

# SHANGHAI JIAO TONG UNIVERSITY



# **BACHELOR'S THESIS**



# 论文题目:<u>利用 PandaX 500 公斤液体氙探</u> 测器寻找无中微子双贝塔衰变

学生姓名:	赖奕辉	
学生学号:	515072910041	
专业:	物理学	
指导教师:	韩柯	
学院(系): _	物理与天文学院	



# 变

### 摘要

在过去的几十年中,科学界对中微子这种微小且难以捉摸的粒子越来越感兴趣,且在 中微子及其性质的研究中获得了许多非常重要的结果。中微子振荡实验证实中微子具有质 量,这是目前在其他地方都很成功的粒子物理"标准模型"唯一预测失败的地方。目前,还有 很多关于中微子性质的问题没有弄清楚:中微子是狄拉克还是马约拉纳粒子?它们的质量 序是多少?而寻找无中微子双贝塔衰变现象是回答这些问题最直接的方法。搜寻无中微子 双贝塔衰变将有助于我们理解这个复杂的世界,包括揭示宇宙中物质与反物质不对称的奥 秘。

目前国际上有多个粒子物理实验合作组正在紧锣密鼓的推进无中微子双贝塔衰变现象的搜寻。而 PandaX-II 是中国的暗物质探测实验,虽然其主要科学目标是直接捕捉暗物质候选粒子弱相互作用大质量粒子和氙核的可能碰撞,但是其中含有的大量<sup>136</sup>Xe元素使得PandaX-II 探测器也可以用于寻找无中微子双贝塔衰变。

本文的主要工作就是利用 PandaX-II 在过去几年时间里积累到的大量数据,进行无中微 子双贝塔衰变的搜寻。工作重点在于重建无中微子双贝塔衰变感兴趣区域的能谱以及构建 本底模型。通过合理的数据筛选方法提高探测器的分辨率,并且进行误差分析,最终得到 PandaX-II 对于无中微子双贝塔衰变的最佳探测结果和半衰期限制。

关键词: 无中微子双贝塔衰变, 双相型液氙探测器, 稀有事件探测



# NEUTRINO-LESS DOUBLE BETA DECAY SEARCH WITH PANDAX 500 KILOGRAM LIQUID XENON DETECTOR

## ABSTRACT

In the past few decades, scientists have become more and more interested in the elusive neutrino particles, and they have made many important achievements in the study of neutrinos and their properties. Neutrino oscillations confirm that neutrinos have mass, which is exactly what the standard model of particle physics failed to predict. At present, many questions remain about the nature of neutrinos: are they Dirac or Majorana? What is their mass order? Finding neutrino-less double beta decay ( $0\nu\beta\beta$ ) is the most straightforward way to answer these questions. Searching for  $0\nu\beta\beta$  will help us understand the complex world, including revealing the asymmetry between matter and antimatter in the universe.

At present, some of the world's frontier particle physics experimental cooperation team is carrying on the  $0\nu\beta\beta$  phenomenon research. PandaX-II is a dark matter detection experiment in China. Although its main scientific goal is to directly capture the possible collisions of the dark matter candidate particles WIMP (weakly interacting massive particles) and the xenon nucleus, the large amount of <sup>136</sup>Xe isotope contained in it makes it possible to detect  $0\nu\beta\beta$ .

The main work of this paper is to use the large amount of data of PandaX-II to search for the  $0\nu\beta\beta$ . The focus of the work is to reconstruct the energy spectrum of the  $0\nu\beta\beta$  energy region and to fit the spectrum with a background model. The best detection result and half-life limit of PandaX-II on  $0\nu\beta\beta$  are obtained by reasonable data analysis techniques.

Key words: neutrino-less double beta decay, liquid xenon detector, rare event detection



第一章	绪论	1
1.1	中微子简介	1
1.2	双贝塔衰变	1
	1.2.1 双中微子双贝塔衰变(2νββ)	1
	1.2.2 无中微子双贝塔衰变(0νββ)	3
	1.2.3 探测方法	4
	1.2.4 研究现状	5
1.3	本文结构安排	7
第二章	PandaX-II 实验简介	8
2.1	基本介绍	8
2.2	0vββ分析可行性	8
2.3	时间投影室	8
	2.3.1 探测原理	9
	2.3.2 发光机制	10
	2.3.3 电子漂移与萃取	10
2.4	PandaX-II 的数据积累	10
第三章	PandaX-II 数据分析	12
3.1	数据处理流程	12
3.2	数据筛选	13
	3.2.1 基础事例筛选和信号质量筛选	13
	3.2.2 筛选后的能谱	17
3.3	本章小结	18
第四章	PandaX-II 0vββ 探测结果	19
4.1	本底模型建立	19
	4.1.1 拟合函数	19
	4.1.2 <sup>222</sup> Rn 和 <sup>60</sup> Co 对于本底的贡献	20
4.2	能谱拟合	20
	4.2.1 贝叶斯表述	20
	4.2.2 程序结构	21
	4.2.3 0vββ 拟合结果	22
4.3	系统误差	23
4.4	0νββ半衰期下限	24
	4.4.1 从最佳拟合到 0vββ 半衰期下限	24
4.5	0νββ 灵敏度	25
	4.5.1 灵敏度预估	26
	4.5.2 利用 Toy MC 计算	26
4.6	本章小结	27
第五章	结论	28
参考文	献	29





# 第一章 绪论

中微子可能是人类最难以想象却又真实存在的粒子了,它数量庞大却又难以与物质发生 相互作用,因此也被戏称为"幽灵粒子"。自1930年中微子被提出以来,关于中微子性质 的研究不断,而20世纪末中微子振荡现象的发现更是激起了物理学家对这个"幽灵粒子"的 兴趣。因为依据粒子物理的标准模型<sup>[1]</sup>,中微子的静质量为零,而中微子的振荡现象恰恰表 明中微子的静质量不为零。自从希格斯粒子被发现以来,标准模型被认为是很完善的,而中 微子静质量不为零这一事实无疑是对标准模型提出的巨大挑战。针对目前存在的疑问,寻找 无中微子双贝塔衰变(0νββ)是直击靶心的一种手段。一旦这个现象被发现,中微子将被证 实为马约拉纳粒子,同时标准模型所描述的轻子数守恒也被违反,这些都将对解答目前的疑 问以及提出新的理论有极大的帮助。

# 1.1 中微子简介

1930年,为了解释  $\beta$  衰变实验上出现的统计问题,泡利(Pauli)引入了一种新的电中 性粒子,当时他把这种粒子叫做"中子",这便是中微子最早的出现。而在中子被发现后,泡 利提出的这个粒子则被费米(Fermi)更名为中微子。1933年,费米提出了  $\beta$  衰变是原子核 内一个中子通过弱相互作用衰变成一个质子,一个电子,一个中微子。1956年,Cowan 和 Reines 通过探测核反应堆里的反  $\beta$  衰变放出的电子和反中微子第一次直接观测到中微子的 存在。对于电子、μ子以及  $\tau$  子来说,它们都有对应的中微子:电子中微子(v<sub>e</sub>)、μ子中微子 (v<sub>µ</sub>)以及  $\tau$  子中微子(v<sub>r</sub>),目前它们的质量排序和绝对质量都还是未知的。在 1962年,Lederman, Steinberger 和 Schwartz 发现了 v<sub>µ</sub>,而 v<sub>r</sub>则是直到 2001年才在 DONUT 实验中被发现。

