# 清华大学

# 综合论文训练

# 题目: 新型微通道板型光电倍增管的迭代刻度

- 系 别:未央书院
- 专 业:数理基础科学+工程物理
- 姓 名:刘逸祺

指导教师:续本达 副教授

2024 年 10月 15日

# 关于学位论文使用授权的说明

本人完全了解清华大学有关保留、使用学位论文的规定,即:学校有权保留学 位论文的复印件,允许该论文被查阅和借阅;学校可以公布该论文的全部或部分 内容,可以采用影印、缩印或其他复制手段保存该论文。

(涉密的学位论文在解密后应遵守此规定)

签 名: 刘边祺 导师签名: 32 年达 日 期: \_2024.6.17\_

# 中文摘要

中微子实验依赖 PMT 捕获物理事例产生的少量光子,来对事例进行位置与时间重建,因此 PMT 的增益刻度对提高探测器的能量分辨率具有重要意义。包括JUNO 在内的部分中微子探测器使用由北方夜视生产的 MCP-PMT 作为主要的光电探测器件。该新型号的 MCP-PMT 在使用 ALD 镀膜增大光电子在 MCP 表面的二次倍增系数与 PMT 的收集效率的同时,也为单光电子电荷响应带来了与众不同的"长尾"的结构,为该种 PMT 的增益刻度与实际使用带来了困难。

本研究将于 8 寸 MCP-PMT 提出的 Gamma-Tweedie 混合电荷模型用于 20 寸 型号的刻度,验证了其具有相似的性质,并对刻度方法加入了光强参数项,提高了 对高电荷道址区域的拟合准确度,使得在高光强工作环境下拟合该 MCP-PMT 电 荷谱变得可能。

本研究的电荷基于波形分析方法 FSMP 获得,并寻找了由刻度结果从 Gamma-Tweedie 混合电荷模型在极大似然意义下的多高斯混合模型近似,使得波形分析的 方法迭代变为现实,并也将提高增益刻度的准确度,最终有效提高 JUNO 的能量 分辨率。

关键词:光电倍增管;微通道板;增益刻度;电荷模型;高斯混合模型

# ABSTRACT

Neutrino experiments rely on PMTs to capture photons generated by physical events, aiming to reconstruct the location and the time. Therefore the gain calibration of PMT is of great significance to improve the energy resolution of the detector. Some neutrino detectors, including JUNO, use the MCP-PMT produced by North Night Vision as the main detecting force. This brand-new type of MCP-PMT applies ALD coating to increase the multiplication factor of photoelectrons on the MCP surface, as well as the collection efficiency of PMT. It also brings a unique 'long tail' structure to the single photoelectron charge response spectrum, which makes it difficult to calibrate and use.

In this study, the Gamma-Tweedie mixture charge model proposed with 8-inch MCP-PMT is applied to 20-inch ones and indicated similar behaviors. Furthermore, light intensity is taken into consideration of the calibration method, thus improving the fitting accuracy of the high charge area entries, also making possible utilization under intensive light circumstances.

The charge spectrum of this work is obtained from waveform analysis method FSMP, and the best Gaussian Mixture Model is found as a maximum-likelihood approximation to the Gamma-Tweedie mixture charge model. The iterative method of waveform analysis becomes a reality, and the accuracy improvement of gain calibration could be expected, finally the energy resolution of JUNO as well.

Keywords: PMT; MCP; gain calibration; charge model; Gaussian Mxiture Model

_	<u>``</u>
	~~

第1章 引言	1
1.1 中微子实验	1
1.2 江门中微子实验与 OSIRIS 探测器	2
1.3 新型光电倍增管的特性	5
第 2 章 MCP-PMT 的带光强电荷模型	9
2.1 MCP-PMT 的单光电子电荷模型	9
2.1.1 MCP-PMT 中的物理过程	9
2.1.2 单光电子电荷响应1	1
2.2 分区间数据的拟合方法1	13
2.2.1 分区间方法1	13
2.2.2 最小二乘法1	14
2.2.3 极大似然法1	15
2.3 傅里叶变换带光强电荷模型1	16
2.3.1 任意光电子数电荷谱1	17
2.3.2 复合 Simpson 数值积分1	18
2.3.3 傅里叶变换的准确度与性能1	19
2.4 联合刻度方法2	20
2.4.1 电荷谱截断预处理2	20
2.4.2 联合刻度	22
2.4.3 刻度模型选择 2	23
第3章 结合波形分析的迭代刻度方法2	25
3.1 快速随机匹配追踪算法2	25
3.2 迭代刻度2	27
3.2.1 时间窗筛选电荷 2	27
3.2.2 刻度结果	30
3.2.3 电荷模型多高斯分解: EM 算法 3	30
3.2.4 电荷模型多高斯分解结果 3	35

3.2.5 迭代结果对比	
第4章 结果	
4.1 工作总结	
4.2 问题与改进方向	
插图索引	
表格索引	
参考文献	40
致 谢	42
声 明	
附录 A 外文资料的调研阅读报告	44

# 主要符号表

JUNO	江门地下中微子观测站(Jiangmen Underground Neutrino Obser-
	vatory)
OSIRIS	在线闪烁体内部放射性调查系统(Online Scintillator Internal Ra-
	dioactivity Investigation System)
IBD	逆 <i>β</i> 衰变反应 (inverse beta decay)
LAB	线性烷基苯
PPO	2,5-二苯基噁唑
bis-MSB	1,4-双(2-甲基苯乙烯基)苯
PE	光电子 (photoelectron)
PMT	光电倍增管(photomultiplier tube)
ALD	原子沉积涂层(atomic layer deposition)
МСР	微通道板(microchannel plate)
MCP-PMT	微通道型光电倍增管(microchannel plate photomultiplier tube)
TT	渡越时间(transit time)
TTS	渡越时间展宽(transit time spread)
QE	量子效率(quantum efficiency)
CE	收集效率(collection efficiency)
DFT	有限离散傅里叶变换(Discrete Fourier Transformation)
IDFT	有限离散逆傅里叶变换(Inverse Discrete Fourier Transformation)
FFT	快速傅里叶变换(Fast Fourier Transformation)
IFFT	快速逆傅里叶变换(Inverse Fast Fourier Transformation)
FSMP	快速随机匹配追踪算法(Fast Stochastic Matching Pursuit)
BIC	贝叶斯信息准则(Bayesian information criterion)

# 第1章 引言

#### 1.1 中微子实验

中微子是一种电中性、自旋量子数为 1/2 且几乎不与任何物质发射相互作用的一种轻子,只参与弱相互作用与引力相互作用。作为建立在量子色动力学等学科基础上的粒子物理理论体系,标准模型对粒子历史上许多重要的发现做出了成功的解释或预言。欧洲核子研究中心(CERN)大型正负电子对撞机(LEP)其中的一个粒子探测器 ALEPH 在  $Z^0$  共振态附近测量了单光子事例的反应截面<sup>[1]</sup>,基于正负电子对撞中唯一稳定的弱相互作用粒子为中微子的前提,得出中微子味道种类  $N_v = 3.14 \pm 0.24$  (stat.) ± 0.12 (syst.).在持续观测与其他探测器交叉验证后,可以认为自然界只存在三种味道的中微子,分别为对应轻子  $e, \mu, \tau$  的  $v_e, v_\mu, v_\tau$ ,及其反粒子。



Neutrino oscillation between three generations

图 1.1 中微子振荡

在标准模型中,中微子应当没有质量,然而在太阳中微子、大气中微子等中微子实验中均观测到了中微子振荡现象,即中微子在传播过程中存在味的转化,例如超级神冈在 90% 置信程度下测量了  $v_e \rightarrow v_\mu$  相角与两个质量本征态的质量平方  $\pm^{[2]}$ ,这意味着中微子必须存在质量。

存在幺正矩阵庞蒂科夫-牧-中川-坂田矩阵 (Pontecorvo-Maki-Nakagawa-Sakata matrix或PMNS matrix)描述中微子味与 质量本征态的关系<sup>[3]</sup>:

$$\begin{bmatrix} v_1 & v_2 & v_3 \end{bmatrix} = \begin{bmatrix} v_e & v_\mu & v_\tau \end{bmatrix} U.$$
(1.1)

$$U = \begin{bmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{bmatrix} \times \begin{bmatrix} e^{i\alpha_1/2} & 0 & 0 \\ 0 & e^{i\alpha_2/2} & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$
(1.2)

其中引入混合角 $\theta_{ij}$ ,  $i < j 与 s_{ij} = \sin \theta_{ij}$ ,  $c_{ij} = \cos \theta_{ij}$ ,  $\delta$ 为 CP 破坏相角。(1.2) 除了用来描述尚且未知是否成立的马约拉纳性的  $\alpha_1, \alpha_2$  部分与卡比博-小林-益川矩阵 (Cabibbo–Kobayashi–Maskawa matrix 或 CKM matrix)形式一致,剩余部分可以 写成:

$$U_{res} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{bmatrix} \begin{bmatrix} c_{13} & 0 & s_{13}e^{-i\delta} \\ 0 & 1 & 0 \\ -s_{13}e^{i\delta} & 0 & c_{13} \end{bmatrix} \begin{bmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$
 (1.3)

*θ*<sub>12</sub>, *θ*<sub>23</sub> 都较大而 *θ*<sub>13</sub> 较小,然而正是 *θ*<sub>13</sub> 的存在,中微子震荡中的 CP 破坏才能够显现。太阳中微子与大气中微子分别能够测量质量平方差:

$$\Delta m_{\rm sol}^2 = m_2^2 - m_1^2 \tag{1.4}$$

$$\Delta m_{\rm atm}^2 = \left| m_3^2 - m_1^2 \right| \approx \left| m_3^2 - m_2^2 \right| \tag{1.5}$$

目前 sin  $2\theta_{13}$  与  $\delta$  还需要精细测量,且只有 (1.4)确定了两个质量相近的质量 本征态 (2 与 1)的质量顺序,(1.5)并不能确定  $m_3$  质量的排序位置,还存在正质 量顺序  $m_3 > m_2 > m_1$  与倒序  $m_2 > m_1 > m_3$  两种可能。不同的质量顺序与质量 平方差参数对能谱的形状有直接的影响,为了精细测量参数并确定中微子质量顺 序,需要积累足够多的反中微子事例,使用能谱拟合得到置信度到达  $3 \sim 5\sigma$  的物 理结果。

#### 1.2 江门中微子实验与 OSIRIS 探测器

江门地下中微子观测站(Jiangmen Underground Neutrino Observatory,以下简称 JUNO)<sup>[4]</sup> 是位于地表 700 米以下的装载两万吨液体闪烁体的球形探测器,预期



图 1.2 混合角、质量参数与质量顺序对反中微子能谱的影响<sup>[4]</sup>

达到  $\sigma_E/E = 3.02\%/\sqrt{E(\text{MeV})}$  的能量分辨率,主要物理目标包括:

- •利用台山与阳江核电站的反应堆反中微子来确定中微子质量顺序;
- 探测地源与地外源的中微子与反中微子事例,例如太阳中微子、大气中微子、 地球中微子、超新星中微子等;
- •寻找质子衰变  $p \to K^+ + \bar{\nu}$ 、暗物质湮灭致中微子发射等新物理。 JUNO 主要探测的物理反应包括:

$$\overline{\nu}_e + p \to e^+ + n \tag{1.6}$$

$$v + e \to v + e \tag{1.7}$$

$$\nu + p \to \nu + p$$
 (1.8)

(1.6) 是逆 β 衰变反应(inverse beta decay,以下简称 IBD) 最主要的探测事例,(1.7) 与(1.8) 分别为中微子和电子与质子的弹性散射(elastic neutrino-electron/proton scattering 或 eES 与 pES)。

JUNO 中心装载两万吨液体闪烁体的球体称为中央探测器,使用 5000 支打拿 级光电倍增管(dynode photomultiplier tube,以下简称 dynode PMT)与 12612 支微 通道型光电倍增管(microchannel plate photomultiplier tube,以下简称 MCP-PMT) 捕获中微子相互作用伴随的微弱闪烁光,周围使用超纯水包裹并设置水切伦科夫 探测器,顶端使用塑料闪烁体阵列。

