

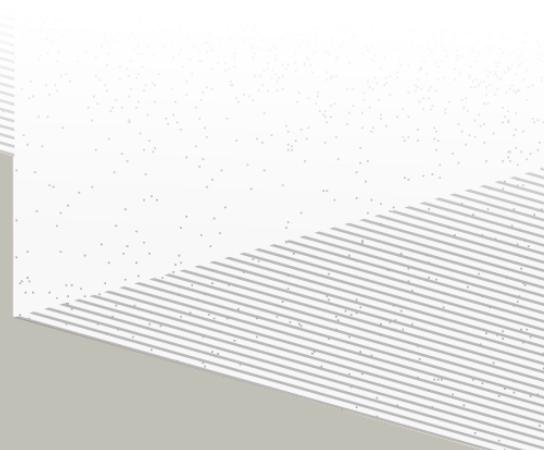
1

2

THÈSE DE DOCTORAT DE

L'UNIVERSITÉ DE NANTES

ÉCOLE DOCTORALE N°596
Matière, Molécules, Matériaux
Spécialité : *Physique des particules*



Par

Léonard Imbert

**Precision measurement of solar neutrino oscillation parameters
with the JUNO small PMTs system and test of the unitarity of the
PMNS matrix**

Thèse présentée et soutenue à Nantes, le Too soon and too early at the same time
Unité de recherche : Laboratoire SUBATECH, UMR 6457

Rapporteurs avant soutenance :

Someone	Very high position, Somewhere
Someone else	Another high position, Somewhere else

Composition du Jury :

Président :	Jacques Chirac	French president, France
Examinateurs :	Yoda	Master, Dagobha
Dir. de thèse :	Frédéric Yermia	Professeur, Université de Nantes
Co-dir. de thèse :	Benoit Viaud	Chargé de recherche, CNRS

Invité(s) :

Victor Lebrin Docteur, CNRS/SUBATECH

³ Contents

⁴	Contents	1
⁵	Remerciements	3
⁶	Introduction	5
⁷	1 Neutrino physics	7
⁸	1.1 Standard model	7
⁹	1.1.1 Limits of the standard model	7
¹⁰	1.2 Historic of the neutrino	7
¹¹	1.3 Oscillation	7
¹²	1.3.1 Phenomologies	7
¹³	1.4 Open questions	7
¹⁴	2 The JUNO experiment	9
¹⁵	2.1 Neutrinos physics in JUNO	10
¹⁶	2.1.1 Reactor neutrino oscillation for NMO and precise measurements	10
¹⁷	2.1.2 Other physics	13
¹⁸	2.2 The JUNO detector	14
¹⁹	2.2.1 Detection principle	15
²⁰	2.2.2 Central Detector (CD)	16
²¹	2.2.3 Veto detector	19
²²	2.3 Calibration strategy	20
²³	2.3.1 Energy scale calibration	20
²⁴	2.3.2 Calibration system	20
²⁵	2.4 Satellite detectors	21
²⁶	2.4.1 TAO	22
²⁷	2.4.2 OSIRIS	22
²⁸	2.5 Software	23
²⁹	2.6 State of the art of the Offline IBD reconstruction in JUNO	24
³⁰	2.6.1 Interaction vertex reconstruction	24
³¹	2.6.2 Energy reconstruction	28
³²	2.6.3 Machine learning for reconstruction	31
³³	2.7 JUNO sensitivity to NMO and precise measurements	34
³⁴	2.7.1 Theoretical spectrum	34

35	2.7.2 Fitting procedure	34
36	2.7.3 Physics results	35
37	2.8 Summary	35
38	3 Machine learning and Artificial Neural Network	37
39	3.1 Boosted Decision Tree (BDT)	37
40	3.2 Artificial Neural Network (NN)	38
41	3.2.1 Fully Connected Deep Neural Network (FCDNN)	39
42	3.2.2 Convolutional Neural Network (CNN)	39
43	3.2.3 Graph Neural Network (GNN)	41
44	3.2.4 Adversarial Neural Network (ANN)	42
45	3.2.5 Training procedure	42
46	3.2.6 Potential pitfalls	45
47	4 Image recognition for IBD reconstruction with the SPMT system	47
48	4.1 Motivations	47
49	4.2 Method and model	48
50	4.2.1 Model	48
51	4.2.2 Data representation	50
52	4.2.3 Dataset	51
53	4.2.4 Data characteristics	52
54	4.3 Results	53
55	4.3.1 J21 results	55
56	4.3.2 Combination of classic and ML estimator	58
57	4.4 Prospect	58
58	4.5 Conclusion	58
59	5 Graph representation of JUNO for IBD reconstruction with the LPMT system	59
60	6 Reliability of machine learning methods	61
61	7 Joint fit between the SPMT and LPMT spectra	63
62	8 Conclusion	65
63	List of Tables	67
64	List of Figures	71
65	List of Abbreviations	73
66	Bibliography	75

Remerciements

⁶⁷

68 Introduction

⁶⁹ **Chapter 1**

⁷⁰ **Neutrino physics**

⁷¹ *The neutrino, or ν for the close friends, a fascinating and invisible particle. Some will say that dark matter also have those property but at least we are pretty confident that neutrinos exists.*

⁷² **1.1 Standard model**

⁷³ **1.1.1 Limits of the standard model**

⁷⁴ **1.2 Historic of the neutrino**

⁷⁵ **First theories**

⁷⁶ **Discovery**

⁷⁷ **Milestones and anomalies**

⁷⁸ **1.3 Oscillation**

⁷⁹ **1.3.1 Phenomologies**

⁸⁰ **1.4 Open questions**

Decrire le m
Regarder th
Kochebina
Limite du r
Interessant,
les neutrino
CP ? Pb des

⁸¹ **Chapter 2**

⁸² **The JUNO experiment**

⁸³ “*Ave Juno, rosae rosam, et spiritus rex*”. It means nothing but I found it in tone.

⁸⁴ The first idea of a medium baseline (~ 52 km) experiment, was explored in 2008 [1] where it was
⁸⁵ demonstrated that the Neutrino Mass Ordering (NMO) could be determined by a medium baseline
⁸⁶ experiment if $\sin^2(2\theta_{13}) > 0.005$ without the requirements of accurate knowledge of the reactor
⁸⁷ antineutrino spectra and the value of Δm_{32}^2 . From this idea is born the Jiangmen Underground
⁸⁸ Neutrino Observatory (JUNO) experiment.

⁸⁹ JUNO is a neutrino detection experiment under construction located in China, in Guangdong prov-
⁹⁰ ing, near the city of Kaiping. Its main objectives are the determination of the mass ordering at the
⁹¹ $3\text{-}4\sigma$ level in 6 years of data taking and the measurement at the sub-percent precision of the oscillation
⁹² parameters Δm_{21}^2 , $\sin^2 \theta_{12}$, Δm_{32}^2 and with less precision $\sin^2 \theta_{13}$ [2].



⁹³ FIGURE 2.1 – **On the left:** Location of the JUNO experiment and its reactor sources in
⁹⁴ southern China. **On the right:** Aerial view of the experimental site

⁹⁵ For this JUNO will measure the electronic anti-neutrinos ($\bar{\nu}_e$) flux coming from the nuclear reactors
⁹⁶ of Taishan, Yangjiang, for a total power of 26.6 GW_{th} , and the Daya Bay power plant to a lesser
⁹⁷ extent. All of those cores are the second-generation pressurized water reactors CPR1000, which is a
⁹⁸ derivative of Framatome M310. Details about the power plants characteristics and their expected flux
⁹⁹ of $\bar{\nu}_e$ can be found in the table 2.1. The distance of 53 km has been specifically chosen to maximize
¹⁰⁰ the disappearance probability of the $\bar{\nu}_e$. The data taking is scheduled to start early 2025.

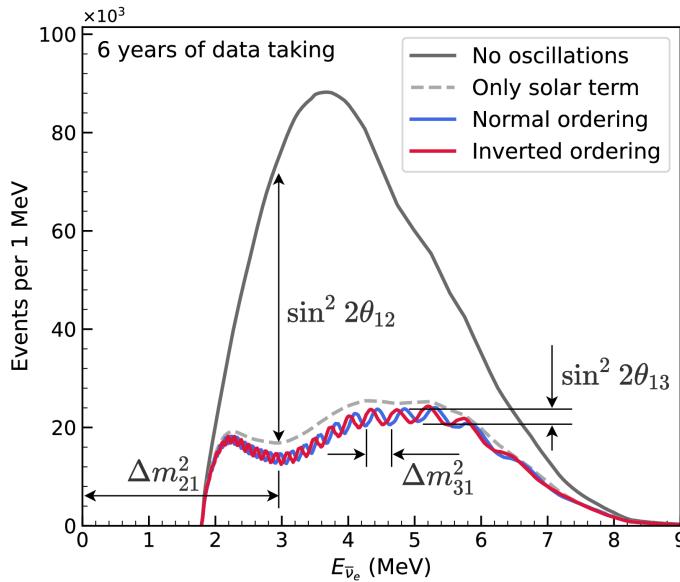


FIGURE 2.2 – Expected number of neutrinos event per MeV in JUNO after 6 years of data taking. The black curve shows the flux if there was no oscillation. The light gray curve shows the oscillation if only the solar terms are taken in account (θ_{12} , Δm_{21}^2). The blue and red curve shows the spectrum in the case of, respectively, NO and IO. The dependency of the oscillation to the different parameters are schematized by the double sided arrows. We can see the NMO sensitivity by looking at the fine phase shift between the red and the blue curve.

99 2.1 Neutrinos physics in JUNO

100 Even if the JUNO design detailed in section 2.2 was optimized for the measurement of the NMO, its
 101 large detection volume, excellent energy resolution and background level and understanding make it
 102 also an excellent detector to measure the flux coming from other neutrino sources. Thus the scientific
 103 program of JUNO extends way over reactor antineutrinos. The following section is an overview of
 104 the different physics topic JUNO will contribute in the coming years.

105 2.1.1 Reactor neutrino oscillation for NMO and precise measurements

Previous works [1, 3] shows that oscillation parameters and the NMO can be observed by looking at the $\bar{\nu}_e$ disappearance energy spectrum coming from medium baseline nuclear reactor. This disappearance probability can be expressed as [2] :

$$P(\bar{\nu}_e \rightarrow \bar{\nu}_e) = 1 - \sin^2 2\theta_{12} c_{13}^4 \sin^2 \frac{\Delta m_{21}^2 L}{4E} - \sin^2 2\theta_{13} \left[c_{12}^2 \sin^2 \frac{\Delta m_{31}^2 L}{4E} + s_{12}^2 \sin^2 \frac{\Delta m_{32}^2 L}{4E} \right]$$

106 Where $s_{ij} = \sin \theta_{ij}$, $c_{ij} = \cos \theta_{ij}$, E is the $\bar{\nu}_e$ energy and L is the baseline. We can see the sensitivity
 107 to the NMO in the dependency to Δm_{32}^2 and Δm_{31}^2 causing a phase shift of the spectrum as we can
 108 see in the figure 2.2. By carefully adjusting a theoretical spectrum to the data, one can extract the
 109 NMO and the oscillation parameters. The statistic procedure used to adjust the theoretical spectrum
 110 is reviewed in more details in the section 2.7. To reach the desired sensitivity, JUNO must meet
 111 multiple requirements but most notably:

- 112 1. An energy resolution of $3\%/\sqrt{E(\text{MeV})}$ to be able to distinguish the fine structure of the fast
 113 oscillation.
- 114 2. An energy precision of 1% in order to not err on the location of the oscillation pattern.
- 115 3. A baseline of 53 ± 0.5 km to maximise the $\bar{\nu}_e$ oscillation probability.
- 116 4. At least $\approx 100,000$ events to limit the spectrum distortion due to statistical uncertainties.

117 **$\bar{\nu}_e$ flux coming from nuclear power plants**

118 To get such high measurements precision, it is necessary to have a very good understanding of the
 119 sources characteristics. For its NMO and precise measurement studies, JUNO will observe the energy
 120 spectrum of neutrinos coming from the nuclear power plants Taishan and Yangjiang's cores, located
 121 at 53 km of the detector to maximise the disappearance probability of the $\bar{\nu}_e$.

Reactor	Power (GW _{th})	Baseline (km)	IBD Rate (day ⁻¹)	Relative Flux (%)
Taishan	9.2	52.71	15.1	32.1
Core 1	4.6	52.77	7.5	16.0
Core 2	4.6	52.64	7.6	16.1
Yangjiang	17.4	52.46	29.0	61.5
Core 1	2.9	52.74	4.8	10.1
Core 2	2.9	52.82	4.7	10.1
Core 3	2.9	52.41	4.8	10.3
Core 4	2.9	52.49	4.8	10.2
Core 5	2.9	52.11	4.9	10.4
Core 6	2.9	52.19	4.9	10.4
Daya Bay	17.4	215	3.0	6.4

TABLE 2.1 – Characteristics of the nuclear power plants observed by JUNO. The IBD rate are estimated from the baselines, the reactors full thermal power, selection efficiency and the current knowledge of the oscillation parameters

122 The $\bar{\nu}_e$ coming from reactors are emitted from β -decay of unstable fission fragments. The Taishan
 123 and Yangjiang reactors are Pressurised Water Reactor (PWR), the same type as Daya Bay. In those
 124 type of reactor more the 99.7 % and $\bar{\nu}_e$ are produced by the fissions of four fuel isotopes ^{235}U , ^{238}U ,
 125 ^{239}Pu and ^{241}Pu . The neutrino flux per fission of each isotope is determined by the inversion of the
 126 measured β spectra of fission product [4–8] or by calculation using the nuclear databases [9, 10].

127 The neutrino flux coming from a reactor at a time t can be predicted using

$$\phi(E_\nu, t)_r = \frac{W_{th}(t)}{\sum_i f_i(t) e_i} \sum_i f_i(t) S_i(E_\nu) \quad (2.1)$$

128 where $W_{th}(t)$ is the thermal power of the reactor, $f_i(t)$ is the fraction fission of the i th isotope, e_i its
 129 thermal energy released in each fission and $S_i(e_\nu)$ the neutrino flux per fission for this isotope. Using
 130 this method, the flux uncertainty is expected to be of an order of 2-3 % [11].

131 In addition to those prediction, a satellite experiment named TAO[12] will be setup near the reactor
 132 core Taishan-1 to measure with an energy resolution of 2% at 1 MeV the neutrino flux coming from
 133 the core, more details can be found in section 2.4.1. It will help identifying unknown fine structure
 134 and give more insight on the $\bar{\nu}_e$ flux coming from this reactor.

135 One the open issue about reactor anti-neutrinos flux is the so-called neutrino anomaly [13], an
 136 unexpected surplus of neutrino emission in the spectra around 5 MeV. Multiples scientists are trying
 137 to explain this surplus by advanced recalculation of the nuclei model during beta decay [14, 15] but
 138 no consensus on this issue has been reached yet.

139 **Background in the neutrinos reactor spectrum**

140 Considering the close reactor neutrinos flux as the main signal, the signals that are considered as
 141 background are:

- The geoneutrinos producing background in the $0.511 \sim 2.7$ MeV region.
- The neutrinos coming from the other nuclear reactors around Earth.

144 In addition to all those physics signal, non-neutrinos signal that would mimic an IBD will also be
 145 present. It is composed of:

- The signal coming from radioactive decay (α , γ , β) from natural radioactive isotopes in the material of the detector.
- Cosmogenic event such as fast neutrons and activated isotopes induced by muons passing through the detector, most notably the spallation on ^{12}C .

150 All those events represent a non-negligable part of the spectrum as shown in figure 2.3.

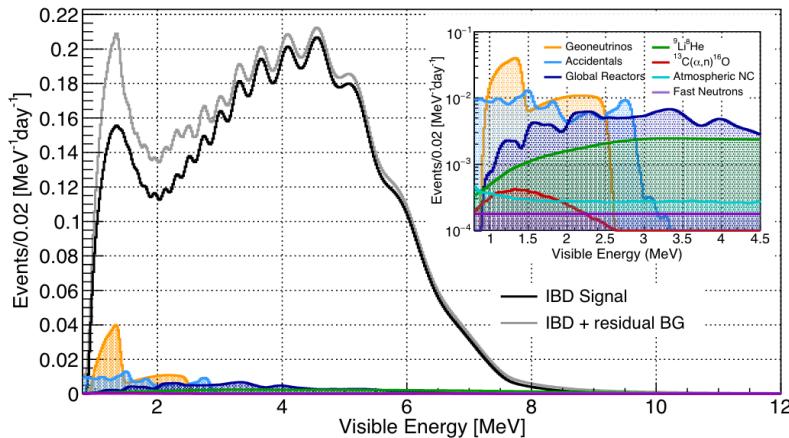


FIGURE 2.3 – Expected visible energy spectrum measured with the LPMT system with (grey) and without (black) backgrounds. The background amount for about 7% of the IBD candidate and are mostly localized below 3 MeV [11]

151 **Identification of the mass ordering**

152 To identify the mass ordering, we adjust the theoretical neutrino energy spectrum under the two
 153 hypothesis of NO and IO. Those give us two χ^2 , respectively χ^2_{NO} and χ^2_{IO} . By computing the
 154 difference $\Delta\chi^2 = \chi^2_{\text{NO}} - \chi^2_{\text{IO}}$ we can determine the most probable mass ordering and the confidence
 155 interval: NO if $\Delta\chi^2 > 0$ and IO if $\Delta\chi^2 < 0$. Current studies shows that the expected sensitivity
 156 the mass ordering would be of 3.4σ after 6 years of data taking in nominal setup[2]. More detailed
 157 explanations about the procedure can be found in the section 2.7.

158 **Precise measurement of the oscillations parameters**

159 The oscillations parameters θ_{12} , θ_{13} , Δm_{21}^2 , Δm_{31}^2 are free parameters in the fit of the oscillation
 160 spectrum. The precision on those parameters have been estimated and are shown in table 2.2. Wee
 161 see that for θ_{12} , Δm_{21}^2 , Δm_{31}^2 , precision at 6 years is better than the reference precision by an order of
 162 magnitude [11]

	Central Value	PDG 2020	100 days	6 years	20 years
$\Delta m_{31}^2 (\times 10^{-3} \text{ eV}^2)$	2.5283	± 0.034 (1.3%)	± 0.021 (0.8%)	± 0.0047 (0.2%)	± 0.0029 (0.1%)
$\Delta m_{21}^2 (\times 10^{-3} \text{ eV}^2)$	7.53	± 0.18 (2.4%)	± 0.074 (1.0%)	± 0.024 (0.3%)	± 0.017 (0.2%)
$\sin^2 \theta_{12}$	0.307	± 0.013 (4.2%)	± 0.0058 (1.9%)	± 0.0016 (0.5%)	± 0.0010 (0.3%)
$\sin^2 \theta_{13}$	0.0218	± 0.0007 (3.2%)	± 0.010 (47.9%)	± 0.0026 (12.1%)	± 0.0016 (7.3%)

TABLE 2.2 – A summary of precision levels for the oscillation parameters. The reference value (PDG 2020 [16]) is compared with 100 days, 6 years and 20 years of JUNO data taking.

2.1.2 Other physics

While the design of JUNO is tailored to measure $\bar{\nu}_e$ coming from nuclear reactor, JUNO will be able to detect neutrinos coming from other sources thus allowing for a wide range of physics studies as detailed in the table 2.3 and in the following sub-sections.

Research	Expected signal	Energy region	Major backgrounds
Reactor antineutrino	60 IBDs/day	0–12 MeV	Radioactivity, cosmic muon
Supernova burst	5000 IBDs at 10 kpc	0–80 MeV	Negligible
DSNB (w/o PSD)	2300 elastic scattering		
Solar neutrino	2–4 IBDs/year	10–40 MeV	Atmospheric ν
Atmospheric neutrino	hundreds per year for ${}^8\text{B}$	0–16 MeV	Radioactivity
Geoneutrino	hundreds per year	0.1–100 GeV	Negligible
	≈ 400 per year	0–3 MeV	Reactor ν

TABLE 2.3 – Detectable neutrino signal in JUNO and the expected signal rates and major background sources

Geoneutrinos

Geoneutrinos designate the antineutrinos coming from the decay of long-lived radioactive elements inside the Earth. The 1.8 MeV threshold necessary for the IBD makes it possible to measure geoneutrinos from ${}^{238}\text{U}$ and ${}^{232}\text{Th}$ decay chains. The studies of geoneutrinos can help refine the Earth crust models but is also necessary to characterise their signal, as they are a background to the mass ordering and oscillations parameters studies.

Atmospheric neutrinos

Atmospheric neutrinos are neutrinos originating from the decay of π and K particles that are produced in extensive air showers initiated by the interactions of cosmic rays with the Earth atmosphere. Earth is mostly transparent to neutrinos below the PeV energy, thus JUNO will be able to see neutrinos coming from all directions. Their baseline range is large (15km \sim 13000km), they can have energy between 0.1 GeV and 10 TeV and will contain all neutrino and antineutrinos flavour. Their studies is complementary to the reactor antineutrinos and can help refine the constraints on the NMO [2].

Supernovae burst neutrinos

Neutrinos are crucial component during all stages of stellar collapse and explosion. Detection of neutrinos coming for core collapse supernovae will provide us important informations on the mech-

184 anisms at play in those events. Thanks to its 20 kt sensible volume, JUNO has excellent capabilities
 185 to detect all flavour of the $\mathcal{O}(10 \text{ MeV})$ postshock neutrinos, and using neutrinos of the $\mathcal{O}(1 \text{ MeV})$
 186 will give informations about the pre-supernovae neutrinos. All those informations will allow to
 187 disentangle between the multiple hydro-dynamic models that are currently used to describe the
 188 different stage of core-collapse supernovae.

189 Diffuse supernovae neutrinos background

190 Core-collapse supernovae in our galaxy are rare events, but they frequently occur throughout the
 191 visible Universe sending burst of neutrinos in direction of the Earth. All those events contributes to
 192 a low background flux of low-energy neutrinos called the Diffuse Supernovae Neutrino Background
 193 (DSNB). Its flux and spectrum contains informations about the red-shift dependent supernovae rate,
 194 the average supernovae neutrino energy and the fraction of black-hole formation in core-collapse su-
 195 pernovae. Depending of the DSNB model, we can expect 2-4 IBD events per year in the energy range
 196 above the reactor $\bar{\nu}_e$ signal, which is competitive with the current Super-Kamiokande+Gadolinium
 197 phase [17].

198 Beyond standard model neutrinos interactions

199 JUNO will also be able to probe for beyond standard model neutrinos interactions. After the main
 200 physics topics have been accomplished, JUNO could be upgraded to probe for neutrinoless beta
 201 decay ($0\nu\beta\beta$). The detection of such event would give critical informations about the nature of
 202 neutrinos, is it a majorana or a dirac particle. JUNO will also be able to probe for neutrinos that
 203 would come for the decay or annihilation of Dark Matter inside the sun and neutrinos from putative
 204 primordial black hole. Through the unitary test of the mixing matrix, JUNO will be able to search
 205 for light sterile neutrinos. Thanks to JUNO sensitivity, multiple other exotic can be performed on
 206 neutrino related beyond standard model interactions.

207 2.2 The JUNO detector

208 The JUNO detector is a scintillator detector buried 693.35 meters under the ground (1800 meters
 209 water equivalent). It consist of Central Detector (CD), a water pool and a Top Tracker (TT) as showed
 210 in figure 2.4a. The CD is an acrylic vessel containing the 20 ktons of Liquid Scintillator (LS). It is
 211 supported by a stainless steel structure and is immersed in that water pool that is used as shielding
 212 from external radiation and as a cherenkov detector for the background. The top of the experiment
 213 is partially covered by the Top Tracker (TT), a plastic scintillator detector which is use to detect the
 214 atmospheric muons background and is acting as a veto detector.

215 The top of the experiment also host the LS purification system, a water purification system, a ven-
 216 tilation system to get rid of the potential radon in the air. The CD is observed by two system of
 217 Photo-Multipliers Tubes (PMT). They are attached to the steel structure and their electronic readout
 218 is submersed near them. A third system of PMT is also installed on the structure but are facing
 219 outward of the CD, instrumenting the water to be cherenkov detector. The CD and the cherenkov
 220 detector are optically separated by Tyvek sheet. A chimney for LS filling and purification and for
 221 calibration operations connects the CD to the experimental hall from the top.

222 The CD has been dimensioned to meet the requirements presented in section 2.1.1:

- 223 — Its 20 ktons monolithic LS provide a volume sizeable enough, in combination with the ex-
 224 pected $\bar{\nu}_e$ flux, to reach the desired statistic in 6 years. Its monolithic nature also allow for a
 225 full containment of most of the events, preventing the energy loss in non-instrumented parts
 226 that would arise from a segmented detector.

- 227 — Its large overburden shield it from most of the atmospheric background that would pollute
228 the signal.
229 — The localization of the experiment, chosen to maximize the disappearance with a 53km base-
230 line and in a region that allow two nuclear power plant to be used as sources.

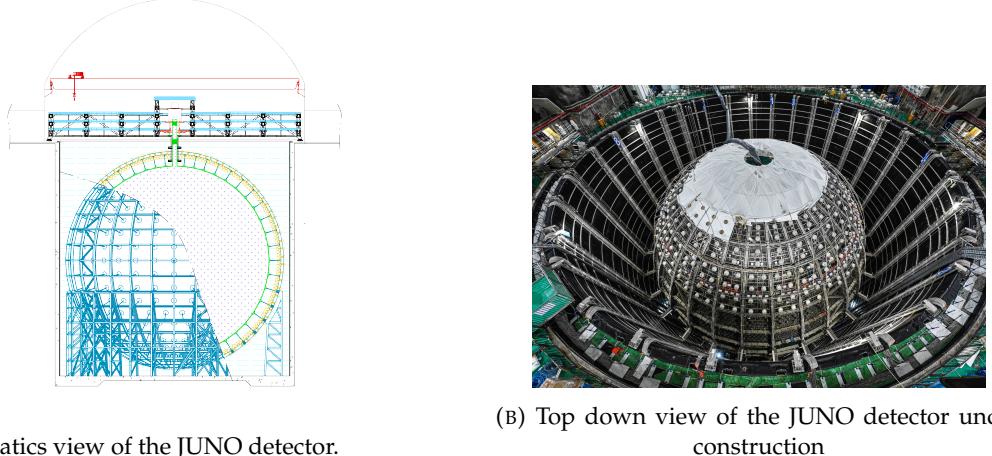


FIGURE 2.4

231 This section cover in details the different components of the detector and the detection systems.

232 2.2.1 Detection principle

The CD will detect the neutrino and measure their energy mainly via an Inverse Beta Decay (IBD) interaction with proton mainly from the ^{12}C and H nucleus in the LS:

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

233 Kinematics calculation shows that this interaction has an energy threshold for the $\bar{\nu}_e$ of $(m_n + m_e -$
234 $m_p) \approx 1.806$ MeV [18] where m_λ is the mass of the λ particle. This threshold make the experiment
235 blind to very low energy neutrinos. The residual energy $E_\nu - 1.806$ MeV is be distributed as kinetic
236 energy between the positron and the neutron. The energy of the emitted positron E_e is given by [18]

$$E_e = \frac{(E_\nu - \delta)(1 + \epsilon_\nu) + \epsilon_\nu \cos \theta \sqrt{(E_\nu - \delta)^2 + \kappa m_e^2}}{\kappa} \quad (2.2)$$

237 where $\kappa = (1 + \epsilon_\nu)^2 - \epsilon_\nu^2 \cos^2 \theta \approx 1$, $\epsilon_\nu = \frac{E_\nu}{m_p} \ll 1$ and $\delta = \frac{m_n^2 - m_p^2 - m_e^2}{2m_p} \ll 1$. We can see from this
238 equation that the positron energy is strongly correlated to the neutrino energy.