在 20 世纪 60 年代末, Homestake 实验观察到了太阳中微子的意外缺失,测量到的太阳 中微子只有标准太阳模型预测的 1/3,这成为了中微子振荡的第一个证据。现在我们知道, 中微子具有两组不同的本征态:质量本征态和味道本征态。味道本征态是指前面提到的三种 不同的"味道":  $v_e$ 、 $v_{\mu}$ 、 $v_{\tau}$ 。在中微子的传播过程当中,其相因子与质量本征态有关,因 此味道本征态在传播过程中是变化的。一个中微子在初始可能是  $v_e$ ,但在传播了一段的距离 之后可能转化为了  $v_{\mu}$ 或  $v_{\tau}$ 。尽管标准模型内中微子被认为是没有质量的粒子,但中微子振 荡的事实显著的说明中微子不是无质量粒子。

研究中微子振荡的实验给出了关于这种现象的清晰图像,但他们同时也提出了几个关于 中微子质量的问题:为什么中微子质量如此之小?中微子是狄拉克还是马约拉纳粒子?中微 子质量序是什么?发现 0vββ 衰变,或者只是给出更严格的限制将有助于解决所有这些关于 中微子物理的未知问题。

## 1.2 双贝塔衰变

1.2.1 双中微子双贝塔衰变(2νββ)

β 衰变是一种在自然界很普遍的核衰变过程,是放射性原子核内的一个中子转化为一个 质子,同时放射电子和反电子中微子而转变为另一种核的过程:

$$(Z, A) \to (Z+1, A) + e^- + \bar{\nu}_e$$
 (1-1)

第1页共30页



而双 β 衰变是一种二阶弱相互作用过程,在 1935 年,Maria Goeppert-Mayer 首次推测 存在双中微子双 β 衰变 ( $2\nu\beta\beta$ ),在这个过程中原子核内原子数不变,两个中子同时转化为 质子并且释放出两个 $e^{-}\bar{\nu}_{e}$ 对:

$$(Z, A) \to (Z + 2, A) + 2e^{-} + 2\bar{\nu}_e$$
 (1-2)

图 1.1(a)中的费曼图描绘了这个过程。但是可以容易预料到的是,这个过程会被高度抑制。因为如果一个原子核可能发生 2νββ,那它往往更加可能发生的是β衰变。在一个可能发生β衰变的不稳定的原子核中,2νββ实际上是不存在的。



图 1-1 费曼图:(a) 双中微子双贝塔衰变和(b) 无中微子双贝塔衰变

但 2vββ 的发生并不是完全没有希望,在一些"偶-偶"原子核(质子和中子数均为偶数) 中就有一些有趣的现象发生。对于"偶-偶"核,如果发生式 1-1 所示的 β 衰变,则会转变为 一个"奇-奇"核(质子和中子数均为奇数)。然而,由于配对效应,偶数个的质子和中子具有 更高的结合能。因此,"偶-偶"核往往比它对应的"奇-奇"核,或者质量数相似但为奇数的原 子核更稳定。如图 1-2(b)中的核质量等压线所示,有奇数质量数的同位素更加容易发生β衰 变而不是 2vββ。同理,质量数为偶数的同位素,如果它的质子数和中子数都是奇数,也可 以很容易地发生β衰变而变为更稳定的"偶-偶"核。

但对于质子数和中子数均为偶数的同位素,如果它发生β衰变,则会衰变到能量更高的态。根据能量守恒定律,这种转变是被禁止。也就是说,如果"偶-偶"核要发生衰变,它必须得通过 2vββ 才能实现到更稳定的状态的转变,如图 1-2(a)所示。



# 图 1-2 同位素核质量等压线示意图,具有(a)偶数和(b)奇数质量数。 对于一些具有偶数质子和中子数的原子核,仅允许双β衰变通道,如图(a)所示

尽管 2νββ 还是有机会发生的,但通过上面的分析,2νββ 的发生几率将是非常小。在具有奇数质量数的核中,相对于β衰变,2νββ 被抑制了约10<sup>21</sup>倍。而在那些2νββ 是唯一可能



的衰变途径的同位素中,两个核子的同时转换也依旧是不太可能发生的。事实上在 1987 年 人们第一次通过时间投影室观察到  $^{82}$ Se 的 2νββ 时,它是当时实验上测到的最慢的自然过程 了。

#### 1.2.2 无中微子双贝塔衰变 (0vββ)

正如之前讨论过的,中微子不是标准模型所预期的无质量粒子,而中微子如何获得它们的质量也依然是未解之谜。这里面一个核心的问题就是中微子的质量是否来源于马约拉纳质量机制,也就是说,我们得搞清楚中微子是不是马约拉纳费米子。在实验上,要决定一个粒子是不是马约拉纳费米子,得通过发现或者排除一些只有马约拉纳费米子才会发生的物理过程。

在 1939 年, Wendell Furry 发现如果中微子是马约拉纳粒子, 那么只放出电子的双 β 衰 变就有可能发生:

$$(Z, A) \to (Z + 2, A) + 2e^{-}$$
 (1-3)

这个反应过程便是 0vββ,如图 1.1(b)所示。鉴于马约拉纳费米子是它自己的反粒子,我 们可以把这种衰变当成是特殊的 2vββ 衰变,在这当中,中微子恰巧与其自身湮灭。而标准 模型是禁止这样的反应发生的,因为它违反轻子数守恒,同样如果中微子是狄拉克费米子的 话,0vββ 也是被禁止的。换言之,0vββ 是用来验证中微子是否为马约拉纳费米子的一个很 好的方法。

就其科学意义而言, 0vββ 在许多方面都很吸引人。首先这个过程直接表明了中微子的 马约拉纳性。其次,这个过程中的轻子数不守恒将违反标准模型,这个反应可以为轻子创世 论(Leptogenesis)提供更多的支持,一种用来解释宇宙中物质比反物质多的理论。因为不 管导致 0vββ 发生的机制是什么,这个过程都违反了轻子数守恒。观察到 0vββ 意味着轻子数 不拥有自然的对称性,这表明轻子在宇宙中物质和反物质不对称的产生中发挥了重要作用。 再者,0vββ 将解释中微子质量的来源和性质,因为它对中微子的绝对质量特别敏感。如果 0vββ 得以发现,基础物理理论将会有重大的突破。

依据轻马约拉纳中微子交换(light Majorana neutrino exchange)理论,通过实验可观察 到的 0vββ 衰变速率(或半衰期)可以用中微子的有效马约拉纳质量< $m_{\beta\beta}$  >来表示:

$$T_{1/2}^{0\nu}{}^{-1} = G^{0\nu} |M^{0\nu}|^2 \frac{\langle m_{\beta\beta} \rangle^2}{m_e^2}$$
(1-4)

$$T_{1/2}^{0\nu} = \frac{\ln 2}{r_{0\nu}} \tag{1-5}$$

其中 $\Gamma_{0\nu}$ 是衰变速率,相对应的 $T_{1/2}^{0\nu}$ 是  $0\nu\beta\beta$  半衰期, $G^{0\nu}$ 是衰减相空间因子(decay phase-space factor), $M^{0\nu}$ 是核矩阵元素, $m_e$ 是指电子质量。 $G^{0\nu}$ 与 $Q_{\beta\beta}$ 的5次方成正比,所以 $Q_{\beta\beta}$ 越高意味着衰变发生的概率越大( $Q_{\beta\beta}$ 指  $0\nu\beta\beta$ 放出的能量)。不同于 $G^{0\nu}$ 是可以精确计算的, $M^{0\nu}$ 里的矩阵元素是理论上最大的不确定量,而且不能在实验上测得。 $M^{0\nu}$ 的计算是一个复杂的多体问题,但是通过一些理论近似,这个值的计算已经得到了很大的改善。