液体闪烁体中主要检测介质为线性烷基苯(Linear alkylbenzene 或 LAB),并



图 1.3 JUNO 探测器侧视示意图<sup>[4]</sup>

掺杂 2,5-二苯基噁唑(以下简称 PPO)与1,4-双(2-甲基苯乙烯基)苯(以下简称 bis-MSB)。液体闪烁体的放射性纯度与透明度直接影响噪声水平,决定了探测灵 敏度的极限,例如对于反应堆中微子的放射性纯度要求为1×10<sup>-15</sup> g/g,对太阳中 微子则为1×10<sup>-16~17</sup> g/g。为了检查液体闪烁体是否达标,在罐装入中心探测器前, 需要经过加压氧化铝纯化塔与超纯水萃取塔,掺杂 PPO 与 bis-MSB 后通过水萃取 与反萃取系统,并以 220 nm 与 50 nm 两级过滤器过滤灰尘以及 <sup>238</sup>U 与 <sup>232</sup>Th 的 放射性本底,最终输入在线闪烁体内部放射性调查系统(Online Scintillator Internal Radioactivity Investigation System,以下简称 OSIRIS) 探测器<sup>[5]</sup>进行采数考察。

OSIRIS 在液体闪烁体循环提纯系统期间,对单批次液体闪烁体展开连续观察,通过研究 U/Th 衰变链中<sup>214</sup>Bi –<sup>214</sup>Po 与<sup>212</sup>Bi –<sup>212</sup>Po 的快符合衰变进行纯度分析,直至液体闪烁体纯度达到探测 IBD 反应与太阳中微子的要求。在 JUNO 中心探测器罐装液体闪烁体期间,OSIRIS 将使用从容纳液体闪烁体的亚克力容器顶部添加新液体闪烁体、从底部引出监测完毕的液体闪烁体的方式,对其纯度进行连续地分析,灵敏度将达 IBD 反应探测要求。



图 1.4 液体闪烁体处理系统的流程图

OSIRIS 为直径与高度均为 9.4 m 的圆柱形探测器,中心使用直径与高度均为 3 m 的亚克力容器灌装液体闪烁体,外围框架使用不锈钢并使用水填充。在亚克力 容器周围,安装有 64 个 20 寸 MCP-PMT,在顶部与底部另有 12 个 20 寸 MCP-PMT 用以排除 µ 子。同时,它也具有皮秒脉冲激光系统、放射源与 LED 作为光源的刻 度系统,其激光系统的激光脉冲展宽约 80 ps,光纤长度几乎一致,故光运行时间 差异相较于 MCP-PMT 自身的渡越时间展宽 (transition time spread,以下简称 TTS) 可以忽略。

OSIRIS 的 24 个激光扩散器遍布在不锈钢钢架上,其中 8 个向内安装在亚克 力容器赤道平面上,两侧各 4 个安装在亚克力容器的顶部与底部用来照亮对侧的 PMT,另有 8 个安装在探测器外部框架上。

OSIRIS 探测器已于 2024 年 3 月完成液体闪烁体灌装,完成了 4 次亚克力容器内部液体闪烁体循环,至今积累了近 3 个月的实际数据,其中包括两轮共计 5 天的激光刻度数据。该工作环境与 JUNO 探测器未来的 PMT 工作环境完全一致,因此该刻度数据对 PMT 的刻度具有重要意义。

### 1.3 新型光电倍增管的特性

光电倍增管是大型中微子实验的核心器件,它主要用来捕获光信号,来对中 微子在液体闪烁体或者水中发生相互作用后释放的光子进行探测。PMT 相较于半 导体探测器等其他探测器件,具有灵敏度高的特点:由于产生载流子需要的能量 并不高,且 PMT 的多级放大能够有效倍增载流子的数量,即使光子数量非常少,



图 1.5 OSIRIS 探测器示意图

也能够有效地产生大量电子,在 PMT 读出电路上接收到显著的电压信号。

为了解决依赖国外公司 PMT 进口的高额成本问题,中国科学院高能物理研究 所、北方夜视技术股份有限公司等单位合作设计、研究、改进与生产了新型 MCP-PMT,在其量子效率(quantum efficiency,以下简称 QE)系数与收集效率(collection efficiency,以下简称 CE)系数等方面均做出了改进<sup>[6]</sup>并有效降低了成本。



图 1.6 20 寸 MCP-PMT 与环绕的 4 个 3 寸 MCP-PMT

该 MCP-PMT 的突出特点包括:

- 相较于常见的 dynode PMT,使用具有微小倾角的斜长微通道板(microchannel plate,以下简称 MCP)取代了分离式的多级打拿级,因此可以发生倍增物理 过程的接触点变得连续,PMT 外接高压引线结构得到简化,电子的收集效率 也较高;
- 相对于其他种类的 MCP-PMT,该 PMT 在 MCP 表面引入了主要成分为复合 Al<sub>2</sub>O<sub>3</sub> – MgO 的原子沉积涂层(atomic layer deposition,以下简称 ALD),该 ALD 涂层具有高二次电子倍增系数的特点<sup>[7]</sup>,即电子在该表面激发二次电 子数目的期望较大;
- 单位数目的电子在从光阴极入射该 MCP-PMT 后,能够产生较其他类型 PMT 更多的倍增电子并在阳极被收集(比值定义为 CE),因此理论意义上具有更 优秀的峰谷比与分辨率;
- 4. 由于二次倍增电子的存在,收集到大电荷信号的概率较其他 PMT 高,表现 在电荷谱具有"长尾"结构。



图 1.7 MCP-PMT 与 dynode PMT 单光电子电荷谱对比<sup>[8]</sup>

其中 2. 与 3. 为该 MCP-PMT 的突出优点,并在测试<sup>[9]</sup> 中得到了验证。同时, 相较于传统 PMT 可以使用高斯分布描述的对称型单光电子响应电荷谱, 4. 呈现的 非对称长尾结构为使用该 PMT 带来了困难:

- 分压实验研究<sup>[10]</sup> 中发现 MCP 对低能电子的增益远小于高能电子,揭示了 单光电子电荷谱不能够认为只有单一的增益模式;
- 只截取主峰部分使用高斯分布拟合,则不能够充分利用长尾部分的信息,对 提高统计量与能量分辨率没有帮助,且在没有充分物理认知的前提下直接使

用主峰均值定义增益不具有坚实的理论依据;

 在光强较强时,电荷谱将有多个峰结构,大电荷道址区域由电荷长尾与整数 倍主峰共同贡献,直接拟合峰结构使得光电子数量估计有偏,进而引入能量 偏差。

JUNO 采购了 15000 支 MCP-PMT 与 5000 支 dynode PMT,其中 20 寸 PMT 的 光阴极覆盖率为 75.2%,在先行探测器 OSIRIS 上安装的 76 个大 PMT 也均为 20 寸 MCP-PMT,因此 20 寸 MCP-PMT 的电荷增益刻度对提高探测器的能量分辨率 具有重大的意义。



相较于其他应用 dynode PMT 的中微子探测器,JUNO 与 JNE 等探测器需要 对新 PMT 展开仔细地增益刻度,尤其需要利用适合的电荷模型,以增进对该类型 PMT 的理解与经验,为 PMT 的信号读出提供坚实的物理基础,进而达到提升能量 分辨率的目标。

# 第2章 MCP-PMT 的带光强电荷模型

#### 2.1 MCP-PMT 的单光电子电荷模型

#### 2.1.1 MCP-PMT 中的物理过程

PMT 的探测效率取决于两个关键因素: QE 系数与 CE 系数。其中, QE 为 PMT 光阴极发射的光电子数量与入射光子数量的比值, CE 为经过倍增后, 阳极收集到 的电子数量与入射光电子数量的比值。

为了提高 QE 系数,在 JUNO 中使用的 20 寸 MCP-PMT 具有以下特征:

- 上半椭圆球面装配发射光阴极,下半椭圆球面装配反射光阴极,实现接近4π
   立体角光电转换范围
- 上下具有两块平行的 MCP



图 2.1 MCP-PMT 的结构设计图<sup>[6]</sup>

在光子在 MCP-PMT 内部光阴极与材料发生相互作用产生光电子后,电子在 椭圆形玻璃球壳内的电场作用下,向 MCP 迁移。PMT 需要使用屏蔽线圈来抵消 地磁场的影响,使得光电子能以尽可能高的效率到达 MCP。MCP 使用加高压的长 通道实现电子加速,并以一定的小角度倾斜,使得电子容易与壁材料发生相互作 用实现倍增,又不使二次电子飞行轨迹与壁夹角过大以致电子难以逸出。MCP 的 长通道截面积占横截面的比例称为开口比 *A*。

MCP-PMT 的主要噪声来源为离子反馈<sup>[12]</sup>。即使 MCP-PMT 需要内部近似真 空的工作条件,但仍存余少许的气体分子。气体分子被电离,在高压电场中与倍

9



增电子运动方向相反,既可能撞击壁产生二次电子被阳极收集,也可能反向加速 直至撞击光阴极,由于从产生到被收集的时间(渡越时间)较电子长,故形成后脉 冲,且还能够对光阴极材料造成损伤,如图 2.2 所示。如果延长加速通道或者提高 高压,这些脉冲也将得到放大,为了有效提高信噪比,常常使用两层 MCP 并使它 们的通道倾角对称。即如图 2.2 中 MCP 1 与 MCP 2 的组合。由于在运动方向上引 入了反角的设计,速度较小的反馈离子更难以进入与产生处对称的另一级 MCP, 从而有效提高信噪比,同时对电子倍增的削弱作用不显著,保证仍然具有较大的 增益。

为了提高 CE 系数,最初在 MCP 入射处使用涂层(镍铬电极),使得入射的光电子可能能够在该表面进行倍增。在后来的设计中,采用了  $Al_2O_3 - MgO$  复合涂层作为 ALD 涂层材料的方案,起到提高二次倍增电子数  $\delta_{ts}$ 与延长寿命的作用。

在先前基于 Furman 模型<sup>[13]</sup>的模拟<sup>[14]</sup>中,光电子与涂层发生的相互作用被分为三类:

- 在涂层表面发生弹性散射;
- 进入涂层原子,发生散射而脱离涂层;
- •与涂层作用产生若干个能量较低的二级电子,亦称为真二次发射电子。

在三种模式中,发生前两种散射模式的单个光电子均只能产生一个光电子,而 发生二次倍增的电子则可以产生自然数个真二次发射电子,其个数 $\delta_{ts}$ 与入射电子 能量 E、与涂层表面法线夹角  $\theta$  存在依赖,由 Furman 模型给出:

$$\delta_{ts}(E,\theta) = \hat{\delta}(\theta) D\left[\frac{E}{\hat{E}(\theta)}\right].$$
(2.1)

其中<sup>+</sup>代表 · 函数项的极值,  $D(\cdot)$  为一个满足 D(1) = 1, D'(1) = 0 的尺度函



图 2.3 光电子与 MCP-PMT 的三种相互作用形式

数, Furman 等人选择:

$$D(x) = \frac{sx}{s - 1 + x^s}.$$
 (2.2)

其中 *s* > 1 为可调节变量。而对于二次倍增电子数服从的分布, Furman 等人给出了泊松分布与二项式分布两种选择,其中二项式分布需要依据材料,选择对应的随机变量上限 *M*,在实际未知材料性质的情况下使用泊松分布是更合适的。

#### 2.1.2 单光电子电荷响应

在 8 寸 MCP-PMT 的单光电子电荷响应研究<sup>[8]</sup>中,电子进入 MCP 倍增的电荷 响应采用 Gamma 分布描述。这是因为在每一次倍增中,倍增电子的个数均服从泊 松分布,并依次级联,这种级联泊松最终可以用 Gamma 分布相较于其他常用分布 (例如高斯分布)更好地近似。相较于常用的高斯分布,Gamma 分布的另一个好处 是在非正区间上概率密度严格为 0。

