to the star explosion. A leading hypothesis to explain the shock wave revival is the neutrino-heating scenario: matter behind the shock wave is heated by neutrino absorption which provides new energy for the wave to pass into the star envelope, where energy dispersion is negligible. If the shock wave is energetic enough to blow off the outer layers of matter, an explosion occurs. If not, matter continues to fall onto the core and a black hole is formed. The timescale of the explosion is about 1 s after the bounce. Neutrinos from core collapse supernovae emerge at energies of about 10 MeV for over 10 s, in comparison to photons, emitted with few eV energies for several months.

- *Cooling (12)*: if an explosion occurs, the star core and remnants keeps cooling by emitting neutrinos of all flavors and becomes a neutron star or a black hole depending on the initial mass and the level of metallicity of the progenitor.

1.6.2 Supernova relic neutrinos

Supernovae, being the last phase of the life of massive stars, are characteristic events of the evolution of the universe since its early beginning. Therefore, it is reasonable to assume that all neutrinos originating from supernova explosions throughout the history of the universe have dispersed in free space and accumulated to form a diffuse background [42].

Detecting the SRN flux would provide unique insight into the star formation history and the dynamics of supernovae. Moreover, studies on baryogenesis, leptogenesis and Big Bang nucleosynthesis would benefit from a better understanding of SRN features [43].

The shape of the accumulated flux of supernova relic neutrinos strongly depends on neutrino emission spectra while the overall normalization is linked to the supernova rate, in turn depending on the cosmic star formation rate. Mass and properties of the progenitor star as well as the fraction of stars failing to explode and collapsing into black holes, also impact the spectral shape. In particular, for black-hole-forming supernovae the accretion phase lasts longer, and higher-energy neutrinos are emitted. In addition, oscillation effects impact the spectrum shape in different ways depending on the mass hierarchy. A prediction of the SRN differential flux can be given by [44]:

$$\frac{dF_\nu}{dE_\nu} = c \int_0^{z_\infty} (1+z) R_{SN}(z) \frac{dN_\nu(E'_\nu)}{dE_\nu} \frac{dt}{dz} dz \tag{1.27}$$

with c being the speed of light, R_{SN} the supernova rate at redshift z, and dN_ν/dE_ν being the neutrino energy spectrum —assumed independent of progenitor mass— emitted from a supernova explosion with energy at production $E'_\nu = (1+z)E_\nu$, where E_ν is the detected energy. t is the cosmic time as defined in the Friedmann equation:

$$\frac{dz}{dt} = -H_0(1+z)\sqrt{\Omega_m(1+z)^3 + \Omega_\Lambda} \tag{1.28}$$

where H_0 is the Hubble constant, Ω_m is the cosmic matter density, and Ω_Λ is another cosmological constant.

A variety of models have been formulated to predict the SRN flux: the supernova rate was historically calculated from cosmic chemical evolution measurements combined by stellar evolution theory and, more recently, with star formation rate models [22, 44]. The results obtained from various experimental and theoretical studies are accurate enough to give the rate of core collapses. In addition, they are supplemented by direct measurements of the optical supernova rate, which have lower precision but show good agreement. On the other hand, there are still open questions related to neutrino spectrum: the many models proposed show slight variations related to the uncertainties on neutrino behaviours in very dense materials. In fact, even though the only available supernova neutrino data from SN 1987A cannot provide high-precision predictions, considered the very low statistics (only 24 ν where detected) [35], it had been crucial to determine the spectrum main properties. In particular, time and energy spectra as well as luminosity from SN 1987A allowed to improve and confirm simulation results. Neutrino flavor changes and self-interactions in dense material, a vivid subject in actual neutrino physics, can add the distortion effects differentiating the various models. Figures 1.9 shows SRN $\bar{\nu}_e$ fluxes for various models [39]. The variability of the models is strongly linked to the parameter dependence of the SRN flux described in Eq. 1.27. More details on representative SRN models and predictions are summarized in Ref. [39, 45, 46].

1.6.3 Experimental search

Supernova bursts emit neutrinos and anti-neutrinos of all flavors with energies ranging from few MeV to tens of MeV, see Figure 1.9. The privileged detection method depends on the type of experiment and especially on the cross-section values of the possible interactions. An important ingredient to take into account to estimate detection probability is the effective area of the detector, namely the neutrino-proton cross section multiplied by

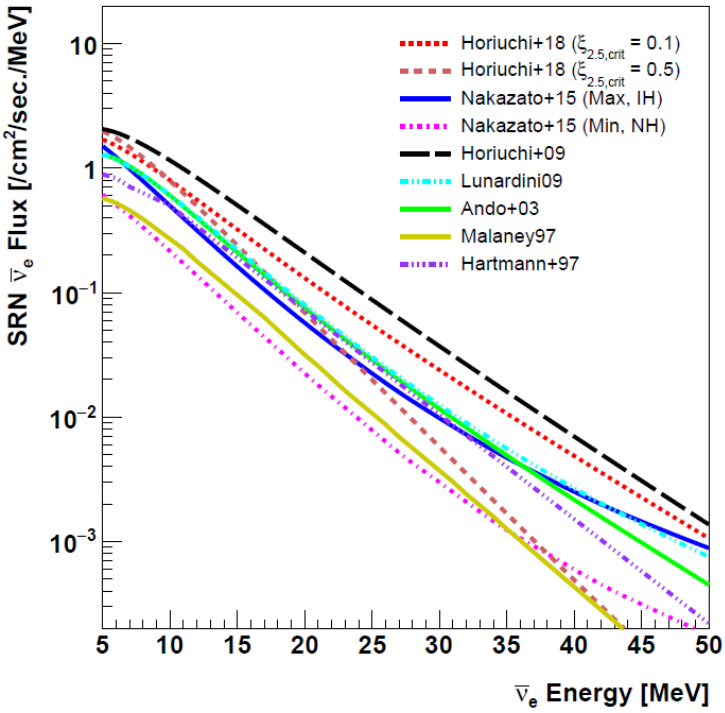


FIGURE 1.9: Supernova Relic Neutrino fluxes for different theoretical models. Details on listed models are presented in Ref. [39], where the Figure is taken from.

the total number of target protons. Given the variety of experiments and the impracticability of treating them all exhaustively in this work, we decided to focus only on water Cherenkov detectors since most of the topics presented along this thesis are linked to the Super-Kamiokande experiment. The dominant neutrino interaction in water Cherenkov detectors is the inverse beta decay of $\bar{\nu}_e$: as illustrated in Figure 1.2, the IBD cross section is at least one order of magnitude larger than those of other interactions in the SRN energy range, making IBD the only channel considered for the SRN search. Details on the IBD channel are summarized in Section 1.2.1: we remind that the advantage of this reaction is the presence of two signals which can be detected in coincidence, a prompt one from the positron Cherenkov radiation and a delayed 2.2 MeV γ from neutron capture on H. In the following, a short description of Cherenkov process is presented followed by details on SRN search in Super-Kamiokande (SK), the most sensitive Cherenkov detector for supernovae neutrinos. A brief state of the art of the SRN search is also introduced.

Cherenkov radiation

A relativistic particle, propagating in a medium with a velocity larger than the speed of light in the material produces a shock wave accompanied by light emission, called Cherenkov radiation. The direction of Cherenkov light can be calculated from classical physics principles: atoms and molecules excited along the charged particle trajectory induce coherent radiation which propagates in the medium. In fact, radiation is not emitted by the charged particle itself, but from the atoms in the detector medium, which remain polarized if the charged particle passes in their vicinity. For relativistic particles, with a speed greater than the speed of light in the medium, there is a symmetric configuration of polarized atoms around the particle track and these particles, after receiving a local

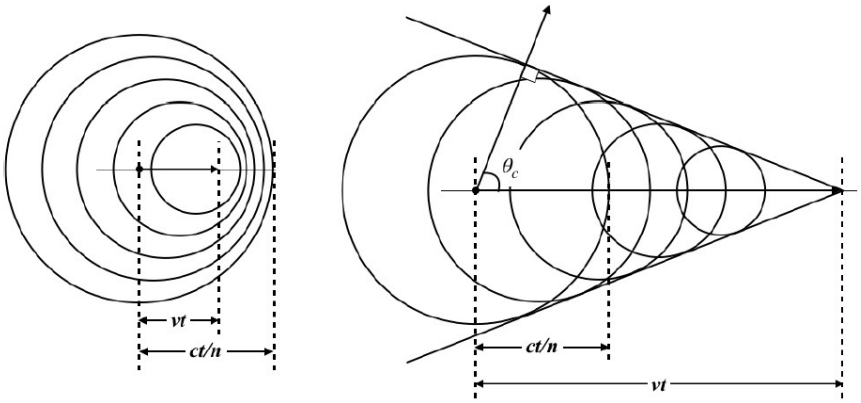


FIGURE 1.10: Illustration of coherent radiation for charged particles with $\beta < c/n$ (left) and $\beta > c/n$ (right, Cherenkov radiation case). A particles moves from left to right during a time t in the medium. It can cross the distance $x = vt$. The radiation travels radially at a distance $x_c = ct/n$, therefore the angle between the direction of Cherenkov photons and particle propagation is $\theta_C = x/x_c$. From [47].

electromagnetic boost, produce light in the form of photons of the wavelength peaked at $\sim$ 400 nm in water, in order to de-excite to their normal state. A Cherenkov light cone is therefore produced by a cascade of photons that constructively interfere with each other, where the constructive interference increases with radiation intensity. The emitted light is only a small fraction of the charged particle energy loss. The Cherenkov threshold, namely the kinetic energy value above which a particle begins to emit Cherenkov light, depends on the refractive index of the material, and on the type of particle. The Cherenkov threshold is given by:

$$E_{thr} = \frac{m}{\sqrt{1 - n^{-2}}} \tag{1.29}$$

where m is the particle mass and n the refractive index [23, 46]. Therefore, in pure water (refractive index is about 1.33 at 589 nm), threshold is 0.76 MeV for electrons, 116 MeV for muons, 156 MeV for charged pions, and 1.04 GeV for protons [39]. Cherenkov photons are mainly emitted in the ultra-violet/visible range, their multiplicity decrease with increasing wavelength and their total yield is in general lower than the radiation produced by the same particles in scintillating detectors. The emitted radiation propagates in a cone with axis along the particle track as illustrated in Figure 1.10 (both relativistic and non-relativistic particles are shown for comparison). Its opening angle θ_C, namely the angle between the velocity of the propagating particle and the emission direction of the Cherenkov photons, can be computed as

$$cos\theta_C = \frac{1}{\beta n(\lambda)} \tag{1.30}$$

with β being the ratio between speed of the Cherenkov particle in the medium and speed of light in vacuum, $n(\lambda)$ the refractive index in the material, which is a function of the wavelength. In pure water, θ_C for ultra relativistic particles ($\beta \sim 1$) is about 42 degrees [48]. At the energies we are considering, electrons and positrons can be considered ultra-relativistic, while heavier particles such as μ and π radiate in cones with lower opening angle. Thanks to the specific properties of the above described processes, a water

Cherenkov detector can infer the energy and time of the radiating particle from the number of Cherenkov photons emitted and their arrival times. In addition, directional information can be obtained by looking at the Cherenkov cone angle, knowing that it is ~ 42 degrees for ultra-relativistic particles. This informations is useful in neutrino detection if the recoil particles produced by neutrino interaction maintain directional informations, as for electron neutrino scattering used in solar analysis. Given that, in Super-Kamiokande the SRN analysis exclusively targets the detection of electron antineutrinos via inverse β-decay processes, the directional information is not useful for signal detection. In fact, in the case of IBD events, the positron produced does not contain informations of incident neutrino direction. However, incident angle reconstruction is an important quantity for background rejection in SRN searches: for example solar neutrino background removal cuts, being solar neutrino flux dominant for energies below 19 MeV, exploit the fact that neutrinos strongly concentrate to the opposite direction of the Sun, in contrast to the isotropic distribution IBD. More details on SRN background sources are presented in the following.

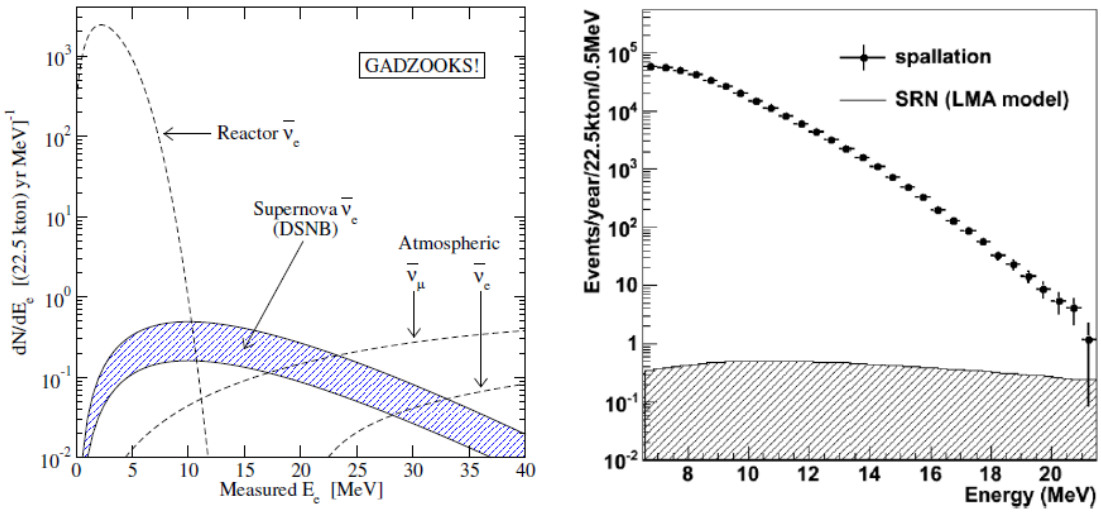
SRN background sources at Super-Kamiokande

The major sources of background for the SRN signal depend on the detection channel and the detection medium of the experiment considered. Since we are focusing on water Cherenkov detectors, with a specific interest in Super-Kamiokande experiment, we will focus on the background sources associated with these detectors. SRN searches in SK focuses on the detection of IBD signals in the energy range from a few MeV to tens of MeV. The strategy is to identify the prompt positron light emission and pair it with the neutron delayed signal. Given the rarity and weakness of the SRN signal, understanding the background sources is crucial. We anticipate, before treating the subject in Chapter 2, that Super-Kamiokande is sufficiently large and deep underground to keep a moderate background rate. Nonetheless its modelization and reduction is imperative for extremely challenging signals like diffuse supernova neutrinos. For reference, Figure 1.11 shows relevant “IBD-like" backgrounds in SK in the left part and spallation contamination on the right. Within the favored energy range of about 10-30 MeV, the lower part of the spectrum suffers from reactor neutrinos contamination while a wide range of astrophysical processes dominate the high energy region (over about 20 MeV). Atmospheric background includes both IBD-like signals, shown in Fig. 1.11 and reactions where gammas are produced and mis-identified as positrons. At very low energies, below about 6 MeV, intrinsic radioactivity dominates: this background is controlled with accurate cuts, purification of water with special water circulation systems to minimize radon ingress, detector material choices, and reduction algorithms, like fiducial volume cuts (more details in Chapter 2). Conversely, between 6 and about 20 MeV spallation backgrounds from the β and $\beta\gamma$ decays of isotopes produced through cosmic muon interactions largely dominates. Backgrounds for the SRN search in Super-Kamiokande are listed below: [22, 39, 44, 50, 51]:

- **Atmospheric neutrino interactions**: a CC process of the type

$$\bar{\nu_\mu} + p \rightarrow \mu^+ + n \tag{1.31}$$

  where the antineutrino comes from atmospheric sources, can mimic an IBD signal if the muon is invisible, namely below Cherenkov threshold, and decays producing a Michel positron. If these conditions stand, this background is irreducible. Analogously, atmospheric $\bar{\nu}_e$ CC interactions, despite their energy range is higher than SRN search energies, are indistinguishable from IBD processes due to their identical



(a) Neutrino-induced background. (b) Muon spallation background.

FIGURE 1.11: SRN background sources. Left panel: spectra of low energy IBD neutrino events. It is assumed full efficiencies, and it includes energy resolution and neutrino oscillations.The upper edge of the SRNs band is set by the SK limit [22], and the lower edge by modern models. Left figure taken from [49]. Right panel: event rate as a function of electron energy for spallation and SRN events. The expected SRN spectrum from LMA model is plotted in shadow. Right figure taken from [46].

signature; however, these processes typically involve neutrinos that are more energetic than SRNs. Their spectra is predicted by theoretical models. Neutral current events also contributes to SRN atmospheric background. NC events are dominated by quasi-elastic scattering with nuclei of the form

$$\nu + {}^{16}O \rightarrow \nu + {}^{15}O + \gamma + n \tag{1.32}$$

The knocked-out nucleon, after interacting with the atmospheric ν, leaves the nucleus in an excited state and a subsequent gamma emission is possible. In addition, the initial nucleons can propagate in water initiating secondary processes. A Cherenkov angle cut can partially remove this background: NCQE interactions emit in fact multiple photons whose overlapping Cherenkov cones are reconstructed as a single cone with particularly large opening angle.

- **Reactor neutrinos**: actual $\bar{\nu}_e$ from nuclear reactors in Japan and South Korea that undergo IBD and produce the same exact signal as SRN. Their spectrum, below 10 MeV, may overlap the SRN spectrum at lower energies due to energy resolution effects.

- **Spallation background**: despite its location 1000 m underground, the Super-Kamiokande's natural shielding is not enough to stop all cosmic ray muons penetrating the Earth's surface. Therefore, the muon rate at SK depth is about 2 Hz [52]. Although it is relatively simple to identify muon tracks, thanks to their huge energy deposits, reactions induced by muon interaction in water may lead to SRN-like signals. In fact, energetic muons can collide and break up oxygen nuclei eventually leading to the creation of excited isotopes, called "spallation isotopes", whose β, γ or n decay products fall into the energy range of SRN search. Many spallation events can be

removed with time and geometrical cuts around reconstructed muon tracks, however long-lived radioactive isotopes still form a serious background source in the energy range below 20 MeV, namely the maximum reconstructed energy of beta spectrum for spallation products [53]. The left panel of Figure 1.11 shows the energy spectrum of spallation events in SK and predictions from an optimistic SRN model [46]: the SRN signal is completely drowned by the spallation background, which is more than 5 orders of magnitude larger. Several types of isotopes are produced from spallation reactions in SK: they include β decaying isotopes like ^{11}Be and ^{16}N, whose high yield and long lifetime make them a source of background if their signal is detected in coincidence with fake hits, due to PMT noise or radioactivity. In addition, isotopes like ^{9}Li and ^{8}He, whose yield is one order of magnitude lower, still represent a possible fake signal because their β decay is accompanied with a n, resulting is a signal indistinguishable from IBD [39, 53]. Since a simple extension of geometrical spallation cuts can veto events within seconds and meters from muon tracks, expanding experimental dead time and wasting precious signal, a more complex rejection method must be developed to remove spallation events. Advances in spallation cuts are in progress and imply a better knowledge of spallation patterns as well as an improvement of neutron tagging techniques. The recent upgrade of Super-Kamiokande, SK-Gd, aims to increase neutron capture efficiency, first to 50% and then to 90% in the future, with the addition of a Gd salt. Its high capture cross section and increased photon energy makes neutron captures considerably more visible than in water [54]. Finally, it is estimated that, if background sources could be very accurately modeled and removed, SK has already the potential sensitivity to detect the SRN signal [44]. Therefore, it is extremely important to improve our ability to exclude spallation events, as it is one of the most serious source of background. This is the reason why spallation reduction is one of the major topics of research in the SRN domain.

Conclusively, supernova relic neutrinos are extremely rare events with an upper limit of few IBD events per year [18, 50, 51, 55]. As show in Figure 1.11, their search is made extremely challenging for the presence of overwhelming background whose modelization and reduction is indispensable for the progress of the SRN analysis.

Status of the SRN search

To date, no conclusive evidence for SRN has been found, despite the huge experimental efforts, principally by the Super-Kamiokande detector [50, 51] and the KamLAND liquid scintillator experiment [56]. Upper limits on the SRN flux for both experiments are shown in Figure 1.12.

The first search at SK was performed, with 1496 live days, with a spectrum fit which included signal and background predictions. It resulted in an upper limit on SRN $\bar{\nu}_e$ flux with and energy threshold of $E_\nu > 19.3$ MeV [42]. Improvements on the analysis technique and an higher statistics, collected in 2855 days, allowed to reach an energy range starting from E_ν>17.3 MeV [51]. A major upgrade to the electronics of the detector enabled the acquisition of data after the IBD primary events, making possible the inclusion of neutron tagging techniques [50]. As a consequence, a better background reduction, especially for signals without neutrons, like spallation decays, was achieved in the low energy region and energy threshold was lowered to 13.3 MeV. Neutron tagging is still subjected to high incertitudes due to its very low efficiency [39].

KamLAND is a scintillator detector of 0.6 kiloton fiducial volume. Unlike Cherenkov detectors, neutron tagging efficiency is almost 100% due to the high light yield and low

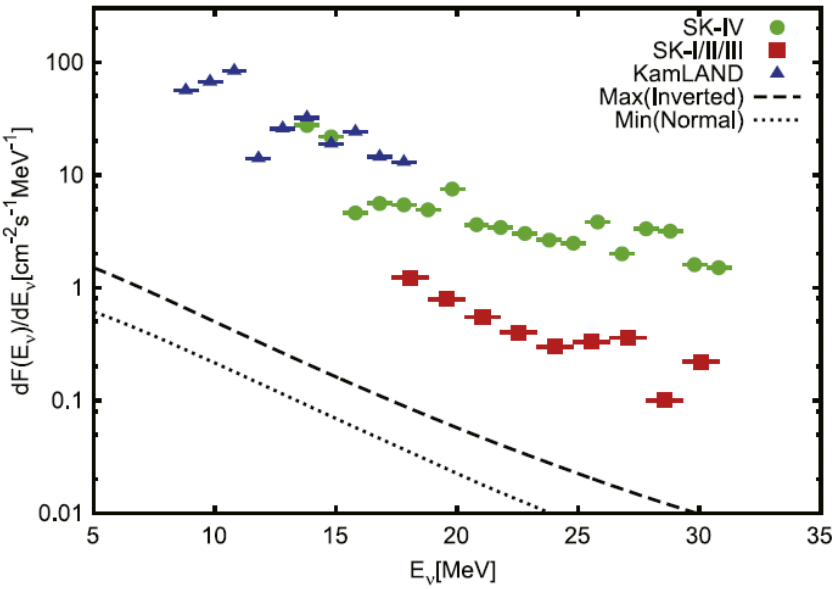


FIGURE 1.12: Upper limits on the $\bar{\nu}_e$ flux from searches in the Super-Kamiokande and the KamLAND experiments. Label "SK-IV" corresponds to results from [50],"SK-I/II/III" from [51]. KamLAND results are from Ref. [56]. Black solid lines correspond to the maximum and minimum predictions in models described in Ref. [57].

detection threshold, KamLAND sensitivity in the low energy region is competitive to SK and it gave the best upper limit of SRN flux below 17.3 MeV.

1.6.4 The challenge of the SRN search

The very low event rate expected in SK for supernovae relic neutrino signals makes its search the most challenging of the experiment. A deep understanding and accurate modelization of background sources is essential if we aim to finally isolate this extremely rare signal. In particular, in the crucial spectral region below 20 MeV, spallation background is widely dominant, being 10^6 times larger than the weak SRN flux. Although other low energy signals suffer from spallation contamination, for none of them is this background as overwhelming as in the case of SRN: in this analysis cosmic induced radioactivity needs to be drastically and carefully reduced by several order of magnitude. Since many isotopes decay with the emission of a single β ray, neutron tagging becomes a powerful technique to reduce spallation contamination. Despite the advantages of this technique, accidental pairing of radioactive decay signals with fake neutrons due to PMT dark noise or intrinsic radioactivity can mimic an IBD event. Moreover, a fraction of spallation isotopes decay emitting a β accompanied by a neutron: these constitute an irreducible source of background that cannot be removed with neutron identification methods. Among these (β+n)-decaying nuclei, ^{9}Li has the highest production rate, an half-life of 0.18 s and endpoint energy of about 10 MeV. A sophisticated model describing muon induced radioactivity and predicting (β+n)-decaying isotopes is required, making the study of cosmic muon reactions a priority for the future of SK. One more problem is the fact that signals from neutron captures on hydrogen have a very low detection efficiency. Despite the performances of the rejection technique improve with the addition of gadolinium, with the actual concentration the efficiency of Gd captures is still lower than 50%. Last but not

least, deeper understating of the spallation background is crucial for next-generation water detectors. Among others, Hyper-Kamiokande (HK) [58], which is under construction at a topological site with 600 m of mountain overburden, is expected to be contaminated by a larger fraction of muon-induced radioactivity. Also, it is not yet decided if Gd will be dissolved in HK. Finally, spallation studies are a top priority if we aim to detect supernova relic neutrinos both for SK-Gd and Hyper-Kamiokande. Although many searches can greatly benefit from accurate models of muon reactions in water, SRN discovery will only be possible if this background will be deeply understood. For all these reasons, the work of this thesis is centered on the construction of an accurate simulation tool able to give interpretations and predictions of the spallation background as well as being functional for reduction technique improvements. The goal is to provide current and future analysis with valuable calculations that can lead toward future discoveries.

In the next Section, a first theoretical description of spallation processes is presented.

1.7 Muon-induced spallation processes

We described in Section 1.6 how spallation events induced by cosmic ray muons constitute an extremely challenging background for the supernova relic neutrino search in Super-Kamiokande. Other low-energy analyses, like solar and reactor neutrinos, are affected by spallation contaminations. Both solar and reactor neutrinos have energy ranges lower than supernova neutrinos, between about 4 and 10 MeV, but they are not as rare (about 15 solar ν_e per day compared to an upper limit of few events per year for SRN) [59]. Nonetheless, induced radioactivity from cosmic-ray muons, dominating in the energy range from about 6 to 20 MeV, constitutes the major background for these additional searches. Especially for solar searches, where the absence of neutrons among the final products precludes the application of neutron tagging techniques for background reduction, the removal of radioactive isotope contamination is crucial. In order to isolate rare signals, sophisticated techniques derived from deeper theoretical understanding of muon spallation reaction in water are required. In this Section the main physics principles inducing spallation reactions are summarized, together with an overview of cosmic muon production and energy losses.

1.7.1 Cosmic ray muons

Particles accelerated at astrophysical sources, commonly called “primary cosmic rays", are continuously bombarding the Earth’s surface. Their interaction with Earth’s atmosphere, namely the collisions of primary cosmic ray particles with air nuclei, results in cascades of secondary particles, in turn leading to the creation of tertiaries and their subsequent showers. The energy splits and redistributes among all the participants at each new cascade step and the secondaries start acquiring transverse momenta while the shower develops longitudinally. Muons constitute an important fraction of the shower particles: they are predominantly generated by pions and kaons, through the following decay channels:

$$\pi^{\pm} \rightarrow \mu^{\pm}\nu_{\mu}(\bar{\nu}_{\mu}) \qquad (\sim 100\%) \tag{1.33}$$

$$K^{\pm} \rightarrow \mu^{\pm}\nu_{\mu}(\bar{\nu}_{\mu}) \qquad (\sim 63.5\%) \tag{1.34}$$

Several factors make atmospheric muons very conspicuous, not only at sea level, but also underground: these particles lose a constant and small amount of energy per track length ($\sim$ 2 GeV totally in the atmosphere [60]). In addition, they have a relatively long lifetime (2.197 $\times 10^{-6}$ s) and a small interaction cross section. Since π are typically produced at altitudes of about 15 km and decay relatively fast, muons are largely produced in higher

atmosphere. Their travel through air is characterized by two main compelling processes: interaction and decay. The dominance of one process over the other depends on the lifetime value of the particle and changes as a function of the kinetic energy and on the local density of the air in the atmosphere, which is in turn related to the altitude. The average muon energy at gound surface is about 4 GeV and the spectrum shape follows directly the convolution of meson source spectrum, energy losses along the path and decay probability [60]. It is important to mention that vertical muon intensity at sea level is not a fixed quantity: meaningful fluctuations can be induced by geomagnetic position, altitude, solar activity and specific conditions of the atmosphere.

Since several muons can be generated in the same air cascade from a common parent cosmic ray, they are expected to arrive almost at the same time in the plane perpendicular to the shower axis. Experimental studies estimated an arrival time spread smaller than few ns between the different muons [61]. When this situation occurs, the multiple muon tracks are referred as “muon bundles". Given that secondary particles in cosmic ray induced showers have in general small transverse momenta ($\approx$ 300 MeV/c), high energy muons are globally collimated near the cascade axis. Several meters of lateral displacements for energetic muons (at TeV scale energies) can be typically obtained only deep underground. Despite muon bundle parameterizations, both at surface and for a given depth, exist [61, 62], most frequently, many models of atmospheric muon flux use the assumption of single muons only.

Due to their small energy loss and long lifetime, muons represent the most penetrating component of cosmic rays: despite being shielded, due to their deep position under the sea level, underground neutrino telescopes observe a large flux of very energetic atmospheric muons. The composition and thickness of the material overburden, as well as the shape of the higher part of the energy spectrum and the types of energy losses contributes to define the energy and angular distributions and the intensity of penetrating muons.

Muon energy losses

For muons, two main types of energy losses can be observed: continuous and discrete. Ionization is responsible for the former: it weakly depends on the primary muon energy and it is generally assumed to be constant for relativistic particles. Delta-ray emission contributes to the continuous part of energy deposition although hard collisions induce higher fluctuations. Ionizing energy loss processes are dominant for muon below 1 TeV in water, at this critical value discrete interactions such as bremsstrahlung, pair production and reactions with atomic nuclei start gaining importance. The average muon energy loss can be parametrized as [63]:

$$-\frac{dE}{dx} = \alpha(E) + \beta(E)E \tag{1.35}$$

Where α is the ionization energy loss, with the typical value of 2 MeV cm^2 g^{-1}, and β is the sum of pair production, bremsstrahlung, and photo nuclear contributions. Ionization can be further divided into restricted ionization, characterized by soft collisions with small fluctuations and delta-ray production, namely hard collisions with atomic electrons, leading to large fluctuations in the energy transfer. Pair production and bremsstrahlung dominates the radiative part of the energy loss for muons at hundreds of GeV, whereas the number of photo nuclear interactions starts growing at higher energies [53]. When the interaction is radiative, the energy is no longer deposited continuously, instead large fluctuations can be observed along the muon path. Processes with large energy transfers, notably bremsstrahlung and photonuclear interactions, have a stochastic nature leading to a considerable range straggling. Larger losses are in general more probable in bremsstrahlung then pair production. In addition, in the case of radiative processes,

cosmogenic muons passing through matter produce a flux of secondary particles, primarily of electromagnetic nature (gamma rays and electrons) which in turn generate an electromagnetic shower. Showers generated by muons can also have an hadronic nature, frequently producing pions, neutrons and other particles. In what follows, we will refer to showers induced by cosmic muon interactions with oxygen as electromagnetic or hadronic depending on whether they are initiated by electromagnetic particles or by hadrons. Occasionally, if these daughters, or the muon itself, are energetic enough a spallation nuclear break-up can occur leading to the creation of lighter isotopes. Nuclear fragments can in turn be unstable and subsequently decay. More details on spallation processes are presented in the following.

1.7.2 Physics principles of spallation reactions

Spallation processes are one branch of nuclear reactions. A spallation reaction occurs when an energetic particle incurs a collision with a complex target nucleus. The projectile energy can have values from 100 MeV to 3000 MeV. If the transferred energy is high enough to exceed the binding energy between nucleons, the nucleus disintegrates into lighter fragments. Thus, nucleons from the original target are "spalled" or knocked-out after the bombardment of the energetic projectile. Residual fragments results in smaller mass and atomic numbers [64, 65]. The disintegration process may lead to different secondary particles: simple atomic constituents as proton and neutrons, light or heavy nuclei depending on the target nature and elementary particles.

Spallation reactions are commonly described as a two stage process [66]. As a first step, the projectile particle interacts with nucleons, neutrons and protons, inside the target nucleus. This collision leads to the creation of a nuclear cascade where energetic (greater than 20 MeV) protons, neutrons and pions start developing within the nucleus as a consequence of nucleon-nucleon collisions [67]. Among the hadrons created during this stage, some manage to escape as secondary particles while others deposit their kinetic energy inside the nucleus itself which in turn reaches an excited state. The above described process, intervening between incident particle and individual nuclear component, is referred to as intra nuclear cascade (INC). The second stage, called nuclear de-excitation, consists in the statistical de-excitation of exited nuclei that relax through two main competing processes: evaporation and fission [68]. In case of evaporation, which consists in the emission of low energy particles (less than 20 MeV), mostly neutrons, protons and alphas, the remaining nucleus can conserve a radioactive nature and gamma rays or beta particles can be ejected. Figure 1.13 illustrates the two processes involved in spallation reactions in a schematic way. More details will be given in the following.

Intra nuclear cascade

INC is a very fast process through which nucleons composing the target nucleus increase their kinetic energy as a result of hard collisions with the projectile particle, heating up the entire nucleus. The initial nuclear cascade take place within 10^{-22} s, the typical time needed by incident particle to cross the nucleus. The reason why this stage is treated as a series of individual nucleon collisions derives from the fact that the De Broglie wavelength of the projectile is smaller than the diameter of the target nucleus. As an example, a proton with 1 GeV kinetic energy has a De Broglie wavelength value $\lambda = h/\sqrt{2mE} \approx 10^{-14}$ cm, smaller than heavy nuclei dimensions [69]. As a consequence of the large energy transfer, nuclear components may leave the nucleus and interact with other targets inducing in turn other spallation reactions. These processes are called inter nuclear cascade, in opposition to the interaction occurring inside the nucleus, and mostly involve neutrons,

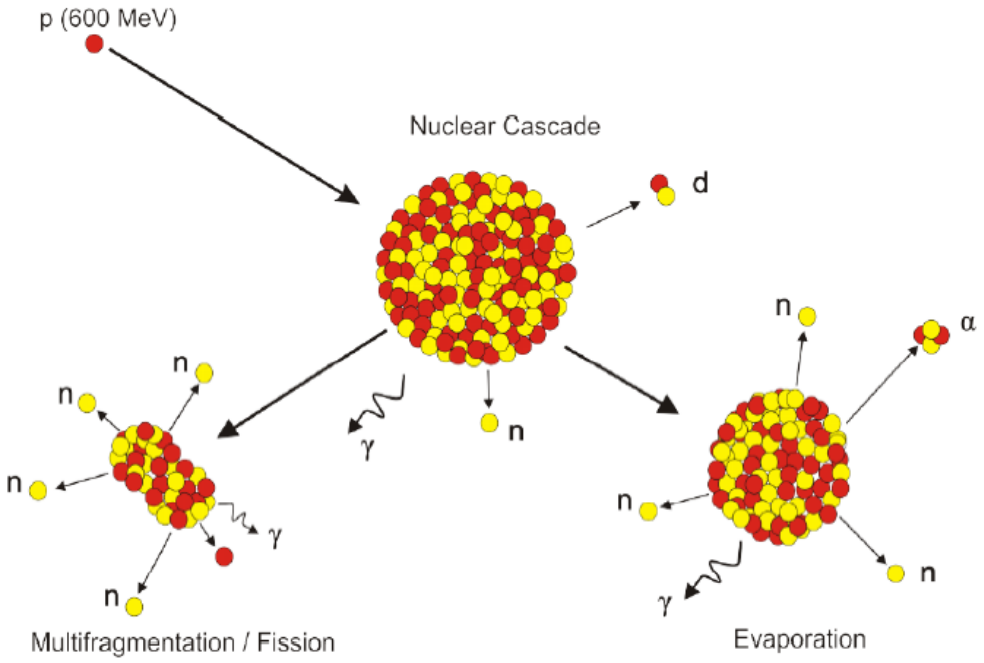


FIGURE 1.13: Scheme of a spallation reaction triggered by a 600 MeV proton. The two stage de-excitation process is illustrated. This figure is taken from [64].

protons and secondary charged pions. Figure 1.14 schematically represents the processes above described.

As we discussed, intra nuclear cascade processes involve multi-scattering interactions of individual nuclear components, each of them separated in space and time. This situation, referred as “collision regime", occurs only if the De Broglie wavelength of the projectile particle is smaller than the nucleus dimension, which depends both on the type of target nucleus and incident particle but most of all on the kinetic energy of the colliding particle. If the incident energy is very low, the De Broglie wavelength is almost as large as the nucleus and the dynamics of the process will be dominated by mean field effects. In the case of an incident energy larger than 20 MeV, the projectile particle mean free path becomes larger than the inter nucleon distances (λ is less or of the order of this length) and mean field approach starts to smoothly evolve into collision regime [70]. Processes occurring at these energies are called “nucleon cascade", in opposition to interactions that arise with growing incident particle energy: particle production thresholds begin to be crossed and new hadrons can be produced consequently to inelastic nucleon-nucleon interaction or collision between newly produced hadrons. This stage is called “hadronic cascade". Finally, at energies between 10 and 200 GeV, thus at De Broglie wavelength smaller than nuclear components dimensions, processes called “partonic cascades" start to dominate: incident particles start to interact with substructures of the nucleon itself [65, 70]. It is important to mention that the transition from one dynamics to the other is not sharp but rather a continuous passage where the different situations are highly mixed. The main consequence of INC is a forward-directed component of high energy hadrons. The cascade extinguishes when the energy of secondary hadrons or nucleons falls below threshold for initiating new INC processes. When INC ends the nucleus is in a highly excited state.

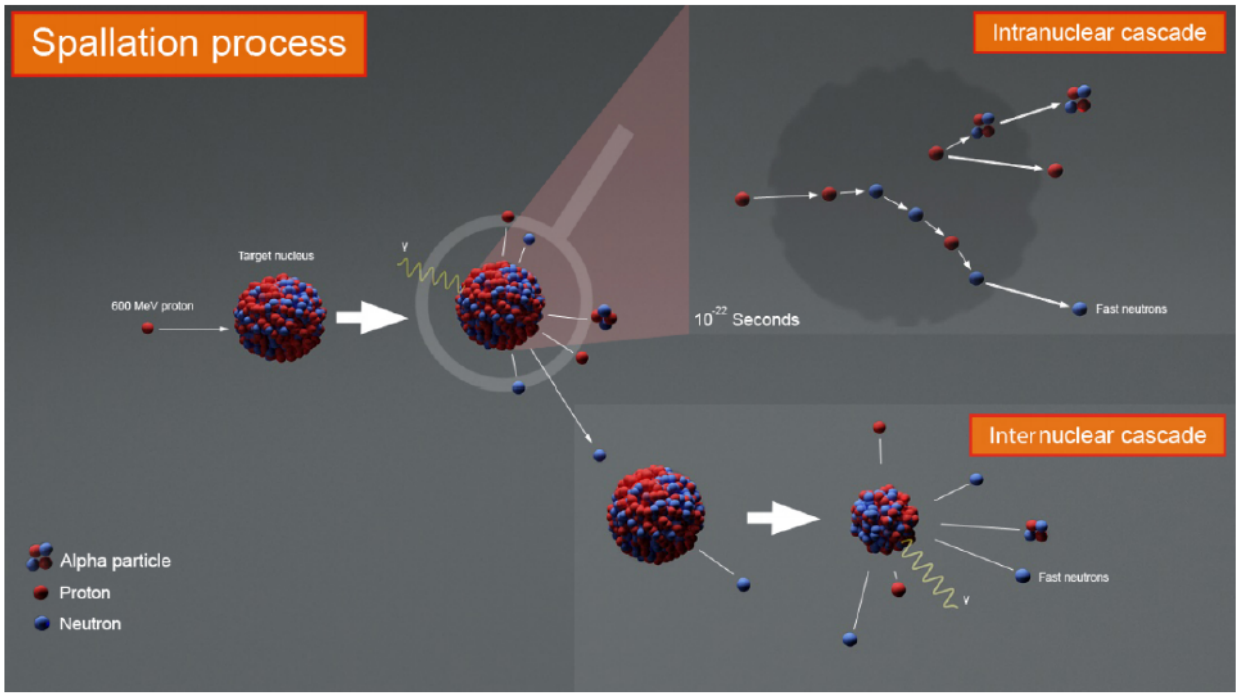


FIGURE 1.14: Schematic representation of nuclear cascades in a spallation reaction, specifically intra nuclear cascades and inter nuclear cascades triggered by a 600 MeV proton are illustrated. Nucleons or nucleon clusters, such as α, are ejected with prompt gammas. Figure from [64].

Pre equilibrium stage

An intermediate state between INC and nuclear de-excitation is commonly referred as "pre equilibrium stage". During this process the energy, first transferred to few nucleons, is evenly distributed throughout the nucleus via a series of internal collisions. This stage, which lasts about 10^{-18} s, can involve the emission of further fast particles or fragments. The ejected particles are normally more energetic than the ones emitted during the de-excitation state and highly contribute to decreasing the energy of nucleon constituents. Pre equilibrium state is, more than a distinct dynamics occurring during spallation processes, a transition phase which has been introduced in order to set the initial conditions for the statistical de-excitation that follows. It greatly helped to improve the agreement between experimental data and simulations [64].

Statistical de-excitation

As a consequence of intra nuclear cascade and pre equilibrium processes the target nucleus is left in a thermally excited state with non zero angular momentum. It is during the following state, generically referred as statistical de-excitation, that the energetic equilibrium is accomplished throughout different and competing processes: evaporation, fission and multi-fragmentation.

Evaporation consists in a relaxation process of the nucleus performed through the emission of low energy (less than 20 MeV) neutrons, protons alpha particles or other light fragments [67]. The vast majority of particles ejected through evaporation are low energy neutrons [65]. This process occurs if the excitation energy of the single nucleons is greater than the binding energy of the compound nucleus. A very detailed description of evaporation stage can be found in reference [71]: with a statistical approach the authors

illustrate emission processes from excited heavy nuclei with equations analogous to the ones for evaporation from a hot system [64].

A second common de-excitation method is fission: fragmentation into two heavy nuclei. Since the available energy accumulated during INC, after the collision with the projectile particle, can be quite large, fission can occur even for nuclei that are naturally not easily fissile. In contrast to evaporation, fission phenomenon is remarkable only for heavy nuclei (more than 75 protons) and when accumulated energies are large. For energies higher than 50 MeV, this process is referred to as "hot fission". In this case, fragments generated throughout fission processes can in turn evaporate emitting lighter fragments [64].

Similarly to fission, the excited nucleus can disintegrate in several fragments (more than two) emitted simultaneously. This multi-body process, called "multi-fragmentation", occurs when energies are comparable to the total binding energy and mechanical instabilities spread to the nucleus breaking it apart. Multi-fragmentation is an highly asymmetric fission process where the multiplicity of fragments is proportional to the energy accumulated in spallation collisions while their size decreases with growing number of fission products. Fragments have atomic number ranging from 2 to 20 and are emitted in a very short time (order of 10^{-9} s). The process in described in Reference [72] as a very unstable system generated from the increase of thermal pressure in the hot nucleus. Subsequently, density fluctuations inside the nucleus lead to a multi-body explosion due to Coulomb repulsion.

The final step of nuclear de-excitation, after emission of nucleons or nucleon clusters, takes place when the residual energy is lower than the binding energy of nucleus components and fragments (typically below about 8 MeV) [65]. In this situation, hot nuclei reach a stable configuration through the emission of gammas. A cascade of consecutive photons characterizes this de-excitation stage and can be considered as the final part of evaporation which terminates the dissipation of residual energy. More precisely, gamma emission manifests even all along the above mentioned de-excitation processes, however it presents a rather small branching ratio. Only when the energy of the compound nucleus is smaller or comparable to the threshold for neutron emission, photon cascades start to be abundantly generated until the ground state is reached [73]. Therefore, gamma emission is less important for the decay of highly excited nuclei which instead tend to undergo fission or multi-fragmentation whose energy dissipation is larger [74].

Chapter 2

The Super-Kamiokande detector

Super-Kamiokande is the largest pure-water Cherenkov detector ever constructed. Its cylindrical tank, with a height of 41 m and radius of 39 m, filled with ultra-pure water, is located in an excavated chamber within the Kamioka mine in Mt. Ikenoyama, Japan [75] (an illustration in Figure 2.1). This chapter describes the structure of the detector

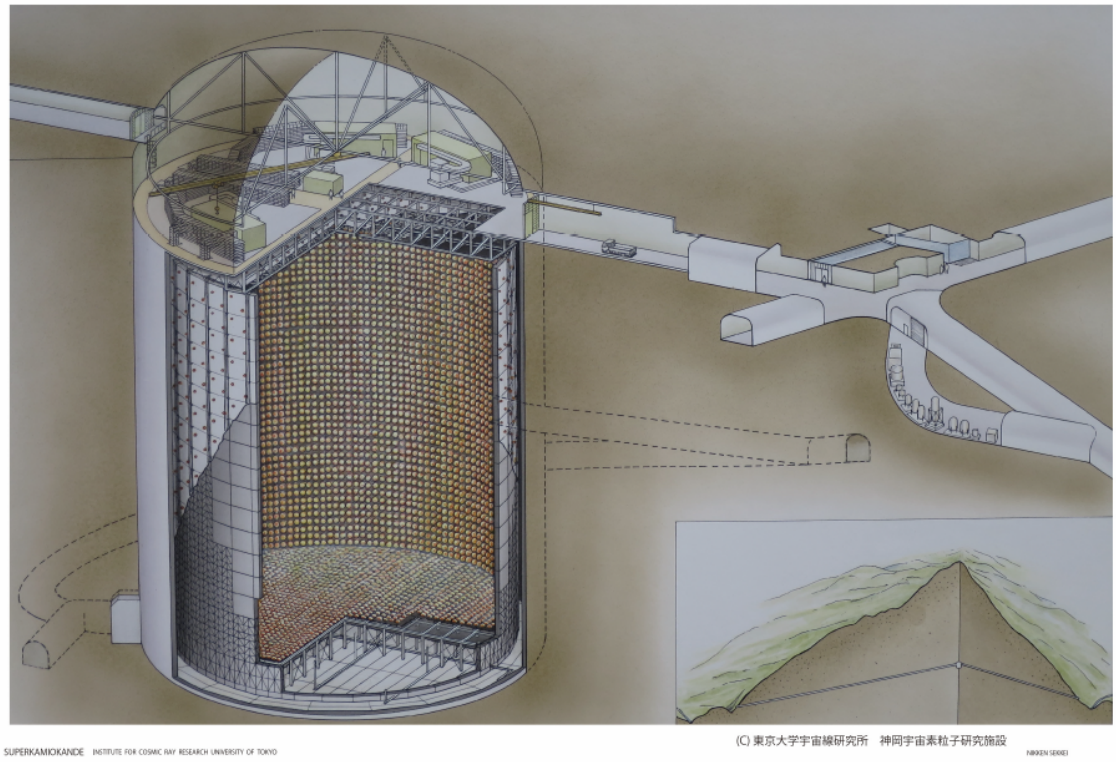


FIGURE 2.1: Schematic illustration of the Super-Kamiokande detector.

and of the trigger system, as well as their modeling in simulations. An overview on the structure, the main hardware components employed to detect signal and key systems to operate and maintain the experiment are presented first, followed by a discussion on detector simulation and event reconstruction methods.

2.1 Overview of the Super-Kamiokande experiment

The original purpose of the Super-Kamiokande experiment (SK) was to search for proton decay and study ν from atmosphere, Sun, and astrophysical sources, including supernovae. In addition SK plays the role of far detector for the Tokai-to-Kamioka (T2K)

long-baseline neutrino experiment [76]. Built between 1991 and 1996 as a successor of the historical Kamiokande experiment, Super-Kamiokande is shielded by 1000 m of rock overburden, (2700 m.w.e.) in order to reduce cosmic ray contaminations. The residual flux is only composed of cosmic ray muons which reach SK location with a rate of ~ 2 Hz. Figure 2.2 shows cosmic muon flux for SK and, for comparison, other underground experiments. The experimental hall is surrounded by a 40-50 cm concrete wall to further shield from natural radioactivity. As it can be seen from Figure 2.1, a large dome is built on top of the actual experiment structure for electronic storage and in-situ calibration holes for the injection of measurement devices or sources. The detector is optically separated

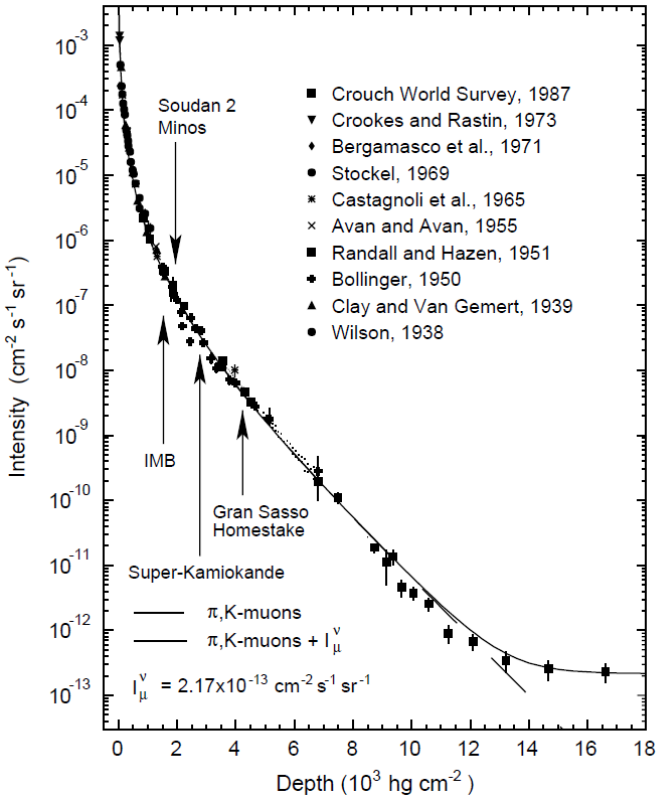


FIGURE 2.2: Cosmic-ray muon flux as a function of depth underground. The unit 10^3 hg cm^{-2} is equivalent to km.w.e. for standard rock. Dashed curve represents the π and K-muons. The solid curve adds the neutrino-induced muon background. Depths of present and past underground experiments are also shown. From [75].

into two concentric cylinders, inner detector (ID) and outer detector (OD) by a layer of black polyethylene terephthalate sheet placed at 2.7 m from tank wall. SK detector globally measures 41.4 m in height and 39.3 m in diameter and contains 50 kilotons of ultra pure water. A fiducial volume (FV) is defined as the region within 2 m from the ID walls. The inner detector alone accommodates 32 kiloton of the total water weight. 20-inch and 8-inch photomultiplier-tubes (PMTs), facing inwards and outwards respectively, are implemented to detect Cherenkov light emitted from charged particles in each region. The total number of SK PMTs inside the inner detector is 11 129 and they represent the main detection device, while the 1 885 OD PMTs work as a veto for cosmic muons. PMTs with their supporting structures are shown in Figure 2.3.

SK started its operations in April 1996, from there on, six running phases can be distinguished, from SK-I to SK-VI (also called SK-Gd). From the first run in 1996 to July

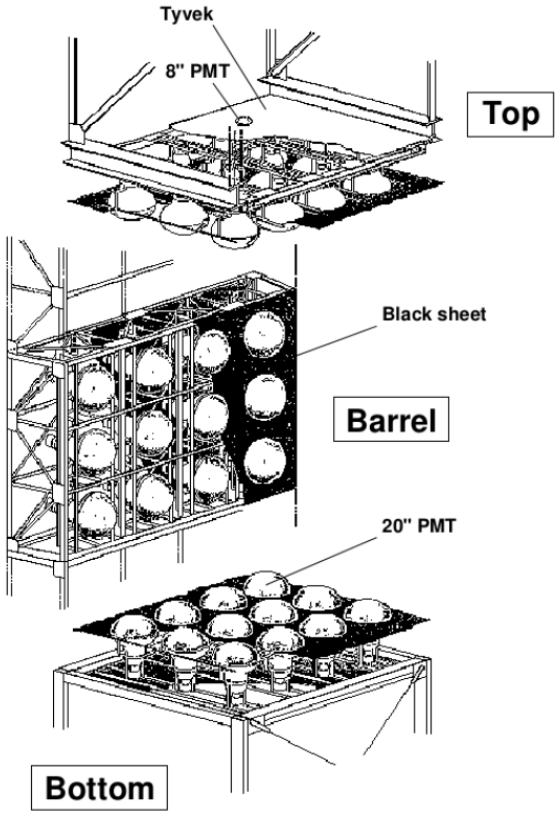


FIGURE 2.3: Schematic illustration of the PMT support structure in SK. From [75].

2001 Super-Kamiokande had $\sim$ 40% photo coverage. A shutdown for maintenance and PMT replacement followed this period, called "SK-I". During the works a catastrophic accident, caused by the implosion of an ID PMT and a subsequent chain reaction which triggered a shock wave impacting the whole detector, resulted in the destruction of more than half of the ID PMTs. SK-II phase, resumed in 2002, operated with only $\sim$5000 ID PMTs and a photo coverage of $\sim$ 19%. An acrylic cover and a fiberglass reinforced plastic (FRP) was added to newly commissioned PMTs to increase isolation and protection against further accidents. After a new tank opening for PMT installation in October 2005, SK operation continued until September 2008 with photo coverage back at 40% (SK-III phase). The fourth and longest period of Super-Kamiokande, SK-IV, ran from August 2008 to June 2018. It started after a major upgrade of electronics and acquisition system which allowed to extend data taking longer after primary events and activate neutron tagging techniques. In 2018, SK tank underwent a refurbishment work to fix a water leak issue in view of the next SK phase, known as SK-Gd. During maintenance, ID and OD PMTs were replaced, a major cleaning of the detector walls was carried out and water recirculation system was upgraded too. Details on the upgrade of SK-V are described in Chapter 3. SK-V started running in January 2019 and represented the final phase of pure water as detector target. Since August 2020, SK-Gd has begun: gadolinium has been dissolved in pure water with a concentration of 0.01% in mass. Gd has one of the largest neutron capture cross-section among natural elements. Moreover, neutron capture on Gd leads to the emission of a gamma-ray cascade with a 8 MeV endpoint, making neutron captures far more visible than in pure water. Hence, Gd doping considerably enhances SK's ability to detect antineutrino signals, and notably SRNs. A summary of the different

TABLE 2.1: Properties of SK operation periods [47].

Phase	SK-I	SK-II	SK-III	SK-IV	SK-V	SK-Gd
Start time	Apr. 1996	Oct. 2002	Jul. 2006	Sep. 2008	Jan. 2019	Aug. 2020
End time	Jul. 2001	Oct. 2005	Sep. 2008	Jun. 2018	May 2020	Running
ID PMTs	11 146	5182	11 129	11 129	11 129	11 129
OD PMTs	1885	1885	1885	1885	1885	1885
Coverage	40%	19%	40%	40%	40%	40%
Electronics	ATM	ATM	ATM	QBEE	QBEE	QBEE
Trigger	Hardware	Hardware	Hardware	Software	Software	Software
Neutron tag	No	No	No	Yes	Yes	Yes
Threshold	4.5 MeV	6.5 MeV	4.0 MeV	3.5 MeV	3.5 MeV	3.5 MeV

SK phases can be found in Table 2.1.

2.2 Detection methods

In Super-Kamiokande neutrinos indirectly generates a signal when they interact with water, producing charged particles above Cherenkov threshold. We discussed in details the main characteristics of Cherenkov radiation in Section 1.6.3, among them the typical light cone emitted along the particle propagation direction. A ring hit pattern is observed on SK wall as a consequence of the projection of the Cherenkov cone on the PMT array on the detector walls, see Figure 2.4. Cherenkov photons hit the surface of photomultiplier tubes, located all around the detector, and are converted into an output charge signal. Photon arrival time and integrated charge at each PMT are crucial information to reconstruct main variables of the generating events, such as vertex position and deposited energy. The type of particle could also be inferred from ring shape analysis: the electron-induced ring is more likely to be fuzzier than the muon-induced ring because of the lighter mass and the higher probability of electromagnetic scattering. The energy of the charged particles and their tracks within the detector play an other crucial role for event identification: as an example, cosmic ray muons traveling downwards and generating a large amount of Cherenkov light are thus distinguishable from neutrino-induced muons that show short and less energetic paths. An example of Cherenkov ring seen from SK event display is shown in Figure 2.5. For reference, SK uses a coordinate system whose origin of axes is set at the center of the ID, z axis runs along the cylinders center with upward direction and x axis points to the west.

2.2.1 ID PMTs

Photomultiplier tubes are evacuated glass chamber which collect light and convert it into an output charge signal. In particular, ID PMTs, whose structure is illustrated in Figure 2.6, are 20 inch R3600 PMTs (about 51 cm) developed by Hamamatsu Photonics in order to meet Super-Kamiokande requirements [75, 77]. A photon hitting the PMT's outer face has a given probability to penetrate and excite the photocatode layer made of bailkali (Sb-K-Cs) and located on the glass surface. The sensitive wavelength range is between 300 and 600 nm, which make these PMTs ideal for the detection of Cherenkov radiation that peaks at wavelength of 400 nm. Photoelectrons (p.e.), generated through photoelectric effect in the cathode, are emitted into the vacuum chamber, focused and accelerated by the electric field present in the PMT. The acceleration continues through multiple dynodes where a multiplication via secondary photoelectron emission triggers an electric cascade. There

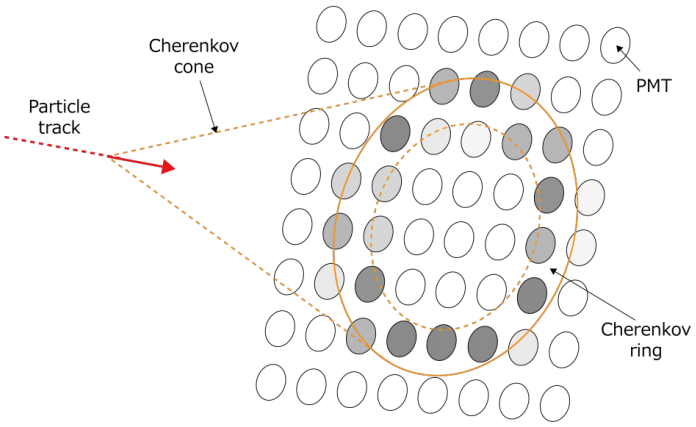


FIGURE 2.4: Illustration of Cherenkov ring hit pattern in SK.

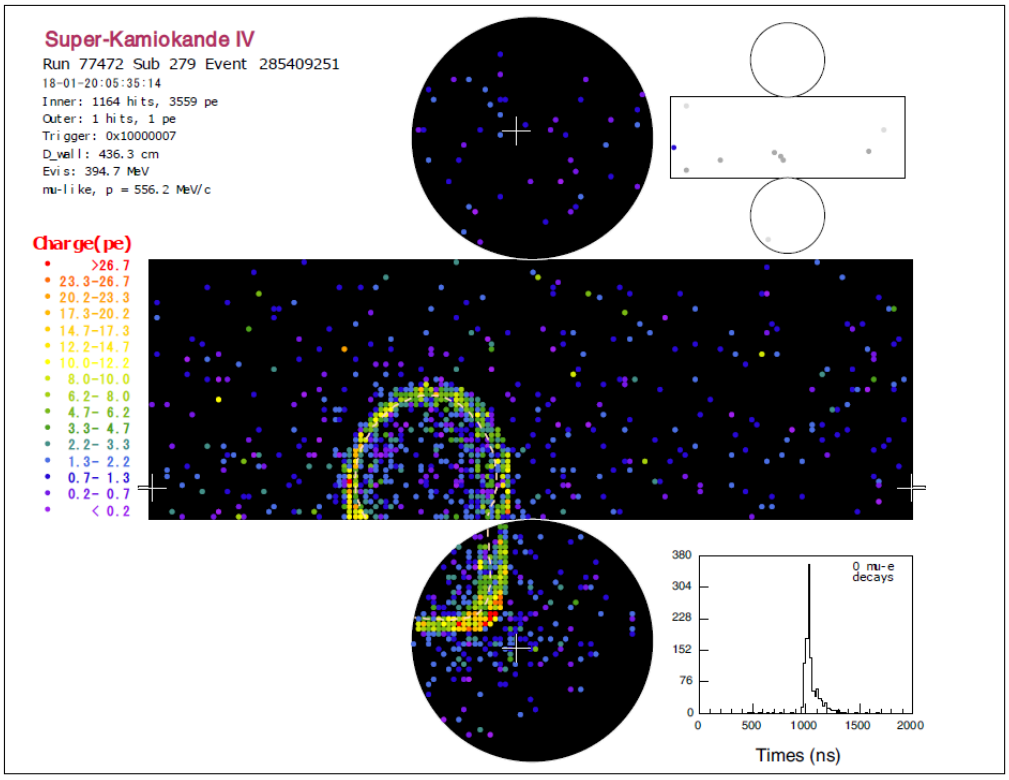


FIGURE 2.5: Example of the Cherenkov ring pattern in SK. Each pixel corresponds to one PMT and the color scale denotes the detected charge deposited.

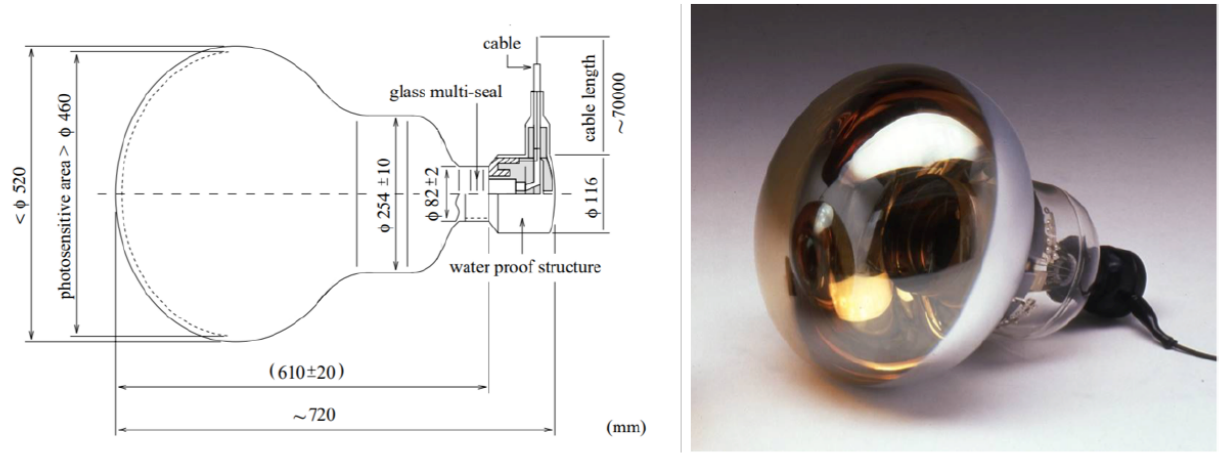


FIGURE 2.6: Schematic view of the cross section of SK 20 inch PMT (Hamamatsu R3600). From [75]. A photograph on the right.

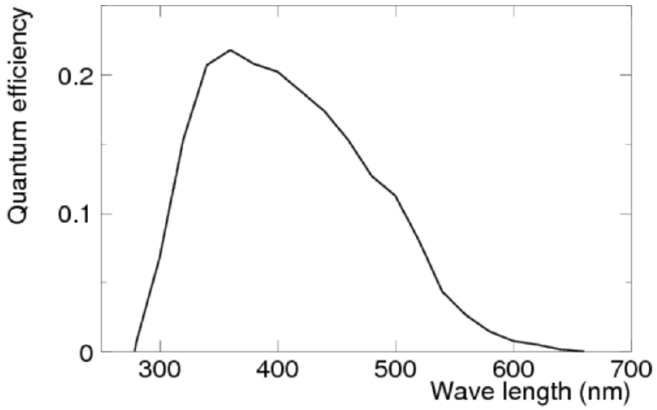


FIGURE 2.7: The quantum efficiency of SK ID-PMT as a function of wavelength [77].

are up to 19 dynodes where electrons are multiplied more than hundred times, leading to a collection of a measurable charge current at the final anode. Quantum efficiency (QE) expresses the probability that a photon impinging on the PMT is transmitted to the cathode, absorbed and converted in an photoelectron. The triggering probability is the probability that the generated photoelectron successfully initiates a self-sustaining avalanche process and thus an output current pulse and it depends on the average collection efficiency at the first dynode, which is about 70% and relatively uniform across the whole PMT surface. QE is a function of the wavelength and angle of incoming photons, as shown in Figure 2.7. It has a maximum value of ∼20% for SK ID-PMTs leading to a total gain factor of $\sim 1 \times 10^7$ with a voltage of $\sim$ 2000 V. The transit time spread is of about 2.2 ns. In SK, the gain of ID PMTs, i.e. the pcatodehoton multiplication factor, was set to the order of 10^7 so that they would be sensitive to a single photoelectron signal, produced by low energy event such as supernova neutrinos. In order to avoid unexpected photoelectron drift from the cathode to the first dynode due to geomagnetic field, a system of Helmholtz coils surrounding the detector wall generates a counterbalancing magnetic field. Therefore, photoelectrons see a total magnetic field reduced from 450 mG to 50 mG. PMT dark noise, namely photoelectron triggered when no real photon is hitting PMT glass, was ∼

3.5 kHz at a threshold of 0.25 p.e. for SK-I and increased to ∼5.2 kHz from SK-II as a consequence of the intrinsic radioactivity of the acrylic cover applied to protect PMTs.

2.2.2 OD PMTs

8-inch PMTs are used for the outer detector. Among them, 591 PMTs of the model R1408, comes from IMB experiment [36] while the remaining 1293 of R5912 type were commissioned and installed after SK-II phase. Although it introduces a time resolution loss, a wavelength shifter is needed to better match the incoming light emission spectrum and the best value of QE for PMTs: photons are absorbed and then re-emitted by an acrylic plate installed on the PMT surface. The final maximum QE lies around $\sim$ 50%, relatively higher than ID-PMTs in order to compensate the lower coverage. Given that OD primarily functions as a veto, keeping a higher QE to observe extra photons is more important than minimizing the time resolution (15 ns). In fact, the primary role of SK outer detector is to spot incoming cosmic ray muons or neutrino interactions where the particle produced, typically muons, enter and/or exit the ID. Moreover, OD could occasionally serve as a substitute of the inner detector in case of malfunctioning or saturation, for example when extremely high energy events from cosmic rays blind the ID-PMTs for a short amount of time. An other important function of the OD is the separation of cosmic ray muons stopping inside the ID from through-going muons.

2.3 Water and air system

Water circulation system

Once emitted, Cherenkov photons need to propagate through the detector to reach the PMT array and be recorded. In order to avoid undesired absorption and loss of light or non-uniformities in signal detection, water purity must be kept at an high and constant level. Especially for low energy interactions, where only few Cherenkov photons are emitted, scattering and absorption would lead to insufficient statistics and bad time response. A sophisticated water purification system and constant monitoring of water properties is crucial for the good functioning of Super-Kamiokande detector. It should be also able to filter radioactive isotopes and impurities from detector materials or surrounding rocks which could form a significant background in the low energy analysis. SK water is collected from two streams within Kamioka mine that are naturally filled by rain and snow melt in the surrounding mountains. Fresh water needs to be subsequently purified and filtered through a complex system that continuously circulates water with a flow of $\sim$ 60 tons/hour, as shown in Figure 2.8 (left). With the system of Figure 2.8 (left) water is supplied in the bottom of the tank and drained on the top, the specific locations of the pumps are accurately calculated to minimize convection and therefore reduce non-uniformities in water purity. Below - 11 m, if the origin of the coordinates is the ID center, convection still exists, making the bottom of the detector the region with higher level of contamination. Further efforts to reduce convection consist in the installation of heat exchangers to maintain constant water temperature in the supply chains. Average temperature is about 13°C, with a variation of 0.01°C, as illustrated in Figure 2.8 (right), with a resulting variation of 5% within the ID region [75]. The low temperature also helps suppressing potential bacterial growth. Temperature is constantly monitored by permanent thermometers installed inside the tank. Absorption and scattering coefficients are measured in a daily basis by calibration lasers systems. Any remaining bacteria are monitored and killed by a UV sterilizer. Finally, thanks to this complex water system, transparency of ultra-pure water is ensured for about 90 m [47, 78].