而中微子的有效马约拉纳质量和它的质量本征态的关系如下:

$$m_{\beta\beta} = |\sum_{i} m_{i} U_{ie}^{2}| \tag{1-6}$$

其中*U<sub>ie</sub>*是电子中微子混合矩阵的元素,表示各中微子中间的转换,中微子振荡实验可以测出矩阵各元素的值。

总的来说,我们只要知道了中微子的质量差和 0vββ 半衰期,就能通过式 1-4 推导出中 微子的绝对质量。

#### 第3页共30页



1.2.3 探测方法

寻找 0vββ 不是一种新的研究热点,而是已经超过 80 年的持续努力。在 19 世纪 30 到 40 年代,弱相互作用下的宇称破缺还没被发现,因此计算出来 0vββ 发生的概率远高于 2vββ, 其半衰期范围错误的被估计在 10<sup>15</sup> ~ 10<sup>16</sup> 年之间。在接下来的十几年内,物理学家进行了一 些尝试,试图找到 0vββ,但都以失败告终。接着在 1957 年,弱相互作用下宇称破缺的发现 表明即使中微子是马约拉纳费米子,由于手征性不变的原因 0vββ 还是被禁止发生的,因此 0vββ 不能被用来分辨中微子为哪种费米子,相关的实验进入了低潮期。直到中微子振荡现 象的发现说明中微子质量不为零并且不同味道的中微子之间存在转换关系,这个事实直接挑 战了标准模型,使得中微子的手征性变化变成可能,物理学家又重新寻找起 0vββ 来。

为了使 0vββ 探测实验的灵敏度最大化,我们必须针对其独特的特征优化实验条件。因为两个未探测到的中微子带走了任意比例的"发射能量",所以 2vββ 的能谱为从 0 到 $Q_{\beta\beta}$ (完全衰变能量)之间的钟形连续谱。而 0vββ 仅输出两个电子,因此它们的总能量必须始终等于 $Q_{\beta\beta}$ 。所以,0vββ 的能谱特征是在 $Q_{\beta\beta}$ 值处有一个单能峰。两者的区别如图 1-3 所示。





图 1-3 中的双 β 衰变能谱表示的只是理想情况,而不是探测器实际得到的能谱。在现实 生活中,我们必须考虑信号统计和背景噪声等因素。实验中得到的能谱往往是一个 2vββ 能 谱的拖尾上有一个很小的 0vββ 鼓包,再考虑到有背景噪声,我们就更难观测到 0vββ 事例了。 因此,这要求我们对于候选同位素和探测技术有谨慎的选择,一是提升灵敏度,二是降低相 对的本底。图 1-4 总结了可能发生 0vββ 的同位素的*Q<sub>ββ</sub>*值和自然丰度。表 1-1 给出了 4 种在 实验上使用最多的同位素的更多信息。通常情况下,实验上选择同位素时会尝试最大化这两 个参数。



图 1-4 0νββ 候选同位素总结: Q<sub>ββ</sub>与自然丰度

第4页共30页



通常在 2.5 MeV 以下的能量区间存在一系列铀和钍的 γ 峰以及相应的康普顿散射谱, 所 以高*Q<sub>ββ</sub>*值能使得我们的感兴趣区域(region of interest)处于本底较低的区域。而低本底对 于这样一个罕见事件探测的实验是非常重要的,因为在最后我们很可能只会在*Q<sub>ββ</sub>*值处看到 一两个事例。如果本底太大,信号很容易就被淹没而找不到了。同样因为事件发生率很低, 并且由于实验预算往往有限,我们希望使用自然丰度高的同位素。即使我们可以对自然丰度 低的原料进行浓缩,我们也希望最初的自然丰富度足够高。

表1-1 常用0vββ候选同位素的 $Q_{BB}$ 值,自然丰度,0vββ半衰期下限和2vββ半衰期测量值

同位素	$Q_{\beta\beta}$ (keV)	a <sub>I</sub> (%)	$T_{1/2}(0v\beta\beta)(10^{25} y)$	$T_{1/2} (2\nu\beta\beta) (10^{21} y)$
<sup>76</sup> Ge	2039	7.61	>8.0	1.65
$^{100}$ Mo	3034	9.63	>0.11	0.0071
<sup>130</sup> Te	2528	34.17	>1.5	0.69
<sup>136</sup> Xe	2458	8.87	>11	2.19

此外,探测器性能也是非常重要的一个因素。比如在真实世界中,我们实验仪器的分辨 能力肯定是有限的, *Q<sub>ββ</sub>*处的单能峰的展宽会使得 0vββ 与 2vββ 无法分清。一般而言,固体 探测器的分辨率会很好,例如高纯锗探测器、低温量能器。而液体探测器的分辨率要比固体 探测器差的多。同时,一些探测技术也可以实现本底的剔除,使得更有可能探测到 0vββ。 而一些有着拓扑特性识别能力的探测器更是可以去除大部分本底。

最终,对于 0νββ 探测,所有实验都会是朝着优化式 1-7 的方向努力。

$$T_{1/2}^{0\nu} \propto a_I \sqrt{\frac{m_d \cdot t}{\Delta E \cdot b}} \tag{1-7}$$

式中a<sub>1</sub>是同位素的自然丰度, m<sub>d</sub>是总质量, t 是实验时长, b 是本底事例率。ΔE 是感兴趣区域, 但是由于这个值是和能量分辨率有关的, 我们可以把它简单的当成能量分辨率来理解。这个式子告诉我们 0vββ 的半衰期与曝光量(同位素质量与测量时间的乘积)的平方根成正比, 与本底事例率和能量分辨率的平方根成反比。

我们也能从这个式子中感受到为什么 0vββ 探测实验要有一个大的改进是如此困难。事 实上,实验可以起作用的所有因素(质量,暴露时间,能量分辨率和本底事例率)都是在平 方根号内,如果想要将实验的灵敏度提升 10 倍,我们必须建立一个比原来大 10<sup>2</sup> 倍的探测 器,而且能量分辨率和本底事例率都不能变差。

1.2.4 研究现状

0vββ 的探测器技术和同位素选择是密不可分的。例如,一般而言固态的实验材料要比 液态或气态方便控制。以氙举例, 氙是一种惰性气体, 所以要想用固态<sup>136</sup>Xe 做实验几乎是 不可能的, 因此一般以氙为目标同位素的实验都是用气体或者液体做实验, 这就很影响探测 器的搭建等各种考量了。近年来, 世界各地都在计划开展吨级实验寻找 0vββ, 而寻找这一 罕见衰变的代价巨大, 不仅仅在于金钱方面, 更加在于技术层面。

下面我将介绍一些知名的 0νββ 实验以及它们的优点。

### (1) 能量分辨率高: GERDA<sup>[2]</sup>, CUORE<sup>[3]</sup>

GERDA(GERmanium Detector Array)和 CUORE(the Cryogenic Underground Observatory for Rare Events)都是位于意大利格兰萨索地下实验室的实验。GERDA 实验使用高纯度<sup>76</sup>Ge 晶体进行实验,它使用了 37 个浓缩探测器,<sup>76</sup>Ge 的同位素丰度最高可达 87%,探测器总质 量为 35.6 千克,在 95%的置信区间内将 0vββ 的半衰期延长到了 3×10<sup>25</sup>年。