同时,真二次发射电子与其他未在涂层表面倍增而进入 MCP 倍增的电子相 比,除了能量与到达阳极时间外没有本质区别,其电荷响应使用 Gamma 分布描 述。在 8 寸的仿真中,真二次发射电子的电荷被发现服从同一个 Gamma 分布 Γ(α<sub>ts</sub>, β<sub>ts</sub>)。在 2.1.1 节中, Furman 模型给出真二次发射电子数目服从泊松分布的建 议,因此真二次发射电子的电荷响应部分可以表示为:

$$\left.\begin{array}{l}
\left. Q_{ts} = \sum_{i=1}^{\delta_{ts}} Q_{i} \\
\left. \delta_{ts} \sim \pi(\lambda) \\
Q_{i} \sim \Gamma(\alpha_{ts}, \beta_{ts}) \end{array}\right\} \implies Q_{ts} \sim \operatorname{Tw}_{\xi}\left(\mu(\theta_{ts}), \sigma^{2}(\theta_{ts})\right), \quad \theta_{ts} = (\lambda, \alpha_{ts}, \beta_{ts}). \quad (2.3)$$

即服从泊松分布个数的 Gamma 分布之和等价于指数参数 $\xi \in (1,2)$ 的 Tweedie 分布。Tweedie 分布是一种特殊的指数族分布,由 $\xi$ 调节分布的种类,由 $\mu$ 代表期 望,由 $\phi$ 线性调节方差:

$$Y \sim \operatorname{Tw}_{\xi}(\mu, \sigma^2) \implies \begin{cases} \operatorname{E}[Y] = \mu \\ \operatorname{var}[Y] = \phi \mu^{\xi} \end{cases}$$
 (2.4)

(2.3)对比(2.4),可得映射关系:

$$\xi = \frac{2 + \alpha_{ts}}{1 + \alpha_{ts}} \tag{2.5}$$

$$\mu = \lambda \cdot \frac{\alpha_{ts}}{\beta_{ts}} \tag{2.6}$$

$$\phi = (1 + \alpha_{ts}) \left[ \lambda \alpha_{ts} \beta_{ts}^{(1 + \alpha_{ts})} \right]^{-\frac{1}{1 + \alpha_{ts}}}$$
(2.7)

对于只在表面弹性散射或进入涂层材料原子核但被散射出涂层的电子,以及 没有与涂层相互作用而直接进入 MCP 进行倍增的电子,它们没有在涂层表面进行 倍增,彼此的区别在于飞行径迹与 TT 不同,从而对 TTS 产生贡献。它们在 MCP 中的倍增行为相同,电荷增益可以使用同一个 Gamma 分布  $\Gamma(\alpha, \beta)$  来描述。该分 布作为单光电子进入 MCP 倍增的电荷结果,直接表示了 MCP-PMT 对单光电子的 增益水平,因此可以定义增益  $Q_1 \equiv \frac{\alpha}{\beta}$ 与分辨率  $\eta \equiv \frac{\sqrt{\alpha/\beta^2}}{\alpha/\beta} = \frac{1}{\sqrt{\alpha}}$ 。

至此, MCP-PMT 的单光电子响应模型可以表示为:

 $S(q) = p_0 \Gamma(\alpha, \beta) + (1 - p_0) \operatorname{Tw}_{\xi} \left( \mu(\theta_{ts}), \sigma^2(\theta_{ts}) \right), \quad \theta_{ts} = (\lambda, \alpha_{ts}, \beta_{ts}).$ (2.8)

其中 *p*<sub>0</sub> 代表主峰的占比,它可以认为是所有未在涂层上进行真二次电子发射的原始电子的比例,它与开口比、涂层材料性质等物理因素直接相关,可以视为级联伯努利事件的概率密度:

- 1. 电子以 p<sub>ch</sub>的概率进入 MCP 倍增, 以 1-p<sub>ch</sub>的概率撞击 MCP 表面 ALD 涂层
- 撞击 MCP 上表面的电子以 p<sub>ts</sub> 的概率进行真二次电子发射,以1 p<sub>ts</sub> 的概 率进入 MCP 进行倍增

最终可以认为  $p_0 = (1 - p_{ch}) \cdot (1 - p_{ts})$ 。1. 中  $p_{ch}$  应满足  $p_{ch} < A$ ,因为存在 电场透镜效应,电场线在平面上的孔隙处呈扩散状,对电子的运动有"发散"的作用,使得电子进入孔隙的概率变小<sup>[15]</sup>,如图 2.4 所示。

8 寸研究中,指出 2. 中有  $p_{ts} \sim 90\%$  且在典型参数下应有  $p_{ts} > 75\%$ 。依据经验应有  $p_0 < A$ ,关于 MCP-PMT 延迟脉冲的研究<sup>[11]</sup>指出在 A = 65% 的情况下,应有  $p_0 = 55\%$ 。



图 2.4 孔隙对电子束的负透镜效应[15]

在8寸研究中,对真二次倍增电子相对主峰的均值与标准差做出了限制建议:

$$\mu_{ts} = \frac{\alpha_{ts}}{\beta_{ts}} \in [0.3, 0.7] Q_1 \tag{2.9}$$

$$\sigma_{ts} = \sqrt{\frac{\alpha_{ts}}{\beta_{ts}^2}} \in [0.05, 0.3] Q_1 \tag{2.10}$$

本研究中拟合均采取这一建议区间。

# 2.2 分区间数据的拟合方法

由于一次激光刻度取数中,电荷计数能够达到数万甚至数十万,直接对每个 计数进行拟合的内存消耗与效率极不理想。如果能够以尽可能接近电荷模型原始 概率分布的分区间方式,将需要计算的计算点减少至数十至数百个,将有约三个 数量级的效率提升。

#### 2.2.1 分区间方法

常用的分区间包括平方根方法、Sturges 公式、Rice 准则、Scott 准则、Terrell-Scott 准则、Doane 公式、Freedman-Diaconis 准则等。

在表 2.1 中, Freedman-Diaconis 准则既在系数部分使用了四分位距 IQR,考虑 了随机变量分布的偏度而避免像 Doane 公式一样直接计算偏度与其标准差,降低 了计算难度,又在幂次上同样渐进极小均方误差的积分<sup>[16]</sup>,即  $k \propto n^{1/3} \simeq h \propto$ 

方法	区间数	区间宽	优势与适用范围
平方根方法	$k = \lceil \sqrt{n} \rceil$		简单
Sturges 公式	$k = \lceil \log_2 n \rceil + 1$		正态分布; 样本数量适中
Rice 准则	$k = \lceil 2\sqrt[3]{n} \rceil$		简单
Scott 准则		$h = \frac{3.49\hat{\sigma}}{\sqrt[3]{n}}$	简单; 正态分布
Terrell-Scott 准则	$k = \sqrt[3]{2n}$		渐进极小均方误差的积分; 适用于非正态分布
Doane 公式	$k = 1 + \log_2(n) + \log_2\left(1 + \frac{ g_1 }{\sigma_{g_1}}\right)$ $g_1 \ \text{@B}, \ \sigma_{g_1} = \sqrt{\frac{6(n-2)}{(n+1)(n+2)}}$		基于 Sturges 公式; 适用于非正态分布
Freedman-Diaconis 准则	$ \circ_1$ $\mathbf{v}$ $(n+1)(n+3)$	$h = 2 \frac{IQR(x)}{\sqrt[3]{n}}$ IQR 四分位距	对离群值不敏感; 适用于非正态分布

表 2.1 多种分区间方法的比较

*n*<sup>-1/3</sup>,确保在样本统计量大的情况下能够以最优的幂次还原原始概率分布。综上所述,本研究中使用 Freedman-Diaconis 准则来对电荷计数进行分区间,再进行拟合。

### 2.2.2 最小二乘法

不妨约定 q 为电荷自变量,  $q_i$  表示第 i 个直方图区间的电荷闭区间,  $n_i$  表示第 i 个直方图区间的实际计数,或称为频数,总事例数  $N = \sum_i n_i$ ;  $e_i$  表示第 i 个直方图区间的预期计数,或称为预期频数,可由下式给出:

$$e_i(\boldsymbol{\theta}) = N \int_{\inf q_i}^{\sup q_i} \mathrm{d}q \cdot f(q; \boldsymbol{\theta})$$
(2.11)

其中  $f(q; \theta)$  代表在给定参数集  $\theta$  时电荷谱的条件概率密度函数。 将直方图的每一个区间都视为一次测量,测量结果为频数,可以定义  $\chi^2(\theta)$ :

$$\chi^{2}(\theta) \equiv \sum_{i} \frac{\left[n_{i} - e_{i}(\theta)\right]^{2}}{\sigma_{i}^{2}}$$
(2.12)

其中  $\sigma_i^2$  代表第 *i* 个直方图区间的频数的方差。为了计算方差,使用预期频数  $e_i$  替代方差  $\sigma_i^2$  或使用频数  $n_i$  替代方差  $\sigma_i^2$  是两种可行的做法。

如果每个区间的计数  $n_i$  相较总事例数 N 占比小,可以认为渐进总事例数 N无穷、发生概率 p 无穷小而 Np 有限的泊松分布  $\pi(n_i; Np)$ ,因此有近似  $\sigma_i^2 = n_i$ :

$$\chi^{2}(\boldsymbol{\theta}) = \sum_{i} \frac{\left[n_{i} - e_{i}(\boldsymbol{\theta})\right]^{2}}{n_{i}}$$
(2.13)

如果将每个区间视为各自独立且具有较大数学期望的计数结果,则可近似认

为服从高斯分布,因其误差可以认为仅仅由统计涨落贡献,故具有方差与数学期 望相等的特征,即 $\sigma_i^2 = e_i$ :

$$\chi^{2}(\theta) = \sum_{i} \frac{\left[n_{i} - e_{i}(\theta)\right]^{2}}{e_{i}}$$
(2.14)

根据大数定律,这两种近似在样本量趋近无穷时,具有相同的极限,但它们 在有限样本量下的取舍则有所不同:

- (2.13) 需要每个区间事例数占比都较小;
- (2.14) 需要每个区间事例数足够多, 尤其不能为空, 否则不能良定义。

在本研究中,发现按照 Freedman-Diaconis 准则划分的直方图中,计数最多的 区间计数可达总计数的 20%,因此(2.13)假设并不成立,且由于每个区间计数较 多,选用(2.14)作为极小化的函数是合适的。

关于拟合结果的误差分析,最准确的办法是对一系列的自变量  $\theta$  做 Monte Carlo 模拟,得到  $\chi^2$  的概率密度分布,从而得到拟合得到  $\chi^2$  的置信区间,作为 该模型的合理性定量依据。因为实际使用中使用蒙卡验证较为困难,故可以使用  $\chi^2$ /ndf 作为评判拟合优度的标准。

定义自由度 ndf 为数据点数与模型参数个数之差。(2.12) 指出  $\chi^2$  代表数据点 偏差平方与方差的比值,它与自由度 ndf 的比值代表平均意义上每个自由度拟合 的相对偏差,在数据点较多的情况下也可近似为每个数据点拟合的相对偏差。一般可以认为  $\chi^2$ /ndf  $\leq$  1 为极优的拟合,  $\chi^2$ /ndf  $\leq$  3 ~ 10 为较好的拟合。

#### 2.2.3 极大似然法

同样约定 q 为自变量,  $q_i$  表示第 i 个直方图区间的自变量闭区间, 共有 m 个区间,  $n_i$  表示第 i 个直方图区间的频数,总事例数  $N = \sum_{i=1}^{m} n_i$ ,  $n = (n_1, n_2, \dots, n_m)$ 代表各区间的频数向量,  $e_i$  表示第 i 个直方图区间的预期频数,各区间预期频数使用 e 表示,可由下式给出:

$$e_i(\theta) = N \int_{\inf q_i}^{\sup q_i} dq \cdot f(q;\theta)$$
(2.15)

可见存在函数满足参数集 $\theta$ 至预期频数e之间的双射。各区间事例数n可以

视为总数为 N 的 m 次多项式分布结果,在该处的概率分布密度为:

$$P(\boldsymbol{n}, \boldsymbol{e} | \boldsymbol{\theta}) = P(\boldsymbol{n} | \boldsymbol{\theta})$$

$$= \frac{N!}{n_1! \cdots n_m!} \left[ \frac{e_1(\boldsymbol{\theta})}{N} \right]^{n_1} \cdots \left[ \frac{e_m(\boldsymbol{\theta})}{N} \right]^{n_m}$$

$$= \frac{N!}{n_1! \cdots n_m!} \cdot \frac{\prod_{i=1}^m e_i^{n_i}(\boldsymbol{\theta})}{N^N}$$
(2.16)