239 The positron and the neutron will then propagate in the detection medium, the Liquid Scintillator
240 (LS), loosing their kinetic energy by exciting the molecule of the LS (more details in section 2.2.2).
241 Once stopped, the positron will annihilate with an electron from the medium producing two 511
242 KeV gamma. Those gamma will themselves interact with the LS, exciting it before being absorbed
243 by photoelectrical effect. The neutron will be captured by an hydrogen, emitting a 2.2 MeV gamma
244 in the process. This gamma will also deposit its energy before being absorbed by the LS.

245 The scintillation photons have frequency in the UV and will propagate in the LS, being re-absorbed
246 and re-emitted by compton effect before finally be captured by PMTs instrumenting the acrylic
247 sphere. The analog signal of the PMTs digitized by the electronic is the signal of our experiment.

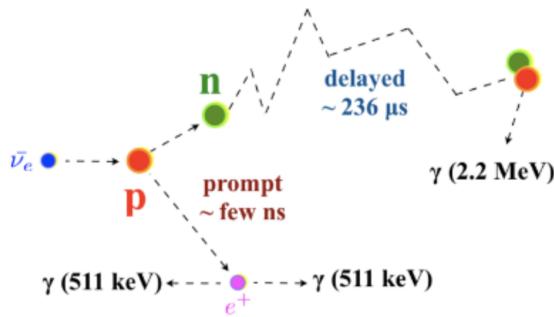


FIGURE 2.5 – Schematics of an IBD interaction in the central detector of JUNO

248 The signal produced by the positron is subsequently called the prompt signal, and the signal coming
 249 from the neutron the delayed signal. This naming convention come from the fact that the positron
 250 will deposit its energy rather quickly (few ns) where the neutron will take a bit more time ($\sim 236 \mu\text{s}$).

251 2.2.2 Central Detector (CD)

252 The central detector, composed of 20 ktons of Liquid Scintillator (LS), is the main part of JUNO. The
 253 LS is contained in a spherical acrylic vessel supported by a stainless steel structure. The CD and
 254 its structural support are submerged in a cylindrical water pool of 43.5m diameter and 44m height.
 255 We're confident that the water pool provide sufficient buffer protection in every direction against the
 256 rock radioactivity.

257 Acrylic vessel

258 The acrylic vessel is a spherical vessel of inner diameter of 35.4 m and a thickness of 120 mm. It is
 259 assembled from 265 acrylic panels, thermo bonded together. The acrylic recipes has been carefully
 260 tuned with extensive R&D to ensure it does not include plasticizer and anti-UV material that would
 261 stop the scintillation photons. Those panels requires to be pure of radioactive materials to not
 262 cause background. Current setup where the acrylic panels are molded in cleanrooms of class 10000,
 263 let us reach a uranium and thorium contamination of <0.5 ppt. The molding and thermoforming
 264 processes is optimized to increase the assemblage transparency in water to >96%. The acrylic vessel
 265 is supported by a stainless steel structure via supporting node (fig 2.6). The structure and the nodes
 266 are designed to be resilient to natural catastrophic events such as earthquake and can support many
 267 times the effective load of the acrylic vessel.

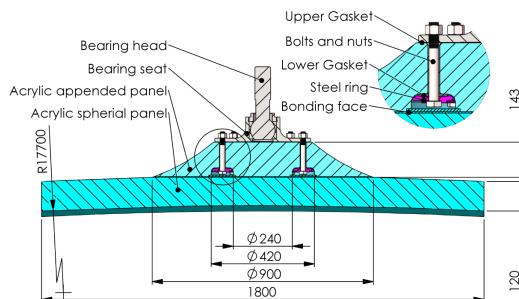


FIGURE 2.6 – Schematics of the supporting node for the acrylic vessel

268 **Liquid scintillator**

269 The Liquid Scintillator (LS) has a similar recipe as the one used in Daya Bay [19] but without gadolinium
 270 doping. It is made of three components, necessary to shift the wavelength of emitted photons to
 271 prevent their reabsorption:

- 272 1. The detection medium, the *linear alkylbenzene* (LAB). Selected because of its excellent transparency,
 273 high flash point, low chemical reactivity and good light yield. Accounting for \sim 98% of the LS, it is the main component with which ionizing particles and gamma interact.
 274 Charged particles will collide with its electronic cloud transferring energy to the molecules,
 275 gamma will interact via compton effect with the electronic cloud before finally be absorbed
 276 via photoelectric effect.
- 277 2. The second component of the LS is the *2,5-diphenyloxazole* (PPO). A fraction of the excitation
 278 energy of the LAB is transferred to the PPO, mainly via non radiative process [20]. The
 279 PPO molecules de-excites in the same way, transferring their energy to the bis-MSB. The PPO
 280 makes for 1.5 % of the LS.
- 281 3. The last component is the *p-bis(o-methylstyryl)-benzene* (bis-MSB). Once excited by the PPO, it
 282 will emit photon with an average wavelength of \sim 430 nm (full spectrum in figure 2.7) that
 283 can be detected by our photo-multipliers systems. It amount for \sim 0.5% of the LS.

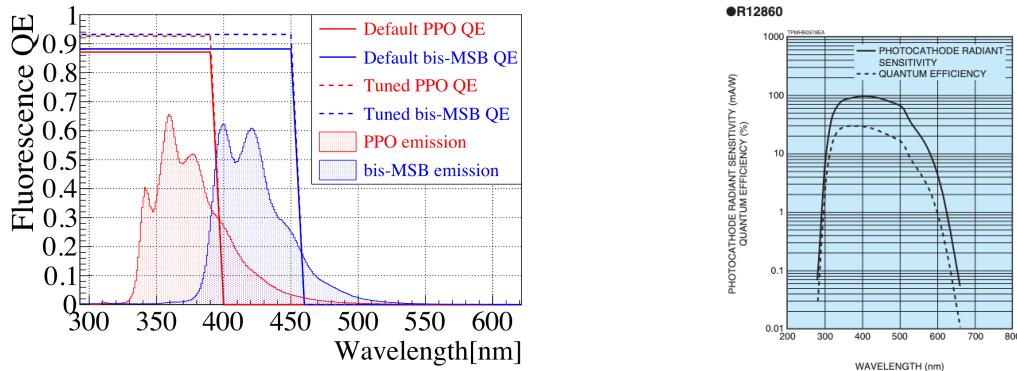


FIGURE 2.7 – On the left: Quantum efficiency (QE) and emission spectrum of the LAB and the bis-MSB [19]. On the right: Sensitivity of the Hamamatsu LPMT depending on the wavelength of the incident photons [21].

285 This formula has been optimized using dedicated studies with a Daya Bay detector [19, 22] to reach
 286 the requirements for the JUNO experiment:

- 287 — A light yield / MeV of the amount of 10^4 photons to maximize the statistic in the energy
 288 measurement.
- 289 — An attenuation length comparable to the size of the detector to prevent losing photons during
 290 their propagation in the LS. The final attenuation length is 25.8m [23] to compare with the CD
 291 diameter of 35.4m.
- 292 — Uranium/Thorium radiopurity to prevent background signal. The reactor neutrino program
 293 require a contamination fraction $F < 10^{-15}$ while the solar neutrino program require $F <$
 294 10^{-17} .

295 The LS will frequently be purified and tested in the Online Scintillator Internal Radioactivity In-
 296 vestigation System (OSIRIS) [24] to ensure that the requirements are kept during the lifetime of the
 297 experiment, more details to be found in section 2.4.2.

298 Large Photo-Multipliers Tubes (LPMTs)

299 The scintillation light produced by the LS is then collected by Photo-Multipliers Tubes (PMT) that
 300 transform the incoming photon into an electric signal. As described in figure 2.8, the incident photons
 301 interact with the photocathode via photoelectric effect producing an electron called a Photo-Electron
 302 (PE). This PE is then focused on the dynodes where the high voltage will allow it to be multiplied.
 303 After multiple amplification the resulting charge - in coulomb [C] - is collected by the anode and
 304 the resulting electric signal can be digitalized by the readout electronics from which the charge and
 305 timing can be extracted.

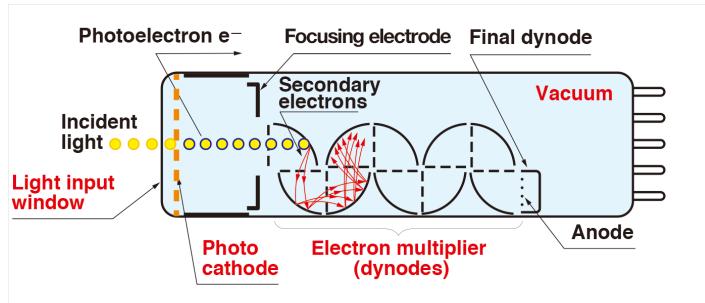


FIGURE 2.8 – Schematic of a PMT

306 The Large Photo-Multipliers Tubes (LPMT), used in the central detector and in the water pool, are
 307 20-inch (50.8 cm) radius PMTs. ~ 5000 dynode-PMTs [21] were produced by the Hamamatsu[®]
 308 company and ~ 15000 Micro-Channel Plate (MCP) [25] by the NNVT[®] company. This system is
 309 the one responsible for the energy measurement with a energy resolution of $3\%/\sqrt{E}$, resolution
 310 necessary for the mass ordering measurement. To reach this precision, the system is composed of
 311 17612 PMTs quasi uniformly distributed over the detector for a coverage of 75.2% reaching ~ 1800
 312 PE/MeV or $\sim 2.3\%$ resolution due to statistic, leaving $\sim 0.7\%$ for the systematic uncertainties. They
 313 are located outside the acrylic sphere in the water pool facing the center of the detector. To maintain
 314 the resolution over the lifetime of the experiment, JUNO require a failure rate $< 1\%$ over 6 years.

315 The LPMTs electronic are divided in two parts. One "near", located underwater, in proximity of the
 316 LPMT to reduce the cable length between the PMT and early electronic. A second one, outside of the
 317 detector that is responsible for higher level analysis before sending the data to the DAQ.

318 The light yield per MeV induce that a LPMT can collect between 1 and 1000 PE per event, a wide
 319 dynamic range, causing non linearity in the PMT response that need to be understood and calibrated,
 320 see section 2.3 for more details.

321 Small Photo-Multipliers Tubes (SPMTs)

322 The Small PMT (SPMTs) system is made of 3-inch (7.62 cm) PMTs. They will be used in the CD
 323 as a secondary detection system. Those 25600 SPMTs will observe the same events as the LPMTs,
 324 thus sharing the physics and detector systematics up until the photon conversion. With a detector
 325 coverage of 2.7%, this system will collect ~ 43 PE/MeV for a final energy resolution of $\sim 17\%$.
 326 This resolution is not enough to measure the NMO, θ_{13} , Δm_{31}^2 but will be sufficient to independently
 327 measure θ_{12} and Δm_{21}^2 .

328 Due to the low PE rate, SPMTs will be running in photo-counting mode in the reactor range and thus
 329 will be insensitive to non-linearity effect. Using this property, the intrinsic charge non linearity of
 330 the LPMTs can be measured by comparing the PE count in the SPMTs and LPMTs [26]. Also, due
 331 to their smaller size and electronics, SPMTs have a better timing resolutions than the LPMTs. At

332 higher energy range, like supernovae events, LPMTs will saturate where SPMTs due to their lower
 333 PE collection will to produce a reliable measure of the energy spectrum.

334 The Data Acquisition System (DAQ) is designed to support the event rate of IBD, background, dark
 335 noise and supplementary storage buffers are present in the LPMT electronics to withstand the event
 336 rate during supernovae burst.

337 2.2.3 Veto detector

338 The CD will be bathed in constant background noise coming from numerous sources : the radioac-
 339 tivity from surrounding rock and its own components or from the flux of cosmic muons. This
 340 background needs to be rejected to ensure the purity of the IBD spectrum. To prevent a big part
 341 of them, JUNO use two veto detector that will tag events as background before CD analysis.

342 Cherenkov in water pool

343 The Water Cherenkov Detector (WCD) is the instrumentation of the water buffer around the CD.
 344 When high speed charged particles will pass through the water, they will produced cherenkov
 345 photons. The light will be collected by 2400 MCP LPMTs installed on the outer surface of the CD
 346 structure. The muons veto strategy is based on a PMT multiplicity condition. WCD PMTs are
 347 grouped in ten zones: 5 in the top, 5 in the bottom. A veto is raised either when more than 19
 348 PMTs are triggered in one zone or when two adjacent zones simultaneously trigger more than 13
 349 PMTs. Using this trigger, we expect to reach a muon detection efficiency of 99.5% while keeping the
 350 noise at reasonable level.

351 Top tracker

352 The JUNO Top Tracker (TT) is a plastic scintillator detector located on the top of the experiment (see
 353 figure 2.9). Made from plastic scintillator from OPERA [27] layered horizontally in 3 layers on the
 354 top of the detector, the TT will be able to detect incoming atmospheric muons. With its coverage,
 355 about 1/3 of the of all atmospheric muons that passing through the CD will also pass through the 3
 356 layer of the detector. While it does not cover the majority of the CD, the TT is particularly effective
 357 to detect muons coming through the filling chimney region which might present difficulties from the
 other subsystems in some classes of events.

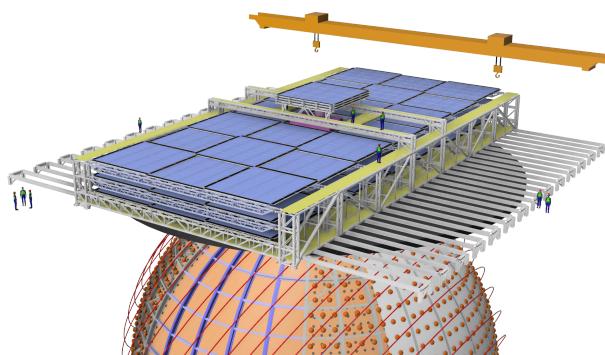


FIGURE 2.9 – The JUNO top tracker

359 2.3 Calibration strategy

360 The calibration is a crucial part of the JUNO experiment. Because we are looking at civil reactor
 361 neutrino it might be impossible to run measurement without signal, it would need to shut down
 362 every reactor from the Taishan and Yangjiang power plants which is realistically impossible. Because
 363 of this continuous rate, low frequency signal event, we need high frequency, recognisable sources in
 364 the energy range of interest : [0-12] MeV for the positron signal and 2.2 MeV for the neutron capture.
 365 It is expected that the CD response will be different depending on the type of particle, due to the
 366 interaction with LS, the position on the event and the optical response of the acrylic sphere (see
 367 section 2.6). We also expect a non-linear energy response of the CD due to the LS properties [19] but
 368 also due to the saturation of the LPMTs system when collecting a large amount of PE [26].

369 2.3.1 Energy scale calibration

370 While electrons and positrons sources would be ideal, for a large LS detector thin-walled electrons
 371 or positrons sources could lead to leakage of radionucleides causing radioactive contamination.
 372 Instead, we consider gamma sources in the range of the prompt energy of IBDs. The sources are
 373 reported in table 2.4.

Sources / Processes	Type	Radiation
^{137}Cs	γ	0.0662 MeV
^{54}Mn	γ	0.835 MeV
^{60}Co	γ	1.173 + 1.333 MeV
^{40}K	γ	1.461 MeV
^{68}Ge	e^+	annihilation 0.511 + 0.511 MeV
$^{241}\text{Am-Be}$	n, γ	neutron + 4.43 MeV ($^{12}\text{C}^*$)
$^{241}\text{Am-}^{13}\text{C}$	n, γ	neutron + 6.13 MeV ($^{16}\text{O}^*$)
$(n, \gamma)p$	γ	2.22 MeV
$(n, \gamma)^{12}\text{C}$	γ	4.94 MeV or 3.68 + 1.26 MeV

TABLE 2.4 – List of sources and their process considered for the energy scale calibration

374 For the ^{68}Ge source, it will decay in ^{68}Ga via electron capture, which will itself β^+ decay into ^{68}Zn .
 375 The positrons will be absorbed by the enclosure so only the annihilation gamma will be released. In
 376 addition, (α, n) sources like $^{241}\text{Am-Be}$ and $^{241}\text{Am-}^{13}\text{C}$ are used to provide both high energy gamma
 377 and neutrons, which will later be captured in the LS producing the 2.2 MeV gamma.

378 From this calibration we call E_{vis} the "visible energy" that is reconstructed by our current algorithms
 379 and we compare it to the true energy deposited by the calibration source. The results shown in figure
 380 2.10 show the expected response of the detector from calibration sources. The non-linearity is clearly
 381 visible from the $E_{\text{vis}}/E_{\text{true}}$ shape. See [28] for more details.

382 2.3.2 Calibration system

383 The non-uniformity due to the event position in the detector (more details in section 2.6) will be
 384 studied using multiples systems that are schematized in figure 2.11. They allow to position sources
 385 at different location in the CD.

- 386 — For a one-dimension vertical calibration, the Automatic Calibration Unit (ACU) will be able
 387 to deploy multiple radioactive sources or a pulse laser diffuser ball along the central axis of
 388 the CD through the top chimney. The source position precision is less than 1cm.

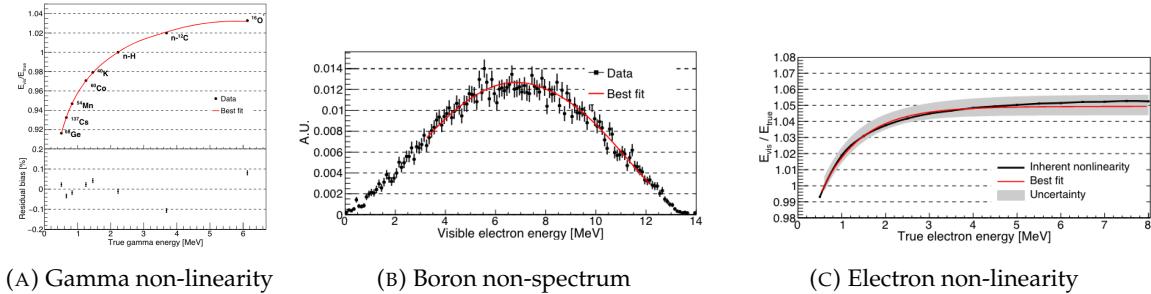


FIGURE 2.10 – Fitted and simulated non linearity of gamma, electron sources and from the ^{12}B spectrum. Black points are simulated data. Red curves are the best fits

- For off-axis calibration, a calibration source attached to a Cable Loop System (CLS) can be moved on a vertical half-plane by adjusting the length of two connection cable. Two set of CSL will be deployed to provide a 79% effective coverage of a vertical plane.
- A Guiding Tube (GT) will surround the CD to calibrate the non-uniformity of the response at the edge of the detector
- A Remotely Operated under-LS Vehicle (ROV) can be deployed to desired location inside LS for a more precise and comprehensive calibration. The ROV will also be equipped with a camera for inspection of the CD.

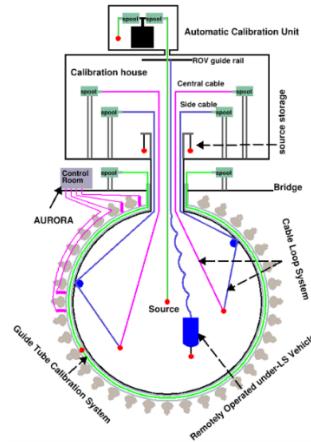


FIGURE 2.11 – Overview of the calibration system

- 397 The preliminary calibration program is depicted in table 2.5.

2.4 Satellite detectors

- 399 As introduced in section 2.1.1 and section 2.2.2, the precise knowledge and understanding of the
400 detector condition is crucial for the measurements of the NMO and oscillation parameters. Thus two
401 satellite detectors will be setup to monitor the experiment condition. TAO to monitor and understand
402 the $\bar{\nu}_e$ flux and spectrum coming from the nuclear reactor and OSIRIS to monitor the LS response.

Program	Purpose	System	Duration [min]
Weekly calibration	Neutron (Am-C)	ACU	63
	Laser	ACU	78
Monthly calibration	Neutron (Am-C)	ACU	120
	Laser	ACU	147
	Neutron (Am-C)	CLS	333
	Neutron (Am-C)	GT	73
Comprehensive calibration	Neutron (Am-C)	ACU, CLS and GT	1942
	Neutron (Am-Be)	ACU	75
	Laser	ACU	391
	^{68}Ge	ACU	75
	^{137}Cs	ACU	75
	^{54}Mn	ACU	75
	^{60}Co	ACU	75
	^{40}K	ACU	158

TABLE 2.5 – Calibration program of the JUNO experiment

403 2.4.1 TAO

404 The Taishan Antineutrino Observatory (TAO) [12, 29] is a ton-level gadolinium doped liquid scin-
 405 tillator detector that will be located near the Taishan-1 reactor. It aim to measure the $\bar{\nu}_e$ spectrum at
 406 very low distance (< 30m) from the reactor to measure a quasi-unoscillated spectrum. TAO also aim
 407 to provide a major contribution to the so-called reactor anomaly [13]. Its requirement are to the level
 408 of 2 % energy resolution at 1 MeV.

409 Detector

410 The TAO detector is close, in concept, to the CD of JUNO. It is composed of an acrylic vessel
 411 containing 2.8 tons of gadolinium-loaded LS instrumented by an array of silicon photomultipliers
 412 (SiPM) reaching a 95% coverage. To efficiently reduce the dark count of those sensors, the detector
 413 is cooled to -50 °C. The $\bar{\nu}_e$ will interact with the LS via IBD, producing scintillation light, that will
 414 be detected by the SiPMs. From this signal the $\bar{\nu}_e$ energy and the full spectrum reconstructed. This
 415 spectrum will then be used by JUNO to calibrate the unoscillated spectrum, most notably the fission
 416 product fraction that impact the rate and shape of the spectrum. A schema of the detector is presented
 417 in figure 2.12a.

418 2.4.2 OSIRIS

419 The Online Scintillator Internal Radioactivity Investigation System (OSIRIS) [24] is an ultralow back-
 420 ground, 20 m³ LS detector that will be located in JUNO cavern. It aim to monitor the radioactive
 421 contamination, purity and overall response of the LS before it is injected in JUNO. OSIRIS will
 422 be located at the end of the purification chain of JUNO, monitoring that the purified LS meet the
 423 JUNO requirements. The setup is optimized to detect the fast coincidences decay of $^{214}\text{Bi} - ^{214}\text{Po}$
 424 and $^{212}\text{Bi} - ^{212}\text{Po}$, indicators of the decay chains of U and Th respectively.

425 Detector

426 OSIRIS is composed of an acrylic vessel that will contains 17t of LS. The LS is instrumented by
 427 a PMT array of 64 20 inch PMTs on the top and the side of the vessel. To reach the necessary

background level required by the LS purity measurements, in addition to being 700m underground in the experiment cavern, the acrylic vessel is immersed in a tank of ultra pure water. The water is itself instrumented by another array of 20 inch PMTs, acting as muon veto. A schema of the detector is presented in figure 2.12b.

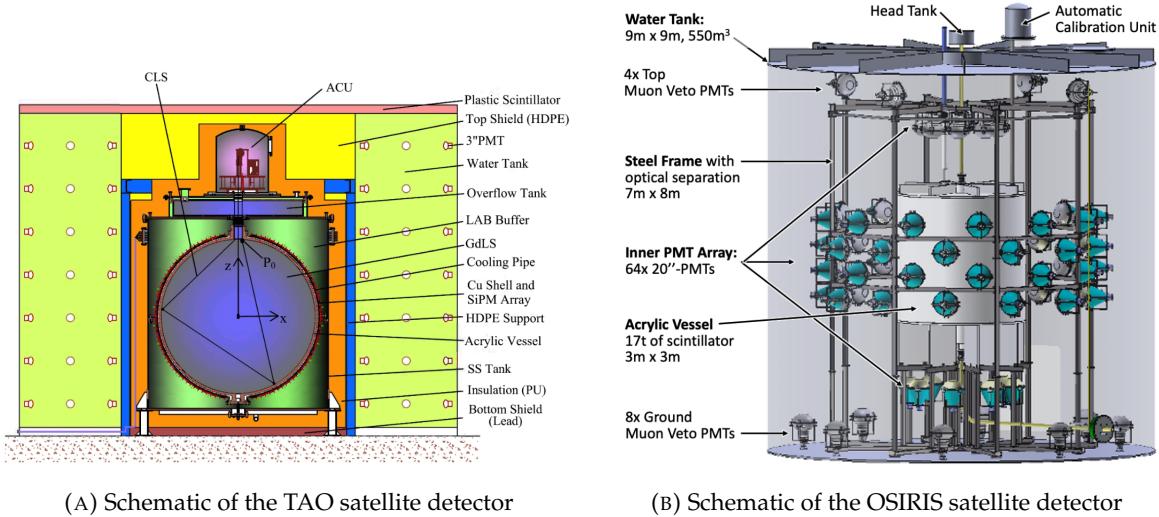


FIGURE 2.12

2.5 Software

The simulation, reconstruction and analysis algorithms are all packaged in the JUNO software, subsequently called the software. It is composed of multiple components integrated in the SNiPER [30] framework:

- Various primary particles simulators for the different kind of events, background and calibration sources.
- A Geant4 [31–33] Monte Carlo (MC) simulation containing the detectors geometries, a custom optical model for the LS and the supporting structures of the detectors. The Geant4 simulation integrate all relevant physics process for JUNO, validated by the collaboration. This step of the simulation is commonly called *Detsim* and compute up to the production of photo-electrons in the PMTs. The optics properties of the different materials and detector components have been measured beforehand to be used to define the material and surfaces in the simulation.
- An electronic simulation, simulating the response waveform of the PMTs, tracking it through the digitization process, accounting for effects such as non-linearity, dark noise, Time Transit Spread (TTS), pre-pulsing, after-pulsing and ringing if the waveform. It's also the step handling the event triggers and mixing. This step is commonly referenced as *Elecsim*.
- A waveform reconstruction where the digitized waveform are filtered to remove high-frequency white noise and then deconvoluted to yield time and charge informations of the photons hits on the PMTs. This step is commonly referenced as *Calib*.
- The charge and time informations are used by reconstruction algorithms to reconstruct the interaction vertex and the deposited energy. This step is commonly reported as *Reco*. See section 2.6 for more details on the reconstruction.
- Once the singular events are reconstructed, they go through event pairing and classification to select IBD events. This step is named Event Classification.

