

中国科学技术大学

博士学位论文



大亚湾反应堆中微子实验离线软件和中
子本底研究

作者姓名: 张一纯

学科专业: 粒子物理与核物理

导师姓名: 张子平 教授
王贻芳 研究员

完成时间: 二零一二年五月

University of Science and Technology of China
A dissertation for doctor degree



Development of the Daya Bay Neutrino Offline Software
and Studies on the Neutron Background

Author : Yichun Zhang

Speciality : Particle and Nuclear Physics

Supervisor : Prof. Ziping Zhang

Prof. Yifang Wang

Finished Time : May, 2012

中国科学技术大学学位论文原创性声明

本人声明所呈交的学位论文, 是本人在导师指导下进行研究工作所取得的成果。除已特别加以标注和致谢的地方外, 论文中不包含任何他人已经发表或撰写过的研究成果。与我一同工作的同志对本研究所做的贡献均已在论文中作了明确的说明。

作者签名: _____ 签字日期: _____

中国科学技术大学学位论文授权使用声明

作为申请学位的条件之一, 学位论文著作权拥有者授权中国科学技术大学拥有学位论文的部分使用权, 即: 学校有权按有关规定向国家有关部门或机构送交论文的复印件和电子版, 允许论文被查阅和借阅, 可以将学位论文编入《中国学位论文全文数据库》等有关数据库进行检索, 可以采用影印、缩印或扫描等复制手段保存、汇编学位论文。本人提交的电子文档的内容和纸质论文的内容相一致。

保密的学位论文在解密后也遵守此规定。

公开 保密 _____ 年

作者签名: _____ 导师签名: _____

签字日期: _____ 签字日期: _____

摘 要

中微子味混合和振荡是目前唯一得到诸多实验结果支持的超出粒子物理标准模型的现象，被认为是发现新物理的突破口。经过半个多世纪的进展，中微子实验已经从发现振荡现象进入到精确测量振荡参数的时代。中微子振荡参数 θ_{13} 是参数化振荡矩阵的六个基本参数之一，其值对未来的中微子实验，如测量混合相位角 δ 或者中微子质量等级具有重要意义。大亚湾反应堆中微子实验的物理目标是精确测量中微子混合角参数 $\sin^2 2\theta_{13}$ 到 0.01 或更好。按照这一物理目标的要求，大亚湾反应堆中微子实验的远近三个实验点共有 8 个全同的探测器模块，每个模块系统误差小于 0.38%，靶质量大于 20 吨，能量分辨好于 $15\%/\sqrt{E}$ 。

作为国内第一个反应堆中微子物理实验，为了配合实验物理数据的特点，大亚湾的在线实时数据获取和离线数据分析软件系统采用了新的技术设计和实现，如面向对象的软件设计模式、粒子物理实验先进软件框架 Gaudi 的使用，动态弱语言 python 作为脚本配置语言的应用，在软件重用、模块化的设计和开发以及数据质量监测，软件系统维护等多方面带来了新的用户体验。

访问原始数据作为离线数据处理的第一环节，是整个离线物理数据处理的基础。我们开发了基于 Gaudi 的原始数据转换服务，它为用户隐藏了原始数据格式相关的复杂细节，而提供了一致的数据访问接口。它主要用于完成对来自 DAQ 系统的二进制原始数据的解码，将电子学压缩信息还原为灵敏探测器击中信息的时间和电荷信号，并存储为离线系统中标准的数据对象，供后续用户分析使用。利用原始数据转换服务模块，我们完成了对探测器性能调试数据的分析工作，完成了对探测器电子学通道台阶和阈值的刻度。

反应堆中微子实验数据中含有大量带有时间关联的符合与反符合信号，时间关联分析是反应堆中微子物理分析中的重要部分，这一点尤其不同于对碰撞物理的实验数据。为产生可用于物理分析的蒙特卡洛模拟样本，需要按照物理事例的时间分布和探测器对物理事例的响应时间分布特点，重新构建各模拟样本

中的时间关联信息。为此，我们扩展了高能物理通用软件框架 Gaudi 中的标准数据 I/O 方式，添加了多格式，多数据流并行输入的接口，重新设计了带有事例回溯的数据循环机制，开发了事例抽样混合的高性能“Interleaving”算法，改进了电子学模拟的整个架构：包括灵敏探测器击中信息的划分和电子学信号在不同模拟模块中的传输方式，提升了蒙特卡洛模拟的效率和可靠性。在上述基础上，完成了在探测器击中级别 (SimHit) 和电子学读出级别 (Readout) 上对不同模拟样本的混合，产生了几十 TB 可用做物理分析的蒙特卡洛数据，并以此完成了初步的信号事例挑选和本底分析。

大亚湾反应堆中微子实验作为低本底、低能量、高精度的实验，需要对各种可能本底进行仔细研究，其中，宇宙线 μ 子产生的一系列关联本底尤其重要。本论文基于线性外推法对大亚湾三个月数据中，由宇宙线 μ 子产生的快中子本底进行了估算，并根据挑选得到的快中子样本，证明了线性外推法的可行性。我们还通过对来自宇宙线 μ 子的散裂中子的研究，计算得出大亚湾三个实验大厅测得的“ μ 子致散裂中子”的产额与 μ 子能量的关系，并与之前的其他地下实验和唯象模型进行了对比。

关键词： 中微子振荡，原始数据转换，时间关联分析，抽样混合，快中子本底，中子产额

ABSTRACT

Neutrino flavor mixing and oscillation is the only evidence of physics beyond the standard model in particle physics supported by various experimental results and considered as the key to search for new physics. After the development in more than half a century, neutrino experiments has come to a era of precise measurements of the oscillation parameters from the discovery of neutrino oscillation phenomena. Among the six parameters of the neutrino mixing matrix, the value of θ_{13} is important to the future neutrino experiments such as the measurement the mixing phase angle δ or neutrino mass hierarchy. The physics goal of Daya Bay reactor neutrino experiment is to measure $\sin^2 2\theta_{13}$ to a precision of 0.01 or better. To reach this physics goal, there are totally 8 identical detector modules with systematic error less than 0.38% and target mass lager than 20 tons at three experiment sites in Daya Bay reactor neutrino experiment. The energy resolution of each module is better than $15\%/\sqrt{E}$.

As the first reactor neutrino experiments in China, in order to adapt to the properties of the physics data, the Daya Bay online data acquisition and offline data analysis software system are implemented by new technology, such as design patterns in object-oriented software, advanced Gaudi framework for particle physics experiments, and the python language in configuration scripts, which brought new client experiences in call of software modules, modular design and development, data quality monitoring and maintenance of software system.

As the first step of the off-line data processing, the raw data access is the basis of the entire offline physics data analysis. We have developed raw data conversion service in Gaudi, which hid the trivial details about the raw data format to clients and provided uniform interfaces for data accessing. This service was mainly used to decode the binary raw data from the DAQ system, restored the compressed

information into time and charge for the hits in sensors. And these information are stored as the standard data object in offline system for subsequent analysis. Using the raw data conversion service module, we analyzed the data from the anti-neutrino detector commissioning, calibrated the pedestal and threshold of each channel in FEEs for the AD modules.

Lots of events are time correlated in the data of reactor neutrino experiment. Analysis of time correlation between different events is an important task in physics data analysis, which is different from data analysis in collision physics experiments. To get the MC samples suitable for physics analysis, time correlations from different MC samples should be reconstructed according to the time distributions of the physics events. With regards to this, we have upgraded the standard data I/O module in Gaudi framework by adding interfaces for the input of multi-format, multi-stream data in parallel and redesigned new data cycling mechanism based on event backtracking. We finally developed the "Interleaving" algorithm with high performance for MC events mixing. To enhance the efficiency and reliability of the MC simulation, we improved the architecture of the digitization simulation by reorganizing the hits from sensors and changing the transmission way for events between different modules. Based on the above works, we mixed different MC samples at both SimHit level and Readout level, and generated $\sim 10^2$ TB MC data for physics analysis. Based on the data, the preliminary events selection and background estimation were finished.

All possible backgrounds should be investigated carefully in Daya Bay reactor neutrino experiment for such a low background, low energy, high precision experiment. Among all the backgrounds, series related backgrounds caused by cosmic ray are particularly important. In this thesis, we have used linear extrapolation method to estimate the fast neutron background from cosmic μ based on three month data. We have checked the energy spectrum of fast neutron events to confirm the linear extrapolation method. We also have calculated the spallation neutron yielded from cosmic ray μ as the function of μ energy at the three different experimental site, and compared the results with that of other underground

ABSTRACT

experiments and the expectation of phenomenological models.

Keywords: neutrino oscillation, raw data conversion, time correlation analysis, re-sampling and mixing, fast neutron, neutron yield

目 录

摘 要	I
ABSTRACT	III
目 录	VII
表 格	XI
插 图	XVI
第一章 中微子物理和中微子物理实验	1
1.1 基本粒子及相互作用	1
1.2 中微子味道混合和振荡	3
1.2.1 中微子的发现和性质	3
1.2.2 中微子质量和味道混合	6
1.3 中微子实验	9
1.3.1 太阳中微子实验	10
1.3.2 大气中微子实验	12
1.3.3 反应堆中微子实验	14
1.3.4 中微子物理亟待解决的问题	16
1.4 大型粒子物理实验数据处理软件系统	17
1.5 论文选题的物理及软件背景和论文结构	19
1.5.1 物理及软件背景	19
1.5.2 论文结构	19

第二章 Daya Bay 反应堆中微子实验	21
2.1 大亚湾实验简介	21
2.1.1 大亚湾中微子实验站布局	23
2.1.2 大亚湾中微子探测器系统简介	24
2.2 大亚湾实验离线软件系统	27
2.2.1 离线数据处理平台	28
2.2.2 大亚湾离线软件框架 (NUWA)	29
2.2.3 大亚湾物理分析轻量级框架 (LAF)	33
2.3 大亚湾实验的初步结果	34
2.3.1 $\text{inverse}\beta$ 信号事例挑选和本底事例估计	34
2.3.2 物理灵敏度计算	37
2.4 本章小结	38
第三章 大亚湾 DAQ 数据到离线数据对象转换	41
3.1 NuWa 框架中的数据持久化机制	41
3.2 二进制原始数据格式和离线数据对象	43
3.2.1 二进制原始数据格式	43
3.2.2 Offline 数据对象模型	46
3.3 原始数据转换服务	48
3.3.1 原始数据输入与二进制解码	49
3.3.2 离线数据对象格式转换	51
3.4 原始数据转换器功能应用举例	53
3.4.1 FEE 电子学板的电子学台阶刻度	54
3.4.2 FEE 电子学板的电子学阈值刻度	55
3.5 原始数据转换器功能扩展	59
3.6 本章小结	60

第四章 Event-Mixing 软件框架	61
4.1 MC 事例混合的物理背景	61
4.1.1 大亚湾反应堆实验物理数据时间结构	62
4.1.2 NUWA 框架蒙卡数据模拟流程和混合方式	66
4.2 事例混合框架的软件设计	70
4.2.1 NUWA 框架事例循环机制和 I/O 模块	71
4.2.2 多样本并行输入流接口	72
4.2.3 事例混合核心算法实现及改进	73
4.2.4 事例回溯机制	80
4.2.5 基于“EvtSelector”机制的各混合模块的软件构建	81
4.2.6 电子学读出事例级混合算法的验证和应用	82
4.3 事例混合优化及应用扩展	85
4.3.1 探测器 SimHit 级别事例混合	85
4.3.2 NUWA 模拟流程改进	89
4.3.3 混合真实数据	91
4.4 本章小结	96
第五章 Daya Bay 实验宇宙线 μ 子致散裂中子产额及快中子本底研究	99
5.1 Daya Bay 反应堆中微子实验中的本底及分类	99
5.1.1 非关联本底	100
5.1.2 关联本底	101
5.2 Daya Bay 实验 μ 子致散裂中子产额估算	104
5.2.1 宇宙线 μ 子产生次级中子的物理过程	104
5.2.2 宇宙线 μ 子致散裂中子产额估算	106
5.3 线性外推法估算快中子本底	114
5.3.1 快中子挑选及结果	115
5.3.2 快中子本底估算	126
5.4 本章小结	127

第六章 总结与展望	131
6.1 总结和展望	131
6.1.1 论文总结	131
6.1.2 展望	132
参考文献	133
致 谢	141
在读期间发表的学术论文与取得的研究成果	145

表格

1.1	四种相互作用的性质。	3
1.2	大亚湾实验误差的要求 [50]	15
1.3	中微子信号和本底事例率的比值	16
2.1	国际上正在运行测量中微子混合振荡参数 θ_{13} 的实验比较。 . . .	21
2.2	实验岩石覆盖	24
2.3	信号和本底总结. 本底和 IBD 事例率被 $\epsilon_{\mu} \cdot \epsilon_m$ 效率所修正 [56] .	37
2.4	绝对效率, 关联和非关联系统误差汇总 [56]	38
3.1	PMT 击中电荷信息二进制编码。	45
3.2	PMT 击中时间信息二进制编码。	45
4.1	大亚湾反应堆实验三个实验大厅的反 β 衰变事例率	65
4.2	大亚湾实验大厅天然放射性事例率, $E > 1\text{MeV}$	65
4.3	大亚湾实验大厅宇宙线 μ 子事例率	66
4.4	混合算法以及 I/O 效率测试结果	97
5.1	μ 子致长寿命同位素	103
5.2	中微子探测器靶物质中, μ^- 在不同原子核上的俘获参数 [90] . .	106
5.3	计算中子产额的参数及误差	110
5.4	spallation 中子中的本底	112
5.5	大亚湾实验测得的宇宙线 μ 子致散裂中子产额	112
5.6	挑选得到的快中子样本中各种“本底”的信噪比估算结果。 . . .	126
5.7	快中子本底汇总	128

插图

1.1	基本粒子的分类。	2
1.2	Higgs 粒子和基本粒子	4
1.3	SNO 实验得到的中微子通量分布	11
1.4	KamLAND 实验得出的中微子存活几率	12
1.5	Super-K 实验得到的中微子事例随天顶角变化	13
1.6	Super-K 实验得到的中微子事例消失和传播距离的关系	14
1.7	历史上的反应堆中微子实验	15
2.1	大亚湾实验布局	23
2.2	大亚湾反中微子探测器	25
2.3	大亚湾反符合系统	27
2.4	离线软件架构	30
2.5	大亚湾模拟过程	31
2.6	大亚湾混合过程	32
2.7	远近点能谱比较和 θ_{13}	39
2.8	测量信号和预测无振荡信号比较	39
3.1	NuWa 框架事例持久化机制 (参考 Gaudi 手册)	42
3.2	二进制原始数据文件存储逻辑	44
3.3	二进制原始数据事例存储逻辑	44
3.4	电子学读出事例离线数据对象	46
3.5	TES 树形组织目录	47
3.6	二进制原始数据中间缓存对象模型	48
3.7	原始数据转换过程图	48
3.8	文件格式解析, 事例和数据格式解析工作流程图	51

3.9	数据转换管理服务中定义的数据管理器“树”	52
3.10	NuWa 框架数据持久化过程	53
3.11	FEE 前端电子学板通道的 Pedestal 分布	54
3.12	FEE 电子学板通道 Pedestal 中心值漂移	55
3.13	PMT 谱形全模型拟合 (无阈值) [73]	56
3.14	PMT 谱形的拟合 (横轴为 ADC 道数)	57
3.15	低阈值 PMT 谱形的拟合	58
3.16	误差函数拟合触发阈值	59
4.1	inverse β 衰变快慢信号及第一类时间关联	62
4.2	NUWA 框架蒙特卡洛数据模拟标准流程	67
4.3	NUWA 框架下产生子级事例混合流程	68
4.4	NUWA 框架下探测器击中级事例混合流程	69
4.5	不同物理事例的灵敏探测器击中的 overlap	69
4.6	NUWA 框架中 Readout 电子学读出级别事例混合	70
4.7	Gaudi 框架“算法—数据分离”设计思想实现。	71
4.8	并行“Interleaving”算法示意图。	74
4.9	并行“Interleaving”算法流程图。	75
4.10	混合过程中,可能遇到的不同物理事例电子学读出事例之间的时间重叠。	76
4.11	串行 interleaving 算法处理电子学读出事例时间重叠的示意。	78
4.12	串行“interleaving”算法流程图。	79
4.13	串行“interleaving”算法中,核心算法模块效率检测结果。	80
4.14	适配器模式示意图	81
4.15	混合算法软件模块结构示意图。	81
4.16	混合流程 UML 简化交互图。	83
4.17	混合算法逻辑性检验	84
4.18	MC 混合数据中 AD 能谱和真实数据 AD 能谱对比。	85
4.19	SimHit 级别事例混合的模块设计及数据流向,其中,灰色箭头为逻辑上的数据流向。	87

4.20	SimHit 级别事例混合的框架结构。	88
4.21	ElecSim 模块处理来自物理事例的 SimHit 的方式。	89
4.22	划分物理事例 SimHit 的方式。	90
4.23	两种电子学模拟划分物理事例后事例关联的比较。	90
4.24	SimHit 划分后电子学模拟真实数据之间的 ADC 和 TDC 分布比较	92
4.25	两种电子学模拟 ADC 结果对比	93
4.26	混合真实数据的算法模式设计	94
4.27	在“串行 Interleaving”算法基础上改进的适用于混合真实数据的 “InsertMix”算法	96
5.1	快中子事例形成的物理机制示意	101
5.2	快中子快信号 MC 模拟能谱和实验能谱分布	102
5.3	$^8\text{He}/^9\text{Li}$ 同位素衰变产生的快信号能谱图 [87]	103
5.4	挑选得到的 μ 子致散裂中子能谱	109
5.5	各地下实验宇宙线 μ 子致散裂中子产额与模拟对比	113
5.6	由 IWSMuon 标记的 μ 子致快中子的事例初步挑选示意图	116
5.7	初挑选得到的快中子信号能谱分布	116
5.8	初挑选得到的快中子信号顶点分布	117
5.9	初步挑选得到的“快中子”样本慢信号与 IWSMuon 信号时间间隔分布图。	117
5.10	去除了 μ 子衰变信号之后的快慢信号能谱分布，左图是快信号， 右图是慢信号	118
5.11	去除了 μ 子衰变信号之后的快慢信号顶点分布。上图是快信号顶 点分布，下图是慢信号顶点分布	118
5.12	探测器体积 cut 之后得到的本底快慢信号	119
5.13	探测器体积 cut 之后得到的快中子快慢信号	120
5.14	“快中子”样本中的快信号与 IWSMuon 信号时间间隔分布图	120
5.15	快中子样本快信号与 IWSMuon 信号时间间隔分布对比	121
5.16	快中子信号与其他本底信号分离对比图	122
5.17	最终得到的快中子信号顶点分布	123

5.18	快中子慢信号能谱拟合	124
5.19	米歇尔电子顶点分布	124
5.20	中子俘获时间模拟结果	125
5.21	大亚湾近点实验大厅, 宇宙线 μ 致散裂中子多重度 MC 模拟结果。	125
5.22	EH1AD2 和 EH2AD1 两个反中微子探测器进行快中子事例挑选的结果。左列是 EH1AD2 的结果, 右列是 EH2AD1 的结果。 . .	127
5.23	快中子本底的线性外推法估算结果	128

第一章 中微子物理和中微子物理实验

对中微子的研究是最近几年来粒子物理和天体物理领域内的前沿课题之一。中微子是构成物质世界的基本粒子，在微观的粒子物理世界和宏观的宇宙起源及演化中起着重大的作用，特别是发现了中微子存在不为零的静止质量和质量—味道本征态的混合之后，物理学家认为是发现超出标准模型的新物理的突破口。中微子在传播过程中，其味道本征态会发生改变的现象被称之为“中微子振荡现象”。太阳中微子实验 [1]、大气中微子实验 [2]、反应堆中微子实验 [3] 和加速器中微子实验 [4] 都证实了这一现象。这一现象的本质来源于中微子静止质量不为零的事实 [5]，也是迄今为止唯一被实验所证实的，超出标准模型的物理现象。由于中微子是已知的最轻的粒子，是所有粒子衰变的最终产物之一，宇宙中存在着巨量自大爆炸开始以来产生的中微子，它们可以构成宇宙中的热暗物质，对宇宙结构的形成有重要影响。中微子振荡中的 CP(空间与电荷共轭反演) 破缺可能与宇宙中的物质—反物质不对称现象有着深刻联系。

1.1 基本粒子及相互作用

在描述微观世界“基本粒子”运动规律最成功的标准模型 (Standard Model) 框架中，基本粒子按照它们参与的各种相互作用可以分为三类 [6]: 夸克、轻子和传播子。夸克和轻子都分别有三类，自旋均为 $\frac{1}{2}$ ，被认为是类点粒子，遵从 Pauli 不相容原理和费米—狄拉克统计。第一代是上夸克 (u)、下夸克 (d)、电子 (e) 和电子中微子 (ν_e)；第二代是粲夸克 (c)、奇夸克 (s)、 μ 子和 μ 子中微子；第三代为顶夸克 (t)、底夸克 (b)、 τ 子和 τ 子中微子。每一代的电荷总和相等。传播子都是规范矢量玻色子，遵从玻色—爱因斯坦统计。微观领域内的三种基本相互作用由它们传递，分别是：光子传递长程的电磁相互作用，胶子传递短程的强相互作用， W^\pm 和 Z^0 粒子传递弱相互作用。标准模型在几十年内历经不断的发展完善和世界范围内高能物理领域的大量实验检验，

与所有已知实验结果高度一致，精度达到 10^{-5} [7]。基本粒子的性质和四种相互作用的特点如图 1.1和表格1.1所示。

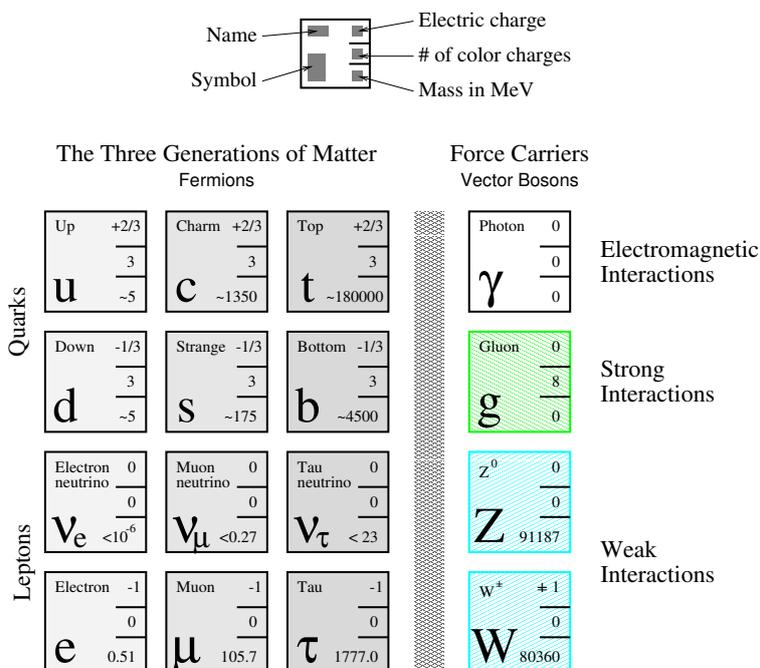


图 1.1 基本粒子的分类。

目前，标准模型检验中最热门的前沿议题是关于规范玻色子获得质量的 Higgs 机制中所预言的 Higgs 粒子的寻找，如图1.2所示。理论上预言其质量最有可能在 115.5 ~ 127GeV 之间。举世瞩目，坐落在欧洲核子中心的大型强子对撞机 (LHC) 上的 ATLAS 和 CMS 实验组在 2012 年 7 月 4 日公布了利用 2011 和 2012 年 $p - p$ 对撞的数据对 Higgs 质量研究的最新结果。CMS 组在 4.9σ 的置信水平下，宣称找到了质量为 125.3GeV 的新玻色子 [8]。而 ATLAS 则宣称新发现的玻色子质量为 126.5GeV，比 CMS 的结果稍高，置信水平为 5σ [9]。对于新发现的玻色子是否就是 Higgs 粒子，还有待于进一步的验证。人类对物质世界的认识和探索不会止步，可以想见，在粒子物理领域，无论是关于在高能量，高精度条件下对标准模型做精密检验，还是在对超越标准模型的尝试，探索新物理方面，都会有更进一步的实验和理论。

	电磁相互作用	弱相互作用	强相互作用	引力相互作用
源	电荷	弱荷	色荷	质量
作用强度	$\frac{e^2}{\hbar c} \simeq \frac{1}{137}$	$(\frac{m_p c}{\hbar})^2 \frac{G}{\hbar c} \sim 10^{-5}$	$\frac{g^2}{\hbar c} \simeq 10$	$\frac{G m_p^2}{\hbar c} \sim 10^{-38}$
力程 (m)	长程 ∞	短程 $\sim 10^{-18}$	短程 10^{-15}	长程 ∞
传递者	光子 γ	中间玻色子 W^\pm, Z^0	胶子 g	引力子 G
自旋宇称	1^-	$1^?$	1^-	2^+
质量 (GeV/c ²)	$M_\gamma = 0$	$M_{W^\pm} = 80.4$ $M_{Z^0} = 91.2$	$M_g = 0$	$M_G = 0$
理论描述	量子电动力学 (QED)	量子味动力学 (QFD)	量子色动力学 (QCD)	几何动力学 (广义相对论)

表 1.1 四种相互作用的性质。

1.2 中微子味道混合和振荡

1.2.1 中微子的发现和性质

1930 年, W.Pauli 面对原子核 β 衰变时能量似乎不守恒的困惑, 提出了一个猜想, 认为有一种当时无法探测的中性粒子带走了 β 衰变中的一部分能量 [10]。这种难以探测的粒子随后被 E.Fermi 命名为“中微子”, 意思是: 电中性的小家伙, 以区别于中子。

上世纪 50 年代, F.Reines 和 C.Cowan 试图用质子反衰变来直接探测中微子 [11]。他们采用了反应堆作为中微子源, 并于 1956 年, 十分清楚地直接探测到了中微子, F.Reines 因此获 1995 年诺贝尔奖。此后中微子实验研究开始得到很大发展。1960 年代初, J.Steinberger、Schwartz 和 Lederman 发现存在两种中微子: 电子中微子和 μ 子中微子 [12], 因此获得 1988 年诺贝尔奖。2000 年, τ 子中微子被发现 [13]。电子中微子、 μ 子中微子和 τ 子中微子分别对应电子、 μ 子和 τ 子。

20 世纪 50 年代初, 费曼和盖尔曼等提出了弱相互作用中的 V-A 理论, 这一理论认为中微子静止质量为零, 随后发展起来更为完善的电弱统一理论继承了这一论断, 并成为粒子物理标准模型的一部分。中微子静止质量是否为零会影响轻子数守恒、质子衰变、宇宙中的热暗物质以及宇宙结构等, 因此, 中微子静止质量问题一直是现代粒子物理和天体物理学中前沿领域内的热点问题。但由于中微子只与物质发生轻微的弱相互作用, 探测中微子及其质量是一件非

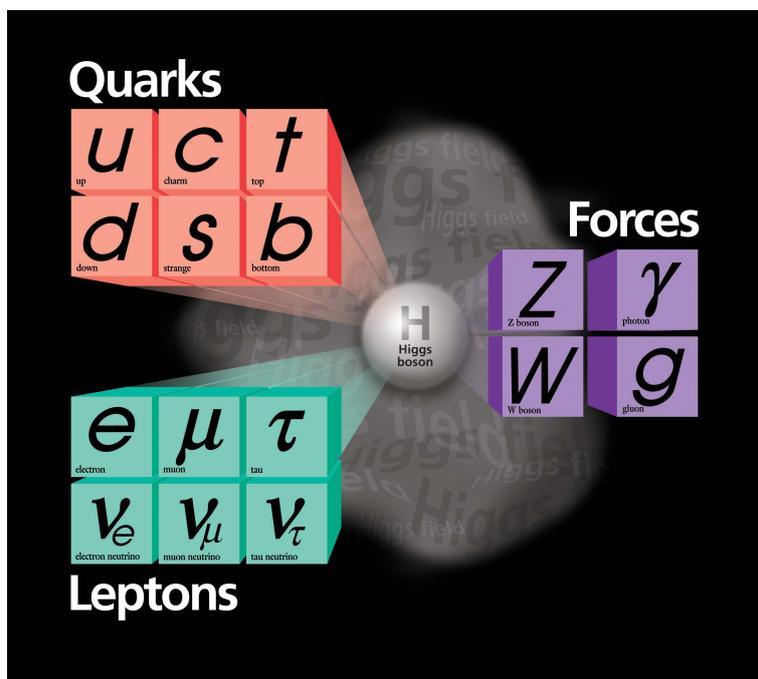


图 1.2 Higgs 粒子和基本粒子

常困难的事情，需要非常先进的探测手段。迄今，测量中微子静止质量的最可信的结果是通过 7 年 WMAP 观测得到的数据，并结合 BAO 和 SN 的数据，给出各代中微子质量之和 $\sum m_\nu$ 在 95% 置信度下小于 0.58eV[14]。

20 世纪 60 年代初，Maki、Nakagawa、Sakata 和 Pontecorvo 提出，如果中微子存在绝对质量，则其质量本征态很可能并非其味道本征态。由量子力学的态迭加原理，表象变换和传播理论可以推导得出，中微子在传播过程中，可以从一种味道变为另外一种味道，可以将这一现象形象地称之为中微子“振荡”。由于有三代中微子，其振荡形式可以由一个 3×3 矩阵进行描述，称作 PMNS(Maki-Nakagawa-Sakata-Portecorvo) 矩阵 [15]，与夸克理论中的 CKM 矩阵非常类似。中微子振荡现象很快引起物理学界的高度关注，因为它提供了一种能够间接证明中微子静止质量不为零的方法，但是，它却只能得到不同中微子质量平方差的绝对值。PMNS 矩阵中存在一个 CP 破缺项，目前，物理学界倾向认为这种 CP 破缺与宇宙中的物质—反物质不对称有关 [16]，之前 CKM 矩阵中的 CP 破缺项太小，不足以解释宇宙中为什么不存在一个反物质世界，按照宇宙的“大爆炸”模型，一个自然结果就是宇宙中对应存在相同数量的物

质与反物质。PMNS 矩阵中的 CP 破缺项似乎可以解决这个难题，这是中微子振荡引起物理学界高度关注的另一个重要原因。

同样在 20 世纪 60 年代初, J.Bahcall 提出了一个完整的太阳模型, 称之为标准太阳模型 (SSM)[17]。根据这一模型, 可以定量计算出太阳上发生的各种聚变反应所辐射出来的能量 (包括中微子携带的能量)。为了验证这一标准太阳模型, 美国物理学家 R.Davis 领导了 Homestake 实验组 [18] 来测量太阳核聚变中产生的 ν_e 中微子, Homestake 实验组于 1968 年首次发表了实验结果, 即观测到的 ν_e 中微子只有标准太阳模型预言的理论值的 1/3, 此结果在物理学界造成了轰动性的影响, 称之为“太阳中微子失踪”之谜。随后的一系列实验 Kamiokande[19], SAGE[20], GALLEX[21], SuperK[22] 和 GNO[23] 都证实了这一问题确实存在。面对“太阳中微子失踪”现象, 物理学界曾经给出了多种理论解释, “中微子振荡”引起来自太阳的 ν_e 中微子“丢失”, 只是其中的一种解释。上述太阳中微子实验结果仅仅能确认探测到的来自太阳的 ν_e 中微子比理论预言的少, 却无法确定“丢失”的中微子是由于“振荡”还是由于标准太阳模型理论本身的缺陷所致。这一问题的真正解决出现在 2001 年与 2002 年: 加拿大的 SNO[1] 实验通过带电流过程、中性流过程以及弹性散射过程同时测量来自太阳的三种中微子的通量, 获得了 ν_e 消失及 ν_μ 和 ν_τ 出现的证据, 且三种中微子通量之和与 J.Bahcall 的标准太阳模型预言一致, 证明了标准太阳模型对中微子产生机制描述基本正确, 说明一种味道中微子“丢失”现象可以用中微子振荡理论来解释。

1978 年, Wolfenstein 发现当中微子在物质中传播时其振荡行为会受到中微子与物质中的电子以及核子之间的相干向前弹性散射的影响 [24]。该影响相当于在中微子有效哈密顿量中增加一个有效势 V 。 V 的出现会改变中微子在物质中的有效质量和味混合角。1985 年, Mikheev 和 Smirnov 发现中微子在变密度的物质中传播时, 其振荡振幅可以被共振放大, 这便是 MSW 效应 [25]。该效应可以很好地解释太阳中微子由太阳中心传播到太阳表面的过程中发生味转化的问题。

“大气中微子反常”现象给出了中微子振荡的另外一个实验证据。根据理论计算, 主要来自宇宙线 μ 子衰变而来的大气中微子应满足 $\nu_\mu:\nu_e$ 的比值为 2。然而, 20 世纪 70 年代以来, Kamiokande[19]、IMB[26]、Soudan-2[27] 和

MACRO [28] 等大气中微子实验观测到这一比值约为 1, 与理论值不符, 即大气中微子反常。日本物理学家小柴昌俊主持的 SuperK 实验 [22] 在 1998 年发现 ν_μ 的减少与中微子的能量和传播距离有很强的依赖关系, 而且, 与中微子的振荡形状完全吻合, 从而首先给出大气中微子振荡的证据。

无论是太阳中微子实验还是大气中微子实验, 其中微子源都来源于大自然, 这一事实使得实验结果物理的诠释在某种程度上依赖于人们对大自然的理解, 如标准太阳模型、宇宙线 μ 子在大气中的簇射过程等。利用已知的人工中微子源中来探测中微子振荡, 是对中微子振荡现象的最佳证明。20 世纪 90 年代末期, 日本的 KEK 实验室用质子打靶产生中微子, 并由超级神冈探测器探测。2003 年 K2K 加速器中微子振荡实验证实了大气中微子振荡 [29], 同为加速器中微子实验的 MINOS[4] 实验也进一步证明振荡现象的存在。

通过观测反应堆裂变过程所产生的 $\bar{\nu}_e$ 的数目, 来研究中微子振荡现象, 称之为反应堆中微子实验。自 20 世纪 80 年代初开始, 已进行了好几代, 但是, 直到发展至 CHOOZ[30] 和 Palo Verde[31] 实验, 也都未能观测到明显的中微子振荡现象, 只给出了中微子混合角的上限。2001 年全面建成的 KamLAND 探测器在 2002 年报告了重要结果 [3]: 约 40% 的反应堆中微子消失了, 或者说振荡为另一种中微子, 而这种新的中微子没有被探测到。中微子是反应堆与太阳唯一共同的东西, 可以确定, 中微子消失现象只能用中微子本身的性质来解释, 同时, 中微子振荡理论可以定量解释所有观测到的数据, 这就证明了中微子的静止质量不为零, 至此中微子振荡得到最终的确认。

1.2.2 中微子质量和味道混合

标准模型认为: 中微子只有三代。这个结论得到 LEP 上的四个实验 L3[32]、DELPHI[33]、OPAL[34]、ALEPH[35] 的实验数据支持。中微子代数 N_ν 的综合结果是 $N_\nu = 2.984 \pm 0.008(\text{stat})$ [36]。综合这些实验的结果, 如果存在第四种中微子, 那么它可能非常重 (大于 Z^0 质量的一半), 或者不参与弱相互作用, 通常被称为惰性中微子 (sterile neutrino) [37]。加速器中微子实验 LSND[38] 在探测 $\bar{\nu}_\mu \rightarrow \bar{\nu}_e$ 振荡的时候, 发现中微子质量平方差为 $\Delta m_{LSND}^2 \sim 1\text{eV}$, 比太阳、大气中微子的质量平方差分别大了 5 个和 2 个量级。如果 LSND 实验被证实, 则说明或存在惰性中微子, 或者 CPT 破坏 [39], 或者存在非标准中微子的

相互作用。由于与 LSND 实验类似的 KARMEN 实验 [40] 并没有观测到 $\bar{\nu}_e$ 出现的信号，所以 LSND 实验被大多数人怀疑。为了验证这个结果，MiniBooNE 中微子实验 [41]，与 LSND 实验有着相似的 L/E_ν ，通过 $\nu_\mu \rightarrow \nu_e$ 过程来观测电子中微子的产生。2007 年和 2008 年的观测结果，在 98% 的置信水平上与 LSND 的结果不一致，但是 MiniBooNE 实验观测到的低能电子中微子信号偏多，与预言不一致，影响了结果的可信度。2010 年 MiniBooNE [42] 发表了反中微子振荡的结果，在 2 倍标准偏差下看到了类似 LSND 的振荡。2011 年，WMAP 公布的 7 年观测数据以 68% 的置信区间给出了对中微子代数的约束： $N_{eff} = 4.34^{+0.86}_{-0.88}$ [14]，看来是否存在惰性中微子还有很长的路要走。

在三代中微子的理论模型下，标记中微子的质量本征态为 ν_1, ν_2, ν_3 ，味道本征态为 ν_e, ν_μ, ν_τ 。根据量子力学的态迭加原理：中微子的弱相互作用本征态可以表示为质量本征态的线性组合。描述这种线性组合的系数矩阵称为 (PMNS) 矩阵。用方程表示为：

$$|\nu_\alpha\rangle = \sum_{j=1}^3 U_{\alpha j} |\nu_j\rangle \quad (\alpha = e, \mu, \tau) \quad (1.1)$$

其中 U 是 PMNS 混合矩阵元，由于存在三代轻子，U 中可以自然的包含 CP 破坏相位。粒子物理手册上提供了一种参数化 PMNS 混合矩阵的方式，称之为“标准参数化” [43]。

$$U = \begin{pmatrix} 1 & 0 & 0 \\ 0 & C_{23} & S_{23} \\ 0 & -S_{23} & C_{23} \end{pmatrix} \begin{pmatrix} C_{13} & 0 & \hat{S}_{13}^* \\ 0 & 1 & 0 \\ -\hat{S}_{13} & 0 & C_{13} \end{pmatrix} \begin{pmatrix} C_{12} & S_{12} & 0 \\ -S_{12} & C_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} e^{i\phi_1} & & \\ & e^{i\phi_2} & \\ & & 1 \end{pmatrix} \\ = \begin{pmatrix} C_{12}C_{13} & C_{13}S_{12} & \hat{S}_{13}^* \\ -S_{12}C_{23} - C_{12}\hat{S}_{13}S_{23} & C_{12}C_{23} - S_{12}\hat{S}_{13}S_{23} & C_{13}S_{23} \\ S_{12}S_{23} - C_{12}\hat{S}_{13}C_{23} & -C_{12}S_{23} - S_{12}\hat{S}_{13}C_{23} & C_{13}C_{23} \end{pmatrix} \begin{pmatrix} e^{i\phi_1} & & \\ & e^{i\phi_2} & \\ & & 1 \end{pmatrix} \quad (1.2)$$

其中， $C_{jk} = \cos \theta_{jk}$ ， $S_{jk} = \sin \theta_{jk}$ ， $\hat{S}_{13} = e^{i\delta} \sin \theta_{13}$ ， δ 一般被称为 Dirac 相位角， ϕ_1 和 ϕ_2 是 Majorana 相位角。“标准参数化”后的中微子混合矩阵可以对一些物理现象给出较为直观的解释。

- 方程1.2的最后一个矩阵是一个对角阵。这一项对中微子的振荡几率没有影响，中微子振荡实验中不能测量出 ϕ_1 和 ϕ_2 这两个参数，这两个参数的物理意义在于描述 Majorana 中微子（中微子与它相应的反粒子是同一种

粒子), 它们产生的效应可以在无中微子发射的双衰变 ($0\nu\beta\beta$) 中观测到。

- 可以用 Jarlskog 不变量表示 [44] 描述轻子 CP 破坏效应。从 U 矩阵的正性可知, 在三种味道中微子振荡模式中, 只有一个 Jarlskog 不变量。按照以上的参数化方法, Jarlskog 变量定义为:

$$\mathcal{J} = S_{12}C_{12}S_{23}C_{23}S_{13}C_{13}^2 \sin \delta, \quad (1.3)$$

这个关系表明, 中微子振荡实验中的 CP 破坏效应正比于 $\sin \delta$, 方程1.2中的是 CP 相角。通过中微子振荡实验, 可以确定 CP 破坏的大小。从上面的公式还可以看出: CP 相角和中微子混合角共同决定 CP 破坏的大小, 表明确定 δ 存在参数简并。

- 中微子振荡与六个参数有关: 三个混合角 θ_{12} 、 θ_{13} 和 θ_{23} , 两个独立的质量平方差 $\Delta m_{21}^2 \equiv m_2^2 - m_1^2$ 和 $\Delta m_{32}^2 \equiv m_3^2 - m_2^2$, 一个 CP 相位角 δ 。大亚湾实验之后, 对于这六个参数还有一个半没有确定, 一个是 δ , 还有就是 Δm_{32}^2 的正负号。如果 $\Delta m_{32}^2 > 0$, 质量等级称为“Normal Hierarchy(NH)”, 反之, 则称之为“Inverted Hierarchy(IH)”。全局拟合 (global fit) 的结果 [45] 如下所示:

$$\begin{aligned} \Delta m_{21}^2 &= 7.59 \pm 0.20 \left({}^{+0.61}_{-0.69} \right) \times 10^{-5} eV^2 \\ \Delta m_{32IH}^2 &= -2.36 \pm 0.11 (\pm 0.37) \times 10^{-3} eV^2 \\ \Delta m_{32NH}^2 &= 2.46 \pm 0.12 (\pm 0.37) \times 10^{-3} eV^2 \\ \theta_{12} &= 34.5 \pm 1.0 \left({}^{+3.2}_{-2.8} \right)^\circ \\ \theta_{23} &= 42.8 {}^{+4.7}_{-2.9} \left({}^{+10.7}_{-7.3} \right)^\circ \\ \theta_{13} &= 5.1 {}^{+3.0}_{-3.3} (\leq 12.0)^\circ \\ \sin^2 \theta_{13} &= 0.008 {}^{+0.012}_{-0.007} (\leq 0.043) \\ \delta_{CP} &\in [0, 360] \end{aligned}$$

根据这一结果, 我们可以知道:

- 质量平方差 Δm_{32}^2 的符号未确定, 中微子质量等级未知。

– 当前中微子振荡实验无法测得 CP 相角。

综合上面的结果，表明 CP 破坏程度的 Jarlskog 不变量 \mathcal{J} 是正比于 $\sin \theta_{13}$ 的，根据大亚湾反应堆中微子实验关于 θ_{13} 的测量结果，我们可以预见，在不久的将来，CP 相角非常可能通过中微子振荡实验得到。

1.3 中微子实验

考虑从源处发出一束动量为 P 的 ν_α ，其中 $\alpha = e, \mu, \tau$ ，在真空中传播一段距离 L 后到达探测器，实验上观测到它们转化成 ν_β 的概率为：

$$P(\nu_\alpha \rightarrow \nu_\beta) = |\langle \nu_\beta | \nu_\alpha(L) \rangle|^2 = \left| \sum_{i=1}^3 U_{\alpha i}^* U_{\beta i} e^{-iE_i L} \right|^2 \quad (1.4)$$

其中 $E_i = \sqrt{P^2 + m_i^2}$ ，假设中微子的动量远大于质量，所以中微子可以看作是极端相对论粒子，即速度近似等于光速，则可以用传播距离代替时间，且有： $E_i = \sqrt{P^2 + m_i^2} \approx P + \frac{m_i^2}{2P}$ ，因此公式1.4 还可以写为：

$$\begin{aligned} P(\nu_\alpha \rightarrow \nu_\beta) = & \delta_{\alpha\beta} \\ & -4 \sum_{i>j} \text{Re}(U_{\alpha i}^* U_{\beta i} U_{\alpha j} U_{\beta j}^*) \sin^2 \left[1.27 \Delta m_{ij}^2 \left(\frac{L}{E} \right) \right] \\ & +2 \sum_{i>j} \text{Im}(U_{\alpha i}^* U_{\beta i} U_{\alpha j} U_{\beta j}^*) \sin \left[2.54 \Delta m_{ij}^2 \left(\frac{L}{E} \right) \right] \end{aligned} \quad (1.5)$$

式中， $\Delta m_{ij}^2 \equiv m_i^2 - m_j^2$ ，单位是 eV^2 ，L 的单位是 km，E 的单位是 GeV，常数 1.27 来自于单位制的换算。从公式1.4或者1.5可以看出，Majorana 相位自动消去而不会影响中微子振荡。因此中微子振荡实验不能区分中微子是 Dirac 粒子还是 Majorana 粒子。通过轻子数破坏过程，如无中微子的双 β 衰变，这一问题有可能得到确认。

当中微子振荡实验的 L/E 满足关系：

$$\frac{\Delta m^2 L}{E} = o(1) \quad (1.6)$$

时，三代中微子振荡近似变成两种中微子的振荡。在方程1.2所示的 PMNS 矩阵参数化条件下，公式1.5近似变成如下形式：

$$P(\nu_\alpha \rightarrow \nu_\beta) = \sin^2 2\theta \sin^2[1.27\Delta m^2 L/E] \quad (1.7)$$

当 $\beta \neq \alpha$ 时，和

$$P(\nu_\alpha \rightarrow \nu_\alpha) = 1 - \sin^2 2\theta \sin^2[1.27\Delta m^2 L/E] \quad (1.8)$$

方程1.7描述中微子的产生 (appearance) 几率，方程1.8描述中微子的消失 (disappearance) 几率，它们是实验中经常用到的两个公式。

根据不同振荡参数的测量方法， Δm_{21}^2 和 θ_{12} 被称为太阳中微子振荡参数，或者记为 Δm_{sol}^2 和 θ_{sol} ； Δm_{32}^2 和 θ_{23} 被称为大气中微子振荡参数，有时记为 Δm_{atm}^2 和 θ_{atm} 。

下面我们按照历史顺序，介绍与中微子振荡现象相关的几个著名实验，以便对中微子振荡机制有进一步深入的理解。

1.3.1 太阳中微子实验

1968 年，Homestake 实验发现测量到的太阳中微子 ν_e 约为标准太阳模型预言值的三分之一。此后，一些辐射化学实验和水切伦科夫实验也成功地发现了太阳中微子丢失现象。2002 年，SNO 实验用 1000 吨重水探测 ^8B 衰变产生的太阳中微子。测量了中微子的带电流过程 (CC: $\nu_e + d \rightarrow e^- + p + p$)、中性流过程 (NC: $\nu_x + d \rightarrow \nu_x + p + n$)、弹性散射过程 (ES: $\nu_x + e^- \rightarrow \nu_x + e^-$)。其中带电流过程仅对 ν_e 中微子敏感，而中性流过程对所有的活性中微子探测是平等的，弹性散射过程可以探测三种活性中微子，但是对 ν_μ, ν_τ 不是很敏感。2005 年 SNO 公布的新结果为 [46]:

$$\begin{aligned} \phi_{CC}^{SNO} &= (1.68 \pm 0.06_{-0.09}^{+0.08}) \times 10^6 \text{cm}^{-2}\text{s}^{-1}, \\ \phi_{ES}^{SNO} &= (2.35 \pm 0.22 \pm 0.15) \times 10^6 \text{cm}^{-2}\text{s}^{-1}, \\ \phi_{NC}^{SNO} &= (4.94 \pm 0.21_{-0.34}^{+0.38}) \times 10^6 \text{cm}^{-2}\text{s}^{-1}, \end{aligned} \quad (1.9)$$