对 (2.16) 取对数得:

$$\ell = \ln P(\boldsymbol{n}|\boldsymbol{\theta})$$
  
=  $\sum_{i=1}^{m} n_i \ln e_i(\boldsymbol{\theta}) + \ln N! - \sum_{j=1}^{m} \ln n_j! - N \ln N$   
=  $\sum_{i=1}^{m} n_i \ln e_i(\boldsymbol{\theta}) + C$  (2.17)

其中 C 代表与自变量 θ 无关的常数,在极大似然过程中对结果没有影响。极 大似然法相较于最小二乘法,不需要做出近似假设,且对各区间频数 n<sub>i</sub> 没有要求, 始终具有良定义。它的缺陷在于对数似然函数没有办法得到直观的拟合优度数据。 本研究暂时使用最小二乘法,但同样实现了极大似然法拟合。

### 2.3 傅里叶变换带光强电荷模型

出于三个原因,考虑光强的刻度方法有被引入的必要性:

- 由于本研究所考虑的 MCP-PMT 相较于传统打拿级 PMT,具有电荷长尾的效应,实际电荷谱上较主峰能道高的区域由单光电子模型中 Tweedie 成分与多光电子的主峰共同贡献;
- 对于低光强的刻度数据,大统计量的0电荷计数被排除在单光电子谱刻度 外,没有得到充分利用;
- 3. 在 PMT 的刻度与实际运行中,可能遇到光强较强以至于多个光电子情形不可忽略甚至占主要成分的情况。

拟合引入 0 电荷计数,能够极大提高统计量,得到相对误差较小的光强,从 而在电荷谱的能道较高范围内尽可能消除多光电子的贡献对长尾参数的影响,提 高目标 MCP-PMT 的增益刻度准确性。

#### 2.3.1 任意光电子数电荷谱

光电子个数服从期望为μ的泊松分布,即:

$$P(PE = n) = \pi(n; \mu) = \sum_{i=0}^{+\infty} \frac{\mu^n}{n!} \cdot e^{-\mu}$$
(2.18)

定义光学量 occupancy 为由至少一个 PE 贡献的:

occupancy = 
$$\frac{\mathbf{N}_{\geq 1PE}}{\mathbf{N}_{\text{total}}}$$
  
=  $\frac{e^{\mu} - 1}{e^{\mu}}$  (2.19)  
=  $1 - e^{-\mu}$ 

当光电子个数为0时,考虑到没有光电子时的电荷应该严格为0,电荷谱密度 可以认为是:

$$S_{\text{OPE}}(q) = N_{\text{OPE}} \cdot \delta(q) = \pi(0; \mu) \cdot \delta(q)$$
(2.20)

当光电子个数为正整数  $k \in \mathbb{N}^+$  时,电荷谱密度等效于 k 个单光电子谱卷积:

$$S_{k}(Q = \sum_{i=1}^{k} q_{i}) = \pi(k; \mu) \cdot \underbrace{\int_{-\infty}^{+\infty} dq_{1} \cdot S(q_{1}) \cdots \int_{-\infty}^{+\infty} dq_{k} \cdot S(q_{k})}_{k}$$

$$= \pi(k; \mu) \cdot \underbrace{S(q) \otimes \cdots \otimes S(q)}_{k}.$$
(2.21)

其中 *S*(*q*) 为单光电子电荷谱概率密度函数。傅里叶变换 *F* 能够将卷积变为另一个域上的直积:

$$\mathcal{F}[f(t) \otimes g(t)] = \mathcal{F}[f(t)] \cdot \mathcal{F}[g(t)].$$
(2.22)

因此对(2.21)做傅里叶变换并对 k ∈ N<sup>+</sup> 求和即得正整数个光电子时的电荷 谱密度为:

$$\tilde{S}_{\geq 1\text{PE}}(p) = \sum_{k=1}^{\infty} \pi(k;\mu) \cdot \tilde{S}^{k}(p).$$
(2.23)

其中 ~(p) 代表函数 · 经过 F 后在另一个数域 {p} 上的数值。考虑 (2.23) 补全

0个光电子时的无穷级数求和:

$$\begin{split} \tilde{S}_{\ge 1\text{PE}}(p) &= \sum_{k=0}^{\infty} \pi(k;\mu) \cdot \tilde{S}^{k}(p) - \pi(k;\mu) \\ &= \sum_{k=0}^{\infty} \frac{e^{-\mu}}{k!} \cdot \left(\mu \tilde{S}(p)\right)^{k} - e^{-\mu} \\ &= e^{\mu(\tilde{S}(p)-1)} - e^{-\mu} \\ &= e^{-\mu} (e^{\mu \tilde{S}(p)} - 1). \end{split}$$
(2.24)

通过傅里叶变换,能够将任意多个光电子的卷积分布转化为另一个定义域上 的直积,并借助泊松分布的性质实现解析求和,既不需要截断无穷级数求和,也 不需要计算卷积,是完全无损失的理论方法,因此得到的参数应是无偏的。

考虑 0 电荷处的概率密度则需要综合考虑 0 个光电子与单光电子中 Tweedie 成分的贡献:

$$S_{\text{total}}(q = 0) = S_{\ge 1\text{PE}}(q = 0) + S_{0\text{PE}}(q = 0)$$
  

$$= \sum_{k=1}^{\infty} \pi(k; \mu) \cdot S^{k}(q = 0) + \pi(0; \mu)$$
  

$$= \sum_{k=0}^{\infty} \pi(k; \mu) \cdot S^{k}(q = 0)$$
  

$$= \sum_{k=0}^{\infty} \frac{e^{-\mu}}{k!} \cdot [\mu S(q = 0)]^{k}$$
  

$$= \sum_{k=0}^{\infty} \frac{e^{-\mu}}{k!} \cdot [\mu(1 - p_{0})e^{-\lambda}]^{k}$$
  

$$= e^{\mu \left[ (1 - p_{0})e^{-\lambda} - 1 \right]}.$$
  
(2.25)

# 2.3.2 复合 Simpson 数值积分

借助 2.3.1 所述的方法,一组参数已经能够利用 FFT 与 IFFT 映射到一系列采样 点上的概率密度函数值。为了得到原电荷谱上特定区间内的预测频数,需要使用数 值积分方法。常用的数值积分方法包括:矩形法,Lagrange 插值法与 Newton-Cotes 插值法等,在此不展开介绍各方法的应用场景与变形。

复合 Simpson 积分法是 Newton-Cotes 插值法的特殊形式,特别适用于间隔给

定的特定区间上的数值积分。其具体计算方式如下:

$$\int_{a}^{b} f(x)dx = \sum_{i=1}^{n/2} \frac{h}{3} \left[ f(x_{2i-2}) + 4f(x_{2i-1}) + f(x_{2i}) \right] - \frac{f^{(4)}(\xi_i)}{90} h^{5}$$

$$= \frac{h}{3} \left[ f(x_0) + 4f(x_1) + 2f(x_2) + \dots + 2f(x_{n-2}) + 4f(x_{n-1}) + f(x_n) \right] - \frac{h^2}{90} \sum_{i=1}^{n/2} f^{(4)}(\xi_i)$$

$$= \frac{h}{3} \left[ f(x_0) + 4 \sum_{i=1}^{n/2} f(x_{2i-1}) + 2 \sum_{i=1}^{n/2-1} f(x_{2i}) + f(x_n) \right] - \frac{h^2}{90} \sum_{i=1}^{n/2} f^{(4)}(\xi_i).$$
(2.26)

其中  $x_i$  为等间距 h 的序列满足  $a = x_0 < x_1 < \cdots < x_n = b$ 。考虑连续函数  $f^{(4)}(\xi)$  的介值定理, 应有:

$$\frac{2}{n}\sum_{i=1}^{n/2} f^{(4)}(\xi_i) = f^{(4)}(\xi), \quad \xi \in \left[\min(\xi_i), \max(\xi_i)\right].$$
(2.27)

依此给出(2.26)的误差项为  $-\frac{nh^5}{180}f^{(4)}(\xi) = -\frac{(b-a)h^4}{180}f^{(4)}(\xi)$ , 与原函数的四阶导数成正比,因此当且仅当对于三次及以下多项式,其误差严格为0,或称为代数精度为3。

复合 Simpson 积分法只需要使用特定点的函数值,在细分区间较多、每段区间较平滑的函数上,它的表现良好,因此在本实验中采样点函数值到原电荷谱区间频数的预测均使用复合 Simpson 积分法计算。为了尽可能减小误差,需要采样点间隔 h 尽可能小。Freedman–Diaconis 准则给出的区间宽 h ∝ n<sup>-1/3</sup>,对于统计量越大的电荷刻度数据,原电荷谱的分区间数越大、间距越小,复合辛普森积分法效果越好。

#### 2.3.3 傅里叶变换的准确度与性能

计算机并不能够对连续的 *S*<sub>total</sub>(*q*) 函数进行傅里叶变换,只能够使用有限离散 傅里叶变换(Discrete Fourier Transformation,以下简称 DFT)对其进行近似,并使用 逆有限离散傅里叶变换(Inverse Discrete Fourier Transformation,以下简称 IDFT)。 应当从准确度与性能方面进行考量。

DFT 与 IDFT 结合的方法能够在离散距离趋近于无穷小时的极限为连续傅里 叶变换与其逆变换,因此当采样间隔足够小时,即可实现理想的近似。合适的间 隔应当由定理 2.1 决定:

**定理 2.1 (Nyquist-Shannon 采样定理):** 对于具有带宽限制(即傅里叶变换后频率 在有限区域以外为零)例如频率上限为 *f*<sub>s</sub> 的连续信号,使用周期冲激序列进行采 样,若要完全还原初始连续信号的信息而不发生混叠,则应以不低于 2f<sub>s</sub>的频率进行采样。

在实际中,由于连续函数的傅里叶变换常常不具有带宽限制,可以设定频率 阈值下限,只需要满足采样频率至少高于该阈值对应频率的二倍,原始连续信号 的主要信息就能够被还原。在本研究中,将依据 Freedman–Diaconis 准则划分的电 荷谱区间继续细分至 32 份后,再增加区间已经没有可以显著观测的变化,可以认 为已经理想地还原了原始连续电荷谱的信息。

快速傅里叶变换(Fast Fourier Transformation,以下简称 FFT,逆算法同理简称 IFFT) 是一种 DFT 的高效算法,利用了 DFT 中系数的对称性与周期性,能够将朴素 DFT 算法的 *O*(*n*<sup>2</sup>) 时间复杂度降低到 *O*(*n* log *n*),对于本研究中每个 MCP-PMT 数万至数十万的统计量表现出显著的优势。

综上所述,傅里叶变换刻度方法的本质优势为避免卷积与无穷级数求和的时间复杂度与精度损失,而使用较多的采样点,同时满足数值积分、FFT 与 IFFT 准确度的要求,只需要付出较为廉价的内存代价,是理想的升级。

### 2.4 联合刻度方法

#### 2.4.1 电荷谱截断预处理

PMT 中存在部分巨大脉冲,是由 μ 子穿过 PMT 玻璃时发射的切伦科夫光与 天然放射性本底造成的<sup>[17]</sup>。具体地,根据脉冲幅值,可以将划分为 PMT 的噪声 计数主要来源分为三类<sup>[18]</sup>:

- 0~10mV, 主要由热致电子发射(暗噪声)贡献;
- 10~100 mV, 主要由天然放射性贡献;
- 100~500 mV, 主要由 µ 子致切伦科夫光贡献。

在实际电荷谱中,除了主要事例(暗噪声、激光等)对应的电荷响应,还存在 由 μ子与天然放射性本底贡献的电荷,且集中在电荷较大的区域。由于其对主峰 后能道较高区域有贡献,因此需要引入一个电荷谱拟合时的上限截断。

在还原山体、玻璃器材中包含的<sup>238</sup>U/<sup>232</sup>Th/<sup>40</sup>K含量与未灌装水之前的实验 条件下,能够得到模拟结果,如图 2.5 所示。

对于不同的数据集应当采取不同的策略。对于暗噪声采数,暗噪声计数率约 10~50 kHz,但由于阈值较高(通常有至少3个或7个或20个PMT同时过阈才记 录波形的设置),按照暗噪声50 kHz、偶然符合时间窗50 ns 的假设进行估计,四



