

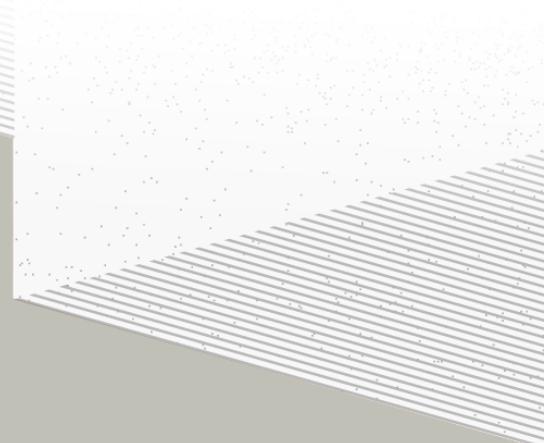
1

2

# THÈSE DE DOCTORAT DE

L'UNIVERSITÉ DE NANTES

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



Par

**Léonard Imbert**

**Deep learning methods and Dual Calorimetric analysis for high precision neutrino oscillation measurements at JUNO**

Thèse présentée et soutenue à Nantes, le 2 Decembre 2024

Unité de recherche : Laboratoire SUBATECH, UMR 6457

## Rapporteurs avant soutenance :

Christine Marquet	Directrice de recherche, LP2I Bordeaux
David Rousseau	Directeur de recherche, IJCLab

## Composition du Jury :

Président :	Barbara Erazmus	Directrice de recherche, Subatech
Examinateurs :	Juan Pedro Ochoa-Ricoux	Professor, University of California, Irvine
	Yasmine Amhis	Directrice de recherche, IJCLab
	Christine Marquet	Directrice de recherche, LP2I Bordeaux
	David Rousseau	Directeur de recherche, IJCLab
Dir. de thèse :	Frédéric Yermia	Professeur, Université de Nantes
Co-dir. de thèse :	Benoit Viaud	Chargé de recherche, CNRS



# <sup>3</sup> Contents

<sup>4</sup>	<b>Contents</b>	<sup>1</sup>
<sup>5</sup>	<b>Remerciements</b>	<sup>5</sup>
<sup>6</sup>	<b>Introduction</b>	<sup>7</sup>
<sup>7</sup>	<b>1 Neutrino physics</b>	<sup>9</sup>
<sup>8</sup>	1.1 Standard model . . . . .	<sup>9</sup>
<sup>9</sup>	1.1.1 Limits of the standard model . . . . .	<sup>9</sup>
<sup>10</sup>	1.2 Historic of the neutrino . . . . .	<sup>9</sup>
<sup>11</sup>	1.3 Oscillation . . . . .	<sup>9</sup>
<sup>12</sup>	1.3.1 Phenomologies . . . . .	<sup>9</sup>
<sup>13</sup>	1.4 Open questions . . . . .	<sup>9</sup>
<sup>14</sup>	<b>2 The JUNO experiment</b>	<sup>11</sup>
<sup>15</sup>	2.1 Neutrinos physics in JUNO . . . . .	<sup>12</sup>
<sup>16</sup>	2.1.1 Reactor neutrino oscillation for NMO and precise measurements . . . . .	<sup>12</sup>
<sup>17</sup>	2.1.2 Other physics . . . . .	<sup>15</sup>
<sup>18</sup>	2.2 The JUNO detector . . . . .	<sup>16</sup>
<sup>19</sup>	2.2.1 Detection principle . . . . .	<sup>17</sup>
<sup>20</sup>	2.2.2 Central Detector (CD) . . . . .	<sup>18</sup>
<sup>21</sup>	2.2.3 Veto detector . . . . .	<sup>22</sup>
<sup>22</sup>	2.3 Calibration strategy . . . . .	<sup>23</sup>
<sup>23</sup>	2.3.1 Energy scale calibration . . . . .	<sup>23</sup>
<sup>24</sup>	2.3.2 Calibration system . . . . .	<sup>24</sup>
<sup>25</sup>	2.3.3 Instrumental non-linearity calibration . . . . .	<sup>24</sup>
<sup>26</sup>	2.4 Satellite detectors . . . . .	<sup>25</sup>
<sup>27</sup>	2.4.1 TAO . . . . .	<sup>25</sup>
<sup>28</sup>	2.4.2 OSIRIS . . . . .	<sup>26</sup>
<sup>29</sup>	2.5 Software . . . . .	<sup>27</sup>
<sup>30</sup>	2.6 State of the art of the Offline IBD reconstruction in JUNO . . . . .	<sup>28</sup>
<sup>31</sup>	2.6.1 Interaction vertex reconstruction . . . . .	<sup>28</sup>
<sup>32</sup>	2.6.2 Energy reconstruction . . . . .	<sup>32</sup>
<sup>33</sup>	2.6.3 Machine learning for reconstruction . . . . .	<sup>34</sup>
<sup>34</sup>	2.7 JUNO sensitivity to NMO and precise measurements . . . . .	<sup>37</sup>

35	2.7.1	Theoretical spectrum . . . . .	37
36	2.7.2	Fitting procedure . . . . .	38
37	2.7.3	Physics results . . . . .	39
38	2.8	Summary . . . . .	39
39	<b>3</b>	<b>Machine learning and Artificial Neural Network</b>	<b>41</b>
40	3.1	Boosted Decision Tree (BDT) . . . . .	41
41	3.2	Artificial Neural Network (NN) . . . . .	42
42	3.2.1	Fully Connected Deep Neural Network (FCDNN) . . . . .	43
43	3.2.2	Convolutional Neural Network (CNN) . . . . .	43
44	3.2.3	Graph Neural Network (GNN) . . . . .	45
45	3.2.4	Adversarial Neural Network (ANN) . . . . .	47
46	3.2.5	Training procedure . . . . .	47
47	3.2.6	Potential pitfalls . . . . .	49
48	<b>4</b>	<b>Image recognition for IBD reconstruction with the SPMT system</b>	<b>53</b>
49	4.1	Motivations . . . . .	53
50	4.2	Method and model . . . . .	54
51	4.2.1	Model . . . . .	54
52	4.2.2	Data representation . . . . .	56
53	4.2.3	Dataset . . . . .	57
54	4.2.4	Data characteristics . . . . .	58
55	4.3	Training . . . . .	60
56	4.4	Results . . . . .	61
57	4.4.1	J21 results . . . . .	62
58	4.4.2	J21 Combination of classic and ML estimator . . . . .	63
59	4.4.3	J23 results . . . . .	66
60	4.5	Conclusion and prospect . . . . .	67
61	<b>5</b>	<b>Graph representation of JUNO for IBD reconstruction</b>	<b>69</b>
62	5.1	Motivation . . . . .	69
63	5.2	Data representation . . . . .	70
64	5.3	Message passing algorithm . . . . .	72
65	5.4	Data . . . . .	74
66	5.5	Model . . . . .	75
67	5.6	Training . . . . .	76
68	5.7	Optimization . . . . .	76
69	5.8	Results . . . . .	77
70	5.9	Conclusion . . . . .	78
71	<b>6</b>	<b>Reliability of machine learning methods</b>	<b>81</b>
72	<b>7</b>	<b>Joint fit between the SPMT and LPMT spectra</b>	<b>83</b>
73	7.1	Motivations . . . . .	84

74	7.1.1	Discrepancies between the SPMT and LPMT results . . . . .	84
75	7.1.2	Charge Non-Linearity (QNL) . . . . .	84
76	7.2	Approach . . . . .	85
77	7.2.1	Data production . . . . .	85
78	7.2.2	Individual fits . . . . .	87
79	7.2.3	Joint fit . . . . .	88
80	7.2.4	Data and theoretical spectrum generation . . . . .	89
81	7.2.5	Limitations . . . . .	89
82	7.3	Fit software . . . . .	90
83	7.3.1	IBD generator . . . . .	90
84	7.3.2	Fit . . . . .	92
85	7.4	Technical challenges and development . . . . .	93
86	7.5	Results . . . . .	93
87	7.5.1	Validation . . . . .	93
88	7.5.2	Covariance matrix . . . . .	97
89	7.5.3	Statistical tests . . . . .	100
90	7.6	Conclusion and perspectives . . . . .	100
91	<b>8</b>	<b>Conclusion</b>	<b>103</b>
92	<b>A</b>	<b>Calculation of optimal <math>\alpha</math> for estimator combination</b>	<b>105</b>
93	A.1	Unbiased estimator . . . . .	105
94	A.2	Optimal variance estimator . . . . .	105
95	<b>B</b>	<b>Charge spherical harmonics analysis</b>	<b>107</b>
96	<b>C</b>	<b>Over-gaussian smearing of spectrum</b>	<b>115</b>
97	<b>List of Tables</b>		<b>117</b>
98	<b>List of Figures</b>		<b>124</b>
99	<b>List of Abbreviations</b>		<b>125</b>
100	<b>Bibliography</b>		<b>127</b>



<sup>101</sup> **Remerciements**



<sup>102</sup> **Introduction**



<sup>103</sup> **Chapter 1**

<sup>104</sup> **Neutrino physics**

<sup>105</sup>

*The neutrino, or  $\nu$  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.*

<sup>106</sup> **1.1 Standard model**

<sup>107</sup> **1.1.1 Limits of the standard model**

<sup>108</sup> **1.2 Historic of the neutrino**

<sup>109</sup> **First theories**

<sup>110</sup> **Discovery**

<sup>111</sup> **Milestones and anomalies**

<sup>112</sup> **1.3 Oscillation**

<sup>113</sup> **1.3.1 Phenomologies**

<sup>114</sup> **1.4 Open questions**

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



<sup>115</sup> **Chapter 2**

<sup>116</sup> **The JUNO experiment**

<sup>117</sup> “*Ave Juno, rosae rosam, et spiritus rex*”. It means nothing but I found it in tone.

<sup>118</sup> The first idea of a medium baseline ( $\sim 52$  km) experiment, was explored in 2008 [1] where it was  
<sup>119</sup> demonstrated that the Neutrino Mass Ordering (NMO) could be determined by a medium baseline  
<sup>120</sup> experiment if  $\sin^2(2\theta_{13}) > 0.005$  without the requirements of accurate knowledge of the reactor  
<sup>121</sup> antineutrino spectra and the value of  $\Delta m_{32}^2$ . From this idea is born the Jiangmen Underground  
<sup>122</sup> Neutrino Observatory (JUNO) experiment.

<sup>123</sup> JUNO is a neutrino detection experiment under construction located in China, in Guangdong prov-  
<sup>124</sup> ing, near the city of Kaiping. Its main objectives are the determination of the mass ordering at the  
<sup>125</sup>  $3\text{-}4\sigma$  level in 6 years of data taking and the measurement at the sub-percent precision of the oscillation  
<sup>126</sup> parameters  $\Delta m_{21}^2$ ,  $\sin^2 \theta_{12}$ ,  $\Delta m_{32}^2$  and with less precision  $\sin^2 \theta_{13}$ [2].

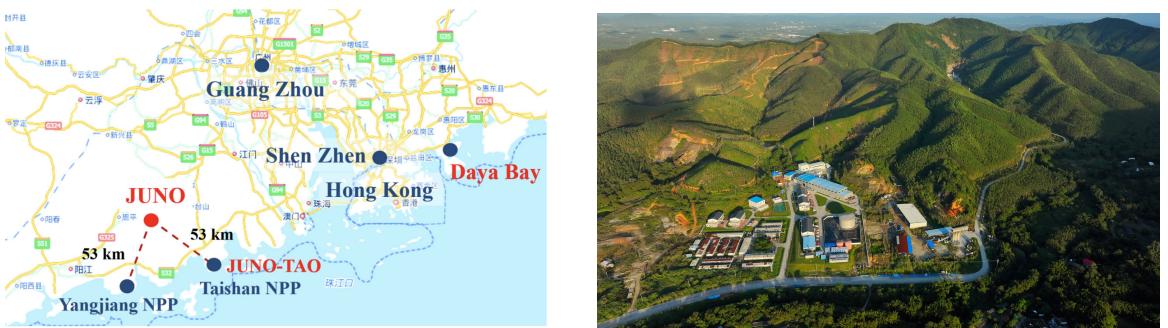


FIGURE 2.1 – On the left: Location of the JUNO experiment and its reactor sources in southern China. On the right: Aerial view of the experimental site

<sup>127</sup> For this JUNO will measure the electronic anti-neutrinos ( $\bar{\nu}_e$ ) flux coming from the nuclear reactors  
<sup>128</sup> of Taishan, Yangjiang, for a total power of  $26.6 \text{ GW}_{th}$ , and the Daya Bay power plant to a lesser  
<sup>129</sup> extent. All of those cores are the second-generation pressurized water reactors CPR1000, which is a  
<sup>130</sup> derivative of Framatome M310. Details about the power plants characteristics and their expected flux  
<sup>131</sup> of  $\bar{\nu}_e$  can be found in the table 2.1. The distance of 53 km has been specifically chosen to maximize  
<sup>132</sup> the disappearance probability of the  $\bar{\nu}_e$ . The data taking is scheduled to start early 2025.

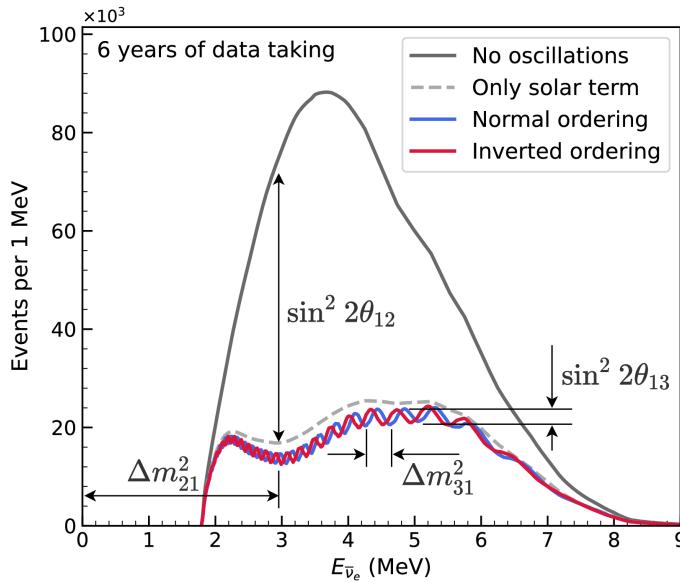


FIGURE 2.2 – Expected number of neutrinos event per MeV in JUNO after 6 years of data taking. The black curve shows the flux if there was no oscillation. The light gray curve shows the oscillation if only the solar terms are taken in account ( $\theta_{12}$ ,  $\Delta m_{21}^2$ ). The blue and red curve 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 blue curve.

## <sup>133</sup> 2.1 Neutrinos physics in JUNO

<sup>134</sup> Even if the JUNO design detailed in section 2.2 was optimized for the measurement of the NMO, its  
<sup>135</sup> large detection volume, excellent energy resolution and background level and understanding make it  
<sup>136</sup> also an excellent detector to measure the flux coming from other neutrino sources. Thus the scientific  
<sup>137</sup> program of JUNO extends way over reactor antineutrinos. The following section is an overview of  
<sup>138</sup> the different physics topic JUNO will contribute in the coming years.

### <sup>139</sup> 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] :

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

<sup>140</sup> 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  
<sup>141</sup> to the NMO in the dependency to  $\Delta m_{32}^2$  and  $\Delta m_{31}^2$  causing a phase shift of the spectrum as we can  
<sup>142</sup> see in the figure 2.2. By carefully adjusting a theoretical spectrum to the data, one can extract the  
<sup>143</sup> NMO and the oscillation parameters. The statistic procedure used to adjust the theoretical spectrum  
<sup>144</sup> is reviewed in more details in the section 2.7. To reach the desired sensitivity, JUNO must meet  
<sup>145</sup> multiple requirements but most notably:

- 146 1. An energy resolution of  $3\%/\sqrt{E(\text{MeV})}$  to be able to distinguish the fine structure of the fast  
147 oscillation.
- 148 2. An energy precision of 1% in order to not err on the location of the oscillation pattern.
- 149 3. A baseline between 40 and 65 km to maximise the  $\bar{\nu}_e$  oscillation probability. The optimal  
150 baseline would be 58 km and JUNO baseline is 53 km.
- 151 4. At least  $\approx 100,000$  events to limit the spectrum distortion due to statistical uncertainties.

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

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

Reactor	Power (GW <sub>th</sub> )	Baseline (km)
Taishan	9.2	52.71
Core 1	4.6	52.77
Core 2	4.6	52.64
Yangjiang	17.4	52.46
Core 1	2.9	52.74
Core 2	2.9	52.82
Core 3	2.9	52.41
Core 4	2.9	52.49
Core 5	2.9	52.11
Core 6	2.9	52.19
Daya Bay	17.4	215
Huizhou	17.4	265

TABLE 2.1 – Characteristics of the nuclear power plants observed by JUNO.

157 The  $\bar{\nu}_e$  coming from reactors are emitted from  $\beta$ -decay of unstable fission fragments. The Taishan  
158 and Yangjiang reactors are Pressurised Water Reactor (PWR), the same type as Daya Bay. In those  
159 type of reactor more than 99.7 % and  $\bar{\nu}_e$  are produced by the fissions of four fuel isotopes  $^{235}\text{U}$ ,  $^{238}\text{U}$ ,  
160  $^{239}\text{Pu}$  and  $^{241}\text{Pu}$ . The neutrino flux per fission of each isotope is determined by the inversion of the  
161 measured  $\beta$  spectra of fission product [4–8] or by calculation using the nuclear databases [9, 10].

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

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

163 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  
164 thermal energy released in each fission and  $S_i(e_\nu)$  the neutrino flux per fission for this isotope. Using  
165 this method, the flux uncertainty is expected to be of an order of 2-3 % [11].

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

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

174 **Background in the neutrinos reactor spectrum**

175 Considering the close reactor neutrinos flux as the main signal, the signals that are considered as  
 176 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.

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

- The signal coming from radioactive decay ( $\alpha$ ,  $\gamma$ ,  $\beta$ ) from natural radioactive isotopes in the  
 181 material of the detector.
- Cosmogenic event such as fast neutrons and activated isotopes induced by muons passing  
 182 through the detector, most notably the spallation on  $^{12}\text{C}$ .

183 All those events represent a non-negligable part of the spectrum as shown in figure 2.3.  
 184

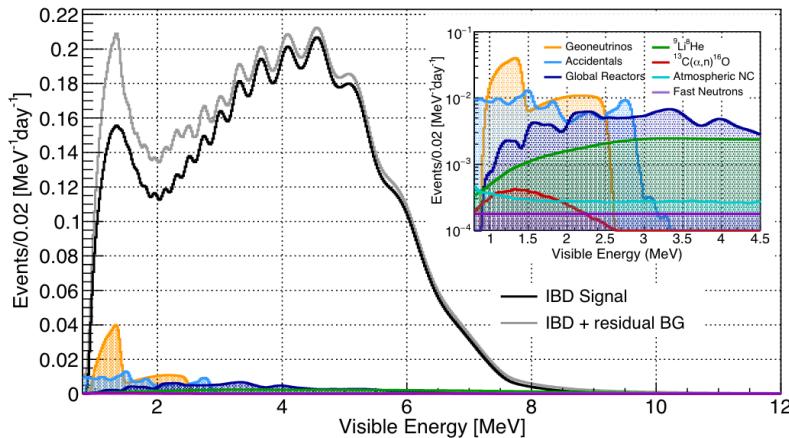


FIGURE 2.3 – Expected visible energy spectrum measured with the LPMT system with (grey) and without (black) backgrounds. The background amount for about 7% of the IBD candidate and are mostly localized below 3 MeV [11]

186 **Identification of the mass ordering**

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

193 **Precise measurement of the oscillations parameters**

194 The oscillations parameters  $\theta_{12}$ ,  $\theta_{13}$ ,  $\Delta m_{21}^2$ ,  $\Delta m_{31}^2$  are free parameters in the fit of the oscillation  
 195 spectrum. The precision on those parameters have been estimated and are shown in table 2.2. Wee  
 196 see that for  $\theta_{12}$ ,  $\Delta m_{21}^2$ ,  $\Delta m_{31}^2$ , precision at 6 years is better than the reference precision by an order of  
 197 magnitude [11]

	Central Value	PDG 2020	100 days	6 years	20 years
$\Delta m_{31}^2 (\times 10^{-3} \text{ eV}^2)$	2.5283	$\pm 0.034$ (1.3%)	$\pm 0.021$ (0.8%)	$\pm 0.0047$ (0.2%)	$\pm 0.0029$ (0.1%)
$\Delta m_{21}^2 (\times 10^{-3} \text{ eV}^2)$	7.53	$\pm 0.18$ (2.4%)	$\pm 0.074$ (1.0%)	$\pm 0.024$ (0.3%)	$\pm 0.017$ (0.2%)
$\sin^2 \theta_{12}$	0.307	$\pm 0.013$ (4.2%)	$\pm 0.0058$ (1.9%)	$\pm 0.0016$ (0.5%)	$\pm 0.0010$ (0.3%)
$\sin^2 \theta_{13}$	0.0218	$\pm 0.0007$ (3.2%)	$\pm 0.010$ (47.9%)	$\pm 0.0026$ (12.1%)	$\pm 0.0016$ (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 $\nu$
Atmospheric neutrino	hundreds per year for ${}^8\text{B}$	0–16 MeV	Radioactivity
Geoneutrino	hundreds per year	0.1–100 GeV	Negligible
	$\approx 400$ per year	0–3 MeV	Reactor $\nu$

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

### 202 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.

### 208 Atmospheric neutrinos

Atmospheric neutrinos are neutrinos originating from the decay of  $\pi$  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].

### 216 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-

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

#### 224 Diffuse supernovae neutrinos background

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

#### 233 Beyond standard model neutrinos interactions

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

#### 242 Proton decay

243 Proton decay is a potential unobserved event where the proton decay by violating the baryon num-  
 244 ber. This violation is necessary to explain the baryon asymmetry in the universe and is predicted  
 245 by multiple Grand Unified Theories which unify the strong, weak and electromagnetic interactions.  
 246 Thanks to its large active volume, JUNO will be able to take measurement of the potential proton  
 247 decay channel  $p \rightarrow \bar{\nu}K^+$  [18] thanks to the timing resolution of the SPMT system. Studies show  
 248 that JUNO should be competitive with the current best limit at  $5.9 \times 10^{33}$  years from Super-K. This  
 249 studies show that JUNO, considering no proton decay events observed, would be able to rules a  
 250 limit of  $9.6 \times 10^{33}$  years at 90 % C.L.

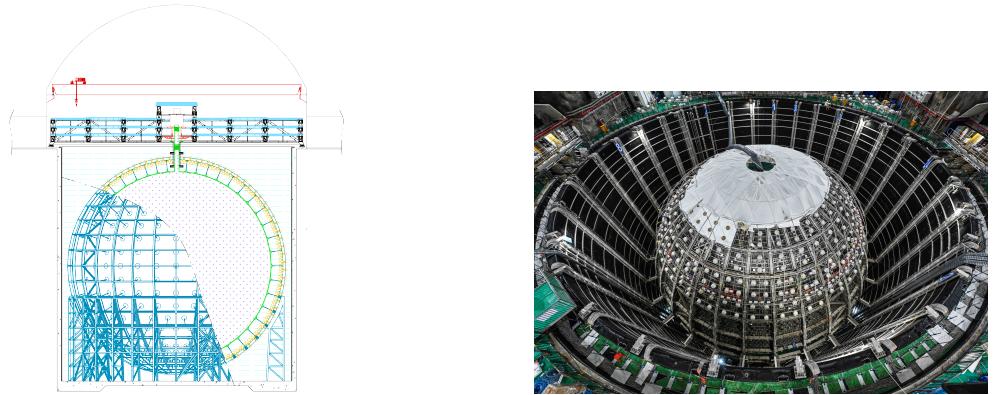
## 251 2.2 The JUNO detector

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

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

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

- Its 20 ktons monolithic LS provide a volume sizeable enough, in combination with the expected  $\bar{\nu}_e$  flux, to reach the desired statistic in 6 years. Its monolithic nature also allow for a full containment of most of the events, preventing the energy loss in non-instrumented parts that would arise from a segmented detector.
- Its large overburden shield it from most of the atmospheric background that would pollute the signal.
- The localization of the experiment, chosen to maximize the disappearance with a 53km baseline and in a region that allow two nuclear power plant to be used as sources.



(A) Schematics view of the JUNO detector.

(B) Top down view of the JUNO detector under construction

FIGURE 2.4

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

### 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  $^{12}\text{C}$  and H nucleus in the LS:

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

Kinematics calculation shows that this interaction has an energy threshold for the  $\bar{\nu}_e$  of  $(m_n + m_e - m_p) \approx 1.806$  MeV [19]. This threshold make the experiment blind to very low energy neutrinos. The residual energy  $E_\nu - 1.806$  MeV is be distributed as kinetic energy between the positron and the neutron. The energy of the emitted positron  $E_e$  is given by [19]

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

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 equation that the positron energy is strongly correlated to the neutrino energy.

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

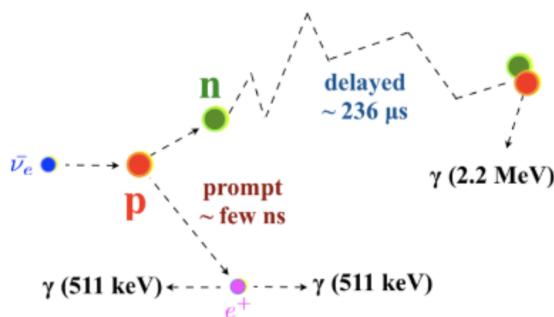


FIGURE 2.5 – Schematics of an IBD interaction in the central detector of JUNO

The scintillation photons have frequency in the UV and will propagate in the LS, being re-absorbed and re-emitted by compton effect before finally be captured by PMTs instrumenting the acrylic sphere. The analog signal of the PMTs digitized by the electronic is the signal of our experiment. The signal produced by the positron is subsequently called the prompt signal, and the signal coming from the neutron the delayed signal. This naming convention come from the fact that the positron will deposit its energy rather quickly (few ns) where the neutron will take a bit more time ( $\sim 236 \mu\text{s}$ ).

## 2.2.2 Central Detector (CD)

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

### Acrylic vessel

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

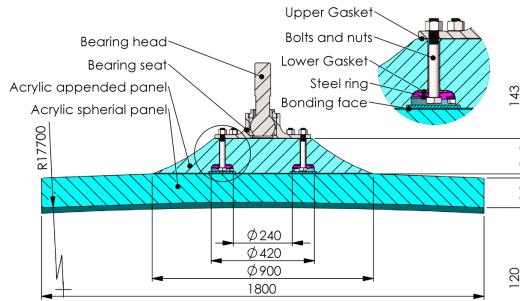


FIGURE 2.6 – Schematics of the supporting node for the acrylic vessel

312 **Liquid scintillator**

313 The Liquid Scintillator (LS) has a similar recipe as the one used in Daya Bay [20] but without gadolinium  
 314 doping. It is made of three components, necessary to shift the wavelength of emitted photons to  
 315 prevent their reabsorption and to shift their wavelength to the PMT sensitivity region as illustrated  
 316 in figure 2.7:

- 317 1. The detection medium, the *linear alkylbenzene* (LAB). Selected because of its excellent trans-  
 318 parency, high flash point, low chemical reactivity and good light yield. Accounting for  $\sim$   
 319 98% of the LS, it is the main component with which ionizing particles and gamma interact.  
 320 Charged particles will collide with its electronic cloud transferring energy to the molecules,  
 321 gamma will interact via compton effect with the electronic cloud before finally be absorbed  
 322 via photoelectric effect.
- 323 2. The second component of the LS is the *2,5-diphenyloxazole* (PPO). A fraction of the excitation  
 324 energy of the LAB is transferred to the PPO, mainly via non radiative process [21]. The  
 325 PPO molecules de-excites in the same way, transferring their energy to the bis-MSB. The PPO  
 326 makes for 1.5 % of the LS.
- 327 3. The last component is the *p-bis(o-methylstyryl)-benzene* (bis-MSB). Once excited by the PPO, it  
 328 will emit photon with an average wavelength of  $\sim$  430 nm (full spectrum in figure 2.7) that  
 329 can thus be detected by our photo-multipliers systems. It amount for  $\sim$  0.5% of the LS.

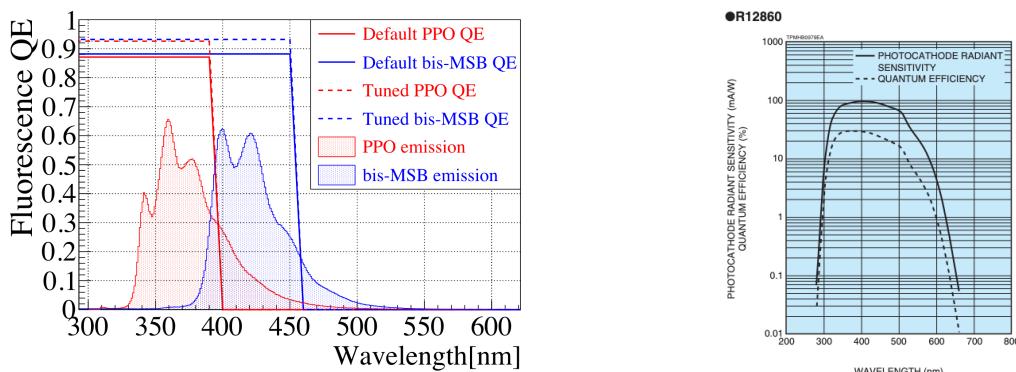


FIGURE 2.7 – On the left: Quantum efficiency (QE) and emission spectrum of the LAB and the bis-MSB [20]. On the right: Sensitivity of the Hamamatsu LPMT depending on the wavelength of the incident photons [22].

330 This formula has been optimized using dedicated studies with a Daya Bay detector [20, 23] to reach  
 331 the requirements for the JUNO experiment:

- 332 — A light yield / MeV of the amount of  $10^4$  photons to maximize the statistic in the energy  
 333 measurement.

- An attenuation length comparable to the size of the detector to prevent losing photons during their propagation in the LS. The final attenuation length is 25.8m [24] to compare with the CD diameter of 35.4m.
- Uranium/Thorium radiopurity to prevent background signal. The reactor neutrino program require a contamination fraction  $F < 10^{-15}$  while the solar neutrino program require  $F < 10^{-17}$ .

The LS will frequently be purified and tested in the Online Scintillator Internal Radioactivity Investigation System (OSIRIS) [25] to ensure that the requirements are kept during the lifetime of the experiment, more details to be found in section 2.4.2.

### 343 Large Photo-Multipliers Tubes (LPMTs)

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

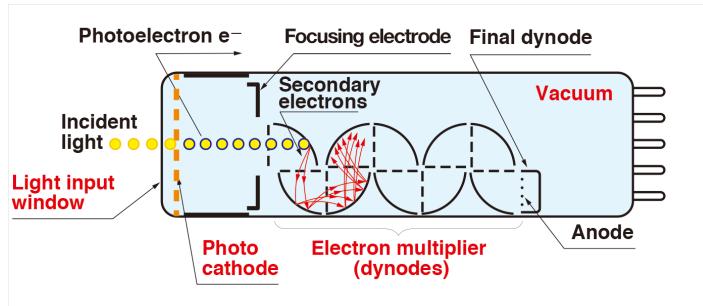


FIGURE 2.8 – Schematic of a PMT

The Large Photo-Multipliers Tubes (LPMT), used in the central detector and in the water pool, are 20-inch (50.8 cm) radius PMTs.  $\sim 5000$  dynode-PMTs [22] were produced by the Hamamatsu<sup>®</sup> company and  $\sim 15000$  Micro-Channel Plate (MCP) [26] by the NNVT<sup>®</sup> company. This system is the one responsible for the energy measurement with a energy resolution of  $3\%/\sqrt{E}$ , resolution necessary for the mass ordering measurement. To reach this precision, the system is composed of 17612 PMTs quasi uniformly distributed over the detector for a coverage of 75.2% reaching  $\sim 1800$  PE/MeV or  $\sim 2.3\%$  resolution due to statistic, leaving  $\sim 0.7\%$  for the systematic uncertainties. They are located outside the acrylic sphere in the water pool facing the center of the detector. To maintain the resolution over the lifetime of the experiment, JUNO require a failure rate  $< 1\%$  over 6 years.

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

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

Before performing analysis, the analog readout of the LPMT need to be amplified, digitised and packaged by the readout electronics schematized in figure 2.9. This electronic is splitted in two parts: *wet* electronic that are located near the LPMTs, protected in an Underwater Box (UWB) and the *dry* electronics located in deicated rooms outside of the water pool.

370 The LPMTs are connected to the UWB by groups of three. Each UWB contains:

- 371 — Three high voltage units, each one powering a PMT.
- 372 — A global control unit, responsible for the digitization of the waveform, composed of six analog-digital units that produce digitized waveform and a Field Programmable Gate Array (FPGA)
- 373 — that complete the waveform with metadatas such as the local timestamp trigger, etc... This
- 374 — FPGA also act as a data buffer when needed by the DAQ and trigger system.
- 375 — Additional memory in order to temporally store the data in case of sudden burst of the input
- 376 — rate (such as in the case of nearby supernovae).
- 377

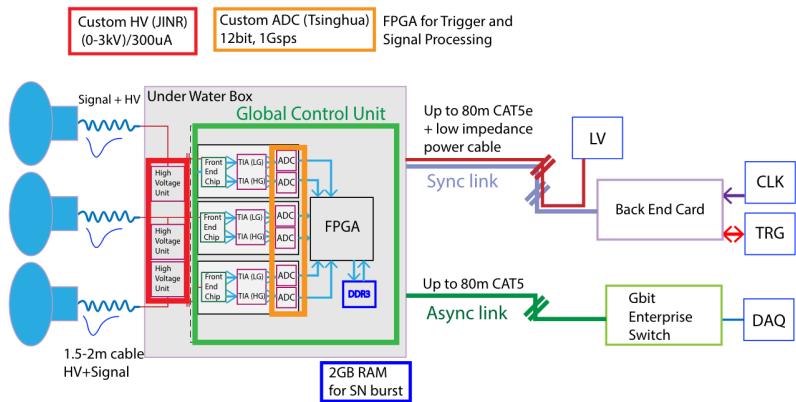


FIGURE 2.9 – The LPMT electronics scheme. It is composed of two part, the *wet* electronics on the left, located underwater and the *dry* electronics on the right. They are connected by Ethernet cable for data transmission and a dedicated low impedance cable for power distribution

378 The *dry* electronic synchronize the signals from the UWBs abd centralise the information of the CD  
379 LPMTs. It act as the Global Trigger by sending the UWB data to DAQ in the case if the LPMT  
380 multiplicity condition is fulfilled.

### 381 Small Photo-Multipliers Tubes (SPMTs)

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

388 The benefit of this second system is to be able to perform another, independent measure of the  
389 same events as the LPMTs, constituting the Dual Calorimetry useful for calibrationa and, as it we  
390 will explore in this thesis, for physics analysis. Due to the low PE rate, SPMTs will be running in  
391 photo-counting mode in the reactor range and thus will be insensitive to LPMT intrinsic effect (see  
392 section 2.3). Using this property, the intrinsic charge non linearity of the LPMTs can be measured by  
393 comparing the PE count in the SPMTs and LPMTs [27]. Also, due to their smaller size and electronics,  
394 SPMTs have a better timing resolutions than the LPMTs. At higher energy range, like supernovae  
395 events, LPMTs will saturate where SPMTs due to their lower PE collection will to produce a reliable  
396 measure of the energy spectrum.

397 The SPMTs will be grouped by pack of 128 to an UWB hosting their electronics as illustrated in figure  
398 2.10. This underwater box host two high voltage splitter boards, each one supplying 64 SPMTs, an

399 ASIC Battery Card (ABC) and a global control unit.

400 The ABC board will readout and digitize the charge and time of the 128 SPMTs signals and a FPGA  
 401 will joint the different metadata. The global control unit will handle the powering and control of the  
 402 board and will be in charge of the transmission of the data to the DAQ.

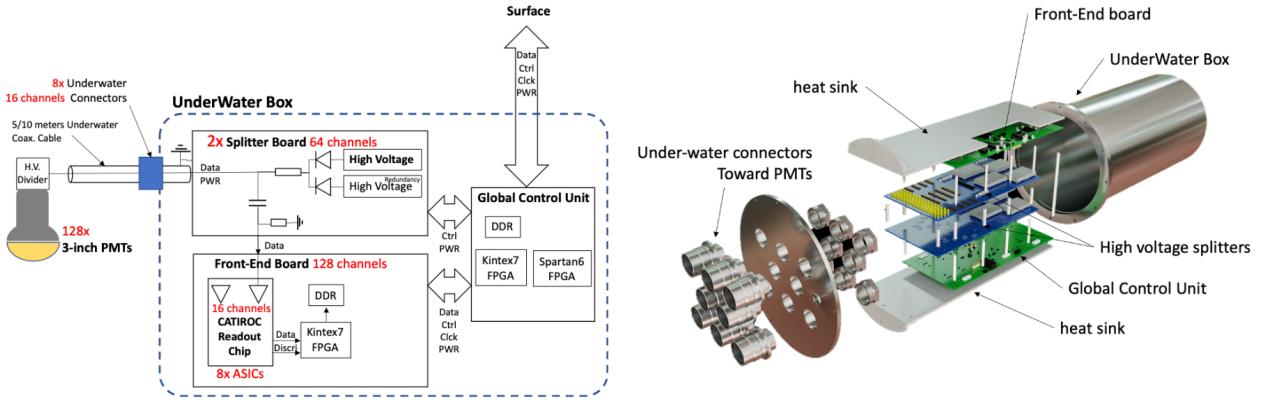


FIGURE 2.10 – Schematic of the JUNO SPMT electronic system (left), and exploded view of the main component of the UWB (right)

### 403 2.2.3 Veto detector

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

#### 408 Cherenkov in water pool

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

#### 417 Top tracker

418 The JUNO Top Tracker (TT) is a plastic scintillator detector located on the top of the experiment (see  
 419 figure 2.11). Made from plastic scintillator from OPERA [28] layered horizontally in 3 layers on the  
 420 top of the detector, the TT will be able to detect incoming atmospheric muons. With its coverage,  
 421 about 1/3 of the of all atmospheric muons that passing through the CD will also pass through the 3  
 422 layer of the detector. While it does not cover the majority of the CD, the TT is particularly effective  
 423 to detect muons coming through the filling chimney region which might present difficulties from the  
 424 other subsystems in some classes of events.

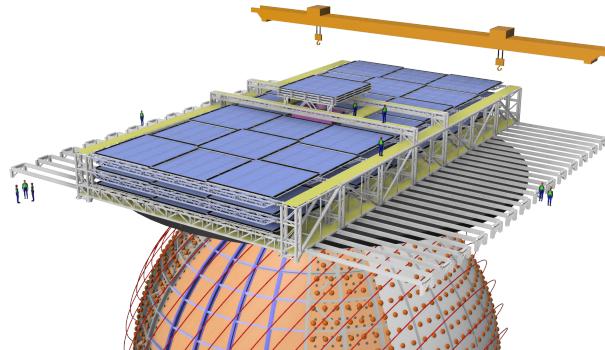


FIGURE 2.11 – The JUNO top tracker

## 425 2.3 Calibration strategy

426 The calibration is a crucial part of the JUNO experiment. The detector will continuously bath in  
 427 neutrinos coming from the close nuclear power plant, from other sources such as geo neutrinos,  
 428 the sun and will be exposed to background noise coming from atmospheric muons and natural  
 429 radioactivity. Because of this continuous rate, low frequency signal event, we need high frequency,  
 430 recognisable sources in the energy range of interest : [0-12] MeV for the positron signal and 2.2 MeV  
 431 for the neutron capture. It is expected that the CD response will be different depending on the type  
 432 of particle, due to the interaction with LS, the position on the event and the optical response of the  
 433 acrylic sphere (see section 2.6). We also expect a non-linear energy response of the CD due to the LS  
 434 properties [20] but also due to the saturation of the LPMTs system when collecting a large amount of  
 435 PE [27].

### 436 2.3.1 Energy scale calibration

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

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

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

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

445 From this calibration we call  $E_{vis}$  the "visible energy" that is reconstructed by our current algorithms  
 446 and we compare it to the true energy deposited by the calibration source. The results shown in figure  
 447 2.12 show the expected response of the detector from calibration sources. The non-linearity is clearly  
 448 visible from the  $E_{vis} / E_{true}$  shape. See [29] for more details.

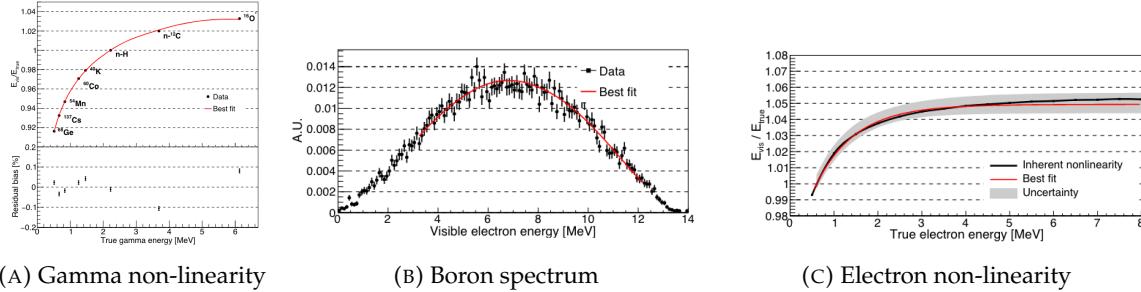


FIGURE 2.12 – Fitted and simulated non linearity of gamma, electron sources and from the  $^{12}\text{B}$  spectrum. Black points are simulated data. Red curves are the best fits. Figures taken from [29].

### 449 2.3.2 Calibration system

450 The non-uniformity due to the event position in the detector (more details in section 2.6) will be  
 451 studied using multiples systems that are schematized in figure 2.13. They allow to position sources  
 452 at different location in the CD.

- 453 — For a one-dimension vertical calibration, the Automatic Calibration Unit (ACU) will be able  
 454 to deploy multiple radioactive sources or a pulse laser diffuser ball along the central axis of  
 455 the CD through the top chimney. The source position precision is less than 1cm.
- 456 — For off-axis calibration, a calibration source attached to a Cable Loop System (CLS) can be  
 457 moved on a vertical half-plane by adjusting the length of two connection cable. Two set of  
 458 CSL will be deployed to provide a 79% effective coverage of a vertical plane.
- 459 — A Guiding Tube (GT) will surround the CD to calibrate the non-uniformity of the response at  
 460 the edge of the detector
- 461 — A Remotely Operated under-LS Vehicle (ROV) can be deployed to desired location inside LS  
 462 for a more precise and comprehensive calibration. The ROV will also be equipped with a  
 463 camera for inspection of the CD.

464 The preliminary calibration program is depicted in table 2.5.

### 465 2.3.3 Instrumental non-linearity calibration

466 As mentioned in the introduction of this section, we expect an instrumental non-linearity due to the  
 467 LPMT system saturating. This results in the LPMT underestimating the number of collected photo-  
 468 electrons. This non-linearity is illustrated in figure 2.14. This non-linearity would consequently  
 469 convolve with the LS non-linearity. To correct this effect, the LPMT are first calibrated to the channel  
 470 level using the dual calorimetry calibration technique which consist of comparing the LPMT and  
 471 SPMT calorimetry calibration using a tunable light source covering the range of 0 to 100 PE per  
 472 LPMT channel.