而 CUORE 实验则是使用 988 块  $5 \times 5 \times 5$  cm<sup>3</sup> 的 TeO<sub>2</sub> 晶体作为探测器,总重 741 千克,



其中<sup>130</sup>Te 的质量达 200 千克,如图 1-5 所示。同时 CUORE 最大的特色是正常运行期间整 个探测器将处于~10 mK 的工作温度,号称是"宇宙最冷的一立方米"。CUORE 的 988 块晶 体上都贴着热敏电阻,随时读出晶体的温度。当晶体内部发生双β衰变时,放出的电子能量 将沉积在晶体内部,使晶体温度升高。通过热敏电阻读出晶体的温度变化,我们就可以反推 回电子沉积的能量大小了。这一原理实现的关键是晶体的温度变化要足够明显,这也是为什 么 CUORE 探测器需要在极低温下工作的核心原因。我们知道晶体在低温下大致遵循德拜 T3 定律,即晶体的比热容与温度的三次方成正比。将晶体温度降到极低时,其比热容也非 常的小,只需要一丁点热量就能使温度发生明显的变化,从而增大探测器的信号。根据 CUORE 在 2018 年的发表结果,在 90%的置信区间内 0vββ 的半衰期下限大于 1.5×10<sup>25</sup> 年。



图 1-5 CUORE 探测器: 19 个 TeO2 晶体塔, 988 道通道阵列

(2) 同位素质量大: KamLAND-Zen<sup>[4]</sup>, SNO+和 EXO

KamLAND-Zen 实验位于日本岐阜县的一个地下 1000 米深的废弃砷矿,使用的是它的前身 KamLAND 的基础设施。KamLAND-Zen 由 13 吨溶解有 383 千克氙的液体闪烁体组成,其中 <sup>136</sup>Xe 占氙的大多数。如图 1-6 所示,这些活性物质位于一个直径 3.08 米的透明尼龙气球中,而这个尼龙气球又位于第二个更大的尼龙气球的中心,周围有 1 千吨的液体闪烁体作为屏蔽体,且时刻都被 1879 个光电倍增管(PMT)监视着。



### 图 1-6 KamLAND-Zen 结构示意图,以品红色突出的部分是含氙的液体闪烁体

尽管它能量分辨率不高,但由于巨大的曝光量,还是获得了很高的 0vββ 衰变下限:  $T^{0v}_{\beta\beta} > 1.07 \times 10^{26}$  yr (90% C.L.)。SNO+和 EXO 也是采用类似的原理,其中 EXO 还在进行 研发,他们希望设计一种独有的 Ba 离子标记技术以降低本底。



(3) 本底超低: NEXT, PandaX-III, SuperNEMO

这类实验大都拥有径迹重建或者拓扑特性识别能力,因此本底可压到极低。NEXT 实验 位于西班牙的 LSC 实验室,使用高压气体氙时间投影室以及电致发光读出器来进行探测。 类似的,PandaX-III 是中国的 0vββ 探测实验,也将采用高压气体氙时间投影室来进行径迹 重现。由于多重散射,电子会在高压氙气里面留下较长的径迹,并且在径迹两端有大的能量 沉积,通过这样鲜明的特点,可去除大部分本底。



#### 图 1-7 SuperNEMO 结构示意图

SuperNEMO 位于法国阿尔卑斯山的 LSM 地下实验室。目前还只有演示模块,在将来 将扩展到更大的模块化探测器。它具有跟踪器-量热计结构,如图 1-7,在探测器之间夹着一 层双β衰变发射同位素,并被量热计包围。通过这样的结构,带电粒子的完整三维轨迹可以 得到重建,能量也同时可以被测量出来。

# 1.3 本文结构安排

本论文围绕 PandaX-II 实验的数据分析展开,使用 PandaX-II 实验的 407.1 天数据,讨 论其 0vββ 的搜索结果,具体结构如下:

第一章:介绍了 0νββ 物理背景,了解其重要物理意义,然后介绍目前国际上主要的 0νββ 实验,对 0vββ 的研究现状和未来目标有充足的认识。

第二章:介绍了 PandaX-II 探测器和 2016-2018 年的数据采集活动,将特别关注 0vββ 相关的探测器功能及原理。

第三章: 概述了 PandaX-II 的数据处理链,同时介绍了数据的筛选和能量重建的流程。 最终展示出 0vββ 感兴趣区域的能谱。

第四章:介绍了通过蒙特卡洛模拟得到的本底模型,拟合能谱后得到的 0vββ 信号强度, 我还展示了我们做能谱拟合及后续分析时编写的程序架构。为了使结果更可信,同时我还进 行了系统误差、半衰期极限等分析。

第五章:总结毕业设计的工作内容,展望了大型液体氙探测器用于 0νββ 探测的前景。



# 第二章 PandaX-II 实验简介

# 2.1 基本介绍

PandaX (Particle AND Astrophysics experiment with Xenon) 实验<sup>[5], [6]</sup>位于中国锦屏实验 室(CJPL)<sup>[7]</sup>, 其全称是粒子与天文氙实验, 其主要科学目标是直接捕捉暗物质候选粒子弱相 互作用大质量粒子 (WIMP) 和氙核的可能碰撞<sup>[8]</sup>。PandaX 是一系列实验的总称, PandaX-I 和 PandaX-II 分别是包含 120 千克, 580 千克有效氙的暗物质探测实验。其下一阶段的 PandaX-III 实验旨在使用高压 <sup>136</sup>Xe 气体时间投影室(TPC)探测 0vββ, 而 PandaX-4T 是 PandaX-II 的延续, 是使用 4 吨液体氙的下一代暗物质探测实验。

本章讨论 PandaX-II 探测器的基本原理,以及过去3年间的数据采集活动。

## 2.2 0vββ 分析可行性

正如上一节介绍的那样,随着实验的升级,PandaX-4T 所用的液氙质量将达到吨级,同时国际上一些其他的使用液氙进行暗物质探测的实验,例如 XENON 也达到吨级。这些实验所消耗的液氙越来越多,而液氙的特殊性质使得它能够用于 0vββ 探测。越来越多的合作组在考虑结合暗物质探测与 0vββ 探测这两个物理目标,搭建多功能的探测器,这样不仅可以节省很多经费,同时也增强实验的竞争力。而本论文进行的 PandaX-II 实验的 0vββ 分析能够让我们对于液氙探测器的性质有进一步的了解,同时也有利于我们在下一代的升级实验中为了这一目标而进行探测器的优化。

氙是一种被广泛的用来进行探测实验的惰性气体,而且它是惰性气体中除氡外最重的元 素,原子序数为 54。氙在空气中含量很低,要得到实验上能使用的氙还需要经过一系列的 提纯,繁琐的过程,这使得其价格昂贵。尽管如此,国际上也还是有很多实验组使用氙来进 行实验,我们可以从中一窥其迷人的魅力。

用氙来进行 0vββ 探测有很多优点,首先作为一种候选同位素,<sup>136</sup>Xe 的自然丰度不低,可以借助大质量的暗物质实验同时开展 0vββ 探测。即使不做提纯,大质量的液体氙探测器 也含有质量可观的<sup>136</sup>Xe用于 0vββ 探测。以PandaX-II 为例,<sup>136</sup>Xe 的自然同位素丰度为8.9%, PandaX-II 实验内的 580kg 自然氙含有 51.6kg 的<sup>136</sup>Xe。