图 2.5 Dynode PMT 与 MCP-PMT 各本底计数率<sup>[17]</sup>

重及以上偶然符合计数率有 f<sub>coin</sub> ≪1 Hz,实际触发率较低,因此暗噪声的电荷计数主要由三重及以下符合贡献。依据图 2.5 可认为在 4 倍单光电子(photoelectron 或 PE)水平下暗噪声触发率接近其他本底的 20 倍左右,可以认为成分较为单一。

对于激光等光源采数,例如 OSIRIS 现场激光刻度触发率为 1.5 kHz,单脉 冲的期望水平为  $\mu = 0.01 \text{ PE}^{[5]}$ ,对于 1.2 节中位于赤道平面上的激光调制位于 75% ~ 78% 强度,光电倍增管平均光子探测效率为 29.1%<sup>[4]</sup>,可以估计单激光通 道下 PMT 接受 PE 频率期望为 3.3 Hz,即 PE 光子数服从强度为 3.3 的时齐泊松过 程,这与现场单激光通道采数 ~ 2 × 10<sup>3</sup> PE / 12 m 的采数率符合。此时暗噪声也 属于本底,但信噪比相较暗噪声采数仍较优,此时预计 5 PE 较为合适。

已经完成的模拟研究中对于 PMT 工作环境的考虑并不完全,仍然需要考虑灌装水与灌装液体闪烁体后的其他本底水平。灌装后水能够进一步屏蔽周围岩石、玻璃的放射性传播,因此天然放射性项在已经灌装水的探测器中影响将被削弱。

液闪中的 U/Th 等元素及其衰变链的活度受到严格控制,但在循环过程中会引入氡,经过 3 次液体闪烁体循环后,OSIRIS 内部 <sup>214</sup>Bi – <sup>214</sup> Po 特征事例的事例率 ~ 2×10<sup>-3</sup> Hz / 20 m<sup>3</sup> 升至 ~ 1×10<sup>-1</sup> Hz / 20 m<sup>3</sup>,在 OSIRIS 探测器中相较于激光刻度采数的频率有 30 倍差距,但在 JUNO 中央探测器中这一比例将更微弱,因此尤其是 JUNO 中央探测器需要在液体闪烁体循环后氡几乎消耗殆尽时再进行激光采数。

本研究中使用 OSIRIS 的激光刻度数据,综合考虑各本底,选取较简单的5倍 主峰截断来近似5倍 PE 截断。

#### 2.4.2 联合刻度

实际刻度中发现对于同一通道(PMT),不同光强与激光入射方向的采数,均 有相似的形状与 PMT 参数,而这种特性随通道(PMT)而异,如图 2.6 与 2.7 所示, 因此做出了光电子电荷响应与光强与激光入射方向耦合变量之间弱相关的猜想。













考虑到每一次激光采数所开启的激光扩散器几何位置均不同,则每一次激光 采数中各通道的光强与激光入射角度两组变量为各不相同的耦合变量组,可以考 虑该耦合变量组与光电子响应电荷谱的相关性。将 2.1.2 节中 PMT 参数  $\theta_{PMT} = (p_0, \alpha, \beta, \lambda, \alpha_{ts}, \beta_{ts})$ 分为每个采数 run 各自独立(称为 separated fit)与所有 run 均使 用同一组(称为 joint fit),分别做电荷谱刻度。

对于 separated fit,每一组电荷谱需要使用 7 个自由变量拟合;对于 joint fit, 所有电荷谱需要同时拟合,共需要 6 个 PMT 参数与每个 run 光强之和的自由度。 对于本研究中使用的 OSIRIS 激光刻度,存在 30 个 run 的数据,故直接拟合需要 36 个自由度,为了有效改善拟合速度与质量,将 joint fit 流程改进为:

- 波形分析得到的存在 PE 波形数与所有波形数比值作为式 (2.19) 中的 occupancy 初值输入;
- 对于所有电荷谱,使用 1.中的光强值拟合得到最优的  $\theta_{PMT}$ ;
- 对于每一个 run 的电荷谱, 使用 2. 中的  $\theta_{PMT}$  优化各个光强;
- 将 3. 中光强作为 2. 的输入,循环该过程,直至收敛。

经由该过程,2.中拟合自由度为6,3.中自由度为1,实现了极大的自由度降低,显著提高了拟合效果并改善了拟合效果。

#### 2.4.3 刻度模型选择

贝叶斯信息准则(Bayesian information criterion,以下简称 BIC)是一种广泛 用于模型选择的信息参数:

$$BIC = k \ln n - 2\ell \tag{2.28}$$

其中 k 为模型自由度, n 为样本数, ℓ 为对数似然函数。BIC 使用对数似然函数, 用于在有限模型中进行模型选择, BIC 最小的模型在模型集中最优<sup>[19]</sup>。

注意到式(2.28)使用系数 k 使得对模型自由度有与样本数 n 呈对数的惩罚关系,因此能够有效减少过拟合情形的发生。对于 PMT 单光电子电荷响应谱与光强与激光入射角度耦合变量的相关性,可以使用 BIC 判定。基于节 2.4.2 中提出的两种对比拟合方法,BIC 结果如图 2.8 所示。

观察到对于所有 channel,联合刻度的 BIC 基本相似,相对偏差均在 10<sup>-3</sup> 量级以内,这直接意味着同一通道 (PMT) 在不同 run 中使用同一组参数 (即减少了自由度)没有显著的影响,进一步意味着:

1. joint fit 已经足够有效地捕捉了电荷谱的特征;

- 2. separated fit 可能存在过度拟合的问题,导致 BIC 准则的惩罚项较大,没有得 到似然函数的改善;
- 3. joint fit 具有较好的泛化能力,在实际应用中很可能已经足够。

该结果表明,20 寸 MCP-PMT 的单光电子电荷响应与光强与激光入射角度的 组合相关性较弱。



图 2.8 joint fit 与 separated fit 的 BIC 对比

# 第3章 结合波形分析的迭代刻度方法

#### 3.1 快速随机匹配追踪算法

探测器中发生事件的光学信号依靠 PMT 读出,当短时间(例如1 TTS 时间窗) 内只有一个光电子到达 PMT,则在相同的工作条件下,不同 PMT 具有相同的增益 数量级,都能输出形状相似的负极性脉冲。然而,当至少两个光电子到达时,各光 电子响应的电压波形将堆积,使得光电子的个数与各自的能量、到达时间分辨难 度提高。面对该问题,有两种常用的技术手段:

- 以光电子波形到达前的一段波形计算本底信号的输出电压水平(基线),并 将光电子波形积分得到电荷信息,使用经验换算关系得到光电子数并反推事 例能量,时间信息则取波峰的特定上升沿(通常为10%)时刻为到达时间;
- 使用经验的单光电子波形,与去除基线的波形做反卷积,得到各光电子的到达时间。

对于方法 1,由于前一个波形采数时间窗口的靠后位置可能存在信号脉冲,而 脉冲的后延电压恢复基线水准较慢,电荷积分的结果可能存在偏差。同时,每个 光电子倍增后的电荷也具有概率分布,而方法 1 只能通过经验关系换算得到光电 子数,无法精细地得到其方差,因此对提高能量分辨率帮助较为有限。而对于时 间分辨率,由于不同光电子波形的堆叠,方法 1 没有办法准确地区分出各个光电 子的到达时间,从而在事例重建的位置分辨率上也存在客观缺陷。

对于方法 2,对于不同的 PMT 与不同的工作电压,由于电场、渡越速度、噪 声本底的水平不同,单光电子波形的幅值与展宽均有差异。如果全部使用经验单 光电子波形进行反卷积,得到的光电子数量与时间将分别由幅值与展宽的有偏估 计引入偏差,从而减低事例重建的能量分辨率与位置分辨率。

为了实现事例重建的能量分辨率与位置分辨率提升,需要一个贝叶斯方法,使 得能够在不同的单光电子数量、到达时间、电荷样本空间(这些空间是维度可变 的)与单光电子响应波形找到最优解并给出误差分析。

快速随机匹配追踪算法(Fast Stochastic Matching Pursuit,以下简称 FSMP)<sup>[20]</sup> 是一种可逆跳跃的马尔可夫链蒙特卡罗方法(Reversible Jump Markov Chain Monte Carlo),能够在不同维度的样本空间采样,利用 Metropolis-Hastings 等采样方法, 实现采样的跳跃(马尔可夫链状态的转移),最终收敛至马尔可夫链的稳态分布。

25



在 PMT 中,光电子的序列,在考虑增益与单电子响应波形卷积后形成波形; FSMP 能在不同样本空间采样,最终得到与波形最匹配的参数对。

图 3.1 FSMP 波形分析方法

针对 dynode PMT,该方法认为单光电子电荷服从高斯分布,光电子数服从泊 松分布,因此波形中电荷的概率分布模型为复合泊松-高斯分布;针对 MCP-PMT, 该方法将单光电子电荷模型分解为若干个高斯分布线性之和,用以表示光电子在 涂层表面不同行为的分类,每一种高斯分布具有各自独立的期望与方差。使用高斯 混合电荷模型的优势是具有解析的形式,从而使得该方法有显著的运算速率提升。



先前采用 4 个整数倍高斯代表 1 至 4 个倍增电子增益的混合,现在已经升级到了非整数 倍高斯混合

#### 图 3.2 FSMP 电荷模型

该方法于模拟数据集与 8 寸 MCP-PMT 激光测试的实际数据集上均得到了应 用与验证,已经获得了能量与时间分辨率的提升。该工作给出结论,在兆电子伏特 的液体闪烁体中微子探测中,该方法在理想情况下,可以提高可见能量的分辨率 12%。

#### 3.2 迭代刻度

FSMP 波形分析方法可以得到光电子携带的电荷信息来作为刻度的电荷谱输入,因此刻度需要依赖波形分析。同时,如 3.1 中所述,FSMP 同样依赖电荷模型作为采样依据,因此波形分析同样需要依赖刻度。在波形分析与刻度结果相互依赖的客观条件下,为了使得刻度与波形分析的结果同时趋近真值,需要波形分析与刻度交替进行并相互作为输入,直至物理结果收敛,则认为收敛到了贝叶斯后验解。

在没有进入迭代时,使用整数倍高斯拟合 MCP-PMT 的电荷谱作为波形分析 的电荷模型输入,已经在实际激光波形上取得了理想的波形拟合效果,得到的 TTS 从(1.719±0.001) ns 减小到(1.703±0.007) ns<sup>[20]</sup>。在具有充分物理动机的 Gamma-Tweedie 混合电荷模型刻度下,预期将能更好地拟合实际电荷谱,获得更准确的 MCP-PMT 参数。在此基础上,将刻度结果作为波形分析的电荷模型输入,借助波 形分析与刻度迭代进行的工作方法,将有助于准确描述和利用 MCP-PMT 特性,发 挥 FSMP 的能量与时间分辨率优势。

#### 3.2.1 时间窗筛选电荷

FSMP 需要一系列相似发光曲线来实现迭代,因此暗噪声与放射源刻度为理想的刻度源。如节 2.4.1 所述,激光信噪比较暗噪声高,且持续时间短、脉冲波形形状相似,是理想的迭代数据集。经过对 FSMP 分析得到的光电子到达时间频次直方图做生成率归一化后能够得到光变曲线。

由于光源与目标 PMT 相对位置的不同,光电子到达时间峰的时间存在偏差。 为了保证得到的 PE 电荷来源尽可能单一,需要对时间窗筛选,本研究中选择与传 统波形预处理相似的 10% 峰值决定上升下降时间,以此为 PE 到达时间的接受区 间,选择该部分光电子的电荷作为电荷谱进行拟合。

对于大多数的激光刻度数据,各 PMT 光变曲线在峰值对齐后与基线分离的时刻相似,如图 3.4 所示,表现出相似的光电子渡越物理过程。同时,也在大量单通道的光变曲线中发现了微小的上升结构,如图 3.5 所示,在主峰后约 200 ns 处存在峰结构。