473 Within such range, the SPMT serve as an approximate linear reference since SPMT operate primarily  
 474 operate in photo-counting mode in this range. Using this technique, the residual non-linearity in the  
 475 LPMT response due to the saturation effect is under 0.3 %.

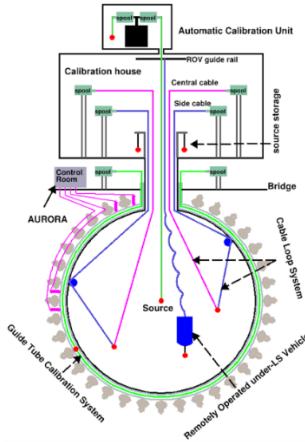


FIGURE 2.13 – Overview of the calibration system

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
	$^{68}\text{Ge}$	ACU	75
	$^{137}\text{Cs}$	ACU	75
	$^{54}\text{Mn}$	ACU	75
	$^{60}\text{Co}$	ACU	75
	$^{40}\text{K}$	ACU	158

TABLE 2.5 – Calibration program of the JUNO experiment

## 476 2.4 Satellite detectors

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

### 481 2.4.1 TAO

482 The Taishan Antineutrino Observatory (TAO) [12, 30] is a ton-level gadolinium doped liquid scin-  
 483 tillator detector that will be located near the Taishan-1 reactor. It aim to measure the  $\bar{\nu}_e$  spectrum at  
 484 very low distance (44m) from the reactor to measure a quasi-unoscillated spectrum. TAO also aim to  
 485 provide a major contribution to the so-called reactor anomaly [13]. Its requirement are to the level of  
 486 2 % energy resolution at 1 MeV.

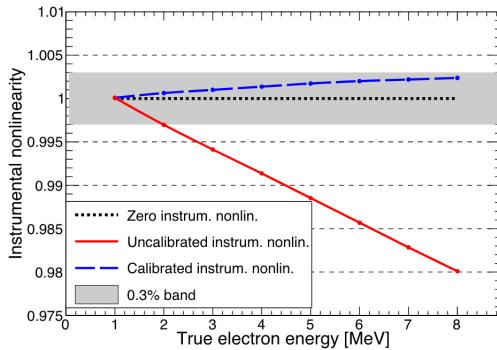


FIGURE 2.14 – Event-level instrumental non-linearity, defined as the ratio of the total measured LPMT charge to the true charge for events uniformly distributed in the detector. The solid red line represents event-level non-linearity without the channel-level correction, with position non-uniformity obtained at 1 MeV applied, in an extreme hypothetical scenario of 50% non-linearity over 100 PEs for the LPMTs. The dashed blue line represents that after the channel-level correction. The gray band shows the residual uncertainty of 0.3%, after the channel-level correction. Figure taken from [29].

#### 487      Detector

488      The TAO detector is close, in concept, to the CD of JUNO. It is composed of an acrylic vessel  
 489      containing 2.8 tons of gadolinium-loaded LS instrumented by an array of silicon photomultipliers  
 490      (SiPM) reaching a 95% coverage. To efficiently reduce the dark count of those sensors, the detector  
 491      is cooled to -50 °C. The  $\bar{\nu}_e$  will interact with the LS via IBD, producing scintillation light, that will  
 492      be detected by the SiPMs. From this signal the  $\bar{\nu}_e$  energy and the full spectrum reconstructed. This  
 493      spectrum will then be used by JUNO to calibrate the unoscillated spectrum, most notably the fission  
 494      product fraction that impact the rate and shape of the spectrum. A schema of the detector is presented  
 495      in figure 2.15a.

#### 496      2.4.2 OSIRIS

497      The Online Scintillator Internal Radioactivity Investigation System (OSIRIS) [25] is an ultralow back-  
 498      ground, 20 m<sup>3</sup> LS detector that will be located in JUNO cavern. It aim to monitor the radioactive  
 499      contamination, purity and overall response of the LS before it is injected in JUNO. OSIRIS will  
 500      be located at the end of the purification chain of JUNO, monitoring that the purified LS meet the  
 501      JUNO requirements. The setup is optimized to detect the fast coincidences decay of  $^{214}\text{Bi} - ^{214}\text{Po}$   
 502      and  $^{212}\text{Bi} - ^{212}\text{Po}$ , indicators of the decay chains of U and Th respectively.

#### 503      Detector

504      OSIRIS is composed of an acrylic vessel that will contains 17t of LS. The LS is instrumented by  
 505      a PMT array of 64 20 inch PMTs on the top and the side of the vessel. To reach the necessary  
 506      background level required by the LS purity measurements, in addition to being 700m underground  
 507      in the experiment cavern, the acrylic vessel is immersed in a tank of ultra pure water. The water is  
 508      itself instrumented by another array of 20 inch PMTs, acting as muon veto. A schema of the detector  
 509      is presented in figure 2.15b.

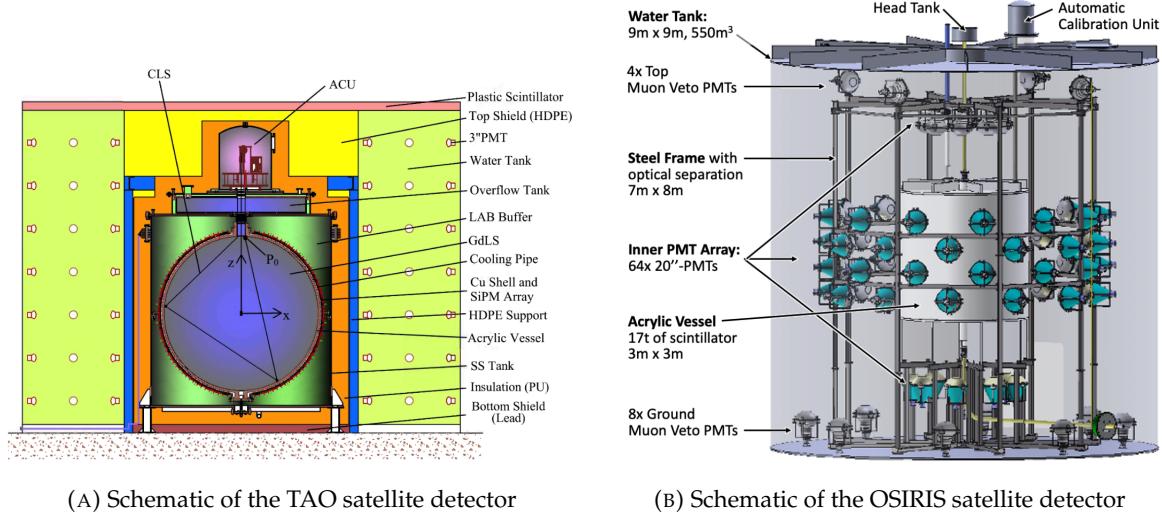


FIGURE 2.15

## 2.5 Software

510 The simulation, reconstruction and analysis algorithms are all packaged in the JUNO software,  
 511 subsequently called the software. It is composed of multiple components integrated in the SNiPER  
 512 [31] framework:

- 514 — Various primary particles simulators for the different kind of events, background and calibra-  
 515 tion sources.
- 516 — A Geant4 [32–34] Monte Carlo (MC) simulation containing the detectors geometries, a custom  
 517 optical model for the LS and the supporting structures of the detectors. The Geant4 simulation  
 518 integrate all relevant physics process for JUNO, validated by the collaboration. This step of the  
 519 simulation is commonly called *Detsim* and compute up to the production of photo-electrons  
 520 in the PMTs. The optics properties of the different materials and detector components have  
 521 been measured beforehand to be used to define the material and surfaces in the simulation.
- 522 — An electronic simulation, simulating the response waveform of the PMTs, tracking it through  
 523 the digitization process, accounting for effects such as non-linearity, dark noise, Time Trans-  
 524 it Spread (TTS), pre-pulsing, after-pulsing and ringing if the waveform. It's also the step  
 525 handling the event triggers and mixing. This step is commonly referenced as *Elecsim*.
- 526 — A waveform reconstruction where the digitized waveform are filtered to remove high-frequency  
 527 white noise and then deconvoluted to yield time and charge informations of the photons hits  
 528 on the PMTs. This step is commonly referenced as *Calib*.
- 529 — The charge and time informations are used by reconstruction algorithms to reconstruct the  
 530 interaction vertex and the deposited energy. This step is commonly reported as *Reco*. See  
 531 section 2.6 for more details on the reconstruction.
- 532 — Once the singular events are reconstructed, they go through event pairing and classification  
 533 to select IBD events. This step is named Event Classification.
- 534 — The purified signal is then analysed by the analysis framework which depend of the physics  
 535 topic of interest.

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

## 541 2.6 State of the art of the Offline IBD reconstruction in JUNO

542 The main reconstruction method currently run in JUNO is a data-driven method based on a like-  
 543 lihood maximization [35, 36] using only the LPMTs. The first step is to reconstruct the interaction  
 544 vertex from which the energy reconstruction is dependent. It is also necessary for event pairing and  
 545 classification.

### 546 2.6.1 Interaction vertex reconstruction

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

#### 549 Charge based algorithm

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

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

551 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  
 552 reconstructed interaction position.  $a$  is a scale factor introduced because a weighted average over  
 553 a 3D sphere is inherently biased. Using calibration we can estimate  $a \approx 1.3$  [37]. The results in  
 554 figure 2.16b shows that the reconstruction is biased from around 15m and further. This is due to the  
 555 phenomena called “total reflection area” or TR Area.

556 As depicted in the figure 2.16a the optical photons, given that they have a sufficiently large incidence  
 557 angle, can be deviated of their trajectories when passing through the interfaces LS-acrylic and water-  
 558 acrylic due to the optical index difference. This cause photons to be lost or to be detected by PMT  
 559 further than anticipated if we consider their rectilinear trajectories. This cause the charge barycenter  
 560 the be located closer to the center than the event really is.

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

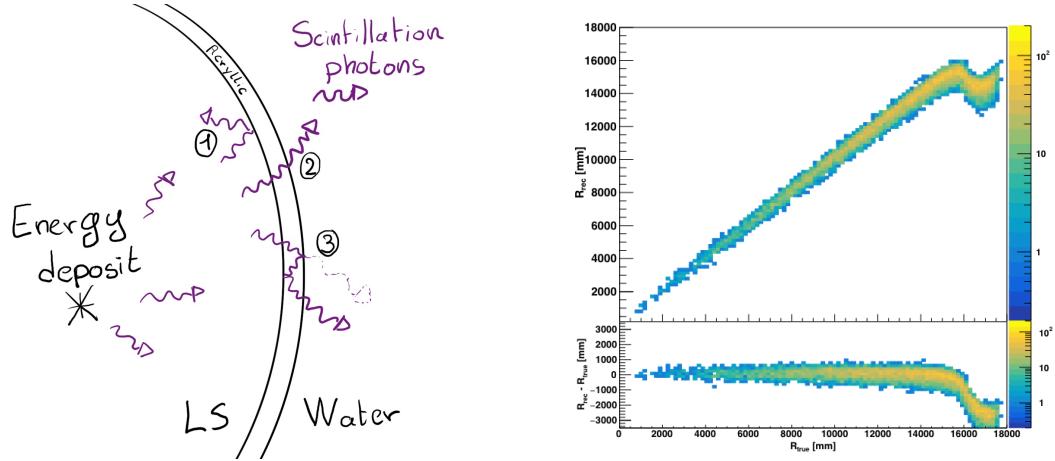
#### 564 Time based algorithm

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

- 567 1. Use the charge based algorithm to get an initial vertex to start the iteration.
- 568 2. Calculate the time of flight correction for the  $i$ th PMT using

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

569 where  $j$  is the iteration step,  $t_i$  is the timing of the  $i$ th PMT, and  $\text{tof}_i$  is the time-of-flight of the  
 570 photon considering an rectilinear trajectory and an effective velocity in the LS and water (see  
 571 [37] for detailed description of this effective velocity). Plot the  $\Delta t$  distribution and label the  
 572 peak position as  $\Delta t^{\text{peak}}$  (see fig 2.17a).



(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 2.16

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

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

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

578 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  
579 step 2.

580 However because the earliest arrival time is used,  $t_i$  is related to the number photoelectrons  $N_i^{pe}$   
581 detected by the PMT [38–40]. To reduce bias in the vertex reconstruction, the following equation is  
582 used to correct  $t_i$  into  $t'_i$ :

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

583 The parameters  $(p_0, p_1, p_2)$  were optimized to (9.42, 0.74, -4.60) for Hamamatsu PMTs and (41.31,  
584 -12.04, -20.02) for NNVT PMTs [37]. The results presented in figure 2.17b shows that the time based  
585 algorithm provide a more accurate vertex and is unbiased even in the TR area. This results  $(\vec{r}_0, t_0)$  is  
586 used as initial value for the likelihood algorithm.

### 587 Time likelihood algorithm

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

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

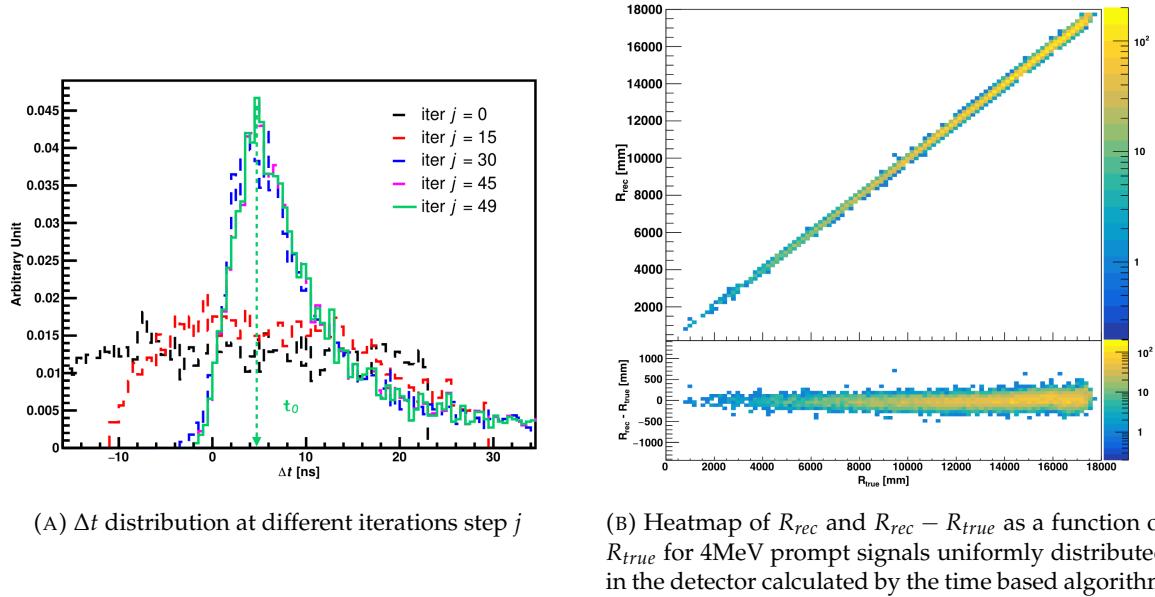


FIGURE 2.17

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

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

592 as the sum of the  $k$  component that emit light in the LS each one characterised by it's decay time  $\tau_k$   
 593 and intensity fraction  $\rho_k$ . The TTS component is expressed as a gaussian convolution

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

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

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

597 The distribution of the residual time  $t_{res}$  of an event can then be compared to  $p(t_{res})$  and the best  
 598 fitting vertex  $\vec{r}_0$  and  $t_0$  can be chosen by minimizing

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

599 The parameter of Eq. 2.10 can be measured experimentally. The results shown in figure 2.18 used  
 600 PDF from monte carlo simulation. The results shows that  $R_{rec} - R_{true}$  is biased depending on the  
 601 energy. While this could be corrected using calibration, another algorithm based on charge likelihood  
 602 was developed to correct this problem.

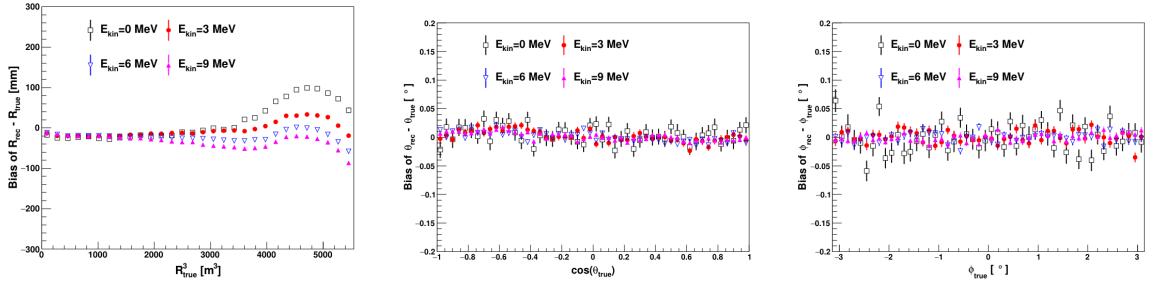


FIGURE 2.18 – Bias of the reconstructed radius  $R$  (left),  $\theta$  (middle) and  $\phi$  (right) for multiple energies by the time likelihood algorithm

### 603 Charge likelihood algorithm

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

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

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

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

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

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

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

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

613 where  $Y$  is the energy scale factor,  $\Omega(\vec{r}_0, r_i)$  is the solid angle of the  $i$ th PMT,  $\epsilon_i$  is its detection  
 614 efficiency,  $f(\theta_i)$  its angular response,  $\zeta_m$  is the attenuation length in the materials and  $\delta_i$  the expected  
 615 number of dark noise.

616 However Eq. 2.14 assume that the scintillation light yield is linear with energy and describe poorly  
 617 the contribution of indirect light, shadow effect due to the supporting structure and the total reflec-  
 618 tion effects. The solution is to use data driven methods to produce the pdf by using the calibra-  
 619 tions sources and position described in section 2.3. In the results presented in figures 2.19, the PDF was  
 620 produced using MC simulation and 29 specific calibrations position [37] along the Z-axis of the  
 621 detector. We see that the charge likelihood algorithm show little bias in the TR area and a better  
 622 resolution than the time likelihood. The figure 2.20 shows the radial resolution of the different  
 623 algorithm presented for this section, we can see the refinement at each step and that the charge  
 624 likelihood yield the best results.

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

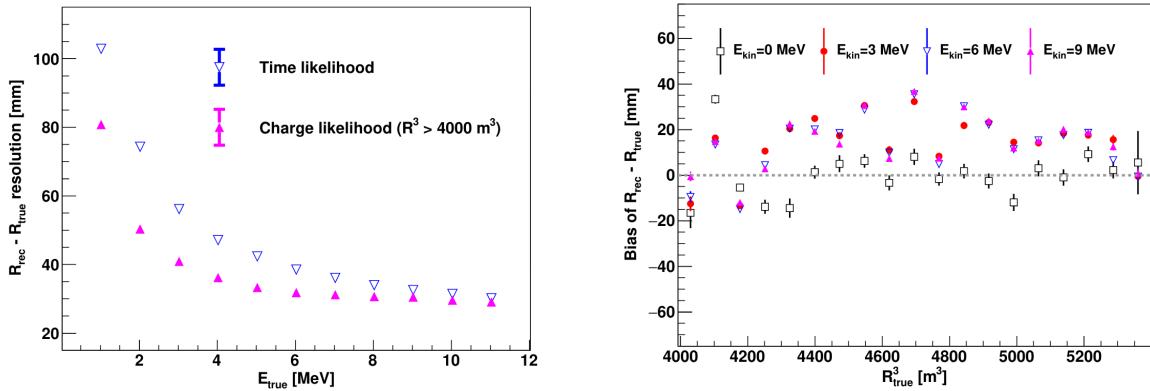


FIGURE 2.19 – On the left: Resolution of the reconstructed  $R$  as a function of the energy in the TR area ( $R^3 > 4000 \text{ m}^3 \equiv R > 16 \text{ m}$ ) by the charge and time likelihood algorithms. On the right: Bias of the reconstructed  $R$  in the TR area for different energies by the charge likelihood algorithm

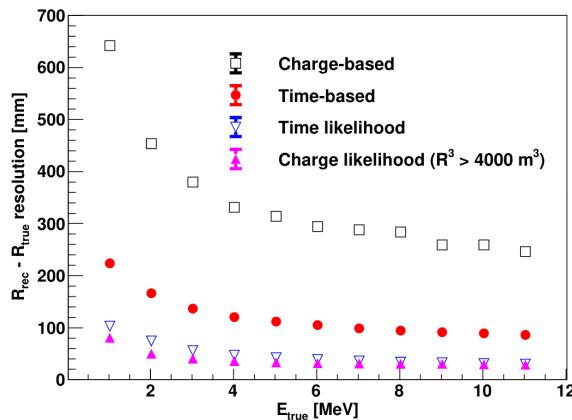


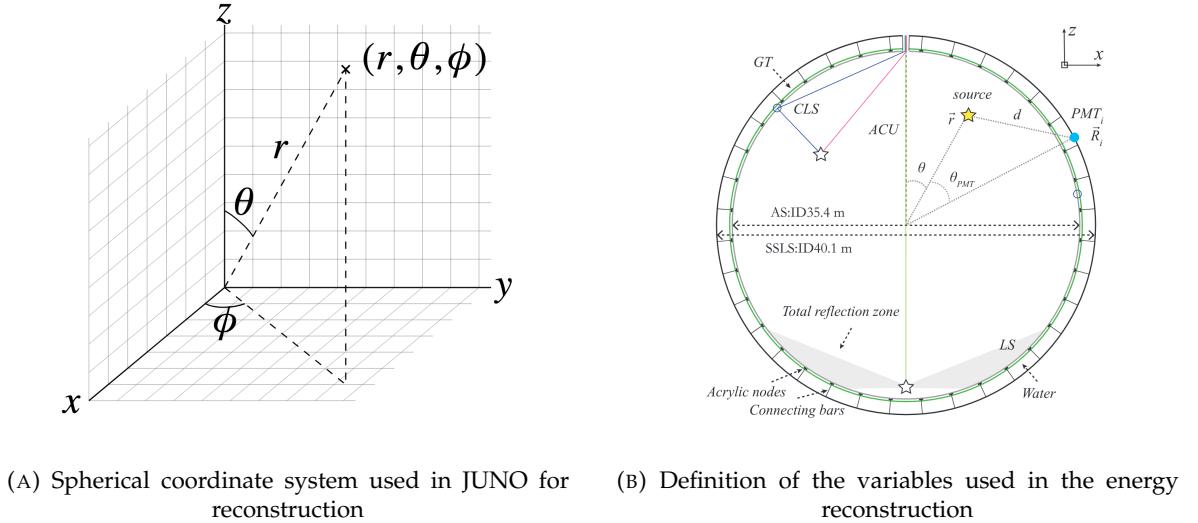
FIGURE 2.20 – Radial resolution of the different vertex reconstruction algorithms as a function of the energy

## 627 2.6.2 Energy reconstruction

628 As explained in section 2.1.1, energy resolution is crucial for the NMO and oscillation parameters  
 629 measurements. Thus the energy reconstruction algorithm should take into consideration as much  
 630 detector effect as possible. The following method is a data driven method based on calibration  
 631 samples inspired by the charge likelihood algorithm described above [41].

### 632 Charge estimation

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



(A) Spherical coordinate system used in JUNO for reconstruction

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

FIGURE 2.21

638 using  $(r, \theta, \theta_{pmt})$  as defined in figure 2.21b.

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

639 where  $i$  runs over the PMTs with the same  $\theta_{pmt}$ ,  $\text{DE}_i$  is the detection efficiency of the  $i$ th PMT.  $\mu_i^D$   
640 is the expected number of dark noise photoelectrons in the time window  $L$ . The time window have  
641 been optimized to  $L = 280$  ns [41].  $\bar{q}_i$  is the average recorded photoelectrons in the time window  
642 and  $\hat{Q}_i$  is the expected average charge for 1 photoelectron. The  $N_{pe}$  map is constructed following the  
643 procedure described in [36].

#### 644 Time estimation

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

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

647 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  
648 by a gaussian

$$t_{TT} = \mathcal{N}(\bar{\mu}_{TT} + t_d, \sigma_{TT}) \quad (2.17)$$

649 where  $\mu_{TT}$  is the mean transit time in PMTs,  $\sigma_{TT}$  is the Transit Time Spread (TTS) of the PMTs and  $t_d$   
650 is the delay time in the electronics. The effective refraction index of the LS is also corrected to take  
651 into account the propagation distance in the detector.

652 The timing PDF  $P_T(t_r | r, d, \mu_l, \mu_d, k)$  can now be generated using calibration sources [41]. This PDF  
653 describe the probability that the residual time of the first photon hit is in  $[t_r, t_r + \delta]$  with  $r$  the radius  
654 of the event vertex,  $d = |\vec{r} - \vec{r}_{PMT}|$  the propagation distance,  $\mu_l$  and  $\mu_d$  the expected number of PE  
655 and dark noise in the electronic reading window and  $k$  is the detected number of PE.

656 Now let denote  $f(t, r, d)$  the probability density function of "photoelectron hit a time t" for an event

657 happening at  $r$  where the photons traveled the distance  $d$  in the LS

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

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

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

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

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

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

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

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

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

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

667 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  
668 charge pdf for  $k$  PE. And we can also construct the time likelihood referred as TMLE:

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

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

670 Merging those two likelihood give the charge-time likelihood QTMLE

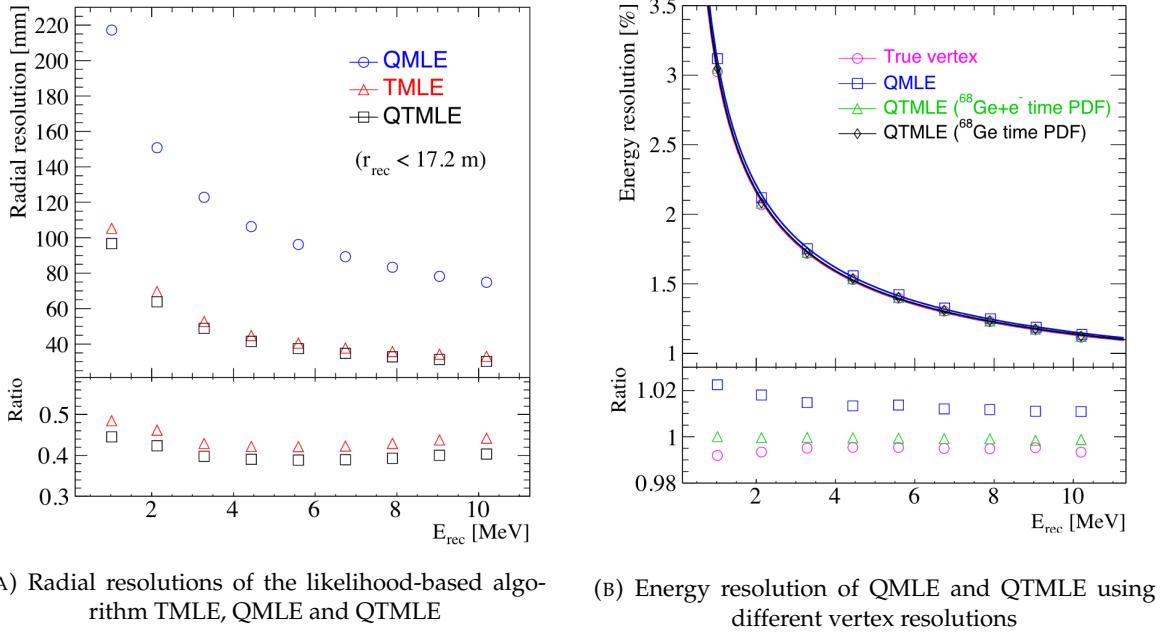
$$\mathcal{L}(q_1, q_2, \dots, q_N; t_{1,r}, t_{2,r}, \dots, t_{N,r} | \vec{r}, t_0, E_{vis}) = \mathcal{L}(q_1, q_2, \dots, q_N | \vec{r}, E_{vis}) \cdot \mathcal{L}(t_{1,r}, t_{2,r}, \dots, t_{N,r} | \vec{r}, t_0) \quad (2.24)$$

671 The radial and energy resolutions of the different likelihood are presented in figure 2.22 (from [41]).  
672 We can see the improvement of adding the time information to the vertex reconstruction and that  
673 an increase in vertex precision can bring improvement in the energy resolution, especially at low  
674 energies.

675 Data driven methods prove to be performant in the energy and vertex reconstruction given that we  
676 have enough calibrations sources to produce the PDF. In the next section, we'll see another type of  
677 data-driven method based on machine learning.

### 678 2.6.3 Machine learning for reconstruction

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



(A) Radial resolutions of the likelihood-based algorithm TMLE, QMLE and QTML

(B) Energy resolution of QMLE and QTML using different vertex resolutions

FIGURE 2.22

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

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

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

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

### 697 Boosted Decision Tree (BDT)

On of the most classic ML method used in physics in last years is the Boosted Decision Tree (see chapter 3.1). They have been explored for vertex reconstruction [42] et for energy reconstruction [42, 43].

For vertex and energy reconstruction a BDT was developed using the aggregated informations presented in 2.6.

Its reconstruction performances are presented in figure 2.24.

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

AccumCharge	$ht_{5\% - 2\%}$
$R_{cht}$	$pe_{mean}$
$z_{cc}$	$J_{cht}$
$pe_{std}$	$\phi_{cc}$
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

704 A second and more advanced BDT, subsequently named BDTE, that only reconstruct energy use a  
 705 different set of features [43]. They are presented in the table 2.7

## 706 Neural Network (NN)

707 The physics have shown a rising for Neural Network (NN) in the past years for event reconstruction,  
 708 notably in the neutrino community [44–47]. Three type of neural networks have explored for event  
 709 reconstruction in JUNO Deep Neural Network (DNN), Convolutional Neural Network (CNN) and  
 710 Graph Network (GNN). More explanation about those neural network can be found in chapter 3.

711 The CNN are using 2D projection of the detector representing it as an image with two channel, one  
 712 for the charge  $Q$  and one for the time  $t$ . The position of the PMTs is structurally encoded in the pixel  
 713 containing the information of this PMT. In [42], the pixel is chosen based on a transformation of  $\theta$   
 714 and  $\phi$  coordinates to the 2D plane and rounded to the nearest pixel. A sufficiently large image has  
 715 been chosen to prevent two PMT to be located in the same pixel. An example of this projection can  
 716 be found in figure 2.23. The performances of the CNN can be found in figure 2.24.

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

721 One of the way to present structurally the sphericity of JUNO to a NN is to use a graph: A collection  
 722 of objects  $V$  called nodes and relations  $E$  called edges, each relation associated to a couple  $v_1, v_2$   
 723 forming the graph  $G(E, V)$ . Nodes and edges can hold informations or features. In [42] the nodes,  
 724 are geometrical region of the detector as defined by the HealPix [48]. The features of the nodes are  
 725 aggregated informations from the PMTs it contains. The edges contains geographic informations of  
 726 the nodes relative positions.

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

730 The neural network then process the graph using Chebyshev Convolutions [49]. The performances  
 731 of the GNN are presented in figure 2.24.

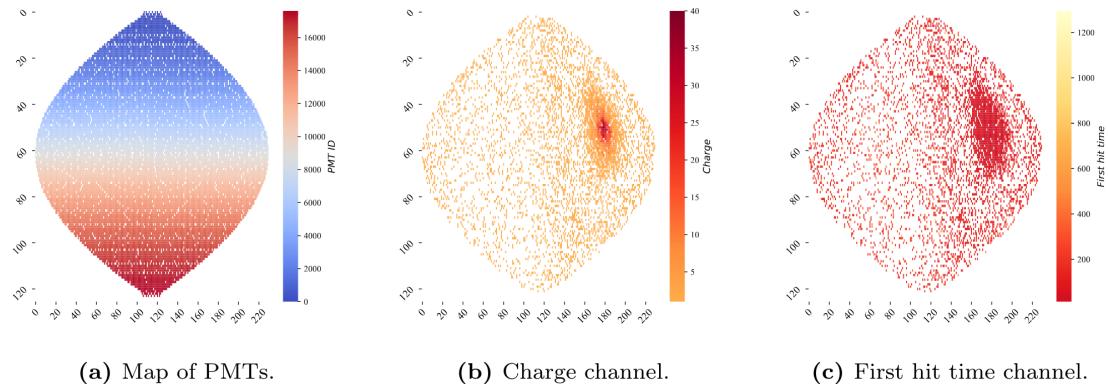


FIGURE 2.23 – Projection of the LPMTs in JUNO on a 2D plane. (a) Show the distribution of all PMTs and (b) and (c) are example of what the charge and time channel looks like respectively

Overall ML algorithms show similar performances as classical algorithms in term of energy reconstructions 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 produced by the calibration, the input from TAO and adjusted Monte Carlo simulations:

- The absolute flux and the fission product fraction yield 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

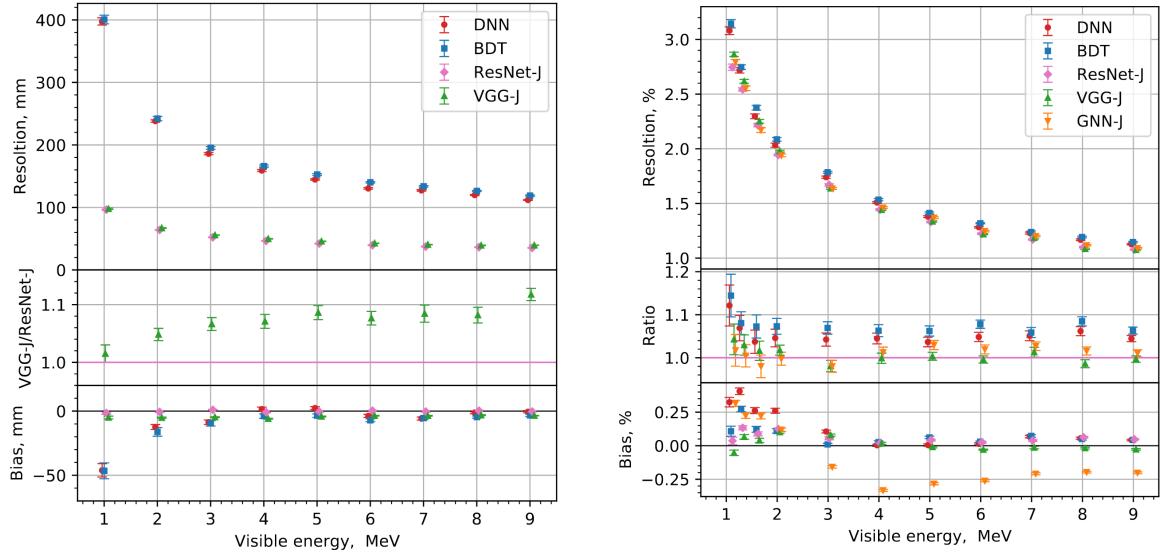


FIGURE 2.24 – Radial (left) and energy (right) resolutions of different ML algorithms. The results presented here are from [42]. DNN is a deep neural network, BDT is a BDT, ResNet-J and VGG-J are CNN and GNN-J is a GNN.

interest  $\theta_{13}, \theta_{12}, \Delta m_{21}^2, \Delta m_{31}^2$ . Noise parameters can be included in the parameters spectrum such as the earth density  $\rho$  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  $\chi^2$  minimization where  $\chi^2$  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  $\sigma_i$  is the uncertainty of this bin. Two classic statistic test exist Pearson and Neyman where the difference is the estimation of  $\sigma_i$  parameters.

This  $\sigma_i$  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 ration  $\frac{1}{N_{th}^i}$  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

<sup>779</sup> statistic test.

<sup>780</sup> The  $\chi^2$  is minimized by exploring the parameter phase space via gradient descent.

<sup>781</sup> **2.7.3 Physics results**

<sup>782</sup> The oscillation parameters are directly extracted from the minimization procedure and the error can  
<sup>783</sup> be estimated directly from the procedure. For the NMO, the data are fitted under the two assumption  
<sup>784</sup> of NO and IO. The difference in  $\chi^2$  give us the preferred ordering and the significance of our test.  
<sup>785</sup> Latest studies show that the precision on oscillation parameters after six year of data taking will be  
<sup>786</sup> of 0.2%, 0.3%, 0.5% and 12.1% for  $\Delta m_{31}^2$ ,  $\Delta m_{21}^2$ ,  $\sin^2 \theta_{12}$  and  $\sin^2 \theta_{13}$  respectively [11]. The expected  
<sup>787</sup> sensitivity to mass ordering is  $3\sigma$  after 6.5 years [50].

<sup>788</sup> **2.8 Summary**

<sup>789</sup> JUNO is one the biggest new generation neutrino experiment. Its goal, the measurements of oscil-  
<sup>790</sup> lation parameters with unprecedented precision and an NMO preference at the 3 sigma confidence  
<sup>791</sup> level, needs an in depth knowledge and understanding of the detector and the physics at hand. The  
<sup>792</sup> characterisation and calibration of the detector are of the utmost importance and the understanding  
<sup>793</sup> of the detector response in its resolution and bias is capital to be able to correctly carry the high  
<sup>794</sup> precision physics analysis of the neutrino oscillation.

<sup>795</sup> In this thesis, I explore the usage of data-driven reconstruction methods to validate and optimize the  
<sup>796</sup> reconstruction of IBD events in JUNO in the chapters 4, 5 and 6 and the usage of the dual calorimetry  
<sup>797</sup> in the detection of possible mis-modelisation in the theoretical spectrum 7.



<sup>798</sup> **Chapter 3**

<sup>799</sup> **Machine learning and Artificial  
Neural Network**

<sup>801</sup>

*"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*

<sup>802</sup> Machine Learning (ML) and more specifically Neural Network (NN) are families of data-driven <sup>803</sup> algorithms. They are used to model complex distributions from a finite dataset to extract a generalist <sup>804</sup> behavior. They learn, adapt their intrinsic parameters, interactively by computing their performances <sup>805</sup> or loss on those datasets. They take advantage of simple microscopic operations such as *if condition* <sup>806</sup> or non-continuous but differentiable function like *ReLU* in heavy numbers to model macroscopic <sup>807</sup> complex and precise behaviours.

<sup>808</sup> They are now widely used in a wide variety of domain including natural language processing, <sup>809</sup> computer vision, speech recognition and, the subject of this thesis, scientific studies.

<sup>810</sup> We found them in particle physics, either as the main algorithm or as secondary algorithm, for event <sup>811</sup> reconstruction, event classification, waveform reconstruction, etc..., domains where the underlying <sup>812</sup> physic and detector processes are complex and highly dimensional. Physicists have traditionally <sup>813</sup> been forced to use simplifications or assumptions to ease the development of algorithms or equations <sup>814</sup> (a good example is the algorithm presented in section 2.6) where machine learning could refine <sup>815</sup> and take into account those effects, provided that they have enough data and computing power.

<sup>816</sup> This chapter present an overview of the different kind of machine learning methods and neural <sup>817</sup> networks that will be discussed in this thesis.

<sup>818</sup> **3.1 Boosted Decision Tree (BDT)**

<sup>819</sup> One of the most classic machine learning algorithm used in particle physics is Boosted Decision Tree <sup>820</sup> (BDT) [51] (or more recently Gradient Boosting Machine [52]). The principle of a BDT is fairly simple <sup>821</sup> : based on a set of observables, a serie of decisions, represented as node in a tree, are taken by the <sup>822</sup> algorithm. Each decision point, or node, takes its decision based on a set of trainable parameters <sup>823</sup> leading to a subtree of decisions. The process is repeated until it reach the final node, yielding the <sup>824</sup> prediction. A simplistic example is given in figure 3.1.

<sup>825</sup> The training procedure follow a simple score reward procedure. During the training phase the <sup>826</sup> prediction of the BDT is compared to a known truth about the data. The score is then used to

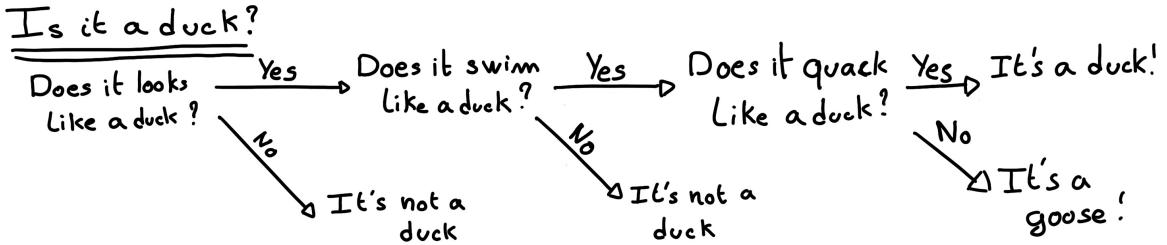


FIGURE 3.1 – Example of a BDT that determine if the given object is a duck

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

## 830 3.2 Artificial Neural Network (NN)

831 One other big family of machine learning algorithm is the artificial Neural Networks (NN). The idea  
 832 of developing automates which component mimic, in a simplistic way, the behavior of biological  
 833 neurons emerge in 1959 with the paper “*What the Frog’s Eye Tells the Frog’s Brain*” [53]. They develop  
 834 an automate where each component possess an *activation function*. Each one of those components  
 835 then transmit its information to the other following a certain efficiency or *weight*. Those works  
 836 influenced scientist and notably Frank Rosenblatt who published in 1958 what is considered the  
 837 first neural network model the Perceptron [54].

838 Modern neural network still nowadays use the neuron metaphor to represent neural network, but  
 839 approach them as a graph where the nodes are neurons possessing an activation function and edges  
 840 holding the weights, or *parameters* in modern literature, between those nodes. Most of the modern  
 841 neural network work with the principle of neurons layers. Each neurons belong to a layer and takes  
 842 input from the preceding layer and forward it result to next layer. For example the most basic couple  
 843 of layers is the fully connected layers where each neurons of the input layer is connected to every  
 844 other neurons of ouput layer. All the neurons posses the same activation function  $F$ . The connection  
 845 between the two layers is expressed as a tensor  $T_j^i$  where  $i$  is the index of the precedent layer and  $j$   
 846 the index of the next layer. The propagation from the layer  $I$  to  $J$  is then described as

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

847 where the learning parameters are the tensor  $T_j^i$  and the bias tensor  $B_j$ . This is the fundamental  
 848 component of the Fully Connected Deep NN (FCDNN) family presented in section 3.2.1. Most of the  
 849 modern neural networks use gradient descent to optimize their parameters, i.e. the gradient of the  
 850 parameter  $\theta$  in respect of the loss function  $\mathcal{L}$  is subtracted each optimisation step

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

851  $i$  being the training step index. This induce  $\mathcal{L}$  needs to be differentiable with respect to  $\theta$ , thus the  
 852 layers and their activation functions also need to be differentiable. This simple gradient descent,  
 853 designated as Stochastic Gradient Descent (SGD), can be extended with first and second order mo-  
 854 mentums like in the Adam optimizer [55] (more details in section 3.2.5).

855 This description of neural networks as layers introduced the principles of *depth* and *width*, the num-  
 856 ber of layers in the NN and the number of neurons in each layer respectively. Those quantities that

not directly used for the computation of the results but describes the NN or its training are designated as *hyperparameters*.