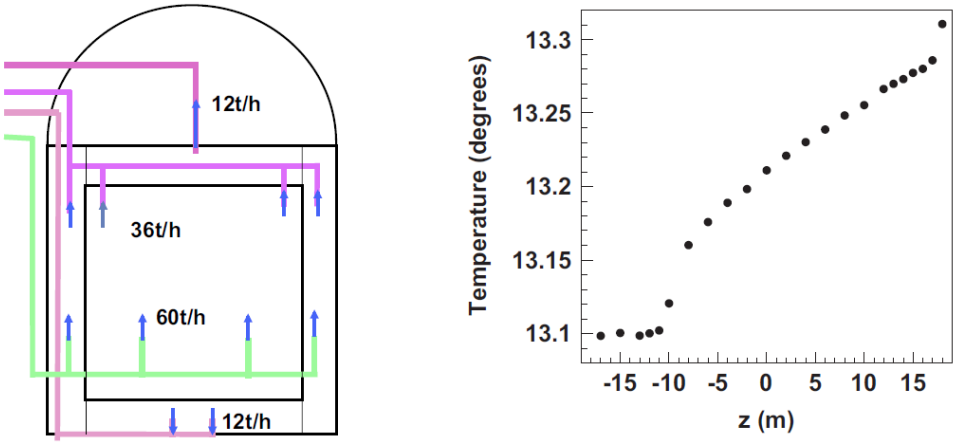


FIGURE 2.8: Left: schematic illustration of water circulation in SK. Right: water temperature as a function of the vertical position in the detector (z). From [75].

Air purification

Mt. Ikenoyama rock is quite rich in natural uranium whose decay chain includes radioactive isotopes of radon: it is not only a source of radioactive impurities in water, causing background signal for MeV energy analysis, but also an air contaminant that carries potential health risks if inhaled. Seasonal variations induce different level of contamination: ~ 30 Bq/m^3 in winter and ~ 1500 Bq/m^3 in summer, making essential a special air supply system with powerful filters and fresh air blower incorporated. This keeps Rn concentration below ~ 50 mBq/m^3 in air and below ~ 2 mBq/m^3 in water [39].

2.4 Electronics and Data acquisition

Four huts accommodating high voltage power supplies and electronic systems as well as a central hut hosting a sophisticated trigger system are located in the dome on top of SK tank, shown in Figure 2.1. We remind that the data acquisition system (DAQ) from SK-I to SK-III consisted of analog timing modules (ATM) which were replaced, in SK-IV phase, by new electronics, QBEE, namely charge to time converters (QTC) electronics with Ethernet. Major improvements were reduction of dead time, extension of dynamic range (namely the ratio between the maximum and minimum signal that is acquired) and low energy threshold. Description of the current DAQ system will be presented in the following.

2.4.1 Front-end electronics

The new front-end electronics consists of QBEE modules which read input signals from 24 PMTs each. These boards allow the readout of single PMT signal every 17 μs controlled by a 60 kHz internal clock. The two main components of the QBEE module are the QTC (charge-to-time converter) and TDC (time-to-digital converter), the first converts PMT charge into time and the second transforms pulse widths and hit times into digital signals. An extension of the dynamic charge range (0.2 to 2500 pC) is achieved via digitalization with three gain values: the PMT charge is simultaneously integrated in QTC

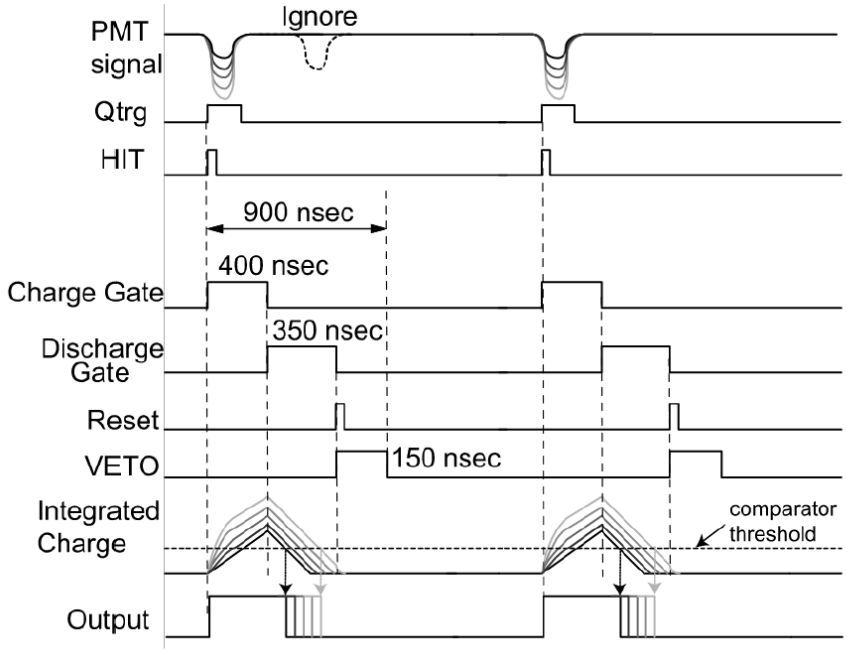


FIGURE 2.9: Timing chart for signal and trigger in a QTC. From [80].

modules with relative gains of 1, 1/7, 1/49 in order to improve charge resolution. Following the gain stage, charge signal, if it passes discriminator threshold, is integrated with a capacitor: a start timing signal is generated (HIT trigger), followed by charging, discharging and veto gates (see Figure 2.9) for a total processing time of 900 ns, during which all other input signal are ignored. QBEE great stability and high-speed digitalization allows the implementation of a software-based trigger as a replacement of the hardware trigger used in SK until 2008 [79]. The main difference is that all data recorded by the PMTs are kept as input to the front-end electronics, which converts and passes them to a merger PC, located in the central hut, for selection, filter and trigger treatments. The transmission to online computers is carried out via Ethernet cables.

2.4.2 Software triggers

After digitalization, signal from PMTs are sent to 20 front-end PCs and subsequently to 10 merger PCs for event building, at this stage software triggers are applied to select events for different types of physics analysis and save them on disks for offline treatments. Data Merger PC combine 1344 consecutive data reads (17 μs each for a total $\sim$ 23 ms window of data) in one data-block which is the one processed by online triggers [81]. The new software trigger scans the data in blocks and searches for events that satisfy some specific trigger conditions. In effect, different types of trigger are applied: according to the event properties, such as the number of PMT hits within 200 ns sliding window, different trigger widths and thresholds are imposed to efficiently use disk space. A summary of SK trigger system is illustrated in Table 2.2, it consists of Super Low Energy (SLE), Low Energy (LE), High Energy (HE), Super High Energy (SHE), Outer Detector (OD) an T2K triggers. SLE requires 31 PMT hits (trigger rate of 3.0-3.4 kHz [82]), which correspond to the minimum SK threshold of about 3 MeV for electrons. SLE events are extremely frequent and therefore their acquisition time width is reduced to 1.5 μs. On the contrary LE, HE and SHE gate expands to 40 μs (5 μs before and 35 μs after trigger timing, respectively to catch electrons from muon decays). Moreover, since SHE trigger was specially introduced in SK-IV

TABLE 2.2: A summary of QBEE triggers [39].

SK-IV Triggers	Threshold [Hits/200 μs]	Time window [μs]
SLE	31	[-0.5, +1]
LE	47	[-5, +35]
HE	50	[-5, +35]
SHE	58	[-5, +35]
OD	22 (only in OD)	[-5, +35]
AFT	SHE + no OD	[+35, +535]
T2K	Beam on	[-500, +535]

to tag neutron captures and SHE events are rare enough, an After Trigger (AFT) is introduced when OD trigger is not issued (namely, not for incoming cosmic muons), enabling the extension of data storing up to 535 μs from SHE trigger start. SHE+AFT triggers are of a special importance for SRN searches to tag prompt and delayed events, as described in Sec. 1.2.1. T2K trigger stores data for about 1 ms when the beam is activated. Trigger thresholds depend on SK periods, values for SK-IV are illustrated in Table 2.2.

Wide-band Intelligent Trigger (WIT)

Software triggers described above are primarily conceived to emulate hardware triggers of SK-I, SK-II, SK-III phases. Therefore DAQ thresholds are not drastically changed compared to previous SK periods: SLE trigger threshold is 3.49 MeV of electron kinetic energy while it was about 4.6 MeV for previous operation phases [47]. Moreover SLE triggering efficiency below 3.99 MeV lowers from more than 99% to about 84% [81]. Further reduction of hit threshold would lead to even lower efficiency. The trigger threshold results from the fact that standard software triggers simply count the number of PMT hits within the 200 ns sliding window: SLE has a threshold of 31 hits/200 ns and an average dark noise rate/200 ns of 11 hits. The structure of these online triggers exclude every event below threshold, among which 2.2 MeV gammas from neutron capture on hydrogen. The only possible configuration that enables neutron tagging is when an AFT trigger is issued, however a prompt event above SHE threshold is required. Neutron captures not accompanied by more energetic events are not seen. These events could play a crucial role in shower identification: neutrons are very abundant within hadronic showers induced by cosmic muons in water and could be used as a warning for spallation background contamination [53]. In principle, SHE+AFT trigger could be used: cosmic muons deposit tens of MeV in the detector and therefore easily pass SHE requirements (about 8 MeV threshold), the AFT window could be exploited to store the large part of eventual γ from neutron captures. However, not only the identification and reconstruction efficiency for 2.2 MeV events would be really low as a consequence of the small amount of Cherenkov light produced (about 7 p.e. detected), but also it would require an excessive strain of the software triggers and undue occupation of disk space as a consequence of the high rate of cosmic muons in SK. Hence AFT trigger is deactivated after cosmic muons.

The solution comes from the implementation of the so-called Wideband Intelligent Trigger (WIT), a new software algorithm which runs in real time, in parallel to other online triggers. WIT, installed in 2015, performs online conversion of digitized time and charge followed directly by vertex reconstruction [81, 82]. It receives a copy of all PMT signals, evaluates independently each data block and applies two-stage vertex reconstruction as a selection criteria [81]. WIT runs on completely dedicated computers with sufficient

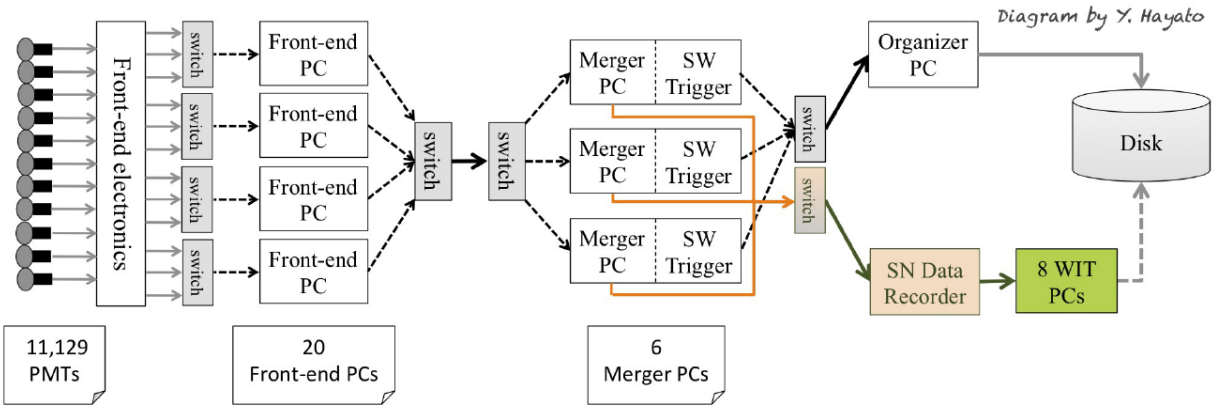


FIGURE 2.10: Super-Kamiokande DAQ. PMT hits in windows of 17 μs are read by the front-end electronics with an internal clock. Merger PCs assemble 1344 consecutive data reads which are sent to software trigger system. A buffer machine (SN Data Recorder) transmits raw data to WIT machines in real time. Events passing software triggers, after going the though Organizer PC, or events reconstructed by WIT computers are then sorted and stored into disk. From there data will be further analyzed offline. From [81].

power to process all the raw un-triggered data. It is able to extract, in real time, low energy events down to 2.5 MeV (recoil electron energy), and it has demonstrated a non negligible efficiency to trigger 2.2 MeV gamma rays. We remind that, at such low energies, a photon typically confers its energy to a single electron by Compton scattering, which only produces a faint Cherenkov signal in the detector. On average, a 2.2 MeV event turns on just 7 PMTs, therefore WIT triggers has to rely on up fluctuation of this value. All 1344 data blocks (17 μs each, called “hardware trigger" for a total of 23 ms) are transmitted via 10-Gigabit Ethernet cables to 10 WIT computers. Each data block has an overlap of 1.1 ms (64 hardware triggers) with neighbor blocks in order to avoid edge effects. Nine WIT computers receive, trigger and reconstruct raw data blocks while the tenth computer takes care of sorting processed events by time and assembling them into “subruns" of 4000 blocks ($\sim$ 90s of data). Reconstructed, time sorted and assembled data are then transmitted to offline disk for further reduction and analysis. A schematic illustration of the SK data acquisition system that includes the WIT trigger is shown in Fig. 2.10.

The WIT trigger criterion starts with a pre-trigger requirement: a coincidence of 11 PMT hits above dark noise (usually around 12 hits) within a 230 ns window. 230 ns is roughly the time a relativistic particle would need to cross the inner detector. Since SK observes about 6 PMT hits per MeV, pre trigger requires roughly 2 MeV (electron kinetic energy) above noise to isolate signals. Following the pre-trigger, a second filter, called STORE (Software Triggered Online Reconstruction of Events), searches for a single vertex for the hits within the 230 ns. Since electrons below 20 MeV in water only travel few centimeters from their production location, their Cherenkov light can be approximated to a point source and the reconstruction of a single vertex is only based on relative time of the PMT hits. A necessary, but not sufficient, condition is that hit times from any selected pair of PMTs must obey $\delta x_{ij}/c > \delta t_{ij}$ with δt_{ij} (δx_{ij}) the time (spatial) difference between hit in PMT i and PMT j, namely the speed-of-light separation of the two photo-multiplier tubes considered must be greater or equal to the hits time coincidence. In order to find vertex candidates, all PMT hits within the 230 ns window satisfying this condition are added to a list, subsequently four-hit combinations for which a possible vertex position would justify

hit times, given PMT positions, are extracted. The requirement is that time residuals of all four hits of the combination must be zero for a given vertex (with time uncertainty $\sigma = 5$ ns). STORE will eventually calculate several candidate vertices: those located outside the fiducial volume or showing an insufficient level of coincidence with all other hits are discarded. Specifically, the level of coincidence of the vertices pre-reconstructed with STORE has to be larger than 6.6, which represents 60% of the eleven hits above the dark noise. This procedure works as a preliminary filter that effectively reject background in a much faster way than a proper vertex reconstruction algorithm. STORE greatly speeds up the the vertex fitting algorithm and the background rejection.

All hits selected by STORE are then passed to two vertex fits: Clusfit [83], a fast algorithm for an initial selection, followed by BONSAI, a slower but more precise one (more details will be presented later in this Chapter, in Section 2.6) [84]. Clusfit filters out low energy events and isolated hits arising from dark noise effects or scattered and reflected light. If Clusfit vertex is further than 1.5 m from PMT wall, BONSAI is applied on all PMT hits. BONSAI is the standard vertex reconstruction algorithm for low energy analysis in Super-Kamiokande, it searches for a vertex within a 1.5 μs window around the trigger time (500 ns before and 1000 ns after). If the candidate vertex is inside the fiducial volume and if the number of hits, whose time residuals are included in a window range of -6 ns and +12 ns, is larger than ten, the event is formed and stored. WIT trigger efficiency for 2.2 gamma ray detection is around 13% over the entire detector and 17% if only FV is considered.

Since triggering and reconstruction performed by WIT are both in real time, this multi-step reconstruction procedure largely reduce the processing time and data storing size. Moreover it allows to detect super low energy electrons as far as the limit of stable and reliable SK event reconstruction (2.49 kinetic MeV). Lowering the energy threshold with WIT allows to investigate muon-induced showers through the detection of neutrons.

2.5 Super-Kamiokande detector simulation

Most of the analysis as well as calibration and systematics estimation requires a detailed Monte Carlo simulation of the detector and its response. The official detector simulation for Super-Kamiokande, referred as SKDetSim, is based on GEANT 3.21 [85]. Details can be found in Ref. [47], while Tab. 2.3 summarizes the main physics processes reproducible with SKDetSim for different particles. Many of these physics reactions are simulated with GEANT 3 software or specialized integrated generators (.i.e. NEUT for neutrino interactions [86] and GCALOR for hadronic interactions [87]), while some others, specifically related to the SK detector characteristics, are specially tuned with calibration data. These processes include scattering in water, reflections on PMT glasses or other surfaces, response of PMTs and electronics. More generally, SKDetSim can model in detail: the particle propagation in water, the emission of Cherenkov photons, reflection and absorption of light on materials, photo-electron production and electronic response.

Charged particles above Cherenkov threshold emit photons according to the equation below:

$$\frac{d^2N}{dxd\lambda} = \frac{2\pi\alpha}{n\lambda^2}\left(1 - \frac{1}{n^2\beta^2}\right) \tag{2.1}$$

with N the number of generated photons with wavelength λ, x the traveling length, n the refractive index in water, α the fine structure constant, and β the velocity of charged particles divided by the speed of light in vacuum. This equation is implemented in SKDetSim, where n is a measured value at specific water temperature and pressure of SK. Only photons within the PMT sensitive range are produced (300 - 700 nm). Once generated, each

photon propagates through SK with a certain probability of being scattered or absorbed by water itself or impurities which depends on the level of water transparency. This is represented by the attenuation length factor which includes absorption coefficient, Rayleigh scattering and Mie scattering and is constantly monitored by calibration systems [78]. Finally, given that a photon impinging on PMT glass surface can be reflected, transmitted or absorbed by the cathode, the probability of the three processes are included in SKDetSim based of experimental measurements. Different libraries including detector calibration coefficients for various SK periods can be loaded and built into SKDetSim, each of them is based on data from various calibration sources.

Simulated events are subjected to the same treatment as real data and the same reduction algorithms are applied in order to enable MC-data comparison: charge and time response of PMTs and electronic system are simulated resulting in the same structure as observed data. For the same purpose, particles simulated in SKDetSim are recorded in trigger windows analogous to the ones applied to data, each trigger window has a fixed and well-defined length and represents what is called a MC event (or SKDetSim event). In general, for each event, the detector dark rate is also simulated and can be adapted to the value measured during different SK run periods, this can be added or not to the outputs depending on the specific needs.

2.6 Event reconstruction

Dedicated softwares are used to extract energy, time and position of a physical event from informations recorded by SK PMTs. These algorithms depend on the total charge deposited, which divides events in two categories: low energy and high energy signals. Here only event reconstruction dedicated to low energy signals and muons is presented.

2.6.1 Low energy event reconstruction

In Super-Kamiokande a signal is defined to be a low energy event if less than 1000 PMT hits are recorded (typically below 80 MeV). The reconstruction in the low energy region is mostly based on timing informations and hit pattern, unlike for higher energy ranges where charge informations are important as well. The reason comes from the fact that most low energy events generate, on each hit PMT, only a single photoelectron signal. Therefore the incident particle energy can be approximated to be proportional to the number of PMT hits. Vertex, direction, energy and Cherenkov angle reconstruction will be presented in the following.

Vertex and direction reconstruction

The vertex of a low energy event is reconstructed with an SK customized fitter, called BONSAI (Branch Optimization Navigating Successive Annealing Iterations) [84]. BONSAI selects four-hit combinations and generates test vertex for each quadruplet. The best fit vertex of all quadruplets is then extracted by maximum likelihood of time of flight (ToF) subtracted timing residuals:

$$\Delta t = t_i - ToF(x) - t_0 \tag{2.2}$$

with t_i the actual hit time on the i-th PMT, x and t_0 the position and emission time of the candidate vertex while $ToF(x)$ is the time of flight for a photon traveling from that vertex, x, to the PMT in consideration. As already mentioned, electron and positrons typically only travel for few cm at low energies (i.e. 10 cm at 20 MeV) therefore, given that vertex

TABLE 2.3: A summary of physics processes simulated in SKDetSim. From [47].

Particle type	Physics process
γ	Rayleigh scattering
	Mie scattering
	Absorption on black sheet
	Absorbtion on acrylic case
	Absorption on PM surface
	Reflection on black sheet
	Reflection on acrylic case
	Reflection of PMT surface
	(e^+, e^-) pair production
	Compton scattering
	Photoelectric effect
$e^\pm$	Multiple scattering
	Ionization and δ-rays production
	Bremsstrahlung
	Annihilation of positron
	Generation of Cherenkov radiation
$\mu^\pm$	Decay in flight
	Multiple scattering
	Ionization and δ-rays production
	Bremsstrahlung
	Direct (e^+, e^-) pair production
	Nuclear interactions
	Generation of Cherenkov radiation
Hadrons	Decay in flight
	Multiple scattering
	Ionization and δ-rays production
	Haronic interactions
	Generation of Cherenkov radiation

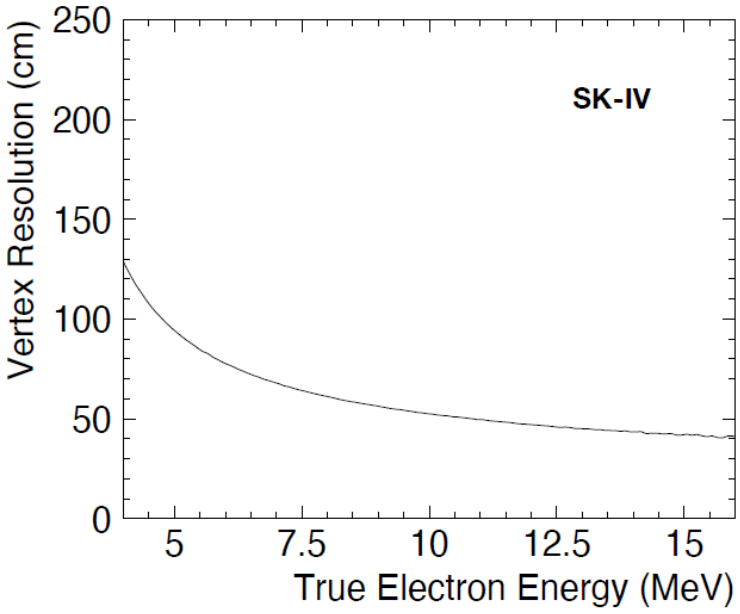


FIGURE 2.11: Vertex resolution as a function of electron true energy for SK-IV phase. From [39].

resolution is about 50 cm, Cherenkov light from these events is considered to be point-like. The probability density function (PDF) for Δt is obtained from calibration measurements while dark noise is assumed to be flat. BONSAI likelihood is:

$$L(x, t_0) = \sum_{i=1}^{N_{hit}} logPDF(\Delta t) \tag{2.3}$$

Vertex resolution, defined as 1σ difference between true and reconstructed vertices, increases with energy as shown in Figure 2.11. BONSAI, in addition of extracting the best vertex position and time, gives an estimation of the quality of the fit (referred as "vertex goodness", g_{vtx}) based on the time residual coincidence::

$$g_{vtx} = \sum_{i}^{N_{hit}} \frac{\sum_i W_i G(\Delta t_i|\sigma)}{\sum_i W_i} \tag{2.4}$$

The formula is constructed as a summation over all PMT hits of weighted sum of Gaussian functions for time residuals of the form $G(\Delta t_i|\sigma) = \exp\left[-0.5\,(\Delta t_i/\sigma)^2\right]$ with $\sigma = 5$ ns, namely the effective timing resolution expected for Cherenkov events. Gaussian weights are defined as $W_i = -\frac{1}{2}\left(\frac{\Delta t_i}{\omega}\right)^2$ with $\omega = 60$ ns, value also based on the hit time residuals but with a much wider effective time resolution. W_i is the weight for each hit PMT to reduce the dark noise contamination [39]. g_{vtx} tends to increase when residual time distribution becomes sharper.

The best vertex calculated with BONSAI is used as input parameter for the estimation of the particle direction. We know that Cherenkov angle distribution for relativistic particles peaks at $\sim 42^\circ$, although it can be diluted as a consequence of multiple scattering. Therefore Cherenkov ring pattern is used to estimate the direction via the maximization of a likelihood function. A typical resolution of 25° is found at 10 MeV. A goodness to evaluate the accuracy of direction reconstruction is constructed with the Kolmogorov-Smirnov

test using the spatial uniformity of hit PMTs:

$$g_{dir} = \frac{\max[\phi_i^{uni} - \phi_i^{data}] - \min[\phi_i^{uni} - \phi_i^{data}]}{2\pi}$$
$$\phi_i^{uni} = \frac{2\pi i}{N_{hit}} \tag{2.5}$$

Where ϕ^{uni} is the azimuthal angle of the i-th hit PMT in a toy simulation where the uniform distribution is assumed (evenly spaced hits around the event direction) and ϕ^{data} is the actual hit angle around event direction. g_{dir} decreases with increasing uniformity of hit space distribution. A quality parameter to discriminate low energy background can be defined with g_{vtx} and g_{dir}:

$$ovaQ = g_{vtx}^2 - g_{dir}^2 \tag{2.6}$$

When $ovaQ$ is large the hit pattern is more uniformly distributed and peaked in time.

Cherenkov angle plays a crucial role for particle identification at low energies. Ultra relativistic particles, like positrons from IBD for SRN search or single γ events, show a Cherenkov angle distribution that peaks at $\sim 42^\circ$, on the contrary, low energetic muons and pions, due to their larger mass, tend to have smaller angles. Therefore, in the analysis where it is needed, Cherenkov angle can reconstructed if needed and can be used to identify the particle type.

Energy reconstruction

Since, for low energy processes, the photon yield is sufficiently low to assume single photoelectron incidence for most PMTs, the energy reconstruction can rely only on the number of PMT hits. An effective value, N_{eff}, is computed taking into account corrections due to various effects related to detector conditions.

$$N_{eff} = \sum_{i=1}^{N_{50}} \left[(X_i + \epsilon_{tail} - \epsilon_{dark}) \times \frac{N_{all}}{N_{normal}} \times \frac{R_{cover}}{S(\theta_i, \phi_i)} \times exp(\frac{r_i}{\lambda(run)}) \times \frac{1}{QE_i} \right] \tag{2.7}$$

With N_{50} being the number of hits in a 50 ns window. Moreover:

- Occupancy (X_i): takes into account the fact that for higher energies the number of hits in a single PMT is no more necessary one. Since when this happens surrounding PMTs are likely to be activated too, X_i is computed checking the hits in a 3×3 patch of PMTs around the candidate one. X_i estimates the probability for the i-th PMTs to generate a certain number of photo electrons.

- Late hits (ϵ_{tail}): scattering, absorption/re-emission or reflection can delay (more than 50 ns) the arrival time of some Cherenkov photons. ϵ_{tail} takes into account this effect.

- Dark noise (ϵ_{dark}): dark noise subtraction from occupancy term. It is estimated using an off-time window and depends on the run period and PMT position.

- Correction for dead PMTs (N_{all}/N_{normal}): a ratio between all PMTs and correctly operating ones to account for possible malfunctions.

- Correction for PMT coverage ($R_{cover}/S(\theta_i, \phi_i)$): R is the overall photo-coverage and S the effective PMT area, which depends on the incident angles θ_i and ϕ_i.

- Correction for water transparency ($r_i/\lambda(run)$): r_i is the distance between i-th PMT and the reconstructed vertex while $\lambda(run)$ is the actual water transparent coefficient, periodically measured and dependent on the run.
- Correction for quantum efficiency ($1/QE_i$): measured during calibration, PMT quantum efficiency value is included in this correction term.

Finally, incident particle energy is reconstructed as a polynomial function of N_{eff} with coefficient extracted by calibration measurements. Energy resolution is about 14%, for 10 MeV electrons [39].

2.6.2 Muon fitter

Super-Kamiokande experiences $\sim$ 2 Hz rate of cosmic-ray muons crossing the detector. As we described in Section 1.7, muons induce one of the main backgrounds in 6-20 MeV region by inducing the production of radioactive isotopes. Since, spallation events are tightly correlated in space and time to the preceding muon, a precise characterization of muon tracks and energy deposition is required for background rejection. Muon reconstruction is done with a fitting algorithm called Muboy [88, 89]. Cosmic ray muons are likely to generate an high amount of Cherenkov light in the detector, thus all the events with more than 1000 p.e. ($\sim$ 140 MeV [52]) are fitted by Muboy. It uses PMT hit informations to reconstruct muon tracks and classify muons through the maximization of a likelihood function dependent on the expected Cherenkov light pattern (more details in Ref. [88]). As a first step, Muboy identifies the patch of nine tubes with the highest energy deposition and earliest hit time and uses the central PMT position as a candidate entry point. Since the Cherenkov light is emitted along the track into progressively smaller rings, it is easy to locate the exit point as the center of the region with highest charge of the event. Causality arguments are used to reject noise hits or identify multiple muons. Once candidate entry and exit point are selected, hence a track is reconstructed, these parameters are varied in a grid and used for the likelihood maximization. Muboy is also able to classify four categories of muons depending on the hit pattern and charge. An illustration is shown in Figure 2.12.

- *Single through-going muon*: individual muons crossing entirely the detector. They represent the default fit category as well as the most abundant one ($\sim$82% [90]).
- *Stopping muons*: individual muons stopping inside the detector, namely that leave an entering signal in the OD but do not exit afterwards ($\sim$7%). Muons are classified as stopping if the amount of charge deposited around the exit position is < 300 p.e., the stopping position is estimated looking for a decrease in the energy deposition along the path [90].
- *Multiple muons*: as described in Section 1.7, the same primary cosmic particle can generate a shower where multiple muons travel in parallel tracks ($\sim$7%). These could reach SK and cross it simultaneously depositing a huge amount of light that is inconsistent with the single muon case. Muboy is able to fit multiple tracks though a specially developed software trigger: a muon crossing the detector activate HE trigger which opens a 40 μs window. Once the HE window is open, other possible unrelated muons can cross the detector but will not be tagged (with an estimation of about 100/day [22]). The way Muboy looks for secondary tracks is in the 40 μs time gate is: if 45 hit tubes have a distance from the entry point which would require a travel speed of more than $\sim$ 33 cm/ns, they are associated to secondary tracks for causality reasons. Although in general, tracks fitted above the second or third are

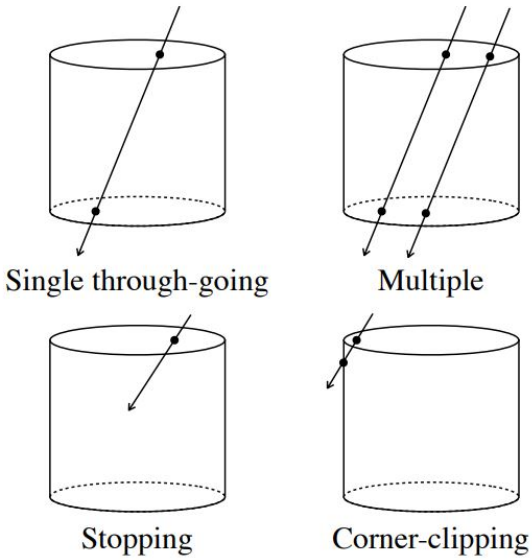


FIGURE 2.12: Illustration of the characteristics of differen type of muons.

not very trustworthy, up to 10 tracks can be fitted. In scenarios where additional tracks cannot be fitted, a muon can be classified as a bundle with only one fitted track. More often in the physics analysis, if multiple tracks are found, only the first tagged track, namely the one with the best fit, is used as a reference to identify subsequent events.

- *Corner clipping muons*: single through-going muons with a track inside the ID with a length of less than 7 m and occurring near the top or bottom corner of the detector ($\sim$4%).

The categorization is highly valuable for spallation rejection since different type of muons have different probabilities to develop showers within the ID and can be associated to distinct processes (i.e. isotopes from stopping muons can derive from μ^- capture on oxygen and can be easily removed with a spherical cut around the stopping point [91]). The quality of the track fit is evaluated and expressed in a parameter called Muboy goodness. Studies presented in Ref. [90] show a resolution of about 100 cm on the entry point for all muon types (up to third track in muon bundles). Direction resolution is found to be 6 degrees.

There is a small fraction ($\sim$ 1.2% [52]) of events that Muboy fails to fit due to particularly strange hit pattern or low number of hits. In this case, a Brute Force Fitter (BFF) performs a fit over a grid of entry and exit points. Although BFF is much slower than Muboy, it manages to reconstruct about 75% of the remaining muons. In addition to the total charge deposited by a muon crossing the detector, the energy deposition per track length is an interesting information for background rejection. In fact, the emission profile, dE/dx, can peak around the point where a shower, possibly containing spallation isotopes, is generated. By mean of ToF informations, light hitting each PMT is therefore assigned to a 50 cm segment along the reconstructed track taking into account water transparency and PMT coverage effects (50 cm is roughly the vertex resolution for events in an energy rage of [3.5, 20] MeV). A likelihood calculation is subsequently applied to smooth dE/dx spectrum. An example of light emission profile is shown in Figure 2.13.

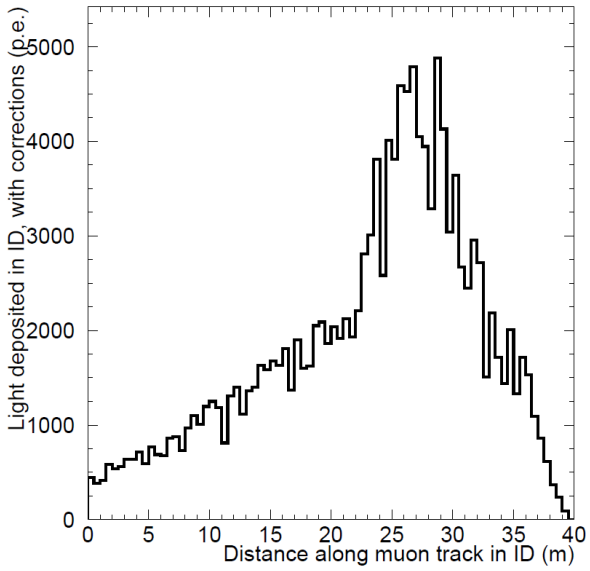


FIGURE 2.13: Example of dE/dx profile for a through-going muon [90].

Chapter 3

Calibration at Super-Kamiokande

At the end of SK-IV phase, the longest running period of the experiment, started in 2008 until early 2018, Super-Kamiokande went through a full refurbishment in view of the transition to SK-Gd phase. The upgrade required major renovation works in the detector in order to repair both structure and hardware and to make sure they are safely usable with gadolinium. A photograph of activities during refurbishment period is shown in Figure 3.1. A complete drain of the ultra-pure water of the tank was necessary as well as an entire recalibration of the detector. This Chapter will describe the refurbishment process and calibration measurements for SK-V, the author actively participated in both works.

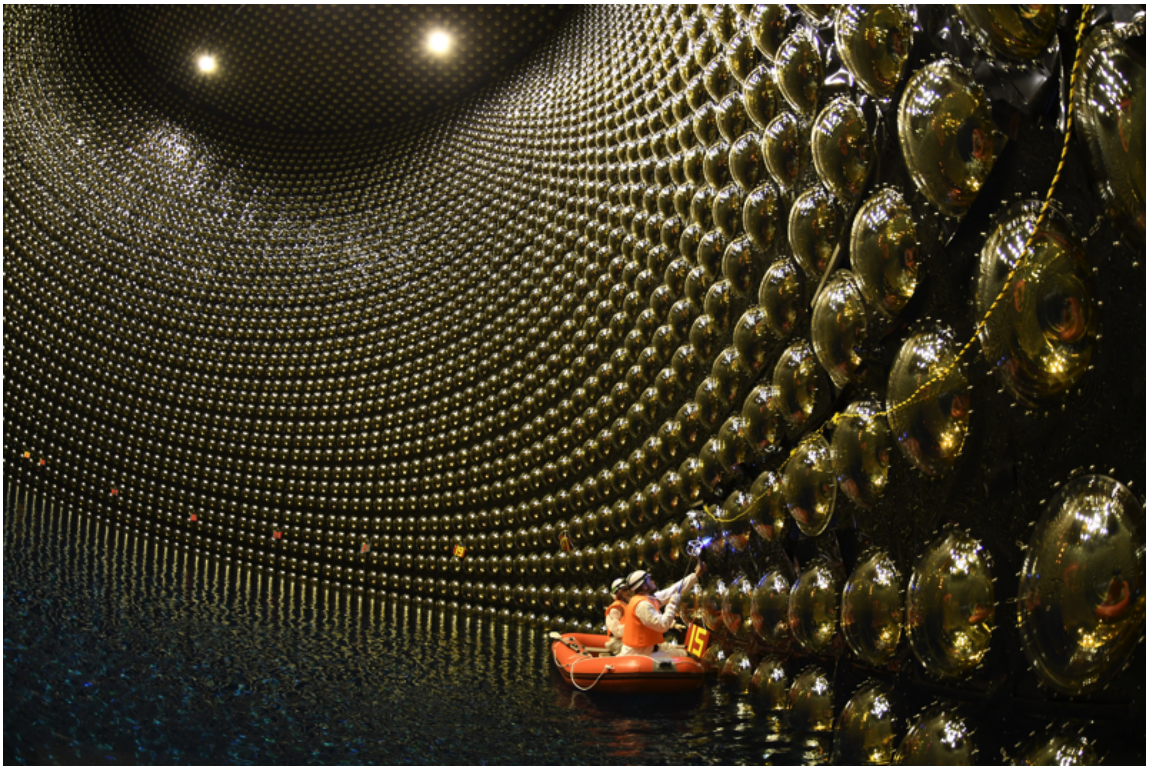


FIGURE 3.1: Photo of the SK inner detector from the refurbishment period, during PMT measurements. Photograph taken by M. Nakahata.

3.1 Super-Kamiokande refurbishment

After 12 years since the last time Super-Kamiokande was open, in June 2018 a major refurbishment work started. SK was officially shut down on May 31st at 9 am and resumed full

operations on January 29th, 2019. Between June and December 2018, upgrade and repair works proceeded tirelessly with a first phase of water draining, until September, followed by a filling period which lasted about three months. In this Section an overview of the motivations behind the refurbishment and a brief summary of the tasks accomplished are presented.

3.1.1 Motivation

The recent upgrade of Super-Kamiokande, SK-Gd, implied a gradual addition of a highly soluble gadolinium sulfate, $Gd_2(SO_4)_3$, in pure water. The goal of SK-Gd project is to reach in the next years a final Gd concentration, in mass, of 0.2%, while the current status presents 0.021% of $Gd_2(SO_4)_3$ (equivalent to 0.01% of Gadolinium). In fact, given its large thermal neutron capture cross-section, Gadolinium is the ideal candidate to increase neutron identification efficiency of the experiment: a photon cascade with energies of about 8 MeV is emitted after a capture on Gd atoms, which makes the event much easier to be identified if compared with the very low energy (2.2 MeV) single photon emitted after capture of Hydrogen. Hence, Gd-doping greatly enhances SK's background rejection capability, making it sensitive to particularly elusive signals. One of the most important prospects concerns the diffuse supernova neutrino background: Super-Kamiokande with Gadolinium dissolved should be able to reach the sensitivity to make the world's first observation of SRN signal [54]. As discussed in Section 1.6, reactor and atmospheric neutrinos are the dominant irreducible background respectively below 10 MeV and above 30 MeV, the window 10-30 MeV is open for measurement but still overwhelmed with background, notably spallation [55]. The current SK search for DSNB neutrinos relies on the detection of a prompt positron event followed by a delayed signal from 2.2 MeV photon after neutron capture of protons. The detection efficiency for low energy photons is currently poor due to the low Cherenkov light emitted, making it challenging to identify neutrons and reduce background. The addition of Gadolinium can be a turning point for SRN analysis: thanks to a capture cross section about 5 orders of magnitude larger than neutron capture on protons ($\sim$ 49 000 barns) and the 8 MeV photon cascade, the observation of the two key signals from inverse beta decays becomes much easier overcoming the poor efficiency of neutron tagging in water [49]. With the target concentration of 0.2% of Gd we will expect 90% efficiency for neutron capture on Gadolinium.

A parallel independent experiment, EGADS (Evaluating Gadolinium's Action on Detector System), was built to test Gadolinium effect on the detector and water system requirements. It consists of a 200 ton cylindrical water tank located in an independent cavern within Kamioka mine. Built in 2010, it started dissolving $Gd_2(SO_4)_3$ in November 2014, after a test period with pure water only. EGADS worked not only as a precursor of SK-Gd, but also a backup detector during the refurbishment period in order to ensure Kamioka observatory live time for an eventual supernova explosion [92]. After the dissolving of 0.2% of gadolinium sulfate it was able to test water circulation system and measure water trasparency and neutron capture efficiency ($\sim$ 84% [92]). Based on the success of EGADS, refurbishment work was planned and accomplished: all the needed modification of the detector were implemented. SK-Gd project has successfully started in 2020.

3.1.2 Refurbishment general tasks

During the refurbishment period a vast spectrum of activities were carried out to prepare SK detector to the upcoming upgrade.

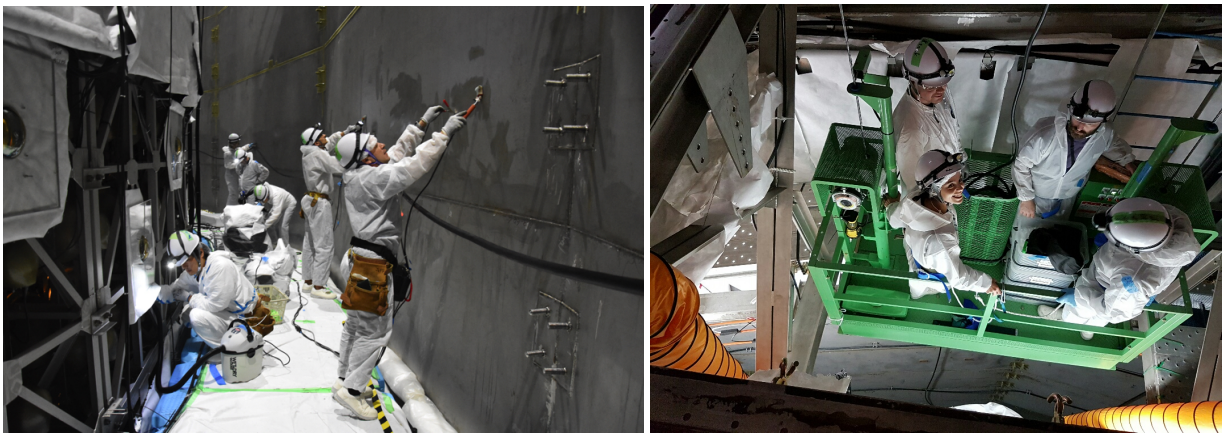


FIGURE 3.2: Pictures taken during refurbishment work. Left: OD repair works on the wall barrel and Tyvek sheets replacement performed every day on a floating platform. Right: gondola to reach water level.

First of all a Gadolinium compatibility testing was performed on hardware parts of the detector: it was essential to ensure conservation of the equipment integrity as well as maintenance of water quality. Soak tests were performed on hardware components as described in [54, 93, 94]. Equipments that were found incompatible were replaced or redesigned from different materials, metallic parts of the detector were examined to catch sight of rust as well as plastic ones to test leaching of dye into the water or to find any change in appearance.

The drain of water in the tank was performed in stages: the detector was depleted of about 2 meters of water every day to make each PMT layer accessible for repair works by mean of floating platforms reachable with gondolas, see Figure 3.2.

Since spots of rust were found over the PMT carrying structure and other metallic components, an extensive cleaning was necessary to avoid water transparency deterioration. This arduous task included a replacements of the Tyvek sheets optically separating inner and outer detector: Tyvek is a white thick reflective material ($\sim$ 80% reflectivity in OD PMTs wavelength range) that allows multiple reflections of Cherenkov light within the OD, contributing to the minimization of light loss due to sparse PMT coverage. It helps increasing the overall detection efficiency. The ID-facing side of Tyvek sheets are black to enhance optical separation between ID and OD. Tyvek also covers tank walls. Together with cleaning and Tyvek replacement, dead OD and ID PMTs were replaced or restored from the moment when the top of OD became accessible. With each water drain, the refurbishment could progress further and subsequent PMT layers became reachable both in ID and OD. All new ID and OD PMTs were tested before being installed as a substitute of a dead or old PMTs. In the case of ID, 136 dead ID channels were replaced by new generation photo-multipliers, called Hyper-K box and line (B&L) 50 cm PMTs (Hamamatsu R12860) with bigger cathode surface resulting in larger photon coverage. These new PMTs (“HK-type PMTS") have an advanced dynode structure as well as optimized glass curvature and focusing electrode. The improvements lead to higher charge resolution. Single photon detection is also improved by a factor 2 (best QE is about 30% versus $\sim$ 20% of the former PMT version while collection efficiency is increased from 73% to 95%) and timing is 2 ns better. They are designed for Hyper-Kamiokande (HK), the next generation of large underground water Cherenkov experiment, making their operation in SK the very first step of realizing HK in the near future. For further details see Ref. [95]. In addition, in the OD, 213 new PMTs were replaced.

Moreover, during the period of upgrade work some calibration sources were replaced

and new systems were installed to monitor in real time detector conditions, such as lasers of the light injector system to monitor water transparency. Measurements on PMT response and detector properties were also carried out.

Finally, the most important goal of the refurbishment was to locate and repair a water leak that had been measured during the last running period of SK. Finding the leak and seal it was critical for SK-Gd, since dispersion of Gadolinium in the environment and water within the mountain would have been dangerous. The tank was repaired during the refurbishment works and all weld lines and support features of the detector barrel wall were sealed with a poly-urea based glue to prevent any future leak. Water level was constantly monitored during filling operation and water loss was found to be constant and about 40-50 l/day showing consistency with expected evaporation and gasifier rate: no sign of significant water leak was found.

The detector upgrade was successfully terminated by January 2019.

3.2 Calibration for SK-V phase

After the 29th of January 2019, a complete calibration of the upgraded Super-Kamiokande detector was required. This date also marked the beginning of SK-V phase.

Calibration is necessary to quantitatively measure detector performances and provide crucial parameters and systematic uncertainties for analysis and simulations. All calibration measurements can be divided in two main categories: those investigating the underlying detector properties, such as PMT response calibration and water transparency monitoring, and those related to physics quantities like vertex resolution and energy scale calibration. Super-Kamiokande employs various calibration techniques for continuous and periodic monitoring of the detector properties in order to keep key parameters up to date and thus improve quality of data taking. A comprehensive summary, even though not exhaustive, of the calibration program fulfilled at the beginning of SK-V phase will be presented in this Section, subjects not treated here can be found in Ref. [78]. We will reserve for the next Section a more detailed description of ID PMT calibration with the presentation of the new methods introduced for the beginning of SK-V: the author was directly involved in the study of PMT response, in the related data taking and hardware work as well as the analysis tasks (discussed in Section 3.3).

3.2.1 Water transparency measurements

It is extremely important to monitor water properties in a large scale detector like SK since changes in water quality and transparency could be sign of a contamination or asymmetries in the detector performances. Various quantities related to water transparency need to be measured in order to have well tuned parameters for analysis and Monte Carlo simulation: scattering and absorption of photons in water are quantified independently with a N_2 laser, while effective attenuation length for low energy physics is estimated through decay electrons from cosmic-ray muons.

In order to determine scattering and absorption coefficient in-situ measurements are performed using mono-directional light sources with well defined wavelengths. Laser beam is injected into SK tank via optical fibers located in different positions. Time of arrival and spacial distribution are then compared to a set of MC prediction samples with varying absorption and scattering coefficients. Water parameters are extracted from the best match MC-Data as a function of time [75, 78]. Based on this water calibration method, for a wavelength value of λ = 400 nm, the water transparency is stable within 15% and approximately 120m [39, 94].

In parallel to laser measurements, an effective water transparency is evaluated on a daily basis directly using Cherenkov photons: since water quality keeps evolving, due to constant water circulation, and contaminations or changes may arise, Michel electrons from cosmic-ray muons are used as a calibration sample for water parameters. Michel electrons spectrum is well known and can be a suitable candidate to keep detector properties under constant monitoring.

3.2.2 Energy calibration

In Super-Kamiokande it is fundamental that the number of hit PMTs are translated into a measurement of energy. Therefore, for low energy events, where only few PMTs are hit, an absolute energy calibration is required. This energy scale in SK is calibrated using three independent methods: LINAC [96], DT generator [91] and decay electrons. Results are used to tune SK MC detector simulation.

LINAC is an electron linear accelerator located on top of SK tank. It produces a mono energetic electron beam, up to 19 MeV energy, that is injected vertically down going at different position of the detector. A germanium detector is able to measure electron energy within about 20 keV precision. LINAC measurements show an accuracy in absolute energy determination better than 1%.

DT is a deuterium-tritium neutron generator that emits isotropically 6.1 MeV γ rays and is permanently deployed into the tank with a moving crane for calibration sources. As LINAC measurements, DT calibration is performed periodically during running phases of Super-Kamiokande. DT not only represents a cross check for LINAC energy scale measurements but is also able to provide direction and position dependence estimations, not possible with LINAC that only have down going beam. Directional bias is found to be less than 1%.

Finally decay electrons from cosmic ray muons represents a cross check for energy scale up to 60 MeV [39].

3.3 ID PMT response

Time and charge informations recorded by SK ID PMTs are the crucial variables to reconstruct energy and position of physics events. The charge collected by PMTs can be parametrized with two factors: quantum efficiency (QE) and PMT gain. We define here QE as the ratio of the photoelectrons collected at the first dynode to the number of incident photons. This interpretation slightly differs from the common definition, also used in Section 2.2.1: QE is normally the ratio between photons absorbed by the cathode and converted to p.e. to the number of photons hitting the PMT surface. However, since once PMTs are installed in the detector it is not possible to isolate the collection efficiency of the first dynode, the probability of the p.e. to reach, hit and be absorbed by the first dynode is incorporated in QE definition. Instead gain is the conversion factor between number of photoelectrons to the PMT charge output, expressed in pico Coulomb (pC). QE plays a crucial role for low energy analysis where most PMT are hit by a single Cherenkov photon, while for high energy events gain calibration becomes crucial to identify the exact number of photoelectrons on single PMTs given its charge response.

In this Section ID PMT gain tuning and relative response calibrations conducted at the beginning of SK-V will be described with a particular emphasis on the new techniques introduced with respect to previous works [78]. In fact, over the period of SK-III and SK-IV data taking, a gradual change of PMT response was observed, most notably a PMT gain increase. However for the last $\sim$ 10 years, most of the operation conditions, in particular

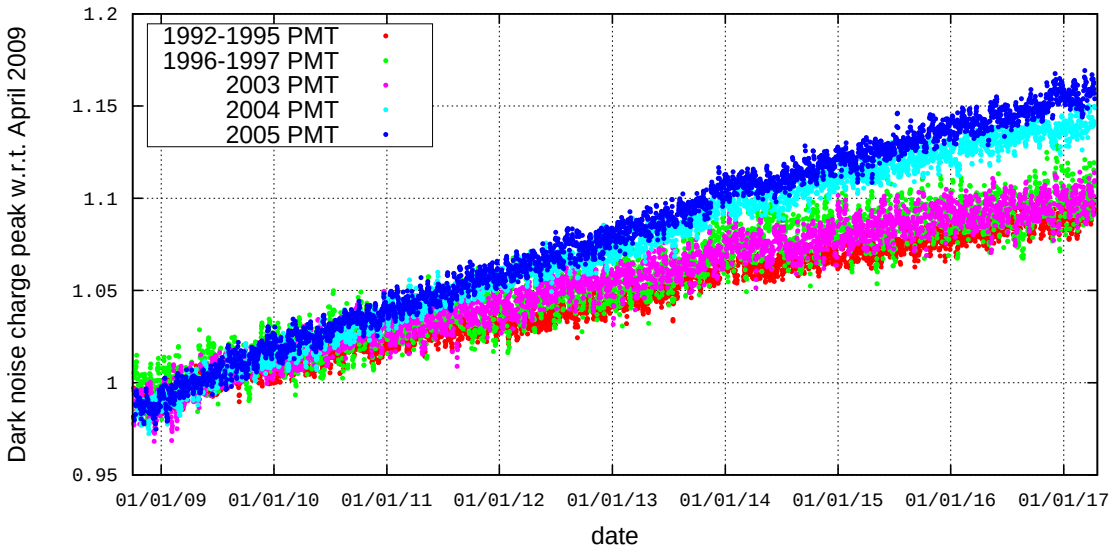


FIGURE 3.3: Relative gain values with respect to gain in April 2009 for ID PMTs as a function of time during the SK-IV period. Data points with different colors corresponds to the averaged gain for PMTs produced in different production years.

PMT high voltage (HV), and ID PMT calibration constant were kept unchanged, only correction factors were applied to the analysis [22, 39]. The deterioration we observed on PMT gain, as well as the fact that 136 HK-type PMTs with improved performances were newly installed in SK, raised the necessity to re-tune PMT high voltage to improve detector uniformity and recalibrate all the basic characteristics of ID PMTs.

3.3.1 Gain drift and its impact during the SK-IV period

The gain of ID PMTs continuously increased during SK-IV phase and the reason why this happened still remains unknown. The time variation of gain values can be measured comparing the peak of dark noise charge spectra for different periods as shown in Figure 3.3. In fact, even in absence of Cherenkov light, a photomultiplier tube can still produce an output current as a consequence of thermal effects within the material, called dark current, that can fake a photoelectric signal. Dark noise, namely the rate of this current, tends to form an isotropic hit pattern and can be measured in region where signal is absent. Although the dark rate remained quite stable in SK-IV, and no contaminations or water quality deteriorations were observed, the peak of dark noise charge kept increasing. Moreover, the rise was not constant all over the detector but it highly depended on the PMT production year. The spread of gain values for PMTs of different production years is also shown in Figure 3.4. On top on this, channel-by-channel variations in charge response were also observed. The main problem arising from gain increase is that the QBEE threshold for PMT signals is effectively reduced: more events, especially dark noise, are now potentially able to pass trigger threshold leading to an higher number of hits recorded. This can have a drastic effect on low energy analysis where number of hits are crucial for event reconstruction. Even with empirical correction factors to account for gain increase, unresolved issues persist: the most significant problem is the non uniformity in the detector response for which the increase of number of hits cannot be simply solved rising QBEE threshold. Therefore, in order to solve the above described problem, PMT gain was completely re-tuned by an adjustment of PMT high voltage. The goal is to globally lower the gain in order, not only to align PMT responses, but also to reduce dark

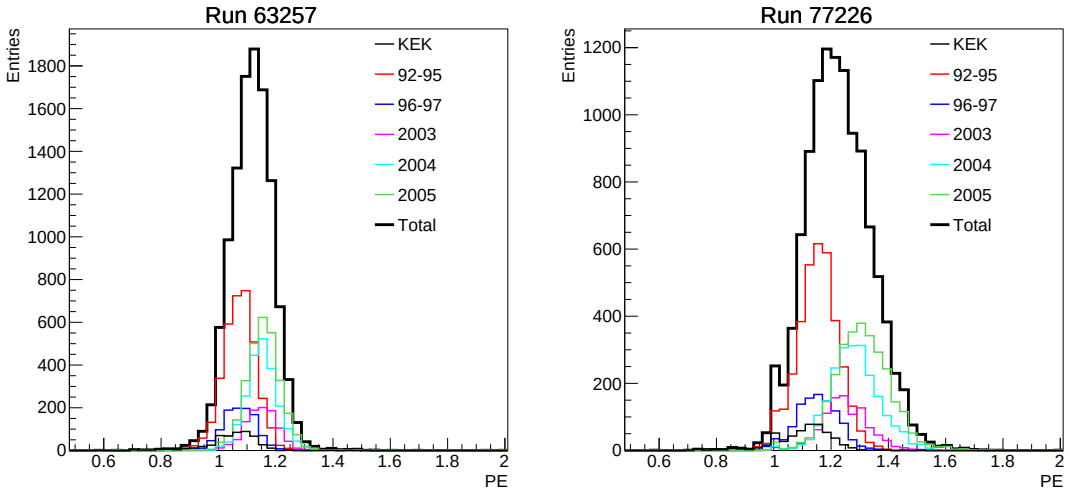


FIGURE 3.4: Distribution of single photoelectrons peak values for calibration source run 63257 taken in Feb. 2009 (left) and for run 77226 taken in Dec. 2017. Lines with different colors show distributions separated by production years.

rate, which could in turn enable to lower trigger threshold, and to retain the dynamic range for high charge events.

3.3.2 Principle of gain adjustment

A new method to tune PMT high voltage was studied and applied to SK-V. A brief summary of the calibration measurement performed in previous SK phases [78] is necessary to understand the reasons why the new method was chosen in 2019.

In order to ensure similar output charge for the same incident light and therefore guarantee uniformity in detector response, PMTs must work under well tuned and relative uniform high voltage conditions. For the beginning of SK, HV calibration was performed with an isotropic multi p.e. light source: a scintillator ball permanently located at the center of SK tank is hit by light from a xenon lamp (Xe) transmitted through an optical fiber [75, 78]. Since geometric acceptance effects could influence PMT response at different positions, 420 reference PMTs, accurately calibrated individually before installation, were distributed uniformly in the detector and used to adjust the HV of neighboring PMTs by aligning charge spectra. An illustration of the previous HV calibration method is shown in Figure 3.5. This calibration technique basically proceeds by making $QE \times gain$ uniform within the detector, in fact it uses an alignment of charge distributions which already incorporate quantum efficiency and gain parameters in an indistinguishable way.

For SK-V an alternative method based on alignment of single photoelectron (single-p.e.) peak positions varying HV values is used and a weaker light source other than Xenon lamp is chosen. In this case the measurements are independent of QE since only one-hit signals are used to reconstruct single-p.e. spectra: this is particular important if we consider that new HK-type PMTs have different performance with respect to the old ones. We list below the advantages and disadvantages of the new and old methods.

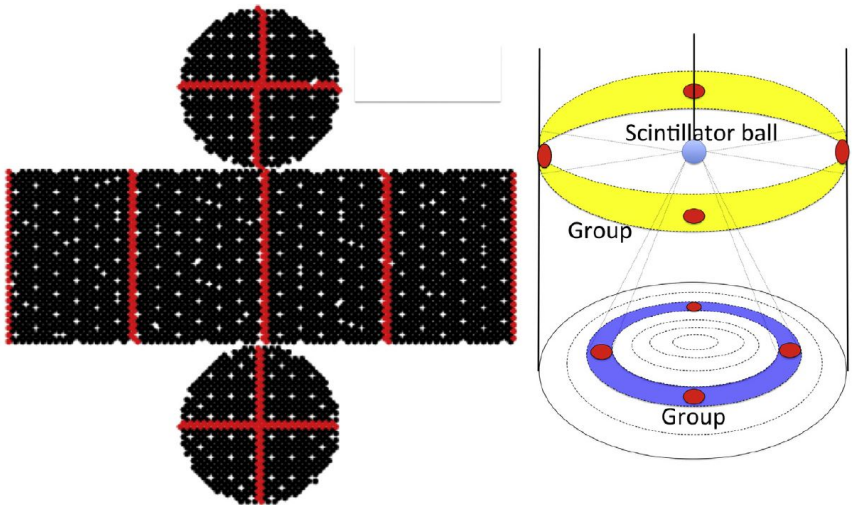


FIGURE 3.5: An illustration of the high voltage calibration using a scintillator ball triggered by a Xe lamp. The red dots represent positions of the pre-calibrated reference PMTs while the coloured bands are the neighboring PMTs whose HV is adjusted according to their reference PMT [78].

Old method [78]:

Pros

- Uniform charge output for multi photoelectron source is guaranteed.
- It is not necessary to fit single-p.e. distributions for each PMTs, which in some cases could not be visible or difficult to identify due to poor charge resolution.

Cons

- Gain is anti-correlated with QE, therefore hit efficiency due to the application of QBEE threshold is not uniform but depends on the performance of the single PMT: a PMT with higher performances, thus better QE (like the new HK-type PMTs), will produce an higher number of photoelectrons at the cathode for the same Cherenkov light. As a consequence it will be tuned with a lower HV, namely lower gain, and the QBEE threshold, which is applied to the PMT charge spectrum, would result more stringent than for PMTs with poorer QE.
- The scintillator ball need to be placed at the exact center of the detector to

New method:

Pros

- QBEE hit threshold results aligned with respect to the single-p.e. peak for each PMT. This ensures uniform hit efficiency which does not depend on QE of different PMTs.
- Calibration results do not depend on position of the light source nor water quality.

Cons

- PMTs with poor or invisible single-p.e. distribution cannot be calibrated with this method.
- The fit process can introduce possible biases or additional uncertainties.
- The response to multiple p.e. light is no longer uniform: a precise calibration of QE for each PMT is required for multi hits events.

exploit the symmetry of PMT disposition.

- Water quality during the period of data taking needs to be perfectly known.

The new method was preferred for SK-V because setting QBEE threshold at the same fraction with respect to the single-p.e. shape allows to better control gain shifts, independently from QE. More practically, the device to install the diffuser ball at the center of the tank was no longer available as well there was no more a group of reference PMT since their gain had changed along with all the others throughout the SK-IV phase and thus they can no longer be used as a guidance.

All this said, gain was tuned using a low intensity source, specifically a laser diode Hamamatsu PLP-10 with peak wavelength at 404 nm (SN-LD), injected into the detector through a diffuser ball with a NIM clock to trigger its signal. The raw output signal from activated PMTs is ADC (analog-to-digital) counts, a count-to-charge conversion is automatically applied using a polynomial function, which depends on the charge range, as the first step of data reduction step in SK. The factor converting charge into number of photoelectron is extracted with calibration measurements and it will be described in the following Section. Single-p.e. peaks were used as reference for PMT gain and a voltage scan was applied to find the target HV. Note that the peak position is extracted from a fit applied to each PMT distribution and, since SK and HK PMTs have different charge responses, fitting functions are also distinct. Moreover, inside SK there are actually other two PMT generations in addition to HK PMTs: the oldest one, SK2 PMTs were produced in the late 1990s while SK3 PMTs were produced after the accident in 2000s. All the three generations have different charge response and therefore different voltage dependence of gain, as shown in Figure 3.6. Tubes with bad fitting results or requiring a target voltage beyond SK safe range were kept with SK-IV voltage, they were 62 in total. The initial target gain was chosen to match the conditions at the beginning of SK-III (1.65×10^7). The first measurements with new HV setting showed that the averaged gain was lowered by a factor 14.3% with a channel-by-channel difference reduced by 26.3% thus restoring uniformity conditions that had been compromised during SK-IV decade [95].

3.3.3 Relative gain

As we anticipated, gain is a conversion factor between the output charge from the PMTs and the observed number of photoelectrons. A unique value that describes medium properties of the detector is essential for event reconstruction, however, despite HV tuning described above, gain value might slightly differ from channel to channel. The determination of single PMT gain is done in two steps: firstly measurements are performed to derive relative gain corrections, namely residual gain variations across PMTs (described here). Subsequently an independent calibration determines the average gain over the whole detector (described in the next section). Once average gain and deviation from it are known for each tube, tables with gain values for each ID PMTs can be derived and implemented in simulations and analysis. This section describes the derivation of the relative gain corrections for all ID PMTs.

A correction factor $G(i)$, for each PMT i, is defined such that

$$Q^{pe}(i) = \frac{Q^{pc}(i)}{G(i) \times G^{abs}}, \tag{3.1}$$

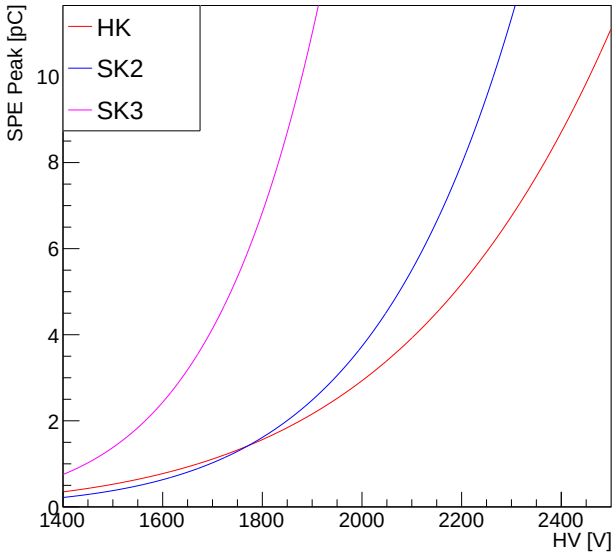


FIGURE 3.6: The voltage curve dependence of gain value, quantified by single-p.e. peak position, for SK2 (blue), SK3 (pink), and Hyper-K PMTs (red). On x axis there are high voltage (HV) values while single photoelectron peak position (SPE peak) is on the y axis.

where Q^{pe} is the charge measured in photoelectrons, the actual variable used in event reconstruction, Q^{pc} is the charge read out by the electronics, and G^{abs} is the absolute global gain averaged over all PMTs whose measurement will be described in the next section.

The principle of relative gain measurement is to employ a ratio method between high and low intensity calibration lights flashing from the same source and same position in the detector. The light source is a laser system that sends photons both to a diffuser ball through optical fibers and to a monitor PMT, used in the analysis for selecting laser events, as depicted in Figure 3.7. Neutral density filters are remotely controlled to attenuate the light to the desired intensity. The high intensity light illuminates all PMTs simultaneously providing a large number of photons for each PMT, such that an average charge per event $Q^{obs}(i)$ can be measured for each PMT i. The low intensity light is tuned such that it fires only a few PMTs per each event ensuring they are mostly single-p.e. hits. The corresponding average hit rate per event is measured to be $N^{obs}(i)$[1]. In principle, if the intensity is known and constant for fixed-position light source, the following equations hold:

$$Q^{obs}(i) \propto I^H \times a(i) \times \epsilon^{qe}(i) \times G(i), \tag{3.2}$$

$$N^{obs}(i) \propto I^L \times a(i) \times \epsilon^{qe}(i), \tag{3.3}$$

where I^H and I^L are the intensities of the high and low intensity light, respectively, $a(i)$ is the geometrical acceptance, and ϵ^{qe} is the quantum efficiency. Since QBEE discriminator

[1] The observed hit rate or probability of observing a laser hit in a single PMT in an event is given by $P_{hit}^{laser} = P_{hit}^{ontime} - P_{hit}^{offtime}$, where the ontime and offtime windows are defined later in this section, and the latter gives an estimate of the dark hit contribution. An occupancy correction based on Poisson probability, $P(x) = \frac{e^{-\mu}\mu^x}{x!}$, is applied to estimate the true mean hit rate $\mu := N^{obs}(i)$ by assuming the probability of observing no hit is $P(0) = 1 - P_{hit}^{laser}$ with $P(0) = e^{-\mu}$, resulting in $N^{obs}(i) = -\ln P(0)$.