其中第一项是统计误差，第二项是系统误差。上式的第三个式子是三种味

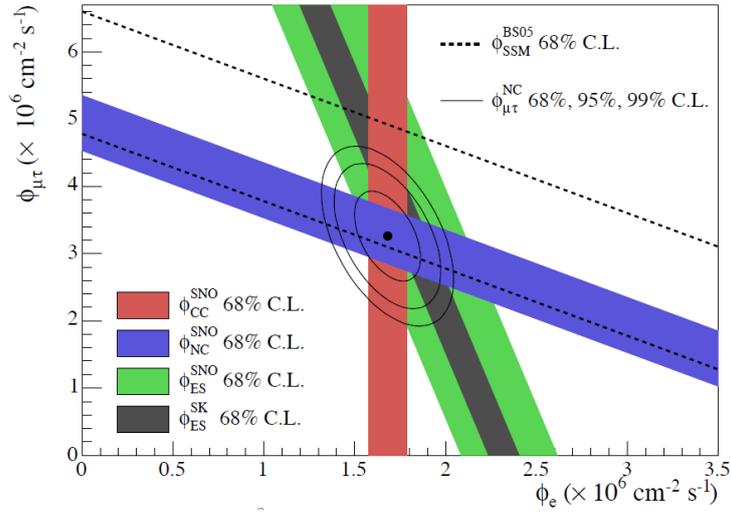


图 1.3 从 SNO 实验三种相互作用过程中得到的中微子通量，带状结构代表 1σ ，等高线代表 68%，95%，99% 置信区间的通量分布，虚线是标准模型预言的结果。

道的中微子总通量。它和标准模型预言的太阳 ^8B 中微子的通量一致。图1.3是电子中微子 (ν_e) 通量 ϕ_e 和非电子中微子 $\phi_{\mu\tau} = \phi_\mu + \phi_\tau$ 通量的二维分布，其中也列出了太阳模型的预言和 Super-K 实验弹性散射的结果。三种过程分别用不同颜色的带表示，它们相交于同一点，说明实验结果是自洽的，而且中性流的结果与标准太阳模型的预言相符。非零的 $\phi_{\mu\tau}$ 是中微子味道转换的强有力证据。中微子味道转换使得中微子总通量在不同味道之间重新分配。

在 SNO 实验后的一段时间里，除了中微子振荡，其他几种能够解释中微子味道转换的机制也被考虑。包括受太阳内部物质效应影响的不同振荡形式、和一些非标准的方案（例如中微子衰变 (Decay)[47] 等模型）。在 SNO 实验之前，通过整体拟合所有的太阳中微子实验数据发现质量平方差和混合角的参数空间有简并情况：“小角混合解” (SMA)，“大角混合解” (LMA)，“低质量平方差解” (LOW)，“真空振荡解” (VO)。SNO 的实验结果在 3σ 可信度下排除了 SMA，但是并没有排除其它情况。如果 LMA 和 CPT 守恒，根据中微子和反中微子的性质，日本的 KamLAND 实验应该看到反中微子有一部分消失，如果与物质效应有关的其他物理机制，那 KamLAND 实验不应该看到中微子减少。KamLAND 在 99.95% 的置信区间内给出了反电子中微子丢失。随着统计量的增加，KamLAND 不仅观察到反电子中微子能谱的变化，同时第一次给出了反

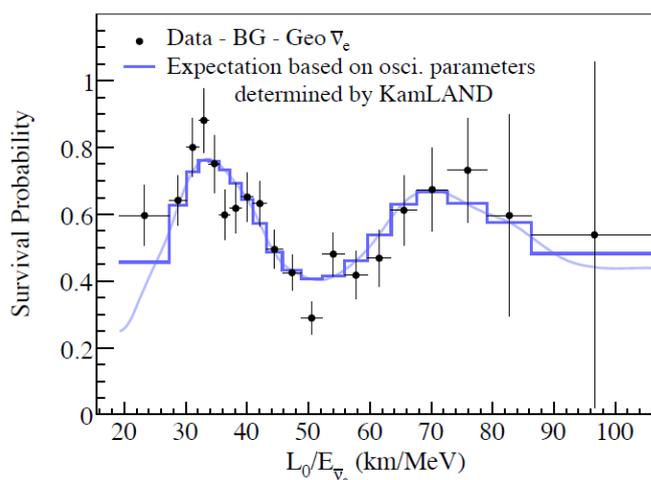


图 1.4 KamLAND 实验给中微子存活几率随 L_0/E 的变化，其中 L_0 为 180km。实线是中微子振荡的最佳拟合曲线。

中微子的存活几率随 L/E 的变化 [48]，如图1.4。自此，中微子振荡的 LMA 是太阳中微子味道转换的最合理机制。

位于意大利 Gran Sasso 的 Borexino 实验组 [49] 从 2007 年以来一直致力于太阳 ${}^7\text{Be}$ 中微子的探测。Borexino 利用 300 吨极纯的液体闪烁体通过 ν_e 和电子的弹性散射过程，探测单能的 0.862MeV 的中微子，探测低能的中微子不仅是对标准太阳模型的检验，更可进一步了低能中微子振荡在太阳中的共振放大物质效应 [25]。

1.3.2 大气中微子实验

上个世纪七八十年代，利用水切伦科夫探测器寻找质子衰变的很多实验，如日本的 Kamiokande 和美国的 IMB 实验，，此时大气中微子成了他们实验的主要本底，并因此发现了大气中微子反常。后来 Super-Kamiokande 采用了五万吨纯水，测量中微子与水反应后产生的带电轻子，带电轻子在水中产生的切伦科夫光被水四周的光电倍增管探测到，同时可以通过切伦科夫光环的清晰程度判断是 ν_e 还是 ν_μ 。图1.5给出了 Super-Kamiokande 观测到的 sub-GeV ($<1.33\text{GeV}$) 和 multi-GeV ($>1.33\text{GeV}$) 的电子型和 μ 子型事例随天顶角的变化。 $\cos\theta = 1$ 代表向下穿， $\cos\theta = -1$ 代表向上穿。

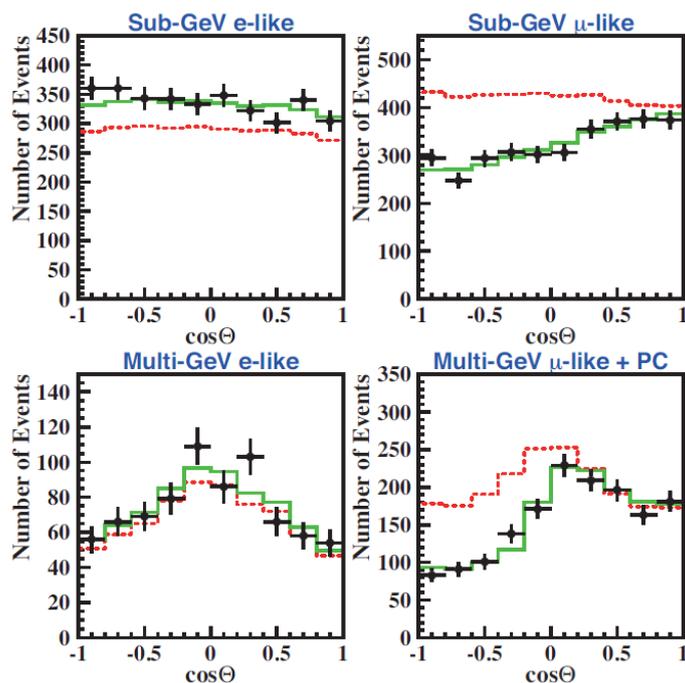


图 1.5 Super-Kamiokande 观测到的 sub-GeV ($<1.33\text{GeV}$) 和 multi-GeV ($>1.33\text{GeV}$) 的电子型和 μ 子型事例随天顶角的变化。虚线代表没有振荡时的 MC 模拟事例，实线代表中微子振荡的最佳拟合曲线。

如果中微子不发生味转换，则向上和向下穿过探测器的中微子总数是相等的，在 $\cos\theta = 0$ 两侧观测到的事例数应该是对称的。然而从图 1.5 可以看出，向上和向下穿过的电子中微子数基本相等，而且是关于 $\cos\theta = 0$ 对称的，说明 ν_e 基本没发生变化。但是 ν_μ 在穿过地球的过程中数量减少了。说明在经过地球的这段距离中发生了 $\nu_\mu(\bar{\nu}_\mu) \rightarrow \nu_\tau(\bar{\nu}_\tau)$ 的振荡。Super-Kamiokande 实验组分析了中微子事例数对 L/E 的依赖关系如图 1.6，实验数据显示的低谷只能由中微子振荡来解释，这给出了大气中微子味道转换的背后物理一定是中微子振荡。

日本的 K2K 实验 [29] 用 KEK 实验室的 12GeV 质子打靶产生 μ 中微子，并利用 120km 远处的 Super-K 探测器进行探测，进一步证实了 μ 中微子的振荡现象。美国的 MINOS 实验利用费米实验室 120GeV 质子产生的 μ 中微子进一步研究了加速器中微子振荡现象，探测器位于 735km 远的一个矿井中，但是最终测量结果与 K2K 的结果基本一致，得到的大气中微子振荡参数的精度更高。

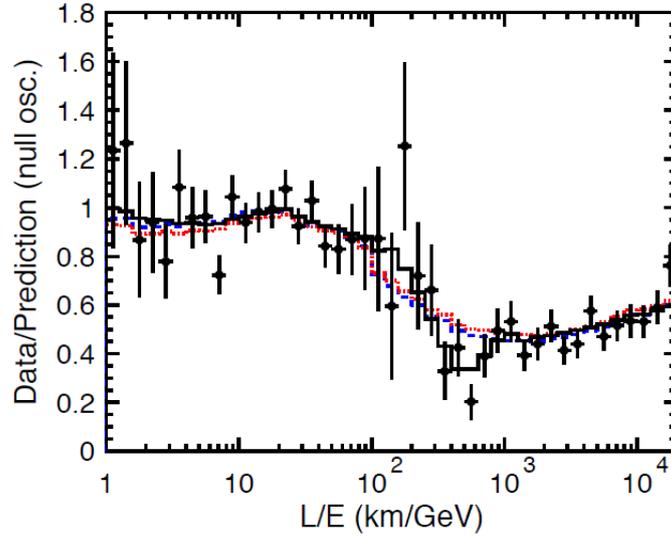


图 1.6 Super-Kamiokande 实验中 μ 中微子事例数对 L/E 的变化关系。点代表实验结果与蒙卡模拟未振荡的事例数比值，实线对应 $\nu_\mu \rightarrow \nu_\tau$ 振荡的最佳拟合值，虚线和点线分别表示中微子衰变和退相干模型的最佳拟合值。

1.3.3 反应堆中微子实验

θ_{13} 的测量可以通过加速器实验和反应堆实验得到。但是在加速器实验中，振荡概率 $P(\nu_\mu \rightarrow \nu_e)$ 依赖于 $\sin^2(2\theta_{13})\sin^2(\theta_{23})\delta_{CP}$ ，质量等级及物质效应。而短基线的反应堆实验可以给出干净的 θ_{13} 值。

通过观测反应堆产生的反电子中微子的存活几率可以得到 θ_{13} 的信息。基线一般在 1km 到 2km 之间，这时候有 $1.27\Delta m_{21}^2 L/E \sim 0$ ，利用近似关系 $\Delta m_{32}^2 \sim \Delta m_{31}^2$ ，得到公式 1.10:

$$P(\bar{\nu}_e \rightarrow \bar{\nu}_e) \simeq 1 - \sin^2(2\theta_{13})\sin^2(1.27\Delta m_{32}^2 L/E) \quad (1.10)$$

图 1.7 显示了几十年来反应堆中微子实验探测结果，其中 KamLAND 实验基线较长，测量出的是 θ_{12} 引起的振荡，其他实验由于探测器质量不够大（最重的 Palo Verda 实验也才 10 吨，或者探测器系统误差太大，或者基线不够长 ($< 1\text{km}$)，都没有测量到 θ_{13} 引起的振荡。

CHOOZ 实验 [30] 给出了最好的结果为 $\sin^2(2\theta_{13}) < 0.17$ ， $\Delta m_{12}^2 < 8 \times$

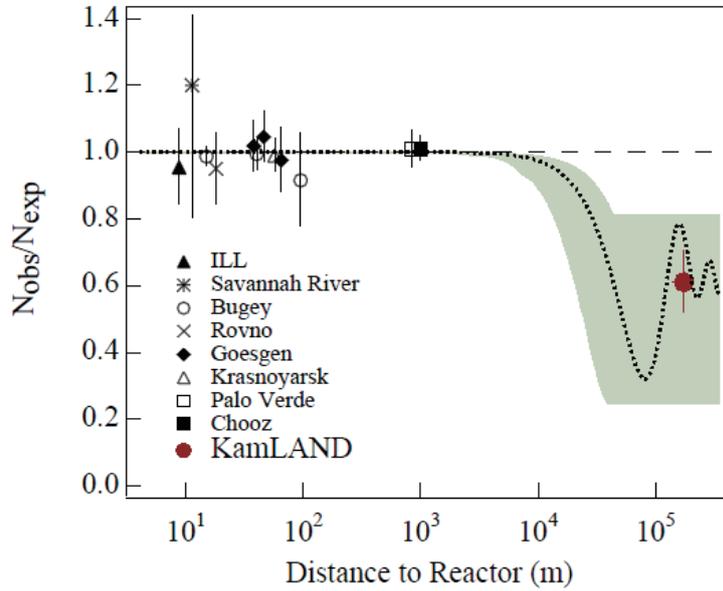


图 1.7 历史上的反应堆中微子实验

$10^{-4}eV^2$ 。

通过蒙特卡洛模拟研究，大亚湾实验要使灵敏度达到 0.01，实验的误差要达到表1.2中的要求。表1.2中包含了三类最基本的系统误差，根据各个探测器之

误差要求	近点	远点
统计误差	0.07%	$\leq 0.24\%$
探测器相关误差	$\leq 0.38\%$ /模块	
反应堆相关误差	$\leq 0.2\%$	
本底相关误差	0.3%	0.2%

表 1.2 大亚湾实验误差的要求 [50]

间或反应堆之间的误差是非独立的还是独立的，系统误差可以分为关联误差和非关联误差。

根据大亚湾实验的本底误差，可以对信噪比提出要求。表1.3列出了各种本底与信号事例率的比值。

在本论文最后一章的快中子研究中，将会估算真实数据中快中子本底的信噪比。

	大亚湾近点	岭澳近点	远点
中微子信号 (/天/模块)	930	760	90
天然放射性/信号	<0.2%	<0.2%	<0.1%
快中子/信号	0.1%	0.1%	0.1%
$^8\text{He}/^9\text{Li}$ 本底/信号	0.3%	0.2%	0.2%

表 1.3 中微子信号和本底事例率的比值 [50]

1.3.4 中微子物理亟待解决的问题

自从人类发现中微子以来，七十多年间，粒子物理学界就对中微子不断进行深入的研究。一系列的太阳、大气、反应堆、以及加速器中微子实验都表明中微子存在非零的静止质量，不同味的中微子之间存在着混合，这已经成为了有坚实实验证据支持的超出标准模型的新物理，中微子质量之谜极有可能是发现更基本理论的突破口。从唯象的角度研究中微子质量与味混合可以帮助我们揭开其背后物理的神秘面纱，从而最终使人们对自然界的认识达到一个新的深度。

中微子物理是人类通往新物理之路的指路灯，但是同时也存在着许多难题亟待解决。这些难题的解答无一例外地需要依靠未来的中微子实验。目前中微子物理领域内面临的问题主要包括：

- 中微子的代数，是否存在惰性中微子？
- 中微子是 Majorana 粒子还是 Dirac 粒子？
- 质量平方差 Δm_{32}^2 的符号，即中微子质量等级。
- 宇宙中物质与反物质的不对称之谜能否通过轻子的 CP 破坏效应给出解释？
- 中微子和反中微子的性质是否破坏 CPT 守恒？
- 中微子的绝对质量和中微子绝对质量的起源。

1.4 大型粒子物理实验数据处理软件系统

近年来，粒子物理领域的大型科学项目越来越依赖于计算机技术。反应堆中微子实验的高事例率、海量数据处理和复杂仪器设备管理对计算机集群的可靠性、高效性都提出了更高的要求。粒子物理实验数据处理可分为在线数据获取和离线数据处理两大部分 [51]。

在线数据获取是指通过计算机系统配合实验探测器运行，对实验数据进行实时的记录和转换，也包括对探测器运行状况和获取的数据质量进行实时监测。具体来说，它的任务主要包括：

- 对整个实验探测器系统进行运行控制。根据反应堆运行状况和当前取数系统设置，并按照各子探测器的管理要求，由计算机系统按一定的程序逻辑实现对整个探测器系统运行的控制。
- 收集实验数据。物理事例在探测器中产生信号，经过前端电子学和触发系统的信号转换和收集，再由在线数据获取系统 (DAQ) 将电子学信息 (时间信息，电荷信息) 以既定的格式组装成事例数据，最后把数据存入磁盘或磁带等外部存储设备中。
- 监测探测器状态。各探测器系统的工作参数由计算机系统预先设置并实时监测，如果出现异常情况，计算机系统会将异常状态显示在监视器上，并发出错误或警告信息，出现严重错误时，要按预先规定的程序执行连锁动作，以避免事故发生。
- 数据质量在线监测。在探测器运行期间，会对在线系统获取的数据进行实时分析，观察当前获取的数据的质量是否正常，以便发现问题并及时解决。此外，还会依照物理数据分析需求对特殊物理事例进行在线标记。

离线数据处理指原始数据存入外部存储设备之后，物理学家根据相应的物理目的，在计算机终端上对数据进行各种数据处理和分析，在深入了解探测器中发生的物理过程之后，最终得到有明确物理意义的结果。这一阶段的主要工作包括：

- 探测器参数刻度。灵敏探测器及其电子学和触发系统在工作期间受外界环境变化等因素的影响，各项物理参数会发生变化，进而影响到性能指标，结果是对现代高灵敏度探测器的测量精度产生影响。因此需要根据探测器运行的具体情况和环境，经常对探测器系统以软件的方法进行刻度，以减小这种系统误差。刻度参数存放在数据库中，以供重建算法等使用。
- 能量重建。原始数据中保存的是粒子和探测器靶物质作用产生的击中信息，能量重建就是利用一系列击中信息中的的时间、电荷和击中分布等信息，还原出当前物理事例的能量。能量重建允许我们根据重建得到的能量进行一定程度的粒子鉴别。
- 蒙特卡罗模拟。依照精确设计的模拟方法，能够比较详尽地反映各种随机物理过程的统计性质。离线数据处理过程中，蒙特卡罗模拟通常用于确定探测器对信号和本底的探测效率、分布特性等。
- 物理数据分析处理。在径迹重建和蒙特卡罗模拟的基础上，通过对事例能量和时间关联特性的分析，能够挑选出一定研究目标下让我们感兴趣的事例。然后进一步使用各种算法对某些物理参量进行深入处理，最终根据特定物理参量的数值或分布特征得出有物理意义的结论。

在线数据获取和离线数据处理过程的顺利完成，离不开可靠完善的数据处理软件系统的支持。现代粒子物理实验往往是个结构复杂，规模庞大的大科学项目，相应配套的软件系统的规模也越来越大，为了高效地对海量数据进行处理，很多现代软件技术被广泛使用在粒子物理实验中。例如，传统的面向过程的编程语言正被面向对象的语言替代，面向接口的设计理念渐渐融入到粒子物理实验软件的设计层面上来，并成为主流；软件工程的理论和方法被广泛接受；为了适应多人合作的工作模式，不少用于软件版本控制或管理的工具，如 CMT[52] 和 SVN[53] 以及相应的规范也被采纳。

1.5 论文选题的物理及软件背景和论文结构

1.5.1 物理及软件背景

描述中微子味道混合与振荡的 MNSP 矩阵参数化理论虽然可以定量地解释中微子在传播过程中发生的味道转换，然而实验上却由于种种原因始终未能完全确定 MNSP 矩阵中的六个参数，尤其是至关重要的混合振荡参数角 θ_{13} 。因此，通过反应堆中微子实验精确测量中微子混合振荡参数角 θ_{13} ，对于完善中微子振荡理论具有重大意义，同时 θ_{13} 的大小也决定着未来中微子实验的设计和实现，如果 θ_{13} 的值不是很小，就能够利用反应堆实验测得 MNSP 矩阵中的 CP 相角，有望解开宇宙中“物质与反物质”不对称之谜，这也是当前粒子物理和宇宙学交叉领域内的前沿热点之一。

大亚湾反应堆中微子实验是国内首个反应堆物理实验，它的显著特点就是事例率高（包括信号事例，天然放射性本底事例和来自宇宙线 μ 子的事例），数据量巨大，对离线数据分析软件系统提出了更高的要求。为了能够可靠地对这些海量数据进行高效处理，离线软件系统采用了新的技术设计和实现，移植了欧洲核子中心（CERN）为对撞物理离线数据处理而开发的通用软件框架 Gaudi[54]，并根据反应堆物理事例的自身特点，对框架结构进行改进和重构，以适应对反应堆物理数据进行特殊处理的需要。

1.5.2 论文结构

第一章简单介绍了中微子物理与中微子相关的各类实验取得的巨大进步，包括中微子静止质量不为零、中微子存在味道—质量本征态的混合、中微子振荡现象的提出和发现。最后简要介绍了中微子物理领域内仍然存在的需要解决的问题以及现代大型粒子物理实验中的软件系统的现状和要求。

第二章是对大亚湾反应堆中微子实验的概述。主要包括大亚湾反应堆中微子实验的科学目标和整体方案布局、中心探测器和反符合探测器的设计、大亚湾实验中的主要本底和误差、以及实验中离线数据分析流程及其软件框架 NUWA。

第三章讲述离线软件原始数据的访问和转换服务，详细介绍了原始数据转

换服务基本功能的实现，对它的扩展和优化，以及原始数据转换服务模块在探测器性能调试过程中的具体应用。

第四章介绍了大亚湾实验 MC 模拟数据混合的物理背景和算法实现。重点讲述反应堆实验真实物理数据的特点，模拟数据混合的物理动机、实现模拟数据混合的两种算法：基于抽样的算法和基于“有限状态自动机”的算法、以及两种物理混合方案的实现：探测器电荷击中级混合和电子学读出级的混合。为适应数据混合，对离线软件框架 NUWA 中的标准数据 I/O 接口的改进以及对电子学模拟架构的重构，也做了相应介绍。

第五章讲述了大亚湾中微子实验中的快中子本底的挑选和分析，详细介绍了利用水切伦科夫探测器进行标记，挑选真实物理数据由宇宙线 μ 子产生的快中子事例，给出了快中子事例的快、慢信号的能谱和时间间隔分布、利用线性外推法估算了快中子本底在反 β 衰变事例中的比例及误差、最后利用真实数据，计算了大亚湾实验三个实验大厅中，“ μ 子致散裂”的中子的产额与 μ 子能量的关系，并与之前其他地下实验测得的中子产额进行了对比。

第六章总结与展望，对大亚湾正常取数以来得到的物理结果的总结和对以后工作的展望。

第二章 Daya Bay 反应堆中微子实验

2000 年以来，国际上陆续出现了多个利用反应堆对中微子振荡混合参数进行精确测量的实验。近年来，利用短基线反应堆中微子实验测量中微子混合参数角 $\sin^2 2\theta_{13}$ 成为中微子研究中的前沿热门课题，目前国际上正在运行的可精确测量这一参数的有三个实验组，分别是中国大亚湾，法国 Double Chooz 和韩国的 RENO[55]，如表2.1所示。由于具有较大的反应堆功率，较厚的山体覆盖和较大的探测器靶质量，大亚湾中微子实验是三个实验中灵敏度最高的。大亚湾反应堆中微子实验是一个国际合作实验组，包括中国，美国，俄罗斯和捷克等地 30 多所高等院校和科研机构，200 多名科研人员。2012 年上半年，大亚湾合作组正式发布探测器运行取数以来利用三个月的数据得到的物理结果，以 5 倍 σ 的置信概率确认混合振荡参数角 θ_{13} 不为零 [56]。

Experiments	热功率 (GW)	距离 近/远 (m)	山体高度 近/远 (mwe)	靶质量 (tons)	sensitivity
Double Chooz	8.5	400/1050	115/300	10/10	0.03
RENO	16.4	290/1380	120/450	16/16	0.02
Daya Bay	11.6(17.4)	363(500) /1985(1615)	260/910	40×2/80	0.01

表 2.1 国际上正在运行测量中微子混合振荡参数 θ_{13} 的实验比较。

2.1 大亚湾实验简介

“标准参数化”得到的 MNSP 矩阵用六个参数描述中微子传播过程中的振荡现象，三个振荡混合角 (θ_{12} , θ_{23} , θ_{13}) 和两个独立的质量平方差 ($\Delta m_{21}^2 \equiv m_2^2 - m_1^2$ 和 $\Delta m_{32}^2 \equiv m_3^2 - m_2^2$)。来自反应堆芯裂变的 $\bar{\nu}_e$ 经过 L 的传播距离之后，其存活几率是：

$$P_{sur} \approx 1 - \sin^2 \theta_{13} \sin^2(1.267 \Delta m_{31}^2 L/E) \quad (2.1)$$

其中, Δm_{31}^2 的单位是 eV^2 , E 是中微子能量, 单位是 MeV 。 L 的单位是 km , 大亚湾反应堆中微子实验被设计在 90% 的置信度以 $\sin^2 2\theta_{13} < 0.01$ 的灵敏度测量混合角 θ_{13} 。 实验用来探测中微子的探测器是 8 个全同的反中微子探测器 (Anti-neutrino Detector), 分布在大亚湾, 岭澳反应堆群周围 2000m 范围内。

大亚湾探测器利用 $\bar{\nu}_e$ 在掺钆液体闪烁体 (Gd-LS) 中发生的 inverse β 衰变反应 (IBD) 探测中微子, $\bar{\nu}_e + p \rightarrow e^+ + n$ 。 inverse β 衰变产生的 e^+ 和 e^- 湮灭和中子在钆核上俘获的的快慢信号符合, 可以用来确定中微子信号。 预期探测的中微子数目可以由下式确定 [57]:

$$N_{det} = \frac{N_p}{4\pi L^2} \int \epsilon \sigma P_{sur} S dE \quad (2.2)$$

其中,

- N_p 是探测器中自由质子的数目
- ϵ 是中微子探测的效率
- σ 是 inverse β 衰变的的积分反应截面
- S 是中微子的微分能谱

在反应堆中微子实验中, 若仅用一个位置固定的探测器来测量 $\sin^2(2\theta_{13})$, 为了得到物理结果, 来自反应堆的中微子的绝对流量, 中微子在探测器靶物质中发生 inverse β 衰变的绝对反应截面, 及探测器的绝对探测效率和各种挑选条件的 cut 效率都必须得到充分且精确的了解。 这种情况下, 想要把最终结果的误差降低到到 1% 以下量级, 是非常困难的, 尤其是对来自反应堆的误差的估算。

Mikaelyan 和 Sinev 指出, 若能够将两个探测器分别放在优化之后不同位置上, 实验结果的系统误差将会得到极大地降低 [58]。 离反应堆比较近的探测器用来测量中微子的流量和能谱, 这就减少了对反应堆内部过程细节的了解。 基于这种实验设计, 我们可以通过比较远点和近点的中微子的流量和能谱差异来得到最终的 $\sin^2(2\theta_{13})$ 。

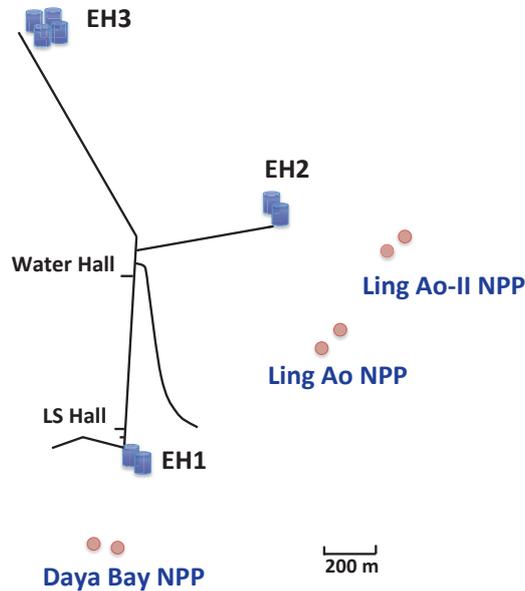


图 2.1 大亚湾实验布局。反应堆堆芯用红色圆点表示，8 个探测器被安装在 3 个实验大厅 (EH1 ~ EH3) 里面。

2.1.1 大亚湾中微子实验站布局

大亚湾核反应堆堆群坐落在中国广东省深圳市南海岸。反应堆堆群包括三对反应堆，每一对被称为一个 NPP(nuclear power plant)，大亚湾堆芯 (Daya Bay NPP)，岭澳堆芯 (Ling Ao NPP) 和岭澳二期堆芯 (Ling Ao-II NPP)。如图2.1所示。反应堆堆群东南面面海，西北毗邻高山。

每个 NPP 都包括一对反应堆，这六个反应堆在设计上是全同的，堆芯功率都是 2.9 GW，最后一个堆芯在 2011 年 8 月份投入商业运行。大亚湾中微子实验站的设施包括地面设备，三个地下实验大厅，以及两个地下功能大厅：液闪大厅 (LS Hall) 和水净化大厅 (Water Hall)。地面设施包括一个用来装配中心探测器的地面装配大厅 (SAB)，一个控制中心以及一个通风设备。三个地下实验大厅通过水平的隧道相连接。

按照实验规划，1 号大厅 (EH1) 和 2 号大厅 (EH2) 各安装两个反中微子探测器，在 3 号大厅 (EH3) 安装四个反中微子探测器。3 号大厅被建设在中微子振荡极大的基线位置。三个实验厅的岩石覆盖、宇宙线 μ 子计数率以及到各个堆芯的平均距离如表2.2 所示。

	岩石覆盖	R_μ	E_μ	大亚湾	岭澳	岭澳二期
EH1	281	1.21	55	364	857	1307
EH2	299	0.90	58	1348	480	528
EH3	884	0.05	140	1912	1540	1548

表 2.2 三个实验大厅的岩石覆盖 (m.w.e), μ 子计数率 R_μ (Hz/m²), 平均的 μ 子能量 E_μ (GeV), 及它们到大亚湾堆芯, 岭澳堆芯以及岭澳二期堆芯的平均距离 (m)。

2.1.2 大亚湾中微子探测器系统简介

反应堆中的燃料核素经过裂变反应, 产生次级不稳定子核, 这些子核通过 β 衰变产生的中微子。这些中微子几乎全部来源于反应堆中的 ^{235}U 、 ^{238}U 、 ^{239}Pu 和 ^{241}Pu 四种核素的裂变 [59][60][61][62], 再由探测器探测得到。为了屏蔽周围环境的天然放射性并排除宇宙线 μ 子产生的其他本底, 大亚湾中微子实验的探测系统进行了一系列防护设计, 分为反中微子中心探测器和由水切伦科夫探测器和 RPC 气体阻性板构成的宇宙线 μ 子反符合系统。

2.1.2.1 反中微子中心探测器

大亚湾中微子实验中用来探测 $\bar{\nu}_e$ 中微子的探测器称之为 AD (Anti-Neutrino Detector), 其结构大小如图2.2所示。

整个探测器的大致结构, 是从外到内的 3 层的同心圆柱容器。最外层是一个圆柱形的 5m 直径, 5m 高的不锈钢罐 (SSV), 里面安装有两个紫外透射的有机玻璃圆柱形罐子。

- 最内层是一个直径 3.1m 的有机玻璃罐 (称为 IAV), 内装有 20 吨钆含量为 0.1% 液体闪烁体, 是探测中微子的有效靶质量。掺钆液闪比普通液闪的好处是: I. 根据来自 KamLAND 和 CHOOZ 实验的经验, 掺钆液闪能够使中子的俘获时间常数由在普通液闪中的 $210\mu\text{s}$ 减少到在含钆液闪中的 $30\mu\text{s}$, 缩短了 inverse β 衰变得到的快慢符合信号之间的时间间隔, 就能够有效减少偶然符合带来的本底。II. 中子在钆的两种同位素 ^{155}Gd 和 ^{157}Gd 的俘获截面也远远大于在 H 上的俘获截面, 能够有效提高中子的俘获效率。III. 中子在钆核上被俘获, 将会放出 3—4 个 γ 光子, 总能量在 8MeV

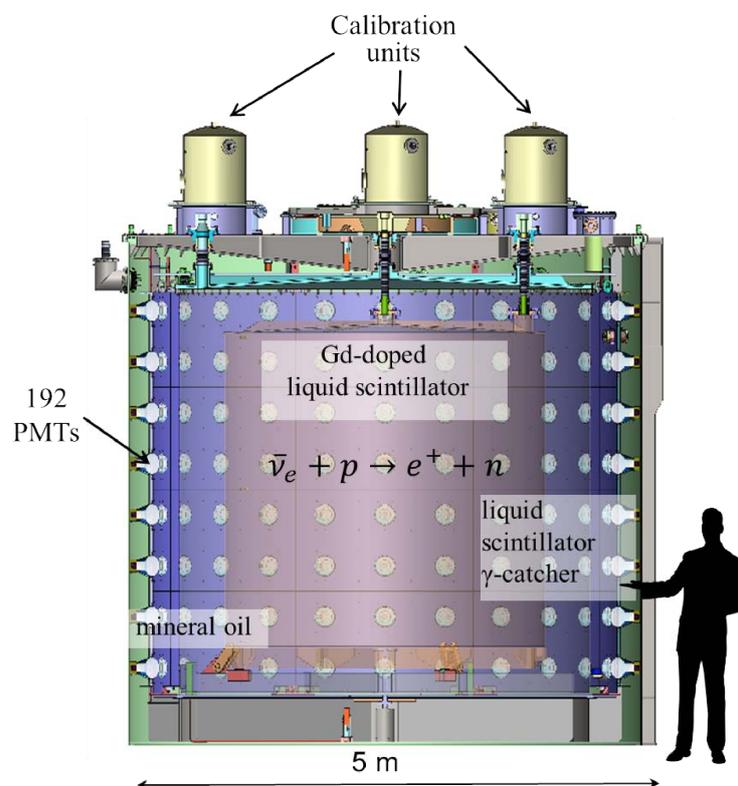


图 2.2 大亚湾反中微子探测器结构示意图。

左右，远高于诸如天然放射性本底的能量，这样可以大大减少本底的事例率，提高信噪比。

- 中间层是一个直径 4m 的外有机玻璃罐（OAV），装有 20 吨的普通液体闪烁体，主要是用于辅助收集掺钆液闪内中子俘获发出 γ 光子的能量。
- 4m 罐和钢罐中间装有 37 吨的矿物油作为缓冲层，用以屏蔽来自光电倍增管，钢罐及其支撑结构和其它来自探测器外部的本底。
- 在矿物油最外层靠近钢罐的地方，以 8 圈 \times 24 列的形式，安装有 192 个面朝向液闪中心的光电倍增管，其型号均为 Hamamatsu R591。光阴极有效覆盖率为 12%，用来接收液体闪烁体发出的闪烁光。光电倍增管背面装有涂有黑色涂层的锥形罩，可以吸收光电倍增管背面的杂散光。
- 探测器上下安装了基于 ESR 的反射板，在节省光电倍增管（成本）的前

提下，增强了探测效率和探测器的均匀性。

- 此外，钢罐顶端盖子上安装有三个自动刻度装置 (ACU)，能够远程控制一个 LED，一个 ^{68}Ge 源，以及一个 $^{241}\text{Am-C}$ 和 ^{60}Co 的混合源，使它们沿着竖直方向 (z 轴) 移动进入探测器内部，每隔一段时间进行的刻度的能精确的监测 AD 能量响应。

2.1.2.2 宇宙线 μ 子反符合探测系统

尽管大亚湾实验大厅有足够厚度的山体覆盖，但由于实验高精度的要求，在实验探测器系统设计中仍然需要对进入探测器中的宇宙线 μ 子进行标记，这是由 μ 子反符合系统完成的。 μ 子反符合系统位于反中微子探测器的外围，根据大亚湾实验的设计要求，反中微子探测器浸泡在灌满高纯水的水池中，探测器周围任一方向上都保证有至少厚度 2.5 米的水层作为屏蔽。在大亚湾实验中，用来屏蔽外部放射性和标记宇宙线 μ 子的水切伦科夫探测器分为内水池 (IWS) 和外水池 (OWS) 两层。近点和远点实验大厅的水切伦科夫探测器分别安装了 288 个和 384 个 8 英寸光电倍增管，内外水池对宇宙线 μ 子的联合探测效率高于 99.5%，误差低于 0.25%。

在水切伦科夫探测器上方，覆盖有 4 层阻性板 (RPC) 用来进一步对宇宙线 μ 子进行探测，近点和远点分别使用了 54 和 81 个 RPC 模块。中微子探测器，水切伦科夫探测器和 RPC 探测器结构如图 2.3 所示。

2.1.2.3 电子学和触发系统

反中微子探测器和水切伦科夫探测器中的 PMT 与前端电子学板 (Front-End Electronics) 相连，前端电子学板上的每一个电子学道对应一个 PMT，以记录来自这个 PMT 的击中电荷以及时间信息，同时还记录信号之前，阈值之下的 100ns 时间窗内的平均的 ADC 值，作为这一电子学道的台阶并在后续处理的过程中加以减除。物理取数时，AD 的触发系统采用基于过阈道数的多重性触发 (NHit) 以及基于电荷求和的能量求和触发 (ESum)。具体到数值，前者定义为 PMT 的着火数大于 45 或者 PMT 电荷之和大于 $\sim 65\text{PE}$ (相应于能量约 $\sim 0.4\text{MeV}$)。这种触发条件使得每个 AD 的触发效率约为 280Hz，同时保

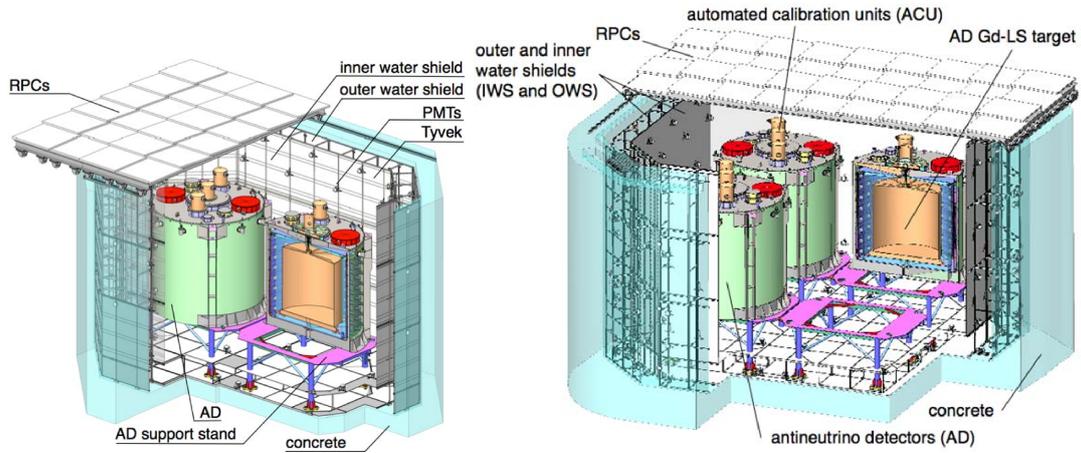


图 2.3 左图为大亚湾近点水切仑柯夫探测器，反中微子探测器和 RPC，右图为大亚湾远点水切仑柯夫探测器，反中微子探测器和 RPC。

证了 IBD 事例的丢失效率可以忽略不计。数据获取系统 (DAQ) 则选择基于 VME 以及 TimeSys 的系统，并根据 BESIII 和 ATLAS 的运行经验，结合大亚湾的实际情况，在其模型基础上进行了修改。

水切仑柯夫探测器的触发系统相对简单，内水池和外水池的 PMT 的着火数大于 6 的触发被归类为“水池 μ 子事例”，即 μ_{WS} 。考虑到 μ 子可能会同时穿过反中微子探测器，标记 AD 中沉积能量大于 20MeV(2500PE) 的信号被识别为“AD μ 子事例”，即 μ_{AD} 。实验中还发现 PMT 存在打火 (flash) 现象，即一个 PMT 发射出光子，被其他 PMT 接收到并触发整个探测系统，这一部分本底可以通过“象限电荷法”进行有效的排除。

2.2 大亚湾实验离线软件系统

Daya Bay 离线软件系统主要实现两部分功能：I. 真实物理数据和 MC 模拟数据的重建，刻度以及各种分析处理；II. 高能物理领域各种常用软件库的分类，管理和调用。离线软件系统包含了：主框架系统、外部软件工具包系统、MC 数据模拟，刻度和重建系统，物理分析轻量级框架系统构成。并借鉴了国际高能物理实验在过去几十年的运行中积累的软件设计经验和最新成果。尤其重要的是：大亚湾中微子实验作为国内首个非对撞物理实验，虽然借鉴了欧洲核子中心的对撞机物理软件框

架 Gaudi, 但是在使用过程中, 根据反应堆物理数据特点需要, 对框架最关键的模块 (数据流 IO 模块、数据缓存机制) 做了调整和改进; 在物理数据分析的过程中, 针对反应堆带有时间关联信息的物理事例的特点, 开发出了运行更可靠, 效率更高, 使用更加方便灵活的轻量级物理分析框架 LAF (Light weight Analysis Framework), 并以此框架为基础, 完成了基于大亚湾实验前三个月数据的最初的一系列物理分析, 也是对并为非对撞物理实验软件框架的设计的一次成功尝试, 适应未来非对撞物理实验的发展要求。

2.2.1 离线数据处理平台

大亚湾离线数据处理平台主要是为离线数据处理提供可靠运转的硬件设备和稳定高效的软件环境, 以保证数据处理过程中 CPU 的逻辑运算, 磁盘等外部存储设施的数据 I/O, 以及数据的网络传递准确性。

如前所述, 大亚湾实验共有三个实验大厅, 其中大亚湾近点在线总事例率为 $\sim 1.4kHz$, 岭澳近点事例率为 $\sim 0.8kHz$, 远点在线事例率为 $\sim 0.6kHz$ 。岭澳近点和远点实验大厅, 当前各有一个 AD 没有安装, 当前三个实验大厅每天物理数据量为 $\sim 300GBytes$, 三个月数据为大约 30TBytes, 预计三年总数据量会达到 360TBytes。这些数据将存储在大亚湾实验现场的磁盘阵列上, 并通过网络传输到高能所、BNL 等合作单位, 进一步进行后续数据处理, 数据传送的网络带宽要达到 $\sim 4000 Mbps$, 更重要的是: 该网络系统还需具备极好的独立性和稳定性。

基于以上离线数据处理对计算环境的要求, 大亚湾离线数据处理采用以下技术平台来保证工作的顺利进行:

- CPU 型号: 采用 Intel 或 AMD 多处理器主板的 PCCluster 或 Grid 高性能计算结构, 最多可允许多达 248 个并发进程
- 数据存储: 应用磁盘阵列和磁带库存储虚拟化技术, 基于 lustre 文件系统, 建立分级存储体系 HSM (Hierarchical Storage Management)
- 网络 I/O 服务软件: 采用 spade 软件从大亚湾现场发送数据, 接收端用 ingest 软件进行数据接收, 并建立 SAN 第二网, 将主要的数据通路和传

统网络分开；计算节点机均采用 100TX/1000TX 双网卡，其中 1000TX 仅提供 NFS 服务，以提高 NFS 的能力。

- 操作系统：操作系统基于 RedHat Distribution 的 Scientific Linux 64 位操作系统；中间刻度数据存储和管理选用 MySQL 数据库系统；使用 PBS 批处理系统进行作业管理。

2.2.2 大亚湾离线软件框架 (NUWA)

大亚湾离线数据分析系统的任务是将大亚湾各个探测模块中产生并由 DAQ 获取到的物理信号在离线计算终端上进行处理，包括根据信号的能量、快慢信号时间间隔特性等物理量的分布进行事例挑选、粒子鉴别等工作。大亚湾离线数据处理软件包括二进制原始数据的格式转换、MC 模拟数据的产生、真实物理数据的刻度和重建以及最后的物理分析。按照功能划分，大亚湾离线软件主要有三大部分：数据分析处理、数据存储和传输，软件模块管理。世界高能物理实验领域已经存在很多先进软件包和通用工具，为了充分利用这些先进资源并尽可能做到和国际接轨，大亚湾离线软件总框架选择 Gaudi 框架和面向对象编程技术，并对数据库进行了有限度地使用以优化对实验参数的管理和更新。大亚湾模拟数据产生基于 GEANT4 模拟软件，以便深入研究探测器性能，准确了解探测器中发生的物理过程，提高物理数据分析的精度和可靠性。海量物理数据使用国际通用的物理数据存储和分析工具 ROOT。

大型软件系统的高效构建依赖于软件框架，软件框架提供一套标准化的接口定义、模式化的模块间通信协议和统一化的应用软件模块实现规则以规范软件开发人员的代码行为，从而达到高效整合软件组不同开发人员的工作的目的。经过深入调研，反复测试，我们开发了基于 Gaudi Framework 的 NUWA 软件框架系统。NUWA 框架的主要体系结构如图2.4所示，它的主要组成部分包括算法模块 (Algorithm)，应用管理器 (Application Manager)，瞬态事例缓存 (Transient Event Store, TES)，服务模块 (Service) 以及转换器 (Converters) 等。

NUWA 框架的体系结构的核心是：以内存中瞬态事例缓存为基础，利用转换器和转换服务模块实现内存和外部存储设备之间数据持久化机制。这一套机

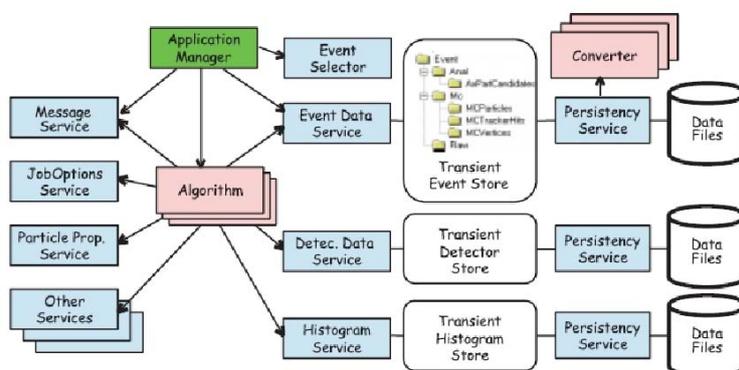


图 2.4 NUWA 框架体系结构 (参考 Gaudi 手册) 示意。

制使得数据和框架中处理数据的算法模块相对独立，对于各种算法和服务模块，框架中提供抽象的虚基类接口，模块之间的各种操作依赖面向对象程序设计中的多态特性，通过标准的抽象接口和模版编程实现。

与 Gaudi 框架不同，NUWA 框架中，在应用程序初始化阶段，用于配置各功能模块的脚本是通过 Python 语言来实现的。

2.2.2.1 大亚湾模拟和混合系统

大亚湾早期的探测器模拟由探测器模拟软件 G4dyb[63] 完成。G4dyb 是基于 Geant4[64] 开发的面向对象的模拟软件。利用 G4dyb 可以全面构建大亚湾探测器的几何结构和物质特性，并模拟粒子与探测器物质的相互作用，最终给出粒子在灵敏探测器中的击中信息，输出的数据以 root 格式存储并带有 MC truth 信息，可供后续用户分析使用。G4dyb 中除了提供高能物理领域常用的相互作用模型，包括电磁作用、强相互作用和可见光相互作用以外，还根据需要，加入了液体闪烁体的猝灭效应和对中子在液闪俘获过程的修正。G4dyb 的模拟结果得到大亚湾小模型探测器实验数据的验证，结果基本可靠。

2008 年，G4dyb 移植到了大亚湾软件框架 NUWA[65] 中，GiGa 软件包被引入作为 Geant4 和软件框架的接口。使用 DetDesc 软件包并采用基于 XML 的统一几何描述标识文件 (GDML)，来重新构建原 G4dyb 中的利用 Geant4 中提供的类构建的探测器几何，探测器模拟使用 DetSim 软件包来模块化 G4dyb 各部分的功能。