— The purified signal is then analysed by the analysis framework which depend of the physics topic of interest.

The steps Reco and Event Classification are divided into two category of algorithm. Fast but less accurate algorithms that are running during the data taking designated as the *Online* algorithms. Those algorithm are used to take the decision to save the event on tape or to throw it away. More accurate algorithms that run on batch of events designated *Offline* algorithms. They are used for the physics analysis. The Offline Reco will be one of the main topic of interest for this thesis.

2.6 State of the art of the Offline IBD reconstruction in JUNO

The main reconstruction method currently run in JUNO is a data-driven method based on a likelihood maximization [34, 35] using only the LPMTs. The first step is to reconstruct the interaction vertex from which the energy reconstruction is dependent. It is also necessary for event pairing and classification.

2.6.1 Interaction vertex reconstruction

To start the likelihood maximization, a rough estimation of the vertex and of the event timing is needed. We start by estimating the vertex position using a charge based algorithm.

Charge based algorithm

The charge-based algorithm is basically base on the charge-weighted average of the PMT position.

$$\vec{r}_{cb} = a \cdot \frac{\sum_i q_i \cdot \vec{r}_i}{\sum_i q_i} \quad (2.3)$$

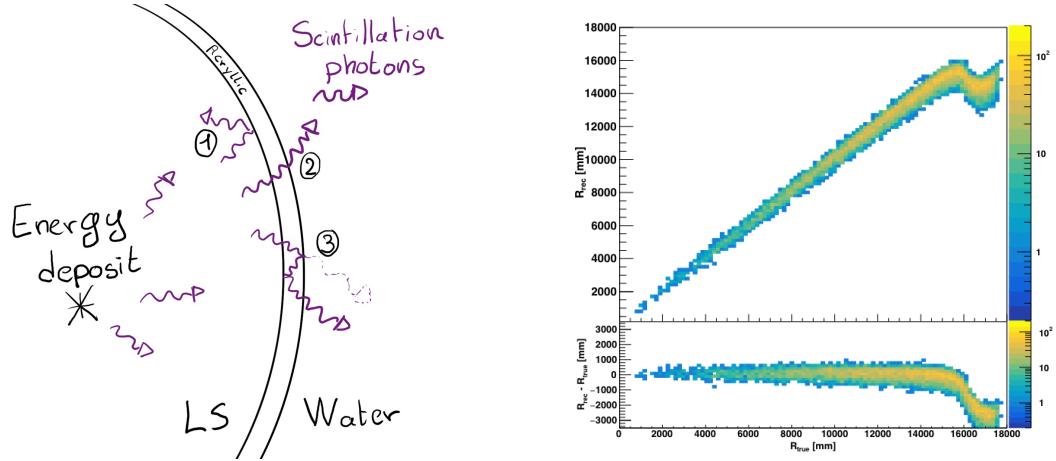
Where q_i is the reconstructed charge of the pulse of the i th PMT and \vec{r}_i is its position. \vec{r}_0 is the reconstructed interaction position. a is a scale factor introduced because a weighted average over a 3D sphere is inherently biased. Using calibration we can estimate $a \approx 1.3$ [36]. The results in figure 2.13b shows that the reconstruction is biased from around 15m and further. This is due to the phenomena called “total reflection area” or TR Area.

As depicted in the figure 2.13a the optical photons, given that they have a sufficiently large incidence angle, can be deviated of their trajectories when passing through the interfaces LS-acrylic and water-acrylic due to the optical index difference. This cause photons to be lost or to be detected by PMT further than anticipated if we consider their rectilinear trajectories. This cause the charge barycenter to be located closer to the center than the event really is.

It is to be noted that charge based algorithm, in addition to be biased near the edge of the detector, does not provide any information about the timing of the event. Therefore, a time based algorithm needs to be introduced to provide initial values.

Time based algorithm

The time based algorithm use the distribution of the time of flight corrections Δt (Eq 2.4) of an event to reconstruct its vertex and t_0 . It follow the following iterations:



(A) Illustration of the different optical photons reflection scenarios. 1 is the reflection of the photon at the interface LS-acrylic or acrylic-water. 2 is the transmission of the photons through the interfaces. 3 is the conduction of the photon in the acrylic.

(B) Heatmap of R_{rec} and $R_{rec} - R_{true}$ as a function of R_{true} for 4MeV prompt signals uniformly distributed in the detector calculated by the charge based algorithm

FIGURE 2.13

489 1. Use the charge based algorithm to get an initial vertex to start the iteration.

490 2. Calculate the time of flight correction for the i th PMT using

$$\Delta t_i(j) = t_i - \text{tof}_i(j) \quad (2.4)$$

491 where j is the iteration step, t_i is the timing of the i th PMT, and tof_i is the time-of-flight of the
492 photon considering an rectilinear trajectory and an effective velocity in the LS and water (see
493 [36] for detailed description of this effective velocity). Plot the Δt distribution and label the
494 peak position as Δt^{peak} (see fig 2.14a).

495 3. Calculate a correction vector $\vec{\delta}[\vec{r}(j)]$ as

$$\vec{\delta}[\vec{r}(j)] = \frac{\sum_i \left(\frac{\Delta t(j) - \Delta t^{\text{peak}}(j)}{\text{tof}_i(j)} \right) \cdot (\vec{r}_0(j) - \vec{r}_i)}{N^{\text{peak}}(j)} \quad (2.5)$$

496 where \vec{r}_0 is the vertex position at the beginning of this iteration, \vec{r}_i is the position of the i th
497 PMT. To minimize the effect of scattering, dark noise and reflection, only the pulse happening
498 in a time window (-10 ns, +5 ns) around Δt^{peak} are considered. N^{peak} is the number of PE
499 collected in this time-window.

500 4. if $\vec{\delta}[\vec{r}(j)] < 1\text{mm}$ or $j \geq 100$, stop the iteration. Otherwise $\vec{r}_0(j+1) = \vec{r}_0(j) + \vec{\delta}[\vec{r}(j)]$ and go to
501 step 2.

502 However because the earliest arrival time is used, t_i is related to the number photoelectrons N_i^{pe}
503 detected by the PMT [37–39]. To reduce bias in the vertex reconstruction, the following equation is
504 used to correct t_i into t'_i :

$$t'_i = t_i - p_0 / \sqrt{N_i^{\text{pe}}} - p_1 - p_2 / N_i^{\text{pe}} \quad (2.6)$$

505 The parameters (p_0, p_1, p_2) were optimized to (9.42, 0.74, -4.60) for Hamamatsu PMTs and (41.31,
506 -12.04, -20.02) for NNVT PMTs [36]. The results presented in figure 2.14b shows that the time based

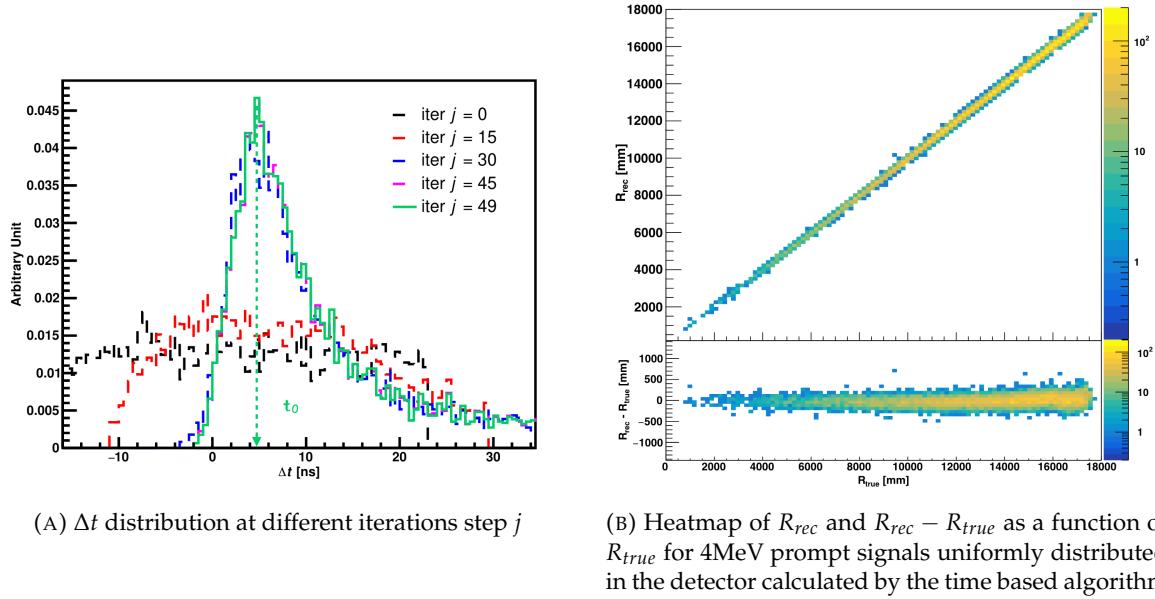
(A) Δt distribution at different iterations step j (B) Heatmap of R_{rec} and $R_{rec} - R_{true}$ as a function of R_{true} for 4MeV prompt signals uniformly distributed in the detector calculated by the time based algorithm

FIGURE 2.14

507 algorithm provide a more accurate vertex and is unbiased even in the TR area. This results (\vec{r}_0, t_0) is
 508 used as initial value for the likelihood algorithm.

509 Time likelihood algorithm

510 The time likelihood algorithm use the residual time expressed as follow

$$t_{\text{res}}^i(\vec{r}_0, t_0) = t_i - \text{tof}_i - t_0 \quad (2.7)$$

511 In a first order approximation, the scintillator time response Probability Density Function (PDF) can
 512 be described as the emission time profile of the scintillation photons, the Time Transit Spread (TTS)
 513 and the dark noise of the PMTs. The emission time profile $f(t_{\text{res}})$ is described like

$$f(t_{\text{res}}) = \sum_k \frac{\rho_k}{\tau_k} e^{-\frac{t_{\text{res}}}{\tau_k}}, \sum_k \rho_k = 1 \quad (2.8)$$

514 as the sum of the k component that emit light in the LS each one characterised by it's decay time τ_k
 515 and intensity fraction ρ_k . The TTS component is expressed as a gaussian convolution

$$g(t_{\text{res}}) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(t_{\text{res}}-\nu)^2}{2\sigma^2}} \cdot f(t_{\text{res}}) \quad (2.9)$$

516 where σ is the TTS of PMTs and ν is the average transit time. The dark noise is not correlated with any
 517 physical events and considered as constant rate over the time window considered T . By normalizing
 518 the dark noise probability $\epsilon(t_{\text{res}})$ as $\int_T \epsilon(t_{\text{res}}) dt_{\text{res}} = \epsilon_{\text{dn}}$, it can be integrated in the PDF as

$$p(t_{\text{res}}) = (1 - \epsilon_{\text{dn}}) \cdot g(t_{\text{res}}) + \epsilon(t_{\text{res}}) \quad (2.10)$$

519 The distribution of the residual time t_{res} of an event can then be compared to $p(t_{\text{res}})$ and the best

520 fitting vertex \vec{r}_0 and t_0 can be chosen by minimizing

$$\mathcal{L}(\vec{r}_0, t_0) = -\ln \left(\prod_i p(t_{\text{res}}^i) \right) \quad (2.11)$$

521 The parameter of Eq. 2.10 can be measured experimentally. The results shown in figure 2.15 used
 522 PDF from monte carlo simulation. The results shows that $R_{\text{rec}} - R_{\text{true}}$ is biased depending on the
 523 energy. While this could be corrected using calibration, another algorithm based on charge likelihood
 524 was developed to correct this problem.

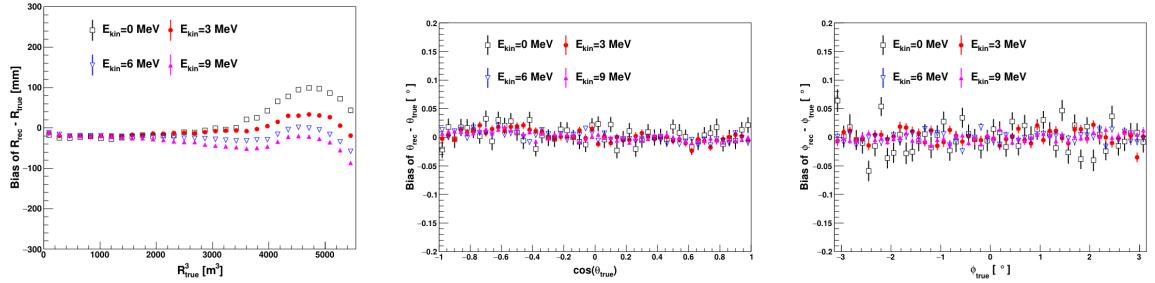


FIGURE 2.15 – Bias of the reconstructed radius R (left), θ (middle) and ϕ (right) for multiple energies by the time likelihood algorithm

525 Charge likelihood algorithm

526 Similarly to the time likelihood algorithms that use a timing PDF, the charge likelihood algorithm
 527 use a PE PDF for each PMT depending on the energy and position of the event. With $\mu(\vec{r}_0, E)$ the
 528 mean expected number of PE detected by each PMT, the probability to observe N_{pe} in a PMT follow
 529 a Poisson distribution. Thus

- 530 — The probability to observe no hit ($N_{pe} = 0$) in the j th PMT is $P_{\text{nohit}}^j(\vec{r}_0, E) = e^{-\mu_j}$
- 531 — The probability to observe $N_{pe} \neq 0$ in the i th PMT is $P_{\text{hit}}^i(\vec{r}_0, E) = \frac{\mu^{N_{pe}} e^{-\mu_i}}{N_{pe}!}$

532 Therefore, the probability to observe a specific hit pattern can be expressed as

$$P(\vec{r}_0, E) = \prod_j P_{\text{nohit}}^j(\vec{r}_0, E) \cdot \prod_i P_{\text{hit}}^i(\vec{r}_0, E) \quad (2.12)$$

533 The best fit values of \vec{R}_0 and E can then be calculated by minimizing the negative log-likelihood

$$\mathcal{L}(\vec{r}_0, E) = -\ln(P(\vec{r}_0, E)) \quad (2.13)$$

534 In principle, $\mu_i(\vec{r}_0, E)$ could be expressed

$$\mu_i(\vec{r}_0, E) = Y \cdot \frac{\Omega(\vec{r}_0, r_i)}{4\pi} \cdot \epsilon_i \cdot f(\theta_i) \cdot e^{-\sum_m \frac{d_m}{\zeta_m}} \cdot E + \delta_i \quad (2.14)$$

535 where Y is the energy scale factor, $\Omega(\vec{r}_0, r_i)$ is the solid angle of the i th PMT, ϵ_i is its detection
 536 efficiency, $f(\theta_i)$ its angular response, ζ_m is the attenuation length in the materials and δ_i the expected
 537 number of dark noise.

538 However Eq. 2.14 assume that the scintillation light yield is linear with energy and describe poorly
 539 the contribution of indirect light, shadow effect due to the supporting structure and the total reflec-

tion effects. The solution is to use data driven methods to produce the pdf by using the calibrations sources and position described in section 2.3. In the results presented in figures 2.16, the PDF was produced using MC simulation and 29 specific calibrations position [36] along the Z-axis of the detector. We see that the charge likelihood algorithm show little bias in the TR area and a better

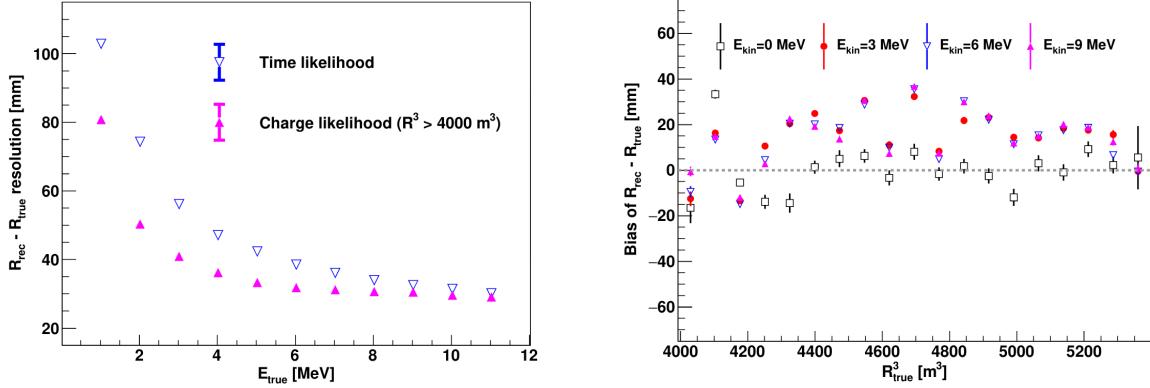


FIGURE 2.16 – On the left: Resolution of the reconstructed R as a function of the energy in the TR area ($R^3 > 4000 \text{ m}^3 \equiv R > 16 \text{ m}$) by the charge and time likelihood algorithms. On the right: Bias of the reconstructed R in the TR area for different energies by the charge likelihood algorithm

resolution than the time likelihood. The figure 2.17 shows the radial resolution of the different algorithm presented for this section, we can see the refinement at each step and that the charge likelihood yield the best results.

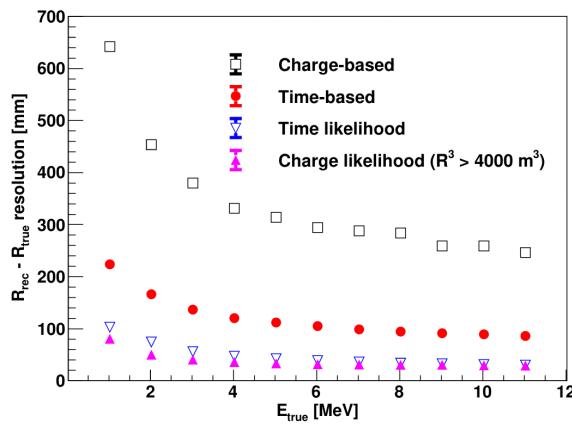


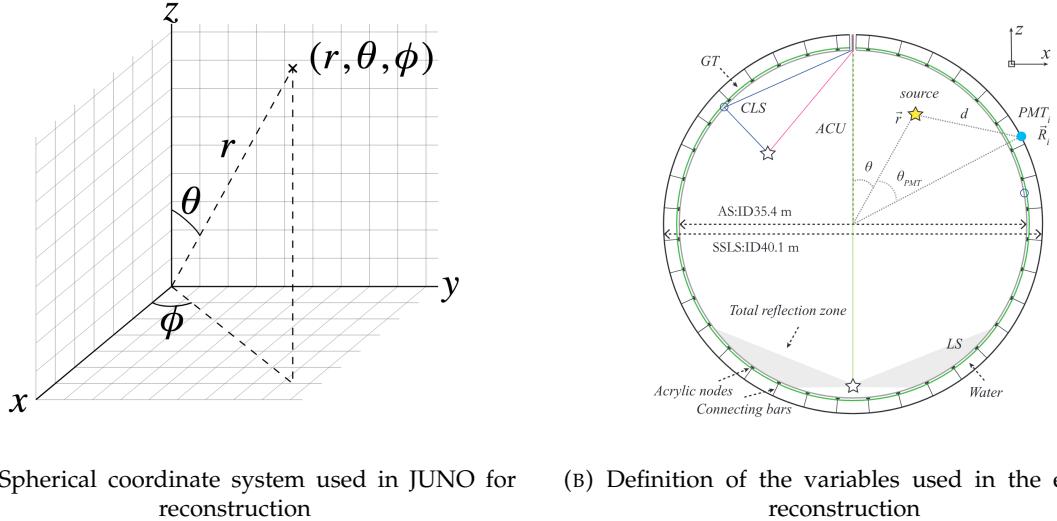
FIGURE 2.17 – Radial resolution of the different vertex reconstruction algorithms as a function of the energy

The charge based likelihood algorithms already give use some information on the energy as Eq. 2.13 is minimized but the energy can be further refined as shown in the next section.

2.6.2 Energy reconstruction

As explained in section 2.1.1, energy resolution is crucial for the NMO and oscillation parameters measurements. Thus the energy reconstruction algorithm should take into consideration as much

552 detector effect as possible. The following method is a data driven method based on calibration
 553 samples inspired by the charge likelihood algorithm described above [40].



(A) Spherical coordinate system used in JUNO for reconstruction

(B) Definition of the variables used in the energy reconstruction

FIGURE 2.18

554 Charge estimation

555 The most important element in the energy reconstruction is $\mu_i(\vec{r}_0, E)$ described in Eq. 2.14. For
 556 realistic cases, we also need to take into account the electronics effect that were omitted in the
 557 previous section. Those effect will cause a charge smearing due to the uncertainties in the N_{pe}
 558 reconstruction. Thus we define $\hat{\mu}^L(\vec{r}_0, E)$ which is the expected N_{pe}/E in the whole detector for an
 559 event with visible energy E_{vis} and position \vec{r}_0 . The position of the event and PMTs are now defined
 560 using $(r, \theta, \theta_{pmt})$ as defined in figure 2.18b.

$$\hat{\mu}(r, \theta, \theta_{pmt}, E_{vis}) = \frac{1}{E_{vis}} \frac{1}{M} \sum_i^M \frac{\bar{q}_i - \mu_i^D}{\text{DE}_i}, \quad \mu_i^D = \text{DNR}_i \cdot L \quad (2.15)$$

561 where i runs over the PMTs with the same θ_{pmt} , DE_i is the detection efficiency of the i th PMT. μ_i^D
 562 is the expected number of dark noise photoelectrons in the time window L . The time window have
 563 been optimized to $L = 280$ ns [40]. \bar{q}_i is the average recorded photoelectrons in the time window
 564 and \hat{Q}_i is the expected average charge for 1 photoelectron. The N_{pe} map is constructed following the
 565 procedure described in [35].

566 Time estimation

567 The second important observable is the hit time of photons that was previously defined in Eq. 2.7. It
 568 is here refined as

$$t_r = t_h - \text{tof} - t_0 = t_{LS} + t_{TT} \quad (2.16)$$

569 where t_h is the time of hit, t_{LS} is the scintillation time and t_{TT} the transit time of PMTs that is described
 570 by a gaussian

$$t_{TT} = \mathcal{N}(\overline{\mu_{TT} + t_d}, \sigma_{TT}) \quad (2.17)$$

571 where μ_{TT} is the mean transit time in PMTs, σ_{TT} is the Transit Time Spread (TTS) of the PMTs and t_d
 572 is the delay time in the electronics. The effective refraction index of the LS is also corrected to take
 573 into account the propagation distance in the detector.

574 The timing PDF $P_T(t_r|r, d, \mu_l, \mu_d, k)$ can now be generated using calibration sources [40]. This PDF
 575 describe the probability that the residual time of the first photon hit is in $[t_r, t_r + \delta]$ with r the radius
 576 of the event vertex, $d = |\vec{r} - \vec{r}_{PMT}|$ the propagation distance, μ_l and μ_d the expected number of PE
 577 and dark noise in the electronic reading window and k is the detected number of PE.

578 Now let denote $f(t, r, d)$ the probability density function of "photoelectron hit a time t" for an event
 579 happening at r where the photons traveled the distance d in the LS

$$F(t, r, d) = \int_t^L f(t', r, d) dt' \quad (2.18)$$

580 Based on the PDF for one photon $k = 1$, one can define

$$P_T^l(t|k = n) = I_n^l [f_l(t) F_l^{n-1}(t)] \quad (2.19)$$

581 where the indicator l means that the photons comes from the LS and I_n^l a normalisation factor. To this
 582 pdf we add the probability to have photons coming from the dark noise indicated by the indicator d
 583 using

$$f_d(t) = 1/L, F_d(t) = 1 - \frac{t}{L} \quad (2.20)$$

584 and so for the case where only one photon is detected by the PMT ($k = 1$)

$$P_T(t|\mu_l, \mu_d, k = 1) = I_1[P(1, \mu_l)P(0, \mu_d)f_l(t) + P(0, \mu_l)P(1, \mu_d)f_d(t)] \quad (2.21)$$

585 where $P(k_\alpha, \mu_\alpha)$ is the Poisson probability to detect k_α PE from $\alpha \in \{l, d\}$ with the condition $k_l + k_d = k$.
 586

587 Now that we have the individual timing and charge probability we can construct the charge likelihood
 588 referred as QMLE:

$$\mathcal{L}(q_1, q_2, \dots, q_N | \vec{r}, E_{vis}) = \prod_{j \in \text{unfired}} e^{-\mu_j} \prod_{i \in \text{fired}} \left(\sum_{k=1}^K P_Q(q_i|k) \cdot P(k, \mu_i) \right) \quad (2.22)$$

589 where $\mu_i = E_{vis}\hat{\mu}_i^L + \mu_i^D$ and $P(k, \mu_i)$ is the Poisson probability of observing k PE. $P_Q(q_i|k)$ is the
 590 charge pdf for k PE. And we can also construct the time likelihood referred as TMLE:

$$\mathcal{L}(t_{1,r}, t_{2,r}, \dots, t_{N,r} | \vec{r}, t_0) = \prod_{i \in \text{hit}} \frac{\sum_{k=1}^K P_T(t_{i,r}|r, d, \mu_i^l, \mu_i^d, k) \cdot P(k, \mu_i^l + \mu_i^d)}{\sum_{k=1}^K P(k, \mu_i^l + \mu_i^d)} \quad (2.23)$$

591 where K is cut to 20 PE and hit is the set of hits satisfying $-100 < t_{i,r} < 500$ ns.

592 Merging those two likelihood give the charge-time likelihood QTMLLE

$$\mathcal{L}(q_1, q_2, \dots, q_N; t_{1,r}, t_{2,r}, \dots, t_{N,r} | \vec{r}, t_0, E_{vis}) = \mathcal{L}(q_1, q_2, \dots, q_N | \vec{r}, E_{vis}) \cdot \mathcal{L}(t_{1,r}, t_{2,r}, \dots, t_{N,r} | \vec{r}, t_0) \quad (2.24)$$

593 The radial and energy resolutions of the different likelihood are presented in figure 2.19 (from [40]).
 594 We can see the improvement of adding the time information to the vertex reconstruction and that
 595 an increase in vertex precision can bring improvement in the energy resolution, especially at low
 596 energies.