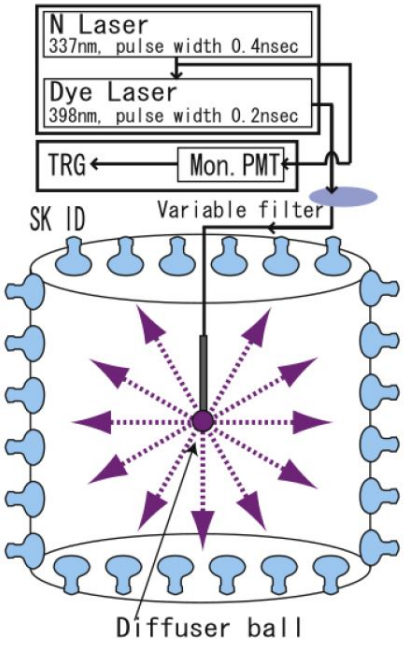


FIGURE 3.7: Diagram of relative gain calibration setup showing a laser split between a diffuser ball deployed into the SK tank and a monitor PMT.

threshold is low (a small fraction of p.e.), N^{obs} is not impacted by percent-level variations in gain, namely N^{obs} does not depend on $G(i)$: studies in Reference [78] proved that 10% gain change impacts $N^{obs}(i)$ with only 1.5% change. Then taking the ratio of Eq. 3.2 and 3.3:

$$G(i) \propto \frac{Q^{obs}(i)}{N^{obs}(i)} \tag{3.4}$$

gives the gain of each PMT modulo a normalization factor common to all ID PMTs. The relative gain for each PMT is obtained by dividing out this normalization by the average $G(i)$ over all PMTs.

Relative gain factors grouped in different PMT types are shown in Figure 3.8. The larger relative gain spread (RMS) seems to be dominated by the older PMTs (SK2 PMTs): this is due to their larger single-p.e. distributions.

3.3.4 Global gain

"Global gain" is the averaged conversion factor between the charge read out by the electronics and input number of photoelectrons, G^{abs} in Eq. 3.1. G^{abs} value is also referred as "pc2pe factor", namely pico Coulomb to photoelectrons. Pc2pe factors for SK-III and SK-IV were 2.243 and 2.658 respectively, reflecting the gain increase.

The gain conversion factor can be generally obtained from the mean value of the single photoelectron distribution: since we are interested in the average response of the detector, the global gain will be extracted from the cumulative single-p.e. distribution, i.e. the sum of all the single-p.e. spectra for all the ID-PMTs. Indeed, once the HV settings have been determined as described at the beginning of this Section and applied to all the PMTs, the single-p.e. distributions of all the ID PMTs are expected to be aligned, therefore we are able to obtain a meaningful single-p.e. distribution by adding them up. In order to build the cumulative single-p.e. spectrum data were taken with a low light intensity source and charge informations recorded by each PMT were collected: first, a custom program

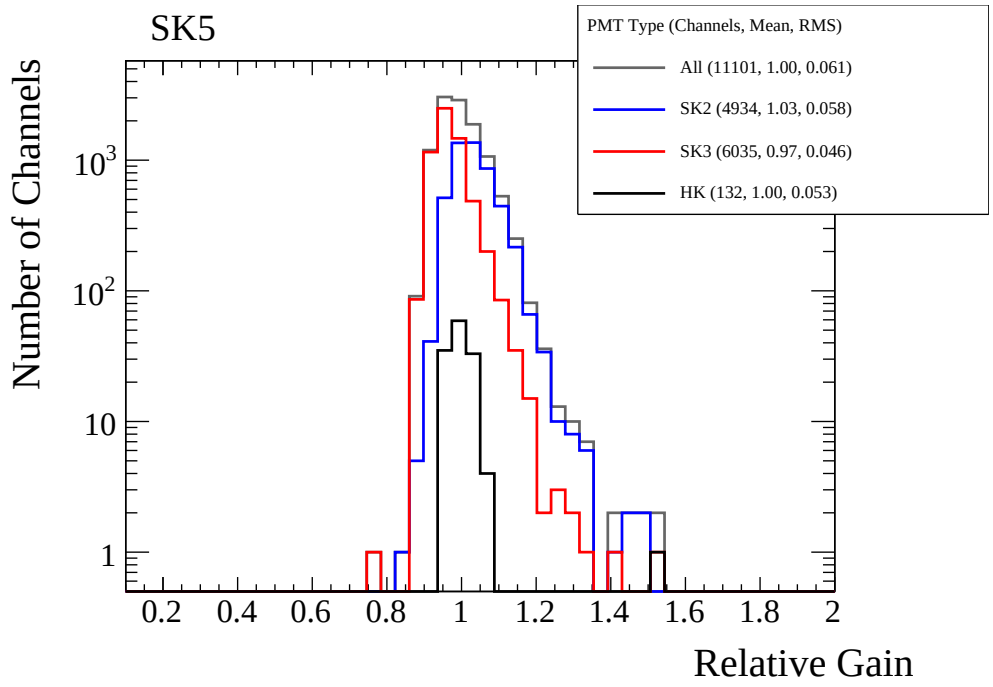


FIGURE 3.8: Relative gain distribution separated by PMT type.

is applied to extract time and charge from the raw data, including application of ADC to pico Coulomb conversion and calibration maps to ensure the hits from all channels are well clustered in time (more details will be described later in this Section within timing calibration description) and second, vertex reconstruction is performed. It is important to clarify from the beginning that, due to the presence of an intrinsic electronic threshold, the distribution resulting from the data taking needs some manipulation before extracting the pc2pe factor: indeed measured spectrum is not completely reconstructed but has a cutoff in the low charge region that would introduce uncertainties in the evaluation of the whole single-p.e. distribution shape close to 0 pico Coulomb, unavoidable if the goal is the evaluation of the mean value. In fact, differently from HV tuning, global gain calibration requires to extract the mean of single-p.e. distribution, not simply the peak position. As a matter of fact, in case of single-p.e. light the mean amount of charge deposited would represent the average response for a single photon hitting PMTs and can be used to scale charge to number of photoelectron conversions for multiple p.e. signals. For this reason, reconstruction of the charge spectrum required, in addition to data with the standard SK-V HV setting, some ancillary measurements with an increased HV ($\sim \times 1.5$) in order to study the single-p.e. distribution behaviors at low charge values. Note that increasing the high voltage means taking data with an higher PMT gain.

Data sets

For the global pc2pe calibration, a uniform and stable source of single photoelectron level light needs to be used: a Nickel+Californium source which isotropically emits $\sim$ 9 MeV gamma rays fulfills those requirements. For this reason, global gain measurement is often refereed as the “Nickel calibration". As shown in Figure 3.9, a ^{252}Cf neutron source was inserted in the center of a $\oslash$ 16 cm Ni ball and held there by a brass rod; the system was then positioned in the SK tank by means of a new automated deployment system (ASDS) through the ID central calibration hole at (35.3, -70.7, 0.0) cm. ^{252}Cf has an half life of 2.65 years and decays through α decay (96.9%) or spontaneous fission (3.1%). In case of

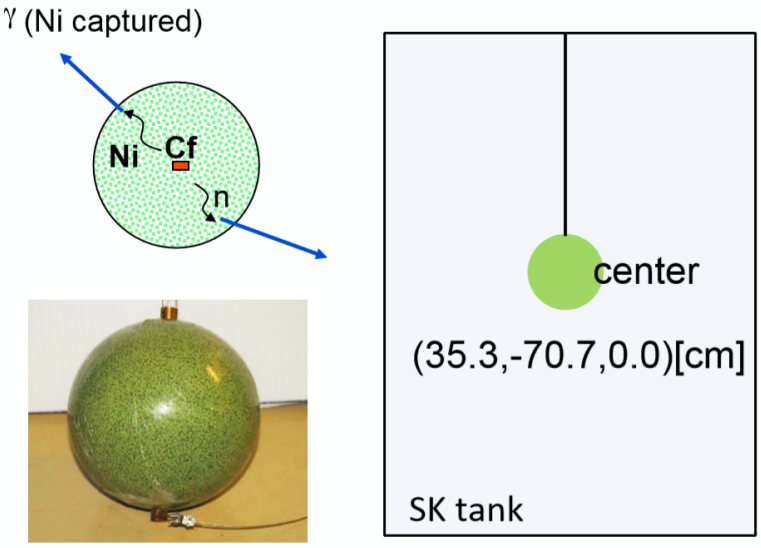


FIGURE 3.9: Schematic view of the Ni+Cf source (top left) and its positioning inside the SK tank (right). Picture of the Ni ball (bottom left).

fission, an average of 3.76 neutrons of about 2 MeV each are emitted, thermalized and subsequently captured by the surrounding Ni mostly through ^{58}Ni$(n,\gamma)^{59}$Ni with γ of $\sim$ 9 MeV[97]. Other less abundant Ni isotopes can emit γ or γ-cascades with lower energies. The spherical shape of Ni ball ensures uniformity of light emission, moreover this system is preferred to a low intensity laser because it better represents low energy events in SK in terms of energy spectra and wavelength without introducing additional filters. The resulting γs undergo Compton scattering or pair production inducing Cherenkov light from the electrons above the Cherenkov threshold: from Poisson statistics, more than 99% of the observed light is due to single photoelectrons since the average charge produced by every event is only 0.004 p.e./PMT. As mentioned, both Ni data with “normal gain", i.e. with HV tables tuned fo SK-V, and “increased gain" were taken. In this last setting, QBEE threshold had also been lowered in order to investigate the shape of the charge distribution at low charge values. Table 3.1 summarizes the calibration runs analyzed and their setting parameters.

TABLE 3.1: Ni data taking runs and settings. Note that QBEE threshold of 0.69 mV roughly corresponds to 0.25 p.e.

Ni calibration	Run 80493	Normal gain	QBEE threshold = 0.69 mV
Ni calibration	Run 80497	Increased gain	QBEE threshold = 0.40 mV

Lastly, in order to compare these data sets with different HV and QBEE thresholds, data with a multi photoelectron level light were taken with same settings applied: we used Xenon laser data where Xe light is injected into the tank through optical fibers and passes to a scintillator ball. The goal is to compute a scaling factor for each channel that quantifies the impact that increasing HV have on the charge distribution, namely how much the single-p.e. spectrum has been shifted for the runs in Table 3.1. This factor is not equal for all PMTs, not only because different PMT types have different dependence on HV changes, as shown in Figure 3.6, but also because variations on a PMT-by-PMT basis may exist. In addition, the advantage of using an high intensity light source ($\sim$ 60 pC at peak) is that the mean value of charge distribution is not affected by QBEE threshold: low

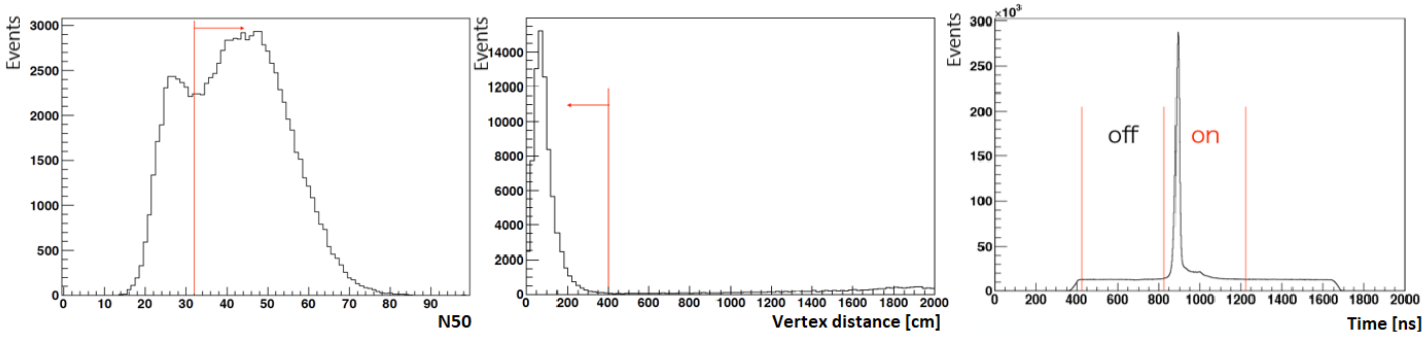


FIGURE 3.10: Distribution of number of hits in a 50 ns window (N50) and cut value (left). The first peak, below cut value, represents background hits above QBEE threshold. Distance between true and reconstructed vertex and cut value (center). Definition of on-time and off-time regions (right). y axis represents the number of events for each one of the three plots.

charge events that do not pass electronic threshold are negligible compared to the total amount of charge deposited and can be ignored. In this way, we have a clear estimation of how much the mean charge value changes when the gain is increased, for each PMT. Xe data are summarized in the Table 3.2.

TABLE 3.2: Xe data taking settings.

Xe calibration	Run 80526	Normal gain	QBEE threshold = 0.69 mV
Xe calibration	Run 80528	Increased gain	QBEE threshold = 0.40 mV

Data analysis

Given their different single-p.e. distributions SK-type and HK-type PMTs are treated separately even though primary selection conditions are common: event reconstruction quality, as expressed in Eq. 2.4, is required to be > 0.4, a cluster of at least 32 hits within a 50 ns window is necessary to remove accidental background which tends to have smaller number of hits and distance from the real position of the source and the one reconstructed from BONSAI has to be less than 4 m. See Figure 3.10 for illustration of selection requirements. Moreover, since there is not an integrated system to trigger calibration events, signal hits are distinguished from background selecting a specific time window in the ToF subtracted time distribution of the recorded hits: as it is shown in Figure 3.10 two regions had been defined, one corresponding to the hits generated from the Ni ball ("on-time window") and a second one related to events due to background and dark noise ("off-time window"). In particular the on-time window corresponds to $(825 < t < 1225)$ ns, where t is the hit time after time of flight subtraction, while the off-time region is $(425 < t < 825)$ ns. Finally the charge distribution of Nickel data is constructed as a subtraction of the charge distributions of on-time hits minus the charge distribution of the off-time ones. See Figure 3.11.

Both Ni data with normal and increased gain are subjected to those selection processes. Single-p.e. distribution for the run with higher gain allows to have an hint of the shape it assumes at low charge, this region is invisible in the data with official gain settings because it is cut out by the effect of QBEE threshold. See Figure 3.12 on the left. In order to compare normal and increased gain a rescaling factor is computed for each channel

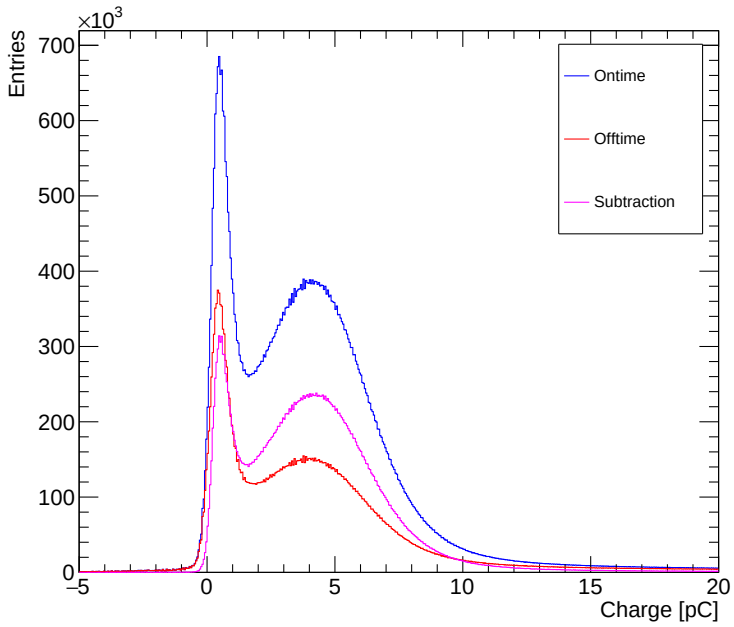


FIGURE 3.11: On-time (blue), off-time (red), and on-time minus off-time (magenta) charge distributions for run 80497 of Ni data. The effect of the subtraction largely reduces the ratio between peak and dark rate pedestal height.

using Xe runs:

$$\text{For each i-PMT} \qquad \textit{scaling factor(i)} = \frac{Q_{mean}\ \text{(nominal gain)}}{Q_{mean}\ \text{(increased gain)}} \tag{3.5}$$

Where Q_{mean} is the mean value of the charge distribution of the Xe data for the two different HV settings (nominal and increased). This factor represents the true increased/normal gain ratio for each cable and it will be applied on a PMT-by-PMT basis on Ni data. Selections for Xe data are quite different from the ones applied to nickel runs: since Xe events are automatically triggered during the data taking there is no need to apply cuts on number of hits. Once Xe scaling factors are applied on the Ni data the single-p.e. distributions with different gains are superimposed and comparable as shown in right Figure 3.12. This procedure is applied separately for SK and HK PMTs: because of the different shape of the charge distributions, visible in Figure 3.13, the computation of the pc2pe factor is performed with two different methods. New HK PMTs have a narrower distribution thanks to their better performances (higher charge resolution, superior photon detection efficiency described in Section 3.1.2) and QBEE threshold has a more efficient effect in reducing dark noise effect in the low part of the spectrum. In the case of SK PMTs, the higher gain data set show a more pronounced exponential peak around 0 that helped the extrapolation below the electronic threshold. However, in the case of Hyper-K PMT no clear peak was seen around 0 even after the gain was increased.

Pc2pe determination: SK-type PMTs

The goal is to reconstruct the full shape of the single-p.e. distribution: by mean of Ni data with higher gain we obtained a better knowledge of the charge region below the threshold, Figure 3.13, however, in order to extend the spectra till 0 pC we need to proceed with

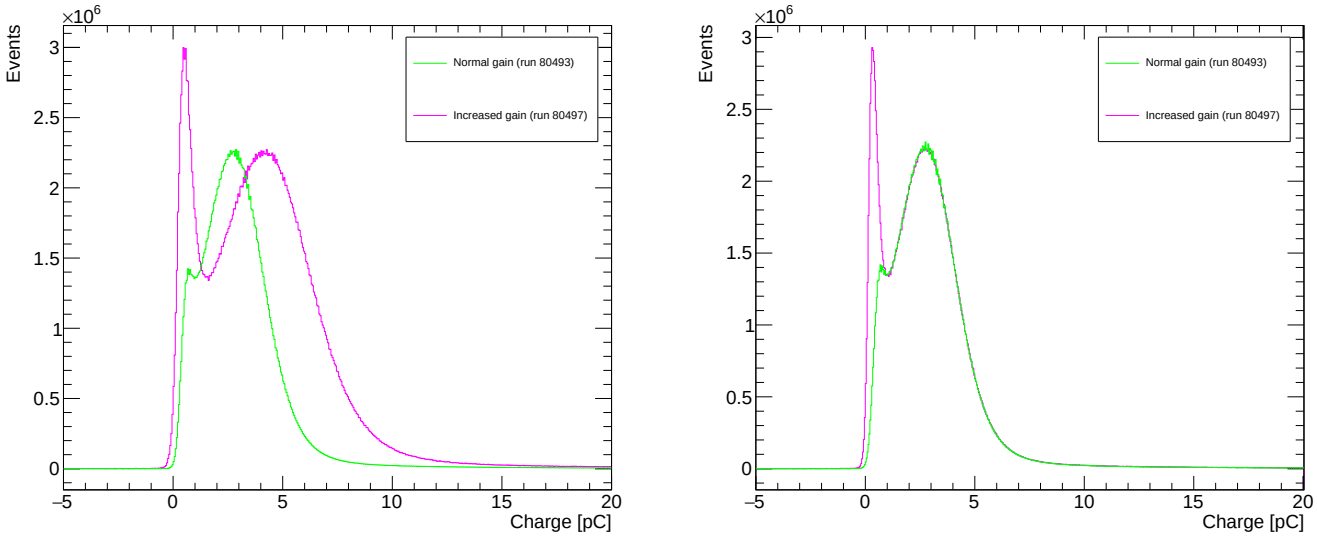


FIGURE 3.12: Left: charge distribution of the Ni data with normal (green) and increased (magenta) HV, i.e. gain (only SK-type PMTs). It is clear how the events taken when the HV is higher and QBEE threshold lower contain details on the charge behavior close to threshold. This is due to the fact that increased gain shifts the single-p.e. peak at higher pC values and a larger number of low charge hits pass the QBEE threshold. Right: same histograms after Xe scaling factors are applied to increased gain Ni data.

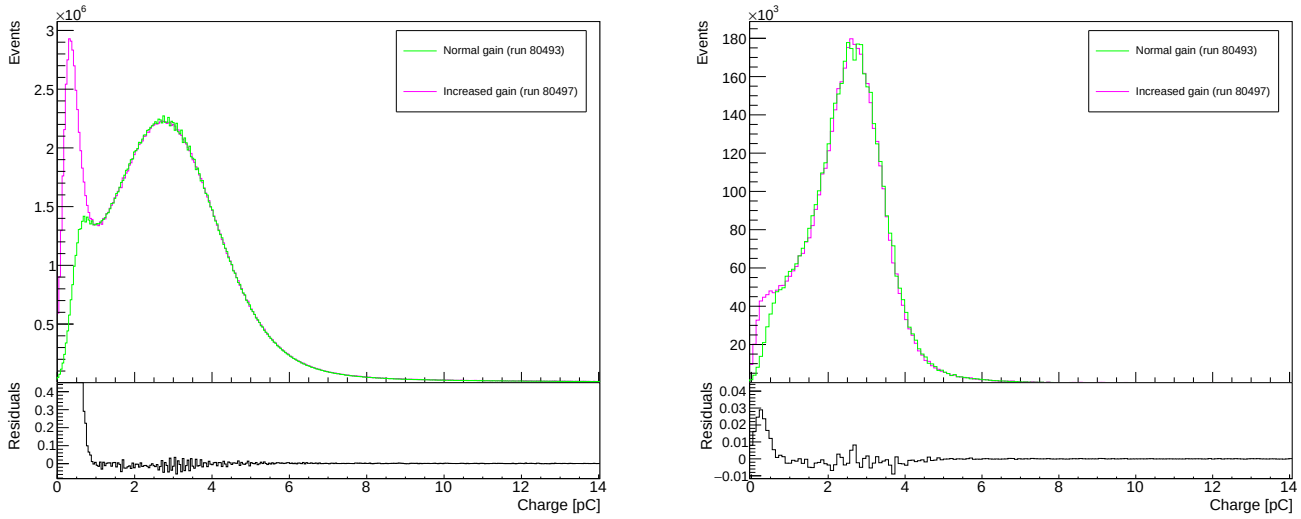


FIGURE 3.13: Charge distribution of the Ni data with normal (green) and increased (magenta) HV after Xe corrections. Residuals are shown in the bottom pad. Left: SK-type PMTs. Right: HK-type PMTs.

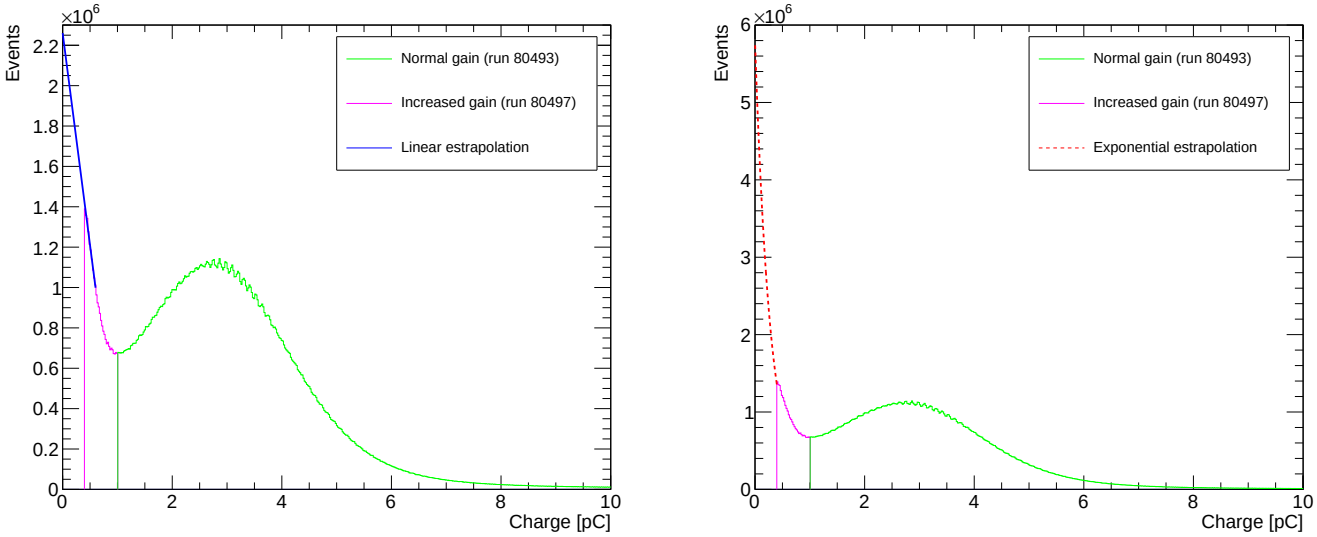


FIGURE 3.14: Single-p.e. charge distribution for the Ni data with linear (left) and exponential (right) extrapolation to zero pC. The Pc2pe values obtained are 2.63 and 2.41 respectively, if mean charge value is calculated from 0 to 100 pC.

an extrapolation. Firstly, the method used for calibration of SK-III phase [78] was applied for consistency. A linear fit is used to extract the parameters for the extrapolation in a the fit region of [0.4, 0.7] pC. Unlike in Ref. [78] where the systematic uncertainty derived from extrapolation was negligible, the assumption of a linear behavior of the charge shape close to zero results ambiguous: indeed, in parallel to the linear extrapolation, an exponential + gaussian fit analogous to what was used to tune the HV setting showed satisfactory results, as presented in Figure 3.14. The reason of this ambiguity can be explained by the fact that Ni data with increased gain were taken with with $\sim \times 1.5$ gain and $\sim \times 0.6$ QBEE threshold for SK-V instead of $\sim \times 2$ gain and $\sim \times 0.5$ QBEE threshold for SK-III therefore leading to a poorer reconstruction of low charge spectrum. This is due to a degradation of detector condition from SK-III to SK-IV, such as an increase of dark rate, a worse charge resolution of QBEE electronics with respect to old ATM system and global PMT aging, as well as problems encountered during the data taking (hit-rate and trigger-rate became too high for the DAQ). The old method would eventually lead to very different pc2pe values, namely mean value of the full charge distribution, depending on the extrapolation hypothesis.

To overcome the ambiguity a new method was implemented: instead of relying on the extrapolation to 0 pC, it was decided to compare the single-p.e. shape from SK-V data with the one encoded in the detector simulation, determined at the beginning of the SK-III phase. Thus, for the final evaluation of the global pc2pe factor proceeded in this way:

- SK-III single photoelectron distribution is assumed as a reference. This assumption is supported by the fact that behavior of charge distribution at low charge was better known during SK-III and the extrapolation to 0 pC was reliable.

- Reference single-p.e. distributions, expressed in “number of photoelectrons" units, that were separated for SK2- and SK3-type of PMTs, are weighted by the number of PMTs of each category present during SK-V and combined to reconstruct the cumulative charge distribution. To do that, different values of pc2pe conversion factors are applied to obtain the distribution in pC units: these charge distribution

predictions with different pc2pe assumptions, from a pc2pe = 2.41 to 2.60 at pace of 0.01, are compared with SK-V data+extrapolation distributions.

- Histograms of the ratio between prediction and data distributions are fitted with a straight line in order to extract the pc2pe values from SK-III template that better adapts to SK-V data. The value for which the linear fit gives a slope closer to zero is taken as pc2pe factor for SK-V. Figure 3.15 shows the overlay of the distributions and the fitting results for the best pc2pe. Outcomes of the scan for different pc2pe conversion factors are presented in Figure 3.16.

To conclude, global pc2pe value for SK-type ID PMTs in SK-V is found to be 2.46. The corresponding absolute gain value, G^{abs}, is evaluated by dividing pc2pe factor with the elementary charge and results to be 1.54 $\times 10^7$. Relative gain corrections for each channel are applied to extract charge to photoelectron conversion on a PMT-to-PMT basis and gain tables are implemented into SKDetSim. Figure 3.17 shows the cumulative single-p.e. distribution used in the simulation for SK-V. While for low energy studies the number of hits is sufficient, for simulations of multiple photons in ID-PMTs, charge values from this distribution are drawn and summed to model the PMT response. The fact that in Super-Kamiokande, the charge response of each PMT is converted by the average gain in the tank and corrected by the relative difference of each individual tube, determined separately by independent calibrations, ensures that uncertainties in the voltage adjustment results are covered and detector response is well modeled.

Pc2pe determination: HK-type PMTs

Unlike SK-type PMTs, Hyper-K box and line PMTs were newly installed during refurbishment work during summer 2018 and no templates from SK-III are available. In addition, their charge response is very different from SK-type PMTs as a consequence of their improved characteristics, as shown in Figure 3.13. For this reason, absolute gain calibration has to follow the previously described method used for calibration of SK-III [78]: rely on the extrapolation to zero charge. Therefore, pc2pe factor will be defined as the average of the charge distribution, from 0 to 100 pC, once it has been prolonged to zero pC. One other substantial difference is that, due to the shape of the charge spectra for HK-type PMTs, the fit is an horizontal line. Figure 3.18 shows the reconstructed single-p.e. distribution: pc2pe value for HK-type ID PMTs is 2.45.

Finally, data from Ni source and extrapolation of charge distributions are also used in MC simulation to extract electronic threshold behaviors: in fact, QBEE threshold is not an absolute wall that stops all signals below a certain value, rather more likely it has a certain efficiency depending on the charge signal. The experimentally observed distribution, with threshold effect folded into it, as the green plots in Figure 3.13 are compared with the partially-experimental partially-reconstructed charge distribution (Figures 3.17 and 3.18) and their ratio is implemented to model single-hit threshold behavior, namely to extract a probability density function for a single photon to pass the threshold value.

3.3.5 Quantum efficiency

Quantum efficiency varies from PMT to PMT due to differences in the manufacturing between different PMT types and production years. Therefore it is crucial, especially for low energy analysis, to map QE values for each channel. For this purpose, Nickel+Californium data, already used for absolute gain calibration, are employed as low intensity source that satisfies Eq. 3.3. While the number of hits per PMT can be counted, the real number of

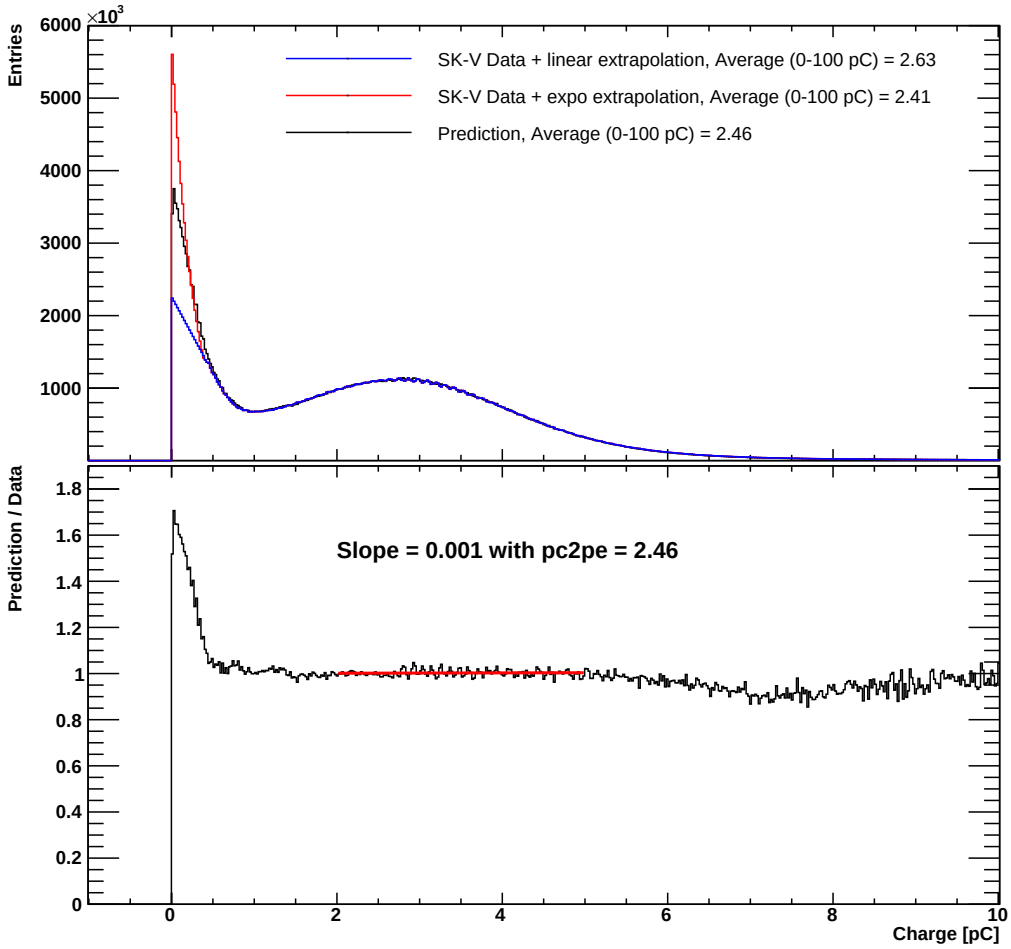


FIGURE 3.15: Comparison between SK-III template single-p.e. distribution (black on the top figure), which is converted in charge units with a pc2pe value of 2.46, and single-p.e. distributions from SK-V Ni data with linear (blue) and exponential (red) extrapolation. The legend box reports the average charge values of the extrapolated spectra from 0 to 100 pC, showing the ambiguity raised from uncertainties in the extrapolation method. In the bottom figure it is shown the ratio between the SK-III template and the single-p.e. distribution with linear extrapolation, a fit is performed and the best pc2pe value correspond to the lower value of the slope parameter (value shown in the bottom figures).

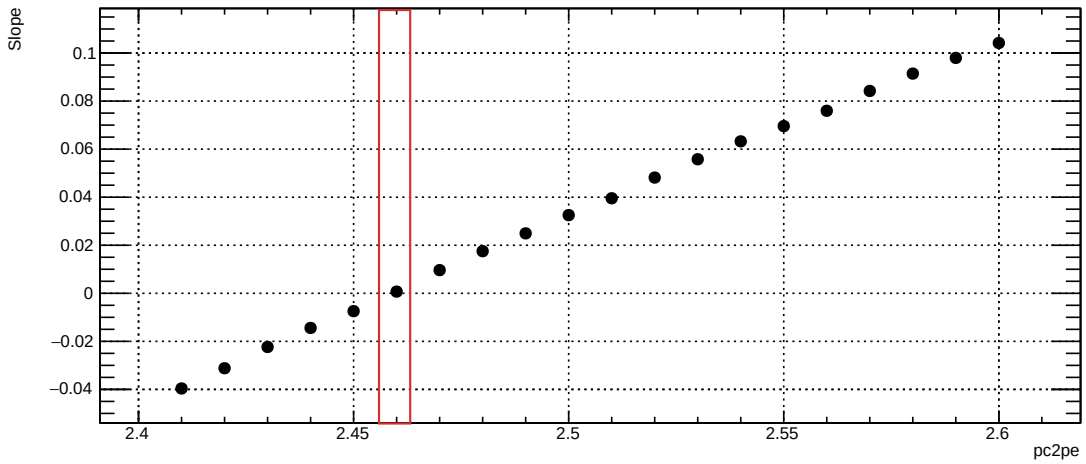


FIGURE 3.16: Straight line slope values obtained from fit of prediction/data ratio for different pc2pe conversion factors applied. The red box shows the value corresponding to the minimum slope, taken as pc2pe value for SK-type ID PMTs.

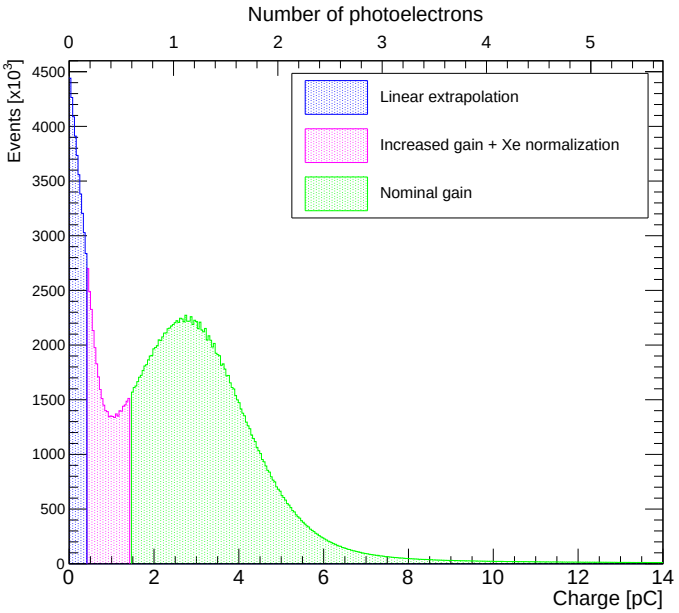


FIGURE 3.17: Cumulative single-p.e. distribution for SK-type ID PMTs. Bottom x-axis shows charge units while the conversion in number of p.e. is illustrated via the top x-axis. This conversion is used for SK-V analysis and simulations.

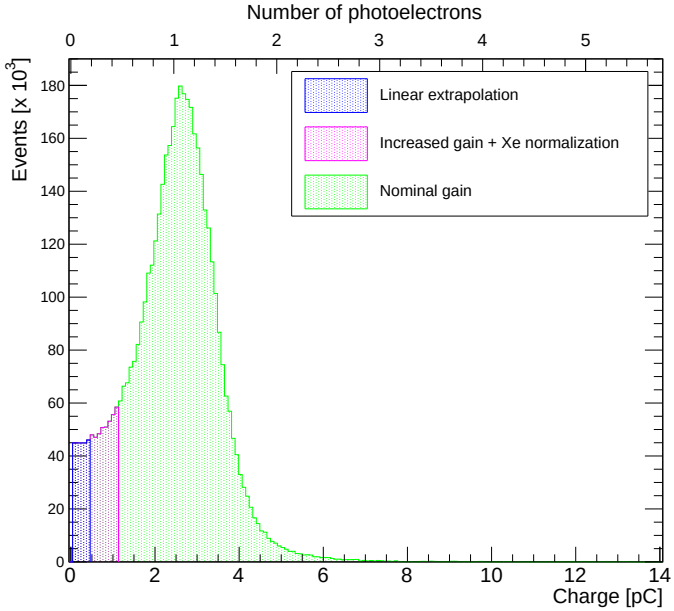


FIGURE 3.18: Cumulative single-p.e. distribution for HK-type ID PMTs. Bottom x-axis shows charge units while the conversion in number of p.e. is illustrated via the top x-axis. This conversion is used for SK-V analysis and simulations.

photons reaching each tube surface needs to be modeled with MC simulations. The ratio of observed and predicted number of hits contains information on QE, once light propagation effects, acceptance and PMT solid angle are corrected. An overall normalization is applied to remove the dependence on the absolute light intensity. The Monte Carlo simulation takes into account water transparency that is measured as described in Section 3.2, reflection, scattering and absorption effects and geometrical acceptance, function of incident angle. Note that HV tuning method described at the beginning of Section 3.3.2, is based on the alignment of QE$\times$gain for each PMT. The differences in QE implies nonalignment of single PMT gain, making relative gain calibration crucial.

3.3.6 Timing calibration

Time information is essential for vertex and direction reconstruction. To understand whether the hits in different PMTs are related to the same interaction, one needs a precise time measurement free of systematic biases. In SK, this is achieved by applying a channel-by-channel correction in time for different signal charges. The PMT time response does not depend only on the photoelectron transit time but also on cable length, electronic processing time and time-walk. The latter is an effect for which for a fixed trigger level the time of threshold crossing depends on pulse amplitude and its rise time: an higher charge signal will rise more rapidly and causes a fake “early arrival" relative to a lower charge pulse. The laser system, with a precise time stamp for signal emission, used for relative gain calibration is employed as light source for timing measurements. In particular, due to the tight correlation between time information and pulse height (i.e. charge), these measurements provide 2-dimensional lists of time and charge response, called TQ tables, used for data treatment and in MC.

To conclude, with the new high voltage tuning, resulting in lower gain, ID PMTs are expected to have a lower dark hit rate that strongly affects the study of low energy physics. This condition may lead to a lowering of SK energy threshold in the future and improve sensitivity to multiple target signals such as solar and supernova neutrinos. In the mean time, the data collected during the calibration described in this Chapter are used to update SKDETSIM to SK-V environment, which helps to improve event reconstruction. Last but no least, the calibration framework and software developed represented a great example for the calibration of SK-Gd phase. All of these improvements combined together, including the major upgrade achieved during the 2018 refurbishment, will help to evolve this 20-year-long experiment to a higher stage of physics studies.

Chapter 4

Simulation of cosmogenic induced spallation

Chapter 4 presents the structure of a complex Monte Carlo simulation built to investigate muon-induced spallation reactions in Super-Kamiokande. The different simulation components, the packages used and their interfaces are described in details as well as relative validation studies.

4.1 Motivation

Despite the fact that Super-Kamiokande is one of the world's largest low-energy neutrino detector, the measured rates of neutrinos from astrophysical sources is modest compared to other signals, making background rejection a crucial part of the analysis. As described in Section 1.7, low energy neutrinos ($\mathcal{O}(10)$ MeV), such as solar and supernova relic neutrinos, suffer from a major background contamination below 20 MeV (electron kinetic energy): radioactivity from spallation induced by cosmic-ray muons interacting within the detector volume. The 1000 m rock shielding given from Mt. Ikenoyama on top of SK assures a drastic reduction of cosmic-ray background, muon rate results of about 2 Hz [50]. However, although muon identification is easy thanks to their large energy deposits and the OD veto system, spallation forms a background that is overwhelming if compared to rare signals like neutrinos from past core collapse supernovae (see Fig. 1.11). Moreover, radioactive isotopes are not only directly generated by the cosmic muons, but, more often, they are daughters of particles in showers initiated by muon interactions with water. These cascades are produced as a consequence of muon energy-loss processes where secondaries can travel for several meters from the initial muon track. In addition, some spallation isotopes have long half-lives, up to tens of seconds, comparable or larger than the time between two successive cosmic muons. The geometrical distance from the muon path, as well as their long life, makes spallation rejection quite challenging: cylindrical cuts around the muon track as well as selections based on the time distance between the primary muon and the candidate spallation event have to be limited to avoid an excessive discard of signal events due to an increase of dead time (the rejection of all low energy events within a cylinder of few meters around the muon track, for few seconds would lead to a $\sim$ 20% dead time increment [53]). In order to avoid a huge signal loss, SK has developed complex rejection methods based on likelihood evaluations with variables like time and transverse distance from muon track as well an empirical variable linked to the muon energy loss [83, 97]. In particular, it was found that significant fraction of spallation events show a spatial link to the point of the muon track from which the highest amount of Cherenkov light originates. Identifying the position of this charge deposition peak, which we now know is associated to the creation of a shower as a consequence of the muon interaction with water atoms, can give a hint on the position of spallation isotopes

[51, 98]. Using this information, SK applies stringent cuts only to a limited fraction of the muon track improving background elimination without increasing the dead time. For solar analysis, these selections result in a spallation rejection efficiency of about 90% but no significant improvement in the dead time reduction is achieved with respect to previous SK analysis [32]. For SRN search, the weak signal is still drown under spallation making DSNB detection extremely difficult below about 16 MeV [50]. It results in a final performance of more than 90% background rejection for 55-90% signal efficiency, depending on reconstructed energy.

It is important to underline that SK cuts to remove spallation background have only been developed from empirical studies rather than simulations and theoretical calculations. In order to further reduce the remaining spallation rate and reach the sensitivity for outstanding discoveries, among all SRN signal, an improvement of background rejection is necessary. A detailed study of spallation processes from cosmic muons using simulation packages can shed light on our understanding of muon energy loss properties and radioactive isotope production. Tracing the physical characteristics of the parent particles generating spallation, the way muon-induced showers are involved in the process as well as their geometrical properties is fundamental to reduce spallation background in Super-Kamiokande. Moreover, a detailed simulation can distinguish different isotopes and enlighten their abundance and their individual properties, impossible with a study only based on data.

Several published studies based on simulation packages such as GEANT [85] and FLUKA [99] are available for muon induced spallation in scintillating material [100, 101]. Both FLUKA and GEANT incorporates experimental results from accelerators and underground experiments. However, only one major calculation of spallation reactions in water, presented in a series of three papers by S.W. Li and J.F. Beacom has been published in the scientific literature [53, 98, 102]. Despite the fact that the total isotope production rate is similar between water and scintillating materials, unstable isotopes are different and much less frequent in water (an order of magnitude difference [53]). In addition, many of them decay invisibly in a water Cherenkov detector. This is the reason why calculations based on scintillator detectors cannot represent spallation in water and Li and Beacom's production remains a unique study. Their work is based on the prediction given by FLUKA package, which is able to simulate the relevant processes involved in radioactive isotope production, such as hadronic, low-energy neutron interactions and photodisintegration. Moreover, FLUKA was extensively used for calculations of muon interactions in underground detectors for its accuracy in modeling muon-nucleus reactions and hadronic processes [53]. Li and Beacom's work presents a comprehensive study in water-based detectors, with a specific attention to Super-Kamiokande. However, it introduces some major simplifications: antimuons are not simulated, the angular dependence of the cosmic muon flux is ignored and the energy spectrum of muons is not independently computed using accurate SK informations. Moreover, details regarding the characteristics of SK detector and event reconstruction are unknown to the authors. Therefore, the goal of this thesis is to reproduce their results and from there build a complex simulation, not only based on FLUKA but using other additional packages, to accurately study muon interaction in SK detector that leads to the creation of spallation isotopes. The goal is to interface simulation results with SK reconstruction tools in order to make possible a direct comparison with data and enable a deep understanding of spallation processes thanks to simulation insights. A meticulous calculation can eventually lead to tailor better rejection cuts for SK, especially now that Gadolinium significantly increases the neutron identification ability of the experiment. Finally, a deeper knowledge of spallation processes in water would have a considerable impact for background rejection in shallower depth new generation experiments, like Hyper-Kamiokande that is currently in construction.

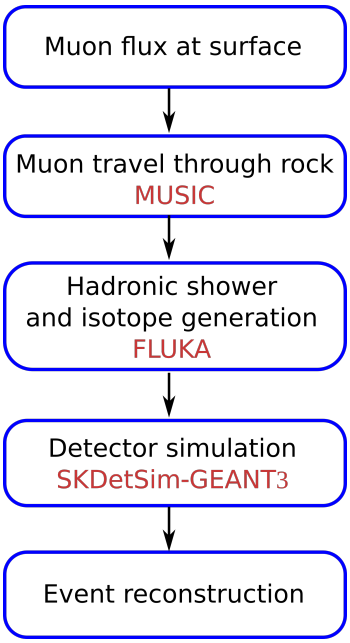


FIGURE 4.1: Schematic description of simulation pipeline, from the generation of cosmic ray muons traveling through the Ikenoyama mountain to event re-construction in SK.

4.2 Simulation pipeline

The full simulation is composed of three main blocks, each of them involves a specific simulation package and need to be sequentially interfaced with the others. The pipeline can be summarized as follows:

- *Model of the cosmic muon flux at surface and propagation of muons through rock to obtain flux at the depth of Super-Kamiokande*: the simulation construction begins with the Monte Carlo muon propagation code MUSIC [103]. Once chosen a parametrization of the muon flux at sea level, the code is used to simulate muon transport through rock and generate the energy and angle distributions of cosmic muons intersecting the detector tank. These are treated as input for the following part.

- *Simulation of muon interactions in water, with particular focus on muon-induced spallation shower generation*: FLUKA [99] Monte Carlo code takes MUSIC results as inputs to generate muons in a cylindrical tank with SK dimensions and study their interactions in water. Spallation reactions are specifically selected and recorded.

- *Reconstruction of muon showering processes in Super-Kamiokande detector*: spallation products simulated with FLUKA can produce a signal in SK. The detector response is modeled by the official Super-Kamiokande Monte Carlo simulation, SKDetSim (described in Section 2.5), reconstruction algorithms are applied afterwards to extract the variables for the analysis.

The three steps briefly illustrated are schematically summarized in Figure 4.1. They represent the basic structure of the spallation simulation we aim to build. Additional details and relative implementations will be described in this Chapter.

4.3 Muon flux at Super-Kamiokande

The angular and energy distributions of cosmic muons reaching Super-Kamiokande tank are crucial inputs that need to be known if we aim to build a detailed simulation of spallation processes inside the detector. While muon directions can be determined from data by mean of track reconstruction algorithms, the measurement of muon energy is not directly feasible in Super-Kamiokande experiment. For this reason, accurate Monte Carlo simulation is the most reasonable alternative to determine it.

4.3.1 MUSIC simulation

MUSIC (MUon SImulation Code) [103–105] is a three-dimensional simulation code for muon transport through large thickness of matter, it is especially designed for underground physics experiments. MUSIC integrates models for all the different types of muon interactions with matter leading to energy losses and deflections, such as: pair production, Bremsstrahlung, ionization and muon-nucleus inelastic scattering. Angular and lateral displacements due to multiple scattering are also taken into account. All the transported muons are considered to be ultra-relativistic and their energy can be up to 10^7 GeV; if, during the propagation, the muon energy becomes less than the muon mass it means that the muon was stopped inside the traversed material.

The main MUSIC code is composed of two separate parts: in an initial phase all the cross sections needed to model the energy losses in a specific material are computed. At this stage parameters such as atomic number, density and radiative length defining the rock need to be specified. Subsequently, the previously calculated cross sections are used to transport a muon with given initial energy and direction cosines to a specific distance in the chosen rock. Finally, MUSIC returns the particle parameters (energy and direction angles) at the end of the transport. Given that the accuracy of MUSIC propagation relies on detailed description of the crossed material, both rock chemical composition and the mountain topology are crucial informations. Together with this, a suitable model of the muon flux at surface is needed. This choice is independent from MUSIC and must be defined by the user.

Muon flux at surface

Most of the muons at sea level are produced high in the atmosphere, typically 15 km, from the interactions of primary cosmic rays and their secondaries in the air, which, for muon production, are principally decays of charged mesons [106]. The shape of their energy and angular distributions at surface reflects a convolution of the production spectra, the energy loss, and the decay probability in the atmosphere. Since the accuracy of MUSIC propagation of muon through rock depends on the parametrization of the surface flux, a suitable choice of the parametrization has to be made. In 1990, Gaisser developed a formula to describe the cosmic-ray muon flux at sea-level [107]:

$$\frac{dI_\mu}{dE_\mu} = 0.14 \left(\frac{E_\mu}{GeV}\right)^{-2.7} \left[\frac{1}{1+\frac{1.1E_\mu cos\theta}{115GeV}} + \frac{0.054}{1+\frac{1.1E_\mu cos\theta}{850GeV}}\right] \tag{4.1}$$

where I_μ is the differential flux in units of $cm^{-2}s^{-1}sr^{-1}$, E_μ is the energy of the muon and θ is the zenith angle. The formula is valid with the assumptions that the curvature of the Earth and the muon decays can be neglected. In 2006, a modified version of the Gaisser equation, optimized for the description of muon fluxes at shallow depth experiments like

Super-Kamiokande, was published [108]. This new formula introduces phenomenological corrections for the low energy region of the spectrum that better fits global experimental data used as a reference. For the present work, the modified Gaisser parametrization has been implemented to model the muon flux before the transport through the mountain. A rejection sampling method is used to implement the modified Gaisser parametrization before MUSIC transport.

Cosmic muons can reach SK depth both alone or in bundles, composed of muons generated from the same primary meson in a very collimated cascade. Since this work is primarily focused on the production of radioactive isotopes, and on the ability of FLUKA to model it, we restrain our studies to single crossing muons only and neglect effects associated with muon bundles. As we will describe throughout the following Chapters, spallation events are rare enough to assume that within a bundle of muons with almost parallel tracks, reaching SK within a very short time, only one would be responsible of generating radioactive isotopes. Therefore, results obtained for single muons can be easily extended to muons in bundles. Differences could arise from track reconstruction uncertainties rather than physical differences in the observables.

Mountain topology

The most important input setting of the muon transport simulation over a mountain profile is the digital map of the surrounding topology. A detailed knowledge of the elevation profile is required since its accuracy directly affect the accuracy of the computation. Elevation data, with a 50 m mesh, were collected by the Geographical Survey Institute of Japan in 1997 around peak of Mt. Nijyugo-yama, mapping a vast region including Mt. Ikenoyama and the nearby area. The same elevation map was used by KamLAND collaboration to evaluate the muon flux in Ref. [100], and it is shown in Figure 4.2. The digital map described in the Figure is also adopted for this work.

Once we are in possession of the elevation data, we can proceed in the construction of a map with the origin at the center of the detector and parametrize the coordinates in terms of (θ, ϕ, h). As illustrated in Figure 4.3, θ represents the zenith angle and ϕ the azimuthal one, which is set to zero when the final muon travels from east to west. Instead, h is the rock thickness a muon has to cross in a given direction to reach the detector.

Figure 4.4 shows the total slant depth for given $cos\theta$ and ϕ. Given that the average value of the rock depth is more than 2000 m, with a minimum value of 986 m, we assumed that the dimensions of the detector are negligible with respect to the distance traveled by the muons: this assumption means that a point-like-detector approximation is applied. For simplicity, at this point of the simulation we consider the muon flux of the detector to be equal to the one at its center, in a subsequent step, corrections need to be considered to account for the fact that the tank, being a cylinder, is not isotropically symmetric in all directions as it would be for a point or a spherical shape.

Rock composition

An other key ingredient of the simulation is the rock chemical composition: this choice affects MUSIC code in the generation of the cross section files used to calculate the muon energy losses. Three different types of rock have been tested:

- Ikenoyama rock: an Inishi type of rock with average atomic and mass number $\langle Z \rangle$ = 10.13 and $\langle A \rangle$ = 20.43 a.u., rock density and radiative length are ρ = 2.70 g/cm^3 and λ = 25.9666 g/cm^2. Details of the chemical composition in elementary percentage can be found in Table II of Ref. [108].

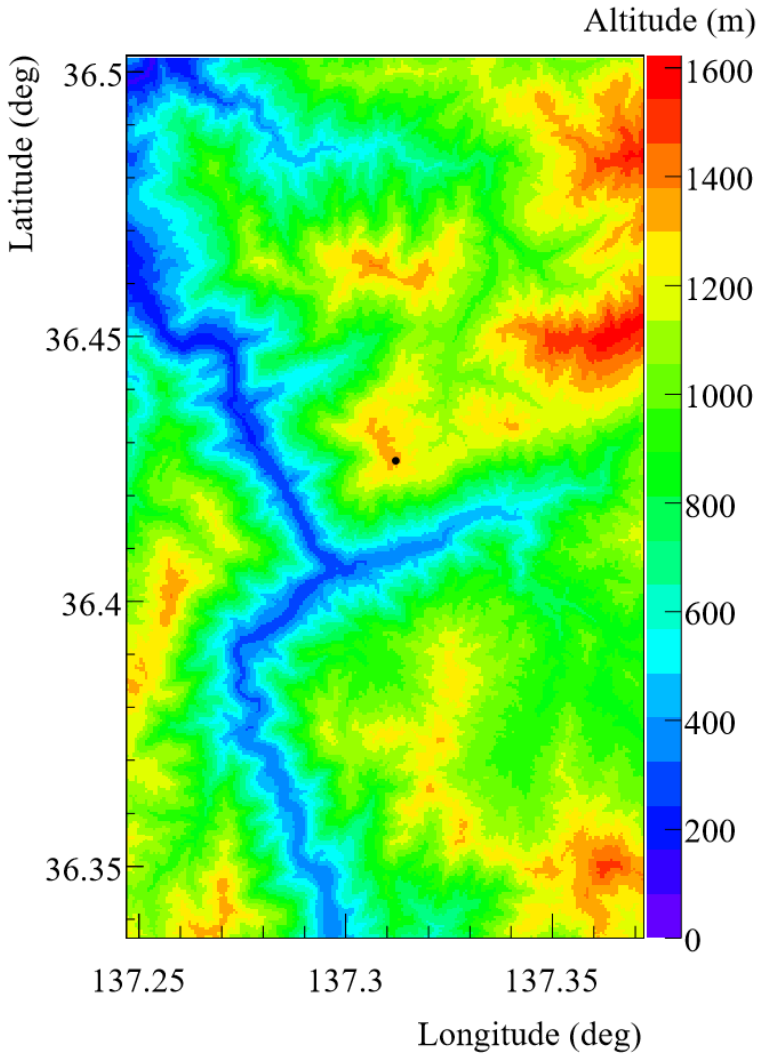


FIGURE 4.2: Ikenoyama topological profile. The black point is the location of KamLAND: it is separated from Super-Kamiokande by 187 m with N66.6E orientation with respect to SK. The bottom of KamLAND experiment is aligned with the top part of SK tank, namely 350 m above sea level [108]. Figure taken from [100]. Note that Super-Kamiokande coordinates are: 36° 25′ 00″ N, 137° 18′ 00″ E.

- Generic skarn: a generic mixture of standard type of rocks found in skarn-type mountains like Mt. Ikenoyama. It is composed of granite (70%) and calcite (30%) by weight with average atomic and mass number $\langle Z \rangle = 10.22$ and $\langle A \rangle = 20.55$ a.u., rock density and radiative length are $\rho = 2.70$ g/cm^3 and $\lambda = 25.411$ g/cm^2 [100].
- Standard rock: material defined in MUSIC libraries, with average atomic and mass number $\langle Z \rangle = 11$ and $\langle A \rangle = 22$ a.u., rock density and radiative length are $\rho = 2.65$ g/cm^3 and $\lambda = 26.48$ g/cm^2.

For the three cases we assumed $\langle Z \rangle$ and $\langle A \rangle$ to be constant with depth and position inside the mountain and we took an average value of the density as input parameter. A stratified approach is not possible due to the lack of a detailed geological profile. For each one of the three types of rock tested, we run a MUSIC simulation until we have collected a final sample with 10^6 muons reaching Super-Kamiokande with non zero energy. Subsequently, for all the three samples, we varied the input density from 2.65 g/cm^3 to 2.75 g/cm^3. As

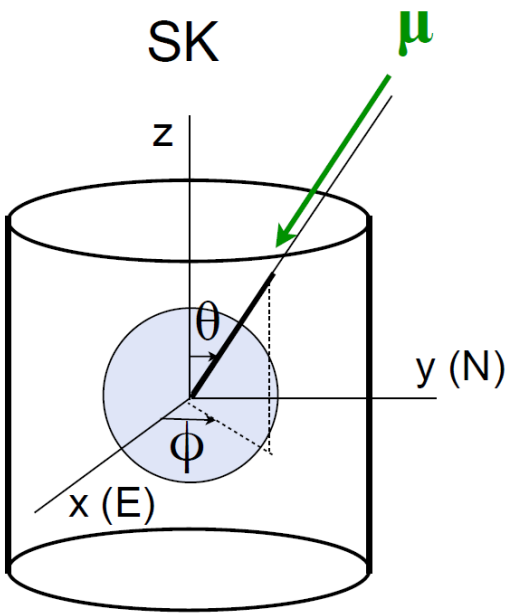


FIGURE 4.3: Coordinated system used to parametrize muon directions: the zero is set at the center of the Super-Kamiokande detector, θ is the zenith angle and ϕ the azimuthal angle.

it can be seen in Figure 4.5, the differences in rock composition does not impact the angle distributions: in fact, as found in Ref. [108], up to 10% changes in $\langle Z \rangle$ and $\langle A \rangle$ do not significantly affect the muon flux. The choice of rock type has negligible effect on the energy distribution too. Moreover, from Figure 4.6 we see how varying densities changes the mean energy value of less than 2%: an energy value of $\bar{E} = 258 \pm 3$ GeV is found, where the uncertainty covers the range of densities for different rock types. This value is in a agreement with KamLAND studies of Ref. [100]: the locations of the two experiments are horizontally separated by 187 m and vertically disaligned of only few tens of meters, making KamLAND results a robust reference to check our simulation outcomes. Finally, from what emerges from results in Figure 4.5 and 4.6, we do not expect the details on the rock composition and density to be a relevant source of uncertainties for the simulation. From now on we will use Ikenoyama rock parameters and $\rho = 2.70$ g/cm^3 as input of the MUSIC code.

In conclusion, despite the huge flux of cosmic muon reaching the Earth's surface, about 6.5×10^5 μ m^{-2} h^{-1} [101], the mountain above Super-Kamiokande strongly attenuates this flux by about a factor 10^6. At surface, the minimum incident value to be able to reach the Super-Kamiokande detector is about 600 GeV. Primary muons with lower energies are absorbed. θ and ϕ distributions, shown in the two-dimensional contour plot of Figure 4.7 (left), reflects the local topology of the Ikenoyama mountain: the flux has angle dependent variations due to thickness differences of the overlaying rock. The muon flux distribution as a function of the energy is represented in Figure 4.7 (right): the value at SK depth is calculated as

$$F_{SK} = \frac{E_{SK}}{E_0} \times F_0 \tag{4.2}$$

where F_0 and E_0 are the values of flux and energy at surface and E_{SK} is the remaining muon energy after MUSIC transport. The spectrum quickly drops at high energies due to the falling shape of the cosmic ray spectrum at sea level. The integrated value of the flux

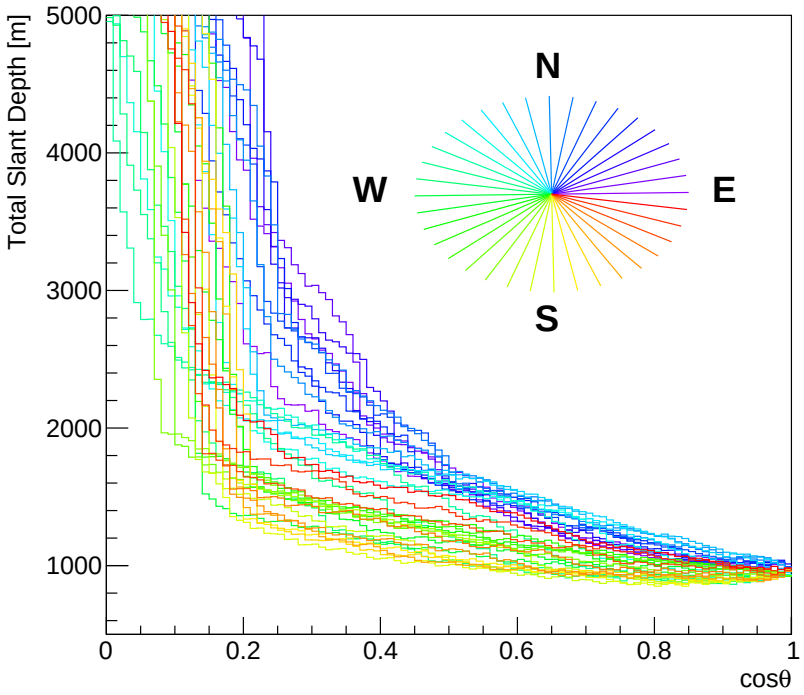


FIGURE 4.4: Total slant depth a muon has to cross for a given $cos\theta$ and ϕ. ϕ equal to zero correspond to east direction.

is 1.54×10^{-7} cm^{-2} s^{-1}, which correspond to a muon rate of 1.97 Hz: this value is within published results [52, 109–111].

Sampling of muon entry points

Once muon distributions are obtained by MUSIC, they need to be implemented as a source for FLUKA simulation. As discussed in Section 4.3.1, MUSIC gives as output flux distributions at the center of the SK detector. Since the tank is a cylinder, the effective cross section of a muon track with the detector depends on its direction, unlike the case of a point-like or spherical detector where it is isotropic. We develop a specific rejection sampling method to choose the muon generation point in order to take into account the flux modifications induced when passing to a cylindrical shape. Here it is described the technique we use to pick the generation points of muons in FLUKA (called "entry points") in order to adjust MUSIC flux distributions when the cylindrical shape of the detector is considered. To find how to sample the distribution of muon entry points in SK we will use the following hypotheses:

1. The spatial distribution of muons in the sky is uniform (also called the flat Earth hypothesis);

2. SK is small and far away from the atmosphere so all muons coming from the same point in the sky can be assumed to have parallel trajectories.

We can hence map a given point in the sky to a single direction (θ, ϕ). Then, as shown in Figure 4.8 for fixed (θ, ϕ) the intersections of the muon trajectories with a plane P perpendicular to the (θ,ϕ) direction will be uniformly distributed. In what follows, we consider a plane P that touches the detector in only one point, at the top. This means that the plan

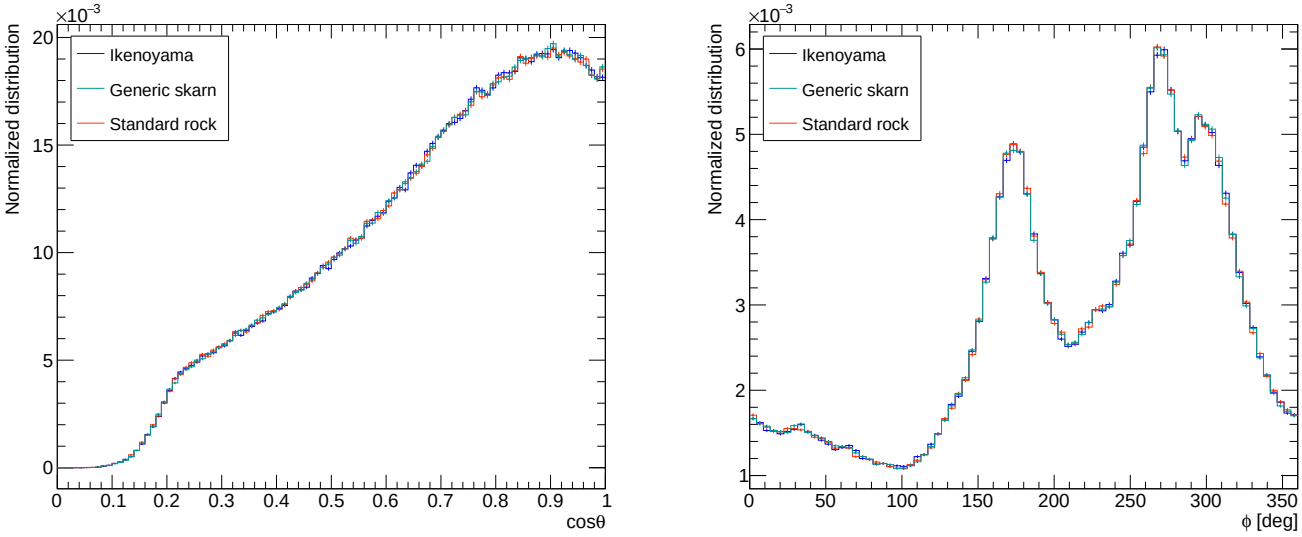


FIGURE 4.5: *cosθ* (left) and ϕ (right) angle distributions for different types of rock: Ikenoyama rock (blue), generic skarn (green) and standard rock (red). Here the density value is 2.70 g/cm^3 for all the three rocks. The results are in very good agreement, proving that the choice of rock composition among the three types considered is not critical.

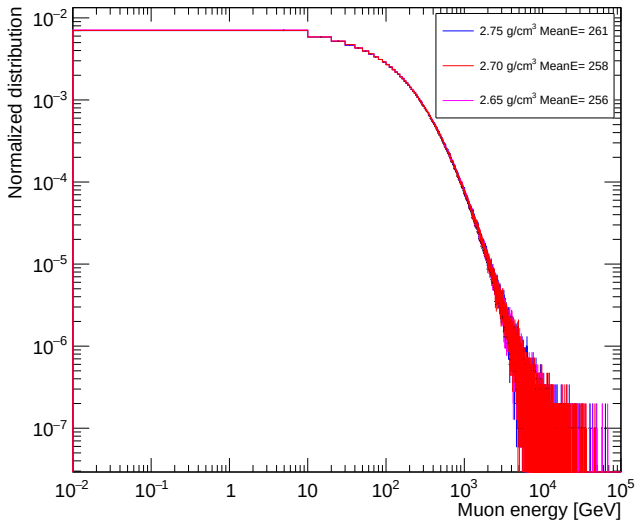


FIGURE 4.6: Energy distributions for different values of density: 2.75 g/cm^3 (blue), 2.70 g/cm^3 (red) and 2.65 g/cm^3 (magenta). In this plots only Ikenoyama type of rock is considered. Similar results are obtained for generic skarn and standard rock.