1.2 节与图 2.2 分别提出了多种脉冲展宽的因素,包括激光脉冲的展宽、在 MCP 表面反射或倍增的电子的滞后脉冲与离子反馈在 MCP 中的反向迁移形成的

27









在 PMT ID 为 1 的通道中,观测到了光变曲线末端具有部分峰的上升结构。红色区间为 10% 上升下降沿筛选电荷的区间。

图 3.5 单通道的光变曲线与时间窗选择示例

延迟脉冲,另外也存在激光在其他 PMT 玻璃表面反射至目标 PMT 的可能,分析如下:

- 激光脉冲展宽 80 ps 且触发频率稳定,对应电子的延迟应远小于 200 ns;
- 被电离的气体分子在 PMT 高压下漂移速度 ~ 10 m/s, 200 ns 对应漂移距离
   ~ 2 μm, 而 MCP 的总厚度 ~ 480 μm;
- 液体闪烁体中光速~2×10<sup>8</sup> m/s, 200 ns 对应~40m。

综上可知,在 OSIRIS 灵敏体积下,光变曲线中主峰后 200 ns 出现的上升沿主要原因很可能不是上述的任意一种原因,最可能归因于图 2.2 中光电子在 MCP 表面反射后重新入射形成的延迟脉冲。

10%的上升下降时间窗筛选具有的优势包括:

- •减少图 2.2 中前脉冲成分,从而在电荷谱上减少未正确放大的指数项计数;
- 减少 2.1.2 节中未考虑的离子反馈电荷成分计数;
- 对其他光路入射的光子产生的光电子电荷计数有一定抑制作用。

如果能够在后续的研究中验证 200 ns 后上升结构主要由反射的光电子贡献, 那么时间窗的后延可以继续向后延申。在目前,由于数量级差异显著,光变曲线 后端的事例损失约小2个数量级,对增益刻度结果的影响较小。

#### 3.2.2 刻度结果

#### 3.2.2.1 带光强增益刻度

由图 3.6 可见,对于该 PMT,在 4 倍主峰后电荷长尾区域双光电子的贡献已 经超越了单光电子的贡献,如果没有使用带光强的拟合方法得到的参数将可能偏 离物理真值。



图 3.6 µ~0.2 多光电子拟合实际电荷谱的成分占比

#### 3.2.2.2 模型参数

拟合优度检验得到了良好的结果,如图 3.7 所示,绝大多数 PMT 拟合得到  $\chi^2/ndf < 10$ ,可以认为是较好的建模,同时部分通道 (PMT)的 joint fit 能够得到 更好的  $\chi^2/ndf$ ,即得到更好的拟合效果。

在 2.1.2 节中提出了描述 MCP-PMT 的新参数 θ<sub>PMT</sub>,在实际刻度中得到各个物 理参数的独立分布与二维联合分布,如图 3.8 与 3.9 所示。

以上结果证明了 20 寸 MCP-PMT 具有与 8 寸相似的行为,且电荷响应能够被 Gamma-Tweedie 混合模型良好地描述。

#### 3.2.3 电荷模型多高斯分解: EM 算法

如 3.1 所述, FSMP 波形分析方法的电荷模型假设为多高斯混合模型。基于 2.1.2 节中所述 Gamma-Tweedie 混合电荷模型,仍需要找到使用若干个高斯分







图 3.8



图 3.9 二次倍增参数一维与散点分布

布线性之和尽可能近似该分布的方法,才能够将第一轮刻度结果作为 FSMP 中采 样电荷的依据,进行第二轮波形分析与刻度。

为了找到使用  $m \in \mathbb{N}^+$  个高斯分布近似 Gamma-Tweedie 混合电荷模型的最优 解,首先使用刻度得到的 Gamma-Tweedie 混合模型采样总样本数 *M* 足够多的随 机变量观测集  $x = \{x_1\}, 1 \in 1, 2, \dots, M$ ,并使用 Expectation Maximization (EM) 算 法得到极大似然的一组高斯分布线性组合解:

$$g_{\theta}(x) = \sum_{j=1}^{m} \lambda_j \mathbf{N}(x; \mu_j, \sigma_j^2), \quad \sum_{j=1}^{m} \lambda_j = 1.$$
(3.1)

给出参数集  $\theta = (\{\lambda_j\}, \{\mu_j\}, \{\sigma_j^2\}),$  还仍需要隐变量集  $Z = \{Z_{ij}\}$  表示样本  $x_i$  是否来自第 *j* 个高斯分布  $N_i$ :

$$Z_{ij} = \begin{cases} 1, & \bar{\mathcal{X}}_i \times \hat{\mathbf{N}}_j \\ 0, & \bar{\mathcal{X}}_{0} \end{cases}$$
(3.2)

至此可以写出似然函数:

$$\ell(\boldsymbol{\theta}; \mathbf{x}, \mathbf{Z}) = \sum_{i=1}^{n} \sum_{j=1}^{m} Z_{ij} \log\left(\lambda_{j} \mathbf{N}(x_{i}; \mu_{j}, \sigma_{j}^{2})\right).$$
(3.3)

实验中只能观测到样本集x,隐变量Z未知,因此只能极大化(3.3)的期望,即 E-step。

#### 3.2.3.1 E-step

在第t步中,在已知 $\theta$ 时的后验概率用 $p_{ij}$ 表示:

$$p_{ij}^{(t)} = \mathbb{P}_{\theta^{(t)}} \left( Z_{ij} = 1 \mid X_i = x_i \right) = \frac{\lambda_j^{(t)} f(x_i; \mathbf{N}(x; \mu_j^{(t)}, \sigma_j^{2(t)}))}{\sum_{r=1}^m \lambda_r^{(t)} f(x_i; \mathbf{N}(x; \mu_j^{(t)}, \sigma_j^{2(t)}))}.$$
 (3.4)

对数似然函数(3.3)的期望可得:

$$Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(t)}) = \mathbb{E}_{\boldsymbol{\theta}^{(t)}}\left[\ell(\boldsymbol{\theta}; \mathbf{x}, \mathbf{Z})\right] = \sum_{i=1}^{n} \sum_{j=1}^{m} p_{ij}^{(t)} \log\left(\lambda_j f(x_i; \mathbf{N}(x; \mu_j, \sigma_j^2))\right).$$
(3.5)

### 3.2.3.2 M-step

(3.5)化简为:

$$Q(\theta|\theta^{(t)}) = \sum_{i=1}^{n} \sum_{j=1}^{m} p_{ij}^{(t)} \log \left(\lambda_{j} f(x_{i}; \mathbf{N}(x; \mu_{j}, \sigma_{j}^{2}))\right)$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{m} p_{ij}^{(t)} \log \left(\lambda_{j} \frac{1}{\sqrt{2\pi\sigma_{j}}} e^{-\frac{(x_{i}-\mu_{j})^{2}}{2\sigma_{j}^{2}}}\right)$$

$$\to \sum_{i=1}^{n} \sum_{j=1}^{m} p_{ij}^{(t)} \left(-\frac{(x_{i}-\mu_{j})^{2}}{2\sigma_{j}^{2}} - \log \sigma_{j} + \log \lambda_{j}\right)$$

$$\to -\sum_{i=1}^{n} \sum_{j=1}^{m} p_{ij}^{(t)} \left(\frac{(x_{i}-\mu_{j})^{2}}{2\sigma_{j}^{2}} + \log \sigma_{j}\right) + \sum_{i=1}^{n} \sum_{j=1}^{m} p_{ij}^{(t)} \cdot \log \lambda_{j}$$
(3.6)

要极大  $Q(\theta|\theta^{(t)})$ , 应有:

$$\boldsymbol{\theta}^{(t+1)} = \operatorname{argmax}_{\boldsymbol{\theta}} Q(\boldsymbol{\theta} | \boldsymbol{\theta}^{(t)}). \tag{3.7}$$

注意到各高斯参数彼此无约束,显然有更新:

$$\lambda_j^{(t+1)} = \frac{1}{n} \sum_{i=1}^n p_{ij}^{(t)}.$$
(3.8)

 $\mu_j, \sigma_j^2$ 参数的更新需要  $\frac{\partial Q(\theta|\theta^{(t)})}{\partial \mu_j}, \frac{\partial Q(\theta|\theta^{(t)})}{\partial \sigma_j} = 0$ :

$$\begin{cases} \frac{\partial Q(\theta|\theta^{(t)})}{\partial \mu_j} = \sum_{i=1}^n \sum_{j=1}^m p_{ij}^{(t)} \cdot \left(\frac{x_i - \mu_j}{\sigma_j}\right) \\ \frac{\partial Q(\theta|\theta^{(t)})}{\partial \sigma_j} = \sum_{i=1}^n \sum_{j=1}^m p_{ij}^{(t)} \cdot \left(-\frac{1}{\sigma_j} + \frac{(x_i - \mu_j)^2}{\sigma_j^3}\right) \end{cases}$$
(3.9)

解 3.9 得到更新的参数:

$$\begin{cases} \mu_{j} = \frac{\sum_{i=1}^{n} p_{ij}^{(t)} \cdot x_{i}}{\sum_{i=1}^{n} p_{ij}^{(t)}} \\ \sigma_{j}^{2} = \frac{\sum_{i=1}^{n} p_{ij}^{(t)} \cdot \left(x_{i} - \mu_{j}^{(t)}\right)^{2}}{\sum_{i=1}^{n} p_{ij}^{(t)}} \end{cases}$$
(3.10)

至此(3.8)与(3.10)构成了完成的 $t \rightarrow t+1$ 步的完整参数更新。EM 算法收敛至极大似然解的稳定性已经得到了充分的论证。EM 算法在存在隐变量的问题中是最优秀的算法之一,主要原因包括:

- (3.8)基于贝叶斯公式,得到的后验概率天然具有非负归一性质;
- (3.10)能确保期望与方差非负,参数更新具有良好的约束。

#### 3.2.3.3 EM 加速: KMeans 算法初值优化

在实际使用中,EM 算法的收敛速度饱受诟病。为了加速 EM 算法的收敛,可 以提供更接近收敛解的初值。同时由于 EM 算法只能保证收敛到局域极值而不一 定能够收敛到全局极值,本节中将兼顾该问题给出借助 KMeans 算法提供多个初 始组的方法:

- 1. 从观测集中随机选出 m 个样本;
- 2. 将与样本 k 距离最小的样本归类为第 k 个聚类;
- 3. 对第 k 个聚类的所有样本取平均, 作为聚类 k 的新均值 (中心);
- 4. 重复 2-3 直至给定步数或均值收敛;
- 5. 计算每个聚类中样本点至聚类中心的距离均方作为该聚类方差。

对于一维至三维数据,距离容易定义,KMeans 是较为简单有效的分聚类算法,适合作为 EM 算法的初值。

#### 3.2.3.4 EM 加速: 向量 Aitken 加速的 Steffenson 形式

$$\boldsymbol{\phi}_{VA}^{(1)} = \boldsymbol{\phi}^{(0)} - \frac{||\boldsymbol{\phi}^{(1)} - \boldsymbol{\phi}^{(0)}||^2 (\boldsymbol{\phi}^{(0)} - 2\boldsymbol{\phi}^{(1)} + \boldsymbol{\phi}^{(2)})}{||\boldsymbol{\phi}^{(0)} - 2\boldsymbol{\phi}^{(1)} + \boldsymbol{\phi}^{(2)}||^2}.$$
(3.11)

Steffenson 方式是一种不动点迭代的方法,将 3.2.3.2 中定义的参数更新步骤记为  $M(\cdot)$ ,由于  $\phi$  有确定的收敛值  $\phi^*$  序列更新  $\phi \to M(\phi) \to \cdots \to \phi^*$  可以视为趋

近不动点  $\phi^*$  的迭代行为,应用于 (3.11) 即得向量 Aitken 的 Steffenson 形式<sup>[21]</sup>:

$$\phi_{VA}^{(2)} = \phi_{VA}^{(1)} - \frac{||M(\phi_{VA}^{(1)}) - \phi_{VA}^{(1)}||^2 (\phi_{VA}^{(1)} - 2M(\phi_{VA}^{(1)}) + M^{(2)}(\phi_{VA}^{(1)}))}{||\phi_{VA}^{(1)} - 2M(\phi_{VA}^{(1)}) + M^{(2)}(\phi_{VA}^{(1)})||^2}.$$
(3.12)