然后是暗物质实验与 0vββ 探测在一些要求上有很多吻合之处,例如两者都需要在低本 底的地下实验室进行,且氙本身就是一种很重的元素,因此在探测器外围的液氙是内部液氙 的绝佳屏蔽体。

对于暗物质实验,特别是具有大量液体氙的探测器一直在迅速发展,在将来液体氙探测 器或能成为暗物质和 0vββ 探测的强大工具。

## 2.3 时间投影室

时间投影室是利用粒子径迹产生电离电子的漂移时间和漂移方向的投影位置确定径迹 三维坐标的探测器,后来基于 TPC 又诞生了双相 TPC 的概念。

#### 第8页共30页



2.3.1 探测原理

PandaX-II 实验采用 500 千克级别的双相氙探测器,其 TPC 由氙、电极、反射层和两个 放置在顶部和底部的 PMT 阵列构成,如图 2-1 所示。



图 2-1 PandaX-II 探测器示意图

氙的一些性质在上一节提到过,作为探测介质,其中的杂质会影响光的传播和电子的 漂移,因此必须时刻对其进行循环提纯。同时通过溢流机制将液位保持在闸门和阳极之间, 使得电子的最大漂移长度约为 600mm。TPC 中的漂移、提取、放大电场由阴极,栅极和阳 极网格施加。由于探测器依赖于光信号进行后期分析,我们希望光信号能够尽量都是被 PMT 吸收,因此提升光接收效率对这个实验有着重要意义。为了提升光接收效率,探测器内设 置了一层反射层以确保大多数光子被 PMT 接收到,实验中使用的是直径为 646 mm 的特氟 龙 (PTFE)反射层。上下两层 PMT 阵列用于接收探测器内发出的闪烁光子 (S1) 和正比 闪烁光子 (S2)。每个 PMT 阵列包括 55 个用于信号读出 PMT 和用于否决外围事件的 24 个 PMT。探测器的敏感体积中含有约 580 千克的自然液氙。整个 TPC 在-94.5 摄氏度左右运行, 并放置在较大的真空室内进行隔热。



图 2-2 双相液氙 TPC 探测原理示意图

当入射粒子与液氙相互作用时,有能量沉积并且产生初始闪烁光信号,我们把它叫做 S1。同时轰击产生的一部分电子在电场作用下向气体氙方向漂移,到气液交界处由萃取电 场将这些自由电子拽入气氙中。由于气体部分有着一个更强的电场,电子加速与氙原子碰



撞产生电致发光(electroluminescent),我们把它叫做 S2。

如图 2-2 所示, S1 的信号一旦发出很快便被 PMT 接收到,其信号宽度短,大约为 10~100 ns。而对于 S2 信号来说,它的宽度主要取决于气氙层的厚度,因此为了更加有效的区分这两个信号,一般将它的宽度控制在 1~10 μs。更加重要的是,通过有效区分 S1 信号和 S2 信号,我们可以计算这两个信号之间相差了多少时间。S1 信号和 S2 信号相差的时间实际上就是电子在液氙内漂移的时间,由于液氙内阻力很大,电子的运动一般是看成匀速运动。通过读出最长的漂移时间我们可以得知液氙的漂移速度是多少,然后每一个事件发生的垂直位置就可以算出来了。另外,通过顶部和底部两个 PMT 阵列不同位置处 PMT 接收到的光子数量,我们可以重建事件发生的水平位置。因此,两相液氙探测器内碰撞发生的三维坐标可以被计算出来。此外,可以很容易想到的是事例的能量越大对应的 S1,S2 信号光子数越多,因此在后期数据处理中我们也是通过信号大小来重建能量,并且得到 0vββ 感兴趣区域的能谱。

对于暗物质探测实验来说,利用两个光信号的意义更体现在本底事例的甄别上。因为 理论上暗物质与氙原子发生的是核反冲,而一般 γ 射线是与氙发生电子反冲。顾名思义, 核反冲是指直接与氙核发生作用,而这类事例由于有更大密度的电离,会有一个比电子反 冲更大的 S1 信号。而更多的能量在发生 S1 碰撞的时候就沉积了,对应的 S2 信号自然就变 小了。在分析过程中,S2 与 S1 的比值被定义为能量分支比,这个地方的 S1,S2 代表的是 光子数。通过分支比我们可以有效区分核反冲和电子反冲,从中挑选可能的暗物质碰撞事 件。

#### 2.3.2 发光机制

正如上一节所提到的,入射粒子与液氙碰撞后在探测器内会引发两种光信号:S1和S2。 S1的来源有两种:(1)入射粒子轰击液氙后,一部分氙原子被激发到能量更高的激发态后 与基态氙原子结合形成受激分子,受激分子退激发发出178 nm的真空紫外光。(2)还有一 部分氙被电离,形成电子-离子对,其中一部分通过复合形成受激分子,然后发生同样的退 激过程。另外一些没有结合的电子在电场作用下逃走了,成为S2产生的来源。由于(1)(2) 这两个过程时间间隔非常短,因此两个S1信号无法区分,形成一个大的S1信号。

S2 信号被称为电致发光,又叫做正比荧光,是由电子激发氙原子放出的。因为气体氙的密度比起液氙要小的多,而且这一区域加了很强的电场,电子的平均自由程变大使得电子在两次碰撞间可以被加速到有足够的能量去激发下一个氙原子。因此,S2 光子数与电子数是正比关系。

#### 2.3.3 电子漂移与萃取

在液氙中,未与离子结合的电子往气氙区漂去,由于液氙密度大,电子自由程短,电场 加大后也不会使电子漂移速度有很明显的提升。对于 PandaX-II,电子的漂移速度在 1.7 mm/μs 左右,而且液氙中的各种杂质会导致电子的损失。当电子达到气液交界处时,在萃取 电场的作用下会脱离液氙,这时候电场强度越大,脱离液氙的电子越多。

## 2.4 PandaX-II 的数据积累

在 PandaX-II 开展实验的过去 3 年间,总共采集了 4 组物理数据,如表 2-1 所示。

第一个数据集(Run 8)是一个调试运行数据集。在此期间,探测器中发现了高含量的 氪污染。随后在 2016 年初对氙进行蒸馏去除污染,之后一段时间的稳定运行(Run 9)<sup>[9]</sup>。 这项工作产生了 WIMP 暗物质截面约束的世界领先结果。在 2016 年晚些时候,实验组使用



CH<sub>3</sub>T 对电子反冲事件进行了一些刻度,但在移除它时遇到了麻烦。于是在现场又对氙进行 蒸馏以解决这个问题。终于,在 2017 年初再次进行数据采集(Run 10 和 Run 11)<sup>[10]</sup>。在本 论文的工作中,只有 Run 9, Run 10, Run 11 的数据被用于 0vββ 半衰期的分析,而且这是 Run11 数据的第一次公布使用, Run11 数据采集期间的情况如图 2-3 所示。

#### 表2-1 PandaX-II实验数据集

数据集	开始时间	结束时间	运行时间 (天)
Run 8	2015年11月21日	2015年12月14日	19.1
Run 9	2016年3月9日	2016年6月30日	79.6
Run 10	2017年4月22日	2017年7月16日	77.1
Run 11	2017年7月17日	2018年8月16日	250.4



图 2-3 Run11 数据采集期间曝光量和电子寿命的变化



# 第三章 PandaX-II 数据分析

## 3.1 数据处理流程

从 PMT 上读出的波形信息最开始都是以二进制的方式储存起来,在后期处理中再提取 出我们需要的信息。下面我将介绍 PandaX-II 所使用的的一整套数据处理流程。