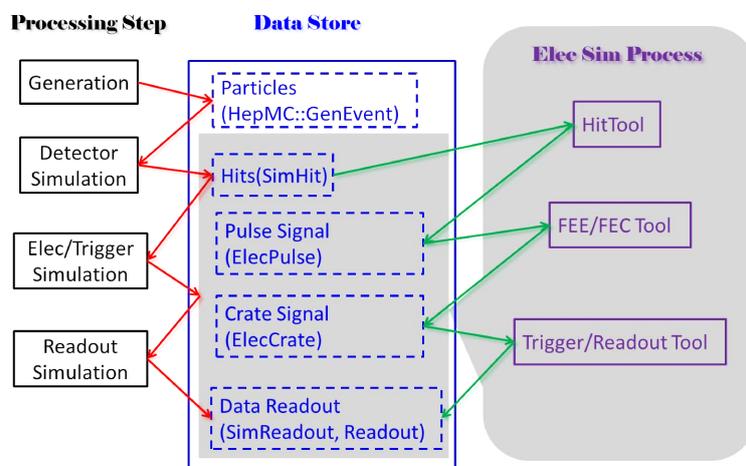


图 2.5 大亚湾模拟框架及数据存储示意。

在 NUWA 中，MC 模拟的整个过程完全按照真实物理世界中的探测器响应过程分阶段完成，分为探测器模拟 (DetSim) 和数字化模拟 (Digitization Simulation) 两部分。数字化模拟过程又分为电子学模拟 (ElecSim)，触发模拟 (TrigSim) 和电子学读出模拟 (ReadoutSim) 三个子部分。每个部分都使用 NUWA 中标准的 Algorithm 模块进行模块化封装，并可以在运行之前，由 python 脚本进行灵活地配置。为统一整个模拟阶段的数据接口，NUWA 框架中定义了标准的 DataModel 用来存储不同模拟阶段产生的模拟数据信息，并提供给下一个模拟阶段使用。整个模拟阶段及数据存储方式，由图 2.5所示。

大亚湾实验数据的显著特点是：不同的物理事例之间带有很强的时间关联信息，这些信息对最终的物理分析至关重要，与此同时，NUWA 框架中单纯的探测器模拟和数字化模拟主要完成对某一特定类型的物理事例在探测器响应的能量沉积的记录和转化，而模拟数据中不同物理事例之间的时间关联并不能直接通过模拟自动产生，需要在模拟完成之后通过被称之为“事例混合”的技术加以构建。不同物理样本的事例混合位于 MC 模拟之后，是产生可用于物理分析的 MC 数据的重要步骤。图2.6不同物理样本事例混合示意。

2.2.2.2 大亚湾离线刻度和重建

大亚湾实验的探测器结构比较简单，因此整个离线刻度和重建系统也比较简单，软件上仅仅是采用 NUWA 框架下标准的 Algorithm 模块和 Tool 工具模

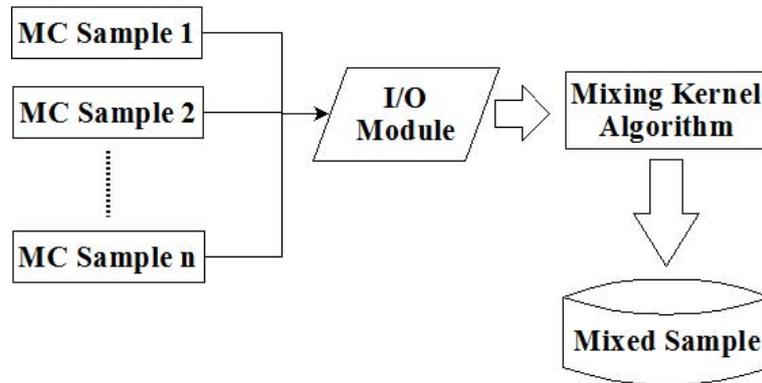


图 2.6 大亚湾物理样本事例混合示意。

块，就可以完成刻度和重建包的构建。

刻度的主要工作是实时得到光电倍增管的增益参数，即每个光电子对应的 ADC。在反中微子探测器的性能调试 DryRun 期间，由于没有灌装液体闪烁体，PMT 使用 LED 光源进行刻度，以仔细深入地研究探测器的能量响应。在探测器运行于物理取数期间，主要是使用放射源进行刻度，考虑到探测器靶体积内的位置不均匀性，这些放射源被设计成为可以通过自动刻度装置 ACU (Automatic Calibration Units) 沿着探测器对称轴 (Z 轴方向) 移动。此外，每个 AD 沿着在径向 (R 轴方向) 的不同位置，安装有三个 ACU，以便了解探测器径向上不同位置的能量响应。除了使用人工光源和放射源进行刻度之外，“宇宙线 μ 致散裂中子”也能够用来对探测器进行刻度，并持续地监测其稳定性。

大亚湾反中微子探测器重建的核心是能量重建，重建主要是基于一个物理事例的总电荷，也就是所有 192 个光电倍增管的电荷之和。AD 的能量刻度常数 $\sim 170\text{PE}/\text{MeV}$ ，是通过放置在 AD 几何中心位置的 ^{60}Co 放射源的 2.506MeV 的峰值能量确定的。顶点重建使用“电荷中心法” (center-of-charge)，即所有 PMT 坐标依照其接收到的电荷的进行加权平均得到的坐标位置。单一的电荷中心法比较粗糙，在此基础上，通过分析 $^{60}\text{Co}(r, z)$ 的数据，修正了电荷中心到顶点的映射。对于 inverse β 衰变的信号事例，主要是通过其快慢信号能量和时间间隔特性进行事例挑选，不依赖顶点重建。对于类似宇宙线 μ 子这类非“Point-like”事例，主要是通过水池和 RPC 进行径迹的联合重建 [66]。通过对 AD 的能量的刻度和重建，最后给出其能量分辨率为 $(7.5/\sqrt{E(\text{MeV})} + 0.9)\%$ 。

2.2.3 大亚湾物理分析轻量级框架 (LAF)

离线软件框架 NUWA 是在欧洲核子中心开发的对撞机物理离线框架 Gaudi 的基础上开发出的离线数据处理软件系统, 它沿袭了 Gaudi 框架的核心体系结构和数据持久化机制, 并根据大亚湾物理目标和探测器系统设计了物理事例的模拟, 刻度和重建等业务逻辑。但是, 从物理上讲, Gaudi 框架中的事例循环机制 EvtSelector 是根据对撞机物理的事例特点加以设计的, 适用于对前后没有时间关联的对撞机物理事例进行逐事例的分析处理; 从计算机技术上讲, Gaudi 框架诞生于上世纪 90 年代, 当时海量数据的存储依赖于磁带, 数据的读取主要依靠顺序读取的磁带机技术而非随机读取的磁盘技术。Gaudi 框架中, 以依赖瞬态事例缓存的数据持久化技术和对内存中瞬态物理事例进行逐次循环读取分析为核心的体系结构, 并不能很好地适应支持随机读取的磁盘机的运行, 面对反应堆物理实验中带有复杂时间关联的物理数据的分析, 则更加显示出不足。

为了克服对撞机框架 Gaudi 核心体系结构中这个不足之处, 更好地应对带有复杂时间关联的物理数据的分析需求。大亚湾实验在 Gaudi 的基础上, 根据反应堆物理事例的特点, 开发了仅用于物理分析的轻量级框架 LAF (Light weight Analysis Framework) [67]。其主要特点有:

- 沿袭了 Gaudi 框架中的三大标准模块: Algorithm, Service 和 Tool 及对每一种模块的管理方式;
- 利用简单文本文件 (.txt 格式), 实现了简单, 高效地对框架内各应用模块进行初始化配置;
- 沿用了 Gaudi 核心体系结构中的逐个物理事例进行循环获取的机制, 但改进了瞬态事例缓存: 主缓存可以缓存一段时间内的物理事例, 时间窗大小能够在模块配置文件中根据具体分析需要进行修改; 除此之外, 还有多个用户缓存以满足不同用户程序的特殊需求;
- 增加了事例“标记—过滤”机制, 使得用户可以根据自身需要, 专注于自己感兴趣的事例, 提高分析效率。

自大亚湾物理数据分析以来, 在 LAF 框架下, 完成了多个共用物理算法工具, 包括: 标记各种物理或非物理的事例 (宇宙线 μ 子事例, PMT Flasher 事

例), 排除宇宙线 μ 子带来的本底的 Veto 算法工具, 计算各探测器运行活时间的算法, 同时物理分析用户也利用 LAF 框架完成了诸如 IBD 事例、各种本底事例 ($^8\text{He}^9\text{Li}$, 快中子) 的挑选和估算程序。

Daya Bay 的离线软件系统仍然在不断的开发和完善过程中, 鉴于本论文篇幅, 有很多未涉及的部分, 包括大亚湾数字化模拟过程 [68]、事例显示软件等, 请参阅相关报告和文档。

2.3 大亚湾实验的初步结果

2012 年 3 月份, 大亚湾反应堆中微子实验国际合作组在北京召开新闻发布会, 公布了利用 2011 年 12 月 24 日到 2012 年 2 月 17 日的数据对中微子混合振荡参数 θ_{13} 的最新测量结果: $\sin^2(2\theta_{13}) = 0.092 \pm 0.016(stat) \pm 0.005(syst)$, χ^2/NDF 的值为 4.26/5, 以 5.2 倍 σ 的置信概率确认了 θ_{13} 不为零。

2.3.1 inverse β 信号事例挑选和本底事例估计

2.3.1.1 信号事例挑选和效率

大亚湾物理分析最重要的一个环节是通过能量和时间信息, 从探测器获取的数据中, 挑选出 inverse β 衰变事例并得到其快信号 (e^+ 和 e^- 湮没) 的能谱分布。大亚湾远近点由于距离反应堆基线不同, 中微子振荡的效应, 可以从 inverse β 衰变快信号的能谱变化中反映出来。

根据大亚湾探测器系统设计和 inverse β 衰变事例的物理特点和对 MC 模拟混合数据的分析研究, 挑选条件如下:

- 正电子湮没信号: $0.7 < E_{prompt} < 12.0\text{MeV}$;
- 中子在 Gd 同位素上的俘获信号: $6.0 < E_{delayed} < 12.0\text{MeV}$;
- 正电子快信号与中子俘获慢信号时间间隔: $1 < \Delta t < 200\mu\text{s}$;
- μ 子带来的本底排除: 如果 $t_{delayed} - t_{\mu_{WS}} < 600\mu\text{s}$, $t_{delayed} - t_{\mu_{AD}} < 1000\mu\text{s}$ 或者 $t_{delayed} - t_{\mu_{sh}} < 1000\mu\text{s}$ 快慢符合信号将会被反符合掉;

- 多重度 cut: 多重度 cut 要求在快信号之前 $200\mu s$ 以及慢信号之后 $200\mu s$ 没有能量大于 0.7 MeV 的信号。

其中, E_{prompt} 是快慢符合中的快信号能量, $E_{delayed}$ 是相应的慢信号能量, $\Delta t = t_{delayed} - t_{prompt}$ 是快慢信号之间的时间差。

影响 IBD 事例的挑选效率有两个因素: 探测器设计时候的探测效率和事例挑选的 cut 条件。如表2.4。但是, 物理分析中没有必要确定绝对效率, 关联误差会因为远近相对测量而被完全消除, 只有非关联误差将会对 θ_{13} 的计算产生影响。

inverse β 衰变产生的中子在钷的两种同位素 ^{155}Gd 、 ^{157}Gd 上俘获的比例 (0.227) 通过拟合中子俘获能谱得出。此外, 部分 inverse β 衰变顶点位于普通液闪之内, 但其中子却能够漂移到掺钷液闪区域被俘获 (*spill-in* 效应), 这部分效应的影响可以通过模拟研究, 相应的, 也有部分在掺钷液闪区域内产生的中子漂移到普通液闪中被 H 核俘获 (*spill-out* 效应), 这部分效应可以通过散裂中子和 Am-C 源的真实数据以及蒙卡模拟中得出。根据模拟结果 *spill-in* 效应要明显大于 *spill-out* 效应, 这是由于中子在 Gd 核上的俘获截面远远大于中子在 H 核上的俘获截面。

inverse β 衰变反应产生的中子在钷上的俘获的能量峰值被重建到 $8.05\text{MeV} \pm 0.5\%$ 。通过比较不同 AD 中的 IBD 中子、散裂中子在钷上俘获的能量峰以及 α 能量沉积可以得到诸 AD 之间的相对能标, 其中非关联能标误差为 0.5% 。通过数据和蒙卡模拟, 由于慢信号 $E_{delayed}$ 能量选择带来的相对效率误差估计为 0.12% 。

μ 子 veto 的效率丢失 ($1 - \epsilon_{\mu}$) 通过对每个 μ 子的 veto 时间累加得到, 这其中考虑了 μ 子 veto 时间窗口重叠效应。因为多重度选择条件造成的效率丢失 ($1 - \epsilon_m$) 则通过考虑在 IBD 事例时间附近偶然出现的非关联信号的概率进行计算。测量的非关联事例率 ($\sim 60\text{Hz}$) 是在进行了 μ 子反符合以及对已知的时间关联信号进行了排除之后估计得到的。在确定 ϵ_{μ} 和 ϵ_m 时考虑了在 AD 中 μ 子事例率和单事例率的和时间相关的晃动。平均的 ϵ_{μ} 和 ϵ_m 值参见表2.3

2.3.1.2 本底估计

大亚湾实验中，需要考虑的本底有如下几种：

- 偶然符合本底：偶然符合本底的事例率估算比较简单，只需要计算恰好落在满足 IBD 事例快慢信号时间间隔 Δt 中的“类 e^+ ”以及“类中子”的信号偶然符合概率。两种信号的事例率由挑选得到的“AD Single”事例在不同能量区间中的事例率得到。大亚湾，岭澳和远点实验大厅每个 AD 中的偶然符合事例率分别为：10 个/天，8 个/天和 3 个/天。偶然符合事例率的误差 ($\sim 1\%$) 是通过减除“类中子”事例过程中引入的系统误差得来的。
- 探测器中由宇宙线 μ 子引致的散裂中子或者能够产生 $\beta - n$ 级联衰变的不稳定长寿命同位素：对于穿过 AD 携带可见能量 $> 20\text{MeV}$ 的 μ 子。内水池 (外水池) 的探测效率为 $99.7\%(\sim 97\%)$ 。 μ 子在内水池中衰变了，以至于外水池有明显的效率降低。 μ 反符合探测系统高效地排除了 μ 子引发的本底。来自 ${}^9\text{Li}/{}^8\text{He}$ 的 $\beta - n$ 级联衰变的关联本底是通过得到这些关联信号距离它们之前，离得最近的 μ 子的时间分布，并进行时间拟合得到的，拟合时用到了已知的这些同位素的衰变寿命。
- 在探测器外部由 μ 子引致的快中子：快中子本底来源于未被 μ 子反符合探测器系统探测到的 μ 子引发的中子和液闪或者掺钆液闪中的质子发生碰撞后，中子又被 Gd 核俘获，这样产生的快慢符合信号类似于 IBD 事例，属于关联本底。如果把 μ 子 veto 条件中， μ 子的能量阈值放开到 50MeV ，并将 IBD 挑选条件中快信号的能量 cut 设置为： $E_{prompt} < 50\text{MeV}$ ，则在快信号能谱中可以观察到由于快中子引起的从 12MeV 到 50MeV 的平坦能谱分布。把所有这些高能的快信号都归结于快中子，并且假设所有被 veto 掉的和未被 veto 掉的 μ 子产生的中子能谱都是相同的，那么在 IBD 信号区域的本底被估计为 $\sim 0.1\%$ 。
- ${}^{13}\text{C}(\alpha, n){}^{16}\text{O}$ 反应产生的本底：在液闪中来自 ${}^{238}\text{U}$, ${}^{232}\text{Th}$, ${}^{227}\text{Ac}$ 以及 ${}^{210}\text{Po}$ 的放射性，通过测量它们的级联衰变或者拟合其能量峰值得到；来

自 $^{13}\text{C}(\alpha, n)^{16}\text{O}$ 反应的本底被探测器中 α 衰变所限制, 这个本底可以通过蒙卡模拟进行估算。

- 由于 ^{241}Am - ^{13}C 中子刻度源引起的关联本底: 来自于 0.5 Hz 的中子源的中子在钢罐上进行非弹性散射以及随后在 Fe、Cr 或者 Mn 上的俘获会产生 γ 光子。如果这两个散射和俘获的信号都进入了闪烁体区域, 它们就非常可能被判定成是 IBD 事例的快慢信号符合。由此产生的类似慢信号的闪光事例得到了测量, 数据表明该类事例率为 ~ 230 个/天。通过模拟的预测, 关联中子事例和单中子事例的相对比率为 $\sim 0.09\%$, 由此可以估计本底事例率为 0.2 ± 0.2 个/天/AD。

表2.3汇总给出了在当前时间取数时, 分析数据得到活时间, 信号和各种本底事例的事例率。

	AD1	AD2	AD3	AD4	AD5	AD6
IBD 候选	28935	28975	22466	3528	3436	3452
DAQ 活时间 (天)	49.5530		49.4971		48.9473	
μ 反符合时间 (天)	8.7418	8.9109	7.0389	0.8785	0.8800	0.8952
$\epsilon_\mu \cdot \epsilon_m$	0.8019	0.7989	0.8363	0.9547	0.9543	0.9538
偶然符合 (/天)	9.82 ± 0.06	9.88 ± 0.06	7.67 ± 0.05	3.29 ± 0.03	3.33 ± 0.03	3.12 ± 0.03
快中子 (/天)	0.84 ± 0.28	0.84 ± 0.28	0.74 ± 0.44	0.04 ± 0.04	0.04 ± 0.04	0.04 ± 0.04
$^9\text{Li}/^8\text{He}$ (/AD/天)	3.1 ± 1.6		1.8 ± 1.1		0.16 ± 0.11	
Am-C 关联 (/AD/天)	0.2 ± 0.2					
$^{13}\text{C}(\alpha, n)^{16}\text{O}$ (每天)	0.04 ± 0.02	0.04 ± 0.02	0.035 ± 0.02	0.03 ± 0.02	0.03 ± 0.02	0.03 ± 0.02
IBD 事例率 (/天)	714.17 ± 4.58	717.86 ± 4.60	532.29 ± 3.82	71.78 ± 1.29	69.80 ± 1.28	70.39 ± 1.28

表 2.3 信号和本底总结. 本底和 IBD 事例率被 $\epsilon_\mu \cdot \epsilon_m$ 效率所修正 [56]

表2.4汇总了有关探测器和信号事例挑选各 cut 的效率和关联以及非关联误差。

2.3.2 物理灵敏度计算

大亚湾实验的方案设计中, 共有大亚湾, 岭澳和远点三个实验大厅, 共 8 个全同的反中微子探测器。近点和岭澳点的探测器可以用来测量中微子事例率, 并作为远点探测器观测到的中微子事例率的预测。通过远点探测器观测到的中微子事例率 M_f 和使用近点探测器预测的无振荡的远点探测器事例率 \bar{N}_f 的比值 $R = \frac{M_f}{\bar{N}_f}$ 可以确认电子型中微子是否减少。

	效率	关联	非关联
靶质子数		0.47%	0.03%
Flasher cut	99.98%	0.01%	0.01%
慢信号能量 cut	90.9%	0.6%	0.12%
快信号能量 cut	99.88%	0.10%	0.01%
多重度 cut		0.02%	<0.01%
俘获时间 cut	98.6%	0.12%	0.01%
钷俘获比例	83.8%	0.8%	<0.1%
Spill-in	105.0%	1.5%	0.02%
活时间	100.0%	0.002%	<0.01%
合计	78.8%	1.9%	0.2%

表 2.4 绝对效率, 关联和非关联系统误差汇总 [56]

大亚湾实验观测到了中微子的减少 [56]:

$$R = 0.94 \pm 0.011(stat) \pm 0.004(syst) \quad (2.3)$$

观测到的远点中微子能谱和由近点预测的能谱的比较如图2.7所示。

测量的六个探测器的 IBD 事例用来估计 θ_{13} 。基于探测到的和预测能谱的差别, 考虑到来源于反应堆的误差, 探测器的系统误差和本底事例及其误差, 构建一个标准的带 pull 项的 χ^2 函数 [56], 可以拟合得到 $\sin^2 2\theta_{13}$ 的最佳值。

图2.8显示了在每个探测器里测量的事例数相对于预期的无振荡事例数的比值。3号实验厅相对于其他两个实验厅的 6.0% 的事例数缺损相当明显, 显示了非零 θ_{13} 的非常明确的证据。振荡的存活几率的最好估计值由图中光滑的曲线给出。右上角显示了 χ^2 函数和 $\sin^2(2\theta_{13})$ 的关系图。

最佳拟合值为 [56]

$$\sin^2(2\theta_{13}) = 0.092 \pm 0.016(stat) \pm 0.005(syst) \quad (2.4)$$

2.4 本章小结

本章主要介绍了大亚湾反应堆中微子实验的概况, 分别介绍了实验的总体布局和工程设计、实验中的中微子探测器以及宇宙线 μ 子反符合探测系统的物理设计和实现; 实验电子学以及触发系统的实现; 较为详细地介绍了大亚湾实

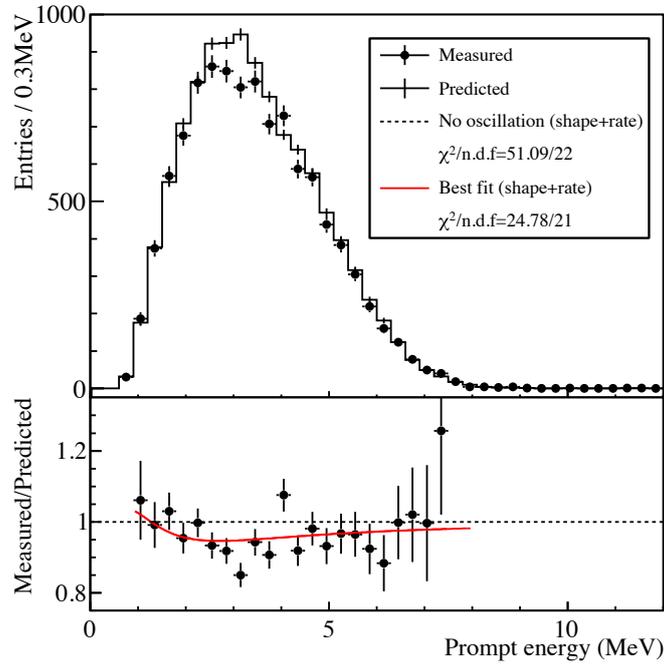


图 2.7 上图：远点实验厅测量的正电子能谱同由近点两个厅的数据预测的无振荡的能谱的比较，仅显示了统计误差。下图：测量能谱同无振荡预测能谱的比值。显示了无振荡假设下的 $\chi^2/n.d.f$ 以及振荡的最佳拟合曲线 (考虑了所有的系统误差) [56]

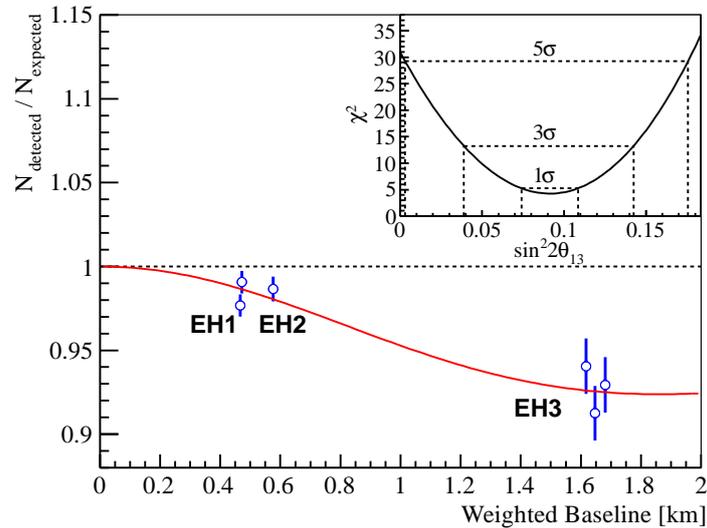


图 2.8 每个探测器中测量信号和预期无振荡信号的比较。反应堆和勘测数据用来计算权重的平均基线。振荡存活几率经过了最佳拟合归一化参数的修正，最佳拟合值由光滑曲线给出。右上角给出了 χ^2 函数和 $\sin^2(2\theta_{13})$ 的关系图 [56]

验的离线软件支撑平台硬件参数和离线数据处理软件框架的基本构建和离线数据模拟，刻度，重建以及分析的基本流程；最后，对大亚湾的实验数据分析方法给出了简明的概括，对最新的实验数据的分析结果，也做了扼要的介绍。

第三章 大亚湾 DAQ 数据到离线数据对象转换

Daya Bay 实验厅内各探测器模块 (内外水切伦科夫探测器, 气体阻性板探测器, 反中微子中心探测器) 的输出信号经过在线数据获取系统 (DAQ) [69] 的收集和编码, 以二进制格式文件存储在磁盘上, 称之为二进制原始数据 (Raw data), 供后续离线数据处理使用。为了高效实现对原始数据的访问, NuWa 离线软件框架中开发了专门用于处理原始数据的数据转换服务。通过应用面向对象编程技术, 数据转换服务隐藏了二进制格式数据到离线数据对象转换的复杂细节, 使得不同应用程序通过统一的接口访问离线数据对象成为可能, 简化了数据获取过程。离线软件框架下的数据转换服务应用了独特的方式对各软件子模块进行管理, 实现了整个数据转换系统模块间的低耦合和高可扩展性, 便于维护和扩充优化。

本章首先简单介绍 NuWa 框架中 I/O 接口中的数据持久化机制; 然后介绍 Daya Bay 原始数据格式和电子学读出事例离线数据对象, 并阐述了原始数据服务的工作原理和管理方式; 最后, 给出了利用二进制数据转换服务在反中微子探测器 Mini Dry Run 和 Dry Run 期间完成的电子学台阶和阈值刻度, 也进行了必要的介绍。

3.1 NuWa 框架中的数据持久化机制

在物理数据离线分析的过程中, 不同阶段的数据文件中的数据格式可能会发生变化。有些中间数据文件需要存储在介质 (如磁盘, 磁带, 闪存) 上。这部分数据一旦存储完毕, 将不再依赖程序的运行和介质的通断电状况, 被称为持久化数据文件 (Persistent Data File)。持久化数据文件中的数据格式称为持久化数据类型, 一般使用较为简单的数据格式, 以节省存储介质的存储空间。

通过适当的配置, 定义良好的持久化数据可以通过特定的软件 (如 ROOT) 直接调入内存中为用户使用。但是在 NuWa 框架中, 为了在数据处理过程中实

现对不同程序模块的数据共享，并简化数据访问接口，一般使用事例数据服务 (Event Data Service) 管理内存中存活的数据块。NuWa 框架在内存中开辟了一段区域，用于存放程序运行期间存活数据块，称为瞬态事例缓存 (Transient Data Store, TES)。TES 中的数据类型称为瞬态数据类型，一般使用较为复杂的数据格式，以满足不同程序模块数据共享的需要。每种持久化数据类型都有相应的瞬态数据类型对应。

持久化数据和瞬态数据之间的交互转换称之为数据持久化。在 NuWa 框架中，数据交互具体是由转换服务 (Conversion Service) 控制完成的。转换服务提供了对特定类型持久化文件的输入/输出 (I/O) 接口，和一组转换持久化数据到对应瞬态数据的转换器 (converter)。读取数据的过程中，数据转换服务由文件中获取持久化类型的数据，交由转换器完成到瞬态数据类型的变换，并将其存放到 TES 中；输出数据的过程正好相反。如图 3.1 所示。

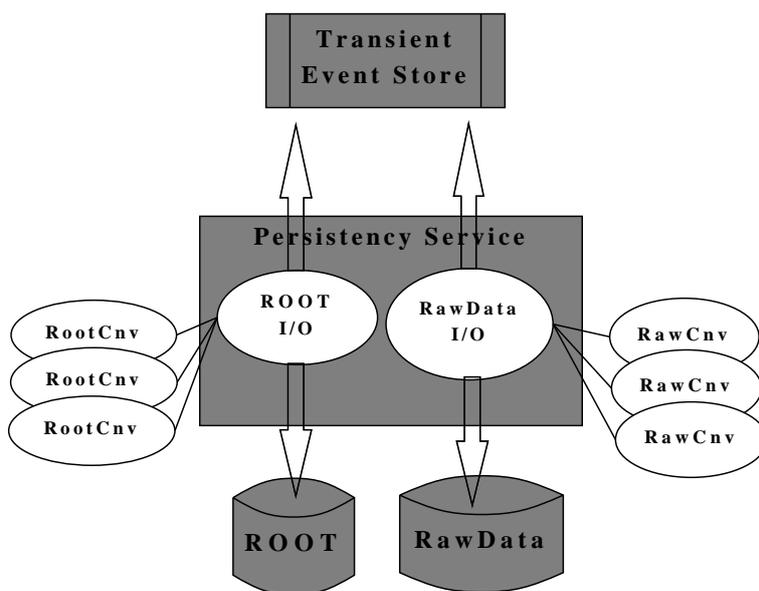


图 3.1 NuWa 框架事例持久化机制 (参考 Gaudi 手册)

从图中可以看出，数据转换器和持久化数据类型是一一对应的。NuWa 框架的数据转换器采用了独特的管理方式，数据转换服务会根据持久化数据的类型，自动选择相应的数据转换器。

在 NuWa 框架中，由于采用了面向对象编程语言，除了二进制格式的原始数据，持久化数据类型和瞬态数据类型全部以 C++ 语言中的类对象来实

现，通称为持久化数据对象和瞬态数据对象。持久化过程中的存储手段运用了 ROOT 中最新的对象持久化技术。

3.2 二进制原始数据格式和离线数据对象

直接从在线数据获取系统得到的二进制原始数据有与其对应的持久化数据对象和瞬态数据对象。本节先对二进制原始数据的存储格式进行介绍，以便于后文对原始数据的存取过程的讲述；然后介绍了二进制数据转换服务中的原始数据缓存对象和在 TES 中对应的瞬态数据对象，即电子学读出事例的离线数据对象。

3.2.1 二进制原始数据格式

Daya Bay 实验的二进制原始数据格式是参考 BES III 实验的数据格式进行设计的，数据格式最小单元片段是 32Bit 的连续存储空间。整个二进制的数据格式自上而下分为三层存储逻辑：文件存储逻辑，事例存储逻辑和数据单元存储逻辑。为提高二进制原始数据的读取效率并提供有效的数据监测手段，在保存了来自每个探测器模块的每个电子学通道的物理信息（时间、电荷信息）之外，为各层的存储逻辑又添加了额外信息。

3.2.1.1 文件存储逻辑

二进制原始数据文件存储逻辑包含文件头块、数据块和文件结束标识块组成 [70]，如图3.2所示。

其中各块分别存储信息如下：

- 文件头包含五条记录，里面保存了原始文件获取时间、文件名、和原始 Run 有关参数、触发类型和刻度参数等信息；
- 文件结束标识包含原始文件结束时间，文件包含事例数和文件大小等信息；
- 事例分隔记录（Data Separator Record）标识一个电子学读出事例结束和新事例的开始，及该事例的编号和事例数据长度等信息；

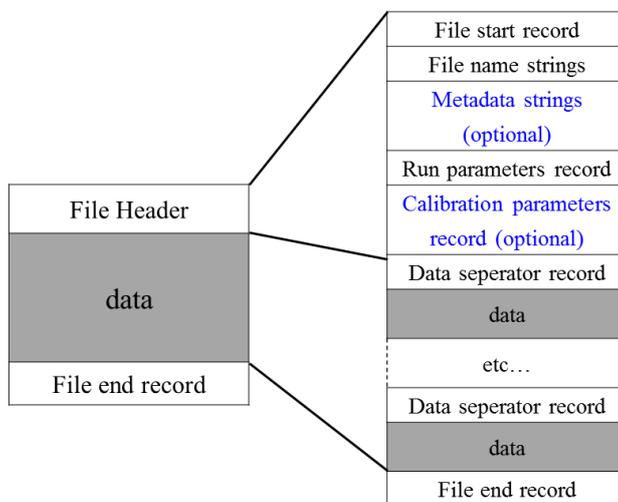


图 3.2 二进制原始数据文件存储逻辑

- 事例数据块中记录了电子学读出事例的具体内容，包含来自每个电子学通道的物理信息。

3.2.1.2 事例存储逻辑

事例存储逻辑应用良好规范的存储格式来整合多个读出模块数据段（Read Out Module Fragment, ROM）。ROM 段是与数据获取系统硬件电子学读出模块相对应的二进制数据结构。简单说来，每个 ROM 段对应硬件上一块连接光电倍增管的前端电子学板（FEE）或者是连接气体阻性板探测器的前端电子学板（FEC）。一个完整事例（Full Event）由多个 ROM 数据段构成，数据获取系统电子学硬件对应为：一个 VME 机箱包含多块前端电子学板。

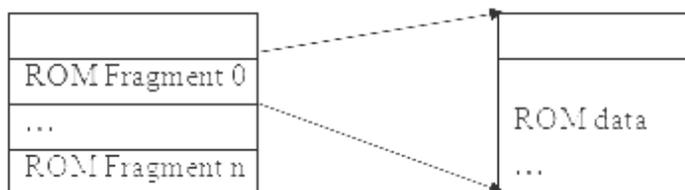


图 3.3 二进制原始数据事例存储逻辑

每个 ROM 数据段自身均会再细分为头部和数据区，数据段头部包含数据

段标记和数据段容量等信息，同时包含此 ROM 段的唯一标识符。多个更低层的数据段组成了一个 ROM 段，这些更低层的数据段称为数据单元逻辑。如图 3.3所示。

3.2.1.3 数据单元存储逻辑

每个电子学通道的读出数据由一个或者多个顺次排列的 32Bit 二进制数据块构成，成为一个数据单元。每个数据单元包含的都是按特定格式压缩编码的电子学通道击中信息。连接不同探测器模块的电子学通道的击中信息所包含的内容不同，编码方式也不相同。

与光电倍增管 (PMT) 对应的电荷时间数据单元编码相对简单 [71]，下面以它们为例进行说明。PMT 击中电荷信息包含有击中的 PMT 对应的电子学通道编号、相应通道电子学噪声台阶、PMT 的工作量程和电荷量。PMT 击中时间信息包含有被击中的 PMT 对应的电子学通道编号、电荷峰值对应的采样周期计数，多次击中计数，时间信息。上述每一项都在一个合理的范围内取值。所以当数据存储的时候，并不是为这些数据项分配独立的数据单元，而仅仅需要根据该数据项的取值范围为其分配恰当的 Bit 位，将多个关联的数据项压缩到同一个 32Bit 长的数据单元中。PMT 的击中信息经过这样压缩编码后，每一个击中的时间，电荷信息正好对应一个数据单元，格式如表 3.1，表 3.2所示。

31	30	29	28-25	24	23-12	11-0
0	1	1	电子学通道编号	量程	电子学噪声台阶	电荷量

表 3.1 PMT 击中电荷信息二进制编码。

31	30	29	28-25	24-21	20-17	16-0
1	0	0	电子学通道编号	峰值所在采样周期计数	多重击中计数	时间

表 3.2 PMT 击中时间信息二进制编码。

有关事例存储编码格式的，相应的技术文档 [70][71] 中有更加详细的介绍。

3.2.2 Offline 数据对象模型

3.2.2.1 电子学读出数据的离线数据对象模型

由于 C++ 面向对象编程技术贯穿了整个 NuWa 离线软件框架的设计理念，在使用 NuWa 进行数据处理的时候，所有算法程序 (Algorithm) 的输入和输出都是组装成 C++ 类形式的数据，称之为数据对象。数据对象模型描述了各种数据对象之间继承、聚集和交互依赖关系，是真实世界对象的抽象和对应。电子学读出事例的离线数据对象模型 (Readout) 是与二进制原始数据对应的瞬态数据对象，展示了数据获取系统电子学硬件的组成方式。按照 UML 图对类和对象关系的描述，简化的电子学读出对象模型如图 3.4所示。从图中可以看出，

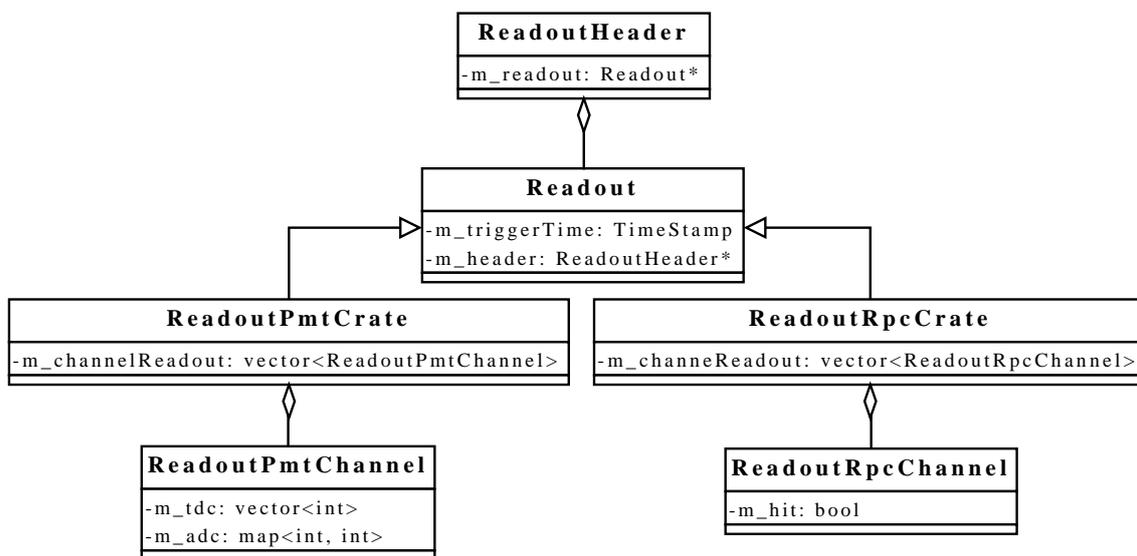


图 3.4 电子学读出事例离线数据对象模型 UML 图。空心三角箭头表示类继承，空心菱形箭头表示对象聚集

电子学读出事例离线数据对象模型是一组对象集合，对象聚集的方式，是真实世界中对应对象的组合模式。Readout 数据基类用于表示来自某一探测器模块电子学机箱的一个事例信号，由它派生得到的各子类型 (ReadoutPmtCrate 和 ReadoutRpcCrate) 分别对应来自 PMT 机箱和 RPC 机箱的信号。每个子类型中聚集了多个表示电子学通道的 Channel 类型，用以存储电子学通道记录的时间电荷信息。

NuWa 框架中，通过将离线数据对象在 TES 中缓存，实现数据在不同的应

用程序间共享。TES 中的数据对象以树形结构进行组织和管理。这些数据对象被归类放置于类似文件系统的路径下。在事例根路径“/Event”下，各种不同逻辑层次的数据通过下级路径识别区分，如“GenHeader”表示模拟事例产生子对象，“Readout”则表示电子学读出事例对象。如图 3.5所示。在每一条路径

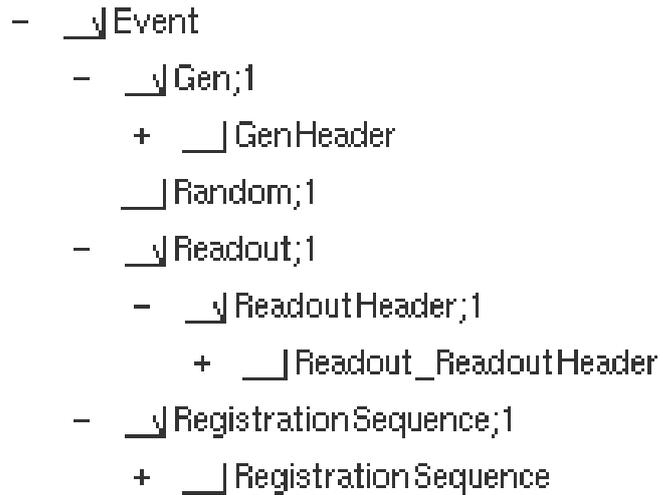


图 3.5 TES 树形组织目录

的最底层，即 TES 目录树的叶节点上，存放最终的离线数据对象，电子学读出事例对象存放在“Readout_ReadoutHeader”路径之下。

3.2.2.2 二进制原始数据的缓存对象模型

无论是二进制原始数据格式，还是与之对应的电子学读出数据对象模型，都是在数据转换服务构建之前预先定义了的。为了增加数据转换服务的使用灵活性，充分利用内存的缓存空间实现事例延迟加载以达到对程序的优化加速。在数据转换服务过程中，另外定义一个与原始数据对应的缓存对象模型 RawDataEvent，作为原始数据转换服务程序中，二进制原始数据和电子学读出数据对象之间的一级缓冲。如图 3.6所示。

RawDataEvent 模型也是对象模型，较电子学读出事例离线数据对象模型为简单，同样是以对象集的聚集方式，缓存由二进制格式解码的临时数据。

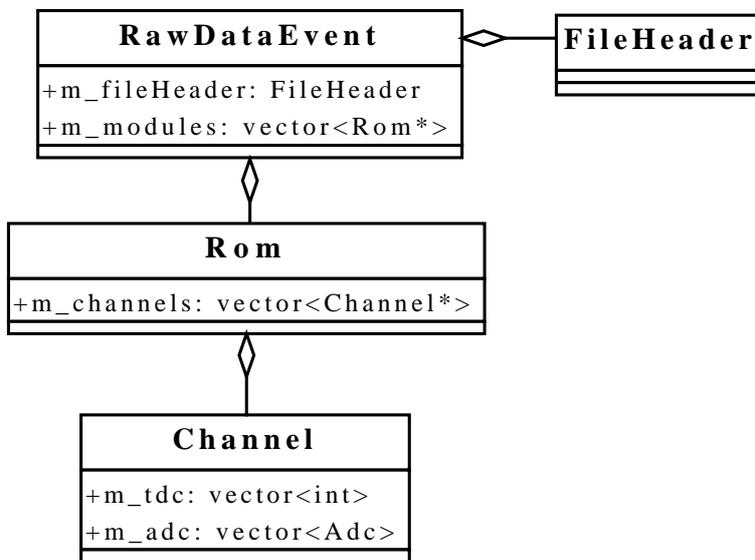


图 3.6 二进制原始数据中间缓存对象模型

3.3 原始数据转换服务

在 NuWa 框架下的原始数据转换过程中。原始数据输入模块 (RawDataInputSvc) 读入二进制原始数据，解析其格式并生成原始数据的缓存对象。原始数据转换管理服务 (RawDataConSvc) 控制转换器转换缓存对象，生成 TES 中的瞬态数据，即电子学读出离线数据对象。原始数据转换管理服务通过独特的方式管理已经添加到转换器列表中的数据转换器，并能根据 TES 中目标数据对象的类型，自动调用相应的数据转换器。整个过程，如图 3.7所示。

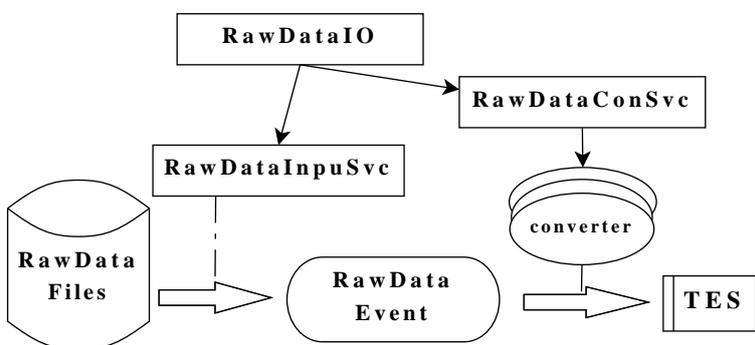


图 3.7 原始数据转换过程图

3.3.1 原始数据输入与二进制解码

RawDataInputSvc 模块负责底层二进制原始数据输入和整个解析过程。与二进制的三个存储逻辑相适应，解析过程包括了对文件格式，事例数据格式和二进制数据格式三个层次的解析。实现数据从二进制格式到其缓存对象模型的转化。

3.3.1.1 二进制文件格式解析

文件格式解析的主要目的是：

- 读取存储在文件头尾中的各条记录，并根据这些记录提供的信息，对文件输入流中的文件进行划分；
- 同一个文件内部，将文件头尾记录部分与具体的事例数据部分分离；
- 同一个文件内部，读取文件中的事例分隔符，将不同的事例数据独立，以进行下一步的事例数据解析。

为方便程序各模块的维护和改进，并尽量做到模块间的低耦合，文件输入流中的文件划分工作封装在 nextInputFile 模块中，而文件头尾各条记录和事例分隔符的解析、识别过程封装在 nextRecord 模块中。两个模块之间并没有直接通信，而是由第三方程序控制协同工作。每次 nextInputFile 模块在输入数据流中打开一个新文件，nextRecord 模块会依次读入文件头中的各条记录并在必要的时候检查其合法性，在此过程中，文件流指针会同步移动到此文件的事例数据区，为事例数据读取和解析做好准备。nextRecord 模块也会识别此文件中的事例分隔符，并依照其中记录的文件块长度信息，在内存中开辟相应大小的缓冲区，并将这一块事例整个读入到事例数据缓冲区，以提高数据读取效率。在文件格式解析过程中，若 nextRecord 模块检测到文件结尾记录，则由 nextInputFile 模块打开文件输入流中的下一个文件继续读取。文件输入流中所有文件处理完毕后，亦由 nextInputFile 模块与框架通信，告知文件输入流处理完毕，整个数据转换过程随之结束。

3.3.1.2 二进制事例和数据格式解析

根据事例分隔符独立开来的数据，在读入事例数据缓冲区之后，需要进一步进行事例数据格式解析，以分离具体物理数据和每个事例块自身头尾的各个记录。这一过程由 `readRawEventBuffer` 模块完成。`readRawEventBuffer` 模块从事例数据缓冲区中读取一个完整事例数据之后，如同文件格式解析过程类似，它先解析其头部以提取事例相关信息，然后以相同的过程顺次向后解析所有的 ROM 数据块。在解析 ROM 数据块的过程中，此模块根据各数据块头部的记录信息，分别创建缓存对象模型中的各次级对象，然后将 ROM 数据块数据区中的每个 32Bit 数据单元进行识别，并按照相应的编码格式进行转换，并存入到相应的次级对象中。最后，缓存对象模型中的各次级对象，聚集成为缓存对象模型定义的格式。经过这个过程，事例数据中的信息也得到分离和解析，从二进制的文件中完全独立出来，并且以缓存对象模型暂时存储在内存中。

3.3.1.3 事例转换接口实现

作为 NuWa 框架数据转换服务中的非默认模块，`RawDataInputSvc` 服务模块在实现具体的文件格式，事例数据格式的解析、转换、对象创建等功能之后，还需要提供简单的接口对复杂的实现细节进行封装，并将接口提供给数据转换服务中的其他默认模块调用。受到 NuWa 框架事例循环 (event by event cycle) 机制的启发，我们在 `RawDataInputSvc` 服务中，将前两节的数据转换过程细节隐藏，封装成为一个简单的 `nextEvent` 接口。每当其他的模块调用 `RawDataInputSvc` 服务的 `nextEvent` 接口的时候，就可以得到一个经过解析、转换、对象创建的 `RawDataEvent` 缓存对象。`nextEvent` 接口在实现了对格式解析和对象创建的细节的封装之外，还负责协调 `nextRecord` 模块和 `nextInputFile` 模块之间、`nextInputFile` 模块和框架之间的通信，以在文件结束或者整个文件输入流结束时，通知相应模块做出正确的响应。上述 `nextRecord`，`readRawEventBuffer`，`nextInputFile` 模块和 `nextEvent` 接口的工作流程，如图 3.8所示。

