上海交通大学博士学位论文

PandaX 实验偶然符合本底研究和⁸B 太阳中 微子的探测

- 博士研究生:阿布都沙拉木•阿布都克力木
- 学 号: 019072210003
- 导 师:刘江来教授
- 申请学位:理学博士
- **学 科:**物理学
- 所在单位:物理与天文学院
- 答辩日期: 2023年6月1日
- 授予学位单位:上海交通大学

Dissertation Submitted to Shanghai Jiao Tong University for the Degree of Doctor

STUDY OF ACCIDENTAL COINCIDENCE BACKGROUND IN THE PANDAX EXPERIMENT AND SEARCH FOR ⁸B SOLAR NEUTRINO

Candidate:	Abdusalam Abdukerim
Student ID:	019072210003
Supervisor:	Prof. Jianglai Liu
Academic Degree Applied for:	Doctor of Philosophy
Speciality:	Physics
Affiliation:	School of Physics and Astronomy
Date of Defence:	June 1, 2023

Degree-Conferring-Institution: Shanghai Jiao Tong University

上海交通大学

学位论文原创性声明

本人郑重声明:所呈交的学位论文,是本人在导师的指导下,独立进行研究工作 所取得的成果。除文中已经注明引用的内容外,本论文不包含任何其他个人或集体已 经发表或撰写过的作品成果。对本文的研究做出重要贡献的个人和集体,均已在文中 以明确方式标明。本人完全意识到本声明的法律结果由本人承担。

学位论文作者签名:

日期: 年 月 日

上海交通大学

学位论文使用授权书

本学位论文作者完全了解学校有关保留、使用学位论文的规定,同意学校保留并 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。

本学位论文属于 □公开论文

□内部论文,□1年/□2年/□3年 解密后适用本授权书。
□秘密论文, ____年(不超过10年)解密后适用本授权书。
□机密论文, ____年(不超过20年)解密后适用本授权书。
(请在以上方框内打"√")

学位论文作者签名:

指导教师签名:

日期: 年月日 日期: 年月日

PandaX 实验偶然符合本底研究和⁸B 太阳中微子的探测

摘要

根据可靠观测结果可知整个宇宙由 4.9% 的普通物质、26.8% 的暗物质和 68.3% 的暗能量组成。其中暗物质是目前粒子物理领域无法通过粒子物理标准模型解释的谜团之一。为了揭开暗物质的真面目,人们设计了各种实验仪器来探测 暗物质。

PandaX 是一项位于中国四川省锦屏地下实验室的,其主要物理目标为探测暗物质的多物理目标实验项目。PandaX 探测器采用的是气-液氙二相性时间投影室技术,其基本探测原理是,通过光敏器件光电倍增管来探测暗物质跟液氙核碰撞产生的微弱信号。自从 2009 年 PandaX 合作组成立以来,PandaX 暗物质探测实验总共经历了三期阶段,它们分别为 PandaX-I、PandaX-II 以及 PandaX-4T。其中PandaX-4T 是正在运行的体量最大、灵敏度最高的探测器。目前为止,PandaX 各期暗物质直接探测实验没有看到预期的暗物质信号。然而,PandaX-II 实验分别在2016 年和 2017 年,PandaX-4T 实验在 2021 年,关于暗物质-核子相互作用的参数空间给出了世界领先的限制。

在整个实验中,数据的获取、处理、重建以及保证数据的质量是非常重要的环节。在本论文中详细讲述了 PandaX-4T 实验的数据处理和重建框架以及流程。为了保证获取的数据质量,我们研发了一套在线监控数据质量的系统。为了方便查看事例特征,研发了一款跨平台的事例可视化软件工具。除此之外,还研发了各种数据质量筛选方法来排除各种非物理事例。

由于暗物质与普通物质的相互作用非常微弱,因此难以将信号从来自于环境、 探测器材料以及其它的本底噪声中区分开来。降低本底噪声以及提高探测灵敏度 是本实验的关键点以及难点。随着探测器体积的增大,在探测器本底中,偶然符 合本底的含量会明显增大。在本论文中,对于 PandaX-II 以及 PandaX-4T 实验偶然 符合本底,利用不同方法进行了精确估计以及验证,并且利用各种数据质量筛选 以及 BDT 等方法进一步压制了它的含量,从而提高了探测器的灵敏度。

暗物质直接探测实验的目标是寻找暗物质与普通物质原子核之间的弹性散射 信号。然而,这种信号有可能会被来自太阳、大气和宇宙射线的中微子所淹没,形 成一个难以克服的本底噪声,称为中微子地板。其中,⁸B 太阳中微子是最主要的 干扰源之一。⁸B 太阳中微子在暗物质探测器中可以通过相中微子-核子相干弹性 散射(CEvNS)产生与暗物质相同信号。目前暗物质直接探测实验的探测灵敏度 已经很接近中微子地板,这也是探测低能暗物质方面的最大的挑战之一。在本论 文中,我们将⁸B 太阳中微子作为信号,进行了寻找⁸B 太阳中微子的盲分析并且 利用各种新颖的方法提高了信噪比。根据揭盲结果,没有观测到明显超出预期的 信号,从而对⁸B 太阳中微子通量以及质量为 3-9GeV/c² 的暗物质-核子相互作用的 参数空间给出了世界领先的限制。

关键词: 暗物质直接探测, PandaX 实验, ⁸B 太阳中微子, 中微子-核子相干弹性 散射, 偶然符合本底, 决策树

STUDY OF ACCIDENTAL COINCIDENCE BACKGROUND IN THE PANDAX EXPERIMENT AND SEARCH FOR ⁸B SOLAR NEUTRINO

ABSTRACT

According to reliable observational results, it is known that the entire universe is composed of 4.9% ordinary matter, 26.8% dark matter, and 68.3% dark energy. Dark matter is one of the mysteries that cannot be explained by the standard model of particle physics. In order to uncover the true nature of dark matter, various experimental instruments have been designed to detect dark matter.

PandaX is a multi-target experimental project located in the Jinping underground laboratory in Sichuan Province, China, with the main physics goal of detecting dark matter. The PandaX detector uses the dual-phase xenon time projection chamber technology, and its basic detection principle is to use photomultiplier tubes to detect the faint signals generated by collisions between dark matter and xenon nuclei. Since the establishment of the PandaX collaboration in 2009, the PandaX dark matter detection experiment has gone through three phases, namely PandaX-I, PandaX-II, and PandaX-4T. Among them, PandaX-4T is the largest in size and highest in sensitivity, and is currently in operation. Currently, none of the PandaX dark matter direct detection experiments have observed the expected dark matter signal. However, the PandaX-II experiment provided world-leading constraints on the parameter space of dark matter-nucleon interactions in 2016 and 2017, and the PandaX-4T experiment did so in 2021.

In the entire experiment, data acquisition, processing, reconstruction, and ensuring data quality are crucial steps. This paper describes in detail the data processing and reconstruction framework and process of the PandaX-4T experiment. In order to ensure the quality of the acquired data, we have developed an online monitoring system for data quality. A cross-platform event visualization software tool has been developed for convenient viewing of event characteristics. In addition, various data quality filtering methods have been developed to exclude various non-physical events.

Due to the extremely weak interaction between dark matter and ordinary matter, it is difficult to distinguish the signal from the background noise coming from the environment, detector materials, and other sources. Reducing background noise and improving detection sensitivity are crucial and challenging points in this experiment. As the detector volume increases, the contribution of accidental coincidences in the detector background becomes more significant. In this thesis, precise estimation and validation of accidental coincidences in the background of PandaX-II and PandaX-4T experiments were conducted using different methods. Various data quality filtering techniques and BDT (Boosted Decision Tree) methods were also employed to further suppress the accidental coincidence background, thus enhancing the sensitivity of the detector.

The goal of direct detection experiments for dark matter is to search for elastic scattering signals between dark matter and ordinary matter nuclei. However, this signal may be overwhelmed by neutrinos from the Sun, atmosphere, and cosmic rays, creating an insurmountable background noise known as neutrino floor. Among them, ⁸B solar neutrinos are one of the main sources of interference. ⁸B solar neutrinos can generate the same signal as dark matter through coherent elastic neutrino-nucleus scattering (CE*v*NS) in dark matter detectors. Currently, the detection sensitivity of direct dark matter detection experiments is approaching the neutrino floor, which is also the biggest challenge in detecting low-energy dark matter. In this paper, we conducted a blind analysis to search for ⁸B solar neutrinos as signals and improved the signal-to-noise ratio using various novel methods. Based on the unblinding results, no significant signal beyond expectations was observed, providing world-leading constraints on the ⁸B solar neutrino flux and the parameter space of dark matter-nucleus interactions with masses ranging from $3-9 \text{ GeV/c}^2$.

KEY WORDS: dark matter direct detection, PandaX experiment, ⁸B solar neutrino, coherent elastic neutrino-nucleus scattering, accidental coincidence background, decision tree

第一章	引言 …		1
1.1	标准模	型	2
1.2	中微子	·	3
	1.2.1	太阳中微子和标准太阳模型	4
	1.2.2	中微子-核子相干弹性散射	8
1.3	暗物质		9
	1.3.1	暗物质存在的依据	10
	1.3.2	暗物质候选粒子	14
1.4	暗物质	的探测	19
	1.4.1	直接探测	20
	1.4.2	间接探测与对撞机实验	23
第二章	PandaX	(实验	26
2.1	中国锦	屏地下实验室	27
2.2	PandaX	(探测器	28
	2.2.1	探测器靶物质-氙	28
	2.2.2	气-液氙二相型 TPC	29
	2.2.3	电子学及数据获取系统	32
2.3	基础设	施	35
	2.3.1	屏蔽体	35
	2.3.2	冷却和循环系统	37
	2.3.3	精馏塔	39
	2.3.4	超纯水屏蔽系统	40
	2.3.5	刻度系统	40
	2.3.6	慢控制系统	42
第三章	数据处于	理与重建	45
3.1	消息队	列	45
3.2	数据格	式	46
3.3	数据库		48
3.4	数据处	理流程	50
3.5	PMT 堦	曾益刻度	53

上海交通大学博士学位论文

3.6	在线数	z据质量监控系统 ······	53
3.7	事例可	「视化工具	56
3.8	PandaX	X 实验数据总结 ······	63
3.9	暗物质	信号区的数据筛选	64
第四章	探测器	的响应参数和刻度	69
4.1	PMT X	双光电子产生率	69
4.2	位置重	建	69
4.3	电荷修	€正 ·····	71
4.4	能量重	建建	73
4.5	探测器	·刻度 ······	74
第五章	暗物质	信号区间的本底估计	79
5.1	电子反	冲本底	79
	5.1.1	CH ₃ T·····	79
	5.1.2	¹²⁷ Xe	80
	5.1.3	探测器材料本底	82
	5.1.4	⁸⁵ Kr ·····	82
	5.1.5	²²⁰ Rn 和 ²²² Rn	82
5.2	核反冲	中本底	83
	5.2.1	中子本底	83
	5.2.2	⁸ B 太阳中微子	85
5.3	表面本	底	87
5.4	本底总	、结	88
第六章	偶然符	合本底	91
6.1	偶然符	合本底的来源	91
6.2	偶然符	F合本底的估计方法	93
6.3	在 Pan	daX-II 实验中的偶然符合本底	93
	6.3.1	孤立 S1 和孤立 S2 信号的选取	93
	6.3.2	偶然符合本底的模拟	100
	6.3.3	偶然符合本底的抑制	100
	6.3.4	BDT 训练结果	103
	6.3.5	BDT 对偶然符合本底的应用	107
6.4	在 Pan	daX-4T 实验中的偶然符合本底	109
	6.4.1	孤立 S1 和孤立 S2 信号的选取	109

	6.4.2	偶然符合本底的抑制	
	6.4.3	死窗口的偶然符合本底	114
第七章	利用 Pa	mdaX-4T 实验数据寻找 ⁸ B 太阳中微子	117
7.1	本底模	型	
7.2	信号模	型	
7.3	BDT 方	7法	124
7.4	统计诠	释	127
第八章	总结与	展望	
参考文献	狱		139
致 谢			155
学术论文	文和科研	₩→→→→→→→→→→→→→→→→→→→→→→→→→→→→→→→→→→→→	

插图索引

图 1-	1 标	准模型粒子 [26]。	3
图 1-	2 pp	-chain 和 CNO-cycle 过程的示意图 [36]。	5
图 1-	3 太	阳中微子通量能谱图 [36]。	6
图 1-	4 在	SNS 中产生中微子的物理过程 [59]。	9
图 1-	5 三	角座星系(Messier 33)的旋转曲线,(黄色和蓝色点),以及	
	从	可见物质分布(灰色线)预测的旋转曲线。这两条曲线之间的	
	差	异可以通过在星系周围添加暗物质晕来解释 [66]。	11
图 1-	6星	系团 Abell 1689 引力透镜的图 [69]。	12
图 1-	7 子	弹星云现象 [71]。	13
图 1-	8 宇	宙微波背景全天图 [72]。	14
图 1-	9 早	期宇宙中 WIMPs 的数密度 [83]。	16
图 1-	10 探	测暗物质的三种通道。	19
图 1-	11 基	于各类探测技术,世界各地的暗物质直接探测实验的总结。	20
图 1-	-12 CI	DMS 探测器原理示意图 [100]。	21
图 1-	-13 CF	RESST 探测器原理示意图 [102]。	22
图 1-	-14 DA	AMA/LIBRA NaI 探测器原理示意图 [106]。	23
图 2-	-1 Pa	ndaX 实验项目运行时间线。	26
图 2-	-2 中	国锦屏地下实验室(CJPL)地理位置[118]。	27
图 2-	-3 CJ	PL-II 的布局图 [119]。	28
图 2-	4 Xe	e、Ge、Ar、I和 Na 等元素跟 WIMP 的弹性散射反冲能事例率	
	谱	,这里假定 WIMP 质量为 100 GeV/c ² , Xe 跟 WIMP 的散射截	
	面	为 10^{-45} cm ² [122]。	30
图 2-	5 氙	对不同类型粒子的阻止本领。其蓝色和红色分别代表电子反	
	冲	和 α 事例,绿色和紫色虚线为核反冲事例 [125]。	31
图 2-	6 气	-液氙二相型 TPC 原理示意图 [127]。	32
图 2-	7 入	射粒子在液氙中能量沉积的物理过程示意图。	33
图 2-	8 通	过 PandaX-II 实验数据重建出来的典型的 S1 和 S2 信号的波形。	33
图 2-	9 Pa	ndaX-4T 探测器 TPC 示意图。	34
图 2-	-10 在	PandaX 实验中使用的一英寸 PMT (左)和三英寸 PMT (右)。	34
图 2-	-11 Pa	ndaX-4T 电子学和数据获取系统示意图。	36

图 2-12	PandaX-4T 实验的各个配套系统的展示图。	37
图 2-13	在 PandaX-II 实验中使用的被动屏蔽体系统。	38
图 2-14	PandaX-4T 探测器的内罐(左)和外罐(右)实拍图。	38
图 2-15	PandaX-4T 冷却总线实拍图(左)和循环提纯系统(右)示意图。	39
图 2-16	PandaX-4T 超纯水罐外观(左边)和内观(右边)实拍图。	40
图 2-17	PandaX-4T 实验的内部放射源刻度注入控制面板。	41
图 2-18	PandaX-4T 实验的外部放射源刻度管(左)和 DD 中子生成器	
	(右)。	42
图 2-19	PandaX-4T 实验的慢控制系统软件部分的结构示意图,一共有三	
	个独立的模块,它们是通过标准的 HTTP 协议来通讯 [135]。	44
图 3-1	在 PandaX-4T 数据处理过程中的消息队列示意图。	47
图 3-2	PandaX-4T 数据库表结构及关系图。	49
图 33	增益刻度前后波形对比(上图为刻度前,下图为刻度后的波形)。	50
图 34	经过 Hit Cluster 算法来组合的 Hit,不同颜色表示来自不同 PMT	
	的 Hit。 ·····	51
图 3-5	PandaX-4T 数据处理流程示意图。	52
图 36	在 PandaX-4T 实验中,通道号为 10115 的 PMT 的 LED 刻度 ADC	
	值的分布。	54
图 37	PandaX-4T 实验通道号为 10200 的 PMT 的 LED 刻度增益随着时	
	间变化的趋势。	54
图 38	PandaX-4T 在线数据质量监控系统。	56
图 39	PandaX-4T 实验在线数据质量监控系统主页面。	57
图 3-10	PandaX-4T 实验在线数据质量监控系统产生的部分数据质量图。…	58
图 3-11	PandaX-4T 实验每个数字化波形采集卡的带宽随着时间变化的	
	趋势,不同颜色表示的是不同的通道。	59
图 3-12	PandaX-4T 事例可视化工具页面。	63
图 3-13	PandaX-4T 实验刻度数据的电子漂移时长跟 S1 信号在顶部和底	
	部 PMT 接收的电荷对称性的二维分布。红色界线之外的部分被	
	排除掉的数据。	68
图 3-14	PandaX-4T 实验刻度数据的电子漂移时长跟 S2 信号宽度的二维	
	分布。红色界线之外的部分被排除掉的数据。	68

上海交通大学博士学位论文

図 / 1	PandaY AT 立验刻度粉据最小 S1 信号(日由一个 Hit 信号组成	
F=1 F	h = h = h = h = h = h = h = h = h = h =	
	的厂的电荷力和《盖它直力图》,红色关线为主向孤古函数,绿 布和粉布公别为苗来由子门及双来由子信早的真斯公布	70
团 4 0	\mathbf{D} and \mathbf{V} \mathbf{T} T	70
图 4-2	ranuax-41 关验中电」有即随有时向文化的趋势。共量已伸红已 公则用 164 koV 的快预修束例和 。 束例计算的结用	72
团12		12
含 4-5	PandaX-41 头短中取小的 $S2$ 的万尔,黑巴思代衣奴据,绿巴以及	
	红巴分别走根据早电于和双电丁峰,用高斯函数米拟合的曲线, 你会具会早期会买款,蓝会头费收入;专买教	74
	红色是全向拟合函数。监巴内资本狄拉兄函数。	74
图 44	PandaX-4T 实验甲各尖甲能峰的 Doke Plot。红线是根据图甲儿	
	个特征峰点米拟合得到的。	75
图 45	根据蒙特卡洛模拟得到的 ²⁴¹ Am- ⁹ Be的两种中子能谱,红色是第	
	一激发态能谱,蓝色是基态能谱。	76
图 46	长寿命放射性元素 ²³² Th的衰变链以及衰变信息。	77
图 4–7	调整(或拟合)后的 NEST 模型产生的 S1 和 S2 分布跟各种刻	
	度数据的比较,带有误差棒的黑色电是数据,红色直方图是由	
	NEST 模型产生的模拟数据。左边三列图分别是经过数据筛选后	
	的 ²⁴¹ Am- ⁹ Be, DD 和 ²²⁰ Rn 的 S1 和 S2 的联合分布,其中粉红	
	色线是 DD 或 ²⁴¹ Am-9Be 数据分布的中位线, 虚线是它们的 95%	
	的分位数线, 蓝色的是 220Rn 数据的中位线, 对应的虚线是 95%	
	的分位数线,紫色虚线是99.5%的核反冲接收线,浅蓝色点是落	
	在核反冲中位线下面的事例。	78
图 5-1	PandaX-4T 实验暗物质能区(0-30 keV _{ee})内的数据能谱,其中蓝	
	色的是 CH ₃ T 的拟合曲线, 橙色的是其他的电子反冲本底拟合曲	
	线, 红色的是总的拟合曲线。	80
图 5-2	¹²⁷ Xe通过电子俘获的方式转变为 ¹²⁷ I的示意图,在此过程处于	
	激发态的 ¹²⁷ I会放出几种γ射线。随后,由于电子俘获的原因,	
	¹²⁷ I 会放出 X 射线或俄歇电子。	81
图 53	PandaX-4T 实验中的 ¹²⁷ Xe 事例率随时间变化的趋势。	81
图 5-4	⁸⁵ Kr的衰变示意图(左)以及在 PandaX-4T 实验中观测到的典型	
	的 β-γ 符合事例的波形 (右)。 ······	83
图 55	长寿命放射性元素 ²³⁸ U的衰变链以及衰变信息。	84

图 56	探测器材料中子本底的产生机制,(a)图是通过(α , n)反应产 生中子本底的机制,(b)图是通过自发裂变的方式产生中子本底	
	的机制。	85
图 57	⁸ B 太阳中微子通过中微子-核相干弹性散射在液氙中产生的核反	
	冲能谱 [170]。	86
图 58	²¹⁰ Po 事例对于不同的水平位置 R ($R = \sqrt{x^2 + y^2}$,其中 x 和 y 是事	
	例的水平面的坐标)的 S1 和 S2 信号大小的二维概率密度分布,	
	其中(a)是当 R=515mm 时的概率密度分布,(b)是当 R=600mm	
	时的概率密度分布。	87
图 59	暗物质质量跟原子核自选无关的弹性散射界面的排除线。其中	
	除了 PandaX 之外,还包括其他的暗物质探测实验的结果 [172-	
	173,175-176]。	89
图 6-1	PandaX 暗物质探测器 TPC 示意图,可能产生孤立 S1 和孤立 S2	
	信号的区域用红色数字来标记。	92
图 62	在 PandaX-II 实验选取孤立 S1 的第二种方法的示意图。	94
图 63	在 PandaX-II 实验中,选取孤立 S1 的第二种方法来选取的孤立	
	S1 跟触发事例 S1 的时间差分布。	95
图 64	在 PandaX-II 实验中,选取孤立 S1 的第三种方法的示意图。	95
图 65	在 PandaX-II 实验中通过第三种选取孤立 S1 的方法来选取的孤	
	立 S1 跟物理事例 S1 的时间差分布,其中(a)为原始的事例分	
	布,(b)为归一化后的分布。	96
图 66	在第三种选取孤立 S1 的方法来选取的物理事例的漂移时间分	
	布,这里选取的物理事例是孤立 S1 跟物理事例 S1 的时间差小	
	于 120µs 的部分。	97
图 67	第三种方法来计算的孤立 S1 信号以及孤立 S2 信号发生的频率	
	随着时间变化的趋势。	98
图 68	2018年3月11日至2018年4月6日期间的孤立S1信号电荷在	
	顶部 PMT 上的分布。	99
图 6–9	第三种方法来得到的孤立 S1(a)和孤立 S2(b)信号的大小分	
	布。	99
图 6-10	通过蒙特卡洛模拟方法来得到的 S1 和 S2 信号的联合分布,其	
	中 a, b, c 分别对应的是 Run 9, 10, 11 的偶然符合本底, d, e,	
	f 分别对应的是 Run 9, 10, 11 的的核反冲刻度数据。1	101

上海交通大学博士学位论文

图 6-11 通过 TMVA 的提供的现成数据来得到的不同算法的本底抑制率
与信号效率("ROC 曲线")的示例,其中 BDT 或 MLP 等方法
的本底抑制率明显高于普通极大似然方法。 102
图 6-12 BDT 训练数据(Run 11 数据集核反冲中位线以下的数据)的输
入特征或变量的分布,蓝色是信号(核反冲事例),红色是本底
(偶然符合事例)。104
图 6-13 训练和测试数据集的 BDT 分数分布。其 K-S 测试值用于衡量过
拟合。
图 6-14 Run 9、10 和 11 的本底抑制率随着信号效率的变化趋势(ROC
曲线)。107
图 6-15 BDT 对电子反冲和核反冲事例的接受效率跟 S1 和 S2 信号的关
系图。其(a, d, g)为 BDT 接收效率随着 S1 信号电荷变化的
趋势。(b, c, e, f, h, i)为 S1 和 S2 信号的二维接收效率分布。108
图 6-16(a)为数据筛选之前的孤立 S1 的分布,(b)为数据筛选之前的
S1 和 S1TBA 的二维分布。
图 6-17 数据筛选之前的孤立 S1 信号的发生频率随着时间变化的趋势。 ··110
图 6-18(a)为 S1 信号质量筛选之后的孤立 S1 的分布,(b)为 S1 信号
质量筛选之后的 S1 和 S1TBA 的二维分布。111
图 6-19 S1 信号质量筛选之后的孤立 S1 信号的发生频率随着时间变化
的趋势。
图 6-20 孤立 S2 事例 (蓝色), 阴极事例 (红色) 以及门电极事例 (绿色)
的 S2 大小, S2TBA 和 S2 信号宽度分布
图 6-21 S2 信号质量筛选之后的孤立 S2 信号的发生频率随着时间变化
的趋势
图 6-22 偶然符合本底的电子漂移时长跟 S2 信号宽度的二维分布, (a)
图是经过电子扩散效应的限制条件之后的分布,(b) 图是经过电
子扩散效应的限制条件之前的分布。113
图 6-23 经过所有数据筛选之后的偶然符合本底的 S1 和 S2 信号联合分
布,其 S1 和 S2 均已修正。红色线是核反冲中位线。114
图 6-24 寻找暗物质数据(数据筛选之前)的电子漂移时长分布,其在
845μs 附近的峰为阴极事例, 红色阴影部分为死窗口偶然符合本
底。

图 6-25	孤立 S2 信号(蓝色)跟死窗口偶然符合本底的 S2 信号(红色)
	分布的比较(数据筛选之前)。
图 7-1	⁸ B 太阳中微子分析中的信号接受效率(红色实线)图。其中包
	括 ⁸ B 太阳中微子(黑色实线)和暗物质(4 GeV/ c^2 和 8 GeV/ c^2
	灰色虚线)反冲能普。除此之外,图中还包括数据重建(蓝色实
	线)、数据筛选(绿色实线),感兴趣区间以及 BDT(粉色实线)
	的信号接受效率。117
图 7-2	相邻能量区间中的2Hits本底(偶然符合,核反冲以及电子反冲
	本底)以及数据分布的比较,(a)为2Hit S1信号的分布,(b)
	为 S2 信号的分布。
图 7-3	感兴趣能量区间中的2Hits本底(偶然符合,核反冲以及电子反
	冲本底)以及数据分布的比较,(a)为2HitS1信号的分布,(b)
	为 S2 信号的分布。
图 74	在 ⁸ B 太阳中微子分析中使用的电产额及光产额跟其他实验中测
	到的实际观测值的比较,本分析中使用的误差是阴影的误差带
	[179]121
图 7-5	波形模拟数据跟刻度数据的 S1(a)和 S2(b)信号波形宽度的
	比较,其红色直方图为波形模拟数据,黑色点为刻度数据。122
图 76	利用扩散模型以及中子刻度数据产生的一批假数据(红色),其
	蓝色为中子刻。123
图 7–7	利用扩散模型产生的假数据(左)跟真实的中子刻度数据(右)
	的变量相关系数。123
图 78	BDT 训练变量过程中敏感度最强的六个变量的分布。125
图 79	训练和测试数据集的 BDT 分数分布。其 K-S 测试值用于判断训
	练模型的过拟合程度。125
图 7-10	根据公式 7-1,对 2 Hit 观测通道的 S2 信号的上下限以及 BDT
	分数界限, 三维扫描的结果。其横轴是 S2 的上限, 纵轴是 S2 信
	号的下限,颜色轴表示 ⁸ B太阳中微子信号观测到的概率。 127
图 7-11	BDT 应用之前(蓝色直方图)和应用之后(红色直方图)的偶然
	符合本底(a)以及波形模拟信号(b)的S2分布。128
图 7-12	通过简单的计数实验来得到的,对于不同 μ (横轴)对应的 CL_S
	以及 CL_{s+b} 值的分布(红色和蓝色虚线),黑色虚线为预期的 CL_s
	值,绿色和黄色区域为一倍和两倍标准误差带。130

- 图 7-14 根据 ⁸B 太阳中微子分析揭盲结果来计算得到的分别对 ⁸B 太阳 中微子通量(上)以及暗物质-核子自旋相关散射截面(下)的 90% 置信水平的上限。上图蓝色线为标准太阳模型预测的 ⁸B 太 阳中微子通量(5.46±0.66)×10⁶/cm²/s[191])。下图中灰色阴影区 域为在不同的曝光量(分别为 1, 10 和 1000 吨. 年)下的中微子 地板 [192]。其中还包括其他实验的结果 [96,176,193]。 ………135

表格索引

表 1-1	四种相互作用力的特征。	2
表 21	惰性气体 Xe, Ar 以及 Ne 的基本物理性质。	29
表 3-1	PandaX-II 实验数据参数总结。第二列对应的是活时间,其第三	
	列为阴极(门电极)对应的负高电压,第四列为最大的电子漂移	
	时长。第五、六、七列对应的参数在第四章中讨论。最后一列为	
	平均的电子寿命,也是在第四章中讨论。	54
表 32	PandaX-4T 实验数据参数总结。其第二列是去掉死时间以及质量	
	有问题的文件之后的活时间。	54
表 51	在 Run 9, 10, 11 数据中的各个本底的估计值以及通过数据筛选	
	之后的寻找暗物质数据的事例数汇总,其中的拟合值是根据暗	
	物质的质量为 400 GeV/c ² 的信号分布来拟合的结果。其单位是,	
	事例数。	39
表 52	在 Set 1、2、3、4 和 5 子数据集中的各个本底的估计值以及通过	
	数据筛选之后的寻找暗物质数据的事例数汇总,其中的拟合值	
	是根据仅本底假设下拟合得到的。其单位是,事例数。 9)0
表 61	在 PandaX-II 实验中,不同数据集的孤立 S1 和孤立 S2 信号的平	
	均发生频率以及它们的误差。)8
表 62	在 PandaX-II 实验中,经过所有数据筛选之后的, Run 9、10 和	
	11 的偶然符合本底含量。其单位是,事例数。10)()
表 63	对 Run 9、10 和 11,最大的信噪比 S 对应的信号接收率 ϵ_s 本底	
	抑制率 $1 - \epsilon_b$ 。 ····································)6
表 64	训练过程中,灵敏度最好的五个变量以及它们的重要性排名。…10)7
表 65	PandaX-4T 实验暗物质信号区间孤立 S1 和孤立 S2 在数据筛选	
	以及数据筛选之后的发生频率。	2
表 66	PandaX-4T 实验暗物质能区各个子数据集偶中经过所有数据筛	
	选之后的然符合本底的含量。其单位是,事例数。11	4
表 7-1	在相邻能量区间(300-800PE)内的各个本底的估计值以及观测	
	到的事例数。其单位是,事例数。11	9
表 72	在感兴趣能量区间(65-300PE)内的本底和信号估计值。其单位	
	是,事例数。	20

表 7–3	对 BDT 以及 S2 上下限优化后的各类本底以及信号估计值。其中				
	包括 BDT 应用之前和之后的对数据的揭盲结果。其单位是,事				
	例数。126				
表 74	计算上限时使用的各类冗余参数(系统误差)以及它们的估计方				

法。 ------133

第一章 引言

粒子物理是物理学的一个分支, 主要研究构成现实世界的最小可观测粒子的 性质和相互作用。粒子物理学的目标是找到能解释粒子行为的基本规律和统一的 理论。粒子物理学的历史可以追溯到古希腊哲学家, 他们试图了解物质的本质和 支配其行为的基本原则。最早的物质理论之一是由德谟克利特(Democritus)提出 的。他认为万物由原子和虚空构成, 原子是不可分割的最小单位。这一概念由亚里 士多德(Aristotle)和伊壁鸠鲁(Epicurus)等其他哲学家进一步发展。然而, 直 到 19 世纪, 原子的概念才得到进一步发展和完善。1803 年, 约翰•道尔顿(John Dalton)提出新的原子论。他认为所有元素都是由原子组成的, 元素的性质取决于 组成它的原子的数量和种类 [1-3]。

20 世纪初人们见证了粒子物理学的许多重大发展。约瑟夫·约翰·汤姆逊 (Joseph John Thomson)在1897年的一次实验中,证明了电子的存在,并测定了电 子的荷质比 [4-5]。他利用了阴极射线管,通过改变电场和磁场的强度,观察阴极射 线的偏转,发现阴极射线是由带负电的粒子组成的,这些粒子的质量比氢原子还要 小得多。他把这种粒子称为电子,开创了原子物理学的新领域。1911年欧内斯特• 卢瑟福 (Ernest Rutherford) 进行了一项著名的实验, 他用 α 粒子轰击一片金箔薄片, 观察到一些粒子被金原子偏转。这导致原子核的发现 [6-7]。1932 年詹姆斯•查德 威克 (James Chadwick) 在研究铍被 α 粒子轰击时产生的一种穿透力很强的射线的 性质时,发现了中子的存在 [8-9]。他利用反冲核的动量,估算出中子的质量几乎 与质子相同,并证明了中子是一种不带电的粒子。那么,是否质子和中子是最基本 粒子呢?答案是否定的。到了上世纪 60 年代,默里•盖尔曼(Murray Gell-Mann) 和乔治•茨威格(George Zweig)提出了夸克模型 [10-11]。他们认为质子和中子 由夸克相互结合形成的复合粒子。在上世纪60年代到70年代,物理学家们在研究 夸克模型、弱电统一理论以及量子色动力学的基础上逐步建立和发展起来了所谓 的标准模型(Standard Model)[12]。它能够很好地解释和描述基本粒子的特性及 相互间的作用,并与量子力学及狭义相对论相兼容。几乎所有的实验的结果都合乎 这套理论的预测。

为了研究基本粒子的特性及其相互作用,粒子物理学家使用了一系列实验技术和仪器。其中包括使用电磁场将粒子加速到高能的粒子以及将粒子以高能碰撞 在一起以产生新粒子并研究其特性的对撞机实验。近年来粒子物理学最重大的发现之一是 2012 年发现的希格斯玻色子,是标准模型预测的基本粒子之一。这种以

-1-

物理学家彼得•希格斯(Peter Higgs)的名字命名的粒子是在欧洲核子研究中心 (Conseil Européen pour la Recherche Nucléaire, CERN)的大型强子对撞机(Large Hadron Collider, LHC)中发现的[13-15]。希格斯粒子的发现对于标准模型来说巨 大的成功。。

1.1 标准模型

标准模型是描述基本粒子和自然力的当前理论框架。这是一个成功且经过充 分检验的理论,可以解释粒子物理学中广泛的实验观察结果。标准模型基于这样 一种观点,即所有物质都由基本粒子组成,基本粒子根据自旋特性(一种量子力 学上的性质)可分为两类:费米子和玻色子。费米子是自旋是半整数(例如1/2, 3/2,5/2等)的粒子,其包括夸克、轻子及其相应的反粒子。玻色子是自旋是整数 (例如0,1,2,3等)的粒子,包括规范玻色子和希格斯玻色子。

标准模型的费米子分为两类: 夸克和轻子。夸克是构成质子和中子的基本粒子,在强相互作用力下结合在一起。它们的自旋为 1/2,有六种味:上、下、粲、奇、顶和底夸克。轻子是一种不参与强相互作用的费米子,它们的自旋是 1/2。轻子包括电子、μ子、τ粒子和与之相应的中微子,以及它们的反粒子。

标准模型的玻色子是一种自旋为整数的基本粒子,它们是物质之间的力的媒介。标准模型的玻色子有四种,分别是光子、胶子、W 玻色子和 Z 玻色子,以及希格斯玻色子。这些力的强度由耦合常数决定,耦合常数是粒子间相互作用强度的度量。表1-1中总结了四种相互作用的特征,包括相对强度、力程、特征时间以及它的媒介子。

相互作用方式	相对强度	力程/m	特征时间/s	媒介子
强	1	10^{-15}	10^{-23}	胶子
电磁	10^{-2}	∞	$10^{-20} \sim 10^{-16}$	光子
弱	10^{-14}	$10^{-18} \sim 10^{-17}$	10^{-10}	W^{\pm}, Z°
引力	10^{-39}	∞		引力子

表 1-1 四种相互作用力的特征。

标准模型的主要挑战之一是它无法解释引力,引力是自然界的四种基本力之一。这是该理论的一个主要局限性,也是寻找量子引力理论的主要动机之一[16],它是种将引力与其他自然力统一起来的理论。标准模型的另外挑战是它无法解释强 CP (Strong CP)问题 [17-18],以及中微子质量问题。标准模型预测中微子没有质量,然而,实验结果表明中微子是有质量的 [19-20]。除此之外,标准模型并

不能解释宇宙中的一些观测现象,例如暗物质(Dark Matter)和[21-22]。这些现 象被认为是由尚未被理解的粒子和力引起的,寻找这些粒子和力是粒子物理学研 究的一个关键领域。这项研究通常被称为"超标准模型",其重点是有助于解释这 些现象的新粒子和相互作用力 [23-25]。



Standard Model of Elementary Particles

图 1-1 标准模型粒子 [26]。

1.2 中微子

中微子也是标准模型预言的一类粒子,属于轻子。它是一种非常微小的中性 粒子,它可以穿透物质而不与之相互作用。中微子的提出和发现是物理学的一个 重要里程碑,它揭示了原子核的结构和衰变的机制,以及宇宙中的一些奇妙现象。

中微子的提出是为了解决一个困扰物理学家的谜题:β衰变。β衰变是一种放 射性衰变,其中一个原子核中的一个中子变成一个质子,同时放出一个电子和一 个反中微子。这个过程看起来很简单,但是物理学家发现,电子的能量并不是一 个固定的值,而是在一个范围内分布的。这似乎违反了能量守恒定律,也就是说, 衰变前后的总能量应该是相等的。

为了解决这个问题,1930年,奥地利物理学家沃尔夫冈·泡利(Wolfgang

E.Pauli)提出了一个大胆的假设: β 衰变中除了电子之外,还有另一种粒子被放出,它是中性的,没有电荷,也没有质量,而且很难被探测到[12]。泡利称这种粒子为"中子",但是后来这个名字被用来指代另一种粒子,即原子核中的中子。因此,1934年,意大利物理学家恩里科•费米将泡利的粒子重新命名为"中微子",意思是"小中性粒子"[27-28]。

中微子的发现是一个艰难的过程,因为它们与物质的相互作用非常微弱,需要极大的灵敏度和精确度才能观测到。1942年,我国物理学家王淦昌第一次提出了中微子探测的具体原理 [29]。1956年,克莱德·科温(Clyde Cowan)、弗雷德里克·莱因斯(Frederick Reines,)等人利用核反应堆首次探测到了中微子 [30]。在核反应堆实验中产生大量的β衰变或电子和反中微子。其反中微子跟质子产生正电子和中子。当正电子和中子入射到探测器内,正电子跟电子发生湮灭产生γ光子,中子被吸收产生高能γ光子。这些光子可以被探测器探测到,从而可以推断出中微子的存在。

目前已发现的中微子有三种类型,分别是电子中微子、μ中微子和τ中微子。 它们与相应的带电轻子(电子、μ子和τ子)相关联,并在某些类型的放射性衰变 中产生。除了这三种中微子外,还有相应的反粒子,称为反中微子。

中微子是宇宙中最多的粒子之一,它们可以来自太阳、恒星、超新星、核反 应堆、宇宙射线等源头。中微子的发现开启了新的物理领域,让我们能够探索太 阳、恒星、超新星、黑洞等天体的内部情况,以及宇宙的起源和演化。

1.2.1 太阳中微子和标准太阳模型

标准太阳模型(Standard Solar Model, SSM)是一种描述太阳内部结构和演化的理论模型,它假设太阳是一个球对称的、处于静力平衡的、由氢和氦组成的气体球 [31-32]。标准太阳模型的历史可以追溯到 19 世纪末,当时 Lane (1869)首次写下了描述气体球静力平衡的方程组 [33]。20 世纪中叶,随着核物理和天体物理的发展,人们开始考虑太阳的能量来源和核反应过程,以及太阳的化学组成和年龄。Schwarzschild、Howard 和 Härm (1957)计算了第一批太阳的演化模型,跟踪了核燃烧核心将氢转化为氦的结构演化过程 [34-35]。Demarque and Percy (1964)引入了校准太阳模型的概念,即调整太阳模型的混合长度参数和氦丰度,使模型具有太阳的观测半径和光度,这相当于我们现在所说的标准太阳模型 [31-32]。从那以后,标准太阳模型不断地根据新的观测数据和理论进展进行修正和改进,例如太阳的化学丰度、太阳的振动模式、太阳的中微子通量等。

太阳的核心是太阳能产生的区域。在核心中,太阳的氢通过称为聚变的过程

转化为氦。这个过程释放出大量能量,其反应为:

$$4p \rightarrow {}^{4}\text{He} + 2e^{+} + 2v_{e} + 26.731\text{MeV}$$
 (1-1)

这个反应被称为4H过程,其4个质子最终转化为1个4He原子核,释放出2 个正电子、2个中微子和大量能量(26.7 MeV)。这里涉及到两种聚变过程,分别为 质子-质子链(Proton-Proton Chain, pp-chain)和碳-氮-氧循环(Carbon-Nitrogen-Oxygen Cycle, CNO-cycle),如图1–2所示。其 pp-chain 是在聚变过程中占主导 (99%),而 CNO-cycle 占很小一部分,这个过程只是在较高温度下很重要[36]。这 些过程中总共有8个反应能够产生中微子,它们的通量能谱在如图1–3中所示。在 低能区域 pp 中微子的通量最高,而在高能区域⁸B 中微子通量比较高。



图 1-2 pp-chain 和 CNO-cycle 过程的示意图 [36]。

尽管 SSM 在预测太阳结构和成分的许多方面取得了成功,但模型的预测与观测之间仍然存在一些差异。其中最值得注意的问题之一是"太阳中微子问题"。最早研究太阳中微子的实验之一是 Homestake 实验,该实验由雷蒙德•戴维斯(Raymond Davis)在 1960年代进行。该实验使用了 10 万加仑的干洗液(四氯乙烯)作为探测器,利用中微子和氯原子的反应来计数中微子 [35,37]。然而,该实验观测到的中微子通量只有理论预测的三分之一左右这个结果与后来的其他实验,如 SAGE 和 GALLEX,也大致一致 [38-39]。这个差异很难用标准太阳模型的变化来





图 1-3 太阳中微子通量能谱图 [36]。

解释。这种差异被称为太阳中微子问题,是 SSM 和我们对中微子的理解存在问题的首批迹象之一。

这个问题最初被认为是 SSM 的问题,但最终因中微子振荡(Neutrino Oscillation)的发现而得到解决。中微子振荡指的是中微子在运动过程中可以在不同的 类型之间转换。上文中提到,中微子是一种没有电荷,质量很小的基本粒子,有 三种类型,这三种类型与三种带电的轻子(电子、 μ 子和 τ 子)相对应,可以通 过弱相互作用产生或消灭。中微子振荡的产生与中微子的质量有关。如果中微子 没有质量,那么它们的类型就是固定的,不会发生变化。但是如果中微子有质量,那么它们的类型就不是与质量相对应的,而是由质量状态的叠加组成的。质量状态是指中微子在真空中以确定的速度运动的状态,它有三种质量状态且质量不相 等,目前还不知道它们的确切值。由于中微子的类型是由质量状态的叠加组成的,当中微子在真空中运动时,不同的质量状态会以不同的相位相互干涉,导致中微 子的类型随着时间和距离的变化而变化。这就是中微子振荡的基本原理。日本的 超级神冈(Super-Kamiokande, Super-K)实验在 1998 年首次报告了中微子振荡的 证据,它发现,来自地球大气层的 μ 中微子的通量随着中微子穿过的地球物质的 长度的变化而变化,这意味着 μ 中微子在运动过程中有一定的概率变成了其他类型的中微子,如 τ 中微子 [40-42]。