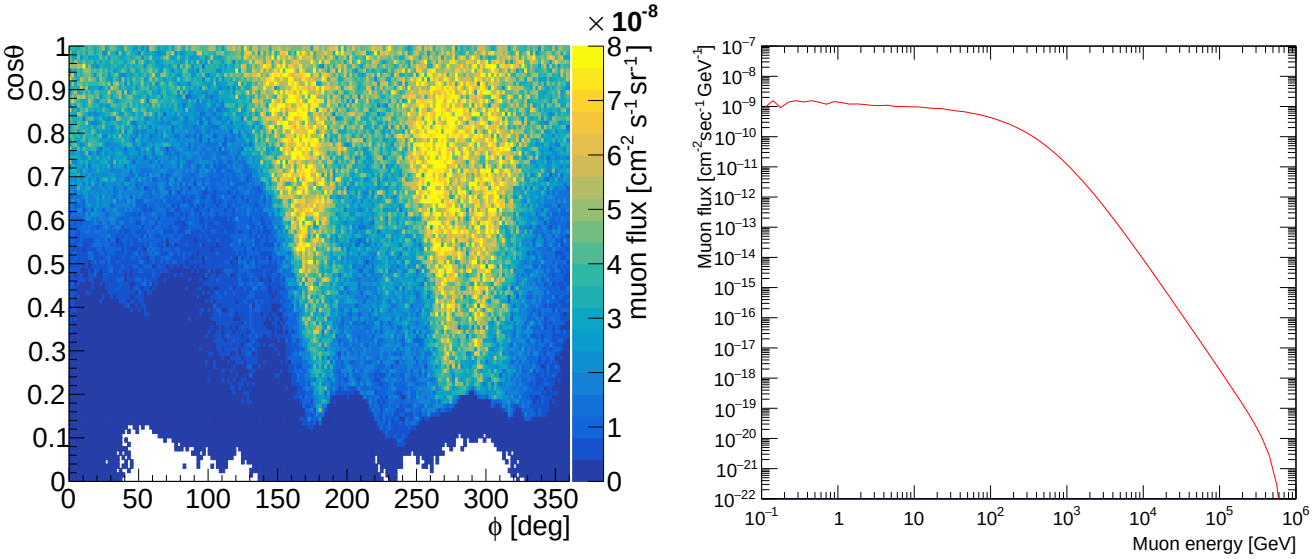


FIGURE 4.7: Left: 2D contour plot of $cos\theta$ (y axis) and ϕ (x axis) angle distributions. Right: muon flux distributions reaching SK detector.

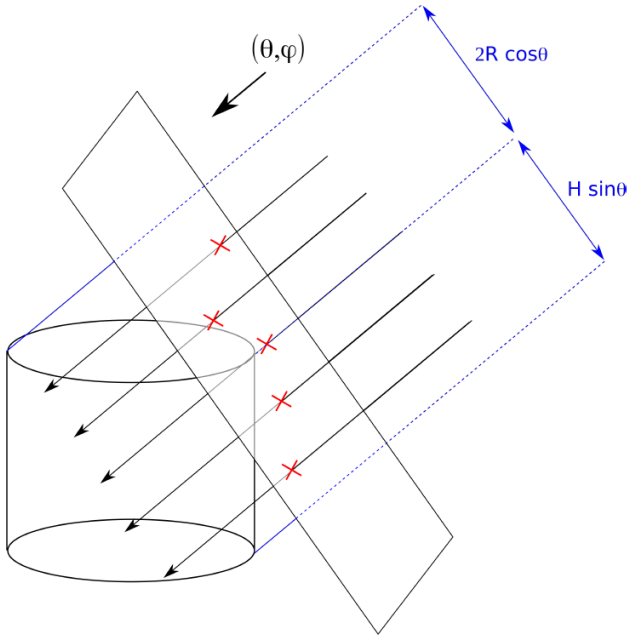


FIGURE 4.8: Muons coming from a single point in the sky. Their directions are all approximately equal to (θ, ϕ). The total height of the detector is H and its radius is R.

has to pass through $\vec{X}$= (R cosϕ, R sinϕ, H/2) (the coordinate system has its zero at the center of the detector). In practice, as shown in the Figure, we can simulate all muons passing through the detector by sampling points uniformly in P, inside a rectangle with the following dimensions:

$$L = 2R + H \tag{4.3}$$

$$l = 2R \tag{4.4}$$

Sampling points uniformly inside a rectangle is simple, however the rectangle in consideration is not centered and it is tilted. So we start sampling points uniformly in a horizontal rectangular surface sitting on top of the detector and positioned as shown in the top panel of Figure 4.9. Then we rotate them around the z axis by ϕ, as shown in the middle panel of the same Figure, and rotate them again around the new y axis by θ as shown in the bottom panel. Lastly, we translate the plane so that it touches the detector in (R cosϕ, R sinϕ, H/2). For a point (x, y, z) in the SK coordinate system, we therefore have the following transformation:

$$(x', y', z') = R \begin{pmatrix} cos\phi \\ sin\phi \\ 0 \end{pmatrix} + \begin{pmatrix} xcos\theta cos\phi + zsin\theta cos\phi - ysin\phi \\ ycos\phi + xcos\theta sin\phi + zsin\theta sin\phi \\ zcos\theta - xsin\theta \end{pmatrix} \tag{4.5}$$

Finally, for each (θ, ϕ) from MUSIC we compute an entry point with the method above described. Since we are sampling in a rectangle some points will still not go through the detector: these entry points, with the associated direction, will be rejected, straightforwardly accounting for geometrical effects. Figure 4.10 shows how the rejection sampling impacts the cosθ and ϕ distributions generated by MUSIC adapting them to represent the flux in the cylinder of Super-Kamiokande detector. While for ϕ the effect is negligible, the sampling modifies the shape of $cos\theta$ distribution, especially for vertical muons. This procedure allows us to convert the direction generated by MUSIC into a sample of entry points distributed on the surface of the inner detector.

4.4 FLUKA simulation

FLUKA [99, 112] is a general purpose Monte Carlo code for the description of interactions and transport of hadrons, ions, and electromagnetic particles from few keV to cosmic ray energies in matter. It is built and frequently upgraded with the aim of maintaining implementations and improvements of sound and modern physical models. FLUKA version 2011.2x.7 is used for this work, together with FLAIR (version 2.3-0), an advanced user interface to facilitate the editing of FLUKA input files, execution of the code and visualization of the output files [113]. For our purpose, FLUKA propagates muons into the Super-Kamiokande detector, simulating all the relevant physics processes that lead to energy losses and creation of secondary particles: ionization and bremsstrahlung, gamma-ray pair production, Compton scattering and muon photonuclear interactions. Hadronic processes such as pion production and interactions, low energy neutron interactions with nuclei and photo-disintegration are also modeled.

4.4.1 Nuclear models and uncertainties

Modern and well tested microscopic models are, for most of the physical problems, the supporting base of FLUKA. The microscopic approach ensures self consistent and solid physical foundation to hadronic and nuclear interactions, enhanced by the fact that results

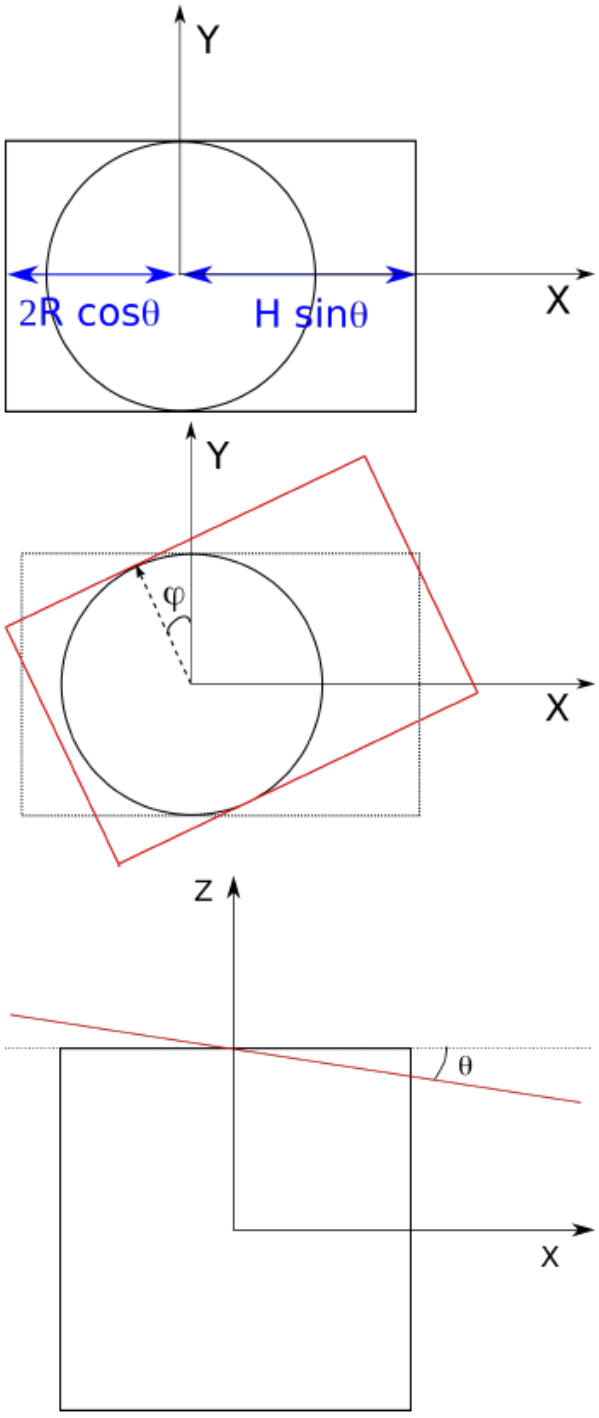


FIGURE 4.9: Sampling method for a rotated rectangular plane. Top: initial sampling in a horizontal plane. Middle: initial rotation by ϕ. Bottom: second rotation by θ. Then we need to translate the plane to make it touch the detector in the right place.

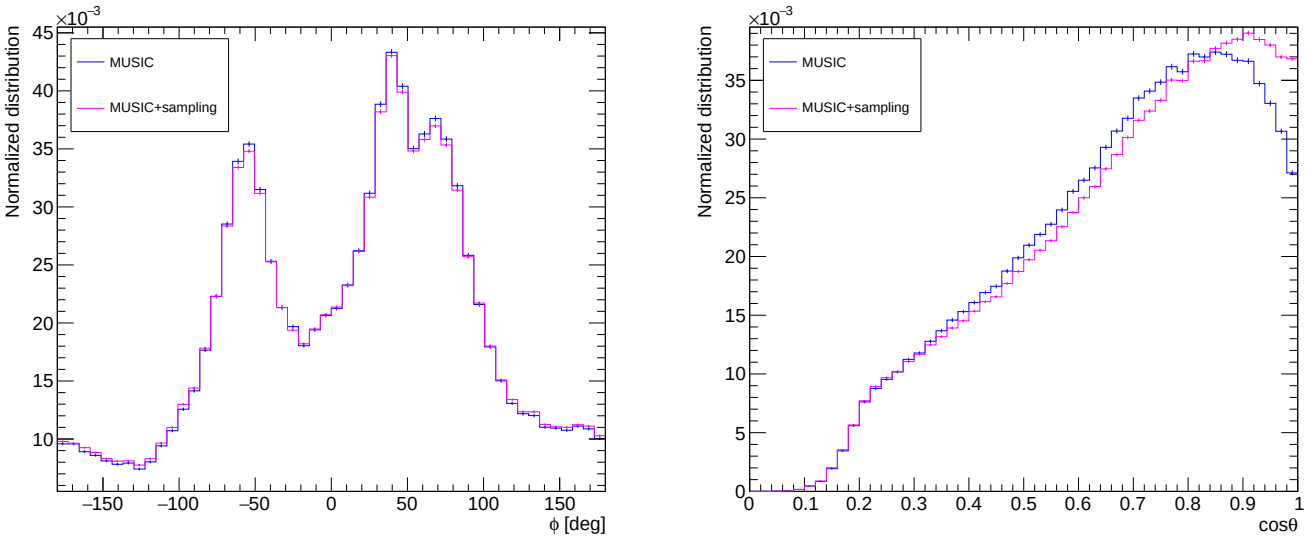


FIGURE 4.10: MUSIC angles distributions before and after applying the sampling method. Right: ϕ. Left: $cos\theta$.

are checked against experimental data at single interaction level [112, 114]. FLUKA incorporates a specific model devoted to nuclear and hadronic interactions called PEANUT [114–116]. Note that the user has no direct access to the model and cannot modify its parameters. Within PEANUT target nuclei are modeled with a local Fermi gas representation [117], while hadron-nucleus interactions are described with a Generalized IntraNuclear Cascade (GINC) approach: a cascade of two-body interactions, where projectile and reaction products are involved. All cross sections implemented in these models are extracted from tabulated data or parametrized fits. GINC model smoothly moves to pre equilibrium stage: the transition occurs until when nucleons are below pion production threshold, 30-100 MeV, and non-nucleons, more frequently pions, are absorbed or decayed. At the end of this reaction chain, there is thermal equilibrium between the nucleus and surrounding environment. The compound nucleus (Z, N) is characterized by its momentum and excitation energy and can evaporate nucleons, fragments or γ rays, or even fission to dissipate the residual energy. Gammas are emitted to release the excitation, when remaining energy is below particle emission threshold. Evaporation, fission and break-up models are the last step of hadron-nucleus interactions and determine the nature of residuals, whose spectrum is strictly related to collisions at the first cascade stages. A series of benchmark results for the physical model included in FLUKA code, which are relevant to the problem of muon-induced backgrounds are presented in Ref. [118].

Uncertainties in hadronic processes are a limiting factor for a precise prediction of both shower particles development and isotopes yields. Several studies on spallation production show different results depending on the isotope type and the order of magnitude of its yield with respect to the others. FLUKA predicted yields for Super-Kamiokande from Ref. [53] agrees with published measurements [52] within few tens of percent, for some isotopes, and within a factor $\approx$ 2-3 for others. Analogously, Borexino [101] find a good agreement between the measured yields and FLUKA predictions: a factor $\approx$ 2-4, which depends on the isotope type. The same data are compared with a GEANT4 simulation where the agreement is within a factor $\approx$ 2-10. An overall factor of 2 precision is considered adequate to estimate uncertainties due to hadronic models [53, 117]. This error is evaluated comparing FLUKA results with data, not varying the model parameters.

4.4.2 Simulation setup

In the following, the main code components we customized to build the FLUKA simulation are described in detail.

FLUKA settings

FLUKA main code fully integrates the most relevant physics models and libraries: several default settings are available and must be chosen at the beginning of the simulation depending on the general physics problem the user is dealing with. In addition to this, FLUKA offers several options to customize the default settings enabling or disabling a certain type of processes or changing the treatment of specific type of interactions.

For this work, FLUKA simulation was built with the default setting PRECISIO(n). All the specifics related to this setting can be found in Reference [99] but we want to enlighten a few that are particularly important for the scope of this study:

- Low-energy neutrons, which are defined to have less than 20 MeV energy, are transported down to thermal energies.
- The absorption is fully analogue for low energy neutrons: in a fully analogue run, each interaction is simulated by sampling each exclusive reaction channel with its actual physical probability, this allows for event-by-event analysis. In general, this is not always the case, especially concerning low-energy neutron interactions.
- Muon photonuclear interactions are activated with explicit generation of secondaries.

Several options are used to complement the default setting, in particular:

- PHOTONUC option: photon and electron interactions with nuclei are activated at all energies.
- MUPHOTON option: controls the full simulation of muon nuclear interactions at all energies and the production of secondary hadrons.
- EVAPORAT(ion) and COALESCE(nce) options: these two are set to give a more detailed treatment of nuclear de-excitations. Despite the related large CPU penalty, it is fundamental to activate these options when isotope production has to be studied. EVAPORAT enables the production of heavy nuclear fragments (A>2) while COALESCE sets the emission of energetic light-fragments.
- IONSPLIT option: used for activating ion splitting into nucleons.
- IONTRANS option: full transport of all light and heavy ions and activation of nuclear interactions.
- RADDECAY option: activate radioactive decay calculations.

Geometry

The very first input for a FLUKA simulation is the detector geometry. The Super-Kamiokande detector is modeled as a cylindric region of radius 19.65 m and height 41.40 m. The inner detector and the fiducial volume are two concentric cylinders with radius 16.90 m and 14.90 m and height 36.20 and 32.20, respectively [119]. The inner detector is, in reality, separated from the outer detector by a layer of photomultiplier tubes, mostly inward-facing: this PMT structure is not simulated in FLUKA, its effect will be treated at the stage

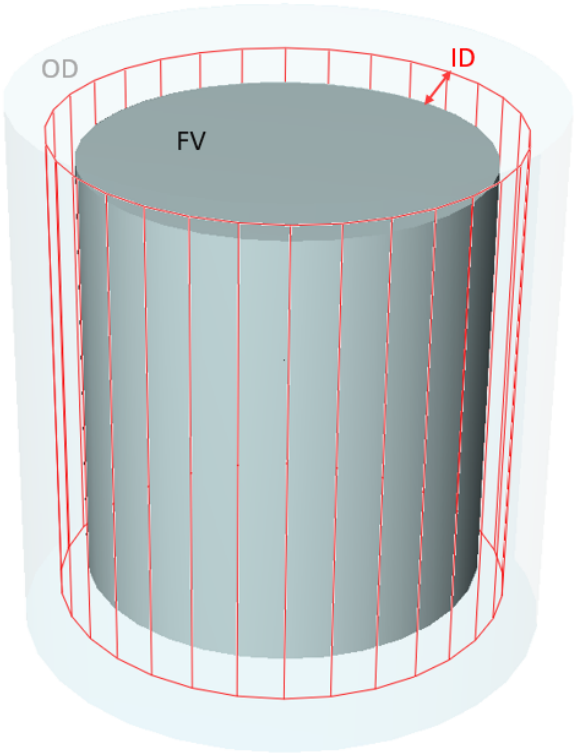


FIGURE 4.11: Illustration of the detector geometry adopted in FLUKA. The three regions are filled in water. Here the colors are chosen to give a graphical intuition of how the cylinders are separated.

of incorporating the FLUKA results in the Super-Kamiokande detector simulation. SK is completely filled with pure water. In Ref. [53] a 2 m thickness of rock was also simulated since the muons can induce showers inside this region and secondary products may reach the active part of the detector. In the same Reference it was proven that this has a minor effect on the results. We decided to omit the 2 m of rock, mainly to save computational time. A schematic illustration of the geometry implemented in the FLUKA simulation can be found in Figure 4.11.

User routines

Due to the complexity of our problem, customized methods, introduced by user-written functions called “user routines", are required in order to incorporate non-standard primary particle distributions, the ones calculated with MUSIC simulation, and to extract event-by-event informations for the shower reconstruction. For our purpose two main routines have been implemented:

- *Source routine*: as its name may suggest, this function is used to sample primary particle properties from distributions. In our particular case, directions and energies of the muons are sampled from the MUSIC outputs. In the same code we have to specify the points in the FLUKA geometrical space where muons are generated. The specific rejection sampling method to choose the muon generation point in order to take into account the flux modifications induced when passing to a cylindrical shape is described in the previous Section (“Sampling of muon entry points”).

- *Scoring routine*: cosmic muon passing through water in SK (called “primaries") lose energy producing a huge amount of secondary particles. Most of them are electrons and photons generated from electromagnetic processes all along the muon path. Since we are only interested in rather rare and energetic interactions leading to spallation production, we select only primaries that lead to the development of showers containing hadrons or generating isotopes, while we neglect muons behaving only as minimum ionizing particles or producing purely electromagnetic showers. This choice is beneficial both to reduce computing time and to ease the output reading and treatment. For all the selected primaries, the positions where they enter and exit the inner detector as well as their energies are written on a file. A similar selection is required when treating particles within the muon-induced shower (also generically referred as “secondaries"): as a standard rule, we neglect the electromagnetic component of the cascades. In fact, our strategy consists in tracking only the specific interactions interesting for spallation studies: inelastic processes of parent muons or daughter particles with nuclei and decays of shower particles or isotopes. Moreover, given that neutrons are not visible in Super-Kamiokande, we only record the 2.2 MeV gammas emitted after a neutron capture on hydrogen. However, it is not impossible that a shower containing an isotope is initiated by a photon, electron or positron. In this specific case we keep track of the electromagnetic interactions triggering the shower. In general, if the primary daughter of the cosmic muon develops a particle cascade we define it “hadronic shower" if the daughter is an hadron and “electromagnetic shower" if it is a photon, electron or positron. Notably, an electromagnetic initiated shower only develops if the parent particle has energy greater than 0.1 GeV [98]: below this value electrons and positrons do not create or accelerate secondaries and gammas could only generate electrons. For this reason, all the electromagnetic particles below 0.1 GeV are systematically ignored.

4.5 SKDetSim - FLUKA interface

With the aim of comparing simulation results to data, it is necessary to carefully model SK detector response. Therefore Super-Kamiokande official detector simulation, SKDetSim, described in Section 2.5, is used to quantify the Cherenkov light generated by the events we are interested to study and give the number of hits that would be recorded together with their timing and pattern. In fact, SKDetSim contains all the details on the geometry, electronics and specific physics processes of the detector, it is extensively tuned on SK data and can very well model Cherenkov photon generation and propagation. In addition, simulated signals are subjected to the same reconstruction algorithm as the real data in order to extract their energy (from number of observed hits), time, position and direction (from hit pattern) while the same reduction techniques are applied for selections (all the details are illustrated in Section 2.6). As already mentioned, particles simulated in SKDetSim are grouped in the same trigger windows that are applied to data.

Since SKDetSim is based on a different simulation package from what we used to generate spallation processes, namely GEANT 3, an essential evaluation of how to propagate FLUKA results in SKDetSim is needed. It is important to avoid that SKDetSim does a parallel generation of muon-induced showers, which would lead to a double counting of isotopes and/or excessive Cherenkov light arising from particles in the cascades. To exclude this situation, different types of interactions have to be deactivated and particular care needs to be taken with the injection of primary muons. In the following it is described how we proceed for the FLUKA-SKDetSim interface through the modification of the SKDetSim input card in order to meet these requirements. The underlying idea is

that FLUKA is used to model shower development in water while SKDetSim takes care of light generation and the detector response. Given that SKDetSim replicates detector conditions, it requires the signals to be stored in trigger windows analogous to the ones where data are stored. Simulated particles will be therefore grouped in fixed and well defined MC events. The following points describe the interfacing between FLUKA and SKDetSim for a typical event, which consists in a primary muon, a shower, potential neutron captures and isotopes decays.

- **Muons**: being two independent Monte Carlo codes, it is not possible to force primary muons in SKDetSim to follow the same exact path, or to generate analogous showers, they do when propagating inside FLUKA. Our strategy consists in generating primary muons in SKDetSim with the same energy, entry point and directions they have in FLUKA. However, an essential change in the input card is the deactivation of muon-nucleus interactions and photofission: the muon does not create hadronic showers in SKDetSim because shower products are already simulated in FLUKA. This means that the muon will in average behave only as a ionizing particle inside the Super-Kamiokande Monte Carlo simulation. Radiative losses through inelastic interactions with atomic nuclei are deactivated but electromagnetic processes such as bremsstrahlung and pair production are still possible and their importance increase with muon energy. This choice is necessary because, as we discussed, the products of purely electromagnetic interactions, if not linked to isotope productions, are not recorded in FLUKA outputs: while they are important to reconstruct the muon track via light generation, they are not relevant to study spallation processes. In principle, muons can have deviations from their initial track after scattering. However, these interactions do not appreciably affect the energy of the muon, nor their trajectory, especially if scattering off nuclei is turned off: the energy loss is small compared to the primary muon energy and the tracks have only minor deflections. Concerning stopping muons, we expect that if a muon enters the detector without enough energy to cross it entirely, it will stop both in FLUKA and SKDetSim. As we will see in the next Chapter, stopping muons that produce spallation products along their track, and not at the end of their path after being captured, are only a minor quantity (less than 0.4%).
  Conclusively, most of the primary muons are treated as minimum ionizing particles with the possibility of emitting electromagnetic radiation but not to interact inelastically with nuclei, each of them is generated in a single MC event, with a fixed trigger window of 40 μs, at t = 0.

- **Shower particles**: unstable isotopes are mainly produced in muon-induced cascades, making their shape and composition a key information for spallation studies. In our simulation structure FLUKA is the software devoted to model shower development. Therefore, we want to keep track of the shower particles generated in FLUKA, avoiding any sort of double counting when their light emission is simulated by SKDetSim. The strategy is to turn off in SKDetSim both the creation of secondary particles and all the photofission processes in hadronic interactions: only the particles already produced in FLUKA will propagate in the detector, without creating new ones. On the contrary, we do not put any constraint on electromagnetic processes: SKDetSim takes care of all the electromagnetic interactions, including the generation of Cherenkov light for charged particles inside the shower. In practice, in SKDetSim muons are free to interact only electromagnetically and produce photons, electrons and positrons. These can in turn only interact electromagnetically producing other γ, e^+, e^-. All other shower particles (gammas, pions, kaons...)

arising from inelastic reactions are taken from FLUKA and injected, with their vertex position and energy, in SKDetSim, where they propagate producing Cherenkov radiation. Their possibility to interact producing secondaries is however discarded. Neutrons are treated separately. The only exception occurs if a radioactive isotope is produced in a shower initiated by EM processes. This can happen for example if an energetic gamma interacts with a nucleus inducing its fission. In this case the photon, electron or positron initiating the shower is also injected into SKDetSim, provided its energy is larger than 0.1 GeV, since no isotope production has been observed below this threshold in previous simulation studies [98]. As we will describe in Chapter 5 this scenario is not dominant and the impact of these extra particles on the muon light pattern is limited.
Finally, in the same MC event where we generated the primary muon we also inject FLUKA shower particles at time t=c/d (typically few ns), where d is the distance of the particles from the muon track entry point. In fact, the light emission from pions and, indirectly, gamma rays is almost coincident in time and space with muon ionization energy losses, making the two indistinguishable and justifying their presence in the same MC event.

- **Neutrons**: being neutral particles, neutrons are invisible in Super-Kamiokande. Therefore, instead of neutron itself, we record the 2.2 MeV gamma that is emitted after propagation, thermalization and capture on hydrogen. The capture time is about 200 μs, namely larger than SKDetSim trigger windows of 40 μs, during data taking neutron signals are searched by specially designed trigger: WIT described in Sec. 2.4.2. Unlike light emitted by charged particles within the shower, which tend to be very close in time and space with the photons created along the muon track, neutron captures produce a delayed signal that can be separated from track light emission. This makes neutrons independently detectable. For this reason each gamma from neutron capture is simulated at t=0 in an independent MC event where the actual capture time is also stored. After a SKDetSim event containing muon and shower we define as many MC events as the number of neutrons generated and captured. This structure is useful in our study to treat neutron captures with algorithms that mimic the format of WIT online trigger applied to data: in Chapter 6 it will be described how SKDetSim events containing neutrons are manipulated to recreate data conditions.

- **Isotope decays**: similar to the neutron case, we identify the isotopes through their decay products. Since the half lives of background isotopes span from few microseconds to tens of seconds, the decay time is saved in dedicated variables while the beta and, eventually, gamma emitted from a single isotope decay are simulated in a SKDetSim event at t=0.

- **Noise**: each of the steps described above requires to model not only the signal, but also the noise in the detector. For muons, shower particles, and isotope decays, we use the modeling of the PMT dark noise from SKDetSim, based on regular measurements made over the SK-IV period. The treatment of neutron capture signals is more complex: this signal in water is particularly weak and the predicted performance of the neutron tagging algorithm is highly sensitive to dark noise modeling. All details on neutron treatments are conducted independently and are described in Chapter 6.

A scheme of the way the different particles from FLUKA simulation are injected in SKDetSim events is shown in Figure 4.12. Table 4.1 summarizes the processes treated in FLUKA

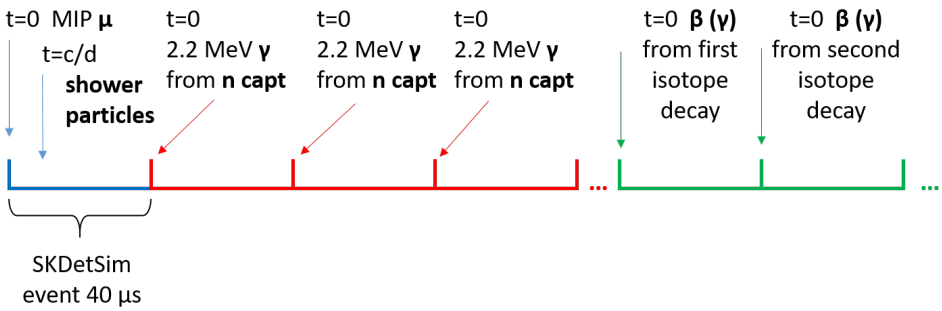


FIGURE 4.12: Scheme of the order of events simulated in SKDetSim within a 40 μs trigger window: muons and showers are contained in the same MC event, followed by neutrons and isotope decays.

SKDetSim	
Deactivated	Activated
Photofission Secondary particle generation in inelastic interactions	EM interactions Cherenkov radiation Decays
FLUKA	
Generated	Ignored
EM show. w/ spall. (if $E > 0.1$ GeV) γ, π, K from inelastic interaction γ from n captures Isotope decay products	EM show. w/o spall.

TABLE 4.1: The Table summarizes the main processes activated and deactivated in SKDetSim and the particles generated or ignored in FLUKA and injected in SKDetSim. More description in the text.

and SKDetSim respectively.

It is important to notice that in this work we will focus on observables that are particularly robust against possible errors in the modeling of shower Cherenkov light pattern due to simulation construction: the yields of spallation-produced isotopes, and the characteristic neutron clusters within the showers. Isotope decays and neutron captures can indeed be simulated in isolation from their parent muons in SKDetSim, and will hence not lead to unwanted interactions. Muon-induced showers still need to be simulated, as they affect the reconstruction of the muon track and therefore neutron and isotope identification (as it will be described in Chapter 6).

Event reconstruction

SK reconstruction softwares are described in Section 2.6. Muboy is used for muon tracks, isotope decay products are identified by BONSAI , and WIT trigger looks for the very low energy gammas from neutron capture.

4.6 Validation

Once the simulation components, their interface and structure are defined, appropriate tests are required to validate both the single simulation segments independently, their concatenation and the well functioning of the whole system. To begin, angle and energy distributions of the muon flux calculated with MUSIC have to be compared with data to confirm that the input spectra used in FLUKA are well describing the actual detector conditions. The impact of SKDetSim light propagation and reconstruction algorithms for simulated muons are tested afterwards. Subsequently, an independent validation of FLUKA outcomes can be achieved with the reproduction of results published in the fundamental work by Shirley W. Li and John F. Beacom [53, 98, 102]: the authors presented one of the very first simulation study of spallation processes in water Cherenkov detector based on FLUKA Monte Carlo code. Their work is the unique reference for spallation simulation in water, making the comparison with their results a good verification of the accuracy of our FLUKA calculations. Finally, results of the global simulation, including all the packages described above, are compared with SK data.

4.6.1 Comparison of MUSIC muon flux with data

As described in 4.3.1, the propagation of cosmic-ray muons from the surface to Super-Kamiokande detector depth is performed by MUSIC Monte Carlo code. In the simulation construction we assumed that the detector dimensions were negligible compared to the distance traversed by the muon inside the mountain. This would give us angle and energy distributions in a point at the center of the detector. To adapt the flux to the cylindrical shape of SK we applied a sampling method for selecting muon entry points in FLUKA, as explained in Section 4.3.1. With this in mind, we can assume that the point-like detector flux from MUSIC (before sampling) is equivalent to what we would find if we take muons passing through a small sphere centered in that same point: unlike the case of the cylinder, a sphere, being symmetric in θ and ϕ, does not distort the intersection cross-section of the muon with the detector and thus does not modify the muon angle distributions. Following this argument, we extracted $cos\theta$ and ϕ distributions from data for a sample of muons intersecting a spherical volume of radius 10 m located at the center of SK and we used the result for a comparison with MUSIC distributions. This comparison is independent of the entry points sampling method we choose. Figure 4.13 shows the results: globally ϕ distribution has a similar shape for data and simulation, there are specific regions with larger

differences, especially for ϕ in (-70,-30) degrees range, that can be explained by uncertainties on the mountain topology map adopted for MUSIC simulation, in Figure 4.3.1. The elevation data were taken in 1997 with a 50 m mesh, the accuracy of the measurement could not be enough to describe local non-uniformities in the geographical features of the ground or variations in the rock densities, assumed constant in MUSIC. Nonetheless, uncertainties in ϕ distribution, here less then 10%, are not expected to have a significant impact on the results of the simulation: the probability of developing a shower is more sensible to changes in muon path length rather than their direction in ϕ plane. The plot on the right in Figure 4.13 represents $\cos\theta$ distribution. The agreement of data and simulation is good. For nearly-vertical muons ($\cos\theta > 0.9$) there is less than 3% difference between data and MUSIC. Analogously, the reason can be possibly found in inaccuracies of the elevation map. Unlike the ϕ angle case, the main effect of these discrepancies could be observed on the muon track length distribution, and as a consequence on shower and isotope yields, since the longer the muon path the higher is the probability to develop a spallation-inducing shower. However, given that we expect high errors on isotope yields due to hadronic uncertainties of the model in FLUKA (order of 100% [53]), the discrepancies in $\cos\theta$ are not expected to have a significant impact on the simulation outcomes. Finally, we consider the results in Figure 4.13 a first proof of validity for MUSIC muon propagation models.

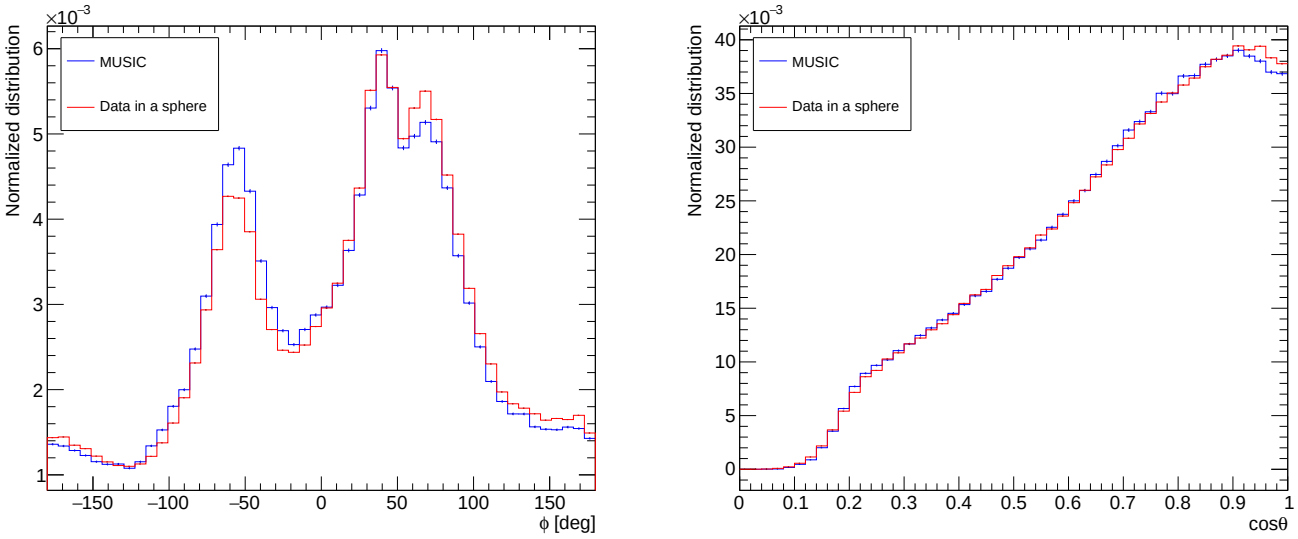


FIGURE 4.13: MUSIC angles distributions compared with data in a sphere of radius 10 m located at the center of the detector. Right: ϕ. Left: $cos\theta$.

4.6.2 Validation of entry points sampling method on reconstructed muons

Once MUSIC results are proven to be accurate, an important check is to evaluate the accuracy of the entry point sampling method described in Section 4.4: the procedure converts MUSIC angle distribution in a sample of entry points distributed on the ID surface, accounting for geometrical effects due to the cylindrical shape of the detector tank. In order to compare simulated muons with data, muon propagation and reconstruction is done by the Super-Kamiokande detector simulation and the fitting algorithms are used afterwards. As described in Section 4.4, SKDetSim is responsible of the whole muon propagation inside the detector, after receiving as input parameters muon entry energies, positions and directions. The only constraint is that muon-nucleus interactions are deactivated so

that muons will often behave only as ionizing particles. The trajectories of the muons simulated in SKDetSim are then reconstructed using Muboy. The fitting algorithm reconstructs muon directions and entry positions as well as the track length of the particle inside the detector inner volume. Charge deposited along the track is also computed.

At this stage, muon directions generated by MUSIC are treated as input for SKDetSim and, for each couple (θ, ϕ) an entry position is computed with the sampling method described in 4.3.1. The reconstructed variables from a sample of 10^6 simulated muons are compared with muons reconstructed from data. While in the data multiple muons can cross the detector at the same time, leading to simultaneous reconstruction of multiple tracks, only single through-going muons are simulated in this work. Thus, since Muboy is able to distinguish single muon from muon bundles, we only select single through-going muons in the data, for a final sample of more than 13×10^6 muons. In Figure 4.14 reconstructed entry points are displayed. The agreement between data and simulation is satisfying, note that the discrete nature of the plots represent the grid used by Muboy to fit the positions. Figure 4.15 shows the reconstructed directions, in terms of (X,Y,Z) unitary vectors, both for simulated and data muons. The distributions of the X,Y,Z components of the muon directions show a reasonably agreement between simulation and data. The discrepancies observed are consistent with the slight mismodeling of the phi distribution from Figure 4.13: these are a consequence of the uncertainties on the mountain elevation map we described in the previous Section. In this work, we consider particularly the impact of the modeling of muon trajectories on the reconstructed muon track length. In fact, the length of the path a muon undertakes inside the detector is directly related to its probability of generating showers: along a longer track the chance of developing at least one energetic muon-nucleus interaction is higher. Figure 4.16 presents the distribution of track lengths for data and simulation: the two peaks are linked to geometrical differences of muons entering the detector from top surface and the ones crossing the barrel. The latter dominantly contribute to the peak at about 3500 cm. The ratio between the higher peak and the shorter one is slightly lower in the simulation, this reflects the deficit of muon with $cos\theta > 0.9$ in Figure 4.13: muons crossing the top part of SK have an average $cos\theta$ value larger than muons crossing the sides. Moreover, it can be observed a small lack of very short tracks (less than 7 m) in the simulation. However, we do not expect this to affect significantly our studies, since muons with short path are very unlikely to create showers or isotopes that are contained in the fiducial volume. Finally, these differences lead to a mean muon track length of 2420 cm for simulation against 2385 cm of the data. As mentioned, we expect significant errors due to hadronic models in FLUKA, a less than 2% difference in track length is assumed to have minor contribution. Last but not least, a global good agreement between simulation and data is also found when comparing the total charge deposited by the muon along its path. The result is shown in Figure 4.17. An important information we have from this plots is that the condition we apply on SKDetSim when propagating cosmic muons, namely the deactivation of muon-nucleus interaction, does not significantly modify the charge deposition: most of the Cherenkov light generated is in fact produced by pure electromagnetic processes which are completely handled by SKDetSim. In practice, as will be seen later in this Chapter, muon reconstruction is not significantly affected by the existence and structure of muon-induced showers. Hence, the reconstruction of simulated muon trajectories will not be sensitive to possible shower mismodeling caused by interfacing FLUKA and SKDetSim.

From the results presented above we can conclude two important things: the muon entry points and trajectories, as well as detector effects, are accurately modeled. Therefore, we conclude that the modelization of muon flux used as input parameter for FLUKA calculation is sufficiently accurate for the scopes of the simulation.

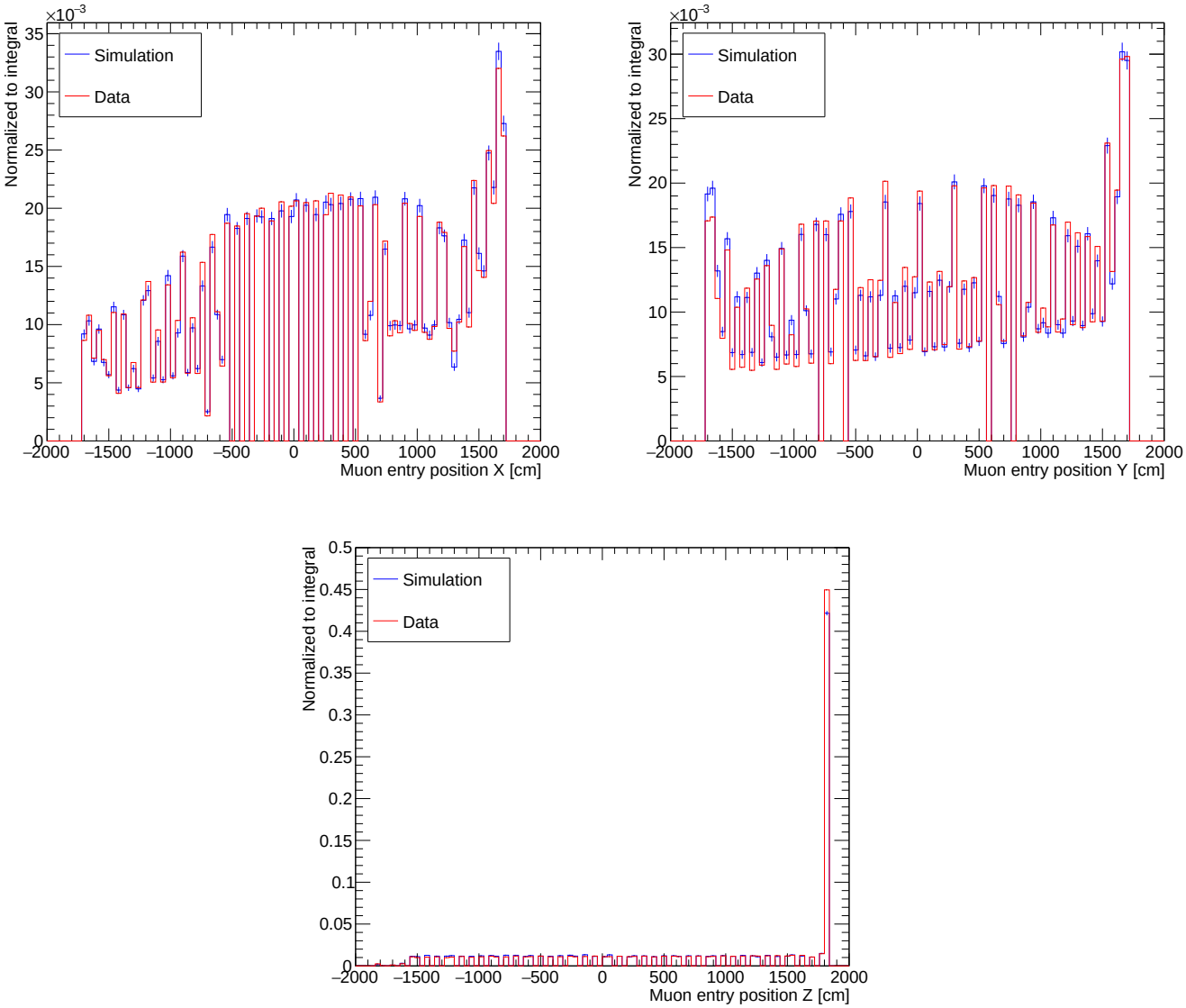


FIGURE 4.14: Reconstructed muon entry positions of data (red) and simulation (blue): X coordinate on the top left plot, Y coordinate on the top right and Z coordinate in the bottom plot. The discrete values reflects the mesh of the grid used in the reconstruction algorithm.

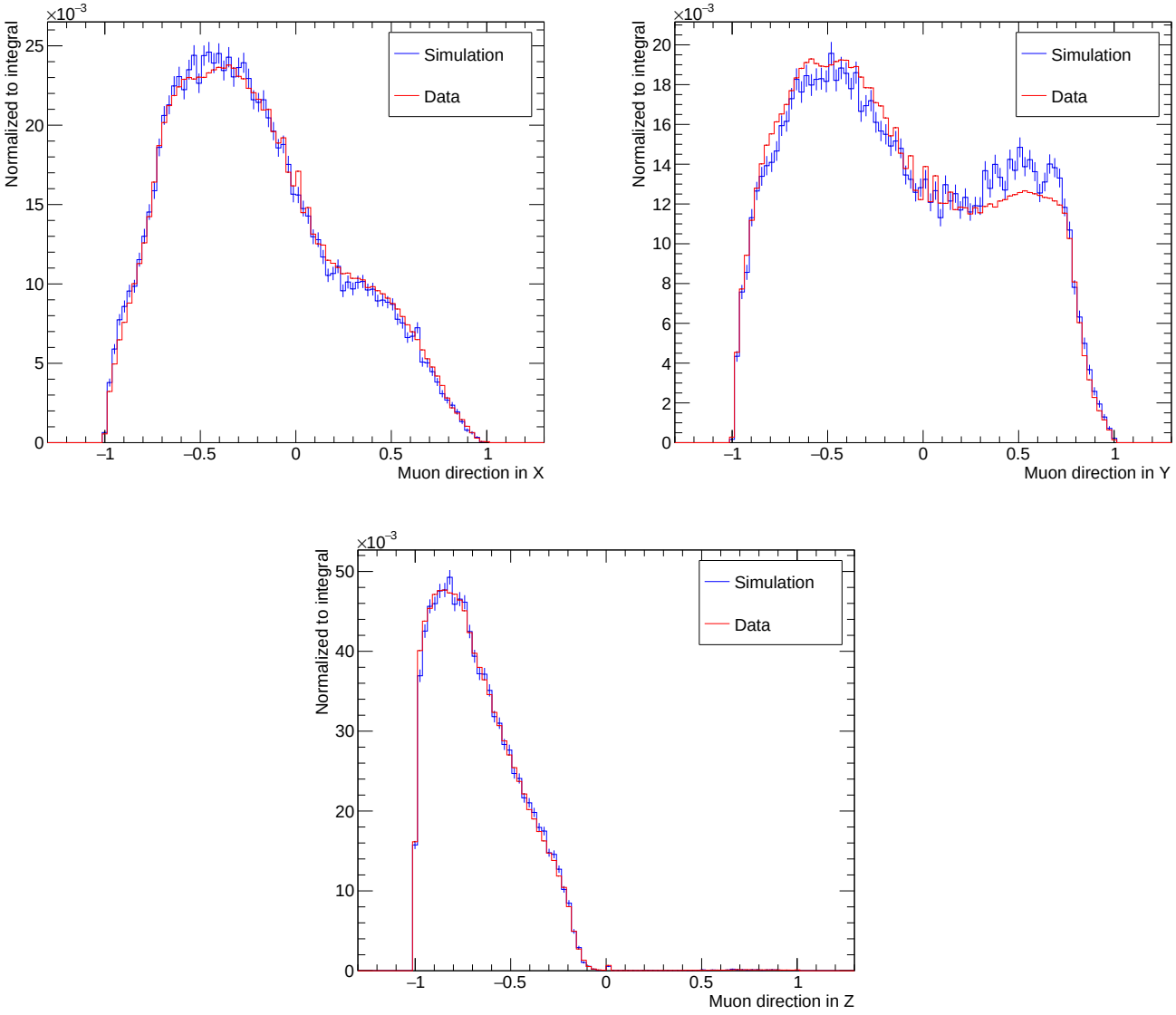


FIGURE 4.15: Reconstructed muon entry directions of data (red) and simulation (blue): X coordinate of the directional unitary vector on the top left plot, Y coordinate on the top right and Z coordinate in the bottom plot.

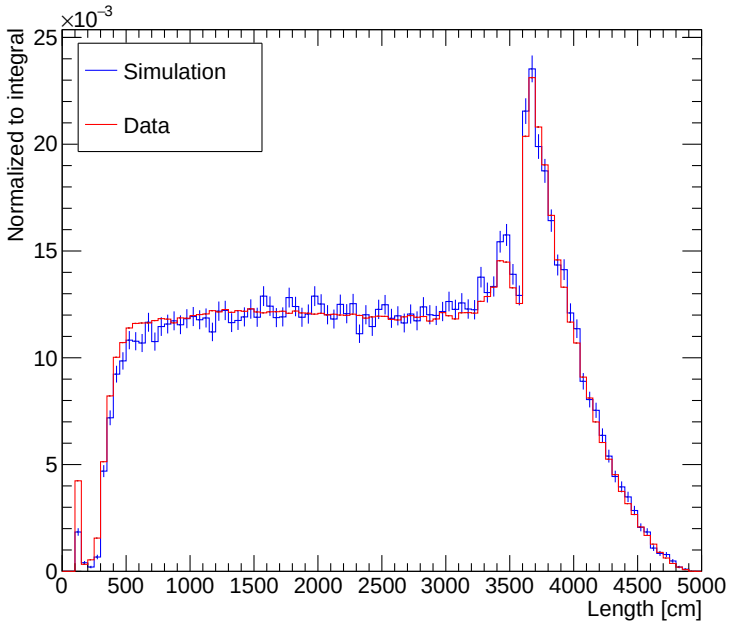


FIGURE 4.16: Reconstructed muon track length distribution for data (red) and simulation (blue).

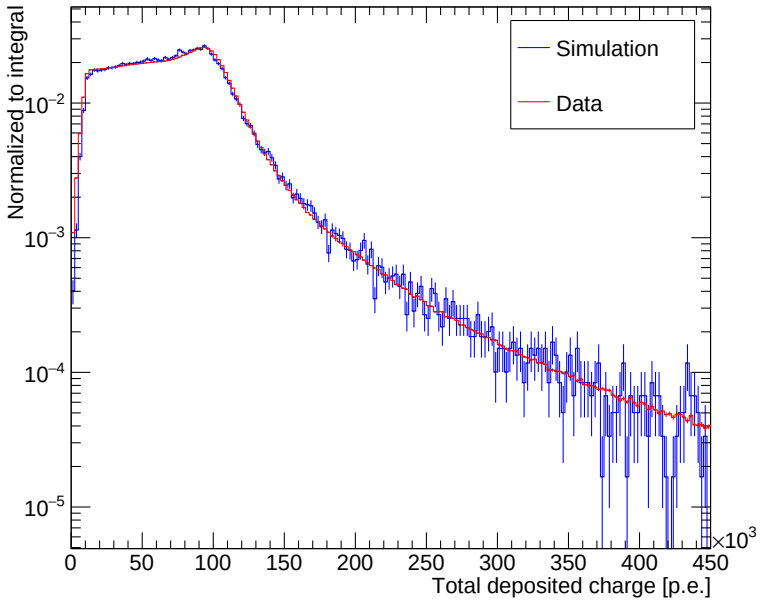


FIGURE 4.17: Reconstructed total charge deposited by muons crossing the ID for data (red) and simulation (blue): the charge value is expressed in number of photoelectrons collected by all the PMTs.

4.6.3 FLUKA validation

The only muon-induced simulation of spallation showers in water Cherenkov detectors is the one published by Li and Beacom in Ref. [53, 98, 102]. Their work is based on FLUKA and presents several results related to muon energy losses in water, shower particle production and isotope decays that are a unique source of comparison to validate the consistency of our FLUKA calculation. Although several simplifications were adopted in Li and Beacom's studies, they reproduced Super-Kamiokande measurement of the spallation isotope yields, presented in [52], within a factor of 2. Moreover, they show a good agreement with SK data for energy spectrum and time profile of spallation background events, proving the accuracy of their calculations. The results are dominated by uncertainties related to hadronic models that govern both shower and isotope production in FLUKA and by the fact that yields for different isotopes vary by orders of magnitude, making the rarest ones difficult to inspect. In this Section, it will be described the procedure we followed to reproduce Li and Beacom's work in order to validate FLUKA results independently from the other parts of our spallation simulation. Replicating their study will also lead us to present some first characteristics of muon interactions in water: the first insights on the properties of spallation events, accessible with simulation, will be presented along this Section before being thoroughly investigated in Chapter 5.

Simulation setup

FLUKA settings described in the previous Chapter, in Section 4.4, are used. Although we tried to reproduce their setup, details related to FLUKA input options as well as the associated user routines may partially differ from Li and Beacom's paper since the specifics of their calculations are not known, in particular customized scoring techniques. The same simplifications for primary muons adopted in [53] are used for consistency:

- All primary muons are vertical, down-going, passing through the center of Super-Kamiokande tank: we do not make use of any angular distribution for cosmic muons. They all intersect the detector with $\cos\theta$=1 and their entry point is the center of the top wall of the outer detector. Angular deflections of the trajectories may be possible due to scattering with water nuclei but, as also proven in [53], they are small compared to the total muon path and can be neglected. Hence, all tracks in the Super-Kamiokande fiducial volume are assumed to be of length 32.20 m.

- Only negative muons are simulated: we do not expect μ^+ and μ^- to develop showers in a different form nor to provide very different isotope yields. In fact, previous studies showed that isotope production differs of only a few percent between the two types of muons [100, 101]. The only process that depends on the muon charge is nuclear capture on oxygen, that occurs only for μ^-: this can lead to the production of isotopes, mostly ^{16}N through $^{16}\mathrm{O}(\mu^-, \nu_\mu)^{16}\mathrm{N}$. For consistency with Li and Beacom's work we neglect the positive muon contribution. Nevertheless, the impact of adding both muon types is studied and presented in the next Chapter.

- Muon energy spectrum is taken from Ref. [53, 108] and shown in Figure 4.18. Logarithmic bins are chosen to underline the importance of high-energy bins: muons with higher energies are more likely to produce showers, and therefore isotopes. The two lines are respectively at: about 6 GeV, namely the order of the energy lost by a minimum ionizing muon crossing Super-Kamiokande fiducial volume vertically, and 1 TeV, energy where radiative losses equal ionization ones [53]. The shape of the energy spectrum has some differences with respect to MUSIC calculations: mean energy is 271 GeV [108] (258 GeV was obtained from MUSIC). We assume that the

discrepancies are mostly due to the fact that two different digital maps of Ikenoyama mountain were adopted for the calculations. This reflects the importance of an accurate modeling of the surrounding topology which lead to visible differences in the resulting spectra. Given that the topological map of Ikenoyama Mountain in Reference [108] is extracted from an old study by M. Nakahata for muon background calculation, while for MUSIC simulation we used actual data from a geographical survey, already implemented by KamLand in [100], we assume the flux, used by Li and Beacom, of Figure 4.18 to be less accurate. Nevertheless, it is adopted here to have better comparison with their results.

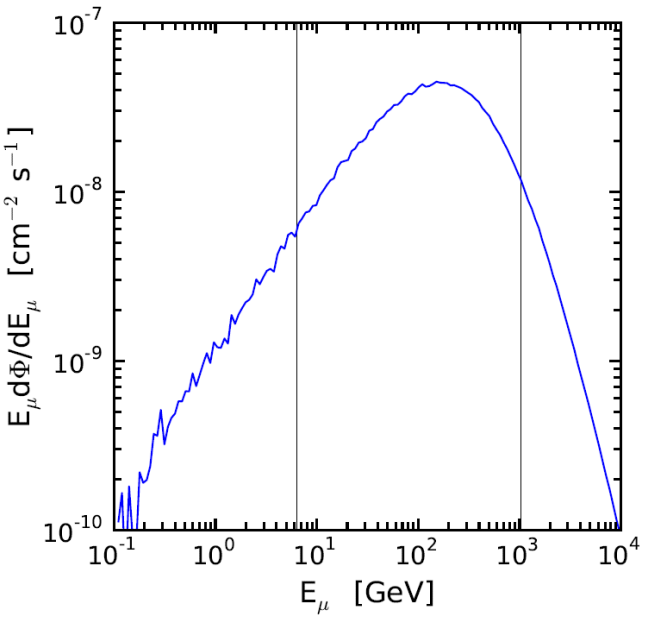


FIGURE 4.18: Energy flux of cosmic-ray muons at Super-Kamiokande, from [53]. The plot has logarithmic bins in the x axis and it is integrated over angles. Two lines separates three categories of muons: below 6 GeV the muons stop inside the detector since this value is the minimum ionization energy loss for a vertical muon passing through the fiducial volume. The central region, between 6 GeV and 1 TeV, contains muons that lose energy mostly through ionization, while the spectrum at higher energies is delimited by the line defining the muon critical energy: 1 TeV is where radiative energy losses start to dominate. The fluctuations due to low statistics (E $<$ 1 GeV) are negligible.

Li and Beacom's result are calculated per muon path length (32.2 m in the fiducial volume) in order to easily rescale them for the actual track length of muons in Super-Kamiokande. In this work, 3×10^7 primary muons are simulated, they are generated at the level of the top barrel of SK detector, namely on the surface of the OD, therefore traveling in water for about 42 m.

Muon energy loss and secondary production

As described in Section 1.7.1, muons lose energy both via continuous processes, such as ionization of atomic electrons, and radiative interactions with atomic nuclei: the last ones are characterized by large energy fluctuations and they can lead to generation of electromagnetic and hadronic showers while ionization does not have significant variations with energy. After the critical energy, 1 TeV for muons in water, radiative processes

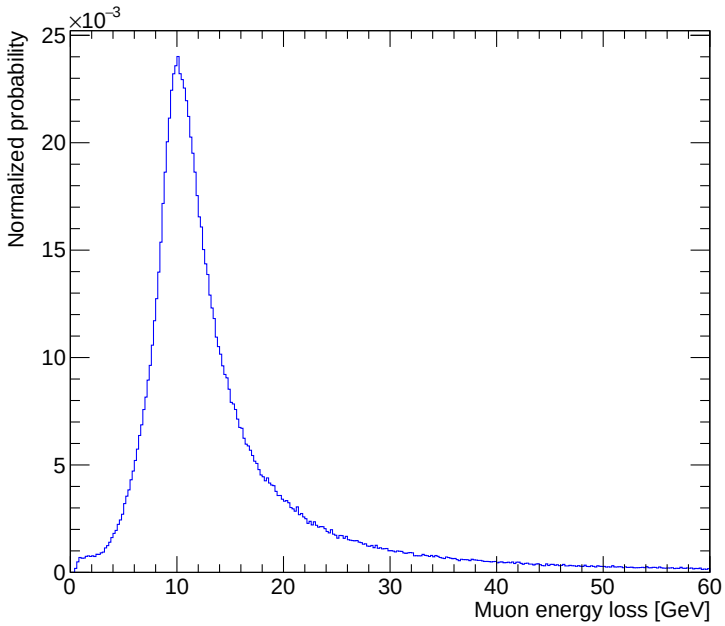


FIGURE 4.19: Probability density function of the energy loss for vertical through-going muons passing through the Super-K outer detector (path length 41.4 m).

become more important than ionization. Figure 4.19 shows the energy loss distribution computed for vertical down-going muons in the Super-Kamiokande outer detector, corresponding to a path length of 41.4 m. The ionization energy loss is about 8 GeV and has very little fluctuations, below this values muons stop inside the detector, about 13% of all the primaries. On average, simulated muons lose 14.4 GeV. The energy loss value is very often negligible compared to total muon energy but in some cases it can be high enough to lead to the creation of numerous secondary particles. The high energy tails in Figure 4.19 are dominated by very energetic bremsstrahlung processes which influence radiative emission at high energies. The distribution in Figure 4.19 well reproduces Li and Beacom's results: they observe an average energy deposition of about 11 GeV inside the FV. Given that our results are calculated for muon tracks crossing from top to bottom of the outer detector, they need to be rescale by the ratio of path length in OD/FV (about 1.29) to allow for a comparison. We obtain an average energy deposition of 11.2 GeV in fiducial volume only.

Tails in Figure 4.19 shows that often the muon energy deposition exceeds the value predicted for minimum ionization, which means there is a large amount of energy available for the creation of secondary particles through radiative processes. Following Li and Beacom's approach, the first method we used to look at these particles and their properties consists in plotting the sum of distances traveled by all the shower particles of a given specie at a certain energy: the relative heights are determined both by the particle multiplicity and the accumulated path length for each energy bin. The results are illustrated in Figure 4.20. As a matter of fact the integrated path length spectra is proportional to the total Cherenkov light emitted for particles above Cherenkov threshold [98]. As shown in Figure 4.20, the dominant secondaries in the path length spectrum are gammas, followed by electrons and positrons. This is expected since the primary ways for muons to lose energy are through electromagnetic processes: other than ionization, there are delta-ray production, pair production, and bremsstrahlung. The production of electrons and

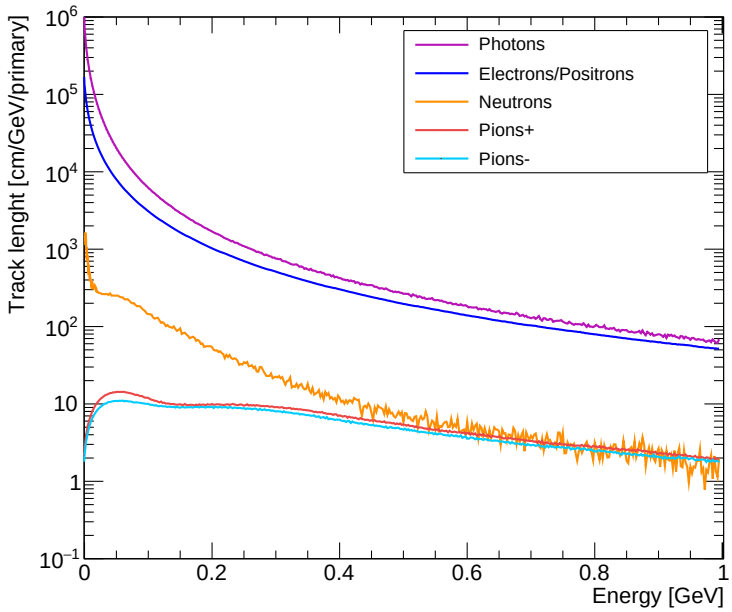


FIGURE 4.20: Secondary particle path length spectra made by cosmic-ray muons in Super-Kamiokande fiducial volume. The y axis is the cumulative path length, namely the sum of the distances traveled by all particles of a given species at each energy, and the x axis is kinetic energy value. The results are normalized per muon, so they are calculated for a single muon path length, here the 32.2 m vertical distance in the FV. This plot is constructed from a FLUKA standard scoring option.

gammas is strictly interconnected, the difference in their spectra in Figure 4.20 is due to the fact that gammas have larger mean free path in water than e^+/e^-. The path length distribution shows remarkable agreement with Li and Beacom's results in Reference [53]. The good agreement is a valuable confirmation of the choice of FLUKA settings: all the relevant processes through which muons lose energy in water are accounted for.

Among the particles generated by FLUKA when muons are propagating through water, some are more relevant if our goal is to study spallation processes. Despite electrons and positrons being the most abundant particles created by muons along their tracks, they are not primarily responsible for the creation of isotopes which are rather generated in hadronic or photonuclear reactions. In addition, the dominant mechanism for isotope production is not even the direct interaction of muons with oxygen nuclei but more often the development of muon-induced showers where secondary particles have enough energy to break up oxygen. Both hadronic and electromagnetic showers are produced and, due to their stochastic nature, they are not uniformly distributed along the muon track. Inside their bulk several types of particles can be created, the most important for our work are neutrons, pions and gammas: they are the primary parents of spallation background isotopes in Super-Kamiokande. In agreement with Ref. [53] we find that only 11% of the muons are directly producing background isotopes. Since the rest is made by shower products, it is important to study the properties of particles in the cascade to investigate spallation characteristics.

In the simple case of vertical muons it is straightforward to extract the distributions of the particle positions with respect to the muon track: information on the parent isotope is crucial to locate spallation events since, according to the type of parent and its mean free path, the isotope can be generated at shorter or larger average distances. In fact, an important variable in the SK low energy analysis used in likelihoods for spallation rejection is the transversee distance of the radioactive decay from the muon track, this corresponds to the distance where the isotope is generated, since, being heavy, it does not move from production position before decaying. The normalized distributions of absorption distance with respect to the muon track for the dominant particles in showers linked to spallation processes is shown in Figure 4.21. The results well reproduce the distributions shown in Figure 4 of Ref. [53], which in turn showed similar features with studies of Borexino, namely in scintillating materials [101]. This further confirms the correct implementation of FLUKA models in our calculation.

The same Figure is a starting point to understand shower mechanisms: given that different species of particles lose energy in different ways, the distance to the muon track varies for neutrons, pions and gammas. Neutrons travel the furthest: almost 95% of them are contained within 3 m of the muon. The mean free path in water is $\approx$ 10 cm for a neutron of few MeV [53] and this value decreases for lower energies, however several scattering steps are needed to thermalize and stop the neutron, making its total travel distance larger than other particles (up to few meters). At the end of their path most of the neutrons are absorbed by capture on hydrogen at non-relativistic energies with the subsequent emission of a 2.2 MeV gamma. The same reaction on oxygen has a negligible contribution. The mean free path for pions at MeV energies is 1 m. They have typically a forward direction and tend to disappear after only one interaction: this explains their transverse distance in Fig. 4.21. Similar to pions, also gammas are mostly forward and do not have large mean free path in water: most of them creates e^+/e^- pairs close to the muon track and their Moliere radius (9.8 cm in water [53]) gives a scale for gamma distances from the primary μ.

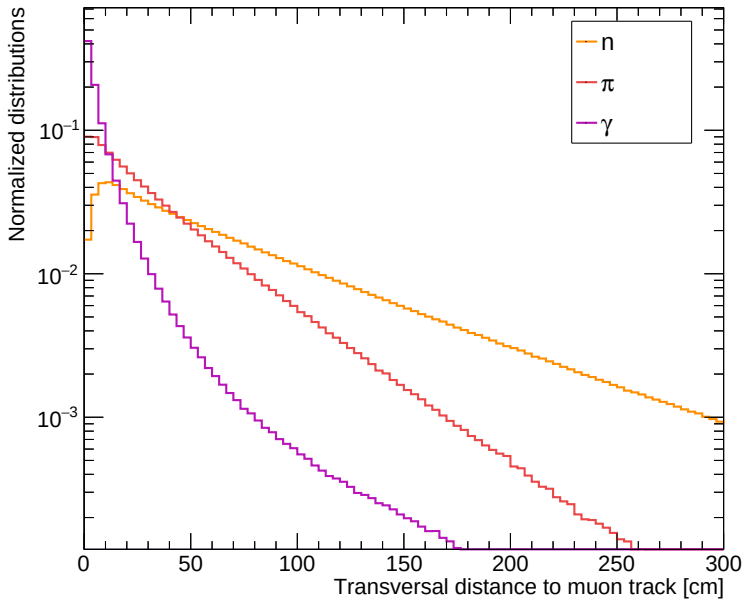


FIGURE 4.21: Absorption distance of secondary particles with respect to the muon track. Here only the transverse distance is considered. The distributions are normalized to their integral: the relative heights should not be taken as a mean of comparison for the importance of distributions.

Isotope production

Several types of spallation isotopes are created by muons in water but only 18 of them β decay in a way that can mimic an interesting neutrino signals in the detector. The rest are stable, have long half-lives, decay invisibly in SK or have a negligible yield and are not taken into consideration. Table 4.2 shows a list of the background isotopes with their half lives and decay products: only β, γ and neutrons are considered since Super-Kamiokande can only detect relativistic charged particles. In particular gammas are visible through pair production or Compton scattering and neutrons, despite the very low detection efficiency [52], emit a 2.2 MeV γ after being captured on hydrogen. Other invisible particles, such as proton, alpha or small mass number isotopes, can be created but are ignored.