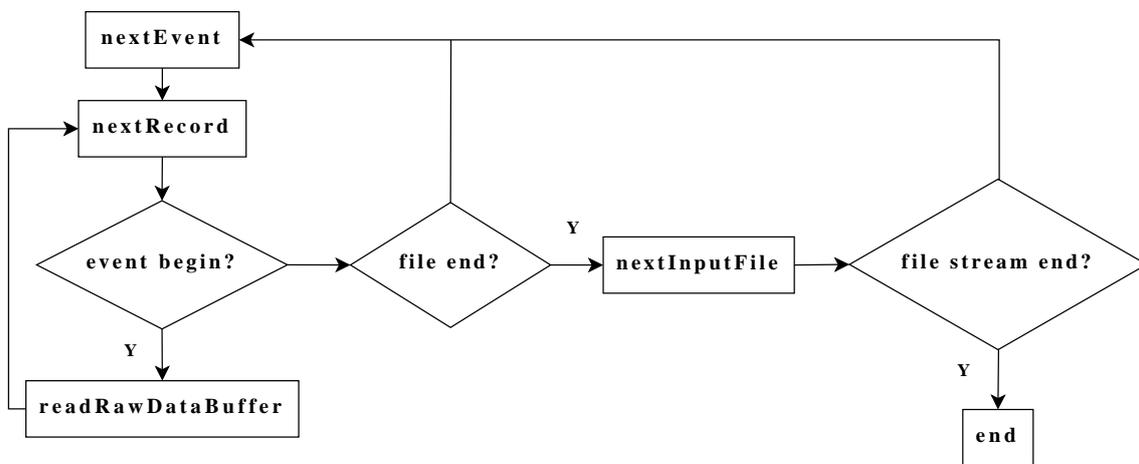


图 3.8 文件格式解析，事例和数据格式解析工作流程图

3.3.2 离线数据对象格式转换

3.3.2.1 转换服务与转换器管理

完成上面三个步骤之后，RawDataInputSvc 已经将位于磁盘上的二进制原始数据中一个完整的事例解码转换为内存中相对应的缓存对象。这个缓存对象需要被再次转换为框架中标准的电子学读出事例离线数据对象，并注册到 TES 中，从而实现数据对象共享。这个过程是由 NuWa 框架中的数据转换服务中定义的标准数据转换器（Converter）完成的。

从电子学读出事例离线数据对象在 TES 中的组织形式可以知道，这一对象位于 TES 根目录以下的若干次级路径下。不同的路径下还有其他的数据对象。数据转换服务不仅要能够实现数据类型的正确转换，还要能够识别所转换的数据对象在 TES 中对应的位置，以便顺利将转换后生成的对象注册到 TES 中。

为了同时实现这两个功能，而又不破坏数据转换器的高内聚性，需要在转换服务模块定义中对这两种功能做适当的功能分解，在转换服务运行时，再将两种功能结合以完成运行任务。因此，原始数据转换管理服务 (RawDataConSvc) 中，引入了与 TES 管理数据相类似的方式，采用树结构对不同的转换器进行加载和管理。如图 3.9所示。

数据转换管理服务通过构建一棵“树”来管理加载到它管理列表中的数据转换器。树的每个叶节点“Leaf”记录了每个具体的数据转换器的唯一标识符

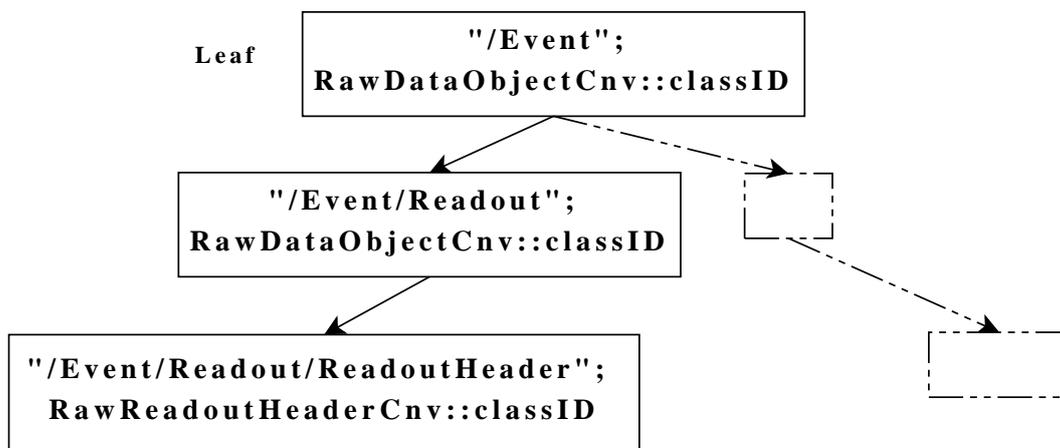


图 3.9 数据转换管理服务中定义的数据管理器“树”

(classID), 及与此数据转换器具有相同标识符的数据对象在 TES 中的存放路径。

在数据转换程序的初始化阶段, 数据转换管理服务负责构建这颗“树”, 将数据转换器和它所负责的数据对象所在的 TES 路径进行绑定。

在数据转换程序的运行阶段, 当用户程序向 NuWa 框架的事例数据服务 (Event Data Service) 发出原始数据转换请求时, 事例数据服务首先在 TES 中查找所要求的离线数据对象, 若找到则返回, 否则请求数据转换管理服务调用转换器做数据转换。转换管理服务则根据用户请求的数据对象类型的 classID, 自动匹配加载到它的管理列表中的转换器。各转换器在进行数据转换时, 首先从数据缓存对象 RawDataEvent 中获取当前事例的物理数据信息, 然后生成对应离线数据对象, 并用缓存对象中的信息给对应的离线数据对象赋值。最后, 数据转换管理服务把生成的离线数据对象注册到 TES 中。这种工作模式, 实现了数据对象和 TES 路径绑定和数据对象转换两种功能的功能分解。同时, 也为新的数据对象的添加预留了充分的空间。

从上述的原始数据转换过程可以抽象出 NuWa 框架中的数据对象的一般转换过程, 即 NuWa 框架中的数据持久化过程。如图3.10所示。

TES 目录的树形结构中, 只有位于最底层的叶节点代表具体的离线数据对象模型, 中间层的叶节点表示目录的路径, 并没有相应的离线数据对象对应。这个时候, 也会为每个路径叶节点定义空的数据转换器, 如图 3.9 中的

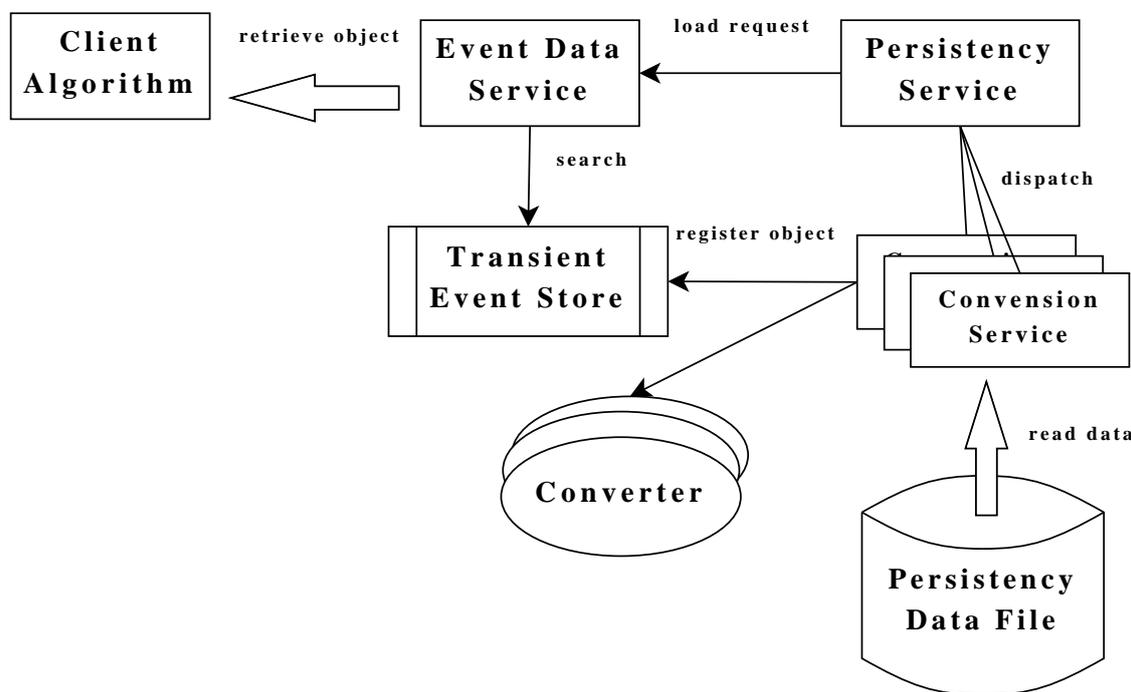


图 3.10 NuWa 框架数据持久化过程

RawDataObjectCnv，以保证 TES 中每个目录树叶节点和数据转换服务中的每个管理树叶节点相对应。

在数据转换过程中，当用户程序向 NuWa 框架的事例数据服务（Event Data Service）发出原始数据转换请求时，事例数据服务首先在 TES 中查找所要求的离线数据对象，若找到则返回，否则请求数据转换服务调用转换器做数据转换。

3.4 原始数据转换器功能应用举例

原始数据转换服务作为 NUWA 框架中的标准模块，应用于来自 DAQ 的二进制数据的转化。在大亚湾反中微子探测器的性能调试期间，发挥了关键性的作用。

大亚湾反中微子探测器在完成机械机构和 PMT 的安装、灌装液闪、掺钷液闪和矿物油并安置在水池中运行取数之前，需要在 SAB（地面装配大厅）进行性能调试。其中灌装液闪之前的性能调试（Dry Run）是对 PMT 及其前段电

子学的检测，尤为重要。调试的方法是将发光二极管（LED）放置于探测器中，以 PMT 接收 LED 发射的光学光子，对探测器的输出数据进行分析处理，以得到 PMT 的单光电子响应、暗噪声计数以及 FEE 板各电子学通道的台阶和阈值等刻度信息。

3.4.1 FEE 电子学板的电子学台阶刻度

在探测器调试数据的分析过程中，RawDataIO 将来自于 DAQ 的二进制数据经过解码和事例组装成为 Readout 类提供给用户使用。在 Pedestal 取数 Run 中，Readout 类中包含了所有电子学读出道的信息，其中 Pedestal 以 ReadoutPMTChannel::Adc 的形式存储，数据分析时，可以很容易得到 Pedestal 的分布信息。

图3.11显示了 AD1 中某一道的 Pedestal 的分布情况。左图是这一道的 Pedestal 的 ADC 的分布，其分布可以用一个高斯分布函数进行描述和刻度。右图是相应的 Channel 的 Pedestal 的 ADC 分布随着时间的变化，纵轴是 ADC 的值，横轴是时间，短时间内 ($< 10^2$ s)，可以认为 Pedestal 仅有随机性变化。

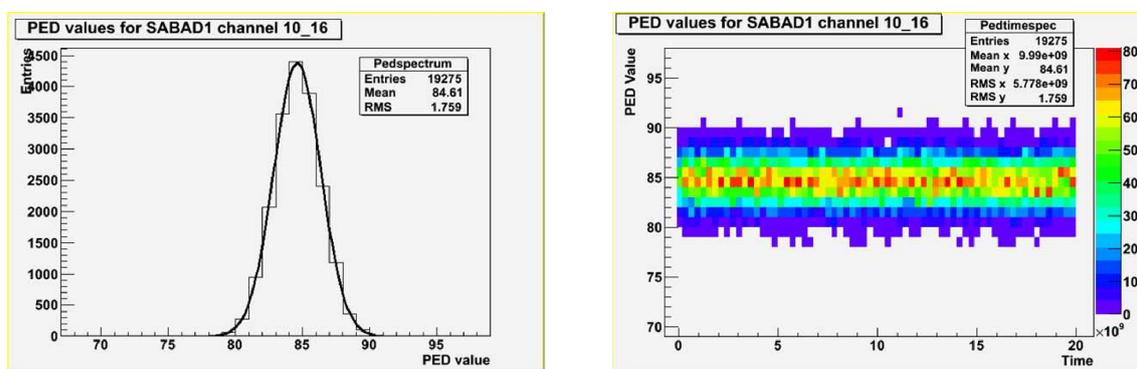


图 3.11 左图为单通道台阶分布，可以用高斯函数进行拟合；右图为 Pedestal 随时间变化，短时间内，可认为 Pedestal 分布没有系统性变化。

图3.12显示了一块 FEE 板上所有电子学通道的 pedestal 在长时间内 (> 24 h) 的变化情况。受环境影响，每一道的 pedestal 在较长时间内都有比较大的起伏，且同一块电子学板上的所有电子学通道的变化趋势几乎一致。

根据对电子学 Pedestal 分布的研究，大亚湾摒弃了通过特殊的 Pedestal Run 的刻度平均值来作为一段时间内的 Pedestal 值的做法，而是采用了每次记

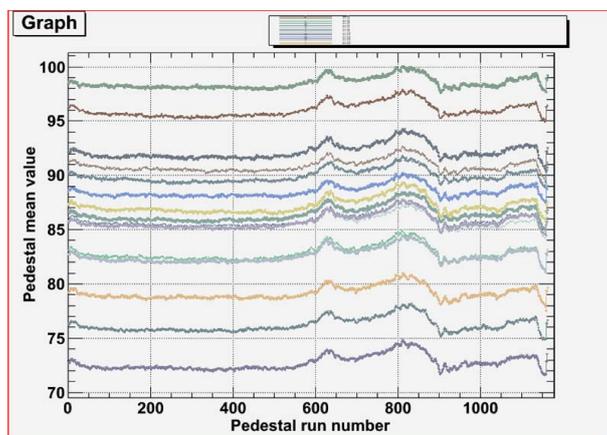


图 3.12 Pedestal 在长时间内的变化，横轴为不同的 Pedestal run 的 run 号，可以近似作为时间轴。纵轴是每个 Pedestal run 刻度出来的 Pedestal 值。

录 ADC 的时候，同时记录在阈值条件之前 100ns 时间窗内的平均的 ADC 值作为 Pedestal，并记为 `ReadoutPMTChannel::PreAdc`，数据分析时，用相应的 `ADC-PreAdc` 值作为真实的物理值的做法，以减小 Pedestal 带来的误差。

3.4.2 FEE 电子学板的电子学阈值刻度

对于触发系统而言，无论是基于过阈值数的多重性触发模式以及基于电荷求和的能量求和触发模式，触发系统都有一定的阈值设定，以减小能量过低的随机噪声带来的本底事例。而对于 FEE 电子学板上的每一道电子学通道，也都设定了一定的阈值，以排除电子学台阶噪声的影响。但在探测器性能调试的时候，预先设置好的电子学通道阈值有两个问题：

- 电子学阈值是 DAC 的电压阈值（大致相当于 PMT 单光电子峰对应的 ADC 道数的 $1/3$ ），但是并没有公认有效的算法，将 DAC 的电压阈值转化成为电荷阈值（ADC 道数）；
- 根据探测器调试取得的数据的结果，ADC 的分布并不是在某一处（约 PMT 单光电子峰 $1/3$ ）以下的部分为零，在 ADC 值较小（约为 PMT 单光电子峰处 ADC 的 $1/10$ ）仍然有分布。

根据以上的描述，可以构建一个如下简单的模型，近似描述 DAC 电压阈值和 ADC 积分电荷阈值之间的关系，来构建刻度算法，将 DAC 设置的电子学

阈值转化成为对应的 ADC 道数阈值，并能够给出其分布。

对于电子学通道没有设置阈值的情况，可以用 PMT 的全模型 [72] 来对 PMT 的谱形进行拟合，得到单光电子峰甚至是台阶等参数，全模型公式如公式3.1所示：

$$SER(x) = \sum_{n=0}^{N_{max}} P(n; \mu) \cdot G_n(x) \otimes B(x) \quad (3.1)$$

公式中，各部分的具体形式和物理意义如下：

- $P(n; \mu) = \frac{\mu^n e^{-\mu}}{n!}$ 是 Poisson 分布，其中 $\mu = mq$ ，m 是击中光阴极的平均光子数，q 是光阴极的量子效率， μ 就是光电倍增管第一打拿极收集到的平均光电子数。 $P(n; \mu)$ 就是当平均光电子数为 μ 的时候，光电倍增管接收到 n 个光电子的概率。
- $G_n(x) = \frac{1}{\sqrt{2n\pi}\sigma} \exp(-\frac{(x-nQ)^2}{2n\sigma^2})$ 是 n 个单光电子高斯分布的卷积，其中，Q 是当光电倍增管的第一打拿极接收到一个光电子时的平均输出电荷， σ 是输出电荷分布的标准偏差。
- $B(x) = \frac{1-w}{\sqrt{2\pi}\sigma_0} \exp(-\frac{(x-Q_0)^2}{2\sigma_0^2}) + w\theta(x)\alpha \exp(-\alpha x)$ ，是电子学通道的台阶分布和指数衰减本底之和，其中， $\theta(x)$ 是阶跃函数。

在不设置电子学通道阈值的情况下，得到的 PMT 单光电子谱形及利用全模型公式进行拟合的结果如图3.13所示：

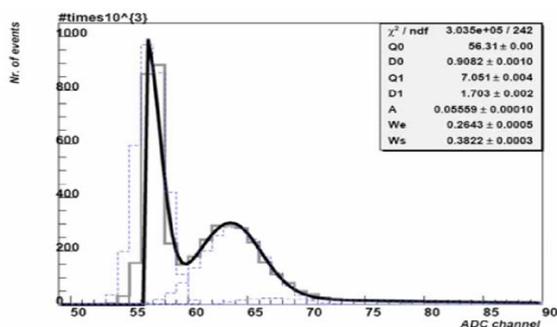


图 3.13 PMT 谱形全模型拟合（无阈值） [73]

但是对于电子学通道设有阈值的情况，PMT 谱形的低端部分受到阈值的影响，则不会有明显的 pedestal 高斯分布，此时，标准的全模型函数则不适合用

来对带有阈值效应的 PMT 谱形进行拟合，需要根据阈值效应，对全模型进行修改。

假设在理想条件下，电子学通道的阈值可以通过 ADC 积分电荷设定并严格按照设定的阈值对信号进行取舍。则整个阈值效应可以通过一个台阶函数来进行描述。而真实情况下，是通过 DAC 的设置电压阈值，高而窄的信号脉冲能够通过阈值而被记录下来，却有着较小的积分电荷，低矮的信号脉冲无法通过电压阈值，但其积分电荷却可以比较大，根据这一事实，可以知道，电压阈值对于积分电荷的 cut 效应并非理想的阶跃函数，但是可以应用近似的误差函数来替代。将误差函数直接与全模型谱形函数相乘，得到的新函数就可以用来对带有阈值效应的 PMT 谱形进行拟合。修正全模型函数如公式3.2所示：

$$CSER(x) = \int_{-\infty}^x G(y)dy \cdot SER(x) \quad (3.2)$$

其中， $G(y) = \frac{1}{\sqrt{2\pi}\sigma} \exp(-\frac{(y-\mu)^2}{2\sigma^2})$ ， μ 是设定的 DAC 阈值转化成为 ADC 积分电荷阈值的中心值，即“等效 ADC 阈值”，而 σ 则可以认为是阈值的“漂移展宽”。

对于设置了电子学通道阈值的情况，用全模型函数和修正全模型函数拟合得到的结果，如图3.14所示。

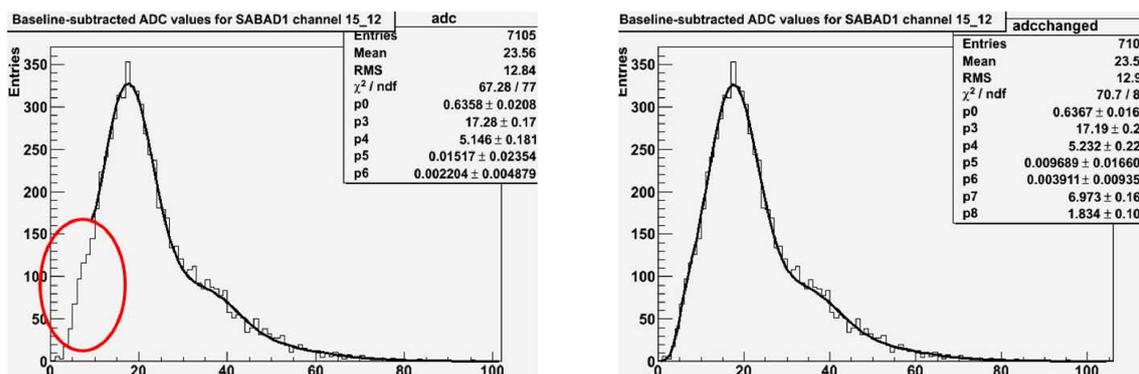


图 3.14 左图是用标准全模型函数拟合 PMT 单光电子谱，靠近阈值的低端 (红圈标识部分) 无法用全模型描述，只能截断；右图是用修正全模型函数拟合，可以进行全谱拟合，同时给出等效的 ADC 积分电荷阈值中心值及其“漂移展宽”。

一般情况下，DAC 电压阈值会设置在 650mA 或者 700mA 之间，根据修正全模型的刻度结果，转化为 ADC 单位，其 μ 值大约相当于 $\frac{1}{4} \sim \frac{1}{3}$ 光电子，

这样的阈值足以使得 Pedestal 的分布不出现在 PMT 的谱形中。当 DAC 电压阈值设置较低的时候，Pedestal 的分布会重新出现在 PMT 的单光电子谱形上，此时，利用修正全模型函数，依然可以对 PMT 的单光电子谱形进行全谱拟合。如图 3.15 所示。

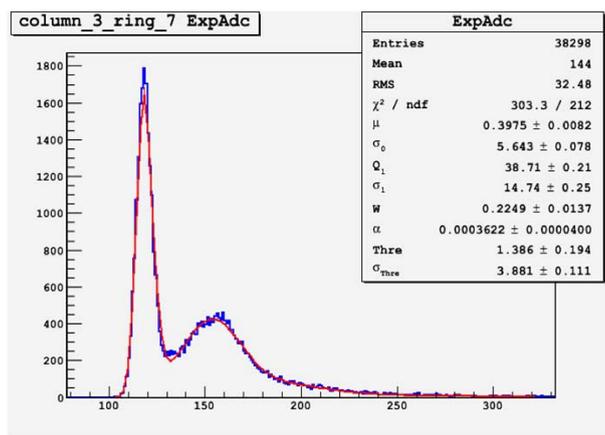


图 3.15 DAC 阈值设置较低时，Pedestal 分布重新出现在 PMT 谱形中，此时，仍然可以用修正全模型进行拟合。

利用误差函数对 PMT 全模型进行修正并拟合来自 PMT 的 ADC 谱，得到“等效 ADC 阈值”以及这个阈值的“漂移展宽”。这个“等效 ADC 阈值”的物理意义可以解释为 FEE 电子学通道的“触发效率”：当某个电子学通道接收到来自 PMT 并且其积分电荷为这个“等效 ADC 阈值”的脉冲信号，则这道电子学通道“触发”并记录此信号的概率为 50%；而此阈值的“漂移展宽”则定量地描述了随着接收到的脉冲信号的积分电荷的变化，信号得到“触发”并被记录的概率的变化情况。

类似大亚湾前段电子学板 FEE 的每个电子学通道设置了电压阈值，整个触发系统也设置了触发阈值，以反中微子探测器探测器为例，常用的触发阈值有两种 ESum(即 AD 内所有 PMT 电荷之和) 和 NHit(即 AD 中 PMT 着火多重度)。对这两种触发模式的触发效率研究发现，触发效率也并非理想的阶跃函数，如图3.16所示，受 FEE 电子学通道阈值刻度方法的启发，可以用误差函数对触发效率函数进行拟合，拟合结果显示，触发效率为 50% 的时候，其触发阈值能量 $E_{th} \sim 0.37\text{MeV}$ [57]。

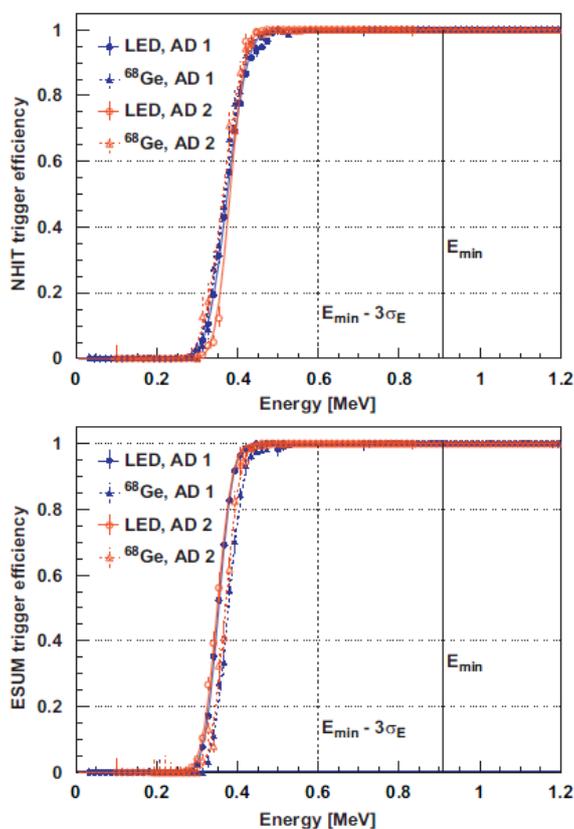


图 3.16 大亚湾反中微子探测器的触发效率随阈值的变化，上图为 NHIT 触发模式，下图为 ESum 触发模式。 E_{min} 是利用 inverse β 衰变事例的 e^+ 信号得到的重建能量的最小值。

3.5 原始数据转换器功能扩展

在离线数据分析过程中，经常要实现多种数据对象之间的相互转换。由于转换两端的数据对象格式没有改变，只是转换方向发生了变化。针对这种需求，NuWa 框架提供了一种简单的解决方案：即每一种数据转换器由转换过程两端的数据类型共同定义；并能同时实现双向转换，具体的转换方向根据用户程序的数据请求由管理数据转换器的数据转换管理服务进行控制。如图3.1所示。

在原始数据转换服务中，每个具体的数据转换器同样预留了从 TES 中定义的离线数据对象类型到二进制原始数据类型的转换接口，体现了数据转换服务的良好扩充性，可以根据需要随时对接口进行功能完善和扩展。

3.6 本章小结

原始数据转换服务在 Daya Bay 中心探测器 AD1, AD2 的 Mini dry run 和 Dry run 测试中, 积累处理了大约 10TB 的原始数据。利用原始数据转换服务, 完成了对 FEE 板电子学通道阈值和台阶的刻度工作, 对整个探测器性能更为了解。在数据处理过程中, 原始数据转换服务运行稳定, 在性能上也没有给用户程序带来严重的磁盘 I/O 瓶颈, 符合实验数据处理需要。原始数据转换服务的功能和性能, 随着 NuWa 离线软件框架的总体升级和 ROOT 工具的版本更新而不断得到扩展和增强, 而整个转换服务模块的设计理念和思路, 却得到了保留和沿袭。

第四章 Event-Mixing 软件框架

Daya Bay 反应堆中微子实验离线数据物理分析中最重要的任务，是完成对中微子混合振荡参数 $\sin^2 2\theta_{13}$ 物理灵敏度的计算。首先需要根据反中微子探测器 (AD) 中的信号时间关联特征和能谱特征，再通过反符合去掉宇宙线 μ 子事例，挑选得到 $\bar{\nu}_e$ 中微子的能谱，并估算其事例率。因此，在大亚湾反应堆实验物理数据分析中，对不同物理事例之间的时间关联分析是重要的组成部分。为了产生可用作物理分析 [74]，尤其是时间关联分析的海量蒙特卡罗模拟数据，我们在 NuWa 框架中开发了用于混合模拟数据的软件服务模块，即事例混合 (event mixing) 模块。

事例混合服务模块可以对已有的多种单一物理样本的模拟数据进行重复抽样，并按照各物理样本的事例率进行混合，得到大量带有时间关联的混合数据，用于初步分析。事例混合服务模块同时具有良好的扩展性，在增加了对多格式，多样本的并行输入接口之后，它可以完成对真实数据和 MC 模拟数据之间的混合。用于精度要求更高的 cut 效率分析研究。

本章首先简单介绍大亚湾反应堆实验事例混合的物理背景、反应堆物理数据时间结构和 NuWa 框架下模拟数据的一般特点；然后介绍事例混合服务模块的软件设计和算法设计；最后阐述为适应事例混合而对 NuWa 模拟子框架的重构和优化，并简要介绍混合真实数据的物理应用和事例混合框架的功能扩展。

4.1 MC 事例混合的物理背景

大亚湾中微子实验中的每个实验大厅，都有多个探测器模块并行运行取数。在线数据获取系统会将来自各个探测器模块的数据流按照事例触发时间进行事例排序，汇总成为前后事例具有时间关联的真实物理数据。事例间带有时间关联，是反应堆中微子数据区别于对撞物理数据的主要特点。

与真实物理数据相适应，大亚湾模拟数据的产生过程中，也需要对事例间

的时间关联加以构建。这一过程，通过“事例混合”的技术来完成。

4.1.1 大亚湾反应堆实验物理数据时间结构

4.1.1.1 第一类时间关联

大亚湾实验中，通过反 β 衰变在中心探测器中产生的快慢信号各自的特征能量和它们之间的特征时间间隔，可以对 $\bar{\nu}_e$ 中微子事例进行挑选，得到来自反应堆 $\bar{\nu}_e$ 中微子的能谱和事例率。这个过程中，快慢信号是相互关联信号，其时间间隔和中心探测器中的掺钆液闪 (Gd-LS) 中的 H 核和 Gd 核的相对比例有关 [57]。如图4.1 所示。

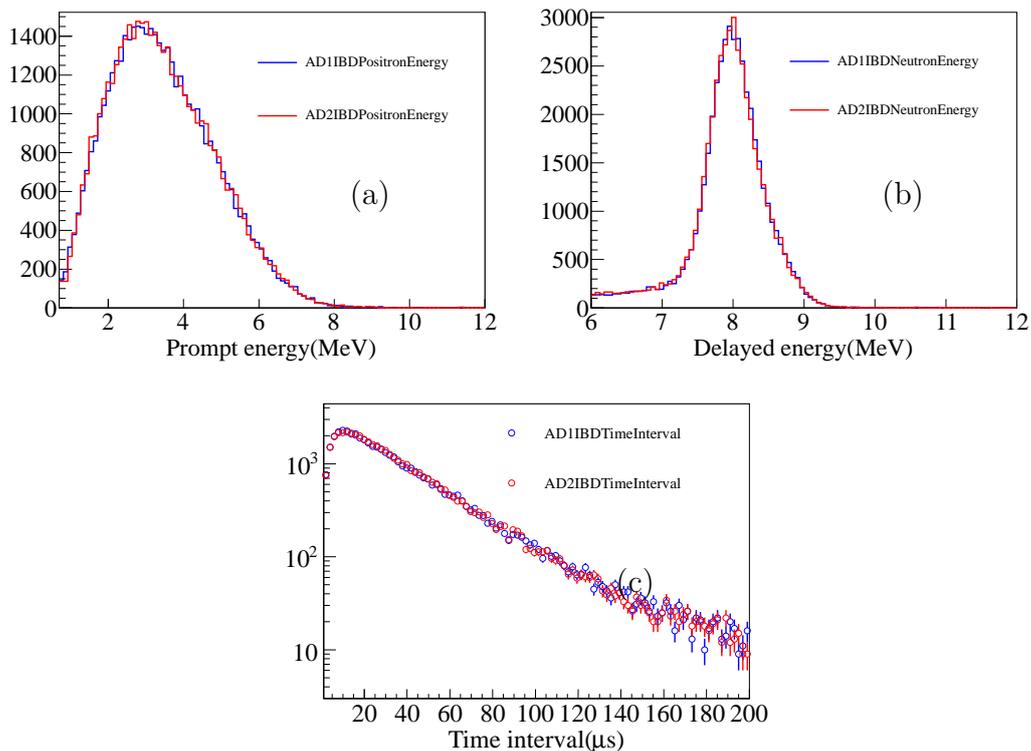


图 4.1 (a) 是大亚湾一号实验厅中心探测器 AD1, AD2 中反 β 衰变产生的快信号能谱; (b) 是中心探测器 AD1, AD2 中反 β 衰变产生的慢信号能谱; (c) 是中心探测器 AD1, AD2 中反 β 衰变产生的快慢信号时间间隔, 属于第一类时间关联

对于来自宇宙线的 muon 而言, 一个 muon 可能会引起多个探测器模块的触发, 得到多个带有时间关联的电子学读出信号, 也可能会在中心探测器模块

中引起多次触发，得到类似反 β 衰变的快慢符合信号，即关联本底。这种由同一个物理事例 (kinematic event) 引起的多个电子学读出信号之间的时间关联，称之为第一类时间关联。它和粒子在探测器中经历的物理过程和读出电子学的数字化过程紧密相连。在模拟中，由探测器模拟软件 GEANT4 和数字化模拟过程共同决定；并决定了离线数据对象模型的设计。

4.1.1.2 第二类时间关联

大亚湾实验中，来自实验大厅周围环境中的天然放射性同位素和宇宙线 μ 子，会在探测器中产生关联本底和非关联本底。对偶然符合本底的估算和和对某些关联本底，如 ${}^8\text{He}/{}^9\text{Li}$ 的研究，需要计算反应堆 $\bar{\nu}_e$ 中微子、天然放射性同位素和宇宙线 μ 子的事例率。特定的物理样本的事例率决定此物理样本中的物理事例在时间轴上的分布，这种时间关联称之为第二种时间关联，可以用样本中前后相邻的物理事例之间的时间间隔分布来描述。

对于任意特定的物理样本，可以假定样本中各物理事例之间相互独立。例如，对于天然放射性同位素，一个原子核是否发生衰变，可以认为和其他原子核的状态没有关系，因此，任意两个天然放射性事例间是相互独立的。对于其他物理过程，如反应堆放出 $\bar{\nu}_e$ 中微子和 $\bar{\nu}_e$ 中微子在探测器中的反 β 衰变过程，宇宙线 muon 在探测器模块内的物理过程，这个假设同样成立。

基于以上认识，可以证明 [75]，对事例率为 λ 的物理样本，在单位时间内出现的物理事例数目 X 的分布服从泊松分布4.1:

$$P(X = n) = \frac{\lambda^n e^{-\lambda}}{n!} \quad (4.1)$$

进一步可以证明 [76]，相邻的物理事例之间的时间间隔 t ，则服从时间常数为 τ 的指数分布4.2:

$$P(t) = \frac{1}{\tau} \exp \frac{-t}{\tau} \quad (4.2)$$

公式中的时间常数 τ 跟特定的物理样本有关，公式4.1和4.2中的事例率和时间常数并满足关系 $\lambda\tau = 1$ 。

假定有两种时间常数分别为 τ_A 和 τ_B 的物理样本 A 和 B ，按照各自的事例率均匀分布之后，混合样本 M 中相邻物理事例之间的时间间隔仍然服从指

数分布，其中的时间常数 τ_M 由如下关系式给出：

$$\frac{1}{\tau_M} = \frac{1}{\tau_A} + \frac{1}{\tau_B} \quad (4.3)$$

公式可以推广到有任意有限多种物理样本的情形4.4，其中 τ_i 是第 i 种待混合样本的时间常数；而 τ_M 和 λ_M 仍然满足倒数关系。

$$\frac{1}{\tau_M} = \sum_i \frac{1}{\tau_i} \quad (4.4)$$

4.1.1.3 大亚湾主要物理样本及其事例率

根据上一小节的结论可知，“事例混合”的关键，是根据事例率为 λ_i 的各物理样本，构建事例率为 λ_M 的混合样本，同时按照各子物理样本的事例率，决定混合样本中各子样本物理事例数的相对比例。

大亚湾实验中，待混合的子物理样本主要有以下几种：

- 探测器靶体积 (LS 和 Gd-LS) 内发生的反 β 衰变 (Inverse Beta Decay) 事例。其事例率和反应堆诸参数、探测器探测效率以及探测器对反应堆张开的立体角都有关。可以根据理论公式 4.5 计算得到 [30]

$$n_\nu(E_\nu) = \frac{\sigma_{ibd}(E_\nu) N_p \varepsilon}{4\pi R^2} \frac{W_{th}}{\langle E_f \rangle} \sum_i \alpha_i S_i(E_\nu) \quad (4.5)$$

其中，

- E_ν ：为反应堆中放出的反中微子能量；
- $\sigma_{ibd}(E_\nu)$ ：反中微子在中心探测器中发生反 β 衰变的截面；
- N_p ：探测器靶体积内的质子数；
- ε ：中心探测器对中子的探测效率；
- R ：从反应堆芯到中心探测器的距离；
- W_{th} ：反应堆的热功率；
- $\langle E_f \rangle$ ：反应堆芯中一次裂变平均吸收的能量；

- α_i : 反应堆中第 i 种燃料的相对比例, 可以通过反应堆模拟得出;
- $S_i(E_\nu)$: 反应堆中第 i 种燃料放出的反中微子的归一化能谱;

通过理论公式, 可以估算出大亚湾三个实验大厅每个中心探测器的反 β 衰变事例率, 如表 4.1 所示。

Site	λ (/day/AD)	τ (s)
DYB site	930	93
LingAo site	760	110
Far site	90	960

表 4.1 大亚湾反应堆实验三个实验大厅的反 β 衰变事例率

- 周围环境及探测器本身的各种天然放射性事例。包括:
 - 实验大厅周围岩层, 水屏蔽层, PMT 玻璃, $GdLS$ 和 LS , 中心探测器焊缝中的 $^{238}U/^{232}Th/^{40}K$;
 - 探测器钢罐及其支撑结构中的 ^{60}Co ;
 - 空气中的 $^{222}Rn/^{85}Kr$

其事例率根据对实验大厅所处环境和探测器所用材料的测量和模拟结果得出 [77], 如表 4.2 所示。

Natural radioation	λ (/day/AD)	τ (s)
in rocks	3.0×10^5	0.29
in water shield	7.3×10^5	0.12
in stainless steel vessel	1.6×10^6	0.057
in PMT glass	6.7×10^5	0.13
in Gd-LS	6.9×10^4	1.3

表 4.2 大亚湾实验大厅天然放射性事例率, $E > 1\text{MeV}$

- 宇宙线 μ 子事例。其事例率可以通过用 MUSIC(μ 子模拟软件包) 模拟 μ 子从大气中产生到地面的传输过程 [78] 和用 GEANT4 模拟 μ 子在实验大厅中各个子探测模块中的响应得到 [79]。如表所示 4.3。

Event type	DYB site $\lambda(/day)$	LingAo site $\lambda(/day)$	Far site $\lambda(/day)$
Muons	3.1×10^7	1.9×10^7	1.0×10^6
Taggable neutrons	5.0×10^2	3.3×10^2	4.7×10^1
Rock neutrons	6.5	4.4	4.2×10^{-1}
${}^8\text{He} + {}^9\text{Li}$	3.7	2.5	2.6×10^{-1}
${}^{12}\text{B} + {}^{12}\text{N}$	3.96×10^2	2.67×10^2	2.75×10^1
${}^9\text{C}$	1.66×10^1	1.12×10^1	1.15
${}^8\text{B}$	2.45×10^1	1.65×10^1	1.71
${}^8\text{Li}$	1.39×10^1	9.3	9.6×10^1

表 4.3 大亚湾实验大厅宇宙线 μ 子事例率

对于由宇宙线 μ 子产生的中子和长寿命同位素，它们和 μ 子之间的时间关联是第二类时间关联，一般将这些事例和引起它们的 μ 子作为一个完整的物理事例进行处理。这些事例的事例率由 μ 子的事例率和 μ 子在各探测器模块中的物理过程共同决定。

4.1.2 NUWA 框架蒙卡数据模拟流程和混合方式

4.1.2.1 NUWA 框架蒙卡数据模拟标准流程

Daya Bay 离线框架 NUWA 中，蒙特卡洛事例模拟主要分为探测器模拟和数字化模拟（电子学模拟，触发模拟和电子学读出模拟）两部分。前者由高能物理探测器模拟软件 GEANT4 完成，后者由 NUWA 框架下开发的数字化模拟软件包完成。通过 NUWA 框架中瞬态数据缓存 TES 实现的数据共享方式和 Python 语言的脚本配置方式，可以将两部分模拟过程相结合成为统一的完整模拟流程；并提供简单的命令行配置接口供用户使用。整个模拟流程，如图 4.2 所示。

依照图中所示的模拟流程得到的最终模拟结果存在如下特点和不足：

- 模拟样本为单一物理样本；
- 源于同一个物理事例的各个电子学读出信号之间的时间关联（第一类时间关联）已经由各电子学读出的触发时间决定；

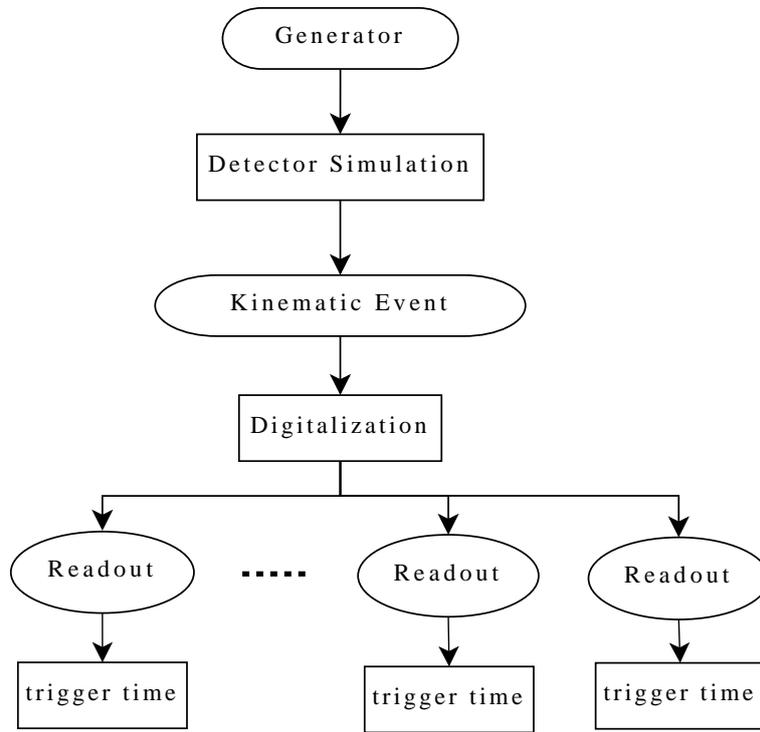


图 4.2 NUWA 框架蒙特卡洛数据模拟标准流程

- 物理事例间的绝对时间间隔（第二类时间关联）并没有意义，无法用于时间关联分析；

为了克服 NUWA 框架中，蒙卡数据模拟标准流程的不足，并产生海量混合数据，需要对模拟标准流程进行改进，完善。

4.1.2.2 产生子级事例混合和 SimHit 级事例混合

改进蒙卡模拟标准流程的一个简单想法是：按照各子物理样本的事例率，均匀混合由模拟框架读入的事例产生子，并在每个事例的产生阶段，按照混合样本总事例率，赋予其有明确意义的绝对时间，以构建第二类时间关联；并在探测器模拟结束，数字化模拟的开始阶段，考虑来不同物理事例中各个灵敏探测器的击中（SimHit）在时间轴上的重叠（overlap）效应。如图4.3所示。

使用这种模拟方式，整个模拟过程比较接近真实物理过程，尤其是考虑了不同物理事例中各个灵敏探测器击中的时间重叠效应。但是，如果要求各物理

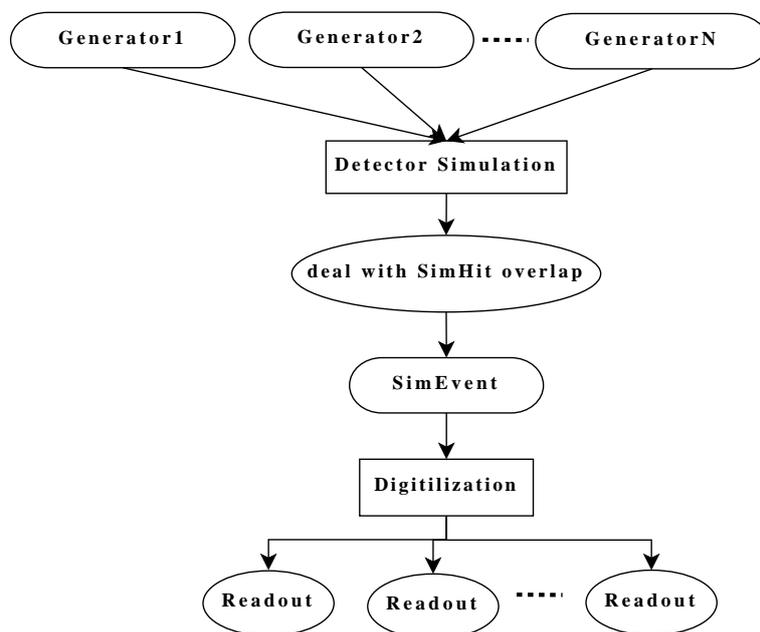


图 4.3 NUWA 框架下产生子级事例混合流程

子样本的事例率改变，重新产生混合样本，则需要按照混合—模拟流程完全重复一遍，因此，会占用大量的 CPU 计算资源和磁盘存储空间。

这种混合模拟方式的一个变体是：先完成对各类物理样本的探测器模拟，再对探测器模拟之后的物理样本进行 SimHit 级的混合，在构建第二类时间关联的同时，考虑来自各物理事例的 SimHits 的时间重叠效应，在此基础上完成数字化模拟过程，产生混合数据。利用这种方法，可以减少探测器模拟带来的 CPU 时间消耗，并能够减少对模拟数据的存储空间需求，但是，当重新产生混合数据时，仍然需要重复数字化模拟过程。模拟流程如图 4.4 所示。

4.1.2.3 电子学读出级事例混合

以上两种方法，在各物理样本的混合过程中，都需要在一定程度上重复数据模拟过程。这加重了 CPU 的运行负担，并需要更多的磁盘存储空间存放待混合样本，不利于快速产生海量数据。

根据大亚湾实验反中微子探测器的读出电子学系统的设计，来自不同的物理事例在灵敏探测器中的击中 (SimHit) 如果产生时间上的重叠 (overlap) 效应，那么在同一个反中微子探测器模块中，这些物理事例产生的电子学读出

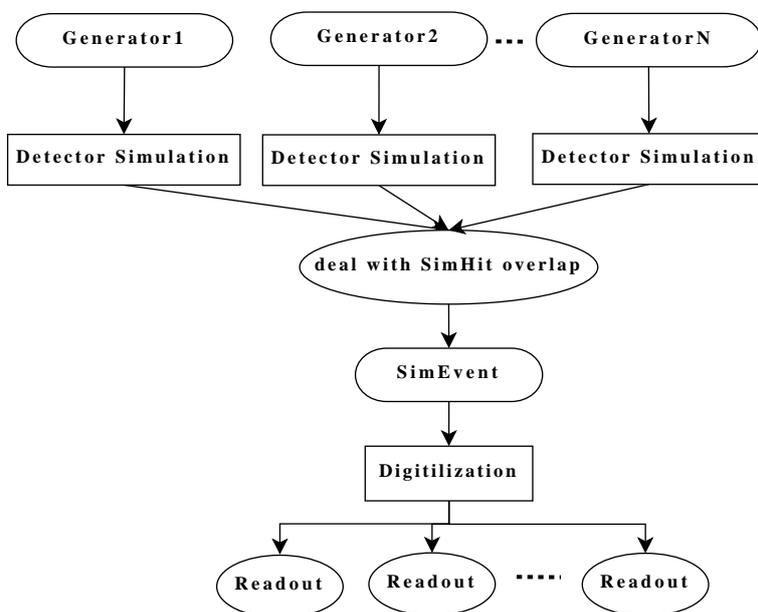


图 4.4 NUWA 框架下探测器击中级事例混合流程

信号之间的时间间隔应该小于 $1.2\mu\text{s}$ ，图4.5中显示了在各物理样本事例率为表 4.1、4.2、4.3 中所示时，其混合总样本中，在同一个反中微子探测器中，相邻电子学读出信号事例之间的时间间隔的分布状况。