Super-K 探测器的发现不仅证明了中微子有质量,而且质量不相等,这与标准 模型的预测不符,也为解决太阳中微子缺失问题提供了一个可能的答案 [40-42]。 然而,为了确定太阳中微子是否真的发生了振荡,需要使用能够同时探测到中微 子的弱荷电流和弱中性流的探测器,如重水切伦科夫探测器。它是一种利用重水 (D₂O)作为中微子的靶和探测介质的探测器,当中微子与重水分子发生弱相互作用时,会产生一个带电的轻子和一个中子,这两个粒子都会在重水中发出切伦科夫辐射,可以被探测器周围的光电倍增管捕获。根据切伦科夫辐射的时间和空间分布,可以判断出中微子的类型和能量。重水切伦科夫探测器的一个优点是,它可以探测到中微子的弱荷电流和弱中性流。这就意味着,重水切伦科夫探测器可以区分出电子中微子和其他类型的中微子,从而更好地探测中微子振荡的效应[40-42]。

重水切伦科夫探测器的一个代表是位于加拿大的苏德伯里中微子观测站 (Sudbury Neutrino Observatory, SNO),它是由加拿大、美国和英国的科学家合 作建造的,位于加拿大的苏德伯里镍矿(Sudbury Nickel Mine)的地下,由一千 吨的重水和九千多个光电倍增管组成[43-44]。SNO 探测器在 2001 年首次报告了 太阳中微子振荡的证据,它发现,观测到的太阳中微子的总通量与理论预测的一 致,但是其中的电子中微子的通量比理论预测的要少得多,这意味着太阳中微子 在到达地球之前,有一部分变成了μ中微子或τ中微子[43-44]。SNO 实验不仅能 够证实中微子振荡的存在,并能够测量⁸B 中微子的通量,这是一种在太阳核心产 生的太阳中微子。该实验还能够测量太阳中微子与重水的中性流(Neutral Current, NC)和带电流(Charge Current, CC)相互作用,提供有关其特性的重要信息。中 性流相互作用是中微子与原子核相互作用而不改变其电荷的过程。这种相互作用 由以下反应方程式描述:

$$\nu + N \to \nu' + N' \tag{1-2}$$

在这个方程中, v 和 v' 代表初始和最终中微子状态, N 和 N' 代表初始和最终 原子核状态。中性流相互作用对所有三种类型的中微子的存在都很敏感,无论它 们的电荷如何。带电流相互作用是中微子在与原子核相互作用时改变其电荷的过 程。这种相互作用由以下反应方程式描述:

$$\nu + N \to \nu' + N' + e^{-} \tag{1-3}$$

在这个方程中,中微子从中性状态(v)变为带电状态(v'),原子核也改变状态(N到N')。CC相互作用一次只对一种类型的中微子敏感,这取决于它的电荷。除了中性流和带电流相互作用外,太阳中微子还可以与探测器中的原子发生弹性散射(Elastic Scattering, ES)。这种相互作用由以下反应方程式描述:

$$\nu + e^- \to \nu' + e^- \tag{1-4}$$

在这个方程式中,中微子散射电子,改变其方向和能量。弹性散射相互作用 对所有三种类型的中微子都很敏感,并提供有关它们特性的重要信息。

关于中微子研究,目前世界各地有不少正在进行的实验。例如,位于意大利的 Borexino (Boron solar neutrino Experiment)实验。该实验使用一大罐有机液体闪烁 体,通过它们与电子的相互作用来探测太阳中微子。Borexino 实验对太阳中微子通 量进行了重要测量,并且还限制了其他类型中微子的存在 [45-47]。正在进行的实 验的另一个实验是位于日本的 KamLAND (Kamioka Liquid Scintillator Antineutrino Detector)实验。该实验使用一大罐液体闪烁体通过与质子的相互作用来探测太阳 中微子。KamLAND 实验也对太阳中微子通量进行了重要测量,也为中微子振荡 提供了证据 [48-50]。除了这些实验之外,还有许多其他正在进行和计划中的实验 在研究太阳中微子。例如,目前中国正在建设中的 JUNO (Jiangmen Underground Neutrino Observatory)实验将使用一个大型液体闪烁体罐,通过它们与电子的相 互作用来探测太阳中微子。该实验将能够对太阳中微子通量进行精确测量,也将 能够研究中微子的振荡 [51-52]。

1.2.2 中微子-核子相干弹性散射

中微子-核子相干弹性散射 (Coherent Elastic Neutrino-Nucleus Scattering, CEvNS) 是一种弱相互作用过程,在该过程中,中微子从原子核中散射出来,但不 会改变原子核的能量或电荷状态。与带电电流散射等其他形式的中微子-核相互作 用不同,CEvNS 是一个相干过程,这意味着原子核表现为单个粒子,而不是单个 质子和中子的集合。CEvNS 是一个中性流过程,这意味着它涉及中微子和原子核 之间称为 Z 玻色子的中性弱力载体粒子的交换。相反,带电流过程涉及称为 W 玻 色子的带电弱力载体粒子的交换。中性流过程很重要,因为它们使我们能够研究 中微子的基本特性,例如它们的质量和混合角,以及原子核的弱荷。

CEvNS 的第一个预测是由 D. Z. Freedman 于 1974 年做出的 [53-55]。在这个预测中, Freedman 证明了 CEvNS 的散射截面与原子核弱荷的平方成正比,它由原子核中的质子数决定核。CEvNS 的散射截面由以下公式给出 [54]:

$$\frac{\mathrm{d}\sigma}{\mathrm{d}T} = \frac{G_F^2 M}{4\pi} \left(1 - \frac{MT}{2E_v^2} \right) Q_w^2 \left[F_w \left(q^2 \right) \right]^2 \tag{1-5}$$

其中 G_F 是费米耦合常数, $T = E_R = q^2/(2M) = E_v - E'_v$ 为核反冲能, M_N 是 核质量, $F_w(q^2)$ 是弱形状因子, $E_v(E'_v)$ 是入射中微子(出射中微子)的动能, M为靶核的质量, Q_w 弱荷。

在 CEvNS 过程中,中微子不会转化为另一种粒子,它只是从目标上反弹并继

续前进。来自散射核的低能反冲是唯一可探测到的特征。直到最近几年,探测器 技术才进步到足以观测这些低能核反冲。CEνNS 是 2017 年首次由 COHERENT 实 验合作组观测到 [54,56]。COHERENT 是一个国际物理学合作项目,旨在利用美 国田纳西州橡树岭国家实验室的中子源(Spallation Neutron Source, SNS)产生的 高通量中微子束。SNS 是一个高能质子加速器,它将质子束打到一个液态汞靶上, 产生大量的中子和 π 介子。π 介子会在很短的时间内衰变成 μ 介子和中微子,而 μ 介子则会停留在靶周围的铝壳中,继续衰变成电子和更多的中微子。这样就形 成了一个高强度、多种(电子中微子、μ 中微子和 τ 中微子)的中微子束 [57],如 图1-4所示。COHERENT 使用了不同类型和大小的探测器,如碘化钠闪烁体、锗 半导体、铯碘闪烁体等,来增加对 CEνNS 信号的探测灵敏度。中微子与原子核发 生弱相互作用时,原子核会以相干和弹性的方式反冲,并释放出一定的能量。

CEvNS 过程经常与"中微子地板"相关讨论,这是寻找暗物质粒子的实验灵 敏度的理论极限 [58]。一般,来自太阳、大气和超新星的中微子可以通过与暗物 质探测器中的靶原子核发生相干弹性散射,且在寻找暗物质的实验中产生可观测 的核反冲信号。它们是只能通过它们的反冲谱来区分。因此,中微子本底限制了 寻找暗物质的实验的灵敏度,它是这些实验设计和分析中需要考虑的重要因素。



图 1-4 在 SNS 中产生中微子的物理过程 [59]。

1.3 暗物质

从哥白尼那里,我们了解到地球不是宇宙的中心,从伽利略那里,我们知道 了世界远比太阳系更大。而哈勃天文望远镜的观测结果告诉我们,银河系是不断 膨胀的宇宙的一小部分。近几十年来的技术进步使物理学家和天文学家能够在宇 宙学尺度上测量宇宙的性质。来自宇宙微波背景(Cosmic Microwave Background, CMB)最新结果表明,我们的宇宙由 4.9% 的普通物质、26.8% 的暗物质和 68.3% 的暗能量组成 [60-62]。寻找以及理解暗能量和暗物质是已成为 21 世纪物理学界 最具挑战性的任务。

1.3.1 暗物质存在的依据

1.3.1.1 星系旋转曲线

天体物理学中暗物质存在的主要证据之一是星系旋转曲线(Galactic Rotation Curve)[63-64]。星系中恒星的旋转运动可以用来推断星系的质量。例如,如果我们考虑一颗围绕星系中心运行的恒星,我们可以使用经典力学定律来确定星系的质量。根据牛顿第二定律以及万有引力定律可知:

$$F = \frac{GM_{\odot}m}{r^2} = \frac{mv^2}{r} \tag{1-6}$$

其中 F 是引力, G 是引力常数, M_{\odot} 是星系的质量, m 是恒星的质量, r 是星系与恒星之间的距离。则 v 是恒星的旋转速度 [65]:

$$v = \sqrt{\left(\frac{GM_{\odot}}{r}\right)} \tag{1-7}$$

旋转曲线方程可用于通过测量星系中恒星的旋转速度来确定星系的质量。然 而,观测到的星系中恒星的旋转速度远高于基于星系可见质量的预期。这表明必 须有额外的看不见的质量,称为暗物质。

例如三角座星系(Messier 33)的旋转曲线如图1-5所示。其横轴轴表示距星 系中心的距离,纵轴表示恒星的自转速度。实线表示观测到的恒星自转速度,而 虚线表示基于星系可见质量的预期自转速度。可以看出,观测到的旋转速度远高 于预期,表明存在暗物质。

1.3.1.2 引力透镜

星系引力透镜(Gravitational Lens)是天体物理学中暗物质存在的另一个关键 证据 [67]。引力透镜是一种由广义相对论预言的现象,它指的是当光线经过一个 大质量天体附近时,由于时空的弯曲而发生偏折。为了理解引力透镜的原理,我 们可以考虑点状光源和点状透镜的情况。在这种情况下,来自光源的光被透镜的 引力场弯曲和聚焦,从而产生扭曲和放大的光源图像。失真和放大的幅度取决于 光源、镜头和观察者之间的距离,以及镜头与观察者之间的质量和距离。引力透



图 1-5 三角座星系(Messier 33)的旋转曲线,(黄色和蓝色点),以及从可见物质分布(灰色线)预测的旋转曲线。这两条曲线之间的差异可以通过在星系周围添加暗物质晕来解释[66]。

镜效应可以使用引力透镜方程进行量化,该方程涉及源、透镜和观察者的位置和 距离。引力透镜方程由以下公式给出 [68]:

$$\beta = \theta - \frac{D_{\rm LS}}{D_{\rm S}}\alpha\tag{1-8}$$

其中 β 是光源相对于镜头的位置, θ 是图像相对于镜头的位置, D_{LS} 是镜头到 光源的距离, D_{S} 是观察者到光源的距离, α 是透镜引力场引起的光线偏转角。

引力透镜方程可用于计算产生给定透镜效应所需的源、透镜和观察者的位置 和距离。例如,如果我们知道光源和透镜的位置和距离,我们就可以使用引力透 镜方程来计算图像的位置和光线的偏转角。同样,如果我们知道光源和图像的位 置和距离,我们可以使用引力透镜方程来计算透镜的位置和光线的偏转角。引力 透镜方程也可以用来确定透镜的质量。由于光线的偏转角取决于透镜的质量,我 们可以利用观察到的透镜效应来推断透镜的质量。这可以通过重新排列引力透镜 方程并求解透镜的质量来完成,如下式所示 [68]:

$$m = \frac{c^2}{4\pi G} \frac{D_{\rm LS} D_{\rm L} \theta}{D_{\rm S}} \tag{1-9}$$

其中 *m* 是透镜的质量, *c* 是光速, *G* 是万有引力常数, *D*_{LS} 是透镜到光源的 距离, *D*_L 是距离观察者与镜头之间的距离, *D*_S 为观察者与光源之间的距离, *θ* 为 光线的偏转角。

通过望远镜和其他仪器可以观察星系的引力透镜效应。例如,星系团 Abell 1689 的引力透镜效应 [69],如图1-6所示。

第一章 引言



图 1-6 星系团 Abell 1689 引力透镜的图 [69]。

引起引力透镜效应所需的引力远大于星系的可见质量。这表明必须有额外的 看不见的质量,即暗物质。例如,星系团 Abell 1689 的引力透镜(如图1-6所示) 已被用于估计该星系团的总质量。星团的质量可以使用引力透镜方程计算。观察 到的透镜效应以及背景和前景星系的位置和距离可用于确定星团的质量。该星团 的估计质量约为 10¹⁵ 个太阳质量,远大于该星团中星系的可见质量。这表明星团 的大部分质量是暗物质。

1.3.1.3 子弹星系团

子弹星系团也是能够表明暗物质存在的天文现象之一 [70]。当两个高速移动 的星系团发生碰撞时,星系团内部的可见星体与不可见的暗物质都会发生碰撞。但 是由于普通物质与普通物质之间碰撞的散射截面要比暗物质之间的散射截面大得 多,可见的普通物质会在星系团与星系团碰撞之后,相对较快的损失动能,并滞 留在碰撞点附近。而暗物质不参与电磁相互作用,只有引力相互作用,可能参与 弱相互作用,在星系团与星系团碰撞的过程中,损失动能较慢,可能会在碰撞点 比较远处滞留。这种普通发光物质和暗物质分离的现象,就好像高速运动的星系 团将内部的暗物质打了出去,如图1-7所示。星系团总体的质量可以通过引力透镜 效应给出,发光星体的质量可以由地球附近的探测器给出。这两种质量的差异说 明了暗物质的存在。


图 1-7 子弹星云现象 [71]。

1.3.1.4 宇宙微波背景辐射

宇宙微波背景辐射是天体物理学中暗物质存在的另一个关键证据 [60-62]。宇 宙微波背景辐射是一种微弱的电磁辐射,可在整个宇宙中观察到。它被认为是大 爆炸遗留下来的残余辐射。宇宙微波背景辐射提供了有关宇宙结构和演化的重要 线索。例如,观察到 CMB 辐射具有黑体光谱,这是热辐射的特征。这表明宇宙微 波背景辐射在发射时处于热平衡状态,并为宇宙在其演化早期阶段的高温、致密 状态提供了证据。

宇宙微波背景辐射也被观测到有小的温度波动,这被认为是宇宙中大尺度结构的"种子",如星系和星系团。观察到这些温度波动具有标度不变的功率谱,这 与宇宙膨胀模型一致。膨胀模型预测宇宙在其演化的早期阶段经历了一段快速膨 胀的时期,这会产生观察到的温度波动。宇宙微波背景辐射的存在也与宇宙中暗 物质的存在相一致。宇宙微波背景辐射具有少量极化,这是由自由电子的光子散 射引起的。观察到的偏振与大量暗物质的存在相一致,暗物质会增大引力并导致 光子散射。

宇宙微波背景辐射可用于估计宇宙中普通物质、暗物质和暗能量的相对数量。 例如,普朗克卫星任务高精度地测量了宇宙微波背景辐射的温度和极化,并利用 这些测量结果来估计宇宙的成分。普朗克任务的最新结果表明,宇宙由 4.9% 的普 通物质、26.8% 的暗物质和 68.3% 的暗能量组成 [60-62]。

总体而言,宇宙微波背景辐射为天体物理学中暗物质的存在提供了有力证据。 宇宙微波背景辐射的温度和极化无法用已知的物理定律来解释,并且需要存在额 外的、看不见的质量,这些质量对宇宙的引力有贡献。这种看不见的物质被称为



图 1-8 宇宙微波背景全天图 [72]。

暗物质,被认为是由不与光相互作用的粒子组成的,因此望远镜和其他仪器看不 到它。

1.3.2 暗物质候选粒子

在介绍暗物质候选者之前,我们将首先讨论暗物质的性质。顾名思义,暗物 质原则上不能跟光子发生任何相互作用,换句话说,它不发光也不吸收光,它跟 普通物质间的相互作用应该很微弱,暗物质不能由标准模型粒子组成(除中微子 之外),暗物质大概率会是"冷",它是稳定的且寿命至少相当于宇宙的年龄。虽然 有些理论家认为暗物质由重子(Baryon)构成的并且试图用重子来解释,但是主 流的理论认为暗物质应该归类为非重子。非重子暗物质根据其运动速度可分为热 暗物质(Hot Dark Matter),温暗物质(Warm Dark Matter),冷暗物质(Cold Dark Matter)[21,73-74]。

热暗物质是指以接近光速的高速运动的暗物质粒子。热暗物质的粒子包括中 微子 [21]。热暗物质能够穿过星系和其他大尺度结构而不与它们相互作用,因此 无法形成我们在宇宙中观察到的结构。温暗物质是指以中等速度运动的暗物质粒 子,温暗物质的例子包括质量在 10-100MeV 范围内的粒子 [21]。其中比较典型温 暗物质候选粒子是惰性中微子(Sterile Neutrinos)[75-76]。冷暗物质是指以低速 运动的暗物质粒子,比光速慢得多。冷暗物质能够形成我们在宇宙中观察到的大 尺度结构,例如星系和星系团。一般大部分主流的暗物质候选粒子都属于冷暗物 质,例如,弱相互作用大质量粒子(Weakly Interacting Massive Particles, WIMPs)、 轴子(Axion)、超中性子(Neutralinos)、重力微子(Gravitinos)、Kaluza-Klein 粒子、暗光子(Dark Photon)等都属于冷暗物质 [21]。

1.3.2.1 WIMPs

WIMPs 是冷暗物质的主要候选者之一。WIMPs 的存在可以解释宇宙微波背 景辐射中观测到的扰动模式。宇宙微波背景辐射是最早期宇宙辐射在经过数十亿 年后留下的遗迹,反映了当时宇宙结构形成的初期条件。如果没有 WIMPs 这样的 冷暗物质存在,那么宇宙微波背景辐射中扰动模式会更加平滑和均匀。WIMPs 的 形成可以通过冻结机制(Freeze-out Mechanism)来解释 [77-78]。在早期高温高密 度的宇宙中,WIMPs 与普通物质之间频繁地发生湮灭和产生反应,并保持热平衡 状态。随着时间推移,由于宇宙膨胀率(Hubble Parameter)*H* 的增加以及温度 *T* 的降低,湮灭反应变得越来越少见,并且无法跟上产生反应所需的速率。因此,在 某个临界时刻 t_f 之后,WIMPs 就从热平衡状态退耦(Decouple),并停止与普通 物质发生有效交互。

这个临界时刻 t_f 可以用玻尔兹曼方程(Boltzmann Equation) 来描述 [79-80]:

$$\frac{dn}{dt} + 3Hn = -\langle \sigma v \rangle (n^2 - n_{\rm eq}^2)$$
(1-10)

其中 *n* 是 WIMPs 的数密度(Number Density), $\langle \sigma v \rangle$ 是 WIMPs 湮灭截面 (Annihilation Cross Section) 与相对速度 *v* 乘积的平均值, n_{eq} 是 WIMPs 在热平衡 状态下达到的数密度。

当 $t > t_f$ 时,右边第二项可以忽略不计,并得到一个简单近似解:

$$n(t) = n(t_f)e^{-3\int_{t_f} tHdt}$$
 (1-11)

由于 H 随时间变化很小,在这个近似下可以认为 $H \approx H(t_f)$,并得到:

$$n(t) = n(t_f)e^{-3H(t_f)(t-t_f)}$$
(1-12)

这个方程表明,在解耦时,WIMPs的数密度随着宇宙膨胀而指数衰减。这个过程称为冻结机制,因为WIMPs的平均能量随着温度降低而不断下降。由于WIMPs的湮灭截面与其质量成正比,而宇宙膨胀率与其质量成反比,因此WIMPs的解耦时刻与其质量有关。一般来说,质量越大的WIMPs解耦得越早,数密度越小。这就导致了一个有趣的现象,被称为WIMPs奇迹(WIMP Miracle)[81]:如果假设WIMPs的湮灭截面与弱相互作用的典型值相当(约 10⁻²⁶ cm³/s),那么在解耦后留下的暗物质丰度(Dark Matter Abundance)恰好与观测值一致(约 0.1)[82]。

暗物质丰度可以用以下公式来定义:

$$\Omega_{\rm DM} = \frac{\rho_{\rm DM}}{\rho_c} \tag{1-13}$$

其中 ρ_{DM} 是暗物质的能量密度(Energy Density), ρ_c 是临界能量密度(Critical Energy Density),即使得宇宙处于平坦状态的最小能量密度。

在解耦后, WIMPs 的遗迹密度(Relic Density)可以用以下公式来估算[21,79]:

$$\Omega_{\rm DM} h^2 \approx \frac{3 \times 10^{-27} {\rm cm}^3/s}{\langle \sigma v \rangle}$$
(1-14)

其中 h 是哈勃常数 H₀ 除以 100 km/s/Mpc 的无量纲值。由于 WIMPs 在解耦 后不再参与有效交互,因此它们的运动速度会随着温度降低而逐渐减小,并在某 个时刻达到动力学解耦(Kinetic Decoupling)。动力学解耦是指 WIMPs 与普通物 质之间弹性散射停止发生的过程。动力学解耦之后,WIMPs 就完全脱离了普通物 质,并形成了自己的结构和分布。



图 1-9 早期宇宙中 WIMPs 的数密度 [83]。

1.3.2.2 轴子

轴子是一种假想的粒子,大约是 1970 年代为了解决强 CP 问题所提出的一个 假想粒子 [84-86]。强 CP 问题是指为什么量子色动力学(QCD)中的强相互作用 不违反 CP 对称性,即物质和反物质之间的差异。Peccei-Quinn 理论是一种解决强 CP 问题的方法,它引入了一个新的对称性和一个新的场,这个场被量子化后就产 生了轴子。轴子是一种 Nambu-Goldstone 玻色子,也就是说,它是由于自发对称性 破缺而产生的无质量或近似无质量粒子。

轴子的质量与其衰变常数成反比,衰变常数又与 Peccei-Quinn 对称性破缺的 能标成正比。因此,轴子可以有很大范围的可能质量,从 10-6 电子伏特到 1012 电子伏特都有可能 [84-86]。目前还没有实验上确定轴子的质量。

轴子可以与其他粒子发生相互作用,其中最重要的两种是与光子和电子的耦合。轴子-光子耦合(Axion-Photon Coupling)指的是在外磁场或者等离子体中,轴子可以转化为光子或者光子可以转化为轴子。这种转化过程叫做 Primakoff 效应。轴子-光子耦合的强度由一个无量纲的耦合常数 gay 决定,它与轴子的质量成正比。轴子-电子耦合(Axion-Electron Coupling)指的是轴子可以与电子交换光子或者与电子发生康普顿散射。轴子-电子耦合的强度由另一个无量纲的耦合常数 gae 决定,它与轴子的质量成反比。

轴子也是一种暗物质的候选者之一,因为它们是中性、稳定、非相对论性和 非碰撞性的粒子。太阳是一种可能的轴子源,因为在太阳内部,高温和高密度导 致了大量的 Primakoff 效应,使得一部分光能转化为轴子动能 [87]。太阳轴子可以 逃逸到太空,并且可以被地球上的探测器捕获。目前有两种主要的实验方法来探 测太阳轴子,一种是利用轴子-光子耦合,另一种是利用轴子-电子耦合。

利用轴子-光子耦合的实验方法是在一个强磁场中,让太阳轴子转化为可见光 或 X 射线,并且用光学或 X 射线望远镜来观测。这种方法的代表性实验有 Axion Dark Matter Experiment (ADMX) [88] 和 CERN Axion Solar Telescope (CAST) [89]。 ADMX 是一个地下实验,它使用了一个超导磁体和一个微波谐振腔来探测太阳轴 子转化为微波光子的信号。CAST 是一个地面实验,它使用了一个废弃的 LHC 磁 体和一个 X 射线望远镜来探测太阳轴子转化为 X 射线光子的信号。

利用轴子-电子耦合的实验方法是在一个强磁场中,让太阳轴子与电子发生康 普顿散射,产生可见光或 X 射线,并且用光学或 X 射线望远镜来观测。这种方法 的代表性实验有 International Axion Observatory (IAXO) [90] 和 Dark Matter Radio (DM Radio) [91]。IAXO 是一个地面实验,它使用了一个大型的磁体和多个 X 射 线望远镜来探测太阳轴子转化为 X 射线光子的信号。DM Radio 是一个地下实验, 它使用了一个超导磁体和一个谐振腔来探测太阳轴子转化为微波光子的信号。

1.3.2.3 暗光子

暗光子是粒子物理学标准模型之外的许多理论所预测的假设粒子。这些粒子 类似于构成光的普通光子,但它们与普通物质的相互作用被认为非常微弱。这使 得它们极难被发现,因此被称为"暗"光子。研究暗光子的一个关键参数是暗光 子与普通光子之间的混合角。这个角度用 θ 表示,描述了暗光子和光子混合在一 起的程度。在许多理论中,这个角度极小,意味着暗光子和光子几乎完全分开。暗 光子的质量是另一个重要参数。在许多理论中,预计暗光子的质量非常小,大约 为几个电子伏特。这远小于普通光子的质量,后者为零。

暗光子的特性对我们对宇宙的理解具有重要意义。例如,如果暗光子具有非 常小的混合角和非常小的质量,它可能是暗物质的潜在候选者,暗物质是构成宇 宙中大部分质量的神秘物质。

寻找暗光子是一个活跃的研究领域,目前正在进行许多实验来检测这些粒子。 例如,杰斐逊实验室的 DarkLight 实验正在使用高强度激光束寻找暗光子 [92]。其 他实验,例如大型强子对撞机的 LHCb 实验,也在寻找高能碰撞中暗光子的迹象 [93]。

1.3.2.4 惰性中微子

惰性中微子是一种假设粒子,类似于更广为人知的活跃中微子,但不参与弱 核力相互作用。这意味着它们不参与普通中微子所做的过程。因此,它们是"惰 性",因为它们不像活跃的中微子那样与其他粒子相互作用。有证据表明惰性中微 子可能存在,但它们的性质和行为尚不清楚。特别是,它们的质量以及它们如何 与其他粒子相互作用目前尚不清楚。一些理论表明惰性中微子可能能够通过一种 新的、尚未发现的力或通过引力与其他粒子相互作用。

惰性中微子引起物理学家的兴趣,因为它们可能为粒子物理学和宇宙学中的 几个未解之谜提供潜在的解释。例如,目前对中微子振荡的理解,即中微子从一 种类型转变为另一种类型的过程,并不能解释所有观察到的现象。惰性中微子的 存在可能有助于解释这些差异。此外,惰性中微子可能为宇宙学中暗物质问题提 供解决方案。一些理论认为惰性中微子可能是一种暗物质,因为它们很难检测到 并且可以解释观察到的引力效应。

有几项正在进行的实验和观察旨在检测和研究惰性中微子。例如,美国费米 实验室的 MiniBooNE 实验旨在通过寻找探测器中中微子与预期行为的偏差来寻 找惰性中微子的证据 [94]。同样,位于南极的冰立方(IceCube)中微子天文台正 在通过观察大量冰中的高能中微子相互作用来寻找惰性中微子 [94-95]。

-18-

1.4 暗物质的探测

科学家们使用几种不同的方法来尝试探测暗物质,包括直接探测、间接探测 和对撞机实验。

直接探测包括寻找暗物质粒子与普通物质的微弱相互作用。这可以使用灵敏的仪器来完成,这些仪器能够探测暗物质粒子与原子核碰撞时释放的微量能量。这些实验通常在地下深处进行,以保护它们免受宇宙射线的影响。到目前为止,还没有使用这种方法发现暗物质粒子的确凿证据。

间接探测涉及寻找暗物质粒子对宇宙中其他现象的影响。例如,一些暗物质 粒子可能衰变或湮灭,产生伽马射线或中微子等可探测粒子。通过研究这些粒子 的分布和性质,科学家们可以推断出暗物质的存在。然而,这种方法具有挑战性, 因为来自暗物质的信号通常微弱并且难以与其他辐射源区分开来。

寻找暗物质的另一种方法是在实验室中使用粒子对撞机产生暗物质。这些强 大的机器将粒子加速到接近光速,然后将它们碰撞在一起,产生可以产生暗物质 粒子的高能碰撞。通过研究这些粒子的特性,科学家可以更多地了解它们的特性 和行为。



图 1-10 探测暗物质的三种通道。

1.4.1 直接探测

暗物质直接探测技术涉及使用能够探测暗物质粒子与普通物质相互作用的传 感器。这种相互作用可以多种方式表现出来,例如原子的电离、闪烁光的产生或 热能的转移。

暗物质直接探测技术涉及使用对暗物质粒子与探测器中的原子相互作用引起 的电子电离或散射敏感的探测器。这些探测器通常由目标材料(例如晶体或液体) 组成,目标材料周围环绕着测量电子电离或散射的灵敏探测器件。当暗物质粒子 通过探测器时,它可能与目标材料中的原子相互作用,导致电子释放或散射。然 后探测器中的灵敏器件能够探测到这些变化,从而可以识别暗物质粒子。

通常情况下,暗物质跟探测器内的普通物质间发生相互作用后,可能会产生三种信号,分别为光、电和热信号。目前世界上大部分的暗物质直接探测实验均利用这三种信号来试图寻找暗物质信号。不过,目前还没有同时能探测三种信号的探测实验,它们最多也能探测两种信号。部分实验还仅利用一种信号来寻找暗物质。比如位于意大利 Laboratori Nazionali del Gran Sasso (LNGS) 地下实验室的 XENON 暗物质直接探测实验能够同时探测光和电信号 [96],同样在 LNGS 的 CRESST (Cryogenic Rare Event Search with Superconducting Thermometers)能够同时探测光信号和热信号 [97],还有位于美国明尼苏达州 Sudan Mine 的 Cryogenic Dark Matter Search (CDMS) 暗物质直接探测实验能够测量电和热信号 [98]。下面将简单介绍几种典型暗物质探测实验。



图 1-11 基于各类探测技术,世界各地的暗物质直接探测实验的总结。

1.4.1.1 CDMS

CDMS 是一种典型的暗物质直接探测实验 [98]。CDMS 实验有多个阶段和地 点,其中最早的一批实验在斯坦福大学进行,后来迁移到明尼苏达州苏丹矿山地 下深处,以减少宇宙线的干扰。目前正在建设的 SuperCDMS 实验将在加拿大安大 略省萨德伯里地下实验室进行。SuperCDMS 实验将使用更多、更灵敏、更先进的 探测器,并采用新型电子反冲判别技术,以提高对低质量 WIMPs 的灵敏度 [99]。 CDMS 实验利用低温探测器来寻找 WIMPs 与原子核散射产生的微弱信号。低温 探测器由锗或硅制成的半导体晶体组成,它们被冷却到接近绝对零度。CDMS 实 验利用了一种叫做"离子化热法"的技术。当 WIMPs 与晶体中的原子核碰撞时, 会在晶体中产生热量和电荷信号,并根据信号的比例来区分 WIMPs 信号和其他本 底信号。晶体探测器被安装在一个超导磁体中,以产生一个强大的磁场,可以抑 制来自宇宙射线的背景信号。晶体探测器被连接到一系列的传感器和放大器,用 于测量热量和电荷信号,并将其转换为可读取的电压信号。传感器和放大器也被 冷却到低温,以减少噪声和热漂移。图1–12展示,经典 CDMS 探测器探测原理示 意图。



图 1-12 CDMS 探测器原理示意图 [100]。

1.4.1.2 CRESST

CRESST 暗物质探测项目 [97], 位于意大利的 LNGS 地下实验室。CRESST 实验使用低温技术来制造高灵敏度的探测器,这些探测器由两部分组成:一个吸收

暗物质粒子能量并发出光信号的闪烁体晶体和一个测量晶体温度变化并放大电信号的超导热计。通过同时记录光信号和电信号,可以区分不同类型的事件,并排除背景噪声。目前,CRESST 实验处于第三阶段(CRESST-III),主要关注降低能量阈值以提高对轻暗物质粒子的灵敏度[101]。在CRESST-III第一阶段,该实验运行了13个探测器模块,并取得了一些的结果。通过采用新的探测器技术和概念,该实验的灵敏度显著提高,尤其是对轻暗物质(Light Dark Matter)。CRESST-III使用了新型闪烁体晶体材料——钙钼酸镧(CaWO₄)和氧化镧铝(LaAlO₃),这些材料具有较低的放射性背景、较高的光产额和较小的能隙。此外,为了降低能量阈值,超导热计也进行了改进,使用了较薄、较小、较轻且具有更高过渡温度(T_c)和更低噪声水平(NEP)的金属铝片作为敏感元件。图1–12展示,CRESST 探测器探测原理示意图。



图 1-13 CRESST 探测器原理示意图 [102]。

1.4.1.3 DAMA/LIBRA

DAMA/LIBRA(DArk MAtter/Large sodium Iodide Bulk for RAre processes)也 是位于意大利 LNGS 地下实验室的,以直接探测暗物质为目标的闪烁探测器,它 只能探测光信号 [103-104]。DAMA/LIBRA 是上一代的 DAMA/NaI 的升级版,探 测器原理以及结构基本上比较类似。该探测器由 25 个高度放射性纯的闪烁铊掺杂 碘化钠 (NaI(Tl)) 晶体组成,这些晶体放置在一个 5 x 5 矩阵中。每个晶体都耦合 到两个低背景光电倍增管。检测器放置在一个密封的铜盒内,该铜盒用高纯氮气 冲洗;为了减少自然环境背景,铜盒周围环绕着低背景多吨屏蔽。此外,由LNGS 岩石材料制成的1米厚混凝土几乎完全包围了这个被动防护罩。

该探测器观察了到了所谓的"年度调制"特征 [103]。年调制是指实验探测到的暗物质粒子数量的季节性变化。这种变化被认为是由于地球绕太阳运行的轨道造成的。因此,实验检测到的暗物质粒子的数量可能会因一年中的不同时间而有所不同。这种年度调制被认为是暗物质的潜在特征。然而 CoGeNT, COSINE-100, SABRE 等类似的实验的结果表示,没有明显的年度调制现象 [105]。对于年度调制的可能解释是,这是有可能数据分析流程导致的。2022 年 8 月, COSINE-100应用了一种类似于 DAMA/LIBRA 使用的数据分析方法,发现了类似的年度调制,表明该信号可能只是一种统计伪像,从此这一假设得到了新的支持 [105]。



图 1-14 DAMA/LIBRA NaI 探测器原理示意图 [106]。

1.4.1.4 XENON、LZ 和 PandaX

XENON(位于 LNGS)[96], LUX-ZEPLIN 或 LZ(位于美国 Sanford 地下实验室)[107]以及 PandaX(位于中国四川省锦屏地下实验室)[108]都采用气-液氙二相性时间投影室技术来探测暗物质信号。三个实验采用的实验技术基本相同,但是实验细节、基础设施以及噪声水平有所区别。这种气-液氙二相性时间投影室同时能探测光和电信号。目前在各种暗物质探测器实验当中此类实验有着很大优势,并且能够探测以及覆盖比较广泛信号区域。关于此类实验在第二章会有更详细的介绍。

1.4.2 间接探测与对撞机实验

暗物质间接探测是一种利用暗物质与其反粒子湮灭或衰变产生的标准模型粒 子来寻找暗物质的方法。这些标准模型粒子包括光子、带电宇宙射线和中微子,它 们可以在不同的波段和能量范围内被探测到。暗物质间接探测的优势是可以利用 已有的天文观测设备和技术,而不需要特别设计新的探测器1。暗物质间接探测 的挑战是要区分暗物质信号和其他天然或人造背景源的干扰,以及要考虑暗物质 分布、湮灭或衰变截面等不确定因素的影响。

目前,有许多实验正在进行或计划进行暗物质间接探测,其中一些主要实验 如下:

- DAMPE: 这是中国主导的一个暗物质探测卫星,叫做"悟空"(Dark Matter Particle Explorer, DAMPE),它于 2015 年 12 月发射升空,是中国科学院空间科学战略性先导专项的第一个任务 [109]。它可以观测从几十 GeV 到几百 TeV 范围内的带电宇宙射线,包括电子、正电子、质子、反质子和重核等。它可以搜索带电宇宙射线能谱中可能存在的异常结构或特征,例如正电子比率增加或断崖等,这些可能由暗物质湮灭或衰变产生。它还可以观测高能伽马射线和 X 射线。
- Fermi-LAT:这是一个安装在费米伽马射线空间望远镜上的大面积望远镜 (Large Area Telescope)[110],它可以观测从 20 MeV 到 300 GeV 范围内的 伽马射线。Fermi-LAT 可以搜索来自银河系中心、卫星星系、星系团等地方 的异常伽马射线信号,这些地方预期有较高密度的暗物质分布。Fermi-LAT 还可以搜索来自太阳或地球方向的伽马射线信号,这些信号可能由被太阳 或地球捕获后湮灭或衰变的暗物质产生。
- H.E.S.S.: 这是一个位于纳米比亚高原上的大气切伦科夫望远镜阵列(High Energy Stereoscopic System)[111],它可以观测从几十 GeV 到几十 TeV 范围内的伽马射线。H.E.S.S. 可以搜索来自银河系中心、卫星星系、星系团等地方的异常伽马射线信号,以及来自超新星遗迹等潜在加强因子(boost factor)区域的信号。
- AMS-02:这是一个安装在国际空间站上的阿尔法磁谱仪(Alpha Magnetic Spectrometer)[112],它可以观测从几百 MeV 到1 TeV 范围内带电宇宙射线,包括电子、正电子、质子、反质子、氦核和反氦核等。AMS-02 可以搜索带电宇宙射线能谱中可能存在的异常结构或特征,例如正电子比率(positron fraction)增加或反核出现等,这些可能由暗物质湮灭或衰变产生。
- IceCube: 这是一个位于南极冰层深处的中微子观测站,由 5000 多个光学传感器组成,它们可以探测到由中微子与冰或附近岩石相互作用产生的带电粒子发出的切伦科夫光 [113]。IceCube 可以搜索来自银河系中心、太阳或地球方向的异常中微子信号,这些信号可能由暗物质湮灭或衰变产生。

对撞机实验是在能量前沿领域进行的一种实验,它通过高速碰撞原子核或亚 原子粒子来产生新的粒子,并用探测器来观测和分析这些粒子。LHC 是目前世界 上最大的对撞机,它可以将两束质子加速到接近光速,并在四个不同的地点进行 碰撞。如果暗物质粒子存在并且足够轻,那么它们有可能在 LHC 中被制造出来。 然而,由于暗物质粒子与普通物质之间很难相互作用,它们会逃逸出探测器而不 留下任何痕迹。

那么如何在对撞机实验中寻找暗物质呢?一个可能的方法是利用所谓的"缺 失能量"信号。当两束质子碰撞时,会释放出大量能量,并转化为各种各样的粒 子。这些粒子会被探测器记录下来,并且可以计算出它们总共携带了多少能量和 动量。如果发现这些能量和动量与碰撞前不相等,那么就意味着有一部分能量和 动量消失了。这可能表明有一些未知的粒子被产生出来,并且带走了一部分能量 和动量。这些未知的粒子有可能就是暗物质粒子 [114]。

第二章 PandaX 实验

PandaX (Particle AND Astrophysical Xenon experiment) 是一项位于中国四川 省锦屏地下实验室 (China Jinping Underground Laboratory, CJPL)的,其主要物理 目标为探测暗物质的多物理目标实验项目。CJPL 是目前全球最深的地下实验室。 该实验项目始于 2009 年,由上海交通大学主导发起,并在 2022 年时已有 20 多个 科研单位参与其中。

PandaX 实验总共分为四期,其中第一期 PandaX-I(运行时间为 2014 年 5 月 至 10 月),第二期 PandaX-II(运行时间为 2016 年 5 月至 2019 年 7 月)以及第四 期 PandaX-4T(运行时间为 2020 年底至今)是以气-液氙二相型时间投影室(Time Projection Chamber, TPC)技术为核心的暗物质以及多物理目标探测实验。PandaX-I、PandaX-II 以及 PandaX-4T 的主要差别在于探测器靶物质氙的质量不同,其质量 分别为 120kg、500kg 以及 3.7 吨。通过不断提高靶物质质量,可以有效提高探测 器灵敏度。第三期的 PandaX-III(正在研发中)的物理目标是探测一种理论预期中 的罕见物理过程,叫做无中微子双β。即使 PandaX-4T 实验的主要物理目标是探测 暗物质,但是由于探测器能探测的信号范围较广,因此可以对更多的理论预测的 新物理事例展开信号搜索。接下来,将详细介绍 PandaX (PandaX-II 和 PandaX-4T) 实验项目。

PandaX = Particle and Astrophysical Xenon Experiments

💷 PANDA X



Phase I: 120kg→ DM 2009-2014



Phase II: 500kg→ DM 2014-2019



Phase III: 1 ton → 0vDBD Future



Phase 4: 3.7 ton → DM 2019-today

图 2-1 PandaX 实验项目运行时间线。

2.1 中国锦屏地下实验室

一般来说,大多数稀有事例寻找实验地点选为深地下。最主要的原因是为了 避免宇宙射线造成的各种本底噪声。这些本底会干扰信号寻找过程。PandaX 也不 例外,实验地点(CJPL)位于中国四川省凉山彝族自治州锦屏山下的深地科学实 验室。该实验室 2009 年由雅砻江水电公司跟清华大学合作建立。CJPL 是首个用 于开展暗物质探测等国际前沿基础研究课题的极深地下实验室,也是世界上最深 的地下实验室,其宇宙射线通量降至地面水平的千万分之一至亿分之一,为相关 精密实验提供洁净的地下实验环境[115]。CJPL 在锦屏二级水电站修建的长达 18 公里的可通驶汽车的隧道中建造, 该隧道上方为厚达 2500 米的山体岩石, 可隔 绝穿透力极强的宇宙射线,环境适宜进行极深地下暗物质实验。另外,与其他部 分由矿井改建而成的地下实验室相比,实验室的交通更为方便,研究人员与科研 设备可乘车进入,不必靠电梯升降。2014年8月1日,清华大学与雅砻江流域水 电开发有限公司签署共同建设锦屏地下实验室二期的合作协议, 二期建设完成后, 实验室总容积将从原来的 4000 立方米扩容至 12 万立方米。2019 年 7 月 20 日, 位于实验室二期的"极深地下极低辐射本底前沿物理实验设施"启动建设。目前 (China Dark Matter Experiment, CDEX) [116], PandaX 以及 (Jinping Underground Nuclear Astrophysics, JUNA) [117] 等实验已入住 CJPL-II, 未来会有更多的实验 会入住 CJPL-II。CJPL 目前已成为全球深地科学研究的重要基地之一。实验室的 建设不仅提高了中国在暗物质探测和深地科学方面的研究能力,也为国际上的相 关研究提供了重要支持。