考虑到(3.11)与(3.12)并不像原始的 EM 算法,不能保证参数更新的合法性,因此在本研究中取  $\phi = (\{\mu_j\}, \{\sigma_j^2\}),$ 相比  $\theta$  需要  $\{\lambda_j\}$  另做合法更新。EM 算法的加速调整如下:

- 1. 利用 *M*(·) 由 *θ*<sup>(0)</sup> 得到 *θ*<sup>(1)</sup>, *θ*<sup>(2)</sup>;
- 3. 判断  $\phi_{VA}^{(1)}$  合法性 (是否非负):
  - 3.1 若合法, 以 $\mu$ 升序重排, (3.8) 计算 $(\lambda^{(0)}, \phi_{VA}^{(1)})$ 下的 $\lambda^{(1)}$ , 初始参数 $\theta_{VA}^{(1)} = (\lambda^{(1)}, \phi_{VA}^{(1)})$  重复 1-3;
  - 3.2 若非法,利用 *M*(·) 更新初始参数  $\theta^{(1)}, \theta^{(2)}, \theta^{(3)}$  重复 2-3;
- 4. 直至  $||\theta_{VA}^{(t+1)} \theta_{VA}^{(t)}||^2 \leq \delta$ 结束迭代。

#### 3.2.4 电荷模型多高斯分解结果

为了决定使用高斯的数量以达到理想的准确度,选取节 2.4.3 中介绍的 BIC 信息准则。

本研究中使用随机数种子 123,使用刻度的分布采样 10000 个随机数并使用 3.2.3 节中所述的方法进行分解,并比较 BIC。本研究对刻度参数集使用 3 至 6 个高斯分布,观察到对于大多数参数集 4 个高斯与 5 个高斯 BIC 接近且优于其他个数,但对于少数情况会出现 5 个高斯在主峰前由单个二次倍增峰贡献处近似较 4 个高斯优的情况,如图 3.10 所示,其中 BIC<sub>4</sub> = 140560.1 > BIC<sub>5</sub> = 140420.4。



图 3.10 利用 BIC 判断 4 个与 5 个高斯分解的优劣

可以得出结论: 5个高斯分解对于主峰前单个二次倍增峰(即 Tweedie 成分) 描述较4个高斯分解略优,较其他个数优。

#### 3.2.5 迭代结果对比

由使用 JWAPtool 粗筛得到的结果与第一轮刻度结果对比:



图 3.11 迭代结果增益对比

这里增益的下降是因为 JWAPtool 粗筛时没有考虑电荷长尾的效应,将所有 电荷取均值作为增益的输入。藉由 Gamma-Tweedie 混合电荷模型的提出,我们将 增益定义为未二次倍增光电子所贡献的主峰的均值,将长尾电荷分离为二次倍增 光电子所贡献的 Tweedie 部分,因此预期得到的均值较简单而没有理论依据的初 筛小。

本研究将继续迭代至前后的物理结果收敛,并认为收敛值为拟合得到的 PMT 最终刻度结果。

# 第4章 结果

### 4.1 工作总结

本论文将 8 寸 MCP-PMT 的电荷模型应用在 OSIRIS 探测器液体闪烁体工作环 境下的 64 支 20 寸 MCP-PMT 的激光刻度数据上,证实该电荷模型具有较好的适 用性,并将继续在 JUNO、JNE 等中微子探测器上开展进一步地刻度,积累 PMT 参数库,以期提高能量分辨率同时,本研究暂时认为该 PMT 的单光电子响应与光 源光强与入射角度相关性较弱,对在大型探测器中的电荷模型应用提供了基于实 际数据的唯象支撑。

为了提高电荷长尾拟合的准确率以及赋予拟合方法在高光强工作条件下的适 用性,论文推导得到了任意光电子数目的电荷谱响应,基于傅里叶变换变换得到 了能够适应不同光强的拟合方法,除了提高激光数据增益刻度的准确性,也为该 MCP-PMT 在高光强场景下的应用提供了理论无偏的拟合工具。

论文基于 FSMP 波形分析方法得到电荷谱,并找到了由电荷模型近似极大似 然高斯混合模型解的方法,在 OSIRIS 探测器上实现了各 MCP-PMT 的增益刻度与 高斯混合电荷模型近似,使得波形分析方法迭代成为现实。

#### 4.2 问题与改进方向

论文还需要进行多轮迭代,直到达到参数收敛的标准。同时拟合所使用的方 法也可以进行调研与改进,以得到更具鲁棒性、有能力找到全局极值的拟合方法。

论文基于 FSMP 波形分析方法得到了光电子到达时间的归一化曲线,即光变曲线,能够以此进行单光电子波形的重新筛选以得到更准确的波形分析结果。以后的工作将着力于实现该想法。

# 插图索引

图 1.1	中微子振荡	1
图 1.2	混合角、质量参数与质量顺序对反中微子能谱的影响 <sup>[4]</sup>	
图 1.3	JUNO 探测器侧视示意图 <sup>[4]</sup>	
图 1.4	液体闪烁体处理系统的流程图	5
图 1.5	OSIRIS 探测器示意图	6
图 1.6	20 寸 MCP-PMT 与环绕的 4 个 3 寸 MCP-PMT	6
图 1.7	MCP-PMT 与 dynode PMT 单光电子电荷谱对比 <sup>[8]</sup>	7
图 1.8	不同中微子探测器所使用的 20 寸 MCP-PMT	
图 2.1	MCP-PMT 的结构设计图 <sup>[6]</sup>	9
图 2.2	MCP-PMT 中物理过程 <sup>[11]</sup>	10
图 2.3	光电子与 MCP-PMT 的三种相互作用形式	11
图 2.4	孔隙对电子束的负透镜效应[15]	13
图 2.5	Dynode PMT 与 MCP-PMT 各本底计数率 <sup>[17]</sup>	21
图 2.6	Channel 11 在不同采数 run 中电荷谱拟合	22
图 2.7	Channel 21 在不同采数 run 中电荷谱拟合	22
图 2.8	joint fit 与 separated fit 的 BIC 对比	24
图 3.1	FSMP 波形分析方法	26
图 3.2	FSMP 电荷模型	26
图 3.3	20 寸 MCP-PMT 在不同数据集上的光变曲线	
图 3.4	峰对齐后的光电子到达时刻曲线形状	
图 3.5	单通道的光变曲线与时间窗选择示例	29
图 3.6	μ~0.2多光电子拟合实际电荷谱的成分占比	
图 3.7	$\chi^2$ /ndf 与不同拟合方式的对比	
图 3.8	主峰参数一维分布	
图 3.9	二次倍增参数一维与散点分布	
图 3.10	利用 BIC 判断 4 个与 5 个高斯分解的优劣	
图 3.11	迭代结果增益对比	

# 表格索引

表 2.1	多种分区间方法的比较	.14	4
-------	------------	-----	---

参考文献

- DECAMP D, DESCHIZEAUX B, LEES J P, et al. Determination of the number of light neutrino species[J/OL]. Phys. Lett. B, 1989, 231(4): 519-529. DOI: 10.1016/0370-2693(89)90704-1.
- [2] FUKUDA Y, HAYAKAWA T, ICHIHARA E, et al. Evidence for Oscillation of Atmospheric Neutrinos[J/OL]. Phys. Rev. Lett., 1998, 81(8): 1562-1567. DOI: 10.1103/PhysRevLett.81.1 562.
- [3] CAHN R N, GOLDHABER G. The Experimental Foundations of Particle Physics[M]. Cambridge, England, UK: Cambridge University Press, 2009.
- [4] JUNO Collaboration. JUNO physics and detector[J/OL]. Progress in Particle and Nuclear Physics, 2022, 123: 103927. DOI: 10.1016/j.ppnp.2021.103927.
- JUNO Collaboration. The Design and Sensitivity of JUNO's scintillator radiopurity pre-detector OSIRIS: arXiv:2103.16900[M]. arXiv, 2021.
- [6] WANG Y, QIAN S, ZHAO T, et al. A new design of large area MCP-PMT for the next generation neutrino experiment[J/OL]. Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, 2012, 695: 113-117. DOI: 10.1016/j.nima.2011.12.085.
- [7] CAO W, WANG B, YANG Y, et al. Secondary electron emission characteristics of the Al2O3/MgO double-layer structure prepared by atomic layer deposition[J/OL]. Ceramics International, 2021, 47(7): 9866-9872. DOI: 10.1016/j.ceramint.2020.12.128.
- [8] WENG J, ZHANG A, WU Q, et al. Single electron charge spectra of 8-inch high-collectionefficiency MCP-PMTs: arXiv:2402.13266[M/OL]. arXiv, 2024. DOI: 10.48550/arXiv.2402. 13266.
- [9] ZHANG A, XU B, WENG J, et al. Performance evaluation of the 8-inch MCP-PMT for Jinping Neutrino Experiment[J/OL]. Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, 2023, 1055: 168506. DOI: 10.1016/j.nima.2023.168506.
- [10] YANG Y, YAN B, LIU S, et al. MCP performance improvement using alumina thin film[J/OL]. Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, 2017, 868: 43-47. DOI: 10.1016/j.nima.2017.06.049.
- [11] CHEN P, TIAN J, GUO L, et al. Photoelectron backscattering in the microchannel plate photomultiplier tube[J/OL]. Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, 2018, 912: 112-114. DOI: 10.1016/j.nima.2017.10.081.

- [12] DINU N, GYS T, JORAM C, et al. Edit 2011 photodetection: Micro Channel plate PMT (MCP-PMT)[C]. Cham, Switzerland: CERN Indico, 2010.
- [13] FURMAN M A, PIVI M T F. Probabilistic model for the simulation of secondary electron emission[J/OL]. Phys. Rev. ST Accel. Beams, 2002, 5: 124404. https://link.aps.org/doi/10.11 03/PhysRevSTAB.5.124404.
- [14] CHEN L, TIAN J, LIU C, et al. Optimization of the electron collection efficiency of a large area MCP-PMT for the JUNO experiment[J/OL]. Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, 2016, 827: 124-130. DOI: 10.1016/j.nima.2016.04.100.
- [15] Electric and Magnetic Field Lenses[M/OL]. Massachusetts Institute of Technology (MIT) 22.09, 2002. https://web.mit.edu/22.09/ClassHandouts/Charged%20Particle%20Accel/CHA P06.PDF.
- [16] FREEDMAN D, DIACONIS P. On the histogram as a density estimator:L 2 theory[J/OL]. Z. Wahrscheinlichkeitstheorie verw Gebiete, 1981, 57(4): 453-476. DOI: 10.1007/BF01025868.
- [17] ZHANG Y, WANG Z, LI M, et al. Study of 20-inch PMTs dark count generated large pulses
   [J/OL]. J. Inst., 2022, 17(10): P10048. DOI: 10.1088/1748-0221/17/10/P10048.
- [18] ZHANG Y, WANG Z, LI M, et al. Dark Count of 20-inch PMTs Generated by Natural Radioactivity: arXiv:2307.15104[M/OL]. arXiv, 2024. DOI: 10.48550/arXiv.2307.15104.
- [19] SCHWARZ G. Estimating the Dimension of a Model[J/OL]. The Annals of Statistics, 1978, 6
   (2): 461 464. https://doi.org/10.1214/aos/1176344136.
- [20] WANG Y, ZHANG A, WU Y, et al. The Fast Stochastic Matching Pursuit for Neutrino and Dark Matter Experiments: arXiv:2403.03156[M]. arXiv, 2024.
- [21] GUO X, LI Q Y, XU W L. Acceleration of the EM algorithm using the Vector Aitken method and its Steffensen form[J/OL]. Acta Math. Appl. Sin. Engl. Ser., 2017, 33(1): 175-182. DOI: 10.1007/s10255-017-0648-3.