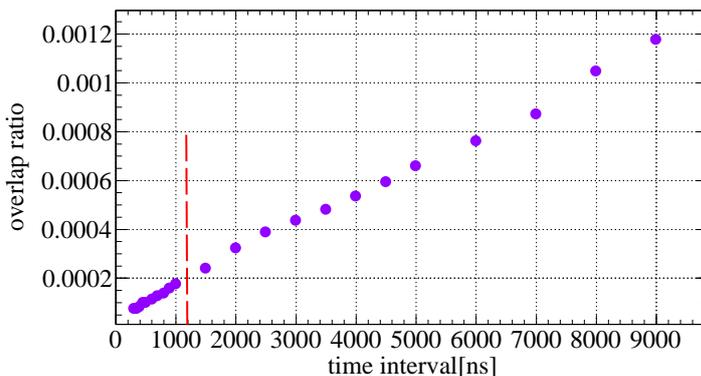


图 4.5 来自同一个 AD 的不同电子学读出之间时间间隔分布状况：横轴是相邻电子学读出之间的时间间隔，纵轴是时间间隔小于相应值的电子学读出占总电子学读出的比例

从图中可以看出，只有大约 0.02% 的电子学读出，其触发时间间隔 $t < 1.2\mu\text{s}$ 。如果不考虑探测器 SimHit 的时间重叠效应，那么在事例混合的时候，则可以先完成每一种物理样本的整个模拟过程，再将模拟得到的数据在电子学

读出级别进行混合，得到混合样本。这样，每当需要改变混合样本中的各子物理样本的相对比例，只需要在配置文件中修改各物理样本的事例率，重新混合就可以了。模拟产生的数据可以用来重复进行多次混合，从而节省大量的 CPU 时间和磁盘存储空间。这种混合过程，如图4.6所示。

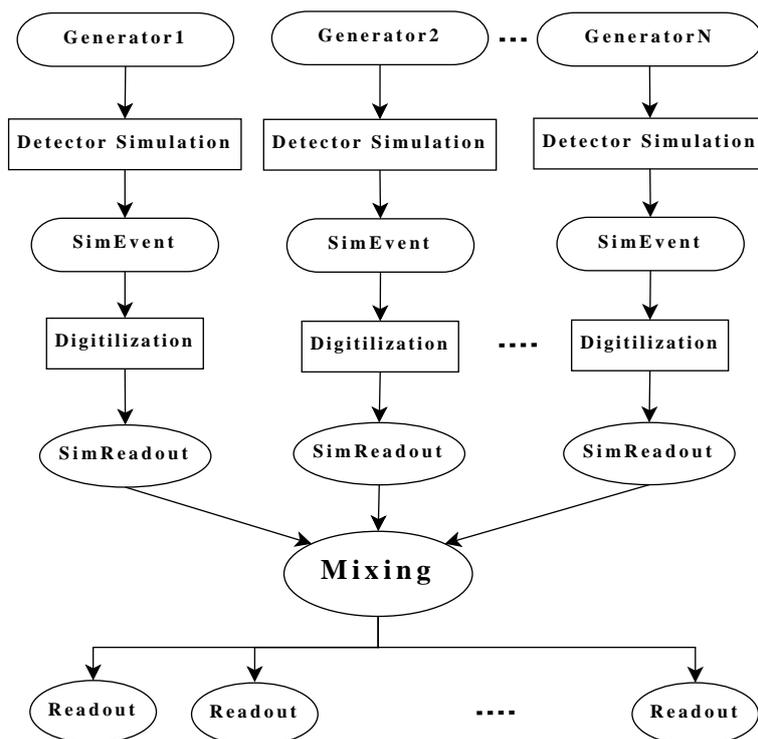


图 4.6 NUWA 框架中 Readout 电子学读出级别事例混合

在利用大亚湾离线软件框架产生海量蒙卡模拟数据的过程中，主要是采用电子学读出级别的混合方式，通过抽样混合产生了数十 TB 的 MC 模拟数据，并以其为基础，完成了最初的 $\text{inverse}\beta$ 事例挑选、Muon 反符合时间窗口研究及偶然符合本底估算等分析工作，其中，很多研究结论被用在了真实物理数据的分析过程中，并收到良好效果。

4.2 事例混合框架的软件设计

整个事例混合服务模块，是在 NUWA 离线框架下设计并实现的，遵循 NUWA 框架中的事例循环机制并利用了框架中标准的 I/O 接口模块。而且根据

混合过程中的特殊需要，对已有的 I/O 模块做了相应的功能扩展，并在混合过程内部添加了特殊的“事例回溯”机制，用于混合过程中多样本，多文件之间的事例查找。这些改进工作，使得原本只能处理串行输入数据流、进行独立事例循环的 NUWA 框架具备了处理并行输入数据流、进行关联事例循环的能力。

4.2.1 NUWA 框架事例循环机制和 I/O 模块

NUWA 框架是大亚湾实验中，以 Gaudi 框架中算法—服务模块之间的相互通信协议和统一的接口实现方式为基础开发出的离线软件框架。它沿袭了 Gaudi 框架中的“算法—数据分离”设计思想来实现数据在不同算法模块之间的共享，并保留了 Gaudi 框架中专为对撞物理设计的事例循环机制。如图4.7所示。

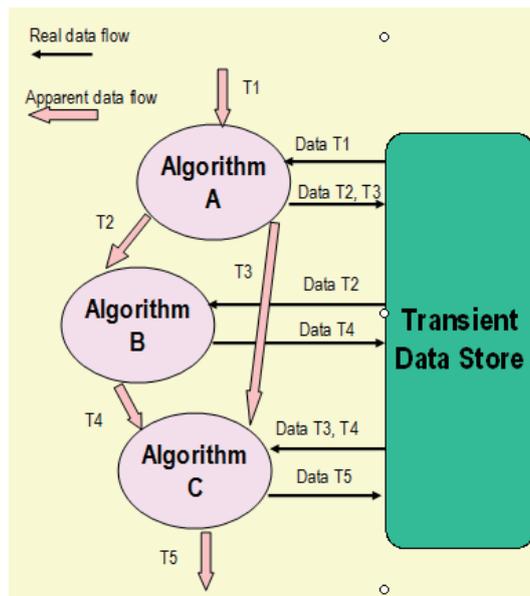


图 4.7 Gaudi 框架“算法—数据分离”设计思想实现。

图4.7中显示了框架对“算法—数据分离”理念的具体实现：

- 不同格式的数据由相应的 I/O 模块读入到瞬态缓存 TES 中；
- 数据处理算法 (Algorithms) 及其子算法 (Sub-Algorithms) 都从 TES 中取得数据，完成对数据的分析处理，并将处理得到的结果数据对象放置到 TES 中的相应位置；

- 最后仍然由相应的 I/O 模块将 TES 缓存中的数据对象持久化到硬盘上，提供给后续离线用户分析使用。

在整个循环过程中，NUWA 框架中的标准 I/O 模块，包括实现二进制数据到离线数据对象模型转换的 RawDataIO 模块和实现离线数据对象持久化的 RootIO 模块都遵循框架中的事例循环机制：

- 每一次的事例循环，最外层的算法模块的 execute() 会执行一次；
- 数据的输入输出，由最外层算法的 execute() 向 I/O 模块发出请求，并由 I/O 模块完成；
- 每次循环，I/O 模块只会输入输出一个事例，在一个循环内部，也只有这一个事例对诸算法是可见的；
- 每次循环结束后，瞬态缓存 TES 会自动清空，为下一个循环预留空间；
- I/O 模块一般只支持单个数据流的串行输入，多数据流的并行输入并没有很好的定义。

这种事例循环机制，从硬件上来说，非常适用于顺序读取的磁带机的工作原理；从物理上来说，适用于前后事例之间并无关联的对撞物理数据分析。对于前后事例之间存在着关联，并需要对关联性加以分析的反应堆物理数据，这种事例循环机制往往不能胜任工作需求，需要加以改进；而支持随机读写的磁盘也为改变这种事例循环机制提供了更多硬件上的便利。

4.2.2 多样本并行输入流接口

对于事例混合来说，NUWA 框架中标准的 I/O 模块并不能完全满足数据获取需求，原因如下：

- 事例混合中的各待混合样本事例率不同，需要加以区分，因此输入端是多样本的并行输入数据流；
- 混合算法需要对不同物理事例之间的时间关联进行重建，并在需要的时候，处理前后物理事例中电子学读出间的时间重叠；

为了满足上述第一点需求，需要对 NUWA 框架中标准的 I/O 模块进行功能扩展：首先，采用类继承的方式，可以保留标准 I/O 模块对输入样本的文件管理方式和数据流处理方式；其次，在继承标准 I/O 模块的基础上，增加了 I/O 模块对多个并行输入样本数据流的接口，I/O 模块可以同时管理多个并行的输入样本数据流，并根据需要在不同样本数据流之间进行切换，多个输入样本各自的信息，如样本名称，样本事例率等，在 I/O 模块中进行单独保存；继承得到的 I/O 模块，在读取事例的时候，会直接将存储在介质中的持久化对象读入到内存中，而不是将持久化对象转换为瞬态缓存对象，因此也不会将读入的持久化对象注册到 TES 中。

通过类继承的方式，对模块进行功能扩展，突破了框架标准 I/O 模块只能处理单一输入样本数据流的局限，也克服了通过瞬态缓存 TES 获取数据必须通过 Algorithm 模块的 `execute()` 触发的限制，为事例混合服务提供了更灵活和高速的事例获取方式，使得更多的关注，可以放到事例混合核心算法的考虑上来。

4.2.3 事例混合核心算法实现及改进

从算法角度看，对多个物理样本中的物理事例按照事例率混合得到混合样本，其本质就是一个多重数组的元素排序问题。典型的排序算法有时间复杂度 [80] 为 $O(N^2)$ 的插入排序，也有时间复杂度为 $O(N\log N)$ 的归并排序 (mergesort) 和快速排序。在混合算法中，由于物理事例的绝对时间并不是之前给定的，而是通过待混合样本的事例率，进行计算得到的，所以，利用这一特性，只需要对物理事例的时间间隔按照公式 4.2 进行抽样，并依次累加得到绝对时间，就可以处理物理事例之间的时间次序，再进一步对物理事例内部的电子学读出进行排序，就可以完成整个混合排序过程。这种对物理事例之间的电子学读出进行交叉排序的过程，类似于扑克牌的洗牌操作，称之为 interleaving 算法。通过适当的算法设计以处理电子学读出的排序问题，我们可以实现时间界小于 $O(N\log N)$ 的更快速的排序算法，以适应产生海量混合数据的需要。

4.2.3.1 事例混合并行“Interleaving”算法初始实现

由于混合算法的输入端是多个待混合物理样本，所以一种直观的想法是将算法设计成独立处理每种待混合样本的结构，即并行 interleaving 算法。这种算法利用在内存中开辟的一块缓存进行对不同样本的物理事例的交叉排序来实现事例混合，如图4.8所示。

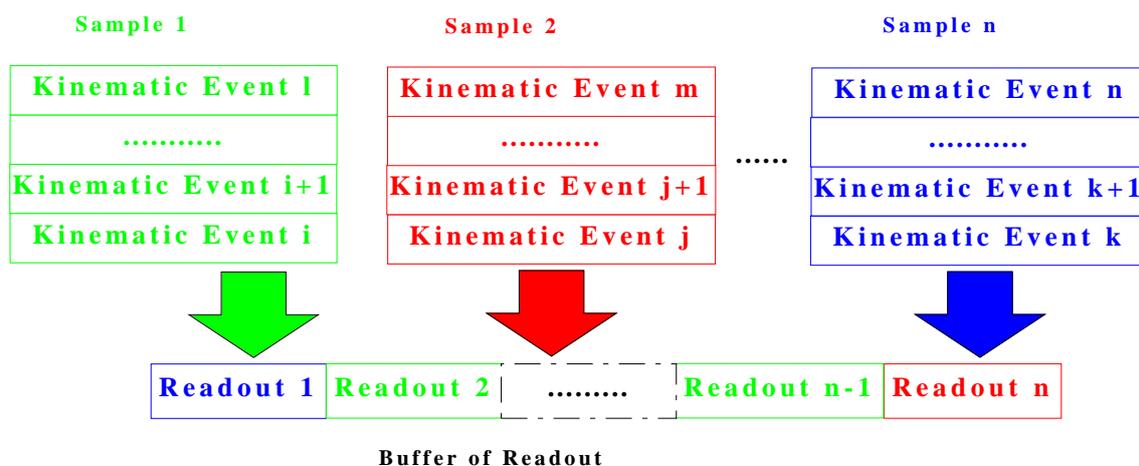


图 4.8 并行“Interleaving”算法示意图。

并行“interleaving”算法执行流程如下：

- 输入：n 种待混合物理样本，每种物理样本的时间常数 τ_i ，每种物理样本的名称；
- 执行：根据公式4.2 计算每种样本中每个物理事例的绝对时间，并在缓存中交叉排序。
 - 1. 从某物理样本中读取一个物理事例到缓存中，按照公式4.2独立计算得到此物理事例的绝对时间；解析各物理事例中的 Readout，根据计算的绝对时间重新设置其触发时间，并保持来自同一个物理事例的各 Readout 触发时间间隔不变；
 - 2. 判断缓存中是否具有来自所有样本的 Readout。如果是，则找到缓存中具有最早触发时间的 Readout；如果不是，则记录缓存中缺少的那种样本，跳回到步骤 1，从缓存中缺少的样本中读取事例。

- 输出：当前缓存中具有最早触发时间的 Readout。

并行 interleaving 算法执行流程图，如图4.9所示。

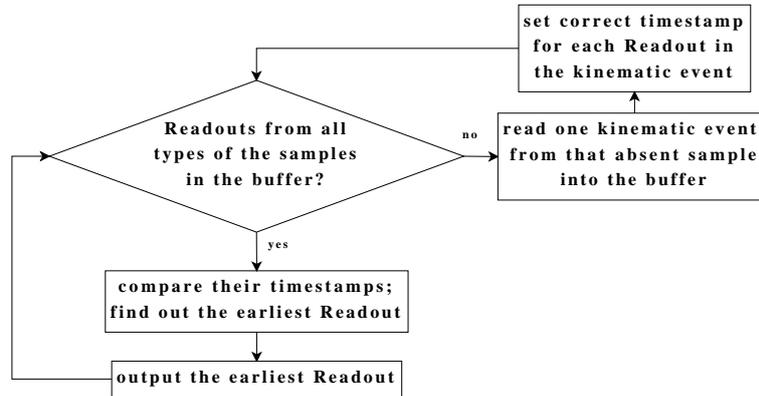


图 4.9 并行“Interleaving”算法流程图。

并行 interleaving 算法的特点在于：每种待混合样本中的物理事例的绝对时间是按照各自样本的事例率独立计算得到。并行 interleaving 算法的空间复杂度很容易估计：若待混合样本种类为 N ，则在混合过程中，缓存平均空间占用是 cN ，其中 c 是单个物理事例占用内存的平均大小，并且这个缓存占用空间随着循环次数的增加只有随机性的涨落却不会有系统性地增加，为常量 $O(1)$ 。

算法的时间复杂度可以这样估算：在每次读入一个新的物理事例之后，要按照计算得到的绝对时间，将这个物理事例插入到缓存中的合适位置，通过选择支持随机插入的容器（如 list）作为缓存，可以将这个插入时间降至常数 c ，从而整个混合过程的时间复杂度为线性 $O(N)$ 。优于多个已序数列排序中常用的归并排序（mergesort）算法的时间复杂度 $O(N\log N)$ 。

4.2.3.2 电子学读出时间重叠（time overlap）

在混合过程中，可能会遇到来自不同物理事例的多个电子学读出事例之间的时间重叠（time overlap），如图4.10所示。

图中用 R 表示一个物理事例内部的电子学读出事例 Readout。混合过程中，读入的物理事例在赋予了绝对时间之后（相邻的物理事例各自的第一个 Readout 事例的时间间隔表征了这两个物理事例之间的时间间隔， $T_2 - T_1$ ），

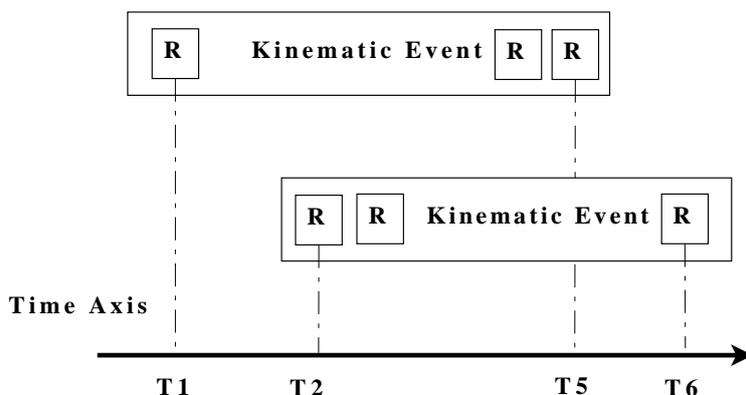


图 4.10 混合过程中，可能遇到的不同物理事例电子学读出事例之间的时间重叠。

前一个物理事例较晚的 Readout 事例的绝对时间要晚于后一个物理事例较早的 Readout 事例的绝对时间（图中的 $T2 < T5$ ）。

按照并行 interleaving 算法的设计，缓存中会对来自不同物理样本的 Readouts 排序，而来自同一种物理样本的物理事例，有且仅有一个。因此，这种算法可以处理来自不同的物理样本的物理事例之间的 Readout 事例重叠，却无法处理来自同一个物理样本的不同物理事例之间的 Readout 事例重叠。

这种电子学读出之间的重叠概率，和各个待混合样本的物理事例率有关，并且能够在一定程度上反映探测器中的物理过程：如一个放射性事例正好处于 inverse β 衰变事例的快慢信号之间，形成一个多重符合事例；或者来自宇宙线 μ 子的散裂中子 (spallation neutron) 和 inverse β 衰变的中子形成多重中子事例。因此，在混合过程中，我们希望能够仔细处理这种重叠效应，使得混合过程尽可能接近真实物理过程。

通过改进并行 “Interleaving” 算法，可以处理图4.10中所示的电子学读出事例的时间重叠效应。但是图4.10 仅仅显示了两个物理事例之间的重叠，理论上来说，存在着无限多重的重叠，即重叠存在无限多状态，仅仅依靠通过对并行 “Interleaving” 这种蛮力算法的改进，无法完成对这种无限状态的处理。

从物理角度考虑，根据各待混合样本的事例率，通过估算，可以认为超过三重的重叠状态概率极小，以至于在实际混合或者真实数据中，不会出现，从而简化重叠问题的复杂度。从软件算法角度考虑，我们则希望通过重新设计巧妙的算法来对任意的重叠状态进行处理，同时，算法本身实现尽可能简单，以

尽可能保持或者降低算法的时间和空间复杂度。

4.2.3.3 基于“有限状态自动机”的串行“Interleaving”算法实现

首先，从软件封装角度的考虑：上节给出的并行 interleaving 算法中，核心算法需要知道不同物理样本的事例率；需要区分来自不同待混物理样本的物理事例并根据相应的事例率计算其绝对时间。这造成了核心算法和 I/O 接口之间一定程度的耦合，使得希望核心算法通过统一的方式处理任意物理事例间任意程度的电子学读出事例 Readout 时间重叠的想法，遭遇到了困难。根据公式4.2和公式4.4可知，一旦在程序运行的配置脚本中，给出了各物理样本的单独的事例率，可以很容易计算出混合后的数据样本中，总事例率，即混合后样本中相邻物理事例的时间间隔分布。而“Interleaving”核心算法只需要根据这个分布进行抽样，重构相邻的物理事例之间的时间间隔，就能够完成整个事例混合工作。按照这种思路，“Interleaving”核心算法就可以首先和 I/O 接口做进一步分离，实现职责分散，降低模块间耦合性。

实现了职责分离之后，由 I/O 模块负责的部分职责有：

- 记录所有待混合样本的名称、事例率；在混合程序执行初期，根据各待混合样本的时间常数，计算混合样本的总时间常数；
- 监听并接受来自核心算法的事例请求消息，根据各待混物理样本事例率 λ_i 进行随机抽样，决定从哪一个物理样本中读入物理事例；
- 按照总事例时间常数的指数分布公式，根据总时间常数 τ_M 抽样得到总混合样本中相邻事例之间的时间间隔，并叠加在上一个物理事例的绝对时间上，得到当前物理事例的绝对时间，并将当前物理事例交给核心算法做进一步处理；

经过上述 I/O 模块的三个步骤的处理，实现了应用统一方式计算各待混合样本的物理事例的绝对时间，之前对各不同物理样本的独立并行处理，转化成不区分物理样本的串行处理。

串行 interleaving 算法的 I/O 模块通过抽样相邻物理事例时间间隔并以此累加来决定每一个物理事例的绝对时间，利用这一事实，可以将处理重叠

的无限状态转移到处理非重叠的有限状态上来。重新设计 interleaving 算法如图4.11所示。

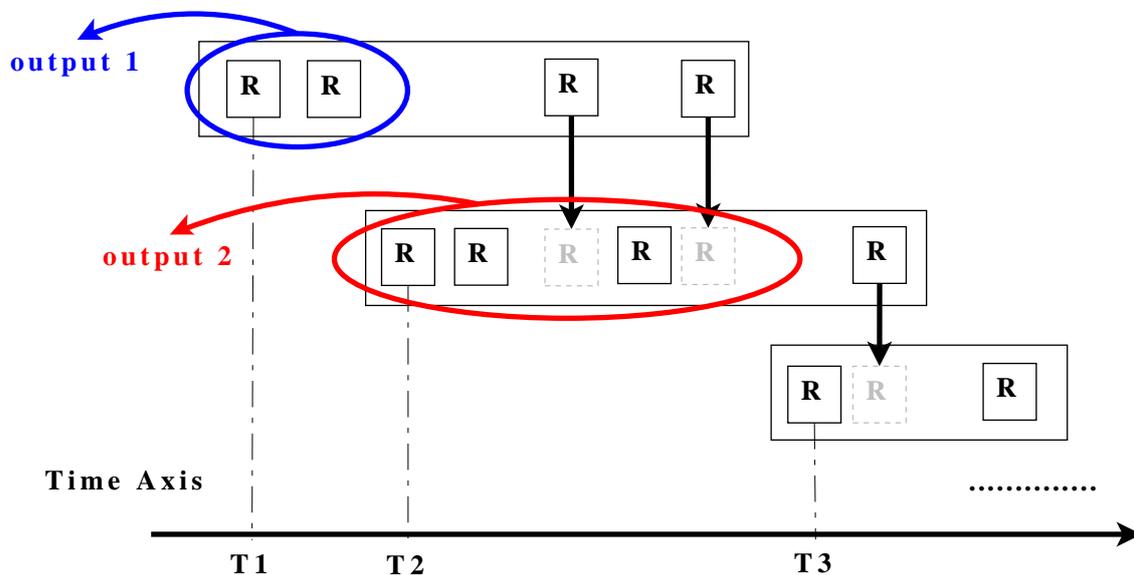


图 4.11 串行 interleaving 算法处理电子学读出事例时间重叠的示意。

核心算法的执行流程解释如下：

- 1. 将图4.11中满足 $T_1 < T < T_2$ 的所有的 Readout 事例（蓝圈）按时间先后顺序依次输出（output1），由于物理事例的绝对时间是由时间累加得到，这部分事例跟后续事例不会有重叠；
- 2. 将前后物理事例的重叠部分的 Readout 重新打包到一个”物理事例”中；如图中以 T_2 为起始时间的”物理事例”；
- 3. 发出事例请求到 I/O 模块，得到新的物理事例，并在 T_2 上进行时间间隔抽样并累加得到 T_3 ，然后重复 1、2 两步的操作（输出红圈部分，合并重叠部分）。

依照这种执行流程，串行 interleaving 算法就可以将理论上存在的无限多种重叠状态转化为两种有限的状态：1. 前后相邻物理事例之间的重叠；2. 之前物理事例重叠部分和下一个物理事例之间的重叠。并对这两种状态采用统一的处理方式加以处理。

算法最终的输出文件格式是以 Readout 为基本事例，并保留了来自同一物理事例的不同 Readout 事例时间间隔信息，与模拟生成的以物理事例为基本事例的文件格式有差异，但是与实验得到的真实数据文件格式保持一致。图4.12为串行 interleaving 算法流程。根据实际代码，增加了一个缓存区 (buffer) 保存每次得到的未重叠部分的 Readout 事例，并对其进行循环输出。

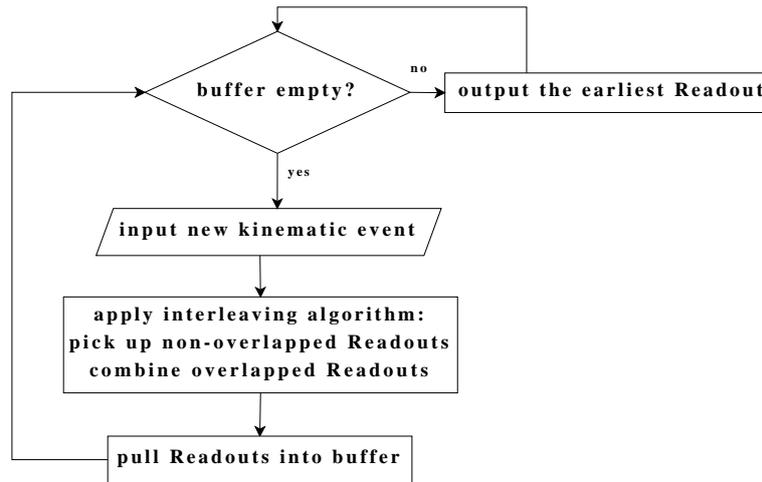


图 4.12 串行 “interleaving” 算法流程图。

有限状态自动机串行 interleaving 算法的空间复杂度很容易得知，由图可知，核心算法所占用的空间最大不超过前后相邻的多个物理事例重叠部分大小之和，实际混合中的物理事例率使得 Readout 时间重叠的概率非常小，故代码运行时，内存空间一般只存放相邻两个物理事例，且不随着循环事例数的增加而有系统性地增加，是常量 $O(1)$ 。

对算法时间复杂度一个简单的分析方法是：关注算法核心部分的一次 interleaving 操作（对相邻两个物理事例中 Readout 事例按绝对时间排序；抽取未重叠部分；合并重叠部分；将未重叠部分填充到缓存中）所耗费的时间。容易看出，一次 interleaving 操作，耗费的时间也是常量 c ，在混合海量数据过程中，核心算法部分时间复杂度是仍然是线性量 $O(N)$ 。

通过各种优化手段，比如操纵内存块的指针而非内存本身 [81]、通过设计精巧的适配器 (adapter) 实现对操纵内存块的指针进行高速排序、通过指针转换实现不同内存块之间的转化。可以将每一次 interleaving 操作耗费的时间常量大大减小。使用 BES III(北京谱仪 III) 离线物理分析软件框架中的时间测试服务

模块 BesTimerSvc 来监测，得到每次 interleaving 操作平均耗时为 $1.05\mu\text{s}$ ，标准差 $\sigma=1.34\mu\text{s}$ ，性能极佳。如图4.13所示。

```

INFO =====BestTimerSvc::Print=====
INFO updateBuffer          mean(t) :    0.910756 ms      rms :    3.60065 ms
INFO performMixing         mean(t) :    0.00105025 ms     rms :    0.00134407 ms
INFO -----
INFO updateBuffer          mean(t) :    0.910756 ms      rms :    3.60065 ms
INFO updateBuffer Execution mean(N) :
INFO performMixing         mean(t) :    0.00105025 ms     rms :    0.00134407 ms
INFO performMixing Execution mean(N) :
INFO =====BestTimerSvc::Print=====

```

图 4.13 串行“interleaving”算法中，核心算法模块效率检测结果。

4.2.4 事例回溯机制

为了减少混合核心算法的内存开销和排序过程中的 CPU 时间消耗，在读入物理事例并在内存中进行时间叠加和排序的过程中，我们只读入一个完整物理事例中的各电子学读出事例到内存中，而不读入与这个物理事例相关联的 MC 真实信息（产生子信息，探测器中径迹、顶点信息）。在输出某个电子学读出事例到 TES 的时候，程序会根据这个电子学读出事例的来源信息（物理样本类型，事例号），通知 I/O 模块读入跟这个电子学读出事例相关的 MC 真实信息到 TES，并将 MC 真实信息和这个电子学读出事例关联，然后一并输出到硬盘上的混合文件中。

NUWA 框架中的标准 I/O 模块和标准数据对象模型都没有支持这种复杂程度的数据处理方式，为此，需要进一步对 I/O 模块和标准对象模型做功能扩展，以满足对数据处理的需求并完成整个混合过程。

NUWA 框架中，标准定义的电子学读出的离线数据对象模型（Readout）并没有记录“样本类型”和“事例号”的成员变量，修改对象模型或者在标准对象模型集合中添加这一类型的子类都是破坏类型封装的不良行为。我们采用了一种称之为“适配器”的对象结构型模式 [82] 来在混合过程中复用框架中提供的标准对象模型，在保持封装完整的情况下，简化这一问题的复杂性；并通过对“适配器”对象的运算符重载，实现了对指针的高效排序，提高了混合核心算法的效率。如图4.14所示。

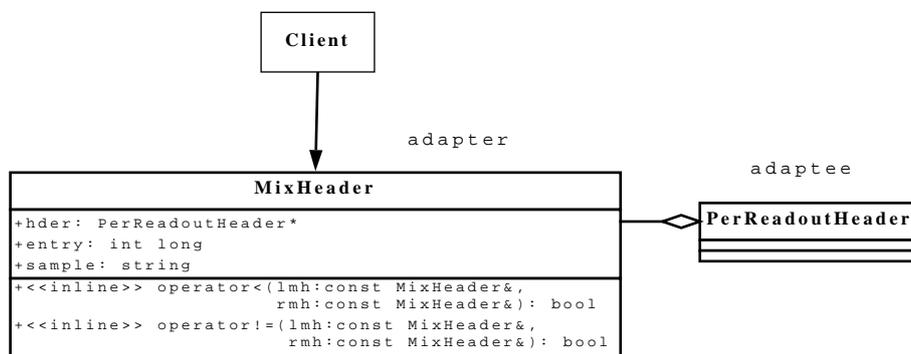


图 4.14 类 MixHeader 适配标准对象类 PerReadoutHeader。

4.2.5 基于“EvtSelector”机制的各混合模块的软件构建

如前所述，NUWA 框架中，主要通过“Service”和“Algorithm”模块来完成具体的功能模块的构建和对其他模块的调用。同时，使用“EvtSelector”模块对基于瞬态事例存储（TES）在每次事例循环中的行为控制。在事例混合模块中，由于要对待混合样本的物理事例实现控制读取和事例回溯，需要同时构建这三个模块（MixRootEvtSelector，MixRootIOcnvSvc，MixingAlgorithm）进行协调工作。如图 4.15 所示。

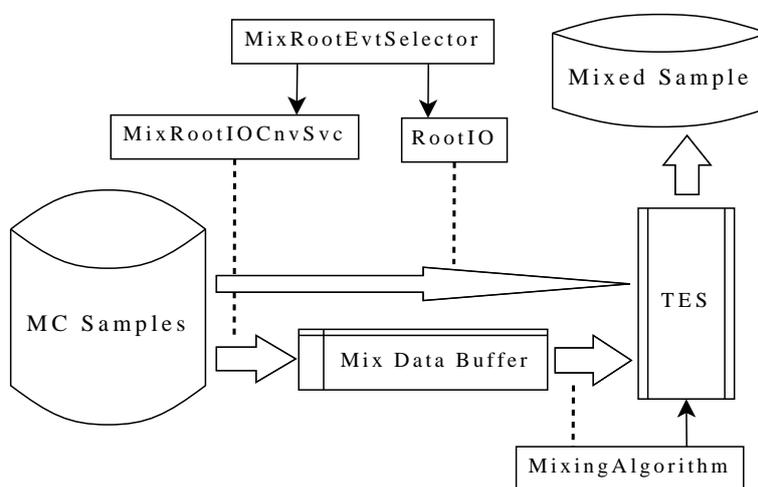


图 4.15 混合算法软件模块结构示意图。

由于抽样读入的物理事例并不是直接读入到瞬态事例缓存 TES 中，MixRootIOcnvSvc 作为 NUWA 中的标准“Service”模块，需要由 MixRootEvtSelector 控制调用。在混合程序的初始阶段，MixRootIOcnvSvc 模块读入所有待

混合样本的种类和各自的时间常数 τ_i 、计算混合样本的总时间常数 τ_M 并设置混合的绝对初始时间。

依照 NUWA 框架的事例循环机制，每一次事例循环过程中，MixRootIOcnvSvc 会按照各待混合样本的事例率进行抽样决定此次循环从哪一个样本中读取物理事例，并到相应的样本中读取一个完整的物理事例到内存中，与此同时，按照时间常数为 τ_M 为时间常数的指数分布进行时间间隔的抽样，并将抽样得到的时间间隔加到上一个物理事例的时间上，作为当前物理事例的绝对时间。

被读入进来的整个物理事例，需要被解析成为一个或者多个按照一定时间排序的电子学读出事例，放入程序内部的混合缓存中 (mixed data buffer)，并根据上节介绍的串行“Interleaving”算法的排序、合并之后，混合缓存中最早电子学读出事例将被传给 MixingAlgorithm 模块，这个最早电子学读出事例的其他标识（属于的样本名称和在这个样本中的事例号）被传给 MixRootEvtSelector。

MixRootEvtSelector 调用标准 I/O 模块，并同时保留着的电子学读出事例的标识传给标准 I/O 模块，RootIO，RootIO 根据得到的标识信息，选中与这个最早电子学事例对应的 MC-truth 信息到事例缓存 TES 中。MixingAlgorithm 模块接收到了传送过来的电子学读出事例之后，会为这个事例安装上事例头，并将 TES 中与这个电子学读出事例相关的 MC-truth 信息进行组装，然后一起输出到混合文件中。事例混合的一次基本循环就此结束。

按照 UML 规范，整个混合流程简化的交互图如图4.16所示：

4.2.6 电子学读出事例级混合算法的验证和应用

在应用 interleaving 算法大规模产生数据之前，必须先对算法的逻辑正确性加以检验。公式4.4 提供了一种简单的检验方法：分别模拟 γ 光子事例样本和 e^+ 事例样本，并将两者按照一定的比例进行混合 ($\tau_{\text{gamma}}=4.32\text{s}$, $\tau_{e^+}=2.16\text{s}$)，混合之后的样本的平均时间间隔的理论值为 $\tau_M=1.44\text{s}$ 。从混合之后的样本中，可以通过 MC 真实信息分别挑选出 γ 光子事例和 e^+ 事例并得到它们各自的相邻事例时间间隔分布，同时混合样本事例时间间隔分布可以直接得到， $\tau_M=1.40\text{s}$ 。如图4.17所示。

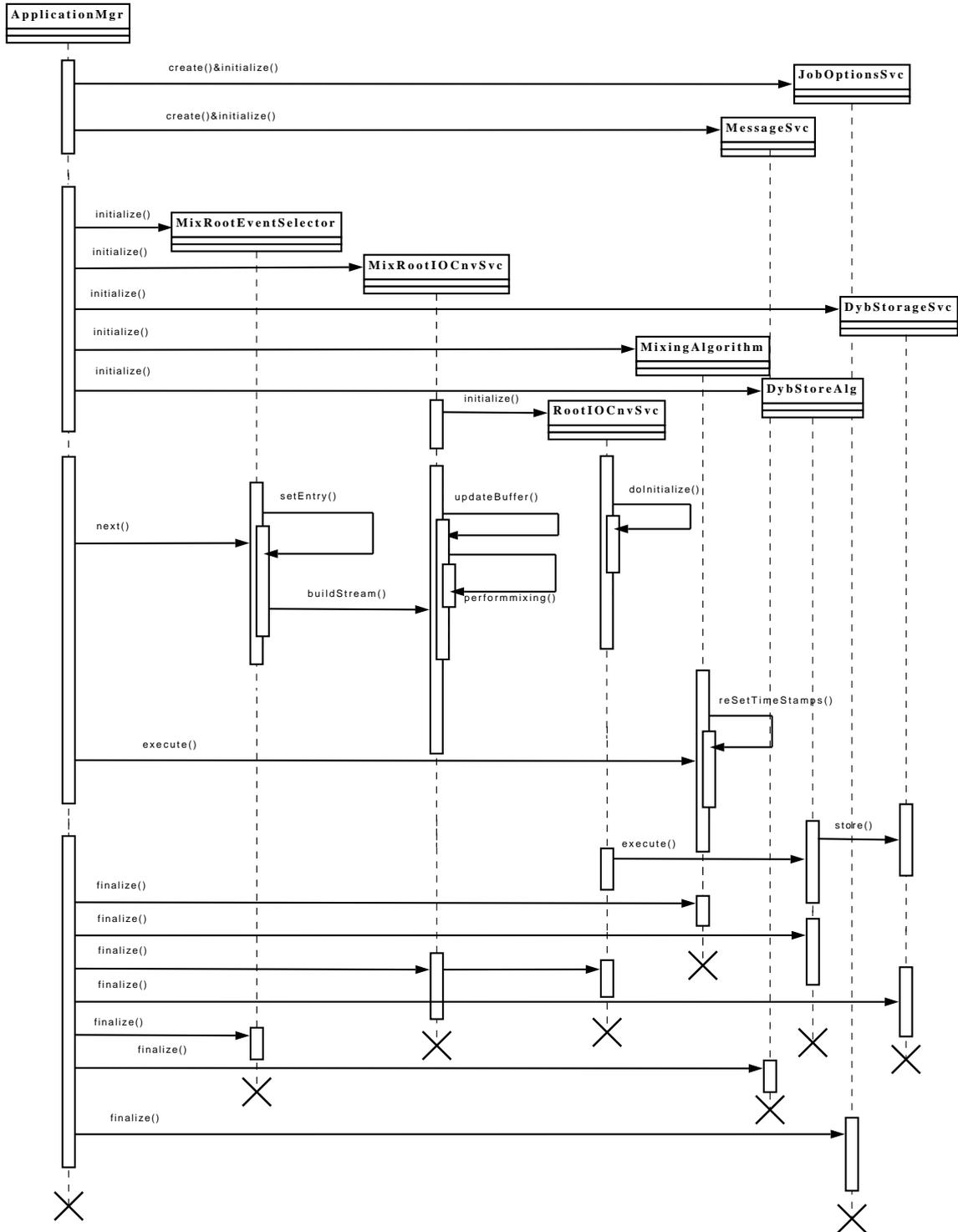


图 4.16 混合流程 UML 简化交互图。

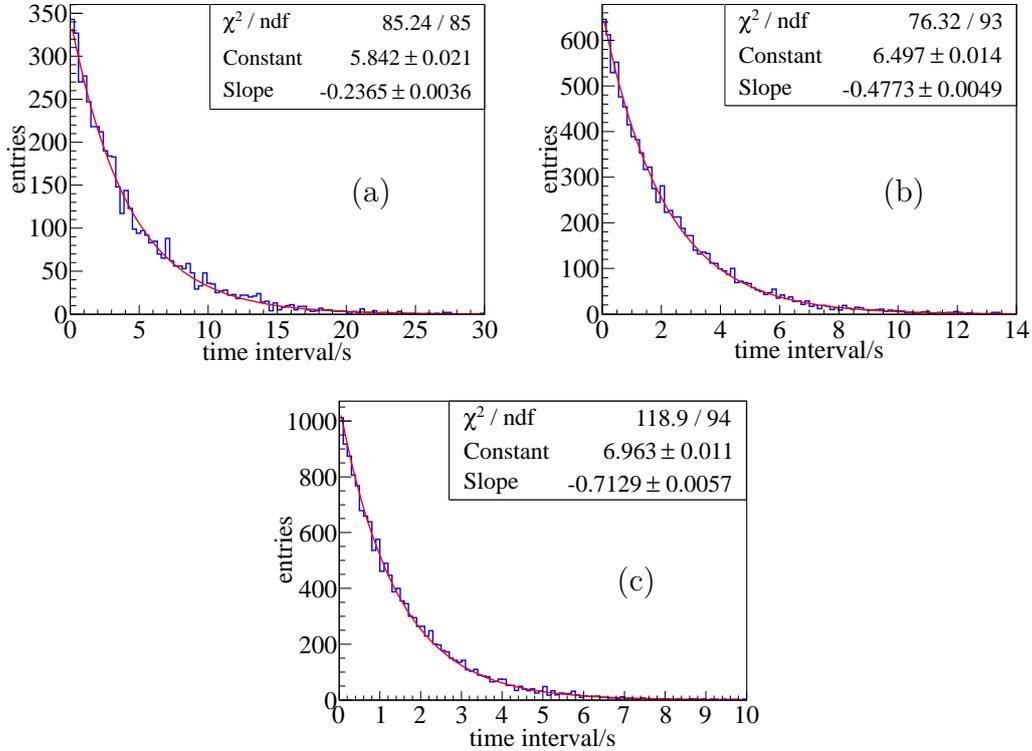


图 4.17 (a) γ 光子相邻事例时间间隔分布, 拟合得到 $\tau_{fg}=4.23\text{s}$, 预设值是 $\tau_{\gamma}=4.32\text{s}$; (b) e^+ 相邻事例时间间隔, 拟合得到 $\tau_{fp}=2.10\text{s}$, 预设值是 $\tau_{e^+}=2.16\text{s}$; (c) 混合样本中相邻事例时间间隔, 拟合得到 $\tau_{fM}=1.40\text{s}$. 理论值是 $\tau_M=1.44\text{s}$

混合算法在产生大亚湾反应堆中微子实验 MC 模拟数据产生的过程中发挥了重要作用: 自正常工作以来, 已产生了数十 TB 的物理数据并利用这些数据完成了最初的信号事例挑选, 宇宙线 μ 子反符合时间窗和本底分析工作。如图 4.18 所示, 是利用串行 interleaving 算法产生的 MC 混合数据和真实数据的 AD 信号能谱对比。

图中 8MeV 处的峰是中子在 AD 内掺钷液闪中钷核上的俘获峰。MC 混合样本根据 8MeV 处的中子俘获峰归一到真实数据; 1GeV 处的峰是来自宇宙线的 μ 子穿过 AD 产生的电离能损; 1GeV 以上的能谱来自宇宙线产生的簇射。5MeV 以下的天然放射性事例的能谱、[10MeV, 20MeV] 之间的 μ 子散裂产物 (spallation products, 如 $^{12}\text{N}/^{12}\text{B}$) 能谱和 μ 子的衰变信号的能谱, MC 数据和真实数据对比不是很好, 预示着 MC tuning 还有相当的工作需要做。此外, 从图上还可以看出, 各待混合样本 (inverse β 衰变, 宇宙线 μ 子和天然放射性事

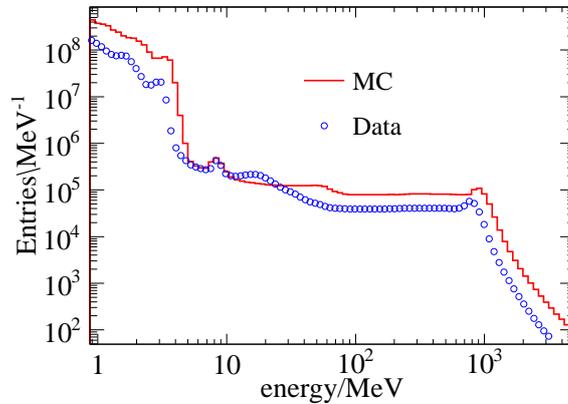


图 4.18 MC 混合数据中 AD 能谱和真实数据 AD 能谱对比。

例) 之间的相对比例, 也需要根据真实数据的研究结果, 进行细致地调整。

4.3 事例混合优化及应用扩展

事例混合服务模块, 由于其特殊的多样本流并行输入和独特的事例循环机制, 可以看做是 NUWA 框架下的功能应用子框架。它具有良好的可扩展, 可移植性, 只要以其特殊的事例循环和并行输入为基础, 对输入输出 I/O 接口和核心业务逻辑算法稍做修改, 便可以进行不同的功能扩展, 满足各种特殊的物理需要。

4.3.1 探测器 SimHit 级别事例混合

在上一节中, 混合算法主要用来对电子学读出事例的 MC 样本进行混合。这种混合方法的缺陷是: 不能够处理探测器级别中, 来自不同的物理事例的光子在灵敏探头上的击中信息 (SimHit) 在时间上的重叠效应, 尤其是考虑到这种重叠效应会引起额外触发的情况。在事例混合框架下, 可以通过读入模拟到探测器击中级别的 MC 样本并修改混合核心算法, 来进行探测器 SimHit 级别混合。与电子学读出事例 Readout 事例混合不同的是, 经过混合的探测器电荷击中 SimHits, 还需要送入到电子学模拟系统中, 进行电子模拟。所以核心算法还需要根据电子学系统的设计要求, 进行修改, 甚至电子学模拟的流程和最后的输出流程, 也需要进行全新的设计或者修改。

根据电子学读出事例的混合的经验和 NUWA 中电子学模拟系统的设计特点，探测器电荷击中 SimHit 级混合的整个框架从逻辑上分离出两部分和 NUWA 框架中五个标准模块来完成：

- 前一部分是输入多样本流和对输入物理事例间 SimHit 时间重叠效应的处理，包括两个模块：
 - MixInputSvc 输入流控制模块，负责记录输入样本流的种类、各样本的物理事例率并按照各事例率对待混合样本进行抽样读取物理事例；
 - PreElecSimSvc 电荷击中处理模块：负责按照电子学系统的设计，对读入的电荷击中 SimHit 按照时间排序并按照 $10.3\mu\text{s}$ 的时间间隙 (time gap) 重新划分成 SimHit 集合并打包。
- 后一部分是将处理之后的 SimHit 集传入电子学模拟系统，并按照混合的特殊需求，进行控制输出，也包括两个模块：
 - DigitalizeAlg 电子学模拟算法模块：负责对接收到的电子学击中 SimHit 集合进行电子学模拟、触发模拟和电子学读出事例，得到完整的模拟事例；
 - ROBufferSvc 缓存服务模块：作为整个混合框架和 NUWA 底层瞬态事例缓存的缓冲，配合 NUWA 框架的事例循环机制，进行控制循环输出；
- MainAlgorithm 主控算法模块：配合 ROBufferSvc 模块，负责在最底层进行事例循环控制，保证整个混合过程按照 NUWA 框架标准事例循环进行。

事例混合框架中，除了 MainAlgorithm 主控算法模块之外的其他模块的主要功能以及逻辑上的数据流传输，如图4.19所示。

由图4.19能够清楚地了解到逻辑上的数据流向，但是物理上真实的数据流向并不是跟逻辑的数据流向完全一致。基于以下两方面原因：

- 电子学模拟算法模块给出最终的模拟数据是以完整的物理数据格式存储的，可能并不只有一个电子学读出事例，但是按照 NUWA 框架使用瞬态

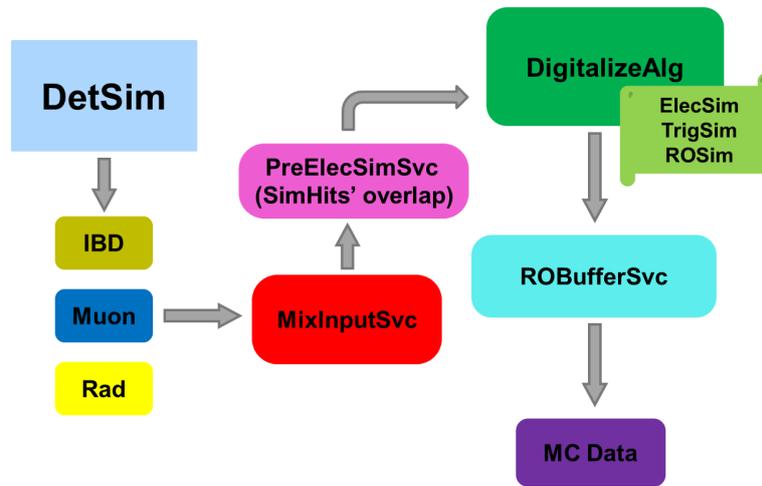


图 4.19 SimHit 级别事例混合的模块设计及数据流向，其中，灰色箭头为逻辑上的数据流向。

事例缓存 TES 进行事例循环的方式，每次循环只能输出一个电子学读出事例，其他的电子学读出事例需要缓存在 ROBufferSvc 模块内部临时开辟的缓存中；

- 对探测器电荷击中按照 $10.3\mu\text{s}$ 的时间间隙进行划分的时候，还同时需要考虑到来自不同物理事例的击中在时间上的重叠效应，而来自下一个物理事例的击中，与之前的在时间轴上排序的 SimHit 可能会形成新的时间间隔，也有可能填补之前 SimHit 已经形成了的时间间隔，考虑并统一处理这些复杂情况，解决思路有点类似电子学读出事例混合中考虑并处理来自不同物理事例的电子学读出在时间上的重叠效应，主要是判断前后若干个物理事例的 SimHit 重叠后的分布，直到找到时间上孤立分布的，并且不会跟来自下一个物理事例的 SimHit 有时间重叠的一段或者几段 SimHits 集合（这段 SimHits 最前面的 SimHit 之前 $10.3\mu\text{s}$ 之内没有 SimHit，最后的 SimHit 之后 $10.3\mu\text{s}$ 内也没有 SimHit），然后将这些时间上孤立的 SimHis 集合打包成 SimHitHeader。在此过程中，由于一次可以打包的 SimHitHeader 可能不仅仅有一个，而后续的电子学模拟系统在一次模拟过程中只能处理一个 SimHitheader，为适应后续电子学模拟的处理需求，同样需要在 PreElecSimSvc 模块中开辟相应的缓存以存放 SimHitHeader