The second column of Table 4.2 lists the values of isotope half lives, taken from [120]. The range of half lives, if we take into account all the spallation isotopes, spans over five orders of magnitude, from 0.008 s (^{13}O) to 13.8 s (^{11}Be). This time scale needs to be compared with the rate of muons in the detector, 2 Hz. Since less than 1% of muons causes spallation that can be a source of background, less than 1 muon per minute is responsible for the creation of background isotopes. Super-Kamiokande energy threshold and detection efficiency have additional impact on this rate. The fourth columns lists the kinetic energy endpoints of the decay spectra.

The fifth and sixth columns show the isotopes yields calculated with FLUKA simulation in this work and from Li and Beacom's study in Ref. [53]. The error included in column four is statistical. We globally find good agreement with the published results: although some of the isotopes with low yield, like ^{17}N, have larger discrepancies, we consider them related to the fact that for the calculations in [53] some approximations have been introduced and their error on the yield values was not shown. Moreover, the range of yields spans five orders of magnitude, the relative importance of the different isotopes is well represented by our FLUKA results and correspond to the one in [53]. Finally, given

Isotope	Half-life[s]	E_{kin}[MeV]	Decay mode	Yields$\times 10^{-3}$ (μ^{-1}g^{-1}cm^3)	Yields[53]$\times 10^{-3}$ (μ^{-1}g^{-1}cm^3)
^{18}N	0.624		β^-	0.007 ± 0.004	0.006
^{17}N	4.173		$\beta^- n$	0.14 ± 0.02	0.19
^{16}N	7.13	4.27+6.13(γ)	$\beta^-\gamma$ (66%)	5.8 ± 0.1	5.8
		10.44	β^- (28%)		
^{16}C	0.747	~ 4	$\beta^- n$	0.01 ± 0.006	0.006
^{15}C	2.449	4.51+5.3(γ)	$\beta^-\gamma$ (63%)	0.27 ± 0.002	0.26
		9.77	β^- (37%)		
^{14}B	0.0138	14.55+6.09(γ)	$\beta^-\gamma$	0.01 ± 0.005	0.006
^{13}O	0.0086	$8 \sim 14$	β^+	0.06 ± 0.01	0.08
^{13}B	0.0174	13.44	β^-	0.59 ± 0.03	0.61
^{12}N	0.0110	16.38	β^+	0.36 ± 0.03	0.41
^{12}B	0.0202	13.37	β^-	3.8 ± 0.09	3.9
^{12}Be	0.0236	11.71	β^-	0.03 ± 0.008	0.03
^{11}Be	13.8	11.51)	β^- (55%)	0.21 ± 0.02	0.26
		9.41+2.1(γ)	$\beta^-\gamma$(31%)		
^{11}Li	0.0085	20.62	$\beta^- n$	0.007 ± 0.004	0.003
^{9}C	0.127	$3 \sim 15$	β^+	0.26 ± 0.02	0.29
^{9}Li	0.178	~ 10	$\beta^- n$ (51%)	0.59 ± 0.03	0.60
		13.6	β^-(49%)		
^{8}B	0.77	13.9	β^+	1.8 ± 0.06	1.9
^{8}Li	0.838	13.0	β^-	4.1 ± 0.09	4.2
^{8}He	0.119	9.67+0.98(γ)	$\beta^-\gamma$ (84%)	0.06 ± 0.01	0.07
		13.6	$\beta^- n$(16%)		

TABLE 4.2: Table of isotope yields of a single muon track in water divided by the water density (1 g/cm^3). Only background isotopes for SK are included. Half-lives listed in the second column, endpoint kinetic energies in the third column are from [121, 122]. Yields in the sixth columns are taken from [53]. The fifth column shows the results of FLUKA simulation with statistical errors.

that we expect systematic errors on isotope yields to dominate the final results, due to uncertainties related to hadronic models in FLUKA that can be up to 100%, we can consider the agreement found from the comparison of our results with Li and Beacom's studies an additional validation for our simulation.

A final distribution we produced from FLUKA results with vertical muons, which shows a good agreement with Li and Beacom's work, is the energy spectrum for the principal types of secondary particle that interacts producing background isotopes. The distributions for the most abundant isotope parents can be found in Figure 4.18. The relative heights on the y axis give the importance of each particle: neutrons dominates the spallation production in Super-Kamiokande, at few MeV energies they are 10 times more important than the other parents. The shape of the spectrum comes from the convolution of neutron nuclear cross-section and its path length shown in Figure 4.20 [53]. In particular, the shape of the distribution at low energies is due to the (n, p) interaction that peaks below 20 MeV, where thermal neutrons interact with hydrogen. The relative importance of π^+, π^- and γ is comparable except for high energies where gamma contribution start decreasing. Once again, the spectra shows how, despite the fact that muons in water

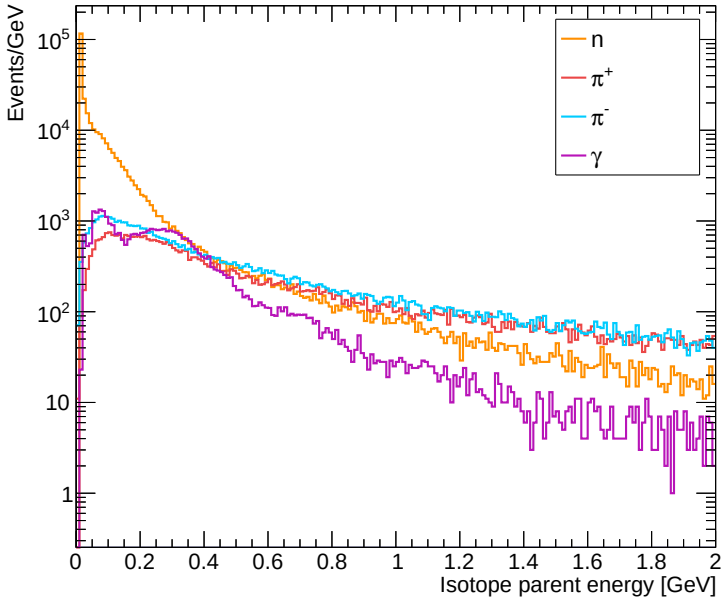


FIGURE 4.22: Isotopes parent particle kinetic energy spectra in SK from FLUKA simulation of vertical muons. The y axis represents the number of events per GeV: the relative height gives an idea of the importance of different parent particles and shape of the curves reflects their interaction processes at different energies.

generate photons and electrons in abundance, the ones that are responsible for isotope creation are mostly hadrons.

All the results presented in this Section are generated with FLUKA by simulating negative vertical muons propagating in SK detector. A good agreement between the simulation we developed and the results published by Li and Beacom in the paper of Ref. [53] is found. Their work, being the first detailed study of spallation processes in water, has been taken as a reference for validating the FLUKA simulation: their results are in good agreement with SK data (within a factor 2). Moreover Li and Beacom performed FLUKA based calculations for scintillator detectors and obtained satisfactory agreement with KamLAND [100] and Borexino [101] published results, demonstrating the accuracy of simulation performance.

In the calculation presented above, several simplifications were adopted: among others, cosmic muon flux is not simulated, antimuon contributions not taken into account and the effect of the detector response is not studied. A more exhaustive simulation, accurately modeling muon interactions in water to deeply study spallation background in SK, requires the construction of a more sophisticated framework. The implementation of angle and energy distribution computed with MUSIC and an interface with the official MC simulation, SKDetSim, to study the detector response are necessary and will be accurately implemented. Despite the simplifications, with the first simulation outcomes exposed in this Section we were able to obtain some insights of the processes occurring in SK. A more accurate calculation will give us additional interesting details and insights: Chapter 5 will be entirely dedicated to the study of spallation using FLUKA. A better understanding of isotopes production with the help of MC simulations could result in major improvements of spallation background rejection.

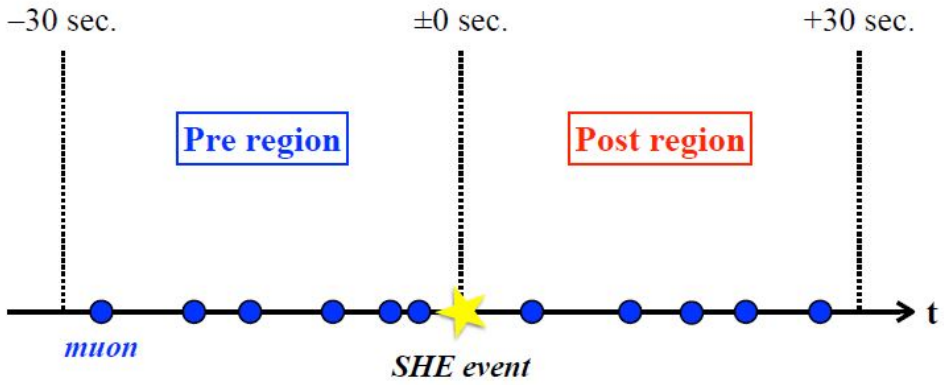
4.6.4 Comparison with SK spallation data

In the previous Sections, studies checking the independent validity and well-functioning of MUSIC and FLUKA simulations were presented. A comparison with SK data containing radioactive isotope candidates can be a powerful confirmation of the capabilities and effectiveness of our spallation simulation in its entirety. As described in Section 4.2, the simulation structure is made of different inter-chained parts which aim to model Super-Kamiokande conditions. From now on, results from this whole pipeline, including MUSIC, FLUKA and SKDetSim as shown in Fig. 4.1, will be used.

4.6.5 Spallation in data

Spallation events used for comparison with MC results are the same set of data used in SRN search. In the analysis these spallation samples are used to tune rejection techniques for this background using likelihood evaluation, as we will discuss later in this section. Their structure and selection will be described in the following. Despite the large fraction of low energy signals below 20 MeV consists of spallation events, isolating a sample of pairs composed by a primary muon and the associated radioactive isotope is quite challenging. In fact, due to the high muon rate and the long half-lives of certain isotopes, it is not possible to unambiguously associate the spallation signal with its parent muon. Therefore each low energy candidate is paired with all muon passing up to 30 s before. This results in the construction of about 60 pairs: among them there could be one correlated muon+isotope association while the others are uncorrelated. This sample of pairs is called “Data sample". A similar sample, “Random sample", is generated pairing the low energy event with muons in a time region of [0, 30] s after the low energy signal: all these pairs are inevitably uncorrelated. Since we cannot have a pure spallation sample, both data and random samples are important to estimate spallation contamination, as it will be explained later. Constructing pairs and for each of them computing a series of observable to investigate the space-time correlations between muon and radioactive decays is important because it is the key strategy for spallation reduction in SK-IV. Cuts on these variables are applied in order to discriminate if the muon+isotope pair is correlated, namely a spallation event, or if it is just an accidental association. However, we need to keep in mind that we are interested in the low energy event, rather than all the pairs it composes with preceding muons: pairs are a necessary tool to decide if the signal is a possible spallation contamination. Indeed if at least one of the pairs the low energy candidate forms with the ~ 60 preceding muons is judged as spallation-like the event is tagged as background and removed.

More specifically, spallation candidate events are extracted from data taken with SHE+AFT triggers in the SK-IV period. Non physical signals, such as PMT noise and calibration events, as well as cosmic muons decays are removed with targeted selections, more details can be found in Ref. [22, 39]. As mentioned, with the removal of non interesting signals, and considered the very low rate of neutrino interactions, the dominant contribution of low energy events in data is given by cosmic muon induced spallation. In order to pair the SHE-triggered event with cosmic muons, muon-like signals are selected and reconstructed with MuBoy algorithm. In particular, events carrying and OD and HE trigger and number of PMT hit greater than 1000 are passed to the muon fitter that computes track informations, energy deposition along the path (dE/dx) and muon status. Since only single through-going muons are simulated, data with status 1 (namely single through going muons) are selected. The two samples constructed with preceding/succeeding muons respectively in the regions in time of [-30, 0] s and [0, 30] s are illustrated in Figure 4.23. Both data and random sample contains pairs of low energy-muon signals, however only



4.23.

FIGURE 4.23: Time regions defined with respect to the SHE-triggered event. "Pre region" defines the data sample while the "post region" defines the random sample. Taken from [39].

the data sample can include pairs with the correct time correlation to be spallation events.

The set of observable computed for pairs to investigate space-time correlations of the muon with the isotopes, often called discriminating variables, and used for spallation event identification are:

- *dt*: time difference between SHE event and the candidate parent muon.
- Q_{peak}: the maximum dE/dx value along the track, within a sliding window of 50 cm.
- *Lt*: distance between the SHE event and the muon track, transverse to the direction of the track.
- *Ll*: longitudinal distance along the muon track from the SHE event vertex and the location of the maximum energy deposit (Q_{peak}).

Figure 4.24 shows a schematic illustration of dt, Q_{peak}, Ll and Lt. Distributions of these variables are made for both samples and used to study the characteristics of spallation events, they are shown in Figure 4.25. In particular, we used low energy events with reconstructed energy above 8 MeV and muon identified as single through-going (since in the simulation no muon bundles are treated). The distributions of data sample (magenta in Fig. 4.25) show clear excesses with respect to the shape of random sample. These contributions are due to spallation: radioactive decays induced by cosmic muons are more likely to be correlated in time and space with the parent muon and to be produced in showers that result in an increase of the local energy deposition pattern. Specifically, we observe in the top left plot of Figure 4.25 a flat time distribution for random pairs while the data sample, which contain spallation pairs, is influenced by the distribution of radioactive isotope half lives. Similarly, isotopes are expected to be close to the muon track (see the excess at low Lt in the bottom left of Figure 4.25) and closely correlated with the position of Q_{peak} (Ll distribution has a peak due to spallation close to zero, bottom right of Figure 4.25). In addition, Q_{peak} value for data sample in the top right of Figure 4.25 suggests that spallation muons locally deposit larger charges: this is the effect of shower development. The observed effect enables us to extract from data sample the distribution of discriminating variables for spallation events: by subtracting random sample distributions (azure in Fig. 4.25) from data we obtain distributions that illustrate spallation properties and can be directly compared with simulation results for radioactive isotopes. The subtracted histograms represent the properties of the so-called "spallation sample".

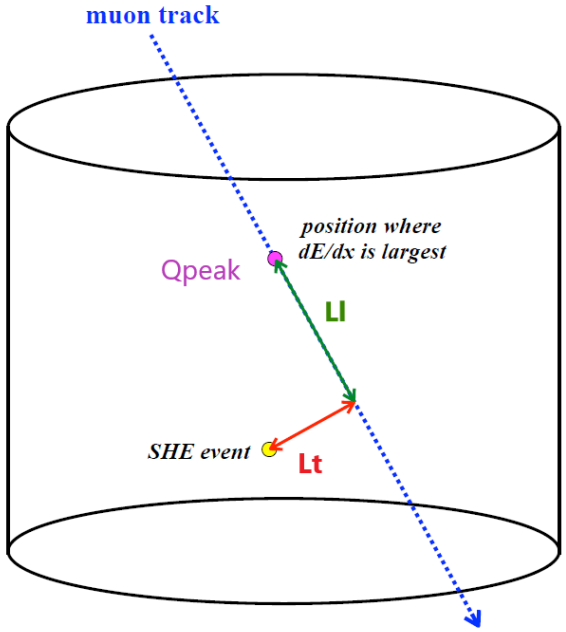


FIGURE 4.24: Schematic illustration of the discriminating variables, Lt, Ll, Q_{peak}.

It is important to mention that the distributions shown in Figure 4.25 are used in the SRN analysis to build likelihood functions for the discrimination of background events: probability density functions (PDFs), both for spallation and random pairs, are prepared for each variable by mean of subtraction and normalization. A likelihood is described as:

$$L_{spall} = log(\prod_{variable_i} \frac{PDF^i_{spall}}{PDF^i_{random}}) \tag{4.6}$$

Cut criteria are defined on the likelihood values and the SHE-triggered event is tagged as spallation if at least one muon in the [-30, 0] s time region gives a likelihood larger than a selected threshold. Additional details on spallation rejection cuts can be found in Ref. [39].

Comparison Data-MC

In order to compare data and MC, we built distributions of the discriminating variables for spallation data and simulated isotopes, the detector response is modeled with SKDetSim and the same reconstruction algorithms are used for MC and data signals. In addition, preliminary selection criteria used to remove noise signals in data are also applied to simulated events.

Muon inducing spallation

A powerful validation of the simulation results comes from the comparison of spallation muons properties with data distributions. As described previously in the case of discriminating variables, the subtraction of data in the post region from the pre region allows to isolate spallation contribution and extract histograms with the associated variables.

Figure 4.26 shows distributions of charge-related and path length observables for muons generating spallation in FLUKA and muons in the spallation sample from SK

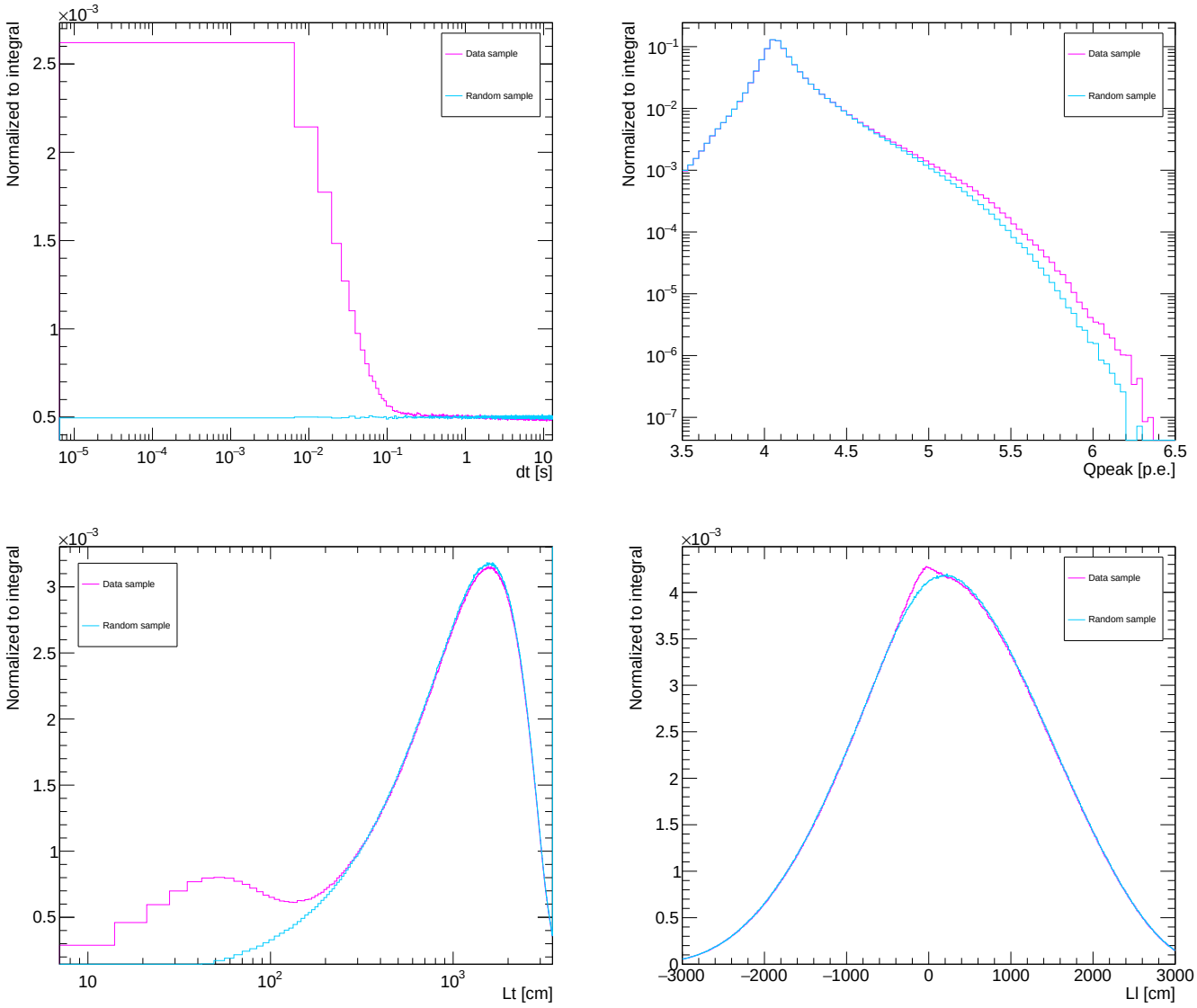


FIGURE 4.25: Distribution of discriminating variables for data (magenta) and random (azure) sample. Both data and random histograms are normalized to the integral of the region where no excess is observed due to spallation events. Note that for the time distribution the absolute value of dt is plotted.

data. Given that the probability of generating a shower containing a background isotope increases with the traveled path, we expect to find that spallation-inducing muons have an average longer track. Indeed, in the top left of Figure 4.26 we observe a drastic reduction of the contribution of short travel distances with respect to Figure 4.16, representing the average of all muons in SK. Analogous results are presented on the top right of the same Figure where the total deposited charge (Q_{tot}) is shown: compared to Figure 4.17 the histogram has larger tails and small events at low charge, underling the higher light produced on average in spallation showers. Both the track length and the total charge spectra from simulation present a good agreement with data, confirming the accuracy of the calculations. In particular, the good agreement of MC-data, especially for what concerns charge variables, not only enlightens that FLUKA models correctly represent the probability of generating spallation isotopes, the parent muon behavior, the shower development and its particle contribution, but also that the way FLUKA results are interfaced with SKDetSim does not introduce significant errors in the light production properties of the muon. In fact, we remind that the structure of our spallation simulation is build in such a way that some processes are treated in SKDetSim while others in FLUKA: pure electromagnetic showers are indeed only generated in the Super-Kamiokande detector simulation while inelastic processes and electromagnetic interactions above 0.1 GeV, if generating a shower containing spallation, are taken from FLUKA (and their propagation and resulting Cherenkov light modeled in SKDetSim). This selection, based on the nature of the interactions, implies a manipulation of GEANT and FLUKA packages in order to activate/deactivate specific processes (as described in Sec. 4.2). Possible errors could come from a double generation or absence of certain type of reactions and the subsequent over/underestimation of Cherenkov light produced. The fact that charge profiles from data and MC have similar shapes excludes that a significant error is introduced as a consequence of incorrect FLUKA-SKDetSim interface. Distributions on the bottom of Figure 4.26 represent the residual charge resQ (left) and Q_{peak}. Q_{peak} is the discriminating variable described above while resQ is defined as:

$$\text{resQ} = Q_{tot} - Q_{ion} \times \mu \text{ track lenght} \tag{4.7}$$

where Q_{tot} is the total charge deposited and Q_{ion} is the charge deposited per cm only by ionizing processes. Therefore resQ represents the light signal only due to radiative processes, among them spallation showers. The good agreement of resQ and Q_{peak} proves that the simulation correctly deal with shower energy depositions. The minor discrepancy in the first bin of resQ suggests that the simulation is slightly overestimating the radioactive contribution in case of very small showers, this effect is less evident in the Q_{peak} distribution.

Spallation isotopes

Quantities in Figure 4.26 are uniquely related to the primary muon and its propagation in water, the characteristic of the low energy event the muon is paired with will be presented in the following. The transverse distance of the isotope decay from the muon track (Lt) is shown in the left of Figure 4.27. It is a purely geometrical (charge-related observables part of Lt definition) variable and the agreement with data is good: this means that the mechanism of isotope generation from a particle within the shower, after it propagates for a given distance, is well modeled by the simulation. On the contrary, the longitudinal distance from Q_{peak} position (Ll) in the right plot of Figure 4.27 does not reproduce the shape of data: the simulation gives an average larger isotope distance from the location of maximum energy deposit. The discrepancy can come from two factors: the simulated showers

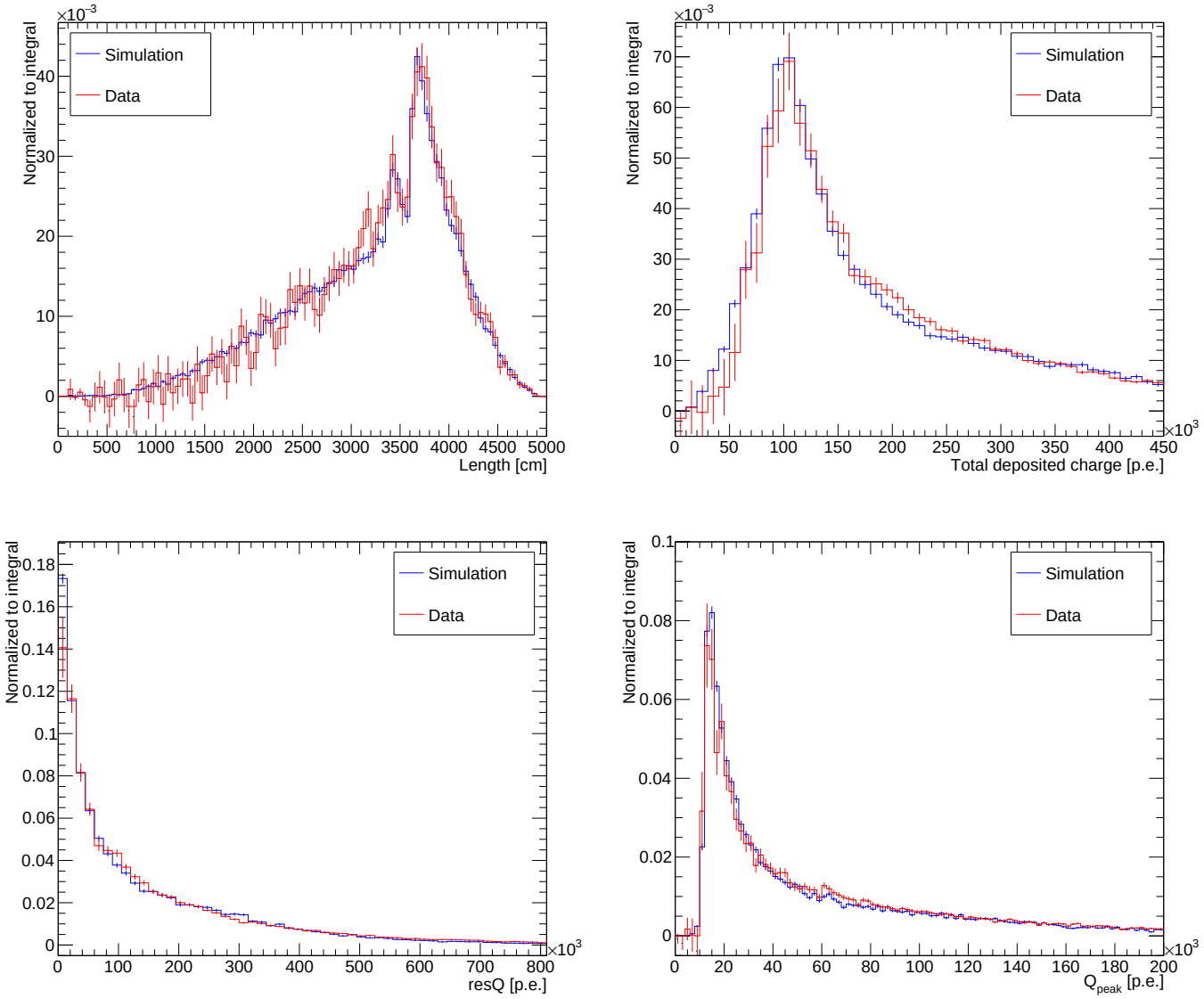


FIGURE 4.26: Distributions from data (red) and simulation (blue) of spallation muon related variables: path length (top left), total deposited charge (Q_{tot} on the top right), residual charge as defined in Eq. 4.7 (bottom left) and Q_{peak} (bottom right). All distributions are normalized to their integral in order to compare MC and data results.

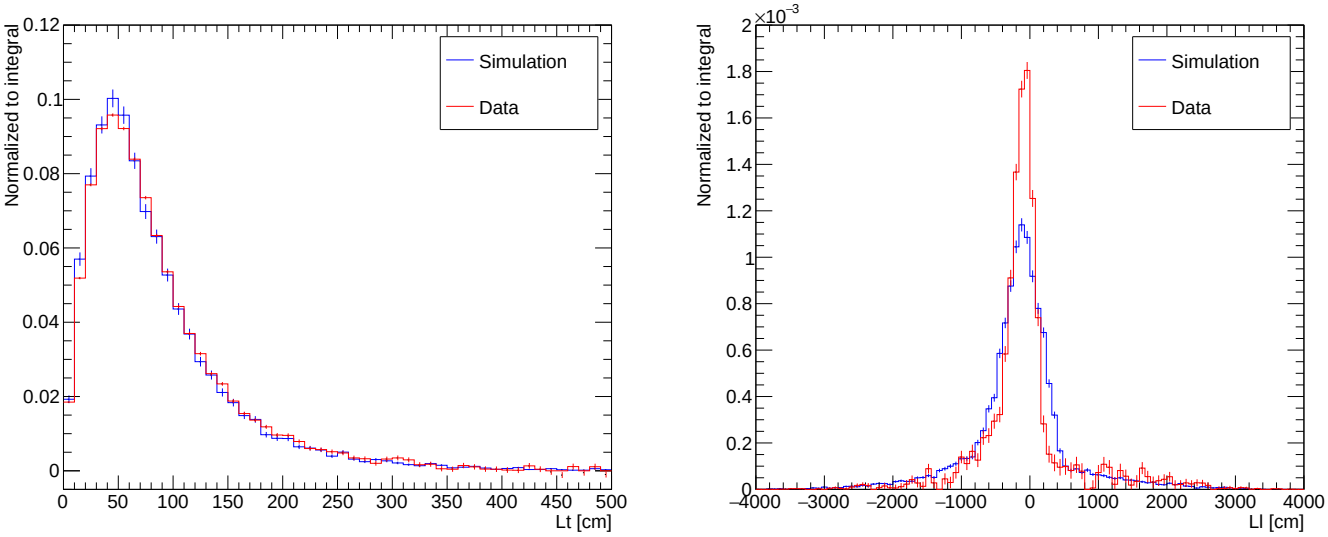


FIGURE 4.27: Distribution of the discriminating variables Lt and Ll for data (red) and simulation (blue).

can have a larger longitudinal extension with respect to what happens in data and/or Q_{peak} localization is not accurate. If the difference is only geometrical, namely related to a different size of showers in MC, we would expect an impact on Lt distribution too. Since this is not the case, the disagreement is most probably due to a mis-reconstruction of Q_{peak} position. A confirmation comes from Figure 4.28: together with Ll for data and MC, the distribution of longitudinal distance of the isotope from the shower barycenter is plotted (black dots). The center of the cascade is computed averaging the vertex positions of all the particles in the shower, if a shower exists. The distribution better matches Ll from data: its width confirms that the effective shower longitudinal extension rarely exceeds 5 m (full width at half maximum is about 4 m), proving that Q_{peak} localization fails to identify the position of spallation showers in our simulation. In fact, although the value of Q_{peak} is well reproduced (bottom right of Figure 4.26), the structure of the simulation can introduce errors in the identification of its position: all the pure electromagnetic showers are only simulated in SKDetSim and their injection position can in principle be different in FLUKA. If we look at variables averaged over the entire track length of different muons (like Q_{peak}, resQ...) the difference is averaged and absorbed. However the situation changes for quantities that rely on the exact position of the shower vertex. As we will see in the next Chapter, despite most of the electromagnetic showers are not very energetic, sometimes EM radiative interactions produce a huge amount of light, in particular through bremsstrahlung processes. It is rare to see two energetic showers, namely competing to be identified as the maximum charge deposition Q_{peak}, along the same track but the fact that shower with isotopes (“spallation showers") are simulated in FLUKA while pure EM are treated uncorrelated by a different MC package, SKDetSim, can increase the probability. Shower generation is in fact a stochastic process, SKDetSim and FLUKA are independent codes, therefore it could happen that while a medium to low energy spallation cascade is simulated in FLUKA, SKDetSim predicts an EM radiative emission for the same muon. In this case where a significant amount of light is deposited in two position along the track, the EM one can result to be more energetic than the spallation shower and it can occur that Q_{peak} position in not associated with the shower containing radioactive isotopes. It is worth mentioning that, at equal initial energy, electromagnetic showers are

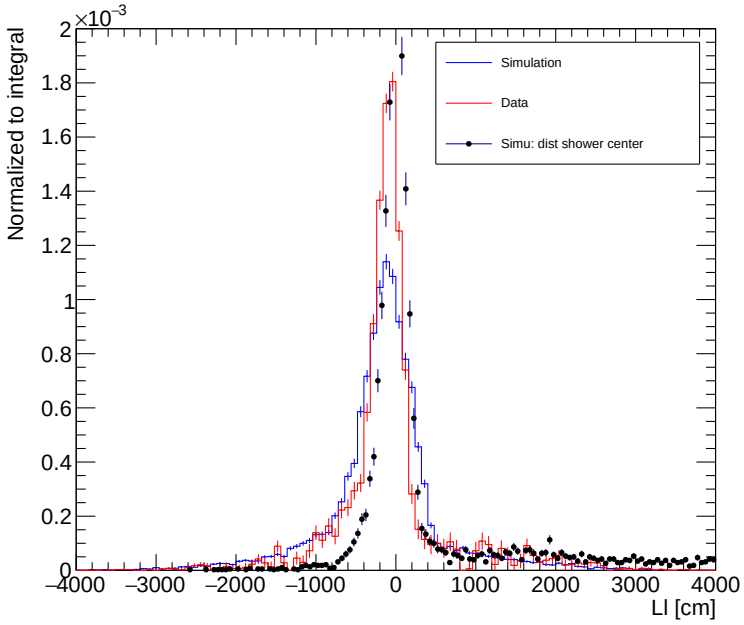


FIGURE 4.28: Red and blue: distribution of Ll discriminating variable, namely longitudinal distance on isotope candidate from Q_{peak} position, for data and MC respectively. The black dots represent the longitudinal distance of the isotope candidate in simulation from the barycenter of the shower (if a shower exists), computed with the reconstructed vertex of the particles within the shower.

in general more efficient in producing Cherenkov light, which is directly related to electron propagation, while hadronic showers produce neutral particles and contain pions and neutrons that becomes non relativistic at higher energies with respect to electrons. We confirm that an effect of this kind do not appreciably affects nor the total energy deposition of the muon, neither resQ and Q_{peak} value, on the contrary the identification of Q_{peak} location is more sensitive to side effects linked to the complexity of simulation structure. Moreover, an additional source of misalignment between data and MC distribution can arise from discrepancies in the isotope composition of the spallation background: we expect up to a factor 2 discrepancies between observed and simulated isotope yields [53, 101]. In the next Chapter we will explain in details how different isotope types are generated by distinct parent particles, each of them having a characteristic mean free path in water impacting on the radioactive decay position. Although the effect observed on Lt is minor, this can still affect geometrical distributions like Ll. Finally, Ll distribution is one of the variables for SRN search but it has been proved that its discriminating power is lower compared to others (dt, Lt) [39]. It will not be used further in this work and all the quantities from simulation related to Q_{peak} will be treated with special care. We will focus on variables proven to be robust against possible mismodeling of the Cherenkov light pattern like isotope yields and geometrical properties of the clusters of neutrons.

In conclusion, this Chapter was dedicated to the description of the simulation of cosmic muon interactions in water and generation of spallation events in SK. Several tests were made to verify the validity of simulated results, both for individual parts of the structure and for the whole simulation. An overall good agreement with data was found, particularly satisfying considering the complexity of the designed structure. Therefore

this simulation is believed to provide sufficiently accurate outcomes to study spallation processes and to develop predictions for data in SK. In addition, thanks to unique insights given by FLUKA calculations, it was possible to investigate details on the main mechanism developing radioactive decays: more specifics on muon propagation in water, shower development and isotope generation will be presented in details in the following Chapter.

Chapter 5

Studies of spallation mechanism with FLUKA simulation

The simulation described in Chapter 4 provides a unique opportunity to shed light on the mechanism of spallation production in water, the properties of primary muons and the time and space correlations with other particles within the same shower. Insights on these processes are extremely difficult to be obtained from data where detection efficiencies and reconstruction uncertainties preclude a detailed description of single microscopic interactions as well as isotope separation. On the contrary, the simulation developed with FLUKA as a core package for shower generation grants comprehensive information for each particle interaction step and can advance our understanding of spallation processes in a water Cherenkov detector such as Super-Kamiokande. These studies can be extremely important to better tune spallation rejection techniques. In particular, clusters of isotopes are produced with and their physical characteristics can be a powerful information to locate spallation background, making the investigation of muon-induced showers essential.

A first description of how spallation isotopes are made was presented in the previous Chapter (Section 4.6.3), where, despite many simplifications adopted in the simulation, a strict bond between the presence of unstable isotopes and muon-induced showers was enlightened. A deeper investigation, introducing angular dependence of cosmic rays and a more accurate calculation of muon fluxes at SK will be presented in this Chapter. Spallation production mechanisms, differences of one isotope from the other, their correlation with preceding muons as well as geometrical properties of the showers will be presented. Muons are classified according to the way they deposit energy and treated separately if they, or a particle in a shower they induced, generate spallation isotopes: the study of their properties provides powerful tools to identify the conditions for the creation of an unstable isotope. These predictions in turn allow to tune spallation cuts for future analyses. Furthermore, all results obtained by FLUKA can be valuable not only for Super-Kamiokande experiment but also for other water-based detector such as the future Hyper-Kamiokande [58]: the depth dependence of isotope yields does not vary very strongly if the experiment has a sufficient rock shielding on top. This is a direct consequence of the moderate variation of isotope production with muon energy ($E_\mu^{0.8-1.1}$ [53]).

5.1 Showers properties

Given the strong link between the development of a shower as a consequence of radiative processes along the muon track and the generation of spallation isotopes, a brief introduction to summarize the main properties of hadronic and electromagnetic showers can be very helpful. Shower mechanism starts with a radiative emission of a given particle that

multiplies through repeated interactions. The initial energy is distributed among secondaries and decreases at each interaction. Particle multiplication continues until the critical energy is reached, this value depends on the nature of the shower, the medium and the particle type. After the critical energy ionizing energy losses start to dominate.

5.1.1 Electromagnetic showers

Electromagnetic showers begin when high-energy photons, electrons or positrons cross a material. When muons travel in water a huge amount of these particles are produced via bremsstrahlung (photons) as well as delta ray (electrons) and pair production from muon interaction with nuclei. In turn, photons and electrons interact with matter via pair production and bremsstrahlung respectively, creating further particles until their energy falls below production threshold and other non radiative energy losses start to dominate. At lower energies, below 1 GeV, Compton scattering is also present. Two variables are used to describe shower development: the radiation length (X_0) and the Moliere radius (r_M). X_0 is defined as the mean distance over which an energetic electron or positron (E > 1 GeV) loses 63% (i.e. 1 - e^{-1}) of its energy. For a given material the radiation length can be approximated with the following formula [2]:

$$X_0 = \frac{716.4A}{Z(Z+1)log(278/\sqrt{Z})} g/cm^{-2} \tag{5.1}$$

where Z is the atomic number and A is mass number of the nucleus. The radiation length is expressed in g/cm^{-2} but results in cm can be obtained by dividing by the density. For water X_0 is equal to $\sim$ 36 cm. The Moliere radius is defined as:

$$r_M = 2.2MeV\frac{X_0}{E_C} \tag{5.2}$$

whit E_C the critical energy of the given material (78.33 MeV for electrons in water), namely the energy at which bremsstrahlung and ionization rates are equal. r_M = 9.9 cm for pure water. Lateral shower extension is almost energy independent and on average 95% of the shower particles are contained in a cylinder with radius $2r_M$. In fact electron and positrons are mostly forward with an average $\cos\theta_Z \approx 0.9$ [98]. On the contrary longitudinal expansion strongly depends on the energy of incident particle: in a 1-10 GeV range, the containment of 95% of the shower requires about 11 - 13 X_0, namely about 4 meters. For higher energies, 100 GeV, the extension can reach 6 - 7 m. Longitudinal and transverse sizes of showers are important because they define the region where isotopes might be contained. Luckily shower dimensions are much smaller than the detector height and customized cuts can remove them without increasing the dead time.

5.1.2 Hadronic showers

Hadronic showers are initiated by charged pions, kaons or other hadrons. These can be created by cosmic muons after interactions with medium nuclei. The physical processes responsible for the propagation of a hadronic shower are considerably different from the processes in electromagnetic showers. A large quantity of the incident hadron energy is shared in multi particle production processes which lead to the creation of new secondaries, often slow pions and neutrons. The most common phenomena which determine the development of the hadronic showers are: hadron production, nuclear de-excitation and pion and muon decays. One of the main difference between electromagnetic and

hadronic processes is that the first ones tend to leave the target nucleus intact while inelastic interactions within hadronic cascades can be very destructive. All hadronic showers have a considerable electromagnetic component, which increases as the shower develops: 1/3 of the produced pions are in fact neutral and lead to the production of gammas as they decay. The longitudinal development of hadronic showers scales with the nuclear interaction length (λ_{int}), which is an empirical quantity and for a given material is always larger than the radiation length (e.g. Water: X_0 = 36.08 cm and λ_{int} = 83.33 cm [123]). As a consequence, hadronic showers take longer to develop than the electromagnetic ones. The probability for an hadron of having an interaction after traveling a distance x in the material is (1 - $e^{-x/\lambda_{int}}$). While lateral shower displacement in electromagnetic showers is mainly due to multiple scattering of electrons and positrons, in the case of hadronic showers it is related to the high transverse momentum transferred during nuclear reactions and is in average larger then the EM case. In general, it takes less material to laterally contain high-energy hadronic showers than low-energy ones. This is a consequence of the fact that the electromagnetic shower fraction increases with energy (more π^0 are created) and EM components develop close to the shower axis. Hadronic showers due to muon propagation in water are a rare process: about 1% of the total shower production [98].

5.2 Muon properties and classification

The simulation we used is built as illustrated in Chapter 4. The main improvement from previous calculations is the accurate determination of muon fluxes at Super-Kamiokande: MUSIC angle and energy distributions are used as inputs for FLUKA Monte Carlo code where the interactions of muons in water are simulated in detail.

5.2.1 Charge ratio

Unlike the setup description in Section 4.6.3, both negative and positive muons are generated. As demonstrated in previous studies [53, 100], we expect the two types of muons to lose energy and develop showers in a very similar manner, with a resulting few percents difference in the isotope yields. The main cause of this difference is the possibility of μ^- to be captured by oxygen nuclei and subsequently create isotopes, reaction that does happen with μ^+. We adopt a muon charge ratio, defined as the number of positive over negative charged muons, of N_{μ^+}/N_{μ^-} = 1.27. This value is obtained from a measurement done by CMS experiment for muons at sea level [124]. On this subject, a study made by MINOS experiment shows that the charge ratio varies depending on the energy of the atmospheric muons [125]. Thus, since only the most energetic muons can penetrate deep underground to reach SK detector, we expect this value to slightly differ from the one at surface. In particular, only muons with surface energies from few hundreds of MeV to 300 GeV maintain a charge ratio of about 1.27, at higher energies several competing processes affect the value found by CMS. More specifically, above few hundreds of GeV the fraction of muons from kaon decays starts to grow, due to the fact that pions have longer lives compared to kaons and become more likely to interact in the atmosphere, instead of decaying with muon production. This consideration does not hold for surface lower energies since most of the less energetic muons are produced at great heights, before any interaction with the atmosphere. Therefore, positive muon fraction is expected to rise at underground sites for two factors: first, kaon decay contribution gains importance with increasing energy, and second, due to differences in strong interaction production channels, muon charge ratio from kaons is greater than that from pion decays. Other processes, like the presence of heavier elements in cosmic rays as the energy increases, could

also affect the muon charge ratio. OPERA, which is located at a depth of 3800 m water equivalent, measured a value of $N_{\mu^+}/N_{\mu^-} = 1.37$ [126]. Since an analogous measurement is not available for Super-Kamiokande depth of 2700 m water equivalent, we evaluate how the choice of charge ratio value impacts our results by running a test simulation with $N_{\mu^+}/N_{\mu^-} = 1.37$, to be compared with results with the nominal value $N_{\mu^+}/N_{\mu^-} = 1.27$. We find that only 7.8% of the isotopes are generated by negative muons captured on oxygen, namely the main interaction which differs between μ^- and μ^+, and among them the large majority, about 70%, lead to the creation of ^{16}N. The ^{16}N total yield varies of less than 1% when taking $N_{\mu^+}/N_{\mu^-} = 1.37$ instead of 1.27. We conclude that the choice of muon charge ratio, without taking into account the effect of mountain depth, does not introduce significant errors in our study. As a confirmation, we observe that ^{16}N isotopes from stopping muons are about 8% of the total yield for the whole isotope species, when $N_{\mu^+}/N_{\mu^-} = 1.27$. Among these stopping muons there are reactions common to both types of muons, like isotopes production along the track, and the contribution of captures on oxygen only possible for negative muons. The simulation values are in good agreement with empirical and theoretical results from [53, 100], confirming the accuracy of our choices.

5.2.2 Muon energy loss and shower production

We underlined that our main goal is to study the characteristics of spallation isotopes inside Super Kamiokande detector and a powerful mean to understand this critical background source is to investigate the processes that lead to the generation of unstable isotopes. For this reason, a specific attention was dedicated to the study of muon-induced showers which contain unstable isotopes that can fake low energy signals in SK (for simplicity often referred as “background showers" or “spallation showers"). More generally, muon energy loss processes in water are studied: muons can simply ionize the medium around the track or radiate secondary particles which in turn can lead to electromagnetic or hadronic showers.

Since only ionizing muons and muons inducing purely electromagnetic cascades do not lead to the production of isotopes we do not record them in FLUKA. Note that we define “purely electromagnetic (purely EM)" a shower that do not contain any hadron. However a generic EM cascade can in principle accommodate hadrons: under some circumstances a high energy gamma within an electromagnetic shower can be absorbed by an atomic nucleus, which after entering an excited state, decays by emitting subatomic particles. Neutrons, protons or alpha particles are knocked out of the nucleus, (γ, n), (γ, p), (γ, α), travel and interact with surrounding water. These reactions are called photonuclear and, due to the large mass of pions and neutrons, require high energies to occur. In this specific case, the informations FLUKA provides is the injection properties of the shower, namely the nature of the primary EM interaction triggering the cascade development and its energy, and all the details on the inelastic interactions within the shower. All the electromagnetic processes, different from the primary one, are neglected. The same applies for the electromagnetic component of hadronic showers: in fact, in the way we construct the simulation, electromagnetic interactions are meant to be simulated in SKDetSim rather than in FLUKA (see Sec. 4.5). In this way, Cherenkov light production is treated uniquely by SKDetSim while FLUKA only investigates hadronic interactions, isotope generation and decays. The only exception is indeed when a photonuclear reaction develops within an EM shower, possibly producing an unstable isotope: the shower triggering process is important to define the characteristics of spallation cascades.

By analyzing various ways muons lose energies we can investigate the peculiar properties of spallation showers: in average we expect isotopes to be contained in showers with a larger extension both in space and time, with a greater initial energy and an higher

fraction of hadronic interactions. A comparison with generic showers can help us to identify the features characterizing the environment where spallation is produced and to be able to better remove them from data. In this Chapter we focus on FLUKA outcomes to deeply interpret spallation production, neglecting for the moment the detector response. Following these considerations, we classify three types of muons according to their energy losses. Each category is treated differently and simulated independently but all of them are representative of a fraction of muons in data. For our scopes, the three muon types are defined as follows:

- *Spallation-inducing muons*: or, for simplicity, “spallation muons". FLUKA propagates muons in water and tracks all the particles created along the path. If an unstable isotope, whose decay products are a possible source of background in SK, is generated, the muon and all the eventual particles created by inelastic interactions within the induced shower are stored. These muons are part of the category called “spallation inducing muons". The way shower components are selected and recorded is widely described in Section 4.4.

- *Hadron-producing muons*: a primary muon inducing a shower containing at least one hadron or isotope, at any interaction level after an initial muon-nucleus process, is considered part of this category. We use this definition to distinguish muons that only lose energy through electromagnetic processes from muons that, directly or through the generation of secondary particles, produce hadrons and isotopes. Spallation showers are included in this category.

- *Electromagnetic only muons*: the large majority of muon crosses the detector without accelerating or creating secondary hadrons anywhere along the track. Most of the times they lose energy as minimum ionizing particles and, more rarely and depending on their primary energy, through purely electromagnetic radiative processes. Muons inducing purely electromagnetic showers are included in this category since they are not relevant for spallation studies. As shown in Section 4.6.2, this class of muons is well simulated by SKDetSim, once muon-nucleus interactions are deactivated, using MUSIC angle and energy distributions. Therefore, it is not necessary to simulate this third category of muons in FLUKA and it will not be treated in this Chapter.

The three muon categories are graphically summarized in Figure 5.1. The signature of cosmic muons in the Super-Kamiokande detector can be fully described with the samples illustrated above. In the next Sections we will focus on the results obtained from FLUKA simulation for spallation and hadron-producing muons in order to investigate properties that can be exploited to identify spallation events. In fact, FLUKA calculations reveal detailed production mechanisms, allow to retrieve the correct parent muon and to associate the secondary isotopes and shower particles, impossible for empirical studies. The detector response is not taken into account in this Chapter: the validity of results can be easily extended to other water based detectors. SKDetSim and the modelization of Cherenkov light within Super-Kamiokande tank will be treated later in this work.

5.3 Spallation inducing muons

1.4×10^8 muons were simulated by mean of FLUKA with an input charge ratio value $N_{\mu^+}/N_{\mu^-} = 1.27$. We find that only less than 0.5% of the simulated muons generate at least one of the 18 isotopes in Table 4.2. This means that about 1.2 muons every 2 minutes are expected to produce a possible source of background in the low energy analysis region

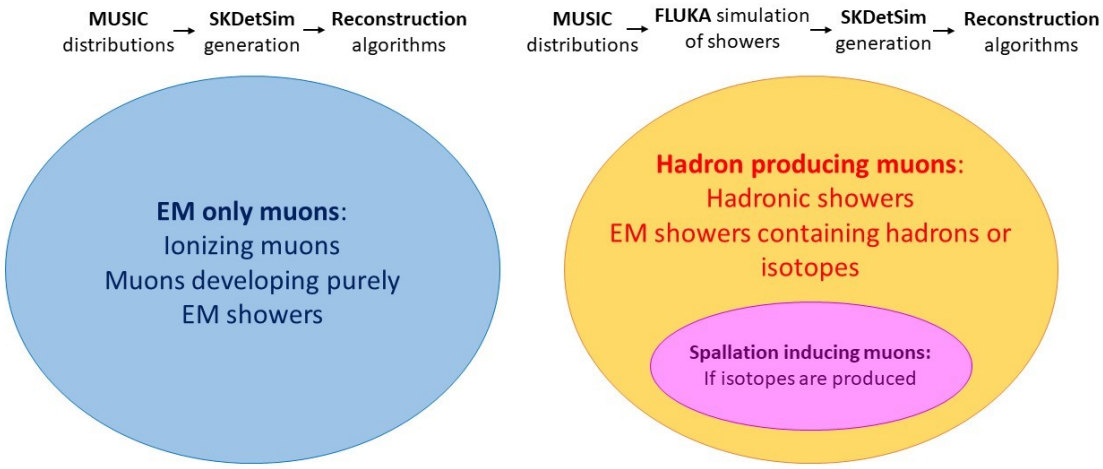


FIGURE 5.1: Illustration of the three muon categories described in the text and the simulation packages involved in the generation.

of Super-Kamiokande. The number of spallation-inducing muons is small compared to the total number of muon crossing the detector (about 2 per second [52]), however it is more than 10^4 times bigger than the expected signal rate for supernovae relic neutrinos. In this Section, the 0.5% fraction of muons are part of the previously defined spallation inducing muon sample, while all the generic 1362×10^5 simulated muons are simply called "primary muons", since they include all the muon categories described before.

Figure 5.2 illustrates the number of background isotopes per muon. It is observed that spallation muons produce in average 2 background isotopes per track and that there is an high correlation between the number of isotopes produced and the muon energy. As a matter of fact, although about 70% create only a single unstable isotope along their path, a small fraction of them manages to induce very energetic showers which contain more than 100 isotopes. This case is very rare, with a fraction below 0.01% of the total showering muons. These events are not problematic for SK because multiple radioactive isotopes from the same parent muon are typically clustered within a few meters, namely the average dimension of showers. Therefore, when several low energy events clustered in time and space are detected, they are identified as spallation and removed. Due to their low rate, it is in fact extremely unlikely to find multiple signals from solar or supernova relic neutrinos in the same region within few seconds. A high number of isotopes produced does not necessarily mean that the global amount of energy deposited is high: there are cases where 100 isotopes are created and about 15 GeV are lost. Figure 4.22 shows that particles generating unstable isotopes do not need to be extremely energetic to break up nuclei, it is a statistical process often occurring at moderate energies ($\mathcal{O}(100)$ MeV). In general, the energy deposited is higher than the light emitted through ionization: this makes possible the detection of a peak in the dE/dx spectrum of the crossing muon which suggests the position of the spallation event. The amount of this localized energy deposition increases with the number of isotopes, which in turn are linked to the dimension of the shower containing them.

Muon energy losses

Among the background muon sample two categories can be distinguished: muons with enough energy to be able to cross the detector and muons stopping inside the tank. Muon bundles are not simulated. Stopping muons represent about 8% of the total and are mainly

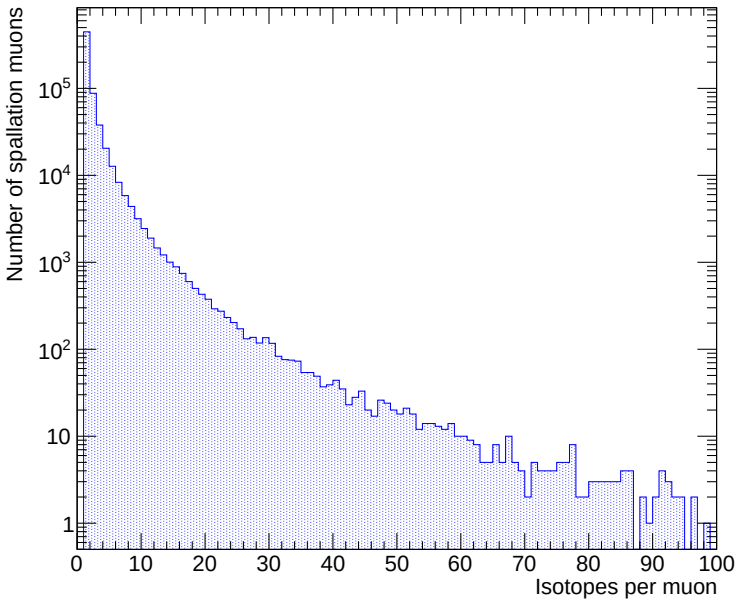


FIGURE 5.2: Number of unstable isotopes that can induce background signals in SK per primary spallation muon.

composed of negative muons that capture on oxygen at the end of their path, leading to nuclear breakup. Stopping muons lose a relative small amount of energy when they cross the detector (in average about 5 GeV in the whole OD) and produce only one isotope along the track, which is ^{16}N for most of the cases, as already discussed. Since Super-Kamiokande can easily distinguish a stopping muon by track reconstruction algorithms and the absence of exit signal in the OD, accurate estimations of ^{16}N rate have been computed with data collected in a sphere centered at the end of muon capture position [127]. These measurements, firstly conducted for calibration studies, report a production rate of about 11 ^{16}N decays per day in a 11.5 kton fiducial volume. If FLUKA results are rescaled for a reduced FV, taking into account the 2 Hz muon rate, in SK we obtain 11.2 ^{16}N per day, in very good agreement with calibration values. Based on similar studies, spallation rejection cuts are implemented in SK to eliminate isotopes from muon captured on ^{16}O by removing a spherical region at the end of the track. Finally, only a 0.4% fraction of isotopes produced by stopping muons are generated along the track, before the stopping location. These events are commonly distributed among μ^+ and μ^-. Figure 5.3 shows muon energy loss spectra for stopping muons and all spallation muons with linear and logarithmic scale, restricted and extended energy range respectively on the left and right plot. In Figure 5.4 negative and positive energy depositions are displayed separately: as expected, negative stopping muons have a larger contribution to spallation production compared to positive ones. From the Figure 5.3 it can be seen how the situation changes when taking into account through-going muons: the energy deposited can vary from tens of GeV to TeV, with a mean value of 78 GeV, leading to the creation of one or multiple isotopes.

Shower properties

As anticipated in Section 4.6.3, it is more common to find an isotope produced by a particle in the bulk of a shower rather than being directly generated by a primary muon. The strict

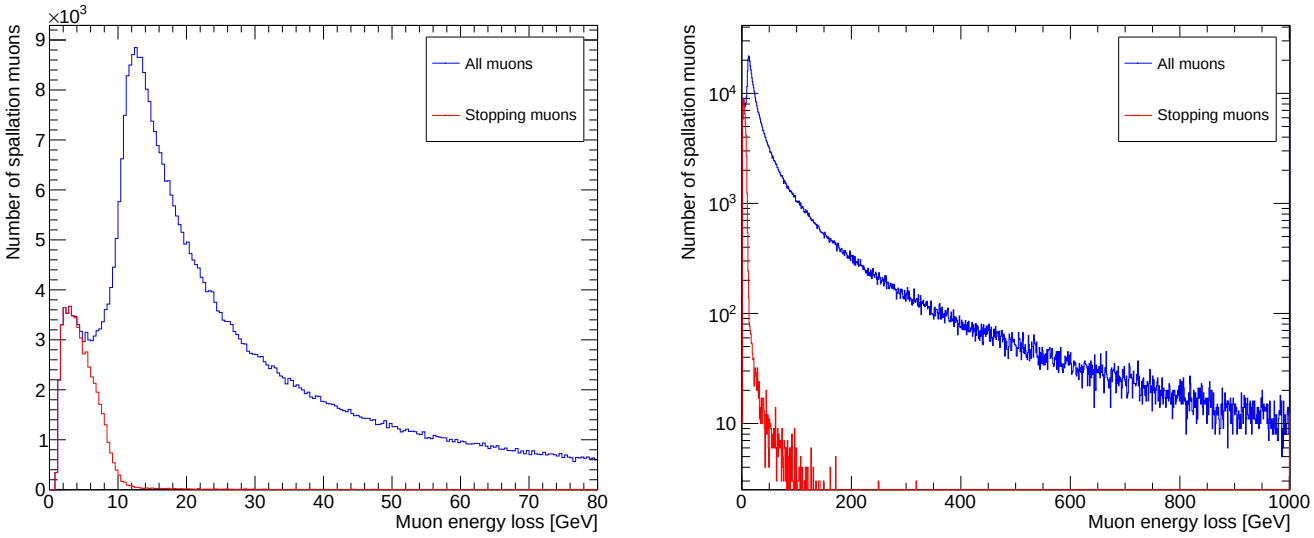


FIGURE 5.3: Energy loss in the whole OD for all muons, most of them are through-going, (blue) and stopping muons (red). Linear scale and [0-80] GeV range on the left plot, logarithmic scale and [0-1000] GeV range on the right plot. These plots refer to spallation muon sample only.

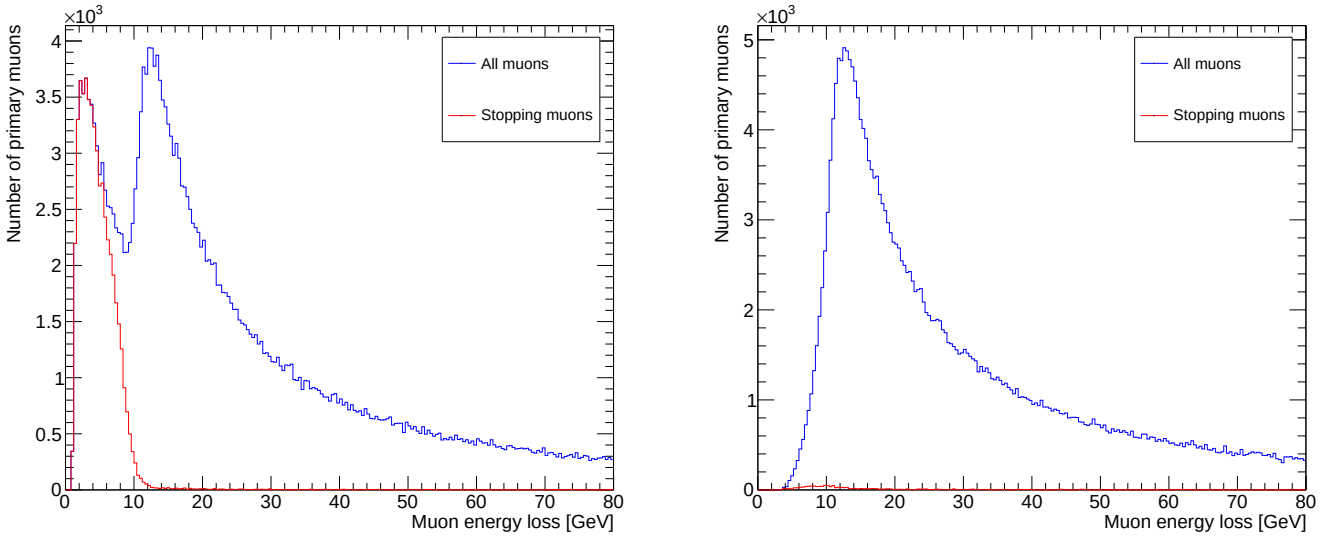


FIGURE 5.4: Energy loss in the whole OD for all muons (blue) and stopping muons (red). Left: negative muons. Right: positive muons. These plots refer to spallation muon sample only.

link between the presence of particle cascades and radioactive isotopes explains large energy deposits often associated with spallation muon travel in water (up to hundreds of GeV as shown in Figure 5.3). Once again, only less than 7% of the isotopes are generated straightly from primary muon interactions with nuclei. In the rest of the cases, muons develops a cascade of particles where daughters or subsequent secondaries are energetic enough to break up oxygen nuclei. Cases with more than two background showers induced by the same muon have never been recorded in our FLUKA simulation: the vast majority of spallation primaries only triggers a single shower, while the case of two background showers per primary μ is a factor 10^3 less probable. In particular, showers can be initiated by electromagnetic or hadronic interactions. Despite the fact that the last ones are dominant, also electromagnetic showers play a key role in spallation production: they represent about 36% fraction of the total.

The number of isotopes created per muon is strongly correlated with the particle multiplicity in showers, see Figure 5.5 (left). In turn, particle multiplicity grows with increasing deposited energy, as shows in Figure 5.5 (right). The plot presents two evident structures: the steeper one is mostly composed of hadronic initiated showers where a large number of hadrons is generated. The second structure includes a significant fraction of electromagnetic showers: these have a lower number of particles produced via inelastic interactions and shower multiplicity looks lower because the particle components of purely electromagnetic interaction products are not counted. Even in this second case, a dependency of the number of shower particles with the energy loss is observed. Figure 5.5 suggests that an higher number of isotopes is in general more likely to be associated with large energy deposits: this is expected since the probability of spallation generation grows with the energy available to produce possible isotope parents. Moreover, very large energy losses (above 2 TeV) are more likely to be associated to electromagnetic showers. These energetic cascades are expected to generate a huge amount of light in SK and therefore are easy to localize and to be removed.

Showers can start from hadronic or electromagnetic interactions and, among these, different primary processes can be distinguished, as illustrated in Figure 5.6. In fact, FLUKA simulation code associates to each interaction a parameter that corresponds to different macro categories of physics processes. Inelastic interactions, delta ray production, pair production and bremsstrahlung are the four types responsible for the creation of background showers. In the case of cosmic muon in water, FLUKA defines "inelastic" every muon-nucleus interaction where an hadron in radiated. Specifically, only inelastic interactions lead to hadronic showers, whereas the rest of the showers are directly linked to electromagnetic radiative processes. The left plot in Figure 5.6 shows the incidence of each type of first interaction while the one on the right displays the total amount of energy deposited through the relative primary process. As we expected, hadronic showers dominate in number, although they are not the most energetic ones: the mean energy deposited per muon by a bremsstrahlung photon triggering a single background showers is almost 10 times bigger than the amount released when hadrons are generated inelastically by a muon impinging on a nucleus (about 340 GeV against 41 GeV). Instead, delta rays and pair production have comparable mean energy depositions (about 33 GeV). The evidence that a large portion of very energetic showers is initiated by bremsstrahlung photon emission, together with the fact that it is less probable for all electromagnetic showers to produce multiple isotopes, given that at least one photonuclear reaction has to occur to cause spallation, explains why the number of isotopes per muon does not grow with the amount of energy deposited in the detector as fast as for the hadronic case (see figure 5.5). On the contrary, hadronic showers, which are larger in number but lower in energy, abundantly contain neutrons, pions, and gammas, all possible direct parents of spallation isotopes. Practically, the probability that hadronic showers contains one or more isotopes

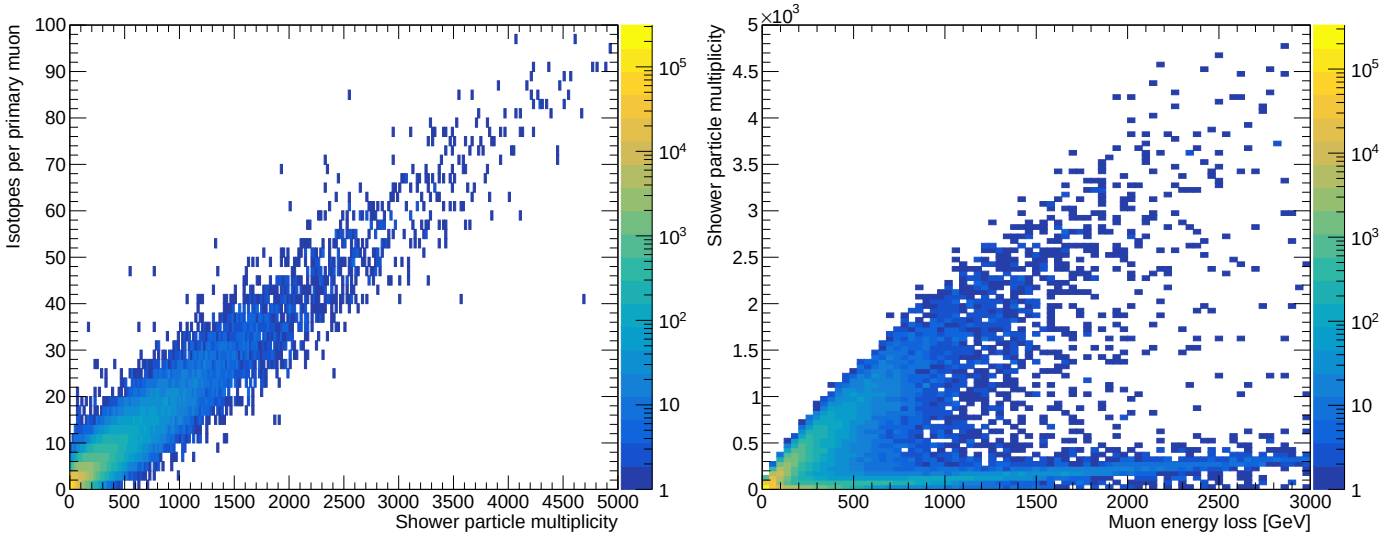


FIGURE 5.5: Left: number of isotopes per primary muon as a function of shower particle multiplicity. Right: map of shower particle multiplicity as a function of muon energy loss. The two structures represent hadronic and electromagnetic shower, the last being the less steep one. Note that the particle components of purely electromagnetic interaction products are not counted because they are not simulated in FLUKA for the reasons explained in Sec. 4.4.These plots refer to spallation muon sample only.

is more than 3 times larger than for electromagnetic cascades.