597 Data driven methods prove to be performant in the energy and vertex reconstruction given that we
 598 have enough calibrations sources to produce the PDF. In the next section, we'll see another type of

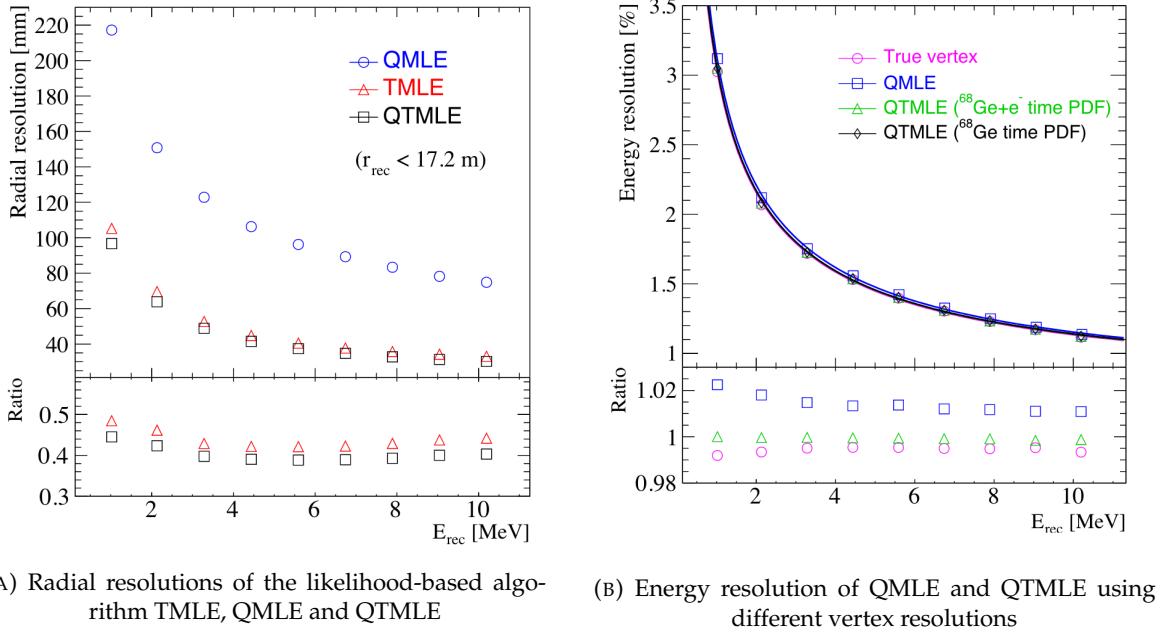


FIGURE 2.19

599 data-driven method based on machine learning.

600 2.6.3 Machine learning for reconstruction

601 Machine learning (ML) is family of data-driven algorithms that are inferring behavior and results
 602 from a training dataset. A overview of methods and detailed explanation of the Neural Network
 603 (NN) subfamily can be found in Chapter 3.

604 The power of ML is the ability to model complex response to a specific problem. In JUNO the
 605 reconstruction problematic can be expressed as follow: knowing that each PMT, large or small,
 606 detected a given number of PE Q at a given time t and their position is x, y, z where did the energy
 607 was deposited and how much energy was it, modeling a function that naively goes:

$$\mathbb{R}^{5 \times N_{pmt}} \mapsto \mathbb{R}^4 \quad (2.25)$$

608 It is worth pointing that while this is already a lot in informations, this is not the rawest representa-
 609 tion of the experiment. We could indeed replace the charge and time by the waveform in the time
 610 window of the event but that would lead to an input representation size that would exceed our
 611 computational limits. Also, due to those computational limits, most of the ML algorithm reduce this
 612 input phase space either by structurally encoding the information (pictures, graph), by aggregating
 613 it (mean, variance, ...) or by exploiting invariance and equivariance of the experiment (rotational
 614 invariance due to the sphericity, ...).

615 For machine learning to converge to performant algorithm, a large dataset exploring all the phase
 616 space of interest is needed. For the following studies, data from the monte carlo simulation presented
 617 in section 2.5 are used for training. When the detector will be finished calibrations sources will be
 618 complementarily be used.

619 **Boosted Decision Tree (BDT)**

620 On of the most classic ML method used in physics in last years is the Boosted Decision Tree (see
 621 chapter 3.1). They have been explored for vertex reconstruction [41] et for energy reconstruction [41,
 622 42].

623 For vertex and energy reconstruction a BDT was developed using the aggregated informations pre-
 624 sented in 2.6.

Parameter	description
$nHits$	Total number of hits
$x_{cc}, y_{cc}, z_{cc}, R_{cc}$	Coordinates of the center of charge
ht_{mean}, ht_{std}	Hit time mean and standard deviation

TABLE 2.6 – Features used by the BDT for vertex reconstruction

625 Its reconstruction performances are presented in figure 2.21.

626 A second and more advanced BDT, subsequently named BDTE, that only reconstruct energy use a
 627 different set of features [42]. They are presented in the table 2.7

628 **Neural Network (NN)**

629 The physics have shown a rising for Neural Network (NN) in the past years for event reconstruction,
 630 notably in the neutrino community [43–46]. Three type of neural networks have explored for event
 631 reconstruction in JUNO Deep Neural Network (DNN), Convolutional Neural Network (CNN) and
 632 Graph Network (GNN). More explanation about those neural network can be found in chapter 3.

633 The CNN are using 2D projection of the detector representing it as an image with two channel, one
 634 for the charge Q and one for the time t . The position of the PMTs is structurally encoded in the pixel
 635 containing the information of this PMT. In [41], the pixel is chosen based on a transformation of θ
 636 and ϕ coordinates to the 2D plane and rounded to the nearest pixel. A sufficiently large image has
 637 been chosen to prevent two PMT to be located in the same pixel. An example of this projection can
 638 be found in figure 2.20. The performances of the CNN can be found in figure 2.21.

639 Using 2D have the upside of encoding a large part of the informations structurally but loose the rota-
 640 tional invariance of the detector. It also give undefined information to the neural network (what is a
 641 pixel without PMT ? What should be its charge and time ?), cause deformation in the representation
 642 of the detector (sides of projection) and loose topological informations.

643 One of the way to present structurally the sphericity of JUNO to a NN is to use a graph: A collection
 644 of objects V called nodes and relations E called edges, each relation associated to a couple v_1, v_2

AccumCharge	$ht_{5\%-2\%}$
R_{cht}	pe_{mean}
z_{cc}	J_{cht}
pe_{std}	ϕ_{cc}
nPMTs	$ht_{35\%-30\%}$
$ht_{kurtosis}$	$ht_{20\%-15\%}$
$ht_{25\%-20\%}$	$pe_{35\%}$
R_{cc}	$ht_{30\%-25\%}$

TABLE 2.7 – Features used by the BDTE algorithm. pe and ht reference the charge
 and hit-time distribution respectively and the percentages are the quantiles of those
 distributions. cht and cc reference the barycenters of hit time and charge respectively

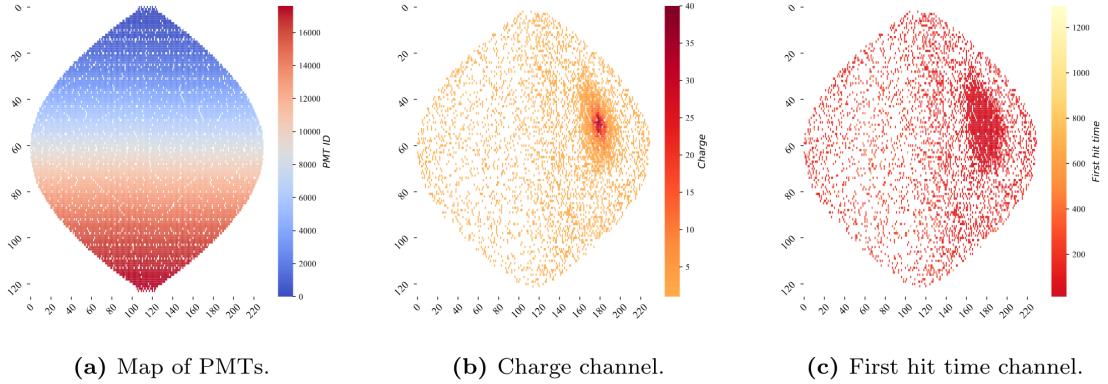


FIGURE 2.20 – Projection of the LPMTs in JUNO on a 2D plane. (a) Show the distribution of all PMTs and (b) and (c) are example of what the charge and time channel looks like respectively

645 forming the graph $G(E, V)$. Nodes and edges can hold informations or features. In [41] the nodes,
 646 are geometrical region of the detector as defined by the HealPix [47]. The features of the nodes are
 647 aggregated informations from the PMTs it contains. The edges contains geographic informations of
 648 the nodes relative positions.

649 This data representation has the advantages to keep the topology of the detector intact. It also permit
 650 the use of rotational invariant algorithms for the NN, thus taking advantage of the symmetries of the
 651 detector.

652 The neural network then process the graph using Chebyshev Convolutions [48]. The performances
 653 of the GNN are presented in figure 2.21.

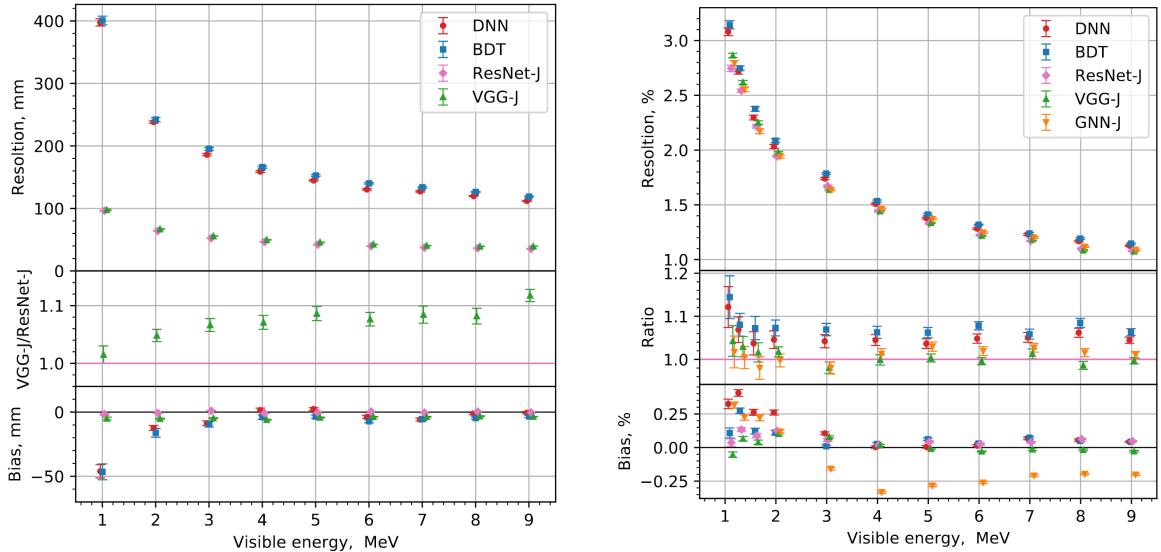


FIGURE 2.21 – Radial (left) and energy (right) resolutions of different ML algorithms.
 The results presented here are from [41]. DNN is a deep neural network, BDT is a BDT,
 ResNet-J and VGG-J are CNN and GNN-J is a GNN.

654 Overall ML algorithms show similar performances as classical algorithms in term of energy recon-
 655 structions with the more complex structure CNN and GNN showing better performances than BDT

and DNN. For vertex reconstruction, the BDT and DNN show poor performance while CNN are on the level of the classical algorithms.

2.7 JUNO sensitivity to NMO and precise measurements

Now that the event have been reconstructed, selected and that the non-IBD background have been rejected, we have access to the measured energy flux from JUNO. We consider two spectra, the one measured by the LPMT system and the one measured by the SPMT system. This give rise to three possible analysis: A LPMT only analysis, a SPMT only analysis and a joint analysis. This joint analysis is the subject of the chapter 7 of this thesis.

The following details about JUNO measurement is common to the three analysis. The details and specific of the joint analysis are detailed in chapter 7.

2.7.1 Theoretical spectrum

To extract the oscillation parameters and the NMO from the measured spectrum, it is compared to a theoretical spectrum. This theoretical spectrum is produced based on the theory of the three flavour oscillation (see section 1.3), the measurements of the calibration and satellite experiments and Monte Carlo simulation:

- The absolute flux and the fission product fraction calibrated by TAO.
- The estimation of the neutrinos flux from other sources, such as the geoneutrinos, by theoretical model.
- The computed cross-section of $\bar{\nu}_e$ and the LS.
- The estimation of mislabelled event, such as fast neutron events from cosmic muons, using Monte Carlo simulation.
- The measured bias and resolution of the LPMT and SPMT system by the calibration.
- The time dependent reactor parameters (age of fuel, instantaneous power of the reactors, etc...)

These systematics parameters come with their uncertainties that need to be taken into account by the fitting framework. This theoretical spectrum will, in the end, depend of the oscillation parameters of interest θ_{13} , θ_{12} , Δm_{21}^2 , Δm_{31}^2 . Noise parameters can be included in the parameters spectrum such as the earth density ρ between the power plants and JUNO.

2.7.2 Fitting procedure

The theoretical and measured spectra are represented as two histograms depending on the energy. The theoretical spectrum is adjusted with the data using a χ^2 minimization where χ^2 is naively defined as

$$\chi^2 = \sum_i \frac{(N_{th}^i - N_{data}^i)^2}{\sigma_i^2} \quad (2.26)$$

where N_{th}^i is the number event in the i th bin of the theoretical spectrum, N_{data}^i is the number of event in the i th bin of the measured spectrum and σ_i is the uncertainty of this bin. Two classic statistic test exist Pearson and Neyman where the difference is the estimation of σ_i parameters.

This σ_i is composed of the systematics uncertainties discussed above but also from the statistic uncertainty of the spectrum. Considering a Poisson process, the statistic uncertainty is estimated as $\sigma_{stat}^i = \sqrt{N^i}$. In a Pearson test, $N^i \equiv N_{th}^i$ whereas in a Neyman test $N^i \equiv N_{data}^i$. Under the assumption that the content of each bin follow a Gaussian distribution (a Poisson with high enough statistic), the two test are equivalent. But studies on Monte Carlo spectrum showed that the Pearson

and Neyman statistic are biased in opposite direction. It is easily visible where, for the same data, Pearson will prefer a higher N_{th}^i to reduce the ratio $\frac{1}{N_{th}^i}$ whereas Neyman will prefer a lower N_{th}^i to reduce the $(N_{th}^i - N_{data}^i)$ term.

This problematic can be circumvented by summing the two test, yielding the CNP statistic test and/or by adding a term

$$\chi^2 = \sum_i \frac{(N_{th}^i - N_{data}^i)^2}{\sigma_i^2} - \ln |\mathbf{V}| \quad (2.27)$$

where V is the covariance matrix of the theoretical spectrum yielding the PearsonV and CNPV statistic test.

The χ^2 is minimized by exploring the parameter phase space via gradient descent.

2.7.3 Physics results

The oscillation parameters are directly extracted from the minimization procedure and the error can be estimated directly from the procedure. For the NMO, the data are fitted under the two assumption of NO and IO. The difference in χ^2 give us the preferred ordering and the significance of our test. Latest studies show that the precision on oscillation parameters after six year of data taking will be of 0.2%, 0.3%, 0.5% and 12.1% for Δm_{31}^2 , Δm_{21}^2 , $\sin^2 \theta_{12}$ and $\sin^2 \theta_{13}$ respectively [11]. The expected sensitivity to mass ordering is 3σ after 6 years [49].

2.8 Summary

JUNO is one the biggest new generation neutrino experiment. Its goal, the measurements of oscillation parameters with unprecedeted precision and an NMO preference at the 3 sigma confidence level, needs an in depth knowledge and understanding of the detector and the physics at hand. The characterisation and calibration of the detector are of the utmost importance and the understanding of the detector response in its resolution and bias is capital to be able to correctly carry the high precision physics analysis of the neutrino oscillation.

In this thesis, I explore the usage of data-driven reconstruction methods to validate and optimize the reconstruction of IBD events in JUNO in the chapters 4, 5 and 6 and the usage of the dual calorimetry in the detection of possible mis-modelisation in the theoretical spectrum 7.

⁷²⁰ **Chapter 3**

⁷²¹ **Machine learning and Artificial
Neural Network**

⁷²³ *"I have the shape of a human being and organs equivalent to those of a human being. My organs, in fact, are identical to some of those in a prostheticized human being. I have contributed artistically, literally, and scientifically to human culture as much as any human being now alive. What more can one ask?"*

Isaac Asimov, *The Complete Robot*

⁷²⁴ Machine Learning (ML) and more specifically Neural Network (NN) are families of data-driven ⁷²⁵ algorithm. They are used to model complex distributions from a finite dataset to extract a generalist ⁷²⁶ behavior. They learn, adapt their intrinsic parameters, interactively by computing its performance ⁷²⁷ or loss on those dataset. They take advantage of simple microscopic operation such as *if condition* or ⁷²⁸ non-continuous but differentiable function like *ReLU*. Through optimizers and the combination of a ⁷²⁹ lot of those microscopic operations, they can obtain complex and precise behaviours.

⁷³⁰ They are now widely used in a wide variety of domain including natural language processing, ⁷³¹ computer vision, speech recognition and, the subject of this thesis, scientific studies.

⁷³² We found them in particle physics, either as the main algorithm or as secondary algorithm, for event ⁷³³ reconstruction, event classification, waveform reconstruction, etc..., domains where the underlying ⁷³⁴ physic and detector process is complex and highly dimensional. Physicists have traditionally been ⁷³⁵ forced to use simplifications or assumptions to ease the development of algorithms or equations ⁷³⁶ (a good example is the algorithm presented in section 2.6) where machine learning could refine and ⁷³⁷ take into account those effects, provided that they have enough data and computing power.

⁷³⁸ This chapter present an overview of the different kind of machine learning methods and neural ⁷³⁹ networks that will be discussed in this thesis.

⁷⁴⁰ **3.1 Boosted Decision Tree (BDT)**

⁷⁴¹ One of the most classic machine learning algorithm used in particle physics is Boosted Decision Tree ⁷⁴² (BDT) [50] (or more recently Gradient Boosting Machine [51]). The principle of a BDT is fairly simple ⁷⁴³ : based on a set of observables, a serie of decisions, represented as node in a tree, are taken by the ⁷⁴⁴ algorithm. Each decision point, or node, takes its decision based on a set of trainable parameters ⁷⁴⁵ leading to a subtree of decision. The process is repeated until it reach the final node, yielding the ⁷⁴⁶ prediction. A simplistic example is given in figure 3.1.

⁷⁴⁷ The training procedure follow a simple score reward procedure. During the training phase the ⁷⁴⁸ prediction of the BDT is compared to a known truth about the data. The score is then used to

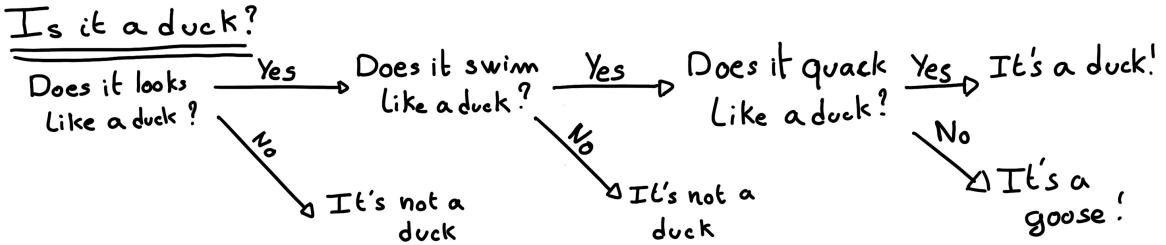


FIGURE 3.1 – Example of a BDT that determine if the given object is a duck

749 backpropagate corrections to the parameters of the tree. Modern BDT use gradient boosting where
 750 the gradient of the loss is calculated for each of the BDT parameters. Following the gradient descent,
 751 we can reach the, hopefully, global minima of the loss for our set of parameters.

752 3.2 Artificial Neural Network (NN)

753 One other big family of machine learning algorithm is the artificial Neural Networks (NN). The idea
 754 of developing automates which component mimic, in a simplistic way, the behavior of biological
 755 neurons emerge in 1959 with the paper “*What the Frog’s Eye Tells the Frog’s Brain*” [52]. They develop
 756 an automate where each component possess an *activation function*. Each one of those component then
 757 transmit its information to the other following a certain efficiency or *weight*. Those works influenced
 758 scientist and notably Frank Rosenblatt who published in 1958 what is considered the first neural
 759 network model the Perceptron [53].

760 Modern neural network still nowadays use the neuron metaphor to represent neural network, but
 761 approach them as a graph where the nodes are neurons possessing an activation function and edges
 762 holding the weights, or *parameters* in modern literature, between those nodes. Most of the modern
 763 neural network work with the principle of neurons layers. Each neurons belong to a layer and takes
 764 input from the preceding layer and forward it result to next layer. For example the most basic set
 765 layer is the fully connected layer where each of its neurons is connected to every other neurons of
 766 the precessing layer. All the neurons posses the same activation function F . The connection between
 767 two the two layers is expressed as a tensor T_j^i where i is the index of the precedent layer and j the
 768 index of the current layer. The propagation from the layer I to J is then described as

$$J_j = F_j(T_j^i I_i + B_j) \quad (3.1)$$

769 where the learning parameters are the tensor T_j^i and the bias tensor B_j . This is the fundamental
 770 component of the Fully Connected Deep NN (FCDNN) family presented in section 3.2.1. Most of the
 771 modern neural networks use gradient descent to optimize their parameters, i.e. the gradient of the
 772 parameter θ in respect of the loss function \mathcal{L} is subtracted to it

$$\theta_{i+1} = \theta_i - \frac{\partial \mathcal{L}}{\partial \theta} \quad (3.2)$$

773 i being the training iteration index. This needs the expression of \mathcal{L} dependent of θ to be differentiable,
 774 thus the layer and their activation function also need to be differentiable. This simple gradient
 775 descent, designated as Stochastic Gradient Descent (SGD), can be completed with first and second
 776 order momentum like with the Adam optimizer [54] (more details in section 3.2.5).

777 This description of neural networks as layer introduced the principle of *depth* and *width*, the number
 778 of layers in the NN and the number of neurons in each layer respectively. Those quantities that not

779 directly used for the computation of the results but describe the NN or its training are designated as
 780 *hyperparameters*.

781 The loss \mathcal{L} described above is a score representing how well the NN is doing. As seen above, it
 782 needs to be differentiable with respect to the parameter of the NN. Depending if we try to minimize
 783 or maximize it, it need to posses a minima or a maxima. For example when doing *regression*, i.e.
 784 produce a scalar result, a common loss is the Mean Square Error (MSE). Let i be our dataset, y_i be the
 785 target scalar, x_i the input data and $f(x_i, \theta)$ the result of the network. The network here is modelled by
 786 f , and its parameter by the set

$$\mathcal{L} := MSE = \frac{1}{N} \sum_i^N (y_i - f(x_i))^2 \quad (3.3)$$

787 Another common loss function is the Mean Absolute Error (MAE)

$$\mathcal{L} := MAE = \frac{1}{N} \sum_i^N |y_i - f(x_i)| \quad (3.4)$$

788 3.2.1 Fully Connected Deep Neural Network (FCDNN)

789 Fully Connected Deep Neural Network (FCDNN) architecture is the natural evolution of the Perceptron.
 790 The input data is represented as a first order tensor I_j and then fed forward to multiple fully
 791 connected layers (Eq 3.1) as presented in the figure 3.2a. Most of the time, the classic ReLU function

$$\text{ReLU}(x) = \begin{cases} x & \text{if } x \geq 0 \\ 0 & \text{otherwise} \end{cases} \quad (3.5)$$

792 is used as activation function. Prelu and Sigmoid are also popular choices:

$$\text{Sigmoid}(x) = \frac{1}{1 + e^{-x}} \quad (3.6) \quad \text{PReLU}(x) = \begin{cases} x & \text{if } x \geq 0 \\ \alpha x & \text{otherwise} \end{cases} \quad (3.7)$$

793 The reasoning behind ReLU and PReLU is that with enough of them, you can mimic any continuous
 794 function as illustrated in figure 3.2b. Sigmoid is more used in case of classification, its behavior going
 795 hand in hand with the Cross Entropy loss function used in classification problems.

796 Due to its simplicity, FCDNN are also used as basic pieces for more complex architectures such as
 797 the CNN and GNN that will be presented in the next section.

799 3.2.2 Convolutional Neural Network (CNN)

800 Convolutional Neural Networks are a family of neural networks that use discrete convolution filters,
 801 as illustrated in an example in figure 3.3, to process the input data, often images. They have the
 802 advantage to be translation invariant by construction, this mean that they are capable of detecting
 803 oriented features independently of their location on the image. The learning parameters are located
 804 in the filters, the network thus learn the optimal filters to extract the desired features. 2D CNN,
 805 where the filters are second order tensors that span over third order tensors, are commonly used in
 806 image recognition [55] for classification or regression problematics.

807 The convolution layers are commonly chained [56], reducing the input dimension while increasing
 808 the number of filters. The idea behind is that the first layers will process local informations and the
 809 latest layers will process more global informations. To try to preserve the amount of information, we
 810 tend to double the numbers of filters for each division of the input data. The results of the convolution

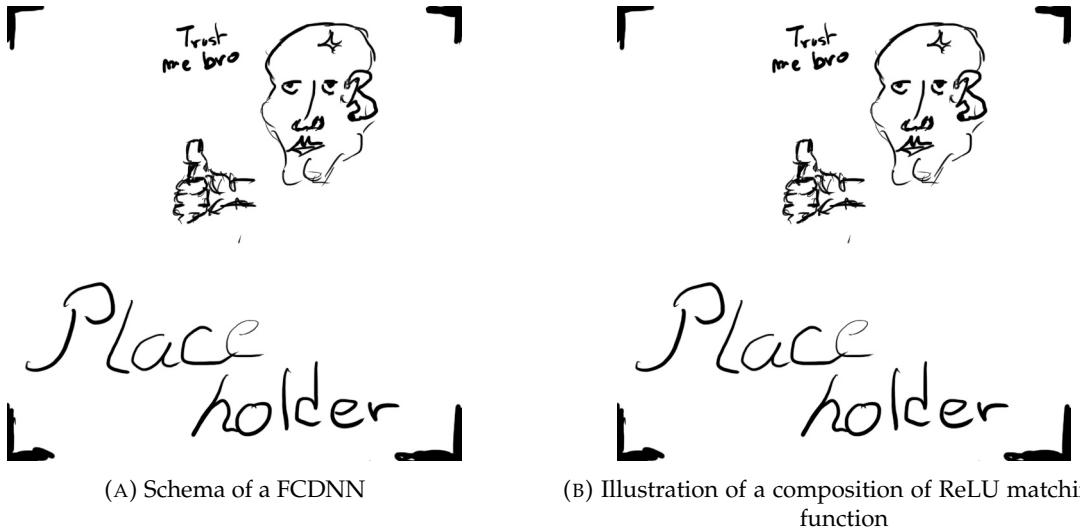


FIGURE 3.2

811 filters is commonly then flattened and feed to a smaller FCDNN which will process the filters results
812 to yield the desired output.