用于 PandaX-II 实验的数据分析框架"统一数据模型"(UDM)是基于 ROOT 和 C++开 发设计的。ROOT 是欧洲核子研究中心(CERN)开发的一款基于 C++的数据处理程序,最 开始是粒子物理学家们为了方便粒子物理的数据分析而设计的,到后来它的功能越来越强大, 可以提供庞大的数据处理、分析,可视化以及存储等功能,应用的领域也越来越广。UDM 包括一个面向对象的数据模块 Bamboo-Shoot,一个用于链接不同数据模块的"桥梁"PandaX Chain,以及一个小工具包 PandaX Tools。

以 Bamboo-Shoot 为框架的基础,我们可以把原始数据转换为中间具有不同风格的数据, 这使我们能够接近事件不同级别的物理信息。而 PandaX Chain 则负责实现不同数据模块间 的转换,执行重建算法。PandaX Tools 可以实现数据模型的可视化,并且有分析人员为自己 定制的一些多样性功能。

图 3-1 中的方块表示在 Bamboo-Shoot 中编码的数据模块, 描述了从原始数据到具有特定用途的物理数据过程中生成的衍生物。这些数据模块包含了: RawData, EsumData (这也是 RawData 类型的一种), HitData, ClusterData, SignalData 以及最后的 Plain root。而这些黄色的箭头其实就是 PandaX Chain 所实现的功能。



#### 图 3-1 UDM 数据模块转换图

最初的二进制文件经过一个基本的转换后变成了 RawData 文件,此时里面含有的主要都还是些很基本的波形信息,然后我们对其进行时间和 PMT 增益的标定。

我们把光子打到 PMT 上形成的一个连续波形称为一个 hit,在 PandaX Chain 中有专门的程序来寻找一段很长的波形片段中的 hit,并且称之为靶击事例辨认 (hit finding)。标定之后的 RawData 文件马上进行的就是 hit finding,在这个过程中,我们找出所有 PMT 波形中的 hit,把它们的信息提取出来后把原始波形丢弃,这一步生成的文件是 HitData。在上一步



的同时,我们把上下两个阵列和外部的 PMT 标定后的波形进行叠加,得到 EsumData,我们可以把它看成只含 3 个 PMT 信息的 RawData。

当同一个事件发出的光子到达 PMT 上,它们的波形应该会有很多相似之处,例如开始时间在理论上应该是一样的。因此我们使用 Clustering 算法,将 hit 进行组合,辨认哪些 hit 是属于同一个集群的。这步会产生不被保存的 ClusterData,将它与之前的 EsumData 结合,我们重建出信号,并且区分 S1 和 S2,这样我们就得到了信号层面的信息 SignalData。最后我们对信号的几何位置,能量进行重建,包括一些其他的操作,然后将含 Bamboo-Shoot 类的 SignalData 转化为方便读取操作纯 root 格式的 AnaData。

# 3.2 数据筛选

在原来的暗物质数据分析中,使用了很多的 cuts 来对数据质量进行把控。而 cuts 是指 在数据分析时,对某些参数做出限制,不满足要求的事例被认为是不好的事例,在后续的分 析中被"砍"掉。对 0vββ 分析来说,关注的能量区间在 MeV 量级,而暗物质区间在 keV 量级,因此我对于原来使用的 cuts 重新进行了分析,得出了适用于 0vββ 分析的 cuts。

我们首先加入了一些基础的 cut 来去除大部分的本底事例,这些 cut 包括能量区间,veto cut 以及有效体积 cut。然后我们还加入了信号质量 cut 来进一步去除可能的低质量事件,这些信号质量 cut 包括一些波形 cut, S1/S2 pattern cut 和信号位置 cut。

但是在研究这些 cut 之前我们还有单次散射的限制来降低我们的本底。

单次散射是相对于多次散射而言的。如果入射粒子与液氙碰撞后能量还很大的话,它很 有可能走一段距离再次发生碰撞,这就是多次散射。但绝大部分 0vββ 都是属于单次散射事 例,这是因为两个电子将沿很小范围的连续轨道沉积其能量,在 S2 波形中呈现出单个脉冲。 在分析中,我们只挑选单次散射的事件。

3.2.1 基础事例筛选和信号质量筛选

(1) 能量区间

液氙的  $Q_{\beta\beta} = 2458 \ keV$ ,能量区间的宽度则取决于探测器的能量分辨率。PandaX-II 在 $Q_{\beta\beta}$ 处的能量分辨率大约为4.2%,100keV 的量级。因此我们最初选用了 2000 – 3000 keV 的区间作为我们的感兴趣区域,但是随着分析的深入,我们希望更多的了解探测器性质,因此能量区间被拓展到了 1000 – 3000 keV。

#### (2) Veto cut

在之前讲探测器构成的时候提到用于否决外围事件的 PMT,这些 PMT 放置在外围,如 果它们看到光信号,那么即使同时其他的内层 PMT 也有信号,我们也把这些信号剔除掉, 因为大部分液氙都在内层,如果发生 0vββ 的话也会在反射层内沉积能量,而不会跑到外层 去。反之,一些从外界穿入的粒子,如μ子则更加可能触发 veto PMT 并且被这个 cut 所剔 除。

(3) 有效体积 cut

由于位置重建,TPC 外罐放射性大等因素,并不是发生在探测器任何位置的事例都是 好的。越靠近TPC 中心的地方位置、能量重建越好,且因为远离外罐,有外层液氙作为屏 蔽,本底也很低。如图 3-2 所示,大多数的事例都集中在顶部和底部。这些事例大部分都是 来自于 PMT 阵列,因为 PMT 里面 <sup>60</sup>Co 和 <sup>40</sup>K 的活度都很高。而另外一方面来看,沿着半 径方向的事例数差异不如纵向大,这说明罐体本身带来的高能本底其实是不如 PMT 多的。



有效体积 cut 是指我们人为对液氙体积做一个限制,这样能有效的降低本底,提升信噪 比。通过对探测灵敏度进行 TPC 的水平方向和垂直方向的扫描,我们最终确定了水平方向 上*r<sup>2</sup>* < 80000 *mm<sup>2</sup>*,垂直方向上233 mm < *z* < 533 *mm*的有效体积,在这个范围外的事例 都不予考虑。而在这个范围内,液氙的质量为 219 千克,这个 cut 是所有 cut 里面限制最强 的。



图 3-2 感兴趣区间(2000 keV < E < 3000 keV)的事例在 PandaX-II TPC 内的分布。单 次散射、信号质量 cut、veto cut 等都加上了,有效体积 cut 如红线所示

(4) 波形 cut

我们应用了两个噪声过滤器来去除 PMT 噪声导致的波形失真,它们是 ripple ratio cut 和 negative charge cut。Ripple ratio 是 S1 波形中正信号部分(过冲信号)的积分除以负信号部分的积分值。大比率的 ripple ratio 表示信号过冲部分很多,更可能是来自电子学噪声干扰。在研究了高能区事件的 ripple ratio 后我在处理数据的时候统一要求 ripple ratio 得小于 0.5%。