Particles in showers

The characteristics of background showers helps us to understand their particle composition, hence the main mechanisms leading to isotope production. As for the case of vertical negative muons described in Section 4.6.3, neutrons are the most abundant particle inside spallation-inducing showers, followed by pions and photons. Figure 5.7 shows the number of neutrons, pions and photons in background showers. Kaons are also included in the Figure for comparison, being the fourth most numerous particle, they are 15 times less important than the other families. In contrast with scintillators, gamma rays primarily produce invisible or stable isotopes in water. Since neutral particles are invisible in SK, neutron capture on hydrogen becomes a very important tool to detect neutrons: the 2.2 MeV gamma emitted after about 200 ns from the capture is the key signal for neutron identification. For this reason Figure 5.7 shows the number of neutron captures instead of the global amount of neutrons generated. As we will extensively discuss in this work, neutron captures lead to the emission of a delayed photon and are hence a powerful tool to identify showers: while other particles produce signals only few ns after the muon and the generated Cherenkov light can be drowned among muon signals emitted along the track, the numerous 2.2 MeV gammas from neutrons are more spaced in time and therefore distinguishable. These properties can be crucial to identify and characterize showers. The abundance of n, π and γ in showers directly reflects the fractions of isotopes produced by interactions of these particles: left graph of Figure 5.8 illustrates the percentage of isotopes generated by different parents. As expected, neutrons dominate the isotope production. On the right of Figure 5.8 isotope perpendicular distances to the muon tracks are plotted for the most common spallation parents. A critical feature, already enlightened in the previous Chapter (see figure 4.21), is the dependence of this distance on the

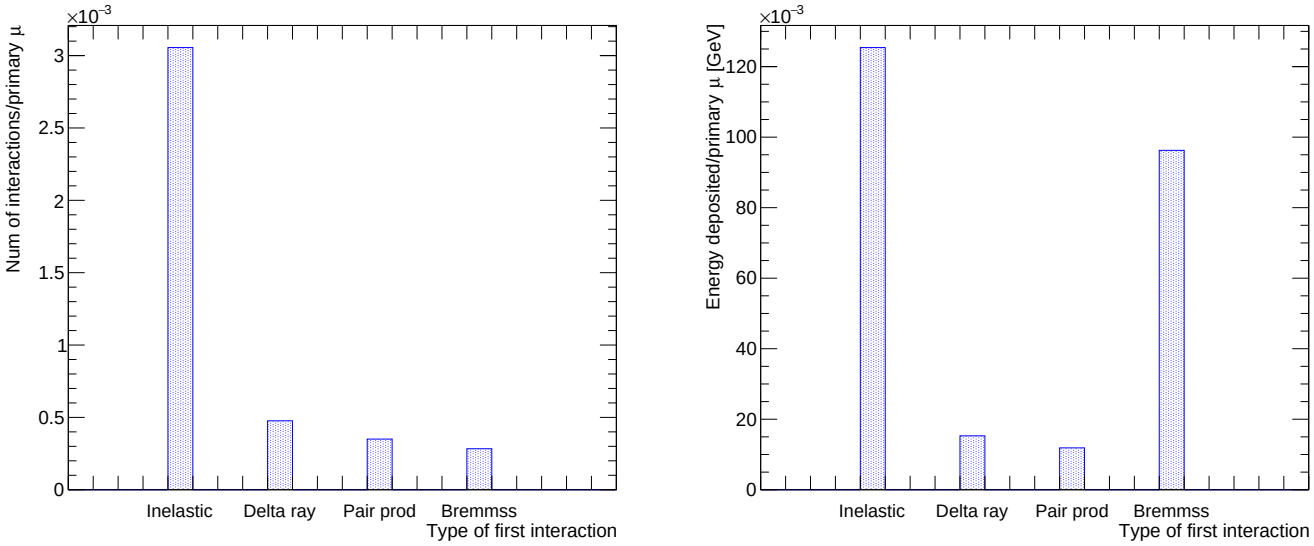


FIGURE 5.6: The plot on the left shows the total number of interactions for the four different processes described along the x axis, while the one on the right displays the total energy, per primary muon, deposited by each of them. Both plots are normalized by the number of total muon simulated. The normalization has been chosen in order to reconstruct the probability of a generic muon to generate a background shower initiated by a certain process. The labels on the x axis refer to four different types of shower injections, tagged by FLUKA code as: inelastic interaction, delta ray production, pair production and bremsstrahlung processes. These plots refer to spallation muon sample only.

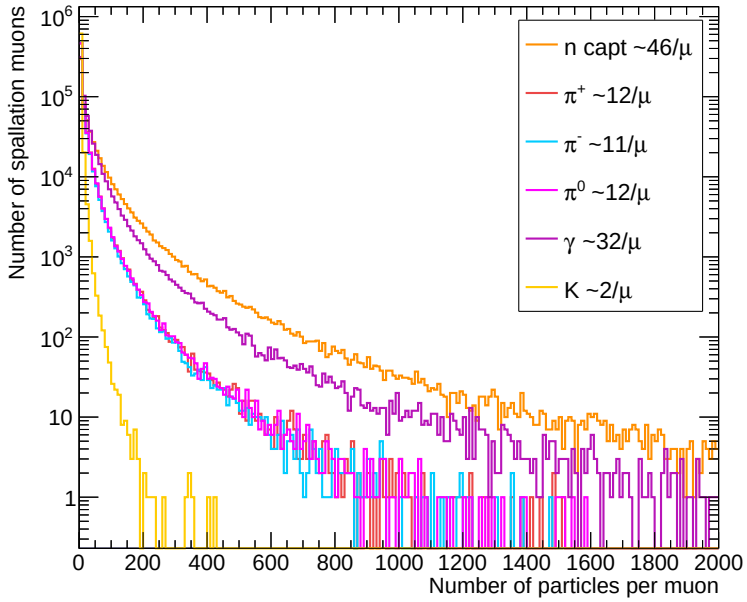


FIGURE 5.7: Number of shower particles per spallation muon for different particle types. Since neutron are only visible if captured on hydrogen, only the number of neutron captures per muon is shown. These plots refer to spallation muon sample only.

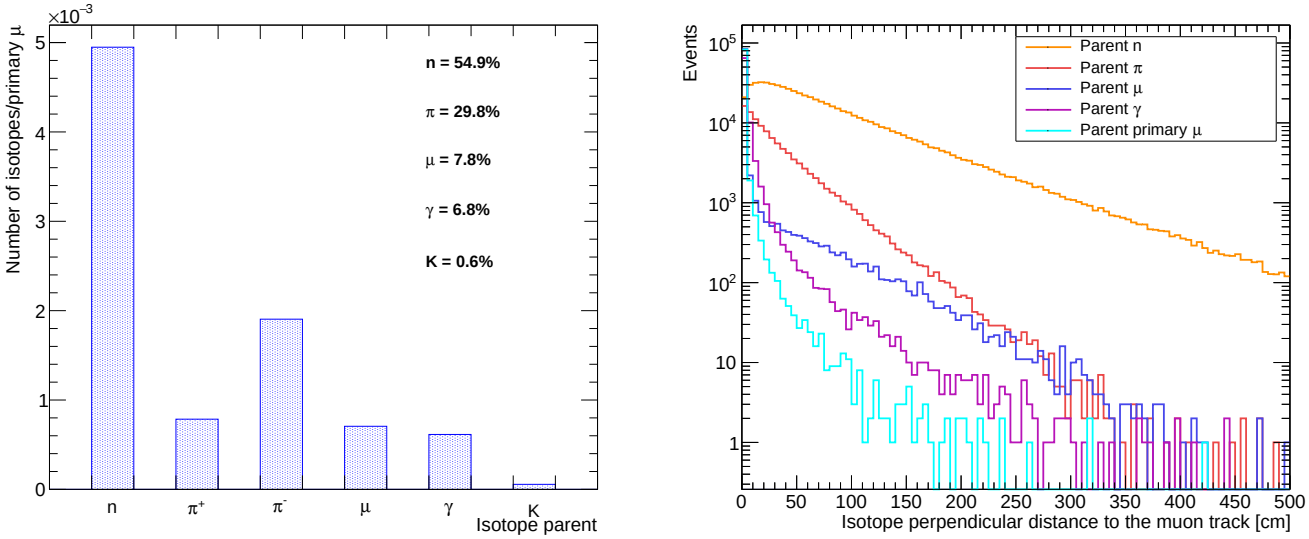


FIGURE 5.8: Left: number of isotopes generated by different type of particle (called "parent" of the isotope) normalized by the total number of muon simulated. The graph can be interpreted as a probability of a generic muon to generate a background isotope within a shower from a specific parent particle. Note that muons include both primary and secondary muons, these last ones are less than 5% of the total. The percentage related to pions includes both π^+ and π^-. Right: isotope perpendicular distance from muon track for different particle families. Primary μ are shown separately in the light blue graph, while the blue one includes both primary and secondary μ.

isotope generation process: isotopes from neutrons are found, in average, at larger distances from muon track with respect to pions or gammas. The mean value is 85, 34 and 8 cm respectively for n, π and γ. Primary cosmic μ generates isotopes very close to the muon track, about 1 cm in average, despite some more rare spallation events can be found at hundreds of cm as a cause of muon scattering inside water that can slightly deflect the track. Secondary muons come from pion decays and thus have a mean distance to the primary track comparable to what is found for pions: isotopes generated by this parent category represents less than 0.5% of the total. Isotope distance from the muon track is a very important variable in the supernova relic neutrino analysis: not only it plays an important role in spallation event rejection but it is also one of the most informative quantities among all discriminating variables. Thus, the capability of the simulation to separating the different contributions of the total transverse distance distribution can provide significant information to tune the background rejection method of low energy analysis. Furthermore, different isotopes have in general a dominant production process (details in Ref. [53]): Figure 5.9, on the left side, shows the contribution of different particles for the generation of each spallation background isotope. The mean transverse distance from a given isotope to the muon track is closely associated with the type of the main hadronic parent. This information, in combination with decay time spectra (illustrated in Figure 5.9), can help develop rejection cuts specifically tuned on an isotope by isotope basis. In addition, Figure 5.9 remarks the relative importance of each isotope type and the wide decay time range of spallation events: both isotope yields and time constants span over numerous orders of magnitudes.

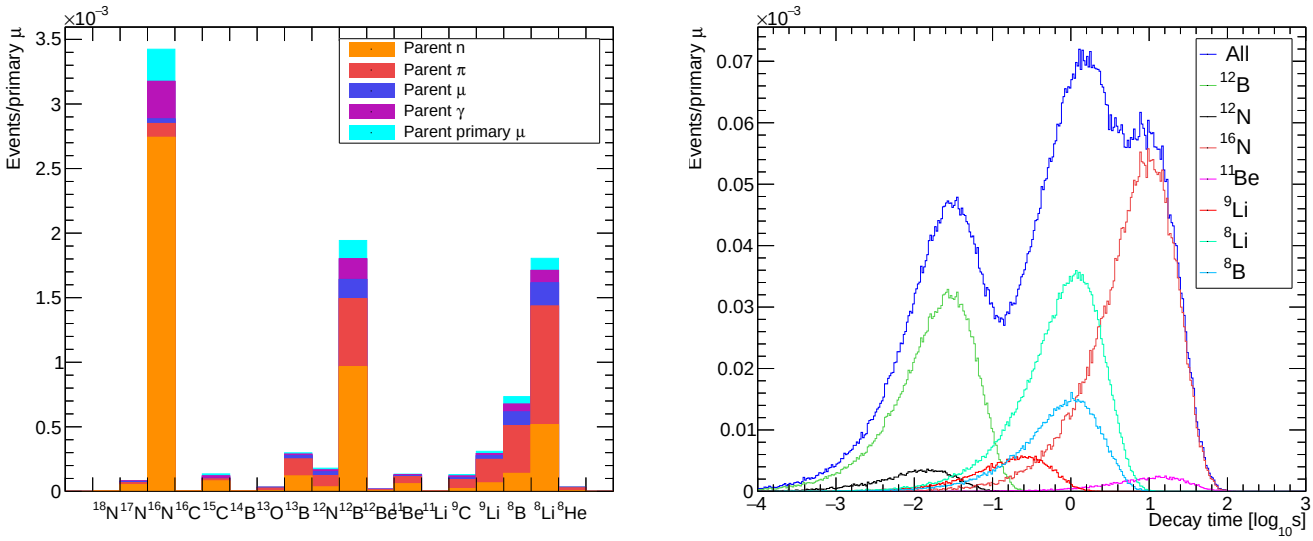


FIGURE 5.9: Left: contributions of different isotopes expressed in terms of parent particle type. The histograms are stacked and the height of different columns represents the relative importance of each isotope. Right: decay time spectrum for all the isotopes (blue line) and contribution of seven among the most abundant isotopes. For both plots normalization is made using the total number of simulated muons.

5.4 Hadron-producing muons

In order to better investigate background showers and their peculiarities, a comparison between spallation-inducing and hadron-producing showers has been made. A sample of 227×10^5 muons are generated in FLUKA and the subsequent interactions and particle production are recorded, in this case the presence of a background isotope is not a requirement. Therefore in all the circumstances where a set of two or more particles from inelastic interactions, photonuclear processes or decays are created, the parent muon falls into the category under study in this Section. As before, purely electromagnetic showers are not considered. One expects showers of this kind to be more frequent and to present lower particle multiplicity with respect to background showers. Indeed, it is found from FLUKA simulation that about 11% of muons develop generic showers with hadrons, which corresponds to about 10 events per minute of this type, to be compared with the more rare case of primary muon leading to background isotopes (less than once each minute).

Muon energy losses

The fraction of stopping muons is only 3.5% and the average energy loss is 18 GeV, almost 4 times less than the case of spallation showers. Figure 5.10 is analogous to Figure 5.3. Here, the decreasing importance of stopping muons reflects the fact that the probability of developing showers increases with muon track length. Given that stopping muons have in general shorter paths and that the fraction of muons captured on oxygen is now sub dominant, the relative contribution from this category is less than half of what was observed for muons producing isotopes. In addition, the contribution of higher energy tails in Figure 5.10 is less important than in Figure 5.3: low energy processes are the most common and the larger part of muons tends to lose a fixed amount of energy, depending on their path, through ionization (vertical muon is about 8 GeV), whereas only few GeV are emitted through radiative processes. Less than 5% of the muons deposit more than

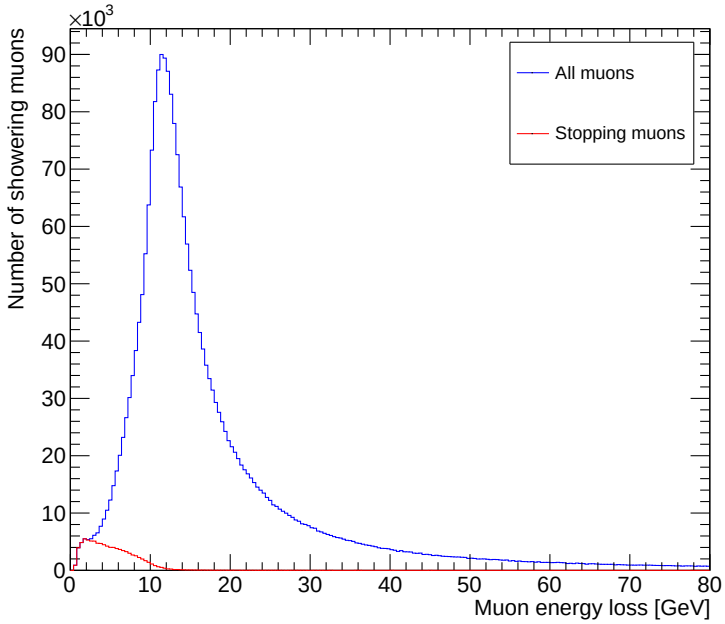


FIGURE 5.10: Energy loss all muons (blue) and stopping muons (red) passing through Super-Kamiokande outer detector. Note that this plot refers to the case of generic showers producing hadrons.

100 GeV of energy, among these there is a large fraction of muons inducing spallation showers.

Shower properties

Not only the amount of energy deposited differs from generic showers and background showers, but also the way this energy is released. Figure 5.11, to be compared with Figure 5.6, clearly indicates delta ray emission as the most probable processes through which muons generate showers, followed by pair production and inelastic interactions. Delta ray generation continues at energies where showering is not possible anymore (0.1 GeV), the energy transfer is small and treated as a continuous. Bremsstrahlung remains a minority despite its huge energy release. This result is reasonable, considering that the shape of muon energy spectrum at SK privileges low energies (see Figure 4.7), together with the fact that bremsstrahlung radiative emissions start gaining in importance only at hundreds GeV of muon energy (the value of critical energy for muons in water is 1 TeV). For hadron-producing showers, not only electromagnetic processes dominates in number, but they also represent more than 83% of energy deposited per muon, it was about 45% for background showers and mostly because of the presence of very few but energetic bremsstrahlung initiated cascades. The importance of electromagnetic energy losses was already expected from preliminary results obtained with FLUKA simulation: photons, electrons and positrons dominate the secondary particle path length spectra in Figure 4.20, indicating that they are the primary particles generated by muons is water. Although this is true for most of the primary muons, the situation drastically changes when considering spallation-inducing muons: the dominant presence of hadronic processes is one of the key features one can exploit to isolate and identify spallation products. Moreover, also shower multiplicity greatly differs when the presence of an isotope is required: Figure 5.12 confronts the number of showers, electromagnetic and hadronic separately, in

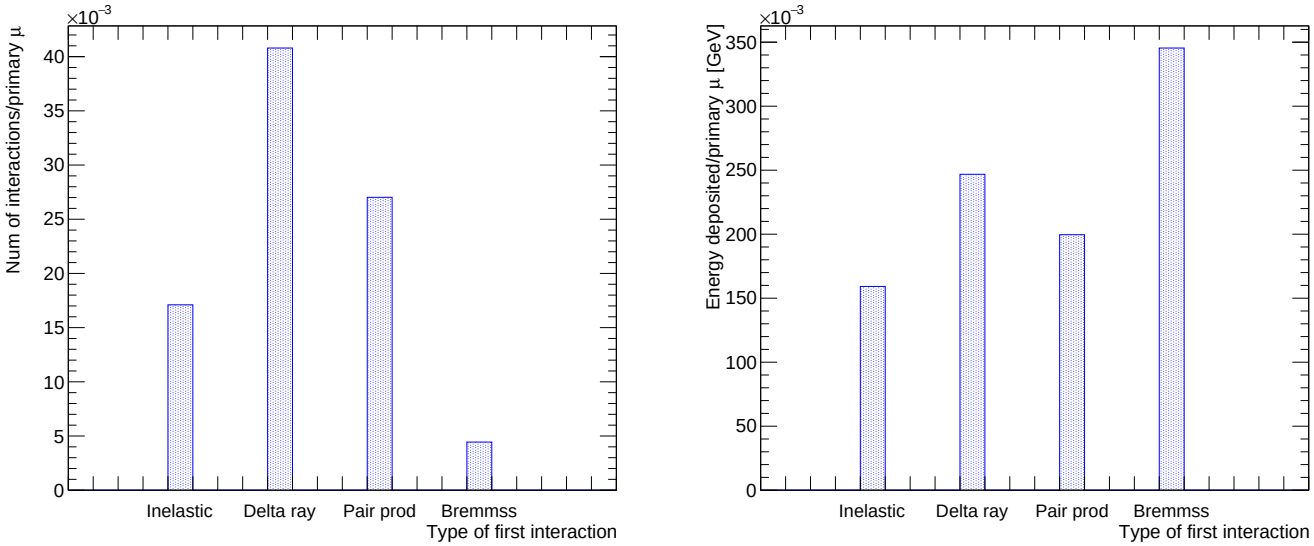


FIGURE 5.11: The plot on the left shows the total number of interactions for the four different processes described along the x axis, while the one on the right displays the total energy deposited by each of them. Both plots are normalized by the number of total muons simulated. The normalization has been chosen in order to reconstruct the probability of a generic muons to generate a background shower initiated by a certain process. The labels on the x axis refer to four different types of shower injections, tagged by FLUKA code as: inelastic interaction, delta ray production, pair production and bremsstrahlung process. Note that this Figures refers to hadron-producing muons.

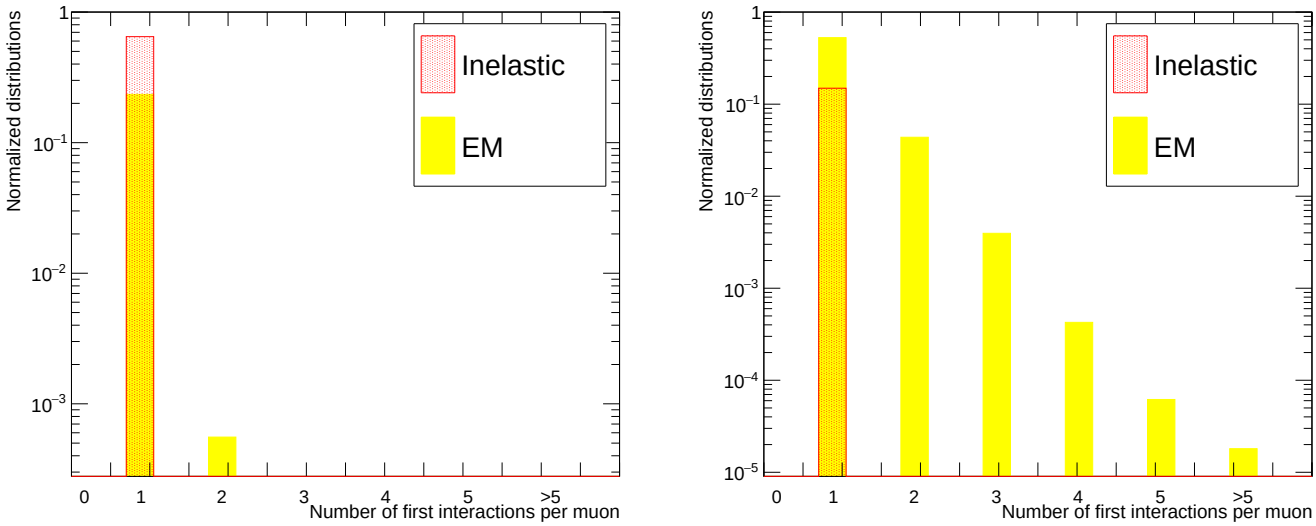


FIGURE 5.12: Number of showers per muon, initiated by inelastic (red) or electromagnetic (yellow) primary interactions. Left: showers containing at least a background isotope. Right: all showers, excluding purely electromagnetic cases. For the normalization, the number of primary muons inducing showers is used, this value is different for left and right histograms since the samples used do not correspond to the same selection and are simulated independently with a different number of primary muons. However, in both cases, the column height represents the probability a primary muon has to generate a certain number of inelastic/electromagnetic shower along its path.

the two situations analyzed. As said, the electromagnetic component is largely enhanced when we do not require a background isotope to be present. The probability of developing several showers along the same muon track increases too. This is explained by the fact that hadronic showers start from more destructive processes involving muon-nuclear interaction with high energy transfers. The probability to observe two hadronic showers along the same muon track can be considered negligible.

Particles in showers

Last but not least, particle composition is an important feature to be studied for generic showers in comparison to background ones: the results from FLUKA simulation are illustrated in Figure 5.13. Neutrons are very widely present, followed by pions and photons. This hierarchy reflects what we found in Fig. 5.7. However, the very crucial difference is the multiplicity: the number of particles in Fig. 5.13 for each different family is suppressed by a factor 10 compared to Fig. 5.7. This very specific property of showers containing background isotopes can be exploited to identify spallation: as mentioned, the presence of numerous neutron capture signals, distinguishable from muon light thanks to its delayed capture time of 200 μs, is a powerful hint of the presence of a a spallation event in the associated shower.

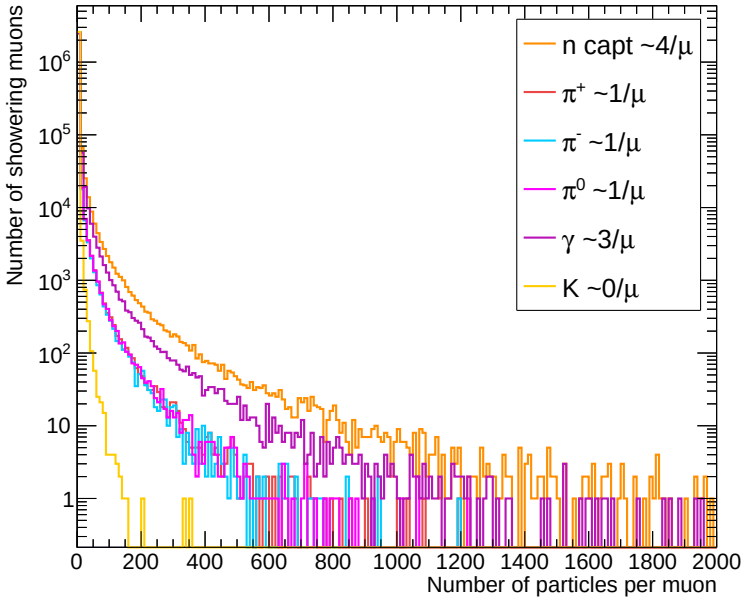


FIGURE 5.13: Number of shower particles per primary muon for different particle types. Since neutron are only visible if captured on hydrogen, only the number of neutron capture per muon is shown. This plots refers to hadron-producing muon case, namely the presence of a background isotope is not required.

5.5 Identifying spallation showers

Results presented in this Chapter are crucial to understand spallation mechanism, the type of particles created in the same shower as a radioactive isotopes, the properties of the cascade and muon energy losses. All these informations can be used to tune better rejection cuts for spallation removal: the evidence that spallation inducing muons show peculiar properties make them distinguishable from average muon behavior. In the following main properties of spallation muons and related showers are summarized:

- Spallation processes are rare: the large majority of muon crosses the detector without inducing radiative emissions and only about 0.5% lead to the creation of spallation isotopes. In fact, it was shown in Ref. [53] that, in contrast with scintillating detectors, the probability of generating unstable isotopes in Super-Kamiokande is low compared to the probability of creating stable or invisibly decaying isotopes.
- Isotopes are frequently found in the bulk of showers rather than being produced directly by the primary muons (only 8% of all muons produce isotopes, only 3% of the isotopes are directly produced by primary muons). They are not generated uniformly along the muon track but they are located and clustered in specific positions, linked to shower generation point.
- Background shower energy is in average 4 times higher than the injection energy for generic showers. Moreover, it is about 10 times greater than the average ionizing energy deposition (if compared with vertical mip muons whose energy deposited through ionizing processes only is about 8 MeV).
- The number of isotopes generated along the muon track increases with the amount of energy deposited and the abundance of particles within the shower.

- The majority of isotopes are part of hadronic showers, the rest, about 36%, are within electromagnetic cascades. In both cases shower charged particles produce Cherenkov light that can be detected in SK.
- It is rare to observe more than one spallation shower along the muon track: this means that isotopes and shower particles are often clustered within few meters (typical dimension of showers in water).
- Neutrons have a particular importance, being the most abundant particles and the main parent of radioactive isotopes (spallation muons produce in average about 46 neutrons per track). Since neutron captures give a delayed signal that can be distinguished from muon light and their mean path in water is longer than for pions and gammas, a study of neutron clusters can provide a good indication of the presence, and the position, of spallation isotopes.

Although these characteristics largely correspond to the average properties of background showers, it is important to underline that there are still cases where this does not hold true: it is still possible to find isolated isotopes, very low energy or low particle multiplicity background showers. Nonetheless, they represent a minority.

Some of the properties listed above are already exploited in SK low energy analysis to localize and remove spallation isotopes: for example, as discussed in Sec. 4.6.4, the amount of charge deposited per 50 cm track segment is studied and the location of its maximal value, Q_{peak} position, gives an indication of the presence of a shower. Since the Cherenkov light emitted per unit of length by a relativistic muon ionizing the medium is constant, the excess of energy must be related to local radiative processes and creation of showers. This information, combined with other variables meant to describe time and space correlation between a low energy candidate event and a muon, are the foundation of the current spallation rejection method for SRN analysis. The limitation of the use of Q_{peak} position as discriminating variable can arise firstly from isotopes created in showers with very little light (or directly by primary muons $\sim 3\%$) for which the charge spectrum does not present a defined peak. Moreover there exists an uncorrelated component formed by isotopes associated with the wrong shower: it can happen in fact that while the isotope is generated in a visible cascade, a bigger charge deposition is produced elsewhere along the muon track and a wrongly correlated Q_{peak} position is associated to the radioactive decay, these events populate the tails of Ll variable distribution in Fig. 4.27. This uncorrelated component is estimated to represent a 10% fraction of all the events [98]. To improve the current rejection efficiency, other features highlighted in this Chapter can be used: the next Chapter will discuss how the abundant presence of neutrons within spallation showers can be a powerful method to identify particle cascades with isotopes and reduce background signal in SK.

Chapter 6

New technique and simulation for spallation background removal

The FLUKA-based simulation framework presented in Chapter 4 is the first complete calculation modeling muon propagation and shower generation in water able to accurately reproduce SK data. Super-Kamiokande conditions and detector response are replicated from cosmic muon fluxes to PMT signals and reconstructed physical variables. We demonstrated the ability of this simulation to successfully infer distributions of the main discriminating variables used in supernova relic neutrino and solar analyses to reject spallation background. The agreement observed between MC and data constitutes a solid proof of the validity of the simulation framework and enables us to exploit the simulation results for a deeper understanding of spallation mechanisms and for making predictions. Chapter 5 investigates the main properties of muon energy losses and generation of radioactive isotopes suggesting new techniques to locate muon-induced hadronic showers and efficiently reject spallation backgrounds. Among these, the large neutron multiplicity associated with spallation showers represents a powerful tool for isotope identification: this Chapter will describe how to use neutrons to better remove spallation and how the FLUKA simulation is able to accurately model neutron geometrical properties and their behavior in SK. Moreover we will demonstrate the capacity of FLUKA calculation to predict isotope yields within hadronic model uncertainties. Moreover we will prove that, thanks to neutron cloud information, it is possible to build spallation-rich samples while keeping the relative fractions of the most abundant isotopes stable. These sample could be very helpful for future studies, for example to model the impact of energy threshold on different spallation isotopes. Simulation results are crucial to compute yields and cut efficiencies for individual isotopes, since predictions from data-driven method are typically averaged over all produced isotopes. The agreement between data and simulation, both for isotope yield predictions and modelization of shower pattern, is crucial for future studies in the recently upgraded phase of Super-Kamiokande, where Gadolinium has been dissolved in pure water: a detailed simulation of muon induced spallation processes is a unique instrument for developing future analysis strategies. In Super-Kamiokande, it has never been possible before the construction of this calculation framework to rely on accurate Monte Carlo predictions for the major background in the low energy analysis. Empirical rejection techniques have shown important limitations due to the complexity of locating the spallation events, therefore the dead time introduced by standard cuts remains high when a large fraction of radioactive isotopes is removed. The new approach described in this Chapter (the “neutron cloud" method), together with precious insights from simulation will help Super-Kamiokande, in its new phase SK-Gd, to gain signal acceptance and approach the sensitivity required for important discoveries like the first observation of supernova relic neutrinos. Promising results have been already observed when introducing neutron-based methods to locate spallation background both in current solar and SRN analysis (for SK-IV data) and a great improvement of this cut performances

is expected for SK-Gd, as presented in Chapter 7.

6.1 Why are neutrons so important?

Previous spallation rejection methods in SK, as described in Section 4.6.4, are based on likelihood cuts defined with probability density functions of a set of discriminating variables [22, 59]. These quantities rely on identifying space and time correlations between the reconstructed muon track and the candidate isotope vertex and the prompt light deposition along the muon path. The pattern of emitted light is the only variable related to a possible muon-induced shower leading to the production of spallation events: dE/dx spectrum can in fact present a peak in the vicinity of the shower initial vertex, if an energetic cascade develops along the muon track. Other properties of the shower, like its extension and particle composition, are not taken into account in the likelihood algorithm.

During the last period of SK-IV, the newly developed online trigger WIT (described in Sec. 2.4.2), made possible the detection of captured neutrons, namely the most produced hadrons in muon-induced showers. This opened a door for a new spallation rejection method based on shower identification through neutron multiplicity. If we just consider the short period of SK-IV where WIT trigger was available to tag neutron captures (only 388 days over 2790 days of total livetime), the application of this new algorithm to the current solar neutrino analysis showed an 12.6% increase of signal efficiency, which correspond to about a year of detector running time. Appreciable improvements were also seen in SRN analysis. However WIT trigger was set up only in 2015, at a late stage of SK-IV, and the low lifetime of the system is a critical limiting factor to the algorithm performance. Another problem comes from the weakness of the signal generated after neutron capture: as we discussed in Section 2.4.2 WIT trigger efficiency for 2.2 MeV gamma rays is around 13% over the entire detector and 17% in FV. The situation is very different in SK-Gd: an enhanced detection efficiency due to gadolinium doping drastically increases the neutron visibility making neutron-based algorithms a key strategy for spallation rejection. Before describing in more details the neutron cloud method, we want to spend some words to summarize why these particles are so crucial for spallation removal and more generally shower identification.

Neutrons are produced in abundance by muons interacting with nuclei in water. Unlike electrons, positrons and gammas, which are created almost uniformly all along the muon track, neutrons tend to populate showers and therefore are clustered in limited regions of space defined by the shower extension. We obtained an average of 46 neutrons per spallation muon by FLUKA simulation (Fig. 5.7), with some rare cases of cascades containing clusters of thousands neutrons. Despite the low detection efficiency of 2.2 MeV gammas, their high multiplicity and delayed light emission (about 200 μs after their creation) make it still possible to identify clusters of neutrons. Moreover neutrons can travel in water for a few meters (95% of them are contained in a radius of 3m from the muon track). They are generated with energies going from a few GeV to tens of MeV, depending on the interaction process. Since their capture cross section (both on H and Gd) is significant only at very low energies, $\mathcal{O}(10)$ eV [128], they thermalize though successive scattering processes before being captured, making their total path in water the longest among all other particles in showers. The path length distribution for neutrons directly reflects the distance from the primary muon track where the neutron-induced spallation isotopes are produced, as illustrated in Figure 5.8: spallation events whose parent is a neutron are found in average 85 cm away from the muon, more than twice the average distance of isotopes from any other parent particle. This properties of neutrons make them particularly useful to estimate the dimensions of a given shower: since they travel

more than any other component, and their multiplicity is in general higher, the maximum distance of the most peripheral neutrons within the cascade is a good evaluation of the shower extension. Another very important point to consider is the fact that more than half ($\sim 55\%$) of the isotopes visible in SK are generated by processes involving neutrons: as an example, the most abundant isotope created in SK, ^{16}N, is mostly produced by neutron interaction with oxygen nuclei (about 76% of the times). With all this in mind, one can tag neutrons and study their spatial distribution in order to locate a region in space where a spallation-inducing shower could be contained. After that, applying customized spherical and elliptical cuts around this neutron cluster (“neutron cloud") instead of traditional cylindrical cuts would help increasing the signal acceptance and better locate where the isotopes are produced.

6.2 Tagging spallation by identifying neutron clouds

As we discussed, neutron clouds associated with muon-induced showers have space and time correlations with possible spallation isotopes within the same cascade, therefore their identification gives very powerful information for spallation background reduction. Neutrons need to be accurately recognized and localized so that their position and multiplicity can be used to tune customized selection cuts on the possible spallation candidates. Given the low signal emitted after neutron capture, WIT trigger is used to select neutron candidates. Subsequently, refined requirements on the observables related to the candidate are applied to reliably discriminate neutron clusters from background noise. The approach used to tag neutron clouds, related variables and selection is presented in the following. In the next Section a comparison between SK-IV data and FLUKA-simulation outcomes will be treated: the good agreement not only gives precious intuitions on underlying mechanisms, but also motivates the use of simulation to optimize future analysis at SK-Gd. Neutron clouds and simulation insights will also be an important ingredient for isotope yields estimation, described in Section 6.5.

Neutron cloud selection

Quality cuts need to be applied on neutron candidates identified by WIT in order to remove events possibly due to radioactivity, PMT flashing and dark noise fluctuations. This selection ensures that neutron clouds are composed of well reconstructed neutrons. A set of variables is used to make quality cuts:

- *Reconstructed vertex:* it is typical in SK low energy analysis to require the signal vertex to be inside the fiducial volume, since radioactivity from the rock or PMT material often occurs close to the detector wall.
- *Reconstructed energy:* the 2.2 MeV gamma emitted when a neutron is captured on H converts into an electron generating Cherenkov radiation. The subsequent reconstructed energy, E_{rec}, is expected to be lower than 5 MeV, if accounting for light yield fluctuations.
- *Position and direction goodness:* the quality of the reconstructed position and direction is expressed by two variables defined by Eq. 2.4 and 2.5. Well-reconstructed events have higher goodness values. Refined cuts on $g_{vtx} - g_{dir}$ space have been tuned and a weight of 0, 1 or 2 is associated to each neutron according to the goodness values. Therefore “good neutrons" will contribute twice in the neutron cloud distributions while poorly reconstructed ones are removed.

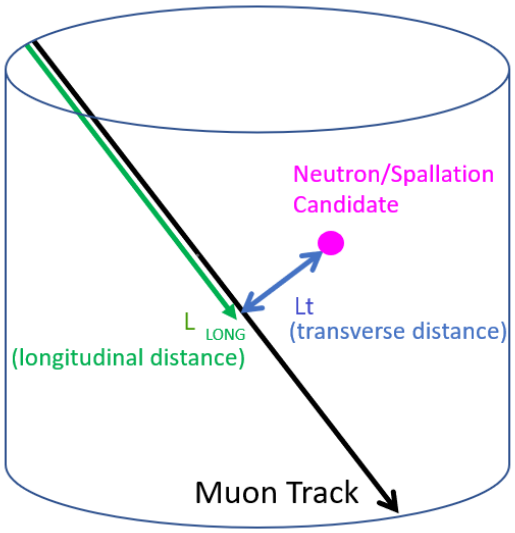


FIGURE 6.1: Graphic illustration of L_{LONG} and L_t.

- *Time correlation with muon:* since neutrons inside muon-induced showers are very likely to show correlations with parent muons, the time difference between neutron candidates and the cosmic parent, Δt, is expected to be of the order of the characteristic capture time of neutrons in water ($\sim$ 204.8 μs). In particular, the large majority of neutrons are assumed to be found within 500 μs after the parent muon. The width of this time window, given the 2 Hz rate of muons in SK, allows to unambiguously associate each neutron cloud to its parent muon. Moreover, events observed less then 20 μs after the muon are neglected to remove PMT afterpulsing signals: delayed pulses due ionization of the residual air in the PMT between the cathode an the first dynode, rare for low energy signals but more frequent when a huge amount of light is created, by e.g. a muon.

- *Space correlations with muon:* transverse distance, L_t, is defined as the distanced between the neutron candidate reconstructed vertex and the muon track, while L_{LONG} is the longitudinal distance from the muon entry point, parallel to the primary track. These two variables, graphically shown in Figure 6.1, characterize the shape and spatial extension of the neutron cloud. All neutron candidates are required to have $L_t < 5$m: we expect the cloud to be contained within a few meters from the track. In parallel, L_t^2 is often used to have a flat phase space, given the cylindrical symmetry of SK. Instead of directly using L_{LONG} it is more informative to consider L_{LONG_i} - $<L_{LONG}>$, where the average longitudinal distance of the cloud is subtracted from L_{LONG_i} for each single ith-neutron.

The observables listed above are used to define good neutron candidates and use them to construct the neutron cloud. The spatial correlation between the neutron cloud and the isotope vertex in turn allows to define spallation rejection cuts. To summarize, a well-reconstructed neutron must lie in the FV with Δt between 20 μs and 500 μs, $E_{rec} < 5$ MeV and $L_t < 5$ m. Then, the fit quality variables , g_{vtx} and g_{dir} defined in Eq. 2.4 and 2.5, attribute a weight to each neutron. When two or more good neutrons are found we consider them to form a neutron cloud, and variables like cloud barycenter position, average transverse and longitudinal extension and multiplicity can be computed. These neutron cloud observables define a spherical or elliptical region in space which corresponds to the area where the muon-induced shower develops: every low energy event within this

region is tagged as spallation and is rejected. In the following Section neutron clouds and their properties will be described and observable distributions will be compared with FLUKA-based simulation outcomes. Section 6.4 will summarize the spallation rejection cuts used in SK-IV solar and SRN analyses based on correlations between neutron clouds and isotope decays.

6.3 Simulated neutron clouds

6.3.1 Simulation samples

The goal is to reproduce, as closely as possible, neutron clouds reconstructed from WIT data. These reconstructed clouds include candidate events corresponding to real muon-induced neutrons but also spurious events with a particular hit pattern that can mimic neutron captures (from radioactivity, dark rate fluctuations...). Therefore, a candidate incorrectly tagged as neutron (also called "fake neutron") can be accidentally associated to a muon which, in reality, does not lead to the generation of any shower. Even if several cuts are applied to select good neutrons and limit the occurrence of mis-tagged events, a fraction of fake neutron is still contaminating the neutron cloud sample. These fake neutrons must be taken into account in the simulation when selecting the events to be compared with data. Specifically, in addition to simulating showering muons with FLUKA and SKDetSim, we need to consider non-showering muons associated with SK noise fluctuations and background events. For this reason, we modeled neutron clouds using the following two samples:

- *Hadron-producing muons:* largely described in the previous Chapter, this sample includes muons leading to the development of a shower where inelastic interactions generate gammas, pions, neutrons, kaons or additional hadrons. This shower could also include a spallation isotope. All the particles in the shower are simulated with FLUKA, taking MUSIC muon distributions as inputs. Muons that induce showers containing spallation isotopes represents a small fraction of this sample (about 5%). Among spallation-inducing muons, about 90% of the showers contain at least one neutron and more than 80% have more than one. We used 2.7×10^5 simulated muons to construct this sample, among these 97.4% are reconstructed by MuBoy as through-going[1].

- *Electromagnetic only muons:* this sample includes all muons that were not classified as hadron-producing. As described in Sec. 5.2, this sample includes: both minimum ionizing muons, which simply cross the detector, and muons only interacting electromagnetically in the detector (possibly inducing EM showers). Indeed, for spallation studies, both categories of muons can be treated in the same way. To construct this sample we simply inject muons in SKDetSim with input energies and directions from MUSIC and entry point sampled as described in Section 4.3.1. We require muon-nucleus interactions to be deactivated in SKDetSim generation. 7.5×10^5 simulated muons are used to construct this sample, among these 86.8% are reconstructed by MuBoy as through-going muons. The fraction of through-going muons is lower than in the hadron-producing muons sample since stopping muons or corner clippers are less likely to develop showers.

The first sample contains real neutrons as well as mistagged ones while the second sample only contributes to the fake neutron fraction. Since the exact fraction of hadron-producing

[1] The remaining 3% are mostly stopping muons. A few muons are also identified as bundles with only one fitted track.

muons in data is unknown and given that, due to large uncertainties related to hadronic models and muon-nucleus interactions in FLUKA, this fraction cannot be precisely obtained from simulation, and should be extracted from data. The procedure used to extract the EM-only muon fraction is described in the following Section. Obtaining a good agreement between simulation results and data is important to estimate the performances of neutron clouds algorithm and motivate the use of FLUKA-based simulation to optimize spallation cuts for SK-Gd.

6.3.2 Neutron tagging procedure for simulation outputs

In Super-Kamiokande, neutron cloud search is performed with the Wideband Intelligent Trigger system described in Chapter 4. In order to be able to compare distributions for data and simulations, FLUKA results need to be adapted, to mimic the detector output, as described below.

1. FLUKA results describing spallation processes are injected into SKDetSim, as described in Sec. 4.2, where detector effects as well as Cherenkov light production and PMT hits are precisely simulated. A muon is generated at the beginning of the first MC event window (40 μs long), together with all shower particles except neutrons. Information related to possible isotopes produced by spallation interactions are recorded separately. After a muon event, the 2.2 MeV gamma-rays from neutron captures are simulated in SKDetSim and each of them is stored in a separate event. As described in Section 4.2 of Chapter 4 the background for neutron captures is not estimated from SKDetSim models but taken from regular measurements and injected into the simulation output.

2. Reconstruction algorithms are applied to muon signals, taking into account the Cherenkov light from shower particles.

3. At this point SKDetSim output files, containing muons and neutrons hit pattern and charge informations, have to be reorganized to mimic WIT format: the neutron search algorithm is designed to process raw data organized in chunks of 17 μs. Thus, the simulated detector output for the muon, all its shower particles, and neutron captures, is stored in a series of 32 WIT-events, each of them 17 μs long, for a total time window of about 500 μs. The PMT hits of the muon and its shower are located at the beginning of each 500 μs window, in the first $\sim$ 2 μs. Neutrons are distributed in the subsequent WIT events according to their capture times. More than 95% of the neutrons are captured within this 500 μs windows. Finally, PMT hits from random trigger data taken throughout the SK-IV period are injected into the simulation results in order to model the background for neutron identification.

4. The simulated WIT output is now reconstructed using the same standard algorithms as with data, described in Sec. 2.6. A series of candidate neutron events are identified and BONSAI reconstruction fitter characterizes them by their reconstructed vertex and time.

5. Neutrons whose reconstructed vertex is within 5 m from the muon track are selected. For each muon event we obtain a structure containing:

    - True neutron clouds: all the true information related to muon, shower particles and neutrons from MC simulation.
    - Reconstructed neutron clouds: MuBoy reconstructed variables for the initial muon, number of tagged neutron candidates per muon, with their positions, timings, energies and fit quality variables.

Selection requirements	Efficiency values (%)
WIT trigger	13.3 ± 0.1
$20\mu s < \Delta t < 500\mu s$	89.4 ± 0.4
L_t <500 cm	76.7 ± 0.3
FV	93.3 ± 0.5
Fit quality	76.9 ± 0.4
$E_{rec} < 5$ MeV	99.9 ± 0.6

TABLE 6.1: Neutron tagging efficiencies for different selection criteria: WIT trigger identifying a neutron candidate, timing cut to remove the primary muon signal and the associated after-pulse, proximity to the muon track, fiducial volume, goodness-of-fit cut, and energy cut. We consider a simulated neutron to be tagged if its true capture time lies within 50 ns and 50 cm of the reconstructed neutron candidate.

6. Finally, selection cuts are applied, as in data, to identify well-reconstructed neutrons.

We remind that only through-going muons are considered: stopping and short track muons have a negligible probability to develop a shower while muon bundles are not yet implemented in the simulation and thus, for the purpose of comparison, are also excluded from data. Bundles make up to 5% of the total number of muons and only one of them is expected to produce a shower, therefore for moderate-size bundles the neutron multiplicity is not expected to change.

Once the neutron tagging procedure is finalized we can apply it to selected FLUKA outputs in order to obtain simulated neutron cloud samples which can be directly compared with the ones from data. The details of the simulation samples we used are described in the following.

6.3.3 Comparison of MC and Data for neutron clouds

The goal is to identify neutron clouds in data and simulation, and compare their properties. In addition, MC results are a key tool to evaluate the performance of neutron tagging algorithms in future analysis. As a first step, the variables described in Section 6.3.2 are extracted both from data and MC and neutron clouds are reconstructed.

Neutron tagging algorithm performances

Trigger efficiency and neutron reconstruction accuracy can be evaluated by comparing the true and reconstructed vertex information in simulated events. If a candidate signal, in the hadron-producing muon sample, is found within 50 ns and 50 cm from a simulated capture, it is considered to be a well-reconstructed neutron. Efficiencies associated with WIT trigger and selection requirements described in Sec. 6.3.2 are illustrated in Table 6.1. The final neutron identification efficiency is around 6.5%. The low value is a consequence of the weakness of the neutron capture signal, often indistinguishable from SK noise (dark rate and radioactivity). An estimation of the mistag rate, namely the number of fake neutrons per primary muon, is extracted by counting the tagged neutrons associated with a muon that does not produce a shower with hadrons, namely using the EM-only muon sample. We find a rate of 0.044 ± 0.001 tagged fake neutrons per EM-only muon. Conversely, the rate of real neutrons tagged per hadron-producing muon is 0.24 ± 0.01.

	Muon rate [Hz] if:		
	0 n tagged	≥ 1 n tagged	≥ 2 n tagged
Hadr. producing muons	0.052	0.0091	0.0023
EM only muons	1.53	0.069	0.0019

TABLE 6.2: Table with muon rates in SK if 0 neutrons are tagged, at least 1 neutron is tagged and at least 2 neutrons are tagged (namely a neutron cloud is found). Muon rates are expressed in Hz and separated for hadron-producing and electromagnetic only muons, where the EM fraction is extracted from χ^2 minimization as described in the main text. Only through-going muons are considered.

Neutron cloud shapes

Both for hadron-producing and EM-only muon samples, neutron-related variables are extracted and neutron clouds are constructed asking that two or more neutrons satisfy requirements described in Sec. 6.3.2. In the simulation these reconstructed clouds represent only about 5% of all real neutron clouds, due to the low neutron identification efficiency.

Since we expect that the majority of muons traverse the detector without inducing showers, the contribution of fake neutrons contaminating the cloud samples is non negligible despite the low neutron mistag rate. Estimating contribution from these fake neutrons requires to determine the fraction of muons not leading to hadronic showers (EM only muons), which, as we said, is determined by nuclear effects that are difficult to model accurately in simulation. Therefore the main variables defining neutron cloud shape, L_{LONG} and L_t, are constructed by varying the relative contributions of the two simulated muon samples. Subsequently, we evaluate the fraction of EM muons by fitting the predicted L_t and L_{LONG_i} - $<L_{LONG}>$ distributions to the SK-IV data. Figures 6.2 and 6.3 show longitudinal and transverse distributions obtained by combining the simulation outputs for hadron-producing and EM muons with different relative weights. In the same Figures, data distributions are superimposed for comparison. For convenience, the relative composition is defined in terms of EM muon fraction with respect to the total number of muons considered, namely:

$$\text{EM fraction} = \frac{\mu \text{ in EM only sample}}{\mu \text{ in hadron-producing sample} + \mu \text{ in EM only sample}} \tag{6.1}$$

This value varies from 0.985 to 0.945 in the Figures. Despite the very high ratio of EM muons, the distributions are dominated by neutrons generated by showering muons: after all selection cuts, 1 neutron out of every 400 EM muons contributes to the distributions in Figures 6.2 and 6.3 while there is about 1 neutron for every 6 showering muons. A separate χ^2 fit is performed for L_t and L_{LONG} to evaluate the robustness of the comparison. Results of the minimization are shown in Figure 6.4. The EM fraction for which simulation distributions best fit data is 96% (χ^2/NDF = 1.1) and 97% (χ^2/NDF = 1.6) for L_t and L_{LONG_i} - $<L_{LONG}>$ respectively. The values obtained are compatible with each other, but larger than the FLUKA prediction of 89%. For the following analysis a fraction of 96.5% is adopted and the difference between the fit results for the two observables is treated as a systematic uncertainty. Table 6.2 shows the resulting muon rates for different tagged neutron multiplicities, when EM fraction is set as 0.965.

Figure 6.5 shows neutron cloud variable distributions from simulation. Contributions of the two different muon samples are displayed. Here the uncertainties combine both the statistical uncertainties and the systematics associated with the different optimal EM

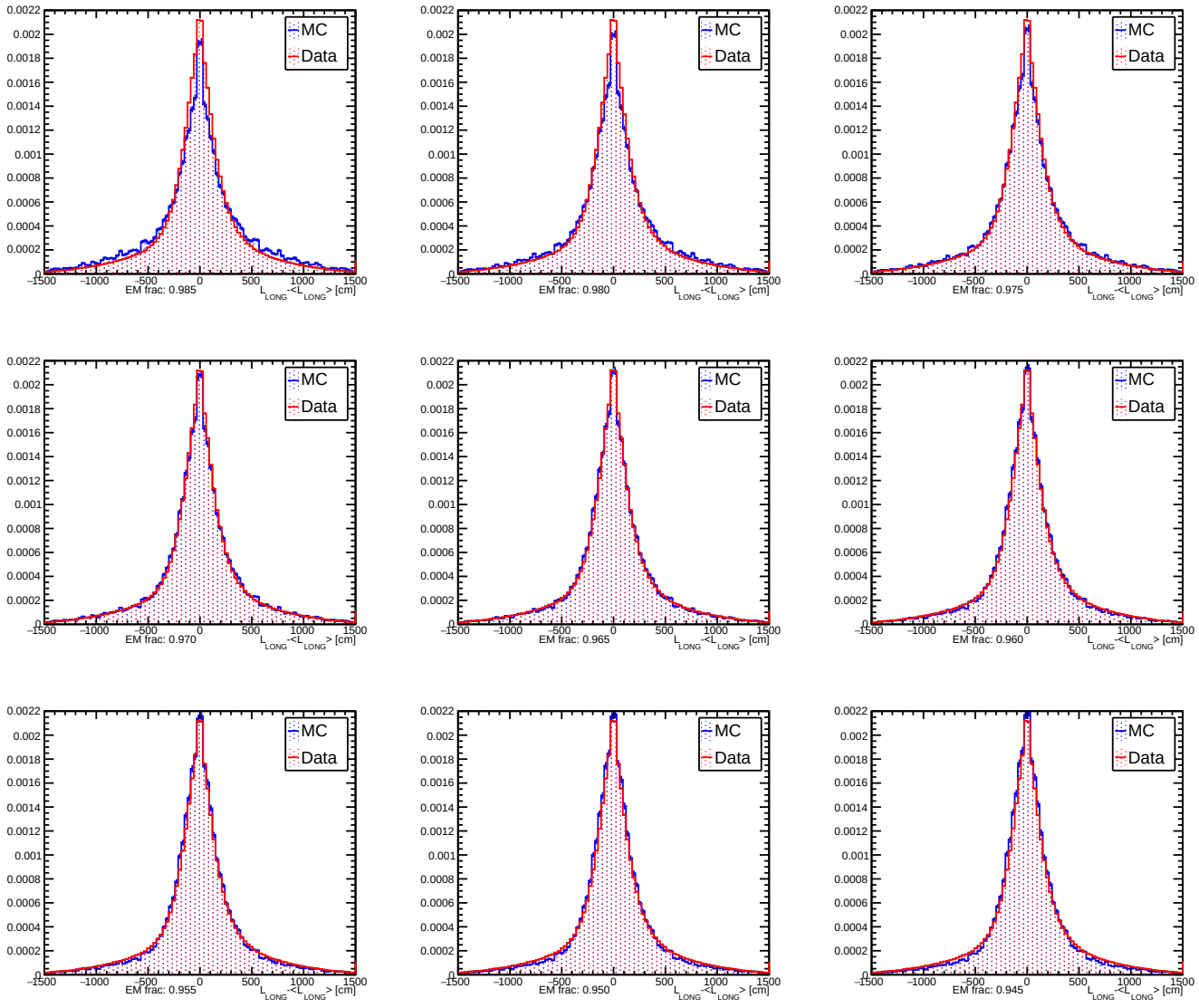


FIGURE 6.2: Longitudinal distance distributions obtained from combining hadron-producing muons and EM muon simulation outputs with different fractions, in blue. In the same Figures data distributions are superimposed for comparison, in red. The label "EM frac", indicated along the x axis, represents the fraction of muons in the EM sample with respect to the total number of muons, as described in Equation 6.1. The distribution is presented as L_{LONG_i} - $<L_{LONG}>$, namely longitudinal distance of each neutron in a cloud after subtraction of its mean value. All the plots are normalized to their integral.

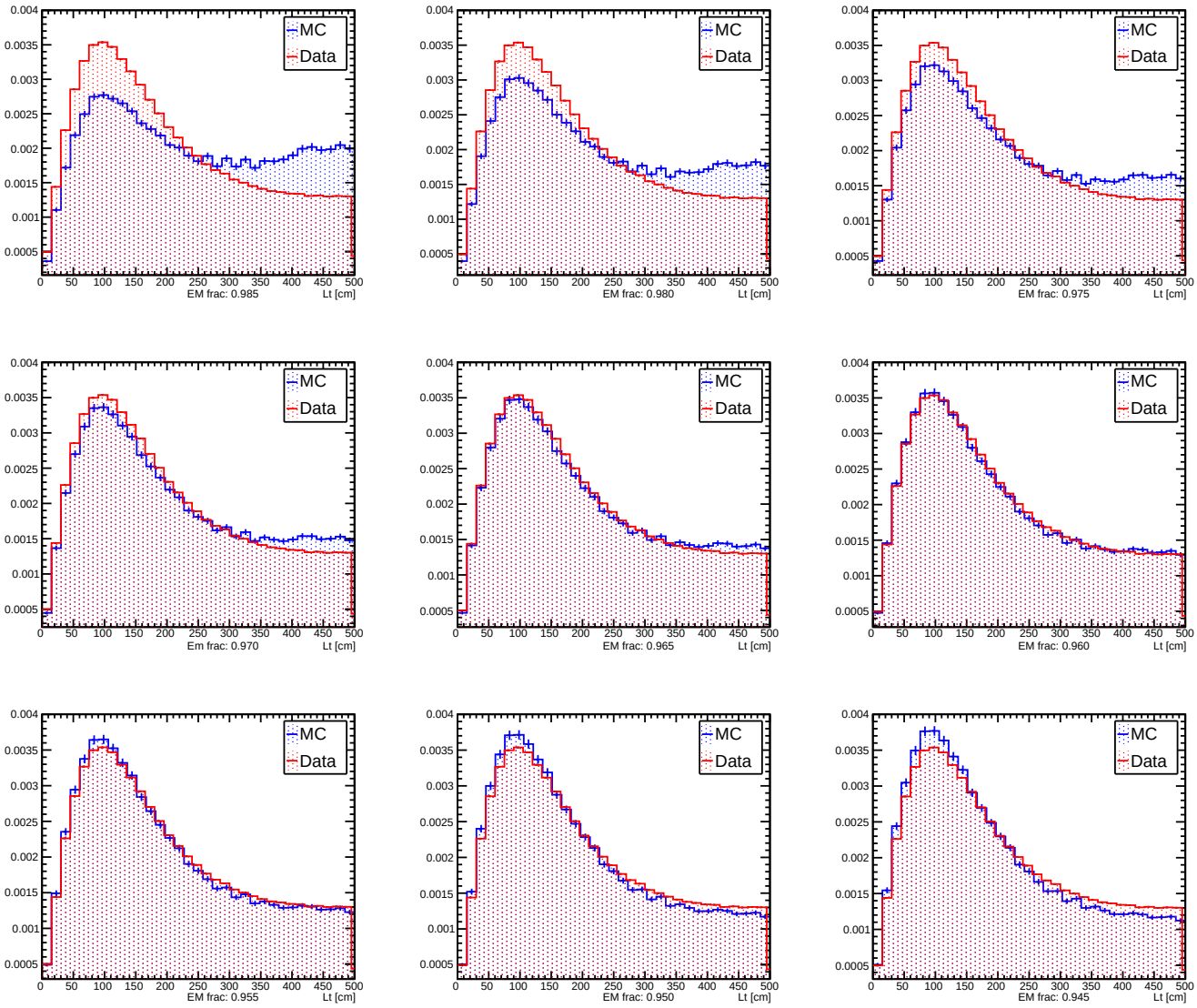


FIGURE 6.3: transverse distance distributions obtained from combining hadron-producing muons and EM muons simulation outputs with different fractions, in blue. In the same Figures data distributions are superimposed for comparison, in red. The label "EM frac", indicated along the x axis, represents the fraction of muon in the EM sample with respect to the total number of muons, as described in Equation 6.1. L_t is shown instead of the variable L_t^2 because differences in shape with data are visually better appreciable. All the plots are normalized to their integral.

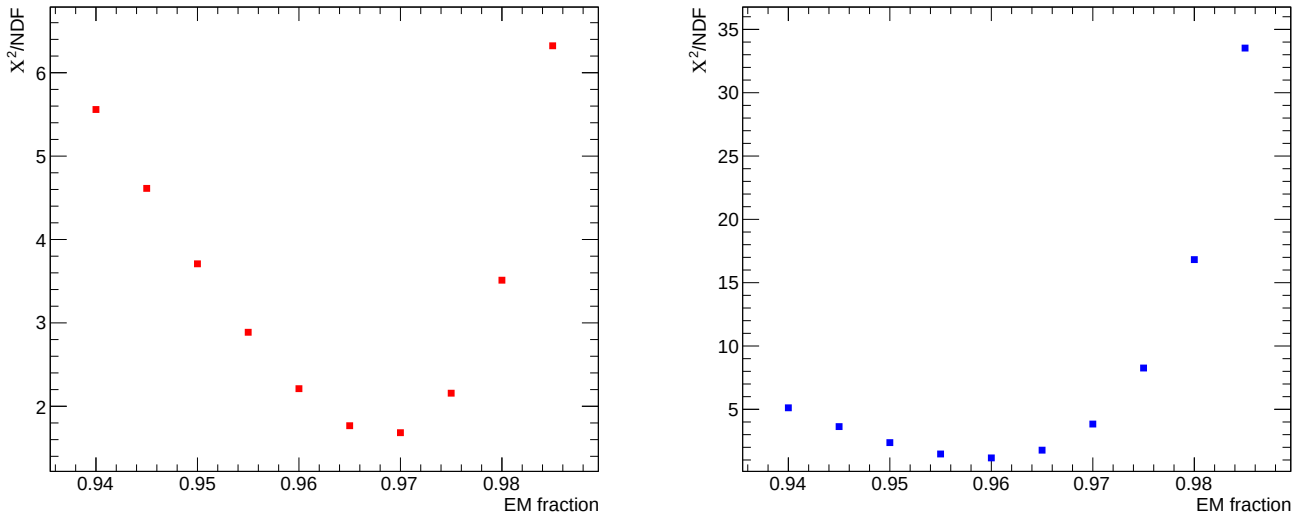


FIGURE 6.4: Values of χ^2/NDF, for different values of EM fraction (as defined in Eq. 6.1), obtained from the comparison of data distributions with simulation plots of L_{LONG_i} - $<L_{LONG}>$ in Figure 6.2 and L_t in Figure 6.3. Left in red for L_{LONG_i} - $<L_{LONG}>$ and right in blue for L_t. The minimum values are respectively for EM percentage of 0.97 and 0.96.

muon fractions described in the previous paragraph. It is evident how neutrons from EM-only muon sample mostly contribute to populate the tails of L_{LONG} total distribution of Figure 6.5. Similarly, fake neutrons dominate large L_t region: they can in fact be observed anywhere within 5 m of the muon track. The real neutron component overwhelmingly dominates for L_{LONG_i} - $<L_{LONG}><$ 5 m and $L_t <$ 3 m.

Figure 6.6 presents final L_{LONG_i} - $<L_{LONG}>$ and L_t^2 distributions from simulation, with EM fraction = 0.965, and SK-IV data. Note that L_t^2 is the variable commonly used in the analysis involving neutron clouds-based cuts. The non squared version, L_t, was used because, due to its shape, the comparison with data as well as the contributions of the two different samples obtained from simulations is visually more evident. Given that the vertex resolution, calculated from MC results, is about 70 cm for 2.2 MeV gammas passing WIT, the distribution shape is dominated by the shower size, not vertex resolution. Longitudinal and transverse distance distributions show the elongated elliptical shape of observed neutron clouds: average L_t and L_{LONG} extensions are about 3 and 5 m respectively. For distances of less than about 5 m, where contributions from real neutrons dominate, the predictions differ from the data by less than 15%. This excellent agreement motivates the use of a FLUKA-based simulation to predict neutron cloud shapes and optimize the associated cuts for future spallation analyses, notably at SK-Gd and Hyper-Kamiokande. At SK-Gd in particular, these simulation-based studies will allow to significantly reshape the spallation reduction procedure, as gadolinium doping will sizably increase the neutron identification efficiency.

Neutron multiplicity

Once the simulation samples are correctly assembled and a good agreement with data is obtained for the geometric variables defining the neutron clouds, L_{LONG_i} - $<L_{LONG}>$ and L_t^2, it is possible to exploit the results in order to study another important feature

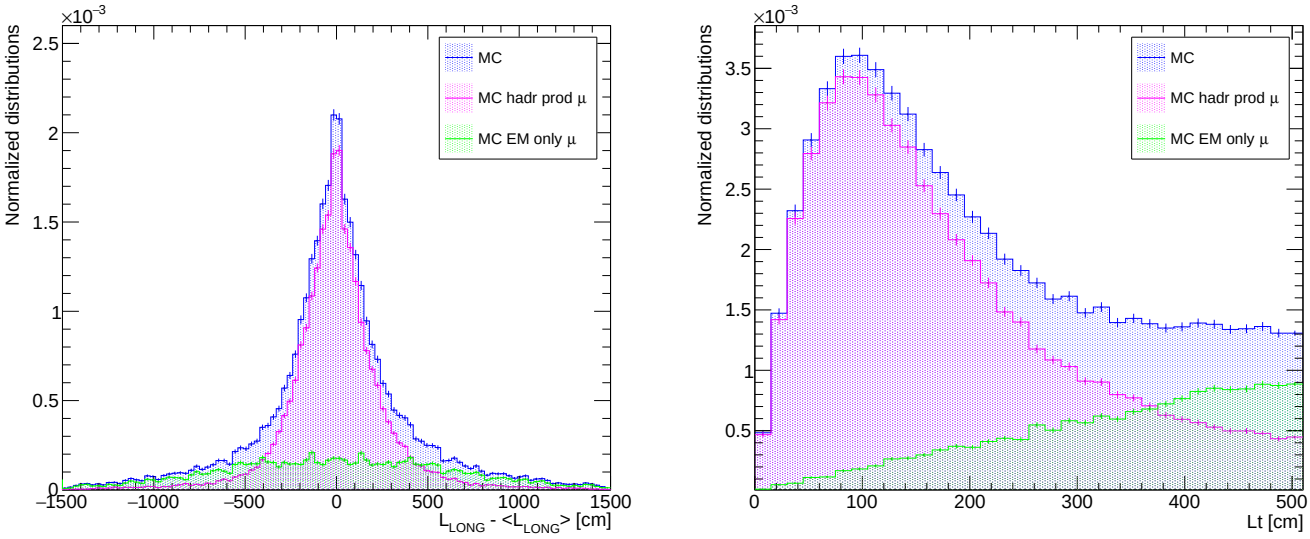


FIGURE 6.5: Longitudinal (left) and transverse (right) distance distributions obtained from combining hadron-producing and EM only muons simulation outputs with EM fraction of 0.97 and 0.96, respectively (defined as in Eq. 6.1). Contributions due to hadron-producing and EM muon sample are shown, respectively in magenta and green. The blue plot represents their sum. All the plots are normalized to their integral.

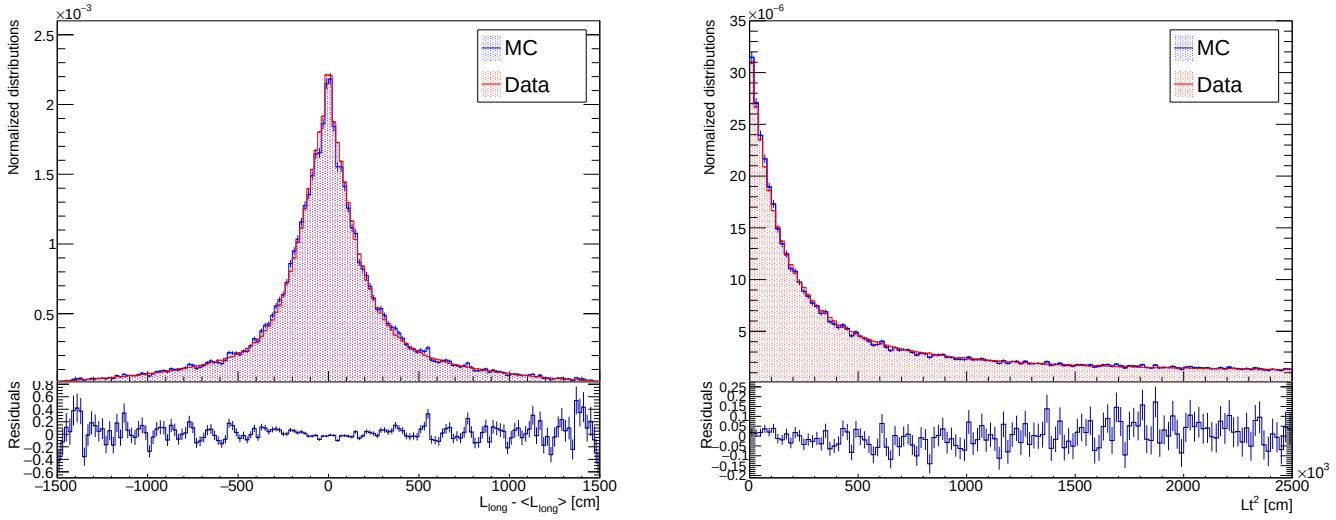


FIGURE 6.6: Longitudinal (left) and squared transverse distance (right) distributions obtained from combining simulation sample outputs with EM fraction of 0.965 (defined as in Eq. 6.1). The blue plot represents simulation results while data are in red. Note that L_t^2 is the variable used for the analysis. All the plots are normalized to their integral. In the bottom rectangle residuals are displayed.