由于有了这两块事例缓存的存在，因此，在程序模块设计的时候，是通过由位于逻辑流程最后的 ROBufferSvc 对模块内部的电子学读出事例的缓存循环输出，当缓存为空时，由 ROBufferSvc 向前一个电荷击中划分模块 PreElecSimSvc 发送出数据请求，PreElecSimSvc 同样是对内部的 SimHitHeader 的缓存进行循环输出，当缓存为空时，再由 PreElecSimSvc 向前一个多样本输入流模块 MixInputSvc 模块发送数据请求，由 MixInputSvc 模块从硬盘上读入完整的物理事例、PreElecSimSvc 负责对读入的物理事例进行解析、划分处理，最后打包发送给 DigitalizeAlg 模块进行电子学模拟并由 ROBufferSvc 进行输出。真实的物理数据流向和各模块之间的相互通讯，如图4.20所示。

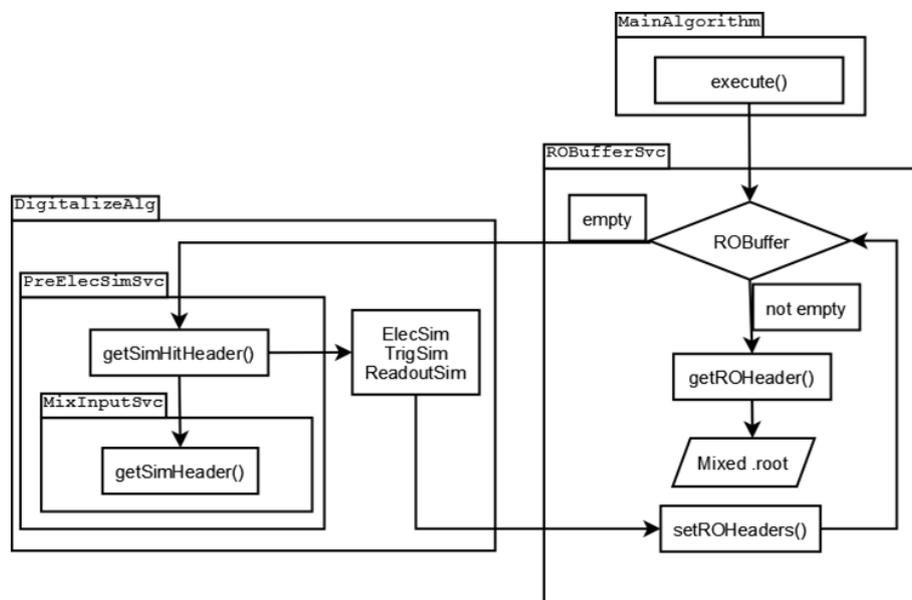


图 4.20 SimHit 级别事例混合的框架结构。

在探测器电荷击中级别进行事例混合，理论上可以尽可能接近真实探测器中发生的物理过程，但是对电荷击中进行复杂的划分，重叠和打包过程，使得模拟最终得到的电子学读出事例与其对应的 MC-truth 信息的对应，变得极其复杂，尤其是在 NUWA 这种以“单事例循环”为基础的框架中，经历了一系列复杂的处理过程，很难做到能够保证最原始的 MC-truth 信息的完整性。其次，SimHit 之间的时间重叠仅仅占到 0.02%，对于物理分析来说，不会产生显著的影响；综合考虑物理需求和效率，利用大规模产生可供物理分析的海量 MC 数据，还是以电子学读出事例的混合为主。

4.3.2 NUWA 模拟流程改进

尽管海量 MC 数据的产生，仍然是以电子学读出事例的混合为主，但是上一节中，对于探测器电荷击中 SimHit 在时间轴上的分布的深入理解，仍然给我们以很多启发：在电子学模拟之前，可以对探测器模拟生成的 SimHit 进行预先划分，再传入经过改进的电子学模拟框架，可以提高模拟的效率和成功率。

在 NUWA 框架下标准的电子学模拟过程中，由电子学模拟模块 ElecSim 接收到来自探测器的电荷击中，并以 12.5ns 为采样周期设置采样点，进而根据采样点以及电荷击中添加电子学脉冲 (pulse)，构造电子学波形。对于具有长时间结构 (10^1-10^2) 的物理事例 (inverse β 衰变事例或者是宇宙线 μ 子事例)，其电荷击中分布往往呈现分段 (trunk) 式的集中分布，绝大多数电荷击中分布在时间跨度很小的时间段内，不同段之间往往会有很长的，但没有电荷击中的时间间隔。ElecSim 模块会对这部分时间间隔也设置采样点，并根据采样点添加电子学噪声脉冲，这样形成的过长的电子学波形会占用太多内存，甚至在模拟过程中，会因为内存耗尽而导致程序不正常中断并退出。如图 4.21 所示。

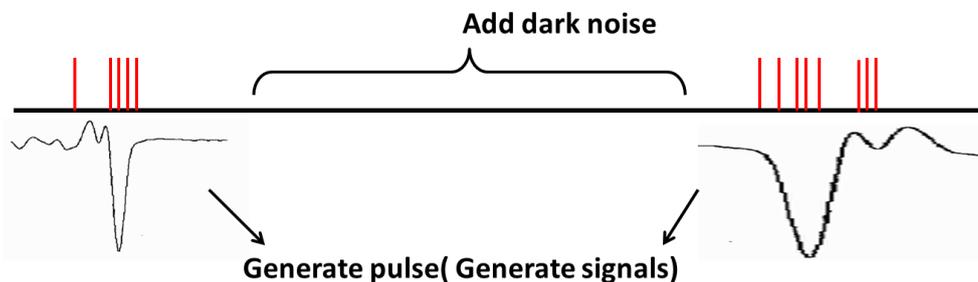


图 4.21 ElecSim 模块处理来自物理事例的 SimHit 的方式。

为了提高电子学模拟的效率和成功率，受上一节 SimHit 级事例混合的启发，我们可以在电子学模拟之前，按照电子学系统的设计，对来自同一个物理事例的 SimHit 按照时间间隔进行划分，同时记录划分之后形成的 SimHit trunk 之间的相对时间，然后仅仅将划分之后得到的 SimHit trunk 传送到下一步，进行电子学模拟，而将 SimHit trunk 之间的空白时间间隔直接丢弃。以节省内存占用和 CPU 计算资源的消耗。

根据大亚湾反中微子探测器的电子学设计，ElecSim 模块中，会按照接收到

的 SimHit trunk 中时间最早的 SimHit 为起点，向前开 300ns 的时间窗，在这段时间窗内添加 PMT 的 pre-pulse 等效应；同时以时间最晚的 SimHit 为起点，向后开 $10\mu\text{s}$ 的时间窗，在这段时间窗内添加 PMT 的 after-pulse、over shoot、ringing 等效应。因此，容易得知，两段能够独立构建波形的 SimHit trunk，它们之间的时间间隔最少要是 $10\mu\text{s} + 300\text{ns}$ ，即 $10.3\mu\text{s}$ 。在电子学处理之前，对 SimHit 的划分，即按照 $10.3\mu\text{s}$ 这个参数为依据。如图4.22所示。

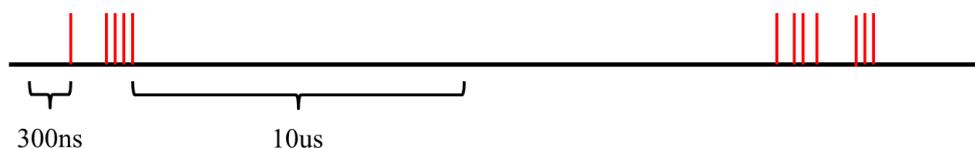


图 4.22 划分物理事例 SimHit 的方式。

对于划分 SimHit 之后，再做电子学模拟的方法，确实能够提高电子学模拟的效率，用 NUWA 开启光学过程模拟的宇宙线 μ 子事例进行测试，对宇宙线 μ 子的电子学模拟的成功率由之前的 60% 提高到 100%。但是划分之后的模拟，却增加了模拟各个阶段的信息对应的复杂程度，尤其是电子学模拟各阶段的信息对应。如图4.23所示。

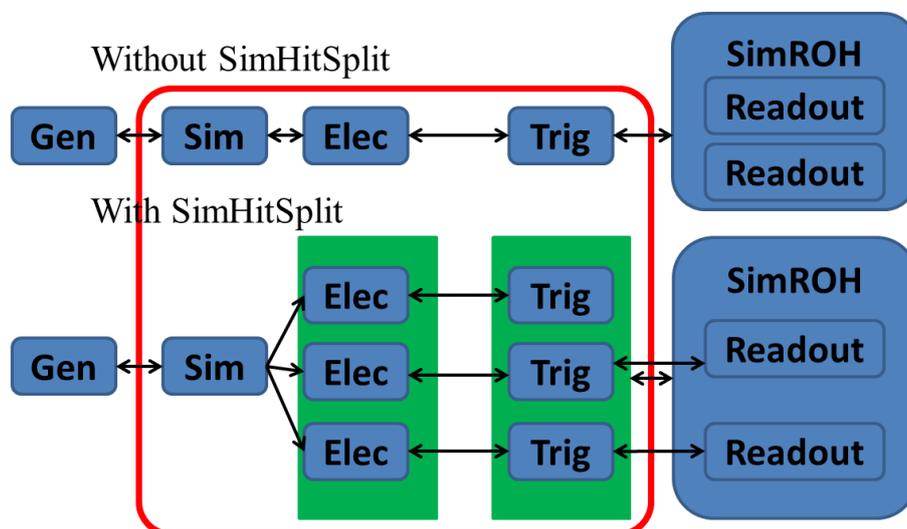


图 4.23 两种电子学模拟划分物理事例后事例关联的比较。

对于 NUWA 框架下这种“单事例循环机制”的架构体系来说，数据对象之间太过复杂的对应关系，无法由瞬态事例缓存中映射到硬盘中。所以，在对

模拟数据的存储过程中，图4.23中红圈所示部分的信息，并没有存储到硬盘上，而只是保留了模拟生成的数据和 MC 产生子信息之间一一对应的关系。真实物理数据中，并没有探测器电荷击中信息和电子学波形等信息，因此，丢弃模拟过程中的相应信息，不会对模拟数据的物理分析造成影响。

利用对 SimHit 进行划分的方法，模拟得到的结果，需要与真实数据和 NUWA 框架下标准的电子学模拟得到的数据进行对比，以保证其正确性。可以从两个方面进行比较:1. 模拟事例与真实事例的 ADC 和 TDC 分布; 2. 两种模拟方式得到的某一种物理事例的 ADC 分布。

图4.24显示了依照 SimHit 划分的方法得到的模拟事例与真实事例的 ADC 和 TDC 的分布对比图。a 图和 b 图是利用大亚湾反中微子探测器 dry run 期间，用 LED 的 double flash 事例模拟 inverse β 衰变事例的快慢信号得到的 ADC 和 TDC 的分布图，a, b, c 三幅图中，主信号都集中在 TDC 为 [900, 1000] 的区间内，并靠近 1000，说明这种模拟与真实的电子学系统符合很好。d 图中显示的是宇宙线 μ 子的 ADC 和 TDC 分布，由于进入反中微子探测器中的宇宙线 μ 子能量一般比较高，会引起比较明显的电子学台阶基线的晃动，d 图中 ADC < 500 的分布显示了这一点，TDC > 1000 的分布，是 PMT 的 after-pulse 和 ringing 等效应的体现。

图4.25则显示了 NUWA 框架中默认的电子学模拟 ElecSim 和引入了 SimHit 划分的方法的 Digitize 数字化模拟对反 β 衰变事例分别模拟的结果对比。图中显示的是位于反中微子探测器的掺钷液闪体积内的信号的 ADC 谱，位于 300 处的峰，是快信号能谱和中子在 H 核上浮获峰的叠加形成的，位于 1300 处的峰，是中子在 Gd 核上的俘获信号。从图上可以看出，两种模拟方法并没有系统性的差别。SimHit 划分的方法没有破坏电子学模拟的逻辑正确性。

4.3.3 混合真实数据

由本章第一节可知：大亚湾反应堆实验模拟数据样本主要分为三类：中微子 inverse β 衰变事例；宇宙线 μ 子事例及其引起的长寿命同位素事例和天然放射性事例。对于 MC 模拟数据来说，三种样本之中，由于天然放射性核素广泛存在于探测器的各种材料和周围环境中，天然放射性各同位素的丰度很难被精确测量，对天然放射性事例的 MC 模拟准确度不高。解决这个问题的方案之一

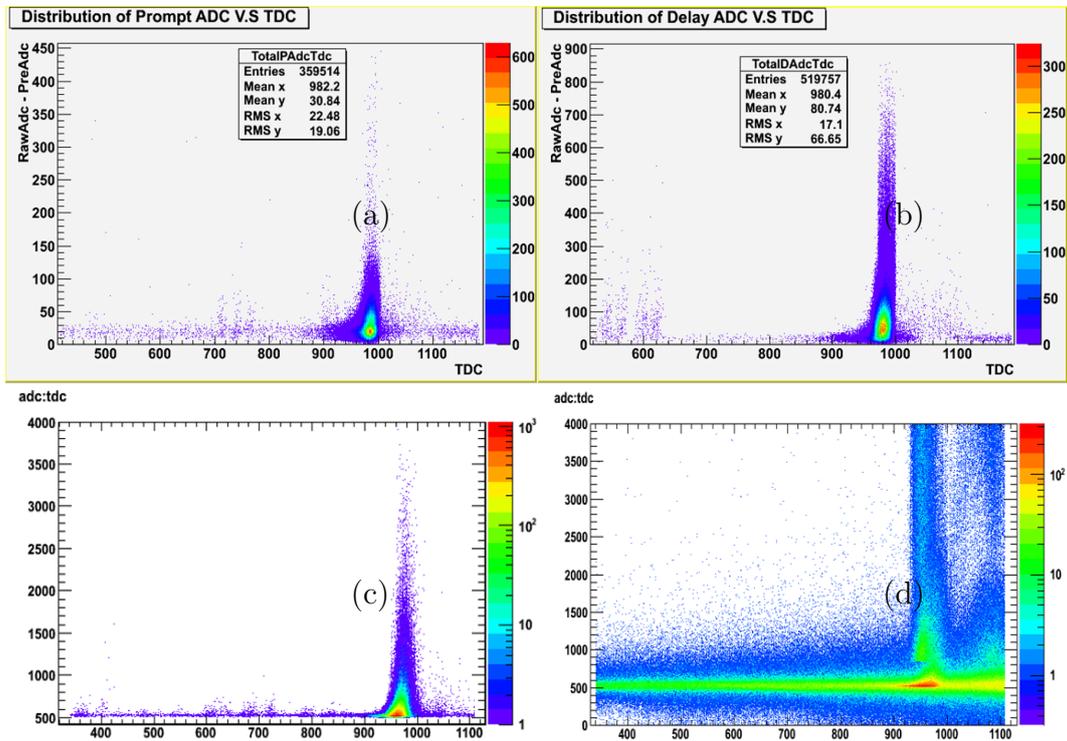


图 4.24 (a) 反中微子探测器的 dry run 期间, LED 的 double flash 的快信号的 ADC 和 TDC 分布: 横轴为 TDC, 纵轴为 ADC; (b) 反中微子探测器的 dry run 期间, LED 的 double flash 的慢信号的 ADC 和 TDC 分布, 横轴纵轴意义同 (a); (c) 按照 SimHit 先划分, 后模拟的数字化模拟方式得到的 $\text{inverse}\beta$ 衰变事例的 ADC 和 TDC 分布; (d) 按照 SimHit 先划分, 后模拟的数字化模拟方式得到的宇宙线 μ 子事例的 ADC 和 TDC 分布;

是: 通过利用大亚湾探测器系统的强制触发 (force trigger) 可以获取天然放射性事例; 再将这部分天然放射性事例和来自 MC 模拟的 $\text{inverse}\beta$ 衰变事例进行混合, 对得到的混合数据做物理分析。此外, 对于 $\text{inverse}\beta$ 事例挑选的各种 cut 条件的效率研究, 也可以通过将 MC 模拟得到的 $\text{inverse}\beta$ 事例按照一定的事例率混合到真实的物理数据中, 然后再对混合样本应用信号事例挑选的 cut 条件, 根据挑选的结果, 去估算各种 cut 条件的挑选效率。

在上述两种过程中, 都对现有的混合框架提出了新的需求——混合真实数据。真实数据是来自探测器系统和电子学触发系统的数据, 在离线软件系统中是以电子学读出事例 Readout 的形式呈现给用户, 所以对真实数据的混合, 能且仅能在电子学读出级别完成, 可以在已经完成的电子学读出级混合框架的基础上, 进行功能扩展和完善得到适合混合真实数据的混合服务模块。

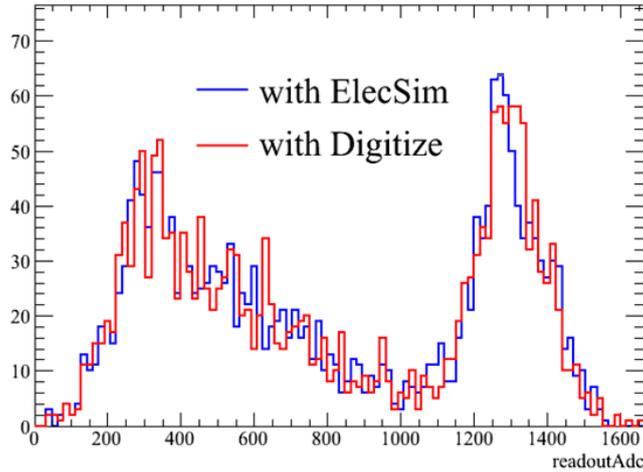


图 4.25 NUWA 框架下标准电子学模拟模块 ElecSim 与引入了 SimHit 划分方法的 Digitize 数字化模块对 $\text{inverse}\beta$ 衰变事例的模拟结果对比。

4.3.3.1 混合真实数据的架构分析与设计

如上所述，混合真实数据有两种方式：

- 1. 通过强制触发取得真实的天然放射性数据，混合到 MC 模拟的信号事例中。这些天然放射性事例的绝对时间没有意义，相互之间并没有关联，混合的时候，可以按照需要重新设置其事例率，并为每一个事例重新赋予绝对时间。这个混合过程中，所有混合样本的事例的绝对时间都要重新设置，可以称之为“PureMix”混合；
- 2. 将 MC 模拟得到的信号事例，混合进入通过物理取数得到的数据中。待混合的 MC 数据的绝对时间，在混合的时候需要重新设置，但是待混合的真实样本中的物理事例的时间有绝对意义，前后物理事例之间也有着关联，在混合过程中，真实样本的时间特性需要保留，只需要把 MC 数据按照特定的事例率插入到真实数据中，可以称之为“InsertMix”混合。

这两种混合真实数据的方式，既有共同点：从输入接口到核心算法的大致流程相同；又有不同点：核心算法和起始，结束的方式不同。要把这两种方式统一到同一个框架下，隐藏其不同的算法细节，却对用户提供统一的使用接口，

就需要使用面向对象的设计模式，对算法模块进行封装。在面向对象设计模式中，有两类行为模式可以用来实现算法模块的封装：

- 1. 对于两种不同的混合算法“PureMix”和“InsertMix”，可以用被称之为“Strategy pattern”的设计模式，进行封装，具体来说，就是定义一系列算法，把它们一个个封装起来，并且使它们在使用的过程中可以相互替换；
- 2. 而对于两种混合算法大致相同的流程架构，可以使用被称之为“Template Method”的设计模式，即定义一个操作中算法的骨架，而将一些步骤延迟到子类中。子类可以不改变算法的结构同时定义该算法的特定步骤。

利用“Strategy”模式和“Template Method”模式构建带有真实数据的混合框架 UML 类图，如图4.26所示。

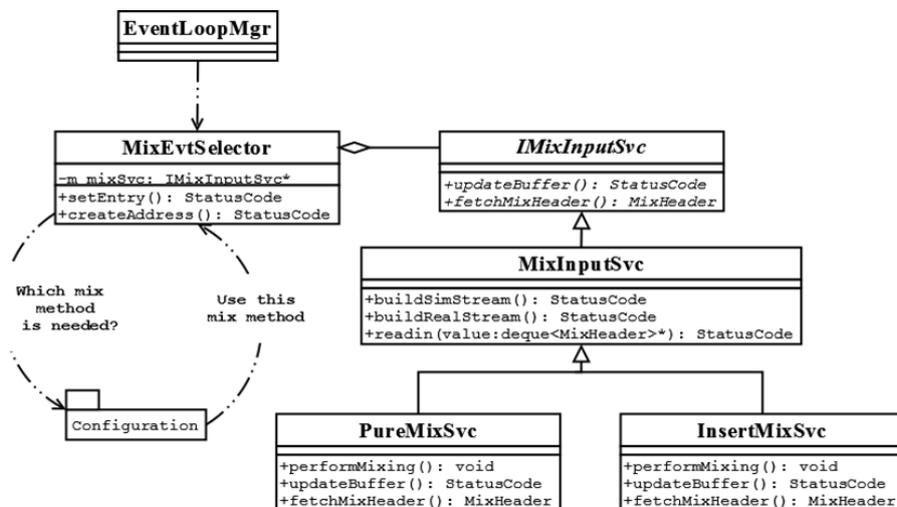


图 4.26 “Strategy”模式和“Template Method”模式合作实现带有真实数据的混合 UML 结构图，类继承部分为“Strategy”模式的应用，利用配置文件实现对不同混合算法类的配置和访问，为“Template Method”模式的应用。

在 NUWA 框架中，“Template Method”模式已经借由框架自带的“Reflex”机制实现。具体表现为：根据程序中的 string 类型给出的类名称和重要参数，自动根据类名称和参数实例化这个类的类对象；NUWA 框架中，表现对配置文件的解析以及根据解析结果实例化用户需求的类对象。

4.3.3.2 混合真实数据的算法设计

对于“PureMix”类型的混合，除了待混合样本的一部分来自于真实取得的物理数据之外，其混合算法跟之前在电子学读出事例级别混和 MC 模拟数据没有任何区别；但是对于“InsertMix”而言，由于其本质，是按照一定的事例率，向来自于真实取得的物理样本中插入 MC 模拟样本，所以，混合的算法需要做一定的改进。改进之后的算法，对于待混合的 MC 模拟样本，仍然全部按照图4.11中所示的“串行 Interleaving 算法”进行混合，对于待混合的真实样本，增加一个独立的缓存来放置从真实样本中读到的电子学读出事例并作如下处理：

- 1. 进行一次标准的“串行 Interleaving 算法”：排序和分离未重叠部分和合并重叠部分；
- 2. 对存放真实样本的电子学读出事例的缓存进行判断，如果缓存为空，则从真实样本中读入一个事例并放置到缓存中；如果缓存不为空，则对缓存进行循环处理；
- 3. 每次循环，将缓存中绝对时间最小的事例 (T_{real}) 与合并之后的那些事例中绝对时间最早 (T_2) 的进行比较，如果 $T_{real} < T_2$ ，则来自真实样本的电子学读出事例将会被放入到先前放置未重叠电子学读出事例事例的缓存中，并与缓存中的电子学读出事例进行排序，直到遇到事例满足条件 $T_{real} > T_2$ ，中断循环；并将未重叠电子学读出事例按照时间顺序依次输出到混合样本中；
- 4. 跳转到步骤 1 进行下一次的 Interleaving 过程；

改进的“InsertMix”算法，如图4.27所示

最后，有一个技巧可以进一步减少“InsertMix”算法中的内存开销和 CPU 计算消耗，考虑到在这种混合过程中，真实数据样本中的物理事例的任何属性都不需要修改（MC 模拟样本中的物理事例的时间需要根据事例率重新设置），MixInputSvc 模块在读入真实样本的事例的时候，并不是真正把一件事例的所有内容全部读入到内存中，而仅仅是在内存中构建一个“虚”的电子学读出事例 Readout(仅仅含有时间信息，样本类型和事例号，而没有相应的

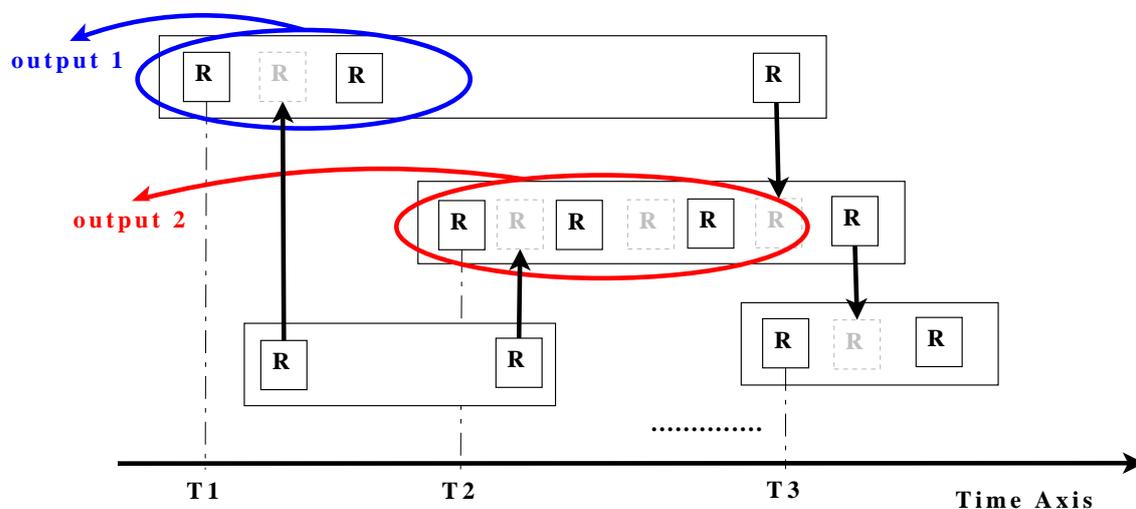


图 4.27 在“串行 Interleaving”算法基础上改进的适用于混合真实数据的“InsertMix”算法

ADC 和 TDC 信息), 在内存中的排序, 占位都用此虚 Readout 来引导。而在 MixAlgorithm 中, 根据虚 Readout 的样本和事例号信息, 会读入相应的真实事例到 TES。这么做的好处, 是减少了一次数据对象的复制传输, 提高效率。这种方法, 在计算机科学中被称为“延迟加载”(“Delay loading”)。

4.4 本章小结

数据混合服务在大亚湾探测器正式物理取数之前混合了海量 (100TB) MC 模拟数据, 并利用混合数据完成了对 inverse β 衰变信号事例的挑选, 提出了信号事例挑选中, 去除偶然符合本底的多重度 cut(multiplicity cut), 并研究了对宇宙线 μ 子事例的反符合时间窗的设置, 天然放射性事例和 μ 子引发的长寿命同位素 $^8\text{He}/^9\text{Li}$ 等本底的估算方法, 这些方法和经验, 得到了保留和优化并用于真实物理数据的分析。在数据混合过程中, 混合核心算法功能稳定, 效率极高, 在对数据混合服务模块的使用过程中, 根据读入的数据结构, 对 I/O 模块进行了修改, 使得混合效率对特定的物理样本 (宇宙线 μ 子) 提高了近一个量级, 如表4.4所示, 对所有的物理样本, 也提高了大约 2.5 倍, 符合实验数据模拟的需要。原有的 MC 数据混合框架经过功能扩展和良好的封装, 可以用来实现真实数据和 MC 模拟数据之间的混合, 在将来的物理分析工作中, 能够

待混合模拟事例类型	混合抽样类型	CPU Time(CPU efficiency)
NUWA data	顺序读取	4:08.31(99%)
NUWA data (without SimHit)	顺序读取	0:43.06(99%)
NUWA data (without SimHit and SimReadout)	顺序读取	0:21.52(99%)
G4dyb data	单文件随机抽样	0:41.23(99%)
G4dyb data	多文件随机抽样	3:49~10:00(15%~5%)

表 4.4 混合算法 CPU 效率测试，对比前三行可知道 I/O 读入是混合速度的一大瓶颈，减少读入数据的容量，可以将效率提高十倍；对比后两行可知道，混合过程中，频繁打开关闭多个文件，将会大幅度提高 CPU 的等待时间，从而降低效率。

对事例挑选 cut 的效率进行更高精度的研究。

第五章 Daya Bay 实验宇宙线 μ 子致散裂中子产额及快中子本底研究

大亚湾实验是一个地下实验，由于其高精度和低本底的要求，实验大厅需要一定厚度的岩石覆盖以屏蔽来自宇宙线的 μ 子，此外，中微子探测器外围也需要 RPC 和水切伦科夫探测器来对穿过的 μ 子进行标记和反符合。利用反中微子探测器、RPC 和内外水切伦科夫探测器，可以在大亚湾远、近点实验大厅，对不同能量（远、近点实验大厅岩石覆盖厚度不同）的 μ 子致散裂中子产额进行测量，并与理论模型和其他地下实验的测量结果 [83] 进行对比。

Daya Bay 实验中的本底来源主要有三个：快中子， $^8\text{He}/^9\text{Li}$ 和天然放射性，其中，快中子和 $^8\text{He}/^9\text{Li}$ 事例为关联本底，因为其快信号和延迟信号都来自于同一个母粒子 μ 子；天然放射性和单个的中子俘获信号则有可能在中微子信号时间窗内偶然符合，产生随机本底。对这种情况，快信号和延迟信号来自于不同的源，形成非关联本底。大亚湾的 μ 子反符合系统具有一定的探测效率，进入探测器系统却未被标识的 μ 子或者是穿过探测器系统周边岩石的 μ 子产生的快中子信号无法被 μ 子探测系统的反符合时间窗 veto 掉，需要在中微子信号事例挑选之后，估算其在挑选得到的信号事例中所占的百分比及误差，并得到其快慢信号的能谱分布。本章的分析中，反中微子探测器中触发信号的能量是 E_{rec} 是使用大亚湾的“AdSimple”重建方式得到的。

5.1 Daya Bay 反应堆中微子实验中的本底及分类

如 2.1 节中所述，大亚湾实验利用来自核反应堆的 $\bar{\nu}_e$ 中微子与靶物质中的 H 核发生 inverse β 衰变的反应，对 $\bar{\nu}_e$ 型中微子进行探测。反 β 衰变的反应释放出的 e^+ 会很快在液闪中通过电离过程沉积能量并最终和 e^- 发生湮灭，放出一对能量为 511keV γ 光子，这个过程极快，时间标度 $\sim 10\text{ns}$ ，正电子由于能量沉积和湮灭，在探测器中的可见能谱可以由反应堆中微子的能谱推导得出，在

0.7—10MeV 间 [63]; 而同时放出的中子则首先在探测器中进行一系列散射, 慢化为热中子之后被液闪中的 Gd 核俘获, 放出平均 3—4 个 γ 光子, 总能量约为 8MeV, 这个时间标度约为 $10^2 \mu\text{s}$, 与 Gd-LS 中的 H 核和 Gd 核的比值有关。inverse β 衰变和中子在钷上的俘获, 如公式 5.1 所示。由于 e^+ 湮灭和中子被俘获两个过程之间有个时间间隔, 所以测量中微子信号需要对这两个相关联的快慢信号进行符合测量, e^+ 湮灭信号被称为快信号 (prompt signal), n 俘获信号是慢信号 (delayed signal)。



因此, 理论上来说, 反中微子探测器中探测到快慢信号, 只要其能量与 inverse β 衰变的快慢信号能量有重叠, 并且快慢信号的时间间隔落在 inverse β 衰变的快慢信号时间窗内, 就有可能通过中微子信号的各种选择条件, 成为信号中的本底。在大亚湾实验中, 为使得对混合振荡参数 $\sin^2(2\theta_{13})$ 的测量精度要求达到 1%, 对于本底控制的要求是比之前的反应堆中微子实验低一个数量级, 即信号噪声比达到 0.5% 的数量级 [50], 有必要对各种本底进行仔细研究, 并通过各种挑选条件, 达到严格控制本底的目的。大亚湾中微子实验中的各种本底, 可以按照快慢信号来源划分为两类: 非关联本底和关联本底, 下面将简要介绍这两类本底。

5.1.1 非关联本底

第一类本底被称为非关联本底 [84], 它的定义为: 形成本底的快慢信号来自于不同的物理过程。非关联本底来源比较广泛, 主要由以下两种:

- 非物理来源: 主要是电子学噪声、光电倍增管打火和暗噪声, 对于电子学噪声, 可以通过对电子学通道和整个触发系统设置一定的阈值来控制, 对于光电倍增管打火, 可以通过相应的 flasher cut 来减少;
- 物理来源: 主要是天然放射性 γ 本底和 μ 子引起的散裂产物偶然符合形成。液闪、光电倍增管的玻璃、探测器钢罐及其支撑结构、探测器周围的空气和岩石中都广泛存在着天然放射性。在大亚湾实验中, 尽管通过一定

厚度的矿物油防护层和水屏蔽层来降低天然放射性事例率，但是仍然会有一些计数率的天然放射性事例在反中微子探测器中形成触发信号；同时，宇宙线 μ 子引起的散裂产物（主要是单中子）也有可能是在探测器中形成触发信号；如果这两种信号的时间间隔恰好在 inverse β 衰变事例的快慢信号时间的 $200\mu\text{s}$ 内，就形成了非关联本底，其中天然放射性信号为快信号，而 μ 子致散裂中子为慢信号。这种偶然符合形成的非关联本底是随机的，降低天然放射性和提高对 μ 子的反符合，都可以降低这类非关联本底。

5.1.2 关联本底

第二类本底被称为关联本底，其显著特点是：形成本底的快慢信号来自同一物理过程或者母粒子。与非关联本底相比，关联本底来源较为单一，主要也分成如下两种：

- 快中子本底：快中子本底来源于宇宙线 μ 子和物质发生作用产生的散裂中子，到达液闪的散裂中子会与反中微子探测器液闪中的氢核发生碰撞，反冲的氢核在探测器靶物质体积内沉积动能，如图5.1所示。

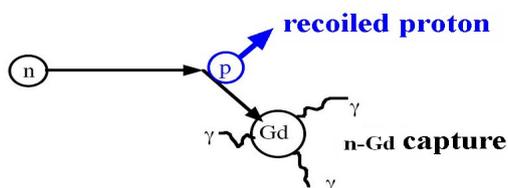


图 5.1 快中子事例形成的物理机制示意

此过程能量沉积过程较快，其时间标度大约在 10^2ns 左右，而沉积能量的能谱在 10^2MeV 内大致呈均匀分布 [79][85][86]，在 inverse β 衰变形成的快信号的能量重叠区，形成快信号；然后中子被慢化并在 Gd 上俘获，这个时间标度在几十 μs ，同样放出总能量为 8MeV 的几个 γ ，这就是慢信号。当产生散裂中子的 μ 母粒子没有被 μ 子系统反符合时间窗排除掉，则快中子的快慢信号就有可能通过 inverse β 事例的挑选条件，形成本底。快中子本底的快信号能谱的模拟和实验结果，如图5.2所示。

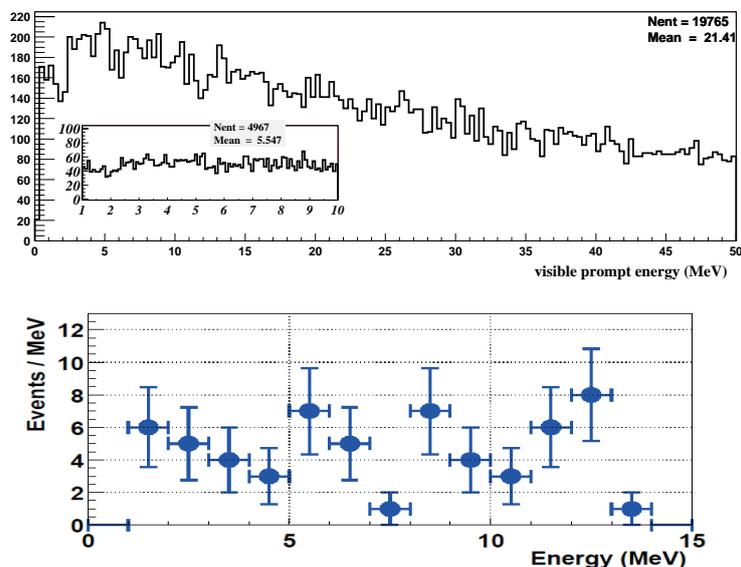
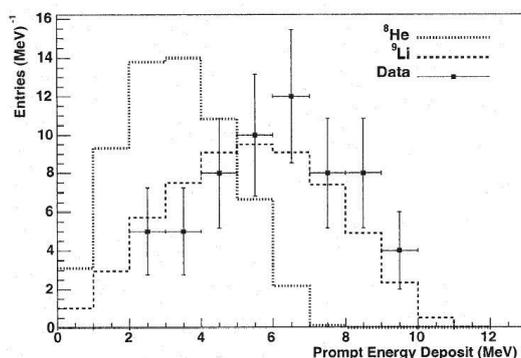
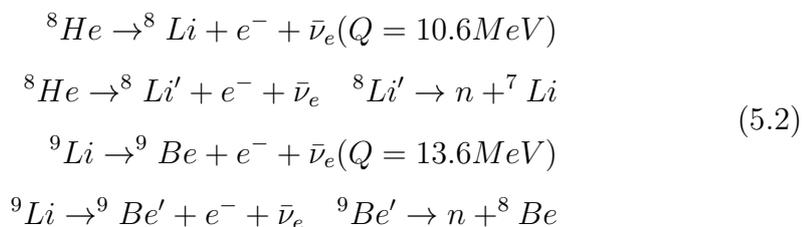


图 5.2 上图是大亚湾早期快中子 MC 模拟数据给出的快中子快信号，即反冲质子在探测器中沉积能量的分布。小图显示了反冲质子动能沉积在反 β 衰变的快信号能区内的分布，可以认为是均匀分布。下图是 KamLAND 实验测到的快中子快信号能谱图 [86]

- 宇宙线 μ 子致长寿命同位素：主要是 $^8\text{He}/^9\text{Li}$ 。宇宙线中高能量的 μ 子在通过反中微子探测器的过程中，会和探测器中的靶物质（主要是 ^{12}C 原子核）发生非弹性碰撞，并产生各种放射性同位素，如表 5.1 所示，这些同位素的半衰期从几十 ms 到几十天不等，发生衰变的模式以 β^+/β^- 衰变为主。其中以同位素 ^8He 、 ^9Li 、 ^{11}Li 最为特殊。这三者的特点是：发生 β 衰变时，伴随中子的级联发射 (β -neutron cascade)，如公式 5.2 所描述。因此，这几种放射性同位素的衰变信号在反中微子探测器中能够产生时间上有关联的快慢信号对：快信号是 β 衰变信号，其能谱如图 5.3 所示；慢信号是级联发射的中子在慢化后被 Gd 或者 H 核俘获放出的 γ ，这个过程产生的信号在中微信号能区和时间窗口内，可以形成本底。同时，这几种放射性同位素的半衰期较长，在反中微子探测器中的衰变和它们的母粒子 μ 子之间的时间间隔较远，使得它们不能通过 μ 子反符合时间窗进行反符合排除，是一种很难处理的底。


 图 5.3 ${}^8\text{He}/{}^9\text{Li}$ 同位素衰变产生的快信号能谱图 [87]


Decay mode	Isotopes	$T_{1/2}$	$E_{max}(\text{MeV})$
β^-	${}^{12}\text{B}$	0.02s	13.4(β^-)
	${}^{11}\text{Be}$	13.80s	11.5(β^-)
	${}^{11}\text{Li}$	0.09s	20.8(β^-)
	${}^9\text{Li}$	0.18s	13.6(β^-)
	${}^8\text{Li}$	0.84s	16.0(β^-)
	${}^8\text{He}$	0.12s	10.6(β^-)
	${}^6\text{He}$	0.81s	3.5(β^-)
β^+, EC	${}^{11}\text{C}$	20.38min	0.96(β^+)
	${}^{10}\text{C}$	19.03s	1.9(β^+)(+0.72MeV γ , 98.53%)
	${}^9\text{C}$	0.13s	16.0(β^+)
	${}^8\text{B}$	0.77ss	13.7(β^+)
	${}^7\text{Be}$	53.3d	0.478(β^+)(γ ,10%)

 表 5.1 宇宙线 μ 子致长寿命同位素 [88]

本章的重点是研究关联本底中的快中子本底，因为宇宙线 μ 子致长寿命同位素 ${}^8\text{He}$ 、 ${}^9\text{Li}$ 是重核粒子，只有宇宙线 μ 子穿过中心探测器并且在中心探测

器中产生这些同位素，它们才会产生本底，从位置上来说，研究起来相对简单。但在我们目前探测器的设计下，我们很可能不能通过位置的重建来去除本底，可以通过能量来判断 μ 子是否穿过液闪，采取较长的反符合时间窗口来减少同位素本底；而快中子本底相对复杂一些，宇宙线 μ 子无论在哪里，特别是在岩石中产生的快中子，都能够漂移进入反中微子探测器而形成本底。

5.2 Daya Bay 实验 μ 子致散裂中子产额估算

大亚湾实验利用反中微子探测器和 μ 子反符合系统，可以有效探测地下不同深度的 μ 子致散裂中子产额随 μ 子能量变化情况。在此之前，需要对宇宙线 μ 子产生散裂中子的物理机制做详细了解。

5.2.1 宇宙线 μ 子产生次级中子的物理过程

宇宙线 μ 子主要通过以下几个过程产生次级中子 [89]:

1. μ 子通过交换虚光子的相互作用过程使物质中的原子核解体，这个过程通常被称为： μ 子致散裂 (muon spallation)；
2. μ 子与束缚在原子核内的中子发生弹性散射；
3. 伴随 μ 子产生的光核作用 (photonuclear reaction) 过程；
4. 伴随以上过程产生次级粒子再产生的中子的过程；
5. μ 子被原子核俘获后产生次级中子的过程。

1、2 和 3 主要是高能 μ 子直接和物质强相互作用的物理过程，包含了 μ 子直接产生中子的过程；4 是 μ 子间接产生中子的物理过程。理论上对过程 2 和 3 的了解比较清楚。对于过程 1 和 4，理论上的理解存在很大的不确定性。这四种过程往往发生在 μ 子运动过程 (neutron generated by through-going muon)。对于这类中子，使用经验公式 [89] 来进行估计：

$$N_n = 4.14E_\mu^{0.74} \times 10^{-6} \quad (\mu \cdot g/cm^2) \quad (5.3)$$

$$\frac{dN}{dE_n} = A \left(\frac{e^{-7E_n}}{E_n} + (0.52 - 0.58e^{-0.0099E_\mu}) e^{-2E_n} \right) \quad (5.4)$$

$$\frac{dN}{d \cos \theta} = \frac{A}{(1 - \cos \theta)^{0.6} + 0.699E_\mu^{-0.136}} \quad (5.5)$$

其中, 公式5.3用来计算 μ 子穿过单位物质厚度的物质过程中产生的次级中子个数。中子数目与物质密度, μ 子能量, 和 μ 子径迹长度三个因素有关; 公式5.4 表示能量为 E_μ 的 μ 子产生的次级中子能量分布; 公式 5.5 表达的是这些次级中子的角分布, 只与 μ 子能量有关。

发生过程 5 的 μ 子并不是在运动过程中产生中子, 而是在探测器中停止下来并且被探测器物质中的原子核俘获并放出中子 (neutron generated by stopped muon)。低能量的 μ 子在物质中停止运动时, μ^+/μ^- 会发生不同的反应过程。 μ^+ 只能通过衰变生成 e^+ 以及相应的中微子; 而 μ^- 发生的反应则可能有两种: μ^- 有可能发生衰变反应生成 e^- 和相应的中微子, 也有可能被原子核俘获形成 μ 子原子 (muonic atom)。因此, μ^+ 和 μ^- 在物质中的寿命不再相同 (在真空中, μ^+ 和 μ^- 具有相同的寿命, $\tau_{\mu^+}/\tau_{\mu^-} = 1.00002 \pm 0.00008$ [90])。 μ^- 的寿命会比 μ^+ 的短, 同时, μ^+/μ^- 寿命之间的差别在不同的物质中也不一样, 参考资料 [91] 中列出了 μ^- 在不同物质中的寿命。

对 μ^- 而言, 可以用 $\Lambda_{capture}$ 表示 muon 在物质中发生核俘获过程的反应率, 用 Λ_{decay} 表示 muon 在物质中发生衰变过程的反应率, 相应的 Λ_{total} 表示 μ^- 在物质中发生反应的总反应率, 则存在以下的关系 [90]:

$$\Lambda_{total} = \Lambda_{capture} + \Lambda_{decay} \quad (5.6)$$

$$\Lambda_{decay} = Q\tau_{\mu^+}^{-1} \quad (5.7)$$

$$\Lambda_{total}^{-1} = \tau_{\mu^-} < \tau_{\mu^+} = 2.19703\mu s \quad (5.8)$$

式中, 其中 Q 是霍夫系数 [92](Huff factor)。霍夫系数描述如下效应: μ^- 束缚在原子轨道上, 造成 μ^- 的衰变率略低于在真空中的值, 故可知 $Q \leq 1$ 。反应式5.9和5.10 显示了 H 核和碳核俘获 μ^- 之后发生的弱相互作用过程。





μ^- 在原子核中发生核俘获反应的平均激发能量 [93] 大约是 15-20MeV, 这个能量远大于核子挣脱原子核束缚的能量阈值。 μ^- 俘获反应伴随的次级反应中, 中子发射过程的几率是很高的。 μ^- 核俘获过程具体发出多少个中子, 不同的原子核有不同的值。

表5.2中列出了反中微子探测器中, μ^- 在不同原子核上的俘获参数。其中, 核俘获过程几率是通过 μ 子寿命和核俘获率的乘积得到。由表可知, 在反中微子探测器中, μ^- 子的主要的核俘获过程发生在碳核、氧核和铁核上, 在氢核上俘获概率相对较低。

元素	μ^- 寿命 (ns)	核俘获率 (s^{-1})	核俘获过程几率 (%)	平均中子个数 (/ 反应)
C	2026.3	0.388×10^5	7.85	1
H	2194.9	0.420×10^3	0.11	1
O	1795.4	1.026×10^5	18.43	0.98
Fe	201	45.30×10^5	91.08	1.12

表 5.2 中微子探测器靶物质中, μ^- 在不同原子核上的俘获参数 [90]