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

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

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

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

### 3.2.1 Fully Connected Deep Neural Network (FCDNN)

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

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

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

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

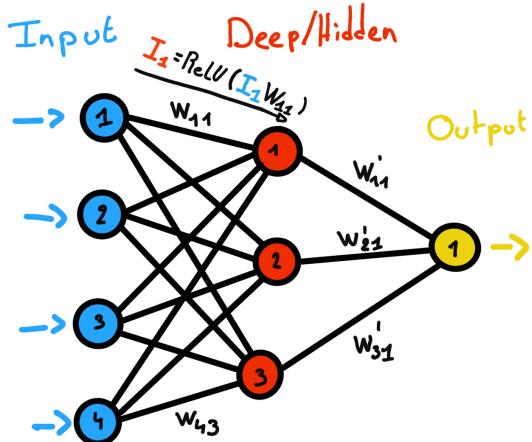
The reasoning behind ReLU and PReLU is that with enough of them, you can mimic any continuous function as illustrated in figure 3.2b. Sigmoid is more used in case of classification, its behavior going hand in hand with the Cross Entropy loss function used in classification problems.

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

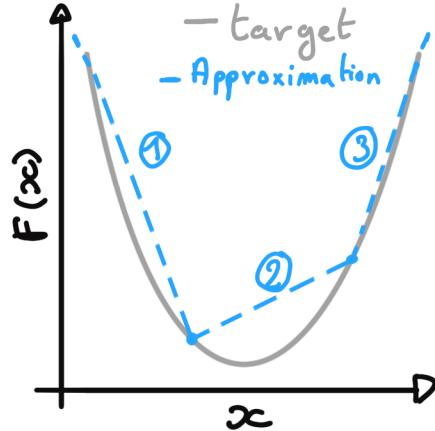
### 3.2.2 Convolutional Neural Network (CNN)

Convolutional Neural Networks are a family of neural networks that use discrete convolution filters, as illustrated in an example in figure 3.3, to process the input data, often images. They have the advantage to be translation invariant by construction, this mean that they are capable of detecting oriented features independently of their location on the image. The learning parameters are located in the filters, the network thus learn the optimal filters to extract the desired features. 2D CNN, where the filters are second order tensors that span over third order tensors, are commonly used in image recognition [56] for classification or regression problematics.

The convolution layers are commonly chained [57], reducing the input dimension while increasing the number of filters. The idea behind is that the first layers will process local informations and the latest layers will process more global informations. To try to preserve the amount of information, we tend to grow the numbers of filters for each division of the input data. The results of the convolution



(A) Schema of a FCDNN



(B) Illustration of a composition of ReLU “approximating” a function. (1) No ReLU is taking effect (2) One ReLU is activating (3) Another ReLU is activating

FIGURE 3.2

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

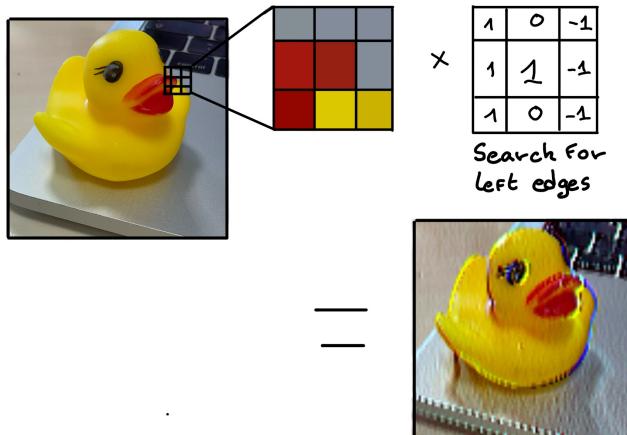


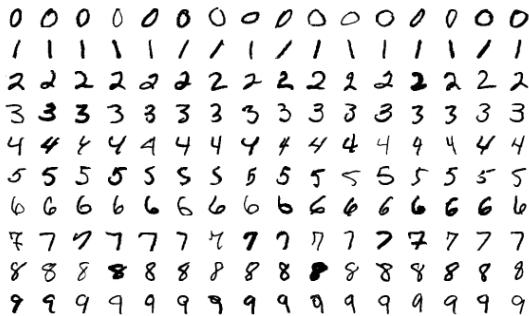
FIGURE 3.3 – Illustration of the effect of a convolution filter. Here we apply a filter with the aim do detect left edges. We see in the resulting image that the left edges of the duck are bright yellow where the right edges are dark blue indicating the contour of the object. The convolution was calculated using [58].

As an example, let’s take the Pytorch [59] example for the MNIST [60], a dataset of black and white images of handwritten digits. Those images are  $28 \times 28$  pixels with only one channel corresponding to the grey level of the pixel. Example of images from this dataset are presented in figure 3.4a

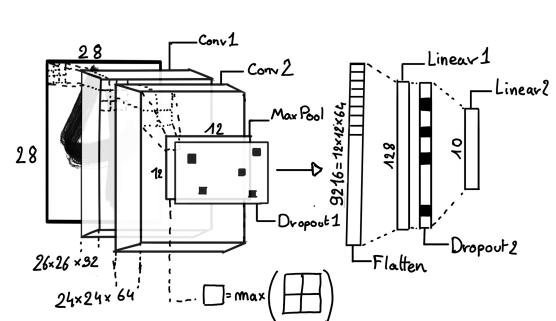
A schema of the CNN used in the Pytorch example is presented in figure 3.4b. Using this schema as a reference, the trained network is made of:

1. A convolutional layer of  $(3 \times 3)$  filters yielding 32 channels. A bias parameter is applied to each channel for a total of  $(32 \cdot (3 \times 3) + 32) = 320$  parameters. The resulting image is  $(26 \times 26 \times 32)$  (26 per 26 pixels with 32 channels). The ReLU activation function is applied to each pixel.

- 901     2. A second convolutional layer of  $(3 \times 3)$  filters yielding 64 channels. This channel also posses  
 902       a bias parameter for a total of  $(64 \cdot (3 \times 3) + 64) = 640$  parameters. Resulting image is  $(24 \times$   
 903        $24 \times 64)$ . This channel also apply a ReLU activation function.
- 904     3. Then comes a  $(2 \times 2)$  max pool layer with a stride of 1 meaning that for each channel the max  
 905       value of pixels in a  $(2 \times 2)$  block is condensed in a single resulting pixel. The resulting image  
 906       is  $(12 \times 12 \times 64)$ .
- 907     4. This image goes through a dropout layer which will set the pixel to 0 with a probability of  
 908       0.25. This help prevent overtraining the neural network (see section 3.2.6 for more details).
- 909     5. The data is the flattened i.e. condensed into a vector of  $(12 \times 12 \times 64) = 9216$  values.
- 910     6. Then comes a fully connected linear layer (Eq. 3.1) with a ReLU activation that output 128  
 911       feature. It needs  $(9216 \cdot 128) + 128 = 1'179'776$  parameters.
- 912     7. This 128 item vector goes through another dropout layer with a probability of 0.5
- 913     8. The vector is then transformed through a linear layer with ReLU activation. It output 10  
 914       values, one for each digit class  $(0, 1, 2, \dots, 9)$ . It need  $(128 \cdot 10) + 128 = 1408$  parameters.
- 915     9. Finally the 10 values are normalized using a log softmax function  $\text{LogSoftmax}(x_i) = \log \left( \frac{\exp(x_i)}{\sum_j \exp(x_j)} \right)$ .
- 916       Each of those values are the probability of the input image to be a certain digit.



(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

917     The final network needs 1'182'144 parameters or, if we consider each parameters to be a double  
 918       precision floating point, 9.45 MB of data. To gives a order of magnitude, such neural network is  
 919       considered "simple", train in a matter of minutes on T4 GPU [61] (14 epochs) and reach an accuracy  
 920       in its prediction of 99%.

### 921     3.2.3 Graph Neural Network (GNN)

922     Graph neural network is a family of neural network where the data is represented as a graph  $G(\mathcal{N}, \mathcal{E})$   
 923       composed of vertex or node  $n \in \mathcal{N}$  and edges  $e \in \mathcal{E}$ . The edges are associated to two nodes  $(u, v) \in$   
 924        $\mathcal{N}^2$ , "connecting" them. The node and the edges can hold features, commonly represented as vector  
 925        $n \in \mathbb{R}^{k_n}$ ,  $e \in \mathbb{R}^{k_e}$  with  $k_n$  and  $k_e$  the number of features on the nodes and edges respectively. We can  
 926       thus define a graph using two tensors  $A_e^{ij}$  the adjacency tensors that hold the features  $e \in [0, k_e]$  of  
 927       the edge connecting the node  $i$  and  $j$  and the tensor  $N_v^i$  that hold the features  $v \in [0, k_n]$  of a node  $i$ .

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

historically in “Neural Message Passing for Quantum Chemistry” [62]. In this algorithm, with each layer of message passing a new set of features is computed for each node following

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

where  $\phi_u$  is a differentiable update function,  $\square_j$  is a differentiable aggregation function and  $\phi_m$  is a 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 nodes  $n_j$  from which it exist an edge  $e_{ij} \rightarrow (n_i, n_j)$ .  $k$  is the layer on which the message passing algorithm is applied.  $\square$  need also a few other property if we want to keep the graph property, most notably the permutational invariance of its parameters (example: mean, std, sum, ...).

The edges features can also be updated, either by directly taking the results of  $\phi_m$  or by using another message function  $\phi_e$ .

To explain this process, let’s take the situation presented in figure 3.5. We start with an input graph on left, in this case the message passing algorithm is mixing the color on each nodes and produce nodes of mixed color. For simplicity, the  $\phi_m$  and  $\phi_u$  function are the identity, they take a color and output the same color.

Let’s look at what’s happening in the node 4. It has 3 neighbours and is a neighbour of itself. The four resulting  $\phi_m$  extract the color of each nodes and then feed them to the  $\square$  function. The  $\square$  function just equally distribute the color in the node. Finally the  $\phi_u$  function just update the node with the output of  $\square$ .

Interestingly we see that the new node 4 does not have any yellow, the color of node 1. But if we were to run the message passing algorithm again, it would get some as node 2 is now partially yellow. If color here represent information, we see that multiple step are needed so that each node is “aware” of the informations the other nodes possess.

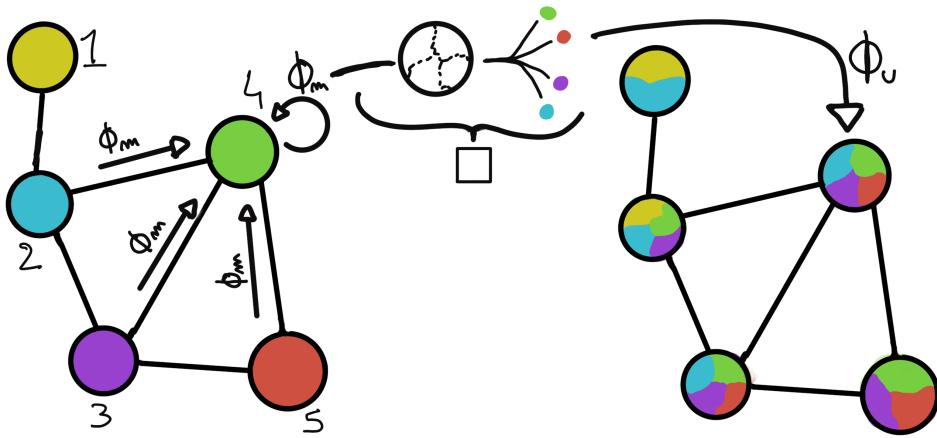


FIGURE 3.5 – Illustration of the message passing algorithm. The detailed explanation can be found in section 3.2.3

Message passing is a very generic way of describing the process of GNN and it can be specialized for convolutional filtering [49], diffusion [63] and many other specific operation. GNN are used in a wide variety of application such as regression problematics, node classification, edge classification, node and edge prediction, ...

It is a very versatile but complex tool.

### 3.2.4 Adversarial Neural Network (ANN)

The adversarial machine learning, Adversarial Neural Networks (ANN) in the case of neural network, is a family of unsupervised machine learning algorithms where the learning algorithm (generator) is competing against another algorithm (discriminator). Taking the example of Generative Adversarial Networks, concept initially developed by Goodfellow et al. [64], the discriminator goal is to discriminate between data coming from a reference dataset and data produced by the generator. The generator goal, on the other hand, is to produce data that the discriminator would not be able to differentiate from data from the reference dataset. The expression of duality between the two models is represented in the loss where, at least a part of it, is driven by the results of the discriminator.

### 3.2.5 Training procedure

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

#### Training lifecycle

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

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

#### The optimizer

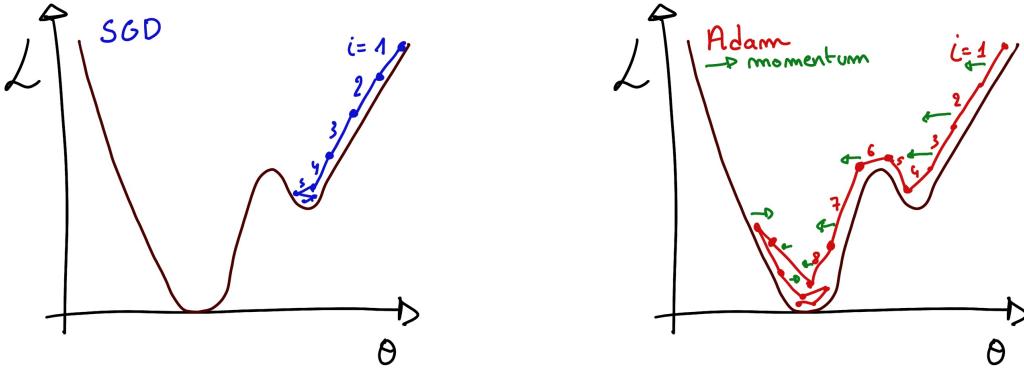
As briefly introduced section 3.2, the parameters of the neural network are optimized using the gradient descent method. We calculate the gradient of the mean loss over the batch with respect of each parameters and we update the parameters in accord to minimize the loss. The gradient is computed backward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

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

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

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

1. SGD (Stochastic Gradient Descent). This is the simplest optimizer, it depend on only one



(A) Illustration of SGD falling into a local minima

(B) Illustration of the Adam momentum allowing it to overcome local minima

FIGURE 3.6

hyperparameter, the learning rate  $\lambda$  (LR) and update the parameters  $\theta$  following

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

where  $t$  is the step index. It is a powerful optimizer but is very sensible to local minima of the loss in the parameters phase space as illustrated in figure 3.6a.

2. Adam [55]. The concept is, in short, to have and SGD but with momentum. Adam possess two momentum  $m(\beta_1)$  and  $v(\beta_2)$  which are respectively proportional to  $\frac{\partial \mathcal{L}}{\partial \theta}$  and  $(\frac{\partial \mathcal{L}}{\partial \theta})^2$ .  $\beta_1$  and  $\beta_2$  are hyperparameters that dictate the moment update at each optimization step. The parameters are then upgraded following

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

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

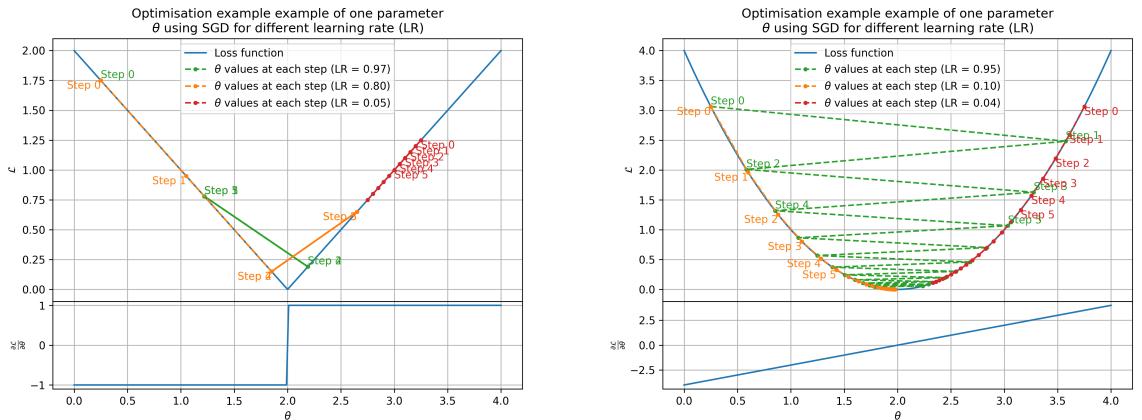
$$\theta_{t+1} = \theta_t - \lambda \frac{m_{t+1}}{\sqrt{v_{t+1}} + \epsilon} \quad (3.13)$$

where  $\epsilon$  is a small number to prevent divergence when  $v$  is close to 0. These momentums allow to overcome small local minima in the parameters phase space as illustrated in figure 3.6a.

The LR is a crucial parameter in the training of NN, as illustrated in figure 3.7. To prevent possible issues, we setup scheduler policies.

## 1000 Scheduler policies

Sometimes we want to update our hyperparameters or take a set of action during the training procedure. We use for this scheduler policies, for example a common policy is a decrease of the learning rate after each epochs. The reasoning is that if the learning rate is too high, the optimizer will continuously miss the minimum and oscillate around it (figure 3.7a). By reducing the learning



(A) Illustration of the SGD optimizer on one parameter  $\theta$  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  $\theta$  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 3.7 – 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.

1005 rate, we allow it to make more fine steps in the parameters phase space, hopefully converging to the  
1006 true minima.

1007 Another policy that is often use is the save of the best model. In some situation, the loss value after  
1008 each epoch will strongly oscillate or can even worsen. This policy allow us to keep the best version  
1009 of the model attained during the training phase.

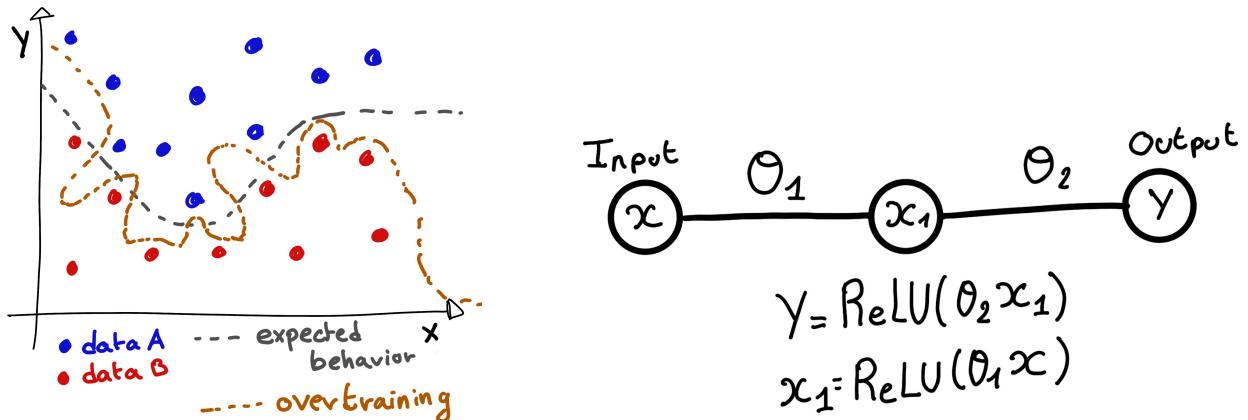
### 1010 3.2.6 Potential pitfalls

1011 Apart from being stuck in local minima, there is also other behaviors and effects we want to prevent  
1012 during training.

#### 1013 Overtraining

1014 This happen when the network learn the specificities of the training dataset instead of a more general  
1015 representation of the underlying data distribution. This can happen if there is not enough data  
1016 in comparison to the number of learning parameters, if the data contains some specific signatures  
1017 specific to the training dataset or if trains for too long on the same dataset. This behavior is illustrated  
1018 in figure 3.8a. Overtraining can be fought in multiple ways, for example:

- 1019 — **More data.** By having more data in the training dataset, the network will not be able the  
1020 specificities of every data.
- 1021 — **Less parameters.** By reducing the number of parameters, we reduce the computing and  
1022 learning capacities of the network. This will force it to fallback to generalist behaviours.
- 1023 — **Dropout.** This technique implies to randomly set part of the neural network to 0. By doing  
1024 this, we force the redundancy in its computing capability and, in a way, modify the data  
1025 decreasing the possibility for specific learning.



(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 3.8

— **Early stopping.** During the training we monitor the network performance over a validation dataset. The network does not train on this dataset and thus cannot learn its specificities. If the loss on the training dataset diverge too much from the loss on the validation dataset, we can stop the training earlier to prevent it from overtraining.

### 1030 Gradient vanishing

1031 Gradient vanishing is the effect of the gradient being so small for the upper layer that the parameters  
 1032 are barely updated after each step. This cause the network to be unable to converge to the minima.

1033 This comes from the way the gradient descent is calculated. Imagine a simple network composed of  
 1034 three fully connected layers: the input layer, a intermediate layer and the output layer. Let  $L$  be the  
 1035 loss,  $\theta_1$  the parameter between the input and the intermediate layer and  $\theta_2$  the parameter between  
 1036 the intermediate and output layer. This network is schematized in figure 3.8b.

1037 The gradient for  $\theta_1$  will be computed using the chain rule presented in equation 3.9. Because  $\theta_1$   
 1038 depends on  $\theta_2$ , if the gradient of  $\theta_2$  is small, so will be the gradient of  $\theta_1$ . Now if we would have  
 1039 much more layer, we can see how the subsequent multiplication of small gradients would lead to  
 1040 very small update of the parameters thus “vanishing gradient”.

1041 Multiple actions can be taken to prevent this effect such as:

- 1042 — **Batch normalization:** In this case we apply a normalization layer that will normalize the data  
 1043 so that, let  $D$  be the data,  $\langle D \rangle = 0$  and  $\sigma_D = 1$ . This help the weight of the network to  
 1044 maintain an appropriate scale.
- 1045 — **Residual Network (ResNet)** [65]: Residual network is a technique for neural network in  
 1046 which, instead of just sequentially feeding the results of each layer to the next one, you ask  
 1047 each layer to calculate the residual of the input data. This technique is illustrated in figure 3.9.

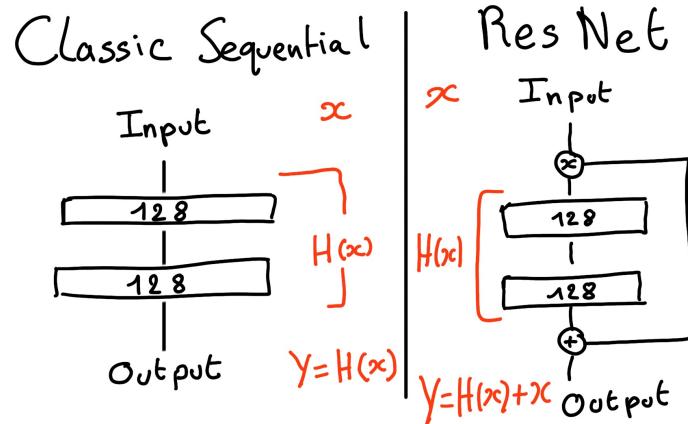


FIGURE 3.9 – Illustration of the ResNet framework

1048 **Gradient explosion**

Gradient explosion happens when the consecutive multiplication of gradient cause exponential grow in the parameter value or if the training lead the network in part of the parameter space where the gradient is significantly higher than usual. For illustration, consider that the loss dependency in  $\theta$  follow

$$\mathcal{L}(\theta) = \frac{\theta^2}{2} + e^{4\theta}$$

$$\frac{\partial \mathcal{L}}{\partial \theta} = \theta + 4e^{4\theta}$$

1049 The explosion is illustrated in figure 3.10 where we can see that the loss degrade with each step of  
 1050 optimization. In this illustration it is clear that reducing the learning rate suffice but this behaviour  
 1051 can happens in the middle of the training where the learning rate schedule does not permit reactivity.

1052 There exist solutions to prevent this explosions:

- 1053 — **Gradient clipping:** Is this case we work on the gradient so that the norm of gradient vector  
 1054 does not exceed a certain threshold. In our illustration in figure 3.10 the gradient for  $\theta > 0$   
 1055 could be clipped at 3 for example.
- 1056 — **Batch normalization:** For the same reasons as for gradient vanishing, normalizing the input  
 1057 data help reduce erratic behaviour.

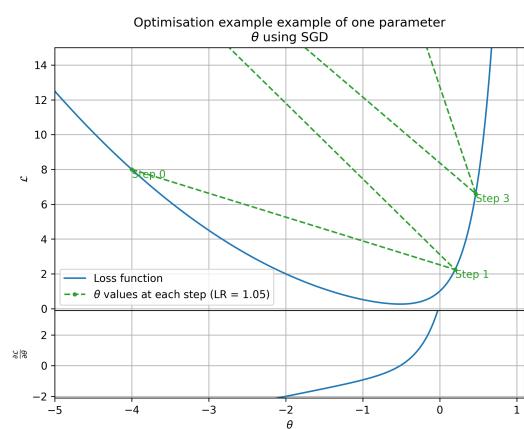


FIGURE 3.10 – Illustration of the gradient explosion. Here it can be solved with a lower learning rate but its not always the case.

1058 **Chapter 4**

1059 **Image recognition for IBD  
reconstruction with the SPMT system**

1061

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.

Dave - What are you talking about, HAL?

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

1062 As explained in chapter 2, JUNO is an experiment composed of two systems, the Large Photomultiplier (LPMT) system and the Small Photomultiplier (SPMT) system. Both of them observe the same 1063 physics events inside of the same medium but they differ in their photo-coverage, respectively 75.2% 1064 and 2.7%, their dynamic range (see section 2.2.2), a thousands versus a few dozen, and their front-end 1065 electronics (see section 2.2.2).

1066 They are complementary in their strengths and weaknesses and support each other, this is what 1067 we call *Dual Calorimetry*. One important point is their differences in expected resolution, the LPMT 1068 system outperform largely the SPMT system but is subject to effects such as charge non linearity [29] 1069 that could bias the reconstruction. Effects that the SPMT system is impervious to. This topic will 1070 be studied in more detail in chapter 7. Also, due to the dynamic range of the LPMT, in case of high 1071 energy and high density event such as core-collapse supernova, the LPMT system could saturate and 1072 the lower photo-coverage become a benefit.

1073 Thus, although event reconstruction algorithm and physics analysis combines both LPMT and SPMT 1074 systems, individual approach are key studies to understand the detector and ensure their reliability. 1075 This topic will also be studied in more details in chapter 7. The subject of this chapter is to propose 1076 a machine learning algorithm for the SPMT reconstruction based on Convolutional Neural Network 1077 (CNN). 1078

1079 **4.1 Motivations**

1080 As explained in chapter 3, Machine Learning (ML) algorithms shine when modeling highly dimensional 1081 data from a given dataset. In our case, we have access to complete monte-carlo simulation of 1082 our detector to produce arbitrary large datasets that could represent multiple years of data taking. 1083 Ideally ML algorithms would be able to consider the entirety of the information in the detector and 1084 converge on the best parameters to yield optimal results, while classical methods could be biased by 1085 the prior knowledge of the detector and physics processes. To study this potential phenomena, we

1086 will compare our machine algorithm to a classical reconstruction method developed for energy and  
 1087 vertex reconstruction [66].

1088 We have access to a very detailed simulation of the detector (section 2.5) that will allow us to simulate  
 1089 arbitrary large dataset while giving access to all the physics parameters of the event. Those  
 1090 parameters include the target of our reconstruction algorithms: the vertex and energy of our event.  
 1091 As introduced above, we hope that the ML algorithm will be able to use all the informations in the  
 1092 event, but that could lead that potential mismodelings in our simulation could be exploited by the  
 1093 algorithm. This specific subject will be studied in chapter 6.

## 1094 4.2 Method and model

1095 One of simplest way to look at JUNO data is to consider the detector as an array of geometrically  
 1096 distributed sensors on a sphere. Their repartition is almost homogeneous, on this sphere surface  
 1097 providing an almost equal amount of information per unit surface on this sphere. It is then tempting  
 1098 to represent the detector as a spherical image with the PMTs in place of pixels. Two events with two  
 1099 different energy or position would produce two different images.

1100 The most common approach in machine learning for image processing and image recognition is the  
 1101 Convolutional Neural Network (CNN). It is widely used in research and industry [57, 67–69] due to  
 1102 its strengths (see section 3.2.2) and has proven its relevance in image processing.

1103 Some CNN are developed to process spherical images [70] but for the sake of simplicity and as a  
 1104 first approach we decided to go with a planar projection of the detector, approach that has proven its  
 1105 efficiency using the LPMT system (see section 2.6.3). The details about this planar projection will be  
 1106 discussed in section 4.2.2.

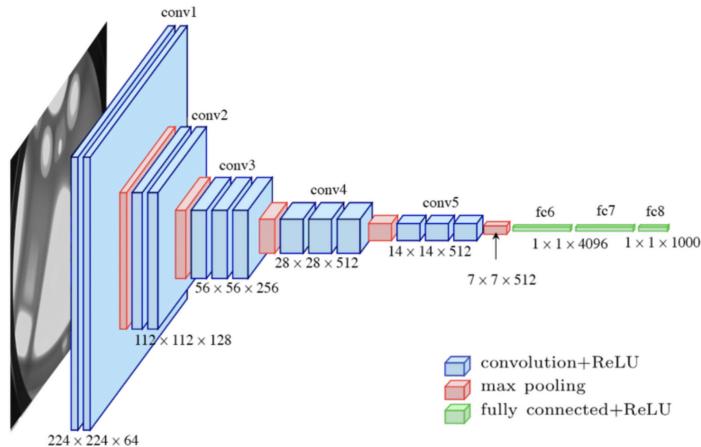


FIGURE 4.1 – Graphic representation of the VGG-16 architecture, presenting the different kind of layer composing the architecture.

### 1107 4.2.1 Model

1108 The architecture we use is derived from the VGG-16 architecture [57] illustrated in figure 4.1. We  
 1109 define a set of hyperparameters that will define the size, complexity and computational power of the  
 1110 NN. The chosen hyperparameters are detailed below and their values are presented in table 4.1.

- 1111 —  $N_{\text{blocks}}$ : the number of convolution blocks, a block being composed of two convolutional  
 1112 layers with  $3 \times 3$  filters using ReLU activation function, a  $3 \times 3$  max-pooling layer (except for  
 1113 the last block).
- 1114 —  $N_{\text{channels}}$ : The number of channels in the first block. The number of channels in the subsequent  
 1115 blocks is computed using  $N_{\text{channels}}^i = i * N_{\text{channels}}, i \in [1..N_{\text{blocks}}]$ .
- 1116 — **FCDNN configuration:** The result of the last convolution layer is flattened then fed to a  
 1117 FCDNN. Its configuration is expressed as a sequence of fully connected linear layer using  
 1118 the PReLU activation function. For example  $2 * 1024 + 2 * 512$  is the sequence of 2 layers  
 1119 with a width of 1024 followed by 2 other layers with a width of 512. Finally the last layer  
 1120 is a 4 neurons wide linear layers without activation function. Each neurons of the last layer  
 1121 represent a component of the interaction vertex: Energy, X, Y, Z.
- 1122 — **Loss:** The loss function. In this work we study two different loss function  $(E + V)$  and  $(E_r +$   
 1123  $V_r)$  detailed below.

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

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

1124 where  $R$  is the radius of the CD. With the energy in MeV and the distance in meters, we use the factor  
 1125 0.85 and 10 to equilibrate the two term of the loss function so they have the same magnitude.

- 1126 — The loss function  $(E + V)$  is close to a simple Mean Squared Error (MSE). MSE is one of the  
 1127 most basic loss function, the derivative is simple and continuous in every point. It is a strong  
 1128 starting point to explore the possibility of CNNs.
- 1129 —  $(E_r + V_r)$  can be seen as a relative MSE.

1130 The idea is that: due to the inherent statistic uncertainty over the number of collected Number of  
 1131 Photo Electrons (NPE), the absolute resolution  $\sigma(E - E_{\text{true}})$  will be larger at higher energy than at  
 1132 low energy. But we expect the *relative* energy resolution  $\frac{\sigma(E - E_{\text{true}})}{E_{\text{true}}}$  to be smaller at high energy than  
 1133 lower energy as illustrated in figure 2.22. Because of this, by using simple MSE the most important  
 1134 part in the loss come from the high energy part of the dataset whereas with a relative MSE, the  
 1135 most important part become the low energy events in the dataset. We hope that by using a relative  
 1136 MSE, the neural network will focus on low energy events where the reconstruction is considered the  
 1137 hardest.

1138 Each combination of those hyperparameters (for example  $(N_{\text{blocks}} = 2, N_{\text{channels}} = 32, \text{FCDNN} =$   
 1139  $(2 * 1024), \text{Loss} = (E + V))$ ), subsequently designated as configurations, is then tested and compared  
 1140 to each other over an analysis sample.

1141 On top those generated models, we define 4 hand tailored models:

- 1142 — “gen\_0”:  $N_{\text{blocks}} = 4, N_{\text{channels}} = 64$ , FCDNN configuration:  $1024 * 2 + 512 * 2$ , Loss :=  $E + V$
- 1143 — “gen\_1”:  $N_{\text{blocks}} = 4, N_{\text{channels}} = 64$ , FCDNN configuration:  $1024 * 2 + 512 * 2$ , Loss :=  $E_r + V_r$
- 1144 — “gen\_2”:  $N_{\text{blocks}} = 5, N_{\text{channels}} = 64$ , FCDNN configuration:  $4096 * 2 + 1024 * 2$ , Loss :=  $E + V$
- 1145 — “gen\_3”:  $N_{\text{blocks}} = 5, N_{\text{channels}} = 64$ , FCDNN configuration:  $4096 * 2 + 1024 * 2$ , Loss :=  $E_r + V_r$

1146 We cannot use the mean loss because we consider multiple loss functions, there is no guarantee that  
 1147 comparison of their numerical value will be meaningful. We use multiple observables to rank the  
 1148 performances of each configuration:

- 1149 — The mean absolute energy error  $\langle E \rangle = \langle |E - E_{\text{true}}| \rangle$ . It is an indicator of the energy bias of our  
 1150 reconstruction.
- 1151 — The standard deviation of the energy error  $\sigma E = \sigma(E - E_{\text{true}})$ . This the indicator on our  
 1152 precision in energy reconstruction.
- 1153 — The mean distance between the reconstructed vertex and the true vertex  $\langle V \rangle = \langle |\vec{V} - \vec{V}_{\text{true}}| \rangle$ .  
 1154 This an indicator of the bias and precision of our vertex reconstruction.

$N_{blocks}$	{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

— 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.

The models were developped in Python using the pytorch framework [59] using NVIDIA A100 [71] and NVIDIA V100 [72] gpus. The A100 was split in two, thus the accessible gpu memory was 20 Gb making it impossible to train some of the architectures due to memory consumption.

The training was monitored in realtime by a custom tooling that was developed during this thesis, DataMo [73].

The training of one model takes between 4h and 15h depending of its size, overall training the full 72 model takes around 500 GPU hours. Even with parallel training, this random search hyper-optimisation was time consuming.

## 4.2.2 Data representation

This data is represented as  $240 \times 240$  images 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  $\theta$  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,  $\phi = 0$  is 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 physics signal 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 image dimension have been 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:

- 1185 — As presented in section 4.2.1, the convolution filter we use are  $3 \times 3$  convolution filter, meaning  
1186 that if SPMTs would be separated by more than one pixel, the first filter would only see one  
1187 SPMT per filter. This behavior would be kind of counterproductive as the first convolution  
1188 block would basically be a transmission layer and would just induce noise in the data.
- 1189 — It keep the network relatively small, while this do not impact the convolution layers, the  
1190 flatten operation just before the FCDNN make the number parameters in the first layer of  
1191 it dependent on the size of the image.
- 1192 — It reduce the number of empty pixel in the image.

1193 The question of empty pixel is an important question in this data representation. There is two kind  
1194 of empty pixels in the data.

1195 The first kind is pixel that contain a SPMT but the SPMT did not get hit nor registered any dark noise  
1196 during the event. In this case, the charge channel is zero, which have a physical meaning but then  
1197 come the question of the time layer. One could argue that the correct time would be infinity (or the  
1198 largest number our memory allows us) because the hit “never” happened, so extremely far from the  
1199 time of the event. This cause numerical problem as large number, in the linear operation that are  
1200 happening in the convolution layers, are more significant than smaller value. We could try to encode  
1201 this feature in another way but no number have any significance due to our time being relative to  
1202 the trigger of the experiment so  $-1$  for example is out of question. Float and Double gives us access  
1203 to special value such as NaN (Not a Number) [74] but the behavior is to propagate the NaN which  
1204 leaves us with NaN for energy and position. We choose to keep the value 0 because it’s the absorbing  
1205 element of multiplication, absorbing the “information” of the parameter it would be multiplied by.  
1206 It also can be though as no activation in the ReLU activation function.

1207 The second kind of pixel is pixel that do not represent parts of the detector such as the corners of  
1208 the image. The question is basically the same, what to put in the charge and the time channel. The  
1209 decision is to set the charge and time to 0 following the above reasoning. It’s important to keep in  
1210 mind the fact that a part of the detector that has not been hit is also an information: There is no signal  
1211 in this part of the detector. This problematic will be explored in more details in chapter 5.

1212 Another problematic that happens with this representation, and this is not dependent of the chosen  
1213 projection, is the deformation in the edges of the image and the loss of the neighbouring information  
1214 in the for the SPMTs at the edge of the image  $\phi \sim 180^\circ$ . This deformation and neighbouring loss  
1215 could be partially circumvented as explained in section 4.5

### 1216 4.2.3 Dataset

1217 In this study we will discuss two datasets of one millions events:

- 1218 — **J21:** The first one comes from the JUNO official mc simulation J21v1r0-Pre2 (released the 18th  
1219 August 2021). This historical version is the one on which the classical algorithm presented in  
1220 [66] was developed. This dataset is used as a reference for comparison to classical algorithm.  
1221 The data in this dataset is *detsim* level (see section 2.5), where only the physic is simulated.  
1222 The charge and time biases and uncertainties are implemented using toy MC adjusted using  
1223 [26, 75]. The time window is not based on a selection algorithm but  $t_0 := t = 0$  is defined as  
1224 the first PMT hit. The window goes up to  $t_0 + 1000$  ns.
- 1225 — **J23:** The second comes from the JUNO official monte-carlo simulations J23.0.1-rc8.dc1 (re-  
1226 leased the 7th January 2024). The data is *calib* level (see section 2.5). Here the charge comes  
1227 from the waveform integration, the time window resolution and trigger decision are all simu-  
1228 lated inside the software. This dataset is more realistic and is used to confirm the performance  
1229 of our algorithm.

1230 To put in perspective this amount of data, the expected IBD rate in JUNO is 47 / days. Taking into  
1231 account the calibration time, and the source reactor shutdown, it amount to  $\sim 94'000$  IBD events  
1232 in 6 years. With this million of event, we are training the equivalent of  $\sim 10$  years of data. With

1233 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  
 1234 event scenarios that could be happening in the detector.

1235 While we expect and hope the monte-carlo simulation to give use a realistic representation of the detector,  
 1236 there could be effect, even after the fine-tuning on calibration data, that the simulation  
 1237 cannot handle. Thus, once the calibration will be available, we will need to evaluate, and if needed  
 1238 retrain, the network on calibration data to establish definitive performances.

1239 The simulated data is composed of positron events, uniformly distributed in the CD volume and in  
 1240 kinetic energy over  $E_k \in [0; 9]$  MeV producing a deposited energy  $E_{dep} \in [1.022; 10.022]$  MeV. This is  
 1241 done to mimic the signal produced by the IBD prompt signal. Uniform distributions are used so that  
 1242 the CNN does not learn a potential energy distribution, favoring some part of the energy spectrum  
 1243 instead of other.

1244 Those events can be considered as “optimistic” as there is no pile-up with potential background or  
 1245 other IBD.

#### 1246 4.2.4 Data characteristics

1247 To delve a bit into the kind of data we will use, you can find in figure 4.2 the repartition of the SPMTs  
 1248 in the image. The color represent the number of SPMTs per pixel.

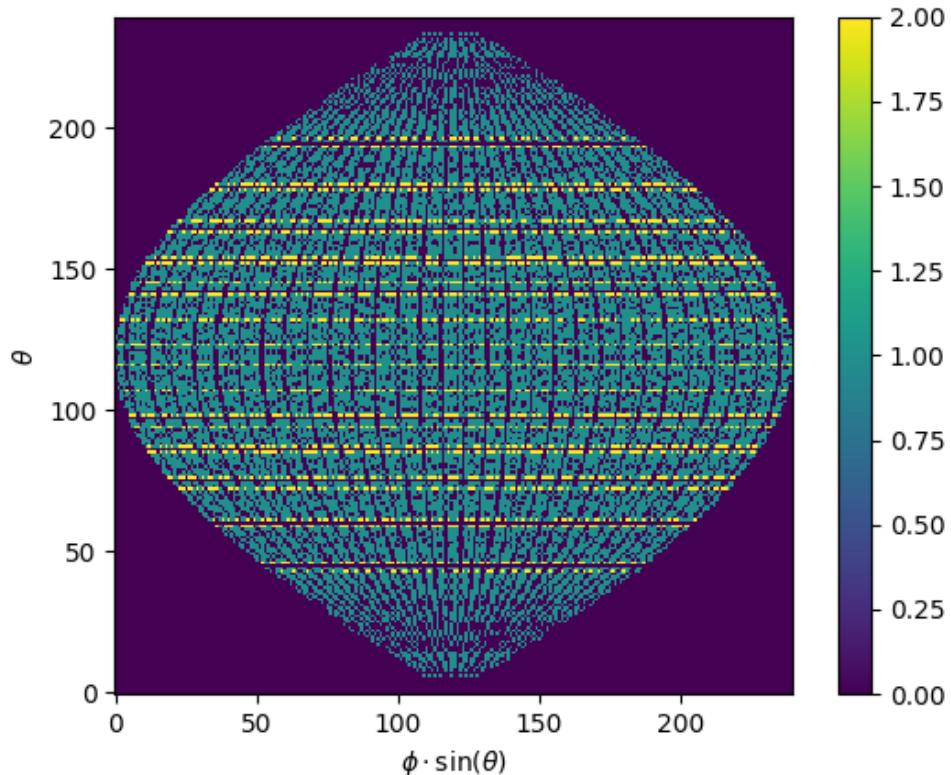


FIGURE 4.2 – Repartition of SPMTs in the image projection. The color scale is the number of SPMTs per pixel

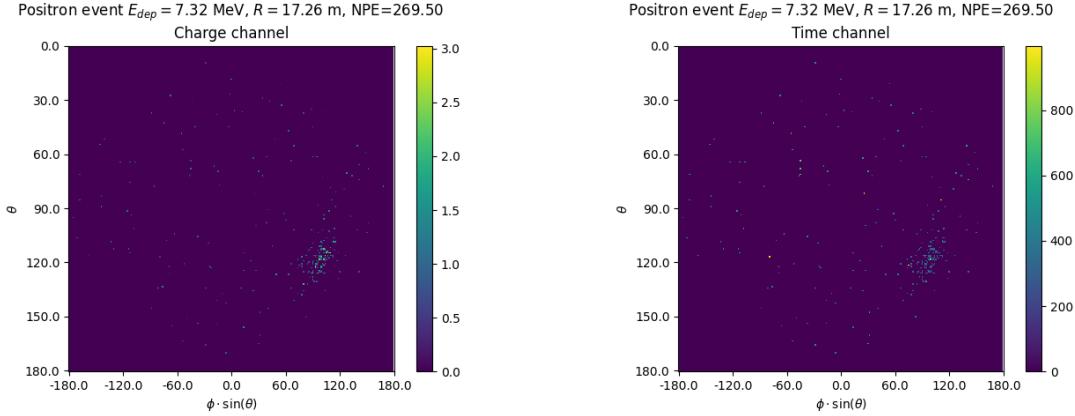


FIGURE 4.3 – Example of a high energy, radial event. We see a concentration of the charge on the bottom right of the image, clear indication of a high radius event. **On the left:** the charge channel. The color is the charge in each pixel in NPE equivalent. **On the right:** The time channel in nanoseconds.