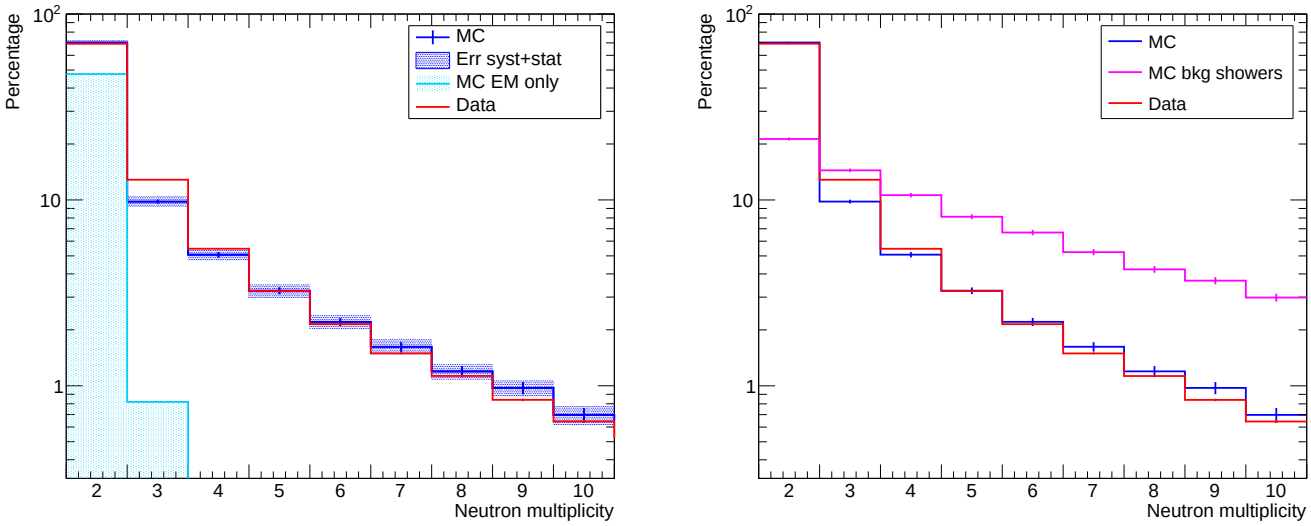


FIGURE 6.7: Neutron multiplicity in percents for data (red) and simulation (blue). Left: systematic and statistical error are indicated for multiplicity distributions from simulation. Systematic uncertainties are calculated varying the EM muon contribution of $\pm$ 0.05 from the best value, 0.965 in order to take into account the difference observed in the χ^2 minimization for L_{LONG} and L_t. The light blue histograms shows the contribution of non-showering muons only. Right: MC and data distributions are as in the right of the Figure, in magenta the distribution of neutron multiplicity for muon induced showers that contains at least a spallation isotope, labeled as "bkg showers".

of the cloud: its neutron multiplicity. Figure 6.7 shows the results for data and simulated clouds. The systematic errors are estimated varying EM muons contribution of $\pm$ 0.05 from the best value, 0.965: this takes into account the difference observed in the χ^2 minimization values due to mismodeling. The light blue histogram in the left Figure represents EM muon sample: its contribution is relevant only for very low multiplicity (clusters with 2 or 3 neutrons). Note that, the abundance of low-multiplicity clouds is due to both fake neutron contributions and the low efficiency of the neutron tagging algorithm. For neutron clouds with multiplicities lower than 10, simulation and data show reasonable agreement. As a comparison, the Figure on the right shows neutron multiplicity for muon-induced showers that contains at least one spallation isotope. As discussed in the previous Chapter, since these background showers are generally bigger in terms of number of particles generated and neutrons are the most abundant ones, capture multiplicity is expected to be higher, as confirmed from Figure 6.7. Finally, Figure 6.8 shows multiplicity distribution up to 45 neutron captures per muon in comparison with data. Despite the good agreement for lower multiplicity values, neutron clouds from simulation are richer in neutrons with respect to data, and FLUKA fails to accurately simulate the tails of the data distribution. Similar results were found by the Borexino collaboration in Ref. [101]. However, clouds with more than 10 reconstructed neutrons represent less than 1% fraction of the total and such large clouds are typically associated with shower producing hundreds, sometimes up to thousands, neutrons. Muons associated with these high-multiplicity showers are not only rare but also deposit a high amount of light in the detector and are hence easier to identify using other observables, such as for example the residual charge Q_{res}. Moreover, neutron multiplicity is a very poorly known quantity in the neutron production problem and models implemented in particle propagation codes

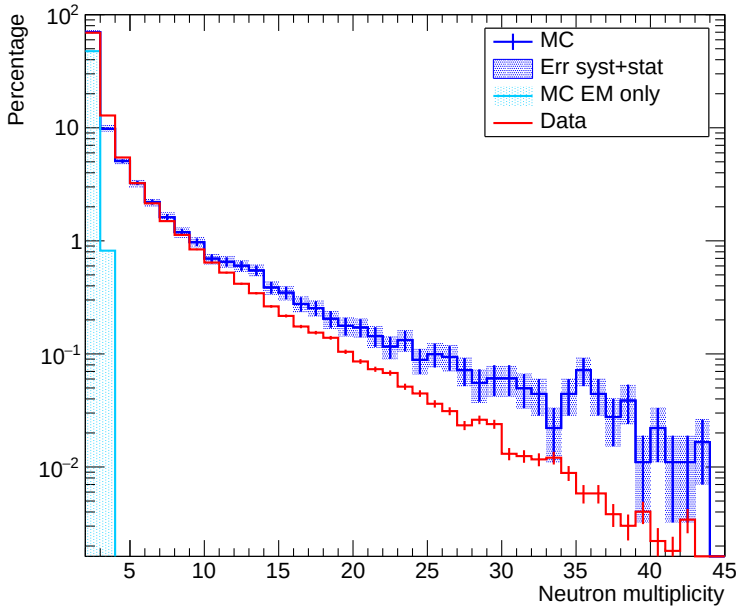


FIGURE 6.8: Neutron multiplicity in percentage for data (red) and simulation (blue). With respect to Figure 6.7, here the range is extended up to 45 neutrons, the contributions of higher values are negligible. Systematic and statistical errors are indicated for multiplicity distributions from simulation. Systematic uncertainties are calculated varying the EM muon contribution of $\pm$ 0.05 from the best value, 0.965 in order to take into account the difference observed in the χ^2 minimization for L_{LONG} and L_t.

like FLUKA are still subject to important uncertainties due to the lack of experimental data [129]. In any case, neutron multiplicity distributions from data can be used to improve the simulation in future.

In conclusion, our results demonstrate the ability of FLUKA-based simulations to accurately model hadronic showers and their properties for the muons inducing cascades potentially containing spallation isotopes. The simulation represents a powerful tool for future SK analyses.

6.4 Spallation rejection cuts

All the studies on neutron clouds and their properties described above are meant to be instructive to develop specific cuts for spallation rejection from SK-IV data [1]. These selections, applied to all the low energy analysis that suffer from spallation contamination, will be briefly described in this Section. Neutron cloud-based cuts are adopted together with likelihood selections described in Sec. 4.6.4: in fact, due to the low neutron detection efficiency in SK-IV, as well as the fact that a small fraction of spallation muons do not produce showers containing neutrons, neutron cloud cuts are not sufficient alone to remove spallation background. However, despite the procedure based on neutron cluster identification cannot be used as standalone for SK-IV, its impact is expected to drastically increase for SK-Gd and it is therefore very important to design a powerful spallation reduction technique that can be optimized for the SK phase with gadolinium dissolved.

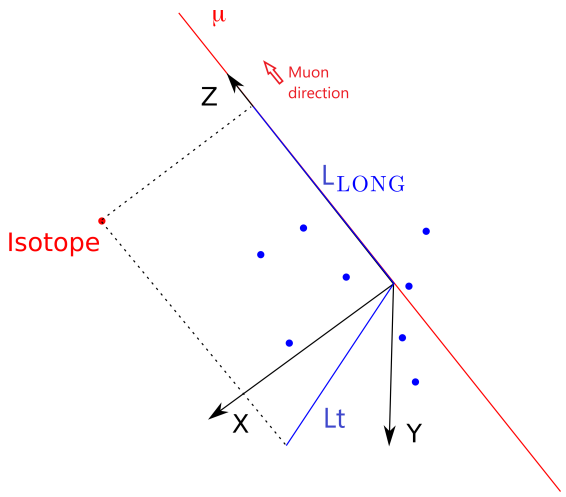


FIGURE 6.9: New coordinate system to tag spallation.

Multiplicity	≥2	≥2	2	3	4-5	6-9	≥10
Δt	0.2	2	30	60	60	60	60
L_{LONG} [cm]	750	500	350	500	550	650	700
L_t [cm]	750	500	200	245	346	447	500

TABLE 6.3: Table showing the different selection conditions for the cloud cut.

Low energy events close in space and time to neutron clouds are most probably spallation signals produced within the same shower. The geometrical variables of neutron clusters define a region of space where this shower possibly develops, guiding the removal of a sizable fraction of spallation events, with minimal harm to the signal efficiency. Neutron clouds are identified with the procedure described in Sec. 6.2 and their barycenter is computed using the weighted high-purity sample of good reconstructed neutrons. Then, cloud average position and time is compared with low energy event properties: firstly, in order to take advantage of the elliptical shape of neutron clusters, we define a new muon-centered coordinate system. The muon track becomes the z axis of the new system while the projection of cloud barycenter on the track is the origin, as shown in Figure 6.9. The advantage of this new coordinate system is that, once L_t and L_{LONG} are computed for the isotope vertex, they can be compared with the distributions of the same observables for the associated neutron cloud to estimate if the radioactive decay signal is located in the vicinity of the neutron cloud. With the new coordinate system, cuts on Δt, L_t and L_{LONG} for each low energy event-neutron cloud pair are defined, depending on the cloud multiplicity. We remind that, due to the 2 Hz muon rate in SK, each low energy candidate is associated to about 60 preceding muons, which could in turn be associated with neutron clouds. The spallation selection cuts must be applied to each of the $\sim$ 60 pairs muon-isotope candidate. All isotopes within 60 s (30 s if neutron multiplicity is equal to 2) and 7.5 m from the primary track are considered. Preliminary spherical cuts in two Δt regions are used to remove events within either 0.2 s and 7.5 m or 2 s and 5 m of cloud center. Then, after defining multiplicity bins (2, 3, 4-5, 6-9, and $\geq$10 neutrons candidates), ellipsoidal cuts on L_t and L_{LONG} are applied. The different cuts are summarized in Table 6.3. If a pair fails the selections, the corresponding low energy event is rejected.

To conclude, the first part of this Chapter introduced a new technique to reduce spallation background in low energy analyses based on neutron cloud identification. Valuable information extracted from the FLUKA-based simulation made possible a deep understating of all the underlying processes: the simulation gives accesses to the fraction of real neutron compared to mistagged ones, the difference in their distributions and their multiplicity. The neutron cloud method has been already implemented for SK-IV data, when WIT trigger was available. Despite the short WIT live time, this cut showed some improvements both in solar and SRN analysis: for the first one a reduction of dead time by a factor of two was observed so that for a signal efficiency of $\sim$ 98.7%, 53% of the spallation isotopes were removed only with neutron cloud cuts [1]. As for the supernova relic neutrino search, with a signal efficiency of more than 99%, 40% of the spallation background is removed [130]. However WIT trigger only has a limited lifetime of 388 live days, therefore the overall impact of neutron cloud cut on the whole SK-IV period is limited: only 10% of the spallation isotopes are removed in the SRN search if we considered the total analysis live time of 2790 days. The performances of the spallation reduction technique described above are mostly limited by the weakness of the signal following a neutron capture on hydrogen. However, neutron cloud cuts are expected to considerably increase their rejection power with SK-Gd, becoming a key tool for spallation reduction. A deeper study of neutron behaviors in presence of Gadolinium and a remodeling and optimization of selection cuts is therefore necessary for future analyses. The proven accuracy of FLUKA simulation, the first realistic framework modeling spallation production in SK, motivates its use for this purpose: next Chapter will present how FLUKA results can be exploited to develop spallation algorithms adapted to SK-Gd. In this context, the simulation presented in this work will become instrumental for designing future analysis strategies: beyond neutrino-antineutrino discrimination, neutron tagging is expected to impact significantly all low energy neutrino searches at SK.

6.5 Isotope yields

Isotope yields in SK were estimated by a study from 2016 with 1890 days of SK-IV data [52]. Here we describe an update of this work using the whole SK-IV live time [1], 2970 days. The results are compared with values obtained with FLUKA simulation: in addition to improvements on the yield estimation method, precious insights from simulation make possible the computation of uncertainties on an isotope-per-isotope basis. Moreover, we will show how we can use neutron captures to isolate a data sample which is rich of spallation events without modifying the relative fractions of the most abundant isotopes.

Isotope yields are extracted from a fit of the Δt distribution. Primary selection cuts on data are required to remove non-spallation events from the sample of muons and isotopes used for the fit: only low energy event vertices within the FV and at least 50 μs after the muons were considered in order to eliminate radioactivity, cosmic $\mu - e$ decays and PMT afterpulsing. We consider events with kinetic energies ranging from 6 to 24.5 MeV, in order to remove the region where radioactivity dominates. All isotope candidates are paired with muons crossing the detector up to 100 s before them, and two high-purity samples are constructed with the following requirements: for the first sample we require $L_t < 200$ cm ("L_t sample"), as it was done in [52], while for the second sample we select muons associated with clouds containing 3 or more neutrons and isotopes verifying the conditions listed in Table 6.3 ("neutron cloud sample").

The fit of time distribution is performed using a sum of exponentials whose characteristic times are associated to the 10 most abundant isotopes:

$$R_{TOT} = \sum_{i}^{n} R_i \cdot e^{(-\Delta t/\tau_i)} + const. \tag{6.2}$$

where R_{TOT} is the total event rate and the τ_i are the isotope decay constants, that are kept fixed to fit the individual production rates R_i. Some isotopes, ^{8}Li and ^{8}B as well as ^{9}C and ^{8}He, are not distinguishable, given that their time constants are within 10% of each other, they are therefore paired together. A final χ^2/NDF of 0.95 is obtained for the L_t dependent sample, showing an overall good agreement of the fit function with data. Isotope yields, Y_i, are obtained by:

$$Y_i = \frac{R_i \cdot FV}{R_\mu \cdot \rho \cdot L_\mu} \tag{6.3}$$

where ρ is the density of water, R_μ is the muon rate (2.00 Hz), L_μ is the average length of reconstructed muon tracks, and FV is the fiducial volume of the detector. However, R_i, the total rate of the i^{th} isotope extracted from fit, needs to be corrected by the efficiencies of selection cuts in order to obtain the total production rates. These rates can in turn be compared with results obtained by FLUKA simulation extracted from a sample of 1.4×10^8 primary muons.

6.5.1 Spallation efficiencies

Obtaining total spallation yields requires computing the efficiencies of the selection cuts applied on the data samples. Unlike in Ref. [52], we now have at our disposal a MC simulation that can greatly help the efficiency computation. However, we need to remind that FLUKA simulation only includes through-going muons, therefore data-based corrections need to be applied to extend efficiency results from simulation to all categories of muons in data.

Firstly, efficiencies related to primary cuts for noise elimination and energy selections for different isotopes are determined with outcomes of simulated radioactive decays passed through SKDetSim, where time dependence of the detector properties and water quality are modeled almost day by day. The decay spectra are simulated using GEANT 4.10.7 for most isotopes, accounting for both beta and gamma emissions. Efficiency values are listed in Tab. 6.4. The same Table shows $L_t < 200$ cm cut efficiencies: simulation only provides values for single through-going muons while data need to be used to complement for other muon categories. L_t selection efficiencies are extracted for each muon status from Δt distributions of data and total efficiency is computed as a weighted average of the observed relative contributions of the different status. The final L_t efficiency from data is (74.5$\pm$0.7)%. The value for single through-going muons only is found to be about 1 σ from MC result. Isotope dependence of the L_t distribution, obtained by MC, is taken into account by defining an individual scaling factor computed dividing each isotope efficiency by the weighted average all isotopes efficiencies (last row of Table 6.4). These values are applied to the L_t efficiency measured from data as a scaling factor for single through-going, multiple and stopping muons yield results. The use of simulation allows to refine the L_t efficiency estimation performed in [52], where the isotope dependence was covered by a global $\sim$ 4% systematics. Isotope yield results are presented in Table 6.5: the estimation of total yields for the most abundant spallation isotopes is determined with a precision of a few percents. Figure 6.10 shows a graphic illustration of the results comparing FLUKA outcomes with all SK-IV data with 2970 days of live time. It is crucial to underline that hadronic models can introduce uncertainties of the order of 100-200% for

Isotope	Primary cuts + E $\leq$ 6 MeV (%)	Lt $\leq$ 200 cm for STG (%)
^{12}N	69.0	94.9
^{12}B	54.3	89.9
^{8}He	23.9	94.6
^{9}C	67.4	94.0
^{9}Li	44.4	91.8
^{8}Li	51.8	91.4
^{8}B	63.2	92.9
^{15}C	37.1	87.7
^{16}N	54.8	88.9
^{11}Be	44.7	86.5
Average		90.3 $\pm$ 2.8

TABLE 6.4: Spallation sample efficiencies from simulation.

the most important isotope yields, as described in Sec. 4.4; taking into account this factor we obtain a satisfactory agreement between fitted data results and yields from FLUKA simulation: all predicted yields lie within a factor 2, or less, from the observed results.

Finally, a similar fit is performed on a sample where muons are associated with a neutron cloud with at least 3 neutrons. Fit results give a χ^2/NDF of 1.04. For the most abundant isotopes, the ratio between the rates obtained by fitting the neutron cloud sample ranges between 60-70% of results from L_t sample. The stability of the ratio between results from two different samples enlightens their correlation: both selections take advantage of the geometrical association of the low energy event with the muon track. The study of isotope yields based on neutron cloud identification demonstrates that it is possible to build a spallation-rich sample, with minor background contamination, exploiting shower properties and, most outstanding, without introducing biases in the relative fraction of the most important isotopes. This can be a meaningful result for SK-Gd where neutron clouds can be used for a more in-depth study of spallation production, for example the sample can be used to train machine learning algorithms for background identification. In particular, in SK-Gd, a more refined study of isotope yields associated with neutron clouds can help understand if there are isotopes more frequently produced in association with neutrons and compare the results with the simulation in order to tune the neutron cloud method on an isotope-per-isotope basis.

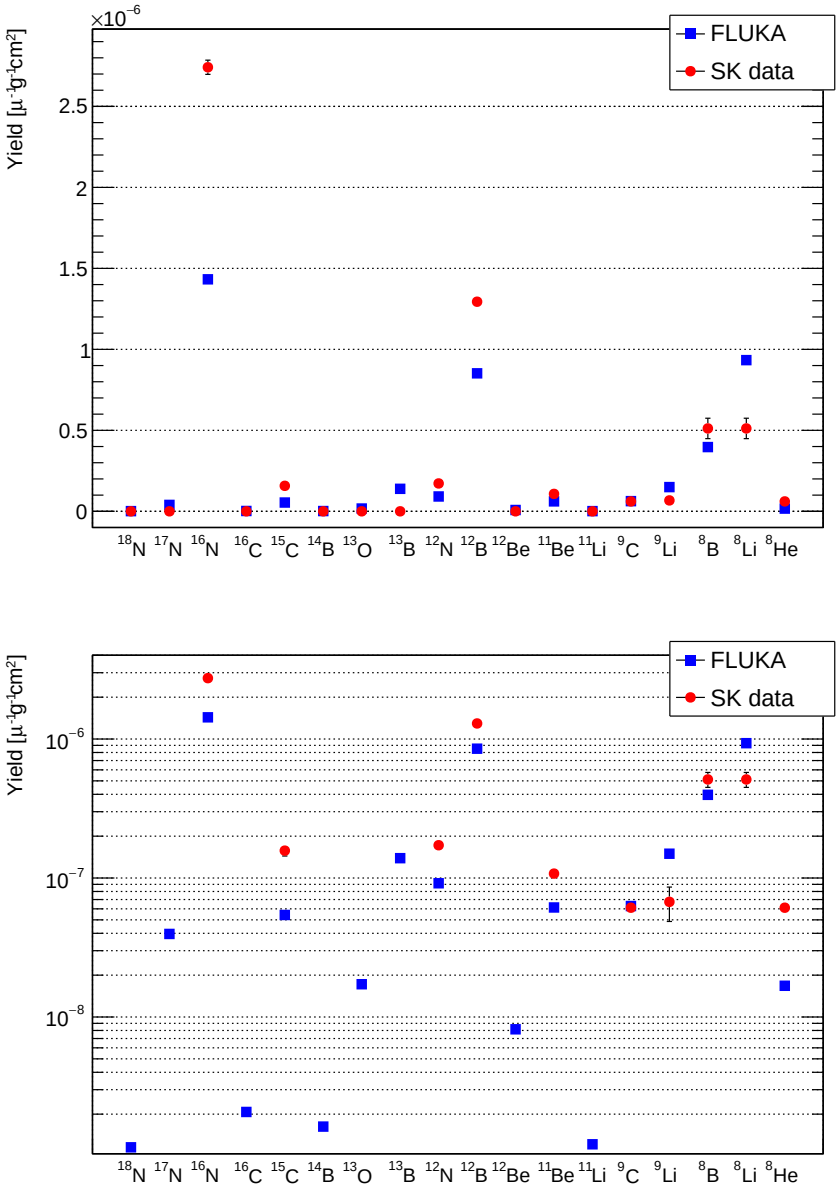


FIGURE 6.10: Isotope yields for main spallation contributions in SK. The top Figure has linear scale in y-axis while the bottom one shows results in logarithmic scale to better distinguish lower yield isotopes. FLUKA results, blue squared markers, are compared with SK data results with the whole SK-IV dataset, red circles. Hadronic model uncertainties can lead to a factor 2 error on simulated yields, here not shown. Note that where the red dot for SK data is not present means that we do not have the value computed for the given isotope because of the low yield.

	Simulation		L_t Data		
Isotope	Yields 10^{-7} ($cm^2\mu^{-1}g^{-1}$)	Fraction	Fit Raw Rate ($kton^{-1}day^{-1}$)	Total Rate ($kton^{-1}day^{-1}$)	Yields 10^{-7} ($cm^2\mu^{-1}g^{-1}$)
^{12}N	0.92	2.3%	1.55	3.04 ± 0.07	1.72
^{12}B	8.6	21.1%	9.19	22.86 ± 0.24	12.9
$^{9}C/^{8}He$	0.8	1.9%	0.11	<1.08	<0.61
^{8}He*	(0.17)	(0.4%)			
^{9}Li	1.5	3.7%	0.39	1.19 ± 0.33	0.67
$^{8}Li/^{8}B$	13.4	32.8%	3.69	9.04 ± 1.1	5.11
^{8}Li*	(9.4)	(23%)			
^{8}B*	(4)	(9.8%)			
^{15}C	0.55	1.3%	0.76	2.78 ± 0.45	1.57
^{16}N	14.5	35.3%	19.64	48.43 ± 0.75	27.3
^{11}Be	0.61	1.5%	0.33	<1.9	<1.05

TABLE 6.5: Measured and predicted spallation rates and yields for data and MC using the L_t cut. Compared to the previous results, observation for ^{15}C has been done and a much stronger constraint has been set on the ^{11}Be measurement. The asterisks for isotopes in the first column list indicate decay contributions with yields not separable from data.

Chapter 7

Optimization of spallation cuts for SK-Gd

The major detector upgrade of summer 2018, described in Chapter 3, successfully prepared Super-Kamiokande to its new, and currently running, phase: SK-Gd. During Summer 2020, 13 tons of gadolinum sulfate octahydrate, $Gd_2(SO_4)_3 \cdot 8H_2O$, were dissolved into the detector's pure water with the goal of enhancing the efficiency of thermal neutron detection. This quantity corresponds to roughly 10% of the proposed target concentration, 0.2% by mass, firstly advised by the authors of Ref. [49].

Although the primary motivation of Gd doping is to increase SK sensitivity to the observation of Supernova relic neutrinos by identifying antineutrinos, the new technique shows a decisive potential in other analysis areas. Among all its advantages, the performances of spallation background rejection techniques based on the identification of neutron clouds, largely described in Chapter 6, greatly benefit from gadolinium properties. In fact, we highlighted in the previous Chapter how the main difficulty with neutron cloud tagging is the low efficiency of the WIT trigger system due to the weakness of the capture signal. Since this efficiency is expected to drastically improve with SK-Gd, we foresee that all the analyses suffering from spallation backgrounds could take advantage from an improved efficiency of the new neutron cloud rejection method. For the SRN analysis, notably, while most spallation backgrounds will be removed by identifying neutron captures following positron-like events (see Section 1.6), dedicated spallation cuts are still needed to remove β+n decays of spallation isotopes such as ^{9}Li. In fact, since about 83% of the ^{9}Li isotopes are produced in spallation-induced showers containing neutrons, increasing the performance of neutron cloud cuts could drastically increase SK's sensitivity to low energy antineutrinos.

With the goal of improving background rejection for SRN analysis as well as other low energy searches, an adjustment and optimization of the SK-IV spallation cuts, in view of an implementation in SK-Gd are necessary. The proven accuracy of the FLUKA-based spallation simulation makes it a powerful tool to study this major background in the presence of gadolinium, and gives unique insights to study future application to low energy analyses. This Chapter will present how the simulation has been upgraded in order to include Gd, which de-excitation models for neutron capture have been tested and implemented and how the results compare to the first SK-Gd data. We then use this new simulation to reevaluate and optimize the performance of neutron-cloud cuts for SK-Gd, providing precious information for future analyses in the MeV energy range.

7.1 Neutron tagging with gadolinium

Among all stable nuclei known in nature, Gadolinium is the one with the largest neutron capture cross-section in water. Table 7.1 shows Gd composition in nature and its

Isotope	Natural abundance (%)	σ (barns)	De-excitation En (MeV)
^{152}Gd	0.20	1050	6.25
^{154}Gd	2.18	85	6.44
^{155}Gd	14.80	**60700**	**8.54**
^{156}Gd	20.47	1.7	6.36
^{157}Gd	15.65	**254000**	**7.94**
^{158}Gd	24.84	2	5.94
^{160}Gd	21.86	0.77	5.64

TABLE 7.1: Gadolinium isotopes with their natural abundances, neutron capture cross sections (σ) and total energy released in the de-excitation through the emission of gamma cascade. From [131].

properties: after reaching thermal energies, ($\mathcal{O}$(eV), typically in few μs and $\sim$ 1 m for MeV neutrons), the neutron can be captured by one of the Gd isotopes, which in turn de-excites emitting a gamma cascade whose total energy depends on the Gd mass number. ^{155}Gd and ^{157}Gd are particularly important due to their large capture cross-sections, since they have a resonance state in the thermal energy region, and release a quite large total energy of about 8 MeV. The multiplicity of photons in the cascade typically ranges from three to ten photons[128].

The amount of gadolinium injected in SK at the beginning of the current phase corresponds to 0.011% concentration, equivalent to 0.021% of gadolinium sulfate $Gd_2(SO_4)_3$. Neutron capture time has been measured with calibration sources and found of to be 115.6 $\pm$ 0.6 μs, relatively stable in time and depth [132]. With the current level of Gd concentration a 45–50% neutron capture efficiency is expected by MC estimations. The remaining neutrons are still captured on hydrogen. A successive phase, planned to begin in 2022, will see the increment of Gd concentration up to 0.03% Gd (0.06% of $Gd_2(SO_4)_3$), which should result in 75% neutron tagging efficiency. The final goal concentration is expected to be around 0.1% of Gd with more than 90% capture efficiency.

7.2 Gadolinium addition in the simulation

Although the general structure of the simulation, described in Chapters 4 and 6, remains unchanged, the simulation pipeline should be modified to account for Gd doping. In FLUKA, after the addition of $Gd_2(SO_4)_3$ in the simulated material at the concentration of 0.021%, neutron captures are recorded. However we observe that photons in a cascade produced after neutron captures are mismodeled: in fact, neutron interactions below 20 MeV are ruled by very complex cross-sections with rich structures difficult to calculate by models. Therefore FLUKA relies on evaluated data files which contains only cumulative spectra and mean multiplicity values: there is no correlation among emitted gamma lines on an event-by-event basis, only average energy is conserved [99, 128]. For this reason, an alternative model to simulate γ cascades need to be implemented in the simulation: neutron capture times and positions are calculated by FLUKA and for each of them, photons from neutron captures are simulated using a separate tool (GEANT4) before passing to SKDetSim, as described in the next paragraph. A capture time of 115.07 $\pm$ 0.05 μs is obtained with FLUKA and is in good agreement with calibration results. This corresponds to a capture efficiency of neutron on Gd of 0.43, as shown in Figure 7.1.

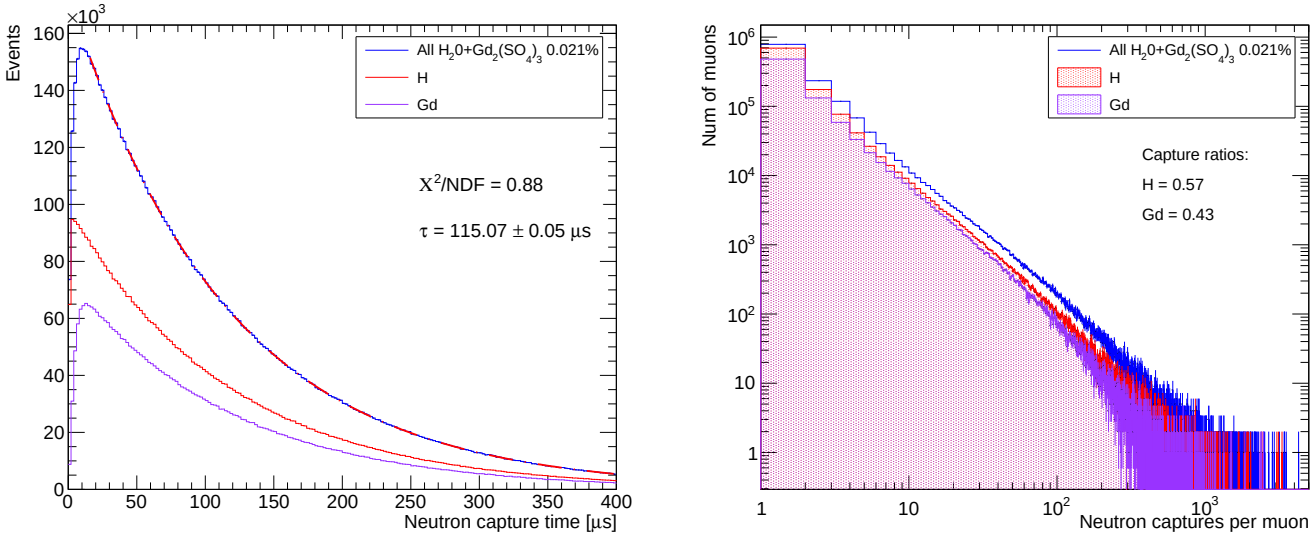


FIGURE 7.1: Neuron capture time spectra with exponential fit results (left) and number of neutron captures per muon (right). Separate contributions for captures on H and Gd are shown.

To simulate neutron capture on gadolinium, two common photofission models have been tested and compared with SK-Gd data: GEANT photofisson [133] and ggarnet models (also called EGLO model) [134]. The first one is integrated into the neutron-fission modeling package of GEANT4 and produces delayed gamma rays with time and energy dependencies consistent with ENDF/B-VII database [128]. The second one is an early version of a model developed using γ-ray energy spectrum measured with a germanium spectrometer, ANNRI detector at JPARC, exposed to a neutron beam source[1]. In order to find the model which best fits the data, we simulated photon cascades in SKDetSim and compared the reconstructed total energy spectrum with data, as shown in Figure 7.3: the region below $\sim$ 4 MeV is not to be considered, it contains in fact a large fraction of fake events due to dark noise or radioactivity, that for simplicity have not been accurately simulated, as well as photons from captures on hydrogen, also not included in this phase. Instead the second peak is the one related to captures on Gd. Both models underestimate the total reconstructed energy, however, since the GEANT photofission model best reproduces SK-Gd data we will implement this one in our simulation. One of the main reasons why simulated cascades differ from data, apart from intrinsic differences in the models, arises from the fact that during the first months of data taking in SK-Gd the detector condition was still under test and not completely stable (for example an increase of dark noise level was observed, probably due to radioactive impurities that came with Gd). Moreover, a dark rate increase with time was observed around the top of the detector, introducing further disuniformities. These resulted in an inexact knowledge of water transparency and energy scale, key parameters in SKDetSim, and a subsequent underestimation of the number of PMT hits in MC with respect to real data. Despite there is an ongoing effort to precisely tune SKDetSim for SK-Gd to improve the data-MC agreement and better results are foreseen in the near future, these uncertainties are expected to impact the results presented in this Chapter.

[1] The main difference between ggarnet model and its new version (ANNRI model) is the photon-strength function used in the calculation, namely the function describing the average response of the nucleus to an electromagnetic probe. Above Cherenkov threshold it is not expected to affect the distributions of the gamma-ray emission [134].

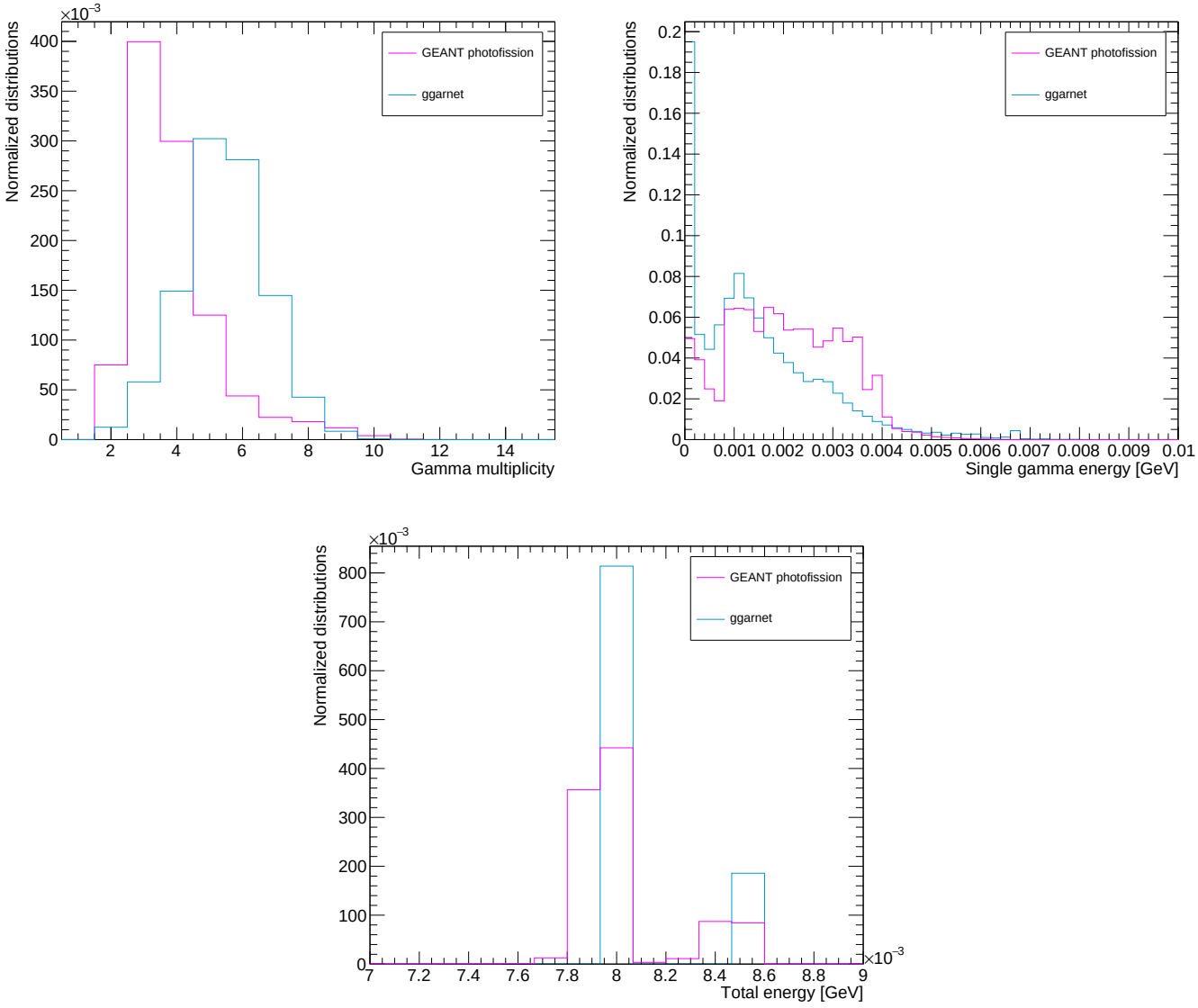


FIGURE 7.2: Multiplicity, single photon energy and total emitted energy spectra for the two different models tested: GEANT photofission in magenta and ggarnet in light blue.

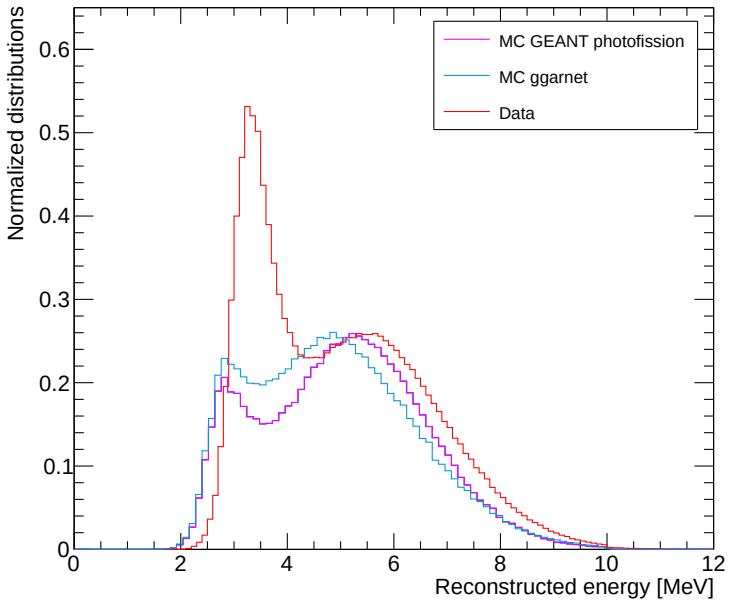


FIGURE 7.3: Reconstructed energy spectra from photon cascades simulated with ggarnet and GEANT photofission models compared to data. The plots are normalized to the height of Gd de-excitation energy peak (around 5 MeV).

7.3 Neutron clouds in SK-Gd

In order to evaluate the accuracy of FLUKA-based simulation to reproduce neutron cloud geometrical properties, the procedure described in Sec. 6.3.3 is replicated: after constructing the two simulated samples, hadron-producing and electromagnetic only muons, their relative fraction is extracted from a fit to data. All selection requirements applied to identify good neutrons are kept unchanged except the reconstructed energy threshold which is extended from 5 MeV to 10 MeV to accommodate more energetic events from Gd de-excitation.

Neutron tagging algorithm performances

The impact of gadolinium doping on neutron tagging performances is remarkable, as a direct consequence of the properties of Gd de-excitation. Table 7.2 shows the efficiency values for all selection criteria applied for neutron cloud identification and it is to be directly compared with the analogous Table for SK with pure water (Tab. 6.1). Neutron detection efficiency by WIT trigger system for SK-Gd is more than 3 times larger than before and largely contributes to the increase of the final neutron identification efficiency which passes from 6.5% to 26.9%. Fit quality efficiency is also improved by a factor $\sim 12\%$ thanks to a more powerful signal arising from photon cascades. Similarly, the L_t cuts at 5 m has a different impact in SK-Gd because the better resolution on neutron candidate position shrinks the transverse distance distribution, as shown in Fig. 7.4: in fact the L_t distribution is generally dominated by the thermalization path of the neutron, which does not change when Gd is added. The only reason why L_t distribution looks thinner in SK-Gd is related to the better reconstruction of photons within the de-excitation cascades.

Selection requirements	Efficiency values (%)
WIT trigger	47.6 ± 0.1
$20\mu s < \Delta t < 500\mu s$	86.2 ± 0.2
L_t <500 cm	82.3 ± 0.1
FV	92.6 ± 0.5
Fit quality	86.6 ± 0.2
$E_{rec} < 10$ MeV	99.5 ± 0.2

TABLE 7.2: Neutron tagging efficiencies for different selection criteria in SK-Gd: WIT trigger identifying a neutron candidate, timing cut to remove the primary muon signal and the associated after-pulse, proximity to the muon track, fiducial volume, goodness-of-fit cut, and energy cut. We consider a simulated neutron to be tagged if its true capture time lies within 50 ns and 50 cm of the reconstructed neutron candidate.

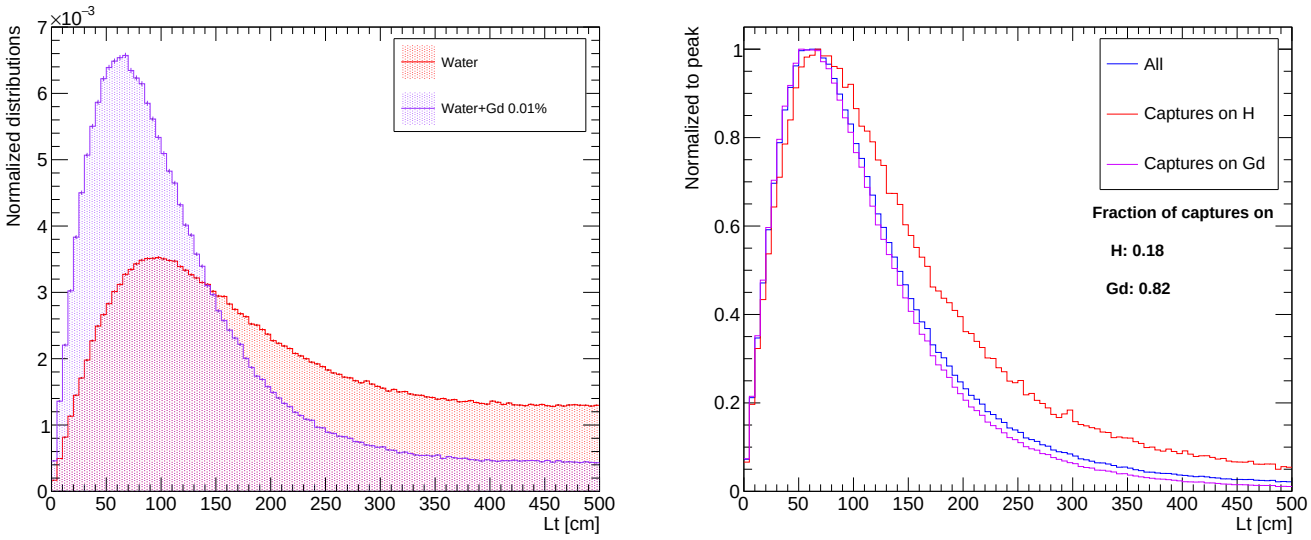


FIGURE 7.4: Left: L_t distribution for pure water and water with Gd dissolved. Note that 0.01% Gd corresponds to about 0.02% $Gd_2(SO_4)_3$. Right: L_t distributions separately for captures on Gd and on H are shown as well as the fraction of detected and well reconstructed signals form the two cases.

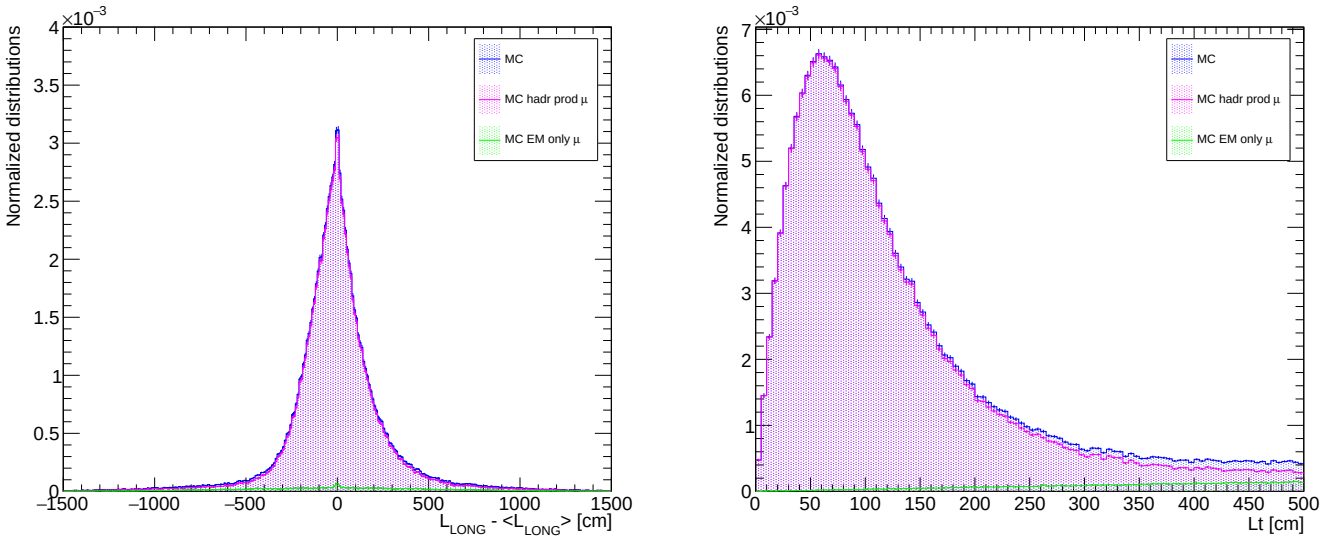


FIGURE 7.5: Longitudinal (left) and transverse (right) distance distributions obtained from combining hadron-producing and EM only muons simulation outputs with EM fraction of 0.78 and 0.76, respectively (defined as in Eq. 6.1). Contributions due to hadron-producing and EM muon sample are shown, respectively in magenta and green. The blue plot represents their sum. All the plots are normalized to their integral.

One other important feature is the mistag rate, namely the number of fake neutrons per primary muon, which passes from 0.044 in SK-IV to 0.084 $\pm$ 0.001 for SK-Gd. The variation is due to the general increase of dark rate observed in the initial phase of SK-Gd. This effect is however accompanied, and counterbalanced, by an augmentation of the number of real neutrons detected per muon, from 0.24 it rises to 1.08 $\pm$ 0.01.

Neutron cloud shapes

Repeating the procedure described in Chapter 6, the distributions of variables defining the neutron cloud shape, L_{LONG} and L_t, are extracted from a fit to data with varying fraction of fake neutron contributions, as expressed in Eq. 6.1. The EM fractions for which simulation distributions best fit data are 76% (χ^2/NDF = 1.7) and 78% (χ^2/NDF = 9.5) for L_t and L_{LONG_i} - $<L_{LONG}>$ respectively. The agreement is less good than in the case of pure water, especially for L_{LONG}, which is most probably due to the fact that Gd de-excitation models simulating gammas emitted from neutron captures give lower visible energies that observed in data. As with the pure water case, a mean value of EM fraction equal to 77% is adopted and the differences between the separate fits are included as a systematic error. The fit value obtained differs from the result of Chapter 6 as a consequence of the increased dark rate at the beginning of SK-Gd, resulting in an higher fake neutron rate. This discrepancy can also be related to fundamental differences between data and FLUKA models: FLUKA calculations tend to give higher neutron multiplicities with respect to data which are more pronounced with Gd, given the higher detection efficiency.

We observe from Figure 7.5 that the contribution of EM-only muon sample is minor in SK-Gd neutron cloud distributions: unlike pure water, the increased detection efficiency leads to a suppression of fake neutron contamination. Figure 7.6 presents final L_{LONG_i} - $<L_{LONG}>$ and L_t^2 distributions from simulation, with EM fraction = 0.77, and SK-Gd first months of data. Finally, multiplicity distributions are shown in Fig. 7.7: a similar trend

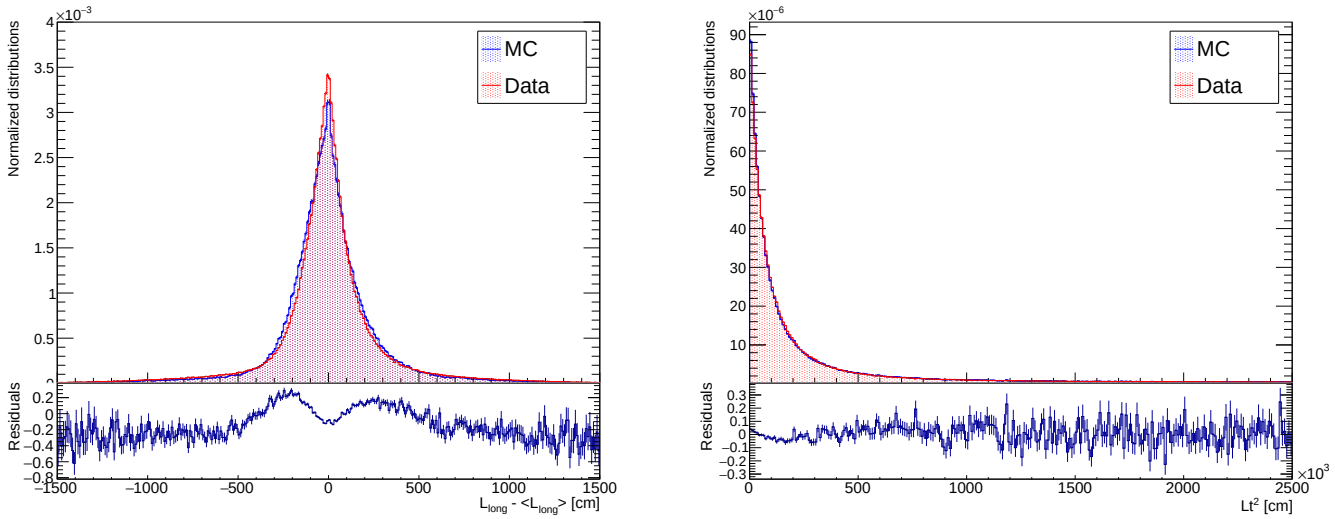


FIGURE 7.6: Longitudinal (left) and squared transverse distance (right) distributions obtained from combining simulation sample outputs with EM fraction of 0.77 (defined as in Eq. 6.1). The blue plot represents simulation results while data are in red. All the plots are normalized to their integral. In the bottom rectangle residuals are displayed.

showing an overestimation of high multiplicity shower from FLUKA simulation was observed in the case of pure water (Fig. 6.8). The average number of neutrons counted per muon is higher than the case of pure water, as expected. Showers with more than 40 detected neutrons are observed with a probability one order of magnitude larger than in SK-IV. Nonetheless showers containing cosmic-induced radioactivity are still the richest in neutrons, making the neutron cloud method particularly useful for spallation rejection (magenta in Fig. 7.7, right). An ongoing effort for a better tuning of calibration parameters included in SKDetSim as well as a reduction of water contamination, and therefore dark rate in SK-Gd phase, are expected to improve the agreement of the FLUKA-based simulation with real data. In the meantime, the sufficiently satisfactory accord obtained from a MC-data comparison motivates the use of the simulation to define an optimization strategy for spallation reduction cuts. Similar studies can be easily implemented for the next phase of Super-Kamiokande with higher Gd concentration. Similarly, given the flexibility of the calculation structure, it can be adapted and exploited for next generation experiments like Hyper-Kamiokande, making the spallation simulation a key tool for future analyses.

7.4 Optimization of neutron cloud cuts for SK-Gd

The general structure of neutron cloud cuts to remove spallation background is described in Sec. 6.4. Selections are based on time difference between the low energy event and the muon, as well as the longitudinal and transverse distance of the isotope to the origin of a cloud-centered coordinate system, and depend on the multiplicity class of the neutron cloud. In fact, within the rotated system illustrated in Fig. 6.9, L_{LONG} and L_t associated with the related neutron cloud define an elliptical region inside which all the low energy events are classified as spallation background and removed. The format of the selection remains unchanged for SK-Gd but all the parameters, included the multiplicity classes

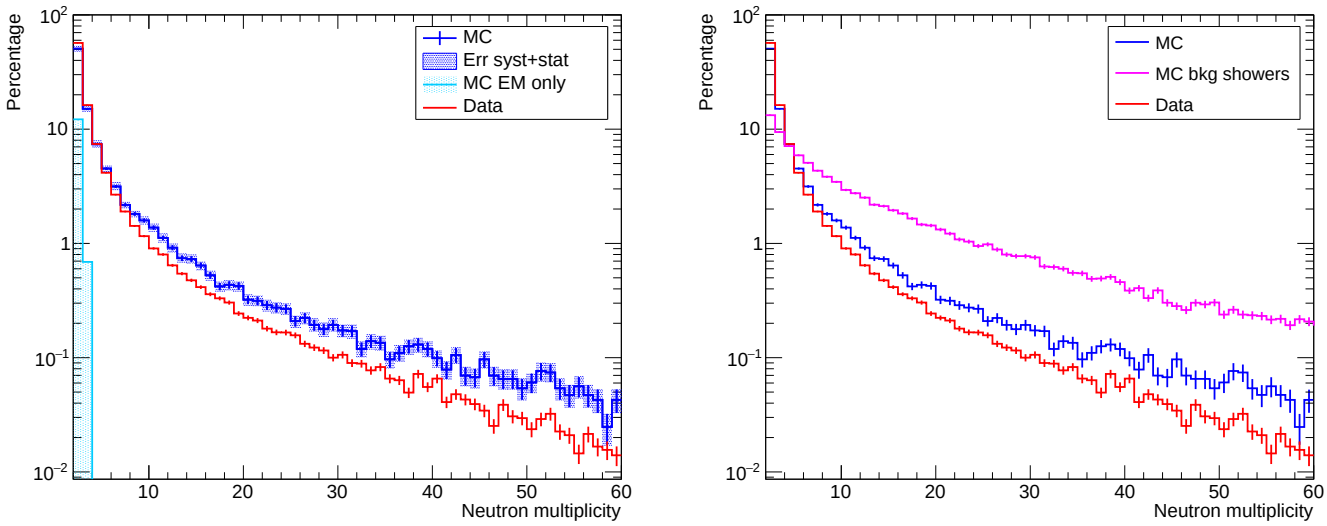


FIGURE 7.7: Neutron multiplicity in percentage for data (red) and simulation (blue). Left: systematic and statistical errors are indicated for multiplicity distributions from simulation. Systematic uncertainties are calculated varying the EM muon contribution of $\pm$ 0.01 from the best-fit value, 0.77 in order to take into account the difference observed in the χ^2 minimization for L_{LONG} and L_t. The light blue histograms shows the contribution of non-showering muons only. Right: MC and data distributions are as in the left part of the Figure, in magenta the distribution of neutron multiplicity for muon-induced showers that contains at least a spallation isotope, labeled as "bkg showers".

have to be updated to better fit the increased detection efficiency. In particular, each parameter defining the neutron cloud method is varied and for each new set of parameter values the rejection method performances are evaluated.

7.4.1 Neutron cloud samples for optimization

The goal is to scan over a grid of cut values for Δt, L_{LONG} and L_t and evaluate in parallel the fraction of spallation background removed and the signal efficiency, namely the amount of low energy events not correlated with preceding muons —that could be signal— surviving the cuts. For each set of cut parameters we want to maximize the spallation rejection while keeping the dead time as low as possible. For this study it is firstly necessary to build two event samples samples which correspond respectively to true spallation-inducing events and to cosmic muons accidentally associated to low energy signals.

- *Spallation sample*: we use results from FLUKA simulation to extract a list of true spallation events. Each of them is associated to a preceding muon which can in turn induce a shower containing a neutron cluster. This set of pairs muon (+ neutron cloud) + spallation event is used to evaluate the rejection power of a given combination of cut parameters, Δt, L_{LONG}, L_t and multiplicity.

- *Random sample*: to construct this sample we extract muons both from the hadron-producing and EM only samples, as described in Sec. 6.3.1, so that all types of primary muons passing through SK are considered. The relative fraction of 77% EM muons, extracted from the fit to data distributions previously described, is adopted. Each of these muons has to be associated to a random vertex which is generated within the SK tank and represents the uncorrelated low energy event. Since, due to the high muon rate, each low energy event can be associated to an average of ~ 60 preceding muons (within 30 s before the event time), each random vertex is paired with more than one muon: the distribution of number of muons associated to a low energy signal is extracted from data and used to randomly sample the number of muon (+ neutron cloud) paired with a given vertex. This set of pairs is necessary to evaluate the signal efficiency, namely the dead time introduced in the analysis as a consequence of spallation removal techniques. It is important to keep in mind that while one low energy event is defined as a unique vertex, it is paired with ~ 60 muons and it contributes therefore to ~ 60 pairs.

After the neutron clouds are identified as described above, each low energy event is paired with all neutron clouds observed up to 60 s before it. Selection cuts similar to the ones described in Section 6.4 are then applied on each pair; if one pair fails the cuts the low energy event is rejected.

7.4.2 Neutron cloud multiplicity classes and cloud extension

Given the increased neutron detection efficiency, the multiplicity bins defined in Table 6.3 for SK-IV are not suitable for SK-Gd. One of the first steps, if we aim to optimize neutron cloud cuts, is to redefine multiplicity ranges in order to accommodate the effect of increased neutron tagging performances without depleting each class from a sufficient statistics to be able to study its specific properties. Figure 7.8 shows a comparison of L_t distributions for different multiplicity classes as defined in pure water (left) and as modified for SK-Gd (right). The first class, which includes 2 to 5 neutron captures, remains the one showing a higher relative fraction of fake neutrons contributions, populating the tails of the L_t distribution.

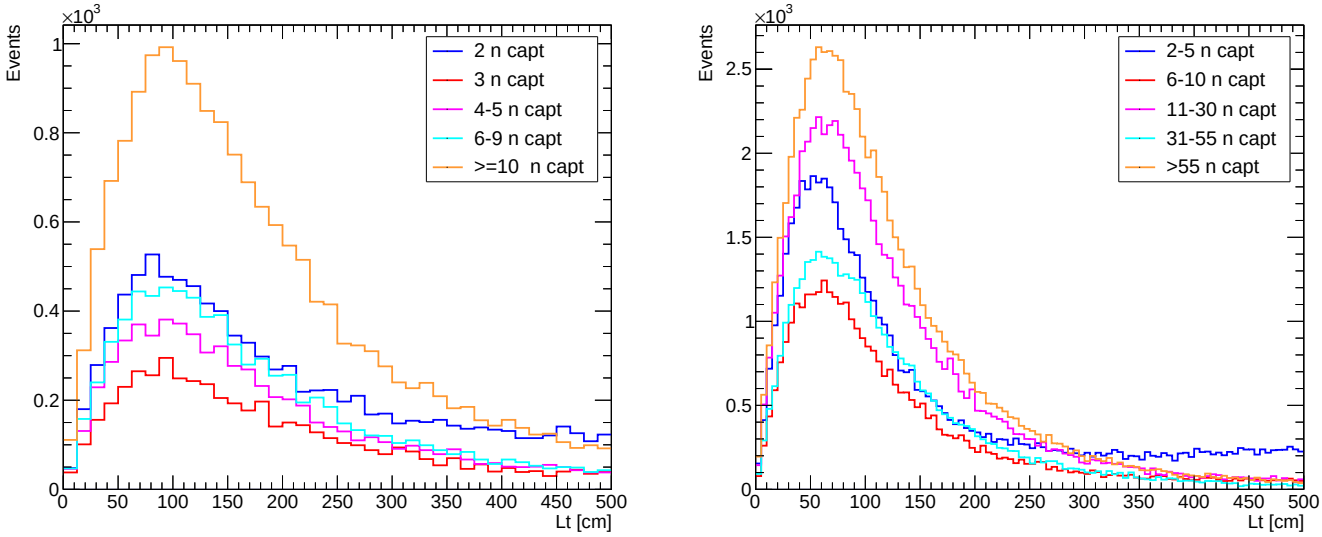


FIGURE 7.8: L_t distribution for different multiplicity classes. Right: multiplicity classes used for pure water phases, as described in Tab. 6.3. Left: redefinition to adapt to SK-Gd performances. Both plots are filled using neutron information from simulation.

Other important parameters to modify are the ones related to neutron cloud extension, due to both the increased resolution of the neutron vertex reconstruction and the redefinition of the neutron multiplicity bins. Since we want to investigate the spallation rejection performances over a grid of cut values, we define a collection of ellipses for each multiplicity category: in the coordinate system of Figure 4.3, L_{LONG} and L_t are taken to be larger and shorter radii of these ellipses. The region contained inside the two radii is where low energy signal will be tagged as spallation. Results from simulation here give us essential information on the position of true spallation events, and enable us to define ellipses which contain a given fraction of radioactive decays. Figure 7.9 shows sets of ellipses of various dimensions: the bigger one contains 99% of the spallation isotopes, the second one 98% and so on until 89%. Smaller ellipses were also tested but are not shown here. It is evident from Fig. 7.9 that the different multiplicity classes show peculiar properties: showers with a high number of neutrons are better localized in L_t and more shifted towards negative values of L_{LONG}. This last behavior is explained by the fact that negative longitudinal distances correspond to a point along the muon track closer to the entry point while the the neutron cloud center is observed later along the track: isotopes can in fact be produced by pion, gammas or other particles which travel less than neutrons and are found closer to the shower injection point.

7.4.3 Cut parameters optimization

Elliptical cuts based on L_t and L_{LONG} values are applied to each low energy event paired with a muon up to 60 s before. As described in Sec. 6.4, these selections are accompanied with Δt-depended cuts which aim to remove signals in a larger/smaller region of space depending if the event is closer/further in time with respect to the primary muon. This technique, together with elliptical cuts, assures a minimum increase of dead time while maintaining an high spallation rejection power. To adapt this part of the neutron cloud reduction method to SK-Gd, two time regions (Δt_{low}, Δt_{high}) need to be defined and for each of them the radius of the spherical region to be removed has to be determined ($r_{\Delta t_{low}}$,

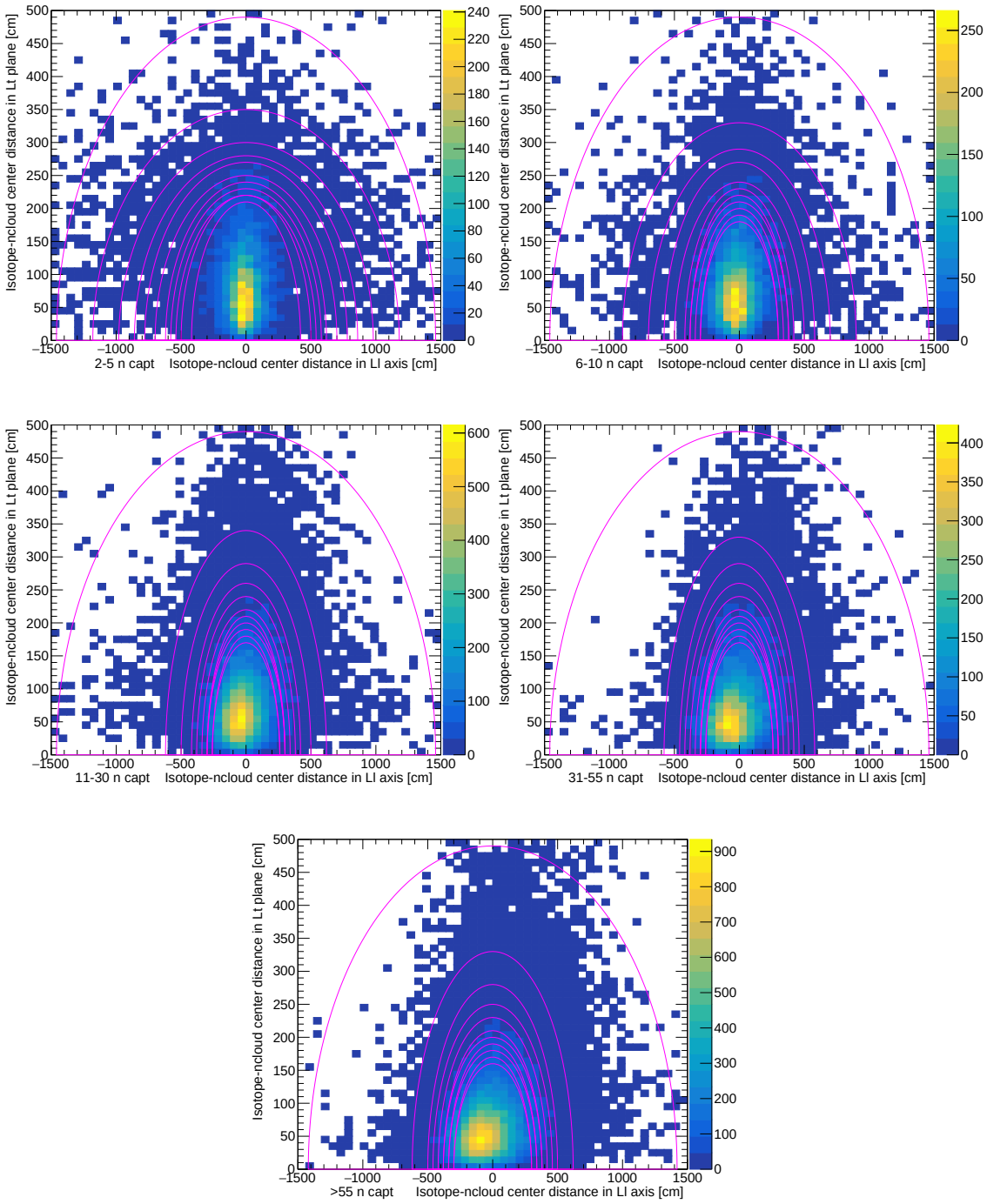


FIGURE 7.9: Each plot in this Figure groups events associated to a given multiplicity class, illustrated on the left of the x axis. The difference between isotope position and the barycenter of the associated neutron cloud is split into its two normal components within the coordinate system of Fig. 4.3: the x axis corresponds to the direction along the muon track and shows L_{LONG} of the isotope (neutron cloud barycenter correspond to the origin of the z axis in Fig. 4.3). The y axis illustrates the transverse position, L_t. In pink, ellipses with different dimensions are superimposed: the bigger one contains 99% of the isotopes, followed by a smaller one with 98% fraction and so on until 89% which corresponds to the smallest ellipse.