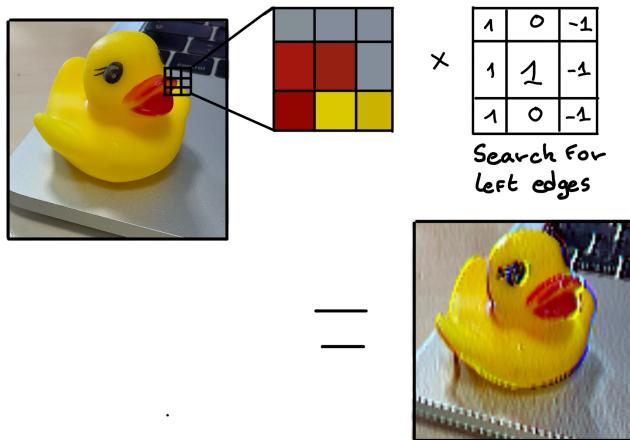


FIGURE 3.3 – Illustration of the effect of a convolution filter. Here we apply a filter with the aim do detect left edges. We see in the resulting image that the left edges of the duck are bright yellow where the right edges are dark blue indicating the contour of the object. The convolution was calculated using [57].

813 As an example, let's take the Pytorch [58] example for the MNIST [59], a dataset of black and white
814 images of handwritten digits. Those images are 28×28 pixels with only one channel corresponding
815 to the grey level of the pixel. Example of images from this dataset are presented in figure 3.4a

816 A schema of the CNN used in the Pytorch example is presented in figure 3.4b. Using this schema as
817 a reference, the trained network is made of:

- 818 1. A convolutional layer of (3×3) filters yielding 32 channels. A bias parameter is applied
819 to each channel for a total of $(32 \cdot (3 \times 3) + 32) = 320$ parameters. The resulting image is
820 $(26 \times 26 \times 32)$ (26 per 26 pixels with 32 channels). The ReLU activation function is applied to
821 each pixel.
- 822 2. A second convolutional layer of (3×3) filters yielding 64 channels. This channel also posses

823 a bias parameter for a total of $(64 \cdot (3 \times 3) + 64) = 640$ parameters. Resulting image is $(24 \times$
 824 $24 \times 64)$. Also with with a ReLU activation function.

825 3. Then comes a (2×2) max pool layer with a stride of 1 meaning that for each channel the max
 826 value of pixels in a (2×2) block is condensed in a single resulting pixel. The resulting image
 827 is $(12 \times 12 \times 64)$.

828 4. This image goes through a dropout layer which will set the pixel to 0 with a probability of
 829 0.25. This help prevent overtraining of the neural network (see section 3.2.6 for more details).

830 5. The data is the flattened i.e. condensed into a vector of $(12 \times 12 \times 64) = 9216$ values.

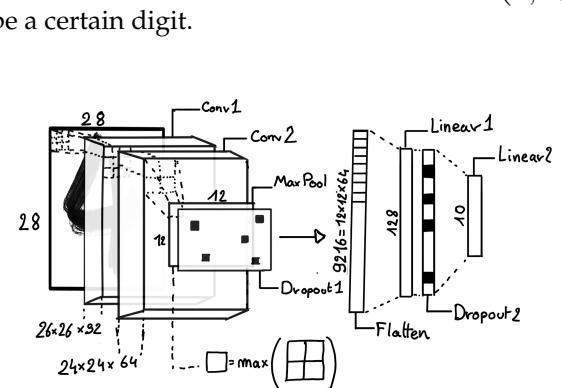
831 6. Then comes a fully connected linear layer (Eq. 3.1) with a ReLU activation that output 128
 832 feature. It needs $(9216 \cdot 128) + 128 = 1'179'776$ parameters.

833 7. This 128 item vector goes through another dropout layer with a probability of 0.5

834 8. The vector is then transformed through a linear layer with ReLU activation. It output 10
 835 values, one for each digit class (0, 1, 2, ..., 9). It need $(128 \cdot 10) + 128 = 1408$ parameters.

836 9. Finally the 10 values are normalized using a log softmax function $\text{LogSoftmax}(x_i) = \log \left(\frac{\exp(x_i)}{\sum_j \exp(x_j)} \right)$
 837 to give the probability of the input image to be a certain digit.

0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3
 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4
 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5
 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6
 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7
 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8
 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9



(A) Example of images in the MNIST dataset

(B) Schema of the CNN used in Pytorch example to process the MNIST dataset

FIGURE 3.4

838 The final network needs 1'182'144 parameters or, if we consider each parameters to be a double
 839 precision floating point, 9.45 MB of data. To gives a order of magnitude, such neural network is
 840 considered "simple", train in a matter of minutes on T4 GPU [60] (14 epochs) and reach an accuracy
 841 in its prediction of 99%.

842 3.2.3 Graph Neural Network (GNN)

843 Graph neural network is a family of neural network where the data is represented as a graph $G(\mathcal{N}, \mathcal{E})$
 844 composed of vertex or node $n \in \mathcal{N}$ and edges $e \in \mathcal{E}$. The edges are associated to two nodes $(u, v) \in$
 845 \mathcal{N}^2 , "connecting" them. The node and the edges can hold features, commonly represented as vector
 846 $n \in \mathbb{R}^{k_n}$, $e \in \mathbb{R}^{k_e}$. We can thus define a graph using two tensors A_e^{ij} the adjacency tensors that hold
 847 the features e of the edge connecting the node i and j and the tensor N_v^i that hold the features v of a
 848 node i .

849 To efficiently manipulate such object we need to structurally encode their property in the neural
 850 network architecture: each node is equivalent (as opposite to ordered data in a vector), each node has
 851 a set of neighbours, ... One of this method is the message passing algorithm presented historically

852 in “Neural Message Passing for Quantum Chemistry” [61]. In this algorithm, with each layer of
 853 message passing a new set of features is computed for each node following

$$n_i^{k+1} = \phi_u(n_i^k, \square_j \phi_m(n_i^k, n_j^k, e_{ij}^k)); n_j \in \mathcal{N}'_i \quad (3.8)$$

854 where ϕ_u is a differentiable update function, \square_j is a differentiable aggregation function and ϕ_m is a
 855 differentiable message function. $\mathcal{N}'_i = \{n_j \in \mathcal{N} | (n_i, n_j) \in \mathcal{E}\}$ is the set of neighbours of n_i , i.e. the
 856 nodes n_j from which it exist an edge $e_{i,j} \rightarrow (n_i, n_j)$. k is the layer on which the message passing
 857 algorithm is applied. \square need also a few other property if we want to keep the graph property, most
 858 notably the permutational invariance of its parameters (example: mean, std, sum, ...).

859 The edges features can also be updated, either by directly taking the results of ϕ_m or by using another
 860 message function ϕ_e .

861 Message passing is a very generic way of describing the process of GNN and it can be specialized
 862 for convolutional filtering [48], diffusion [62] and many other specific operation. GNN are used in a
 863 wide variety of application such as regression problematics, node classification, edge classification,
 864 node and edge prediction, ...

865 It is a very versatile but complex tool.

866 3.2.4 Adversarial Neural Network (ANN)

867 The adversarial machine learning, Adversarial Neural Networks (ANN) in the case of neural net-
 868 work, is a family of unsupervised machine learning algorithms where the learning algorithm (gen-
 869 erator) is competing against another algorithm (discriminator). Taking the example of Generative
 870 Adversarial Networks, concept initially developed by Goodfellow et al. [63], the discriminator goal
 871 is to discriminate between data coming from a reference dataset and data produced by the generator.
 872 The generator goal, on the other hand, is to produce data that the discriminator would not be able to
 873 differentiate from data from the reference dataset. The expression of duality between the two models
 874 is represented in the loss where, at least a part of it, is driven by the results of the discriminator.

875 3.2.5 Training procedure

876 A neural network without the adequate training is like an empty shell. If the parameters are not
 877 optimized they are, most of the time, initialized to random number and so the output will just be
 878 random. The training is a key step in the production of a solid and reliable NN. This section aim to
 879 give an overview of the different concept and tools used in the training of our neural networks.

880 Training lifecycle

881 The training of NN does not follow strict rules, you could imagine totally different lifecycle but I will
 882 describe here the one used in this thesis, the most common one.

883 The training is split into *epochs* during which the NN will train on a set of subsamples called *batch*.
 884 The size of those batch is called *batch size*, a.k.a. the number of data it contains (how many images,
 885 how many events,...). Each process of a batch is called a *step*. At the end of each epochs, the neural
 886 network is evaluated over a validation dataset. This validation dataset is not used for training (no
 887 gradient of the loss is computed) and is used as reference for the network performance and monitor
 888 overtraining (see section 3.2.6). Most of the time, the parameters are updated at each step using the
 889 mean loss over the batch and the optimizer hyperparameters are updated at each epochs.

890 **The optimizer**

891 As briefly introduced section 3.2, the parameters of the neural network are optimized using the
 892 gradient descent method. We calculate the gradient of the mean loss over the batch with respect
 893 of each parameters and we update the parameters in accord to minimize the loss. The gradient is
 894 computed backward from the loss up to the first layer parameters using the chain rule:

$$\frac{\partial \mathcal{L}}{\partial \theta_1} = \frac{\partial \theta_2}{\partial \theta_1} \frac{\partial \mathcal{L}}{\partial \theta_2} = \frac{\partial \theta_2}{\partial \theta_1} \frac{\partial \theta_3}{\partial \theta_2} \frac{\partial \mathcal{L}}{\partial \theta_3} = \frac{\partial \theta_2}{\partial \theta_1} \prod_{i=2}^{N-1} \frac{\partial \theta_{i+1}}{\partial \theta_i} \frac{\partial \mathcal{L}}{\partial \theta_N} \quad (3.9)$$

895 where θ is a parameter, i is the layer index. We see here that the gradient of the first layer is dependent
 896 of the gradient of all the following layers. We thus need to compute the gradient closest to loss first
 897 before computing the gradient of the earlier layers. This is called the *backward propagation*.

898 This update of the parameters is done following an optimizer policy. Those optimizers depends on
 899 hyperparameters. The ones used in this thesis are:

- 900 1. SGD (Stochastic Gradient Descent). This is the simplest optimizer, it depend on only one
 901 hyperparameter, the learning rate λ (LR) and update the parameters θ following

$$\theta_{t+1} = \theta_t - \lambda \frac{\partial \mathcal{L}}{\partial \theta} \Big|_{\theta_t} \quad (3.10)$$

902 where t is the step index. It is a powerful optimizer but is very sensible to local minima of the
 903 loss in the parameters phase space as illustrated in figure 3.5a.

- 904 2. Adam [54]. The concept is, in short, to have and SGD but with momentum. Adam possess
 905 two momentum $m(\beta_1)$ and $v(\beta_2)$ which are respectively proportional to $\frac{\partial \mathcal{L}}{\partial \theta}$ and $(\frac{\partial \mathcal{L}}{\partial \theta})^2$. β_1
 906 and β_2 are hyperparameters that dictate the moment update at each optimization step. The
 907 parameters are then upgraded following

$$m_{t+1} = \beta_1 m_t + (1 - \beta_1) \frac{\partial \mathcal{L}}{\partial \theta} \quad (3.11)$$

$$v_{t+1} = \beta_2 v_t + (1 - \beta_2) \left(\frac{\partial \mathcal{L}}{\partial \theta} \right)^2 \quad (3.12)$$

$$\theta_{t+1} = \theta_t - \lambda \frac{m_{t+1}}{\sqrt{v_{t+1}} + \epsilon} \quad (3.13)$$

908 where ϵ is a small number to prevent divergence when v is close to 0. These momentums
 909 allow to overcome small local minima in the parameters phase space as illustrated in figure
 910 3.5a.

911 The LR is a crucial parameter in the training of NN, as illustrated in figure 3.6. To prevent possible
 912 issues, we setup scheduler policies.

913 **Scheduler policies**

914 Sometimes we want to update our hyperparameters or take a set of action during the training
 915 procedure. We use for this scheduler policies, for example a common policy is a decrease of the
 916 learning rate after each epochs. The reasoning is that if the learning rate is too high, the optimizer
 917 will continuously miss the minimum and oscillate around it (figure 3.6a). By reducing the learning
 918 rate, we allow it to make more fine steps in the parameters phase space, hopefully converging to the
 919 true minima.

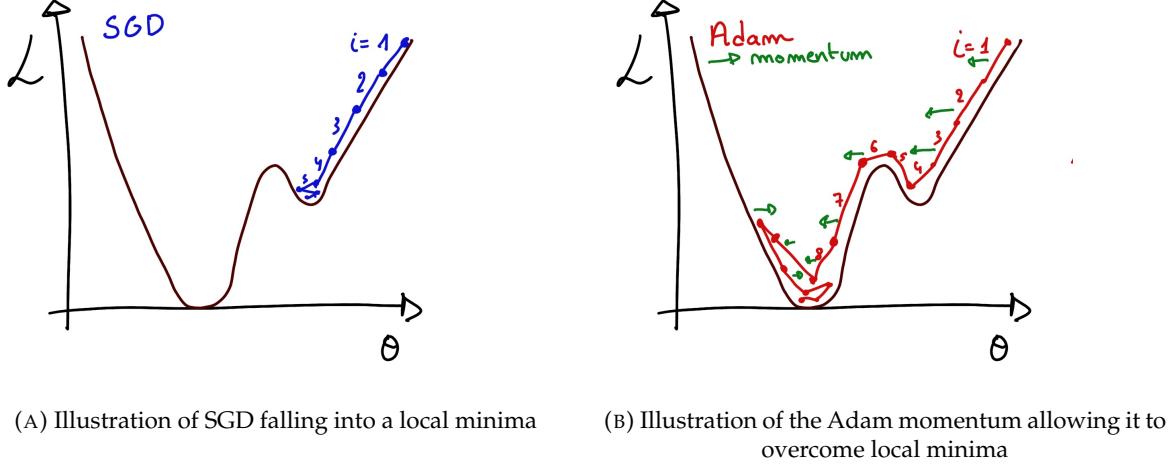


FIGURE 3.5

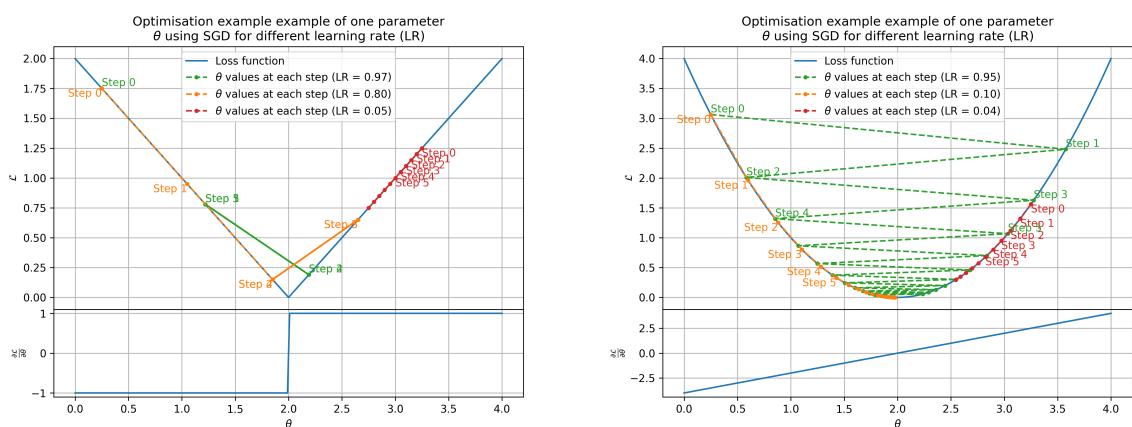
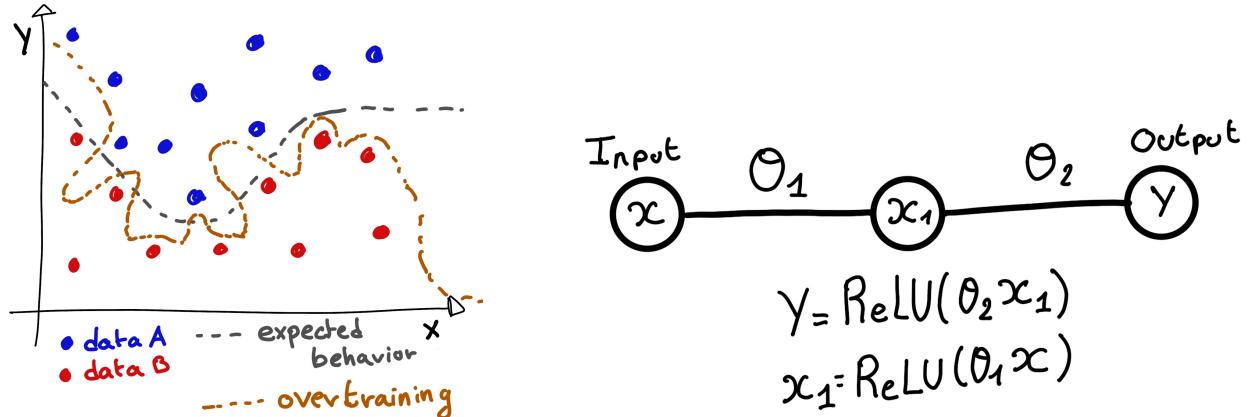
(A) Illustration of the SGD optimizer on one parameter θ on the MAE Loss. We see here that it has trouble reaching the minima due to the gradient being constant.(B) Illustration of the SGD optimizer on one parameter θ on the MAE Loss. We see two different behavior: A smooth one (orange and red) when the LR is small enough and a more chaotic one when the LR is too high.

FIGURE 3.6 – Illustration of the SGD optimizer. In blue is the value of the loss function, orange, green and red are the path taken by the optimized parameter during the training for different LR.



(A) Illustration of overtraining. The task at hand is to determine depending on two input variable x and y if the data belong to the dataset A or the dataset B . The expected boundary between the two dataset is represented in grey. A possible boundary learnt by overtraining is represented in brown.

(B) Illustration of a very simple NN

FIGURE 3.7

916 Another policy that is often used is the save of the best model. In some situations, the loss value after
 917 each epoch will strongly oscillate or even worsen. This policy allows us to keep the best version
 918 of the model attained during the training phase.

919 3.2.6 Potential pitfalls

920 Apart from being stuck in local minima, there are also other behaviors and effects we want to prevent
 921 during training.

922 Overtraining

923 This happens when the network learns the specificities of the training dataset instead of a more general
 924 representation of the underlying data distribution. This can happen if there is not enough data
 925 in comparison to the number of learning parameters, if the data contains some specific signatures
 926 specific to the training dataset or if it trains for too long on the same dataset. This behavior is illustrated
 927 in figure 3.7a. Overtraining can be fought in multiple ways, for example:

- 928 — **More data.** By having more data in the training dataset, the network will not be able to learn the
 929 specificities of every data.
- 930 — **Less parameters.** By reducing the number of parameters, we reduce the computing and
 931 learning capacities of the network. This will force it to fallback to generalist behaviors.
- 932 — **Dropout.** This technique implies to randomly set part of the neural network to 0. By doing
 933 this, we force the redundancy in its computing capability and, in a way, modify the data
 934 decreasing the possibility for specific learning.
- 935 — **Early stopping.** During the training we monitor the network performance over a validation
 936 dataset. The network does not train on this dataset and thus cannot learn its specificities. If
 937 the loss on the training dataset diverges too much from the loss on the validation dataset, we
 938 can stop the training earlier to prevent it from overtraining.

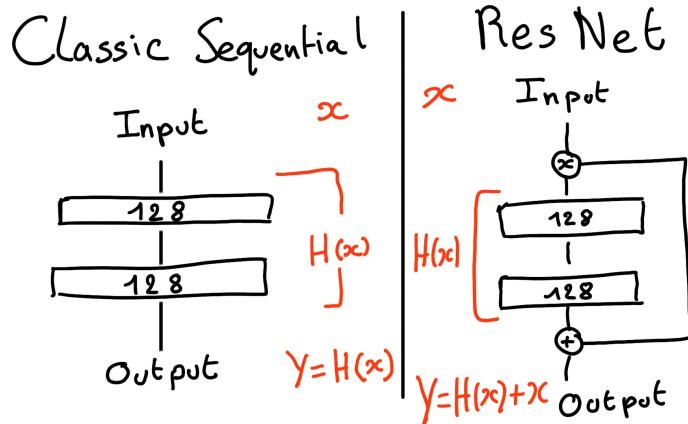


FIGURE 3.8 – Illustration of the ResNet framework

939 **Gradient vanishing**

940 Gradient vanishing is the effect of the gradient being so small for the upper layer that the parameters
 941 are barely updated after each step. This cause the network to be unable to converge to the minima.

942 This comes from the way the gradient descent is calculated. Imagine a simple network composed of
 943 three fully connected layers: the input layer, a intermediate layer and the output layer. Let L be the
 944 loss, θ_1 the parameter between the input and the intermediate layer and θ_2 the parameter between
 945 the intermediate and output layer. This network is schematized in figure 3.7b.

946 The gradient for θ_1 will be computed using the chain rule presented in equation 3.9. Because θ_1
 947 depends on θ_2 , if the gradient of θ_2 is small, so will be the gradient of θ_1 . Now if we would have
 948 much more layer, we can see how the subsequent multiplication of small gradients would lead to
 949 very small update of the parameters thus "vanishing gradient".

950 Multiple actions can be taken to prevent this effect such as:

- 951 — **Batch normalization:** In this case we apply a normalization layer that will normalize the data
 952 so that, let D be the data, $\langle D \rangle = 0$ and $\sigma_D = 1$. This help the weight of the network to
 953 maintain an appropriate scale.
- 954 — **Residual Network (ResNet) [64]:** Residual network is a technique for neural network in
 955 which, instead of just sequentially feeding the results of each layer to the next one, you ask
 956 each layer to calculate the residual of the input data. This technique is illustrated in figure 3.8.

957 **Chapter 4**

958 **Image recognition for IBD
reconstruction with the SPMT system**

959 *Dave - Give me the position and momentum, HAL.*

HAL - I'm afraid I can't do that Dave.

Dave - What's the problem ?

HAL - I think you know what the problem is just as well as I do.

960 *Dave - What are you talking about, HAL?*

HAL - $\sigma_x \sigma_p \geq \frac{\hbar}{2}$

961 As explained in chapter 2, JUNO is an experiment composed of two systems, the Large Photomultiplier (LPMT) and the Small Photomultiplier (SPMT). Both of the system observe the same physics event inside of the same medium but they differ in their photo-coverage, respectively 75.2% and 2.7%, their dynamic range (see section 2.2.2), a thousands versus a few dozen, and their front-end electronics (see section 2.2.2).

966 They are complementary in their strengths and weaknesses and support each other. One important point is their differences in expected resolution, the LPMT system outperform largely the SPMT system but is subject to effects such as charge non linearity [28] that could bias the reconstruction, effect that the SPMT system is impervious to. This topic will be studied in more detail in chapter 7. 967 Also, due to the dynamic range of the LPMT, in case of high energy and high density event such as core-collapse supernova, the LPMT system could saturate and the lower photo-coverage become a benefit.

973 Thus, although event reconstruction algorithm and physics analysis combines both LPMT and SPMT systems, individual approach are key studies to understand the detector and ensure their reliability. 974 This topic will also be studied in more details in chapter 7. The subject of this chapter is to propose 975 a machine learning algorithm for the SPMT reconstruction based on Convolutional Neural Network (CNN). 976 977

978 **4.1 Motivations**

979 As explained in chapter 3, Machine Learning (ML) algorithms shine when modeling highly dimensional data from a given dataset. In our case, we have access to complete monte-carlo simulation of 980 our detector to produce arbitrary large datasets that could represent multiple years of data taking. 981 Ideally ML algorithms would be able to consider the entirety of the information in the detector 982 and converge on the best parameters to yield optimal results, while classical methods where the 983 algorithms could be biased by the prior knowledge of the detector and physics processes. To study 984

985 this potential phenomena, we will compare our machine algorithm to a classical reconstruction
 986 method developed for energy and vertex reconstruction [65].

987 We have access to a very detailed simulation of the detector (section 2.5) that will allow us to simulate
 988 arbitrary large dataset of data while giving access to the all the physics parameters of the event. Those
 989 parameters include the target of our reconstruction algorithms: the vertex and position at with the
 990 event happened. As introduced above, we hope that the ML algorithm will be able to used all the
 991 informations in the event, meaning that potential mismodelings in our simulation could be exploited
 992 by the algorithm. This specific subject will be studied in chapter 6.

993 4.2 Method and model

994 One of simplest way to look at JUNO data is to consider the detector as an array of geometrically
 995 distributed sensors on a sphere. Their repartition is almost homogeneous, on this sphere surface
 996 providing an almost equal amount of information per unit surface on this sphere. It is then tempting
 997 to represent the detector as a spherical image with the PMT in place of pixel. Two events with two
 998 different energy or position would produce two different images.

999 The most common approach in machine learning for image processing and image recognition is the
 1000 Convolutional Neural Network (CNN). It is widely used in research and industry [56, 66–68] due to
 1001 its strengths (see section 3.2.2) and has proven its relevance in image processing.

1002 Some CNN are developed to process spherical images [69] but for the sake of simplicity and as a
 1003 first approach we decided to go with a planar projection of the detector, approach that has proven its
 1004 efficiency using the LPMT system (see section 2.6.3). The details about this planar projection will be
 1005 discussed in section 4.2.2.

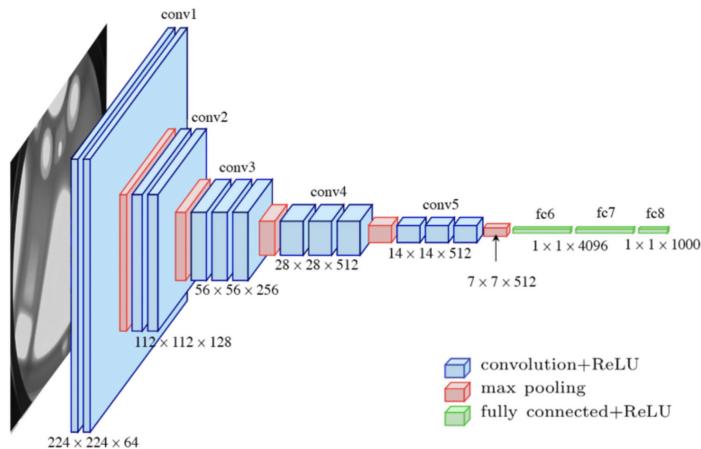


FIGURE 4.1 – Graphic representation of the VGG-16 architecture, presenting the different kind of layer composing the architecture.

1006 4.2.1 Model

1007 The architecture we use is derived from the VGG-16 architecture [56] illustrated in figure 4.1. We
 1008 define a set of hyperparameters that will define the size, complexity and computational power of the
 1009 NN. The chose hyperparameters are detailed below and their values are presented in table 4.1.