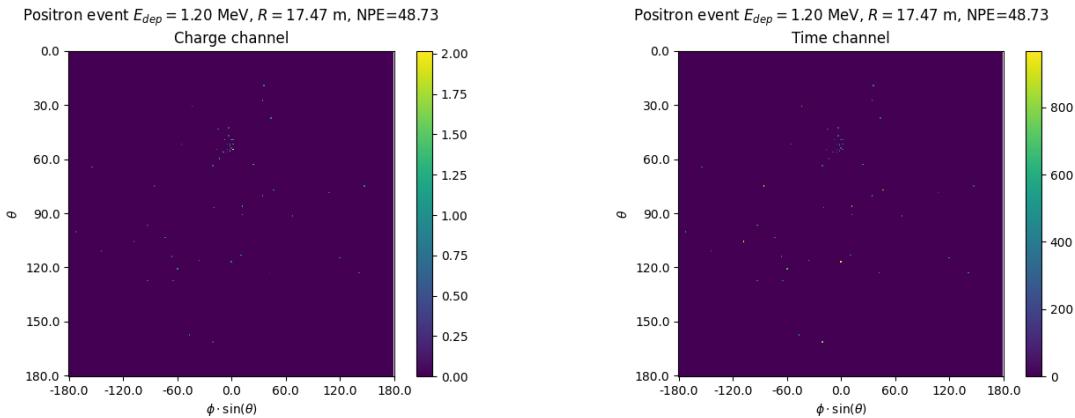


FIGURE 4.4 – Example of a low energy, radial event. The signal here is way less explicit, we can kind of guess that the event is located in the top middle of the image. **On the left:** the charge channel. The color is the charge in each pixel in NPE equivalent. **On the right:** The time channel in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

1249 In figures 4.3, 4.4, 4.5 and 4.6 are presented events from J23 for different positions and energies.  
1250 We see some characteristics and we can instinctively understand how the CNN could discriminate  
1251 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 ratio of NPE per deposited energy. Assuming a linear response of the LS we can model:

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

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

1252 where  $NPE_{tot}$  is the total number of PE detected by the event,  $P_{mev}$  is the mean number of PE detected  
1253 per MeV and  $D_N$  is the dark noise contribution that is considered energy independent. In the case  
1254 where the readout time window is dependent of the energy the dark noise contribution become

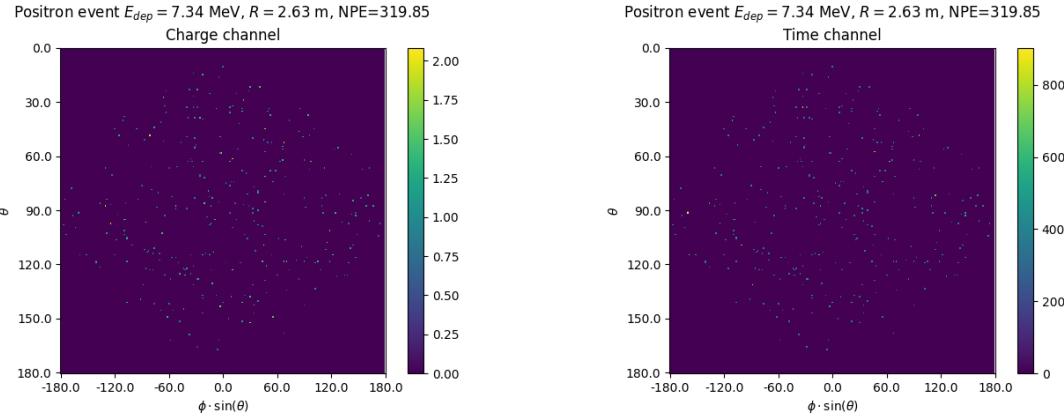


FIGURE 4.5 – Example of a high energy, central event. In this image we can see a lot of signal but uniformly spread, this is indicative of a central event. **On the left:** the charge channel. The color is the charge in each pixel in NPE equivalent. **On the right:** The time channel in nanoseconds.

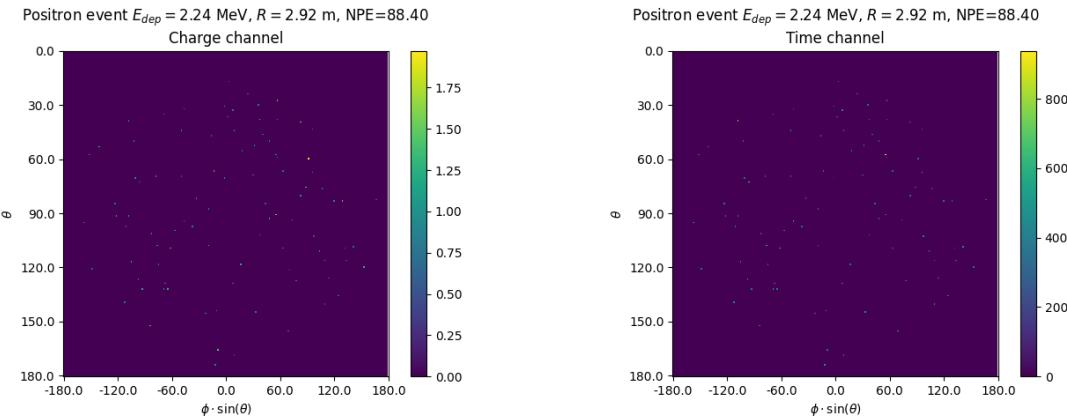
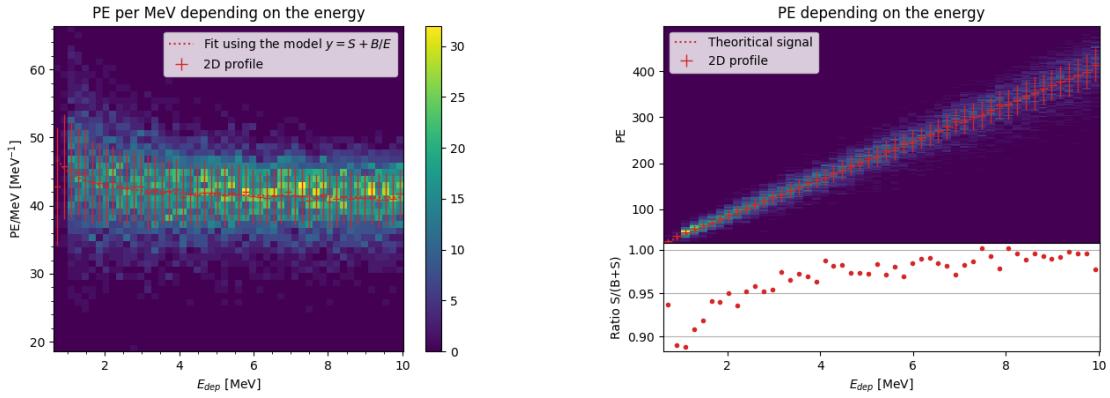


FIGURE 4.6 – Example of a low energy, central event. Here there is no clear signal, the uniformity of the distribution should make it central. **On the left:** the charge channel. The color is the charge in each pixel in NPE equivalent. **On the right:** The time channel in nanoseconds.

1255 energy dependant, also the LS response is realistically energy dependant but figure 4.7a shows that  
1256 we have heavily dominated by statistical uncertainties which is why we are using this simple model.  
1257 The fit shows a light yield of 40.78 PE/MeV and a dark noise contribution of 4.29 NPE. As shown in  
1258 figure 4.7b, the physics makes for 90% of the signal at low energy.

### 1259 4.3 Training

1260 The optimizer used for the training is the Adam [55] optimizer, with a learning rate  $\lambda$  of 1e-3. The  
1261 other hyperparameters were left to their default value ( $\beta_1 = 0.9$ ,  $\beta_2 = 0.999$  and  $\epsilon = 1e^{-8}$ ). The  
1262 learning rate was reduced exponentially during the training at a rate of  $\gamma = 0.95$ , thus  $\lambda_{i+1} = 0.95\lambda_i$   
1263 where  $i$  is the epoch.  
1264 The training was composed of 30 epochs, each epoch constituted of 10k steps using a batch size of 64



(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

1265 events. The validation was computed over a 100 steps on the validation dataset.

## 1266 4.4 Results

1267 Before presenting the results, lets discuss the different observables.

1268 The event are considered point like in this study. The target truth position, or vertex, is the mean po-  
1269 sition of the energy deposits of the positron and the two annihilation gammas. Due to the symmetries  
1270 of the detector, we mainly consider and discuss the bias and precision evolution depending of the  
1271 radius  $R$  but we will still monitor the performances depending of the spheric angle  $\theta$  and  $\phi$ . From the  
1272 detector construction and effect we expect dependency in radius due to the TR area effect presented  
1273 in section 2.6 and the possibility for the positron or the gammas to escape from the CD for near the  
1274 edge events. We also expect dependency in  $\theta$ , the top of the experiment being non-instrumented due  
1275 to the filling chimney. It is also to be noted that the events in the dataset are uniformly distributed in  
1276 the CD, and so are uniformly distributed in  $R^3$  and  $\phi$ . The  $\theta$  distribution is not uniform and we will  
1277 have more event for  $\theta \sim 90^\circ$  than  $\theta \sim 0^\circ$  or  $\theta \sim 180^\circ$ .

1278 We define multiple energy in JUNO:

- 1279 —  $E_\nu$ : The energy of the neutrino.
- 1280 —  $E_k$ : The kinetic energy of the resulting positron from the IBD.
- 1281 —  $E_{dep}$ : The deposited energy of the positron and the two annihilation gammas.
- 1282 —  $E_{vis}$ : 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.
- 1283 —  $E_{rec}$ : 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  $E_{vis}$  while the ones presented in section 2.6.3 reconstruct  $E_{dep}$ .

1284 In this study, we will set  $E_{dep}$  as our target for energy reconstruction. This choice is motivated by the  
1285 ease with which we can retrieve this information in the monte-carlo data while  $E_{vis}$  is less trivial to  
1286 retrieve.  
1287

1290 **4.4.1 J21 results**

1291 Those results comes from the “gen\_30” model, meaning then 30th model generated using the table

1292 [4.1](#) or

1293 — “gen\_30”:  $N_{blocks} = 3$ ,  $N_{channels} = 32$ , FCDNN configuration:  $2048 * 2 + 1024 * 2$ , Loss :=  $E + V$

1294 The performances of its reconstruction are presented in blue in figure [4.8](#). Superimposed in black is

1295 the performances of the classical algorithm from [66].

1296 **Energy reconstruction**

1297 By looking at the figure [4.8a](#) and [4.8b](#), the CNN has similar performances in its energy resolution.

1298 Only at the end of the energy range does the resolution get a little better.

1299 This is explained by looking at the true and reconstructed energy distributions in figure [4.10a](#). We

1300 see that the distributions are similar for energies before 8 MeV but there is an excess of event recon-

1301 structed with energies around 9 MeV while a lack of them for 10 MeV. The neural network seems to

1302 learn the energy distribution and learn that it exist almost no event with an energy inferior to 1.022

1303 MeV and not event with an energy superior to 10 MeV.

1304 The first observation is a physics phenomena: for a positron, its minimum deposited energy is the

1305 mass energy coming from its annihilation with an electron 1.022 MeV. There is a few event with

1306 energies inferior to 1.022 MeV, in those case the annihilation gammas or even the positron escape the

1307 detector. The deposited energy in the LS is thus only a fraction of the energy of the event.

1308 The second observation is indeed true in this dataset but has no physical meaning, it is an arbitrary

1309 limit because the physics region of interest is mainly between 1 and 9 MeV of deposited energy

1310 (figure [2.2](#)). By learning the energy distribution, the CNN pull event from the border of it to more

1311 central value. That’s why the energy resolution is better: the events are pulled in a small energy

1312 region , thus a small variance but the bias become very high (figure [4.8a](#)).

1313 This behavior also explain the heavy bias at low energy in figure [4.8a](#). The energy bias of the CNN if

1314 fairly constant over the energy range, it is interesting to note that the energy bias depending on the

1315 radius is a bit worse than the classical method.

1316 **Vertex reconstruction**

1317 For the vertex reconstruction we do not study  $x$ ,  $y$  and  $z$  independently but we use  $R$  as a proxy

1318 observable. Figure [4.9](#) shows the error distribution of the different vertex coordinates. We see that

1319  $R$  errors and biases are slightly superior to the cartesian coordinates, thus  $R$  is a conservative proxy

1320 observable to discuss the subject of vertex reconstruction.

1321 The comparison of radius reconstruction between the classical algorithm and “gen\_30” are presented

1322 in the figures [4.8c](#), [4.8d](#), [4.8e](#) and [4.8f](#).

1323 Radius reconstruction is worse than the classical algorithms in all configuration. In energy, figure

1324 [4.8c](#), where we see a degradation of almost 20cm over the energy range.

1325 When looking over the true event radius, figure [4.8d](#), we lose between 30 and 45cm of resolution.

1326 The performances are the best for central and radial event.

1327 The precision also worsen when looking at the edge of the image  $\theta \approx 0$ ,  $\theta \approx 2\pi$  respectively the

1328 top and bottom of the image, and when  $\phi \approx -\pi$  and  $\phi \approx \pi$  respectively the left and right side of

1329 the image. This is the confirmation that the deformation of the image is problematic for the event

1330 reconstruction.

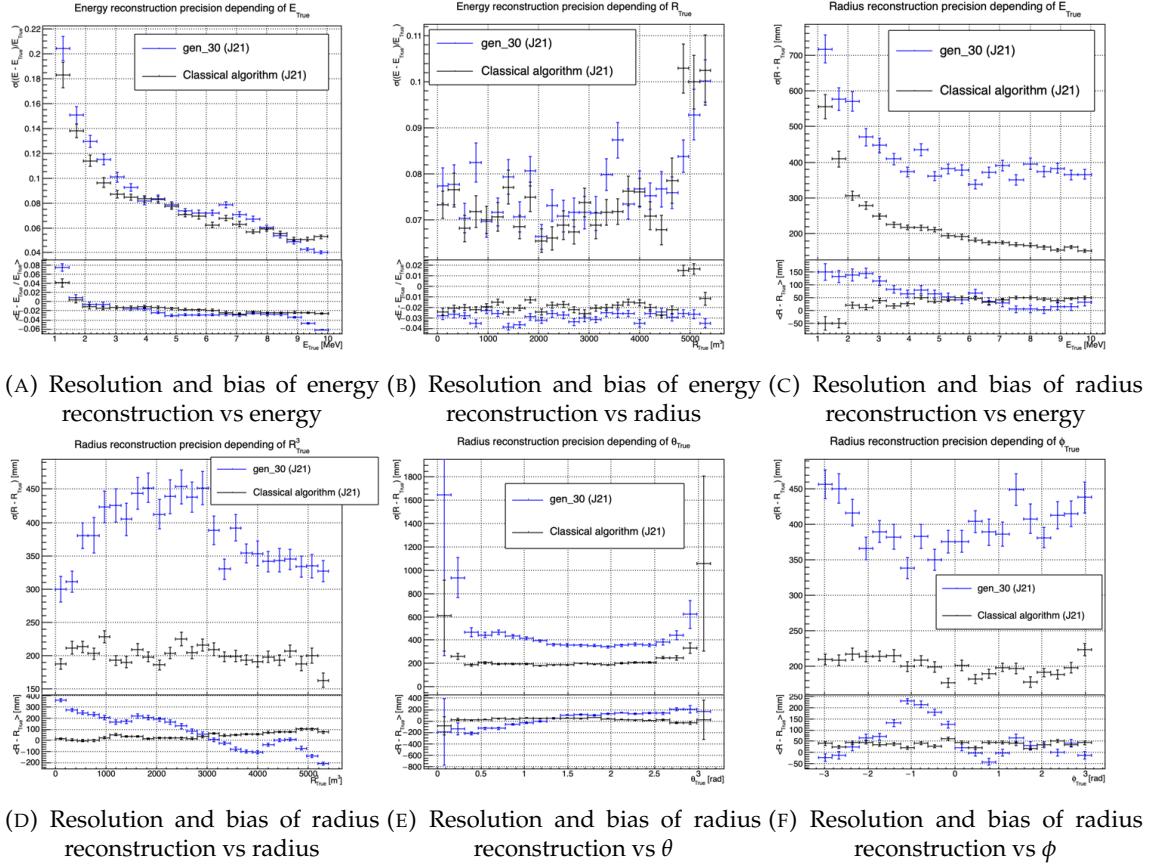


FIGURE 4.8 – Reconstruction performance of the “gen\_30” model on J21 data and its comparison to the performances of the classic algorithm “Classical algorithm” from [66]. The top part of each plot is the resolution and the bottom part is the bias.

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 4.4.2. Over radius,  $\theta$  and  $\phi$  the bias is inconsistent, sometimes event better than the classical reconstruction but can also be 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 other parts but with much less precision.

#### 4.4.2 J21 Combination of classic and ML estimator

As it has been presented in previous section, there is instances where the reconstructed energy and vertex behaves differently between the neural network and the classic algorithm. For instance, if we look at figure 4.8c, we see that while the CNN tend to overestimate the radius at low energy while the classical algorithm seems to underestimate it. Let’s designate the two reconstruction algorithms as estimator of  $X$ , the truth about the event in the phase space  $(E, x, y, z)$ . The CNN and the classical algorithm are respectively designated as  $\theta_N(X)$  and  $\theta_C(X)$ .

$$E[\theta_N] = \mu_N + X; \text{Var}[\theta_N] = \sigma_N^2 \quad (4.7)$$

$$E[\theta_C] = \mu_C + X; \text{Var}[\theta_C] = \sigma_C^2 \quad (4.8)$$

where  $\mu$  is the bias of the estimator and  $\sigma^2$  its variance.

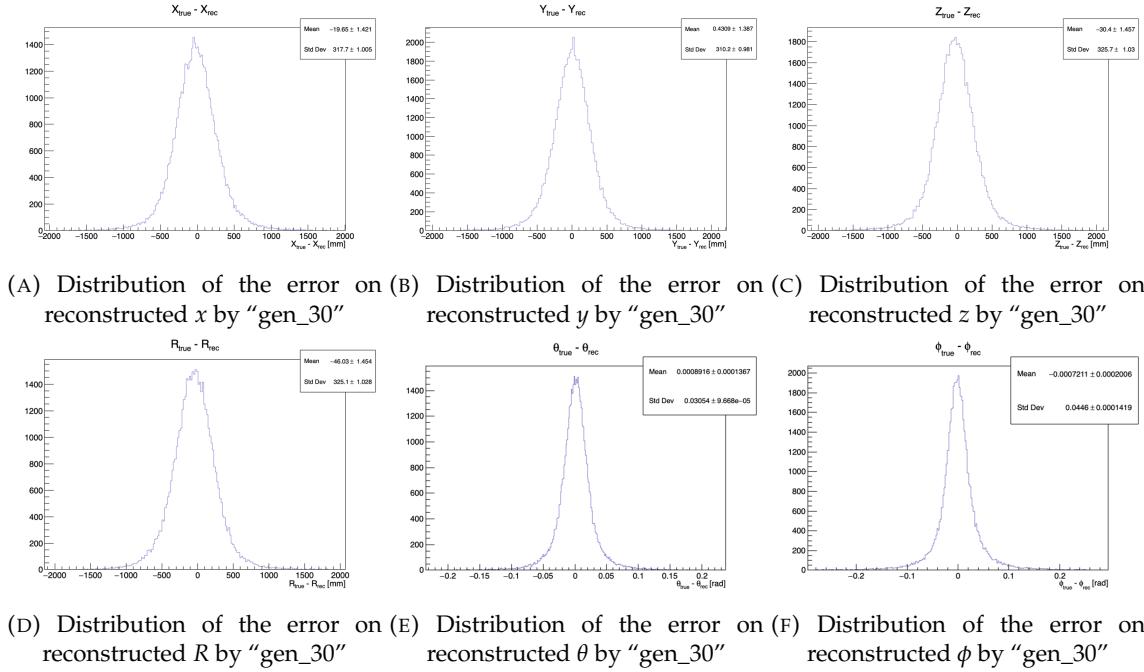


FIGURE 4.9 – Error distribution of the different component of the vertex by "gen\_30". The reconstructed component are  $x$ ,  $y$  and  $z$  but we see similar behavior in the error of  $R$ ,  $\theta$  and  $\phi$ .

<sup>1339</sup> Now if we were to combine the two estimators using a simple mean

$$\hat{\theta}(X) = \frac{1}{2}(\theta_N(X) + \theta_C(X)) \quad (4.9)$$

then the variance and mean would follow

$$E[\hat{\theta}] = \frac{1}{2}E[\theta_N] + \frac{1}{2}E[\theta_C] \quad (4.10)$$

$$= \frac{1}{2}(\mu_N + X + \mu_C + X) \quad (4.11)$$

$$= \frac{1}{2}(\mu_N + \mu_C) + X \quad (4.12)$$

$$\text{Var}[\hat{\theta}] = \frac{1}{4}\sigma_N^2 + \frac{1}{4}\sigma_C^2 + 2 \cdot \frac{1}{4} \cdot \sigma_{NC} \quad (4.13)$$

$$= \frac{1}{4}\sigma_N^2 + \frac{1}{4}\sigma_C^2 + \frac{1}{2} \cdot \sigma_{NC} \quad (4.14)$$

$$= \frac{1}{4}\sigma_N^2 + \frac{1}{4}\sigma_C^2 + \frac{1}{2} \cdot \sigma_N \sigma_C \rho_{NC} \quad (4.15)$$

<sup>1340</sup> Where  $\sigma_{NC}$  is the covariance between  $\theta_N$  and  $\theta_C$  and  $\rho_{NC}$  their correlation.

<sup>1341</sup> We see immediately that if the two estimators are of opposite bias, the bias of the resulting estimator  
<sup>1342</sup> is reduced. For the variance, it depends of  $\rho_{NC}$  but in this case if  $\sigma_C^2$  is close to  $\sigma_N^2$  then even for  
<sup>1343</sup>  $\rho_{NC} \lesssim 1$  then we can gain in resolution.

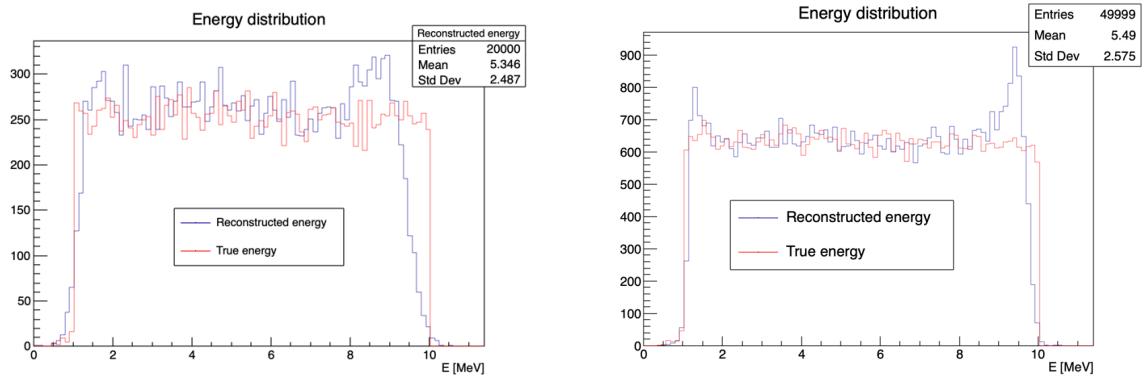
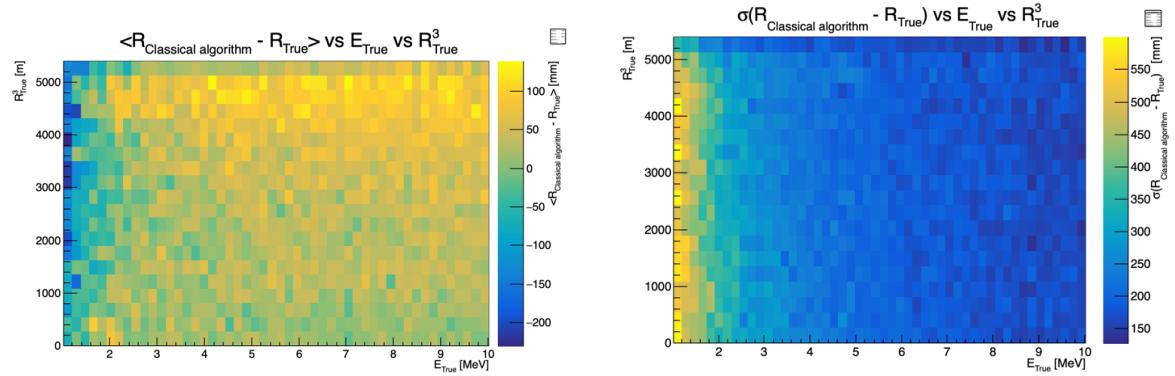


FIGURE 4.10

FIGURE 4.11 – Radius bias (on the left) and resolution (on the right) of the classical algorithm in a  $E, R^3$  grid

1344 By generalising the equation 4.9 to

$$\hat{\theta}(X) = \alpha\theta_N + (1 - \alpha)\theta_C; \alpha \in [0, 1] \quad (4.16)$$

1345 we can determine an optimal  $\alpha$  for two combined estimators. The estimators with the smallest  
1346 variance

$$\alpha = \frac{\sigma_C^2 - \sigma_N\sigma_C\rho_{NC}}{\sigma_N^2 + \sigma_C^2 - 2\sigma_N\sigma_C\rho_{NC}} \quad (4.17)$$

1347 and the estimator without bias

$$\alpha = \frac{\mu_C}{\mu_C - \mu_N} \quad (4.18)$$

1348 See annex A for demonstration.

1349 Its pretty clear from the results shown in figure 4.8 that the bias, variances and correlation are not  
1350 constant across the  $(E, R^3)$  phase space. We thus compute those parameters in a grid in  $E$  and  $R^3$  for  
1351 the following results as illustrated in 4.11.

1352 The map we are using are composed of 20 bins for  $R^3$  going from 0 to  $5400 \text{ m}^3$  (17.54 m) and 50 bins  
1353 in energy ranging from 1.022 to 10.022 MeV. In the case where we are outside the grid, we use the  
1354 closest cell.

The performance of this weighted mean is presented in figure 4.12. We can see that even when the CNN resolution is much worse than the classical algorithm, it can still bring some information thus improving the resolution. This comes from the correlation of the reconstruction error to be smaller than 1 as presented in figure 4.13. We even see some anticorrelation in the radius reconstruction for High radius, high energy, event.

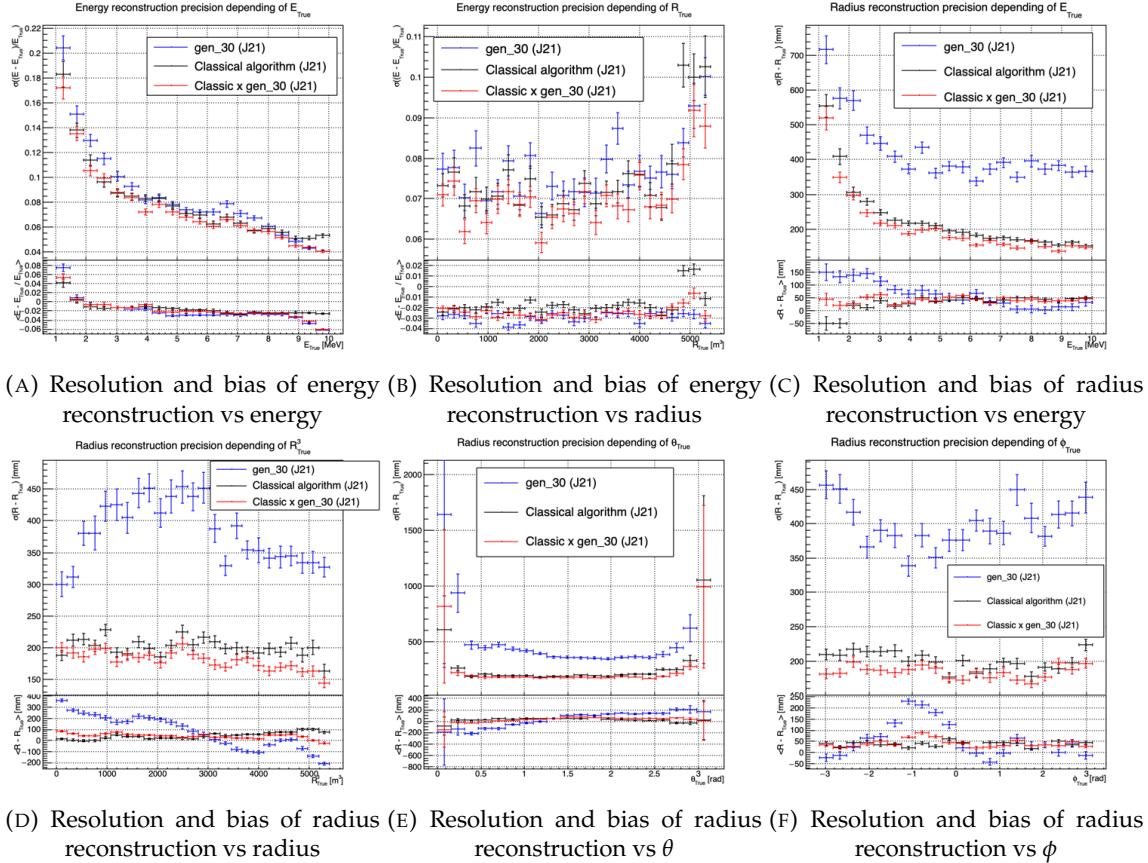


FIGURE 4.12 – Reconstruction performance of the “gen\_30” model on J21, the classic algorithm “Classical algorithm” from [66] and the combination of both using weighted mean. The top part of each plot is the resolution and the bottom part is the bias.

This technique is not suited for realistic reconstruction, we rely too much on the knowledge of the resolution, bias and correlation between the two methods. While this is possible to determine using simulated data or calibration sources, the real data might differ from our model and we would need to really well understand the behavior of the two system. But this is an excellent tool to indicate potential improvements to algorithms and reconstruction methods, showing with this results a potential upper limit to the reconstruction performances.

#### 4.4.3 J23 results

The J21 simulation is fairly old and newer version, such as J23, include refined measurements of the light yield, reflection indices of materials of the detector, structural elements such as the connecting structure and more realistic dark noise. Additionally, the trigger, waveform integration and time window are defined using the algorithms that will ultimately be used by the collaboration to process real physics events.

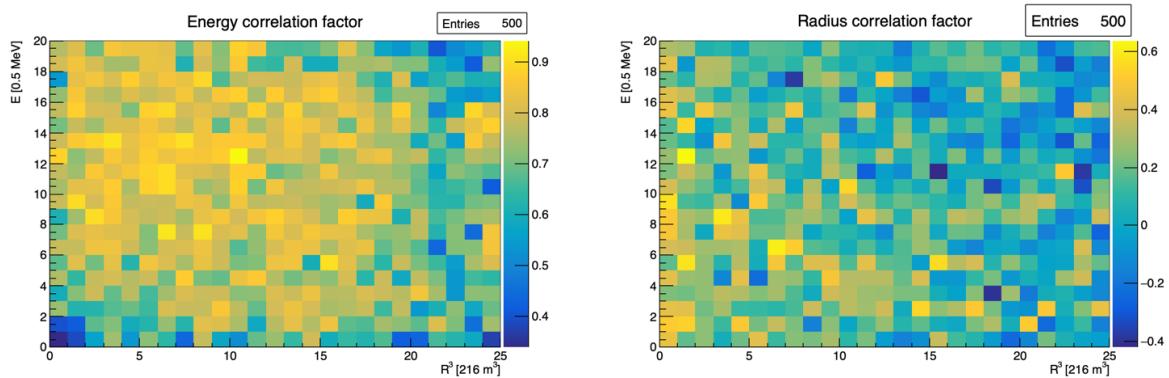


FIGURE 4.13 – Correlation between CNN and classical method reconstruction (on the left) for energy and (on the right) for radius in a  $E, R^3$  grid

We retrained the models defined in 4.2.1 on the J23 data and used the same selection procedure. The results from the best architecture, “gen\_42”, are presented in figure 4.14. Following the table 4.1, “gen\_42” is defined as:

— “gen\_42”:  $N_{blocks} = 3$ ,  $N_{channels} = 64$ , FCDNN configuration:  $4096 * 2$ , Loss :=  $E + V$

### 1376 Energy reconstruction

The results of the energy reconstruction are presented in figures 4.14a and 4.14b. Similarly to what we seen for J21, the resolution is close to the one of the classical algorithm with the exception of the start and end of the spectrum. This come from “gen\_42” learning the shape of the distribution and pulling events from the extreme energies, like 1 and 10 MeV, to more common seen energy, like 2 and 9 MeV as illustrated in figure 4.10b. The bias disappear with the exception of low and high energy events.

### 1383 Vertex reconstruction

The vertex reconstruction, presented in figures 4.14c, 4.14d, 4.14e and 4.14f is not yet to the level of the classical reconstruction but the degradation is smaller than for “gen\_32” being at most a difference of 15cm of resolution and closing to the performance of the classical algorithm in the most favourable condition. “gen\_42” has also very little bias in comparison with the classical method with the exception of the transition to the TR area and at the very edge of the detector.

Unfortunately could not rerun the classical algorithms over the J23 data, as the algorithm was optimised for J21 and was not included and maintained over J23. The combination method need for the two estimators to be run on the same set of event, which was impossible without the classical algorithm being maintained for J23.

Overall the resolution improved over the transition from J21 to J23, effect probably coming from a more complete and rigorous simulation.

## 1395 4.5 Conclusion and prospect

The CNN is a fine tool for event reconstruction in JUNO, and while the reconstruction performances are satisfactory, it show its limitation, the main one concerning the data representation. A lot of

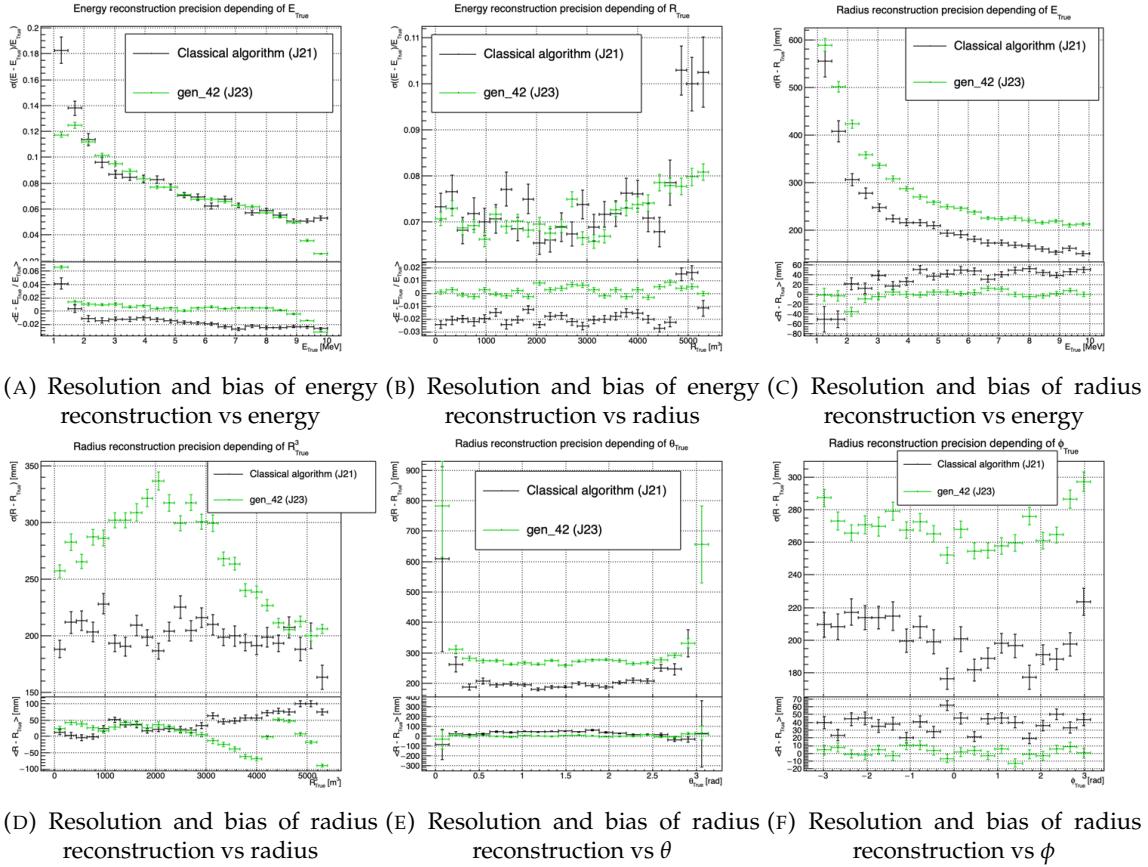


FIGURE 4.14 – Reconstruction performance of the “gen\_42” model on J23 data and its comparison to the performances of the classic algorithm “Classical algorithm” from [66]. The top part of each plot is the resolution and the bottom part is the bias.

1398 training time and resources is consumed going and optimizing over pixel with no physical meaning,  
1399 the NN needs to optimized itself to take into account edges cases such as event at the edge of the  
1400 image and deformation of the charge distribution.  
1401 Those problems could be circumvented, we could imagine a two part CNN where the first part  
1402 reconstruct the  $\theta$  and  $\phi$  spherical coordinates and then rotate the image to locate the event in the  
1403 center of the image. The second part, from this rotated image, would reconstruct the radius and  
1404 energy of the event.  
1405 To overcome the problematic of the aggregation of PMT time information and the meaning of the  
1406 time channel in case of no hit, we could transform this channel into a dimension. This would results  
1407 in an image with multiple charge channels, each one representing the charge sum in a time interval.  
1408 In this thesis, we decided to solve those problem by moving away from the 2D image representation,  
1409 looking into the graph representation and the Graph Neural Network (GNN). This is be the subject  
1410 of the next chapter.

<sup>1411</sup> **Chapter 5**

<sup>1412</sup> **Graph representation of JUNO for  
IBD reconstruction**

<sup>1414</sup> "The Answer to the Great Question of Life, the Universe and  
Everything is Forty-two"

Douglas Adams, *The Hitchhiker's Guide to the Galaxy*

<sup>1415</sup> We previously showed, in chapter 4, that neural networks are relevant as reconstruction tools in  
<sup>1416</sup> JUNO. Even if they show worse performances, the combination with classical estimators could still  
<sup>1417</sup> bring improvements. We discussed the use of Convolutional Neural Network (CNN) in the previous  
<sup>1418</sup> chapter and their limitations, more specifically the limitation of the image as data representation for  
<sup>1419</sup> the experiment.

<sup>1420</sup> In this chapter we propose to use a Graph Neural Network (GNN), a Neural Network specialized to  
<sup>1421</sup> process graph as presented in section 3.2.3, to overcome those limitations.

<sup>1422</sup> **5.1 Motivation**

<sup>1423</sup> As explained in chapter 2 the JUNO sensors, the Large Photomultipliers (LPMT) and Small Photo-  
<sup>1424</sup> multipliers (SPMT), are arranged on a spherical plane. When trying to represent this plane as a  
<sup>1425</sup> 2D image, due to the inherent problem of the projection, some part of the image are distorted and  
<sup>1426</sup> part of the image do not have any physical meaning (see section 4.2.2). A way to represent the data  
<sup>1427</sup> without inducing deformation is the graph, an object composed of a collection of nodes and edges  
<sup>1428</sup> representing the relation between the nodes.

<sup>1429</sup> From this graph representation, we can construct a neural network that will process the data while  
<sup>1430</sup> keeping some interesting properties. For example the rotational invariance, i.e. the energy and  
<sup>1431</sup> radius of the event do change by rotation our referential. For more details see section 3.2.3. Graph  
<sup>1432</sup> representation also has the advantage to be able to encode global and higher order informations.

<sup>1433</sup> An approach was already proposed in JUNO by Qian et al. [42] where each nodes of the graph are  
<sup>1434</sup> like pixels, they represent geometric region of the detector and are connected with their neighbours.  
<sup>1435</sup> The LPMT informations are then aggregated on those nodes. The network then process the data  
<sup>1436</sup> using the equivalent of convolution but on graph [49].

<sup>1437</sup> In this work we want to take a step further in the graph representation by including the SPMT and  
<sup>1438</sup> including a maximum of raw informations.

## 1439 5.2 Data representation

1440 In an ideal world we would like to have every PMTs represented as node in the graph, each PMT  
 1441 being hit is an informations but the fact that PMTs were not hit is also an important information.  
 1442 It's by being aware of the whole of the system that we are able to give meaning to a subpart. As a  
 1443 reminder, in the Central Detector (CD), JUNO will posses 17612 LPMTs and 25600 SPMTs for a total  
 1444 of 43212 PMTs. This amount of information in itself is still manageable by modern computer if it  
 1445 were to be used in a neural network but when defining the relations between the nodes, it become a  
 1446 bit more tricky.

1447 Excluding self relation and considering the relation to be undirected, the edge from  $A$  to  $B$  is the  
 1448 same from  $B$  to  $A$ , the amount of necessary edges is given by  $\frac{n(n-1)}{2}$  which for 43212 PMTs amount  
 1449 for 933'616'866 edges. If we encode an information with double precision (64 bits) in what we call an  
 1450 adjacency matrix, each information we want to encode in the relation would consume 4 GB of data.  
 1451 When adding the overhead due to gradient computation during training, this would put us over the  
 1452 memory capacity of a single V100 gpu card (20 GB of memory). We could use parallel training to  
 1453 distribute the training over multiple GPU but we considered that the technical challenge to deploy  
 1454 this solution was not worth the trouble.

1455 The option of connecting PMTs node only to their neighbours could be tempting to reduce the num-  
 1456 ber of edge, but this solution does not translate well in term of internal representation in memory.  
 1457 Edges of sparsely connected nodes can be stored in efficient manner in a sparse matrix but the  
 1458 calculation in itself would often results in the concretization of the full matrix in memory, resulting  
 1459 in no memory gain during training.

1460 We finally decided of a middle ground where we define three *families* of nodes:

- 1461 — The core of the graph is composed of nodes representing geometric regions of the detector.  
   1462 We call those nodes **mesh** nodes. Those mesh nodes are densely connected to each other. We  
   1463 keep their number low to gain in memory consumption.
- 1464 — All the fired PMTs, that have been hit, will be represented as nodes. We call those node **fired**.  
   1465 Fired nodes are connected to the mesh they geometrically belong.
- 1466 — A final node which will hold global information about the detector and on which we will read  
   1467 the interaction vertex and energy. It's designated as the **I/O** node for input/output. This node  
   1468 will be connected to every mesh nodes.