图 2-2 中国锦屏地下实验室(CJPL)地理位置 [118]。



图 2-3 CJPL-II 的布局图 [119]。

2.2 PandaX 探测器

在上一章内容提到 XENON, LZ 以及 PandaX 都采用气-液氙二相型 TPC 作为 探测器的核心技术来探测信号。下面对具体探测原理展开详细讨论。

2.2.1 探测器靶物质-氙

通常情况下,选择探测器的靶材料时,应尽选择比较稳定、密度高且不容易 发生化学反应的物质。众所周知,元素周期表中最稳定的元素是稀有气体,例如 氙、氩和氖等。表2-1中简要总结了最主流的三种稀有气体的物理性质 [120-121]。 从表中可以看出,无论是哪种惰性气体,它们的各个能级都被电子充满,因此 不容易发生化学反应。其中,氙的原子序数最大,密度也最高,介电系数也 高,这些特点能够提高暗物质与氙发生碰撞的概率。如图2-4所示 [122],在 30 keV 以下的反冲能区,氙与 WIMP 发生弹性散射的概率或事例率比其他元素 高。密度高的另一个好处是,相对于其他惰性气体,使氙具有良好的自屏蔽功 能。自屏蔽功能可以有效隔离液态氙核心区域与探测器其他材料带来的各种放 射性本底。除此之外,氙的沸点较高,大约为 165K,是三种惰性气体中最高 的一个。这种特点可以降低制冷的难度和成本。氙主要有六种同位素,分别为: ¹²⁶Xe(0.09%),¹²⁹Xe(26.44%),¹³¹Xe(21.18%),¹³²Xe(26.89%),¹³⁴Xe(10.44%),¹³⁶Xe(8.87%) [123]。除了 ¹³⁶Xe 之外,没有放射性同位素。还可以 ¹³⁶Xe 来研究无中微子双 β 衰

变 [124]。相比于其他类型的闪烁体,例如在第一章提到的 NaI(TI),氙的闪烁效 率也相当可观。在气-液氙二相型 TPC 中,能够很好地区分电子反冲(Electronic Recoil)和核反冲(Nuclear Recoil)。前一种是指入射粒子(例如 β , γ , Axion 或 ALP,暗光子等)与氙原子的核外电子发生反应。后一种是指入射粒子(例 如中子,中微子,WIMP等)与氙原子核发生反应。为了更好地理解电子反冲和 核反冲,这里首先定义激发-电离数比例 $\alpha = N_{ex}/N_i$,其中 N_{ex} 是激发的原子数 目, N_i 是氙离子-电子对数目。对于不同的入射粒子, α 的值也不同。一般来说, $\alpha_{NR} > \alpha_{ER}$ 。这里主要的原因是电子反冲粒子的入射径迹相对来说比较狭窄又长,这导致氙离子-电子对的重结合概率较低,反而核反冲粒子的入射径迹相对来说比 较宽又短,这导致氙离子-电子对的重结合概率较高。换句话说,对于同样的能量 沉积,电子反冲粒子的 S1 信号相比于核反冲粒子的 S1 较小。图2–5描述了氙对 不同类型粒子的阻止本领能谱,从中能够得出以上结论。

物理性质[单位]	Xe	Ar	Ne
原子序数 Z	54	18	10
相对原子质量	131.3	40.0	20.2
在标准大气压中的沸点 [K]	165.0	87.3	27.1
在标准大气压中的熔点 [K]	161.4	83.8	24.6
在标准大气压中的气态密度 $T_b \begin{bmatrix} s \\ l \end{bmatrix}$	9.99	5.77	9.56
在标准大气压中的液态密度 T _b [<u>s</u>]	2.94	1.40	1.21
液态介电常数	1.95	1.51	1.53
地球大气中的占比 [ppm]	0.09	9340	18.2
闪烁光波长 [nm]	178	129	80
第一电离能 [eV]	12.1	15.8	21.6
对 α 的闪烁能 [eV]	16.3, 17.9	27.1	_

表 2-1 惰性气体 Xe, Ar 以及 Ne 的基本物理性质。

2.2.2 气-液氙二相型 TPC

如图2-6所示,圆柱形探测器内部的大部分被液氙充满,而顶部则被气氙充满。 此外,探测器的顶部和底部都有光电倍增管(Photo Multiplier Tube, PMT)阵列用 于探测光信号。当外部粒子进入探测器并与液氙原子发生碰撞时,入射粒子的动 能被沉积在液氙中。这种能量以三种形式表现出来:电、光和热。由于该 TPC 无 法探测热信号,因此只能探测到光和电信号,这两种信号通道。

当入射粒子与氙原子碰撞时, 氙原子处于激发态。处于激发态的氙原子与周



图 2-4 Xe、Ge、Ar、I和 Na 等元素跟 WIMP 的弹性散射反冲能事例率谱,这里假定 WIMP 质量为 100 GeV/c², Xe 跟 WIMP 的散射截面为 10⁻⁴⁵ cm²[122]。

围的氙原子形成激发态分子(Excimer),这种激发态分子会立即退激发并发出光信号,即 S1 信号。S1 信号实质上是波长为 178nm 的紫外线闪烁光,这种闪烁光由单重态(Singlet)和三重态(Triplet)组成。此外,入射粒子还能够产生氙离子-电子对。在一般条件下,大部分这些氙离子-电子对会重新结合形成处于激发态的氙原子,随后与周围的氙原子形成激发态分子,最终退激发并发出光信号,即 S1 信号。因此,S1 信号不仅来自于初始的激发态氙原子,还来自于氙离子-电子对的重新结合过程,而且 S1 信号的大部分来自于重新结合过程,如图2–7所示。

TPC 内部有向下的漂移电场。入射粒子碰撞产生的氙离子-电子对中的电子在 漂移电场的作用下向上漂移,如图2-6所示。这里的电场只导致部分氙离子-电子 对参与重组合。向上漂移的电子到达气液氙界面时,会被比漂移电场更强的电场 拽到气氙中,这被称为萃取电场。在气氙中,气氙原子受到激发后再退激发产生 光信号。这种在气氙中产生的光信号大小与电子数成正比,因此称为正比光信号, 简称 S2 信号。无论是 S1 信号还是 S2 信号,都可以被安装在 TPC 项部和底部的 PMT 阵列所探测到。此外,S1 信号和 S2 信号的时间差被称为电子漂移时间。换 句话说,这种时间差可以反映出事例发生的纵向位置。通过 S2 信号在上方 PMT 阵列的接收密度,可以确定事例的水平方向位置。利用这种 TPC 技术,我们可以



图 2-5 氙对不同类型粒子的阻止本领。其蓝色和红色分别代表电子反冲和 α 事例,绿色和紫 色虚线为核反冲事例 [125]。

获取事例的三维位置信息。这也是气-液氙二相型 TPC 技术的最重要的优势之一。 如图2--9所示, PandaX-4T 探测器的 TPC 是正二十四边形的类圆柱体,高度 和直径约为 1.2m 左右,这相当于 PandaX-II 探测器 TPC (高度和直径约为 60cm) 的两倍左右。大体积能容纳更多的氙。PandaX-4T 能容纳的液氙约为 3.7 吨,是

PandaX-II(500kg)的7倍多一点儿。 类圆柱体底部和顶部均有直径约为1.3m的圆形铜板。PMT 阵列均安装在铜板上。两块铜板被24块相连的白色聚四氟乙烯((Polytetrafluoroethylene, PTFE) 板来支撑。这些 PTFE 有反光作用,从而提高 S1 和 S2 信号被 PMT 采集的效率。 从底部到顶部,在 PTFE 板外面有 58 圈铜来制作的电场整形环。它的作用是使得 电子在电场中以匀速往上漂移。

在 PandaX-4T 探测器的 TPC 中,有四个电极:阳极(Anode)、阴极(Cathode)(15-20kV 的负高压)、门电极(Gate)(5-6kV 的负高压)以及底部保护电极(Bottom Screening)。其中,阴极到门电极间距离约为1185mm 左右,能形成向上方向的电场,使得电子漂移。门电极到阳极间的距离约为10mm 左右,气液氙的分界面位于它们的正中间。此电场的主要作用是从下面漂移过来的电子拽出到气 氙中以及让电子跟气氙原子发生碰撞而产生 S2 信号。

阳极和保护电极离 PMT 阵列比较靠近,它们均接地,可以避免高压电场击穿 PMT。另外,保护电极跟阴极间的距离约为 100mm 左右。由于阴极的电压比较高,为了避免击穿 PMT,距离相比阳极到顶部 PMT 的距离(46mm)大两倍左右。由

于阴极跟保护电极间的电场方向向上,因此在此区域发生的入射粒子产生的电子 不会拽出到气氙,即不会产生 S2 信号。

在 PandaX-II 和 PandaX-4T 实验中使用的 PMT 均由日本滨松光电子公司生产 [126]。总共有 473 支 PMT,其中 368 支是 3 英寸型号为 R11410-23 的 PMT2-10, 安装在 TPC 项部和底部的中心区域。105 支是 1 英寸型号为 R8520-406 的 PMT, 安装在 TPC 项部和底部的边缘区域。1 英寸的 PMT 在本实验中的主要作用是排 除掉在探测器边缘发生的事例或者多次散射的事例(通常 WIMP 信号是单次散射 事例)。这种 PMT 的使用沿用自 PandaX-II 实验。它对前文提到的波长为 178nm 的紫外闪烁光有较高的量子效率(约为 38% 左右)以及探测效率[126]。



图 2-6 气-液氙二相型 TPC 原理示意图 [127]。

2.2.3 电子学及数据获取系统

在前文中已经提到,无论是 S1 信号还是 S2 信号,都可以被探测器顶部和底部的 PMT 阵列探测到。PMT 是一种电子器件,利用光电效应将光信号转化为电信号。它通常由光电阴极、几层打拿极和阳极组成,被陶瓷或玻璃包围成真空管状。打拿极之间有一定的电势差,入射光撞击光阴极产生电子后,电子被加速到打拿极并产生倍增的二次电子。经过一系列倍增,电子最终到达阳极。PMT 具有高增益、低干扰和高灵敏度对高频信号等优点,因此在高能物理、医学成像和天文等领域得到广泛应用。



图 2-7 入射粒子在液氙中能量沉积的物理过程示意图。



图 2-8 通过 PandaX-II 实验数据重建出来的典型的 S1 和 S2 信号的波形。

PandaX-II 和 PandaX-4T 实验所使用的 PMT 均来自于日本滨松光电子公司,其中 PMT 型号为 R11410-23 (3 英寸)和 R8520-406 (1 英寸)。它们具有耐低温、耐高压和低本底等优点,非常符合 PandaX-4T 实验的需求。通过 PandaX-II 的使用经验,我们发现,虽然过高的电压可以提高电子的放大倍数,但也会增加 PMT 打火的风险。因此,我们需要调整电压以获得合适的倍增倍数,最终调整得到的电压约为 750V 左右。然而,随着探测器的运行,各个 PMT 的状态不一定能保持稳定。有些 PMT 仍可能出现打火,有些 PMT 的后脉冲信号可能会增加。为了获得高质



图 2-9 PandaX-4T 探测器 TPC 示意图。



图 2-10 在 PandaX 实验中使用的一英寸 PMT (左)和三英寸 PMT (右)。

量的数据,我们需要关闭一些异常的 PMT。这会对事例的位置重建和能量重建等 产生一定程度的影响,但我们仍然可以通过各种数据分析方法,一定程度上修正 或补充丢失的数据。

PMT 将收集到的光信号通过专为 PandaX 实验设计的分压器进行读出。 R11410-23 型的 PMT 具有 12 个打拿极,其对应的分压器采用阴极接负高压、阳极 接正高压以及第五个打拿极接地的设计。为了减少高能事例对 PMT 造成的饱和, 探测器中心部分的 14 支 PMT (7 支在顶部,7 支在底部)采用了双读出分压器。 这意味着信号不需要经过所有的打拿极到达阳极,只需到达第八个打拿极就能够 读出信号。这种设计对于高能事例(MeV 级别的反冲能事例)提供了更好的硬件 支持。然而,在真实的取数过程发现,双读出设计,对于高能事例的表现不如预 期,这种设计可能下一轮取数阶段有望重新优化或设计。

被 PMT 探测到的 S1 和 S2 光信号,经过分压器读出后,通过电子学和数据获取系统将光信号转化为数字信号。如图2-11所示,电子学与数据获取系统由三个部分组成,分别是前端电子学、数据获取与采集模块以及数据处理与传输模块。前端电子学包括 PMT 分压器、信号解耦和放大模块。数据获取与采集模块包括 V1725B 波形采集卡、V2495 控制模块、时钟分发板以及 VME 机箱。数据处理与传输模块由服务器、交换机和磁盘阵列组成。

由于 PandaX-4T 探测器比其前任探测器(PandaX-II)体积更大且 PMT 数量 更多,因此电子学和数据获取系统的要求更高。该实验采用 CEAN 的 V1725B 数 字化波形采集卡,其采样率为 250MS/s。相比于以往的 PandaX-II 实验(采样率 为 100MS/s),采样率得到了大幅提升。此次实验采取"自触发"模式来采集数 据,只要过阈值的信号都会被记录下来,几乎没有信号损失。以往的 PandaX-II 或 PandaX-I 实验中采用的是"全局外触发"模式,这种触发模式依赖于 S2 信号以及 只记录 S2 信号前后的固定窗口,因此有可能出现信号损失。PandaX-4T 实验共有 473 支 PMT,即 473 道信号通道。为确保这么多 PMT 和高采样率波形采集卡的数 据高效快速传输,该实验采用并行传输方案,总带宽可达 500MB/s。关于更详细 的电子学和数据采集系统内容,请参考 [128]。

2.3 基础设施

为了保证 PandaX 实验整个探测器正常运行,除了 TPC 之外,还需要诸多的基础设施。在 PandaX-4T 实验中,这些基础设施包括超纯水屏蔽系统、制冷和循环系统、刻度系统、氙存储系统、精馏塔、PMT、电子学以及数据获取系统等,如图2–12所示。由于 PandaX-II 实验规模相对较小,因此其基础设施相对简单。下面将以 PandaX-4T 实验为主,对这些系统进行简要介绍。

2.3.1 屏蔽体

前文中提到, 锦屏地下实验室虽然能够大幅度降低宇宙射线本底, 但在实验 室环境内仍存在一定含量的各种放射性元素导致的本底。为了进一步降低来自实 验室环境的本底, PandaX-II 实验采用了屏蔽体,包括主动的反符合系统及被动的 屏蔽体。其中,主动的反符合系统是一英寸 PMT, 安装在 TPC 的顶部和底部边缘, 通过它以及分析采集到的数据,可以标记并排除掉部分来自环境的放射性元素导





图 2-11 PandaX-4T 电子学和数据获取系统示意图。

致的本底。例如,如果一个事例同时被一英寸和三英寸 PMT 探测到,那么这种事例大概率是多次散射,通常情况下,暗物质信号不能够发生多次散射,因此这种事例被主动屏蔽体排除掉。

被动屏蔽体由多层高密度材料组成,如图2-13所示。外层是 40 厘米厚的 PTFE, 其内是 20 厘米厚的铅层。在铅层内部,是 20 厘米的 PTFE 层,其内是 5 厘米厚 的高纯无氧铜和更靠内的 5 厘米厚的铜罐。PTFE 和铅均是使用相应厚度的块堆 砌而成,而在被动屏蔽体中使用的材料均为低本底材料。实验室中存在着长寿命 的 ²³²Th 和 ²³⁸U 的衰变产物 ²²²Rn 和 ²²⁰Rn (常温下均为气态)。这些元素的衰变链 会产生多种射线,如 α、β 和 γ 射线,同时也会产生中子。中子本底信号与暗物 质信号的性质基本相似,不带电的中子穿透性更强,因此降低中子含量尤为重要。 为此,在 PandaX-II 实验中,采用了引入外界新鲜山风并保持空气循环流的方法来 降低 ²²²Rn 的含量。此外,在实验运行期间,通过往铜屏蔽体与铜罐之间的空隙持 续灌注高纯度(99.99%)氮气,使得空隙内充满高纯氮气。实验室中的 ²²²Rn 含 量通过商用 Rn 探测器 RAD7 来测量 [129],约为 100-200 Bq/m³[130]。需要注意 的是,PandaX-4T 实验没有使用被动屏蔽体。为了更高效地屏蔽实验室环境本底, 上海交通大学博士学位论文



图 2-12 PandaX-4T 实验的各个配套系统的展示图。

PandaX-4T 实验采用了超纯水屏蔽系统。关于 PandaX-II 实验屏蔽体更多细节可参见文献 [130-131]。

2.3.2 冷却和循环系统

为了降低本底噪声(宇宙射线),不仅要将实验室地点选在深地下,还需为探测器创造低温环境,此外,为了得到液态氙,必须进行降温。为了保证氙的纯度, 在运行期间,PandaX 实验独立研发了一套循环系统,该系统由探测器容器、冷却 总线以及循环提纯系统组成。

探测器容器由内罐(Inner vessel)和外罐(Outer vessel)组成。其中,内罐指的是低温容器,高度约为 2.4m,内直径约为 1.4m,TPC 直接安装在内罐中。而外罐指的是常温容器,高度约为 3.1m,内直径约为 1.7m,其内罐放置在外罐中,如图2–14所示。为避免漏热,内罐和外罐之间抽成了水平真空达到 10⁻³Pa[132]。另外,由于内罐和外罐距离探测器 TPC 核心区域较近,因此内罐和外罐必须采用低放射性材料。

氙是一种惰性气体,在常温下以气态存在于大气中。如前所述, 氙是作为探测器靶材料使用的, 其中大部分以液态形式存在。为了获得液态氙, 需要使用冷却设备。PandaX-4T 实验使用的冷却总线装置包括 PC-150、PC-90 和 RDK-500B 等三种模块 [132], 以及液氮和真空泵。在将氙注入 TPC 之前, 需要首先将内罐抽真空。然后, 使用 PC-150、PC-90 和 RDK-500B 等三种模块对氙气进行制冷,将其液化。液态氙通过管道从上到下流入 TPC, 受重力的作用。由于这三种冷却模块



图 2-13 在 PandaX-II 实验中使用的被动屏蔽体系统。



图 2-14 PandaX-4T 探测器的内罐(左)和外罐(右)实拍图。

无法完全满足冷却要求,因此需要使用液氮来辅助冷却。

尽管 PandaX-4T 拿到的氙纯度比较高,但这种纯度无法完全满足实验要求。从 工业中得到的氙里面可能含有一些氧气、氮气、氪、氡以及水等杂质。此外,当 液化后的氙流入 TPC 的过程中也难以避免杂质进入。另外,可能部分杂志吸附在 探测器材料中,比如,PTFE 等。当入射粒子在液氙中释放能量时,产生的电子可 能被这些杂质吸附,从而导致 S2 信号转换率或探测效率降低。因此,清除这些杂

上海交通大学博士学位论文

质是必要的。为此,PandaX-4T 实验开发了一套循环提纯系统,如图2-15所示。整 个系统由两条独立的循环管道组成,其中一个停止工作时,另一个能够正常运行。 每个循环管道由控制器、纯化器和循环泵(KNF泵)组成。KNF泵的作用是为整 个循环管道提供动力。纯化器的作用是净化氙。其中,纯化器的真空腔体内的高 温金属能够与杂质如氧气、水等发生反应,从而达到净化氙的目的。氙的净化流 速是另一个关键参数。如果流速过快,会给 KNF 泵施加压力,使其受损。如果流 速过慢,会增加氙中的杂质。因此,必须找到合适的流速来保证探测器的稳定且 低本底运行。PandaX-4T 实验中,氙的平均流速约为 70 kg/h。关于冷却以及循环 系统参见[132]。



图 2-15 PandaX-4T 冷却总线实拍图(左)和循环提纯系统(右)示意图。

2.3.3 精馏塔

在前文中提到,工业氙中含有氪和氡两种元素,这两种元素无法通过循环系 统有效彻底去除。它们的本底会在暗物质信号区间造成不可避免的影响。因此,有 效降低这两种本底有助于提高探测器的探测灵敏度。氪指的是⁸⁵Kr,半衰期约为 10.76年,均匀分布在探测器中。氡指的是²²²Rn,半衰期约为3.8天,是由²³⁸U衰 变产生的。氙、氪和氡这三种元素具有不同的物理性质,如沸点和熔点等。因此, 可以通过精馏和吸附等方法从氙中分离出氪和氡。在标准大气压下,氪、氡和氙 的沸点分别为120K、211K和165K。据此可知,氪最容易挥发,会在气相中聚集, 而氡则不容易挥发,会分布在液态氙中。利用它们的物理特性可以有效降低氪和 氡的含量。为了去除氪和氡,PandaX-4T实验独立开发了精馏塔系统。该系统氙 气收集效率可达99%,流速可达10kg/h,预计可以将从工业氙中的氪含量降低到 7个数量级,并将氡的含量降低到1.8倍。关于精馏塔系统的更多详细内容,请参 见[133]。

2.3.4 超纯水屏蔽系统

即使 PandaX 实验地点位于 2400m 深的地下,实验室内及周围的岩石和材料 仍然具有放射性,例如自然界中存在的长寿命放射性元素 ²³⁸U 和 ²³²Th 等。这些 元素的衰变链中包含 α、β 和 γ 等衰变,其中 α 粒子也能与其他材料一起产生中 子。此外,²³⁸U 也能通过自发裂变产生中子。如前所述,中子能像暗物质一样产 生核反冲事例,而探测器难以区分这两种事件。因此,排除这些本底对于暗物质 探测非常重要。

为了更有效地屏蔽这些本底, PandaX-4T 实验研发了一套超纯水屏蔽系统。该 系统能够去除自来水中的各种杂质和电解质,将自来水转化为超纯净水。平均每 小时能够产生约 10 吨超纯净水。该系统通过管道连接到高度约 13m、直径约 10m 的圆柱形大水罐,该水罐被称为屏蔽水罐(Water Shield Tank)(如图2–16(左)所 示)。经过净化后,超纯净水通过该管道流入屏蔽水罐。探测器 TPC 被放置于屏蔽 水罐内,即整个探测器完全浸泡在屏蔽水罐中(如图2–16(右)所示)。如图2–12所 示,水罐顶部为实验室的洁净区,因此可以方便地利用其他探测手段(例如高纯 锗探测器、ICP-MS 等)来测量水中各种放射性水平。总之,超纯水系统在抑制本 底方面具有关键作用。



图 2-16 PandaX-4T 超纯水罐外观(左边)和内观(右边)实拍图。

2.3.5 刻度系统

为了更好地理解探测器的性能以及对不同类型事例的响应,需要使用各种放射性物质来检测探测器对它们的响应。一般将这个过程称作探测器刻度。PandaX-4T 实验为此研发了一套刻度系统。探测器刻度是物理实验中不可或缺的一部分, 几乎所有的粒子物理实验室都会对探测器进行刻度,以便了解探测器的性能和对 不同类型事例的响应。PandaX-4T 探测器刻度系统分为内部放射源刻度和外部放射源刻度。

内部放射源指的是放射源能够直接注入到探测器内部的液态氙中。这种放射 源可以通过循环系统的管道注入到探测器中的液态氙中,比如^{83m}Kr 和²²⁰Rn 等寿 命较短的放射性元素。这种放射源的优势是能够让放射源均匀分布在整个探测器 内部。图2–17是内部放射源刻度控制面板,其中包括^{83m}Kr 腔体和²²⁰Rn 腔体,这 些腔体是专门为存放^{83m}Kr 和²²⁰Rn 而设计的。此外,该控制面板还连接到循环系 统的 loop2 管道。



图 2-17 PandaX-4T 实验的内部放射源刻度注入控制面板。

外部放射源是指将放射源放置在探测器外罐内部。这种放射源本身不会直接进入液氙。在 PandaX-4T 实验中,常用的刻度放射源包括⁶⁰Co、²³²Th、¹³⁷Cs、²⁴¹Am-⁹Be 以及 Deuterium-Deuterium(DD)等。除了 DD 源之外,其余的放射源在以往 的 PandaX-II 实验中使用过,都是固体胶囊大小的。虽然这些放射源都是外部放 射源,但它们产生的次级粒子会在探测器内部产生各种类型的事例。除了 DD 源 之外,其他放射源通过超纯水罐顶部的外部放射源端口送到探测器外罐内表面的 放射源细管。如图2–18(左)所示。其中,DD 源与其他外部刻度源有所不同。它 是通过长度约为 50cm 的圆柱形中子发生器产生中子,然后进入探测器内部,如 图2–18(右)所示。



图 2-18 PandaX-4T 实验的外部放射源刻度管(左)和 DD 中子生成器(右)。

2.3.6 慢控制系统

PandaX 实验的另一个重要部分是慢控制系统。为了保证探测器本身及外界环 境的稳定性,通常需要通过某种方式来监控它。为此, PandaX 实验开发了一套慢 速控制系统(Slow Control System, SCS),它能够在检测到异常状态时发出警报, 从而确保 PandaX 实验运行期间的稳定性。慢控制系统的核心部件是各类传感器, 包括温度传感器、液位传感器、高电压传感器以及液体流速传感器等,这些传感器 安装在探测器的各个部位。传感器收集到的实时监控信息通过网页可视化,一旦 出现异常,所有实验室成员会通过传感器收集到的数据及各类通讯设备(邮件或 微信)收到警报信息。慢控制系统的软件部分采用传统的客户端/服务端模式,如 图2-19所示。该系统由三个独立的模块组成,分别为数据采集模块、数据存储模块 以及数据可视化或异常警报模块。它们之间通讯采用标准的 HTTP 协议。数据采 集模块的主要功能是从不同的传感器获取信号,将其转换为数字数据,并将结果 数据发送到数据存储模块。该模块还可以将结果数据发送到其他外部目的地,例 如标准输出或文本文件,用于测试目的。传感器信号可分为两类:模拟信号和数 字信号。DAM-8082 的模拟输入模块具有 24 位 ADC (Analog to Digital Converter), 用于将模拟信号转换为数字化数据。所有数字化数据通过串行接口传输到数据采 集计算机,这里使用的是带有 RS-232 或 RS-485 协议的 USB 转串行适配器。

由于传感器分布在实验厅的不同部位,而且 RS-232 协议的电缆长度限制在 20 m 以下,因此需要部署多台数据采集计算机。在这种情况下,选择 Raspberry Pi 3 Model B (Pi) 作为默认数据采集计算机是合适的。该设备配备了四核 64 位 ARM Cortex A53 CPU (BCM2837), 1 GB RAM,一个以太网端口,机载无线 LAN 和四个 USB 2.0 端口。另外, Raspberry Pi 的丙烯酸包装盒的尺寸仅为 85 × 65 ×

35 mm。使用 Raspberry Pi 可以使 SCS 更经济实惠,其成本(包括内存卡、包装 盒和电源适配器)约为 350 元人民币。此外,操作系统采用 Raspbian,该操作系 统是 Debian GNU/Linux [10]的 ARM 版本,提供了适当的环境来部署数据采集程序。数据采集程序使用 Python 语言来编写,因此可以在 Raspberry Pi 或普通的基于 x86/x64 的计算机上运行,无需进行任何修改。

数据存储模块是一个数据库,用于存储传感器的时间序列数据,采用的是 InfluxDB 开源版本 1.7.3。InfluxDB 是一种专门设计用于存储时间序列数据的数据 库。它引入了"测量"这一概念,类似于传统 SQL 数据库中的"表"。每个测量是 一个包含字段、标签和时间的容器,表示一个时间序列。字段是键值对,其中键 是字符串,值是实际数据。一组字段的集合构成了一个序列,称为一个点。标签 用于记录测量的元数据,也是由键值对组成的。标签被索引,因此标签查询速度 很快。例如,在 PandaX-4T 中测量的温度值位于一个包含"温度"字段的测量中, 并具有标签以表示每个点的位置和设备。当新类型的数据到达时,可以随时创建 新的测量。

PandaX-4T SCS 使用开源版本 1.7.7 的 Chronograf 来实现数据可视化模块,该 模块与 InfluxDB 实例安装在同一服务器中并相连。用户可以通过现代桌面或移动 Web 浏览器直接访问 Chronograf,并通过其数据浏览器界面直接使用 InfluxQL 查 询来对任何测量进行调查。自定义仪表板是访问固定测量集的首选方法,整个系 统的概述页面被实现为自定义仪表板。用户可以通过向仪表板添加大小自定义的 单元格来显示一组测量值,并通过在图形编辑器上配置可视化样式和时间范围来 形成高度表达性的用户界面。默认情况下,仪表板将显示一小时的历史数据,而 查询一天、30 天和一年的历史数据可在 1 秒、不到 5 秒的时间内完成, Chronograf 提供了高效的数据显示界面。

在异常警报模块中,我们使用的是 1.5.2 版本的 Kapacitor,它是一种数据处理 引擎。该软件也安装在 InfluxDB 的同一服务器上,可以通过 Chronograf 接口直接 管理。一旦 Kapacitor 实例连接到 Chronograf,就可以在用户界面的"警报"部分 构建警报规则。每个规则只能选择一个时间序列,并提供了三种警报类型(阈值、 相对、死亡)以便构建不同的数据异常情况规则,并在触发时由选定的数据处理 程序进行处理。在 PandaX-4T 中,我们使用了电子邮件和 post 处理程序来处理警 报,并将警报消息推送到自定义短消息服务(SMS)网关,该网关会向现场工作 的人员发送短信。使用 Chronograf 和 Kapacitor 几乎满足了 PandaX-4T 慢控制系统 的所有要求,系统部署简单,所有配置都可以通过浏览器的用户界面来完成,只 需编写自定义 SMS 网关即可发送短信。关于慢控制系统详细讨论参见 [134-135]。



图 2-19 PandaX-4T 实验的慢控制系统软件部分的结构示意图,一共有三个独立的模块,它 们是通过标准的 HTTP 协议来通讯 [135]。

第三章 数据处理与重建

在2.2.3简要介绍了电子学和数据获取系统的运作方式。本章将介绍数字化后的数据处理过程,包括数据重建和筛选。尽管 PandaX-II 和 PandaX-4T 在数据处理和重建方面大体相似,但本章将主要通过 PandaX-4T 实验来介绍整个 PandaX 实验的数据处理和重建逻辑。

一旦 S1 和 S2 信号通过 PMT 变成模拟信号,就需要经过进一步的处理才能 进行有效的数字化采集。这些处理步骤包括使用解耦器和放大器对模拟信号进行 传输。解耦器将分离不同通道的信号,避免传输过程中的干扰,而放大器对模拟 信号进行放大,以便进行更准确的数字化采集。

接下来,模拟信号通过波形采集卡(V1725B)进行数字化处理。波形采集卡 采用快速数模转换器技术,可将模拟信号高速、高精度地数字化处理。数字化后 的数据通过光纤连接到四台戴尔服务器上。每台服务器都连接着8块波形采集卡, 其各自负责若干 PMT 的数字化信号的采集。

四台服务器上的数据通过 10 Gbps 的光纤交换机传输到另一台服务器。该服 务器将四台服务器上的数据按时间戳排序,并将其打包处理后,保存到磁盘上。这 些数据将在后续的物理分析过程中用于各种不同的分析。

3.1 消息队列

相较于 PandaX-4T 实验数据, PandaX-II 实验数据规模较小。在 PandaX-II 实验中,通常,将最新采集到的数据在实验室会及时进行处理和粗略分析。然而,由于实验室基础设施条件的限制,无法建立数据中心进行大规模处理以及数据分析。因此,采集到的数据会备份到硬盘中,然后通过快递寄送到上海交通大学。上海交通大学拥有足够的存储和计算资源,因此后续的数据分析会在该校的计算集群上进行。此外,为了安全起见,原始数据也会备份到磁盘中。

与 PandaX-II 实验的数据不同, PandaX-4T 实验数据规模更大,采样率更高,因此需要及时保存和处理数据。为了解决数据存储问题,我们在距实验室场地较近的成都市建立了一个独立的数据中心。实验室采集的数据可以通过移动的高速专线网络直接传送到成都数据中心。分析人员可以通过远程登录到成都数据中心进行数据处理和分析。为了自动化整个过程,我们采用了所谓的消息队列管道服务。

正如上文所述,戴尔的四台服务器将数据通过 10Gbps 的光纤交换机传输到

另一台服务器。如图3-1所示,这台服务器通过 Event Builder 程序处理数据,然后 使用 Blocker Sorter 按照时间戳排序。接着,Group Converter 将数据转化为 Group Data,将数据分组以方便后续处理。经过转化后,近期的数据会被保存在服务器 上,而早期数据则被删除。此外,新数据通过 1Gbps 的专属网络直接从锦屏传输 到成都。Group Data 经过一系列处理后,会生成数据质量图并在网页上显示。这 一系列操作都是在一个名为 Kafka 的消息队列系统的帮助下完成的 [136]。

Kafka 是一种高吞吐量的分布式流平台,用于处理实时数据流。由 LinkedIn 公司开发,现已成为 Apache 软件基金会的顶级项目。Kafka 的主要功能是提供可 扩展、高吞吐量和低延迟的数据流平台,可用于记录日志数据、实时数据处理和 流数据聚合。Kafka 采用发布-订阅(Pub-Sub)模型来处理数据流,其中有生产者 (Producer)和消费者 (Consumer)两种角色。生产者负责将数据写入 Kafka,消费 者则负责从 Kafka 中读取数据。这样,生产者和消费者可以独立运行,互不干扰。Kafka 中的数据以 Topic 的形式存储,Topic 是一种逻辑分类,用于将不同类型的 数据隔离开来。每个 Topic 由一组 Partition 组成,Partition 是物理存储单元,负责存储数据并维护数据的顺序。Kafka 集群由若干个 Broker 组成,每个 Broker 是一 个独立的进程,负责存储数据和处理请求。所有 Broker 之间是对等的,它们之间 通过网络进行通信。每个 Broker 都维护一份所有 Topic 的元数据,这样就能知道 数据在哪里和如何复制。当 Producer 向一个 Topic 发布数据时,数据会被写入该 Topic 的所有 Partition 中读取。

在 PandaX-4T 中,使用的 Kafka 消息队列如图3–1所示,共有四个核心的 Topic, 这些 Topic 都在同一个 broker 中。如图3–1所展示, Event Builder 作为一个生产者, 首先将数据处理完的消息发送到第一个 Topic, 然后 Block Sorter 从第一个 Topic 收到消息并将其消费掉,然后开始运行。同样的, Block Sorter 运行完后作为一个 生产者,将其消息送到第二个 Topic,以此类推,程序和程序之间通过这四个 Topic 来进行信息交流,最终完成整套数据处理流程。

3.2 数据格式

PandaX-II 实验之前使用的是以由欧洲核子研究组织(Conseil Européen pour la Recherche Nucléaire, CERN)开发的大型数据分析软件包 ROOT[137] 作为底层独 立开发的名为 Bamboo-Shoot2 的库程序。由于 ROOT 版本更新为频繁,因此遇到 了数据格式和 ABI(Application Binary Interface)不兼容问题。ABI 描述了编译器 生成的二进制代码如何和系统进行交互,并且规定了编译器如何生成代码,以便





图 3-1 在 PandaX-4T 数据处理过程中的消息队列示意图。

与特定的硬件平台和操作系统兼容。例如,ABI 会定义函数调用约定、数据类型的 表示、寄存器的使用等等。ABI 不兼容会引发运行时错误,而这些错误本应在链 接时就被报告,数据格式不兼容则更为麻烦。为了消除这些问题,我们重新开发 了第三代 Bamboo Shoot3[138],并且没有依赖于 ROOT。这个库分为三部分:对象 序列化库(Object Serialization Library),命名为 pbss;文件 IO 库(file IO library), 命名为 pbsf;和支持应用(Support Utilities),命名为 pbsu。

对象序列化库采用开放式数据格式,它支持树形数据结构,并通过对每个字 段进行数字标记来提供自定义结构体字段的前向/后向兼容性。目前,该实现在序 列化和解析方面的速度可达数十亿字节每秒。文件 IO 库也具有稳定的数据格式, 提供简单的顺序访问和基本的随机访问支持。这两个部分都是围绕单个存储对象 可能很大(通常为100kB-10MB),顺序访问是主要访问模式的假设建立的。支持 应用程序库提供了一些遵循函数式编程风格的简单支持应用程序。例如,需要加 新的数据结构 CalibData,可以直接在自定义的伪代码文件里写对应的数据结构即 可:

```
tuple RawPmtSegment {
   uint64 startTime;
   uint32 channelNumber;
   [uint16] adcValue;
};
struct GroupData {
   uint32 runNumber = 1;
   uint32 groupNumber = 2;
   uint64 startTime = 3;
```

```
uint64 endTime = 4;
[RawPmtSegment] segments = 5;
};
tuple CalibPmtSegment {
  uint64 startTime;
  uint32 channelNumber;
  [float] peValue;
};
struct CalibData {
  uint32 runNumber = 1;
  uint32 groupNumber = 2;
  uint64 startTime = 3;
  uint64 endTime = 4;
  [CalibPmtSegment] segments = 5;
};
```

3.3 数据库

PandaX 数据的产生以及处理过程中需要将一些基本信息保存下来,这些信息 在后期的数据处理以及分析中会用到。PandaX-II 以及 PandaX-4T 中使用的都是 是 PostrgeSQL[139]。PostgreSQL 是一个开源关系型数据库管理系统,它具有高稳 定性、高可用性和高扩展性。该系统是一个功能强大的数据库,支持数据库事务、 关系数据模型、SOL 查询语言、表级并发控制、多版本并发控制和全文搜索等。 PostgreSQL 是基于对象关系模型设计的,支持存储对象、继承、表类型、函数类 型、触发器和规则等。该系统支持通过存储过程、触发器和函数等来实现规则和 触发器。PostgreSQL 支持多种数据类型,包括整型、浮点型、字符串型、日期型、 布尔型、数组型、枚举型、JSON、XML、HSTORE、GEOMETRY 和用户自定义数 据类型等。该系统还支持数据类型的转换和运算,并支持类似 SOL 的数据定义语 言。PostgreSQL 还提供了一些高级特性,如联机备份、分布式事务、分布式数据 库支持、分布式锁定、分布式状态管理和分布式查询优化等。PostgreSQL 中最基 本的概念之一是 Schema, 它是一种用于组织数据库对象的逻辑结构。例如, 可以 使用 Schema 来区分不同类型的数据或不同用户的数据。Schema 可以包含表,视 图,索引,序列等。其中表是存储数据的容器。表由列和行组成,列存储数据的 元数据,行存储数据的实际值。PostgreSQL中的表可以有很多类型,例如基本表, 视图,临时表,共享表等。除此之外序列是 PostgreSQL 中的一种特殊对象,用于 生成自动增量的整数。它可以用作主键或其他字段的默认值。

PandaX-II 和 PandaX-4T 实验中,从启动数据获取系统到结束时间段所采集到的数据被称为 run,每个 run 由若干个文件(file)组成。在 PandaX-4T (PandaX-II)
实验中,一般每个正常文件的时长为 90 秒(180 秒)左右,大小约为 1GB(1GB) 左右。当每次启动数据获取系统开始取数时,会自动将 run 以及文件的索引、开始 时间和结束时间保存至 runs 和 files 表格中。此外,还会将 PMT 的各种信息,包括 X-Y 坐标、PMT 编号、通道编号和 PMT 类型等,也都保存在数据中。数据库还可 以用来保存每个 run 随着时间变化的总带宽,包括每个通道的带宽等信息。每周 平均进行一次 PMT 增益刻度信息的记录也需要写入数据库。除了以上提到的内 容外,还有很多重要的数据信息都保存在数据库中。可以看出,数据库在整个数 据处理和分析过程中扮演着非常重要且不可忽略的角色。图3-2展示了 PandaX-4T 数据库表结构及关系图。PandaX 数据共有三个镜像,分别在锦屏、成都和上海。 其中,锦屏是主镜像,其余的都是副镜像。写操作只能在主镜像中进行,副镜像 主要用于读取信息。



图 3-2 PandaX-4T 数据库表结构及关系图。

3.4 数据处理流程

PandaX-II 实验和 PandaX-4T 实验的数据处理流程逻辑基本相似,因此本文只介绍 PandaX-4T 的数据处理流程。有关 PandaX-II 的详细数据处理流程,请参阅 [131]。

首先,数据经过 Group Converter 转换成 Group Data,然后使用 PandaX-4T 开 发的数据处理链项目 P4-Chain 将数据转换为包含事例物理信息的文件。Bamboo-Shoot3 作为底层库,在整个 P4-Chain 数据处理链中被调用。P4-Chain 数据处理 链的第一步是使用 PMT 增益信息将 Group Data 中的模拟信号单位 ADC 值转换 为光电子(Photo Electron, PE)值,并确定基线(Baseline)。然后,使用 ZLE (Zero Length Encoding)的方式对数据进行压缩。通常情况下,基线是根据波形前 40 个样本(Sample,波形时间宽度的最小单位,在 PandaX-II 中一个样本为 10ns, 在 PandaX-4T 中 4ns)的平均值确定的。如果一个 Sample 的值减去基线值后小于 0.25PE,那么该 Sample 的值将被设为 0,这就是所谓的软件 ZLE。经过重新处理 数据,基于基线和 ZLE 模式得到所谓的 Calib Data。它是一种过渡数据,一般不 会被保存下来。



图 3-3 增益刻度前后波形对比(上图为刻度前,下图为刻度后的波形)。

接下来,数据经过寻找脉冲信号(Hit Finder)算法来确定脉冲(Hit)信号,并 计算 Hit 的各种信息,例如 Hit 的面积(电荷)、高度、宽度和类型。为了进行更底 层的数据分析,数据也可以保存为 Hit Data。一旦确定了 Hit 信号,所有 PMT 的 Hit 通过 Hit 聚集(Clustering)算法按照发生时间(起始时间和结束时间)对其进 行组合,形成 Signal。接着对信号进行量化计算,例如计算电荷、高度、宽度、波 形峰位置和 X-Y 位置等,以及一些更复杂的信息,例如半峰宽(Full Width of Half Maximum, FWHM)、峰数(Number of Peaks)和十分之一峰宽(Full Width of Tenth Maximum, FWHM)、峰数(Number of Peaks)和十分之一峰宽(Full Width of Tenth Maximum, FWHM)、峰数(Signal 进行分类并进行标记,例如 S1、S2、Noise、 Sparking 等。在 PandaX-4T 中,对于 Signal 的分类,其分类条件是跟信号波形的 宽度、波形的峰数以及信号在项部和底部 PMT 中的分布等一些信号特征相关。最 终,Signal 以及其相关信息保存为 Signal Data。