- 1010 — N_{blocks} : the number of convolution blocks, a block being composed of two convolutional
1011 layers with 3×3 filters using ReLU activation function, a 3×3 max-pooling layer (except for
1012 the last block) and a dropout layer.
- 1013 — $N_{channels}$: The number of channels in the first block. The number of channels in the subse-
1014 quent blocks are calculated using $N_{channels}^i = 2^i * N_{channels}$, $i \in [1..N_{blocks}]$.
- 1015 — **FCDNN configuration:** The result of the last convolution layer is flattened then fed to a
1016 FCDNN. Its configuration is expressed as a sequence of fully connected linear layer using
1017 the PReLU activation function. For example $2 * 1024 + 2 * 512$ is the sequence of 2 layers
1018 with a width of 1024 followed by 2 other layers with a width of 512. Finally the last layer
1019 is a 4 neurons wide linear layers without activation function. Each neurons of the last layer
1020 represent a component of the interaction vertex: Energy, X, Y, Z.
- 1021 — **Loss:** The loss function. In this work we study two different loss function $(E + V)$ and $(E_r +$
1022 $V_r)$ detailed below.

$$(E + V)(E, x, y, z) = \left\langle (E - E_{true})^2 + 0.85 \sum_{\lambda \in [x, y, z]} (\lambda - \lambda_{true})^2 \right\rangle \quad (4.1)$$

$$(E_r + V_r)(E, x, y, z) = \left\langle \frac{(E - E_{true})^2}{E_{true}} + \frac{10}{R} \sum_{\lambda \in [x, y, z]} (\lambda - \lambda_{true})^2 \right\rangle \quad (4.2)$$

1023 where R is the radius of the CD. With the energy in MeV and the distance in meters, we use the factor
1024 0.85 and 10 to equilibrate the two term of the loss function so they have the same magnitude.

- 1025 — The loss function $(E + V)$ is close to a simple Mean Squared Error (MSE). MSE is one of the
1026 most basic loss function, the derivative is simple and continuous in every point. It is a strong
1027 starting point to explore the possibility of CNNs.
- 1028 — $(E_r + V_r)$ can be see as a relative MSE.

1029 The idea is that: due to the inherent statistic uncertainty over the number of collected Number of
1030 Photo Electrons (NPE), the absolute resolution $\sigma(E - E_{true})$ will be larger at higher energy than at
1031 low energy. But we expect the *relative* energy resolution $\frac{\sigma(E - E_{true})}{E_{true}}$ to be smaller at high energy than
1032 lower energy as illustrated in figure 2.19. Because of this, by using simple MSE the most important
1033 part in the loss come from the high energy part of the dataset whereas with a relative MSE, the most
1034 important become the low energy events in the dataset. We hope that by using a relative MSE, the
1035 neural network will focus on low energy events where the reconstruction is considered the hardest
1036 part of the dataset.

1037 On top those generated models, we define 4 hand tailored models:

- 1038 — “gen_0”: $N_{blocks} = 4$, $N_{channels} = 64$, FCDNN configuration: $1024 * 2 + 512 * 2$, Loss := $E + V$
- 1039 — “gen_1”: $N_{blocks} = 4$, $N_{channels} = 64$, FCDNN configuration: $1024 * 2 + 512 * 2$, Loss := $E_r + V_r$
- 1040 — “gen_2”: $N_{blocks} = 5$, $N_{channels} = 64$, FCDNN configuration: $4096 * 2 + 1024 * 2$, Loss := $E + V$
- 1041 — “gen_3”: $N_{blocks} = 5$, $N_{channels} = 64$, FCDNN configuration: $4096 * 2 + 1024 * 2$, Loss := $E_r + V_r$

N_{blocks}	{2, 3, 4}
$N_{channels}$	{32, 64, 128}
FCDNN configurations	$2 * 1024$ $2 * 2048 + 2 * 1024$ $3 * 2048 + 3 * 512$ $2 * 4096$
Loss	{ $E + V$, $E_r + V_r$ }

TABLE 4.1 – Sets of hyperparameters values considered in this study

1042 Each combination of those hyperparameters (for example ($N_{blocks} = 2$, $N_{channels} = 32$, FCDNN =
1043 $(2 * 1024)$, Loss = $(E + V)$)), subsequently designated as configurations, is then tested and compared

to each other over an analysis sample. We cannot use the mean loss because we consider multiple loss functions, there is no guarantee that comparison of their numerical value will be meaningful. We use multiple observables to rank the performances of each configuration:

- The mean absolute energy error $\langle E \rangle = \langle |E - E_{true}| \rangle$. It is an indicator of the energy bias of our reconstruction.
- The standard deviation of the energy error $\sigma E = \sigma(E - E_{true})$. This the indicator on our precision in energy reconstruction.
- The mean distance between the reconstructed vertex and the true vertex $\langle V \rangle = \langle |\vec{V} - \vec{V}_{true}| \rangle$. This an indicator of the bias and precision of our vertex reconstruction.
- The standard deviation of the distance between the true and reconstructed vertex $\sigma V = \sigma|\vec{V} - \vec{V}_{true}|$. This is an indicator if the precision in our vertex reconstruction.

4.2.2 Data representation

This data is represented as 240×240 images, equivalent to third order tensor, with a charge Q channel and a time t channel. The SPMTs are then projected on the plane as illustrated in figure 4.2. The x position is proportional to θ and the y position is defined by $\phi \sin \theta$ in spherical coordinates. $\theta = 0$ is defined as being the top of the detector and $\phi = 0$ is defined as an arbitrary direction in the detector. In practice, this is the $\phi = 0$ given by the MC simulation.

$$x = \left\lfloor \frac{\theta \cdot H}{\pi} \right\rfloor, \theta \in [0, \pi] \quad (4.3)$$

$$y = \left\lfloor \frac{(\phi + \pi) \sin \theta \cdot W}{2\pi} \right\rfloor, \phi \in [-\pi, \pi], \theta \in [0, \pi] \quad (4.4)$$

where H is the height of the image, W the width of the image and $(0, 0)$ the top left corner of the image.

When two SPMTs are in the same pixel, the charges are summed and the lowest of the hit-time is chosen. The SPMTs being located close to each other, we expect the time difference between two successive physics signals, two photons being collected, to be small. The first hit time is chosen because it can be considered as the relative propagation time of the photons that went the "straightest", i.e. that went under the less perturbation of the two. The only potential problem in using this first time come from the Dark Noise (DN). Its time distribution is uniform over the signal and could come before a signal hit on the other SPMT in the pixel. In that case, the time information in the pixel become irrelevant and we lose the timing information for this part of the detector. As illustrated in figure 4.2 the dimension have been chosen optimized so that at most two SPMTs are in the same pixel while keeping the number of empty pixels relatively low to prevent this kind of issue.

While it could be possible to use larger images (more pixel) to prevent overlapping, keeping image small images gives multiple advantages:

- As presented in section 4.2.1, the convolution filter we use are 3×3 convolution filter, meaning that if SPMTs would be separated by more than one pixel, the first filter would only see one SPMT per filter. This behavior would be kind of counterproductive as the first convolution block would basically be a transmission layer and would just induce noise in the data.
- It keep the network relatively small, while this do not impact the convolution layers, the flatten operation just before the FCDNN make the number parameters in the first layer of it dependent on the size of the image.
- It reduce the number of empty pixel in the image.

The question of empty pixel is an important question in this data representation. Their is two kind of empty pixel in the data.

1085 The first kind is pixel that contain a SPMT but the SPMT did not get hit nor registered any dark noise
 1086 during the event. In this case, the charge channel is zero, which have a physical meaning but then
 1087 come the question of the time layer. One could argue that the correct time would be infinity (or the
 1088 largest number our memory allows us) because the hit “never” happened, so extremely far from the
 1089 time of the event. This cause numerical problem as large number, in the linear operation that are
 1090 happening in the convolution layers, are more signifiant than smaller value. We could try to encode
 1091 this feature in another way but no number have any significance due to our time being relative to
 1092 the trigger of the experiment so -1 for example is out of question. Float and Double gives us access
 1093 to special value such as NaN (Not a Number) [70] but the behavior is to propagate the NaN which
 1094 leaves us with NaN for energy and position. We choose to keep the value 0 because it’s the absorbing
 1095 element of multiplication, absorbing the “information” of the parameter it would be multiplied by.
 1096 It also can be though as no activation in the ReLU activation function.

1097 The second kind of pixel is pixel that do not represent parts of the detector such as the corners of
 1098 the images. The question is basically the same, what to put in the charge and the time channel. The
 1099 decision is to set the charge and time at 0 following the reasoning presented above. Its important
 1100 to keep in mind that the fact that a part of the detector that has not been hit is also an information:
 1101 There is no signal in this part of the detector. This problematic will be explored in more details in
 1102 chapter 5.

1103 Another problematic that happens with this representation, and this is not dependent of the chosen
 1104 projection, is the deformation in the edges of the image and the loss of the neighbouring information
 1105 in the for the SPMTs at the edge of the image $\phi \sim 180^\circ$. This deformation and neighbouring loss
 1106 could be partially circumvented as explained in section 4.4

1107 4.2.3 Dataset

1108 In this study we will discuss two datasets of one millions events:

- 1109 — **J21**: The first one comes from the JUNO official mc simulation J21v1r0-Pre2 (released the
 1110 18th August 2021). This historical version is the one on which the NN the classical algorithm
 1111 presented in [65] was developped. This dataset is used as a reference for comparison to
 1112 classical algorithm. The data in this dataset is *detsim* level (see section 2.5), so where only
 1113 the physic is simulated. The charge and time biases and uncertainties are added using toy
 1114 MC adjusted using [25, 71].
- 1115 — **J23**: The second comes from the JUNO official monte-carlo simulations J23.0.1-rc8.dc1 (re-
 1116 leased the 7th January 2024). The data is *calib* level (see section 2.5). Here the charge waveform
 1117 integration, time window resolution and trigger decision are all simulated inside the software.
 1118 This dataset is more realistic and is used to confirm the performance of our algorithm.

1119 To put in perspective this amount of data, the expected IBD rate in JUNO is 47 / days. Taking into
 1120 account the calibration time, and the source reactor shutdown, it amount to $\sim 94'000$ IBD events
 1121 in 6 years. With this million of event, we are training the equivalent of ~ 10 years of data. With
 1122 this amount we reach a density of $4783 \frac{\text{event}}{\text{m}^3 \cdot \text{MeV}}$, meaning our dataset is representative of the multiple
 1123 event scenarios that could be happening in the detector.

1124 While we expect and hope the monte-carlo simulation to give use a realistic representation of the
 1125 detector, there could be effect, even after the fine-tuning on calibration data, that the simulation
 1126 cannot handle. Thus, once the calibration will be available, we will need to evaluate, and if needed
 1127 retrain, the network on calibration data to establish definitive performances.

1128 The data used during this analysis is monte carlo data using the official JUNO simulation software
 1129 (see section 2.5 for details). The simulated data is composed of positron events, uniformly distributed
 1130 in the CD volume and in kinetic energy over $E_k \in [0; 9]$ MeV producing a deposited energy $E_{dep} \in$
 1131 $[1.022; 10.022]$ MeV. This is done to mimic the signal produced by the IBD prompt signal. Uniform

1132 distribution are used so that the CNN does not learn a potential energy distribution, favoring some
1133 part of the energy spectrum instead of other.

1134 Those events can be considered as “optimistic” as there is no pile-up with potential background or
1135 other IBD.

1136 **4.2.4 Data characteristics**

1137 To delve a bit into the kind of data we will use, you can find in figure 4.2 the repartition of the SPMTs
1138 in the image. The color represent the number of SPMTs per pixel.

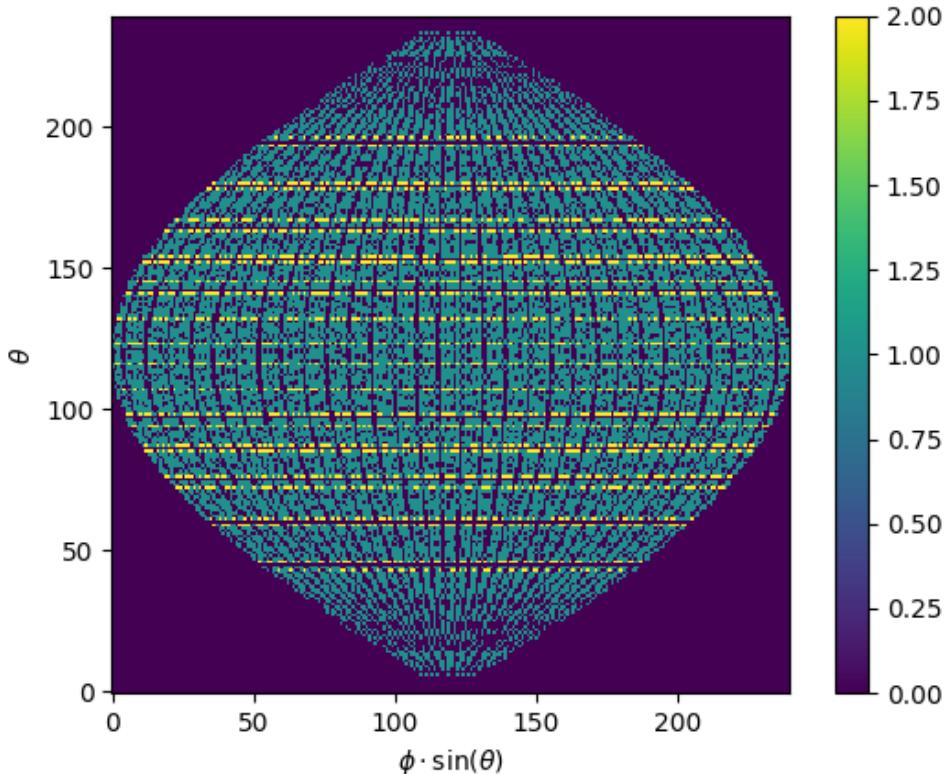


FIGURE 4.2 – Repartition of SPMTs in the image projection. The color scale is the number of SPMTs per pixel

1139 In figures 4.3, 4.4, 4.5 and 4.6 are presented events from J23 for different positions and energies.
1140 We see some characteristics and we can instinctively understand how the CNN could discriminate
1141 different situations.

To give an idea of the strength of the signal in comparison to the dark noise background, figure 4.7a present the distribution of the ration NPE over the deposited energy. Assuming a linear response of the LS we can model:

$$NPE_{tot} = E_{dep} \cdot P_{mev} + D_N \quad (4.5)$$

$$\frac{NPE_{tot}}{E_{dep}} = P_{mev} + \frac{D_N}{E_{dep}} \quad (4.6)$$

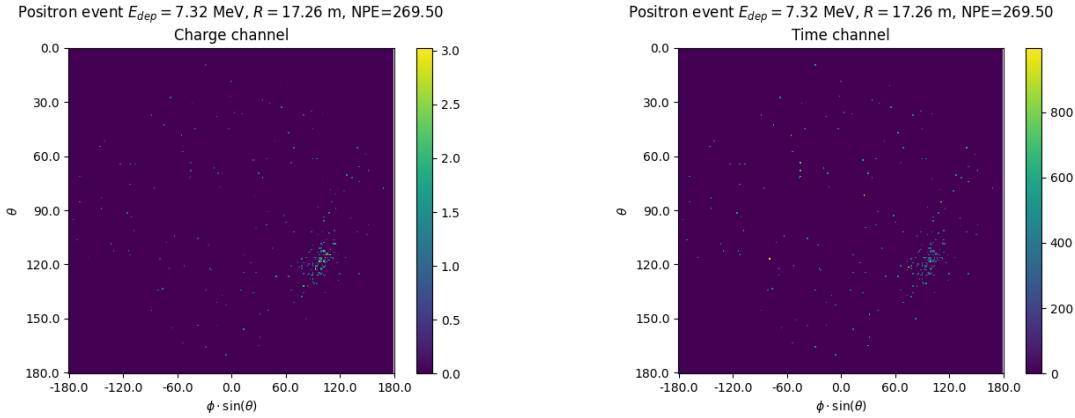


FIGURE 4.3 – Example of a high energy, radial event. We see a concentration of the charge on the bottom right of the image, clear indication of a high radius event. **On the left:** the charge channel. The color is the charge in each pixel in NPE equivalent. **On the right:** The time channel in nanoseconds.

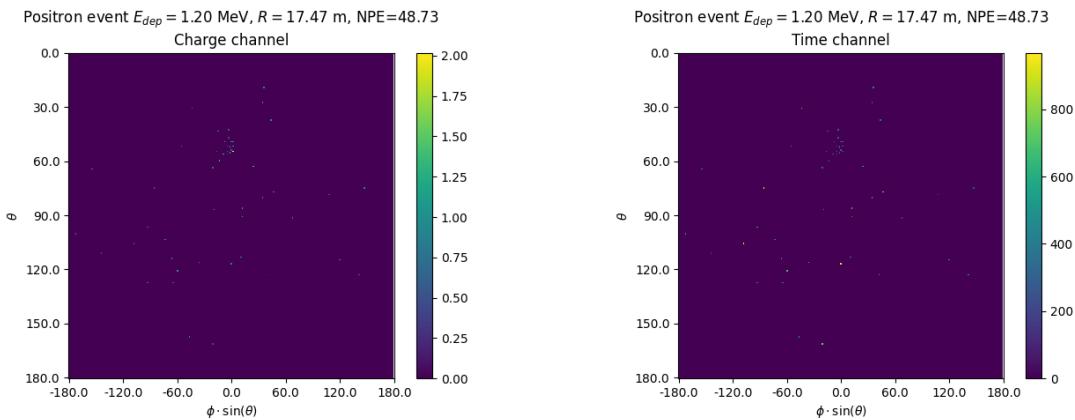


FIGURE 4.4 – Example of a low energy, radial event. The signal here is way less explicit, we can kind of guess that the event is located in the top middle of the image. **On the left:** the charge channel. The color is the charge in each pixel in NPE equivalent. **On the right:** The time channel in nanoseconds.

1142 where NPE_{tot} is the total number of PE detected by the event, P_{mev} is the mean number of PE detected
 1143 per MeV and D_N is the dark noise contribution that is considered energy independent. In the case
 1144 where the readout time window is dependent of the energy the dark noise contribution become
 1145 energy dependant, also the LS response is realistically energy dependant but figure 4.7a shows that
 1146 we have heavily dominated by statistical uncertainties which is why we are using this simple model.

1147 The fit gives shows a light yield of 40.78 PE/MeV and a dark noise contribution of 4.29 NPE. As
 1148 shown in figure 4.7b, the physics makes for 90% of the signal at low energy.

1149 4.3 Results

1150 Before presenting the results, let's discuss the different observables.

1151 The events are considered point like in this study. The target truth position, or vertex, is the mean po-

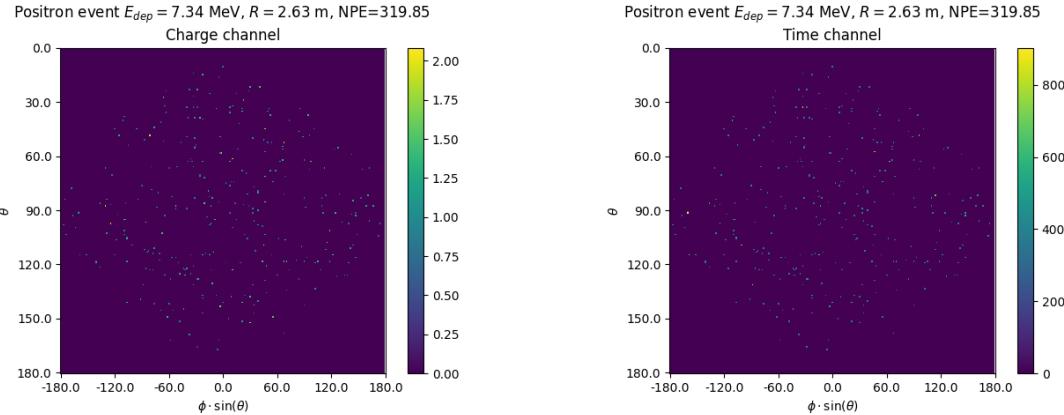


FIGURE 4.5 – Example of a high energy, central event. In this image we can see a lot of signal but uniformly spread, this is indicative of a central event. **On the left:** the charge channel. The color is the charge in each pixel in NPE equivalent. **On the right:** The time channel in nanoseconds.

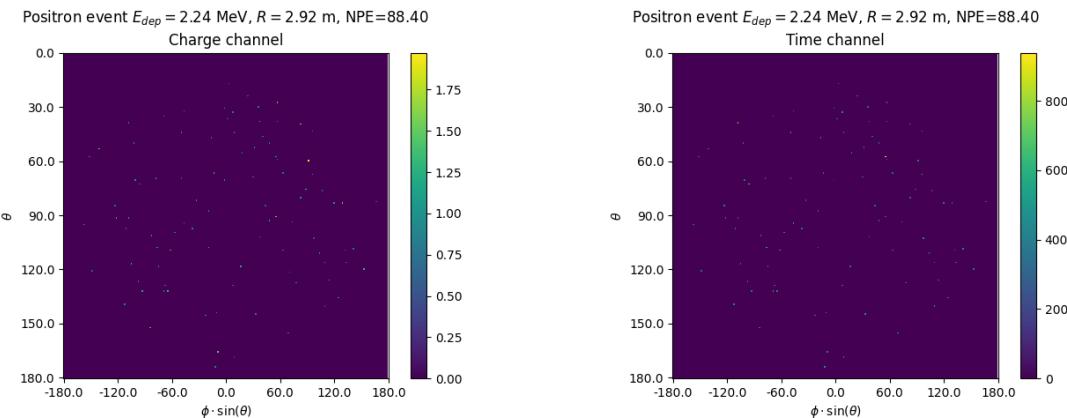
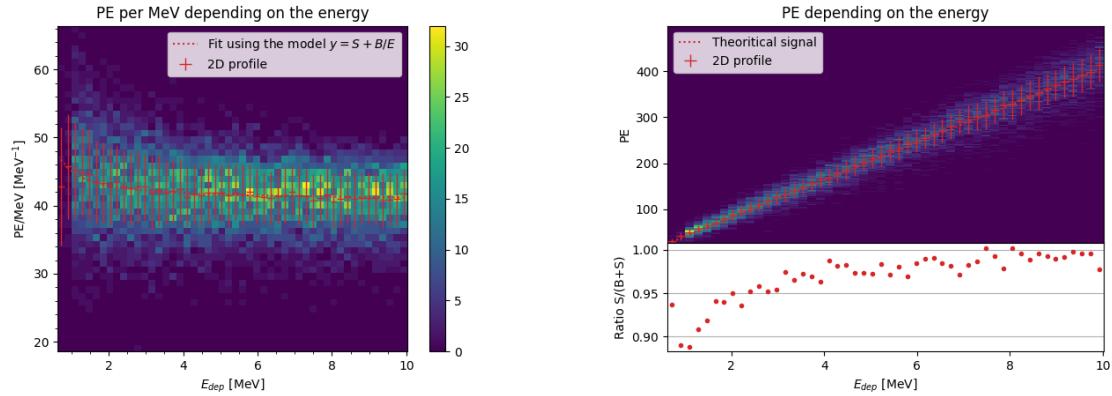


FIGURE 4.6 – Example of a low energy, central event. Here there is no clear signal, the uniformity of the distribution should make it central. **On the left:** the charge channel. The color is the charge in each pixel in NPE equivalent. **On the right:** The time channel in nanoseconds.

1152 sition of the energy deposits of the positron and the two annihilation gammas. Due to the symmetries
 1153 of the detector, we mainly considered and discuss the bias and precision evolution depending of the
 1154 radius R but we will still monitor the performances depending of the spheric angle θ and ϕ . From the
 1155 detector construction and effect we expect relative important dependencies in radius thanks to the
 1156 TR area effect presented in section 2.6 and the possibility for the positron or the gammas to escape
 1157 from the CD for near the edge events. We also expect dependence in θ , the top of the experiment
 1158 being non-instrumented due to the filling chimney. It is also to be noted that the events in the dataset
 1159 are uniformly distributed in the CD, and so are uniformly distributed in R^3 and ϕ . The θ distribution
 1160 is not uniform and we will have more event for $\theta \sim 90^\circ$ than $\theta \sim 0^\circ$ or $\theta \sim 180^\circ$.

1161 We define multiple energy in JUNO:

- 1162 — E_ν : The energy of the neutrino.
- 1163 — E_k : The kinetic energy of the resulting positron from the IBD.
- 1164 — E_{dep} : The deposited energy of the positron and the two annihilation gammas.
- 1165 — E_{vis} : The equivalent visible energy, so E_{dep} after the detector effect such as the absorption of
 scintillation photons by the LS and the LS response non-linearity.



(A) Distribution of PE/MeV in the J23 Dataset. This distribution is profiled and fitted using equation 4.6

(B) On top: Distribution of PE vs Energy. On bottom: Using the values extracted in 4.7a, we calculate the ration signal over background + signal

FIGURE 4.7

— E_{rec} : The reconstructed energy by the reconstruction algorithm. The expected value depend on the algorithm we discuss about. For example the algorithm presented in section 2.6 is reconstructing E_{rec} while the ones presented in section 2.6.3 reconstruct E_{dep} .
 In this study, we will set E_{rec} as our target for energy reconstruction. This choice is motivated by the ease with which we can retrieve this information in the monte-carlo data while E_{vis} is less trivial to retrieve.

4.3.1 J21 results

Those results comes from the “gen_30” model, meaning then 30th model generated using the table 4.1 or
 — “gen_30”: $N_{blocks} = 3$, $N_{channels} = 32$, FCDNN configuration: $2048 * 2 + 1024 * 2$, Loss := $E + V$
 The performances of its reconstruction are presented in blue in figure 4.8. Superimposed in black is the performances of the classical algorithm from [65].

Energy reconstruction

By looking at the figure 4.8a and 4.8b, the CNN has similar performances in its energy resolution with the exception of high and low energy event.

This is explained by looking at the true and reconstructed energy distributions in figure 4.10. We see that the distributions are similar for energies between 3 and 8 MeV but there is an excess of reconstructed energies around 1.5 and 9 MeV while a lack of reconstructed energy for 1 and 10 MeV. The neural network seems to learn the energy distribution and learn that it exist almost no event with an energy inferior to 1.022 MeV and not event with an energy superior to 10 MeV.

The first observation is a physics phenomena: for a positron, its minimum deposited energy is the mass energy coming from its annihilation with an electron 1.022 MeV. There is a few event with energies inferior to 1.022 MeV, in those case the annihilation gammas or even the positron escape the detector. The deposited energy in the LS is thus only a fraction of the energy is deposited.

The second observation is indeed true in this dataset but has no physical meaning, it is an arbitrary limit because the physics region of interest is mainly between 1 and 9 MeV of deposited energy

(figure 2.2). By learning the energy distribution, the CNN pull event from the border of it to more central value. That's why the energy resolution is better: the events are pulled in a small energy region , thus a small variance but the bias become very high (figure 4.8a).