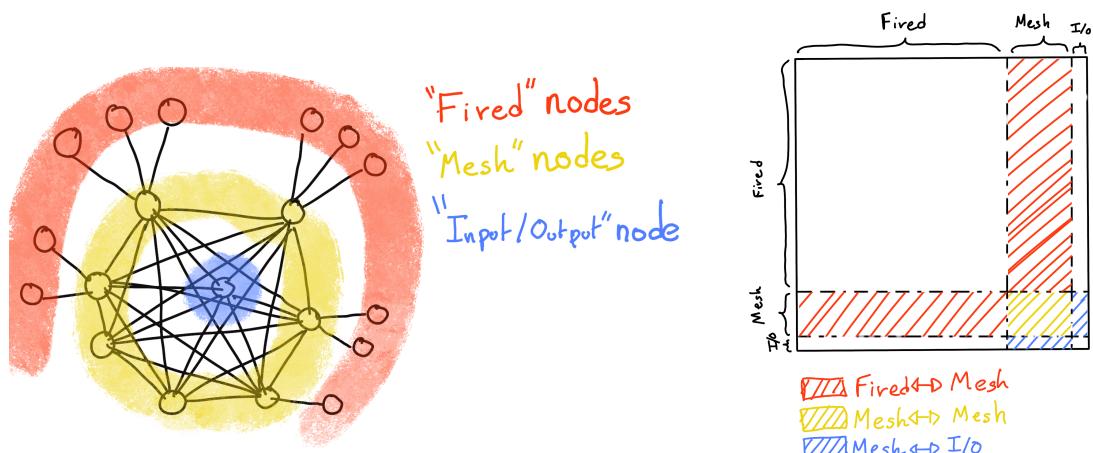
1469 Those nodes and their relations are illustrated in figure 5.1a. From this representation, we end up  
 1470 with three distinct adjacency adjacency matrix

- 1471 — A  $N_{\text{fired}} \times N_{\text{mesh}}$  adjacency matrix, representing the relations between fired and mesh. Those  
   1472 relations are undirected.
- 1473 — A  $N_{\text{mesh}} \times N_{\text{mesh}}$  adjacency matrix, representing the relation between meshes. Those relation  
   1474 are directed.
- 1475 — A  $N_{\text{mesh}} \times 1$  adjacency between the mesh and I/O nodes. Those relations are undirected.

1476 The adjacency matrix representing those relation is illustrated in figure 5.1b.

1477 The mesh segmentation is following the Healpix segmentation [76]. This segmentation offer the  
 1478 advantage that almost each mesh have the same number of direct neighbours and it guarantee that  
 1479 each mesh represent the same extent of the detector surface. The segmentation can be infinitely  
 1480 subdivided to provide smaller and smaller pixels. The number of pixel follow the order  $n$  with  
 1481  $N_{\text{pix}} = 12 \cdot 4^n$ . This segmentation is illustrated in figure 5.2. To keep the number of mesh small, we  
 1482 use the segmentation of order 2,  $N_{\text{pix}} = 12 \cdot 4^2 = 192$ .

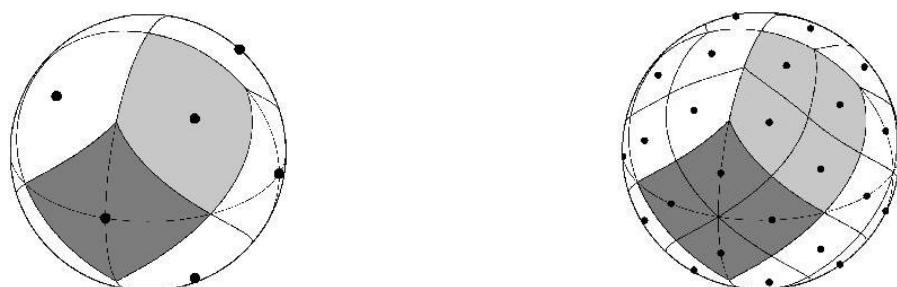
1483 We decided on having the different kind of nodes **mesh (M)**, **fired (F)** and **I/O** have different set of  
 1484 features. The features used in the graph are presented in figure 5.3. Most of the features are low level  
 1485 informations such as the charge or time information but we include some high order features such  
 1486 as



(A) Illustration of the different nodes in our graphs and their relations.

(B) Illustration of what a dense adjacency matrix would look like and the part we are really interested in. Because Fired → Mesh and Mesh → I/O relations are undirected, we only consider in practice the top right part of the matrix for those relations.

FIGURE 5.1

FIGURE 5.2 – Illustration of the healpix segmentation. **On the left:** A segmentation of order 0. **On the right:** A segmentation of order 1

- 1487 1.  $P_l^h$ : Is the normalized power of the  $l$ th spherical harmonic. For more details about spherical  
1488 harmonics in JUNO, see annex B.

2.  $\mathbb{A}$  and  $\mathbb{B}$  are informations that represent the likeliness of the interaction vertex to be on the segment between the center of two meshes.

$$\mathbb{A}_{ij} = (\vec{j} - \vec{i}) \cdot \frac{\vec{l}_1}{D_{ij}} + \vec{i} \quad (5.1)$$

$$\mathbb{B}_{ij} = \frac{Q_i}{Q_2} \left( \frac{\vec{l}_2}{\vec{l}_1} \right)^2 \quad (5.2)$$

$$l_1 = \frac{1}{2}(D_{ij} - \Delta t \frac{c}{n}) \quad (5.3)$$

$$l_2 = \frac{1}{2}(D_{ij} + \Delta t \frac{c}{n}) \quad (5.4)$$

1489 where  $\vec{i}$  is the position vector of the mesh  $i$ ,  $D_{ij}$  is the distance between the center of the meshes  
1490  $i$  and  $j$ ,  $Q_i$  the sum of charges on the mesh  $i$ ,  $\Delta t = t_i - t_j$  where  $t_i$  the earliest time on the mesh  
1491  $i$  and  $n$  the optical index of the LS.  $\mathbb{A}$  is the vertex between center of meshes distance ratio  
1492 between  $i$  and  $j$  based on the time information. For  $\mathbb{B}$ , the charge ratio evolve with the square  
1493 of the distance, so the mesh couple with the smallest  $\mathbb{B}$  should be the one with the interaction  
1494 vertex between its two center.

Nodes			Edges		
Fired	Mesh	I/O	Fired $\rightarrow$ Mesh	Mesh $\rightarrow$ Mesh (1) $\rightarrow$ Mesh (2)	Mesh $\rightarrow$ I/O
Q	$\langle Q_m \rangle$	$\langle x \rangle$	$X - X_m$	$X_{m1} - X_{m2}$	$\langle x \rangle - x_m$
t	$6Q_m$	$\langle y \rangle$	$Y - Y_m$	$Y_{m1} - Y_{m2}$	$\langle y \rangle - y_m$
X	$\min(t_m)$	$\langle z \rangle$	$Z - Z_m$	$Z_{m1} - Z_{m2}$	$\langle z \rangle - z_m$
Y	$\max(t_m)$	$\Sigma Q$	$t - \min(t)$	$\min(t_1) - \min(t_2)$	$\Sigma Q_m / \Sigma Q$
Z	$6t_m$	$P_l^h; l \in [0,8]$	$Q / \Sigma Q_m$	$\langle Q_{m1} \rangle - \langle Q_{m2} \rangle$ $\langle Q_{m1} \rangle + \langle Q_{m2} \rangle$	$\langle t_m \rangle$
LPMT: 1 SPMT: -1	X <sub>m</sub> Y <sub>m</sub> Z <sub>m</sub>			$D_{m1 \rightarrow m2}^{-1}$ $\mathbb{A}$ $\mathbb{B}$	

$Q$  is the charge [nPE]  
 $t$  is the time [ns]  
 $X, Y, Z$  are the coordinates [m]  
 $Q_m, t_m$  are the set of charge and time in a mesh  
 $X_m, Y_m, Z_m$  the coordinates of the center of the mesh  
 $\langle x \rangle, \langle y \rangle, \langle z \rangle$  the position of the charge barycenter.

FIGURE 5.3 – Features held by the nodes and edges in the graph.  $D_{m1 \rightarrow m2}^{-1}$  is the inverse of the distance between two mesh center. The features  $P_l^h$ ,  $\mathbb{A}$  and  $\mathbb{B}$  are detailed in section 5.2

1495 Because our different nodes do not have the same number of features, they live in different spaces.  
1496 Most library and public algorithms available are designed with node living in the same space in  
1497 mind, we thus had to develop a custom message passing algorithm.

### 5.3 Message passing algorithm

1499 As introduced in previous section and in figure 5.3, our graphs nodes and edges will have different  
1500 number of features depending on their nature, meaning that we cannot have a single message passing  
1501 function. We thus need to define a message passing function for each transition inside or outside

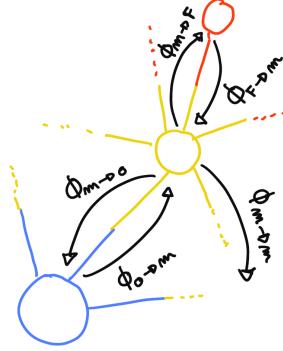


FIGURE 5.4 – Illustration of the different update function needed by our GNN

1502 a family. Using the notation presented in section 3.2.3

$$n_i^{k+1} = \phi_u(n_i^k, \square_j \phi_m(n_i^k, n_j^k, e_{ij}^k)); n_j \in \mathcal{N}_i' \quad (5.5)$$

we need to define

$$\phi_{u;f \rightarrow m} \phi_{m;f \rightarrow m} \quad (5.6)$$

$$\phi_{u;m \rightarrow f} \phi_{m;m \rightarrow f} \quad (5.7)$$

$$\phi_{u;m \rightarrow m} \phi_{m;m \rightarrow m} \quad (5.8)$$

$$\phi_{u;m \rightarrow io} \phi_{m;m \rightarrow io} \quad (5.9)$$

$$\phi_{u;io \rightarrow m} \phi_{m;io \rightarrow m} \quad (5.10)$$

1503 to update the nodes after each layers as illustrated in figure 5.4. We would also need update function  
 1504 for the edges but for the sake of technical simplicity in this work, we will limit ourself to the nodes  
 1505 update. A wide variety of message passing algorithm exists, with different use cases and goal behind  
 1506 them. To stay generalist and to match to the best the specificity of our architecture, we implement  
 1507 the following algorithm:

$$\phi_u := I_{i'}^{n'} = I_i^n A_{i',e}^i W_n^{e,n'} + I_i^n S_n^{n'} + B^{n'} \quad (5.11)$$

1508 using the Einstein summation notation.  $I_i^n$  is the tensor holding the nodes informations with  $i$   
 1509 the node index and  $n$  the feature index.  $n$  represent the features of the previous layer and  $n'$  the  
 1510 features of this layer.  $A_{i',e}^i$  is the adjacency tensor, discussed in the previous section, representing the  
 1511 connection between the node  $i'$  and the node  $i$ , each connections holding the features indexed by  $e$ .  
 1512 The learnable weights are composed of:

- 1513 — The tensor  $W_n^{e,n'}$  which represent the passage from the previous feature domain  $n$ , the previous  
 1514 layer, to the current domain  $n'$ , this layer, knowing the relation  $e$ .
- 1515 —  $B^{n'}$  which is a learnable bias tensor on the new features  $n'$ .
- 1516 —  $S_n^{n'}$  which can be viewed as a self loop relation where the node update itself based on the  
 1517 previous layer informations.

1518 If a node have neighbours in different families, the different  $I_{i'}^{n'}$  coming from the different  $\phi_u$  are  
 1519 summed.

$$I_{i'}^{n'} = \sum_{\mathcal{N}} \phi_{u,\mathcal{N}} \quad (5.12)$$

1520 where  $\mathcal{N}$  are the neighbouring family and  $\phi_{u,\mathcal{N}}$  the update function between the target node family  
 1521 and the neighbour  $\mathcal{N}$  family.

1522 We thus have a  $S$ ,  $W$  and  $B$  for each of the  $\phi_u$  function we defined above. The IAW sum can be seen

as the  $\phi_m$  function and  $IS + B$  as the second part of the  $\phi_u$  function. Interestingly, the number on learnable weight in those layer is independent of the number of nodes in each family and depends solely on the number of features on the nodes and the edges.

The expression above only update the node features. We could update the edges, using the results of  $\phi_m$  for example, but for technical simplicity we only update the nodes and keep the edges constant.

This operation of message passing is the constituent of our message passing layer, designed in this work as *JWGLayer*. To this layer, we can adjoin an activation function such as *PReLU*

$$I_i^{n'} = PReLU \left( \sum_{\mathcal{N}} I_i^n A_{i',e}^i W_n^{e,n'} + I_i^n S_n^{n'} + B^{n'} \right) \quad (5.13)$$

## 5.4 Data

For this study we will be using a 1M positrons event dataset, uniformly distributed in energy with  $E_k \in [0, 9]$  MeV and uniformly distributed in the detector. Those events come from the JUNO official simulation version J23.0.1-rc8.dc1 (released the 7th January 2024). All the event are *calib* level, with simulation of the physics, electronics, digitizations and triggers. 900k events will be used for the training, 50k for validation and loss monitoring and 50k for the results analysis in section 5.8. Each events is between 2k and 12k fired PMTS, resulting in fired nodes being the largest family in our graphs in all circumstances as illustrated in figure 5.5c.

As expected, by comparing the scale between the figure 5.5a and 5.5b we see that the LPMT system is predominant in term of informations in our data. The number of PMT hits grow with energy but do not reach 0 for low energy event due to the dark noise contribution which seems to be around 1000 hits per event for the LPMT system (left limit of figure 5.5a) and around 15 hits per event for the SPMT system (left limit of figure 5.5b) which is consistent with the results show in section 4.2.2.

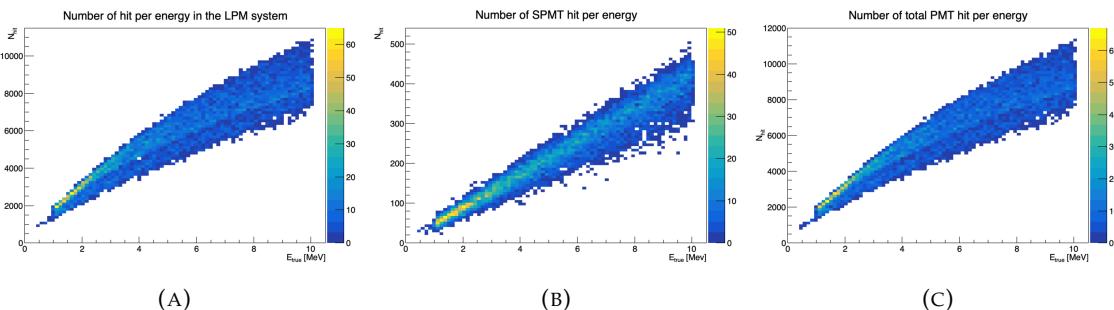


FIGURE 5.5 – Distribution of the number of hits depending on the energy. **On the right:** for the LPMT system. **In the middle :** for the SPMT system. **On the left:** For both system.

The structure seen in the distribution in figure 5.5a comes from the shape of the number of hits depending on the radius as shown in figures 5.6a and 5.6b where the number of hit decrease with radius. It is important to understand that this is not representative of the number of PE per event and the decrease in hits over the radius means that the PE are just more concentrated in a smaller number of PMTs.

No quality cut is applied here, we rely only on the trigger system. It means that event that would not trigger are not present in the dataset but for events that triggered twice, it happens rarely, the two trigger are considered as two separate event.

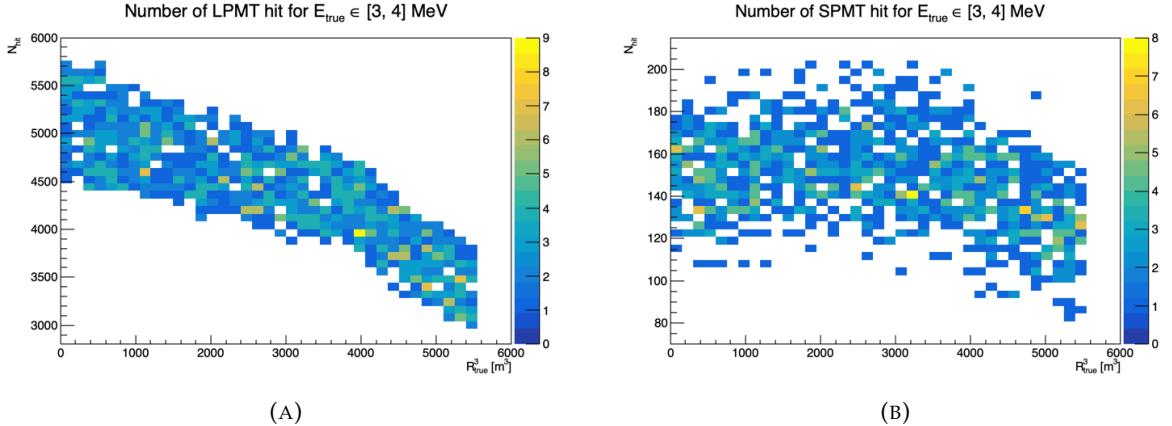


FIGURE 5.6 – Distribution of the number of hits depending on the radius. **On the right:** for the LPMT system. **On the right :** for the SPMT system. To prevent the superposition of structure of different scales we limit ourselves to the energy range  $E_{\text{true}} \in [0, 9]$ .

## 5.5 Model

In this section we'll discuss the different layer composing the final version of the model. As introduced above, each JWGLayer is defined by the number of features on the nodes and edges of the output graph, assuming it takes as input the graph from the precedent layer. For simplicity, when discussing a graph configuration, it will be presented as follow: {  $N_f$ ,  $N_m$ ,  $N_{IO}$ ,  $N_{f \rightarrow m}$ ,  $N_{m \rightarrow m}$ ,  $N_{m \rightarrow f}$  } where

- $N_f$  is the number of feature on the fired nodes.
  - $N_m$  is the number of features on the mesh nodes.
  - $N_{IO}$  is the number of features on the I/O node.
  - $N_{f \rightarrow m}$  is the number of features on the edges between the fired and mesh nodes.
  - $N_{m \rightarrow m}$  is the number of features on the edges between two mesh nodes.
  - $N_{m \rightarrow f}$  is the number of features on the edges between the mesh nodes and the I/O node.
- Because we do not change the number of features on the edges, we can simplify the notation to {  $N_f$ ,  $N_m$ ,  $N_{IO}$  }. As an example, the input graph configuration, following the figure 5.3, is { 6, 8, 13, 5, 8, 5 } or, without the edge features, { 6, 8, 13 }.

The final version of the model, called JWGV8.4.0 is composed of

- An JWGLayer, converting the input graph { 6, 8, 13 } to { 64, 512, 2048 } with a PReLU activation function.
- 3 resnet layers, each of them composed of
  1. 2 JWG layers with a PReLU activation function. They do not change the dimension of the graph
  2. A sum layer that sums the features in the input graph with the one computed from the JWG layers
- A flatten layer that flatten the features of the I/O and mesh nodes in a vector.
- 2 fully connected layers of 2048 neurons with a PReLU activation function.
- 2 fully connected layers of 512 neurons with a PReLU activation function.
- A final, fully connected layer of 4 neurons acting as the output of the network.

A schematic of the model is presented in figure 5.7.

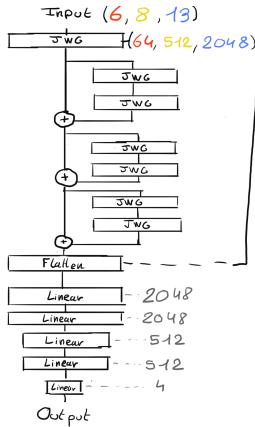


FIGURE 5.7 – Schema of the JWGv8.4.0 architecture, the colored triplet is the graph configuration after each JWG layers

## 1579 5.6 Training

1580 The optimizer used for training is the Adam optimizer and default hyperparameters ( $\beta_1 = 0.9$ ,  
 1581  $\beta_2 = 0.999$  and  $\epsilon = 1e-8$ ) with a learning rate  $\lambda = 1e-8$ . The training last 200 epochs of 800 steps.  
 1582 We use a batch size of 8. The learning rate is constant during the first 20 epochs then exponentially  
 1583 decrease with a rate of 0.99. The model saved is the model with the best validation loss during the  
 1584 training. The validation is computed over a single batch.

## 1585 5.7 Optimization

1586 Due to the extensive training time, up to 90h per training on the more complex architectures, and  
 1587 the heavy memory consumption of the models that would often exceed the 20GB limit of the V100,  
 1588 random search was not a realistic approach to the hyper optimisation. We were able to extend the  
 1589 memory limit to 40GB thanks to a local A100 GPU card available inside the laboratory.

1590 The hyperparameters optimization was thus done “by hand”, by looking at the results of the previous  
 1591 training and tinker hyperparameters that seems to play a role in the training. During this process,  
 1592 the model went into some heavy refactoring. At the start, the message passing algorithm was not  
 1593 the one presented above but each  $\phi_u$  and  $\phi_m$  function were FCDNN. Due to problems of memory  
 1594 consumption and gradient vanishing we pivoted to the message passing algorithm presented above.

1595 Even the features on the graph went under investigation. With the addition of high level observables  
 1596 to the mesh and I/O nodes and edge, there was too much possibility to test everything. We went  
 1597 with the decision to keep the raw observables in the fired and for the higher order observables  
 1598 we tried to take the one that would be difficult for the NN to reconstruct or at least would need  
 1599 multiple layer to reproduce. Basically, because the operation in the JWGLayer are linear operation,  
 1600 any variables dependent on order > 1 of the input would be candidates. This is why we introduce  
 1601 standard deviation,  $A$ ,  $B$  and  $P_l^h$  for example.

1602 Substantial effort went to the data processing process, transforming JUNO files into understandable  
 1603 graphs, before the training. Due to the volatile nature of the graph features during the optimization,  
 1604 the current code do not take preprocessed data and compute the observables, adjacency matrix,  
 1605 etc... on the fly. This data processing is carried out on the CPU, using a worker pool to allow for  
 1606 multiprocess. The raw data are coming from ROOT file produced by the collaboration software,

1607 the Event Data Model (EDM) used internally by the collaboration [77] had to be interfaced to our  
 1608 code, interface maintained through the evolution of the collaboration software. For the harmonic  
 1609 power calculation, we migrated from the Healpix library to Ducc0 [78] for a more fine control of the  
 1610 multithreading.

1611 Over the course of the project, the model went over more than 60 different configurations to end on  
 1612 the one presented in this chapter.

## 1613 5.8 Results

1614 The reconstruction performance of “JWGv8.4” are presented in figure 5.9 and compared to the “Omlil-  
 1615 rec” algorithm, the official IBD reconstruction algorithm in JUNO. Omlilrec is based on the QTMLR  
 1616 reconstruction method that was presented in section 2.6.

1617 We also present the results of the optimal variance combination of the two algorithm labelled as  
 1618 “JWG 8.4 x Omlilrec” where the reconstructed target  $\hat{\theta}_{\text{target}}$  is the weighted sum of the result of the  
 1619 two estimator JWGv8.4  $\theta_J$  and Omlilrec  $\theta_O$ .

$$\hat{\theta} = \alpha\theta_J + (1 - \alpha)\theta_O; \alpha \in [0, 1] \quad (5.14)$$

1620 For more details about the combination and the computation of  $\alpha$ , refer to annex A.2.

1621 One thing that need to be addressed before discussing results is that the Omlilrec algorithm do not  
 1622 reconstruct the deposited energy  $E_{\text{dep}}$  but reconstruct the visible energy  $E_{\text{vis}}$ . The difference between  
 1623 those two different observables comes from the event-wise and channel-wise non-linearity, presented  
 1624 in 2.3. The multiples energy observables are already discussed in section 4.4. For the following  
 1625 results, the systematic bias of Omlilrec that appear due, to the comparison to  $E_{\text{true}}$  instead of  $E_{\text{vis}}$  is  
 1626 corrected using a 5th degree polynomial

$$\frac{E_{\text{true}}}{E_{\text{rec}}} = \sum_{i=0}^5 P_i E_{\text{true}}^i \quad (5.15)$$

1627 The fitted distribution and the corresponding fit is presented in figure 5.8. The value fitted for this  
 1628 correction are presented in table 5.1.

$P_0$	$1.24541 +/- 0.00585121$
$P_1$	$-0.168079 +/- 0.00716387$
$P_2$	$0.0489947 +/- 0.00312875$
$P_3$	$-0.00747111 +/- 0.000622003$
$P_4$	$0.000570998 +/- 5.7296e-05$
$P_5$	$-1.72588e-05 +/- 1.98355e-06$

TABLE 5.1 – Parameters of the 5th degree polynomial used to correct Omlilrec  
 reconstructed energy.

1629 Overall, energy and radius resolutions are not on par with Omlilrec. We see from the energy de-  
 1630 pendent energy resolution in fig 5.9a that our resolution is a bit more than twice the resolution of  
 1631 Omlilrec and the combination brings no improvements. Same observation for the energy resolution  
 1632 depending on the radius.

1633 The radius resolution, presented in the figures 5.9c, 5.9d, 5.9e and 5.9f is much worse than the  
 1634 Omlilrec one. This comes a bit as a surprise, as the energy reconstruction is dependent on the  
 1635 vertex reconstruction to correct for the non-uniformity and non-linearity effect. This mean that  
 1636 either the GNN could outperform the classical methods if the vertex was correctly reconstructed,

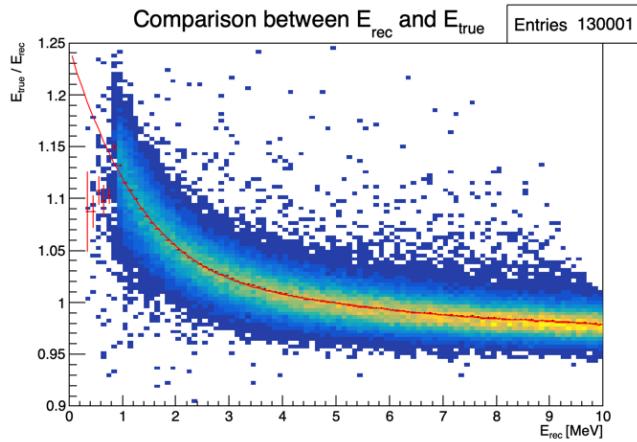


FIGURE 5.8 – Comparison between Omilrec  $E_{rec}$  and the true energy  $E_{true}$ . The profile of the distribution  $E_{true} / E_{rec}$  vs  $E_{rec}$  is fitted with a 5th degree polynomial.

1637 or that somewhere the GNN reconstruct the vertex correctly but has trouble to formulate it in x,y,z  
1638 coordinates on the latest layer.

1639 The GNN behaviours are close to Omilrec, indicating that the same information is used in the same  
1640 way by both algorithms, just that the GNN seems to be less fine-tuned than Omilrec. If the precedent  
1641 reasoning is true, it would mean that by adding more parameters, more layer or a higher pixelisation  
1642 of the Healpix representation, the GNN could reach Omilrec performances.

## 1643 5.9 Conclusion

1644 In this chapter, I present a proposition for a GNN architecture to reconstruct the energy and position  
1645 of the prompt signal of an IBD interaction. The GNN is not competitive in terms of resolution with  
1646 the more classical method Omilrec, which is the state of the art reconstruction method for IBD in the  
1647 JUNO collaboration, but show encouraging results that could be exploited by going further in the  
1648 optimisation of the hyper parameters. The message passing algorithm is still pretty naive and could  
1649 probably be refined for JUNO's need.

1650 Another possible improvement is to find a way to increase the Healpix pixelisation. Through our  
1651 different work on reconstruction and by looking at the different classical methods, it seems that  
1652 the time information is crucial for the vertex reconstruction, and thus for the energy reconstruction.  
1653 While we are keeping every raw informations about the fired PMTs, it is possible that the aggregation  
1654 on mesh nodes could cause the information loss and it has been noticed that allowing more channels  
1655 to the hidden layer mesh nodes improve the resolution. This observation can be compared to the  
1656 convolutional GNN presented section 2.6.3 that has similar performance with the classical method  
1657 with an order 5 Healpix segmentation resulting in 3072 pixels, comforting the need of a finer pixeli-  
1658 sation, or more parameters dedicated to aggregation through an increase of channels on the mesh  
1659 nodes. Both of those improvements require some heavy memory optimisations, distributed training  
1660 or more powerful hardware to address the memory consumption issue.

1661 A final possible improvement would be to go further in the proximity of raw information. The charge  
1662 and time used in the PMTs are extracted from a waveform, we could imagine a world where the full  
1663 PMT waveform in the trigger window would be set of channels on the PMT node.

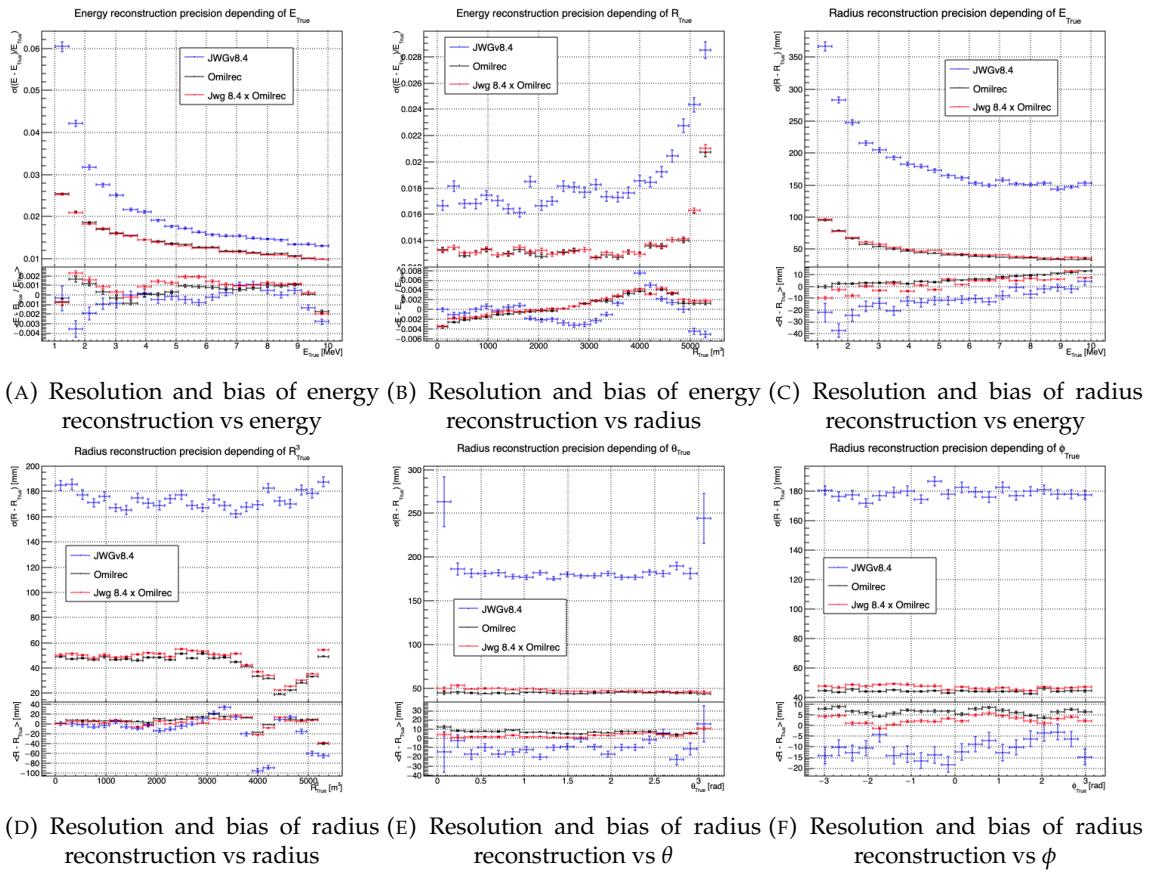


FIGURE 5.9 – Reconstruction performance of the Omilrec algorithm based on QTMLE presented in section 2.6, JWGv8.4 presented in this chapter and the combination between the two as presented in section 4.4.2. The top part of each plot is the resolution and the bottom part is the bias.



<sup>1664</sup> Chapter 6

<sup>1665</sup> **Reliability of machine learning  
methods**

<sup>1666</sup>

*"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*

<sup>1667</sup>



<sup>1668</sup> **Chapter 7**

<sup>1669</sup> **Joint fit between the SPMT and LPMT  
spectra**

<sup>1671</sup> *"We demand rigidly defined areas of doubt and uncertainty!"*

*Douglas Adams, The Hitchhiker's Guide to the Galaxy*

<sup>1672</sup> JUNO is an experiment of precise measurements, where we try to observe small fluctuation in the en-  
<sup>1673</sup> ergy spectrum and have the ambition to achieve sub-percent precision on the oscillation parameters  
<sup>1674</sup> measurement. It is crucial to understand extremely well the reconstruction and the effect we are deal-  
<sup>1675</sup> ing with. The challenge reside in the standard technology used in the detector, scintillator observed  
<sup>1676</sup> by PMT, but in a scale never seen before, for scintillator volume as for its PMT. Understanding every  
<sup>1677</sup> effect that goes in the detector can become extremely complicated. Being able to compare the results  
<sup>1678</sup> of the same experiment with two systems is thus precious, this is the origin the dual calorimetry with  
<sup>1679</sup> the LPMT and SPMT system.

<sup>1680</sup> The resolution and bias of the reconstruction needs to be extremely well characterized: the target  
<sup>1681</sup> resolution of 3% [50] is unprecedented as is required to be able to distinguished between Normal  
<sup>1682</sup> Ordering (NO) and Inverse Ordering (IO). The non-linearity uncertainty needs to be contained under  
<sup>1683</sup> 1% or the risk appear to measure the wrong ordering [27].

<sup>1684</sup> One of the possible non-linearity source, that will be used as a reference source in this chapter, is the  
<sup>1685</sup> charge non-linearity (QNL) that will be detailed in next section. The dual calorimetry will be able to  
<sup>1686</sup> solve this problem, using calibrations method and measurements that will be used to correct it [27].

<sup>1687</sup> More generally, comparing the results of the two system will allow to detect potential issues on  
<sup>1688</sup> the calibration or reconstruction. This is done in this thesis by comparing directly the spectra and  
<sup>1689</sup> oscillation parameters measurements of the two system.

<sup>1690</sup> The study of the independents results of the two system can bring some informations [79] but this is  
<sup>1691</sup> missing the important correlation that should be present between the two system: they see the same  
<sup>1692</sup> events, in the same scintillator, their bound to be correlated. We explore in this chapter a preliminary  
<sup>1693</sup> study of the impact of those correlation via multiple methods and the impact of QNL at various  
<sup>1694</sup> degree.

<sup>1695</sup> In the next section will discuss the motivations behind this study. In section 7.2, I present the ap-  
<sup>1696</sup> proaches and assumptions in this study. In section 7.3 I present the fit framework used and following  
<sup>1697</sup> in section 7.4 the technical improvement brought and the difficulties faced during the development.  
<sup>1698</sup> To end this chapter I present the results in 7.5 and discuss the conclusion and perspectives in 7.6.

## 1699 7.1 Motivations

### 1700 7.1.1 Discrepancies between the SPMT and LPMT results

1701 As discussed in the introduction of this chapter, the SPMT system and the LPMT system will observe  
 1702 the same events. This mean that, after calibration, if the two system show significant differences  
 1703 in their results this is the signal of potential overlook of an effect or problem. Being able to detect  
 1704 such differences is thus crucial as, as discussed above, even the smallest deviation from our model  
 1705 could lead to the impossibility to measure the Mass Ordering (MO) or even worse, wrong our  
 1706 measurement.

1707 The two system are expected to have the same sensitivity to the oscillation parameters  $\theta_{12}$  and  $\Delta m_{21}^2$   
 1708 [11]. We will thus rely on the measurement of those two parameters to detect potential discrepancies.

1709 We could just look at the value and compare them to the estimated independent error of the two  
 1710 system, but we believe and will demonstrate in this chapter that the independent study of the two  
 1711 system is missing a lot of informations, and that, by taking into account the statistic and systematic  
 1712 correlations between the two systems, we can produce much more powerful statistical test.

1713 Our work in this chapter is to develop such tools. The first step is, of course, to verify that in the  
 1714 case of no discrepancies, the results are coherent with the independent analysis. This will give us  
 1715 the distribution of those statistical test in absence of discrepancies. When we will have real data, we  
 1716 will be able to compare it to those distribution to obtain a p-value characterizing the absence of those  
 1717 potential discrepancies.

1718 To evaluate the power of our methods, we need to simulate a concrete discrepancy. We decide to  
 1719 study a plausible effect, the Charge Non-Linearity (QNL) that is detailed next section, but the goal of  
 1720 those tools is to be discrepancy agnostic, as those discrepancies could come from a variety of source  
 1721 (calibration issue, insufficient simulation tuning, etc...)

### 1722 7.1.2 Charge Non-Linearity (QNL)

1723 The CD energy response is subject of two kind on non-linearity, the first one is the LS response  
 1724 non-linearity where the LS photo-production is not linear with the deposited energy as illustrated  
 1725 in figure 2.12a. The second one is the LPMT response non-linearity where the charge read from the  
 1726 LPMT is not linear in respect to the number of collected Photo-Electron (PE) (see section 2.3).

1727 The LS non-linearity comes from physics sources. Particles interaction in the LS will produce mainly  
 1728 scintillation light as discussed in section 2.2.2 but will also produce some cherenkov light (< 10% of  
 1729 the collected light). Both mechanisms posses intrinsic non-linearity, for the Cherenkov emission it's  
 1730 dependent on the velocity of charged particle velocity while the scintillation photon-yield follow a  
 1731 so-called Birk's law with a "quenching" effect depending on the energy and type of particle [16]. This  
 1732 results in a event-wise QNL.

1733 The LPMT response non-linearity can come from sheer saturation when subject to a high rate of  
 1734 photon inducing a gain non-linearity or come from readout effects such as the electronic noise, the  
 1735 overshoot, the integration time window and even the waveform algorithm. All of those effects results  
 1736 in a channel-wise QNL.

1737 Precedent studies [27] suggest a model to emulate this non-linearity response that will be used in this  
 1738 work. We thus define the channel wise non-linearity that would be applied to each LPMT readout

$$\frac{Q_{rec}}{Q_{true}} = \frac{-\gamma_{qnl}}{9} Q_{true} + \frac{\gamma_{qnl} + 9}{9} \quad (7.1)$$

1739 where  $Q_{rec}$  is the reconstructed number of PE by the PMT,  $Q_{true}$  is true number of PE that hit the  
 1740 PMT, and  $\gamma_{qnl}$  is a factor representing the amplitude of the non-linearity.

1741 We also define an event-wise non-linearity characterized by

$$\frac{E_{vis}}{E_{true}} = \frac{-\alpha_{qnl}}{9} E_{true} + \frac{\alpha_{qnl} + 9}{9} \quad (7.2)$$

1742 where  $E_{vis}$  is the visible energy that can be collected by the detector and  $E_{true}$  is the true deposited  
 1743 energy. An example of the effect of such event-wise QNL is presented in figure 7.1.

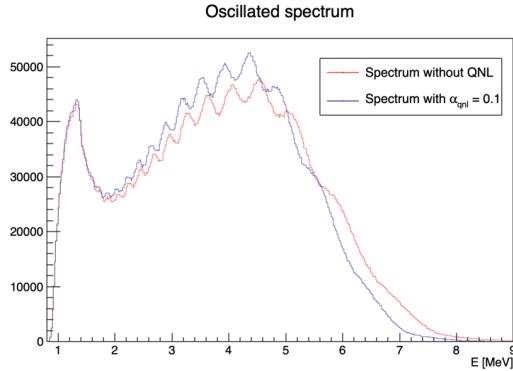


FIGURE 7.1 – Two oscillated spectrums of  $1e7$  event expected in JUNO. In red the spectrums without supplementary QNL. In blue the same spectrum but where an event-wise QNL  $\alpha_{qnl} = 10\%$  is added.

1744 Using 1M from the JUNO official simulation J23.0.1-rc8.dc1 (released the 7th January 2024), we  
 1745 simulated events up to the photon collection in LPMTs and simulated an additional channel-wise  
 1746 QNL by applying the equation 7.1 to the number of collected photons.

1747 In figure 7.2a we show the distribution of the ratio  $\frac{Q_{rec}}{Q_{true}}$  for central events ( $R < 4m$ ) and different  
 1748 value of  $\gamma_{qnl}$ . In figure 7.2a we show the mean of this distribution as function of the energy. We show  
 1749 effective  $\alpha_{qnl}$  for each value of  $\gamma_{qnl}$ . We see that using the event-wise QNL is equivalent to the mean  
 1750 behavior of using channel-wise QNL.

1751 When using channel-wise non-linearity, we need to simulate a number of PE per LPMT, the process  
 1752 can be quite tedious if we want a realistic simulation so in this study we are only using event-wise  
 1753 non-linearity to make the process simpler. This event-wise non-linearity will be characterized by  
 1754  $\alpha_{qnl}$ .

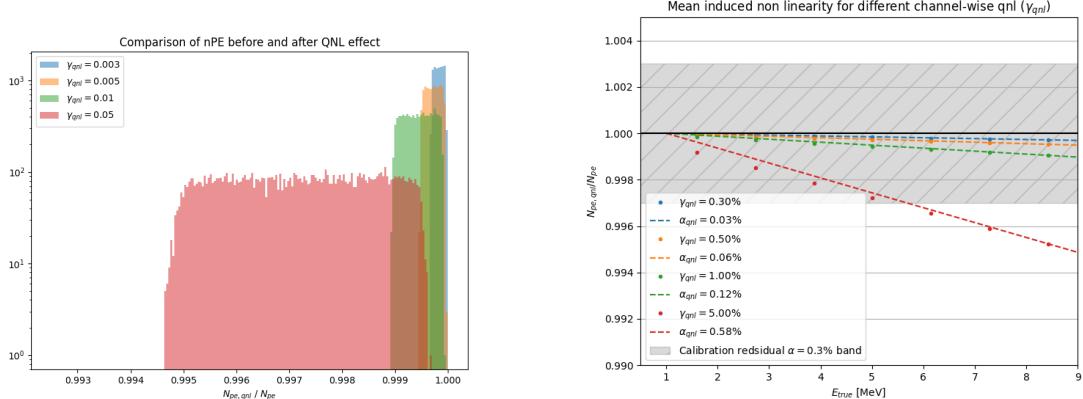
## 1755 7.2 Approach

1756 In this section we detail the testing procedure behind each of our tools.

### 1757 7.2.1 Data production

#### 1758 IBD spectra

1759 The first step is the generation of data on which we run our tools. In this study we are producing  
 1760 monte-carlo toy. In each toy we generate an  $\bar{\nu}_e$  energy spectrum coming from the Taishan, Yangjiang  
 1761 and Dayabay reactor power plant, the reactor used as source for the NMO analysis. The parameters



(A) Distribution of ratio of collected nPE after the additional QNL over the number of nPE that would be collected for different  $\gamma_{qnl}$ . We select event with an interaction radius  $R < 4\text{m}$  to not be affected by the non-uniformity.