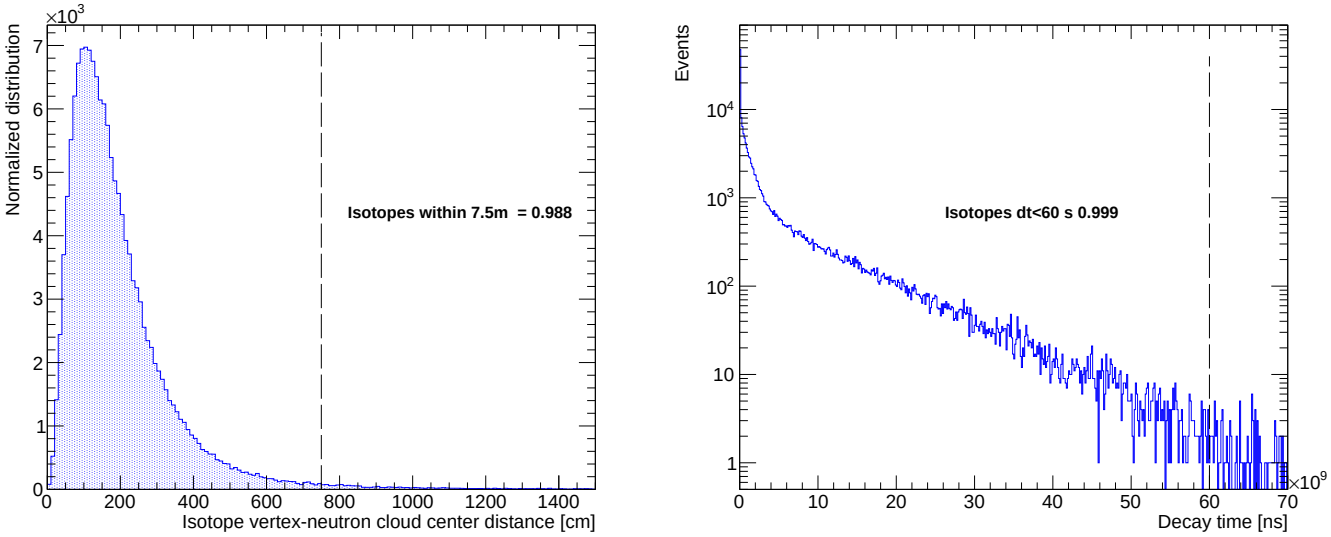


FIGURE 7.10: Left: distribution of the distance between neutron cloud barycenter and isotope vertex. Right: isotope decay time. Vertical dashed lines shows the chosen cut values. Efficiencies of the selections are also written in the Figures. For the plot on the left, only isotopes from simulation truly associated with a muon-induced neutron cloud are considered.

$r_{\Delta t_{high}}$). For each multiplicity class and each L_t-L_{LONG} pair defining one of the ellipses of Fig. 7.9, we build a grid of parameter scanning values as follows:

- Δt_{low} from 0.1 to 0.9 s and $r_{\Delta t_{low}}$ = 7.5 m
- Δt_{high} from 1 to 10 s and $r_{\Delta t_{low}}$ from 1 to 7 m

Before scanning over the grid of parameters, preliminary requirements are necessary to reduce computational time by removing accidental events: first of all a neutron cloud with at least two neutron candidates need to be present. Only 7.2% of the radioactive decays induced by primary muons passing through SK are not associated to the production of neutron clusters and cannot be removed with cloud identification techniques: among them we observe isotopes directly produced by muons or isotopes produced within small showers where less than two neutrons are captured. In addition, all isotope candidates are required to be within 7.5 m from the neutron cloud center and 60 s from the muon track. Efficiencies of these selections, shown in Figure 7.10, are sufficiently high to ensure that the vast majority of spallation events associated with a neutron cloud are considered.

Given a set of parameters (multiplicity, L_t, L_{LONG}, Δt_{low}, Δt_{high}, $r_{\Delta t_{low}}$, $r_{\Delta t_{high}}$), the goal is to predict the performance of neutron cloud rejection method by estimating the fraction of remaining real spallation events and uncorrelated signals removed, namely the amount of dead time introduced by the cuts. These values are extracted from the two samples, spallation and random sample, described in Sec. 7.4.1 and evaluated for all the points within the grid of parameters constructed. The results are shown in Figure 7.11: up to $\sim$ 85% of spallation isotopes could be removed at the cost of a 16% dead time. This performance is not that far from the one of the spallation cut used for the old solar analysis ($\sim$ 90% spallation rejection, 20% deadtime) [32]. This previous spallation cut was likelihood-based so the fact that, despite we are neglecting muon bundles contribution, we could achieve a similar performance with neutron clouds alone is very promising. The values obtained in

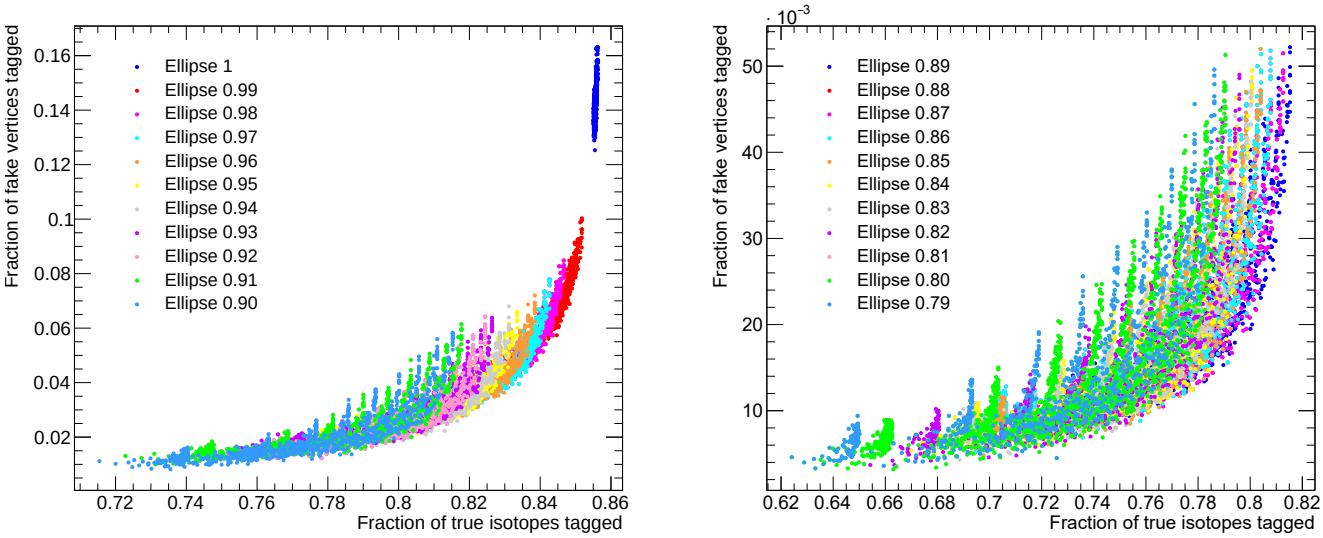


FIGURE 7.11: Both plots in the Figure shows the fraction of real isotopes correctly tagged as spallation on the x axis and the fraction of low energy events accidentally associated with previous muons which are wrongly identified as spallation signal on the y axis. The y axis corresponds to the fraction of dead time introduced by the neutron cloud cut. Each dot is associated to a given set of parameters: multiplicity, L_t, L_{LONG}, Δt_{low}, Δt_{high}, $r_{\Delta t_{low}}$, $r_{\Delta t_{high}}$. Different colors distinguish ellipses of various sizes: on the left plot are the ellipses as defined in Fig. 7.9 while on the right a set of smaller ellipses is considered and the relative fraction of isotopes included within them is listed in the color legend.

Figure 7.11 are used to evaluate neutron cloud performances for SK-Gd in terms of signal efficiency and fraction of spallation background removed.

7.5 Neutron cloud cut performances for SK-Gd

The results obtained from the scan over a grid of cut parameters shown in Fig. 7.11 need to be refined taking into account the fact that a real spallation event can be identified as background by the algorithm even if it is accidentally paired with a generic random muon. Given that each spallation candidate is associated with 60 preceding muons, even if only one is possibly its real parent, the random efficiency of forming an uncorrelated pair tagged as spallation by the algorithm is non negligible and in fact slightly increases the probability of a radioactive event to be removed. This effect need to be included in the curves of Fig. 7.11 in order to obtain conclusive results. The final curves showing the dead time (namely 1 - signal efficiency) as a function of the fraction of residual spallation are shown in Figure 7.12 for different sets of cuts. The dashed lines indicate the points in the cut parameter space for which the neutron cloud method shows better performances: indeed along these lines each dead time value is associated with the minimum fraction of residual spallation among all cuts. Each analysis in the low energy range of SK that suffers from spallation background contamination can take advantage of these results to set neutron cloud cut values according to the performances required for the analysis.

As a comparison with SK-IV results, in the solar search, keeping the same signal efficiency as with the pure water analysis, the fraction of isotopes removed passes from

53% to about 79%. Analogously, for SRN studies, for a signal efficiency of about 90% the amount of residual spallation is almost a factor 2 less.

At the maximum reduction power with the largest ellipse tested, we obtained that up to 88% of the isotopes can be removed with neutron cloud cut only. This results already shows the increased efficicncy that this reduction algorithm has with Gd-enhanced neutron detection. Although a non negligible fraction of spallation background is not produced in association with neutron clusters (7.2%), thus requiring the implementation of additional selections, neutron clouds within SK-Gd become the most efficient way to clean the signal from cosmogenic contaminations. Moreover, there is still a residual $\sim$ 5% of radioactive decays associated with a neutron cloud but not identified through this method. Most of these neutron are in fact still invisible in SK as a consequence of the low energy released after the capture. With an increased concentration of gadolinium, and therefore an even higher detection efficiency, this fraction of invisible neutrons can be further reduced and the algorithm rejection power will increase.

7.6 Additional isotopes contributions

Given the presence of new heavy nuclear species in pure water, one might wonder if additional unstable β-decaying isotopes, other than the ones listed in Tab. 4.2, can contaminate SK low energy signals. Since FLUKA simulation gives access to all the isotopes produced by the interactions of cosmic muons in water and their production rates, we can compare the spallation yields obtained in pure water with the case of Gd, and investigate whether new contributions have to be taken into account. The isotope production yields from simulations with 0.02% $Gd_2(SO_4)_3$ and the future concentration value of 0.2% are shown in Figure 6.10. Note that the yields of the isotopes already considered in SK-IV do not change when Gd is added. Moreover, for the current Gd concentration there are no additional β-decay contamination: only Gd-induced isotopes with production rate equivalent or lower of the ones listed in Tab. 4.2 are recorded. On the contrary, for the future expected concentration of 0.2% $Gd_2(SO_4)_3$ some of the heavy nuclei start to contribute to spallation backgrounds, even if in minor fraction. Table 7.3 lists the decay types, half lives and endpoint energies of these isotopes. Given the rate of water recirculation (120 tons per hour) in SK is extremely unlikely that isotopes with half lives of the order of hours, or higher, could represent a source of spallation background. The yields of ^{161} Gd, ^{143}Sm and ^{137}Nd are very low, making these nuclei a negligible source of contamination. As for ^{31}S and ^{30}P their yield starts growing with the increase of concentration of Gd and their endpoint energy is sufficiently high to make them visible; they are however among the isotopes with the lower production rates and their contribution is estimated to be minor. In conclusion, introducing gadolinium powder in pure water is not expected to introduce additional isotope contributions to the spallation background.

The ability of the FLUKA-simulation to predict the performances of neutron cloud cuts for SK-Gd is just one example of the potentials of this calculation framework. Being able to have a powerful tool that can guide the interpretation of such an important background as muon-induced spallation is remarkably valuable for SK. For extremely challenging searches like SRN, where major discoveries are expected in the next few years, spallation contamination has to be deeply understood, accurately modeled and drastically reduced. Spallation simulation firstly helped to design a new background removal method, neutron cloud identification, which is not only very efficient in SK-IV but will probably be the key technique for background removal in the current SK phase and in the future. These

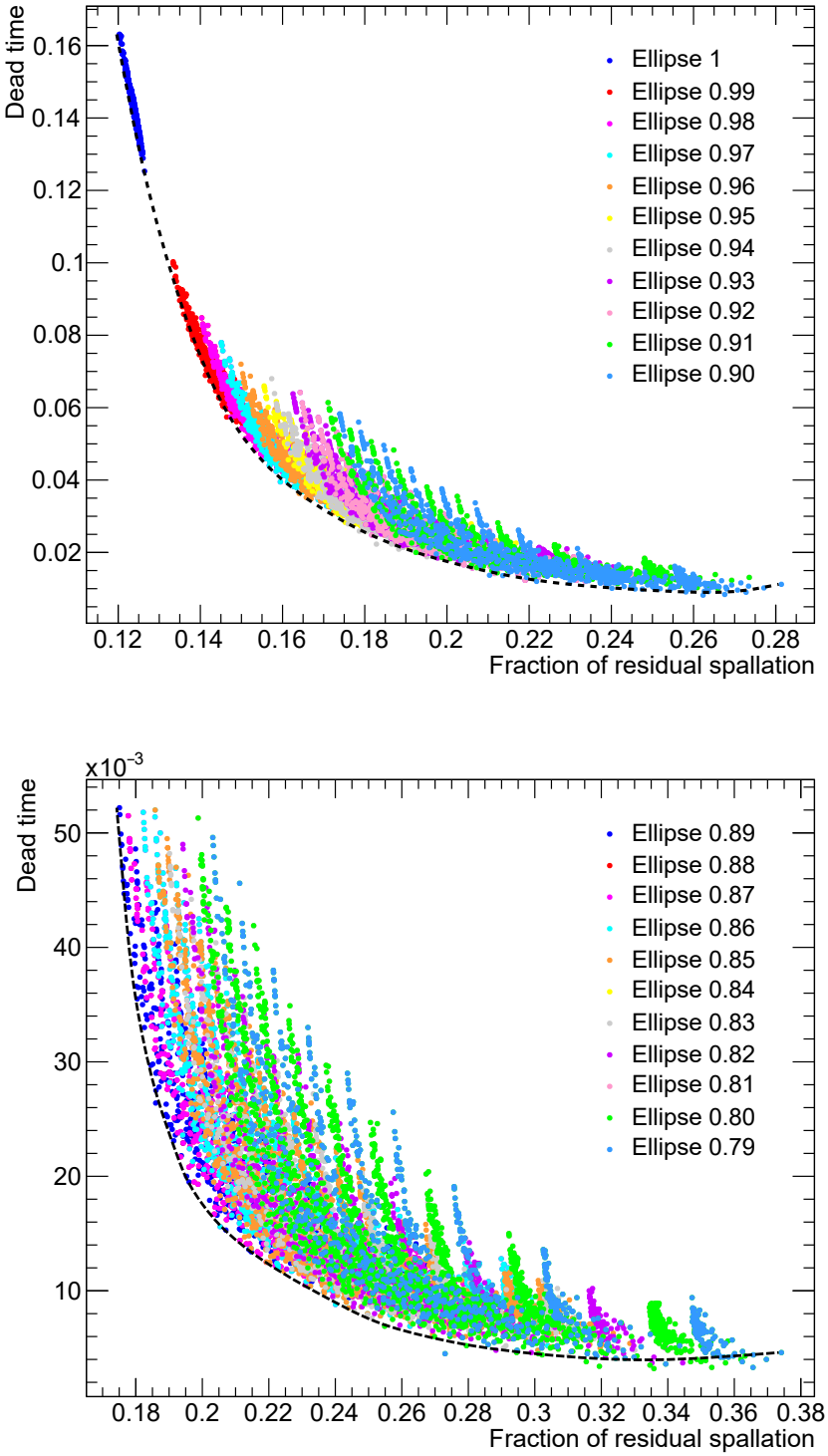


FIGURE 7.12: Both plots in the Figure shows the fraction of residual spallation on the x axis and the corresponding dead time introduced on the y axis. Each dot is associated to a given set of parameters: multiplicity, L_t, L_{LONG}, Δt_{low}, Δt_{high}, $r_{\Delta t_{low}}$, $r_{\Delta t_{high}}$. Different colors distinguish ellipses of various sizes: on the top plot are the ellipses as defined in Fig. 7.9 while on the bottom one a set of smaller ellipses is considered and the relative fraction of isotopes included within them is listed in the color legend. The dashed lines indicates the points in the cut parameters space with the best performances.

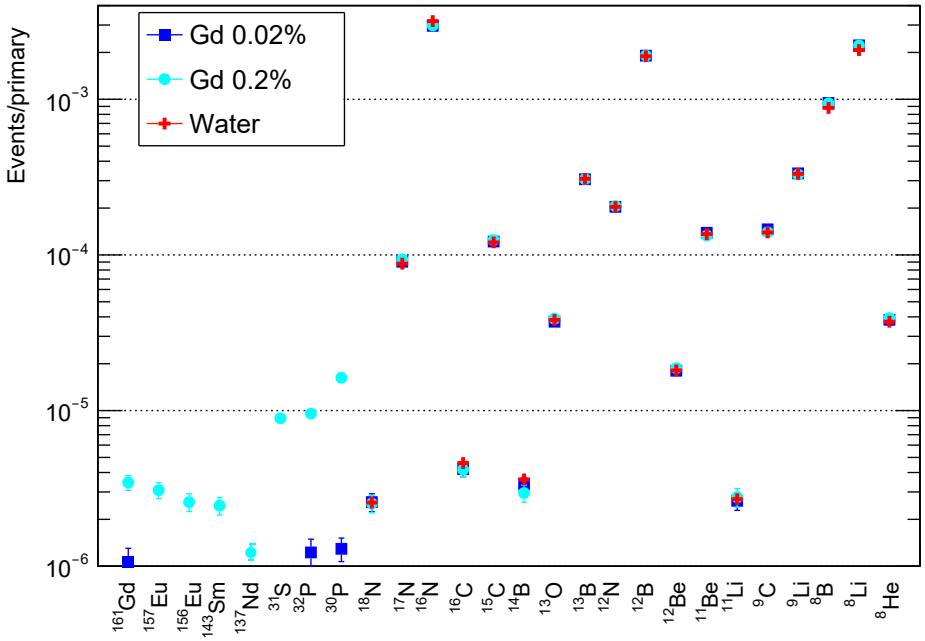


FIGURE 7.13: Isotope yields in SK. Contributions of the nuclei from Tab. 4.2, namely associated to muon-induced spallation in pure water, are shown in red. Possible contributions from heavier isotopes in the two concentrations of 0.02% $Gd_2(SO_4)_3$ and 0.2% $Gd_2(SO_4)_3$ are respectively in blue and cyan.

Isotope	Decay type	Half life	Endpoint En [MeV]
^{161}Gd	β^-	3.6 min	1.9
^{157}Eu	β^-	15.1 d	1.4
^{156}Eu	β^-	15.1 d	2.4
^{143}Sm	β^+	8.7 min	3.4
^{137}Nd	β^+	38.5 min	3.7
^{31}S	β^+	2.5 s	5.4
^{32}P	β^-	14.0 d	1.7
^{30}P	β^+	2.4 min	4.2

TABLE 7.3: Properties of the possible additional isotopes produced in FLUKA simulation by muons in water with Gd.

insights will help to shape the analysis strategies not only for SK-Gd but also for Hyper-Kamiokande and other future experiments.

Conclusions

Super-Kamiokande detector has the potential to provide invaluable knowledge about the Sun, galactic and extra-galactic supernovae and neutrino properties. One of its prominent goals is the detection of the supernova relic neutrino signal: this flux, originating from one of the most powerful and complex phenomena in the Universe, is decisive to unravel crucial details on cosmic history and stellar evolution. The extremely low rate of SRN signals, as well as the presence of important backgrounds that are difficult to model, makes SRN search extremely challenging. Muon-induced spallation is the dominant contamination below 20 MeV, about four orders of magnitude larger than the SRN flux. Despite the elaborated cuts introduced in SK, the residual background rate is still large. As a matter of fact, the current selections, which are essentially developed from empirical studies, can only remove a significant fraction of radioactive contamination at the price of significant signal loss. To not overly discard SRN rare events, more sophisticated cuts, based on a deeper understanding of spallation processes, need to be introduced. This is the reason why a major effort was undertaken to elaborate a complex Monte Carlo simulation accurately describing muon-induced spallation: this thesis presents the first complete and detailed simulation study dedicated to spallation background in Super-Kamiokande. A theoretical calculation of this kind, using state-of-the-art models and software, is able to shed light on the isotope production mechanisms, their properties and distributions, and points towards new identification techniques. In addition, a major advantage of the simulation is the possibility to distinguish different isotopes contributions: refined cuts can be implemented on an isotope-per-isotope basis and how the yields of different isotopes depend on energy threshold applied can be estimated. Especially for SRN analysis, where even with neutron tagging, some spallation isotopes decays yield IBD-like signatures, the ability to study individual isotopes is crucial.

The core of the simulation developed in this work is the FLUKA Monte Carlo package. It has been integrated with MUSIC code to predict cosmic muon flux at SK depth and SKDetSim for modeling the detector response. Extensive studies have been dedicated to the validation of simulation results against SK data: a good agreement was found when comparing cosmic muon distributions, their energy depositions as well as the geometrical characteristics of spallation events. The accuracy of our calculation was further confirmed by the estimation of isotope yields: results extracted from SK-IV data, appropriately corrected with selection efficiencies derived from simulation studies, agrees with FLUKA values within hadronic model uncertainties. Future progress in theoretical models could further improve the performances of the simulation. In particular, improving the agreement between the simulation and SK data would benefit from an ongoing effort to investigate muon bundles by e.g. adapting the MUPAGE code (originally designed for ANTARES) to SK studies [62].

Guided by simulation insights we better understand the production and properties of muon-induced showers and decays of radioactive isotopes, we revealed the different components of the cascades and the correlation between isotopes, their parent particles and the parent muon, details that were formerly unavailable on an isotope-by-isotope basis. These studies can be informative not only for Super-Kamiokande but in general for all water-Cherenkov detectors. An important characteristics of spallation showers is the

high multiplicity of neutrons. This motivates the introduction of a new method to locate muon-induced radioactivity: the identification of neutron clusters greatly helps to tag and remove spallation events.

The new spallation reduction technique described in Chapter 6 has been implemented in SK-IV SRN and solar analysis. However, its main limitation results from the poor detection efficiency of neutron capture signals, a 2.2 MeV gamma in water, and the short livetime of the neutron tagging trigger. This situation drastically changes in SK-Gd where the neutron capture visibility largely increases as a consequence of gadolinium doping. In the recently started phase of SK, with 0.01% of Gd concentration, the $\sim$ 45% neutron capture efficiency on Gd nuclei enhances the capabilities of neutron cloud algorithms, making this reduction technique a key component of spallation rejection. FLUKA simulation has played a central role in the optimization of these cuts for SK-Gd. In fact, it has been instrumental to adapt the neutron identification approach in the upgraded detector and to investigate the increased cut performances. The deeper understanding of shower development and the possibility to have a pure spallation sample thanks to the simulation allow to optimize the neutron cloud cut ability to remove a major fraction of radioactive isotopes while keeping high signal efficiency. It also instructs how the cut parameters can be varied to obtain a given result in terms of spallation rejection rate and dead time. An empirical optimization would have been more complicated due to the difficulties of pairing parent muons and induced radioactive decays, but also because a dark rate increase and asymmetry was observed in the first months of data in SK-Gd.

The detailed information provided by the simulation is not only valuable for SRN search but also for all the low energy analyses affected by spallation backgrounds. Therefore, other than neutrino-antineutrino separation, the increased neutron tagging efficiency in SK-Gd will significantly impact all low energy searches. Moreover, outcomes from simulation can help us to identify the main characteristics of residual isotopes, guiding the development of removal strategies for future analyses. This is especially important for the current SK-Gd phase, where the Gd concentration is limited and a large fraction of neutrons are still captured on protons.

The capabilities of FLUKA simulation can be further exploited. For SRN spectral studies, namely the analysis based on the fit of data with signal and background spectral shapes [130], obtaining isotope-per-isotope information, only accessible with simulation calculations, is extremely important: even with perfect neutron tagging, the impact of spallation cuts on ^{9}Li spectrum still need to be accurately modeled. In addition, the dependence of this specific isotope on the shower characteristics, on the muon energy depositions or the neutron multiplicity can be studied to develop targeted cuts. Moreover, with the aim of lowering more and more the energy threshold for SRN search, the analysis can greatly benefit from details on the individual isotope dependence on the energy cut, provided by the simulation.

Furthermore, thanks to the flexibility of the simulation framework, the calculation can be adapted for other future experiments. In particular, in Hyper-Kamiokande, where the mountain overburden is considerably less than in SK, spallation removal techniques will be even more crucial for SRN observation: this groundbreaking discovery depends on our ability to handle this background.

To conclude, the work of this thesis provided a unique and powerful tool that can be of great help not only for several analysis in SK-Gd but also for future experiments. Understanding and modeling spallation is the first step towards an improvement of our background reduction strategies, which would in turn open the door for new discoveries.

Bibliography

[1] Kamiokande Collaboration et al. *New Methods and Simulations for Cosmogenic Induced Spallation Removal in Super-Kamiokande-IV*. 2021. arXiv: 2112.00092 [hep-ex].

[2] K. Zuber. *Neutrino Physics*. 2012.

[3] J. Chadwick. "Verh. der Deutschen Physikalischen Ges. 16 383". In: (1914).

[4] J. Chadwick. "Versuch einer Theorie der Strahlen I Zeitschrift fr Physik". In: 88, Issue 3, page 161 (1934).

[5] B. Pontecorvo. "Chalk River Lab, PD-205 report". In: (1946).

[6] Cowan C. Reines F. "The Neutrino, Nature, Volume 178, page 446". In: (1956).

[7] M. Goldhaber et al. "Helicity of Neutrinos". In: *Phys. Rev.* 109 (3 1958), pp. 1015–1017. DOI: 10.1103/PhysRev.109.1015. URL: https://link.aps.org/doi/10.1103/PhysRev.109.1015.

[8] G. Danby et al. "Observation of High-Energy Neutrino Reactions and the Existence of Two Kinds of Neutrinos". In: *Phys. Rev. Lett.* 9 (1 1962), pp. 36–44. DOI: 10.1103/PhysRevLett.9.36. URL: https://link.aps.org/doi/10.1103/PhysRevLett.9.36.

[9] M. L. Perl et al. "Evidence for Anomalous Lepton Production in $e^+ - e^-$ Annihilation". In: *Phys. Rev. Lett.* 35 (22 1975), pp. 1489–1492. DOI: 10.1103/PhysRevLett.35.1489. URL: https://link.aps.org/doi/10.1103/PhysRevLett.35.1489.

[10] K. Kodama et al. "Observation of tau neutrino interactions". In: *Physics Letters B* 504.3 (2001), pp. 218 –224. ISSN: 0370-2693. DOI: https://doi.org/10.1016/S0370-2693(01)00307-0. URL: http://www.sciencedirect.com/science/article/pii/S0370269301003070.

[11] "Electroweak parameters of the Z0 resonance and the standard model". In: *Physics Letters B* 276.1 (1992), pp. 247 –253. ISSN: 0370-2693. DOI: https://doi.org/10.1016/0370-2693(92)90572-L. URL: http://www.sciencedirect.com/science/article/pii/037026939290572L.

[12] Sakata S. Maki Z. "Remarks on the unified model of elementary particles". In: *Prog. Theor. Phys. 28 870-880.* (1962).

[13] Raymond Davis. "Solar Neutrinos. II. Experimental". In: *Phys. Rev. Lett.* 12 (11 1964), pp. 303–305. DOI: 10.1103/PhysRevLett.12.303. URL: https://link.aps.org/doi/10.1103/PhysRevLett.12.303.

[14] Raymond Davis et al. "Search for Neutrinos from the Sun". In: *Phys. Rev. Lett.* 20 (21 1968), pp. 1205–1209. DOI: 10.1103/PhysRevLett.20.1205. URL: https://link.aps.org/doi/10.1103/PhysRevLett.20.1205.

[15] Y. Fukuda et al. "Evidence for Oscillation of Atmospheric Neutrinos". In: *Physical Review Letters* 81.8 (1998), pp. 1562–1567. ISSN: 1079-7114. DOI: 10.1103/physrevlett.81.1562. URL: http://dx.doi.org/10.1103/PhysRevLett.81.1562.

[16] Q. R. Ahmad et al. "Measurement of the Rate of $\nu_e + d \rightarrow p + p + e^-$ Interactions Produced by ^8B Solar Neutrinos at the Sudbury Neutrino Observatory". In: *Phys. Rev. Lett.* 87 (7 2001), p. 071301. DOI: 10.1103/PhysRevLett.87.071301. URL: https://link.aps.org/doi/10.1103/PhysRevLett.87.071301.

[17] J. A. Formaggio et al. "From eV to EeV: Neutrino cross sections across energy scales". In: *Reviews of Modern Physics* 84.3 (2012), pp. 1307 –1341. ISSN: 1539-0756. DOI: 10.1103/revmodphys.84.1307. URL: http://dx.doi.org/10.1103/RevModPhys.84.1307.

[18] John F. Beacom. "The Diffuse Supernova Neutrino Background". In: *Annual Review of Nuclear and Particle Science* 60.1 (2010), pp. 439–462. DOI: 10.1146/annurev.nucl.010909.083331. eprint: https://doi.org/10.1146/annurev.nucl.010909.083331. URL: https://doi.org/10.1146/annurev.nucl.010909.083331.

[19] G. Fogli et al. "Probing supernova shock waves and neutrino flavor transitions in next-generation water Cerenkov detectors". In: *Journal of Cosmology and Astroparticle Physics* 2005 (Dec. 2004). DOI: 10.1088/1475-7516/2005/04/002.

[20] P. Vogel et al. "Angular distribution of neutron inverse beta decay, anti-neutrino(e) + p —> e+ + n". In: *Phys. Rev. D* 60 (1999), p. 053003. DOI: 10.1103/PhysRevD.60.053003. arXiv: hep-ph/9903554.

[21] A Oralbaev et al. "The inverse beta decay: a study of cross section". In: *Journal of Physics: Conference Series* 675.1 (2016), p. 012003. DOI: 10.1088/1742-6596/675/1/012003. URL: https://doi.org/10.1088/1742-6596/675/1/012003.

[22] L Wan. *Experimental Studies on Low Energy Electron Antineutrinos and Related Physics*. Tsinghua University, December 2018.

[23] C. Giunti et al. "Fundamentals of Neutrino Physics and Astrophysics". In: 2007.

[24] E W Otten et al. "Neutrino mass limit from tritium beta decay". In: *Reports on Progress in Physics* 71.8 (2008), p. 086201. DOI: 10.1088/0034-4885/71/8/086201. URL: https://doi.org/10.1088/0034-4885/71/8/086201.

[25] Ettore Majorana et al. "A symmetric theory of electrons and positrons". In: *Ettore Majorana Scientific Papers: On occasion of the centenary of his birth*. Ed. by Giuseppe Franco Bassani. Berlin, Heidelberg: Springer Berlin Heidelberg, 2006, pp. 201–233. ISBN: 978-3-540-48095-2. DOI: 10.1007/978-3-540-48095-2_10. URL: https://doi.org/10.1007/978-3-540-48095-2_10.

[26] Ashley Michael Timmons. "Introduction to Neutrino Physics". In: *Search for Sterile Neutrinos with the MINOS Long-Baseline Experiment*. Cham: Springer International Publishing, 2017, pp. 1–23. ISBN: 978-3-319-63769-3. DOI: 10.1007/978-3-319-63769-3_1. URL: https://doi.org/10.1007/978-3-319-63769-3_1.

[27] Dusan Budjas et al. "Pulse shape discrimination studies with a Broad-Energy Germanium detector for signal identification and background suppression in the GERDA double beta decay experiment". In: *Journal of Instrumentation* 4 (Sept. 2009). DOI: 10.1088/1748-0221/4/10/P10007.

[28] J. Argyriades et al. "Measurement of the double-beta decay half-life of Nd150 and search for neutrinoless decay modes with the NEMO-3 detector". In: *Physical Review C* 80.3 (2009). ISSN: 1089-490X. DOI: 10.1103/physrevc.80.032501. URL: http://dx.doi.org/10.1103/PhysRevC.80.032501.

[29] L. Wolfenstein. "Neutrino oscillations in matter". In: *Phys. Rev. D* 17 (9 1978), pp. 2369–2374. DOI: 10.1103/PhysRevD.17.2369. URL: https://link.aps.org/doi/10.1103/PhysRevD.17.2369.

[30] A. Y. Smirnov. "The Mikheyev-Smirnov-Wolfenstein (MSW) Effect". In: *International Conference on History of the Neutrino: 1930-2018*. 2019. arXiv: 1901.11473 [hep-ph].

[31] John N. Bahcall et al. "Solar Models: Current Epoch and Time Dependences, Neutrinos, and Helioseismological Properties". In: *The Astrophysical Journal* 555.2 (2001), pp. 990–1012. ISSN: 1538-4357. DOI: 10.1086/321493. URL: http://dx.doi.org/10.1086/321493.

[32] Yokozawa T. *Precision solar neutrino measurements with Super-Kamiokande-IV*. Doctoral thesis. University of Tokyo, Dec. 2012.

[33] Nakano Y. ^{8}*B solar neutrino spectrum measurement using Super-Kamiokande IV*. Doctoral thesis. University of Tokyo, Feb. 2016.

[34] E. G. Adelberger et al. "Solar fusion cross sections. II. The pp chain and CNO cycles". In: *Rev. Mod. Phys.* 83 (1 2011), pp. 195–245. DOI: 10.1103/RevModPhys.83.195. URL: https://link.aps.org/doi/10.1103/RevModPhys.83.195.

[35] K. Hirata et al. "Observation of a neutrino burst from the supernova SN1987A". In: *Phys. Rev. Lett.* 58 (14 1987), pp. 1490–1493. DOI: 10.1103/PhysRevLett.58.1490. URL: https://link.aps.org/doi/10.1103/PhysRevLett.58.1490.

[36] R. M. Bionta et al. "Observation of a neutrino burst in coincidence with supernova 1987A in the Large Magellanic Cloud". In: *prl* 58.14 (Apr. 1987), pp. 1494–1496. DOI: 10.1103/PhysRevLett.58.1494.

[37] E. N. Alekseev et al. "Detection of the Neutrino Signal From SN1987A in the LMC Using the Inr Baksan Underground Scintillation Telescope". In: *Phys. Lett. B* 205 (1988), pp. 209–214. DOI: 10.1016/0370-2693(88)91651-6.

[38] G. A. Tammann et al. "The Galactic Supernova Rate". In: *apjs* 92 (June 1994), p. 487. DOI: 10.1086/192002.

[39] Yosuke ASHIDA. *Measurement of Neutrino and Antineutrino Neutral-Current Quasielastic-like Interactions and Applications to Supernova Relic Neutrino Searches*. Doctoral thesis. High Energy Physics Group Division of Physics and Astronomy, Kyoto University, Feb 2020.

[40] Kei Kotake et al. "Explosion mechanism, neutrino burst and gravitational wave in core-collapse supernovae". In: *Reports on Progress in Physics* 69.4 (2006), pp. 971–1143. ISSN: 1361-6633. DOI: 10.1088/0034-4885/69/4/r03. URL: http://dx.doi.org/10.1088/0034-4885/69/4/R03.

[41] T. Padmanabhan. *Theoretical Astrophysics*. Vol. 2. Cambridge University Press, 2001. DOI: 10.1017/CBO9780511840159.

[42] M. Malek et al. "Search for Supernova Relic Neutrinos at Super-Kamiokande". In: *Physical Review Letters* 90.6 (2003). ISSN: 1079-7114. DOI: 10.1103/physrevlett.90.061101. URL: http://dx.doi.org/10.1103/PhysRevLett.90.061101.

[43] Amol Dighe et al. *SISSA/EP/98/132 hep-ph/9812244 The Physics of Relic Neutrinos*. 1998.

[44] Shinichiro Ando et al. "Relic neutrino background from cosmological supernovae". In: *New Journal of Physics* 6 (2004), pp. 170–170. ISSN: 1367-2630. DOI: 10.1088/1367-2630/6/1/170. URL: http://dx.doi.org/10.1088/1367-2630/6/1/170.

[45] Zhang H. *Study of low energy electron anti-neutrinos at Super-Kamiokande IV*. Doctoral thesis. Beijing, Tsinghua University, 2012.

[46] Zhang Y. *Experimental studies of supernova relic neutrinos at Super-Kamiokande-IV*. Doctoral thesis. Beijing, Tsinghua University, 2015.

[47] Hagiwara Kaito. *Search for Astronomical Neutrinos from Blazar TXS0506+056 in Super-Kamiokande*. Doctoral thesis. Okayama University, March 2020.

[48] Yoichiro Suzuki. "The Super-Kamiokande experiment". In: *The European Physical Journal C* 79 (Apr. 2019). DOI: 10.1140/epjc/s10052-019-6796-2.

[49] John Beacom et al. "Antineutrino Spectroscopy with Large Water Cherenkov Detectors". In: *Physical review letters* 93 (Oct. 2004), p. 171101. DOI: 10.1103/PhysRevLett.93.171101.

[50] H. Zhang et al. "Supernova Relic Neutrino search with neutron tagging at Super-Kamiokande-IV". In: *Astroparticle Physics* 60 (2015), pp. 41–46. ISSN: 0927-6505. DOI: 10.1016/j.astropartphys.2014.05.004. URL: http://dx.doi.org/10.1016/j.astropartphys.2014.05.004.

[51] K. Bays et al. "Supernova relic neutrino search at super-Kamiokande". In: *Physical Review D* 85.5 (2012). ISSN: 1550-2368. DOI: 10.1103/physrevd.85.052007. URL: http://dx.doi.org/10.1103/PhysRevD.85.052007.

[52] Y. Zhang et al. "First measurement of radioactive isotope production through cosmic-ray muon spallation in Super-Kamiokande IV". In: *Phys. Rev. D* 93.1 (2016), p. 012004. DOI: 10.1103/PhysRevD.93.012004. arXiv: 1509.08168 [hep-ex].

[53] Shirley Li et al. "First calculation of cosmic-ray muon spallation backgrounds for MeV astrophysical neutrino signals in Super-Kamiokande first calculation of cosmic-ray muon". In: *Physical Review C* 89 (Feb. 2014). DOI: 10.1103/PhysRevC.89.045801.

[54] Hiroyuki Sekiya. "The Super-Kamiokande Gadolinium Project". In: *PoS* ICHEP2016 (2016), p. 982. DOI: 10.22323/1.282.0982.

[55] Shunsaku Horiuchi et al. "Diffuse supernova neutrino background is detectable in Super-Kamiokande". In: *Phys. Rev. D* 79 (8 2009), p. 083013. DOI: 10.1103/PhysRevD.79.083013. URL: https://link.aps.org/doi/10.1103/PhysRevD.79.083013.

[56] A. Gando et al. "SEARCH FOR EXTRATERRESTRIAL ANTINEUTRINO SOURCES WITH THE KamLAND DETECTOR". In: *The Astrophysical Journal* 745.2 (2012), p. 193. ISSN: 1538-4357. DOI: 10.1088/0004-637x/745/2/193. URL: http://dx.doi.org/10.1088/0004-637X/745/2/193.

[57] Kenichiro Nakazato et al. "SPECTRUM OF THE SUPERNOVA RELIC NEUTRINO BACKGROUND AND METALLICITY EVOLUTION OF GALAXIES". In: *The Astrophysical Journal* 804.1 (2015), p. 75. ISSN: 1538-4357. DOI: 10.1088/0004-637x/804/1/75. URL: http://dx.doi.org/10.1088/0004-637X/804/1/75.

[58] K. Abe et al. *Letter of Intent: The Hyper-Kamiokande Experiment — Detector Design and Physics Potential —*. 2011. arXiv: 1109.3262 [hep-ex].

[59] K. Abe et al. "Solar neutrino results in Super-Kamiokande-III". In: *Phys. Rev. D* 83 (5 2011), p. 052010. DOI: 10.1103/PhysRevD.83.052010. URL: https://link.aps.org/doi/10.1103/PhysRevD.83.052010.

[60] S. Cecchini et al. "Atmospheric muons: experimental aspects". In: *Geoscientific Instrumentation, Methods and Data Systems* 1.2 (2012), pp. 185–196. DOI: 10.5194/gi-1-185-2012. URL: https://gi.copernicus.org/articles/1/185/2012/.

[61] Y. Becherini et al. "A parameterisation of single and multiple muons in the deep water or ice". In: *Astroparticle Physics* 25.1 (2006), pp. 1 –13. ISSN: 0927-6505. DOI: https://doi.org/10.1016/j.astropartphys.2005.10.005. URL: http://www.sciencedirect.com/science/article/pii/S092765050500157X.

[62] G. Carminati et al. "Atmospheric MUons from PArametric formulas: a fast GEnerator for neutrino telescopes (MUPAGE)". In: *Computer Physics Communications* 179.12 (2008), pp. 915 –923. ISSN: 0010-4655. DOI: https://doi.org/10.1016/j.cpc.2008.07.014. URL: http://www.sciencedirect.com/science/article/pii/S001046550800266X.

[63] J. Beringer et al. "Review of Particle Physics". In: *Phys. Rev. D 86* (July 2012). DOI: 10.1103/PhysRevD.86.010001.

[64] Nikhil Vittal Shetty. *Study of Particle Transport in a High Power Spallation Target for an Accelerator Driven Transmutation System*. PhD thesis. RWTH Aachen University, 2013.

[65] Giovani Dalla Valle Garcia. *Calculation of cosmic-ray muon spallation backgrounds with Geant4 for Supernova Relic Neutrino signals in Super-Kamiokande*. Master thesis. Laboratoire Leprince-Ringuet Ecole polytechnique, 2020.

[66] R. Serber. "Nuclear Reactions at High Energies". In: *Phys. Rev.* 72 (1947), pp. 1114–1115. DOI: 10.1103/PhysRev.72.1114.

[67] G.J. Russel. "Spallation physics. An Overview". In: *ICANS-XI International Collaboration on Advanced Neutron Sources* (1991).

[68] V.S. Barashenkov et al. "Fission and decay of excited nuclei". In: *Nuclear Physics A* 206.1 (1973), pp. 131 –144. ISSN: 0375-9474. DOI: https://doi.org/10.1016/0375-9474(73)90611-8. URL: http://www.sciencedirect.com/science/article/pii/0375947473906118.

[69] G.S. Bauer. "Physics and technology of spallation neutron sources". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 463.3 (2001). Accelerator driven systems, pp. 505 –543. ISSN: 0168-9002. DOI: https://doi.org/10.1016/S0168-9002(01)00167-X. URL: http://www.sciencedirect.com/science/article/pii/S016890020100167X.

[70] Joseph Cugnon. "Cascade Models and Particle Production: A Comparison". In: *Particle Production in Highly Excited Matter*. Ed. by Hans H. Gutbrod et al. Boston, MA: Springer US, 1993, pp. 271–293. ISBN: 978-1-4615-2940-8. DOI: 10.1007/978-1-4615-2940-8_12. URL: https://doi.org/10.1007/978-1-4615-2940-8_12.

[71] V. Weisskopf. "Statistics and Nuclear Reactions". In: *Phys. Rev.* 52 (4 1937), pp. 295–303. DOI: 10.1103/PhysRev.52.295. URL: https://link.aps.org/doi/10.1103/PhysRev.52.295.

[72] V.A. Karnaukhov et al. "Multifragmentation and nuclear phase transitions (liquid-fog and liquid-gas)". In: *Nuclear Physics A* 734 (2004), pp. 520 –523. ISSN: 0375-9474. DOI: https://doi.org/10.1016/j.nuclphysa.2004.01.095. URL: http://www.sciencedirect.com/science/article/pii/S0375947404001150.

[73] J. Robb Grover et al. "Dissipation of Energy and Angular Momentum by Emission of Neutrons and Gamma Rays". In: *Phys. Rev.* 157 (4 1967), pp. 814–823. DOI: 10.1103/PhysRev.157.814. URL: https://link.aps.org/doi/10.1103/PhysRev.157.814.

[74] E. Strumberger et al. "A more detailed calculation of particle evaporation and fission of compound nuclei". In: *Nuclear Physics A* 529.3 (1991), pp. 522 –564. ISSN: 0375-9474. DOI: https://doi.org/10.1016/0375-9474(91)90584-S. URL: http://www.sciencedirect.com/science/article/pii/037594749190584S.

[75] S. Fukuda et al. "The Super-Kamiokande detector". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 501.2 (2003), pp. 418–462. ISSN: 0168-9002. DOI: https://doi.org/10.1016/S0168-9002(03)00425-X. URL: https://www.sciencedirect.com/science/article/pii/S016890020300425X.

[76] K. Abe et al. "The T2K experiment". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 659.1 (2011), pp. 106–135. ISSN: 0168-9002. DOI: https://doi.org/10.1016/j.nima.2011.06.067. URL: https://www.sciencedirect.com/science/article/pii/S0168900211011910.

[77] A. Suzuki et al. "Improvement of 20 in. diameter photomultiplier tubes". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 329.1 (1993), pp. 299–313. ISSN: 0168-9002. DOI: https://doi.org/10.1016/0168-9002(93)90949-I. URL: https://www.sciencedirect.com/science/article/pii/016890029390949I.

[78] K. Abe et al. "Calibration of the Super-Kamiokande detector". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 737 (2014), pp. 253–272. ISSN: 0168-9002. DOI: 10.1016/j.nima.2013.11.081. URL: http://dx.doi.org/10.1016/j.nima.2013.11.081.

[79] S. Yamada et al. "Commissioning of the new electronics and online system for the Super-Kamiokande experiment". In: *IEEE Trans. Nucl. Sci.* 57 (2010). Ed. by Sascha Marc Schmeling, pp. 428–432. DOI: 10.1109/TNS.2009.2034854.

[80] H. Nishino et al. "High-speed charge-to-time converter ASIC for the Super-Kamiokande detector". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 610.3 (2009), pp. 710–717. ISSN: 0168-9002. DOI: https://doi.org/10.1016/j.nima.2009.09.026. URL: https://www.sciencedirect.com/science/article/pii/S0168900209017495.

[81] Giada Carminati. "The new Wide-band Solar Neutrino Trigger for Super-Kamiokande". In: *Physics Procedia* 61 (2015). 13th International Conference on Topics in Astroparticle and Underground Physics, TAUP 2013, pp. 666–672. ISSN: 1875-3892. DOI: https://doi.org/10.1016/j.phpro.2014.12.068. URL: https://www.sciencedirect.com/science/article/pii/S1875389214006816.

[82] Muhammad Elnimr. "Low Energy 8 B Solar Neutrinos with the Wideband Intelligent Trigger at Super-Kamiokande". In: *Journal of Physics: Conference Series* 888 (Sept. 2017), p. 012189. DOI: 10.1088/1742-6596/888/1/012189.

[83] J. Hosaka et al. "Solar neutrino measurements in Super-Kamiokande-I". In: *Physical Review D* 73.11 (2006). ISSN: 1550-2368. DOI: 10.1103/physrevd.73.112001. URL: http://dx.doi.org/10.1103/PhysRevD.73.112001.

[84] Smy M. "Low Energy Event Reconstruction and Selection in Super-Kamiokande III". In: volume 5 (3-11 July 2007). Proceedings, 30th International Cosmic Ray Conference (ICRC 2007): Merida, Yucatan, Mexico, pp. 1279–1282.

[85] R. Brun et al. "GEANT3". In: (Sept. 1987).

[86] Y. Hayato. "A Neutrino Interaction Simulation Program Library NEUT". In: *Acta Physica Polonica B* 40.9 (Sept. 2009), p. 2477.

[87] C. Zeitnitz et al. "The GEANT-CALOR interface and benchmark calculations of ZEUS test calorimeters". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 349.1 (1994), pp. 106–111. ISSN: 0168-9002. DOI: https://doi.org/10.1016/0168-9002(94)90613-0. URL: https://www.sciencedirect.com/science/article/pii/0168900294906130.

[88] Conner Z. *A study of solar neutrinos using the Super-Kamiokande detector.* University of Maryland, 1997.

[89] Shantanu Desai. "High energy neutrino astrophysics with Super-Kamiokande". PhD thesis. Boston U., 2004.

[90] Kirk Ryan Bays. "Search for the Diffuse Supernova Neutrino Background at Super-Kamiokande". PhD thesis. University of California Irvine, 2012.

[91] E Blaufuss et al. "16N as a calibration source for Super-Kamiokande". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 458.3 (2001), pp. 638–649. ISSN: 0168-9002. DOI: https://doi.org/10.1016/S0168-9002(00)00900-1. URL: https://www.sciencedirect.com/science/article/pii/S0168900200009001.

[92] Chenyuan Xu. "Current status of SK-Gd project and EGADS". In: *Journal of Physics: Conference Series* 718 (May 2016), p. 062070. DOI: 10.1088/1742-6596/718/6/062070.

[93] Takaaki Mori. "Status of the Super-Kamiokande gadolinium project". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 732 (2013). Vienna Conference on Instrumentation 2013, pp. 316–319. ISSN: 0168-9002. DOI: https://doi.org/10.1016/j.nima.2013.06.074. URL: https://www.sciencedirect.com/science/article/pii/S0168900213009078.

[94] Lauren Helen Victoria Anthony. *A model independent determination of the π^0 background in Super-Kamiokande in the T2K ν_e appearance measurement.* Doctoral thesis. Oliver Lodge Laboratory, University Of Liverpool, Feb. 2020.

[95] Xia Junjie. "Upgrade and Calibration of Super-Kamiokande Inner Photodetectors". PhD thesis. University of Tokyo, 2019.

[96] M Nakahata et al. "Calibration of Super-Kamiokande using an electron LINAC: The Super-Kamiokande Collaboration". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 421.1 (1999), pp. 113–129. ISSN: 0168-9002. DOI: https://doi.org/10.1016/S0168-9002(98)01200-5. URL: https://www.sciencedirect.com/science/article/pii/S0168900298012005.

[97] Yusuke Koshio. "Study of Solar Neutrinos at Super-Kamiokande". PhD thesis. University of Tokyo, 1998.

[98] Shirley Li et al. "Spallation Backgrounds in Super-Kamiokande Are Made in Muon-Induced Showers". In: *Physical Review D* 91 (Mar. 2015). DOI: 10.1103/PhysRevD.91.105005.

[99] A. Ferrari et al. *FLUKA: A multi-particle transport code*. DOI: CERN-2005-10(2005),INFN-TC-05-11,SLAC-R-773..

[100] S. Abe et al. "Production of Radioactive Isotopes through Cosmic Muon Spallation in KamLAND". In: *Phys. Rev. C* 81 (2010), p. 025807. DOI: 10.1103/PhysRevC.81.025807. arXiv: 0907.0066 [hep-ex].

[101] G.B. Bellini et al. "Cosmogenic Backgrounds in Borexino at 3800 m water-equivalent depth". In: *Journal of Cosmology and Astroparticle Physics* 2013 (Apr. 2013). DOI: 10.1088/1475-7516/2013/08/049.

[102] Shirley Li et al. "Tagging Spallation Backgrounds with Showers in Water-Cherenkov Detectors". In: *Physical Review D* 92 (Aug. 2015). DOI: 10.1103/PhysRevD.92.105033.

[103] P. Antonioli et al. "A three-dimensional code for muon propagation through the rock: MUSIC". In: *Astroparticle Physics* 7.4 (1997), pp. 357 –368. ISSN: 0927-6505. DOI: https://doi.org/10.1016/S0927-6505(97)00035-2. URL: http://www.sciencedirect.com/science/article/pii/S0927650597000352.

[104] V.A. Kudryavtsev et al. "Narrow muon bundles from muon pair production in rock". In: *Physics Letters B* 471.2 (1999), pp. 251 –256. ISSN: 0370-2693. DOI: https://doi.org/10.1016/S0370-2693(99)01378-7. URL: http://www.sciencedirect.com/science/article/pii/S0370269399013787.

[105] V.A. Kudryavtsev. "Muon simulation codes MUSIC and MUSUN for underground physics". In: *Computer Physics Communications* 180.3 (2009), pp. 339 –346. ISSN: 0010-4655. DOI: https://doi.org/10.1016/j.cpc.2008.10.013. URL: http://www.sciencedirect.com/science/article/pii/S0010465508003640.

[106] Thomas K. Gaisser et al. *Cosmic Rays and Particle Physics*. 2nd ed. Cambridge University Press, 2016. DOI: 10.1017/CBO9781139192194.

[107] Thomas K. Gaisser et al. "High-energy cosmic rays". In: *Nuclear Physics A* 777 (2006). Special Isseu on Nuclear Astrophysics, pp. 98 –110. ISSN: 0375-9474. DOI: https://doi.org/10.1016/j.nuclphysa.2005.01.024. URL: http://www.sciencedirect.com/science/article/pii/S0375947405000540.

[108] Alfred Tang et al. "Muon simulations for Super-Kamiokande, KamLAND, and CHOOZ". In: *Physical Review. D, Particles Fields* 74.5 (Sept. 2006). ISSN: 0556-2821. DOI: 10.1103/PHYSREVD.74.053007.

[109] Y. Gando et al. "Search for $\bar{\nu}_e$ from the Sun at Super-Kamiokande-I". In: *Physical Review Letters* 90.17 (2003). ISSN: 1079-7114. DOI: 10.1103/physrevlett.90.171302. URL: http://dx.doi.org/10.1103/PhysRevLett.90.171302.

[110] Y. Fukuda et al. "Evidence for Oscillation of Atmospheric Neutrinos". In: *Physical Review Letters* 81.8 (1998). ISSN: 1079-7114. DOI: 10.1103/physrevlett.81.1562. URL: http://dx.doi.org/10.1103/PhysRevLett.81.1562.

[111] A. Habig et al. "An Indirect Search for WIMPs with Super-Kamiokande". In: (July 2001).

[112] G Battistoni et al. "The FLUKA code: description and benchmarking". In: *AIP Conference Proceedings* 896.1 (Mar. 2007). ISSN: 0094-243X. DOI: 10.1063/1.2720455.

[113] Vasilis Vlachoudis. "FLAIR: A POWERFUL BUT USER FRIENDLY GRAPHICAL INTERFACE FOR FLUKA". In: Apr. 2009.

[114] A. Fasso et al. In: *Proceedings of the Specialists Meeting on Shielding Aspects of Accelerators, Targets and Irradiation Facilities* (1994).

[115] A. Ferrari et al. In: *Proceedings of Workshop on Nuclear Reaction Data and Nuclear Reactors Physics, Design and Safety* (1998).

[116] Alfredo Ferrari et al. "Nuclear Reactions in Monte Carlo Codes". In: *Radiation protection dosimetry* 99 (Feb. 2002), pp. 29–38. DOI: 10.1093/oxfordjournals.rpd.a006788.

[117] Guanying Zhu et al. "Developing the MeV potential of DUNE: Detailed considerations of muon-induced spallation and other backgrounds". In: *Physical Review C* 99.5 (2019). ISSN: 2469-9993. DOI: 10.1103/physrevc.99.055810. URL: http://dx.doi.org/10.1103/PhysRevC.99.055810.

[118] A. Empl et al. *Study of Cosmogenic Neutron Backgrounds at LNGS*. 2012. arXiv: 1210.2708 [astro-ph.IM].

[119] S. Fukuda et al. "The Super-Kamiokande detector". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 501.2 (2003), pp. 418 –462. ISSN: 0168-9002. DOI: https://doi.org/10.1016/S0168-9002(03)00425-X. URL: http://www.sciencedirect.com/science/article/pii/S016890020300425X.

[120] National Nuclear Data Center. *NuDat 2 database*. http://www.nndc.bnl.gov/nudat2/. Online version 2.6.

[121] *http://www.nndc.bnl.gov*.

[122] *http://www.tunl.duke.edu/nucldata/GroundStatedecays/*.

[123] P.A. Zyla et al. "Review of Particle Physics". In: *PTEP* 2020.8 (2020), p. 083C01. DOI: 10.1093/ptep/ptaa104.

[124] V. Khachatryan et al. "Measurement of the charge ratio of atmospheric muons with the CMS detector". In: *Physics Letters B* 692.2 (2010), pp. 83 –104. ISSN: 0370-2693. DOI: https://doi.org/10.1016/j.physletb.2010.07.033. URL: http://www.sciencedirect.com/science/article/pii/S0370269310008725.

[125] P. Adamson et al. "Measurement of the Atmospheric Muon Charge Ratio at TeV Energies with MINOS". In: *Phys. Rev. D* 76 (Sept. 2007). DOI: 10.1103/PhysRevD.76.052003.

[126] N. Agafonova et al. "Measurement of the atmospheric muon charge ratio with the OPERA detector". In: *Eur. Phys. J. C* 67 (2010), pp. 25–37. DOI: 10.1140/epjc/s10052-010-1284-8. arXiv: 1003.1907 [hep-ex].

[127] E Blaufuss et al. "16N as a calibration source for Super-Kamiokande". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 458.3 (2001), pp. 638 –649. ISSN: 0168-9002. DOI: https://doi.org/10.1016/S0168-9002(00)00900-1. URL: http://www.sciencedirect.com/science/article/pii/S0168900200009001.

[128] A.C. Kahler et al. "ENDF/B-VII.1 Neutron Cross Section Data Testing with Critical Assembly Benchmarks and Reactor Experiments". In: *Nuclear Data Sheets* 112.12 (2011). Special Issue on ENDF/B-VII.1 Library, pp. 2997–3036. ISSN: 0090-3752. DOI: https://doi.org/10.1016/j.nds.2011.11.003. URL: https://www.sciencedirect.com/science/article/pii/S0090375211001141.

[129] Y. F. Wang et al. "Predicting neutron production from cosmic-ray muons". In: 64.1, 013012 (July 2001), p. 013012. DOI: 10.1103/PhysRevD.64.013012. arXiv: hep-ex/0101049 [hep-ex].

[130] Kamiokande Collaboration et al. *Diffuse Supernova Neutrino Background Search at Super-Kamiokande*. 2021. arXiv: 2109.11174 [astro-ph.HE].

[131] Pablo Fernandez Menendez. *Neutrino Physics in Present and Future Kamioka Water-Cherenkov Detectors with Neutron Tagging*. PhD thesis. Faculty of Sciences of University Autonomous of Madrid.

[132] Super-Kamiokande collaboration. *First Gadolinium Loading to Super-Kamiokande*. In process of publication.

[133] Jiawei Tan et al. "Geant4 Modifications for Accurate Fission Simulations". In: *Physics Procedia* 90 (2017). Conference on the Application of Accelerators in Research and Industry, CAARI 2016, 30 October - 4 November 2016, Ft. Worth, TX, USA, pp. 256–265. ISSN: 1875-3892. DOI: https://doi.org/10.1016/j.phpro.2017.09.005. URL: https://www.sciencedirect.com/science/article/pii/S1875389217301645.

[134] Kaito Hagiwara et al. *Gamma Ray Spectrum from Thermal Neutron Capture on Gadolinium-157*. Sept. 2018.

Résumé en français

Les neutrinos ont joué un rôle clé en astrophysique, de la caractérisation des processus de fusion nucléaire dans le Soleil à l'observation de la Supernova SN1987A et de multiples événements extra-galactiques. L'expérience Super-Kamiokande (SK), le plus grand détecteur souterrain de neutrinos au monde, a joué un rôle majeur dans ces études astrophysiques en étudiant les neutrinos de basse énergie ($\mathcal{O}(10)$ MeV). Il a notamment contribué à caractériser le spectre des neutrinos solaires ^{8}B et présente actuellement la meilleure sensibilité pour l'observation des neutrinos reliques de supernova (SRN). Ce signal provient du flux cumulé de neutrinos provenant de tous les événements de supernovae qui se sont produits dans tout l'Univers. La détection de ce fond cosmologique est actuellement l'une des découvertes les plus attendues en physique des neutrinos.

A partir de fin 2020, avec le début de SK-Gd, la phase upgradée de Super-Kamiokande avec l'injection de sel de gadolinium dans l'eau ultrapure du détecteur, une augmentation remarquable de l'efficacité de détection des neutrons est attendue et ainsi une étendue de l'expérience atteint le flux de neutrinos de supernovae actuellement non observé. Un effort majeur de réparation du détecteur ainsi que la mise en œuvre de nouvelles méthodes d'étalonnage ont été déterminants pour la transition de l'expérience et les premiers mois de prise de données ont été couronnés de succès. Le dénominateur commun des activités de travail présentées dans cette thèse est la préparation de cette nouvelle phase passionnante du Super-Kamiokande.

Les recherches à basse énergie en SK, comme l'analyse SRN, sont confrontées à des défis importants en raison des bruits de fond importants de la spallation induite par des rayons cosmiques : les muons atteignant SK peuvent briser les noyaux d'oxygène, produisant des isotopes instables dont les désintégrations radioactives sont une contamination majeure en dessous de 20 MeV. La réduction de cette contamination majeure est fondamentale afin d'atteindre la sensibilité pour la première observation de neutrinos reliques de supernova. Bien que les techniques de réduction actuelles, au prix d'une perte de signal importante, suppriment une grande partie des désintégrations de spallation, ce bruit de fond reste dominant à des énergies cinétiques de 6 à 20 MeV. C'est la raison pour laquelle dans ce travail une nouvelle méthode a été développée mettant en œuvre des algorithmes de marquage neutronique de pointe, ainsi qu'une caractérisation approfondie des mécanismes induisant la spallation. Une nouvelle et unique simulation basée sur FLUKA, incorporant les modèles nucléaires les plus avancés, a joué un rôle clé dans l'étude des processus de spallation et, à son tour, dans la mise au point de nouvelles méthodes de rejet. Grâce à la simulation, nous disposons désormais non seulement d'une étude approfondie des fonds de spallation mais également d'un outil précieux pour optimiser les stratégies d'analyse des futures recherches à basse énergie dans les détecteurs eau-Chérenkov. C'est en vue de SK-Gd que la nouvelle méthode de réduction de la spallation, soigneusement optimisée grâce à des informations uniques issues de la simulation, prouve son réel potentiel, montrant une efficacité sans précédent dans l'identification de la spallation et ouvrant la porte à de nouvelles découvertes importantes.

Cette thèse est organisée en 7 Chapitres : Le Chapitre 1 présente brièvement les principales propriétés des neutrinos, y compris un aperçu des sources de neutrinos et des processus d'interaction. Il contient également un résumé concis de la dynamique d'explosion

de la supernova et des caractéristiques du flux SRN. Des détails sur la recherche de SRN sont présentés et suivis d'une description des processus de spallation. Le Chapitre 2 est consacré au détecteur Super-Kamiokande : sa structure, son électronique, son système d'eau ainsi que les algorithmes de simulation et de reconstruction Monte Carlo personnalisés sont présentés. Super-Kamiokande a récemment été modernisé en vue de sa nouvelle phase avec du gadolinium. Les grands travaux de rénovation se sont achevés à l'été 2018, suivis d'une période d'étalonnage au cours de laquelle des efforts importants ont été consentis pour introduire de nouvelles techniques de mesures. Ces études ont abouti au chargement réussi de Gd dans l'eau pure et au début de SK-Gd en août 2020. Ces sujets seront décrits dans le Chapitre 3. Le Chapitre 4 explique la construction de la simulation pour modéliser les fonds de spallation, les modèles adoptés, les principaux composants et leurs interfaces. La seconde moitié de ce Chapitre est consacrée à la validation de cette simulation à l'aide ded donnés d'SK, étape essentielle dans le développement d'un cadre de calcul précis. Les premières informations sur les processus de spallation obtenues avec la nouvelle simulation sont présentées au Chapitre 5 : une compréhension plus approfondie des pertes d'énergie des muons, du développement des gerbes et une meilleure compréhension de la production d'isotopes radioactifs est cruciale pour améliorer les techniques d'élimination de la spallation. Par la suite, le Chapitre 5 illustre la nouvelle technique développée pour identifier les événements de spallation dans SK-IV. La simulation reproduit avec précision les données et s'avère être un outil clé non seulement pour interpréter les résultats de l'analyse actuelle, mais aussi pour concevoir des stratégies futures. Enfin, le Chapitre 6 éclaire la capacité de la simulation de spallation à adapter et à optimiser les coupes de rejet pour SK-Gd, en fournissant des prévisions de performances et en guidant de nouvelles approches d'analyse.

Certaines des études présentées dans cette thèse sont présentées dans une publication intitulée " New Methods and Simulations for Cosmogenic Induced Spallation Removal in Super-Kamiokande-IV", qui peut être trouvé dans la Réf. [1].

ECOLE DOCTORALE

Titre : Nouvelles études sur le fond cosmogénique de spallation pour la recherche de neutrinos reliques de Supernova dans l'expérience Super-Kamiokande.

Mots clés : neutrinos, Super-Kamiokande, spallation

Résumé : The Super-Kamiokande experiment had a major part in astrophysical studies by investigating low energy neutrinos. It currently exhibits the best sensitivity to the diffuse neutrino background from distant supernovae. Low energy searches however face significant challenges due to important backgrounds from cosmic muon spallation: the reduction of this major contamination is fundamental in order to reach the sensitivity for a first observation of supernova relic neutrinos. Despite current reduction techniques, at the cost of a signal loss, remove a large fraction of spallation decays, the background rate remains dominant at 6-18 MeV kinetic energies. This is the reason why a new method has been developed implementing state-of-the-art neutron tagging algorithms, as well as a thorough characterization of spallation-inducing mechanisms. A unique FLUKA-based simulation, incorporating the most advanced nuclear models, played a key role in the investigation of spallation processes and in turn the tuning of new rejection methods. Thanks to the simulation we now have, not only an in-depth study of spallation backgrounds, but also a valuable tool for optimizing the analysis strategies of future low energy searches in water-Cherenkov detectors. Starting from the end of 2020, with the beginning of SK-Gd, the upgraded phase of Super-Kamiokande with the injection of gadolinium salt into the detector's otherwise ultrapure water, an increase of neutron detection efficiency is expected and thereby an extent of the experiment reach to the currently unobserved supernova relic neutrinos. A major effort in repairing the detector as well as an implementation of new calibration methods had been instrumental for the transition of the experiment and the first months of data taking have been successful. It is in this context that the new spallation reduction method, carefully optimized thanks to unique insights from simulation, proves its potential showing an unprecedented efficiency in spallation identification and opening the door for new discoveries.

Title : New studies on cosmogenic induced spallation background for Supernova relic neutrino search in the Super-Kamiokande experiment.

Keywords : neutrinos, Super-Kamiokande, spallation

Abstract : L'expérience Super-Kamiokande a joué un rôle majeur dans les études astrophysiques en étudiant les neutrinos de basse énergie. Elle présente actuellement la meilleure sensibilité au bruit de fond diffus des neutrinos des distantes supernovae. Les recherches à basse énergie sont cependant confrontées à des défis importants en raison des bruits de fond significatifs de la spallation cosmique des muons : la réduction de cette contamination majeure est fondamentale afin d'atteindre la sensibilité pour une première observation de neutrinos reliques. Malgré les techniques de détection actuelles qui éliminent une grande partie des désintégrations de spallation avec un coût d'une perte de signal importante, le taux de fond reste dominant aux énergies cinétiques de 6 à 18 MeV. C'est la raison pour laquelle une nouvelle méthode a été développée mettant en œuvre des algorithmes de neutron tagging de pointe, ainsi qu'une caractérisation approfondie des mécanismes induisant la spallation. Une simulation originale basée sur FLUKA, incorporant les modèles nucléaires les plus avancés, joue un rôle clé dans l'étude des processus de spallation et dans la mise au point de nouvelles méthodes de rejet. Grâce à la simulation, nous disposons désormais non seulement d'une étude approfondie des fonds de spallation, mais aussi d'un outil précieux pour optimiser les stratégies d'analyse des futures recherches à basse énergie dans les détecteurs water-Tcherenkov. À partir de fin 2020, avec le début de SK-Gd, la phase améliorée de Super-Kamiokande avec l'injection de gadolinium dans l'eau ultrapure du détecteur, une augmentation remarquable de l'efficacité de détection des neutrons est attendue pour la recherche du fond diffus de neutrinos de supernova actuellement non observé. Un effort majeur de réparation du détecteur ainsi que la mise en œuvre de nouvelles méthodes de calibration ont été déterminants pour la nouvelle phase de l'expérience et. Les premiers mois de prise de données ont été couronnés de succès. C'est dans ce contexte que la nouvelle méthode de réduction du bruit de fond de spallation, soigneusement optimisée grâce à des informations issues de la simulation, prouve son réel potentiel montrant une efficacité sans précédent dans l'identification de la spallation et ouvrant la porte à de nouvelles découvertes.

Institut Polytechnique de Paris
91120 Palaiseau, France