另外一个 cut, negative charge cut 则是对 qS1Esum/qS1 的限定。qS1Esum 是从所有 PMT 的波形相加后的总和波形中计算出的光子数, qS1 是从单个 PMT 波形中计算出光子数再将 所有 PMT 光子数相加的值。一般来说, 波形上面计算光子数的时候我们只会对负信号部分 进行积分, 因此正值是不会被考虑在内的。但由于算法上的区别, 在计算 qS1Esum 时我们 先对所有的 PMT 波形进行了相加, 如果某几个 PMT 波形有正值的话, 很有可能相加后在 总波形上还是负值, 这样的话我们仍然使用了这部分区间来计算光子数, 相当于我们考虑到 了正值部分, 因此 qS1Esum 会比 qS1 小。因此, 我们要求 qS1Esum/qS1 大于 0.985。较小 的比率意味着信号内可能存在一些意外的过冲。这两个 cut 的效果见图 3-3, 红线即为我们 的 cut。



第 14 页 共 30 页



(3-1)

图 3-3 左图: ripple ratio cut 右图: negative charge cut

#### (5) S1/S2 pattern cut

这部分是由"small" S2 cut 和 S1 charge asymmetry cut 组成的。

在实验中我们给罐体加了高压以保证电子的漂移,这样的话免不了电极上会因打火而发 光。源于电极上的打火的噪声信号很容易被"small" S2 cut 分辨。与正常情况下信号沉积 能量相比,打火电离产生的 S2 信号波形更窄。因此,基于 S2 和 S1 的比值,这些坏的事例 很容易被挑选出来。根据之前的暗物质数据分析,这类事件通常有更低的log<sub>10</sub>(<sup>S2</sup>/<sub>S1</sub>)值,约为

1.5。如图 3-4,我们画出 $log_{10}\left(\frac{S2}{S1}\right)v.sS1$ 的二维图,将有较低比值的事件给"砍"掉。



 $log_{10}\left(\frac{S2}{S1}\right) > 1.5 - 0.595e^{-\frac{x}{10000}}$ 

S1 charge asymmetry cut 的定义是A'<sub>TB</sub> = (S1T – S1B)/S1,即为顶部接收到的 S1 光子数 减去底部接收到的 S1 光子数的和除以 S1 总光子数。如图 3-5 所示,我画出了A'<sub>TB</sub>为纵轴,电子漂移时间为横轴的二维分布图。电子漂移时间越长,事件发生的越靠底部,所以顶部 PMT 接收到的光子数更小,继而A'<sub>TB</sub>值变小,所以我们可以在图上看到一个斜向下的分布。

对于同一位置发生的不同能量的事件,我们有一个假设是光子的分布大致都是相同的, 好的事件会很靠近彼此。因此图 3-5 中出现的很多散点很有可能是不好的事例,例如顶部某 个光电管突然打火就会使得比值异常的高。图中的红线表示我们对事例进行的 cut,红线之 间的事例得到保留,红线外的事例被扔掉。cut 的具体表达式为:

$$\frac{S1T - S1B}{S1} < -1.8 \times 10^{-6} \times dT + 0.05 \tag{3-2}$$

$$\frac{S1T - S1B}{S1} > -1.8 \times 10^{-6} \times dT - 0.3 \tag{3-3}$$



(6) 信号位置质量 cut

信号位置质量 cut 会将位置重建质量很差的事例给丢掉。这个 cut 基于的是位置重建好 坏以及不同重建算法之间的位置差。

位置重建的质量体现在χ<sup>2</sup>不能很大。如图 3-6 所示,具体形式见下面的式子:





同时我们要求用 PAF 算法和 NN 算法得出的事件位置之间的水平距离不能太大。PAF 和 NN 都是 PandaX-II 数据处理中会用到的位置重建算法,分别是"光接收函数"和"神经 网络"的缩写。如果用两种不同的方法重建出来发现位置相差太多,我们就有理由怀疑这个 事例是有些问题的。从图 3-7 沿半径方向事例的位置差分布可以看出半径越大, 重建出来的 位置不确定性越大。

$$D_{PAF.NN.} < 40 - 1.43 \times 10^{-4} r_{PAF}^2 + 2.5 \times 10^{-9} r_{PAF}^4 + 2.2 \times 10^{-13} r_{PAF}^6 - 1.2 \times 10^{-18} r_{PAF}^8$$
(3-6)



#### 3.2.2 筛选后的能谱

PandaX-II 探测器的信号探测效率被认为是 100%,因为高能信号发出的光很强,不应该 会被漏掉。但是我加入的各类 cut 会消减事例数,相当于降低了探测效率。因此,对筛选效 率要有一个合理的计算。

这些连续的数据筛选的影响已经显示在表 3-1 和图 3-8 中了。在表 3-1 中,我们显示了 存活事件的数量和相关 cut 的信号效率估计值。在我们的分析链中通过单次散射 cut 在信号 预处理期间丢弃了大量多次散射的事件。基于蒙特卡罗模拟我们估计出这个 cut 有 6.6%的 可能剔除 0vββ 的多次散射信号。

同时由于我们选择的感兴趣区间很宽,所有的 0vββ 都会落在这个范围,因此能量区间 cut 的效率忽略不计。由于我们只探测单次散射的事例,Veto cut 的效率也不需要计算。而有 效体积的效率则是直接反映在了我们最终拟合时的氙的质量上,因此也不需要考虑。

通过比较应用相应 cut 之前和之后的事件数,我们评估了这些信号质量 cut 的效率。它 们的值是 3 个 Run 的平均值,误差则是 3 个 Run 的标准差。通过结合这些 cut,我们得到最 终的探测效率是 (91.4 ± 0.71)%。

#### 表 3-1 事件选取和信号筛选效率,探测效率表示能通过 cut 的 0vββ 事件

Cut 名字	事件数	探测效率(%)
单次散射 cut	38429692	93.4
能量区间 cut	891140	~100
Veto cut	499915	-
有效体积 cut	35759	-
波形 cut	35490	$99.2 \pm 0.69$
S1/S2 Pattern cut	35416	$99.8 \pm 0.03$
信号位置质量 cut	34992	$98.8 \pm 0.15$
总计	-	91.4 ± 0.71





图 3-8 3000keV 以下的能谱,通过比较能谱的变化分析各个 cut 带来的影响,可以看到 低能区的本底被极大的压制了

# 3.3 本章小结

在介绍完 0vββ 的物理背景和 PandaX-II 实验后,本章主要讲了 PandaX-II 数据处理的流程,以及我对于其中数据质量 cut 的修改。熟悉 PandaX-II 数据处理的流程是一切分析的基础,因为无论是各阶段的信息提取还是 cut 的添加都是这个流程中的一步。对于任何一个高能物理的实验来说,定义合适的 cut 都是非常重要的,因为探测器接收到的信号是巨大的,我们所有想要分析的东西,想要找到的新物理都隐藏在庞大的数据当中,做 cut 这一步就好比是慢慢揭开它神秘的面纱。好的 cut 能帮助我们有效的去除本底,提高能量分辨率,是在探测器造好,数据采集完后我们还能优化实验结果的一个主要手段。在对 PandaX-II 数据的高能区分析中,我对于之前所使用的暗物质信号质量 cut 整体进行了一遍分析,剔除许多不合适的 cut,并且对剩下的作了修改。通过这个操作,保证了能谱的准确,一定程度的提高了能量分辨率,对后续的能谱拟合有重要意义。



# 第四章 PandaX-II 0vββ 探测结果

在完成了信号质量 cut 重定义的工作后,我们得到了最后用于分析物理结果的能谱。因此,我紧接着的工作就是建立本底模型,并且使用合理的统计学方法对能谱进行拟合。拟合结果得出后,我还进行了系统误差的计算,通过比较系统误差与探测结果使我们的结果更加可信。