图 3-4 经过 Hit Cluster 算法来组合的 Hit,不同颜色表示来自不同 PMT 的 Hit。

Signal Data 只是把数据文件中的所有 Signal 按照时间顺序排列在一起。为了 获得事例,需要对在某一时间窗口内的 Signal 进行配对。在 P4-Chain 数据处理链 中,有一个专门的算法称为 Physical Data Builder,用于对 Signal 进行配对和事例 切分。处理完毕后,生成的数据称为 Physical Event Data 或简称 PE Data。PE Data 中有多个信号,其中包括 S1、S2、Noise、Sparking、Unknown 等,还可能存在多 个 S1 或多个 S2。通常情况下,最大的 S1 和最大的 S2 信号是物理上相关的,而 且 S1 必须出现在 S2 之前。除此之外,对于 S1 的选择还有一些更细微的要求,例 如,在 S1 信号附近不能出现其他信号,或者 S1 信号的大小必须比附近的信号至 少大三倍。满足这些条件后,对标记的最大 S1 和最大 S2 信号进行更多变量的计 算,例如使用更优的位置重建算法(见4.2)对最大的 S2 进行 X-Y 坐标重建,以 及对最大的 S1 和最大的 S2 之间的信号信息进行更精细的计算。这些变量在后续 的数据筛选中具有至关重要的作用。

经过以上计算之后将 PE Data 转化为 ROOT 格式的数据文件 [137]。通常, ROOT 文件是用于存储和分析大量数据的常用格式。这些文件是压缩过的二进制 文件,包含一个或多个 TTree 对象,每个 TTree 包含多个 Branch,每个 Branch 存 储一种特定类型的数据,如整数、浮点数、数组等。ROOT 文件非常适合存储和 分析大量数据,因为它支持高效的数据压缩和快速的数据检索。PandaX 的所有数 据最终也被转化为 ROOT 格式的 TTree 对象,以便进行快速检查和可视化。

ROOT 格式的 PE Data 包含了事例的所有信息,它的容量很大。在 WIMP 信号区间,为了更方便地进行分析,PE Data 进一步压缩成 Analysis Data,简称 Ana Data。Ana Data 是对 PE Data 进行初步筛选后的结果,例如去除了死时间和限定了 S1 和 S2 的取值范围等条件。由于探测器尺寸较大,信号的光产额和电产额不均匀,因此在 PE Data 筛选过程中还对 S1 和 S2 大小做了非均匀性修正。

整个数据处理流程的示意图如图3-5所示。此外,数据处理的每个步骤都需要 访问数据库以获取所需的信息,例如 PMT 增益、PMT 位置坐标等。



图 3-5 PandaX-4T 数据处理流程示意图。

3.5 PMT 增益刻度

在上文中提到,将数据从 Group Data 转化为 Hit Data 或者 Signal Data 之前, 首先将 ADC 值转化为更有物理意义的 PE 值。一般将其称之为电荷 (Charge)。这 是通过 PMT 增益刻度来实现的。该刻度的意义是将 PMT 接受的光子与模拟信号 ADC 之间的关系进行参数化。无论是 PandaX-II 还是 PandaX-4T 实验,都使用蓝光 LED 对所有的 PMT 进行刻度。在 TPC 顶部的外侧放置 3 个(PandaX-II) 或 4 个 (PandaX-4T) LED,并将其电压设置在 1-10V 范围内。在 LED 刻度时,每个 PMT 尽可能只收到一个光子,此时 PMT 光阴极收到的光子数服从泊松分布。由于 LED 与各个 PMT 之间的距离不同且 LED 位置固定,因此对于每个 PMT,调整 LED 电 压以找到最合适的光强。一般情况下,无论是 PandaX-II 实验还是 PandaX-4T 实 验,每周都会对 PMT 进行一次 LED 刻度,并将结果写入数据库。之后采集的数 据都将根据此结果进行处理和分析。在收集 LED 刻度数据后,对每个 PMT 的信 号波形面积进行积分,并通过直方图获得每个 PMT 的 ADC 面积分布,如图3-6所 示。通常情况下,如果没有光子达到 PMT 光阴极,则采集到的是 PMT 本身造成 的围绕基线的暗噪声,因此此时的 ADC 面积积分在零附近浮动,对应于图中的第 一个峰。一般情况下,一个光子有一定的几率产生另一个光子(该几率称为量子 效率),且该几率相对较大。如果 LED 光子的能量足够大,则可能导致光阴极的 功函数增加两倍或更多,从而可以产生两个光电子。图3-6中的第二和第三个峰分 别对应单光子和双光子。这里使用三个高斯函数来拟合整个 ADC 面积分布:

$$f(q) = p_0 \times \text{Gaus}(q, \mu_0, \sigma_0) + p_1 \times \text{Gaus}\left(q, \mu_0 + \mu_1, \sqrt{\sigma_0^2 + \sigma_1^2}\right) + p_2 \times \text{Gaus}\left(q, \mu_0 + 2\mu_1, \sqrt{\sigma_0^2 + 2\sigma_1^2}\right)$$
(3-1)

其中 μ_1 是第二个高斯的均值,即 PMT 的增益值。这里还可以得到它的分辨率 $R_{SEG} = \sigma_1/\mu_1$ 。根据每周一次 LED 刻度的结果来看,正常的 PMT 的增益约为 140 ADC 左右,如图3–7所示。

3.6 在线数据质量监控系统

无论是 PandaX-II 还是 PandaX-4T 实验, 在整个数据获取阶段都有可能出现异常情况, 比如 PMT 突然打火或电极打火等各种问题。为了及时调整探测器并保证数据质量, PandaX-4T 实验开发了在线数据质量监控系统。该系统在数据获取阶段能够自动生成各种数据质量相关的图表。该系统采用的是 RESTful (Representational State Transfer)架构的 API (Aplication Programming Interface) [140-141]。它是一种应用程序接口的架构风格, 它遵循 RESTful 架构原则, 将资源表示为 URI,



图 3-6 在 PandaX-4T 实验中, 通道号为 10115 的 PMT 的 LED 刻度 ADC 值的分布。



PMT gain evolution (10200)

图 3-7 PandaX-4T 实验通道号为 10200 的 PMT 的 LED 刻度增益随着时间变化的趋势。

并使用 HTTP 方法(如 GET、POST、PUT 等,简称 CRUD 操作)来操作资源。 RESTful API 的核心思想是将资源表示为 URI,并使用 HTTP 方法来操作资源。例 如,使用 GET 方法读取资源,使用 POST 方法新建资源,使用 PUT 方法更新资源, 使用 DELETE 方法删除资源。这样做的好处是,开发者可以通过简单的 HTTP 方 法来操作资源,而无需关心底层实现细节。

PandaX-4T的在线数据质量监控系统包括客户端(前端)和服务端(后端)两个

部分。客户端指的是用户可以直接接触的网页,它是采用 ReactJS 来开发的 [142]。 它是一种 JavaScript 库,用于构建用户界面,由 Facebook 开发并于 2013 年开源。 ReactJS 的目标是提供一种简单、灵活的方法来构建和维护复杂的用户界面。ReactJS 的核心思想是组件化,将用户界面分解成若干独立的组件,每个组件都具有自 己的状态和行为,从而提高代码的可维护性和可重用性。ReactJS 使用虚拟 DOM 技术来渲染组件,将虚拟 DOM 与真实 DOM 进行比较,并只更新需要更新的部分, 提高了性能并减少了不必要的渲染。在 PandaX-4T 的在线数据质量监控系统中,使 用了一些由第三方提供的现成的组件包,如 Material-UI、Material-Table、BizCharts 等 [143-145],这些组件包提供了现成的组件,如各种按钮、输入框、表格、图标、 选项卡、栅格以及各类型的画图模块等。ReactJS 还提供了一种叫做状态的机制, 用于管理组件的状态。对于组件间的状态分享,采用了 React-Redux[146],它是一 个专门为 ReactJS 应用程序设计的绑定层,它提供了一种方便的方法来连接组件 与 Redux 中的总状态 (Store)。使用 React-Redux 可以使代码更加简洁、易于维护。 数据质量监控主页面如图3–9所示。

在线监控数据质量系统的服务端使用的是 Go 语言的 Echo[147]。它是一个高性能、简洁、高度可扩展的 Go 语言 Web 框架。它提供了许多强大的特性,如路由、中间件、模板渲染和请求验证,可以帮助使用者快速开发 Web 应用程序。数据质量监控时需要对数据库访问且从中获取各种信息,因此在服务端,对于 PandaX-4T的数据库查询我们使用的是 Go 语言的 Go-Pg 库。它是是一个 Go 语言的 ORM (对象关系映射)库,用于连接和操作 PostgreSQL 数据库。它是一个非常轻量级的库,可以轻松地在 Go 应用程序中使用 PostgreSQL。Go-Pg 提供了一组简单而强大的API,可以轻松地进行 CRUD 操作。它还支持高级功能,如事务、查询构建器、模型关联、模型钩子等。

PandaX-4T 数据质量图的生成是独立于服务端和客户端的。这些图片是由 PandaX-4T 数据处理链中的 Data Quality Plots Generator 程序生成的。在前文中提 到,整个数据处理链通过消息队列和管道进行通讯。一旦 Group Data 生成,数据 将被传输到成都,并通过第三个消息队列分别转化为 Hit Data、Signal Data 和 PE Data。接着,Data Quality Plots Generator 通过第四个消息队列生成一批数据质量 图。该程序采用面向对象编程,所有图形都使用 ROOT 进行绘制并保存到指定目 录下。在网页端点击特定的 run 时,会向服务端发送图片对应的 URL 请求,从而在 网页端显示图片。图3-8展示了在线数据质量监控系统的结构。Data Quality Plots Generator 的输入文件分别为 Hit Data、Signal Data 和 PE Data,因此数据质量图 也大致上分为三种类型。在网页端共有 50 多张数据质量图,包括每个通道的 Hit

发生频率、不同类型信号的 PMT 电荷分布、不同信号的位置分布和电子寿命等。 图3-10展示了部分数据量图。此外,网页端还能显示各个数字化波形采集卡对应 的每个通道的带宽随着时间变化的趋势,如图3-11所示。

在线数据质量监控系统的功能不仅仅是显示图片。一般来说,现场取数人员 使用该系统对每个运行类型做标记,并将标记自动写入数据库中。此外,该系统还 可以标记质量差的数据文件以及异常或坏掉的 PMT。它还可以显示每个 PMT 增 益随时间变化的趋势,以及每个 PMT 通道的波形。通过网页端,可以计算和显示 各种数据的采集时间。总之,PandaX-4T 的在线数据质量监控系统在日常取数阶 段扮演着不可或缺的关键作用。



图 3-8 PandaX-4T 在线数据质量监控系统。

3.7 事例可视化工具

在 PandaX 实验数据分析中,通常需要频繁查看事例波形以及事例在 PMT 阵列上的电荷分布。在 PandaX-II 实验中,直接使用 ROOT 来画出事例的波形以及它在 PMT 阵列上的电荷分布。然而,一般 ROOT 画出的图都是静态的,无法进行进一步的灵活操作,而且查看事例波形的效率较低。为了解决这个问题,在 PandaX-4T 实验数据分析中开发了一款跨平台的事例可视化工具,命名为 P4-Event-Viewer。 P4-Event-Viewer 同在线监控数据质量系统一样,也使用了 RESTful 框架的 API。客户端部分同样使用 ReactJS,但打包以及部署部分使用了 ElectronJS[148]。它是 Runs Files Rates Channels Gain Duration Help **Run Information & Data Quality** Enter Run ID Enter Description Keyword Enter Run Type : gate, junk, 6.3KV or AmBe etc e.g. : 1122 or 1122-1130 e.g. : HW_DEBUG or hw_debug Run Information Q Search ⊎ Drag headers here to group by Start Time Duration Run Type Description Actions Run ID Number of File Cathode -16kV, Gate ľ 3788 27 2021/03/30 11:27:16 41m49s PHYSICS DM -5kV, bkg Cathode -16kV, Gate ï 3787 83 2021/03/30 09:16:28 2h6m43s PHYSICS_DM -5kV, bkg Cathode -16kV, Gate ľ 3786 58 2021/03/30 07:33:18 1h29m2.4s PHYSICS_DM -5kV, bkg Cathode -16kV, Gate ľ 3785 38 2021/03/30 06:25:44 59m56s PHYSICS_DM -5kV, bkg ľ 3784 3 2021/03/30 06:16:03 4m50s UNDETERMINED Cathode -16kV, Gate 3783 133 2021/03/30 01:52:18 3h29m20s PHYSICS_DM -5kV, bkg

上海交通大学博士学位论文

图 3-9 PandaX-4T 实验在线数据质量监控系统主页面。

一个开源框架,它能够使用 JavaScript、HTML 和 CSS 构建跨平台的桌面应用程序。 它提供了一个可以访问桌面系统 API 的完整环境。ElectronJS 的最初目的是为了为 GitHub 的 Atom 编辑器构建一个跨平台的框架,但是现在,它已经发展成为一个十 分流行的开发框架,被许多开发者用来构建桌面应用程序。ElectronJS 的最大优势 之一是它能够让开发者使用同一套 Web 技术来构建桌面应用程序,而不需要学习 一系列新的技术。它能够节省开发者花费在学习新技术上的时间,提高开发效率, 换句话说,开发者只要熟悉 Web 技术,就能轻松的使用 ElectronJS 来搭建一款桌 面应用。ElectronJS 还具有跨平台性,可以让开发者在 Windows, macOS 和 Linux 系统上发布相同的应用程序,而无需重新编写代码。ElectronJS 为开发者提供了一 个完整的开发环境,可以让开发者轻松打包和部署应用程序,从而节省开发者的 时间和精力。除此之外,在数据可视化部分使用的是 EChartsJS[149]。EChartsJS (Enterprise Charts) 是一个由百度开发的 JavaScript 图表库,用于在 web 应用程序中 创建各种各样的图表和图形。ECharts 支持大量的图表类型,如柱状图、饼图、折 线图、散点图、雷达图、热力图、地图等。同时它支持数据的动态更新,可以实现



图 3-10 PandaX-4T 实验在线数据质量监控系统产生的部分数据质量图。

图表的实时监控和数据可视化。EChartsJS 还支持丰富的交互功能,如支持鼠标悬 浮显示数据、支持点击事件、支持拖拽和缩放等。这些交互功能可以让图表更加 直观和互动,提高用户体验。并且它还支持多种数据格式,如 CSV、JSON、Excel 等。这使得它可以从不同的数据来源中读取数据,并且可以方便地与其他数据分 析工具集成。EChartsJS 的另外一个重要功能是 Down Sampling。Down Sampling 是 指在图表中对数据进行采样的功能。这个功能可以帮助开发者处理大量数据时的 性能问题,并且可以使图表更加平滑。EChartsJS 提供了两种 Down Sampling 的方 式,第一种是 Sampling by data: 是根据数据点的数量来进行采样,只保留一定数量 的数据点。这种方式可以保证图表上的数据点数量不会过多。第二种是 Sampling by time: 这种方式是根据时间来进行采样的,只保留一定时间间隔内的数据点。这 种方式可以保证图表上数据点的时间间隔不会过短。使用 Down Sampling 功能可 以更好地展示大量数据,并且可以使图表更加平滑。一般波形数据点可达上百万 个,因此经过使用 EChartsJS 的 Down Sampling (Sampling by time)的方式,可以 对波形可视化性能以及显示效果进行自动优化。

值得一提的是 EChartJS 是纯 Javascript 库,无法直接在 ReactJS 中使用它。因此 在项目中使用的是 ECharts-for-React。它是是一个基于 React 框架的 EChartsJS 库, 可以在 React 应用程序中使用 EChartsJS 创建图表。ECharts-for-React 将 EChartsJS 的功能封装成了 ReactJS 组件,开发者可以使用这些组件在 ReactJS 应用程序中创 建各种图表。

在前文中提到了 P4-Event-Viewer 采用 RESTful 框架的 API,因此必然存在服





图 3-11 PandaX-4T 实验每个数字化波形采集卡的带宽随着时间变化的趋势,不同颜色表示的是不同的通道。

务端。服务端会根据客户端发送的请求将请求的内容发送给客户端,这里发送的 内容是事例数据。下面介绍获取事例数据的流程。

在三中已经介绍了 PandaX-4T 实验的标准数据处理流程,但在日常数据分析 中通常需要更改重建算法并计算新的数据变量。然而,如果根据更新后的算法重 新处理数据,会非常耗时(至少需要 4-5 天)并占用大量计算资源,从而影响其他 日常数据分析。在日常数据分析中,通常只需要分析一部分筛选过的数据,而不需 要对所有数据进行分析。因此,在 PandaX-4T 数据分析中,开发了一个名为 Group Picker 的独立程序。它的功能是仅对某一部分筛选过的数据进行重新处理。这种 处理对计算服务器非常友好,不会占用计算资源,而且处理速度也非常快(一般 在 1 小时之内可以处理 5000 个左右的事例)。这种数据处理方式与标准的数据处 理方式之间的主要区别在于 Group Data 部分。标准的数据处理方式是根据原始数 据进行切分的。而通过 Group Picker 处理的数据是根据事例的信息(run number, file number, event number, event start time, event end time)从 Group Data 版本的事 例数据中提取并合并到一起,并保存到一个文件。随后的数据处理与标准的数据 处理相同(Group Data -> Signal Data -> PE Data -> ROOT file)。此外,它们之间 的另一个区别是原始数据没有索引,而 Group Picker 处理后的数据带有索引,这 主要是为了更方便地查询事例而引入的功能。Group Picker 处理的数据本质上是 反向数据处理过程。因为原始的 Group Data 中不存在事例的数据结果,事例在 PE Data 中存在。因此,第一次处理必须按照标准的数据处理方式处理数据,才能获 取事例的基本信息并进行筛选。

有了事例数据以后,通过服务端程序可以对事例数据进行读取,处理以及计算,然后根据客户端请求把对应的数据反馈到客户端。在服务端,我们使用了一个叫做 yhirose/cpp-httplib 的 C++ HTTP/HTTPS 库 [150]。它是一个用 C++ 11 来开发编写的库,由 Yutaka Hirose 开发的,并在 GitHub 上公开发布。它库提供了简单易用的 API,可以让开发者在 C++ 项目中轻松实现 HTTP 客户端和服务器功能。它支持 HTTP/1.1 和 HTTPS,并且支持常见的 HTTP 方法,如 GET、POST、PUT 和 DELETE。它还支持 Windows、macOS 和 Linux 等平台,并且兼容 C++11 和 C++14 等版本以及许多其他功能,如文件上传、Cookie 支持、压缩支持、连接池等。它还支持许多 C++11 和 C++14 特性,如 lambda 表达式、移动语义等。整个库只有一个头文件组成,在程序只包含头文件 httplib.h 即可。使用这个库可以很容易地实现一个简单的 HTTP 服务器,例如:

```
#include <httplib.h>
int main() {
    httplib::Server svr;
    svr.Get("/", [](const httplib::Request& req, httplib::Response& res) {
        res.set_content("Hello, World!", "text/plain");
    });
    svr.listen("localhost", 1234);
}
```

这段代码实现了一个简单的 HTTP 服务器,当客户端发送 GET 请求到服务器 的根路径时,服务器会返回"Hello, World!" 的文本。

虽然客户端和服务端间的请求/响应可以以各种各样的数据格式来进行,比如 TXT,CSV,XML,HTML以及JSON等。通常情况下,JSON(JavaScript Object Notation)格式其中最为流行的数据格式。JSON是一种轻量级的数据交换格式,它 具有简洁明了的语法和良好的可读性。它常用于在网络应用程序中进行数据交换, 因为它可以在各种编程语言中很容易地解析和生成。因此在服务端程序中使用了 一个叫做 nlohmann/json 的 C++ JSON 库 [151]。它是由 Niels Lohmann 开发的,并 在 GitHub 上公开发布。nlohmann/json 提供了一个简单易用的 API,可以让开发人 员在 C++项目中轻松实现 JSON 的读写操作。它提供了一个 JSON 类型,可以表示 JSON 数据,并且支持常见的 JSON 数据类型,如整型、浮点型、字符串、布尔型、 数组和对象。它还提供了一系列输入输出运算符和函数,可以实现 JSON 数据的 读写操作。nlohmann/json 还提供了许多其他功能,如 JSON 数据的遍历、排序、过 滤等,使用它可以实现自定义数据类型的序列化和反序列化。此外, nlohmann/json 还支持 JSON 数据的格式化和验证功能。它也同样地只由一个头文件来组成。它 在 GitHub 上有着高达几千次点赞,并且已经成为了许多 C++项目的首选以及标 配。关于更多的 JSON 数据操作,请访问 [151]。

在服务端程序中,总共创建了9个 API 的 URL,其中5个是事例(Event)相关的,2个是信号(Signal)相关的,还有2个是关于 PMT 通道相关的。例如事例相关的 API 是,/event,/event/next,/event/previous 等等。这些 API 已请求形式发送到服务端程序后,服务端程序将对应的 JSON 数据格式发送客户端。例如在/event的 API 中以 JSON 格式包含事例的波形数据以及所有 PMT 的电荷数据。为了减轻客户端的计算压力以及内存资源,所有的数据计算都在服务端进行,这种方式的可视化由于频繁发送请求在操作中可能会造成一定的延迟。但是它能够提高整个工具的性能。

通常情况下,该应用的客户端被安装在本地电脑上,而服务端被安装在远程服务器上。当然,服务端也可以在本地运行,但需要安装 Bamboo-Shoot3 和许多其他相关库。由于不同用户使用不同的操作系统,因此服务端的安装对电脑的兼容性有相对较高的要求。

然而,远程服务器上已经有了可用的程序运行环境,因此用户通常会在远程 服务器上安装并编译程序,或直接使用已编译的程序。但是,客户端的请求无法 直接发送到远程服务器上的服务端程序。可以通过使用 SSH 的端口映射功能来解 决这个问题。

SSH 的端口映射,也称为端口转发,是一种在两台计算机之间建立网络连接的技术。它可以将一台计算机上的某个网络端口映射到另一台计算机上的另一个网络端口,从而实现远程访问。例如,可以将本地计算机上的端口 22 映射到远程计算机上的端口 22,以便在本地计算机上使用 SSH 协议连接远程计算机。

这样就可以在本地计算机上使用 SSH 客户端软件连接远程计算机,并在远程 计算机上执行命令或传输文件。SSH 的端口映射可以通过加密来保证数据传输的 安全性。这样就可以在不需要直接连接远程计算机的情况下,进行远程控制和访 问。这对进行远程管理、远程数据备份和远程监控等操作非常有用。

SSH 端口映射的命令是 ssh -L 或 ssh -R,其中-L 表示本地端口映射,-R 表示 远程端口映射。例如,使用命令 ssh -L 8080:192.168.1.100:80 root@192.168.1.100 可 以将本地计算机上的端口 8080 映射到远程计算机 192.168.1.100 上的端口 80,从 而在本地计算机上通过浏览器访问远程计算机上的 web 服务器。

此外,还可以使用-N参数来禁用远程命令执行,这样可以使用 SSH 连接作为 端口转发工具。例如,使用命令 ssh -L 8080:localhost:80 -N user@remotehost 可以 将本地计算机上的端口 8080 映射到远程计算机上的端口 80,并且禁止远程命令 执行,只作为端口转发工具。另外,用户还可以使用-f 参数将 SSH 连接放入后台 运行。例如,执行以下命令: ssh -f -L 8080:localhost:80 user@remotehost。这样,在 连接到远程主机后,SSH 进程将会在后台运行,这样用户就可以在终端中执行其 他操作。

图3-12展示的是事例可视化工具的事例的信号页面。可以看到,在Logo下方 有一个输入框,用于输入请求和响应的端口号,所有请求和响应都通过端口映射 实现。输入端口号并按下回车键后,可以进入事例页面。在端口号下方的选项卡 中,可以切换事例和信号页面。如前文所述,事例由多个信号组成,其中包含 S1、 S2 和 Noise 等。因此,当点击选项卡中的信号时,会显示该事例的所有信号相关 内容。事例页面与信号页面基本相似,唯一的区别在于左侧的表格。事例页面的 表格包含事例的索引(run number, file number, event number),而信号页面的表 格包含该事例的信号索引、信号类型和信号总电荷(单位为 PE)。

在选项卡下方,有一个类似手风琴的组件,单击该组件会显示该事例的基本 信息(事例索引、信号数目、总电荷等)或该信号的基本信息(类型、电荷、宽 度、高度和位置等)。该组件下方是事例或信号的波形,使用 EChartsJs 库实现。通 过单击不同图例按钮(Top/Bottom/Veto/Total),可以显示或隐藏不同的波形成分。 由于波形使用 Down sampling 显示,无论数据点有多少,可以通过滚动条灵活地 缩放波形。此外,通过单击右上角的下载按钮,可以以 SVG 格式下载整个波形。 波形下方是 PMT 的电荷分布图,每个圆圈表示一个 PMT,每个 PMT 都有唯一的 编号或通道号。当将鼠标悬停在相应的圆圈上时,会自动显示该通道号和该 PMT 的坐标。

PMT 的颜色表示该 PMT 接收到的光的电荷量,颜色越接近红色,表示电荷越大;越接近蓝色,表示电荷越小。当将鼠标悬停在某个 PMT 上时,该 PMT 的 电荷量会在颜色轴上自动显示。PMT 电荷分布下方的输入框可以输入 PMT 的通 道号,按下回车键可以显示该 PMT 通道号对应的波形。

上海交通大学博士学位论文

S P4 Event Viewer

Enter port number 11180

EVENTS SIGNALS



图 3-12 PandaX-4T 事例可视化工具页面。

3.8 PandaX 实验数据总结

实验数据的质量和采集时长直接影响分析结果和探测器灵敏度。因此,在保证数据质量的前提下,采集更多的数据可以更方便地进行数据分析。PandaX-II实验中有三个高质量的数据集,分别为Run 9、Run 10和Run 11。Run 9是PandaX-II 实验中第一批物理数据,持续了79.6天。为了对探测器进行电子反冲刻度, 氚化 甲烷被注入到探测器中(详见[130])。但是,刻度结束后,无法完全去除氚化甲烷, 这对暗物质能区的分析带来了困难。因此,通过离线精馏探液氙的方式去除了大 部分氚化甲烷。精馏完成后,进入了Run 10数据集的采集阶段,持续了77.1天。通 过对Run 9和Run 10数据集进行暗物质分析,获得了世界最佳的探测结果。随后, 继续采集了持续244.2天的数据,形成了Run 11数据集。表3-1总结了PandaX-II 实验中的数据参数。PandaX-4T实验尚未发布正式的物理数据集结果。自2020年 12月开始试运行,经过95天运行,获得了第一批试运行数据集。于2021年8月首 次公布试运行数据集并发布寻找暗物质信号的结果。试运行是了解探测器的过程, 需要对探测器运行条件进行适当调整。因此,根据探测器运行条件,将PandaX-4T 实验试运行数据集分为5个子数据集。各个子数据集的主要区别在于探测器 TPC 电极的电压不同。表3-2总结了PandaX-4T 实验试运行数据集的具体参数。

Dataset	Live time	Cathode (Gate) HV	Max drift time	PDE	EEE	SEG	$ au_e$
	(day)	(kV)	(µs)	(%)	(%)	(PE/e^{-})	(µs)
Run 9	79.6	-29 (-4.95)	350	11.50	46.34	24.36	623
Run 10	77.1	-24 (-4.95)	360	12.05	50.78	23.69	850
Run 11	244.2	-24 (-4.95)	360	11.99	47.49	23.53	1367

表 3-1 PandaX-II 实验数据参数总结。第二列对应的是活时间,其第三列为阴极(门电极)对 应的负高电压,第四列为最大的电子漂移时长。第五、六、七列对应的参数在第四章中讨论。 最后一列为平均的电子寿命,也是在第四章中讨论。

Dataset	Live time	Cathode (Gate) HV	Max drift time	PDE	EEE	SEG _b	$ au_e$
	(day)	(kV)	(μs)	(%)	(%)	(PE/e^{-})	(µs)
Set 1	1.81	-20 (-4.9)	800	9.0	90.2	3.8	800.4
Set 2	12.28	-18.6 (-4.9)	810	9.0	90.2	3.8	939.2
Set 3	5.13	-18 (-5)	817	9.0	92.6	4.6	833.6
Set 4	32.98	-16 (-5)	841	9.0	92.6	4.6	1121.5
Set 5	33.84	-16 (-5)	841	9.0	92.6	4.6	1288.2

表 3-2 PandaX-4T 实验数据参数总结。其第二列是去掉死时间以及质量有问题的文件之后的 活时间。

3.9 暗物质信号区的数据筛选

在3.4中指出, PE Data 转换为 ROOT 格式后,可以对其进行进一步处理得到 Ana Data。这是对 PE Data 进行初步筛选后得到的数据,包括死时间、S1 和 S2 的

取值范围以及事例的发生位置等。除了这些筛选,还需对 Ana Data 进行进一步的数据筛选,因为其中仍包含许多非物理的或非常明显的本底。

PandaX-II 和 PandaX-4T 暗物质数据分析中的数据筛选有很多相似之处,但是由于 PandaX-4T 实验的数据量庞大且复杂,采用的筛选条件更多。尽管 PandaX-4T 的数据筛选条件大部分能够覆盖 PandaX-II 的数据筛选条件,但是为了选取高质量的物理事例,PandaX-4T 实验采用了更多的筛选条件。因此,下面主要介绍 PandaX-4T 的数据筛选条件。

- 文件质量筛选:由于探测器运行不稳定,例如 PMT 或电压等问题,会产生 各种打火事例,导致数据文件中会出现更多的孤立 S1 或 S2 事例,增加了 偶然符合本底的数量。一般情况下,这些文件被放弃使用。关于偶然符合 本底,第六章中进行详细讨论。
- 死时间条件:通常,当高能物理事例(MeV级别)在液态氙中沉积能量或 产生大量电子时,这些电子在电场的作用下上升,但由于电子数量很多,不 能同时转化为S2光信号,一部分电子会间歇性地转化为S2光信号,从而 产生许多非物理事例。通常使用死时间参数去除这些事例,但这样做会丢 失一部分采集数据的时间。
- 对 S1 和 S2 信号电荷的限制:这种限制应根据信号特征来确定。例如,在 PandaX-II 实验中,对于暗物质信号来说,S1 信号的电荷应在 3 至 45 PE 之 间,而 S2 信号的电荷应在 100 至 10000 PE 之间。而在 PandaX-4T 实验中, S1 信号的电荷应在 2 至 135 PE 之间,S2 信号的电荷应在 80 至 20000 PE 之间。
- 4. S1 信号点亮的 PMT 数量下限:由于 S1 信号相对于 S2 信号较小,特别是 对于低能事例,S1 信号的幅度可能很小,因此被 S1 信号点亮的 PMT 数量 有时很少。但为了排除其他打火事例造成的 S1 信号或暗噪声造成的偶然 符合 S1 信号,例如,在 PandaX-II 实验中要求被 S1 点亮的 PMT 数目不少 于三个,而在 PandaX-4T 实验中要求不少于两个。
- 5. 对 S1 信号波形中峰数目的限制:由于部分单电子或双电子的 S2 信号在波 形中表现为若干个小信号,因此它们中的一个可能被软件算法标记为 S1 信号,而且这种假的 S1 信号可能具有许多峰。因此,根据峰的数量来排除 这种情况。
- 6. 对 S1 信号在顶部和底部 PMT 接收的信号大小对称性的限制:通常情况下, 在液氙产生的光,通过气液氙表面能够,发生全反射,这时正常物理事例的 S1 信号大部分分配在底部 PMT,因此当沉积能量较低时,顶部 PMT 可

能探测不到信号。如果 S1 信号只被顶部 PMT 探测到,说明这种事例大部 分是非物理事例。这种对称性与 S1 信号大小密切相关。因此,通过 S1 信号大小和 S1 信号在顶部和底部 PMT 的对称性来排除部分非物理事例。从 图3–13中可以看出,大部分只被顶部 PMT 探测到的 S1 信号事例都被排除 掉。

- 7. 对 S1 信号在顶部或底部 PMT 中的分布的弥散程度的限制:有时 S1 信号 绝大部分集中在某一个 PMT,这种信号可能是 PMT 本身打火的原因而造 成的。因此,可以根据 S1 信号在顶部或底部 PMT 中的分布的弥散程度来 排除这种事例。
- 对 S1 信号波形附近的其他信号大小的限制: 部分 S1 信号附近可能出现各 种杂乱的信号,而且这种信号数量也有可能较多,为了保证 S1 信号的物 理来源性,这种事例也会被排除掉。
- 9. 对电子漂移时间的限制:一般情况下,在探测器顶部(门电极)或底部(阴极)部分存在各种明显的本底事例。例如阴极事例、气体事例以及各种电压造成的打火事例。排除这种事例的最直接方法是利用电子的漂移时间,换句话说,根据S1和S2信号发生的时间差来确定这种事例发生的位置。根据探测器的大小以及电子的漂移速度,限制条件也不一样。例如,在PandaX-II 实验中(其电子漂移速度1.7mm/µs),这种条件被限定为50-350µs,对于PandaX-4T 实验(其电子漂移速度为1.4mm/µs),40-800µs。
- 对于水平方向位置的限制:大部分与真实信号相似的信号通常发生在探测器中心区域,而探测器边缘部分则会受到很多材料相关的放射性元素发出的各种射线的影响,这些显然来自于本底信号可以通过水平方向位置来排除。
- 11. 根据电子的扩散效应对事例进行筛选:随着探测器体积的增大,电子漂移 距离变长,因此在 PandaX-4T 探测器相比于 PandaX-II,电子团漂移时容易 出现扩散现象。这种扩散现象导致 S2 信号的波形变宽,并且扩散效应与事 例在垂直方向上的位置有关。当事例发生的位置靠近底部时,扩散效应越 明显。相反,当事例位置较远离底部时,扩散效应不太明显。这是物理事 例的一个非常重要的特性。由于 PandaX-II 探测器相对于 PandaX-4T 探测 器较小,因此在 PandaX-II 实验中,这种扩散效应不太明显。在 PandaX-4T 实验中,我们可以根据事例的这种特点排除绝大部分偶然符合本底(关于 偶然符合本底见六)。因为偶然符合本底是由无关联的 S1 信号和 S2 信号 偶然符合形成的,因此我们看不到这种扩散效应。电子的扩散效应可以根

据 S2 信号的宽度和电子漂移时长来描述。如图3-14所示,红色线上方和 下方的事例大部分是偶然符合事例。

- 12. 对于 S2 信号波形的对称性的限制:一般,由于探测器核心区域电场相对 于来说较均匀,在此区域发生的正常物理事例导致的 S2 信号波形通常具 有类似高斯分布一样的对称性。如果波形对称性较差,则说明这些事例可 能受到强电场的影响,因此这种事例可能来自于电极。通过 S2 信号波形 的对称性可以排除这种事例。
- 13. 对于 S2 信号在顶部和底部 PMT 接收到的信号大小对称性的限制:低能事例中,S2 信号与 S1 信号在顶部和底部 PMT 上的分布不同,因为 S2 信号 产生的位置在气氙中,因此大部分 S2 信号被顶部 PMT 探测到。通过这种 特征可以排除一些在 PMT 分布上异常的 S2 信号事例。
- 14. 对于不同位置重建算法重建的位置坐标之差的限制: PandaX 实验中采用 了多种位置重建方案(详见[152])。如果某个事例通过不同位置重建算法 重建出的位置坐标之差较大,则说明该事例的 S2 信号在顶部 PMT 上的分 布不均匀,此外,该事例的水平方向上的位置信息误差也较大。因此,可 以排除这种事例。
- 15. 对于 S2 信号波形形状的限制,大部分来自于物理事例的正常 S2 信号形状 与高斯分布相似。如果波形形状过于"瘦"或过于"胖",则说明该 S2 信 号可能来自于电极(门电极或阴极)。可以通过 S2 信号大小与其宽度和高 度之比作为分布变量来排除这种事例。
- 16. 根据非 S1 和非 S2 信号的占比来对事例进行筛选:一般的事例除了 S1 和 S2 信号之外,还有很多其他的非 S1 和 S2 信号。为了保证 S1 和 S2 信号 的环境更加"干净",我们通过计算 S1 和 S2 信号在整个事例中的占比来 排除一些质量较差的事例。

以上提到的都是 PandaX 实验暗物质能区内使用的最主要的数据筛选方法。一 般会利用刻度数据来确定每个筛选的具体参数。虽然这些筛选方法能够去除大部 分非物理或质量较差的物理事例,但也会对真正的信号造成一定程度的损失。因 此,在计算信号的事例率时,需要考虑这些数据筛选方法对真实信号损失率的影 响。



图 3-13 PandaX-4T 实验刻度数据的电子漂移时长跟 S1 信号在顶部和底部 PMT 接收的电荷 对称性的二维分布。红色界线之外的部分被排除掉的数据。



图 3-14 PandaX-4T 实验刻度数据的电子漂移时长跟 S2 信号宽度的二维分布。红色界线之外的部分被排除掉的数据。

第四章 探测器的响应参数和刻度

在本章节中,我们首先讨论各种探测器的响应参数,包括 PMT 双光电子产生 率、最小的单电子信号的电荷、探测器探测效率、电子的拽出效率等。除此之外, 还讨论探测器的不均匀性修正参数、能量重建以及位置重建等。最后,我们讨论 探测器刻度。

4.1 PMT 双光电子产生率

在3.5中介绍,通过 LED 光来进行 PMT 增益刻度。如图3-3所示,虽然在 LED 光会产生一定比例的双光电子,但是相对于真实的 178nm 紫外光,LED 光的波长 较大(约为 390nm),PMT 的光阴极产生双光电子的几率较小,甚至可以被忽略。为了得到真实的双光电子产生几率,需要从真实物理信号中寻找最小的 S1 信号。为此,在 PandaX 实验中选取了刻度数据或暗物质数据。对于刻度数据,其中一 个是 ²⁴¹Am-⁹Be 数据(核反冲刻度),另一个是 ^{83m}Kr 数据(电子反冲刻度)。之 所以选取刻度数据,是因为这些数据信号大部分都来自于真实物理事例。选取 S1 的条件是,该信号只由一个 Hit 来组成,之后通过直方图来得到其电荷分布,如 图4-1所示。这里使用双高斯函数来拟合数据,拟合后得到的结果(PandaX-4T 实 验)是,从 ²⁴¹Am-⁹Be 数据的分布中计算得到 22.8%,而从 ^{83m}Kr 数据的分布中计算得到 23.9%。这中微小差别可以当作两种数据的系统误差。

4.2 位置重建

2.2.2中对 TPC 介绍时提到, PandaX 实验能够对探测器内部的物理事例进行 三维位置重建,这是它们的最明显的优势之一。在垂直方向(Z方向),位置重建 是根据电子的漂移时长来确定的,也就是 S1 信号和 S2 信号之间的时间差。例如, 在 PandaX-4T 实验中,阴极和门电极之间的距离约为 1185mm,最靠近阴极的事 例(称为阴极事例)产生的电子漂移时间约为 840μs,由此可得电子的平均漂移速 度约为 1.4mm/μs。

PandaX-4T 实验中,我们以三种算法来计算了物理事例的水平方向(X-Y)位置。第一种算法比较直观,是根据物理事例 S2 信号在顶部 PMT 阵列被接收的电荷分布来确定,这种算法被称之为 Center of Gravity (COG):



图 4-1 PandaX-4T 实验刻度数据最小 S1 信号(只由一个 Hit 信号组成的)的电荷分布(蓝色 直方图),红色实线为全局拟合函数,绿色和粉色分别为单光电子以及双光电子信号的高斯 分布。

$$x_{\text{COG}} = \sum_{i} \left(X_{i} \times q_{i} \right) / \sum_{i} q_{i},$$

$$y_{\text{COG}} = \sum_{i} \left(Y_{i} \times q_{i} \right) / \sum_{i} q_{i},$$
(4-1)

其中 i 是 PMT 的编号, x_{COG} 和 y_{COG} 分别表示第 i 个 PMT 的 X 和 Y 坐标, q_i 是被第 i 个 PMT 接收的电荷。此算法虽然直观且简单,但是对于靠近边缘区域的 事例有一定的误差,因为边缘区域的 PMT 数目相对来说较少,因此在这种区域重 建的位置比实际位置更偏向内侧。第二种算法叫做 Template Matching (TM)。它 首先是通过模拟 (Geant4 模拟) 来得到光子在不同位置被顶部 PMT 阵列接收的 电荷密度分布以及该事例的确切位置信息,然后通过最大似然法将数据的电荷密 度分布跟模拟的电荷密度分布比较,从而匹配最佳的位置信息。第三种算法叫做 Photon Acceptance Function (PAF),这种算法利用的也是最大似然函数法,但是对 于电荷密度分布,它利用的是根据数据得到的解析函数。计算 PAF 位置前,必须 提前算出它的 COG 位置,因为 COG 位置是 PAF 位置算法的初始输入值。关于此 方法更详细的内容请参见 [152]。

4.3 电荷修正

PandaX-II 或 PandaX-4T 探测器的体积较大,因此无法保证探测器各项参数和性能的稳定性。例如,TPC内的电场存在一定的不均匀性,气液氙表面的水平度也存在不均匀性。此外,在TPC内发生物理事例时,PMT的光学立体角不同,因此能够影响能量重建的准确性。因此,无论是 PandaX-II 还是 PandaX-4T 实验,都需要修正 S1 和 S2 大小。通常情况下,可以通过根据探测器内均匀分布的某种已知的单能峰来修正上述提到的不均匀性。这种单能峰可以是液氙内部的本底,也可以是从外部注入的放射源。此外,由于暗物质事件主要发生在低能区域,因此单能峰的能量不能过高,并且需要有足够多的统计量。

在 PandaX-4T 实验中,为了修正 S1 和 S2 的大小,使用放射源^{83m}Kr 采集数据,它具有 41.6 keV 的β衰变,统计量约为6万,满足上述要求。由于 PMT 阵列的光学立体角不同,导致 S1 信号的空间非均匀性。因为无法使用一个简单的拟合函数对 S1 信号的大小进行三维修正,因此采用三维直方图的方法对 S1 信号的大小进行三维修正。首先,将整个 TPC 在三个(X-Y-Z)方向分割为若干个小块,但同时也要保证每个小块有足够的统计量。然后,对每个小块使用高斯函数进行拟合,并得到均值。在得到均值后,通过插值的方法对所有小块进行平滑化,最终将数值存储在三维直方图中,并生成关于 S1 的三维修正映射关系文件。每次对其他物理事例进行修正时,可以根据该物理事例的电荷和位置信息以及映射文件,计算出其修正后的 S1 大小。这就是对 S1 大小进行修正的核心逻辑。

对于修正 S2 大小的方法,同样使用了^{83m}Kr。这种修正方法与修正 S1 大小的 方法类似,但是对于 S2 信号来说,这种映射关系只在水平方向(X-Y)上进行修 正,而垂直方向的修正是根据电子寿命进行的。