(B) Ratio of collected nPE after the additional QNL over the number of nPE that would be collected at different energies. We select event with an interaction radius  $R < 4\text{m}$  to not be affected by the non-uniformity. The dots represent the mean of the distribution in figure 7.2a and the dashed line are the equivalent event-wise non-linearity from eq 7.2. The hatched zone is the residual non-linearity expected after calibration [29].

FIGURE 7.2

of the reactors come from JUNO official database, shared among all physics analysis, the JUNO common inputs. This is the starting point of the two observed spectrum, the LPMT and SPMT spectrum. On those spectrum, we add the physics effect such as the LS non-linearity etc... (more details in section 7.3.1). On top on those effect, we apply the reconstruction resolution of each systems to their respective spectra. We end up with the two LPMT and SPMT spectrum.

We will study the effect of exposure on our methods at different threshold: 100 days, 1 year, 2 year and finally 6 years which is the nominal data taking period for the NMO analysis.

Those spectra are generated for different QNL,  $\alpha_{qnl} = 0$ , no spectrum distortion, and for  $\alpha_{qnl} \in \{0.01, 0.005, 0.003, 0.002, 0.001\}$ . As a reminder, the calibration guarantee a residual event-wise non-linearity  $\alpha_{qnl} \leq 0.003$  [29].

The first test does not require any fitting and we are just comparing the LPMT and SPMT spectra using the expected statistical correlation matrix in the case  $\alpha_{qnl} = 0$ , for details about the generation of this correlation matrix, refer to section 7.5.2. This test is the spectrum  $\chi^2$  test or  $\chi^2_{spe}$  test. In this test we produce a  $\chi^2$  representing show the incompatibility between the LPMT and SPMT spectra:

$$\Delta_i = h_{L,i} - h_{S,i} \quad (7.3)$$

$$U = AVA^T \quad (7.4)$$

$$\chi^2_{spe} = \vec{\Delta}^T U^{-1} \vec{\Delta} \quad (7.5)$$

Where  $h_{L,i}$  and  $h_{S,i}$  are the content of the  $i$ th bin of the LPMT and SPMT bin respectively.  $V$  is the covariance matrix of the LPMT + SPMT spectrum.  $A$  is a transfers matrix such as:

$$A_{ij} = \frac{\partial \Delta_i}{\partial h_j} = \frac{\partial (h_{L,i} - h_{S,i})}{\partial h_j} \quad (7.6)$$

1774 Thus  $A_{ij} = 1$  if  $i = j$  and  $A_{ij} = -1$  if  $j$  is the SPMT bin corresponding to the  $i$  LPMT bin.

1775 This  $\chi^2_{spe}$  is minimal when the statistic between the bins of the LPMT and SPMT spectra follow the  
 1776 covariance matrix  $V$ . By looking at the distribution of this  $\chi^2_{spe}$  when  $\alpha_{qnl} = 0$  we can produce  
 1777 p-values for the values found when  $\alpha_{qnl} \neq 0$ .

1778 **Background spectra**

1779 The JUNO common inputs provides only LPMT background spectra. Those background spectra are  
 1780 already smeared by the LPMT resolution and thus needs to be regenerated to be smeared following  
 1781 the SPMT resolution. Luckily the SPMT resolution is greater than the LPMT, allowing us to oversmear  
 1782 the spectrum using

$$S(E) = L(E) * \frac{1}{\sqrt{|\Delta\sigma^2|}\sqrt{2\pi}} e^{-\frac{E^2}{2|\Delta\sigma^2|}}; |\Delta\sigma^2| = \sigma_L^2 - \sigma_S^2 \quad (7.7)$$

1783 Where  $S(E)$  is the SPMT spectrum,  $L(E)$  the LPMT spectrum,  $\sigma_L$  and  $\sigma_S$  the LPMT and SPMT resolution  
 1784 respectively. This formula is valid under the assumption that the LPMT and SPMT smearing are  
 1785 gaussian and that the LPMT and SPMT have the same bias. Those two assumption are valid in the  
 1786 context of the IBD spectrum production as detailed in section 7.3.1. The demonstration of equation  
 1787 7.7 can be found in annex C.

1788 **7.2.2 Individual fits**

1789 Each of the spectra, LPMT and SPMT, are then fitted individually with and without the presence of  
 1790 QNL over multiples toys. The results allow us to compute the correlation between the oscillations  
 1791 parameters measured by both of the systems when there is no QNL allowing us to compute a  $\chi^2$   
 1792 representing the compatibility between the oscillation parameters measured by the SPMT and by the  
 1793 LPMT. Because the SPMT system is not sensible to the oscillation parameters  $\Delta m_{31}^2$  and  $\theta_{13}$ , the test  
 1794 is only done on the oscillation parameters  $\theta_{12}$  and  $\Delta m_{21}^2$ . We can thus produce the individual chi  
 1795 square  $\chi^2_{ind}$

$$\Delta_\lambda = \lambda_L - \lambda_S \quad (7.8)$$

$$\vec{\Delta} = [\Delta_{\theta_{12}} \Delta_{\Delta m_{21}^2}] \quad (7.9)$$

$$U = A V A^T \quad (7.10)$$

$$\chi^2_{ind} = \vec{\Delta}^T U^{-1} \vec{\Delta} \quad (7.11)$$

1796 where  $\lambda_L$  and  $\lambda_S$  are the measured parameters by the LPMT and SPMT systems respectively. The  
 1797 different  $\lambda$  considered are  $\theta_{12}$  and  $\Delta m_{21}^2$ .  $V$  here is the  $4 \times 4$  covariance matrix between the param-  
 1798 eters  $\theta_{12,L}$ ,  $\Delta m_{21,L}^2$ ,  $\theta_{12,S}$  and  $\Delta m_{21,S}^2$ .  $A$  is the transfer matrix that allow to compute the covariance  
 1799 matrix de  $\vec{\Delta}$  from  $V$  following

$$A_{ij} = \frac{\partial \Delta_i}{\partial j}; i \in \{\theta_{12}, \Delta m_{21}^2\}; j \in \{\theta_{12,L}, \Delta m_{21,L}^2, \theta_{12,S}, \Delta m_{21,S}^2\} \quad (7.12)$$

1800 Same as described above, by comparing the distribution of this  $\chi^2_{ind}$  when  $\alpha_{qnl} = 0$  and  $\alpha_{qnl} \neq 0$  we  
 1801 can compute the power of this test in term of p-values.

$\sin^2(2\theta_{12})$	$\Delta m_{21}^2$	$\Delta m_{31}^2$	$\sin^2(2\theta_{13})$
$0.851^{+0.020}_{-0.018}$	$7.53 \pm 0.18 \times 10^{-5} \text{ eV}^2$	$2.5283 \pm 0.034 \times 10^{-3} \text{ eV}^2$	$0.8523 \pm 0.00268$

TABLE 7.1 – Nominal PDG value [16]. All value are reported assuming Normal Ordering.

### 7.2.3 Joint fit

#### Standard joint fit

The final step is to produce a joint fit between the two spectra. In this case we adjust our model, the oscillated spectrum, over two spectrum at the same time. We minimize a  $\chi^2_{joint}$  defined over the two spectra, the LPMT and SPMT one

$$\Delta_i = D_i - T_i \quad (7.13)$$

$$\chi^2_{joint} = \vec{\Delta}^T V^{-1} \vec{\Delta} \quad (7.14)$$

where  $D_i$  is the content of the  $i$ th bin measured, from the data, and  $T_i$  is the theoretical number of event in this bin.  $V$  is the covariance matrix of our spectrum.

$T$  is the fitted function and depend on multiple parameters

- The oscillation parameters  $\theta_{12}$ ,  $\Delta m_{21}^2$ ,  $\theta_{13}$  and  $\Delta m_{31}^2$ . Those parameters can be free, have a pull term or be fixed during the fit.
- We take into account in the data production the matter effect and parametrize them by the parameter  $\rho$  which is the effective rock density between the reactors and the experiment. Same as the oscillation parameters, this parameter can be free, pulled or fixed.
- The exposure of the considered data which is just a normalization factor in front of the theoretical spectrum. This parameter is fixed at the start of the fit.

In the standard joint fit, the free parameters are  $\sin^2(2\theta_{12})$ ,  $\Delta m_{21}^2$  and  $\Delta m_{31}^2$ .  $\sin^2(2\theta_{13})$  is fixed to the PDG favorite value. For simplicity, we refer to  $\sin^2(2\theta_{12})$  and  $\sin^2(2\theta_{13})$  as  $\theta_{12}$  and  $\theta_{13}$  respectively. Both of the LPMT and SPMT system are sensible to  $\theta_{12}$  and  $\Delta m_{21}^2$ , thus the terms are totally free and the starting value are the PDG nominal. Only the LPMT system is sensible to  $\Delta m_{31}^2$ , we let it free so we can observe the effect of the deformation on it while the solar parameters  $\theta_{12}$ ,  $\Delta m_{21}^2$  are constrained by the SPMT system. To prevent  $\Delta m_{31}^2$  to take absurd value, we add a pull term using the PDG nominal value and error. The PDG nominal values used in this study can be found in table 7.1.

$$\chi^2_{joint} = \vec{\Delta}^T V^{-1} \vec{\Delta} + \frac{\Delta m_{31}^2 - \Delta m_{31,PDG}^2}{\sigma_{31,PDG}} \quad (7.15)$$

$\theta_{13}$  is the parameter on which we are least accurate. It's fixed to nominal value to prevent degeneracy (table 7.1).

The covariance matrix is produced from a correlation matrix  $C$

$$V_{ij} = \sigma_i \sigma_j C_{ij} \quad (7.16)$$

where  $\sigma_i$  is the uncertainty on the number of event in the  $i$ th bin. We consider in this study that the content of each bin follow a Poisson statistic and that thus the uncertainty is  $\sigma_i = \sqrt{N_i}$  where  $N_i$  is the content of the  $i$ th bin. The content of each bin can come from two sources the data and the theoretical spectrum  $\sigma_i = \sqrt{D_i}$  (Pearson test) and  $\sigma_i = \sqrt{T_i}$  (Neyman test). Precedent studies have show that

<sup>1829</sup> both Pearson and Neyman tests show bias at low statistic, we thus use the Pearson V test where

$$\chi_{joint}^2 = \vec{\Delta}^T V^{-1} \vec{\Delta} + \frac{\Delta m_{31}^2 - \Delta m_{31,PDG}^2}{\sigma_{31,PDG}} + \ln|V| \quad (7.17)$$

<sup>1830</sup> and the covariance matrix  $V$  is computed using the data spectrum for the uncertainty.

<sup>1831</sup> The estimation of the covariance is crucial in this study as the strength of this test rely on the sys-  
<sup>1832</sup> tematic and statistical correlations between the LPMT and SPMT spectrum. The generation methods  
<sup>1833</sup> and results of this matrix is detailed in section 7.5.2.

### <sup>1834</sup> Delta joint fit

<sup>1835</sup> Using the same structure we define a second joint fit, the Delta joint fit where, in addition to every-  
<sup>1836</sup> thing that was discussed above, we add two other parameters  $\delta\theta_{12}$  and  $\delta\Delta m_{21}^2$  and split the theoretical  
<sup>1837</sup>  $T(\theta_{12}, \Delta m_{21}^2, \dots)$  spectrum in two

$$\begin{aligned} T_{LPMT} &\equiv T(\theta_{12} + \delta\theta_{12}, \Delta m_{21}^2 + \delta\Delta m_{21}^2, \dots) \\ T_{SPMT} &\equiv T(\theta_{12}, \Delta m_{21}^2, \dots) \end{aligned} \quad (7.18)$$

<sup>1838</sup> If the there is no additional effect between the LPMT and the SPMT spectra, the fit should converge  
<sup>1839</sup> to  $\delta\theta_{12} = \delta\Delta m_{21}^2 = 0$ . By observing the dispersion of those parameter we can define the probability  
<sup>1840</sup>  $P(\alpha_{qnl} = 0 | (\delta\theta_{12}, \delta\Delta m_{21}^2))$  and use the median value of  $(\delta\theta_{12}, \delta\Delta m_{21}^2)$  when  $\alpha_{qnl} \neq 0$  to define a  
<sup>1841</sup> sensibility.

<sup>1842</sup> The last test we explore in this thesis is to fit the same spectrum with the standard joint fit that we  
<sup>1843</sup> consider as the hypothesis without distortion  $H_0$  and the hypothesis where there can be a distortion,  
<sup>1844</sup> the delta joint fit designed as the  $H_1$  hypothesis. By looking at the dispersion of  $\chi_{joint,H_0}^2 - \chi_{joint,H_1}^2$   
<sup>1845</sup> we can extract a sensibility to potential distortion.

#### <sup>1846</sup> 7.2.4 Data and theoretical spectrum generation

<sup>1847</sup> To implement the joint fit, we have technically two data spectra and two theoretical spectra. The  
<sup>1848</sup> data in this study are produced using an IBD generator *IBD gen*, see section 7.3.1. The theoretical  
<sup>1849</sup> spectrum are produced the same way as data spectrum but with much higher statistics,  $10^7$  events to  
<sup>1850</sup> compare with the  $\approx 10^5$  events after 6 years. The two spectrum, that we get as a collection of event,  
<sup>1851</sup> are binned in two histograms from 0.8 to 9 MeV of reconstructed energy with bins of 0.02 MeV each  
<sup>1852</sup> resulting in 410 bins per spectrum. An illustration of the theoretical spectrum can be found in figure  
<sup>1853</sup> 7.3. The low number of events in the tail of the spectrum can cause instability due to the low statistic,  
<sup>1854</sup> we thus cut the spectrum at 7.5 MeV / 335 bins for the fit.

<sup>1855</sup> All the IBD spectra presented and used in this study are produced assuming Normal Ordering using  
<sup>1856</sup> the PDG nominal value [16] for oscillation parameters. Those values are reported in table 7.1.

#### <sup>1857</sup> 7.2.5 Limitations

<sup>1858</sup> In this work we are only working considering the statistical errors. We can ignore systematic effects,  
<sup>1859</sup> such as effects that would affect the neutrino spectrum or the background spectrum, as they are  
<sup>1860</sup> entirely correlated between the two systems. The details of those systematic effects can be found in  
<sup>1861</sup> [11].

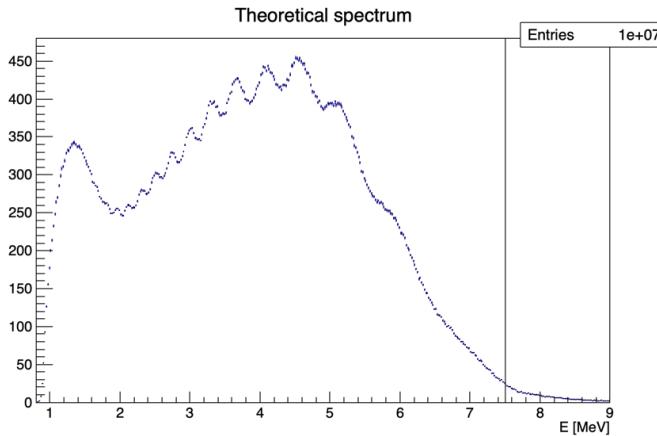


FIGURE 7.3 – Theoretical LPMT spectrum at nominal oscillation values binned using 410 from 0.8 to 9 MeV. It is rescaled to correspond to the amount of event expected in 6 years of data taking. The black line represent the 335 bin cut

1862 Most of our results assume decorrelated detection effects between the SPMT and LPMT system.  
 1863 Their respective reconstruction effect are simulated using simple gaussian drawing on the resolution,  
 1864 independent from their position. This approach was used in previous sensibility and precision  
 1865 studies [11, 80]. The potential effect of those reconstruction effects and a first attempt to take them  
 1866 into account are explored in section 7.5.2.

1867 Even if the goal of this work is to propose deformation agnostic tools, the QNL we use in this study is  
 1868 simplistic as we consider event-wise, position uniform deformation. We show in figure 7.2a and 7.2b  
 1869 that event-wise QNL is equivalent to the mean behaviour of channel-wise QNL but a more complete  
 1870 study would simulate channel-wise deformation for each event.

### 1871 7.3 Fit software

1872 In this section, I describe the ft framework that was used in this study. The software is composed  
 1873 of two part as illustrated in figure 7.4: A standalone part composed of ROOT [81] macros, and the  
 1874 Avenue framework.

1875 The Avenue framework is responsible for the spectrum and configuration reading, transforming the  
 1876 raw collection of event into spectra, managing the physics effect such as the oscillation, computing  
 1877 and minimizing the  $\chi^2$  with the help of the RooFit library. The macros are invoking, if necessary,  
 1878 the Avenue framework and are the entry point for fitting, generating the necessary inputs quantity  
 1879 such as the spectra and correlation matrix, analysing the fit results and managing jobs for distributed  
 1880 computing.

1881 In this section we will focus on the IBD generator in section 7.3.1 and the fit macro in itself in section  
 1882 7.3.2.

#### 1883 7.3.1 IBD generator

1884 The IBD generator is a standalone generator used to produce oscillated and non oscillated spectrum  
 1885 as seen in the JUNO experiment. It takes as inputs physics parameters and a collection of histograms,  
 1886 values and function provided by JUNO to its analysis groups referred as the JUNO common inputs.

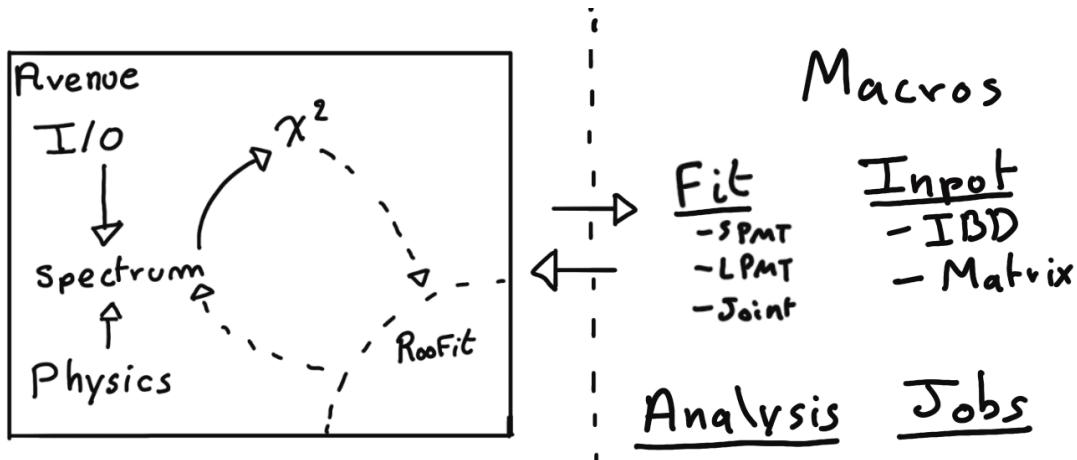


FIGURE 7.4 – Schematic description of the fit framework

1887 The arguments allow to enable or disable effect such as non-uniformity and non-linearity. It finally  
 1888 take as an argument the number of event to generate  $N_{evt}$ . Optionally, we generate an effective  
 1889 number of event  $N$  by drawing in a Poisson distribution of mean  $N_{evt}$ .

1890 Then for each event we

- 1891 1. Choose randomly following the reactor power fraction the reactor from which the neutrino  
 1892 come.
- 1893 2. Generate a random position in the detector following a uniform distribution in the detector.
- 1894 3. Draw a random neutrino energy  $E_\nu$  from the expected neutrino emission spectrum of every  
 1895 reactor. This spectrum is calculated by:
  - 1896 (a) Computing the power spectrum of each isotopes  $^{235}\text{U}$ ,  $^{238}\text{U}$ ,  $^{239}\text{Pu}$ ,  $^{241}\text{Pu}$  using the Huber-  
 1897 Mueller model [5, 8].
  - 1898 (b) Summing the contribution of each isotopes following the respective fission fraction [0.58,  
 1899 0.07, 0.30, 0.05] as reported in [82].
  - 1900 (c) The power of each reactor is then adjusted by their distance from the detector, the detector  
 1901 efficiency and their mean duty cycle (11 of 12 month).
  - 1902 (d) The total spectrum is then finally adjusted by taking into account the correction of the Day  
 1903 Bay bump [83], adjustment due to spent nuclear fuel and due to the non-equilibrium.
- 1904 4. (Optional) Compute the survival probability due to oscillation at nominal oscillation parame-  
 1905 ters value. If the neutrino does not survive, the event is rejected and start from step (1).
- 1906 5. Compute the emitted positron energy  $E_{pos}$  from the mass difference. If the neutrino does not  
 1907 have enough energy reject the event and start from step (1).
- 1908 6. Compute the deposited energy  $E_{dep}$  by incrementing  $E_{pos}$  by 511 keV to account for the positron  
 1909 annihilation. We do not consider cases where some of the energy leak outside of the detector.
- 1910 7. Correct the deposited energy with the expected event-wise non-linearity from [29] to obtain  
 1911 the visible energy  $E_{vis}$ .
- 1912 8. (Optional) Add a custom non-linearity as described in section 7.1.2. This non linearity is  
 1913 characterized by  $\alpha_{qnl}$  to obtain  $E_\alpha$ .
- 1914 9. Finally, using the expected resolution of the LPMT and SPMT system provided in the JUNO  
 1915 common inputs, we draw from a gaussian characterized by those resolution the reconstructed  
 1916 energy  $E_{rec}$  or  $E_{lpmt}$  and  $E_{spmt}$  for each system. The resolutions are provided as ABC para-  
 1917 meters using

$$\frac{E_{vis}}{\sigma E_{vis}} = \sqrt{\left(\frac{A}{\sqrt{E_{vis}}}\right)^2 + b^2 + \left(\frac{c}{E_{vis}}\right)^2} \quad (7.19)$$

1918 The relative and absolute resolutions of the LPMT and SPMT systems are illustrated in figure  
 1919 7.5.

1920 The events are stored as n-tuples and are not yet bin at the end of this step.

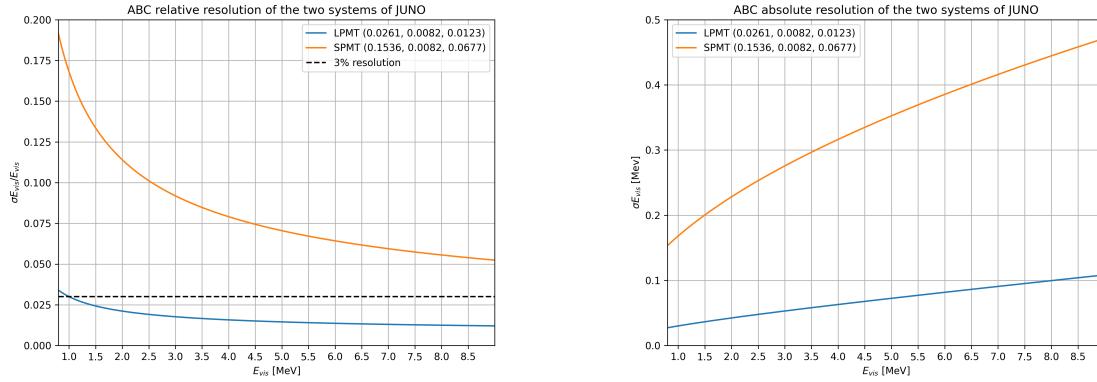


FIGURE 7.5 – Relative (On the left) and absolute (On the right) resolution of the LPMT and SPMT systems used in this study. The number in parenthesis are the parameter A, B and C respectively for each systems.

### 1921 7.3.2 Fit

1922 The fit macro is the core of this fitting procedure. This macro is responsible for loading the fit  
 1923 configuration and setup the Avenue framework. Using Avenue, it will setup the data files, theoretical  
 1924 spectrum, choose the binning,  $\chi^2$ , etc... It also have the possibility to generate toys on the fly based  
 1925 on the theoretical spectrum. Given this theoretical spectrum we can randomize the bin content either  
 1926 by:

1. Drawing the bin content in a Poisson distribution with the bin content as parameter.
2. Drawing the bin content in a Gaussian distribution with the bin content as mean and variance.  
 The bin content is then rounded to the nearest integer.
3. Drawing the bin difference following a given covariance matrix using the Choleski decomposition.  
 This matrix is at least the statistical covariance matrix but can also contain systematic uncertainties.

$$V = LL^T \quad (7.20)$$

$$\mathbf{R} \sim \mathcal{N}(0, 1) \quad (7.21)$$

$$\tilde{\mathbf{h}} = \lceil \mathbf{h} + L\mathbf{R} \rceil \quad (7.22)$$

$$(7.23)$$

1930 where  $V$  is covariance matrix used to produce the fluctuations,  $\mathbf{R}$  is a draw in a multinomial  
 1931 distribution of mean 0 and variance 1,  $\mathbf{h}$  the bin content of the theoretical spectrum and  $\tilde{\mathbf{h}}$  the  
 1932 bin content of the generated toy.

1933 The first two methods allow for the fast production of statistics toys while the third allow for the  
 1934 production of statistical and systematical dependent toys. Unfortunately, none of those methods are  
 1935 fitted to produce toy with a QNL different from the theoretical spectrum. The uncertainty on the  
 1936 reconstructed energy  $\sigma E_{rec}$  being dependent on  $E_{vis}/E_\alpha$  makes that we would need to deconvolute  
 1937 the reconstruction effect from the theoretical spectrum. It is much easier to just produce those toys  
 1938 from the IBD generator.

## 1939 7.4 Technical challenges and development

1940 The fit framework Avenue was already partially developed with multispectrum fitting in mind but  
 1941 a lot technical development was necessary to allow for a joint fit. The first step was to migrate the  
 1942 framework from ROOT5 (last release in March 2018) to ROOT6 (v6.26.06 released in July 2022) to  
 1943 ensure compatibility with the data coming from the JUNO collaboration and benefit of the improve-  
 1944 ment and corrections that came with ROOT6. This allow us to upgrade the C++ standard from C++11  
 1945 to C++17. A substantial effort has been done to modernize the code, generalizing the function via  
 1946 templating to help readability and using smart pointer to prevent possible memory leaks.

1947 The Avenue framework had to be adapted, notably on the chi-square calculation and spectrum  
 1948 generation to correctly take into account the correlation between the SPMT and LPMT spectrum.  
 1949 The delta joint fit requiring two more parameters over a spectrum twice as large as before with  
 1950 LPMT takes much more time, around 15h for 6 years exposure, than the single LPMT fit. Thus the  
 1951 framework and the fit macro had to be updated with ditributed calculation in mind. Notably the  
 1952 aggregation of fit results could be done in a single file instead of managing a file per fit. In case of  
 1953 numerous toy, the hard drive access time could lead to long analysis time.

1954 While the IBD generator was already able to generate LPMT and SPMT spectrum, it was not designed  
 1955 for generating correlated spectrum. As detailed in section 7.3.1, up to the reconstruction effect, the  
 1956 two spectrum need to share the same generation else the two spectrum would be decorrelated and it  
 1957 would be like we would run two different experiment.

## 1958 7.5 Results

### 1959 7.5.1 Validation

1960 The first step is to confirm that the updated fit framework is able to reproduce existing results and  
 1961 that the joint fit behave as expected, meaning

- 1962 — Without QNL, the individual (*LPMT* and *SPMT*) fit converge to the parameter nominal value  
 1963 and that their errors in the ones reported in existing analysis such as [11].
- 1964 — The standard joint fit with an independent covariance matrix (*Indep Standard joint*), meaning  
 1965 that the covariance between the LPMT and SPMT spectra is 0, believe to have twice as much  
 1966 information and thus believe to have a grater precision than the the individual fits.
- 1967 — The standard joint (*Standard joint*) fit but with a correlated covariance matrix has errors similiar  
 1968 to the LPMT individual fit as the LPMT drive the precision on  $\theta_{13}$  and  $\Delta m_{31}^2$  and that the LPMT  
 1969 as SPMT are expected to have the same precision on  $\theta_{12}$  and  $\Delta m_{21}^2$ .
- 1970 — The delta joint (*Delta joint*) fit with covariance matrix have the same resolution as the standard  
 1971 joint fit. The supplementary parameter  $\delta\theta_{12}$  and  $\delta\Delta m_{21}^2$  should not bring supplementaty  
 1972 precision.

1973 The italicized name are the name used in the results reports to identify each fit. We also look into the  
 1974 *Indep Delta joint* which is the fit with the supplementary delta parameters but the covariance between  
 1975 the LPMT and SPMT spectra is 0 and the *Weighted* results where

$$\frac{1}{\sigma_{\text{Weighted}}^2} = \frac{1}{\sigma_{\text{LPMT}}^2} + \frac{1}{\sigma_{\text{SPMT}}^2} \quad (7.24)$$

1976 We expect the weighted resolution to be similar to the *Indep Standard joint* as, in both of those test, we  
 1977 do not consider the correlation between the SPMT and LPMT results.

1978 **Asimov studies**

We ran the tests presented above on the updated framework, the results are reported in table 7.2. All those test are ran considering statistics error only, 6 years exposure with all backgrounds, Pearson  $\chi^2$  (covariance is estimated using data spectrum) and  $\theta_{13}$  fixed to nominal value. For the SPMT fit  $\Delta m_{31}^2$  is fixed at nominal value as the SPMT system is net expected to be sensitive to this parameter.

	$\Delta m_{21}^2$ error	$\delta \Delta m_{21}^2$ error	$\theta_{12}$ error	$\delta \theta_{12}$ error	$\Delta m_{31}^2$ error	$\chi^2$
LPMT	1.29936e-07		1.33852e-03		4.39399e-06	3.23088e-18
SPMT	1.38297e-07		1.38653e-03			2.87502e-18
Indep Standard joint	9.48731e-08		9.86765e-04		4.39212e-06	6.10592e-18
Standard joint	1.29723e-07		1.18342e-03		4.39287e-06	3.38055e-18
Weighted	9.46966e-08		9.63002e-04			
Delta joint	1.35780e-07	3.43529e-08	1.38236e-03	1.46865e-04	4.39309e-06	3.38055e-18
Indep Delta joint	1.38297e-07	1.89391e-07	1.38653e-03	1.87830e-03	4.39241e-06	6.10592e-18
Fixed $\Delta m_{21}^2$ and $\Delta m_{31}^2$						
Indep Standard joint			9.33082e-04			4.82955e-26
LPMT			1.27032e-03			2.58849e-26
SPMT			1.31070e-03			2.24106e-26
Weighted			9.12193e-04			
Fixed $\Delta m_{31}^2$ and $\theta_{12}$						
Indep Standard joint	8.97117e-08					6.10617e-18
SPMT	1.30734e-07					2.87522e-18
LPMT	1.23319e-07					3.23095e-18
Weighted	8.97066e-08					

TABLE 7.2 – Results of the asimov studies of the updated framework. All results are asimov fit, considering 6 years exposure,  $\theta_{13}$  is fixed to nominal value,  $\chi^2$  is pearson meaning that he error is esitmated using the data spectrum

In every cases presented above, the fit converges to the parameters nominal value thus only the errors are presented.

We observe, as expected, that  $\sigma_{\text{Weighted}} \approx \sigma_{\text{Indep Standard joint}}$  with the exception of  $\sigma_{\theta_{12}}$ . This could from the slight difference in statistic between the SPMT and LPMT spectra. Indeed, due to a larger smearing in energy resolution, event that would be in the spectra range [0.8, 7.5] MeV are smeared outside the spectrum. This deficit is partially compensated by event outside the spectrum coming back in this range but we expect very few event outside the spectrum in comparison to event at the edges of it. Thus the event deficit is not totally compensated.  $\theta_{12}$  being mainly driven by the amplitude of the spectrum (see illustration 2.2), that's why we think this the origin of the difference.

The second observation is that  $\sigma_{\text{Standard joint}} \approx \sigma_{\text{LPMT}}$ . Once the covariance matrix between the LPMT and SPMT is correctly introduced, the fit “understand” that it does not have supplementary information and the LPMT system, which have the best precision, dominate the resolution.

Finally for the *Delta* fit, the error on  $\delta \theta_{12}$  and  $\delta \Delta m_{21}^2$  are of the same order of magnitude than the errors on  $\theta_{12}$  and  $\Delta m_{21}^2$  in the absence of the covariance matrix. As the LPMT and SPMT spectra are not connected through the covariance matrix, the delta parameters are unconstrained thus the similar errors. Once the covariance matrix is introduced, the delta are much more constrained and present errors an order of magnitude smaller than the error om their respective parameters.

Overall, the asimov studies are satisfactory. The joint fit behave as expected and the errors on the delta parameters are significantly smaller than the error on their respective parameters, indicating great potential if the converge to value different from 0.

2003 **Toy studies**

2004 Once we validated that the asimov study is yielding coherent results, we study the behaviour of  
 2005 toy studies. The above asimov study is using the Pearson  $\chi^2$  (Eq. 7.13) without pull parameter, we  
 2006 show in figure 7.6 the effect of using a simple Pearson  $\chi^2$ . We see that  $\sin^2(2\theta_{12})$  (reported as  $\theta_{12}$  for  
 2007 simplicity) is biased of about  $0.5\sigma$  and  $\Delta m^2_{21}$  biased of about  $0.1\sigma$ . When introducing the PearsonV  
 2008  $\chi^2$  (Eq. 7.17) the bias disappear as reported in figure 7.7.

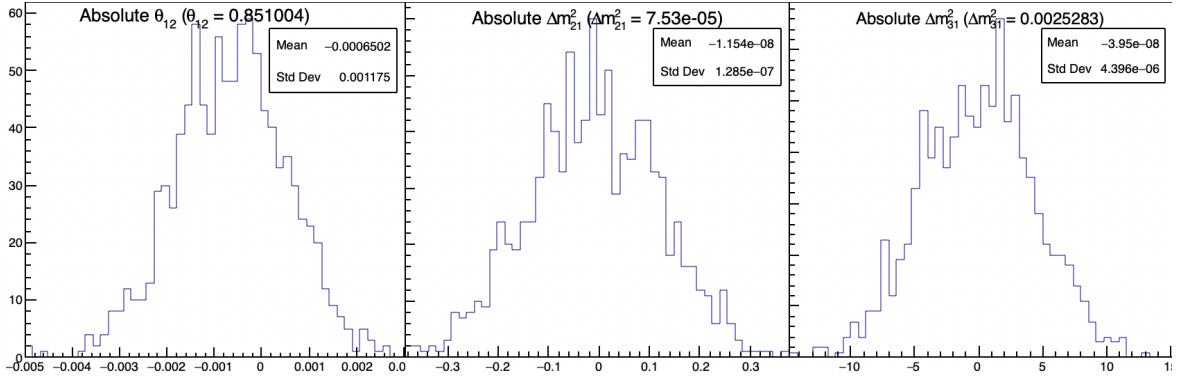


FIGURE 7.6 – Distribution of BFP - nominal value for 1000 toy Standard joint fit. 6 years exposure, all background, Pearson  $\chi^2$ ,  $\theta_{13}$  fixed.

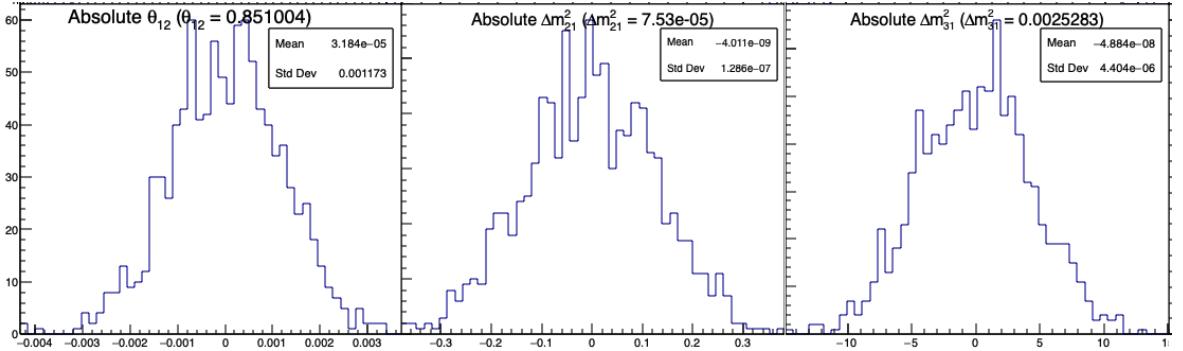


FIGURE 7.7 – Distribution of BFP - nominal value for 1000 toy Standard joint fit. 6 years exposure, all background, PearsonV  $\chi^2$ ,  $\theta_{13}$  fixed.

2009 When the supplementary parameters are introduced in the Delta Joint fit, the fit is stable as shown  
 2010 in the results figure 7.8. The resolutions on the oscillation parameters are slightly worse in the Delta  
 2011 joint fit due to the supplementary freedom. As seen in the asimov studies, the resolution of the  $\delta$   
 2012 parameters is an order of magnitude smaller than their respective parameters, indicating that they  
 2013 can be powerful tools to detect discrepancies between the SPMT and LPMT spectra.

2014 **Effect of supplementary QNL on the LPMT spectrum**

2015 Now that we know that the framework and joint fit behave correctly on unbiased data, we test the  
 2016 effect of introducing the QNL, as presented in Eq. 7.2, in the LPMT spectrum. To test the effect, we  
 2017 consider a QNL  $\alpha_{qnl} = 1\%$ . For reference, this is about three time the expected residual QNL after  
 2018 calibration ( $\alpha_{qnl} = 0.3\%$  [29]). The background had to be removed as JUNO provide them already  
 2019 smeared, thus the introduction of supplementary QNL is not trivial, the resolution being dependent  
 2020 of  $E_{vis}$  which is affected by the QNL. We use a covariance matrix assuming no QNL. The effect of

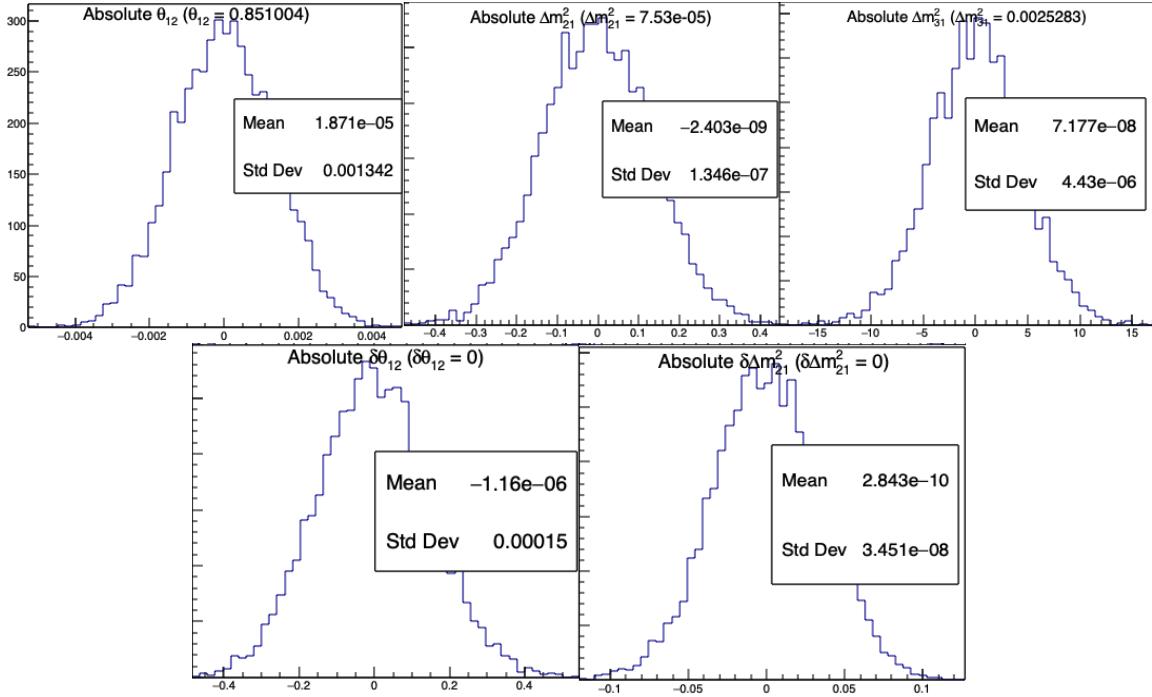


FIGURE 7.8 – Distribution of BFP - nominal value for 5000 toy Delta joint fit. 6 years exposure, all background, PearsonV  $\chi^2$ ,  $\theta_{13}$  fixed.

2021 this QNL on the spectrum is illustrated in figure 7.9. In table 7.3 we report the results of the different  
2022 scenarios.

Mean (std dev)	$\theta_{12} [10^{-3}]$	$\Delta m^2_{21} [10^{-7} \text{ eV}^2]$	$\Delta m^2_{31} [10^{-6} \text{ eV}^2]$	$\delta\theta_{12} [10^{-3}]$	$\delta\Delta m^2_{21} [10^{-7} \text{ eV}^2]$
LPMT	-1.569 (1.171)	-0.957 (0.989)	-8.235 (3.898)	Irrelevant	Irrelevant
SPMT	-0.164 (1.191)	-0.603 (1.054)	Not sensitive	Irrelevant	Irrelevant
Indep Standard	-0.880 (1.174)	-0.786 (1.004)	-8.195 (3.900)	Irrelevant	Irrelevant
Standard	-8.106 (1.423)	-2.483 (1.018)	-6.649 (4.008)	Irrelevant	Irrelevant
Indep Delta	-0.169 (1.190)	-0.598 (1.054)	-8.234 (3.899)	-1.397 (0.259)	-0.361 (0.366)
Delta	-0.163 (1.183)	-1.532 (1.036)	-8.193 (3.934)	-1.441 (0.193)	0.654 (0.303)

TABLE 7.3 – Results of the different fit scenarios on QNL distorted data  $\alpha_{qnl} = 1\%$ .

The mean value are reported subtracted from their nominal value. For SPMT  $\Delta m^2_{31}$  is fixed at nominal value. The  $\chi^2$  is PearsonV. The correlation matrix used to fit suppose no QNL in the spectrum.

2023 The results in table 7.3 are subtracted from their nominal value, themselves reported in table 7.1.  
2024 We clearly see the bias induced by  $\alpha_{qnl} = 1\%$  when comparing the SPMT and LPMT results. The  
2025 Indep Standard is, as expected, the mean value between the SPMT and LPMT: the fit having no  
2026 informations about the correlation between the spectrum think it have two uncorrelated experiment  
2027 thus report an in between value. When introducing the relationship between the LPMT and SPMT  
2028 spectra in the Standard fit, the joint fit cannot find a clean minima, it thus converge to a completely  
2029 incorrect value.

2030 Introducing the  $\delta$  without the correlation in Delta Indep remove the bias and converge to the SPMT  
2031 minima, the  $\delta$  absorbing the deformation of the LPMT spectra.

2032 Finally, with the  $\delta$  and the covariance matrix,  $\theta_{12}$  is unbiased,  $\delta\theta_{12}$  absorbing the deformation.  $\delta\Delta m^2_{21}$   
2033 is still heavily biased, even more than LPMT only, for the same reason than the Standard fit: the  
2034 correlation make it difficult to converge to the nominal value.