对于宇宙线 μ 子产生的这两种快中子本底, 由以前的基于 geant3 模拟结果显示 [94], μ 子被原子核俘获后产生的中子本底相对于 μ 子在运动中产生的散裂中子本底非常小, 主要原因有两点: 在探测器中停留的 μ 子的只占穿过探测器的总 μ 子数的 $\sim 5\%$; 停留的 μ 子只有 μ^- 能够被原子俘获形成 μ 原子, 之后, 一部分 μ^- 子会衰变, 只有一部分 μ^- 能够发生核俘获过程, 并产生中子。模拟中仅给出了上限, 在中子总的本底贡献中可以忽略。因此之后的分析中将不再考虑 μ 子被原子核俘获后产生的中子本底的影响。

5.2.2 宇宙线 μ 子致散裂中子产额估算

5.2.2.1 μ 子致中子产额估算过程及结果

如第二章所述, 大亚湾实验采用了 8 个高 5m, 直径 5m 的反中微子探测器来探测反电子型中微子, 探测器靶物质为掺钆 $\sim 0.1\%$ 的液体闪烁体, 为屏蔽宇

宙线 μ 子和来自外部的天然放射性事例，反中微子探测器周围 4π 立体角范围内都有至少 2.5m 厚的水层屏蔽。

为了对大亚湾三个实验大厅的 μ 子致散裂中子产额进行初步估算，我们可以将掺钆液闪作为目标体积，研究宇宙线 μ 子在其中的中子产生情况，根据大亚湾实验的 μ 子反符合系统和反中微子探测器的设计，可以比较容易地对通过探测器系统的 μ 子进行标记，并挑选在 Gd 核上俘获的中子，得到穿过掺钆液闪的 μ 子个数 N_μ 以及在掺钆液闪内产生并在 Gd 核上俘获的中子个数 $N_{neutron}$ ；考虑到 μ 子致中子产额的定义 (μ 穿过物质单位长度上产生的平均中子个数) 以及事例挑选 cut 条件带来的效率 ξ ，可以得到形式上如下的公式：

$$N_{neutron} = N_\mu * X_\mu * Y_{neutron} * \xi \quad (5.11)$$

公式 (5.11) 中各参数的物理意义如下：

- $N_{neutron}$ 表示穿过掺钆液闪的 μ 子在掺钆液闪中产生的中子数；
- N_μ 表示穿过掺钆液闪的 μ 子数；
- X_μ 表示 μ 子穿越掺钆液闪的平均长度，其单位是 g/cm^2 ；
- $Y_{neutron}$ 定义为中子产额；
- ξ 代表事例挑选的 cut 条件带来的总效率；

公式5.11可以改写成为另外一种形式，以提供一种估算中子产额的简单方法：

$$Y_{neutron} = \frac{N_{neutron}}{N_\mu * X_\mu * \xi} \quad (5.12)$$

按照公式提供的思路，按照大亚湾对 μ 子标记和对中子进行挑选的标准方法，可以按照如下步骤进行中子产额公式进行进一步细化分析对 5.12的估算：

- 标记穿过探测器并发生簇射过程的 μ 子，这类 μ 子被定义成为 ADShowerMuon，按照标准 cut 条件，ADShowerMuon 的定义是：在反中微子探测器中触发信号能量 $>2.5\text{GeV}$ ，并且与 IWSMuon 或者 OWSMuon 有关联的事例；并以这类 μ 子的触发时间为起点，向后开一定时间的 veto 时

间窗口，主要是去除 ADShowerMuon 产生的次级粒子和长寿命同位素的影响；

- 标记穿过掺钆液闪的 μ 子，这类 μ 子被定义成为 ADMuon，根据大亚湾标准 cut 条件，ADMuon 的定义是：反中微子探测器中触发信号能量 $>20\text{MeV}$ ，并且与 IWSMuon 或者 OWSMuon 有关联的事例。由于估算中子产额时候用的是穿过掺钆液闪的 μ 子，所以，在按照标准 ADMuon 定义的基础上挑选得到的 μ 子个数 $N_{select\mu}$ 需要利用修正因子 $Q_{AD\mu}$ 和 $Q_{GdLS\mu}$ 估算穿过掺钆液闪体积内的 μ 子个数，如公式 5.13 所示。此外，为了避免挑选过程中对中子的重复计数，要求挑选得到的 ADMuon 样本中，相邻 μ 子之间的时间间隔大于 $200\mu\text{s}$ 。

$$N_{\mu} = N_{select\mu} * Q_{AD\mu} * Q_{GdLS\mu} \quad (5.13)$$

公式中，修正因子 $Q_{AD\mu}$ 代表穿过 AD 的 μ 子个数和按照标准定义挑选得到的 ADMuon 个数的比例，而 $Q_{GdLS\mu}$ 表示穿过 AD 的 μ 子个数和穿过掺钆液闪的 μ 子个数的比例。这两者都需要根据 Monte Carlo 模拟得出： $Q_{AD\mu} = 1.01 \pm 0.01$ (有 $\sim 1\%$ 的 ADMuon 在 AD 中的径迹比较短，沉积能量 $<20\text{MeV}$ ，故这个效率 >1)， $Q_{GdLS\mu} = 0.61 \pm 0.015$ 。

- 同样的，由于部分中子在普通液闪中或者是水中产生，修正因子 Q_{ADn} 和 Q_{GdLSn} 需要加上以得到在掺钆液闪中产生的中子个数，其中， Q_{GdLSn} 中包含了 *spill-in* 效应：即在掺钆液闪外部产生的中子，会漂移到掺钆液闪中并在钆核上俘获。如公式 5.14 所示：

$$N_{neutron} = N_{selectn} * Q_{ADn} * Q_{GdLSn} \quad (5.14)$$

修正因子 Q_{ADn} 和 Q_{GdLSn} 能够通过 MC 模拟得出，通过对 1 号实验大厅的模拟，得到结果： $Q_{ADn} = 0.90 \pm 0.01$ ， $Q_{GdLSn} = 0.952 \pm 2\%$ 。

挑选穿过掺钆液闪的 ADMuon 在掺钆液闪中产生的散裂中子。首先要求挑选的中子信号和之前最近的被标记的 ADMuon 之间的时间间隔 Δt 满足条件： $10\mu\text{s} < \Delta t < 200\mu\text{s}$ ，同时，中子俘获的能量 E 满足： $6\text{MeV} < E <$

12MeV。这两个 cut 条件引入了 cut 效率: $\xi_{\Delta t}$ 和 ξ_{6MeV} , 由于在掺钷液闪中, 仍然有部分中子会在氢核上俘获, 这一部分俘获的中子同样会引入效率: $\xi_{Gdcapture}$, 此外, 在掺钷液闪中产生的中子还能够通过 *spill-out* 效应进入普通液闪中, 这一部分中子的丢失, 通过效率 $\xi_{spill-in}$ 进行描述。故总效率可以写成:

$$\xi = \xi_{\Delta t} * \xi_{6MeV} * \xi_{Gdcapture} * \xi_{spill-out} \quad (5.15)$$

通过 MC 模拟, 可以得到上述四个效率及其误差分别为 $\xi_{\Delta t}=0.80\pm 0.01$, $\xi_{6MeV}=0.9086\pm 0.23\%$, $\xi_{Gdcapture}=0.8383\pm 0.03\%$ 。根据 G4dyb 模拟结果 [63], 掺钷液闪中的中子 *spill-out* 效应大概是 *spill-in* 效应的 24%。*spill-in* 效应的效率 $\sim 105.0\%$, *spill-in/out* 效应的相对误差 $\sim 0.02\%$ [95], 可以估算得到 $\xi_{spill-out}=0.988\pm 0.0002$ 。

挑选得到的散裂中子, 其能谱如图 5.4 所示。从图中的能谱分布和尝试拟

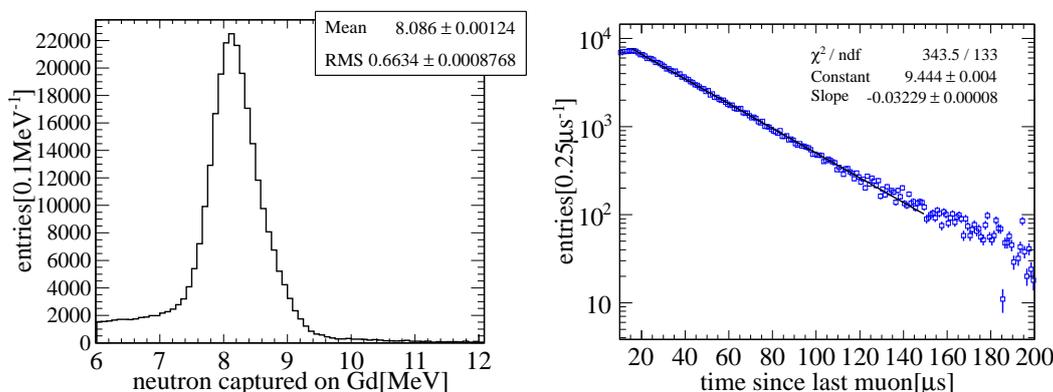


图 5.4 左上图为挑选得到的近点实验大厅 (EH1) 中的 μ 子致散裂中子在钷上的俘获能谱图, 低能端的分布主要是因为掺钷液闪边缘区域的能量泄漏引起。从图中可以看出, 挑选得到的散裂中子中非中子本底很少, 中子个数误差主要来源于其统计误差。右上图是散裂中子距离最近的 μ 子的时间间隔, 拟合结果得到时间常数为 $30.96\mu s$, 较 *inverse β* 衰变产生的初始能量较低的中子俘获时间稍长。

合可知, 散裂中子挑选结果中的非中子本底非常少, 作为散裂中子产额的初期研究, 暂时可以不考虑非中子本底; 对于其他中子本底, 主要来源是 *inverse β* 衰变产生的中子, 下文会构建一个简单的模型进行估算。

- 最后, μ 子在掺钷液闪中的平均长度 X_μ 通过 MC 模拟得到的 L_μ 和

测量得到的掺钆液闪的密度 ρ_{GdLS} 的乘积计算得到： $X_\mu = L_\mu * \rho_{GdLS} = 177.68 \pm 2.07g/cm^2$;

综上所述，计算宇宙线中子产额的各参数及其误差如表5.3所示。其中， Q_{ADn} 对于三个实验大厅，各不相同，原因在于实验取数期间，各实验大厅的反中微子探测器安装进度不同，以至于各实验大厅中通过整个探测系统的 μ 致散裂中子和的被反中微子探测器探测到的 μ 致散裂中子比例不同。

	EH1		EH2	EH3		
	AD1	AD2	AD3	AD4	AD5	AD6
$N_{select\mu}$	3606400	3601670	4135700	2359080	2371760	2356470
$N_{selectn}$	22680	23264	28019	26157	26805	26005
ρ_{GdLS}	0.86±0.01g/cm ³					
L_μ	206.6±0.04cm					
X_μ	177.68±2.07g/cm ²					
$\xi_{\Delta t}$	0.80±0.01					
ξ_{6MeV}	0.9086±0.23%					
$\xi_{Gdcapture}$	0.8383±0.03%					
$\xi_{spill-out}$	0.988±0.02%					
ξ	0.6020±0.0076					
$Q_{AD\mu}$	1.01±0.01					
$Q_{GdLS\mu}$	0.61±0.015					
Q_{ADn}	0.90±0.01		0.83±0.01		0.97±0.01	
Q_{GdLSn}	0.952±2%					

表 5.3 宇宙线中子产额计算中用到的六个反中微子探测器的参数及误差：除一，二行挑选得到的 μ 子和 neutron 个数之外，其余各参数的误差通过 MC 模拟得到

对表5.3中前两行数值的误差作如下考虑：大亚湾内外水池探测器对 μ 子的联合探测效率高达 99.7%，对于 $E_\mu > 20MeV$ ，反中微子探测器对 μ 子的探测效率接近 100%，因此，对于 μ 事例挑选而言，其主要误差是挑选的统计误差，对近点 $\sim 0.05\%$ ，对远点 $\sim 0.065\%$ 。

至于 $N_{selectn}$ 的误差，事例挑选的系统误差由表中的各 cut 条件给出，挑选得到的事例个数可以给出统计误差 $\sim 0.6\%$ 。对于本底，首先考虑其他过程产生的中子和类中子本底。inverse β 衰变引入的中子本底，可以构建如下简单模型计算。由于是在挑选得到的 μ 子之后开一个 $\sim 200\mu s$ 的时间窗，并在这个时间

窗口内部挑选中子。跟宇宙线 μ 子没有时间关联的 inverse β 衰变产生的中子，有一定的偶然符合概率落进这个时间窗内，被当做是散裂中子形成本底。较为严格的模型需要考虑每一个 μ 子之后的时间窗内落进 ibd 中子的概率，在这里给出一个简化模型：由于 ibd 事例的绝对时间是时间轴上的均匀分布，可以计算出为挑选中子而开辟的总时间窗长度，再根据 ibd 事例率近似计算得到这一段总时间窗口内的 ibd 事例个数，同时，认为其中 ibd 中子个数与 ibd 事例个数相等，由于 ibd 事例产生的中子，不一定会全部落入到这个时间窗内。以大亚湾近点实验大厅 AD1 为例，挑选散裂中子所开的时间窗为 $3606400 \times 200/10^6\text{s} = 721.28\text{s}$ ，大亚湾近点每个反中微子探测器中的 inverse β 衰变的事例数 ~ 714 个，在这段时间窗内，可能的 ibd 中子的个数为 $714 \times \frac{721.28\text{s}}{86400\text{s}} \approx 5.96$ 。对于由岩石中子引起的本底，可以用岩石中子在反中微子探测器中的事例率估算得到，岩石中子引发的探测器中的快中子和单中子的总事例率：对近点实验大厅 $\sim 13.7/\text{day}$ ，对远点大厅 $\sim 1.1/\text{day}$ ，可以知道，这部分本底很少，可以忽略不计。对于类中子本底，主要是由两部分：和宇宙线 μ 子相关联的长寿命同位素事例，以及和宇宙线无关的 AmC 刻度源中子。后者由于和宇宙线无关，可以通过其事例中和挑选中子的时间窗的乘积来做近似估算，由大亚湾的初步结果，近点的 Am-C 中子事例率 $\sim 260/\text{day}$ ，远点的 Am-C 事例率 $\sim 200/\text{day}$ ，可以计算得到这一部分本底。对于 $^{13}\text{C}^{16}\text{O}$ 引发的关联本底，可以通过其和宇宙线 μ 子在 $200\mu\text{s}$ 内的偶然符合计数率得到，近点的宇宙线 μ 子 $\sim 200\text{Hz}$ ，远点的宇宙线 μ 子 $\sim 20\text{Hz}$ ， $^{13}\text{C}^{16}\text{O}$ 事例率 $\sim 12.87\text{Hz}$ ，可以计算得到其本底事例个数。对于和宇宙线 μ 子相关联的长寿命同位素，可以认为，挑选散裂中子的时间窗仅仅局限在探测器 μ 子后面 $200\mu\text{s}$ ，这些时间段相对于长寿命同位素的衰变时间很短，总效应相当于在这段时间段内，长寿命同位素可以近似认为均匀分布，由大亚湾的本底研究结果，可以知道三个实验大厅的类中子长寿命同位素的事例率，进而计算出散裂中子中的本底，如表 5.4 所示：

对于非中子本底，最主要的贡献来自于宇宙线 μ 子在掺钷液闪中被原子核俘获形成的“stop muon”，然后衰变产生的 Michel 电子形成的本底，根据对 Michel 电子能谱的研究结果 [96]，在图 5.4 中没有看到明显的 Michel 电子能谱，当前的研究工作，对这部分本底可以暂时忽略不计。

	EH1		EH2		EH3	
	AD1	AD2	AD3	AD4	AD5	AD6
$N_{select\mu}$	3606400	3601670	4135700	2359080	2371760	2356470
$\sigma_{select\mu}$	1803.2	1800.8	2067.9	1533.4	1541.6	1531.7
Time Windows(s)	721.28	720.33	827.14	471.82	474.35	471.29
σ_{TW} (s)	0.36	0.36	0.41	0.39	0.38	0.38
N_{ibd}	5.96	5.98	5.10	0.49	0.49	0.49
$\sigma_{N_{ibd}}$	0.003	0.003	0.003	3.2×10^{-4}	3.2×10^{-4}	3.2×10^{-4}
$Rate_{isotopoes}(/day/AD)$	745 \pm 27.7		541 \pm 19.1		57 \pm 6.7	
$N_{isotopoes}$	6.22	6.21	5.17	0.31	0.31	0.31
$Rate_{Am-C}(/day/AD)$	262	268	266	208	201	192
N_{Am-C}	2.17	2.17	2.19	1.08	1.08	1.08
$Rate_{^{13}C^{16}O}(Hz/AD)$				12.87		
$N_{^{13}C^{16}O}$	371.6	371.6	371.6	24.2	24.2	24.2

表 5.4 挑选得到的散裂中子样本中的本底估算

5.2.2.2 结果及讨论

通过计算,可以得到表格5.5 所示的大亚湾三个实验大厅对 μ 子致散裂中子产额的最终估算结果。

实验大厅	μ 子平均能量	散裂中子产额 ($/(\mu \cdot g \cdot cm^{-2})$)
EH1	57GeV	$Y_{nAD1} = 8.148 \pm 0.323 \times 10^{-5}$
		$Y_{nAD2} = 8.332 \pm 0.331 \times 10^{-5}$
		$Y_{EH1} = 8.240 \pm 0.231 \times 10^{-5}$
EH2	58GeV	$Y_{EH2} = 8.128 \pm 0.323 \times 10^{-5}$
		$Y_{nAD1} = 1.543 \pm 0.061 \times 10^{-4}$
EH3	126.5GeV	$Y_{nAD2} = 1.566 \pm 0.062 \times 10^{-4}$
		$Y_{nAD3} = 1.534 \pm 0.061 \times 10^{-4}$
		$Y_{EH3} = 1.547 \pm 0.035 \times 10^{-4}$

表 5.5 计算得到的大亚湾三个实验大厅, 6 个反中微子探测器掺钆液闪中的 μ 子致散裂中子产额, 误差由表5.3中的各项误差进行误差传递得到。

大亚湾三个实验站点测得的宇宙线 μ 子致散裂中子产额, 与之前其他地下实验的结果和唯象模型的模拟结果的对比, 如图 5.5所示:

由图5.5中结果可以得到如下结论: 在不考虑簇射 μ 子 ("shower muon") 的

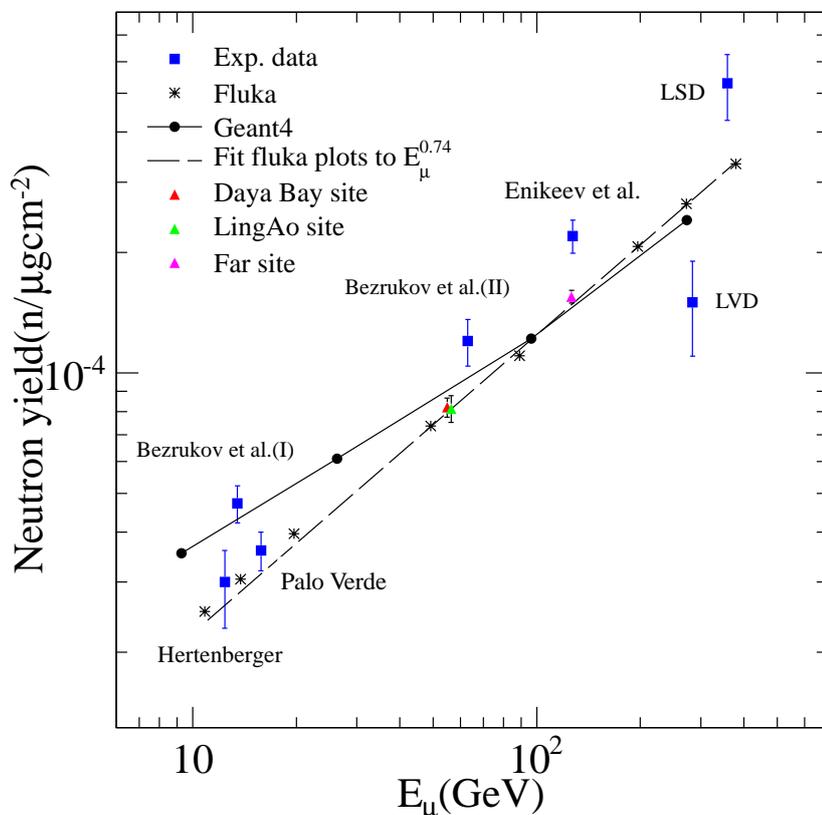


图 5.5 图中显示了大亚湾三个实验大厅测得的 μ 子致散裂中子产额与之前其他地下实验以及模拟结果的对比 (Hertenberger[97], Bezrukov et al.(I)[98], Palo Verde[99], Bezrukov et al.(II)[98], Enikeev et al.[100], LVD[101], LSD[102] 和 Geant4[103] 以及 Fluka[89] 的模拟结果)。大亚湾的结果测得的结果, 误差很小, 图中的误差只是近似给出了三个大厅之间的相对比例。

情况下, 宇宙线 μ 子引发的散裂中子产额与 μ 子平均能量之间的近似满足 $E_{\mu}^{0.74}$ 的关系, 文献 [97] 的结果和 fluka 模拟结果的比照也能够支持这一点。”shower muon” 会引入比较复杂的本底, 给这一部分 μ 子的散裂中子的挑选带来了相当大的不确定度, 影响实验测量结果的精度。

从图上也能够看出, 大亚湾实验的误差较之以前的地下实验较小, 这可以从下面几个原因上加以分析讨论:

- 大亚湾实验的反中微子探测器具有较大的靶质量, 使用的掺钷液闪, 缩短了中子的俘获时间并提高了中子俘获信号的能量, 进一步提高了对中子的

挑选效率，降低了中子挑选条件的系统误差；在每个实验大厅，设置多个反中微子探测器独立测量，以更好地消除系统误差；

- 大亚湾反中微子探测器的电子学系统响应速度快，没有在探测器中发生簇射过程的 μ 子所产生的多个中子，都能够被很好的记录，而以往的实验，则需要分别考虑单中子，双中子甚至甚至是多中子的情况，需要考虑多中子情形下中子丢失的效率（使用 $\epsilon_{i,j}$ 来标识每一种情况）以及多中子引起的探测器 DAQ 活时间的损失（ ϵ_i^{daq} ）情况，加大了中子计数的不确定度，有较大的系统误差，而大亚湾实验中，需要考虑的中子计数相关的参数主要是各种挑选条件的效率及其误差，很容易通过 MC 模拟得到，精度相对更高。
- 大亚湾的 μ 子系统，对于宇宙线 μ 子的标记效率高达 99.7%，由于计算中子产额的时候，用的是穿过反中微子探测器的 μ 子，在能量满足 $E > 20\text{MeV}$ 的情况下， μ 子反符合系统和反中微子探测器对 μ 子的联合探测效率可以近似认为是 100%，对于能量低于 20MeV 的 μ 子，通过 MC 模拟得到的修正因子，可以对这部分事例做出合理的修正，最终挑选得到的 μ 子样本非常干净，可以认为主要误差来源于统计误差，并且由于 μ 子样本的大统计量，因而统计误差较小，这也是最终结果误差较小的主要原因之一。

大亚湾实验中的中子产额估算，在未来还能够作如下研究：

- 考虑“shower muon”引起的中子的产额：尽管在反中微子探测器中发生簇射过程的 μ 子只占总数的一小部分，但是很可能会对中子产额有着比较明显的影响，簇射 μ 子，产生的长寿命同位素产物也能够产生次级中子，从而对中子产额有贡献，但是，从真实数据中得到较为准确的次级中子信号是研究过程中的需要解决的问题

5.3 线性外推法估算快中子本底

如本章第一节中介绍，大亚湾中微子实验中，一类很重要的本底是由宇宙线 μ 子致散裂快中子。快中子本底属于关联本底，其快慢信号均和来自 μ 子的

中子有关。这一节中，我们先用内水池 (IWS) 标记到的 muon 挑选出快中子的样本，得到其快慢信号的能谱和时间间隔分布，然后用一种简单的，不需要依赖于 MC 模拟的线性外推法得到快中子本底的事例率及其信噪比。

5.3.1 快中子挑选及结果

在大亚湾离线物理数据中，首先要根据一定的标准 cut 条件对噪声事例、PMT 的 flasher[104] 事例进行排除，并对 μ 子事例有如下标记 (tag):

- 水池 μ 子 (IWSMuon/OWSMuon) 定义为水池 PMT 触发多重度 ≥ 12 ;
- 探测器 μ 子 (ADMuon) 定义为 AD 中的触发信号能量 $\geq 100\text{MeV}$ ，且此信号与水池 μ 子关联 (位于其前后 $2\mu\text{s}$ 内)， $2\mu\text{s}$ 的时间窗是为了弥补不同的探测器的触发时间的漂移;
- 探测器簇射 μ 子 (ADShowerMuon) 定义为 AD 中的触发信号能量 $\geq 2.5\text{GeV}$ ，并且此信号与 IWSMuon 或者 OWSMuon 相关联;

根据如上对 μ 子的标记，我们可以先挑选出一部分“好的” IWSMuon 样本，在此样本基础上再进行快中子事例挑选。“好” IWSMuon 样本的挑选条件如下:

- 此 IWSMuon 不与 ADMuon 或 ADShowerMuon 关联;
- 以此 IWSMuon 为时间基点，时间窗 $[-202\mu\text{s}, 402\mu\text{s}]$ 内，没有另一个 ADMuon，ADShowerMuon 或者 IWSMuon，这样是为了避免 ADMuon 或 ADShowerMuon 的影响，和另外临近 IWSMuon 引发的中子重复计数;

根据挑选得到的“好” IWSMuon 样本，进一步挑选其中的快中子信号，挑选条件如下:

- 以每个 IWSMuon 为时间基点，开时间窗 $[-2\mu\text{s}, 200\mu\text{s}]$ ，在这段时间内寻找反中微子探测器中的触发信号，快信号能量在区间 $[0.7\text{MeV}, 100\text{MeV}]$ 内;

- 多重度 cut: 在 $[-2\mu\text{s}, 200\mu\text{s}]$ 时间窗内, 每个反中微子探测器中只能有 2 个信号, 并且快信号之前 $200\mu\text{s}$ 、慢信号之后 $200\mu\text{s}$ 的时间窗内, 没有其他信号 (去除多重符合事例);

事例挑选的时间关联, 如图5.6所示:

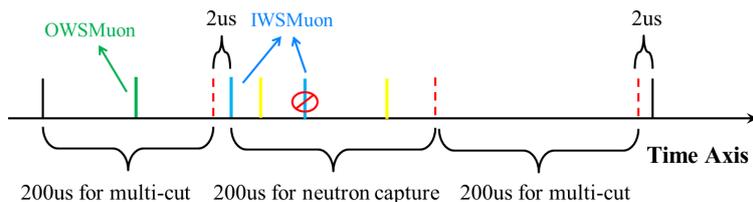


图 5.6 由 IWSMuon 标记的 μ 子致快中子的事例初步挑选示意图

经过 IWSMuon 样本挑选和初步的快中子信号挑选, 得到的事例快慢信号的能谱分布如图5.7 所示:

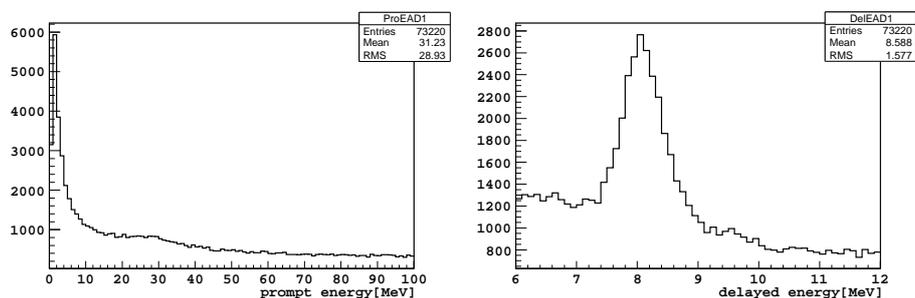


图 5.7 左图是经过事例初步挑选得到的快慢信号对快信号能谱分布; 右图是慢信号能谱分布

从快信号能谱可知, 大多数信号的能量主要集中在 5MeV 以下的区域, 而从慢信号的能谱可以清楚的看出, 中子在在钷上的俘获峰, 叠加在本底之上。同时, 快慢信号的顶点分布, 如图5.8 所示:

从顶点分布图上可以看出, 绝大多数快信号和慢信号集中在掺钷液闪之外、普通液闪之内的区域。为了研究这些信号来源, 我们可以进一步观察慢信号和其母粒子 IWSMuon 信号之间的时间间隔分布。如图5.9 所示:

由图中所示, IWSMuon 信号与慢信号时间间隔在 $10\mu\text{s}$ 之内的分布较为集中, 并呈现指数分布, 用指数函数对这部分分布进行拟合, 得到的时间常数 τ_f

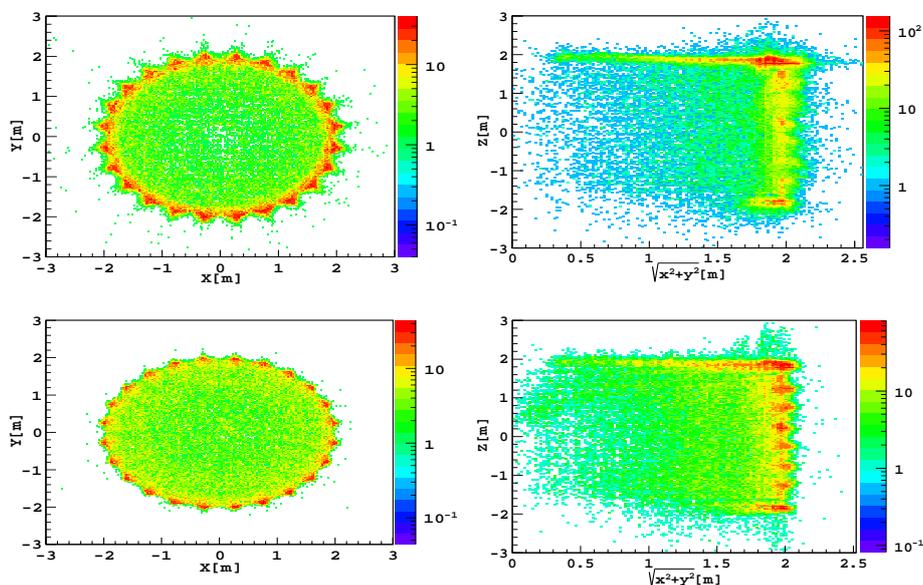


图 5.8 图中坐标 X, Y, Z 是以圆柱形反中微子探测器的几何中心为坐标原点, 以探测器旋转对称轴为 Z 轴形成的直角坐标系。上图显示了快信号的顶点分布, 下图则是慢信号顶点分布, 左列是事例顶点在 X-Y 平面上的分布, 右图是事例顶点在 R-Z 平面内的分布

$=2.27\mu\text{s}$, 和 μ 衰变的平均时间 $\tau_\mu = 2.197\mu\text{s}$ 非常接近, 这一部分事例绝大多数来源于停止在探测器中的宇宙线 μ 子的衰变, 为了去除这一部分 μ 衰变事例, 可以设置 cut 要求慢信号的事例距跟它关联的 IWSMuon 时间间隔 $\geq 12\mu\text{s}$, 根据 μ 衰变的时间分布可知, 这样做可以去除 $\sim 99.58\%$ 的 μ 衰变信号。

去除了 μ 子衰变事例的样本, 其快慢信号的能谱分布如图 5.10 所示。快信号能谱和图 5.7 相比, 在能量区间 $[15\text{MeV}, 30\text{MeV}]$ 的事例明显减少, 减少的部分可能是能量较低的宇宙线 μ 子在探测器中形成的“stop muon”事例, 慢信号

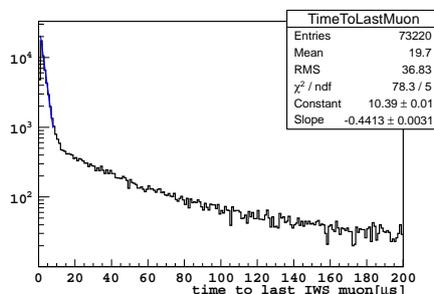


图 5.9 初步挑选得到的“快中子”样本慢信号与 IWSMuon 信号时间间隔分布图。

的本底也大幅降低。

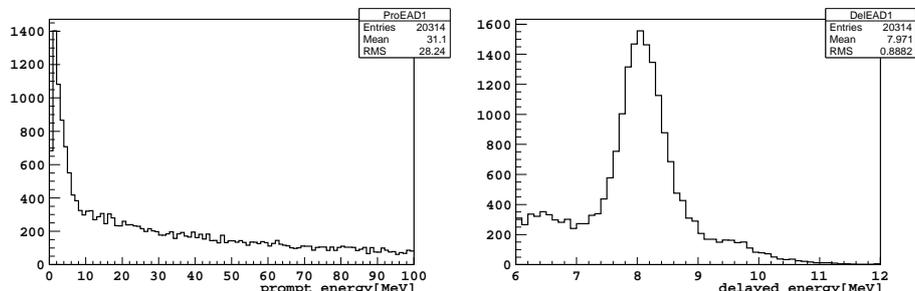


图 5.10 去除了 μ 子衰变信号之后的快慢信号能谱分布，左图是快信号，右图是慢信号

去除 μ 子衰变事例之后的样本，其快慢信号顶点分布如图5.11 所示，可以看出，对慢信号，其顶点在掺钎液闪之外，普通液闪内仍有聚集。

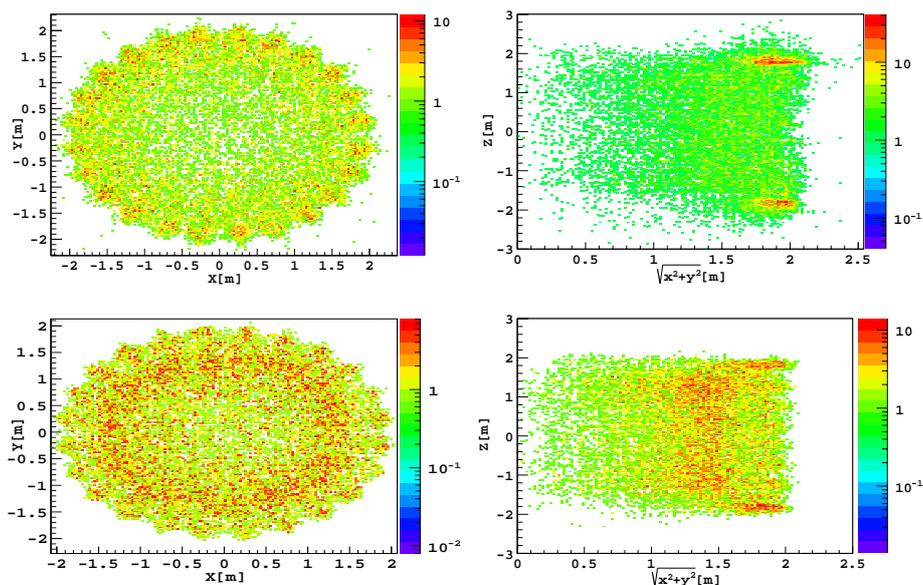


图 5.11 去除了 μ 子衰变信号之后的快慢信号顶点分布。上图是快信号顶点分布，下图是慢信号顶点分布

由于对 inverse β 衰变的事例，通过能量 cut 要求其慢信号仅仅是中子在钎核上的俘获信号。只有在钎上俘获的快中子，才有可能通过反 β 衰变的事例挑选条件，形成本底。所以，我们对样本进行进一步检查，挑选慢信号位于掺钎液闪之外的事例：对慢信号要求 $R > 1.5\text{m}$ 且 $|Z| > 1.5\text{m}$ ，得到这部分事例的快慢信号能谱和快慢信号时间间隔分布。如图5.12 所示。

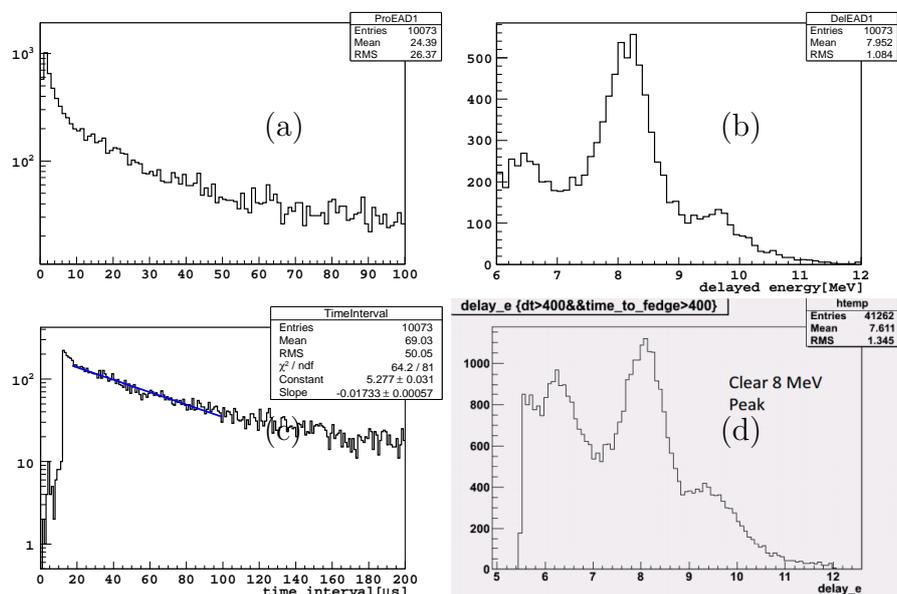


图 5.12 慢信号位于掺钆液闪之外的快慢信号能谱和时间间隔：(a) 图为快信号能谱分布：(b) 图为慢信号能谱分布：(c) 图为快慢信号时间间隔：(d) 图为自动刻度装置的中子在钢罐上的俘获能谱

通过对比慢信号能谱和自动刻度装置中的放射源产生的中子能谱，图5.12中 (d) 图和 NNDC (National Nuclear Data Center) 提供的数据，可以知道这部分事例的慢信号主要是中子在探测器不锈钢罐 ($\sim 20\%Cr$, $\sim 70\%Fe$, Ni , Mn) 上的俘获能谱。8MeV 左右的峰，主要是中子在 ^{52}Cr ($\sim 7.9MeV$)、 ^{60}Ni ($\sim 7.8MeV$) 上的俘获峰；6MeV \sim 7MeV 之间的峰，则是中子在 ^{56}Fe 的不同能级上俘获的能量 (5.9MeV \sim 7.2MeV)。9MeV 和 10MeV 中间的峰，是中子在 ^{58}Ni ($\sim 9.0MeV$) 和 ^{61}Ni 的不同能级上俘获的能量 (9.4MeV \sim 10.5MeV) 而这部分符合事例快慢信号时间间隔并非呈良好的指数分布。如前所述，只有在掺钆液闪内的快中子信号才有可能通过 inverse β 事例的挑选条件，成为本底。为了研究这部分能够成为本底的快中子事例，我们可以通过要求慢信号 $R < 1.5m$, $|Z| < 1.5m$, 快信号 $R < 2.0m$, $|Z| < 2.0m$, 将这一部分中子在钢罐上俘获的偶然符合事例去除。

去除了快中子在掺钆液闪之外的事例之后的快慢信号能谱和时间间隔分布如图 5.13 所示。

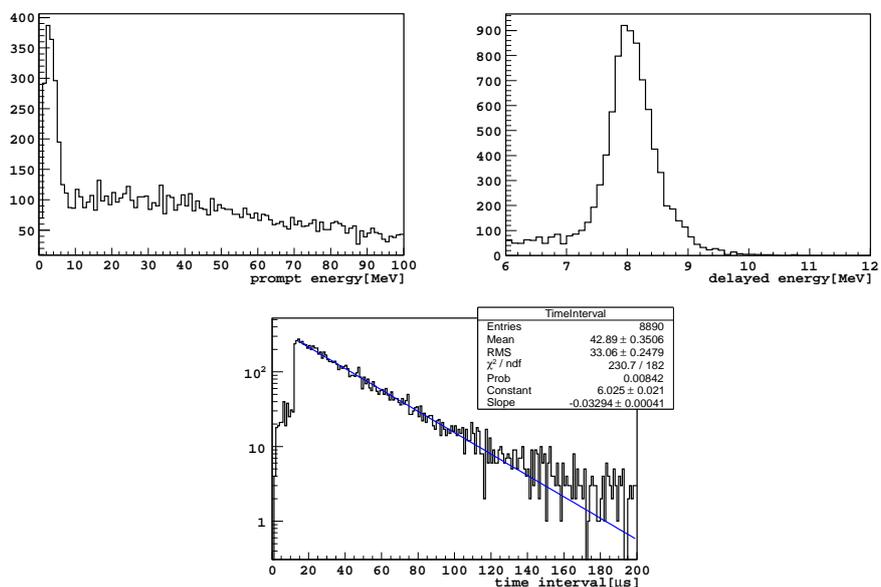


图 5.13 图中显示了做了 fiducial cut 之后得到的样本事例快慢信号能谱和时间间隔，从快信号能谱和快慢信号时间间隔上看，挑选得到的样本中，仍然有不少 IBD 事例（快信号 0.7MeV~12MeV 之间的能量分布）和偶然符合信号（快慢信号时间间隔 $<12\mu\text{s}$ 的部分）。

由快中子产生的物理机制和模拟可知，快中子信号的快信号：反冲质子在靶物质中沉积动能的时间标度一般小于 100ns。可以利用这个条件，将真正的快中子与 IBD 和其他偶然符合事例区分开来。通过研究样本中快信号和 IWSMuon 之间的时间间隔分布，可以得到较为优化的 cut 条件，以将快中子信号和 inverse β 衰变信号和偶然符合本底区分开来。样本中快信号和 IWSMuon 之间的时间间隔分布，如图 5.14 所示。

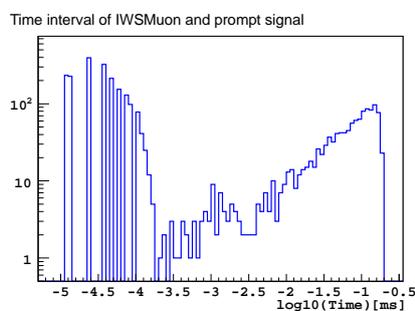


图 5.14 “快中子”样本中的快信号与 IWSMuon 信号时间间隔分布图

从图中可以看出，时间间隔的分布明显表现出两部分，设置 cut： $\log_{10}(|\text{time interval}|) < -3.7$ ，即 $|\Delta T(IWS-Prompt)| < 200\text{ns}$ 可以将两部分物理机制不同的事例划分开来。

图5.15显示了横轴中，在大于 -3 的区间内，由于开时间窗和多重度 cut 的不同，造成 inverse β 衰变事例分布和美方成员的挑选结果完全不同。对和 muon 关联的中子在掺钷液闪上俘获的平均时间常数 $28\mu\text{s}$ 来说， $200\mu\text{s}$ 的时间窗包含了 $\sim 99.91\%$ 的中子俘获信号，而且可以明显减小与 μ 子非关联的信号 (inverse β 衰变事例和偶然符合事例)。

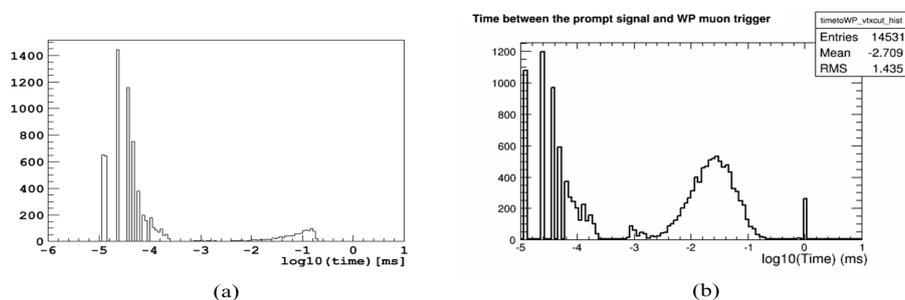


图 5.15 “快中子”样本中的快信号与 IWSMuon 信号时间间隔分布图对比，右图是美方成员进行快中子事例挑选得到的分布，由于用的时间窗不同，样本中含有大量 inverse β 衰变和偶然符合事例，左图则显示了良好的时间窗选择可以去除这些与 IWSMuon 无关的事例

利用上述的时间关联 cut，分离 inverse β 衰变事例和快中子事例如图5.16所示。

由上图左列的快信号能谱和快慢信号时间间隔来看，这部分事例以 inverse β 衰变事例为主，还有少量偶然符合事例。而右列则是最终挑选得到的快中子信号。

而得到的快中子快慢信号各自的顶点分布则如图5.17所示：

在图5.16中，可以注意到在快中子慢信号能谱的 $[8.5\text{MeV}, 9\text{MeV}]$ 处有个平台分布，和 inverse β 衰变的中子信号的能谱分布有明显区别。考虑到快中子的中子俘获信号主要是集中在反中微子探测器掺钷液体闪烁体的外缘，如图5.17所示。靠近掺钷液闪边缘部分的中子俘获事例，会有比较明显的能量泄漏，通过对中子在钷上的俘获峰的拟合，可以很好地解释快中子慢信号能谱在 $[8.5\text{MeV}, 9\text{MeV}]$ 区间内的平台分布。为了描述中子在钷的两种同位素 ^{155}Gd

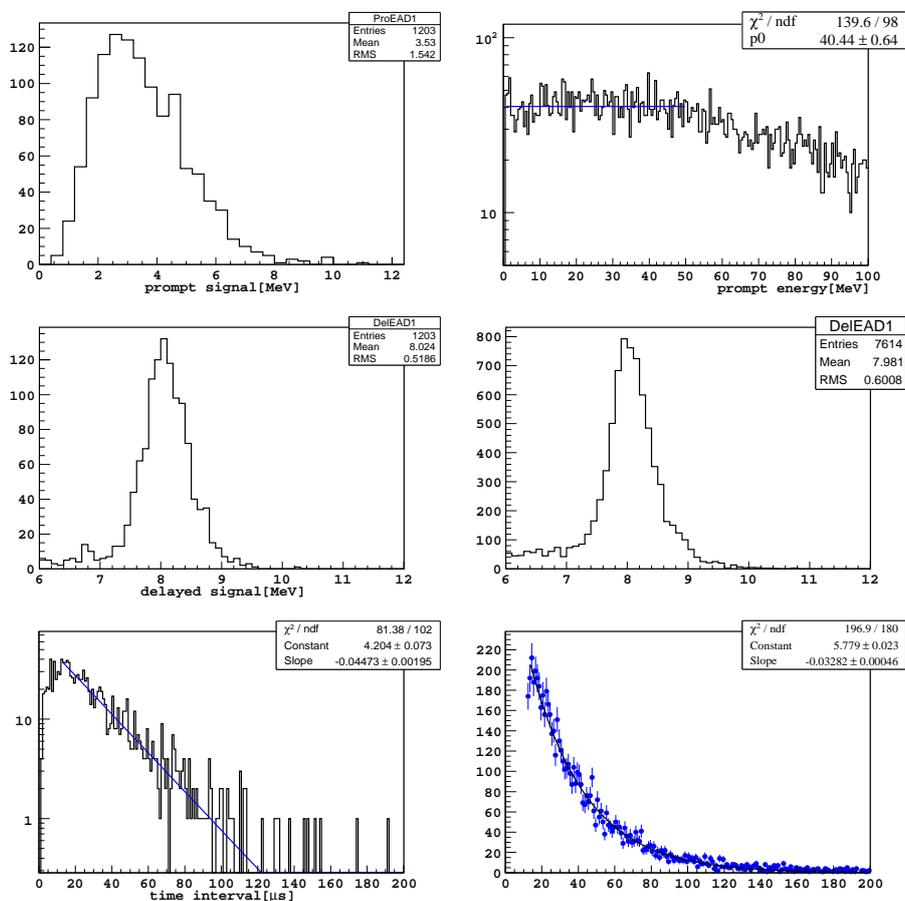


图 5.16 图中显示了利用时间关联 $|\Delta T(IWS-Prompt)|$, 分离得到的两部分不同事例。左列是 $|\Delta T(IWS-Prompt)| > 200ns_{cut}$ 后得到的快慢信号能谱和时间间隔分布, 右列是 $|\Delta T(IWS-Prompt)| < 200ns$ 得到的相应分布