首先,我们需要介绍电子寿命的概念。在2.3.2中提到,液氙中的各种杂质通 过循环系统逐步去除。然而,循环系统刚开始运行时无法去掉所有的杂质。例如, 在刚开始采集数据时,液氙中仍然含有一些电负性气体。随着时间的推移,这些 杂质的含量逐渐减少。当物理事件的能量沉积在探测器中,产生的电子在电场的 作用下向上漂移。在这个过程中,部分电子可能会被这些电负性杂质吸收。这种 吸收与事例发生的位置有关。事例发生的位置越靠近阴极,被吸收的电子就越多, 类似于放射性元素的衰变。这个过程服从指数函数:

$$S2_{\text{detected}} = S2_{\text{original}} \times e^{-t_{\text{drift}}/\tau_e}, \qquad (4-2)$$

其 $S2_{detected}$ 和 $S2_{original}$ 分别为探测到的和原始的 S2 信号(电荷)大小, t_{drift} 和 τ_e 分别为电子的漂移时长以及电子寿命。

为了计算电子寿命,需要在 TPC 内部均匀分布的单能峰,同时还要保证其寿命相对较长。在 PandaX-4T 实验中,我们选择 ^{131m}Xe、²²²Rn 和 ²¹⁸Po 作为放射性单能峰来计算电子寿命。其中,^{131m}Xe 是中子的刻度导致的核激发产生的亚稳态同位素。它的退激发会产生 164 keV 的单能峰,其半衰期约为 11.9 天。探测器液 氙内的一个重要本底是 ²²²Rn,它来自于 ²³⁸U 的衰变链,能放出能量为 5.5 MeV 的 *α* 粒子,同时衰变为 ²¹⁸Po。²¹⁸Po 也能放出能量为 6 MeV 的 *α* 粒子,同时衰变为 ²¹⁴Pb。由于两种 *α* 粒子的能量比较接近,因此来自 ²²²Rn 和 ²¹⁸Po 的 *α* 粒子也可以 用来计算电子寿命。

计算电子寿命后,根据电子寿命值,通过指数函数对其他物理事件进行 S2 垂 直方向的修正。图4-2展示了两种电子寿命计算结果,其差别很小。正如前文所述, 由于循环系统的作用,液态氙中电负性杂质的浓度逐渐减少,因此电子寿命的数 值也在不断上升。



图 4-2 PandaX-4T 实验中电子寿命随着时间变化的趋势。其蓝色和红色分别用 164 keV 的特 种峰事例和 α 事例计算的结果。

4.4 能量重建

修正 S1 和 S2 信号电荷之后,可以使用以下公式对电子反冲事例进行能量重建:

$$E_{\rm ee} = 0.0137 \text{keV} \times \left(\frac{S1}{g_1} + \frac{S2_b}{g_{2b}}\right).$$
 (4-3)

这里的 Eee 是等效电子反冲能,需要注意的是,由于核反冲产生了额外的热 能,所以在该公式中需要添加 Lindhard 因子 [153]。其中 S2, 是指 S2 信号中只有 底部 PMT 阵列接收到的部分。因为当物理事件的能量沉积较高时,顶部 PMT 很 容易出现饱和现象,因此在此只考虑由底部 PMT 接收的信号。0.0137 keV 是氙原 子激发所需要的平均能量 [154], g1 是光子探测效率(Photon Detection Efficiency, PDE),该因子考虑到了 PMT 和探测器本身的各种因素。 g_{2b} = EEE × SEG_b,其中 EEE 和 SEG 分别代表电子拽出效率(Electron Extraction Efficiency, EEE)和单电 子增益(Single Electron Gain, SEG)。SEG是指电子从气氙中拽出后产生的光电子 信号的大小, SEG 的值是通过选取数据中的最小 S2 信号来计算的。如图4-3所示, 黑色点是数据中选取的最小 S2 信号大小的分布, 该分布由费米-狄拉克和双高斯 函数拟合而成,费米-狄拉克函数考虑了硬件和数据重建造成的最小单电子 S2 信 号的探测效率,第一和第二个高斯分布分别代表单电子和双电子的光电子峰。单 电子高斯峰的均值即为 SEG 的值,根据图4-3的结果, SEG 的值约为 19.3 ± 0.1。 同样地,对于较大的能量沉积,PMT 很容易出现饱和现象,从而导致 SEG 的误差 较大。因此, PandaX-4T 实验专门利用底部 PMT 来计算 SEG_b,约为 SEG 的 20%。 EEE 是在特定电压下,从液氙中拽出到气氙的电子所占的百分比,一般这个比例 与电压成正比。

通过所谓的 Doke Plot(光电反相关图)[155],可以确定 g_1 和 g_{2b} 。在探测器运行期间,存在固有或者外部注入的各种单能峰,例如来自^{131m}Xe 的 164 keV,来自^{129m}Xe 的 236 keV,来自¹²⁷Xe 的 408 keV 以及来自^{83m}Kr 的 40 keV等。通过数据筛选可以得到这些单能峰对应的 S1 和 S2 的分布。图4-4展示了这些单能峰对应的光产额和电产额。对于不同的能量沉积,电产额和光产额是不同的。图中显示,光电产额表现为反线性关系。通过这些单能峰的数据以及拟合直线函数,可以得到 g_1 和 g_{2b} 。直线的斜率和截距分别对应 g_1 和 g_{2b} 。表3-1和3-2中总结了,PandaX-II 以及 PandaX-4T 实验的 PDE, EEE 以及 SEG 值。



图 4-3 PandaX-4T 实验中最小的 S2 的分布,黑色点代表数据,绿色以及红色分别是根据单 电子和双电子峰,用高斯函数来拟合的曲线,红色是全局拟合函数。蓝色为费米狄拉克函数。

4.5 探测器刻度

探测器刻度是暗物质探测实验中非常重要的一个环节。通过刻度,我们可以 了解探测器对不同类型事件的响应,微调光产额、电产额以及涨落等参数。最重 要的是,通过刻度数据,能够建立信号模型(核反冲信号模型)以及本底模型(电 子反冲本底模型)。建立模型时需要用到半经验的物理模型 NEST (Noble Element Simulation Technique) [156-157]。NEST 能够对惰性元素中的闪烁、电离和电致发 光过程进行较准确和全面的模拟。除了在暗物质探测实验中, NEST 还有其他实 验中的应用,包括双 β 衰变搜索、PET 扫描和一般辐射探测技术。NEST 已经根 据过去的许多实验结果,对其模拟结果进行了验证。NEST 是一个独立的 C++ 包, 它生成一个用于链接的编译库(对外部应用程序有用)以及用于快速计算的工具。 此外,用户可以方便地创建自己的自定义探测器,以生成快速模拟结果,同时考虑 各种影响,例如探测器效率、探测器本身参数、位置依赖性、光学和电场变化。目 前 NEST 开发组推出了最新的 NEST 2.0 版本,与以前的版本不同,此版本不需要 GEANT4[158] 或 ROOT 库。然而, 它允许跟 GEANT4 和 ROOT 集成在一起。虽 然现有的 GEANT4 模拟包可以对惰性元素中的闪烁进行简单的计算,但 NEST 半 经验模型提供了更为稳健的光产额和电产额的计算。NEST 还考虑了能量和场的 相关性,以及内在复杂的物理过程。有了 NEST 以及刻度数据,根据 PandaX 实验



图 4-4 PandaX-4T 实验中各类单能峰的 Doke Plot。红线是根据图中几个特征峰点来拟合得到的。

探测器参数,可以对 NEST 模型中的自由参数进行调整,包括对 g₁、g₂ 以及 SEG 等。通过参数调整,可以得到对应的信号模型以及本底模型。无论是信号模型还 是本底模型,模型是根据给定的任何能量沉积谱,能够产生对应的 S1 和 S2 的分 布。换句话说,我们可以利用调整后的模型和任何理论预测的信号能够产生 S1 和 S2 的分布。模型误差可以根据刻度数据跟模型之差算出来。

在2.3.5中提到,在 PandaX 实验中有很多放射性刻度源。PandaX-4T 利用²⁴¹Am-⁹Be 和 DD 来对核反冲事例进行刻度 [159]。其中²⁴¹Am-⁹Be 源是根据如下的 (α, n)反应来产生中子 [160]:

$${}^{9}\text{Be} + \alpha \to n + {}^{12}\text{C}^* \tag{4-4}$$

如4-5图所示,中子有两种衰变通道,分别为基态和第一激发态,它们的比例分别约为40%以及60%。其中第一激发态退激发时会放出4.4 MeV的γ射线。如2.3.5中描述,由于²⁴¹Am-⁹Be 是固体,因此通过安装在探测器外罐内侧的刻度 管道来进行刻度。除此之外,PandaX-4T 实验使用 DD 中子来进行核反冲刻度。 在2.3.5中已提到,DD 中子是通过专门设计的 DD 发生器以及 DD 管道来抵达探测 器内部,中子由以下反应来产生 [161]:

$$D + D \rightarrow {}^{3}He + n$$
 (4–5)

PandaX-4T 实验采用 220 Rn 作为电子反冲刻度源。由于地球的年龄约为 46 亿



图 4-5 根据蒙特卡洛模拟得到的²⁴¹Am-⁹Be 的两种中子能谱,红色是第一激发态能谱,蓝色 是基态能谱。

年,地球上半衰期较短的核素已经衰变殆尽,现在地球上的放射性核素来源于三 个长期平衡的放射系,分别为²³²Th(半衰期约为1.405×10¹⁰年)、²³⁸U(半衰期 约为4.468×10⁹年)和²³⁵U(半衰期约为7.04×10⁸年)。刻度源²²⁰Rn 来自于处 理过且纯度较高的²²⁸Th,而²²⁸Th则来源于长寿命的²³²Th。图4-6展示了²³²Th的 衰变链,可以看出²²⁰Rn 的子核的半衰期较短,最长的也仅有约10小时左右,最 终衰变为稳定的²⁰⁸Pb。在衰变过程中,会产生各种 α 、 γ 和 β 射线,并且在低能 区域的能谱形状是平坦的。通过这种平坦的能谱,我们可以理解电子反冲事例的 各种物理行为。

根据刻度数据拟合 NEST 模型可以得到核反冲信号模型以及电子反冲信号模型。这里所说的拟合指的是通过调整 NEST 模型中的自由参数,使得模拟结果与数据相符。图4-7展示了经过数据筛选后的刻度数据(包括²⁴¹Am-⁹Be, DD 和²²⁰Rn),以及 NEST 模型模拟出来的分布,从图4-7中可以看出 S1 和 S2 的模拟结果与数据相符较好。这说明根据刻度数据调整的 NEST 模型能够较好地描述核反冲或电子反冲事例。NEST 模型的能量区间是根据暗物质反冲能区间来选取的。图4-7左列的三张图是三种刻度数据的 S1 和 S2 的联合分布,其中粉红色(²⁴¹Am-⁹Be 和 DD)或蓝色(²²⁰Rn)实线是整个分布的中位线,虚线是它们的 95% 分位数线,紫色的线表示 99.5% 的核反冲接收线。通过这条线,可以去除一些来自 TPC 表面的极低能本底。从图中可以看出,除了个别事例之外(浅蓝色圆点),所有的电子反



图 4-6 长寿命放射性元素 232Th 的衰变链以及衰变信息。

冲事例几乎都在核反冲事例中位线的上方。这说明核反冲与电子反冲之间的区分 度也较好,这对提高探测器的灵敏度有一定的好处。



图 4-7 调整(或拟合)后的 NEST 模型产生的 S1 和 S2 分布跟各种刻度数据的比较,带有误 差棒的黑色电是数据,红色直方图是由 NEST 模型产生的模拟数据。左边三列图分别是经过 数据筛选后的²⁴¹Am-⁹Be, DD 和 ²²⁰Rn 的 S1 和 S2 的联合分布,其中粉红色线是 DD 或 ²⁴¹Am-⁹Be 数据分布的中位线,虚线是它们的 95% 的分位数线,蓝色的是 ²²⁰Rn 数据的中位 线,对应的虚线是 95% 的分位数线,紫色虚线是 99.5% 的核反冲接收线,浅蓝色点是落在核 反冲中位线下面的事例。

-78-

第五章 暗物质信号区间的本底估计

PandaX 暗物质探测实验中最重要的关键点是对本底的精确估算以及抑制本 底含量,因为精确估计本底对于发现暗物质至关重要。在粒子物理实验中,通常 会剔除不在暗物质能区的数据,例如在暗物质实验中,我们寻找 100 GeV/c² 左右 的弱相互作用粒子,会将能量过高或过低事例视为本底而排除,以提高信号的灵 敏度。在此过程中,人们往往会一边查看数据,一边调整数据筛选条件,以使实 验结果更符合预期结果。然而,这种做法会在调整过程中引入主观因素,从而对 实验结果产生不客观的影响。为避免这种人为因素导致的偏见,需要利用刻度数 据和各种估计方法对数据进行筛选和对本底进行精确估算,而不查看暗物质信号 区间数据的情况下进行分析,这种方法称为盲分析(Blind Analysis)。一旦确定了 筛选条件和本底估计值,就可以查看暗物质信号区间的数据,但不能再对筛选条 件和估算值进行修改,这种做法被称为揭盲(Unblinding)。在 PandaX-II 实验的最 终分析中,我们采用了盲分析方法。在 PandaX-4T 实验中由于第一批数据是在探 测器调试阶段采集,因此为了对探测器进一步了解没有进行盲分析。

在 PandaX 实验中,暗物质信号区间的本底,根据本地来源以及性质,主要分为四种:电子反冲本底、核反冲本底、TPC 表面本底和偶然符合本底。除了偶然符合本底外,本章中对其他的本底来源以及估计方法进行大致介绍。

5.1 电子反冲本底

在暗物质信号区间内,大部分是电子反冲本底。这些电子反冲事例主要分为 两种:一种是具有特殊能谱形状的,例如氚化甲烷(CH₃T,能够产生β射线)以 及¹²⁷Xe(会产生不同特征峰的X射线)等。另一种电子反冲本底是总的能谱形状 较平,例如来源于²³⁸U等长寿命元素衰变链的²²²Rn等,液氙中含有的少量⁸⁵Kr 以及探测器材料本身产生的各类γ射线等。

5.1.1 CH₃T

PandaX-II 实验中 2016 年取完暗物质寻找数据以后,将 CH₃T 作为电子反冲刻 度来注入到探测器内部。刻度结束后,试图利用循环系统来去除它,虽然在此过程 去除了大部分 CH₃T,但是依然有一部分 CH₃T 吸附在在探测器 TPC 内部。为了进 一步去除它,将液氙从探测器回收,且进行了精馏。2017 年再次将液氙重新装回探 测器内。通过数据分析,发现了 CH₃T 的含量降低了 100 倍左右。PandaX-II 实验结

束前,2019年同样地,为了获取电子反冲刻度数据,探测器内注入了 CH₃T,刻度 结束后也是通过循环系统以及精馏系统去掉了大部分 CH₃T。由于循环以及精馏 系统达到了纯化极限,液氙中依然存在着微量的 CH₃T。PandaX-4T 实验中使用的 部分氙来自于 PandaX-II 实验,因此,PandaX-4T 实验也是不可避免的从 PandaX-II 实验引入了 CH₃T 本底。CH₃T 放出的β衰变不是单一的特征峰,它在暗物质能区 能够产生连续能谱的β衰变,因此无法通过其他的间接放法来估算它的含量。如 图5–1所示,只能是通过对暗物质寻找数据进行拟合估计。在估计其他本底以及得 到它们的能谱的情况下,根据暗物质寻找数据来进行自由拟合就能得到 CH₃T 的 形状。



图 5-1 PandaX-4T 实验暗物质能区(0-30 keV_{ee})内的数据能谱,其中蓝色的是 CH₃T 的拟合曲线, 橙色的是其他的电子反冲本底拟合曲线,红色的是总的拟合曲线。

5.1.2 ¹²⁷**Xe**

天然的氙里面有约为 0.09% 的 ¹²⁶Xe,如果 ¹²⁶Xe 遇到中子,它会通过中子俘 获的方式转化为不稳定的 ¹²⁷Xe。通常情况下,液氙运输过程受到宇宙射线的原因, 宇宙线照射材料可能产生中子,因此 ¹²⁶Xe 通过中子俘获转变为 ¹²⁷Xe。由于 ¹²⁷Xe 本身也不稳定,通过电子俘获的方式转变为激发态的 ¹²⁷I(半衰期约为 36 天)。处 于激发态的 ¹²⁷I 通过放出 γ 射线(主要是 203 keV 和 375 keV 的两种 γ 射线),或 内转换电子。除此之外,由于电子俘获的原因,外层轨道的电子会跳到内层空穴 位,从而放出 X 射线或俄歇电子,如图5–2所示。主要放出的特征 X 射线能量分 别为 33 keV 和 5 keV,它们产生的比例分别为 83% 以及 13% 并且数据中能够选出 33 keV 对应的部分。根据 33 keV 和 5 keV 的比例以及它们在探测器分布密度,可以推算出 5 keV 的在低能区域的贡献。图5-3中可以看出,¹²⁷Xe 事例率随时间不断下降,通过指数函数拟合可得半衰期约为 35.86 天,跟理论值基本上符合。



图 5-2 ¹²⁷Xe 通过电子俘获的方式转变为¹²⁷I的示意图,在此过程处于激发态的¹²⁷I 会放出 几种 γ 射线。随后,由于电子俘获的原因,¹²⁷I 会放出 X 射线或俄歇电子。



图 5-3 PandaX-4T 实验中的 127 Xe 事例率随时间变化的趋势。

5.1.3 探测器材料本底

任何一种材料中会有各种放射性元素,包括在 PandaX 探测器内。例如长寿命的²³⁸U,²³²Th,²³⁵U,⁶⁰Co,¹³⁷Cs等。这些元素的衰变链中元素放出的γ射线,在 探测器内会贡献到电子反冲本底。这种本底的一个重要特点是,由于长寿命元素, 半衰期极长,可达数亿年,因此对于探测器来说含量固定,无法通过各种方法去 除。虽然部分的材料本底可以根据一英寸的光电管来识别出来,但是大部分材料 本底是无法区分。一般,在探测器组装之前,我们通过在锦屏地下实验室的高纯 锗探测器来对此进行测量。除此之外,通过 Geant4 对探测器材料的放射性进行的 模拟并且根据测量数据进行拟合,从而得到探测器材料的γ射线在低能区域的贡 献。关于材料本底测量更为详细的内容参见 [162]。

5.1.4 ⁸⁵Kr

通常情况下,在工业中提取的氙中含有部分少量的 ⁸⁵Kr 元素,探测器循环系 统无法去掉它,由于它的沸点跟氙的沸点不同,因此可以通过 PandaX 实验的精馏 系统来除去 ⁸⁵Kr。即使通过精馏系统来去掉大部分 ⁸⁵Kr,由于精馏系统本身存在一定的精馏极限,因此需要对剩余的 ⁸⁵Kr 本底进行估计。⁸⁵Kr 的丰度约为 11×10⁻¹¹,半衰期约为 11 年。它有两种 β 衰变通道,分别为 99.56% 的 680 keV 和 0.43% 的 173.4 keV。这里的 β 衰变导致的反冲能是一个连续的能谱。其中 0.43% 的 β 衰变之后,⁸⁵Kr 会变成 ^{85m}Rb,半衰期为 1 μ s,如图5-4 (左)所示。由于 β 事例和 γ 事例发生的时间间隔只有 1 μ s,这种事例的波形种会有时间间隔约为 1 μ s的两个 S1 信号,如图5-4 (右)所示。因此通过寻找第二种通道造成的 β - γ 符合事例以及根据两种衰变通道的比例,可以确定 ⁸⁵Kr 在探测器中含量。在 PandaX-II 实验最终的分析中总共看到了 49 个 β - γ 偶然符合事例,而在 PandaX-II 实验下降了 10 倍左右。

5.1.5 ²²⁰Rn 和 ²²²Rn

在4.5中介绍到,²²⁰Rn 被当作电子反冲刻度源来使用,它的衰变链中的各个元素的半衰期较短,对后期的采集的低本底数据影响很小,因此它算是一种理想电子反冲刻度源。然而5.1.3中说到几乎所有的材料中都含有²³⁸U,²³²Th 和²³⁵U等长寿命放射性元素,²²⁰Rn 来源于其中的²³²Th 的衰变链,因此探测器必有²²⁰Rn 本底。然而被当作刻度来使用的²²⁰Rn 是通过特殊处理的纯度较高的²²⁸Th。如图4-6所示,来自于²³²Th 的衰变链的²²⁰Rn 本底主要是由它的衰变链上的²⁰⁸Tl 和²¹²Pb 元



图 5-4 ⁸⁵Kr 的衰变示意图(左)以及在 PandaX-4T 实验中观测到的典型的 β - γ 符合事例的波形(右)。

素发射出来的 β 射线产生的。类似于 ⁸⁵Kr, ²²⁰Rn 的衰变链上存在 ²²⁰Rn-²¹⁶Po 的 α - α 符合事例以及 ²¹²Bi-²¹²Po 的 β - α 符合事例,因此 ²²⁰Rn 的含量可以根据这种符 合事例来计算, ²²⁰Rn 含量的系统误差可以根据两种符合事例计算出来的差别来得 到。

跟 ²³²Th 一样, ²³⁸U 也是长寿命放射性元素,其中衰变链中 ²²²Rn 是在暗物质 能区的另外一个 Rn 的本底。它产生的本底主要是来自于衰变链中的 ²¹⁴Pb 产生的 β 射线并且它能衰变为 ²¹⁴Po,随后 ²¹⁴Po 衰变为 ²¹⁰Pb,其半衰期约为 164 μ s。同样 地,通过 ²¹⁴Bi-²¹⁴Po 的 β-α 符合事例,能够估计 ²²²Rn 的含量。关于 Rn 本底详细 内容参见 [163]。

5.2 核反冲本底

PandaX 暗物质探测实验中核反冲本底含量相对于电子反冲事例较小,但是核 反冲本底跟暗物质信号具有相同的性质,它们都是核反冲事例,因此核反冲本底 的压制以及精确估计直接影响到探测器的灵敏度。探测器中的核反冲本底主要分 为两种,中子本底以及 ⁸B 太阳中微子本底。下面对这两种核反冲本底进行介绍。

5.2.1 中子本底

优越的实验室(中国锦屏地下实验室)环境,使得探测器远离宇宙射线的影响 并且水系统等基础设施能够避免岩石表面的放射出来的中子达到探测器内部。然 而上文中提到的²³²Th,²³⁵U和²³⁸U等长寿命元素产生的α粒子能够在探测器中

第五章 暗物质信号区间的本底估计



图 5-5 长寿命放射性元素 ²³⁸U 的衰变链以及衰变信息。

的其他材料中通过 (α , n) 反应产生中子。 α 粒子打到能产生中子的大部分材料的 原子序数较小,例如 PMT 中的 Al 元素以及 PTFE 中的 F 元素。除此之外,中子还 可以通过 ²³⁸U 的自发裂变的放射被发射出来,如图5-6所示。中子信号跟暗物质 物质信号最主要的区别是,中子有可能发生多次散射。利用中子的这种特点能识 别出部分中子。中子的估计方法大体上分为两种;一种是基于蒙特卡洛模拟 [164], 另一种是基于数据。对于蒙特卡洛模拟方法来说,首先探测器安装之前,通过在 锦屏地下实验室的高纯锗探测器来对探测器材料进行测量,随后它的放射性活度 跟能谱形状之间关系是通过 Sources-4A 来得到 [165]。最终将得到的能谱形状作 为输入,通过 Geant4 模拟就能得到中子本底的含量。这种方法的一个缺点是过于 依赖于模拟,无法预测真正的误差。其中 Sources-4A 对中子的处理较简单,尤其 是在发生 (α , n) 反应时只考虑中子,而忽略此过程产生的 γ 射线,即使在 Geant4 模拟中改进了此缺陷,但是无法预测估计值的有效误差。因此需要通过另外一种 方式是来估计中子本底,即基于数据的估计方法。当中子产生且能量沉积后,有可 能被氙的同位素 ¹³¹Xe 来俘获,从而转变为 ¹³²Xe。由于 ¹³²Xe 不稳定处于激发态, 通过放出高能 γ 射线而退激发,因此产生高能 γ 事例信号。除此之外中子第一次
能量沉积后有可能继续多次能量沉积,从而产生多次散射事例信号。利用中子在数据中的这种特征可以选出对应的信号(多次散射信号或者高能γ信号)。值得一提的是,无论是多次散射或高能γ事例,它们跟单次散射的比例是通过 Geant4 模拟来得到的。因此最终的中子本底是根据以下公式来算的:

$$n_{\text{neutron}} = \frac{n_{\text{feature}}}{r_{\text{MC}}} \tag{5-1}$$

其中 *n*_{feature} 是在数据中筛选出来的多次散射事例或者高能 γ 事例数目, *r*_{MC} 是 通过 Geant4 蒙特卡洛模拟来得到的单词散射事例跟多次散射事例或高能 γ 事例 的比例, *n*_{neutron} 是最终算出来的中子本底的估计值。中子本底的系统误差可以根 据基于蒙特卡洛模拟方法和基于数据的方法的差别来算出来。关于 PandaX 中子 本底更为详细内容请参见 [166-167]。



图 5-6 探测器材料中子本底的产生机制, (a) 图是通过 (α, n) 反应产生中子本底的机制, (b) 图是通过自发裂变的方式产生中子本底的机制。

5.2.2 ⁸B 太阳中微子

根据标准太阳模型可知,由于太阳内部复杂的聚变过程,大量的中微子不断 地产生。近年来,XENON、PandaX 等暗物质直接探测实验的探测器靶物质已经发 展到了吨级,根据理论模型,它们现在的探测器灵敏度能够达到了通过中微子-核 相干弹性散射来探测太阳中微子的程度。由于 PandaX-II 实验探测器体量相比于 PandaX-4T 实验探测器较小,其灵敏度也较弱,因此在 PandaX-II 暗物质寻找分析 中可以忽略它的含量。而在 PandaX-4T 实验中中微子是不可忽略的本底。如图1– 3中所示,在 pp-chain 太阳中微子的所有来源中,由于 15MeV 的Q值,在⁸B 的β 衰变中产生的中微子最有可能被探测到:

$${}^{8}B \rightarrow {}^{8}Be^{\star} + e^{+} + v_{e} \tag{5-2}$$

根据几个中微子探测实验的测量,地球上⁸B 太阳中微子的通量约为 5×10⁶ cm⁻²s⁻¹[168-169]。与暗物质事例率计算类似,中微子事例率是通过中微子-核子散射截面与中微子通量的卷积来计算 [170]:

$$\frac{dR_{\nu}}{dE_{r}} = MT \times \sum_{A} f_{A} \int_{E_{\nu}^{\min}} \frac{dN}{dE_{\nu}} \frac{d\sigma \left(E_{\nu}, E_{r}\right)}{dE_{r}} dE_{\nu}$$
(5-3)

其 *M* 为靶物质质量, *T* 为采集数据时长, f_A 为靶物质元素的同位素分数, E_v 为中微子动能, E_r 为反冲能, $\frac{dN}{dE_v}$ 为中微子通量, $\frac{d\sigma(E_v,E_r)}{dE_r}$ 为微分中微子-核散射截面。根据计算可知,将氙作为靶物质,核反冲能阈值为 1 keV 探测器,采集 1 吨 ·年的数据之后,将通过相干中微子-核散射能够探测到大约 100 个 ⁸B 的太阳中微 子事例 [170]。

图5-7所展示,各类太阳中微子在液氙中的核反冲能谱,在大于 0.1 keV 区域, 主要是由⁸B 的能谱占主导。不难看出,中微子-核相干弹性散射中产生的核反冲信 号极其微弱,大部分信号主要在 1 keV 以下。这种微弱核反冲信号跟低质量暗物质 (质量为 6 GeV/c²,散射截面为 4.7×10⁻⁴⁵ cm⁻²)的反冲信号是无法区分的 [170]。 根据 PandaX-4T 探测器目前的灵敏度来说,理论上可以探测 ⁸B 太阳中微子在液氙 中的相干弹性散射信号。因此寻找暗物质分析中需要考虑它的含量。利用图5-7中 的 ⁸B 太阳中微子核反冲能谱以及 NEST 模型,最终在 PandaX-4T 实验中算出来的 ⁸B 太阳中微子本底含量约为 0.6 ± 0.3 个。



图 5-7 8B 太阳中微子通过中微子-核相干弹性散射在液氙中产生的核反冲能谱 [170]。

5.3 表面本底

在 PandaX 实验中,另外一种比较重要的本底是表面本底。所谓的表面指的 是 TPC 内部的 PTFE 材料。4.3中提到,物理事例的能量沉积后电子往上漂移的 过程中受到电负性杂质的影响电子的数量逐步减少,如果事例发生在 TPC 内部的 PTFE 材料附近,有一部分电子会吸附在 PTFE 材料,最终达到气氙的电子数量会 更少。虽然这种事例通过在3.9中提到的对事例水平方向位置的数据筛选方式排除 掉,但是有可能极少部分事例落入到暗物质能区,尤其是位置重建算法对低能事 例带来的不确定性会导致部分的表面本底被选为暗物质候选事例。其中,来源于 ²²²Rn 的 ²¹⁰Pb 发出的 β 射线(能量为 65 keV)和 γ 射线(能量为 47 keV)极有可 能产生表面事例。PTFE 表面吸附的还有来源于 222Rn 的 210Po, 但是 210Po 能放出 能量为 5.3MeV 的 α 射线,因此数据中可得到该事例的分布(位置分布以及 S1 和 S2 信号的分布) 且能够构造概率密度分布。利用这种概率密度分布可以估计落入 暗物质能区的表面本底的含量。如5-8图中展示,²¹⁰Po 事例对于不同水平位置 R $(R = \sqrt{x^2 + y^2})$,其中 x 和 y 是事例的水平面的坐标),S1 和 S2 信号大小的二维概 率密度分布也不一样,除此之外,大部分表面本底在极低能区,尤其是聚集在核 反冲事例 S1 和 S2 信号大小二维分布 99.5% 分位数线的下面。因此表面本底的有 效去除且精确估计对于提高探测器低能区域的灵敏度至关重要。关于表面本底详 细讨论参见 [171]



图 5-8 ²¹⁰Po 事例对于不同的水平位置 R ($R = \sqrt{x^2 + y^2}$,其中 x 和 y 是事例的水平面的坐标) 的 S1 和 S2 信号大小的二维概率密度分布,其中 (a) 是当 R=515mm 时的概率密度分 布,(b) 是当 R=600mm 时的概率密度分布。

5.4 本底总结

以上除了偶然符合本底以外,对 PandaX 暗物质信号区间内所有本底进行了 总体介绍。对于偶然符合本底,在下一章进行详细介绍。

表5-1和5-2中汇总了 PandaX-II 以及 PandaX-4T 实验暗物质信号区间内经过 所有数据筛选之后的总本底含量[172-173]。不难看出,大部分本底是来自于电子 反冲事例,其 PandaX-II 实验中在核反冲中位线以下,接近一半的贡献来自于材料 表面本底。表5-1最后两列和5-2的最后两行结果来看,总的估计值跟实际观测数 据在误差范围内基本一致。表5-1中的最佳拟合值是根据暗物质信号+本底信号假 设下拟合得到的结果,其暗物质质量假设为400 GeV/c²。表5-2中的最佳拟合值是 根据仅本底假设下得到的。总结以上,在暗物质信号区间内没有观测到明显超出 预期的疑似暗物质事例。因此,根据已知的本底以及信号模型,对暗物质的物理参 数空间进行限制。一般来说,在粒子物理实验领域,这种限制是通过统计学中的假 设检验来进行。其采用的统计检验量为剖面似然比(Profile likelihood ratio, PLR) 方法 [174]。这是在粒子物理或高能物理领域内非常主流的统计检验量(在7.4中, 对此展开详细讨论)。图5-9展示的是暗物质质量跟原子核自选无关的弹性散射界 面的排除线,其中包括其他的暗物质探测实验的结果。它的物理意义是在90%的 置信水平下,对暗物质质量和它跟核子的弹性散射截面的二维参数空间进行排除, 换句话说,图中的红色线上方是已被间实验数据结果排除掉的二维参数空间,只 能从红色或黑色线下面寻找暗物质。由于不同实验的具有不同的本底水平以及探 测器体量,因此通过它们得到的结果有所不同。

上海交通大学博士学位论文

		A 1 1 1		0 0	T + 1 C++ 1	TT (1 1 1
	EK	Accidental	Neutron	Surface	Total fitted	Iotal observed
Run 9	381.5	2.20	0.77	2.13	387 ± 23	384
Below NR median	2.7	0.46	0.37	2.12	5.6 ± 0.5	4
Run 10	141.7	1.08	0.48	2.66	145.9 ± 16	143
Below NR median	1.7	0.24	0.22	2.65	4.8 ± 0.6	0
Run 11, span 1	216.5	1.04	0.60	6.24	224 ± 22	224
Below NR median	4.2	0.32	0.32	6.22	11.1 ± 1.1	13
Run 11, span 2	448.2	1.60	0.92	9.58	460 ± 35	469
Below NR median	8.26	0.50	0.50	9.54	18.8 ± 1.7	21
Total	1187.9	5.9	2.77	20.6	1217 ± 60	1220
Below NR median	16.8	1.52	1.42	20.5	40.3 ± 3.1	38

表 5-1 在 Run 9, 10, 11 数据中的各个本底的估计值以及通过数据筛选之后的寻找暗物质数据的事例数汇总,其中的拟合值是根据暗物质的质量为 400 GeV/c² 的信号分布来拟合的结果。其单位是,事例数。



图 5-9 暗物质质量跟原子核自选无关的弹性散射界面的排除线。其中除了 PandaX 之外,还 包括其他的暗物质探测实验的结果 [172-173,175-176]。

Rn	6.9 ± 3.8	42.8 ± 23.5	22.7 ± 12.5	162.0 ± 88.9	112.1 ± 61.5	346.5 ± 190.2	1.42 ± 0.78	ı
Kr	$1.1 {\pm} 0.7$	$7.7{\pm}4.9$	3.2 ± 2.1	20.4 ± 13.1	20.9 ± 13.4	53.3 ± 34.2	0.21 ± 0.13	I
Material	$0.8 {\pm} 0.1$	$5.7 {\pm} 0.7$	$2.4{\pm}0.4$	15.2 ± 1.9	$15.6 {\pm} 1.9$	$39.7{\pm}5.0$	$0.16 {\pm} 0.02$	I
solar v	$0.8 {\pm} 0.2$	5.4 ± 1.1	2.3 ± 0.5	14.3 ± 2.9	14.6 ± 2.9	37.4 ± 7.5	$0.16 {\pm} 0.03$	I
¹³⁶ Xe	$0.7{\pm}0.1$	4.6 ± 0.9	$1.9{\pm}0.4$	11.8 ± 2.4	12.1 ± 2.4	31.1 ± 6.2	0.05 ± 0.01	I
Flat ER (data)	4.0±2.9	54.5 ± 10.5	12.2±4.9	240.5 ± 21.8	180.9 ± 18.9	492.1±31.2	2.06 ± 0.14	509.6 ± 22.8
CH_3T	17±5	88±11	21±6	258 ± 24	148 ± 17	532±32	5.1 ± 0.3	532±32
¹²⁷ Xe	0.19 ± 0.04	1.08 ± 0.25	0.96 ± 0.22	3.99 ± 0.92	1.91 ± 0.44	8.13 ± 1.07	0.12 ± 0.02	8.41 ± 2.08
Neutron	0.02 ± 0.01	$0.15 {\pm} 0.08$	0.07 ± 0.03	0.45 ± 0.22	0.46 ± 0.23	1.15 ± 0.57	0.69 ± 0.35	0.82 ± 0.41
${ m B}_8$	0.01 ± 0.01	0.05 ± 0.03	0.03 ± 0.02	0.26 ± 0.13	0.29 ± 0.15	0.64 ± 0.32	0.62 ± 0.31	0.61 ± 0.17
Surface	$0.01 {\pm} 0.01$	0.07 ± 0.02	0.03 ± 0.01	$0.18 {\pm} 0.05$	$0.18{\pm}0.05$	0.47 ± 0.13	0.42 ± 0.12	$0.44{\pm}0.11$
Accidental	0.04 ± 0.01	$0.32 {\pm} 0.05$	0.03 ± 0.01	$0.99 {\pm} 0.18$	1.05 ± 0.21	2.43 ± 0.47	$0.80 {\pm} 0.15$	2.31 ± 0.45
Sum	21±6	144 ±15	34±8	504±32	333 ± 25	1037 ± 45	9.8 ±0.6	1054 ± 39
Data	21	148	34	496	359	1058	6	

据仅本底假设下拟合得到的。其单位是,事例数。

第五章 暗物质信号区间的本底估计

第六章 偶然符合本底

在前一章中,除了偶然符合本底外,介绍了暗物质信号区间内的所有本底。这些本底的共同特点是它们都有明确的物理来源。换句话说,这些本底的 S1 和 S2 信号都源自于同一个物理事件。与这些本底不同的是,PandaX 暗物质探测实验中还有一种本质不同的本底,即偶然符合本底。偶然符合本底本质上是由一个孤立的 S1 信号和一个孤立的 S2 信号偶然地出现在同一个事例波形窗口而形成的本底。换言之,形成偶然符合本底的 S1 和 S2 信号并不来自同一个物理来源,因此它是一种非物理来源的本底。偶然符合本底是探测器里最主要的四种本底之一,因此通过精细估计它的含量以及排除它,能够有效提高探测器灵敏度。

6.1 偶然符合本底的来源

偶然符合本底是由孤立的 S1 信号和孤立的 S2 信号组成的。一旦这些孤立信 号出现在同一个波形窗口中,它们很可能被软件重建为正常的物理信号,如果该 事例落在暗物质能区,极有可能被误判为暗物质信号。因此,有效去除且准确估 计偶然符合本底,对于提高探测器的灵敏度至关重要。

为了去除或识别偶然符合本底,首先需要了解形成偶然符合本底的孤立 S1 和 S2 信号的来源。对于孤立 S1 信号的来源来说,它本身可能是物理来源或非物理 来源,以下是孤立 S1 信号可能的物理来源:

- 在 TPC 电极上的打火事例,由于电极上的电压高,容易产生打火事例,这些打火事例被软件误判为 S1 信号,但是它一般不产生电子;
- 在阴极和底部 PMT 阵列的屏蔽电极之间的物理事例(图6-1中的区域 2), 由于电场方向向下,因此电子不能漂移到气氙中,因此不能产生 S2;
- 能量沉积很小的物理事例,因为产生的电子特别少而且电子可能电负性杂 质吸附,因此电子无法达到气氙,也不会产生 S2;
- 发生在阳极上方的气氙(图6-1中的区域5)的物理事例,由于该区域没有 电场,因此无法产生S2;
- 单电子信号有可能被软件错误的识别为 S1 信号;
- 在 TPC 外部发生的高能物理事例产生的 S1 或 S2 信号的光有可能漏进 TPC 内,产生假的 S1 信号 (图6-1中的区域 6);

孤立 S1 的主要非物理来源是 PMT 的暗噪声。暗噪声是 PMT 在运行中随机产生的微弱噪声信号,它能产生非常微弱的信号,在每个 PMT 中都会出现暗噪声,

因此它们有一定概率偶然符合,在事例重建过程中可能形成可观的 S1 信号。

对于孤立 S2, 主要以下四种可能的来源:

- 在探测器核心探测区域(图6-1中的区域1)发生的真实物理事例,由于探测效率的原因,微弱的S1没有被探测到或软件识别;
- 在探测器核心探测区域(图6-1中的区域1)发生的真实物理事例,由于离液氙表面太近,导致S1和S2信号叠在一起,被识别为一个S2;
- 在门电极上面且在阳极下面的区域(图6-1中的区域3和4)中发生的物理 事例,如果在区域4中发生,只有S2信号,如果在区域3中发生,S1和S2 信号叠在一起被识别为一个S2,这种信号可能具有更小的波形宽度和不对 称的波形形状;
- 高能物理事例(MeV 量级)的产生的电子可能没有同时拽出到气氙区域,部 分电子可能聚集在液体表面,并逐步释放到气体中,产生较小的 S2 信号。



图 6-1 PandaX 暗物质探测器 TPC 示意图,可能产生孤立 S1 和孤立 S2 信号的区域用红色数 字来标记。

6.2 偶然符合本底的估计方法

偶然符合本底的估计方法相对来说比较直观。前文提到,偶然符合本底是由 孤立的 S1 和 S2 信号组成的。因此,一旦获得它们各自的发生频率,就可以根据 以下计算公式估算偶然符合本底的含量:

$$n_{\rm acc} = \bar{r}_1 \cdot \bar{r}_2 \cdot \Delta t_w \cdot T \cdot \epsilon. \tag{6-1}$$

为了计算孤立 S1 和孤立 S2 信号发生频率,我们将每个数据集(Run9,10 和 11) 切分为等同时长的子数据集,并对每个子数据集的孤立 S1 和 S2 信号单独计算它 们的发生频率,这些计算得到的发生频率的平均值作为各个数据集中信号发生频 率的估计。其中 \bar{r}_1 和 \bar{r}_2 分别为孤立 S1 和孤立 S2 的平均发生频率, Δt_w 是电子在 液氙中的最大漂移时间,*T* 是暗物质寻找数据集的采集时长, ϵ 是数据筛选效率。

6.3 在 PandaX-II 实验中的偶然符合本底

下面对在 PandaX-II 实验中的偶然符合本底进行单独讨论。

6.3.1 孤立 S1 和孤立 S2 信号的选取

为了精确计算孤立 S1 信号的出现频率,需要首先正确识别数据中的这种类型信号。在 PandaX-II 实验中,采用了三种方法来选取孤立 S1 信号并计算其频率。 其中,第一种方法是基于特殊类型的"随机触发"数据集来进行选取,而另外两种方法则基于暗物质数据(极低本底情况下采集的数据)进行选取。下面详细描述这三种方法。

在 PandaX-II 实验中,通常情况下,当 PMT 探测到信号并且信号通过提前设定的信号阈值时,数据获取系统将以该信号开始的时间为起点,记录该信号前后500µs 的窗口(所有事例的波形窗口大小固定且为 1ms)。这是 PandaX-II 实验中的标准采集数据模式。此外还有一种采集数据的模式是"随机触发"模式。在这种采集模式下,数据获取系统并不会设定信号阈值,而是在采集数据的过程中随机记录 1ms 的窗口。之后,从这种模式下采集的数据中选取只有 S1 信号的事例并进行筛选(筛选方式与暗物质寻找数据筛选方式相同),其发生频率为:

$$r_1 = \frac{n_{iS1}}{T},$$
 (6–2)