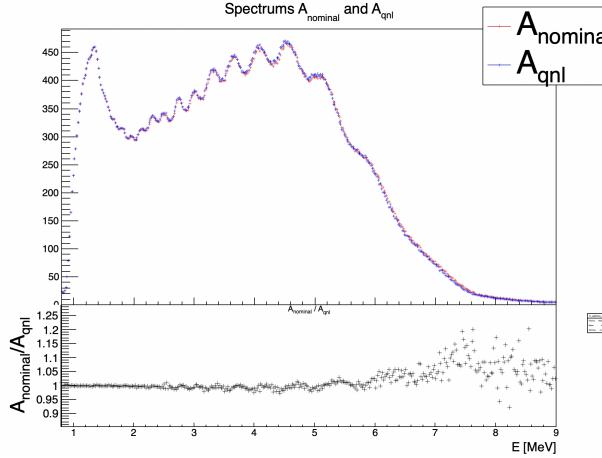


FIGURE 7.9 – **Top:** Theoretical spectrum without QNL (in red) and with  $\alpha_{qnl} = 1\%$  (in blue). **Bottom:** Ratio between the theoretical spectrum with and without QNL.

2035 Overall  $\Delta m_{31}^2$  bias is unchanged as the SPMT spectrum bring no information about the parameter  
2036 and the  $\delta$  are significant, naively up to  $7.46\sigma$  for  $\delta\theta_{12}$  in the Delta fit.

### 2037 7.5.2 Covariance matrix

2038 The covariance matrix between the LPMT and SPMT spectra is at the heart of this study as it  
2039 was already mentioned in section 7.2 and demonstrated in section 7.5.1. In this section we discuss  
2040 the different approaches taken to estimate it. In this work we will mainly discuss the statistical  
2041 covariance matrix between the two spectra, how the number of event in a LPMT bin influence the  
2042 number of bin in the SPMT spectra due to the resolution, but we will still discuss the reconstruction  
2043 effect, mostly due to non-uniformity in the reconstruction resolution.

#### 2044 Analytical method

2045 The first method discussed is the analytical method where we propagate the resolution of the LPMT  
2046 and SPMT spectra over a non-smeared spectrum. Following the approach used in the IBD generation  
2047 in section 7.3.1, we consider the system resolution  $\sigma(E)$  to be only dependent in energy. We do not  
2048 consider the position of the event.

2049 The first step is to compute the statistical uncertainty of the input spectrum while taking into ac-  
2050 count the smearing, considering no uncertainty on the smearing. For this, using the notation of  
2051 section 39.2.5 *Propagation of errors* of PDG2020 [16] and considering an extended spectrum of 820 bins  
2052 following the binning scheme introduced in 7.2.4, the first 410 for the LPMT and the last 410, we  
2053 consider

- 2054 —  $\boldsymbol{\theta} = (\theta_0, \dots, \theta_n); n = 820$  the content of the spectrum bins.
- 2055 —  $\boldsymbol{\eta}(\boldsymbol{\theta}) = (\eta_0(\boldsymbol{\theta}), \dots, \eta_m(\boldsymbol{\theta})); m = 820$  the set of smearing functions representing the PMT resolu-  
2056 tion.

2057  $\eta_m$  can thus be defined as

$$\eta_i = \sum_j^n G(i, \sigma(E_i))(j)\theta_j \quad (7.25)$$

2058 where  $G(i, \sigma(E_i))(j)$  is the smearing function defined as

$$G(i, \sigma(E_i))(j) = \int_{\lfloor E_i \rfloor}^{\lceil E_i \rceil} \frac{1}{\sigma(E_i)\sqrt{2\pi}} e^{-\frac{(E_i - E_j)^2}{2\sigma(E_i)^2}} \quad (7.26)$$

2059 where  $E_i$  is the mean energy in the bin  $i$  and  $\lfloor E_i \rfloor$  and  $\lceil E_i \rceil$  are the lower and higher energy bound of  
2060 the  $i$ th bin respectively.

2061 We can then construct the transfer matrix  $A$  as

$$A_{ij} = \frac{\partial \eta_i}{\partial \theta_j} = G(i, \sigma(E_i))(j) \quad (7.27)$$

2062 and then compute the first part of our covariance matrix

$$U = A V A^T \quad (7.28)$$

2063 where  $V$  is the uncorrelated covariance matrix simply defined, under the assumption of poissonian  
2064 statistic for the bin content

$$V_{ij} = \sqrt{\theta_i \theta_j} \quad (7.29)$$

2065 Now we just need to consider the uncertainty on the smearing  $\sigma\eta_i$ , considering no uncertainty on the  
2066 unsmeared spectrum. From Eq. 7.25, the  $G(i, j) \equiv G(i, \sigma(E_i))(j)$  are independent from each other  
2067  $\forall i, j$ . This mean that this covariance matrix is diagonal, we only need  $\sigma G(i, j)$ . We can derive this  
2068 term from two equation:

- 2069 — The term  $G(i, j)\theta_j$  represent the number of event smeared from the bin  $j$  that end up in the bin  
2070  $i$ . This is a number, we thus assume poissonian statistic so that  $\sigma[G(i, j)\theta_j] = \sqrt{G(i, j)\theta_j}$ .
- 2071 — Using basic error propagation we can say that  $\sigma^2[G(i, j)\theta_j] = \theta_j^2\sigma^2G(i, j) + G(i, j)^2\sigma^2\theta_j$ .

Using  $\sigma\theta_j = \sqrt{\theta_j}$  we derive

$$G(i, j)\theta_j = \sigma^2[G(i, j)\theta_j] = \theta_j^2\sigma^2G(i, j) + G(i, j)^2\theta_j \quad (7.30)$$

$$\Rightarrow \sigma^2G(i, j) = \frac{G(i, j)\theta_j - G(i, j)^2\theta_j}{\theta_j^2} \quad (7.31)$$

$$= \frac{(1 - G(i, j))G(i, j)}{\theta_j} \quad (7.32)$$

2072 By summing the two obtained covariance matrix, we can extract a correlation matrix presented  
2073 in figure 7.10. The correlation between the SPMT and LPMT spectra is greater at the start of the  
2074 spectrum, where the absolute smearing is the smallest up to 5% and diffuse as the bins are further  
2075 from each other and the absolute resolution grow.

## 2076 Empiric method

2077 The second method is the empiric way where we generate toys and just compute the empirical  
2078 correlation between the bin contents.

$$\text{Corr}(\theta_i, \theta_j) = \frac{\mathbb{E}[\theta_i\theta_j] - \mathbb{E}[\theta_i]\mathbb{E}[\theta_j]}{\sigma\theta_i\sigma\theta_j} \quad (7.33)$$

2079 We thus generate  $10^7$  event using the IBD generator presented in section 7.3.1, then produce spectra

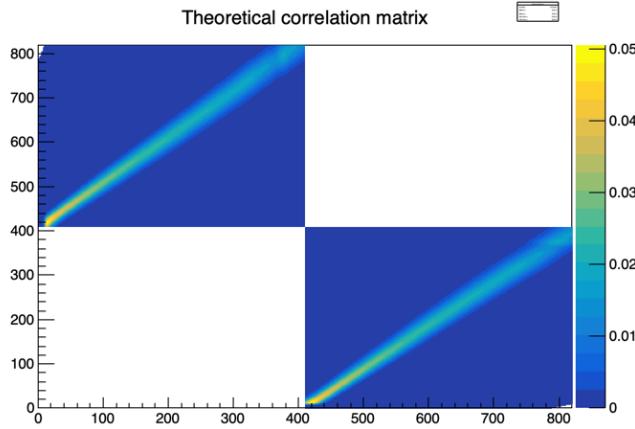


FIGURE 7.10 – Theoretical correlation matrix between the LPMT spectrum (bins 0-409) and the SPMT spectrum (410-819). The diagonal has been set to 0 (it was 1) for readability purpose.

from this finite set of events, meaning we must choose a number  $N$  of toy each composed of  $M$  event in a manner to not have the best estimate.

Due to the nature of our estimator, the estimated correlation coefficient is subject to statistical fluctuation as any estimator. There is no definite formula to compute the standard deviation of the correlation coefficient as suggested in this study [84] but all cited formula depend solely on the number of samples, in our case the number of toy  $N$ , and the correlation coefficient. This indicate that maximizing the number of toy is the right decision, even if each toy posses only one sole event.

To study this rather counter intuitive observation (How can a spectrum with only one event can be representative of the experiment ?), I present in figure 7.11 the upper left corner of the estimated correlation matrix for different configurations of  $N$  and  $M$  in the limit of  $10^7$  total event. We see in figure 7.11a that if the toy number  $N$  is too low, the statistical noise make the correlation pattern almost completely disappear, in figure 7.11b we see clearly the same correlation pattern as in the theoretical matrix in figure 7.10. On the final matrix in figure 7.11c the pattern is clearly visible, but we see a shade of anti-correlation around the spectrum that was not present in the theoretical correlation matrix.

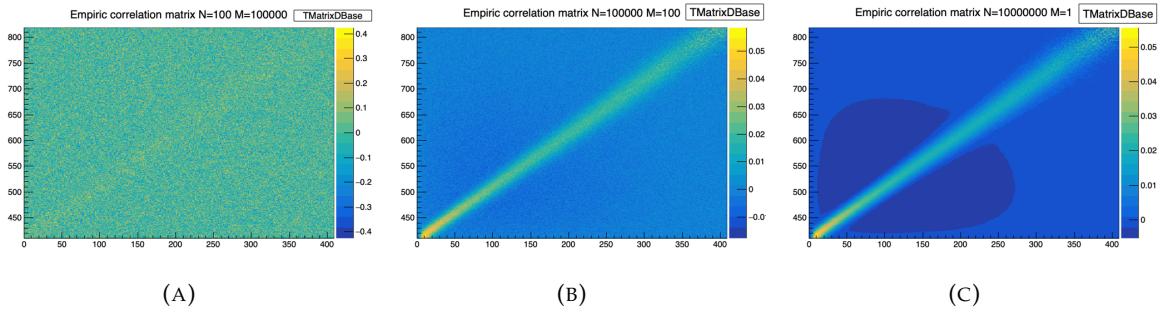


FIGURE 7.11 – Upper left corner of the estimated correlation matrix between the LPMT and SPMT spectrum for different configuration of  $N$  toy with different number of  $M$  events per toy

The difference between the element of the theoretical and the empiric correlation matrix are presented in figure 7.12a. We note that the difference between the two is very small with a bias of  $1.8 \cdot 10^{-3}$  and a standard deviation of  $1.9 \cdot 10^{-3}$  while the interesting correlation are of the order  $10^{-2}$ . Also as

presented in figure 7.12b, the most extreme differences comes from the low end of the spectrum.  
 This low energy difference could be explained as the theoretical does not take into account event that would be smeared from outside the spectrum.  $E < 0.8$ , MeV back inside the spectrum thus missing on the potential correlations.

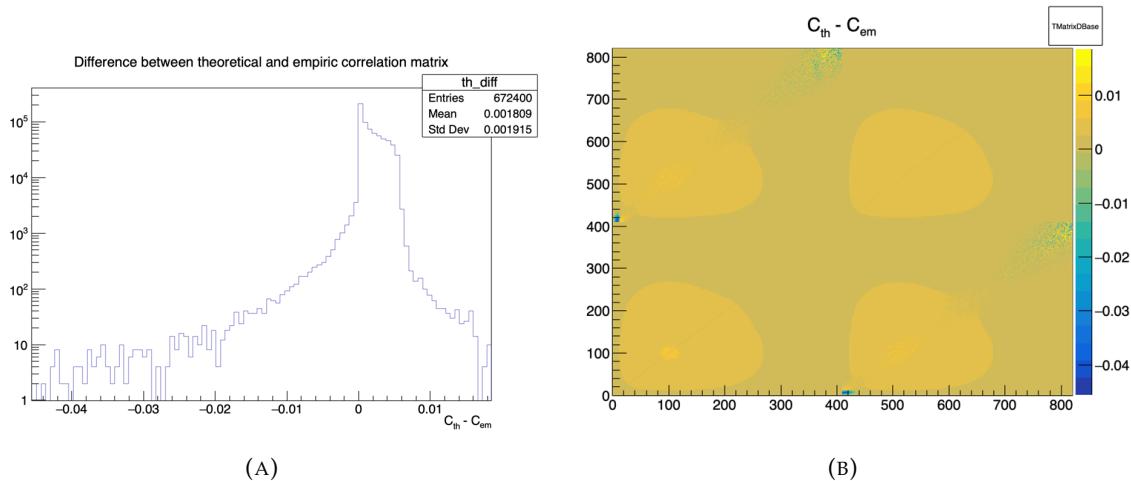


FIGURE 7.12 – Difference between the element of the theoretical and empiric correlation matrix

These empirical correlation matrix still pose an issue: These matrix needs to be inverted for  $\chi^2$  calculation. The framework use the Cholesky decomposition [85] for this, requiring the correlation matrix to be positive definite, which is not guarantee using this empirical methods. Furthermore, we see in those empirical matrix unexpected behavior and spectrum signature in these anti-correlation. For this reason we use the theoretical matrix for the studies presented in this thesis.

### Empirical matrix from fully simulated event

The last study on the correlation matrix between the LPMT and SPMT spectrum consist in simulating and reconstructing full event in the official JUNO simulation framework and constructing an empirical matrix from those events.

### TODO

#### 7.5.3 Statistical tests

In this part, I present the results of the statistical presented in section 7.2.

$\chi^2_{spe}$

The  $\chi^2_{spe}$  is a chi-square representing the compatibility between the LPMT ans SPMT spectra under constraints of the correlation matrix between the two. The results for different exposure can be found in figure 7.13

## 7.6 Conclusion and perspectives

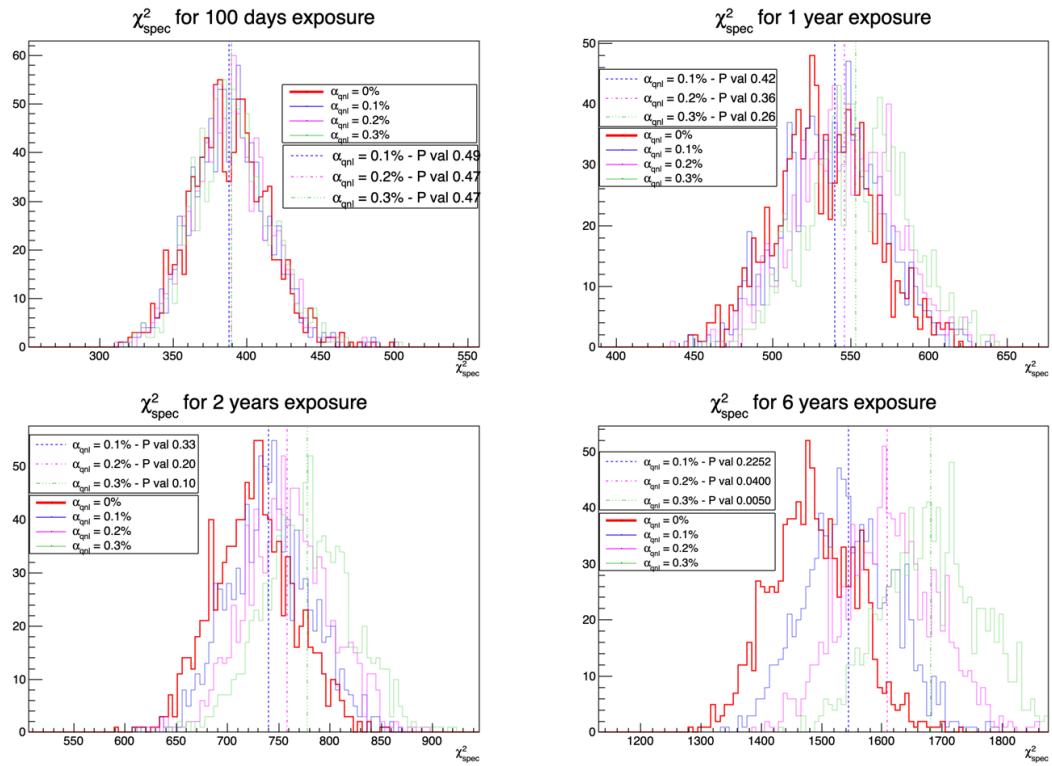


FIGURE 7.13 – Distribution of the  $\chi^2_{\text{spec}}$  for 1000 toys for different exposure. The dashed line represent the median of the distribution and the p-value are the percentage of the  $\alpha_{qnl} = 0$  distribution that are greater than those medians.



<sup>2119</sup> Chapter 8

<sup>2120</sup> Conclusion



<sup>2121</sup> **Appendix A**

<sup>2122</sup> **Calculation of optimal  $\alpha$  for estimator combination**

<sup>2124</sup> This annex the details of the determination of the optimal  $\alpha$  for estimator combination presented in  
<sup>2125</sup> section 4.4.2.

<sup>2126</sup> As a reminder, the combined estimator  $\hat{\theta}$  of  $X$  is defined as

$$\hat{\theta}(X) = \alpha\theta_N + (1 - \alpha)\theta_C; \alpha \in [0; 1] \quad (\text{A.1})$$

<sup>2127</sup> where  $\theta_N$  and  $\theta_C$  are both estimator of  $X$ .

<sup>2128</sup> **A.1 Unbiased estimator**

For the unbiased estimator, it is straight-forward. We search  $\alpha$  such as  $E[\hat{\theta}] = X$

$$E[\hat{\theta}] = E[\alpha\theta_N + (1 - \alpha)\theta_C] \quad (\text{A.2})$$

$$= E[\alpha\theta_N] + E[(1 - \alpha)\theta_C] \quad (\text{A.3})$$

$$= \alpha E[\theta_N] + (1 - \alpha)E[\theta_C] \quad (\text{A.4})$$

$$= \alpha(\mu_N + X) + (1 - \alpha)(\mu_C + X) \quad (\text{A.5})$$

$$X = \alpha\mu_N + \mu_C - \alpha\mu_C + X \quad (\text{A.6})$$

$$0 = \alpha(\mu_N - \mu_C) + \mu_C \quad (\text{A.7})$$

$$(A.8)$$

$$\Rightarrow \alpha = \frac{\mu_C}{\mu_C - \mu_N} \quad (\text{A.9})$$

<sup>2129</sup> **A.2 Optimal variance estimator**

The  $\alpha$  for this estimator is a bit more tricky. By expanding the variance we get

$$\text{Var}[\hat{\theta}] = \text{Var}[\alpha\theta_N + (1 - \alpha)\theta_C] \quad (\text{A.10})$$

$$= \text{Var}[\alpha\theta_N] + \text{Var}[(1 - \alpha)\theta_C] + \text{Cov}[\alpha(1 - \alpha)\theta_N\theta_C] \quad (\text{A.11})$$

$$= \alpha^2\sigma_N^2 + (1 - \alpha)^2\sigma_C^2 + 2\alpha(1 - \alpha)\sigma_N\sigma_C\rho_{NC} \quad (\text{A.12})$$

<sup>2130</sup> where, as a reminder,  $\rho_{NC}$  is the correlation factor between  $\theta_C$  and  $\theta_N$ .

Now we try to find the minima of  $\text{Var}[\hat{\theta}]$  with respect to  $\alpha$ . For this we evaluate the derivative

$$\frac{d}{d\alpha} \text{Var}[\hat{\theta}] = 2\alpha\sigma_N^2 - 2(1-\alpha)\sigma_C^2 + 2\sigma_N\sigma_C\rho_{NC}(1-2\alpha) \quad (\text{A.13})$$

$$= 2\alpha(\sigma_N^2 + \sigma_C^2 - 2\sigma_N\sigma_C\rho_{NC}) - 2\sigma_C^2 + 2\sigma_N\sigma_C\rho_{NC} \quad (\text{A.14})$$

then find the minima and maxima of this derivative by evaluating

$$\frac{d}{d\alpha} \text{Var}[\hat{\theta}] = 0 \quad (\text{A.15})$$

$$2\alpha(\sigma_N^2 + \sigma_C^2 - 2\sigma_N\sigma_C\rho_{NC}) - 2\sigma_C^2 + 2\sigma_N\sigma_C\rho_{NC} = 0 \quad (\text{A.16})$$

$$2\alpha(\sigma_N^2 + \sigma_C^2 - 2\sigma_N\sigma_C\rho_{NC}) = 2\sigma_C^2 - 2\sigma_N\sigma_C\rho_{NC} \quad (\text{A.17})$$

$$\alpha = \frac{\sigma_C^2 - \sigma_N\sigma_C\rho_{NC}}{\sigma_N^2 + \sigma_C^2 - 2\sigma_N\sigma_C\rho_{NC}} \quad (\text{A.18})$$

2131 This equation shows only one solution which is a minima. From Eq. A.18 arise two singularities:

2132 —  $\sigma_N = \sigma_C = 0$ . This is not a problem because as physicists we never measure with an absolute

2133 precision, neither us or our detectors are perfect.

2134 —  $\sigma_N = \sigma_C$  and  $\rho_{CN} = 1$ . In this case  $\theta_C$  and  $\theta_N$  are the same estimator in term of variance thus

2135 any value for  $\alpha$  yield the same result: an estimator with the same variance as the original ones.

<sup>2136</sup> **Appendix B**

<sup>2137</sup> **Charge spherical harmonics analysis**

<sup>2138</sup> When looking at JUNO events we can clearly see some pattern in the charge repartition based on  
<sup>2139</sup> the event radius as illustrated in figure B.4. When dealing with identifying features and pattern on a  
<sup>2140</sup> spherical plane, the astrophysics community have been using, with success, the spherical harmonic  
<sup>2141</sup> decomposition. The principle is similar to a frequency analysis via Fourier transform. It comes to  
<sup>2142</sup> saying that a function  $f(r, \theta, \phi)$ , here our charge repartition of the spherical plane constructed by our  
<sup>2143</sup> PMTs, can be expressed

$$f(r, \theta, \phi) = \sum_{l=0}^{\infty} \sum_{m=-l}^l a_l^m r^l Y_l^m(\theta, \phi) \quad (\text{B.1})$$

<sup>2144</sup> where  $a_l^m$  are constants complex factor,  $Y_l^m(\theta, \phi) = Ne^{im\phi} P_l^m(\cos \theta)$  are the spherical harmonics of  
<sup>2145</sup> degree  $l$  and order  $m$  and  $P_l^m$  their associated Legendre Polynomials. Those harmonics are illustrated  
<sup>2146</sup> in figure B.1. By reducing the problem to the unit sphere  $r = 1$ , we get rid of the term  $r^l$ . The Healpix  
<sup>2147</sup> library [76] offer function to efficiently find the  $a_l^m$  factor from a given Healpix map.

<sup>2148</sup> For the above decomposition, we will define the *Power* of an harmonic as

$$S_{ff}(l) = \frac{1}{2l+1} \sum_{m=-l}^l |a_l^m|^2 \quad (\text{B.2})$$

<sup>2149</sup> and the *Relative Power* as:

$$P_l^h = \frac{S_{ff}(l)}{\sum_l S_{ff}(l)} \quad (\text{B.3})$$

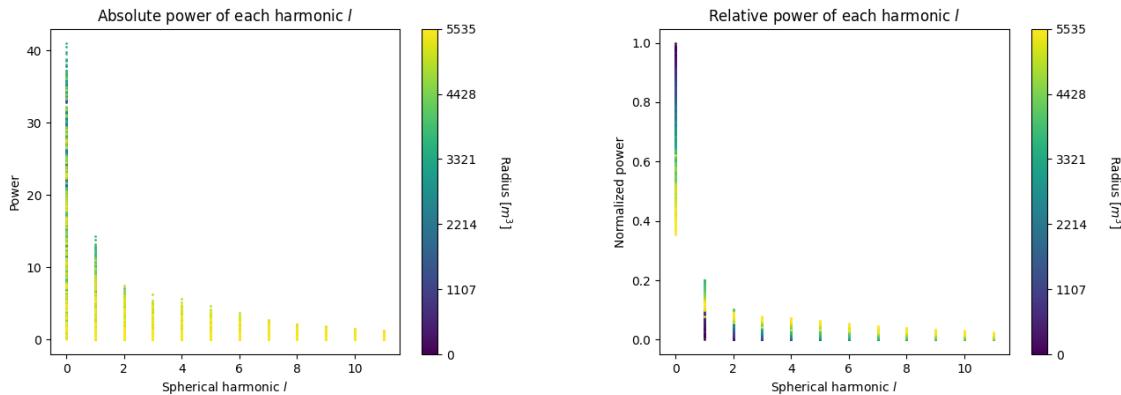
<sup>2150</sup> For this study we will use 10k positron events with  $E_{kin} \in [0; 9]$  MeV uniformly distributed in the  
<sup>2151</sup> CD from the JUNO official simulation version J23.0.1-rc8.dc1 (released the 7th January 2024). All the  
<sup>2152</sup> event are *calib* level, with simulation of the physics, electronics, digitizations and triggers. We first  
<sup>2153</sup> take a sub-set of 1k events and look at the power and relative power distribution depending on the  
<sup>2154</sup> radius and harmonic degree  $l$ . The results are shown in figure B.2. While don't see any pattern in  
<sup>2155</sup> absolute power, it is pretty clear that there is a correlation between the relative power of  $l = 0$  and  
<sup>2156</sup> the radius of the event.

<sup>2157</sup> When applying the same study but dependent on the energy, no clear correlation appear. The results  
<sup>2158</sup> for the  $l = 0$  harmonic are presented in the figure B.5. Thus, in this study we will focus on the radial  
<sup>2159</sup> dependency of the relative power of each harmonic.

<sup>2160</sup> In figures B.6 and B.7 are presented the distribution of the relative power of each harmonic for  $l \in$   
<sup>2161</sup>  $[0, 11]$ . The relation between the radius and the relative power become even more clear, especially  
<sup>2162</sup> for the first harmonics  $l \in [0, 4]$ . After that for  $l > 4$  their relative power is close to 0 for central event,  
<sup>2163</sup> thus loosing power. It also interesting to note the change of behavior in the TR area, clearly visible  
<sup>2164</sup> for  $l = 1$  and  $l = 2$ .

$l:$	$P_\ell^m(\cos \theta) \cos(m\varphi)$	$P_\ell^{ m }(\cos \theta) \sin( m \varphi)$
0 s		
1 p		
2 d		
3 f		
4 g		
5 h		
6 i		
$m:$	6 5 4 3 2 1 0	-1 -2 -3 -4 -5 -6

FIGURE B.1 – Illustration of the real part of the spherical harmonics

FIGURE B.2 – Scatter plot of the absolute and relative power, respectively on the left and right plot, of each harmonic degree  $l$ . The color indicate the radius of the event.

2165 As an erzats of reconstruction algorithm, we fit each of those distribution with a 9th degree polynomial  
 2166 which give us the relation

$$F(R^3) \longmapsto P_l^h \quad (\text{B.4})$$

2167 We do it this way because some of the distribution have multiple solution for a given relative power,  
 2168 for example  $l = 1$ , while each radius give only one power. We now just need to find

$$F^{-1}(P_l^h) \longmapsto R^3 \quad (\text{B.5})$$

2169 Inverting a 9th degree polynomial is hard, if not impossible. The presence of multiple roots for the  
 2170 same power complexify the task even more. To circumvent this problem, we reconstruct the radius  
 2171 by locating the minima of  $(F(R^3) - \hat{P}_l^h)^2$  where  $\hat{P}_l^h$  is the measured power fraction.

2172 To distinguish between multiple possible minima, we use as a starting point the radius given by the  
 2173 procedure on  $l = 0$  that, by looking at the fit in figure B.6, should only present one minima. For  $l > 0$   
 2174 we also impose bound on the possible reconstructed  $R^3$  as  $R^3 \in [R_0^3 - 100, R_0^3 + 100]$  where  $R_0^3$  is the  
 2175 reconstructed  $R^3$  by the harmonic  $l = 0$ .

2176 The minimization algorithm used are the Bent algorithm for  $l = 0$  and the Bounded algorithm for  
 2177  $l > 0$  provided by the Scipy library [86]. We then do the mean of the reconstructed radius from  
 2178 the different harmonics. The reconstruction results are shown in figure B.3. The performance seems  
 2179 correct but we see heavy fluctuation in the bias. To really be used as a reconstruction algorithm, the  
 2180 method needs to be refined as discussed in the next section.

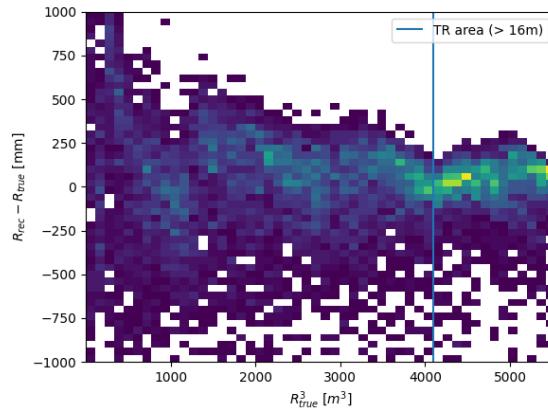


FIGURE B.3 – Error on the reconstructed radius vs the true radius by the harmonic method

## Conclusion

2181 We have clearly shown in this analysis the relevance the of relative harmonic power for radius  
 2182 reconstruction, and provided an erzats of a reconstruction algorithm. We will not delve further in  
 2183 this thesis but if we wanted to refine this algorithm multiple paths can be explored:

- 2184 — No energy signature in the harmonics: This is surprising that there is no correlation between  
 2185 the energy and the amplitude of the harmonics. We know that the energy is heavily correlated  
 2186 with the total number of photoelectrons collected, it would be unintuitive that we see no  
 2187 relation.
- 2188 — Localization of the event: We shown here the relation between the relative power of the har-  
 2189 monic and the radius but don't get any information about the  $\theta$  and  $\phi$  spherical coordinates.  
 2190 This information is probably hidden in the individual power of each order  $m$  of the degree  $l$ .  
 2191 This intuition comes from the figure B.1 where in the higher degree  $l$  we see that the order  $m$   
 2192 are oriented. Intuitively, the order should be able to indicate a direction where the signal is  
 2193 more powerful.
- 2194 — Combination of the degree power: Here we combined the radius reconstructed by the dif-  
 2195 ferent degree via a simple mean but we shown in section 4.4.2 and annex A that this is note  
 2196 the optimal way to combine estimator. A more refined algorithm probably exist to take into  
 2197 account the predicting power of each order.

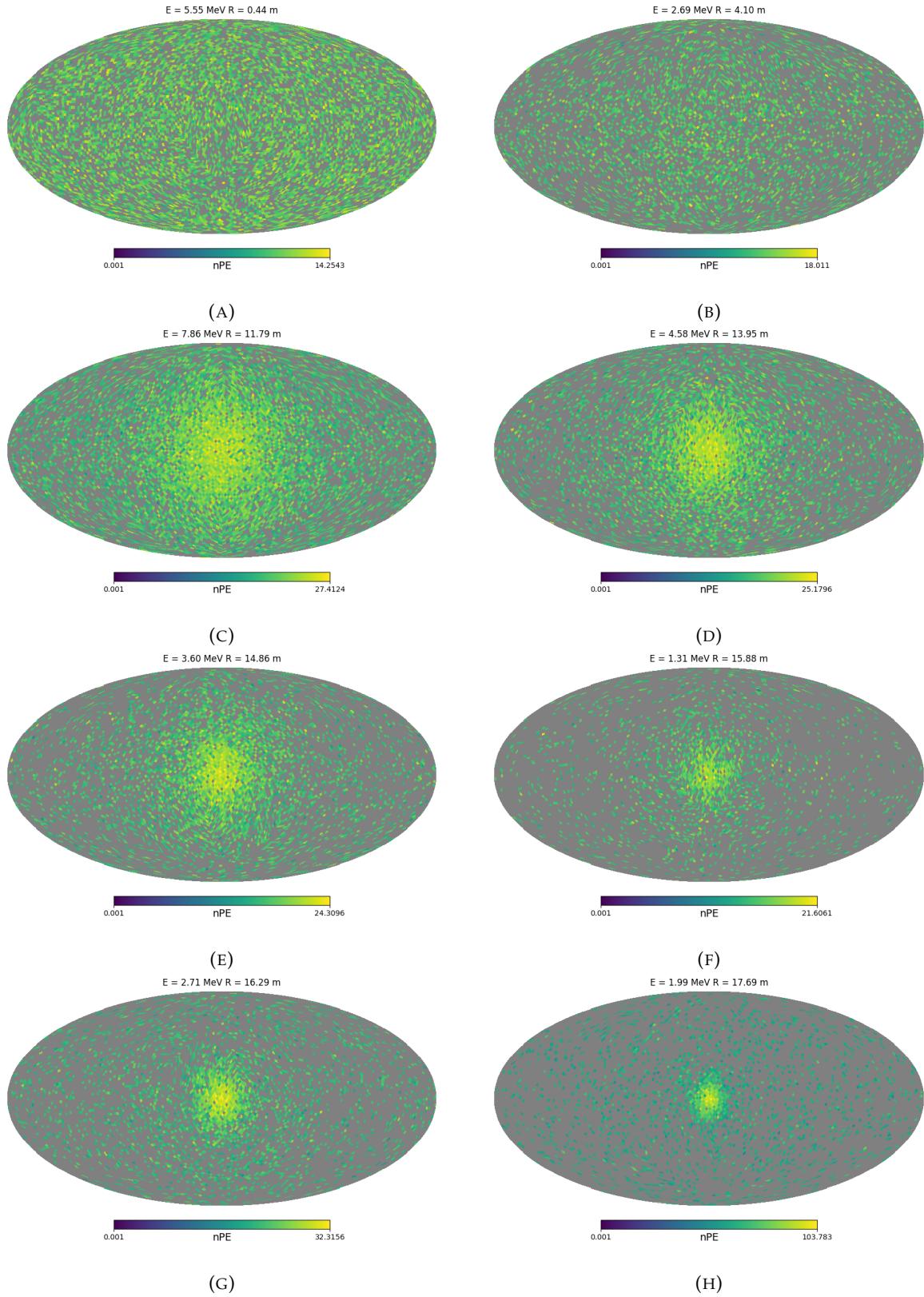


FIGURE B.4 – Charge repartition in JUNO as seen by the Healpix segmentation. Those are Healpix map of order 5 (i.e. 12288 pixels). The color represent the summed charge of the PMTs in each pixels. The color scale is logarithmic. The view have been centered to prevent event deformations.

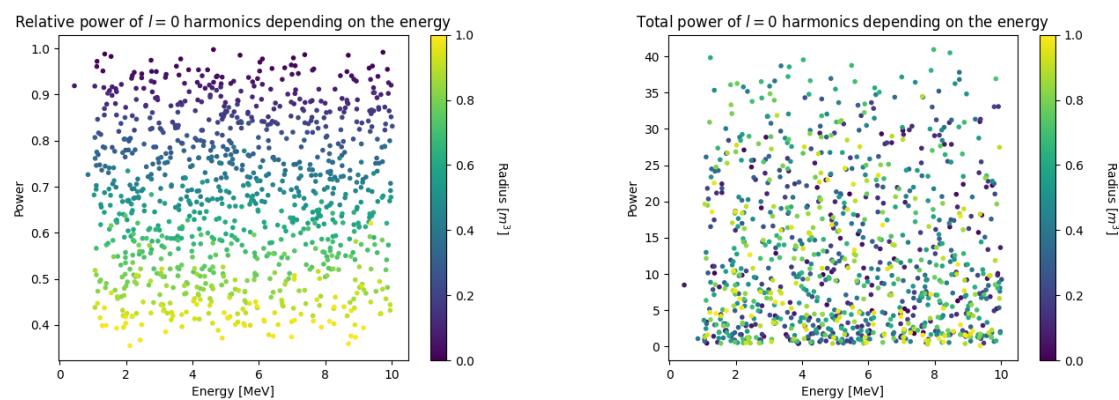


FIGURE B.5 – Scatter plot of the absolute and relative power, respectively on the left and right plot, of the  $l = 0$  harmonic. The color indicate the radius of the event.

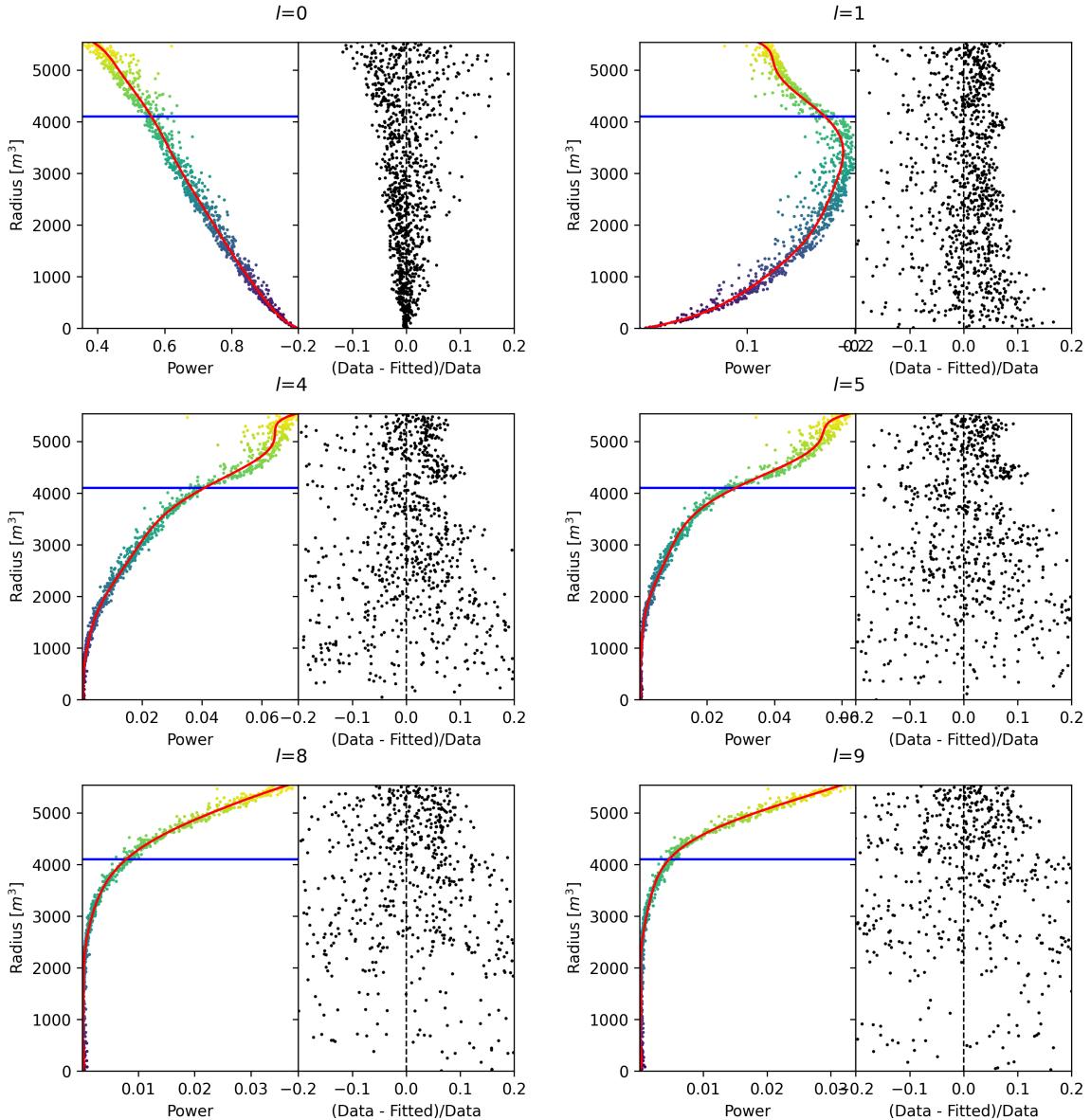


FIGURE B.6 – Plot of the distribution of the relative power of each harmonic dependent on  $R^3$  (on the left). The Total Reflection (TR) area is represented by the horizontal blue line. The distribution are fitted using a 9th degree polynomial (red curve). The relative power error between the distribution and the fit is represented on the left. **Part 1**

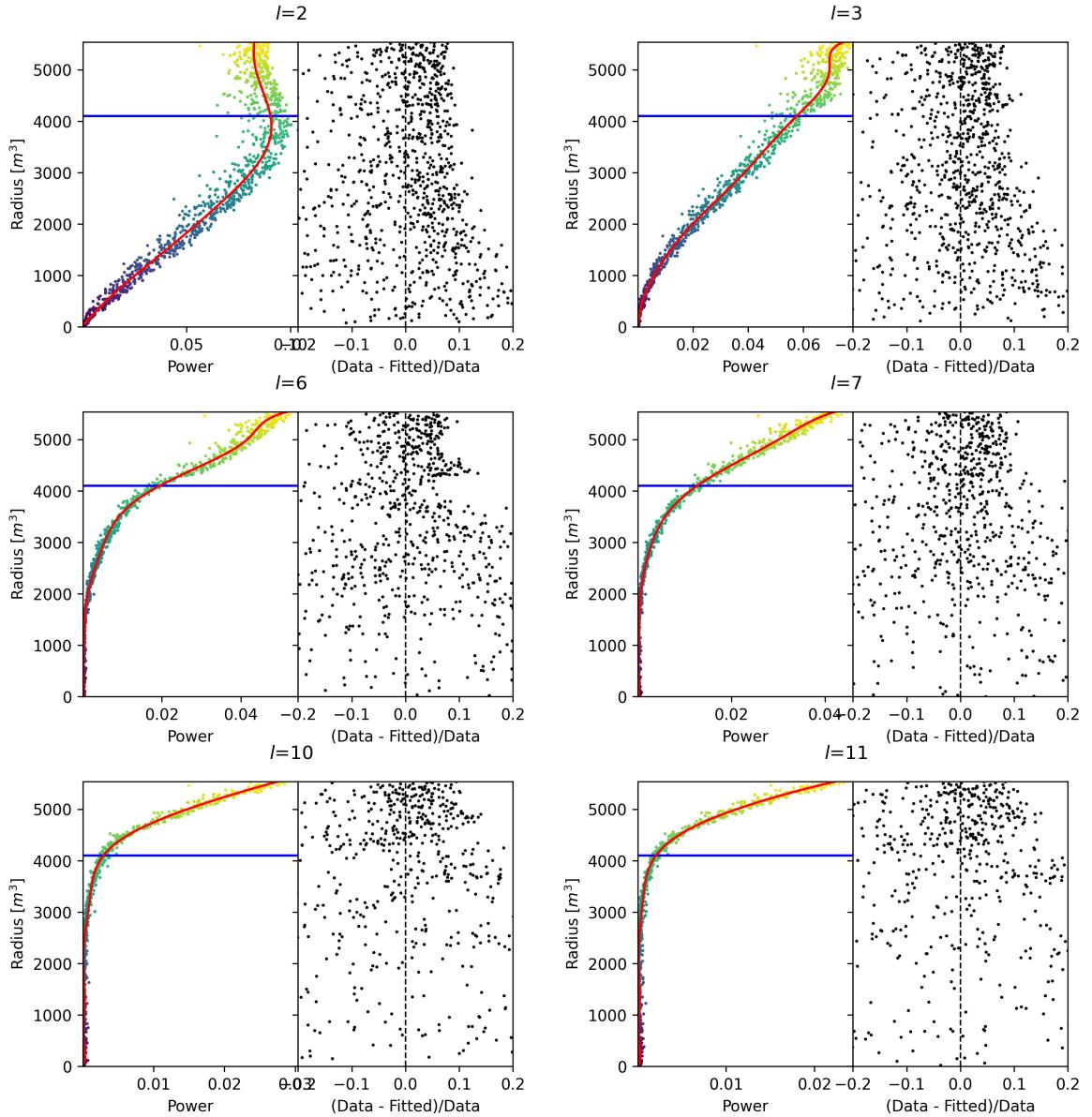


FIGURE B.7 – Plot of the distribution of the relative power of each harmonic dependent on  $R^3$  (on the left). The Total Reflection (TR) area is represented by the horizontal blue line. The distribution are fitted using a 9th degree polynomial (red curve). The relative power error between the distribution and the fit is represented on the left. **Part 2**



<sup>2199</sup> **Appendix C**

<sup>2200</sup> **Over-gaussian smearing of spectrum**

<sup>2201</sup> In this section we demonstrate that a spectrum  $S$  smeared by a gaussian  $G$  parametrized by its  
<sup>2202</sup> varianse  $\sigma_1^2$  can be smeared by a gaussian parametrized by the variance  $\sigma_2^2$  from the the smeared  
<sup>2203</sup> spectrum  $K(E, \sigma_1) = S(E) \star G(E, \sigma_1)$  under the condition that  $\sigma_2^2 > \sigma_1^2$ .