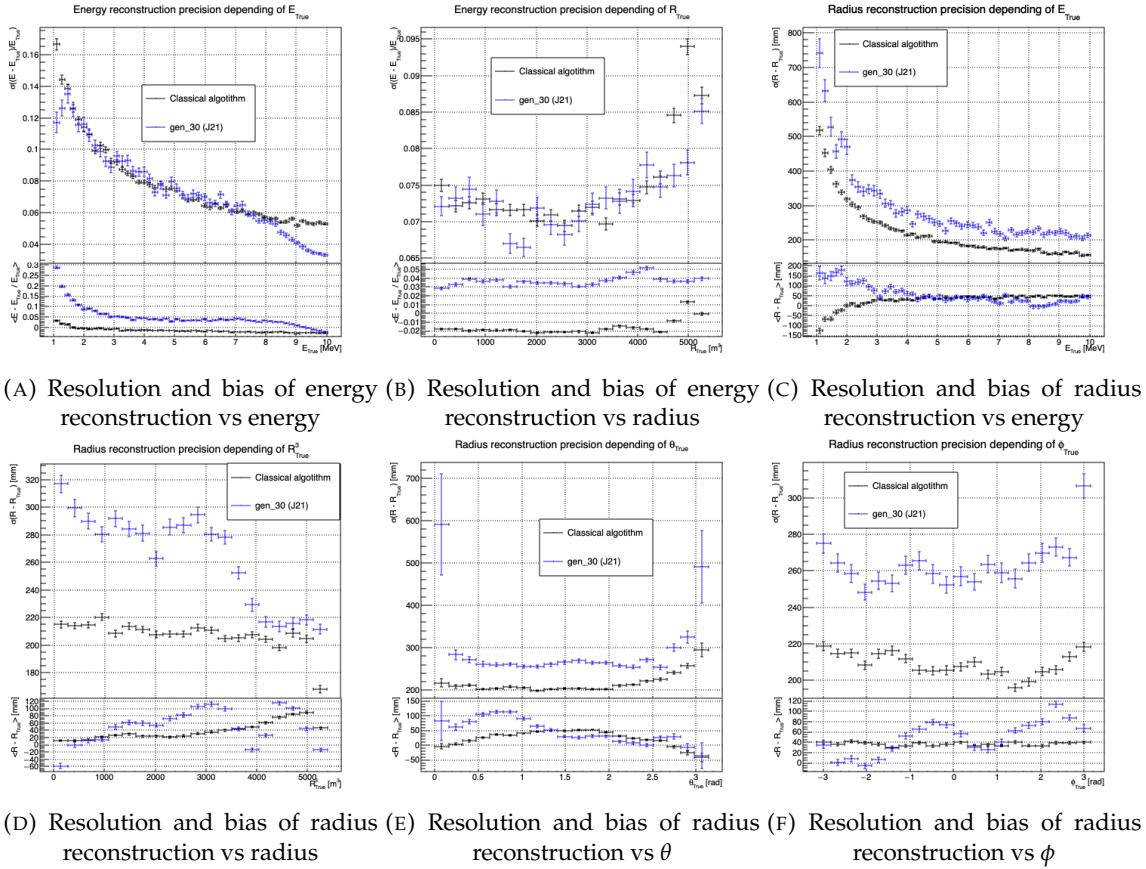


FIGURE 4.8 – Reconstruction performance of the “gen_30” model on J21 data and it’s comparison to the performances of the classic algorithm “Classical algorithm” from [65]. The top part of each plot is the resolution and the bottom part is the bias.

This behavior also explain the heavy bias at low energy in figure 4.8a. The energy bias of the CNN if fairly constant over the energy range, it is interesting to note that the energy bias depending on the radius is equivalent but of opposite sign of the classical algorithm (figure 4.8b). This behavior is studied in more detail in section 4.3.2.

1200 Vertex reconstruction

For the vertex reconstruction we do not study x , y and z independently but we use R as a proxy observable. Figure 4.9 shows the error distribution of the different vertex coordinates. We see that R errors and biases are slightly superior to the cartesian coordinates, thus R is a conservative proxy observable to discuss the subject of vertex reconstruction.

The comparison of radius reconstruction between the classical algorithm and “gen_30” are presented in the figures 4.8c, 4.8d, 4.8e and 4.8f.

Radius reconstruction is worse than the classical algorithms in all configuration, in energy, figure 4.8c, where we see a degradation of almost 25cm at low energy to 4cm at higher energy.

When looking over the event radius, figure 4.8d, we lose between 4 and 10cm of resolution in the central part of the detector $R < 15.2$ m and have a sudden spike of performance where we almost catch up to the classical method for more radial event $R > 15$ m. This is around this threshold that we start to see the effect of the TR area (see section 2.6). In this region, the signal become more localised as shown in figure 4.3 versus figure 4.5 and the reflection can also present pattern dependant on the radius. This locality and specificity of the information must be helping the CNN to localize the event with more precision.

The precision also worsen when looking at the edge of the image $\theta \approx 0, \theta \approx 2\pi$ respectively the top and bottom of the image, and when $\phi \approx -\pi$ and $\phi \approx \pi$ respectively the left and right side of the image. This is the confirmation that the deformation of the image is problematic for the event reconstruction.

The bias in radius reconstruction is about the same order of magnitude depending of the energy but is of opposite sign. As for the energy, this behavior is studied in more details in section 4.3.2. Over radius, θ and ϕ the bias is inconsistent, sometimes event better than the classical reconstruction and but also can much worse than the classical method. This could come from the specialisation of some filters in the convolutional layers for specific part of the detector that would still work “correctly” for close part but with much less precision.

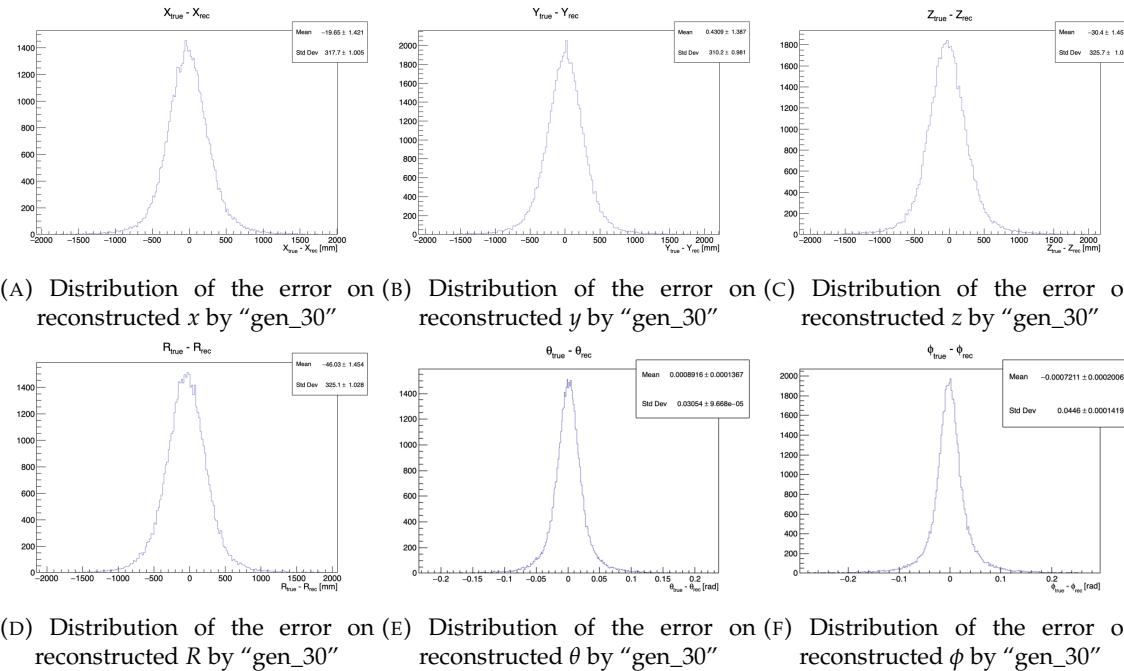


FIGURE 4.9 – Error distribution of the different component of the vertex by “gen_30”.
The reconstructed component are x, y and z but we see similar behavior in the error of R, θ and ϕ .

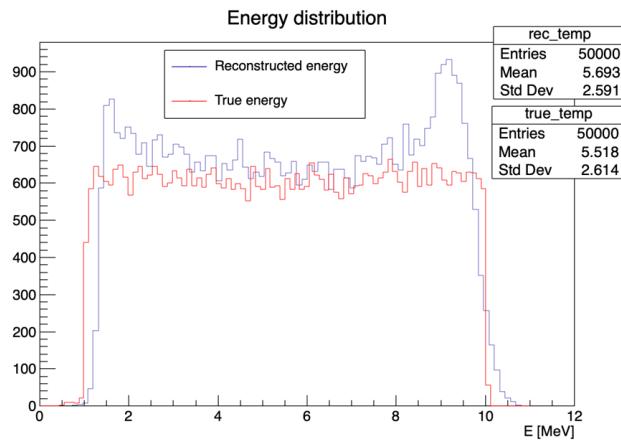


FIGURE 4.10 – Distribution of “gen_30” reconstructed energy and true energy of the analysis dataset (J21)

1226 **4.3.2 Combination of classic and ML estimator**

1227 **4.4 Prospect**

1228 **4.5 Conclusion**

1229 Intoduction next chapter

1230 **Chapter 5**

1231 **Graph representation of JUNO for IBD
reconstruction with the LPMT system**

¹²³³ **Chapter 6**

¹²³⁴ **Reliability of machine learning
methods**

¹²³⁵

"Psychohistory was the quintessence of sociology; it was the science of human behavior reduced to mathematical equations. The individual human being is unpredictable, but the reactions of human mobs, Seldon found, could be treated statistically"

Isaac Asimov, Second Foundation

¹²³⁶

¹²³⁷ **Chapter 7**

¹²³⁸ **Joint fit between the SPMT and LPMT
spectra**

¹²³⁹

1240 Chapter 8

1241 Conclusion

1242 List of Tables

1243 2.1	Characteristics of the nuclear power plants observed by JUNO. The IBD rate are esti- 1244 mated from the baselines, the reactors full thermal power, selection efficiency and the 1245 current knowledge of the oscillation parameters	11
1246 2.2	A summary of precision levels fir the oscillation parameters. The reference value (PDG 1247 2020 [16]) is compared with 100 days, 6 years and 20 years of JUNO data taking.	13
1248 2.3	Detectable neutrino signal in JUNO and the expected signal rates and major back- 1249 ground sources	13
1250 2.4	List of sources and their process considered for the energy scale calibration	20
1251 2.5	Calibration program of the JUNO experiment	22
1252 2.6	Features used by the BDT for vertex reconstruction	32
1253 2.7	Features used by the BDTE algorithm. <i>pe</i> and <i>ht</i> reference the charge and hit-time 1254 distribution respectively and the percentages are the quantiles of those distributions. 1255 <i>cht</i> and <i>cc</i> reference the barycenters of hit time and charge respectively	32
1256 4.1	Sets of hyperparameters values considered in this study	49

1257 List of Figures

1258 2.1	On the left: Location of the JUNO experiment and its reactor sources in southern china. On the right: Aerial view of the experimental site	9
1259 2.2	Expected number of neutrinos event per MeV in JUNO after 6 years of data taking. The black curve shows the flux if there was no oscillation. The light gray curve shows the oscillation if only the solar terms are taken in account (θ_{12} , Δm_{21}^2). The blue and red curve shows the spectrum in the case of, respectively, NO and IO. The dependency of the oscillation to the different parameters are schematized by the double sided arrows. We can see the NMO sensitivity by looking at the fine phase shift between the red and the blue curve.	10
1260 2.3	Expected visible energy spectrum measured with the LPMT system with (grey) and without (black) backgrounds. The background amount for about 7% of the IBD candidate and are mostly localized below 3 MeV [11]	12
1261 2.4	15
1262 a	Schematics view of the JUNO detector.	15
1263 b	Top down view of the JUNO detector under construction	15
1264 2.5	Schematics of an IBD interaction in the central detector of JUNO	16
1265 2.6	Schematics of the supporting node for the acrylic vessel	16
1266 2.7	On the left: Quantum efficiency (QE) and emission spectrum of the LAB and the bis-MSB [19]. On the right: Sensitivity of the Hamamatsu LPMT depending on the wavelength of the incident photons [21].	17
1267 2.8	Schematic of a PMT	18
1268 2.9	The JUNO top tracker	19
1269 2.10	Fitted and simulated non linearity of gamma, electron sources and from the ^{12}B spectrum. Black points are simulated data. Red curves are the best fits	21
1270 a	Gamma non-linearity	21
1271 b	Boron non-spectrum	21
1272 c	Electron non-linearity	21
1273 2.11	Overview of the calibration system	21
1274 2.12	23
1275 a	Schematic of the TAO satellite detector	23
1276 b	Schematic of the OSIRIS satellite detector	23
1277 2.13	25
1278 a	Illustration of the different optical photons reflection scenarios. 1 is the reflection of the photon at the interface LS-acrylic or acrylic-water. 2 is the transmission of the photons through the interfaces. 3 is the conduction of the photon in the acrylic.	25
1279 b	Heatmap of R_{rec} and $R_{\text{rec}} - R_{\text{true}}$ as a function of R_{true} for 4MeV prompt signals uniformly distributed in the detector calculated by the charge based algorithm	25
1280 2.14	26
1281 a	Δt distribution at different iterations step j	26
1282 b	Heatmap of R_{rec} and $R_{\text{rec}} - R_{\text{true}}$ as a function of R_{true} for 4MeV prompt signals uniformly distributed in the detector calculated by the time based algorithm	26

1300	2.15 Bias of the reconstructed radius R (left), θ (middle) and ϕ (right) for multiple energies by the time likelihood algorithm	27
1301		
1302	2.16 On the left: Resolution of the reconstructed R as a function of the energy in the TR area ($R^3 > 4000\text{m}^3 \equiv R > 16\text{m}$) by the charge and time likelihood algorithms. On the right: Bias of the reconstructed R in the TR area for different energies by the charge likelihood algorithm	28
1303		
1304	2.17 Radial resolution of the different vertex reconstruction algorithms as a function of the energy	28
1305		
1306	2.18	29
1307	a Spherical coordinate system used in JUNO for reconstruction	29
1308	b Definition of the variables used in the energy reconstruction	29
1309		
1310	2.19	31
1311	a Radial resolutions of the likelihood-based algorithm TMLE, QMLE and QTMLE	31
1312	b Energy resolution of QMLE and QTMLE using different vertex resolutions	31
1313		
1314	2.20 Projection of the LPMTs in JUNO on a 2D plane. (a) Show the distribution of all PMTs and (b) and (c) are example of what the charge and time channel looks like respectively	33
1315		
1316	2.21 Radial (left) and energy (right) resolutions of different ML algorithms. The results presented here are from [41]. DNN is a deep neural network, BDT is a BDT, ResNet-J and VGG-J are CNN and GNN-J is a GNN.	33
1317		
1318		
1319	3.1 Example of a BDT that determine if the given object is a duck	38
1320	3.2	40
1321	a Schema of a FCDNN	40
1322	b Illustration of a composition of ReLU matching a function	40
1323	3.3 Illustration of the effect of a convolution filter. Here we apply a filter with the aim do detect left edges. We see in the resulting image that the left edges of the duck are bright yellow where the right edges are dark blue indicating the contour of the object. The convolution was calculated using [57].	40
1324		
1325		
1326	3.4	41
1327	a Example of images in the MNIST dataset	41
1328	b Schema of the CNN used in Pytorch example to process the MNIST dataset	41
1329		
1330	3.5	44
1331	a Illustration of SGD falling into a local minima	44
1332	b Illustration of the Adam momentum allowing it to overcome local minima	44
1333	3.6 Illustration of the SGD optimizer. In blue is the value of the loss function, orange, green and red are the path taken by the optimized parameter during the training for different LR.	44
1334		
1335	a Illustration of the SGD optimizer on one parameter θ on the MAE Loss. We see here that it has trouble reaching the minima due to the gradient being constant.	44
1336	b Illustration of the SGD optimizer on one parameter θ on the MAE Loss. We see two different behavior: A smooth one (orange and red) when the LR is small enough and a more chaotic one when the LR is too high.	44
1337		
1338	3.7	45
1339	a Illustration of overtraining. The task at hand is to determine depending on two input variable x and y if the data belong to the dataset A or the dataset B . The expected boundary between the two dataset is represented in grey. A possible boundary learnt by overtraining is represented in brown.	45
1340	b Illustration of a very simple NN	45
1341	3.8 Illustration of the ResNet framework	46
1342		
1343		
1344		
1345		
1346		
1347		
1348	4.1 Graphic representation of the VGG-16 architecture, presenting the different kind of layer composing the architecture.	48
1349		
1350	4.2 Repartition of SPMTs in the image projection. The color scale is the number of SPMTs per pixel	52
1351		

1352	4.3	Example of a high energy, radial event. We see a concentration of the charge on the bottom right of the image, clear indication of a high radius event. On the left: the charge channel. The color is the charge in each pixel in NPE equivalent. On the right: The time channel in nanoseconds.	53
1353			
1354			
1355			
1356	4.4	Example of a low energy, radial event. The signal here is way less explicit, we can kind of guess that the event is located in the top middle of the image. On the left: the charge channel. The color is the charge in each pixel in NPE equivalent. On the right: The time channel in nanoseconds.	53
1357			
1358			
1359			
1360	4.5	Example of a high energy, central event. In this image we can see a lot of signal but uniformly spread, this is indicative of a central event. On the left: the charge channel. The color is the charge in each pixel in NPE equivalent. On the right: The time channel in nanoseconds.	53
1361			
1362			
1363			
1364	4.6	Example of a low energy, central event. Here there is no clear signal, the uniformity of the distribution should make it central. On the left: the charge channel. The color is the charge in each pixel in NPE equivalent. On the right: The time channel in nanoseconds.	54
1365			
1366			
1367			
1368	4.7	54
1369	a	Distribution of PE/MeV in the J23 Dataset. This distribution is profiled and fitted using equation 4.6	55
1370	b	On top: Distribution of PE vs Energy. On bottom: Using the values extracted in 4.7a, we calculate the ration signal over background + signal	55
1371			
1372			
1373	4.8	Reconstruction performance of the “gen_30” model on J21 data and it’s comparison to the performances of the classic algorithm “Classical algorithm” from [65]. The top part of each plot is the resolution and the bottom part is the bias.	55
1374	a	Resolution and bias of energy reconstruction vs energy	56
1375	b	Resolution and bias of energy reconstruction vs radius	56
1376	c	Resolution and bias of radius reconstruction vs energy	56
1377	d	Resolution and bias of radius reconstruction vs radius	56
1378	e	Resolution and bias of radius reconstruction vs θ	56
1379	f	Resolution and bias of radius reconstruction vs ϕ	56
1380			
1381			
1382	4.9	Error distribution of the different component of the vertex by “gen_30”. The reconstructed component are x , y and z but we see similar behavior in the error of R , θ and ϕ	57
1383	a	Distribution of the error on reconstructed x by “gen_30”	57
1384	b	Distribution of the error on reconstructed y by “gen_30”	57
1385	c	Distribution of the error on reconstructed z by “gen_30”	57
1386	d	Distribution of the error on reconstructed R by “gen_30”	57
1387	e	Distribution of the error on reconstructed θ by “gen_30”	57
1388	f	Distribution of the error on reconstructed ϕ by “gen_30”	57
1389			
1390			
1391	4.10	Distribution of “gen_30” reconstructed energy and true energy of the analysis dataset (J21)	58
1392			

¹³⁹³ List of Abbreviations

ACU	Automatic Calibration Unit
BDT	Boosted Decision Tree
CD	Central Detector
CLS	Cable Loop System
CNN	Convolutional NN
DNN	Deep NN
DN	Dark Noise
FCDNN	Fully Connected Deep NN
GNN	Graph NN
GT	Guiding Tube
IBD	Inverse Beta Decay
IO	Inverse Ordering
JUNO	Jiangmen Underground Neutrino Observatory
LPMT	Large PMT
LR	Learning Rate
LS	Liquid Scintillator
MC	Monte Carlo simulation
ML	Machine Learning
MSE	Mean Squared Error
NMO	Neutrino Mass Ordering
NN	Neural Network
NO	Normal Ordering
NPE	Number of Photo Electron
OSIRIS	Online Scintillator Internal Radioactivity Investigation System
PE	Photo Electron
PMT	Photo-Multipliers Tubes
PReLU	Parametrized Rectified Linear Unit
ROV	Remotely Operated under-LS Vehicle
ReLU	Rectified Linear Unit
ResNet	Residual Network
SGD	Stochastic Gradient Descent
SPMT	Small PMT
TAO	Taishan Antineutrino Oservatory
TR Area	Total Reflexion Area
TTS	Time Transit Spread
TT	Top Tracker
WCD	Water Cherenkov Detector

1394 Bibliography

- 1395 [1] Liang Zhan, Yifang Wang, Jun Cao, and Liangjian Wen. "Determination of the Neutrino Mass
 1396 Hierarchy at an Intermediate Baseline". *Physical Review D* 78.11 (Dec. 10, 2008), 111103. ISSN:
 1397 1550-7998, 1550-2368. DOI: [10.1103/PhysRevD.78.111103](https://doi.org/10.1103/PhysRevD.78.111103). eprint: [0807.3203\[hep-ex, physics:hep-ph\]](https://arxiv.org/abs/0807.3203). URL: [http://arxiv.org/abs/0807.3203](https://arxiv.org/abs/0807.3203) (visited on 09/18/2023).
- 1398 [2] Fengpeng An et al. "Neutrino Physics with JUNO". *Journal of Physics G: Nuclear and Particle
 1399 Physics* 43.3 (Mar. 1, 2016), 030401. ISSN: 0954-3899, 1361-6471. DOI: [10.1088/0954-3899/43/3/030401](https://doi.org/10.1088/0954-3899/43/3/030401). eprint: [1507.05613\[hep-ex, physics:physics\]](https://arxiv.org/abs/1507.05613). URL: [http://arxiv.org/abs/1507.05613](https://arxiv.org/abs/1507.05613) (visited on 07/28/2023).
- 1400 [3] Liang Zhan, Yifang Wang, Jun Cao, and Liangjian Wen. "Experimental Requirements to Deter-
 1401 mine the Neutrino Mass Hierarchy Using Reactor Neutrinos". *Physical Review D* 79.7 (Apr. 14,
 1402 2009), 073007. ISSN: 1550-7998, 1550-2368. DOI: [10.1103/PhysRevD.79.073007](https://doi.org/10.1103/PhysRevD.79.073007). eprint: [0901.2976\[hep-ex\]](https://arxiv.org/abs/0901.2976). URL: [http://arxiv.org/abs/0901.2976](https://arxiv.org/abs/0901.2976) (visited on 09/18/2023).
- 1403 [4] A. A. Hahn, K. Schreckenbach, W. Gelletly, F. von Feilitzsch, G. Colvin, and B. Krusche. "Antineutrino spectra from 241Pu and 239Pu thermal neutron fission products". *Physics Letters B*
 1404 218.3 (Feb. 23, 1989), 365–368. ISSN: 0370-2693. DOI: [10.1016/0370-2693\(89\)91598-0](https://doi.org/10.1016/0370-2693(89)91598-0). URL:
 1405 <https://www.sciencedirect.com/science/article/pii/0370269389915980> (visited on
 1406 01/16/2024).
- 1407 [5] Th A. Mueller et al. "Improved Predictions of Reactor Antineutrino Spectra". *Physical Review C*
 1408 83.5 (May 23, 2011), 054615. ISSN: 0556-2813, 1089-490X. DOI: [10.1103/PhysRevC.83.054615](https://doi.org/10.1103/PhysRevC.83.054615).
 1409 eprint: [1101.2663\[hep-ex, physics:nucl-ex\]](https://arxiv.org/abs/1101.2663). URL: [http://arxiv.org/abs/1101.2663](https://arxiv.org/abs/1101.2663)
 1410 (visited on 01/16/2024).
- 1411 [6] F. von Feilitzsch, A. A. Hahn, and K. Schreckenbach. "Experimental beta-spectra from 239Pu
 1412 and 235U thermal neutron fission products and their correlated antineutrino spectra". *Physics
 1413 Letters B* 118.1 (Dec. 2, 1982), 162–166. ISSN: 0370-2693. DOI: [10.1016/0370-2693\(82\)90622-0](https://doi.org/10.1016/0370-2693(82)90622-0).
 1414 URL: <https://www.sciencedirect.com/science/article/pii/0370269382906220> (visited
 1415 on 01/16/2024).
- 1416 [7] K. Schreckenbach, G. Colvin, W. Gelletly, and F. Von Feilitzsch. "Determination of the antineu-
 1417 trino spectrum from 235U thermal neutron fission products up to 9.5 MeV". *Physics Letters B*
 1418 160.4 (Oct. 10, 1985), 325–330. ISSN: 0370-2693. DOI: [10.1016/0370-2693\(85\)91337-1](https://doi.org/10.1016/0370-2693(85)91337-1). URL:
 1419 <https://www.sciencedirect.com/science/article/pii/0370269385913371> (visited on
 1420 01/16/2024).
- 1421 [8] Patrick Huber. "On the determination of anti-neutrino spectra from nuclear reactors". *Physical
 1422 Review C* 84.2 (Aug. 29, 2011), 024617. ISSN: 0556-2813, 1089-490X. DOI: [10.1103/PhysRevC.84.024617](https://doi.org/10.1103/PhysRevC.84.024617). eprint: [1106.0687\[hep-ex, physics:hep-ph, physics:nucl-ex, physics:nucl-th\]](https://arxiv.org/abs/1106.0687).
 1423 URL: [http://arxiv.org/abs/1106.0687](https://arxiv.org/abs/1106.0687) (visited on 01/16/2024).
- 1424 [9] P. Vogel, G. K. Schenter, F. M. Mann, and R. E. Schenter. "Reactor antineutrino spectra and
 1425 their application to antineutrino-induced reactions. II". *Physical Review C* 24.4 (Oct. 1, 1981).
 1426 Publisher: American Physical Society, 1543–1553. DOI: [10.1103/PhysRevC.24.1543](https://doi.org/10.1103/PhysRevC.24.1543). URL:
 1427 <https://link.aps.org/doi/10.1103/PhysRevC.24.1543> (visited on 01/16/2024).
- 1428 [10] D. A. Dwyer and T. J. Langford. "Spectral Structure of Electron Antineutrinos from Nuclear
 1429 Reactors". *Physical Review Letters* 114.1 (Jan. 7, 2015), 012502. ISSN: 0031-9007, 1079-7114. DOI:
 1430 [10.1103/PhysRevLett.114.012502](https://doi.org/10.1103/PhysRevLett.114.012502). eprint: [1407.1281\[hep-ex, physics:nucl-ex\]](https://arxiv.org/abs/1407.1281). URL:
 1431 [http://arxiv.org/abs/1407.1281](https://arxiv.org/abs/1407.1281) (visited on 01/16/2024).