# 致 谢

衷心感谢导师续本达副教授对我的悉心指导,不厌其烦地为我的研究指明方 向,并提供了充足的资源。

同时也要感谢所有对我给予过帮助和支持的人,感谢翁俊、王宇逸等师兄对 待懵懂的我的耐心与关切。

感谢所有对我的课题提出过改进意见的中外老师,是你们让这项研究变得更 加充实。 本人郑重声明: 所呈交的学位论文, 是本人在导师指导下, 独立进行研究工作 所取得的成果。尽我所知, 除文中已经注明引用的内容外, 本学位论文的研究成 果不包含任何他人享有著作权的内容。对本论文所涉及的研究工作做出贡献的其 他个人和集体, 均已在文中以明确方式标明。

签名: \_ 文选祺 日期: \_ 2024.6.17\_\_

# 附录 A 外文资料的调研阅读报告

#### PMT Iterative Calibration

### Contents

A.1	Intr	roduction	44
A.2	MC	CP-PMT Charge Model	45
A.2	2.1	Gaussian Mixture Model	45
A.2	2.2	Gamma Tweedie Mixture Model	45
A.3	Miz	xture Fit Methods	46
A.3	3.1	Non-linear Least Square Methods	46
A.3	3.2	Expectation Maximization Algorithm	47

### A.1 Introduction

Photomultiplier tubes (PMTs) with microchannel plate (MCP) and atomic layer deposition (ALD) has been applied in Jiangmen Underground Neutrino Observatory (JUNO) and China Jinping Underground Laboratory (CJPL)<sup>[1]</sup>. It's proposed to overcome the poor time performance of the existing dynode MCP-PMT, aiming to increase the quantum effeciency of photoelectrons (PE). Lin Chen et al concluded from Monte Carlo simulation that 97.5% of photoelectrons land on the MCP active area, emitting multiple secondary electrons from ALD. The acceptance fraction of impinging PE is nearly 100%<sup>[1]</sup>.

However, these MCP-PMTs possess a unique long tail structure in their charge response spectrum compared to traditional dynode PMTs, from which calibration difficulty arises<sup>[2]</sup>. Traditionally, dynode PMTs' single electron response (SER) charge spectrum resembles Gaussian or Poisson distribution. Therefore, a Tweedie distribution would be enough to calibrate the charge response expectation  $\mu_a$  and variance  $\sigma_a^2$ .

For now, some papers on ALD MCP-PMTs are published. For 8-inch MCP-PMTs with ALD, Aiqiang Zhang et al only fitted local peak by Gaussian model, but didn't cali-

brated the long tail structure and give  $\mu_q$  and  $\sigma_q^{2[2]}$ . For 20-inch MCP-PMTs, H.Q. Zhang used a measured SER of MCP-PMT output and randomly sampled it to simulate the charge response of MCP-PMT. They came to a conclusion that there was no good charge response model for these MCP-PMTs<sup>[3]</sup>.

However, in our waveform analysis workflow, charge model calibration is essential. FSMP is adopted in waveform analysis. It requires a charge model result as prior input and output the best fit charge response, light curve and waveform with PE as final posterior result through Reversible Jump Markov Chain Monte Carlo<sup>[4]</sup>. The accurate fit charge response is required for waveform analysis to achieve more accurate waveform analysis result. By applying FSMP again, we obtain a new set of waveform results, thus obtaining a more accurate charge calibration result. Through several rounds of calibration, we expect to reach a reasonable convergence solution. A poorly fit charge result is most likely to achieve poor optimization, therefore a reasonable MCP-PMT calibration is needed.

#### A.2 MCP-PMT Charge Model

#### A.2.1 Gaussian Mixture Model

Aiqiang Zhang proposed a Gaussian mixture charge model based on his former work<sup>[2]</sup>. The charge response consist of several Gaussian components, each of which is contributed by different integer numbers of true-secondary electrons. Therefore, the components' expectations satisfy  $\mu_k = k\mu_q$  and variances satisfy  $\sigma_k^2 = k\sigma_q^2$ , where  $\mu$  stands for single electron response expectation and  $\sigma^2$  stands for single electron response variance. The total charge response could be represented as:

$$Q_{\rm sum} = \sum_{k=1}^{4} N(k\mu, k\sigma^2) \tag{A.1}$$

Such model is adopted in our early attempt, however not suitable for single PE charge response selected by FSMP, shown in Fig A.1.

#### A.2.2 Gamma Tweedie Mixture Model

Jun Weng proposed a charge model based on MCP and ALD mechanism, suggesting that the number of true-secondary electrons follows Poisson distribution, and single electron charge response follows Gamma distribution. Therefore, the total charge response could be considered as a mixture of Gamma and compound Poisson-Gamma distribu-



Figure A.1 Gaussian Mixture Model Fit

tions<sup>[5]</sup>:

$$Q_{\text{MCP-PMT}} = p_0 \times Q_{\text{peak}} + (1 - p_0)Q_{\text{ts}}$$
  
=  $p_0 \times Q_{\text{peak}} + (1 - p_0) \sum_{n=0}^{\infty} \sum_{i=0}^{n} Q_i$  (A.2)

In equation A.2,  $p_0$  is the probability of PE not emitting secondary electrons,  $Q_{\text{peak}}$  and  $Q_i$  follow Gamma distribution, and the number of true-secondary electrons *n* follows Poisson distribution.

Especially, compound Poisson-Gamma distribution is a special case of Tweedie distribution, whose power parameter statisfies  $p \in (1, 2)$ . Tweedie distribution is widely used in generalized linear regression.

#### A.3 Mixture Fit Methods

#### A.3.1 Non-linear Least Square Methods

Assume the the parameters are  $\theta = (\alpha, \beta, \mu, \phi, p)$ , observed charge data set is  $(y_1, y_2, \dots, y_n)$ , and the mixture distribution function is  $f(y) = \pi f_1(y) + (1 - \pi) f_2(y)$ .

The target parameters are:

$$\psi_{\text{LS}} = \underset{(\theta,\pi)}{\operatorname{argmin}} \sum_{i} \left[ \left( y_i - f(y|\theta) \right)^2 \right]$$
(A.3)

Then  $\mu_q, \sigma_q^2$  are:

$$\mu_q = \pi \cdot \frac{\alpha}{\beta} + (1 - \pi) \cdot \mu$$

$$\sigma_q^2 = \pi^2 \cdot \frac{\alpha}{\beta^2} + (1 - \pi)^2 \cdot \mu^p \phi$$
(A.4)

## A.3.2 Expectation Maximization Algorithm

At the (r + 1) iteration, the posterior probability of  $y_i$  belonging to k th mixture component is:

$$\hat{p}_{ik}^{(r+1)} = \frac{\hat{\pi}_k^{(r)} f_k(y_i \mid \hat{\theta}_k^{(r)})}{\sum_{h=1}^{\mathcal{K}} \hat{\pi}_h^{(r)} f_h(y_i \mid \hat{\theta}_h^{(r)})} \quad (k = 1, 2)$$
(A.5)

Q function is defined as the expectation of log-likelihood (E-step):

$$Q = \sum_{k=1}^{2} \sum_{i=1}^{n} \hat{p}_{ik}^{(r+1)} \log f_k(y_i) + \sum_{k=1}^{2} \sum_{i=1}^{n} \hat{p}_{ik}^{(r+1)} \log \pi_k$$
(A.6)

Gamma distribution  $f_1(y)$  and Tweedie distribution  $f_2(y)$  is irrelevant, therefore could be maximized separately.

Use  $\hat{\theta}_k^{(r)}, \hat{\pi}_k^{(r)}$  to generate new parameters (M-step):

$$\hat{\theta}_{k}^{(r+1)} = \underset{\theta}{\operatorname{argmax}} \sum_{i=1}^{n} \hat{p}_{ik}^{(r+1)} \log f_{k}(y_{i})$$

$$\hat{\pi}_{k}^{(r+1)} = \frac{1}{n} \sum_{i=1}^{n} \hat{p}_{ik}^{(r+1)}$$
(A.7)

 $\hat{\theta}^{(\infty)}, \hat{\pi}^{(\infty)}$  are the final estimated posterior parameters and mixture ratio.

#### A.3.3 Correlation

If the charge model is accurate, meaning residuals  $y_i - f(y) \stackrel{\text{i.i.d}}{\sim} N(\mu, \sigma^2)$ , NLS is exactly the same as EM:

$$\ell(\theta) = \sum_{i} \log P(y_i|\theta) = -\frac{1}{2\sigma^2} \sum_{i} (y_i - f(y))^2 - n\log\sigma - \frac{n}{2}\log 2\pi$$
(A.8)

参考文献

- [1] Lin Chen, Huizhen Yang, Xingchao Wang, Liping Tian, Dongyan Ding, Yunji Wang, Ke Ji, Pengxiang Zheng, Ting Luo, and Chenye She. A new large area MCP-PMT for high energy detection. Sci. Rep., 13(20436):1–8, November 2023.
- [2] Aiqiang Zhang, Benda Xu, Jun Weng, Huiyou Chen, Wenhui Shao, Tong Xu, Ling Ren, Sen Qian, Zhe Wang, and Shaomin Chen. Performance evaluation of the 8-inch MCP-PMT for Jinping Neutrino Experiment. <u>Nucl. Instrum. Methods Phys. Res., Sect. A</u>, 1055:168506, October 2023.
- [3] H.Q. Zhang, Z.M. Wang, F.J. Luo, A.B. Yang, D.R. Wu, Y.C. Li, Z.H. Qin, C.G. Yang, Y.K. Heng, Y.F. Wang, and H.S. Chen. Gain and charge response of 20" mcp and dynode pmts. Journal of Instrumentation, 16(08):T08009, aug 2021.
- [4] D.C. Xu, B.D. Xu, E.J. Bao, Y.Y. Wu, A.Q. Zhang, Y.Y. Wang, G.L. Zhang, Y. Xu, Z.Y. Guo, J.H. Pei, H.Y. Mao, J.S. Liu, Z. Wang, and S.M. Chen. Towards the ultimate pmt waveform analysis for neutrino and dark matter experiments. <u>Journal of Instrumentation</u>, 17(06):P06040, jun 2022.
- [5] Jun Weng, Aiqiang Zhang, Qi Wu, Lishuang Ma, Benda Xu, Sen Qian, Zhe Wang, and Shaomin Chen. Single electron charge spectra of 8-inch high-collection-efficiency MCP-PMTs. (arXiv:2402.13266), May 2024. doi: 10.48550/arXiv.2402.13266.

# 综合论文训练记录表

W. 41. 44. 44	1.1.1 Mar. A. (.)	214			TH //1	1 + + + + + + + + +
学生姓名	刘逸祺	孚亏	20200/3/00	*	址级	水-19001
论文题目	光电倍增管	<b><i><b>的进代</b></i>亥</b>	」度	94 0	9	
・主要内容以及进度安排	对大型中微子 进行电荷刻度 通过用 一 一 一 一 一 一 一 一 一 一 一 一 一 一 一 一 一 一 一	与告告罪 沉迷 花式	中候选的圆户镇腹,形为析方法结合起来, 所为析方法结合起来, 析 化果的准确率。 间 电荷 呵声 的 建 和 一 建 的 理 论 推 导 与 计 图 到 虚 方 法 ; 赌 管 上 开 展 迭 代 刻 度 描 建 一 充 支 提 写 据 赴 理 与 论 文 提 写 指 考核	御月 一 一 一 一 一 一 一 一 一 一 一 一 一 一 一 一 一 一 一	电纸 狮麦 了 这 字字 年降的 工建 与 生化 有 一座子 计	营出了。 新 正 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
中期考核意见		海滨	, 正常进行	组组长3 202⁄	È字: ♀年	年月3日

指导教师评语	记之把微逸画版完电信唱音的脑雨主喜,"忆 违随机区和已逸得。"进行了精丽的问是,有利 Gamma-Tweedic 模型测定了新空信唱音的基本号数, 然后确定了 影文写个启外 模型、显素加快了这等速度。 论文的成果呈家友撑了微遍通瓶完电信馆宫的方用、ひ 中级分离路物馆领试有全部的常不价值。 指导教师签字: <u>3英本达</u> 2024年 6月16日
评阅教师评语	论文将 Gamma-Tweedie 电荷模型用于在JUNO 实验上被使用的 20 才 MCP-PMT上,取得了转好的刻度效果,具有重要的科学意义。论文逻辑清晰、内容释实,是一篇优秀的本科生论文。 评阅教师签字: 张志欣打 2024年6月17日
答辩小组评语	论之工作取得了所能性成果。回答问题很确清晰,写作达到本科论之的爱求。同意通过答辨。 智爱龙、同意通过答辨。 普辨小组组长签字: 手被雷 2024 年6月17日

总成绩: <u>A</u> 教学负责人签字: **本**教祥写 2024 年 6 月 20 日