其中 *r*₁ 为孤立 S1 信号的发生频率, *n*_{iS1} 为通过筛选条件的孤立 S1 信号的数 目, *T* 是数据采集时长。

<u>-93</u>

第二种选取方法是通过暗物质寻找数据来进行的。在暗物质寻找数据中有一部分只有 S1 信号的事例,这种事例被 S1 信号触发并且其前后记录 500μs,然而,后 500μs 窗口中没有出现 S2 信号。其孤立 S1 是从前 500μs 的窗口中寻找(如 图6-2中的 Δt_{AB})。为了避免触发事例的 S1 (S1_{max})跟孤立 S1 有物理关联性,从触发事例的 S1 信号的开始时间到往前 10μs 窗口不允许被寻找孤立 S1 信号。通过第二种方法来计算孤立 S1 信号出现的频率为:

$$r_1 = \frac{n_{iS1}}{n_{tS1}} \cdot \frac{1}{\Delta t_{AB}},$$
 (6-3)

其中 n_{iS1} 为被选取的孤立 S1 信号的数目, n_{tS1} 为触发事例的 S1 信号的数目, Δt_{AB} 为选取孤立 S1 信号的时间窗口。



图 6-2 在 PandaX-II 实验选取孤立 S1 的第二种方法的示意图。

为了验证孤立 S1 信号跟触发事例的 S1 信号间是否有关联性,可以通过它们 间的时间差分布来判度。如图6-3所示事例数目随着时间差的变大而变少,尤其是 在数据 Run 11 较明显。这说明它们间存在某种物理的关联性。有一种可能性是, 这部分事例是来自于在阴极下面发生的²¹⁴Bi-²¹⁴Po 符合事例。由于这部分只能存 在 S1 信号,因此看到事例数随着时间差指数下降的趋势,通过指数函数来拟合得 到的半衰期约为 173.59 ± 12.53 μs,这结果跟²¹⁴Po 的理论半衰期 164 μs 基本符 合。因此直接采用这种方法来得到的结果会导致孤立 S1 的发生频率高估,因此去 除²¹⁴Bi-²¹⁴Po 产生的孤立 S1 部分才能得到真正的孤立 S1 的发生频率。

第三种选取方法也是根据暗物质寻找数据来进行的。选取孤立 S1 信号之前, 首先通过数据筛选的方式选取一批真实的物理事例,除了在3.9中提到的质量筛选 条件之外,对 S1 和 S2 信号的大小有额外的条件,即 S1>100PE 且 S2>10000PE。 主要的原因是为了跟最终使用的暗物质寻找数据区分开。其后,从这种提前选出 来的事例的前 500µs 窗口中寻找孤立 S1 信号,同时要求,物理是事例的 S2 信号 跟孤立 S1 的距离足够远或大于等于电子的最大漂移时间,如图6-4所示。通过这



图 6-3 在 PandaX-II 实验中,选取孤立 S1 的第二种方法来选取的孤立 S1 跟触发事例 S1 的时间差分布。

种选取方式可以保证选出来的孤立 S1 跟物理事例的 S2 信号间没有物理关联性。 其孤立 S1 的出现频率可以根据以下公式来计算:

$$r_1 = \frac{n_{iS1}}{\sum (t_{S2} - \Delta t_{A2})},$$
(6-4)

其 *n*_{iS1} 是选取的孤立 S1 的数目,如图6-4所示 *t*_{S2} 是物理事例的 S2 信号的开始时间, Δ*t*_{A2} 是物理事例的开始时间与电子的最大漂移时间的时间差。



图 6-4 在 PandaX-II 实验中,选取孤立 S1 的第三种方法的示意图。

同第二种方法一样,为了验证物理事例的 S1 和孤立 S1 信号间是否具有关联

性,可以通过它们的时间差分布来判度。如图6-5所示,它们间的时间差分布在 0-350 μ s 之内随着时间差逐步增长,350-500 μ s 之内的分布整体较平。对于 Run 9,最大的电子漂移时间是 350 μ s,而对于 Run 10 和 Run 11 来说它的值为 360 μ s (见表3-1)。由于物理事例的 S2 信号跟孤立 S1 信号间的时间间隔被电子的最大漂移时间限制,因此可知,这种限制条件导致离物理事例 S2 近的地方孤立 S1 信号较少,远的地方较多,而超过最大电子漂移时间的地方分布较平。这种分布特征可以根据蒙特卡洛模拟来得到。具体的模拟方法是,首先从物理事例的 S1 信号后 500 μ s 内随机抽样物理事例的 S2 信号,然后根据最大漂移时间的限制条件从物理事例 S1 信号前 500 μ s 内抽样,这样可以得到如图6-5中的蒙特卡洛模拟结果 (黑色直方图)。假设孤立 S1 信号的发生频率为 500Hz,而模拟结果是 498.9Hz,基本符合理论预期。对于 Run 10 来说,由于采集数据时间较短的原因(见表3-1),这种特征不是很明显。在 Run 11 中 Δt < 120 μ s 的区域,时间差分布不符合模拟结果,孤立 S1 信号的出现频率相对于模拟预期超出了 11.6%。对这部分事例单独查看了它们的漂移时间分布(物理事例的漂移时间)。如图6-6所示,这部分的事例发生在阴极附近,对此具体的物理来源暂时没有结论。



图 6-5 在 PandaX-II 实验中通过第三种选取孤立 S1 的方法来选取的孤立 S1 跟物理事例 S1 的时间差分布,其中(a)为原始的事例分布,(b)为归一化后的分布。

孤立 S2 信号的选取方式相当于来说较直观,在将所有的事例筛选条件加上的 条件下,直接将被 S2 信号触发的且没有跟 S1 配对的事例选取即可。它的发生频 率为:

$$r_2 = \frac{n_{iS2}}{T},$$
 (6–5)

其 r₂ 为孤立 S2 信号的发生频率, n_{iS2} 为通过筛选条件的孤立 S2 信号的数目, T 是采集数据的时长。



图 6-6 在第三种选取孤立 S1 的方法来选取的物理事例的漂移时间分布,这里选取的物理事 例是孤立 S1 跟物理事例 S1 的时间差小于 120µs 的部分。

为了计算孤立 S1 和孤立 S2 信号发生频率以及它们的误差,我们将每个数据集(Run9,10 和 11)切分为等同时长的子数据集,并针对依照前述方法在每个子数据集中分别处理。这些计算得到的发生频率的平均值作为各个数据集中信号发生频率的估计,其无偏差的标准差作为估计值的误差:

$$\bar{r}_1 = \frac{1}{\sum_i T_i} \sum_i r_{1i} \cdot T_i, \qquad (6-6)$$

$$\bar{r}_2 = \frac{1}{\sum_i T_i} \sum_i r_{2i} \cdot T_i, \tag{6-7}$$

其 \bar{r}_1 和 \bar{r}_2 为孤立 S1 和孤立 S2 信号的平均发生频率, T_i , r_{1i} 和 r_{2i} 分别为 子数据集的时长, 孤立 S1 和孤立 S2 信号的发生频率。表格6–1中总结的是, 在 PandaX-II 实验中, 不同数据集的孤立 S1 和孤立 S2 信号的平均发生频率以及它们 的误差。

从表6-1中的结果可知,数据集 Run 9 的孤立 S1 的发生频率较高。其中的主要原因是在 Run 9 的数据采集期间,由于 PMT 的增益相对来说较高,PMT 的暗噪声的频率也较高,因此导致 Run 9 的孤立 S1 信号发生频率相比于其他两个数据集来说高三倍左右。对于 Run 9,10,和11,每个 PMT 的平均的暗噪声频率分别为 1.9,0.17 以及 0.23 kHz。由于 Run 9 采集数据期间没有采集随机触发模式的数据,因此对于 Run 9,没有对应的的结果。对于 Run 10 和 Run 11,通过第一种和第三种方法来计得到的孤立 S1 信号的发生频率在误差范围内比较一致。随机触发

Dun	Duration [days]		r. [H7]		
Kuli D	Duration [days]	Method 1	Method 2	Method 3	
9	79.6	-	1.40 ± 0.83	1.53 ± 0.69	0.0121 ± 0.0007
10	77.1	0.46 ± 0.28	0.27 ± 0.27	0.47 ± 0.08	0.0130 ± 0.0025
11	244.2	0.77 ± 0.31	0.37 ± 0.61	0.69 ± 0.38	0.0121 ± 0.0004

第六章 偶然符合本底

表 6-1 在 PandaX-II 实验中,不同数据集的孤立 S1 和孤立 S2 信号的平均发生频率以及它们的误差。

模式来采集的数据时长比较短,因此它引入的误差相对来说也较大。值得注意的 是,由于第二种方法中对孤立 S1 信号的发生频率存在高估问题,因此在表格中的 结果是均为修正后的结果。孤立 S2 信号的发生频率在三个数据集中都比较稳定。 最终的统计推断分析中,我们采用了第三种方法的结果。



图 6-7 第三种方法来计算的孤立 S1 信号以及孤立 S2 信号发生的频率随着时间变化的趋势。

图6-7中显示第三种方法来计算的孤立 S1 信号以及孤立 S2 信号发生的频率 随着时间变化的趋势。总的来看,孤立 S2 信号发生的频率随着时间比较稳定。而 孤立 S1 信号发生的频率在某些采集数据的阶段出现异常,例如 Run 9 数据集中, 频率随着时间的变化较大。Run 9 采集数据期间的不稳定的电极电压或 PMT 的随 机打火可能导致这种不稳定的变化趋势。在 Run 11 数据采集期间的频率不稳定性 来自于 PMT 本身,如图6-8所示,其中编号为 11001,11002 以及 11003 的 PMT 的对此有主要的贡献。

如图6-9中显示, 孤立 S1 信号的大小分布主要聚集在小于 10PE 的部分。然而



图 6-8 2018 年 3 月 11 日至 2018 年 4 月 6 日期间的孤立 S1 信号电荷在顶部 PMT 上的分布。

数据集 Run 9 的孤立 S1 的大小分布在小于 6PE 的地方明细比其他的数据高了一个数量级。如同上文中提到,由于 Run 9 的 PMT 的增益较大,因此这部分主要是 PMT 的暗噪声偶然符合造成的,而且暗噪声偶然符合形成的 S1 信号号比物理事 例引起的 S1 信号小。对于 Run 11,在 10PE 附近的峰同样地来自于不稳定的 PMT (编号为 11001、11002 以及 11003) 打火。



图 6-9 第三种方法来得到的孤立 S1 (a) 和孤立 S2 (b) 信号的大小分布。

6.3.2 偶然符合本底的模拟

选取孤立 S1 和孤立 S2 信号事例之后,可以进行基于数据的蒙特卡洛模拟并 且能得到偶然符合本底的模拟样本。偶然符合本底的模拟方法相对来说较直观,通 过将孤立 S1 和孤立 S2 信号随机配对的方式来可以得到偶然符合本底的样本,随 机配对的同时,对每一个模拟事例分配随机的漂移时间,水平方向的位置是由孤立 S2 信号来定。除此之外模拟过程中对孤立 S1 和孤立 S2 信号进行修正,同时通过 在3.9中提到的跟漂移时间相关的数据筛选条件来对模拟数据进行筛选。表6–2中 的结果是根据公式6–1计算得到的各个数据集中的偶然符合本底的含量。由于 Run 9 数据集中,孤立 S1 信号的发生频率较高,因此相比于其他数据集,在 Run 9 中 偶然符合本底的含量也较高。其中 ϵ 为数据筛选效率。

上文中多次提到,暗物质信号是核反冲事例,如果一个事例在核反冲中位线 下面,说明这种事例的暗物质信号的可能性更高,因此对在表格6-2中专门对核反 冲中位线下面部分的偶然符合事例,单独计算了它们的含量。

Run	Туре	ϵ	nacc
0	total	21.9%	8.15 ± 3.70
9	below NR median	3.5%	1.31 ± 0.59
10	total	25.6%	3.16 ± 0.81
10 t	below NR median	8.5%	1.06 ± 0.27
11	total	18.2%	9.87 ± 5.45
11	below NR median	5.6%	2.93 ± 1.62

表 6-2 在 PandaX-II 实验中,经过所有数据筛选之后的, Run 9、10 和 11 的偶然符合本底含 量。其单位是,事例数。

图6-10(a、b、c)是通过蒙特卡洛模拟方法来得到的 S1 和 S2 信号的联合分 布。图6-10(d、e、f)是经过同样的数据筛选之后的核反冲刻度数据的 S1 和 S2 信号的联合分布,其中红色线是中位线。在 S1 和 S2 的联合分布中可以看出,由于孤立 S1 和孤立 S2 的信号特征,大部分的偶然符合本底聚集在 S1 和 S2 较小的 区域,跟核反冲事例的有一定的交接部分。因此压制偶然符合本底能够有效的提高探测器灵敏度。

6.3.3 偶然符合本底的抑制

偶然符合本底跟其他本底之间最主要的区别是偶然符合本底的 S1 和 S2 信号 之间没有物理的关联性,尽管模拟得到偶然符合样本经过了所有的数据筛选条件,



图 6-10 通过蒙特卡洛模拟方法来得到的 S1 和 S2 信号的联合分布,其中 a, b, c 分别对应的 是 Run 9, 10, 11 的偶然符合本底, d, e, f 分别对应的是 Run 9, 10, 11 的的核反冲刻度数据。

但还是有一定含量的偶然符合本底没有被完全去除。这种情况下,不能光用 S1 和 S2 信号的联合分布来排除更多的偶然符合本底。因此这里需要用多变量分析来更 有效的去除偶然符合本底。在 PandaX-II 实验中为了去除更多的偶然符合本底,我 们使用了增强型决策树(Boosted Decision Tree, BDT)[177]。决策树(Decision Tree)是一种常见的分类和回归方法,即是一种机器学习算法,用于通过一系列 预测问题来预测一个目标变量的值。决策树是一种树形结构,其中的每个非叶子 节点代表一个特定的预测问题,而叶子节点代表预测的结果。为了决定预测的结 果,决策树从根节点开始,根据每个预测问题的答案向下遍历,直到到达叶子节 点。决策树的优点是它易于理解和解释,同时还具有很高的分类精度。决策树的 缺点是如果遇到统计涨落较大的数据集会容易过拟合。增强型决策树是对决策树 的一种改进,它通过对多个决策树的预测结果进行组合,以提高分类精度。增强 型决策树是通过训练一系列的决策树来实现的,每棵决策树都在上一棵决策树的 基础上进行训练,并针对其预测错误的样本进行改进。最终,所有决策树的预测 结果进行加权组合,以得到最终的预测结果。增强型决策树的优点是它可以有效 地避免过拟合的问题,同时具有很高的分类精度,而且比决策树的计算复杂度要 低。在二分类问题上,增强型决策树的有较大优势,尤其是对于在高能物理领域 的信号本底的区分问题经常使用增强型决策树。

关于增强型决策树,有诸多开源且现成的软件框架。其中 Scikit Learn 是目前 最为主流的用 Python 来写的开源机器学习框架 [178]。它提供了许多机器学习算

法,包括决策树、还有支持向量机、最近邻,朴素贝叶斯以及随机梯度下降等算法。它的接口易于使用,并且提供了大量的文档和示例,方便初学者学习。然而在高能物理领域更为常用的开源的机器学习框架是TMVA(Toolkit for Multivariate Data Analysis)[177]。它是一种为用户提供多种分类器和回归器的工具,并且支持多种数据输入格式。TMVA 是 ROOT 的一部分,支持 C++和 Python 编程语言。它提供了多种分类器,如,决策树、神经网络、支持向量机、Boosting 和 Bagging等,以及多种回归器,如,回归树和神经网络。TMVA 具有强大的性能评估功能,允许用户通过交叉验证、训练和测试数据集和多个性能指标评估分类器的准确性。此外,TMVA 支持将分类器的输出与其他算法或工具结合,以提高分类器的效率。TMVA 还支持对分类器的超参数进行调整,以获得更好的分类效果。它提供了一个简单的用户界面,可以轻松地使用和管理分类器。



图 6-11 通过 TMVA 的提供的现成数据来得到的不同算法的本底抑制率与信号效率("ROC 曲线")的示例,其中 BDT 或 MLP 等方法的本底抑制率明显高于普通极大似然方法。

我们使用了 TMVA 来抑制在 PandaX-II 实验中的偶然符合本底。偶然符合本 底的抑制本质上是一种二分类过程(信号及本底的分类),它的基本原理是使用 某种算法(例如 BDT),来学习从输入特征到输出类别的映射关系。无论是 BDT 的训练学习或其他的机器学习算法的训练学习,首先需要要给它提供训练数据集。 对于压制偶然符合本底,信号是来自于位于核反冲中位线下面的核反冲事例,本 底是来自于位于核反冲中位线下面的偶然符合模拟事例。训练数据的输入特征或 变量是如下:

- 修正之后的 S1 信号 (qS1);
- •修正之后的 S2 信号 (qS2);
- •修正之前的 S1 信号 (qS1R);
- 修正之前的 S2 信号 (qS2R);
- S2 信号的宽度 (wS2);
- S2 信号在波形高度的 10% 的位置的宽度 (wTenS2);
- S1 信号在顶部和底部 PMT 中探测到的比例,可以表示为: $r = \frac{qS1_{top}-qS1_{bottom}}{qS1_{top}+qS1_{bottom}}$ (S1TBA);
- S2 信号被顶部和顶部 PMT 探测到的比例,可以表示为: $r = \frac{qS2_{top}}{qS2_{bottom}}$ (S2TBR);
- S2 信号波形的前半部分跟整个波形的比例,能够反映出波形的对称性 (S2SY1 是对原始波形对称性, S2SY2 是光滑的波形的对称性);
- S1 信号中的局部峰的数目 (S1NPeaks);
- 在底部 PMT 阵列中接收的光最多的 PMT 跟 S1 信号电荷的比例 (S1LargestBCQ);

图6-12是,来自于 Run 11 数据集的 BDT 变量分布。

6.3.4 BDT 训练结果

在训练数据时使用的程序是 TMVA 官方提供的标准分类程序(参考地址)。 BDT 算法有很多可调参数,例如,森林中的决策树数目,决策树的最大深度,树叶 节点中所需训练事例的最小百分比,决策树的增强型等。关于决策树的增强类型, 比较主流的有 Adaptive Boost 和 Gradient Boost。Adaptive Boost,简称 Ada Boost 是 一种在决策树中非常有效的增强类型,通常用于二分类问题,但也可以用于多分 类问题。Ada Boost 的工作原理是通过联合多个弱分类器来实现强分类器的效果。 它通过分配不同的权重给训练样本,来解决分类错误率高的问题。它通过迭代地 对样本进行分类,并通过加大分类错误样本的权重,使分类器更加注重分类错误 样本,从而达到不断提高分类准确率的目的。Ada Boost 的过程可分为5步:1.初 始化所有样本的权重分布。2.训练一个弱分类器,并计算它的分类错误率。3.对 于分类错误的样本,加大其权重分布。4.计算当前弱分类器的权重,并将其与其 他弱分类器的权重相加。5.如果当前弱分类器的分类错误率已经达到最小值,则 终止迭代,否则继续进行第2步。

除了 Ada Boost, Gradient Boost 是另一种常用于 BDT 的增强类型。与 Ada Boost 不同, Gradient Boost 通过不断地拟合残差,使分类器更好地预测样本的标签。随着迭代次数的增加, Gradient Boost 能够构建出一个高度复杂的决策树,从



图 6-12 BDT 训练数据(Run 11 数据集核反冲中位线以下的数据)的输入特征或变量的分布, 蓝色是信号(核反冲事例),红色是本底(偶然符合事例)。

而得到更高的准确率。然而, Gradient Boost 也存在过拟合的风险,因此需要通过 一些技巧来防止过拟合,比如剪枝或者使用正则化等。Gradient Boost 的核心过程 也可以分为5步:1.初始化一个基分类器,如决策树或线性回归。2.计算基分类 器的输出与真实标签的差距(残差)。3.构建一个新的决策树来拟合残差。4.将 新的决策树与基分类器的结果相加,得到更新后的分类器。5.如果模型的准确率 已经达到最大,则终止迭代;否则继续第2步。

这两种增强类型对偶然符合本底有着相同的分类结果,但是由于 Gradient

Boost 需要较长的训练时间,最终的训练中使用的是 Ada Boost。通常情况下, BDT 分类使用的数据分为两个部分,训练集以及测试集。训练数据集是用于训练 BDT 模型的数据集。一般训练结束后 TMVA 会产生一个 XML 模型文件,对于新的位 置数据可以利用 XML 模型文件来进行预测或分类。测试数据是用于评估机器学 习模型性能的数据集。在测试过程中,机器学习模型将使用测试数据集中的数据 进行预测,并将预测结果与测试数据集中真实的结果进行比较,从而评估模型的 准确性和可靠性。在偶然符合本底的分类中,将70%的数据分为训练数据集,30% 数据分为测试数据集。除了增强类型,BDT 的类外一个重要的参数是 NTree,它 是决策树增强迭代的最大数目。如果 NTree 太少,会导致欠拟合,如果 NTree 太 大,会导致过拟合。因此 NTree 值对于分类结果产生较大的影响。对于 NTree 值的 调整,我们使用的是 Kolmogorov-Smirnov(K-S)测试。K-S 测试是一种统计学方 法,用于比较两个分布的相似性。该测试主要用于检验数据是否符合某种理论分 布。如果两个分布的差异较小,则说明数据可能符合理论分布。否则,数据可能不 符合理论分布。K-S 测试的基本思路是计算两个分布的极差 (Maximal Difference), 然后根据极差的大小来判断两个分布是否相似。具体来说,K-S 测试通过计算样 本分布与理论分布的累积分布函数(Cumulative Distribution Function, CDF)之间 的极差来实现:

$$D = \max |F_1(x) - F_2(x)|$$
(6-8)

其中, *F*₁(*x*) 和 *F*₂(*x*) 分别表示样本分布与理论分布的 CDF。*D* 表示两个分布的极差。如果 *D* 较小,说明两个分布较为相似。如果 *D* 较大,说明两个分布较为不同。在偶然符合本底的分类中,通过训练数据集和测试数据集的 BDT 分数分布,可以算出 Kolmogorov-Smirnov (K-S) 测试值,从而判断,BDT 结果是否过拟合。通过若干次的训练,最终得到了如图6–13中的结果。当 NTree 的值为 90 时,其本底(偶然符合本底事例)和信号(核反冲事例)的 BDT 分数分布的相似度最高,其 K-S 测试值分别 0.067 和 0.458。

图6-13中的是 BDT 的分数分布,根据这个分布,可以得到最优的分类界限, 使得信噪比最大化或"性价比最大化"。优化公式是:

$$S = \frac{\epsilon_s n_s}{\sqrt{\epsilon_s n_s + \epsilon_b n_b}},\tag{6-9}$$

其 n_s 和 n_b 分别为本底和信号数目, ϵ_s 和 ϵ_b 是在某一个分类界限下的本底和 信号的排除效率。图6–14所展示,Run 9,10,11的本底抑制率随着信号效率的变 化趋势。从图中可知,随着信号效率的增大,本底的抑制率缓慢减少。对于不同的 Run,表6–3中总结了最大的信噪比S对应的本底抑制率和信号接受率。总体来说,



图 6-13 训练和测试数据集的 BDT 分数分布。其 K-S 测试值用于衡量过拟合。

通过 BDT 算法,能够去除 70% 的偶然符合本底的同时能够保留 70% 的信号。这 结果相比于一般的筛选条件来讲,对探测器灵敏度是很大的提升。在训练的过程 中虽然使用了在上文中提到的所有的输入变量,但其中在表6-4中按照灵敏度来列 出来的五个变量有着最为核心的分类作用。图6-12中可以看出,偶然符合本底 S2 信号的宽度相对于核反冲事例来说较小而且波形对称性也更差,这是因为部分偶 然符合本底的 S2 信号来自于 TPC 电极。对于 S1TBA 的分布来说,多数核反冲事 例的 S1TBA 值为-1,说明项部 PMT 没有探测到 S1 信号,这是因为当 S1 较小时, TPC 液氙面的反光作用,导致所有的光只被底部 PMT 探测到。与此相反,在偶然 符合本底的分布中这种趋势不明显,说明不少 S1 信号产生跟项部电极或 PMT 打 火的有关。因此,表6-4中的变量灵敏度排名结果跟图6-12分布基本符合。

Run	S	ϵ_s	$1 - \epsilon_b$
9	25.9	90.4%	70.2%
10	26.5	91.1%	74.6%
11	26.2	90.7%	73.7%

表 6-3 对 Run 9、10 和 11,最大的信噪比 S 对应的信号接收率 ϵ_s 本底抑制率 1 – ϵ_b 。



图 6-14 Run 9、10 和 11 的本底抑制率随着信号效率的变化趋势(ROC 曲线)。

Variable	Importance
S1TBA	0.135
wS2	0.101
S2SY2	0.093
qS1R	0.089
qS2R	0.081

表 6-4 训练过程中,灵敏度最好的五个变量以及它们的重要性排名。

6.3.5 BDT 对偶然符合本底的应用

BDT 模型训练之后,通过 XML 模型文件对任何事例或样本可以计算它们的 BDT 分数。在表6-2中已给出不考虑 BDT 的偶然符合本底的估计值。通过 XML 模 型文件对偶然符合本底计算 BDT 分数以及最佳信噪比的 BDT 分数限制条件来去 除偶然符合本底之后,能够得到最终的偶然符合本底估计值。BDT 分类之后,对 于 Run 9,10,11,计算出来的偶然符合本底的含量分别为 2.09±0.95,1.03±0.26 以及 2.53±1.39,其在核反冲中位线以下的含量分别为 0.39±0.18,0.27±0.07 以 及 0.77±0.42[172]。

BDT 训练模型不仅用于估计偶然符合本底,还要对其他所有本底事例上应用 还包括信号模型。BDT 分类对不同类型的事例有不同的效果。图6–15(a, d, g) 所展示的是,核反冲或电子反冲事例的 BDT 接收效率随着 S1 信号电荷变化的趋 势,BDT 对电子反冲和核反冲事例,除了小 S1 区域之外,均有较高的接收效率。 图中的曲线被视为效率曲线,在建立信号模型的时会用到。而从图6-15(b, c, e, f, h, i)中的 S1 和 S2 信号的二维接收效率分布来看,对电子反冲事例来讲,BDT 的主要去掉的是 S1<5PE 部分,而核反冲来说,主要被去掉的是 S2 较小的部分。 其最主要的原因是,由于电子反冲事例的 S2/S1 值相比于核反冲较大,这种特征 更像是偶然符合本底,因此 BDT 对 S1 信号较小的部分更为敏感。



图 6-15 BDT 对电子反冲和核反冲事例的接受效率跟 S1 和 S2 信号的关系图。其(a, d, g) 为 BDT 接收效率随着 S1 信号电荷变化的趋势。(b, c, e, f, h, i) 为 S1 和 S2 信号的二维 接收效率分布。

6.4 在 PandaX-4T 实验中的偶然符合本底

随着探测器体积和规模的增大,相应地探测器的灵敏度也会提高,但是在探测器运行过程中,不可控因素也会增多。例如,探测器 TPC 中电场的稳定性、PMT 数量的增加等。这些因素会导致探测器中偶然符合本底的几率越来越大。因此,在 大型探测器中准确估计偶然符合本底对于提高暗物质探测灵敏度至关重要,特别 是对于低能暗物质探测。在 PandaX-4T 实验中,偶然符合本底也是一个重要的本 底之一。接下来将详细讨论 PandaX-4T 实验中的偶然符合本底。

6.4.1 孤立 S1 和孤立 S2 信号的选取

在2.2.3中介绍, PandaX-4T 实验采用了"自触发"模式来采集数据。与 PandaX-II 实验不同的是, PandaX-4T 实验中每一个事例的波形窗口大小不同,因此选取孤立 S1 和孤立 S2 信号的方式也不同。相对于孤立 S2 信号,选取孤立 S1 信号相对较为直观,其筛选条件为:整个事例窗口中不存在任何被判定为 S2 信号的波形。由于 PMT 的数量比 PandaX-II 实验多了约三倍,导致孤立 S1 信号的暗噪声较多。

图6-16显示了 PandaX-4T 实验中孤立 S1 信号在数据筛选之前的 S1 大小分布 以及 S1 信号跟 S1TBA 的二维分布。从图中可以看出,PandaX-4T 实验中孤立 S1 信号的特征基本类似于 PandaX-II 实验中的孤立信号特征,大部分信号集中在小 S1 信号区域。然而,如图6-16(a)所示,在 7PE-15PE 范围内,PandaX-4T 实验 中的孤立 S1 信号存在一个明显的峰。而图6-16(b)则显示,该峰对应的 S1TBA 值大于 0,并且从信号电荷来看,有可能被软件错误地标记为 S1 信号的单电子信 号。通过对大量此类事例波形的查看,我们发现其中大部分是由于单电子信号波 形较为分散所致,其中最大的一部分被错误地标记为 S1 信号。因此,在数据筛选 过程中,我们需要对这类问题进行更为细致的处理。

PandaX-4T 实验中,数据筛选之前,孤立 S1 信号的平均频率约为 18.7Hz,相 比于 PandaX-II 实验高了 40 倍左右。其中最主要的原因是因为 PandaX-4T 实验使 用的 PMT 数量比 PandaX-II 实验更多,这导致暗噪声的增加,从而增加了孤立 S1 信号的发生频率。图6–17显示,孤立 S1 信号的发生频率随时间变化呈现出一定的 趋势。特别地,在 2021 年 1 月 17 日至 2021 年 1 月 24 日期间,孤立 S1 信号的发 生频率比平均频率高出两倍左右。通过数据质量监控系统发现,这是由于顶部的 四个 PMT 不稳定导致的,进而增加了孤立 S1 信号的发生频率。

经过 S1 信号质量筛选后, PandaX-4T 实验中孤立 S1 信号中被错误地标记为 S1 信号的单电子基本被排除,如图6-18所示。此外,孤立 S1 信号的平均发生频率下降为 9.5Hz。从图6-19中可以看出,经过 S1 信号质量筛选后,在高频率部分



图 6-16 (a) 为数据筛选之前的孤立 S1 的分布, (b) 为数据筛选之前的 S1 和 S1TBA 的二 维分布。



图 6-17 数据筛选之前的孤立 S1 信号的发生频率随着时间变化的趋势。

反而比平均频率更低(与图6-17相比)。可能的原因是,当顶部四个 PMT 不稳定 而打火时,质量较好的 S1 信号变为质量更差的信号,因此经过质量筛选后,这部 分事例更多地被排除掉。

在 PandaX-4T 实验中,选取孤立 S2 信号的方式与 PandaX-II 实验类似,即在 波形窗口内不能出现被认为"合格"的 S1 信号。所谓"合格"的 S1 信号是指在 S1 附近相对较"干净"或没有其他跟随信号的信号,为了避免 S1 信号来自单电 子或其他信号的一部分。因此,在选取孤立 S2 信号时,这种筛选条件必须被考虑 在内。在6.3.1中已介绍,孤立 S2 信号的最主要来源之一是 TPC 电极 (阳极、阴 极和门电极)。在 PandaX-4T 实验中,使用空心不锈钢圆环来制作电极,圆环采用 的是低本底的不锈钢。为了保证电场的均匀性,圆环中心覆盖了透光率较高的用 不锈钢丝制成的细网。尽管这种不锈钢材料具有低本底特性,但它仍含有一定量 的放射性元素,这些放射性元素会产生一定程度的本底。此外,网丝在强电场下 会发射电子。图6–20显示了孤立 S2 信号、门电极事例和阴极事例的大小、S2TBA 和宽度分布。门电极事例和阴极事例是根据电子漂移时间选择的,其漂移时间范

-110-



图 6-18 (a)为 S1 信号质量筛选之后的孤立 S1 的分布, (b)为 S1 信号质量筛选之后的 S1 和 S1TBA 的二维分布。



图 6-19 S1 信号质量筛选之后的孤立 S1 信号的发生频率随着时间变化的趋势。

围分别为 0-5μs 和 800-845μs。从 S2 大小和 S2TBA 的分布来看,孤立 S2 信号的 分布形状更像门电极事例,但从 S2 宽度分布来看,孤立 S2 事例相比于门电极或 阴极事例具有较大的宽度。这部分事例很有可能来自气氙部分。如2-9图所示,顶 部 PMT 光阴极有约 750V 的电压,因此它与阳极可以形成向上的电场。如果某种 入射粒子在该区域产生能量沉积,产生的电子将向下漂移并产生 S2 信号。气氙区 域的电场相比于阴极和门电极之间的电场较弱,因此电子到达阳极时会导致较为 明显的扩散效应,并且 S2 信号波形的宽度较大。因此,孤立 S2 事例的宽度分布 中含有较宽的部分。图6-20中各个事例分布的 S2 取值范围为 200-1200PE。此外, 更大的孤立 S2 信号中的大部分来自于大能量信号产生的"尾巴"信号或被算法错 误判断为 S2 的其他噪声。

经过一系列 S2 信号相关的数据筛选后,孤立 S2 信号的发生频率为 0.0045Hz。 图6-21展示了孤立 S2 信号的发生频率随时间变化的趋势。与孤立 S1 信号类似, 2021 年 1 月 17 日至 2021 年 1 月 24 日孤立 S2 信号的发生频率明显低于其它时间 段的频率。这主要是因为,跟孤立 S1 信号一样,不稳定的 PMT 导致了 S2 信号质

-111 -



图 6-20 孤立 S2 事例(蓝色), 阴极事例(红色)以及门电极事例(绿色)的 S2 大小, S2TBA 和 S2 信号宽度分布

量的变差,经过 S2 信号的质量筛选后,更多的 S2 信号被排除。此外,由于孤立 S2 信号的发生频率远低于孤立 S1 信号,因此相比于孤立 S1 信号,它的变化浮动 更大。



图 6-21 S2 信号质量筛选之后的孤立 S2 信号的发生频率随着时间变化的趋势

孤立 S1 和孤立 S2 信号的发生频率的误差是根据频率随着时间变化的浮动来 计算。它们的相对误差分别为 10.5%(孤立 S1)和 12.7%(孤立 S2)。表6-5中总 结了 PandaX-4T 实验暗物质信号区间孤立 S1 和孤立 S2 在数据筛选以及数据筛选 之后的发生频率。

Туре	Live duration [Days]	Isolated S1 rate [Hz]	Isolated S2 rate [Hz]
Pre-selection	86.04	18.74 ± 0.47	0.12 ± 0.1
Post-selection	86.04	9.51 ± 0.96	0.0045 ± 0.0006

表 6-5 PandaX-4T 实验暗物质信号区间孤立 S1 和孤立 S2 在数据筛选以及数据筛选之后的发 生频率。

6.4.2 偶然符合本底的抑制

PandaX-4T 实验中,与 PandaX-II 实验类似,通过选取孤立 S1 和孤立 S2 信 号,利用随机配对的方法来得到偶然符合本底的样本。PandaX-II 实验中,采用一 系列数据筛选和 BDT 方法来去除大部分的偶然符合本底。而在 PandaX-4T 实验 中,仅采用通常的数据筛选方法也能将偶然符合本底抑制到很低水平,因此没有 使用 BDT 方法。在数据筛选中,漂移电子的扩散效应起到了最关键的作用。由于 PandaX-4T 实验探测器相比 PandaX-II 实验探测器更大,漂移电子的扩散效应更加 明显,这主要体现在 S2 信号的宽度和电子漂移时长上。对于正常的物理事例,S1 和 S2 信号间具有关联性, S2 信号的宽度随着电子漂移时长的增加而增大。因此, 物理事例的电子漂移时长和 S2 信号宽度的二维分布具有特征。而对于偶然符合 本底, S1 和 S2 信号没有物理关联性,因此它的电子漂移时长和 S2 信号宽度的二 维分布上没有物理事例的特征。基于蒙特卡罗模拟方法得到的偶然符合本底样本 的电子漂移时长和 S2 信号宽度的二维分布如图6-22所示。在图6-22(a) 中,显示 了未经过电子扩散效应的限制条件的分布。可以看出,S2信号的宽度没有随着电 子漂移时长的增加而增大,表明 S1 和 S2 信号间没有关联性。而在图6-22(b) 中, 显示了经过电子扩散效应的限制条件之后的分布。可以看出,大部分偶然符合本 底可以被这种数据筛选方式排除。综上所述,漂移电子的扩散效应在抑制偶然符 合本底中具有非常关键的作用。



图 6-22 偶然符合本底的电子漂移时长跟 S2 信号宽度的二维分布, (a) 图是经过电子扩散效 应的限制条件之后的分布, (b) 图是经过电子扩散效应的限制条件之前的分布。

如图6-23所示, PandaX-4T 实验的偶然符合本底总体上跟 PandaX-II 实验的偶 然符合本底具有类似的分布特征。从 S1 和 S2 信号的联合密度分布来看,大部分 信号聚集在 S1 信号较小的区域。经过所有数据筛选之后以及根据公式6-1计算得 到的偶然符合本底的含量为 2.43 ± 0.47,其核反冲中位线以下有 0.80 ± 0.15,其 误差根据通过误差传递公式来计算。表6-6中总结了 PandaX-4T 实验暗物质能区各 个子数据集偶中经过所有数据筛选之后的然符合本底的含量。

	Live time	Max drift time	
Dataset	(Days) (μs)		n _{acc}
Set 1	1.81	800	0.04 ± 0.01
Set 2	12.28	810	0.32 ± 0.05
Set 3	5.13	817	0.03 ± 0.01
Set 4	32.98	841	0.99 ± 0.18
Set 5	33.84	841	1.05 ± 0.21
Total	86.04	-	2.43 ± 0.47

表 6-6 PandaX-4T 实验暗物质能区各个子数据集偶中经过所有数据筛选之后的然符合本底的 含量。其单位是,事例数。



图 6-23 经过所有数据筛选之后的偶然符合本底的 S1 和 S2 信号联合分布,其 S1 和 S2 均已 修正。红色线是核反冲中位线。

6.4.3 死窗口的偶然符合本底

在 PandaX-4T 实验中,我们采用了另一种独立的方法来估计偶然符合本底,即 死窗口的偶然符合本底。一般情况下,S1 和 S2 信号相关的物理事例产生的电子 会在最大漂移时长内到达门电极。由于探测器的高度是有限,在 PandaX-4T 中物 理事例最大漂移时长约为840µs 左右。如果漂移时长超过840µs,则说明S1和S2 信号不来自同一个事例,即为偶然符合事例。图6-24中的红色阴影部分表示死窗 口偶然符合本底,其中的事例选取范围为850-1000µs 区间。图中845µs 附近的峰 对应的是阴极事例,而其余部分对应的是物理事例窗口。通过比较物理事例窗口 和死窗口的分布高度,可以得知物理事例窗口中包含相当多的偶然符合本底。因 此,可以通过死窗口偶然符合本底来估计物理窗口中的偶然符合本底含量:

$$n_{\rm phy} = \frac{n_{\rm off} \cdot \Delta t_{\rm phy}}{\Delta t_{\rm off}},\tag{6-10}$$

其 n_{phy} 为物理事例窗口的偶然符合本底估计值, n_{off} 为死窗口内的偶然符合本 底的事例数, Δt_{phy} 和 Δt_{off} 分别为物理事例窗口以及死窗口大小。另外如图6-25所 示,死窗口偶然符合本底的 S2 跟孤立 S2 信号的一维分布也比较一致。然而死窗 口偶然符合事例的统计量较少,如果直接使用 850-1000 μ s 内的事例来估计,经过 数据筛选后估计值的统计涨落非常大。因此为了提高事例统计量,将死窗口扩展 到 5000 μ s。经过数据筛选之后,根据公式6-10计算得到,在物理事例窗口中的偶 然符合本底的含量为 2.48 ± 0.87,其误差来自于统计涨落。其结果在误差范围内 跟6.4.2中的结果(2.43 ± 0.47)基本一致。



图 6-24 寻找暗物质数据(数据筛选之前)的电子漂移时长分布,其在 845µs 附近的峰为阴 极事例,红色阴影部分为死窗口偶然符合本底。



图 6-25 孤立 S2 信号(蓝色)跟死窗口偶然符合本底的 S2 信号(红色)分布的比较(数据筛选之前)。

第七章 利用 PandaX-4T 实验数据寻找 ⁸B 太阳中微子

PandaX 实验的主要目标是探测暗物质。然而,由于探测器可以探测到的信号 范围相对广泛,因此不仅可以探测暗物质,还可以探测中微子。正如在1.2.2中所 述,⁸B 太阳中微子通过 CEvNS 过程,能够在低能区域产生与暗物质信号难以区 分的核反冲事例。换句话说,⁸B 太阳中微子是暗物质寻找分析中的一种本底。因 此,我们通过 PandaX-4T 数据可以寻找⁸B 太阳中微子信号。

通常情况下,在探测器中寻找信号时需要确定信号的能量区域,即感兴趣的 能量区间(Region of Interest)。确定感兴趣的能量区间以后,根据它才能进行本底 估计以及建模。根据图5--7可知,⁸B 太阳中微子核反冲信号大部分都聚集在1keV 以下。因此相比于暗物质分析,此分析中需要将核反冲探测阈值下降到更低的水 平。⁸B 太阳中微子分析中我们将探测阈值从暗物质分析中的 1.33 keV 下降到至 0.95 keV,整个感兴趣的能量区间为 0.95-5 keV。如图7--1所示,⁸B 太阳中微子分 析中的信号接受效率(红色实线)在低反冲能区域比暗物质分析的信号效率(粉 色虚线)更高。



图 7-1 ⁸B 太阳中微子分析中的信号接受效率(红色实线)图。其中包括 ⁸B 太阳中微子(黑 色实线)和暗物质(4 GeV/c² 和 8 GeV/c² 灰色虚线)反冲能普。除此之外,图中还包括数据重 建(蓝色实线)、数据筛选(绿色实线),感兴趣区间以及 BDT(粉色实线)的信号接受效率。

在本章节,我们采用标准的盲分析来寻找⁸B太阳中微子信号。

7.1 本底模型

根据表5-2可知,在暗物质分析中含量最多的是电子反冲本底。而在极低能区, 偶然符合本底变成主导的本底。相比于偶然符合本底,核反冲、电子反冲以及材 料表面本底的含量非常少,甚至可以忽略。因此偶然符合本底的估计及抑制是本 分析的难点也是关键点。

在 PandaX-4T 实验暗物质分析中提到,基于物理事例的电子扩散效应特征将 偶然符合本底抑制到很低水平,因此没有进一步使用 BDT 方法来进一步去除偶然 符合本底。但是 ⁸B 太阳中微子分析中由于偶然符合本底主导,因此需要利用 BDT 方法来去除偶然符合本底。此分析中同样地使用了一些基础的数据筛选。例如,死 时间条件,电子扩散效应条件,一英寸 PMT 的限制条件,单次散射条件,TPC 有 效体积条件以及波形窗口的"干净"条件等等。这些数据筛选条件能够去除大部 分的噪声以及偶然符合本底。另外跟 PandaX-4T 实验暗物质分析不同的是,由于 Set 3 数据集中噪声较多,因此这部分数据在此分析中被放弃使用。由于死时间条 件比 PandaX-4T 实验暗物质分析更严格,5 个子数据集的活时间为 64.7 天。

与暗物质分析不同的是,在⁸B 太阳中微子中分析中感兴趣的能量区间 0.95-5 keV 对应的 S1 信号不是以它的电荷大小来定义,而根据 S1 信号包含的 Hit 信号数目来定义,它可以分为三个观测通道,分别为 1,2 以及 3 Hit。其 S2 信号的取 值范围是 65-300PE。