和 ^{157}Gd 上的俘获和掺钷液闪边缘发生的能量泄漏效应, 拟合函数采用“双 Crystall Ball 函数”, 拟合结果如图 5.18 所示:

上一节中, 对快中子样本进行挑选的目的是得到比较干净的快中子事例, 进而得到比较准确的快中子能谱, 尤其是反冲质子动能沉积引起的快信号, 基于这个原因, 我们不太关心这些 cut 造成的快中子事例的损失, 而更多地关注得到的快中子样本中, 混入的其他事例所占的比例。对于上述的对由内水池标记的 μ 子致散裂快中子挑选条件而言, 能够通过挑选条件影响最终快中子信号能谱的“本底”, 主要有如下三类:

1. 在探测器中的“stop muon”事例, 慢信号是 Michel 电子, 并且 μ 子的衰

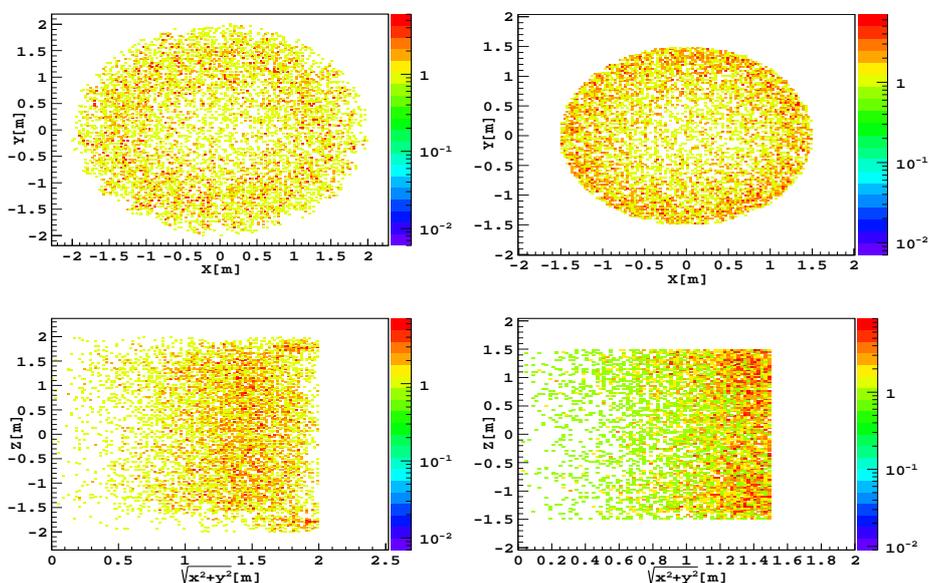


图 5.17 快中子快慢信号的顶点分布。上图是快信号顶点分布，下图是慢信号顶点分布。由于快中子主要由水池 μ 子产生，从外部进入反中微子探测器并进行触发，故中子俘获顶点集中在反中微子探测器外部。

变寿命 $>12\mu\text{s}$ 的残留部分；

2. 内水池标记的 μ 子致双中子事例；
3. 与内水池标记的 μ 子没有关联的其他事例，包括：偶然符合事例和 inverse β 衰变事例，前者是快信号与 IWSMuon 没有关联，而慢信号可能会有关联，后者是快慢信号都与 IWSMuon 没有关联。

对宇宙线 μ 子，缺乏足够准确同时足够统计量的 MC 模拟数据，所以要精确地计算前两种“本底”在快中子样本中所占的比例和误差是一件相当困难的事情，为此，我们可以采用近似手段，估算前两种“本底”通过上述各种挑选条件之后能够残留下来的上限，假如上限所占的比例都很小，那就足以证明挑选得到的快中子样本的纯度。下面，按照上述顺序，对每一种“本底”给出其上限的估算方法和结果。

1. 对“stop muon”事例，如图5.9所示， μ^+ 寿命为 $2.197\mu\text{s}$ ， μ^- 寿命为 $2.026\mu\text{s}$ ，两种 μ 子总寿命为 $2.1\mu\text{s}$ [96]，在要求慢信号距离 IWSMuon 的

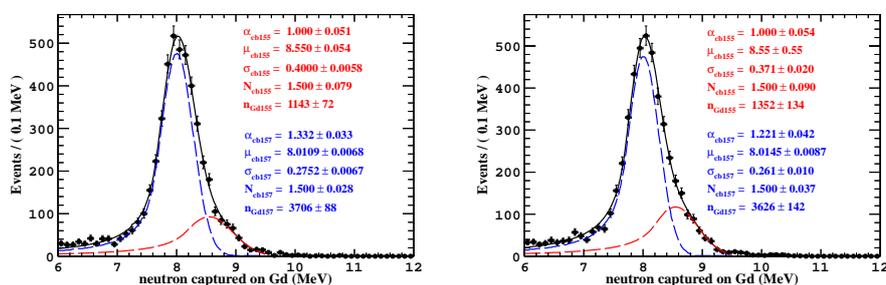


图 5.18 快中子在钆的两种同位素 ^{155}Gd 和 ^{157}Gd 上的俘获能谱拟合结果，对 AD1（左图），在 ^{155}Gd 上俘获的中子占 23.5%，对 AD2（右图），占 24%，由于快中子顶点分布和散裂中子，刻度源中子都不一样，其能量泄漏效应更加明显，需要仔细研究进行进一步细致的拟合

时间间隔 $>12\mu\text{s}$ ，则残存的“stop muon”信号占有所有“stop muon”的比例为 $e^{\frac{12\mu\text{s}}{2.1\mu\text{s}}} \sim 0.33\%$ 。由于是估算上限，则可假定图中 $12\mu\text{s}$ 内全部是“stop muon”事例，所以残留事例为 $(52892/0.9967) * 0.33\% \approx 175$ 。根据大亚湾实验对真实数据中的“stop muon”的慢信号 Michel 电子的顶点分布的研究结果，如图 5.19[96]，对慢信号进行 fiducial cut 之后，残留的“stop muon”上限事例个数 $175 * 1\% \approx 1.8$ 。

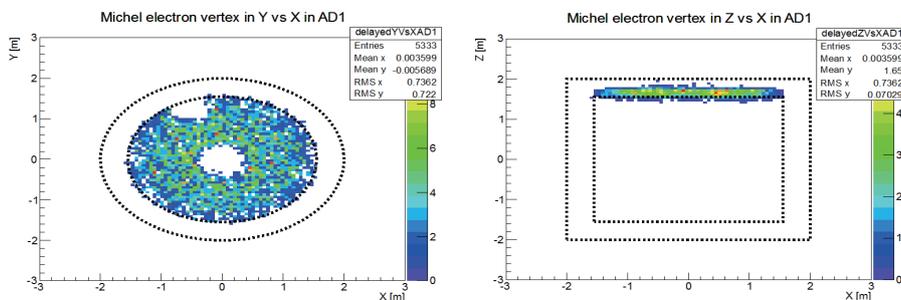


图 5.19 反中微子探测器中“stop muon”产生的慢信号，即米歇尔电子信号的顶点分布。图中的分布对顶点做了 cut: $\sqrt{x^2 + y^2} < 1.5\text{m}$ 之后得出的，可见绝大多数米歇尔电子分布在探测器普通液闪中。左图中的“hole”位于反中微子探测器顶端的 ACU 的位置，主要是因为 ACU 系统中的 stop muon 能量较低，同时在水中的径迹较短，产生的切伦科夫光更弱，无法达到水池 muon 的 cut 要求: $n\text{PMT} > 12$ ，而被排除导致

2. 内水池标记的 μ 子致双中子事例，双中子事例是关联事例，前后两个中子都来自于同一个宇宙线 μ 子事例，由于在最后一步时间关联 cut 中，要求快信号和 IWSMuon 的时间间隔 $< 200\text{ns}$ ，需要估算双中子中第一个中子

200ns 内在掺钆液闪和液闪中的被俘获的比例，根据对中子在掺钆液闪和液闪中俘获过程的模拟，如图5.20 所示，可以得到：中子在掺钆液闪中俘获时间 $<200\text{ns}$ 所占的比例： $4175/1983852 \approx 0.2\%$ ；中子在液闪中的俘获时间 $<200\text{ns}$ 所占的比例： $1 - e^{-\frac{200\text{ns}}{208.2\mu\text{s}}} \approx 0.096\%$ 。

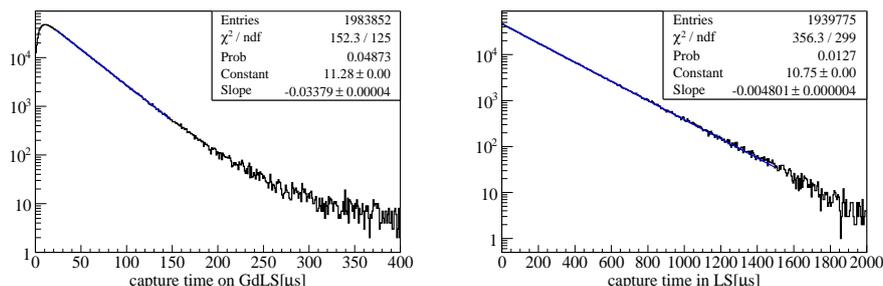


图 5.20 MC 模拟数据给出的中子俘获时间分布：左图是中子在掺钆液闪中俘获时间分布，右图是中子在普通液闪中的俘获时间分布。

同时，根据大亚湾对近点 μ 子致中子多重度的模拟结果，如图5.21 所示，可知，宇宙线 μ 子产生双中子的概率为： $2n/(0n+1n+2n) \approx 2\%$ ，这里没有计入 $3n$ 或者 $3n$ 以上的事例，因为这一部分无法通过事例挑选的多重度 cut。对双中子“本底”的估计从图5.10 开始，同样因为是估算上限，认为图中所示全部是 IWSMuon 与其产生的散裂中子，得到，掺钆液闪中的双中子“本底”： $20326 * 2\% * 0.2\% \approx 0.8$ ；液闪中的双中子“本底”： $20326 * 2\% * 0.096\% \approx 0.4$ ，总共 1.2 个。

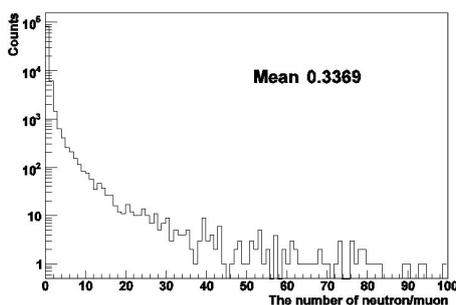


图 5.21 大亚湾近点实验大厅，宇宙线 μ 致散裂中子多重度 MC 模拟结果。

- 与内水池标记的 μ 子没有关联的 inverse β 衰变及其他事例，这一部分“本底”的估算较为简单，如图 5.16所示，由于这一部分“本底”

和 IWSMuon 没有任何关联, 可以用时间窗外推的方法估算, 在距离 IWSMuon $200\mu\text{s}$ 的时间窗内, 有 1203 个这种“本底”, 根据时间窗口外推至距离 IWSMuon 200ns 的时间内, 则这种“本底”大约有 $1203 \times 200\text{ns} / 200\mu\text{s} \approx 1.2$ 个。

表格5.6显示了对快中子信号中的各种“本底”估算结果。

“本底”类型	事例数	信号总数	信噪比 B/S
stop muon	<1.8		<0.024%
double neutron	<1.2	7614	<0.016%
IBD and accidental coincidence	~ 1.2		$\sim 0.016\%$
total	<4.2		<0.055%

表 5.6 挑选得到的快中子样本中各种“本底”的信噪比估算结果。

以上的分析和挑选结果, 是以 1 号厅 AD1 获取的数据为样本, 对 1 号厅的 AD2 和 2 号厅的 AD1 获取的数据应用同样的事例选择条件, 也可以得到这两个探测器中的快中子事例的快慢信号能谱和时间间隔分布, 如图5.22 所示:

5.3.2 快中子本底估算

在上一节通过挑选被内水池标记的 μ 子产生的快中子本底, 得到了快中子本底的快信号能谱, 这一能谱在能区 $[0.7\text{MeV}, 50\text{MeV}]$ 内均匀分布, 与大亚湾模拟结果相符合。这启发我们可以利用基于 inverse β 衰变快信号能谱的“线性外推法”, 对反中微子信号挑选结果中的快中子本底做一估算。为此需要对反中微子信号事例的挑选条件稍作改动:

- 探测器 μ 子 (ADMuon) 定义为 AD 中的触发信号能量 $\geq 50\text{MeV}$;
- inverse β 衰变信号的挑选: flasher cut、多重度 cut, Muon veto 等, 慢信号能量区间, 快慢信号时间间隔不变; 快信号能量区间重新定义为 $[0.7\text{MeV}, 50\text{MeV}]$ 之间

事例挑选所采用的数据集是: 基于在 2011 年 9 月 23 日到 2012 年 2 月 17 日的的数据, 每个探测器都是独立运行取数, 并进行 Muon Veto 得到其 DAQ 活时间。图5.23是线性外推法示意。

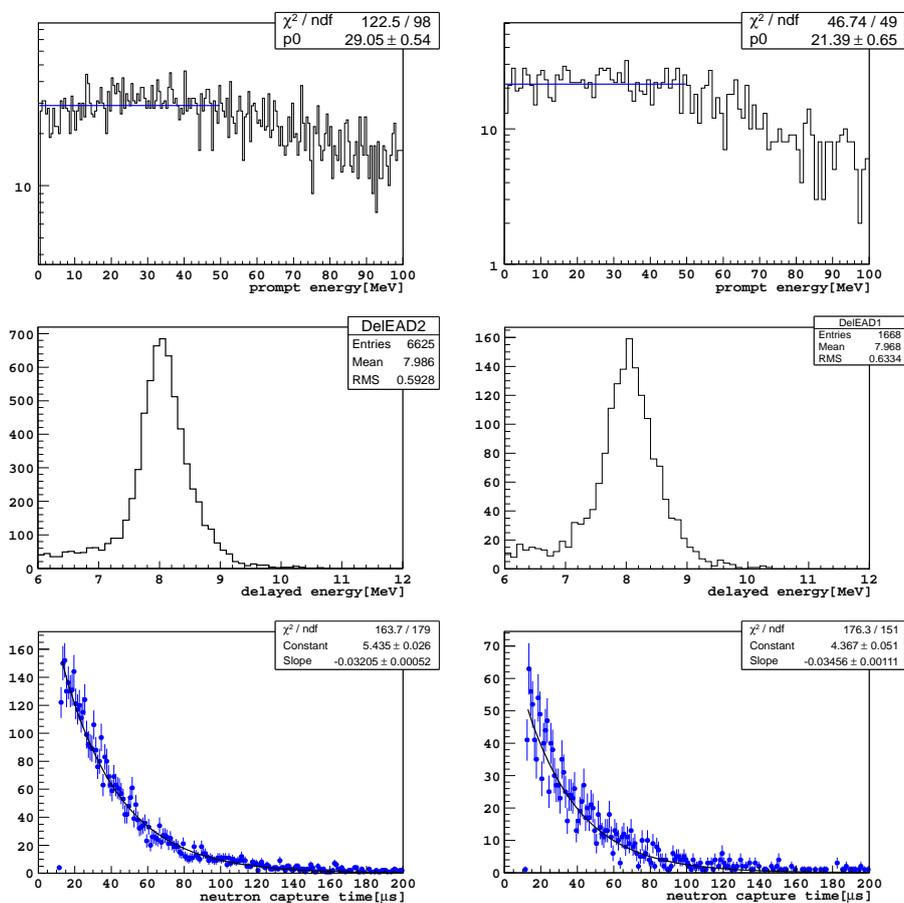


图 5.22 EH1AD2 和 EH2AD1 两个反中微子探测器进行快中子事例挑选的结果。左列是 EH1AD2 的结果，右列是 EH2AD1 的结果。

对于大亚湾远点三个探测器，可以根据上图中快中子的个数，估算出快中子的信噪比和事例率。6 个反中微子探测器的快中子本底结果，如表格 5.7 所示。

5.4 本章小结

对于大亚湾这类位于低能量，低本底和高精度要求的地下实验，对于宇宙线 μ 子的屏蔽是十分重要的，直接关系到整个实验的成败。即使有足够厚度的岩石覆盖，实验中，仍然需要对被反符合探测系统所标记的 μ 子事例进行仔细分析，尽可能排除 μ 子产生的次级粒子对信号事例的影响。本章首先简单介绍了大亚湾实验中的主要本底的产生机制及分类，并回顾了宇宙线 μ 子产生

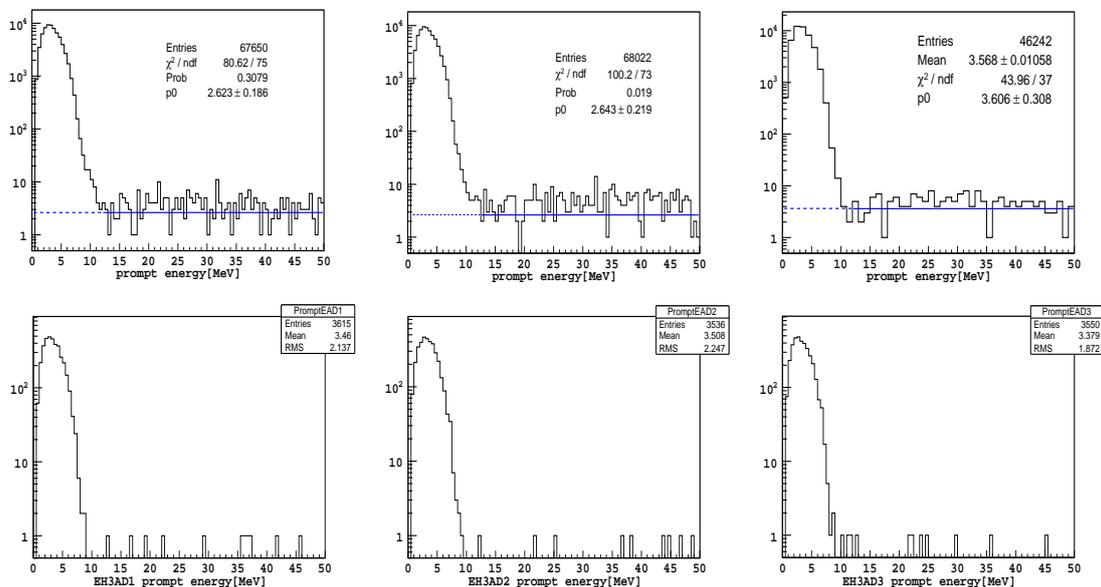


图 5.23 线性外推法估算快中子本底示意图，第一行从左到右分别是 EH1AD1、EH1AD2 和 EH2AD1，以 EH1AD1 为例，线性外推法给出的快中子事例率为 $0.5868 \pm 0.0416(\text{stat}) \pm 0.176(\text{sys})$ ，其中，统计误差由线性拟合结果给出，系统误差暂时放置为结果的 50%，也可以通过不同探测器之间的最大差别给定。第二行从左到右分别是 EH3AD1，EH3AD2 和 EH3AD3，这三个探测器由于目前其数据量太小，没有用零阶多项式进行拟合，而是通过数个的方式，进行外推，得到结果，由于统计量小，误差较大，暂时认为是 100%

探测器	快中子数	信号数	活时间 (day)	B/S(%)	事例率 (/day)
EH1AD1	59.28	63440	~101.2	0.09344 ± 0.00663	$0.5868 \pm 0.0416(\text{stat})$
EH1AD2	59.73	63836	~102.7	0.09357 ± 0.0082	$0.5816 \pm 0.0522(\text{stat})$
EH2AD1	40.75	43585	~79.7	0.09394 ± 0.00802	$0.5111 \pm 0.0437(\text{stat})$
EH3AD1	3.06	3604		0.00849 ± 0.00849	0.0626 ± 0.0626
EH3AD1	2.54	3527	~48.9	0.00721 ± 0.00721	0.0520 ± 0.0520
EH3AD1	2.54	3549		0.00716 ± 0.00716	0.0520 ± 0.0520

表 5.7 大亚湾实验反中微子探测器快中子本底结果，近点三个探测器，仅给出了拟合的统计误差。远点的三个探测器，仅给出估算结果，表中所列 100% 的误差为系统误差和统计误差总和

散裂中子的基本物理过程；利用大亚湾实验的反中微子探测器以及 μ 子反符合系统，对大亚湾三个实验大厅中的 μ 子致散裂中子产额进行了估算，最后，利用实验数据，挑选得到 μ 子产生的快中子信号，得到其快信号能谱分布，并利用其能谱的特点，采用线性外推法，计算得到六个探测器中快中子本底在 inverse β 信号事例中的信噪比。

第六章 总结与展望

6.1 总结和展望

6.1.1 论文总结

本论文的工作主要是在大亚湾实验中离线软件框架下对来自于 DAQ 数据获取系统的二进制原始数据的访问和转换的软件开发并利用相关软件进行探测器测试数据分析,对 MC 模拟数据,MC 模拟数据和真实数据进行混合,为合作组提供可供物理分析的 MC 数据,以及对来自宇宙线 μ 子的散裂中子产额以及快中子本底的估算。

论文包含以下几个方面:

第一部分是对实验二进制原始数据的访问和数据格式转换服务软件的开发与测试,原始数据转换服务在中心探测器 AD1, AD2 的 Mini dry run 和 Dry run 测试中,发挥了重要的作用。利用原始数据转换服务,完成了对 FEE 板电子学通道阈值和台阶的刻度工作以及其他测试数据的分析,使得合作组对反中微子探测器性能更为了解。在数据转换服务的开发过程中,注重了算法的性能和整体架构的设计,使得数据转换服务运行稳定可靠,没有成为后续用户分析程序的性能瓶颈。

第二部分是基于改进了数据流 I/O 接口的 NUWA 框架,实现了“并行 Interleaving 算法”对 MC 模拟物理样本的在电子学读出级的混合,并进一步开发出了性能更高的“串行 Interleaving”算法,提高了样本混合的效率和质量,为合作组提供了数个版本 $\sim 100\text{TB}$ 的混合数据,并以此为基础,完成了初步的事例挑选和本底分析,经过 MC 混合数据“训练”的信号事例挑选条件和宇宙线 μ 子本底的排除条件,大部分在分析真实物理数据中得到了保留和发展。与此同时,还实现了探测器击中级 MC 样本混合,并提出了“SimHitSplit”方法,在保证电子学模拟的逻辑正确性的基础上,提高了电子学模拟的效率和可靠性。最后,采用多种设计模式,实现了未来对于混合 MC 和真实数据之间的

算法要求，为未来更高精度的物理分析打下基础。

第三部分介绍在探测器中宇宙线 μ 子本底。首先介绍宇宙线 μ 子致散裂中子的物理原理和大亚湾实验对宇宙线 μ 子的测量方法，以及散裂中子在大亚湾探测器中形成快中子的物理机制。并根据实验数据，以掺钆液闪为靶体积，分析并估算了大亚湾三个实验大厅中，宇宙线 μ 子致散裂中子的产额。最后从实验数据中挑选出快中子本底，得到其快信号特征能谱并与 MC 模拟数据和其他实验结果进行对比，根据其快信号能量特征，利用线性外推法，对 inverse β 事例中的快中子本底的信噪比和误差，进行了估算。

最后，对工作进行了总结和展望。

6.1.2 展望

大亚湾实验已经运行 8 个月，并且发表了第一个物理结果。在国际竞争激烈的环境下，以 5σ 的置信概率首先确认中微子混合角 θ_{13} 不为零。在整个实验正式开始取数之前，相关的分析工具包括能谱预测的分析已经准备好。

大亚湾实验目前发表的结果是基于事例率的振荡分析。在未来的 2 到 3 年内实验将继续取数，使我们有足够的统计量进行能谱形状的振荡分析，同时使我们能以世界最高的精度精确测量 θ_{13} 。LAF 轻量级物理分析软件框架将在实验未来的时间里将继续持续为实验提供物理分析的软件支持。

除了中微子振荡，反应堆中微子实验还有一些其他物理课题可以深入，比如研究可能存在的惰性中微子等。

参考文献

- [1] Ahmad Q R, et al. [SNO Collaboration]. Measurement of the Rate of $\nu_e + d \rightarrow p + p + e^-$ Interactions Produced by ^8B Solar Neutrinos at the Sudbury Neutrino Observatory. *Phys. Rev. Lett.*, 2001, 87:071301.
- [2] Fukuda Y, et al. [SuperK Collaboration]. Measurement of a small atmospheric ν_μ/ν_e ratio. *Phys. Lett. B*, 1998, 433:9.
- [3] Eguchi K, et al. [KamLAND Collaboration]. First Results from KamLAND: Evidence for Reactor Antineutrino Disappearance. *Phys. Rev. Lett.*, 2003, 90:021802.
- [4] Michael D, et al. [MINOS Collaboration]. Observation of Muon Neutrino Disappearance with the MINOS Detectors in the NuMI Neutrino Beam. *Phys. Rev. Lett.*, 2006, 97:191801.
- [5] Mohapatra R N, Senjanovic G. Neutrino Mass and Spontaneous Parity Nonconservation. *Phys. Rev. Lett.*, 1980, 44:912.
- [6] Perkins D H. *Introduction to High Energy Physics*(4th ed). Cambridge:Cambridge: Printed in the United Kingdom at the University Press, 世界图书出版公司, ISBN: 0-521-62196-8, 1999.
- [7] Amsler C, et al. [Particle Data Group]. Review of Particle Physics. *J. Phys. G:Nucl. Part. Phys.*, 2006, 33:1.
- [8] Incandela J. [CMS Collaboration]. Status of the CMS SM Higgs Search. 2012. July 4.
- [9] Gianotti F. [ATLAS Collaboration]. Status of the Standard Model Higgs Searches in ATLAS. 2012. July 4.
- [10] 许咨宗. *核与粒子物理导论*. 安徽: 中国科学技术大学出版社, 书号: 978-7-312-02310-1, 2009.
- [11] Reines F, Cowan C L. Free Antineutrino Absorption Cross Section. I. Measurement of the Free Antineutrino Absorption Cross Section by Protons. *Phys. Rev.*, 1959, 113:273.
- [12] Danby G, Gaillard J M, Goulianos K, et al. Observation of High-Energy Neutrino Reactions and the Existence of Two Kinds of Neutrinos. *Phys. Rev. Lett.*, 1962, 9:36.
- [13] Kodama K, et al. [DONUT Collaboration]. Observation of tau neutrino interactions. *Phys. Lett. B*, 2001, 504:218.

-
- [14] Komatsu E, et al. Seven-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Interpretation. *Astrophys. J. Suppl.*, 2011, 192:18.
- [15] Maki Z, Nakagawa M, Sakata S. Remarks on the Unified Model of Elementary Particles. *Prog. Theor. Phys.*, 1962, 28:870.
- [16] Fukugita M, Yanagida T. Baryogenesis without grand unification. *Phys. Lett. B*, 1986, 45:174.
- [17] Bahcall J N, Bahcall N A, Shaviv G. Present Status of the Theoretical Predictions for the ^{37}Cl Solar-Neutrino Experiment. *Phys. Rev. Lett.*, 1968, 20:1209.
- [18] Davis R J, Harmer D S, Hoffman K C. Search for Neutrinos from the Sun. *Phys. Rev. Lett.*, 1968, 20:1205–1209.
- [19] Hirata K S, et al. [Kamiokande Collaboration]. Observation in the Kamiokande-II detector of the neutrino burst from supernova SN1987A. *Phys. Rev. D*, 1988, 38:448.
- [20] Abdurashitov J N, et al. [SAGE Collaboration]. Solar neutrino flux measurements by the Soviet-American gallium experiment (SAGE) for half the 22-year solar cycle. *J. Exp. Theor. Phys.*, 2002, 95:181.
- [21] Anselmann P, et al. [GALLEX Collaboration]. GALLEX results from the first 30 solar neutrino runs. *Phys. Lett. B*, 1992, 285:376.
- [22] Fukuda Y, et al. [Super-Kamiokande Collaboration]. Erratum: Measurements of the Solar Neutrino Flux from Super-Kamiokande's First 300 Days. *Phys. Rev. Lett.*, 1998, 81:1158.
- [23] Altmann M, et al. [GNO Collaboration]. GNO solar neutrino observations: results for GNO I. *Phys. Lett. B*, 2000, 490:16.
- [24] Wolfenstein L. Neutrino oscillations in matter. *Phys. Rev. D*, 1978, 17:2369.
- [25] Mikheev S P, Smirnov A Y. Resonance amplification of oscillations in matter and spectroscopy of solar neutrinos. *Sov. J. Nucl. Phys.*, 1985, 41:913.
- [26] Casper D, et al. Measurement of atmospheric neutrino composition with the IMB-3 detector. *Phys. Rev. Lett.*, 1991, 66:2561.
- [27] Allison W W M, et al. [Soudan 2 Collaboration]. Measurement of the atmospheric neutrino flavour composition in Soudan 2. *Phys. Lett. B*, 1997, 391:491.
- [28] Ambrosio M, et al. [MARCO Collaboration]. Measurement of the atmospheric neutrino-induced upgoing muon flux using MACRO. *Phys. Lett. B*, 1998, 434:451.

-
- [29] Ahn M H, et al. [K2K Collaboration]. Indications of Neutrino Oscillation in a 250 km Long-Baseline Experiment. *Phys. Rev. Lett.*, 2003, 90:041801.
- [30] Apollonio M, et al. [CHOOZ Collaboration]. Search for neutrino oscillations on a long base-line at the CHOOZ nuclear power station. *Eur. Phys. J. C*, 2003, 27(3):331.
- [31] Boehm F, et al. [Palo Verde Collaboration]. Search for Neutrino Oscillations at the Palo Verde Nuclear Reactors. *Phys. Rev. Lett.*, 2000, 84:3744.
- [32] Acciarri M, et al. Determination of the number of light neutrino species from single photon production at LEP. *Phys. Lett. B*, 1998, 431:199.
- [33] Abreu P, et al. Searches for neutral Higgs bosons in e^+e^- collisions around $\sqrt{s} = 189$ GeV. *Eur. Phys. J. C*, 2000, 17:53.
- [34] Abbiendi G, et al. [OPAL Collaboration]. Experimental properties of gluon and quark jets from a point source. *Eur. Phys. J. C*, 1999, 11:217.
- [35] Darate D, et al. Single and multi-photon production in e^+e^- collisions at a centre-of-mass energy of 183 GeV. *Phys. Lett. B*, 1998, 429:201.
- [36] Karlen D, et al. [LEP Collaboration and the LEP SUSY working group]. The Number of Light Neutrino Types from Collider Experiments. *J. Phys. G:Nucl. Part. Phys.*, 2006, 478:33.
- [37] Kayser B. Neutrino Mass, Mixing, And Flavor Change. *J. Phys. G:Nucl. Part. Phys.*, 2006, 33:156.
- [38] Aguilar A, et al. [LSND Collaboration]. Evidence for neutrino oscillations from the observation of ν_e appearance in a ν_μ beam. *Phys. Rev. D*, 2001, 64:112007.
- [39] Murayama H, Yanagida T. LSND, SN1987A, and CPT violation. *Phys. Lett. B*, 2001, 520:263.
- [40] Eitel K, et al. [LSND Collaboration]. The search for neutrino oscillations $\bar{\nu}_\mu \rightarrow \bar{\nu}_e$ with KARMEN. *Nucl. Phys. Proc. Suppl.*, 1999, 77:212.
- [41] Aguilar-Arevalo A A, et al. [MiniBooNE Collaboration]. Search for Electron Neutrino Appearance at the $\Delta m^2 \sim 1eV^2$ Scale. *Phys. Rev. Lett.*, 2007, 98:231801.
- [42] Aguilar-Arevalo A A, et al. [MiniBooNE Collaboration]. Event Excess in the MiniBooNE Search for $\bar{\nu}_\mu \rightarrow \bar{\nu}_e$ Oscillations. *Phys. Rev. Lett.*, 2010, 105:181801.
- [43] Amsler C, et al. [Particle Data Group]. Review of Particle Physics. *Phys. Lett. B*, 2008, 667:1.
- [44] Jarlskog C. Commutator of the Quark Mass Matrices in the Standard Electroweak Model and a Measure of Maximal CP Nonconservation. *Phys. Rev. Lett.*, 1985, 55:1039.

-
- [45] Gonzalez-Garcia M C, Maltoni M, Salvado J. Updated global fit to three neutrino mixing: status of the hints of $\theta_{13} > 0$. *JHPE*, 2010, 056:1004.
- [46] Aharmim B, et al. Electron energy spectra, fluxes, and day-night asymmetries of ^8B solar neutrinos from measurements with NaCl dissolved in the heavy-water detector at the Sudbury Neutrino Observatory. *Phys. Rev. C*, 2005, 72:055502.
- [47] Barger V D, et al. Measurement of the High-Mass Drell-Yan Cross Section and Limits on Quark-Electron Compositeness Scales. *Phys. Rev. Lett*, 1999, 82:2640.
- [48] Abe S, et al. Precision Measurement of Neutrino Oscillation Parameters with KamLAND. *Phys. Rev. Lett*, 2008, 100:221803.
- [49] Arpesella C, et al. Direct Measurement of the ^7Be Solar Neutrino Flux with 192 Days of Borexino Data. *Phys. Rev. Lett.*, 2008, 101:091302.
- [50] Guo X H, et al. [Daya Bay Collaboration]. A Precision Measurement of the Neutrino Mixing Angle θ_{13} using Reactor Antineutrinos at Daya Bay. *arXiv:hep-ex/0701029*, 2007..
- [51] 唐孝威等. 粒子物理实验方法. 北京: 人民教育出版社, 第一版, 1982.
- [52] Christian A. Configuration Management Tool(Version v1r18p20051101). 2005.
- [53] Collins-Sussman B, Fitzpatrick B W, Pilato C M. Version Control with Subversion(For Subversion 1.2). Creative Commons, 559 Nathan Abbott Way, Stanford, California 94305, USA: TBA, 2006.
- [54] Barrand G, et al. GAUDI-A software architecture and framework for building HEP data processing applications. *Computer Physics Communication*, 2001, 140(1-2):45–55.
- [55] Ahn J K, et al. [RENO Collaboration]. RENO: An Experiment for Neutrino Oscillation Parameter θ_{13} Using Reactor Neutrinos at Yonggwang. *arXiv:1003.1391[hep-ex]*, 2010..
- [56] An F P, et al. [Daya Bay Collaboration]. Observation of Electron-Antineutrino Disappearance at Daya Bay. *Phys. Rev. Lett.*, 2012, 108:171803.
- [57] An F P, et al. [Daya Bay Collaboration]. A side-by-side comparison of Daya Bay antineutrino detectors. *Nuclear Instruments and Methods in Physics Research A*, 2012, 685:78–97.
- [58] Mikaelyan L A, Sinev V V. Neutrino oscillations at reactors: What is next? *Phys. Atomic Nucl.*, 2000, 63:1002.
- [59] Schreckenbach W G K, Colvin G, Feilitzsch F. Determination of the antineutrino spectrum from ^{235}U thermal neutron fission products up to 9.5 MeV. *Phys. Lett. B*, 1985, 160:325.

- [60] Vonfeilitzsch A F, Schreckenbach K. Phys. Lett. B, 1982, 118:162.
- [61] Hahn A A, et al. Antineutrino spectra from ^{241}Pu and ^{239}Pu thermal neutron fission products. Phys. Lett. B, 1989, 218:365.
- [62] Vogel P, Schenter G K, Mann F M, et al. Reactor antineutrino spectra and their application to antineutrino-induced reactions. II. Phys. Rev. C, 1981, 24:1543.
- [63] 占亮. 大亚湾中心探测器模拟与物理灵敏度研究. 博士论文, 2009, 中国科学院高能物理研究所: 北京.
- [64] Agostinelli S, et al. GEANT4-a simulation toolkit. Nuclear Instruments and Methods in Physics Research A, 2003, 506(3):250–303.
- [65] http://dayabay.bnl.gov/oum/_downloads/OfflineUserManual.pdf..
- [66] 王玲玉. 博士论文, 2011, 中国科学院高能物理研究所: 北京.
- [67] http://dayabay.ihep.ac.cn/cgi-bin/DocDB/ShowDocument?docid=7151/20120108_LAF_workshop.pdf..
- [68] Jetter S, Dwyer D, Jiang W Q, et al. PMT waveform modeling at the Daya Bay experiment. Chinese Physics C, 2012, 36(8):733–741.
- [69] Li F, Ji X L, Li X N. DAQ Architecture Design of Daya Bay Reactor Neutrino Experiment. IEEE TRANSACTIONS ON NUCLEAR SCIENCE, 2011, 58(4):1723.
- [70] http://dayabay.ihep.ac.cn/cgi-bin/DocDB/ShowDocument?docid=3442/The_raw_data_file_format_in_the_Daya_Bay_DAQ_v11.pdf..
- [71] http://dayabay.ihep.ac.cn/cgi-bin/DocDB/ShowDocument?docid=2991/Data_Format_for_FEEV1.0.doc..
- [72] Bellamy E H, Bellettini G, Budagov J, et al. Absolute calibration and monitoring of a spectrometric channel using a photomultiplier. Nuclear Instruments and Methods in Physics Research A, 1994, 339:468–476.
- [73] 王志民. 中微子探测器模型的研究. 博士论文, 2009, 中国科学院高能物理研究所: 北京.
- [74] http://dayabay.ihep.ac.cn/cgi-bin/DocDB/ShowDocument?docid=1564/mc_mixing.pdf..
- [75] 陈希孺. 概率论与数理统计. 合肥: 中国科学技术大学出版社, 2009.
- [76] 李惕培. 实验的数据处理. 北京: 科学出版社, 1980:54-55.

- [77] 李小波. 低本底和液体闪烁体探测器技术研究. 博士论文, 2011, 中国科学院高能物理研究所: 北京.
- [78] Kudryavtsev V A. Muon simulation codes MUSIC and MUSUN for underground physics. *Computer Physics Communications*, 2009, 180(3):339–346.
- [79] 刘金昌. 大亚湾反应堆中微子实验中的中子本底研究. 博士论文, 2008, 中国科学院高能物理研究所: 北京.
- [80] Weiss M A. 数据结构与算法分析—C++ 描述. 北京: 人民邮电出版社, 第三版, 2006.
- [81] Lippman S B, Moo B E, Lajoie J. C++ Primer 中文版. 北京: 人民邮电出版社, 第四版. 特别版, 2011.
- [82] Gamma E, Helm R, Johnson R, et al. 设计模式—可复用面向对象软件的基础. 北京: 机械工业出版社, 1994.
- [83] Wang Y F, Miller L, Gratta G. New approach to background subtraction in low-energy neutrino experiments. *Phys. Rev. D*, 2000, 62:013012.
- [84] Bemproad C, Gratta G, Vogel P. Reactor-based neutrino oscillation experiments. *Rev. Mod. Phys.*, 2002, 74:297–328.
- [85] Shimizu I. An Evidence for Spectral Distortion of Reactor Anti-Neutrinos and A Study of Three Flavor Neutrino Oscillation. Ph.D. thesis, 2004, Department of Physics, Tohoku University:Sendai.
- [86] Toshiyuki I. Measurement of Reactor Anti-Neutrino Disappearance in KamLAND. Ph.D. thesis, 2003, Graduate School of Science, Tohoku University:Sendai.
- [87] <http://dayabay.ihep.ac.cn/DocDB/0005/000543/001/dayabay06.ppt..>
- [88] Hagner T, et al. Muon-induced production of radioactive isotopes in scintillation detectors. *Astr. Part. Phys.*, 2000, 14:33–47.
- [89] Wang Y F, Balic V, Gratta G, et al. Predicting neutron production from cosmic-ray muons. *Phys. Rev. D*, 2001, 64:013012.
- [90] Suzuki T, Measday D F. Total nuclear capture rates for negative muons. *Phys. Rev. C*, 1987, 35:2212–2224.
- [91] Eidelman S, et al. [Particle Data Group]. Gauge and Higgs boson summary table. *Phys. Lett. B*, 2004, 592:31–88.
- [92] Huff R W, et al. Decay rate of bound muons. *Ann. Phys.(N. Y.)*, 1961, 16:288.

- [93] Mukhopadhyay N C, et al. Nuclear Muon capture. *Phys. Rep.*, 1977, 30C:1–144.
- [94] 关梦云. 大亚湾反应堆中微子实验本底研究和模型研制. 博士论文, 2006, 中国科学院高能物理研究所: 北京.
- [95] 安丰鹏. 大亚湾中微子流强计算及对实验灵敏度的影响. 博士论文, 2012, 中国科学院高能物理研究所: 北京.
- [96] http://dayabay.ihep.ac.cn/cgi-bin/DocDB/ShowDocument?docid=7457/muonDecay_12Jan2012.pdf.
- [97] Hertenberger, Chen M, Dougherty B L. Muon-induced neutron and pion production in an organic liquid scintillator at a shallow depth. *Phys. Rev. C*, 1995, 52:3449.
- [98] Bezrukov L B, et al. *Sov. J. Nucl. Phys.*, 1973, 17:51.
- [99] Boehm F, et al. Neutron production by cosmic-ray muons at shallow depth. *Phys. Rev. D*, 2000, 62:092005.
- [100] Enikeev R I, et al. *Sov. J. Nucl. Phys.*, 1987, 46:883.
- [101] Aglietta M, et al. Measurement of the Neutron Flux Produced by Cosmic-Ray Muons with LVD at Gran Sasso. *arXiv:hep-ex/9905047*, 1999..
- [102] Aglietta M, et al. *Nuovo Cimento Soc. Ital. Fis. C*, 1989, 12:467.
- [103] Aglietta M, et al. Muon-induced neutron production and detection with GEANT4 and FLUKA. *Nucl. Instrum. Meth. A*, 2005, 545:398–411.
- [104] http://dayabay.ihep.ac.cn/cgi-bin/DocDB/ShowDocument?docid=7191/about_flasher_summary.pdf.

致 谢

在这篇论文的结尾，回顾五年的科研生活，太多感谢！谨借这个机会向那些在我攻读博士学位期间给予我指导和关心的老师、同学以及支持我的朋友们表示由衷的感谢。

诚挚的感谢我的导师张子平教授，是张老师给了我学习高能物理的机会，并为我创造了良好的科研环境并不断给予方向上的指导。在在高能所交流期间，张老师仍一直密切关注着我的工作进展。五年来，张老师谦和朴实的为人，科研上兢兢业业、孜孜以求的精神都深深地感染和激励着我，是我永远学习的榜样。

诚挚的感谢我的合作导师王贻芳研究员，他指导了我论文中的 MC 模拟数据和真实数据混合以及 μ 子致散裂中子的物理分析部分。王老师有着深厚的中微子物理理论基础和丰富的实验经验，并对许多问题具有敏锐的洞察力和独到的见解，他对软件方法和物理思想的结合让我耳目一新，尤其是在 μ 子致散裂中子物理分析方面选题的让我的学位论文更加完整，整体质量更高，同时也让我接触和了解了许多新知识。在此，我要深深地感谢王老师对我不多而又难得的关心和指导。

感谢高能所曹俊研究员在大亚湾中微子物理离线模拟和探测器方面的指点和帮助，几年来的离线、模拟和分析的工作，都是在曹俊老师的总体安排下完成的；感谢山东大学王萌教授，黄性涛教授在离线软件框架初期工作中的选题和指导，使得我能够在比较短的时间内理解高能物理框架并开展相应工作；感谢高能所软件组李卫东老师在发表软件论文过程中的鼎力支持。

在此特别感谢何苗副研究员，邹佳恒博士，路浩奇博士和陈晓苏硕士。何苗博士和邹佳恒博士精通高能离线软件框架，在我对框架的学习、研究过程中对我细心指导，耐心讲解，一同工作，教会了我很多软件代码分析的思想、技巧和方法，使得我的能力得到了良好的训练和提高。路浩奇博士在 MC 模拟以及混合工作中对我全力支持，与我一起深入检查混合数据质量并一同讨论了大

量物理细节问题，加深了我对大亚湾物理数据的理解，他也带领着我一步步进入中子本底的分析工作，教会了我数据分析的手段和技巧。陈晓苏硕士经常和我一起讨论计算机和软件基础理论，让我的眼界得到开阔，对软件系统的理解更加专业和深刻，提高了我对现代计算机和软件系统的认知。在此，向这四位亦师亦友的同事表示衷心的感谢！

感谢近代物理系高能组赵政国教授，陈宏芳教授，汪晓莲教授，许咨宗教授，王群教授等老师，他们在治学上严格律己，对学生爱护有加，宽容豁达，一直注重学生创造思维的培养，并为学生创造便利的研究环境以及自由活泼的学术氛围。

感谢高能组的李小男研究员，杨长根研究员，张家文研究员，江群老师，白景芝老师，王岚老师，唐晓老师，杨艺老师，庄红林老师，罗小兰老师，李全林老师，刘蕾老师，李秋菊老师，夏远光工程师，徐俊工程师在我在现场工作和值班期间给予的生活上的悉心照顾和工作中的全力帮助。感谢 Jeff 工程师及夫人的盛情款待。

感谢高能所中微子组的各位同学和同事：占亮，温良剑，李小波，安丰鹏，关梦云，钟玮丽，刘金昌，丁雅韵，王志民，张书华，张清民，肖华林，季晓璐，马烈华，王玲玉，蒋文奇，徐吉磊，索仑，田新春，曹国富，夏冬梅，章飞虹，陈晓辉，赵庆旺，吴振忠，钟真，段斯涵，李登杰，于泽源，刘颖彪，宁哲。他们有的我刚来时领我上路，不断在科研工作中给予我指导和鼓励，有的在日常生活中陪伴和交流。感谢大家几年来的辛勤工作和无私奉献。

感谢大亚湾合作组中各合作单位的同行。感谢合作组美方成员钱鑫，凌家杰，魏华佳，Nakajima.Yasuhiro，Dan.Dwyer，J.Pedro，Brett.Viren，上海交通大学刘江来教授、李高嵩博士在数据分析和软件开发工作中的多次有益讨论。大家共同努力创造了一个多元化背景的工作，交流环境，节奏轻快有趣，给我带来了不可多得的人生经历。

感谢科大高能组的徐敏，许依春，黄炳鑫，刘超，张一飞，张云龙，薛镇，李翠，胡继峰，孙振田，姬长胜，许志蕊，周晓蓉，彭涛，李龙科，王雅迪，闫文成，崔相利，达红玉，殷永光，浦实，徐浩洁，庞龙刚，庞锦毅，刘娟，王驰，王志宏，方馨，王小状等人给我支持和帮助。

感谢我的新老朋友们，吴雯辉，祖明，邹冬，王佳，黄大骏，夏丹，韩丹，

致 谢

黄玥。一路走来，给我莫大的精神支持和鼓励。

在此我要感谢我亲爱的父母，感谢他们一直以来给予的伟大的关爱和无尽的支持，是他们为我指明了成长的道路，我取得的每一分进步都有他们的汗水和功劳，没有他们就没有我今天的一切。

最后，向所有关心和帮助过我，但不能在此一一记述的老师、同学、亲人、朋友表示由衷的感谢。

张一纯

2012年11月12日

在读期间发表的学术论文与取得的研究成果

待发表论文:

1. Y. C. ZHANG et al., “Mixing MC evnets in reactor neutrino experiment” , Chin. Phys. C received.
2. 张一纯, 陈晓苏, 何苗等, “大亚湾反应堆中微子实验 Monte Carlo 样本事例中的 Interleaving 算法研究”, 中国科学技术大学学报, 已接收.

合作组论文:

1. F. P. An et al. [Daya Bay Collaboration], “Observation of electron-antineutrino disappearance at Daya Bay” , Phys. Rev. Lett. 108, 171803 (2012).
2. F. P. An et al. [Daya Bay Collaboration], “A side-by-side comparison of Daya Bay antineutrino detectors” , Nucl. Instr. and Meth Phys. 685, 78-97 (2012).

参加国际会议报告:

1. 2009 年 6 月参加香港中文大学大亚湾国际合作组会议, 报告题目 “Raw Data I/O”。