- [11] JUNO Collaboration et al. "Sub-percent Precision Measurement of Neutrino Oscillation Parameters with JUNO". *Chinese Physics C* 46.12 (Dec. 1, 2022), 123001. ISSN: 1674-1137, 2058-6132. DOI: [10.1088/1674-1137/ac8bc9](https://doi.org/10.1088/1674-1137/ac8bc9). eprint: [2204.13249 \[hep-ex\]](https://arxiv.org/abs/2204.13249). URL: <http://arxiv.org/abs/2204.13249> (visited on 08/11/2023).
- [12] JUNO Collaboration et al. *TAO Conceptual Design Report: A Precision Measurement of the Reactor Antineutrino Spectrum with Sub-percent Energy Resolution*. May 18, 2020. DOI: [10.48550/arXiv.2005.08745](https://doi.org/10.48550/arXiv.2005.08745). eprint: [2005.08745 \[hep-ex, physics:nucl-ex, physics:physics\]](https://arxiv.org/abs/2005.08745). URL: <http://arxiv.org/abs/2005.08745> (visited on 01/18/2024).
- [13] G. Mention, M. Fechner, Th. Lasserre, Th. A. Mueller, D. Lhuillier, M. Cribier, and A. Letourneau. "Reactor antineutrino anomaly". *Physical Review D* 83.7 (Apr. 29, 2011). Publisher: American Physical Society, 073006. DOI: [10.1103/PhysRevD.83.073006](https://doi.org/10.1103/PhysRevD.83.073006). URL: <https://link.aps.org/doi/10.1103/PhysRevD.83.073006> (visited on 03/05/2024).
- [14] V. Kopeikin, M. Skorokhvatov, and O. Titov. "Reevaluating reactor antineutrino spectra with new measurements of the ratio between ^{235}U and ^{239}Pu β^- spectra". *Physical Review D* 104.7 (Oct. 25, 2021), L071301. ISSN: 2470-0010, 2470-0029. DOI: [10.1103/PhysRevD.104.L071301](https://doi.org/10.1103/PhysRevD.104.L071301). eprint: [2103.01684 \[hep-ph, physics:nucl-ex, physics:nucl-th\]](https://arxiv.org/abs/2103.01684). URL: <http://arxiv.org/abs/2103.01684> (visited on 01/18/2024).
- [15] A. Letourneau et al. "On the origin of the reactor antineutrino anomalies in light of a new summation model with parameterized β^- transitions". *Physical Review Letters* 130.2 (Jan. 10, 2023), 021801. ISSN: 0031-9007, 1079-7114. DOI: [10.1103/PhysRevLett.130.021801](https://doi.org/10.1103/PhysRevLett.130.021801). eprint: [2205.14954 \[hep-ex, physics:hep-ph\]](https://arxiv.org/abs/2205.14954). URL: <http://arxiv.org/abs/2205.14954> (visited on 01/16/2024).
- [16] Particle Data Group et al. "Review of Particle Physics". *Progress of Theoretical and Experimental Physics* 2020.8 (Aug. 14, 2020), 083C01. ISSN: 2050-3911. DOI: [10.1093/ptep/ptaa104](https://doi.org/10.1093/ptep/ptaa104). URL: <https://doi.org/10.1093/ptep/ptaa104> (visited on 12/04/2023).
- [17] Super-Kamiokande Collaboration et al. "Diffuse Supernova Neutrino Background Search at Super-Kamiokande". *Physical Review D* 104.12 (Dec. 10, 2021), 122002. ISSN: 2470-0010, 2470-0029. DOI: [10.1103/PhysRevD.104.122002](https://doi.org/10.1103/PhysRevD.104.122002). eprint: [2109.11174 \[astro-ph, physics:hep-ex\]](https://arxiv.org/abs/2109.11174). URL: <http://arxiv.org/abs/2109.11174> (visited on 02/28/2024).
- [18] Alessandro Strumia and Francesco Vissani. "Precise quasielastic neutrino/nucleon cross section". *Physics Letters B* 564.1 (July 2003), 42–54. ISSN: 03702693. DOI: [10.1016/S0370-2693\(03\)00616-6](https://doi.org/10.1016/S0370-2693(03)00616-6). eprint: [astro-ph/0302055](https://arxiv.org/abs/astro-ph/0302055). URL: <http://arxiv.org/abs/astro-ph/0302055> (visited on 01/16/2024).
- [19] Daya Bay et al. *Optimization of the JUNO liquid scintillator composition using a Daya Bay antineutrino detector*. July 1, 2020. DOI: [10.48550/arXiv.2007.00314](https://doi.org/10.48550/arXiv.2007.00314). eprint: [2007.00314 \[hep-ex, physics:physics\]](https://arxiv.org/abs/2007.00314). URL: <http://arxiv.org/abs/2007.00314> (visited on 07/26/2023).
- [20] J. B. Birks. "CHAPTER 3 - THE SCINTILLATION PROCESS IN ORGANIC MATERIALS—I". *The Theory and Practice of Scintillation Counting*. Ed. by J. B. Birks. International Series of Monographs in Electronics and Instrumentation. Jan. 1, 1964, 39–67. ISBN: 978-0-08-010472-0. DOI: [10.1016/B978-0-08-010472-0.50008-2](https://doi.org/10.1016/B978-0-08-010472-0.50008-2). URL: <https://www.sciencedirect.com/science/article/pii/B9780080104720500082> (visited on 02/07/2024).
- [21] Photomultiplier tube R12860 | Hamamatsu Photonics. URL: https://www.hamamatsu.com/eu/en/product/optical-sensors/pmt/pmt_tube-alone/head-on-type/R12860.html (visited on 02/08/2024).
- [22] Yan Zhang, Ze-Yuan Yu, Xin-Ying Li, Zi-Yan Deng, and Liang-Jian Wen. "A complete optical model for liquid-scintillator detectors". *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 967 (July 2020), 163860. ISSN: 01689002. DOI: [10.1016/j.nima.2020.163860](https://doi.org/10.1016/j.nima.2020.163860). eprint: [2003.12212 \[physics\]](https://arxiv.org/abs/2003.12212). URL: <http://arxiv.org/abs/2003.12212> (visited on 02/07/2024).
- [23] Hai-Bo Yang et al. "Light Attenuation Length of High Quality Linear Alkyl Benzene as Liquid Scintillator Solvent for the JUNO Experiment". *Journal of Instrumentation* 12.11 (Nov. 27, 2017), T11004–T11004. ISSN: 1748-0221. DOI: [10.1088/1748-0221/12/11/T11004](https://doi.org/10.1088/1748-0221/12/11/T11004). eprint: [1703.04890](https://arxiv.org/abs/1703.04890).

- 1490 01867 [hep-ex, physics:physics]. URL: <http://arxiv.org/abs/1703.01867> (visited on
1491 07/28/2023).
- 1492 [24] JUNO Collaboration et al. *The Design and Sensitivity of JUNO's scintillator radiopurity pre-detector*
1493 OSIRIS. Mar. 31, 2021. DOI: [10.48550/arXiv.2103.16900](https://doi.org/10.48550/arXiv.2103.16900). eprint: [2103.16900](https://arxiv.org/abs/2103.16900) [physics]. URL:
1494 [http://arxiv.org/abs/2103.16900](https://arxiv.org/abs/2103.16900) (visited on 02/07/2024).
- 1495 [25] Angel Abusleme et al. "Mass Testing and Characterization of 20-inch PMTs for JUNO". *The*
1496 *European Physical Journal C* 82.12 (Dec. 24, 2022), 1168. ISSN: 1434-6052. DOI: [10.1140/epjc/s10052-022-11002-8](https://doi.org/10.1140/epjc/s10052-022-11002-8). eprint: [2205.08629](https://arxiv.org/abs/2205.08629) [hep-ex, physics:physics]. URL: [http://arxiv.org/abs/2205.08629](https://arxiv.org/abs/2205.08629) (visited on 02/08/2024).
- 1497 [26] Yang Han. "Dual Calorimetry for High Precision Neutrino Oscillation Measurement at JUNO
1498 Experiment". *AstroParticule et Cosmologie*, France, Paris U. VII, APC, June 2021.
- 1499 [27] R. Acquafredda et al. "The OPERA experiment in the CERN to Gran Sasso neutrino beam".
1500 *Journal of Instrumentation* 4.4 (Apr. 2009), P04018. ISSN: 1748-0221. DOI: [10.1088/1748-0221/4/04/P04018](https://doi.org/10.1088/1748-0221/4/04/P04018). URL: <https://dx.doi.org/10.1088/1748-0221/4/04/P04018> (visited on
1501 02/29/2024).
- 1502 [28] JUNO collaboration et al. "Calibration Strategy of the JUNO Experiment". *Journal of High*
1503 *Energy Physics* 2021.3 (Mar. 2021), 4. ISSN: 1029-8479. DOI: [10.1007/JHEP03\(2021\)004](https://doi.org/10.1007/JHEP03(2021)004). eprint:
1504 [2011.06405](https://arxiv.org/abs/2011.06405) [hep-ex, physics:physics]. URL: [http://arxiv.org/abs/2011.06405](https://arxiv.org/abs/2011.06405) (visited
1505 on 08/10/2023).
- 1506 [29] Hans Th J. Steiger. *TAO – The Taishan Antineutrino Observatory*. Sept. 21, 2022. DOI: [10.48550/arXiv.2209.10387](https://doi.org/10.48550/arXiv.2209.10387). eprint: [2209.10387](https://arxiv.org/abs/2209.10387) [physics]. URL: [http://arxiv.org/abs/2209.10387](https://arxiv.org/abs/2209.10387)
1507 (visited on 01/16/2024).
- 1508 [30] Tao Lin et al. "The Application of SNiPER to the JUNO Simulation". *Journal of Physics: Conference Series* 898.4 (Oct. 2017). Publisher: IOP Publishing, 042029. ISSN: 1742-6596. DOI: [10.1088/1742-6596/898/4/042029](https://doi.org/10.1088/1742-6596/898/4/042029). URL: <https://dx.doi.org/10.1088/1742-6596/898/4/042029> (visited on 02/27/2024).
- 1509 [31] S. Agostinelli et al. "Geant4—a simulation toolkit". *Nuclear Instruments and Methods in Physics*
1510 *Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 506.3 (July 1,
1511 2003), 250–303. ISSN: 0168-9002. DOI: [10.1016/S0168-9002\(03\)01368-8](https://doi.org/10.1016/S0168-9002(03)01368-8). URL: <https://www.sciencedirect.com/science/article/pii/S0168900203013688> (visited on 02/27/2024).
- 1512 [32] J. Allison et al. "Geant4 developments and applications". *IEEE Transactions on Nuclear Science*
1513 53.1 (Feb. 2006). Conference Name: IEEE Transactions on Nuclear Science, 270–278. ISSN: 1558-
1514 1578. DOI: [10.1109/TNS.2006.869826](https://doi.org/10.1109/TNS.2006.869826). URL: <https://ieeexplore.ieee.org/document/1610988?isnumber=33833&arnumber=1610988&count=33&index=7> (visited on 02/27/2024).
- 1515 [33] J. Allison et al. "Recent developments in Geant4". *Nuclear Instruments and Methods in Physics*
1516 *Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 835 (Nov. 1,
1517 2016), 186–225. ISSN: 0168-9002. DOI: [10.1016/j.nima.2016.06.125](https://doi.org/10.1016/j.nima.2016.06.125). URL: <https://www.sciencedirect.com/science/article/pii/S0168900216306957> (visited on 02/27/2024).
- 1518 [34] Wenjie Wu, Miao He, Xiang Zhou, and Haoxue Qiao. "A new method of energy reconstruction
1519 for large spherical liquid scintillator detectors". *Journal of Instrumentation* 14.3 (Mar. 8, 2019),
1520 P03009–P03009. ISSN: 1748-0221. DOI: [10.1088/1748-0221/14/03/P03009](https://doi.org/10.1088/1748-0221/14/03/P03009). eprint: [1812.01799](https://arxiv.org/abs/1812.01799) [hep-ex, physics:physics]. URL: [http://arxiv.org/abs/1812.01799](https://arxiv.org/abs/1812.01799) (visited on
1521 07/28/2023).
- 1522 [35] Guihong Huang et al. "Improving the energy uniformity for large liquid scintillator detectors".
1523 *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers,*
1524 *Detectors and Associated Equipment* 1001 (June 11, 2021), 165287. ISSN: 0168-9002. DOI: [10.1016/j.nima.2021.165287](https://doi.org/10.1016/j.nima.2021.165287). URL: <https://www.sciencedirect.com/science/article/pii/S0168900221002710> (visited on 03/01/2024).
- 1525 [36] Ziyuan Li et al. "Event vertex and time reconstruction in large volume liquid scintillator detector".
1526 *Nuclear Science and Techniques* 32.5 (May 2021), 49. ISSN: 1001-8042, 2210-3147. DOI:
1527 [10.1007/s41365-021-00885-z](https://doi.org/10.1007/s41365-021-00885-z). eprint: [2101.08901](https://arxiv.org/abs/2101.08901) [hep-ex, physics:physics]. URL: [http://arxiv.org/abs/2101.08901](https://arxiv.org/abs/2101.08901) (visited on 07/28/2023).

- [37] Gioacchino Ranucci. "An analytical approach to the evaluation of the pulse shape discrimination properties of scintillators". *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 354.2 (Jan. 30, 1995), 389–399. ISSN: 0168-9002. DOI: [10.1016/0168-9002\(94\)00886-8](https://doi.org/10.1016/0168-9002(94)00886-8). URL: <https://www.sciencedirect.com/science/article/pii/0168900294008868> (visited on 03/07/2024).
- [38] C. Galbiati and K. McCarty. "Time and space reconstruction in optical, non-imaging, scintillator-based particle detectors". *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 568.2 (Dec. 1, 2006), 700–709. ISSN: 0168-9002. DOI: [10.1016/j.nima.2006.07.058](https://doi.org/10.1016/j.nima.2006.07.058). URL: <https://www.sciencedirect.com/science/article/pii/S0168900206013519> (visited on 03/07/2024).
- [39] M. Moszyński and B. Bengtson. "Status of timing with plastic scintillation detectors". *Nuclear Instruments and Methods* 158 (Jan. 1, 1979), 1–31. ISSN: 0029-554X. DOI: [10.1016/S0029-554X\(79\)90170-8](https://doi.org/10.1016/S0029-554X(79)90170-8). URL: <https://www.sciencedirect.com/science/article/pii/S0029554X79901708> (visited on 03/07/2024).
- [40] Gui-Hong Huang, Wei Jiang, Liang-Jian Wen, Yi-Fang Wang, and Wu-Ming Luo. "Data-driven simultaneous vertex and energy reconstruction for large liquid scintillator detectors". *Nuclear Science and Techniques* 34.6 (June 17, 2023), 83. ISSN: 2210-3147. DOI: [10.1007/s41365-023-01240-0](https://doi.org/10.1007/s41365-023-01240-0). URL: <https://doi.org/10.1007/s41365-023-01240-0> (visited on 08/17/2023).
- [41] Zhen Qian et al. "Vertex and Energy Reconstruction in JUNO with Machine Learning Methods". *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 1010 (Sept. 2021), 165527. ISSN: 01689002. DOI: [10.1016/j.nima.2021.165527](https://doi.org/10.1016/j.nima.2021.165527). eprint: [2101.04839\[hep-ex, physics:physics\]](https://arxiv.org/abs/2101.04839). URL: <http://arxiv.org/abs/2101.04839> (visited on 07/24/2023).
- [42] Arsenii Gavrikov, Yury Malyshkin, and Fedor Ratnikov. "Energy reconstruction for large liquid scintillator detectors with machine learning techniques: aggregated features approach". *The European Physical Journal C* 82.11 (Nov. 14, 2022), 1021. ISSN: 1434-6052. DOI: [10.1140/epjc/s10052-022-11004-6](https://doi.org/10.1140/epjc/s10052-022-11004-6). eprint: [2206.09040\[physics\]](https://arxiv.org/abs/2206.09040). URL: <http://arxiv.org/abs/2206.09040> (visited on 07/24/2023).
- [43] R. Abbasi et al. "Graph Neural Networks for low-energy event classification & reconstruction in IceCube". *Journal of Instrumentation* 17.11 (Nov. 2022). Publisher: IOP Publishing, P11003. ISSN: 1748-0221. DOI: [10.1088/1748-0221/17/11/P11003](https://doi.org/10.1088/1748-0221/17/11/P11003). URL: <https://dx.doi.org/10.1088/1748-0221/17/11/P11003> (visited on 04/04/2024).
- [44] S. Reck, D. Guderian, G. Vermarien, A. Domi, and on behalf of the KM3NeT collaboration on behalf of the. "Graph neural networks for reconstruction and classification in KM3NeT". *Journal of Instrumentation* 16.10 (Oct. 2021). Publisher: IOP Publishing, C10011. ISSN: 1748-0221. DOI: [10.1088/1748-0221/16/10/C10011](https://doi.org/10.1088/1748-0221/16/10/C10011). URL: <https://dx.doi.org/10.1088/1748-0221/16/10/C10011> (visited on 04/04/2024).
- [45] The IceCube collaboration et al. "A convolutional neural network based cascade reconstruction for the IceCube Neutrino Observatory". *Journal of Instrumentation* 16.7 (July 2021). Publisher: IOP Publishing, P07041. ISSN: 1748-0221. DOI: [10.1088/1748-0221/16/07/P07041](https://doi.org/10.1088/1748-0221/16/07/P07041). URL: <https://dx.doi.org/10.1088/1748-0221/16/07/P07041> (visited on 04/04/2024).
- [46] DUNE Collaboration et al. "Neutrino interaction classification with a convolutional neural network in the DUNE far detector". *Physical Review D* 102.9 (Nov. 9, 2020). Publisher: American Physical Society, 092003. DOI: [10.1103/PhysRevD.102.092003](https://doi.org/10.1103/PhysRevD.102.092003). URL: <https://link.aps.org/doi/10.1103/PhysRevD.102.092003> (visited on 04/04/2024).
- [47] K. M. Górski, E. Hivon, A. J. Banday, B. D. Wandelt, F. K. Hansen, M. Reinecke, and M. Bartelmann. "HEALPix: A Framework for High-Resolution Discretization and Fast Analysis of Data Distributed on the Sphere". *The Astrophysical Journal* 622 (Apr. 1, 2005). ADS Bibcode: 2005ApJ...622..759G, 759–771. ISSN: 0004-637X. DOI: [10.1086/427976](https://doi.org/10.1086/427976). URL: <https://ui.adsabs.harvard.edu/abs/2005ApJ...622..759G> (visited on 04/04/2024).
- [48] Michaël Defferrard, Xavier Bresson, and Pierre Vandergheynst. *Convolutional Neural Networks on Graphs with Fast Localized Spectral Filtering*. Feb. 5, 2017. DOI: [10.48550/arXiv.1606.09375](https://arxiv.org/abs/1606.09375).

- 1594 eprint: 1606 . 09375[cs , stat]. URL: <http://arxiv.org/abs/1606.09375> (visited on
1595 04/04/2024).
- 1596 [49] JUNO Collaboration et al. "JUNO Physics and Detector". *Progress in Particle and Nuclear Physics*
1597 123 (Mar. 2022), 103927. ISSN: 01466410. DOI: 10 . 1016 / j . ppnp . 2021 . 103927. eprint: 2104 .
1598 02565 [hep-ex]. URL: <http://arxiv.org/abs/2104.02565> (visited on 09/18/2023).
- 1599 [50] Leo Breiman, Jerome Friedman, R. A. Olshen, and Charles J. Stone. *Classification and Regression
1600 Trees*. New York: Chapman and Hall/CRC, Oct. 25, 2017. 368 pp. ISBN: 978-1-315-13947-0. DOI:
1601 10 . 1201 / 9781315139470.
- 1602 [51] Jerome H. Friedman. "Greedy function approximation: A gradient boosting machine." *The Annals of Statistics* 29.5 (Oct. 2001). Publisher: Institute of Mathematical Statistics, 1189–1232. ISSN:
1603 0090-5364, 2168-8966. DOI: 10 . 1214 / aos / 1013203451. URL: <https://projecteuclid.org/journals/annals-of-statistics/volume-29/issue-5/Greedy-function-approximation-A-gradient-boosting-machine/10.1214/aos/1013203451.full> (visited on 04/29/2024).
- 1604 [52] J. Y. Lettin, H. R. Maturana, W. S. McCulloch, and W. H. Pitts. "What the Frog's Eye Tells
1605 the Frog's Brain". *Proceedings of the IRE* 47.11 (Nov. 1959). Conference Name: Proceedings of
1606 the IRE, 1940–1951. ISSN: 2162-6634. DOI: 10 . 1109 / JPROC . 1959 . 287207. URL: <https://ieeexplore.ieee.org/document/4065609> (visited on 05/06/2024).
- 1607 [53] F. Rosenblatt. "The perceptron: A probabilistic model for information storage and organization
1608 in the brain". *Psychological Review* 65.6 (1958). Place: US Publisher: American Psychological
1609 Association, 386–408. ISSN: 1939-1471. DOI: 10 . 1037 / h0042519.
- 1610 [54] Diederik P. Kingma and Jimmy Ba. *Adam: A Method for Stochastic Optimization*. Jan. 29, 2017.
1611 DOI: 10 . 48550 / arXiv . 1412 . 6980. eprint: 1412 . 6980[cs]. URL: <http://arxiv.org/abs/1412.6980> (visited on 05/13/2024).
- 1612 [55] Olga Russakovsky et al. *ImageNet Large Scale Visual Recognition Challenge*. Jan. 29, 2015. DOI:
1613 10 . 48550 / arXiv . 1409 . 0575. eprint: 1409 . 0575[cs]. URL: <http://arxiv.org/abs/1409.0575>
1614 (visited on 05/17/2024).
- 1615 [56] Karen Simonyan and Andrew Zisserman. *Very Deep Convolutional Networks for Large-Scale Image
1616 Recognition*. Apr. 10, 2015. DOI: 10 . 48550 / arXiv . 1409 . 1556. eprint: 1409 . 1556[cs]. URL:
1617 <http://arxiv.org/abs/1409.1556> (visited on 05/17/2024).
- 1618 [57] Anna Allen. *generic-github-user/Image-Convolution-Playground*. original-date: 2018-09-28T22:42:55Z.
1619 July 15, 2024. URL: <https://github.com/generic-github-user/Image-Convolution-Playground> (visited on 07/16/2024).
- 1620 [58] Jason Ansel et al. *PyTorch 2: Faster Machine Learning Through Dynamic Python Bytecode Trans-
1621 formation and Graph Compilation*. Publication Title: 29th ACM International Conference on Ar-
1622 chitectural Support for Programming Languages and Operating Systems, Volume 2 (ASPLOS
1623 '24) original-date: 2016-08-13T05:26:41Z. Apr. 2024. DOI: 10 . 1145 / 3620665 . 3640366. URL:
1624 <https://pytorch.org/assets/pytorch2-2.pdf> (visited on 07/16/2024).
- 1625 [59] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner. "Gradient-based learning applied to document
1626 recognition". *Proceedings of the IEEE* 86.11 (Nov. 1998). Conference Name: Proceedings of the
1627 IEEE, 2278–2324. ISSN: 1558-2256. DOI: 10 . 1109 / 5 . 726791. URL: <https://ieeexplore.ieee.org/document/726791> (visited on 07/16/2024).
- 1628 [60] NVIDIA T4 Tensor Core GPUs for Accelerating Inference. NVIDIA. URL: <https://www.nvidia.com/en-gb/data-center/tesla-t4/> (visited on 07/16/2024).
- 1629 [61] Justin Gilmer, Samuel S. Schoenholz, Patrick F. Riley, Oriol Vinyals, and George E. Dahl. *Neural
1630 Message Passing for Quantum Chemistry*. June 12, 2017. DOI: 10 . 48550 / arXiv . 1704 . 01212.
1631 eprint: 1704 . 01212[cs]. URL: <http://arxiv.org/abs/1704.01212> (visited on 05/22/2024).
- 1632 [62] Yaguang Li, Rose Yu, Cyrus Shahabi, and Yan Liu. *Diffusion Convolutional Recurrent Neural
1633 Network: Data-Driven Traffic Forecasting*. Feb. 22, 2018. DOI: 10 . 48550 / arXiv . 1707 . 01926.
1634 eprint: 1707 . 01926[cs , stat]. URL: <http://arxiv.org/abs/1707.01926> (visited on
1635 05/22/2024).
- 1636 [63] Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil
1637 Ozair, Aaron Courville, and Yoshua Bengio. *Generative Adversarial Networks*. June 10, 2014. DOI:
1638 10 . 1145 / 2623378 . 2623437. URL: <https://dl.acm.org/doi/10.1145/2623378.2623437>

- 1646 10.48550/arXiv.1406.2661. eprint: 1406.2661[cs,stat]. URL: <http://arxiv.org/abs/1406.2661> (visited on 05/29/2024).
- 1647 [64] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. "Deep Residual Learning for Image Recognition". *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). ISSN: 1063-6919. June 2016, 770–778. DOI: 10.1109/CVPR.2016.90. URL: <https://ieeexplore.ieee.org/document/7780459> (visited on 07/17/2024).
- 1650 [65] Victor Lebrin. "Towards the Detection of Core-Collapse Supernovae Burst Neutrinos with the 3-inch PMT System of the JUNO Detector". These de doctorat. Nantes Université, Sept. 5, 2022.
- 1652 URL: <https://theses.fr/2022NANU4080> (visited on 05/22/2024).
- 1654 [66] Dan Cireşan, Ueli Meier, and Juergen Schmidhuber. *Multi-column Deep Neural Networks for Image Classification*. version: 1. Feb. 13, 2012. DOI: 10.48550/arXiv.1202.2745. eprint: 1202.2745[cs]. URL: <http://arxiv.org/abs/1202.2745> (visited on 06/27/2024).
- 1656 [67] R. Abbasi et al. "A Convolutional Neural Network based Cascade Reconstruction for the Ice-Cube Neutrino Observatory". *Journal of Instrumentation* 16.7 (July 1, 2021), P07041. ISSN: 1748-0221. DOI: 10.1088/1748-0221/16/07/P07041. eprint: 2101.11589[hep-ex]. URL: <http://arxiv.org/abs/2101.11589> (visited on 06/27/2024).
- 1658 [68] D. Maksimović, M. Nieslony, and M. Wurm. "CNNs for enhanced background discrimination in DSNB searches in large-scale water-Gd detectors". *Journal of Cosmology and Astroparticle Physics* 2021.11 (Nov. 2021). Publisher: IOP Publishing, 051. ISSN: 1475-7516. DOI: 10.1088/1475-7516/2021/11/051 (visited on 06/27/2024).
- 1660 [69] Taco S. Cohen, Mario Geiger, Jonas Koehler, and Max Welling. *Spherical CNNs*. Feb. 25, 2018.
- 1662 DOI: 10.48550/arXiv.1801.10130. eprint: 1801.10130[cs,stat]. URL: <http://arxiv.org/abs/1801.10130> (visited on 07/13/2024).
- 1664 [70] "IEEE Standard for Floating-Point Arithmetic". *IEEE Std 754-2019 (Revision of IEEE 754-2008)* (July 2019). Conference Name: IEEE Std 754-2019 (Revision of IEEE 754-2008), 1–84. DOI: 10.1109/IEEEESTD.2019.8766229. URL: <https://ieeexplore.ieee.org/document/8766229> (visited on 07/03/2024).
- 1666 [71] Chuanya Cao et al. "Mass production and characterization of 3-inch PMTs for the JUNO experiment". *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 1005 (July 2021), 165347. ISSN: 01689002. DOI: 10.1016/j.nima.2021.165347. eprint: 2102.11538[hep-ex, physics:physics]. URL: <http://arxiv.org/abs/2102.11538> (visited on 02/08/2024).
- 1668
1670
1672
1674
1676
1678
1680