如上文2.2.3所述,在 PandaX-4T 实验暗物质分析中为了避免信号损失,采用 的是"自触发模式",因此每个事例的波形窗口不一样。但是这种触发模式对于孤 立 1 Hit 的偶然符合本底估计带来了偏差。原则上,偶然符合本底的 S1 和 S2 信号 时间差分布是均匀的,由于事例的波形时间窗口是没有限制,这导致时间差约小 偶然符合事例发生的概率越大。因此在⁸B 太阳中微子分析中为了消除这种偏差, 我们采用了跟 PandaX-II 实验类似的固定的波形时间窗口 1ms 模式。但是这是在 软件上实现的。因此通过这种事例重建模式来也能避免更多的信号损失。

在⁸B 太阳中微子分析中,偶然符合本底的估计原理与暗物质分析基本相似。 但是,孤立 S1 和孤立 S2 信号的选取存在一些区别。由于事例重建模式采用了固 定的事例窗口模式,因此基本上找不到孤立 S1 信号。换句话说,在固定事例窗口 模式下,根据 S1 和 S2 信号的配对逻辑,不能存在孤立 S1 或孤立 S2 信号事例。 因此,我们采用了从数据中随机开 1ms 窗口的方式来计算孤立 S1 信号发生的概 率。这种方法类似于 PandaX-II 实验中采用的"随机触发"模式。对于孤立 S2 信 号,我们将固定窗口扩大到 1.5ms,并且选取 S1 和 S2 信号时间差在 0.9-1.5ms 范 围内的数据来计算孤立 S2 的发生频率。换句话说,在扩展的死窗口内发生的事例 率来定义孤立 S2 信号的发生频率。选取孤立 S1 和孤立 S2 信号之后,同样地,基于蒙特卡洛方法对它们进行随机配对,从而得到完整的偶然符合本底信号分布。

由于 ⁸B 太阳中微子分析是盲分析,且偶然符合本底是主要的本底,因此需 要精确估计。在揭盲数据之前,需要使用不同的数据验证估计方法和估计值的有 效性。为此,我们定义了所谓的"相邻能量区间"(Side Band),即不包含目标信 号(⁸B 太阳中微子信号)的能量区间。我们将 S2 信号的 300-800 PE 定义为相邻 能量区间,其 S1 与感兴趣的能量区间相同,包含 1、2 和 3 Hit 等三个观测通道。 表7-1总结了相邻能量区间中 1、2 和 3 Hit 观测通道的各种本底估计值和观测到的 事件数。其中物理本底包括核反冲、电子反冲和材料表面本底。这些本底的估计 方法与暗物质分析相同。可以看出,三个观测通道的估计值和观测值在 30% 的相 对误差范围内基本一致。该相对误差是根据 S1 和 S2 信号的分布形状差异计算的, 如图7-2所示。其紫色阴影部分表示 30% 的相对误差带。

N _{Hit}	Physical	AC	Total	Observation
1	9.4	2060.5	2069.9	2043
2	10.1	33.8	43.9	47
3	6.9	2.2	9.1	7

表 7-1 在相邻能量区间(300-800PE)内的各个本底的估计值以及观测到的事例数。其单位 是,事例数。



图 7-2 相邻能量区间中的 2 Hits 本底(偶然符合,核反冲以及电子反冲本底)以及数据分布 的比较,(a)为 2 Hit S1 信号的分布,(b)为 S2 信号的分布。

根据相邻能量区间的结果来看,偶然符合本底的估计方法以及本底建模是有效的。在感兴趣能量区间内,利用同样的方法估计的偶然符合本底在表7-2中显示。

它的相对误差跟相邻能量区间一样根据 S1 和 S2 信号分布形状差别来计算的,如 图7--3所示,其紫色阴影误差带,即 30%。

N_{Hit}	Physical	⁸ B	AC	Total
1	0.60	5.14	4636.53	4642.27
2	0.32	2.64	71.65	74.61
3	0.15	0.51	1.05	1.71

表 7-2 在感兴趣能量区间(65-300PE)内的本底和信号估计值。其单位是,事例数。



图 7-3 感兴趣能量区间中的 2 Hits 本底(偶然符合,核反冲以及电子反冲本底)以及数据分 布的比较,(a)为 2 Hit S1 信号的分布,(b)为 S2 信号的分布。

7.2 信号模型

在4.5中提到,暗物质信号模型是基于所谓的 NEST 模型来建立的。在⁸B 太阳中微子分析中,我们同样使用了 NEST 模型,由于⁸B 太阳中微子对应的反冲能 更低,因此对于核反冲对应的光产额及电产额的误差,应采用更为精确的值。为 此在⁸B 太阳中微子分析中使用的光产额及电产额对应的误差均来自于最新的结 果 [179],如图7-4所示。

在⁸B 太阳中微子分析中,信号模型跟暗物质分析有所不同。在此分析中,我 们采用了波形模拟技术以及 NEST 模型。波形模拟技术是基于 NEST 模型提供的 S1 和 S2 信号分布,通过重新组合刻度数据中选取的 S1 脉冲信号以及 S2 单电子 信号来生成模拟事例波形和 PMT 阵列接收的信号分布。为了更接近真实数据,需 要考虑各种物理过程,包括 PMT 接收的 S1 脉冲信号波形、单重态/三重态/再结合 概率、液氙中的紫外光散射、单电子波形、液氙表面反射、电子扩散效应、物理事


图 7-4 在⁸B 太阳中微子分析中使用的电产额及光产额跟其他实验中测到的实际观测值的比较,本分析中使用的误差是阴影的误差带[179]

例波形窗口暗噪声发生的概率、PMT的打火、PMT的后脉冲信号、以及由门电极 或阴极不锈钢网丝光致电离效应。通过波形模拟技术以及参数化以上提到的物理 过程,能够产生与真实数据相近的大量波形模拟数据。与真实数据相同,将波形 模拟数据可以直接用于重建效率、筛选效率和 BDT 效率等方面的分析。图7-5中 展示了波形模拟数据和中子刻度数据的 S1 和 S2 信号的波形宽度分布,可以看出 波形模拟数据基本上能够较好的描述真实数据。

波形模拟技术的本质是基于数据中的物理参数,产生一批虚假数据的过程。近年来,基于深度学习的生成模型在生成假数据方面展现出了巨大的潜力,例如生成对抗网络(Generative Adversarial Networks, GANs)、变分自编码器(Variational Autoencoder, VAE)、流模型(Flow Model)和扩散模型(Diffusion Model)等[180-183]。这些生成模型的本质是通过深度学习方法学习数据的分布,从而能够生成符合该分布的新样本。其中,扩散模型是最近备受关注的一种基于时间的生成模型,它使用扩散过程来生成新的样本。在这个过程中,每个时间步都会将前一个时间步中的样本向周围进行扩散,并在周围的样本中进行采样。这个过程可以看



图 7-5 波形模拟数据跟刻度数据的 S1 (a) 和 S2 (b) 信号波形宽度的比较,其红色直方图 为波形模拟数据,黑色点为刻度数据。

作是一个从先前样本的密度向目标密度的不断迭代的过程。通过多次迭代这个过程,我们可以逐渐逼近目标密度,从而得到高质量的样本集合。论文[183]将扩散模型与非平衡热力学相结合,提出了一种新的深度无监督学习方法。这个方法使用扩散模型来学习数据的分布,通过最小化非平衡热力学势函数来进行优化。该势函数包括两个项:一个是负对数似然项,用来学习数据的分布。另一个是非平衡项,用来控制数据生成过程的不确定性。通过最小化这个势函数,可以得到高质量的数据分布,并产生高质量的样本。在实现扩散模型时,通常使用神经网络来模拟扩散过程,该神经网络被称为扩散模型的动力学网络。它的输入是一个初始样本,输出是一个新的样本。在每个时间步中,网络将前一个时间步中的样本向周围进行扩散,并在周围的样本中进行采样。这个过程可以通过反向传播算法进行优化,以最小化扩散模型的目标函数。扩散模型在图像生成、自然语言生成等方面具有广泛的应用。

虽然在⁸B 太阳中微子信号模型中没有使用扩散模型,但是在基于真实数据建 立信号模型方面,扩散模型具有巨大的应用价值,同时扩散模型可以与波形模拟技 术相结合。在这里,利用扩散模型和中子刻度数据,产生了一批假数据,如图7-6所 示。通过比较它们的分布,可以知道通过扩散模型产生的假数据的各类分布基本 与真实中子刻度数据一致。各个变量之间的相关性也符合得较好,如图7-7所示。



图 7-6 利用扩散模型以及中子刻度数据产生的一批假数据(红色),其蓝色为中子刻。



图 7-7 利用扩散模型产生的假数据(左)跟真实的中子刻度数据(右)的变量相关系数。

7.3 BDT 方法

由于数据筛选能够将偶然符合本底抑制到较低水平,在 PandaX-4T 实验的暗物质信号区间,没有进一步使用 BDT 方法来抑制本底。然而,在⁸B 太阳中微子 信号区间,偶然符合本底是主导因素,因此需要使用 BDT 方法。与 PandaX-II 实 验暗物质分析不同的是,在⁸B 太阳中微子分析中,为了减小信号样本带来的统计 误差,BDT 训练所使用的信号样本来自于波形模拟数据。BDT 训练的输入特征或 变量与 PandaX-II 实验相似,但其数量更多。在此介绍一些 PandaX-II 实验暗物质 分析中未曾使用过的输入变量:

- S1 信号的宽度(wS1)。
- 组成 S2 信号的 Hit 电荷的标准偏差(S2HitStd)。
- S2 信号的电荷跟它的宽之比(S2QW),其定义为: S2/wS2。
- S2 信号顶部 PMT 阵列中接收最多光的 PMT 的电荷跟 S2 信号总电荷之比 (S2MaxCHQ)。
- S2 信号的前 1µs 波形的电荷 (S2F)。
- S2 信号的前 1µs 波形的 TBA (S2FTBA)。
- 组成 S2 信号的前 1µs 波形的脉冲信号大小的标准偏差(S2FHitStd)。
- S2 信号的前 1µs 波形的大小跟它的宽度之比 (S2FQW)。

图7-8展示, BDT 训练过程中最为敏感的前 6 个变量分布。它们的灵敏度排名 为: S2HitStd > S2TBA > wS1 > S2FQW > S2MaxCHQ > wS2。根据 S2HitStd 分布 可知,物理事例的 S2 信号的 Hit 信号分布相对于偶然符合分布来说,涨落比较小。 这一结果能够反应在6.3.1中所述偶然符合本底的来源。从其他 S2 信号变量的分 布中也能看出,本底变量分布中跟信号分布不同的部分很有可能电极(门电极或 阴极)事例有关。这结果符合在6.4.1中讨论的结论。在6.4.1中提到,在 PandaX-4T 实验中有相当一部分的单电子或单电子信号碎片被错误地标记为 S1 信号。因此 对于图7-8中本底 S1 信号的宽度分布里较大的部分可能跟单电子或单电子信号中 一些碎片有关。

在上文中提到,偶然符合本底是根据孤立 S1 和孤立 S2 信号通过蒙特卡洛随 机配对的方法来得到。由于孤立 S2 信号的事例数有限,在随机配对时,务必避免 重复使用孤立 S2 信号,重复使用孤立 S2 信号,会导致过拟合以及错误的测试结 果。为了避免这种问题,模拟偶然符合本底时,将孤立 S1 和孤立 S2 信号分为两 个部分,分别为训练数据集以及测试数据集。为了防止过拟合,根据 K-S 测试,将 BDT 的 NTree 参数的值定为 50。图7–9所示,训练和测试数据集的 BDT 分数分布。

跟暗物质分析不同的是,在⁸B太阳中微子分析中,我们采用了盲分析,因此



图 7-8 BDT 训练变量过程中敏感度最强的六个变量的分布。



图 7-9 训练和测试数据集的 BDT 分数分布。其 K-S 测试值用于判断训练模型的过拟合程度。

寻找最优的 BDT 分数界限时,使得观测到 ⁸B 太阳中微子信号的概率最大化。一般来说,在粒子物理实验中单位时间内事例发生的概率服从泊松分布,因此使得最大化 ⁸B 太阳中微子信号观测到的概率的定义是:

$$\sum_{N_{\text{obs}}>N_{\text{mathrmcri}}} \frac{(B+S)^{N_{\text{obs}}}}{N_{\text{obs}}!} e^{-(B+S)},$$
(7-1)

其 $S \in {}^{8}B$ 信号的估计值, B 为本底的估计值。其 N_{cri} 的含义为:在已知的本 底估计下,观测⁸B太阳中微子信号的概率为小于 N_{cri} 的阈值。这里我们取 N_{cri} 为 2.28% (两倍标准偏差的显著性)。上文中提到的 65-300PE 是被定义为初步的 感兴趣区间,而不是最优的感兴趣能量区间。因此我们寻找最优的 BDT 分数界限 的同时还确定了 S2 的上下限,换句话说,通过三维扫描寻找最优值的方法来确定 S2 信号的上下限以及 BDT 分数界限。图7-10所示的是对 3 Hit 观测通道的 S2 信 号的上下限以及 BDT 分数界限的三维扫描的结果。其横轴是 S2 的上线,纵轴是 S2 信号的下限, 颜色轴表示 ⁸B 太阳中微子信号观测到的概率。红色四角星对应的 是其⁸B太阳中微子信号观测到的概率最大的坐标,即65-230PE。因此可知,对于 2 Hit 观测通道来说,最优的 BDT 分数界限为 0.3,最优的 S2 上下限为 65-230PE。 同理,对3Hit的观测通道能够得到对应的最优的BDT分数界限以及最优的S2上 下限。由于1Hit 观测通道相比于本底含量,信号的估计值几乎可以忽略。因此1 Hit 的观测通道被视作为相邻区间。优化所有的参数以后,可以对数据进行揭盲。 表7--3中总结了, 2 Hit 以及 3 Hit 观测通道的优化后的本底, 信号以及观测到的事 例数。根据 BDT 应用之前和之后的结果来看,优化后的 BDT 分数界限将 2 Hit (3 Hit) 观测通道的偶然符合本底抑制到 98% (96%) 的同时保留 39% (31%) 的⁸B 太阳中微子信号。图7-11所示, BDT 应用之前和之后的偶然符合本底以及⁸B 太阳 中微子的 S2 信号分布。从图中不难看出,无论是偶然符合本底或⁸B 太阳中微子 信号,它们对 BDT 分数界限的响应,随着 S2 信号没有明显的改变(BDT 分数界 限跟 S2 信号的大小没有明显的关联性)。最终通过应用 BDT 之后,在 2 Hit 以及 3 Hit 观测通道中观测到数据只有一个(其 S1 = 1.6 PE 以及 S2 = 165 PE), 然而, 总 的估计值(信号+本底)约为3.21。这说明没有明显超出预期的信号,因此跟暗 物质分析类似,可以对 ⁸B 太阳中微子信号或暗物质信号的物理参数空间给限制。

N _{Hit}	S2 range (PE)	BDT	ER	NR	Surface	AC	Total prediction	⁸ B	Observation
2	65 – 230	pre	0.04	0.10	0.14	62.43	62.71	2.32	59
		post	0.02	0.04	0.03	1.41	1.50	1.42	1
3	65 - 190	pre	0.01	0.05	0.08	0.79	0.93	0.42	2
		post	0.00	0.02	0.03	0.02	0.07	0.29	0

表 7-3 对 BDT 以及 S2 上下限优化后的各类本底以及信号估计值。其中包括 BDT 应用之前 和之后的对数据的揭盲结果。其单位是,事例数。





图 7-10 根据公式7-1,对 2 Hit 观测通道的 S2 信号的上下限以及 BDT 分数界限,三维扫描的结果。其横轴是 S2 的上限,纵轴是 S2 信号的下限,颜色轴表示⁸B 太阳中微子信号观测 到的概率。

7.4 统计诠释

在粒子物理实验中,根据本底和信号模型,需要对观测到的数据进行统计上的诠释。通常情况下,有两种情况;一种是发现新的粒子,对于这种情况来说,需要计算它的显著性。另外一种是没有发现新粒子,在这种情况下,对物理模型的物理参数空间设置上限。PLR 是一种在粒子物理领域广泛应用的统计方法,用于检验新物理模型并设置模型物理参数空间的上限。对于发现新粒子,PLR 用于检验数据中存在新粒子信号的假设,其思路是将数据在信号 + 本底假设下的可能性 与仅本底假设下的可能性进行比较。相对于本底预期,如果在数据中观测到的事例数显着过量可能表明存在新粒子信号,并且这种过量的显着性可以通过从 PLR



图 7-11 BDT 应用之前(蓝色直方图)和应用之后(红色直方图)的偶然符合本底(a)以及 波形模拟信号(b)的 S2 分布。

则可以在给定的置信水平下对各种信号模型参数设置上限。无论是计算 P 值或计 算物理参数的上线,整个流程涉及到统计学中的假设检验。在科学研究中,假设是 对一种现象或一组现象的陈述或提出的解释。假设检验涉及收集和分析数据,以 确定观察到的结果是否与假设一致,或者它们是否表明应该拒绝假设。在假设检 验中,第一步是制定零假设,它代表默认假设或现状。零假设通常被选择为已经 过先前测试和确认的完善的理论或模型。备择假设代表感兴趣的假设,它可能是 一个新的理论或对现有理论的修改。下一步是选择检验统计量,它是观察数据的 函数,用于衡量数据与零假设和备择假设的兼容性。检验统计量的选择取决于要 解决的具体问题,但一些常见的选择包括似然比、t 检验和卡方检验。第三步是定 义拒绝区域,这是在零假设下不太可能出现的检验统计值的范围。拒绝区域通常 是根据预先确定的显着性水平选择的,该显着性水平是当零假设为真时拒绝零假 设的概率。显着性水平的常见选择包括 0.05 和 0.01。第四步是计算 P 值, 它是在 零假设下观察到的检验统计量与观察到的值一样或更极端的概率。P 值用于确定 是否拒绝或无法拒绝零假设。如果 P 值小于选定的显着性水平,则拒绝零假设以 支持备择假设。如果 P 值大于所选的显着性水平,则不会拒绝零假设。在粒子物 理实验领域, PLR 是最为常用的检验统计量。大多数情况下, 很难发现新粒子, 例 如⁸B 太阳中微子分析中也没有观测到超出预期的信号,因此需要给模型物理参数 计算上限。计算上限的第一步是构造似然函数。它将观察数据的概率描述为模型 参数的函数。似然函数可以写成:

$$L(\boldsymbol{\theta}|\text{data}) = \prod_{i=1}^{n} f(x_i|\boldsymbol{\theta})$$
(7-2)

-128-

其中 θ 表示模型的参数, $f(x_i|\theta)$ 是观测数据 x_i 的概率密度函数 (Probability Density Function, PDF)。参数的最大似然估计量定义为使似然函数最大化的 θ 的值。PLR 检验统计量用于将零假设(无信号存在)与备择假设(信号存在)进行比较。PLR 被定义为两种假设下的似然比,并列出了干扰参数(影响背景形状的参数):

$$\lambda(\mu) = \frac{L(\mu, \hat{\hat{\theta}})}{L(\hat{\mu}, \hat{\theta})}$$
(7-3)

其中 μ 为信号强度参数, θ 为冗余参数, $L(\hat{\mu}, \hat{\theta})$ 为似然函数在整个参数空间 中的极大值, $L(\mu, \hat{\theta})$ 是给定 μ 值的备择假设下的最大似然。检验统计量 q_{μ} 定义 为:

$$q_{\mu} = \begin{cases} -2\ln\lambda(\mu), & \hat{\mu} \leq \mu\\ 0, & \hat{\mu} > \mu \end{cases}$$
(7-4)

为了得到 q_{μ} 分布,通常使用蒙特卡洛方法在零假设和备择假设下大量模拟生成伪观测数据并为每个伪数据计算检验统计量。零假设下检验统计量的分布用于 计算 P 值。P 值是针对无效假设的证据的度量。它是一个概率值,用于确定观察 到的效果的统计意义。P 值通常表示为介于 0 和 1 之间的数字。P 值越小,反对零 假设的证据就越强。通常,P 值小于或等于 0.05(或 0.10)被认为具有统计学意 义。它可以计算为:

$$P_{\mu} = \int_{q_{\mu,\text{ob}}}^{\infty} f\left(q_{\mu} \mid \mu\right) \mathrm{d}q_{\mu} \tag{7-5}$$

一些实验需要检验的信号强度 μ 的数值非常小,这会导致得到相近的 $f(q_{\mu} | \mu)$ 与 $f(q_{\mu} | 0)$ 的分布,这时在接受某个 μ 值时,其对应的 p 值与信号不存在 ($\mu = 0$) 情况的 P 值相当接近,这样实验在这个 μ 值附近已经缺乏灵敏度。这时,一种 CL_S 的方法被用来对 P 值进行修正 [184]。首先,需要计算仅本底假设下的 P 值:

$$p_{b} = \int_{0}^{q_{\mu,obs}} f(q_{\mu} \mid 0) dq_{\mu}$$
(7-6)

修正后的定义为:

$$CL_S = \frac{p_{s+b}}{1 - p_b} \tag{7-7}$$

对于某一个特定的置信水平 1 – α (95% 或 90%),若根据某个信号强度 μ 的 值计算得到的 CL_s $\leq \alpha$,我们将拒绝这个假设。由于 1 – p_b 的值小于 1, CL_s 的值 通常大于对应的 p_{s+b} ,从而能够得到更加保守的结果。图7–12展示的是通过简单 的计数实验 (Counting Experiment)来得到的,对于不同 μ 对应的 CL_s 以及 CL_{s+b} (p_{s+b})值分布,其中包含一倍和两倍标准误差带 (绿色和黄色区域)。

为了便于横向比较各个实验室得出的结果,2021年主流的暗物质直接探测实验联合发布了关于统一采用暗物质直接探测相关的统计分析方法的白皮书[185]。 在白皮中提到,后续的暗物质直接探测相关的分析中,推荐一致采用 CL_{s+b},因此 在 PandaX-4T 暗物质以及⁸B 太阳中微子分析中,我们统一采用 CL_{s+b}。



图 7-12 通过简单的计数实验来得到的,对于不同 μ (横轴)对应的 CL_s 以及 CL_{s+b} 值的分布 (红色和蓝色虚线),黑色虚线为预期的 CL_s 值,绿色和黄色区域为一倍和两倍标准误差带。

在通常情况下,为了的得到更准确的结果,需要做大量的蒙特卡洛模拟去产 生伪观测数据并计算检验统计量。这做法需要耗很多时间且计算资源。Cowan 等 人为此提出了渐进近似方法 [174]。渐近方法是一种在某些假设下近似检验统计量 分布的技术。在粒子物理学中的假设检验中,通常使用渐近方法来近似零假设下 的轮廓似然比检验统计量的分布。渐近方法的主要假设是样本量足够大,因此测 试统计量的分布近似正态。具体而言,对于给定的测试统计量 *q*,渐近分布由下 式给出:

$$f(q_{\mu}) = \frac{1}{2\sqrt{q_{\mu}}} \exp\left(-\frac{q_{\mu}}{2}\right) \tag{7-8}$$

这种分布被称为具有一个自由度的卡方分布,它提供了在大样本量极限下测 试统计量分布的良好近似。然而,重要的是要注意,在样本量小或似然函数表现 不好的情况下,渐近逼近可能不准确。在这种情况下,蒙特卡洛模拟来估计测试 统计的分布。

对于这种复杂的统计分析,有几种现成且成熟的软件包,例如附属于 ROOT 的 RooStats[186],附属于 Scikit-HEP 软件项目 HepStats 和 PyHF[187-188]。前者 由 C++ 来开发的,后者用纯 Python 来开发的。在 ⁸B 太阳中微子分析中我们使用 的是 RooStats。RooStats 在粒子物理实验领域广泛应用的统计分析的开源软件包。它是跟 ROOT 集成在一起。RooStats 旨在与 RooFit 一起使用,RooFit 是一个数据 建模和拟合工具包,也是跟 ROOT 集成在一起。RooFit 提供了一套的非常实用的 功能,用于将复杂模型拟合到数据,包括定义自定义概率密度函数和同时拟合多 个数据集。RooStats 通具有进行假设检验、构建置信区间和计算 p 值的能力计算 物理参数的上限。RooStats 的优势之一是其灵活性和可扩展性。该工具包设计为 易于扩展和定制,以满足不同应用程序的需求。例如,用户可以定义自己的统计测试或实现用于拟合数据的自定义模型。

为了简化建模部分,这里使用了专门构建概率密度函数的工具 HistFactory[189]。它是一种基于 RooFit/RooStats 框架,采用 XML 文件的和 ROOT 直方 图,构建参数化概率密度函数的工具。该工具采用模块化方法从更原始的概念构 建复杂的概率密度函数。

```
<!DOCTYPE Channel SYSTEM 'HistFactorySchema.dtd'>
<Channel Name="channel1" InputFile="example.root" >
<Data HistoName="data" HistoPath="" />
<StatErrorConfig RelErrorThreshold="0.05" ConstraintType="Poisson" />
<Sample Name="signal" HistoPath="" HistoName="signal">
<OverallSys Name="syst1" High="1.05" Low="0.95"/>
<NormFactor Name="SigXsecOverSM" Val="1" Low="0." High="3." />
</Sample>
<Sample Name="background1" HistoPath="" NormalizeByTheory="True"
HistoName="background1">
<StatError Activate="True" HistoName="background1_statUncert" />
<OverallSys Name="syst2" Low="0.95" High="1.05"/>
</Sample>
<Sample Name="background2" HistoPath="" NormalizeByTheory="True"
HistoName="background2">
<StatError Activate="True" />
<OverallSys Name="syst3" Low="0.95" High="1.05"/>
</Sample>
</Channel>
```

以上是使用 HistFactory 规范来写的 XML 配置代码示例。该代码定义中了一 个名为 channel1 的单一通道,其中包含三个样本: signal、background1,background2。 该通道指定一个位于 example.root 的输入文件,以及一个名为 data 的数据直方图。 该通道还将统计误差类型设置为 Poisson,相对误差阈值为 0.05。信号样本有一个 名为信号的直方图,以及一个名为 syst1 的系统误差,范围为 0.95 到 1.05。它还具 有一个名为 SigXsecOverSM 的归一化因子,值为 1,范围为 0 到 3。background1 样 本有一个名为 background1 的直方图,并按理论归一化。它还具有基于名为 background1statUncert 的直方图统计误差,以及名为 syst2 的系统误差,范围为 0.95 至 1.05。background2 样本具有跟 background1 类似的结构。从整个 XML 配置代码文 件的结构可见,利用 HistFactory 来构造 PDF 及其简单且明了。因此整个 ⁸B 太阳 中微子分析的统计诠释部分通过 RooStats 和 HistFactory 进行的。

由于在⁸B 太阳中微子分析中没有观测到明显超出预期的事例,因此可以对 ⁸B 太阳中微子的物理参数给出限制。在此分析中我们将对⁸B 太阳中微子的通量 给出限制。除此之外,在⁸B 太阳中微子的信号区间,对于该信号区域的暗物质能 区也可以给出质量对应的散射截面的上限。

对揭盲之后的两种观测通道 2 Hit 和 3 Hit,我们利用 PLR 方法,对结果进行 了统计诠释。其似然函数定义为:

$$\mathcal{L} = G(\delta_{\epsilon}) G(\delta_{s}) G(\delta_{b}) G(\delta_{\Phi}) \times \left[\prod_{i} G(\delta_{BDT,s}^{i}) G(\delta_{BDT,b}^{i}) \frac{\lambda_{i}^{N_{i}}}{N_{i}!} e^{-\lambda_{i}} \right],$$
(7-9)

其*i*表示 2 和 3 Hit 的观测通道, δ 和 δ^i 分别为 2 和 3 Hit 之间的独立以及相 关的冗余参数。它们是均值为 0 的高斯分布的标准偏差。 δ_{ϵ} BDT 应用之前对物理 本底以及信号的系统误差,其来自于数据重建以及数据筛选。它的值是根据波形 模拟和刻度数据的差别来算出来的。 δ_b 是偶然符合本底在 BDT 应用之前的系统 误差,其值是基于图7–3中的数据跟偶然符合本底模型的分布之差来计算出来的。 $\delta^i_{BDT,s}$ 为 BDT 应用之后被引入的系统误差,其值是 BDT 对波形模拟以及中子刻 度数据应用之后的效率分布差别来给出的,如图7–13所示。 $\delta^i_{BDT,b}$ 是 BDT 对偶然 符合本底应用之后被引入的系统误差,它是根据 BDT 对在暗物质分析中使用的方 法来得到的偶然符合本底的效率跟在 ⁸B 太阳中微子分析中模拟出来的偶然符合 本底的效率之差来得到的。 δ_{Φ} 是来自于 ⁸B 太阳中微子通量的系统误差。 δ_s 为来 自于 NEST 模型或电产额以及光产额的系统误差,如图7–4所示。值得注意的是 ⁸B 太阳中微子信号,暗物质信号以及中子本底都属于核反冲信号,因此 δ_s 是对于不 同的核反冲信号能谱取不同的值。这里通过 f_i 的转换系数来表示不同信号的相对

- 132 -

误差,例如 f_i^{ν} 和 f_i^{χ} 分别对应的是 ⁸B 太阳中微子信号,暗物质信号。在 2 Hit (3 Hit)观测通道,对于 ⁸B 太阳中微子信号,4 GeV/ c^2 的暗物质信号以及,8 GeV/ c^2 的暗物质信号,对应的 f_i 值分别为 0.29 (0.39), 0.45 (0.60)以及 0.16 (0.24)。在表格7-4中总结了它们对应的值。 λ_i 和 N_i 分别为信号和本底的估计值以及观测到的事例数。对于 ⁸B 太阳中微子信号有:

$$\lambda_{i}^{\nu} = N_{\nu} \left(1 + \delta_{s} f_{i}^{\nu} \right) \left(1 + \delta_{\epsilon} \right) \left(1 + \delta_{\text{BDT},s}^{i} \right) + N_{\text{AC}} \left(1 + \delta_{b} \right) \left(1 + \delta_{\epsilon} \right) \left(1 + \delta_{\text{BDT},b}^{i} \right) + N_{\text{other}} , \qquad (7-10)$$

对暗物质信号有:

$$\lambda_{i}^{\chi} = N_{\chi} \left(1 + \delta_{s} f_{i}^{\chi} \right) \left(1 + \delta_{\epsilon} \right) \left(1 + \delta_{\text{BDT},s}^{i} \right) + N_{\nu} \left(1 + \delta_{s} f_{i}^{\nu} \right) \left(1 + \delta_{\epsilon} \right) \left(1 + \delta_{\text{BDT},s}^{i} \right) \left(1 + \delta_{\Phi} \right)$$
(7-11)
+ $N_{\text{AC}} \left(1 + \delta_{b} \right) \left(1 + \delta_{\epsilon} \right) \left(1 + \delta_{\text{BDT},b}^{i} \right) + N_{\text{other}},$

其 N_ν, N_{AC}, N_{other}, 和 N_χ 分别代表 ⁸B 中微子信号,偶然符合本底,其他本底 (中子本底,电子反冲本底以及表面本底)以及暗物质信号预期估计值。

将⁸B 太阳中微子作为信号的假设下,对于它的通量,90%的置信水平上限为 9.0×10⁶/cm²/s。考虑到 NEST 模型误差,相比于 XENON1T 的 1.4×10⁷/cm²/s 来说 [190],这一结果更为严谨且保守。与此同时,将⁸B 太阳中微子视为本底的假设 下,我们还对质量为 3-9 GeV 的暗物质给出了对应的散射截面的 90% 的置信水平 上限,如图7–14所示。同样地,这一结果在该质量区间内最为严格且保守的结果。 根据表7–3可知,观测到的事例数比估计值更为低,这就导致暗物质质量对应的散 射截面上限在向下方向的涨落,因此图7–14中,排除曲线靠近在一倍标准偏差带 的下方边界。除此之外,排除线已经触碰到中微子地板,虽然这次实验数据中没 有看到中微子信号,但是这种趋势会给暗物质探测带来更大的挑战。

Nuisance parameters	Stde	ev.	Estimated by	
		2-Hit		
pre-BDT eff.	δ_{ϵ}	0.14		WS vs. NR
NR signal rate	$\delta_s f_i$	f_i		NEST uncert
AC rate	δ_b	0.30		Pred. vs Side-band
BDT eff. to signal	$\delta^i_{ ext{BDT},s}$	0.14	0.13	WS vs. NR
BDT eff. to AC	$\delta^i_{{ m BDT},b}$	0.19	0.18	Alter. models
Solar ⁸ B flux δ_{Φ}		0.0)4	Ref

表 7-4 计算上限时使用的各类冗余参数(系统误差)以及它们的估计方法。



图 7-13 对于中子刻度数据(黑色)和波形模拟数据(蓝色), BDT 信号接收效率随着 S2 信 号电荷的变化趋势。它们的差别作为 BDT 模型对信号系统误差来在计算上限时使用。



图 7-14 根据⁸B 太阳中微子分析揭盲结果来计算得到的分别对⁸B 太阳中微子通量(上)以 及暗物质-核子自旋相关散射截面(下)的 90% 置信水平的上限。上图蓝色线为标准太阳模 型预测的⁸B 太阳中微子通量(5.46±0.66)×10⁶/cm²/s[191])。下图中灰色阴影区域为在不同 的曝光量(分别为1,10和1000吨.年)下的中微子地板[192]。其中还包括其他实验的结果 [96,176,193]。

第八章 总结与展望

目前暗物质直接探测是国际粒子物理研究的热门课题之一。PandaX 是一项 位于中国四川省锦屏地下实验室的以探测暗物质为主要目标的多物理目标实验项 目。PandaX 暗物质探测实验总共经历了三期,它们分别为 PandaX-I、PandaX-II 以 及 PandaX-4T。其中 PandaX-4T 是正在运行的体量最大、灵敏度最高的探测器。尽 管 PandaX 各期暗物质直接探测实验到目前为止没有观测到预期的暗物质信号,然 而,对于暗物质-核子相互作用的参数空间分别在 2016、2017 以及 2021 年给出了 世界领先的限制。

在本论文第一章,首先对暗物质和中微子相关理论和实验进行了简要介绍。在 第二章,对 PandaX 实验的探测原理以及对它的基础设施进行了概括性的介绍。在 第三章,对 PandaX 实验以及实验数据重建、处理进行了详细的介绍。除此之外, 还介绍了为了获取较高质量的数据以及数据分析开发的在线数据质量监控系统以 及跨平台的事例可视化软件工具,最后对数据以及数据筛选方法进行了总结。在 第四和第五章对探测相应参数以及各类本底以及它们的估计进行了大致讨论。

对于探测稀有事例信号的实验包括 PandaX 实验来说,有效降低或排除本底 噪声的干扰是该类实验的最大难点。在第六章,对 PandaX 实验中非常重要的本 底之一,即偶然符合,进行了系统性介绍。首先是对偶然符合本底的本质以及性 质进行了讨论,其中包括它的来源、形成机制以及估计方法等。由于 PandaX-II 和 PandaX-4T 实验规模不同,偶然符合本底的估计方法有所区别,尤其是组成偶然符 合本底的孤立 S1 和孤立 S2 信号的选取方式。因此我们对 PandaX-II 和 PandaX-4T 实验的偶然符合本底进行了分开讨论。对于 PandaX-II 实验的偶然符合本底的孤 立 S1 信号,采用了三种独立的选取方法,并且在误差范围内得到了相同的结果。 为了进一步抑制偶然符合本底,除了一般性的数据筛选之外,还采用了 BDT 方法 来进一步排除了偶然符合本底。在暗物质信号区间, BDT 方法在保留 70% 的信号 下,能够排除90%的偶然符合本底。在 PandaX-II 实验中, BDT 对提高探测器灵 敏度方面具有不可或缺的作用。然而,在 PandaX-4T 实验中探测器体量以及采集 数据模式跟 PandaX-II 实验有所不同,因此我们采用了不同的方法来选取了孤立 S1 信号以及孤立 S2 信号。在 PandaX-4T 实验中通过死窗口的偶然符合本底来验 证了其估计值的有效性。PandaX-4T 实验暗物质分析中,通过物理信号的电子扩 散效应,能够有效排除掉大部分偶然符合本底。因此在 PandaX-4T 暗物质分析中, 没有用 BDT 方法来进一步排除偶然符合本底。

随着探测器体量的增大以及灵敏度的提高,探测器能够探测更多的稀有事例 信号。然而,当实验达到一定规模时,他们就会遇到一个极限:来自于太阳、大 气层和宇宙射线的中微子也会与探测器中的原子核或电子发生弹性散射,并产生 类似于暗物质信号的事例。这些中微子信号就会成为一个无法消除的本底,被称 为中微子地板。中微子地板会影响实验对暗物质信号的辨识能力。其中,⁸B 太阳 中微子是最主要的干扰源之一。⁸B 太阳中微子在暗物质探测器中可以通过相中微 子-核子相干弹性散射(CEvNS)产生与暗物质相同类型和能谱分布的信号。目前 PandaX-4T 以及其它类似的实验已经接近了中微子地板。在第七章中,我们进行 了寻找 ⁸B 太阳中微子信号的分析,整个分析是严格遵循盲分析的标准。在 ⁸B 太 阳中微子信号区间,最为核心的本底是偶然符合本底。因此对偶然符合本底的估 计以及抑制相比于暗物质分析更为细致且严格。由于盲分析,对信号区间揭盲之 前,我们通过相邻区间的信号,对本底以模型进行验证以及评估它们的有效性。对 于信号模型,由于⁸B 太阳中微子信号区间,没有足够的信号样本(核反冲刻度数 据),因此我们采用了波形模拟技术来产生了大批量模拟信号,从而避免统计涨 落。随后,基于波形模拟信号以及模拟的偶然符合本底训练了 BDT 模型,并通过 BDT 对偶然符合本底进行了进一步排除。根据最终的揭盲结果,没有观测到明显 超出预期的信号事例。因此根据揭盲结果对⁸B太阳中微子的通量以及该信号区间 的质量为 3-9 GeV/c² 的暗物质-核子相互作用的参数空间给出了世界领先的限制。

在论文中对偶然符合本底分析结果来看,暗物质信号区间内,虽然偶然符合本底含量随着探测器体积增大而增高,由于它的信号特征比较明显,因此通过数据筛选的方式能够排除掉大部分的偶然符合本底。根据偶然符合本底信号的特征来看,它的信号相对来说较小,正因为这种小信号特征,它在⁸B 太阳中微子信号区间,变成了主导的本底。尽管通过各种数据筛选方法以及 BDT 方法来能够抑制偶然符合本底的含量,但是目前的信噪比以及观测事例的统计误差来看,未来需要进行更为精细且全面的分析。对于下一轮要进行的⁸B 太阳中微子分析,有以下方面可以改进:第一,尽量保证探测器的稳定运行,从而避免各种非物理的打火事例进入到信号区间。第二,对偶然符合本底的估计,采用更多的方法以及验证,从而降低估计带来的误差。第三,对偶然符合本底的含量进行更加精确且细致的定量分析。第四,提高采集数据的时长,从而提高⁸B 太阳中微子事例率。第五,需要降低信号模型包括 NEST 模型以及波形模拟带来的误差。第六,由于探测效率以及探测阈值等原因,很多 S1 信号没有被探测到,因此这时需要把仅电子信号(仅 S2 信号)观测通道以及传统的 S1 和 S2 符合观测通道同时考虑,从而提高信号的事例率。

目前 PandaX-4T 实验已采集完成第一批的科学数据,总的活时长约为 150 天。 通过以上提到的改进方案以及整合两批数据,有望在下一轮的中观测到⁸B 太阳中 微子信号。

参考文献

- DALTON J. A New System of Chemical Philosophy[M]. Manchester: S. Russell, 1808: 1-40.
- [2] ROCKE A J. In Search of El Dorado: John Dalton and the Origins of the Atomic Theory[J]. Social Research, 2005, 72(1): 125-158.
- [3] LAIDLER K J, KING M C. The development of the concept of chemical reaction rate from Lavoisier to Arrhenius"[J]. Journal of Chemical Education, 1995, 72(1): 29-34.
- [4] DAVIS E, FALCONER I. J.J. Thomson and the Discovery of the Electron[M]. London: Taylor & Francis, 1997.
- [5] BUCHWALD J Z. The Rise of the Electron: J.J. Thomson and his Critics on Cathode Rays and Atomic Structure, c.1894-1906[J]. Historical Studies in the Physical Sciences, 1985, 15(2): 213-254.
- [6] RUTHERFORD E. Radioactive Transformations[M]. New York: Charles Scribner's Sons, 1914.
- [7] CHADWICK J, GOLDHABER M. A Nuclear Photo-effect: Disintegration of the Diplon by γ Rays[J]. Nature, 1921, 108(2712): 171-172.
- [8] CHADWICK J. Possible Existence of a Neutron[J]. Nature, 1932, 129(3252): 312.
- [9] BROWN A, CHADWICK J. The Neutron and the Bomb: A Biography of Sir James Chadwick[M]. New York: Oxford University Press, 1986.
- [10] GELL-MANN M. A Schematic Model of Baryons and Mesons[J]. Physics Letters, 1964, 8(3): 214-215.
- [11] ZWEIG G. An SU(3) Model for Strong Interaction Symmetry and its Breaking[J]. CERN Report No.8182/TH.401, 1964.
- [12] GRIFFITHS D J. Introduction to Elementary Particles[M]. Weinheim: Wiley-VCH, 2008: 67-70.
- [13] AAD G, et al. Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC[J]. Physics Letters B, 2012, 716(1): 1-29.

[14]	CHATRCHYAN S, et al. Observation of a new boson at a mass of 125 GeV with
	the CMS experiment at the LHC[J]. Physics Letters B, 2012, 716(1): 30-61.