Let  $K'(E, \sigma_2) = S(E) \star G(E, \sigma_2)$  the target spectrum we can expand

$$K'(E, \sigma_2) = S(E) \star G(E, \sigma_1) \star G^{-1}(E, \sigma_1) \star G(E, \sigma_2) \quad (\text{C.1})$$

$$= K(E, \sigma_1) \star G^{-1}(E, \sigma_1) \star G(E, \sigma_2) \quad (\text{C.2})$$

<sup>2204</sup> where  $G^{-1}(E, \sigma_1)$  is defined as  $G(E, \sigma_1) \star G^{-1}(E, \sigma_1) = \delta(E)$ .

By moving into Fourier space we can express

$$G(E, \sigma_1) \star G^{-1}(E, \sigma_1) = \delta(E) \quad (\text{C.3})$$

$$F[G(E, \sigma_1)](\nu) \times F[G^{-1}(E, \sigma_1)](\nu) = 1 \quad (\text{C.4})$$

<sup>2205</sup> with  $F[G(E, \sigma_1)](\nu)$  the fourier transform of  $G$

$$F[G(E, \sigma_1)](\nu) = e^{-\frac{\sigma_1^2(2\pi)^2}{2}\nu^2} \quad (\text{C.5})$$

we have

$$F[G^{-1}(E, \sigma_1)](\nu) = (F[G(E, \sigma_1)](\nu))^{-1} = (e^{-\frac{\sigma_1^2(2\pi)^2}{2}\nu^2})^{-1} \quad (\text{C.6})$$

$$= e^{\frac{\sigma_1^2(2\pi)^2}{2}\nu^2} \quad (\text{C.7})$$

Thus we express

$$F[G^{-1}(E, \sigma_1) \star G(E, \sigma_2)] = e^{\frac{\sigma_1^2(2\pi)^2}{2}\nu^2} \times e^{-\frac{\sigma_2^2(2\pi)^2}{2}\nu^2} \quad (\text{C.8})$$

$$= e^{\frac{(2\pi)^2}{2}(\sigma_1^2 - \sigma_2^2)\nu^2} \quad (\text{C.9})$$

$$= e^{\frac{(2\pi)^2}{2}\Delta\sigma^2\nu^2}; \Delta\sigma^2 = (\sigma_1^2 - \sigma_2^2) \quad (\text{C.10})$$

<sup>2206</sup> We see that  $F^{-1}[F[G^{-1}(E, \sigma_1) \star G(E, \sigma_2)]]$  is solvable if  $\Delta\sigma^2 = (\sigma_1^2 - \sigma_2^2) < 0 \Rightarrow \sigma_2 > \sigma_1$ . In that case

$$G^{-1}(E, \sigma_1) \star G(E, \sigma_2) = \frac{1}{\sqrt{|\Delta\sigma^2|}\sqrt{2\pi}} e^{-\frac{E^2}{2|\Delta\sigma^2|}} \quad (\text{C.11})$$



# List of Tables

2208	<b>2.1</b>	Characteristics of the nuclear power plants observed by JUNO. . . . .	13
2210	<b>2.2</b>	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. . . . .	15
2211	<b>2.3</b>	Detectable neutrino signal in JUNO and the expected signal rates and major background sources . . . . .	15
2212	<b>2.4</b>	List of sources and their process considered for the energy scale calibration . . . . .	23
2214	<b>2.5</b>	Calibration program of the JUNO experiment . . . . .	25
2215	<b>2.6</b>	Features used by the BDT for vertex reconstruction . . . . .	36
2217	<b>2.7</b>	Features used by the BDTE algorithm. <i>pe</i> and <i>ht</i> reference the charge and hit-time distribution respectively and the percentages are the quantiles of those distributions. <i>cht</i> and <i>cc</i> reference the barycenters of hit time and charge respectively . . . . .	36
2220	<b>4.1</b>	Sets of hyperparameters values considered in this study . . . . .	56
2221	<b>5.1</b>	Parameters of the 5th degree polynomial used to correct Omilrec reconstructed energy. . . . .	77
2222	<b>7.1</b>	Nominal PDG value [16]. All value are reported assuming Normal Ordering. . . . .	88
2223	<b>7.2</b>	Results of the asimov studies of the updated framework. All results are asimov fit, considering 6 years exposure, $\theta_{13}$ is fixed to nominal value, $\chi^2$ is pearson meaning that the error is estimated using the data spectrum . . . . .	94
2224	<b>7.3</b>	Results of the different fit scenarios on QNL distorted data $\alpha_{qnl} = 1\%$ . The mean value are reported subtracted from their nominal value. For SPMT $\Delta m_{31}^2$ is fixed at nominal value. The $\chi^2$ is PearsonV. The correlation matrix used to fit suppose no QNL in the spectrum. . . . .	96
2226			
2227			
2228			
2229			



# List of Figures

2231    2.1	<b>On the left:</b> Location of the JUNO experiment and its reactor sources in southern china. <b>On the right:</b> Aerial view of the experimental site . . . . .	11
2232    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 ( $\theta_{12}$ , $\Delta m_{21}^2$ ). 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. . . . .	12
2233    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] . . . . .	14
2234    2.4	. . . . .	17
2235    a	Schematics view of the JUNO detector. . . . .	17
2236    b	Top down view of the JUNO detector under construction . . . . .	17
2237    2.5	Schematics of an IBD interaction in the central detector of JUNO . . . . .	18
2238    2.6	Schematics of the supporting node for the acrylic vessel . . . . .	19
2239    2.7	<b>On the left:</b> Quantum efficiency (QE) and emission spectrum of the LAB and the bis-MSB [20]. <b>On the right:</b> Sensitivity of the Hamamatsu LPMT depending on the wavelength of the incident photons [22]. . . . .	19
2240    2.8	Schematic of a PMT . . . . .	20
2241    2.9	The LPMT electronics scheme. It is composed of two part, the <i>wet</i> electronics on the left, located underwater and the <i>dry</i> electronics on the right. They are connected by Ethernet cable for data transmission and a dedicated low impedance cable for power distribution . . . . .	21
2242    2.10	Schematic of the JUNO SPMT electronic system ( <b>left</b> ), and exploded view of the main component of the UWB ( <b>right</b> ) . . . . .	22
2243    2.11	The JUNO top tracker . . . . .	23
2244    2.12	Fitted and simulated non linearity of gamma, electron sources and from the $^{12}\text{B}$ spectrum. Black points are simulated data. Red curves are the best fits. Figures taken from [29]. . . . .	24
2245    a	Gamma non-linearity . . . . .	24
2246    b	Boron spectrum . . . . .	24
2247    c	Electron non-linearity . . . . .	24
2248    2.13	Overview of the calibration system . . . . .	25
2249    2.14	Event-level instrumental non-linearity, defined as the ratio of the total measured LPMT charge to the true charge for events uniformly distributed in the detector. The solid red line represents event-level non-linearity without the channel-level correction, with position non-uniformity obtained at 1 MeV applied, in an extreme hypothetical scenario of 50% non-linearity over 100 PEs for the LPMTs. The dashed blue line represents that after the channel-level correction. The gray band shows the residual uncertainty of 0.3%, after the channel-level correction. Figure taken from [29]. . . . .	26
2250    2.15	. . . . .	27

2274	a	Schematic of the TAO satellite detector	27
2275	b	Schematic of the OSIRIS satellite detector	27
2276	2.16		29
2277	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.	29
2281	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	29
2283	2.17		30
2284	a	$\Delta t$ distribution at different iterations step $j$	30
2285	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	30
2287	2.18	Bias of the reconstructed radius R (left), $\theta$ (middle) and $\phi$ (right) for multiple energies by the time likelihood algorithm	31
2289	2.19	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	32
2293	2.20	Radial resolution of the different vertex reconstruction algorithms as a function of the energy	32
2295	2.21		33
2296	a	Spherical coordinate system used in JUNO for reconstruction	33
2297	b	Definition of the variables used in the energy reconstruction	33
2298	2.22		35
2299	a	Radial resolutions of the likelihood-based algorithm TMLE, QMLE and QTMLE	35
2300	b	Energy resolution of QMLE and QTMLE using different vertex resolutions	35
2301	2.23	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	37
2303	2.24	Radial (left) and energy (right) resolutions of different ML algorithms. The results presented here are from [42]. DNN is a deep neural network, BDT is a BDT, ResNet-J and VGG-J are CNN and GNN-J is a GNN.	38
2306	3.1	Example of a BDT that determine if the given object is a duck	42
2307	3.2		44
2308	a	Schema of a FCDNN	44
2309	b	Illustration of a composition of ReLU “approximating” a function. (1) No ReLU is taking effect (2) One ReLU is activating (3) Another ReLU is activating	44
2311	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 [58].	44
2315	3.4		45
2316	a	Example of images in the MNIST dataset	45
2317	b	Schema of the CNN used in Pytorch example to process the MNIST dataset	45
2318	3.5	Illustration of the message passing algorithm. The detailed explanation can be found in section 3.2.3	46
2320	3.6		48
2321	a	Illustration of SGD falling into a local minima	48
2322	b	Illustration of the Adam momentum allowing it to overcome local minima	48
2323	3.7	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.	49

2326	a	Illustration of the SGD optimizer on one parameter $\theta$ on the MAE Loss. We see here that it has trouble reaching the minima due to the gradient being constant.	49
2327	b	Illustration of the SGD optimizer on one parameter $\theta$ 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. . . . .	49
2328	3.8	. . . . .	50
2329	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. . . . .	50
2330	b	Illustration of a very simple NN . . . . .	50
2331	3.9	Illustration of the ResNet framework . . . . .	51
2332	3.10	Illustration of the gradient explosion. Here it can be solved with a lower learning rate but its not always the case. . . . .	52
2333	4.1	Graphic representation of the VGG-16 architecture, presenting the different kind of layer composing the architecture. . . . .	54
2334	4.2	Repartition of SPMTs in the image projection. The color scale is the number of SPMTs per pixel . . . . .	58
2335	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. <b>On the left:</b> the charge channel. The color is the charge in each pixel in NPE equivalent. <b>On the right:</b> The time channel in nanoseconds. . . . .	59
2336	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. <b>On the left:</b> the charge channel. The color is the charge in each pixel in NPE equivalent. <b>On the right:</b> The time channel in nanoseconds. . . . .	59
2337	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. <b>On the left:</b> the charge channel. The color is the charge in each pixel in NPE equivalent. <b>On the right:</b> The time channel in nanoseconds. . . . .	60
2338	4.6	Example of a low energy, central event. Here there is no clear signal, the uniformity of the distribution should make it central. <b>On the left:</b> the charge channel. The color is the charge in each pixel in NPE equivalent. <b>On the right:</b> The time channel in nanoseconds. . . . .	60
2339	4.7	. . . . .	61
2340	a	Distribution of PE/MeV in the J23 Dataset. This distribution is profiled and fitted using equation 4.6 . . . . .	61
2341	b	<b>On top:</b> Distribution of PE vs Energy. <b>On bottom:</b> Using the values extracted in 4.7a, we calculate the ration signal over background + signal . . . . .	61
2342	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 [66]. The top part of each plot is the resolution and the bottom part is the bias. . . . .	63
2343	a	Resolution and bias of energy reconstruction vs energy . . . . .	63
2344	b	Resolution and bias of energy reconstruction vs radius . . . . .	63
2345	c	Resolution and bias of radius reconstruction vs energy . . . . .	63
2346	d	Resolution and bias of radius reconstruction vs radius . . . . .	63
2347	e	Resolution and bias of radius reconstruction vs $\theta$ . . . . .	63
2348	f	Resolution and bias of radius reconstruction vs $\phi$ . . . . .	63
2349	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 $R$ , $\theta$ and $\phi$ . . . . .	64
2350	a	Distribution of the error on reconstructed $x$ by “gen_30” . . . . .	64

2378 <b>b</b>	Distribution of the error on reconstructed $y$ by "gen_30" . . . . .	64
2379 <b>c</b>	Distribution of the error on reconstructed $z$ by "gen_30" . . . . .	64
2380 <b>d</b>	Distribution of the error on reconstructed $R$ by "gen_30" . . . . .	64
2381 <b>e</b>	Distribution of the error on reconstructed $\theta$ by "gen_30" . . . . .	64
2382 <b>f</b>	Distribution of the error on reconstructed $\phi$ by "gen_30" . . . . .	64
2383 <b>4.10</b>	. . . . .	65
2384 <b>a</b>	Distribution of "gen_30" reconstructed energy and true energy of the analysis dataset (J21) . . . . .	65
2385 <b>b</b>	Distribution of "gen_42" reconstructed energy and true energy of the analysis dataset (J23) . . . . .	65
2388 <b>4.11</b>	Radius bias ( <b>on the left</b> ) and resolution( <b>on the right</b> ) of the classical algorithm in a $E, R^3$ grid . . . . .	65
2389 <b>4.12</b>	Reconstruction performance of the "gen_30" model on J21, the classic algorithm "Classical algorithm" from [66] and the combination of both using weighted mean. The top part of each plot is the resolution and the bottom part is the bias. . . . .	66
2393 <b>a</b>	Resolution and bias of energy reconstruction vs energy . . . . .	66
2394 <b>b</b>	Resolution and bias of energy reconstruction vs radius . . . . .	66
2395 <b>c</b>	Resolution and bias of radius reconstruction vs energy . . . . .	66
2396 <b>d</b>	Resolution and bias of radius reconstruction vs radius . . . . .	66
2397 <b>e</b>	Resolution and bias of radius reconstruction vs $\theta$ . . . . .	66
2398 <b>f</b>	Resolution and bias of radius reconstruction vs $\phi$ . . . . .	66
2399 <b>4.13</b>	Correlation between CNN and classical method reconstruction ( <b>on the left</b> ) for energy and ( <b>on the right</b> ) for radius in a $E, R^3$ grid . . . . .	67
2400 <b>4.14</b>	Reconstruction performance of the "gen_42" model on J23 data and it's comparison to the performances of the classic algorithm "Classical algorithm" from [66]. The top part of each plot is the resolution and the bottom part is the bias. . . . .	68
2404 <b>a</b>	Resolution and bias of energy reconstruction vs energy . . . . .	68
2405 <b>b</b>	Resolution and bias of energy reconstruction vs radius . . . . .	68
2406 <b>c</b>	Resolution and bias of radius reconstruction vs energy . . . . .	68
2407 <b>d</b>	Resolution and bias of radius reconstruction vs radius . . . . .	68
2408 <b>e</b>	Resolution and bias of radius reconstruction vs $\theta$ . . . . .	68
2409 <b>f</b>	Resolution and bias of radius reconstruction vs $\phi$ . . . . .	68
2410 <b>5.1</b>	. . . . .	71
2411 <b>a</b>	Illustration of the different nodes in our graphs and their relations. . . . .	71
2412 <b>b</b>	Illustration of what a dense adjacency matrix would looks like and the part we are really interested in. Because Fired $\rightarrow$ Mesh and Mesh $\rightarrow$ I/O relations are undirected, we only consider in practice the top right part of the matrix for those relations. . . . .	71
2416 <b>5.2</b>	Illustration of the healpix segmentation. <b>On the left:</b> A segmentation of order 0. <b>On the right:</b> A segmentation of order 1 . . . . .	71
2418 <b>5.3</b>	Features held by the nodes and edges in the graph. $D_{m_1 \rightarrow m_2}^{-1}$ is the inverse of the distance between two mesh center. The features $P_l^h, \mathbb{A}$ and $\mathbb{B}$ are detailed in section 5.2 . . . . .	72
2420 <b>5.4</b>	Illustration of the different update function needed by our GNN . . . . .	73
2422 <b>5.5</b>	Distribution of the number of hits depending on the energy. <b>On the right:</b> for the LPMT system. <b>In the middle :</b> for the SPMT system. <b>On the left:</b> For both system. . . . .	74
2423 <b>a</b>	. . . . .	74
2424 <b>b</b>	. . . . .	74
2425 <b>c</b>	. . . . .	74
2426 <b>5.6</b>	Distribution of the number of hits depending on the radius. <b>On the right:</b> for the LPMT system. <b>On the right :</b> for the SPMT system. To prevent the superposition of structure of different scales we limit ourselves to the energy range $E_{true} \in [0, 9]$ . . . . .	75
2429 <b>a</b>	. . . . .	75

2430	<b>b</b>	75
2431	5.7 Schema of the JWGv8.4.0 architecture, the colored triplet is the graph configuration after each JWG layers	76
2432		
2433	5.8 Comparison between Omilrec $E_{rec}$ and the true energy $E_{true}$ . The profile of the distribution $E_{true}/E_{rec}$ vs $E_{rec}$ is fitted with a 5th degree polynomial.	78
2434		
2435	5.9 Reconstruction performance of the Omilrec algorithm based on QTML presented in section 2.6, JWGv8.4 presented in this chapter and the combination between the two as presented in section 4.4.2. The top part of each plot is the resolution and the bottom part is the bias.	79
2436		
2437	a Resolution and bias of energy reconstruction vs energy	79
2438	b Resolution and bias of energy reconstruction vs radius	79
2439	c Resolution and bias of radius reconstruction vs energy	79
2440	d Resolution and bias of radius reconstruction vs radius	79
2441	e Resolution and bias of radius reconstruction vs $\theta$	79
2442	f Resolution and bias of radius reconstruction vs $\phi$	79
2443		
2444		
2445	7.1 Two oscillated spectrums of $1e7$ event expected in JUNO. In red the spectrums without supplementary QNL. In blue the same spectrum but where an event-wise QNL $\alpha_{qnl} = 10\%$ is added.	85
2446		
2447	7.2	86
2448		
2449	a Distribution of ratio of collected nPE after the additional QNL over the number of nPE that would be collected for different $\gamma_{qnl}$ . We select event with an interaction radius $R < 4m$ to not be affected by the non-uniformity.	86
2450		
2451	b Ratio of collected nPE after the additional QNL over the number of nPE that would be collected at different energies. We select event with an interaction radius $R < 4m$ to not be affected by the non-uniformity. The dots represent the mean of the distribution in figure 7.2a and the dashed line are the equivalent event-wise non-linearity from eq 7.2. The hatched zone is the residual non-linearity expected after calibration [29].	86
2452		
2453	7.3 Theoretical LPMT spectrum at nominal oscillation values binned using 410 from 0.8 to 9 MeV. It is rescaled to correspond to the amount of event expected in 6 years of data taking. The black line represent the 335 bin cut	90
2454		
2455	7.4 Schematic description of the fit framework	91
2456		
2457	7.5 Relative (On the left) and absolute (On the right) resolution of the LPMT and SPMT systems used in this study. The number in parenthesis are the parameter A, B and C respectively for each systems.	92
2458		
2459	7.6 Distribution of BFP - nominal value for 1000 toy Standard joint fit. 6 years exposure, all background, Pearson $\chi^2$ , $\theta_{13}$ fixed.	95
2460		
2461	7.7 Distribution of BFP - nominal value for 1000 toy Standard joint fit. 6 years exposure, all background, PearsonV $\chi^2$ , $\theta_{13}$ fixed.	95
2462		
2463	7.8 Distribution of BFP - nominal value for 5000 toy Delta joint fit. 6 years exposure, all background, PearsonV $\chi^2$ , $\theta_{13}$ fixed.	96
2464		
2465	7.9 Top: Theoretical spectrum without QNL (in red) and with $\alpha_{qnl} = 1\%$ (in blue). Bottom: Ratio between the theoretical spectrum with and without QNL.	97
2466		
2467	7.10 Theoretical correlation matrix between the LPMT spectrum (bins 0-409) ans the SPMT spectrum (410-819). The diagonal has been set to 0 (it was 1) for readability purpose.	99
2468		
2469	7.11 Upper left corner of the estimated correlation matrix between the LPMT and SPMT spectrum for different configuration of N toy with different number of M events per toy	99
2470		
2471	a	99
2472	b	99
2473	c	99
2474		
2475	7.12 Difference between the element of the theoretical and empiric correlation matrix	100
2476		
2477	a	100
2478		
2479		

2482	<b>b</b>	100
2483	<b>7.13</b> Distribution of the $\chi^2_{spe}$ for 1000 toys for different exposure. The dashed line represent 2484 the median of the distribution and the p-value are the percentage of the $\alpha_{qnl} = 0$ 2485 distribution that are greater than those medians.	101
2486	<b>B.1</b> Illustration of the real part of the spherical harmonics	108
2487	<b>B.2</b> Scatter plot of the absolute and relative power, respectively on the left and right plot, 2488 of each harmonic degree $l$ . The color indicate the radius of the event.	108
2489	<b>B.3</b> Error on the reconstructed radius vs the true radius by the harmonic method	109
2490	<b>B.4</b> Charge repartition in JUNO as seen by the Healpix segmentation. Those are Healpix 2491 map of order 5 (i.e. 12288 pixels). The color represent the summed charge of the PMTs 2492 in each pixels. The color scale is logarithmic. The view have been centered to prevent 2493 event deformations.	110
2494	<b>a</b>	110
2495	<b>b</b>	110
2496	<b>c</b>	110
2497	<b>d</b>	110
2498	<b>e</b>	110
2499	<b>f</b>	110
2500	<b>g</b>	110
2501	<b>h</b>	110
2502	<b>B.5</b> Scatter plot of the absolute and relative power, respectively on the left and right plot, 2503 of the $l = 0$ harmonic. The color indicate the radius of the event.	111
2504	<b>B.6</b> Plot of the distribution of the relative power of each harmonic dependent on $R^3$ (on 2505 the left). The Total Reflection (TR) area is represented by the horizontal blue line. The 2506 distribution are fitted using a 9th degree polynomial (red curve). The relative power 2507 error between the distribution and the fit is represented on the left. <b>Part 1</b>	112
2508	<b>B.7</b> Plot of the distribution of the relative power of each harmonic dependent on $R^3$ (on 2509 the left). The Total Reflection (TR) area is represented by the horizontal blue line. The 2510 distribution are fitted using a 9th degree polynomial (red curve). The relative power 2511 error between the distribution and the fit is represented on the left. <b>Part 2</b>	113

# <sup>2512</sup> List of Abbreviations

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



# 2513 Bibliography

- 2514 [1] Liang Zhan, Yifang Wang, Jun Cao, and Liangjian Wen. "Determination of the Neutrino Mass  
 2515 Hierarchy at an Intermediate Baseline". *Physical Review D* 78.11 (Dec. 10, 2008), 111103. ISSN:  
 2516 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).
- 2517 [2] Fengpeng An et al. "Neutrino Physics with JUNO". *Journal of Physics G: Nuclear and Particle  
 2518 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).
- 2519 [3] Liang Zhan, Yifang Wang, Jun Cao, and Liangjian Wen. "Experimental Requirements to Deter-  
 2520 mine the Neutrino Mass Hierarchy Using Reactor Neutrinos". *Physical Review D* 79.7 (Apr. 14,  
 2521 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).
- 2522 [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*  
 2523 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:  
 2524 <https://www.sciencedirect.com/science/article/pii/0370269389915980> (visited on  
 2525 01/16/2024).
- 2526 [5] Th A. Mueller et al. "Improved Predictions of Reactor Antineutrino Spectra". *Physical Review C*  
 2527 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).  
 2528 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)  
 2529 (visited on 01/16/2024).
- 2530 [6] F. von Feilitzsch, A. A. Hahn, and K. Schreckenbach. "Experimental beta-spectra from 239Pu  
 2531 and 235U thermal neutron fission products and their correlated antineutrino spectra". *Physics  
 2532 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).  
 2533 URL: <https://www.sciencedirect.com/science/article/pii/0370269382906220> (visited  
 2534 on 01/16/2024).
- 2535 [7] K. Schreckenbach, G. Colvin, W. Gelletly, and F. Von Feilitzsch. "Determination of the antineu-  
 2536 trino spectrum from 235U thermal neutron fission products up to 9.5 MeV". *Physics Letters B*  
 2537 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:  
 2538 <https://www.sciencedirect.com/science/article/pii/0370269385913371> (visited on  
 2539 01/16/2024).
- 2540 [8] Patrick Huber. "On the determination of anti-neutrino spectra from nuclear reactors". *Physical  
 2541 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).  
 2542 URL: [http://arxiv.org/abs/1106.0687](https://arxiv.org/abs/1106.0687) (visited on 01/16/2024).
- 2543 [9] P. Vogel, G. K. Schenter, F. M. Mann, and R. E. Schenter. "Reactor antineutrino spectra and  
 2544 their application to antineutrino-induced reactions. II". *Physical Review C* 24.4 (Oct. 1, 1981).  
 2545 Publisher: American Physical Society, 1543–1553. DOI: [10 . 1103 / PhysRevC . 24 . 1543](https://doi.org/10.1103/PhysRevC.24.1543). URL:  
 2546 <https://link.aps.org/doi/10.1103/PhysRevC.24.1543> (visited on 01/16/2024).
- 2547 [10] D. A. Dwyer and T. J. Langford. "Spectral Structure of Electron Antineutrinos from Nuclear  
 2548 Reactors". *Physical Review Letters* 114.1 (Jan. 7, 2015), 012502. ISSN: 0031-9007, 1079-7114. DOI:  
 2549 [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:  
 2550 [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  $^{235}\text{U}$  and  $^{239}\text{Pu}$   $\beta^-$  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  $\beta^-$  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] JUNO Collaboration et al. "JUNO Sensitivity on Proton Decay  $p \rightarrow \bar{\nu}K^+$  Searches". *Chinese Physics C* 47.11 (Nov. 1, 2023), 113002. ISSN: 1674-1137, 2058-6132. DOI: [10.1088/1674-1137/ace9c6](https://doi.org/10.1088/1674-1137/ace9c6). eprint: [2212.08502 \[hep-ex, physics:hep-ph\]](https://arxiv.org/abs/2212.08502). URL: <http://arxiv.org/abs/2212.08502> (visited on 08/09/2024).
- [19] 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).
- [20] 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).
- [21] 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).
- [22] Photomultiplier tube R12860 | Hamamatsu Photonics. URL: [https://www.hamamatsu.com/eu/en/product/optical-sensors/pmt/pmt\\_tube-alone/head-on-type/R12860.html](https://www.hamamatsu.com/eu/en/product/optical-sensors/pmt/pmt_tube-alone/head-on-type/R12860.html) (visited on 02/08/2024).
- [23] 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).

- [24] 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: [1703.01867](https://arxiv.org/abs/1703.01867) [hep-ex, physics:physics]. URL: <http://arxiv.org/abs/1703.01867> (visited on 07/28/2023).
- [25] JUNO Collaboration et al. *The Design and Sensitivity of JUNO's scintillator radiopurity pre-detector 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: <http://arxiv.org/abs/2103.16900> (visited on 02/07/2024).
- [26] Angel Abusleme et al. "Mass Testing and Characterization of 20-inch PMTs for JUNO". *The 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> (visited on 02/08/2024).
- [27] Yang Han. "Dual Calorimetry for High Precision Neutrino Oscillation Measurement at JUNO Experiment". AstroParticule et Cosmologie, France, Paris U. VII, APC, June 2021.
- [28] R. Acquaferredda et al. "The OPERA experiment in the CERN to Gran Sasso neutrino beam". *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 02/29/2024).
- [29] JUNO collaboration et al. "Calibration Strategy of the JUNO Experiment". *Journal of High 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: [2011.06405](https://arxiv.org/abs/2011.06405) [hep-ex, physics:physics]. URL: <http://arxiv.org/abs/2011.06405> (visited on 08/10/2023).
- [30] 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> (visited on 01/16/2024).
- [31] 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).
- [32] S. Agostinelli et al. "Geant4—a simulation toolkit". *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 506.3 (July 1, 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).
- [33] J. Allison et al. "Geant4 developments and applications". *IEEE Transactions on Nuclear Science* 53.1 (Feb. 2006). Conference Name: IEEE Transactions on Nuclear Science, 270–278. ISSN: 1558-1578. 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).
- [34] J. Allison et al. "Recent developments in Geant4". *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 835 (Nov. 1, 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).
- [35] Wenjie Wu, Miao He, Xiang Zhou, and Haoxue Qiao. "A new method of energy reconstruction for large spherical liquid scintillator detectors". *Journal of Instrumentation* 14.3 (Mar. 8, 2019), 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> (visited on 07/28/2023).
- [36] Guihong Huang et al. "Improving the energy uniformity for large liquid scintillator detectors". *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, 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).
- [37] Ziyuan Li et al. "Event vertex and time reconstruction in large volume liquid scintillator detector". *Nuclear Science and Techniques* 32.5 (May 2021), 49. ISSN: 1001-8042, 2210-3147. DOI:

- 2663        [10.1007/s41365-021-00885-z](https://doi.org/10.1007/s41365-021-00885-z). eprint: [2101.08901 \[hep-ex, physics:physics\]](https://arxiv.org/abs/2101.08901). URL: <http://arxiv.org/abs/2101.08901> (visited on 07/28/2023).
- 2664
- 2665 [38] 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).
- 2666
- 2667
- 2668
- 2669
- 2670 [39] 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).
- 2671
- 2672
- 2673
- 2674
- 2675 [40] 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).
- 2676
- 2677
- 2678
- 2679 [41] 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).
- 2680
- 2681
- 2682
- 2683 [42] 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).
- 2684
- 2685
- 2686
- 2687
- 2688 [43] 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).
- 2689
- 2690
- 2691
- 2692
- 2693 [44] 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).
- 2694
- 2695
- 2696
- 2697 [45] 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).
- 2698
- 2699
- 2700
- 2701
- 2702 [46] 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).
- 2703
- 2704
- 2705
- 2706 [47] 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).
- 2707
- 2708
- 2709
- 2710 [48] 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).
- 2711
- 2712
- 2713
- 2714

- [49] 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://doi.org/10.48550/arXiv.1606.09375). eprint: [1606.09375\[cs, stat\]](https://arxiv.org/abs/1606.09375). URL: <http://arxiv.org/abs/1606.09375> (visited on 04/04/2024).
- [50] JUNO Collaboration et al. "JUNO Physics and Detector". *Progress in Particle and Nuclear Physics* 123 (Mar. 2022), 103927. ISSN: 01466410. DOI: [10.1016/j.ppnp.2021.103927](https://doi.org/10.1016/j.ppnp.2021.103927). eprint: [2104.02565\[hep-ex\]](https://arxiv.org/abs/2104.02565). URL: <http://arxiv.org/abs/2104.02565> (visited on 09/18/2023).
- [51] Leo Breiman, Jerome Friedman, R. A. Olshen, and Charles J. Stone. *Classification and Regression Trees*. New York: Chapman and Hall/CRC, Oct. 25, 2017. 368 pp. ISBN: 978-1-315-13947-0. DOI: [10.1201/9781315139470](https://doi.org/10.1201/9781315139470).
- [52] 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: 0090-5364, 2168-8966. DOI: [10.1214/aos/1013203451](https://doi.org/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).
- [53] J. Y. Lettvin, H. R. Maturana, W. S. McCulloch, and W. H. Pitts. "What the Frog's Eye Tells the Frog's Brain". *Proceedings of the IRE* 47.11 (Nov. 1959). Conference Name: Proceedings of the IRE, 1940–1951. ISSN: 2162-6634. DOI: [10.1109/JRPROC.1959.287207](https://doi.org/10.1109/JRPROC.1959.287207). URL: <https://ieeexplore.ieee.org/document/4065609> (visited on 05/06/2024).
- [54] F. Rosenblatt. "The perceptron: A probabilistic model for information storage and organization in the brain". *Psychological Review* 65.6 (1958). Place: US Publisher: American Psychological Association, 386–408. ISSN: 1939-1471. DOI: [10.1037/h0042519](https://doi.org/10.1037/h0042519).
- [55] Diederik P. Kingma and Jimmy Ba. *Adam: A Method for Stochastic Optimization*. Jan. 29, 2017. DOI: [10.48550/arXiv.1412.6980](https://doi.org/10.48550/arXiv.1412.6980). eprint: [1412.6980\[cs\]](https://arxiv.org/abs/1412.6980). URL: <http://arxiv.org/abs/1412.6980> (visited on 05/13/2024).
- [56] Olga Russakovsky et al. *ImageNet Large Scale Visual Recognition Challenge*. Jan. 29, 2015. DOI: [10.48550/arXiv.1409.0575](https://doi.org/10.48550/arXiv.1409.0575). eprint: [1409.0575\[cs\]](https://arxiv.org/abs/1409.0575). URL: <http://arxiv.org/abs/1409.0575> (visited on 05/17/2024).
- [57] Karen Simonyan and Andrew Zisserman. *Very Deep Convolutional Networks for Large-Scale Image Recognition*. Apr. 10, 2015. DOI: [10.48550/arXiv.1409.1556](https://doi.org/10.48550/arXiv.1409.1556). eprint: [1409.1556\[cs\]](https://arxiv.org/abs/1409.1556). URL: <http://arxiv.org/abs/1409.1556> (visited on 05/17/2024).
- [58] Anna Allen. *generic-github-user/Image-Convolution-Playground*. original-date: 2018-09-28T22:42:55Z. July 15, 2024. URL: <https://github.com/generic-github-user/Image-Convolution-Playground> (visited on 07/16/2024).
- [59] Jason Ansel et al. *PyTorch 2: Faster Machine Learning Through Dynamic Python Bytecode Transformation and Graph Compilation*. Publication Title: 29th ACM International Conference on Architectural Support for Programming Languages and Operating Systems, Volume 2 (ASPLOS '24) original-date: 2016-08-13T05:26:41Z. Apr. 2024. DOI: [10.1145/3620665.3640366](https://doi.org/10.1145/3620665.3640366). URL: <https://pytorch.org/assets/pytorch2-2.pdf> (visited on 07/16/2024).
- [60] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner. "Gradient-based learning applied to document recognition". *Proceedings of the IEEE* 86.11 (Nov. 1998). Conference Name: Proceedings of the IEEE, 2278–2324. ISSN: 1558-2256. DOI: [10.1109/5.726791](https://doi.org/10.1109/5.726791). URL: <https://ieeexplore.ieee.org/document/726791> (visited on 07/16/2024).
- [61] 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).
- [62] Justin Gilmer, Samuel S. Schoenholz, Patrick F. Riley, Oriol Vinyals, and George E. Dahl. *Neural Message Passing for Quantum Chemistry*. June 12, 2017. DOI: [10.48550/arXiv.1704.01212](https://doi.org/10.48550/arXiv.1704.01212). eprint: [1704.01212\[cs\]](https://arxiv.org/abs/1704.01212). URL: <http://arxiv.org/abs/1704.01212> (visited on 05/22/2024).
- [63] Yaguang Li, Rose Yu, Cyrus Shahabi, and Yan Liu. *Diffusion Convolutional Recurrent Neural Network: Data-Driven Traffic Forecasting*. Feb. 22, 2018. DOI: [10.48550/arXiv.1707.01926](https://doi.org/10.48550/arXiv.1707.01926). eprint: [1707.01926\[cs, stat\]](https://arxiv.org/abs/1707.01926). URL: <http://arxiv.org/abs/1707.01926> (visited on 05/22/2024).

- [64] Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. *Generative Adversarial Networks*. June 10, 2014. DOI: [10.48550/arXiv.1406.2661](https://doi.org/10.48550/arXiv.1406.2661). eprint: [1406.2661\[cs, stat\]](https://arxiv.org/abs/1406.2661). URL: <http://arxiv.org/abs/1406.2661> (visited on 05/29/2024).
- [65] 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](https://doi.org/10.1109/CVPR.2016.90). URL: <https://ieeexplore.ieee.org/document/7780459> (visited on 07/17/2024).
- [66] 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. URL: <https://theses.fr/2022NANU4080> (visited on 05/22/2024).
- [67] 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](https://doi.org/10.48550/arXiv.1202.2745). eprint: [1202.2745\[cs\]](https://arxiv.org/abs/1202.2745). URL: <http://arxiv.org/abs/1202.2745> (visited on 06/27/2024).
- [68] 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](https://doi.org/10.1088/1748-0221/16/07/P07041). eprint: [2101.11589\[hep-ex\]](https://arxiv.org/abs/2101.11589). URL: <http://arxiv.org/abs/2101.11589> (visited on 06/27/2024).
- [69] 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](https://doi.org/10.1088/1475-7516/2021/11/051). URL: <https://dx.doi.org/10.1088/1475-7516/2021/11/051> (visited on 06/27/2024).
- [70] Taco S. Cohen, Mario Geiger, Jonas Koehler, and Max Welling. *Spherical CNNs*. Feb. 25, 2018. DOI: [10.48550/arXiv.1801.10130](https://doi.org/10.48550/arXiv.1801.10130). eprint: [1801.10130\[cs, stat\]](https://arxiv.org/abs/1801.10130). URL: <http://arxiv.org/abs/1801.10130> (visited on 07/13/2024).
- [71] NVIDIA A100 GPUs Power the Modern Data Center. NVIDIA. URL: <https://www.nvidia.com/en-gb/data-center/a100/> (visited on 08/06/2024).
- [72] NVIDIA V100. NVIDIA. URL: <https://www.nvidia.com/en-gb/data-center/v100/> (visited on 08/06/2024).
- [73] Leonard Imbert. *leonard-IMBERT/datamo*. original-date: 2023-10-17T12:37:38Z. Aug. 9, 2024. URL: <https://github.com/leonard-IMBERT/datamo> (visited on 08/09/2024).
- [74] “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](https://doi.org/10.1109/IEEEESTD.2019.8766229). URL: <https://ieeexplore.ieee.org/document/8766229> (visited on 07/03/2024).
- [75] 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](https://doi.org/10.1016/j.nima.2021.165347). eprint: [2102.11538\[hep-ex, physics:physics\]](https://arxiv.org/abs/2102.11538). URL: <http://arxiv.org/abs/2102.11538> (visited on 02/08/2024).
- [76] K. M. Gorski, E. Hivon, A. J. Banday, B. D. Wandelt, F. K. Hansen, M. Reinecke, and M. Bartelman. “HEALPix – a Framework for High Resolution Discretization, and Fast Analysis of Data Distributed on the Sphere”. *The Astrophysical Journal* 622.2 (Apr. 2005), 759–771. ISSN: 0004-637X, 1538-4357. DOI: [10.1086/427976](https://doi.org/10.1086/427976). eprint: [astro-ph/0409513](https://arxiv.org/abs/astro-ph/0409513). URL: <http://arxiv.org/abs/astro-ph/0409513> (visited on 08/10/2023).
- [77] Teng Li, Xin Xia, Xing-Tao Huang, Jia-Heng Zou, Wei-Dong Li, Tao Lin, Kun Zhang, and Zi-Yan Deng. “Design and development of JUNO event data model\*”. *Chinese Physics C* 41.6 (June 2017). Publisher: IOP Publishing, 066201. ISSN: 1674-1137. DOI: [10.1088/1674-1137/41/6/066201](https://doi.org/10.1088/1674-1137/41/6/066201). URL: <https://dx.doi.org/10.1088/1674-1137/41/6/066201> (visited on 08/16/2024).

- 2819 [78] Martin Reinecke. *Ducc0*. original-date: 2021-04-12T15:35:50Z. Aug. 9, 2024. URL: <https://gitlab.mpcdf.mpg.de/mtr/ducc> (visited on 08/16/2024).
- 2820 [79] Anatael Cabrera et al. *Multi-Calorimetry in Light-based Neutrino Detectors*. Dec. 20, 2023. DOI: [10.48550/arXiv.2312.12991](https://doi.org/10.48550/arXiv.2312.12991). eprint: [2312.12991\[hep-ex, physics:physics\]](https://arxiv.org/abs/2312.12991). URL: [http://arxiv.org/abs/2312.12991](https://arxiv.org/abs/2312.12991) (visited on 08/19/2024).
- 2821 [80] Angel Abusleme et al. “Potential to Identify the Neutrino Mass Ordering with Reactor Antineutrinos in JUNO” (May 2024). eprint: [2405.18008](https://arxiv.org/abs/2405.18008).
- 2822 [81] Rene Brun et al. *root-project/root: v6.26/06*. Version v6.26-06. Mar. 3, 2022. DOI: [10.5281/zenodo.3895860](https://doi.org/10.5281/zenodo.3895860). URL: <https://zenodo.org/records/3895860> (visited on 09/05/2024).
- 2823 [82] X. B. Ma, W. L. Zhong, L. Z. Wang, Y. X. Chen, and J. Cao. “Improved calculation of the energy release in neutron-induced fission”. *Physical Review C* 88.1 (July 12, 2013). Publisher: American Physical Society, 014605. DOI: [10.1103/PhysRevC.88.014605](https://doi.org/10.1103/PhysRevC.88.014605). URL: <https://link.aps.org/doi/10.1103/PhysRevC.88.014605> (visited on 09/06/2024).
- 2824 [83] Daya Bay Collaboration et al. “Measurement of the Reactor Antineutrino Flux and Spectrum at Daya Bay”. *Physical Review Letters* 116.6 (Feb. 12, 2016). Publisher: American Physical Society, 061801. DOI: [10.1103/PhysRevLett.116.061801](https://doi.org/10.1103/PhysRevLett.116.061801). URL: <https://link.aps.org/doi/10.1103/PhysRevLett.116.061801> (visited on 09/06/2024).
- 2825 [84] Timo Gnambs. “A Brief Note on the Standard Error of the Pearson Correlation”. *Collabra: Psychology* 9.1 (Sept. 6, 2023). Ed. by Thomas Evans, 87615. ISSN: 2474-7394. DOI: [10.1525/collabra.87615](https://doi.org/10.1525/collabra.87615). URL: <https://doi.org/10.1525/collabra.87615> (visited on 09/10/2024).
- 2826 [85] “Note Sur Une Méthode de Résolution des équations Normales Provenant de L’Application de la Méthode des Moindres Carrés à un Système D’équations Linéaires en Nombre Inférieur à Celui des Inconnues. — Application de la Méthode à la Résolution D’un Système Défini D’éQuations LinéAires”. *Bulletin géodésique* 2.1 (Apr. 1, 1924), 67–77. ISSN: 1432-1394. DOI: [10.1007/BF03031308](https://doi.org/10.1007/BF03031308). URL: <https://doi.org/10.1007/BF03031308> (visited on 09/10/2024).
- 2827 [86] Pauli Virtanen et al. “SciPy 1.0: fundamental algorithms for scientific computing in Python”. *Nature Methods* 17.3 (Mar. 2020). Publisher: Nature Publishing Group, 261–272. ISSN: 1548-7105. DOI: [10.1038/s41592-019-0686-2](https://doi.org/10.1038/s41592-019-0686-2). URL: <https://www.nature.com/articles/s41592-019-0686-2> (visited on 08/14/2024).