- [15] ELLIS J, ENGLERT C, HIGGS P, et al. The Higgs boson turns ten[J]. Nature, 2022, 596(7870): 48-55.
- [16] ESPOSITO G. An Introduction to quantum gravity[C]//. 2011. arXiv: 1108.3269 [hep-th].
- [17] KIM J, CAROSI G. Axions and the strong CP problem[J]. Reviews of Modern Physics, 2010, 82(1): 557-601.
- [18] PECCEI R. The strong CP problem and axions[J]. Lecture Notes in Physics, 2006, 741: 3-17.
- [19] GONZALEZ-GARCIA M, MALTONI M. Phenomenology with massive neutrinos[J]. Physics Reports, 2015, 460: 1-129. DOI: 10.1016/j.physrep.2014.12.005.
- [20] GIUNTI C, KIM C. Fundamentals of Neutrino Physics and Astrophysics[M]. Oxford: Oxford University Press, 2007.
- [21] BERTONE G, HOOPER D, SILK J. Particle dark matter: evidence, candidates and constraints[J]. Physics Reports, 2005, 405: 279-390.
- [22] FRIEMAN J, TURNER M, HUTERER D. Dark energy and the accelerating universe[J]. Annual Review of Astronomy and Astrophysics, 2008, 46: 385-432.
- [23] ELLIS J. Beyond the Standard Model today[J]. The European Physical Journal H, 2016, 41: 341-369.
- [24] LANGACKER P. The standard model and beyond[J]. CRC Press, 2009. DOI: 10.1201/9781420079060.
- [25] CHENG T, LI L. Gauge Theory of Elementary Particle Physics: Problems and Solutions[M]. Oxford: Oxford University Press, 2010.
- [26] Wikipedia. Standard model of particle physics[Z]. https://en.wikipedia.org/wiki /Standard_Model.
- [27] FERMI E. Versuch einer Theorie der β -Strahlen. I[J]. Zeitschrift für Physik A Hadrons and Nuclei, 1934, 88(3-4): 161-177. DOI: 10.1007/BF01351864.
- [28] FERMI E. Nuclear Physics[M]. Chicago: University of Chicago Press, 1954.
- [29] WANG K C. A Suggestion on the Detection of the Neutrino[J/OL]. Phys. Rev., 1942, 61: 97-97. https://link.aps.org/doi/10.1103/PhysRev.61.97. DOI: 10.1103 /PhysRev.61.97.

- [30] COWAN C L, REINES F, HARRISON F B, et al. Detection of the Free Neutrino: a Confirmation[J/OL]. Science, 1956, 124(3212): 103-104. eprint: https://www .science.org/doi/pdf/10.1126/science.124.3212.103. https://www.science.org/do i/abs/10.1126/science.124.3212.103. DOI: 10.1126/science.124.3212.103.
- [31] DEMARQUE P, PERCY J. A Series of Solar Models[J]. The Astrophysical Journal, 1964, 140: 541-551.
- [32] GUENTHER D, DEMARQUE P, PINSONNEAULT M, et al. Standard Solar Models: Results for New Solar Composition Using OPAL and MHD Equations of State[M]. Chicago: The Astrophysical Journal Supplement Series, 2004.
- [33] CLAYTON D D. Principles of Stellar Evolution and Nucleosynthesis[M]. Chicago: University of Chicago Press, 1983.
- [34] SCHWARZSCHILD M, HOWARD R, HÄRM R. Inhomogeneous Stellar Models. V. a. Solar Model with Convective Envelope and Inhomogeneous Interior[J]. The Astrophysical Journal, 1957, 125: 233-250.
- [35] BAHCALL J, PINSONNEAULT M. Standard solar models[M]. Palo Alto: Annual Review of Astronomy, 1989.
- [36] REDCHUK M. Solar neutrino analysis with the Borexino detector[J]. Journal of Physics: Conference Series, 2018, 1056.
- [37] DAVIS JR. R, HARMER D, HOFFMAN K. Search for Neutrinos from the Sun[J]. Physical Review Letters, 1968, 20(21): 1205-1209.
- [38] BAHCALL J, PINSONNEAULT M. Standard solar models, with and without helium diffusion, and the solar neutrino problem[J]. Reviews of Modern Physics, 1994, 64(4): 885-926.
- [39] HAXTON W. The Solar Neutrino Problem[J]. Annual Review of Astronomy and Astrophysics, 1995, 33: 459-503.
- [40] FUKUDA Y, et al. Evidence for oscillation of atmospheric neutrinos[J]. Physical Review Letters, 1998, 81(8): 1562-1567.
- [41] NAKAHATA M, et al. Establishing atmospheric neutrino oscillations with Super-Kamiokande data[J]. Nuclear Physics B, 2016, 908: 366-381. DOI: 10.1016/j.nu clphysb.2016.02.012.
- [42] KAJITA T, et al. The Super-Kamiokande experiment[J]. The European Physical Journal C, 2019, 79(4): 337. DOI: 10.1140/epjc/s10052-019-6796-2.

- [43] AHMAD Q, et al. Measurement of the rate of $v_e + d \rightarrow p + p + e^-$ interactions produced by ⁸B solar neutrinos at the Sudbury Neutrino Observatory[J]. Physical Review Letters, 2001, 87(7): 071301.
- [44] AHMAD Q, et al. Direct evidence for neutrino flavor transformation from neutralcurrent interactions in the Sudbury Neutrino Observatory[J]. Physical Review Letters, 2002, 89(1): 011301.
- [45] BELLINI G, et al. Final results of Borexino Phase-I on low-energy solar neutrino spectroscopy[J]. Physical Review D, 2014, 89(11): 112007. DOI: 10.1103/Phys RevD.89.112007.
- [46] AGOSTINI M, et al. First directional measurement of sub-MeV solar neutrinos with a monolithic liquid-scintillator detector: Borexino Collaboration[J]. ArXiv preprint arXiv:2103.15601, 2021.
- [47] AGOSTINI M, et al. Improved Measurement of Solar Neutrinos from the Carbon-Nitrogen-Oxygen Cycle by Borexino and Its Implications for the Standard Solar Model[J]. Physical Review Letters, 2022, 129(25): 252701. DOI: 10.1103/Phys RevLett.129.252701.
- [48] GANDO A, et al. Measurement of the 8B Solar Neutrino Flux with the Kam-LAND Liquid Scintillator Detector[J]. Physical Review C, 2011, 84(3): 035804. DOI: 10.1103/PhysRevC.84.035804.
- [49] EGUCHI K, et al. First Results from KamLAND: Evidence for Reactor Antineutrino Disappearance[J]. Physical Review Letters, 2003, 90(2): 021802. DOI: 10 .1103/PhysRevLett.90.021802.
- [50] SUZUKI Y, NAKAHATA M, MORIYAMA S, et al. KamLAND: Kamioka Liquid Scintillator Antineutrino Detector[M]. Singapore: World Scientific Publishing Co Pte Ltd, 2004.
- [51] ZHAN L, WANG Y, CAO J, et al. Proceedings of the International Workshop on Prospects of Particle Physics: "Neutrino Physics and Astrophysics" [M]. Singapore: World Scientific Publishing Co Pte Ltd, 2009.
- [52] AN F, et al. Neutrino physics with JUNO[J]. Journal of Physics G: Nuclear and Particle Physics, 2016, 43(3): 030401. DOI: 10.1088/0954-3899/43/3/030401.
- [53] FREEDMAN D. Coherent effects of a weak neutral current[J]. Physical Review D, 1974, 9(5): 1389-1392. DOI: 10.1103/PhysRevD.9.1389.

- [54] AKIMOV D, et al. Observation of coherent elastic neutrino-nucleus scattering[J]. Science, 2017, 357(6356): 1123-1126. DOI: 10.1126/science.aa00990.
- [55] ARISTIZABAL D, et al. CEvNS: Theory overview[J/OL]. Mini-workshop: Theoretical tools for neutrino scattering, 2021. %E2%80%9Chttps://indico.fnal.gov /event/50335/sessions/19119/attachments/146467/187021/aristizabal_theory_to ols.pdf%E2%80%9D.
- [56] CADEDDU M, et al. First measurement of coherent elastic neutrino-nucleus scattering on argon[J]. Physical Review Letters, 2020, 124(8). DOI: 10.1103/PhysR evLett.124.081301.
- [57] SCHOLBERG K. Neutrinos from the Spallation Neutron Source: The COHER-ENT Experiment[M]. Singapore: World Scientific Publishing Co. Pte Ltd, 2019.
- [58] BILLARD J, FIGUEROA-FELICIANO E, STRIGARI L. Implication of neutrino backgrounds on the reach of next generation dark matter direct detection experiments[J]. Physical Review D, 2014, 89(2). DOI: 10.1103/PhysRevD.89.023524.
- [59] COHERENT. How to search CEvNS[Z]. https://coherent.ornl.gov/how-do-you-look-for-cevns/.
- [60] COLLABORATION P. Planck 2015 results. XIII. Cosmological parameters[J]. Astronomy & Astrophysics, 2016, 594(A13). DOI: 10.1051/0004-6361/201525 830.
- [61] RIESS A, et al. Observational evidence from supernovae for an accelerating universe and a cosmological constant[J]. The Astronomical Journal, 1998, 116(3): 1009-1038. DOI: 10.1086/300499.
- [62] PERLMUTTER S, et al. Measurements of Omega and Lambda from 42 highredshift supernovae[J]. The Astrophysical Journal, 1999, 517(2): 565-586. DOI: 10.1086/307221.
- [63] BINNEY J, TREMAINE S. Galactic Dynamics[M]. Princeton: Princeton University Press, 2008: 604-610.
- [64] MCGAUGH S, LELLI F, SCHOMBERT J. Radial Acceleration Relation in Rotationally Supported Galaxies[J]. Physical Review Letters, 2016, 117(20). DOI: 10.1103/PhysRevLett.117.201101.
- [65] SOFUE Y, RUBIN V. Rotation Curves of Spiral Galaxies[J]. Annual Review of Astronomy and Astrophysics, 2001, 39(1): 137-174. DOI: 10.1146/annurev.astr o.39.1.137.

- [66] Wikipedia. Galaxy rotation curve[EB/OL]. https://en.wikipedia.org/wiki/Galax y_rotation_curve.
- [67] SCHNEIDER P, KOCHANEK C S, WAMBSGANSS J. Gravitational Lensing: Strong, Weak and Micro[M]. Berlin: Springer Science & Business Media, 2006.
- [68] PETTERS A O, LEVINE H, WAMBSGANSS J. Singularity theory and gravitational lensing[J]. Progress in Mathematical Physics, 2001, 21. DOI: 10.1007/97 8-1-4612-0199-7.
- [69] UMETSU K, BROADHURST T, ZITRIN A, et al. A Precise Cluster Mass Profile Averaged from the Highest-quality Lensing Data[J]. The Astrophysical Journal, 2010, 714(2): 1470-1490. DOI: 10.1088/0004-637X/714/2/1470.
- [70] CLOWE D, BRADAC M, GONZALEZ A H, et al. A direct empirical proof of the existence of dark matter[J]. Astrophys. J. Lett., 2006, 648: L109-L113. arXiv: astro-ph/0608407. DOI: 10.1086/508162.
- [71] NASA. Bullet Cluster[Z]. https://science.nasa.gov/matter-bullet-cluster.
- [72] AGENCY E S. CMB Maps[Z]. https://wiki.cosmos.esa.int/planck-legacy-archi ve/index.php/CMB_maps.
- [73] FENG J L. Dark Matter Candidates from Particle Physics and Methods of Detection[J]. Annual Review of Astronomy and Astrophysics, 2010, 48(1): 495-545.
 DOI: 10.1146/annurev-astro-082708-101659.
- [74] FREESE K. The Cosmic Cocktail: Three Parts Dark Matter[M]. Princeton: Princeton University Press, 2014.
- [75] ABAZAJIAN K N, ACERO M A, AGARWALLA S K, et al. Light Sterile Neutrinos: A White Paper[J]. ArXiv preprint arXiv:1204.5379, 2012.
- [76] BOYARSKY A, RUCHAYSKIY O, SHAPOSHNIKOV M. The role of sterile neutrinos in cosmology and astrophysics[J]. Annual Review of Nuclear and Particle Science, 2009, 59: 191-214. DOI: 10.1146/annurev.nucl.010909.083654.
- [77] ARCADI G, DUTRA M, GHOSH P, et al. The waning of the WIMP? A review of models, searches, and constraints[J]. The European Physical Journal C, 2018, 78(3): 203. DOI: 10.1140/epjc/s10052-018-5662-y.
- [78] BERNAL N, DUTRA M, MAMBRINI Y, et al. Dark matter production from Goldstone boson interactions: phenomenological constraints[J]. Journal of Cosmology and Astroparticle Physics, 2017, 2017(11): 021. DOI: 10.1088/1475-75 16/2017/11/021.

- [79] BRINGMANN T, HOFMANN S. Thermal decoupling of WIMPs from first principles[J]. Journal of Cosmology and Astroparticle Physics, 2007, 2007(04): 016.
 DOI: 10.1088/1475-7516/2007/04/016.
- [80] CHEN C S, LIN G L, LIN Y H. Kinetic decoupling of WIMPs: Analytic expressions[J]. Physical Review D, 2015, 91(8): 083526. DOI: 10.1103/PhysRevD.91 .083526.
- [81] HOCHBERG Y, KUFLIK E, VOLANSKY T, et al. The Fraternal WIMP Miracle[J]. Journal of High Energy Physics, 2020. DOI: 10.1007/JHEP01(2020)021.
- [82] GRIEST K, KAMIONKOWSKI M. Unitarity Limits on the Mass and Radius of Dark Matter Particles[J]. Physical Review Letters, 1990. DOI: 10.1103/PhysRev Lett.64.615.
- [83] CERN. A schematic view of comoving number density of a stable species as they evolve through the process of thermal freeze-out[Z]. https://cds.cern.ch/record/2 305579/plots.
- [84] PECCEI R, QUINN H. CP Conservation in the Presence of Pseudoparticles[J]. Physical Review Letters, 1977. DOI: 10.1103/PhysRevLett.38.1440.
- [85] WEINBERG S. A New Light Boson?[J]. Physical Review Letters, 1978. DOI: 10.1103/PhysRevLett.40.223.
- [86] WILCZEK F. Problem of Strong P and T Invariance in the Presence of Instantons[J]. Physical Review Letters, 1978. DOI: 10.1103/PhysRevLett.40.279.
- [87] SIKIVIE P. Experimental Tests of the "Invisible" Axion[J]. Physical Review Letters, 1983. DOI: 10.1103/PhysRevLett.51.1415.
- [88] DU N, et al. A Search for Invisible Axion Dark Matter with the Axion Dark Matter Experiment[J]. Physical Review Letters, 2020. DOI: 10.1103/PhysRevLett.124 .101303.
- [89] ZIOUTAS K, ANDRIAMONJE S, ARSOV V, et al. First results from the CERN axion solar telescope (CAST)[J]. Physics Letters B, 584(1-4): 260-271. DOI: 10 .1016/j.physletb.2004.01.066.
- [90] ARMENGAUD E, et al. Physics potential of the International Axion Observatory (IAXO)[J]. Journal of Cosmology and Astroparticle Physics, 2019. DOI: 10.108 8/1475-7516/2019/06/047.

- [91] IRWIN K, et al. Dark Matter Radio: A Quantum Sensor for Ultralight Dark Matter[J]. Annual Review of Nuclear and Particle Science, 2021. DOI: 10.1146/ann urev-nucl-102920-024617.
- [92] BALEWSKI J, et al. The DarkLight Experiment: A Precision Search for New Physics at Low Energies[J]. ArXiv:1412.4717 [hep-ex], 2014.
- [93] ILTEN P, SOREQ Y, THALER J, et al. Inclusive Dark Photon Search at LHCb[J]. Physical Review Letters, 2016. DOI: 10.1103/PhysRevLett.116.251803.
- [94] ARGUELLES C, KATORI T, SALVADO J. A sterile neutrino at MiniBooNE and IceCube[J]. Physics Letters B, 2014, 737: 310-314. DOI: 10.1016/j.physletb .2014.07.016.
- [95] AARTSEN M, others (IceCube Collaboration). Search for unstable heavy neutrinos with IceCube-DeepCore[J]. Physical Review Letters, 2021, 130(2). DOI: 10.1103/PhysRevLett.130.021802.
- [96] APRILE E, others (XENON Collaboration). Dark Matter Search Results from a One Ton-Year Exposure of XENON1T[J]. Physical Review Letters, 2018, 121(11). DOI: 10.1103/PhysRevLett.121.111302.
- [97] ANGLOHER G, others (CRESST Collaboration). Results on light dark matter particles with a low-threshold CRESST-II detector[J]. The European Physical Journal C, 2016, 76(1). DOI: 10.1140/epjc/s10052-016-3877-3.
- [98] AGNESE R, others (CDMS Collaboration). Search for Low-Mass Weakly Interacting Massive Particles with SuperCDMS[J]. Physical Review Letters, 2014, 112. DOI: 10.1103/PhysRevLett.112.241302.
- [99] AGNESE R, others (SuperCDMS Collaboration). Projected Sensitivity of the SuperCDMS SNOLAB Experiment[J]. Physical Review D, 2018, 95(8). DOI: 10.1 103/PhysRevD.95.082002.
- [100] COLLABORATION S. CDMS detector[Z]. https://www.slac.stanford.edu/exp /cdms/.
- [101] ABDELHAMEED A, others (CRESST Collaboration). First results from the CRESST-III low-mass dark matter program[J]. Physical Review D, 2019, 100(10). DOI: 10.1103/PhysRevD.100.102002.
- [102] STRAUSS R, ANGLOHER G, BAUER P, et al. A prototype detector for the CRESST-III low-mass dark matter search[J/OL]. Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and

Associated Equipment, 2017, 845: 414-417. https://www.sciencedirect.com/science/article/pii/S0168900216306039. DOI: https://doi.org/10.1016/j.nima.2016.06.060.

- [103] SAVAGE C, GELMINI G, GONDOLO P, et al. Compatibility of DAMA/LIBRA dark matter detection with other searches[J]. Journal of Cosmology and Astroparticle Physics, 2009. DOI: 10.1088/1475-7516/2009/04/010.
- [104] BERNABEI R, others (DAMA Collaboration). First model independent results from DAMA/LIBRA-phase2[J]. Nuclear Physics A, 2018, 979: 21-39. DOI: 10 .1016/j.nuclphysa.2018.09.002.
- [105] ADHIKARI G, others (COSINE-100 Collaboration). Search for a dark matterinduced annual modulation signal in NaI(Tl) with the COSINE-100 experiment[J]. Physical Review Letters, 2019. DOI: 10.1103/PhysRevLett.123.031 302.
- [106] COLLABORATION D. DAMA/LIBRA detector[Z]. https://arstechnica.com/sc ience/2019/02/more-bad-news-for-controversial-20-year-old-claim-of-dark-m atter-detection/.
- [107] AKERIB D, others (LUX-ZEPLIN Collaboration). Projected WIMP sensitivity of the LUX-ZEPLIN dark matter experiment[J]. Physical Review D, 2020, 101: 052002. DOI: 10.1103/PhysRevD.101.052002.
- [108] CUI X, others (PandaX-II Collaboration). Dark Matter Results From 54-Ton-Day Exposure of PandaX-II Experiment[J]. Physical Review Letters, 2017, 119: 181302. DOI: 10.1103/PhysRevLett.119.181302.
- [109] AMBROSI G, others (DAMPE Collaboration). Direct detection of a break in the teraelectronvolt cosmic-ray spectrum of electrons and positrons[J]. Nature, 2017, 552: 63-66. DOI: 10.1038/nature24475.
- [110] ACKERMANN M, others (Fermi-LAT Collaboration). Searching for Dark Matter Annihilation from Milky Way Dwarf Spheroidal Galaxies with Six Years of Fermi Large Area Telescope Data[J]. Physical Review Letters, 2015, 115: 231301. DOI: 10.1103/PhysRevLett.115.231301.
- [111] ABDALLAH H, others (H.E.S.S. Collaboration). Search for dark matter annihilations towards the inner Galactic halo from 10 years of observations with H.E.S.S.[J]. Physical Review Letters, 2016. DOI: 10.1103/PhysRevLett.117 .111301.

- [112] AGUILAR M, others (AMS Collaboration). Towards Understanding the Origin of Cosmic-Ray Positrons[J]. Physical Review Letters, 2019, 122: 041102. DOI: 10.1103/PhysRevLett.122.041102.
- [113] AARTSEN M, others (IceCube Collaboration). Search for annihilating dark matter in the Sun with 3 years of IceCube data[J]. The European Physical Journal C, 2017, 77: 146. DOI: 10.1140/epjc/s10052-017-4689-9.
- [114] SCHMITT M. Collider Signatures of Dark Matter[J/OL]. SLAC Summer Institute on Particle Physics (SSI), 2007. https://www-conf.slac.stanford.edu/ssi/200 7/talks/schmitt_080207.pdf.
- [115] LI Y F, HAHN R L, HAN K J, et al. Research progress at China Jinping Underground Laboratory[J]. Chinese Physics C, 2014, 38(1): 016001. DOI: 10.1088/1 674-1137/38/1/016001.
- [116] YUE Q, KANG K J, CHENG J P, et al. Introduction of the CDEX experiment[J]. Journal of Physics: Conference Series, 2013, 455(1): 012001. DOI: 10.1088/174 2-6596/455/1/012001.
- [117] YUE Q, KANG K J, CHENG J P, et al. Progress of Jinping Underground laboratory for Nuclear Astrophysics (JUNA)[J]. Journal of Physics: Conference Series, 2014, 515(1):012004. DOI: 10.1088/1742-6596/515/1/012004.
- YUE Q, KANG K, LI J, et al. The CDEX Dark Matter Program at the China Jinping Underground Laboratory[J/OL]. Journal of Physics: Conference Series, 2016, 718(4): 042066. https://dx.doi.org/10.1088/1742-6596/718/4/042066. DOI: 10.1088/1742-6596/718/4/042066.
- [119] CJPL. CJPL-II layout[Z]. https://cjpl.tsinghua.edu.cn/column/About_Facility. 2021.
- [120] KHAN M, KHAN A. Inert gas: A review of its applications in different industries[J]. Journal of Cleaner Production, 2019, 244. DOI: 10.1016/j.jclepro.2019 .118608.
- [121] MASON E, MONCHICK L. The Noble Gases: Introduction and History[M]. Amsterdam: Elsevier Science Ltd, 2004.
- [122] MARRODÁN UNDAGOITIA T, RAUCH L. Dark matter direct-detection experiments[J]. J. Phys. G, 2016, 43(1): 013001. arXiv: 1509.08767. DOI: 10.1088/0 954-3899/43/1/013001.

- [123] BRITANNICA E. Xenon | Definition, Properties, Atomic Mass, Compounds, & Facts[EB/OL]. Encyclopædia Britannica Inc. 2021. https://www.britannica.co m/science/xenon.
- [124] ENGEL J, MENÉNDEZ J. Status and future of nuclear matrix elements for neutrinoless double-beta decay: a review[J]. Reports on Progress in Physics, 2017.
 DOI: 10.1088/1361-6633/aa5bc5.
- [125] APRILE E, DOKE T. Liquid Xenon Detectors for Particle Physics and Astrophysics[J]. Rev. Mod. Phys., 2010, 82:2053-2097. arXiv: 0910. 4956 [physics.ins-det]. DOI: 10.1103/RevModPhys.82.2053.
- [126] XIANG X, WEI T, et AL. C X. Performance of Photosensors in the PandaX-I Experiment[J]. Journal of Instrumentation, 2015, 10(11): P11010. DOI: 10.1088 /1748-0221/10/11/P11010.
- [127] AALBERS J, et al. A next-generation liquid xenon observatory for dark matter and neutrino physics[J]. J. Phys. G, 2023, 50(1): 013001. arXiv: 2203.02309 [physics.ins-det]. DOI: 10.1088/1361-6471/ac841a.
- [128] HE C, et al. A 500 MS/s waveform digitizer for PandaX dark matter experiments[J]. JINST, 2021, 16(12): T12015. arXiv: 2108.11804. DOI: 10.1088/1 748-0221/16/12/T12015.
- [129] COTHERN C, SMITH JR. J. Environmental Radon[M]. Cham: Springer, 2018.
- [130] 王秋红. 利用 PandaX 二期探测器寻找暗物质[D]. 上海: 中国科学院上海应 用物理研究所, 2020.
- [131] 谢鹏伟. 利用 PandaX 一期二期探测器探测自旋不相关暗物质[D]. 上海: 上海交通大学, 2017.
- [132] ZHAO L, CUI X, MA W, et al. The cryogenics and xenon handling system for the PandaX-4T experiment[J]. JINST, 2021, 16(06): T06007. arXiv: 2012.10583. DOI: 10.1088/1748-0221/16/06/T06007.
- [133] CUI X, et al. Design and commissioning of the PandaX-4T cryogenic distillation system for krypton and radon removal[J]. JINST, 2021, 16(07): P07046. arXiv: 2012.02436. DOI: 10.1088/1748-0221/16/07/P07046.
- [134] HU J, GONG H W, LIN Q, et al. Development of the Liquid Level Meters for the PandaX Dark Matter Detector[J]. Chin. Phys. C, 2014, 38(5): 056002. arXiv: 1306.3064. DOI: 10.1088/1674-1137/38/5/056002.

- [135] JI P, CHEN X, TAN A, et al. A low-cost slow control system for the PandaX-4T experiment[J]. Radiat. Detect. Technol. Methods, 2019, 3(3): 53. DOI: 10.1007 /s41605-019-0130-2.
- [136] FOUNDATION A S. Apache Kafka[Z]. https://kafka.apache.org/. 2021.
- [137] TEAM T R. ROOT: analyzing petabytes of data, scientifically.[Z]. https://root.c ern.ch/. 2021.
- [138] Revive. Bamboo-shoot3[Z]. https://github.com/revive/bamboo-shoot3. 2020.
- [139] GROUP T P G D. PostgreSQL: The world's most advanced open source database[Z]. https://www.postgresql.org/. 2021.
- [140] FIELDING R. Architectural Styles and the Design of Network-based Software Architectures[J]. Doctoral dissertation, University of California, Irvine, 2000.
- [141] RICHARDSON L, RUBY S. RESTful Web Services Cookbook: Solutions for Improving Scalability and Simplicity[M]. 2nd. O' Reilly Media, 2013.
- [142] Facebook/Meta. React -A JavaScript library for building user interfaces[Z]. http s://reactjs.org/.
- [143] Material-UI. Mui-org/material-ui: MUI (formerly Material-UI) is the React UI library you always wanted. Follow your own design system, or start with Material Design.[Z]. https://github.com/mui-org/material-ui.
- [144] Material-Table. Mbrn/material-table: Datatable for React based on material-ui's table with additional features[Z]. https://github.com/mbrn/material-table.
- [145] Alibaba. BizCharts: Powerful data visualization library based on G2 and React[Z]. https://github.com/alibaba/BizCharts. 2021.
- [146] TEAM R. React Redux: Official React bindings for Redux[Z]. https://react-redu x.js.org/. 2021.
- [147] LabStack. Echo: High performance, minimalist Go web framework[Z]. https://gi thub.com/labstack/echo. 2021.
- [148] FOUNDATION O. Build cross-platform desktop apps with JavaScript, HTML, and CSS | Electron[Z]. https://www.electronjs.org/. 2021.
- [149] FOUNDATION A S. Features | EChartsJS[Z]. https://www.echartsjs.com/en/fe ature.html. 2021.
- [150] Yhirose. GitHub yhirose/cpp-httplib: A C++ header-only HTTP/HTTPS server and client library[Z]. https://github.com/yhirose/cpp-httplib. 2021.

- [151] Nlohmann. GitHub nlohmann/json: JSON for Modern C++[Z]. https://github.c om/nlohmann/json. 2021.
- [152] ZHANG D, et al. Horizontal position reconstruction in PandaX-II[J]. JINST, 2021, 16(11): P11040. arXiv: 2106.08380. DOI: 10.1088/1748-0221/16/11 /P11040.
- [153] LINDHARD J, SCHARFF M. Energy Dissipation by Ions in the kev Region[J]. Phys. Rev., 1961, 124: 128-130. DOI: 10.1103/PhysRev.124.128.
- [154] DAHL C E. The physics of background discrimination in liquid xenon, and first results from Xenon10 in the hunt for WIMP dark matter[D]. Princeton University, New Jersey, 2009.
- [155] DOKE T, HITACHI A, KIKUCHI J, et al. Absolute Scintillation Yields in Liquid Argon and Xenon for Various Particles[J]. Jap. J. Appl. Phys., 2002, 41: 1538-1545. DOI: 10.1143/JJAP.41.1538.
- [156] SZYDAGIS M, BARRY N, KAZKAZ K, et al. NEST: A Comprehensive Model for Scintillation Yield in Liquid Xenon[J]. JINST, 2011, 6: P10002. arXiv: 1106 .1613. DOI: 10.1088/1748-0221/6/10/P10002.
- [157] SZYDAGIS M, BALAJTHY J, BRODSKY J, et al. Noble Element Simulation Technique[CP]. v2.1.0. Zenodo, 2020. DOI: 10.5281/zenodo.3905382.
- [158] AGOSTINELLI S, et al. GEANT4–a simulation toolkit[J]. Nucl. Instrum. Meth. A, 2003, 506: 250-303. DOI: 10.1016/S0168-9002(03)01368-8.
- [159] YAN B, et al. Determination of responses of liquid xenon to low energy electron and nuclear recoils using a PandaX-II detector[J]. Chin. Phys. C, 2021, 45(7): 075001. arXiv: 2102.09158. DOI: 10.1088/1674-1137/abf6c2.
- [160] KUDRYAVTSEV V A, ZAKHARY P, EASEMAN B. Neutron production in (α, n) reactions[J]. Nucl. Instrum. Meth. A, 2020, 972: 164095. arXiv: 2005.0 2499. DOI: 10.1016/j.nima.2020.164095.
- [161] HE D D, JING S, ZHENG Y L. Design and optimization of thermal neutron device based on deuterium-deuterium neutron generator[J]. Fusion Engineering and Design, 2021, 166: 112289.
- [162] QIAN Z, et al. Low radioactive material screening and background control for the PandaX-4T experiment[J]. JHEP, 2022, 06: 147. arXiv: 2112.02892. DOI: 10.1007/JHEP06(2022)147.

- [163] MA W, et al. Internal calibration of the PandaX-II detector with radon gaseous sources[J]. JINST, 2020, 15(12): P12038. arXiv: 2006.09311. DOI: 10.1088/17 48-0221/15/12/P12038.
- [164] CHEN X, et al. BambooMC A Geant4-based simulation program for the PandaX experiments[J]. JINST, 2021, 16(09): T09004. arXiv: 2107.05935. DOI: 10 .1088/1748-0221/16/09/T09004.
- [165] WILSON W, PERRY R, CHARLTON W, et al. Sources: A code for calculating (alpha, n), spontaneous fission, and delayed neutron sources and spectra[J]. Progress in Nuclear Energy, 2009, 51(4): 608-613. DOI: https://doi.org/10.1016 /j.pnucene.2008.11.007.
- [166] WANG Q, et al. An Improved Evaluation of the Neutron Background in the PandaX-II Experiment[J]. Sci. China Phys. Mech. Astron., 2020, 63(3): 231011. arXiv: 1907.00545. DOI: 10.1007/s11433-019-9603-9.
- [167] HUANG Z, et al. Neutron-induced nuclear recoil background in the PandaX-4T experiment*[J]. Chin. Phys. C, 2022, 46(11): 115001. arXiv: 2206.06087. DOI: 10.1088/1674-1137/ac8539.
- [168] AGOSTINI M, ALTENMÜLLER K, APPEL S, et al. Comprehensive measurement of pp-chain solar neutrinos[J]. Nature, 2018, 562(7728): 505-510.
- [169] AHARMIM B, AHMED S, ANTHONY A, et al. Combined analysis of all three phases of solar neutrino data from the Sudbury Neutrino Observatory[J]. Physical Review C, 2013, 88(2): 025501.
- [170] BILLARD J, STRIGARI L, FIGUEROA-FELICIANO E. Implication of neutrino backgrounds on the reach of next generation dark matter direct detection experiments[J]. Phys. Rev. D, 2014, 89(2): 023524. arXiv: 1307.5458. DOI: 10.1103/P hysRevD.89.023524.
- [171] ZHANG D. Estimating the surface backgrounds in PandaX-II WIMP search data[J]. JINST, 2019, 14(10): C10039. DOI: 10.1088/1748-0221/14/10/C10039.
- [172] WANG Q, et al. Results of dark matter search using the full PandaX-II exposure[J]. Chin. Phys. C, 2020, 44(12): 125001. arXiv: 2007.15469. DOI: 10.1088 /1674-1137/abb658.
- [173] MENG Y, et al. Dark Matter Search Results from the PandaX-4T Commissioning Run[J]. Phys. Rev. Lett., 2021, 127(26): 261802. arXiv: 2107.13438. DOI: 10.11 03/PhysRevLett.127.261802.

- [174] COWAN G, CRANMER K, GROSS E, et al. Asymptotic formulae for likelihoodbased tests of new physics[J]. Eur. Phys. J. C, 2011, 71: 1554. arXiv: 1007.1727. DOI: 10.1140/epjc/s10052-011-1554-0.
- [175] APRILE E, AALBERS J, AGOSTINI F, et al. Dark Matter Search Results from a One Ton-Year Exposure of XENON1T[J]. Phys. Rev. Lett., 2018, 121: 111302.
 DOI: 10.1103/PhysRevLett.121.111302.
- [176] AKERIB D, et al. Results from a search for dark matter in the complete LUX exposure[J]. Phys. Rev. Lett., 2017, 118(2): 021303. DOI: 10.1103/PhysRevLett .118.021303.
- [177] HOCKER A, et al. TMVA Toolkit for Multivariate Data Analysis[J]., 2007. arXiv: physics/0703039.
- [178] PEDREGOSA F, VAROQUAUX G, GRAMFORT A, et al. Scikit-learn: machine learning in Python[EB/OL]. 2011. https://scikit-learn.org/stable/index.html.
- [179] SZYDAGIS M. The Noble Element Simulation Technique (NEST): Recent Updates and Improvements[J]. Bulletin of the American Physical Society, 2022.
- [180] GOODFELLOW I J, POUGET-ABADIE J, MIRZA M, et al. Generative Adversarial Networks[J]. Advances in Neural Information Processing Systems (NIPS), 2014.
- [181] KINGMA D P, DHARIWAL P. Glow: Generative Flow with Invertible 1x1 Convolutions[J]. Advances in Neural Information Processing Systems (NIPS), 2018.
- [182] KINGMA D P, WELLING M. An Introduction to Variational Autoencoders[J]. Foundations and Trends in Machine Learning, 2019. DOI: 10.1561/2200000056.
- [183] SOHL-DICKSTEIN J, WEISS E A, MAHESWARANATHAN N, et al. Deep Unsupervised Learning using Nonequilibrium Thermodynamics[J]. Proceedings of the International Conference on Machine Learning (ICML), 2015.
- [184] 阿布都沙拉木•阿布都克力木谌勋吾尔尼沙•依明尼亚孜. 剖面似然比统 计及 CLS 方法在 PandaX-II 暗物质实验数据分析中的应用[J]. 上海交通大 学误报, 2019(1508-1514). DOI: 10.16183/j.cnki.jsjtu.2019.12.015.
- [185] BAXTER D, et al. Recommended conventions for reporting results from direct dark matter searches[J]. Eur. Phys. J. C, 2021, 81(10): 907. arXiv: 2105.00599 [hep-ex]. DOI: 10.1140/epjc/s10052-021-09655-y.
- [186] TEAM T R. RooStats Tutorial[Z]. https://twiki.cern.ch/twiki/bin/view/RooStats /WebHome. 2020.

- [187] TEAM T S H. Statistics tools and utilities —hepstats[Z]. https://scikit-hep.org/h epstats/. 2021.
- [188] TEAM T S H. Pure-python fitting/limit-setting/interval estimation histfactory[Z]. https://scikit-hep.org/pyhf/. 2021.
- [189] TEAM T R. HistFactory XML Reference[Z]. https://twiki.cern.ch/twiki/bin/vie w/RooStats/HistFactoryXMLReference. 2012.
- [190] APRILE E, AALBERS J, AGOSTINI F, et al. Search for Coherent Elastic Scattering of Solar 8B Neutrinos in the XENON1T Dark Matter Experiment[J]. Physical review letters, 2021, 126(9): 091301.
- [191] VINYOLES N, SERENELLI A M, VILLANTE F L, et al. A new generation of standard solar models[J]. The Astrophysical Journal, 2017, 835(2): 202.
- [192] RUPPIN F, BILLARD J, FIGUEROA-FELICIANO E, et al. Complementarity of dark matter detectors in light of the neutrino background[J]. Physical Review D, 2014, 90(8): 083510.
- [193] APRILE E, et al. Light Dark Matter Search with Ionization Signals in XENON1T[J]. Phys. Rev. Lett., 2019, 123(25): 251801. arXiv: 1907.11485. DOI: 10.1103/PhysRevLett.123.251801.

致 谢

时光荏苒,转眼间我即将要完成博士学业,感谢在这段旅程中给予我指导、支持、帮助的人们。

首先,我要感谢我的导师刘江来教授。在我的博士研究期间,我有幸成为了 刘江来老师的学生,他是一位杰出的粒子物理实验学家。他的勤奋和拼搏精神深 深地感染了我,让我知道科研需要的不仅仅是天赋,更需要毅力和耐心。他的工 作态度让我深受感染,每一次与他的交流,都是我在不断成长和提高自己的过程 中的关键一步。刘江来老师具有极高的责任感和奉献精神,在 PandaX-4T 探测器 搭建过程中,他不仅积极参与所有的实验设计和数据分析,还积极引领我们的研 究方向,让我们更好地融入研究团队,感受到科研的乐趣。他的言传身教,教给 了我很多关于如何做人和如何从事学术研究的宝贵经验和启示,这些都将对我今 后的学术和职业生涯产生积极的影响。他对科学研究的执着和勤奋一直是我学习 的榜样。

在此,我要特别感谢我的另一位导师谌勋老师。作为我的导师,谌老师不仅 在学术上给予了我无微不至的指导,而且在实践方面也给了我极大的帮助,尤其 是在编程方面。谌老师在编程方面的能力非常强,他在我写代码时给了我非常有 用的建议和指导。他不仅仅传授给我知识和技能,还教给我思考和解决问题的方 法,让我在项目中能够快速地定位和解决各种问题。同时,谌老师还鼓励我学习 新技术,让我拥有更广阔的视野和更高的技术水平。除此之外,谌老师还在我的 关于偶然符合本底相关的一篇论文中给了我非常重要的帮助和指导,让我更加深 入地理解了研究课题。谌老师严谨的治学态度和精益求精的工作作风,深深地影 响了我,让我在今后的研究工作中更加注重细节、追求卓越。在生活方面,谌老 师也一直给予了我非常多的关心和照顾。无论是在项目中遇到问题还是个人生活 中遇到困难,谌老师都会积极地给我建议和帮助。在我完成论文的过程中,谌老 师还给了我很多鼓励和支持。谌勋老师是一位敬业专注的科学家,也是一位亲切 友好的导师。我非常感谢他对我学术和生活上的关怀和支持,他对我的帮助将会 伴随我一生。

除了我的两位导师刘江来老师和谌勋老师,我还有幸得到了其他几位老师的 帮助和指导,特别是在数据分析方面,他们为我的研究工作提供了非常有力的支 持。其中,周宁老师是一位同时在理论和实验方面都有丰富经验的专家。他在我 研究的过程中提供了极有价值的数据分析建议,帮助我更加深入地理解研究课题, 为我后续的研究工作提供了宝贵的思路和方向。林箐老师是一位严谨认真、责任 心强的科学家。他在我的实验数据分析过程中给予了我极大的帮助,通过他的指 导,我不仅提高了数据分析的能力,还深刻地认识到了科研工作的严谨性和细致 性。孟月老师是一位非常有爱心、关怀学生的老师。她在我的研究工作中给予了 我极有价值的指导、帮助、鼓励和支持。这里我还非常感谢实验团队的季向东老 师、陈葳老师、王舟老师、肖翔老师、杨勇老师、韩柯老师、张涛老师、赵力老 师、周小朋老师。您们的专业知识、热情和耐心,让我在研究过程中不断进步,真 正体验到了团队合作的力量。

感谢谢鹏伟、谈安迪、王秋红、燕斌斌、崔翔仪、林业、李涛、陶奕、夏经铠 等博士后的悉心指导和支持。他们在我的研究工作中给予了我许多宝贵的经验和 启示,为我今后的科研之路打下了坚实的基础。我会在他们的思想和研究中寻找 灵感和指引,从而更好地解决问题和提高自己的研究能力。我在他们的指导和帮 助下,收获了许多珍贵的经验和成果。

在这里,我想表达对所有同学的感激之情,你们是我在学术路上不断成长的 良师益友。我想特别感谢杜海燕、程晨、申国防、黄迪、张丹、李帅杰、袁影、马 文博、武蒙蒙、黄周、宁旭阳、Nasir、尚晓凤、杨继军、曾鑫宁、司琳、钱志成、 苑哲、顾琳慧、何昶达、薄子豪、李家富、芦晓盈、罗棱尹、王秀丽、张瀛心、樊 英杰、周钰博、颜玺雨、李晨翔等同学。我们一起共同经历了无数次科研的挑战和 胜利。感谢你们对我科研上的支持和鼓励,感谢你们在学术上的指导和帮助。在 这里,我衷心祝愿你们在未来的学术和生活道路上一路顺风,再次感谢你们的陪 伴和帮助。

母亲是世界上最伟大的职业之一。她是我的第一位导师,教会我如何做人,如 何爱人,如何为人处世。母亲对我的影响是深远而持久的,她的爱和关怀贯穿了 我整个人生的每个阶段。她是我心灵上的支柱,是我生活中的精神家园。在攻读 博士期间,母亲也一直在关心我,鼓励我,支持我。即使身在远方,她仍然会每 天给我打电话,关心我的学业和生活。她总是鼓励我,让我有勇气面对困难和挑 战。母亲的关爱和支持是我在攻读博士期间最重要的精神动力之一。我要感谢她 无私的付出和爱心,让我成为今天的自己。

最后,非常感谢我的妻子。我真的非常感激她在我生命中的出现。在我求学和科研的过程中,她一直默默地支持我,从来没有抱怨过。她总是在我需要她的时候出现在我身边,给我无私的关爱和支持。在这充满挑战和压力的博士研究生生涯中,她为我付出了很多,还时常为我加油打气,给我鼓励和支持,帮我减轻研究生活的压力,使我更专注于科研工作。在未来的岁月里,我会继续努力,为
我们的未来奋斗,为她和我所珍视的生活而努力。我也希望我的工作能够取得更 多的成果,为我们的未来打下更加坚实的基础。

在我的博士研究生生涯即将落下帷幕之际,最后,我再次感谢所有支持和帮助过我的人们。你们的付出和支持,让我能够顺利地完成这项艰苦而充实的学术 探索之旅。

学术论文和科研成果目录

学术论文

- [1] 第二作者: First Search for Solar ⁸B Neutrino in the PandaX-4T Experiment using Neutrino-Nucleus Coherent Scattering[J]. Phys.Rev.Lett., 2023, 130(2): 021802.
- [2] 第一作者: Study of Background from Accidental Coincidence Signals in the PandaX-II Experiment[J]. Chin.Phys.C (2022), 46(10): 103001.
- [3] 有重要贡献的文章: Search for Light Dark Matter with Ionization Signals in the PandaX-4T Experiment[J]. Submitted to Phys.Rev.Lett. (arxiv:2212.10067).
- [4] 有重要贡献的文章: Dark Matter Search Results from the PandaX-4T Commissioning Run[J]. Phys.Rev.Lett. 127 (2021) 26: 261802.
- [5] 有重要贡献的文章: Results of Dark Matter Search using the Full PandaX-II Exposure[J]. Chin.Phys.C 44 (2020) 12: 125001.