## 4.1 本底模型建立

整个 0vββ 分析过程最核心的就是感兴趣区域的拟合,而拟合的基础就是建立本底模型。 拟合的最终目标是测试我们会有多大的可能性在 Q-value 处存在多大的信号。根据实际测量 到的各组件放射性和蒙特卡洛模拟,我们对于感兴趣区域的主要本底构成有很清楚的了解, 而我们的模型正是基于这个结果。

4.1.1 拟合函数

在我们的分析中,对能谱做过两段不同区间的拟合,一段是 2000 – 3000 keV,另外一 段是 1000 – 3000 keV。其中我们将前者称为"最佳拟合",后者称为"预拟合"。

(1) 2000 - 3000 keV

对于 2000 到 3000 keV 的区间, 能谱上主要的本底为  $^{232}$ Th 的 2615 keV 峰,  $^{238}$ U 的一系 列衰变链对应的 γ 峰, 碰巧同时吸收了两个  $^{60}$ Co 的 γ 射线后产生的 2506 keV 峰, 一个很平 坦的  $^{222}$ Rn 的部分 β 谱, 和一个平的本底。加上 0vββ 峰, 用公式总结的话就是:

$$f_{ROI}(E) = N_{Th}Th232Func(E) + N_UU238Func(E) + N_{co}Co60Func(E) + N_{Rn}Rn222Func(E) + FlatBkg + N_{0\nu\beta\beta}0\nu\beta\betaFunc(E)$$
(4-1)

(2) 1000 - 3000 keV

在我们将感兴趣区域下限降到 1000 keV 后,不仅原有的各元素的分布变复杂了,能谱上还增加了 <sup>40</sup>K、2νββ 的部分。图 4-1 为模拟的探测器内不同元素对应的能谱。



图 4-1 利用蒙特卡洛模拟得到的各元素对应的能谱: (a)<sup>40</sup>K (b)<sup>238</sup>U (c)<sup>60</sup>Co (d)<sup>232</sup>Th



同样的用公式表示的话就是:

 $f_{ROI}(E) = N_{Th}Th232Func(E) + N_UU238Func(E) + N_{co}Co60Func(E)$ 

+*N<sub>Rn</sub>Rn222Func(E)* + *N<sub>K</sub>K40Func(E)* + *N<sub>2νββ</sub>2vββFunc(E)* (4-2) 注意对于这个区间的拟合我们实际上是忽略了*Q<sub>ββ</sub>*±100 *keV*范围的,所以并不需要加上 0vββ 峰,至于原因会在后面 4.2 讲。

4.1.2<sup>222</sup>Rn 和<sup>60</sup>Co 对于本底的贡献

探测器的本底众多导致拟合参数很多,这给后面的能谱拟合造成了很大的困难。其中我 们不得不着重考虑的是 <sup>222</sup>Rn 和 <sup>60</sup>Co 对本底的贡献,这不仅仅是拟合参数越多耗时越长的问 题,更加是考虑到这两个元素的能量区间离 $Q_{\beta\beta}$ 非常近。如果完全将它们的事例率设置成参 数,得到的结果将会有很大的统计误差,同时也很不靠谱。因此,对 <sup>222</sup>Rn 和 <sup>60</sup>Co 的事例率 进行分析就很有必要了。

其中<sup>222</sup>Rn 的分析是通过一套独立的 Bi-Po 耦合分析得到的,且在最后的最佳拟合中,我们固定了<sup>222</sup>Rn 的值。

而得到<sup>60</sup>Co的事例率则是我们进行预拟合最主要的目的,因为<sup>60</sup>Co的2506 keV 峰是 通过同时吸收另外两个峰1173 keV 和1332 keV 得到的,这两个γ峰是其附近区间内主要的 本底,我们的预拟合区间可以包含这两个峰。假设各元素的衰变链达到了平衡态,对于这3 个峰之间的比例关系我们是从模拟中知道的。通过拟合出这两个峰的高度,我们可以较为准 确的计算出2505 keV 峰的高度。

# 4.2 能谱拟合

在充分了解高能本底组成之后,我们对能谱进行了两次拟合,第一次是预拟合,在 $Q_{\beta\beta}$ 附近 200 keV 的范围内是不拟合的,也就是所谓的"blind analysis",这步的意义正如上一章 所讲,是为了让我们能确定部分本底模型参数的值,使我们进行正式拟合时不必要加入太多 参数。同时 $Q_{\beta\beta}$ 附近不进行拟合可以保证我们这步得到的参数不会有偏向的影响我们正式拟 合的结果。第二次是 2000 – 3000 keV 的拟合,也就是我们的正式拟合,这次拟合中我们同 时加入了 0vββ 信号,拟合结果将会告诉我们信号强度有多大。

4.2.1 贝叶斯表述

我们利用 BAT<sup>[11]</sup>工具包对能谱进行拟合,BAT 的全称是"贝叶斯分析工具包(Bayesian analysis toolkit)",所以先来介绍什么叫做贝叶斯表述。

最基本的贝叶斯定理告诉我们:

$$P(B|A) = \frac{P(B)P(A|B)}{P(A)}$$
(4-3)

如果 B 是由 S 个元素组成的集合,那么我们可以将上式进行推广:

$$P(B_i|A) = \frac{P(B_i)P(A|B_i)}{\sum_{j \in S} P(B_j)P(A|B_j)}$$
(4-4)

如果我们进一步赋予 A、B 实际意义,A 代表我们的实验数据,B 代表我们的物理模型。 以 0vββ 探测为例,每一个不同的 0vββ 半衰期值其实就代表一个不同的物理模型,而根据 式 4-3,我们可以从探测到的数据计算某一个模型的概率,这便是贝叶斯表述。将式 4-3 改 写的更加符合实验表述,如下所示:

$$P(\Gamma_{0\nu}, \theta | E) = \frac{\pi(\Gamma_{0\nu})\pi(\theta)P(E|\Gamma_{0\nu}, \theta)}{\int_{\Omega} \int_{0}^{\infty} \pi(\Gamma_{0\nu})\pi(\theta)P(E|\Gamma_{0\nu}, \theta)d\Gamma_{0\nu}d\theta}$$
(4-5)

第 20 页 共 30 页



其中 $\theta$ 为除了衰变速率外的所有实验参数,  $\Omega$ 为其参数空间,  $\pi$ 表示的是先验概率。最后我们可以通过对 $P(\Gamma_{0\nu}, \theta|E)$ 进行  $\theta$ 空间的积分得到 $P(\Gamma_{0\nu}|E)$ , 也就是从实验数据出发计算出的某个半衰期的概率。

 $P(E|_{\Gamma_{0\nu}}, \theta)$ 也被称为"似然函数",将能谱分成若干个区间,按照模型在每个小区间上的事例数为 $\lambda_i$ ,而实际值为 $E_i$ ,实际值应该是满足以 $\lambda_i$ 为均值的泊松分布,所以似然函数为:

$$P(E|\Gamma_{0\nu},\theta) = \prod_{i} \frac{\lambda_{i}^{E_{i}}}{E_{i}!} e^{-\lambda_{i}}$$
(4-6)

概率最大的模型即为我们的最佳拟合模型。

4.2.2 程序结构

图 4-2 显示了各个文件之间的关系,下面我将一一解释这些文件的作用。



### 图 4-2 用于能谱拟合的 BAT 程序关系图

Dataset.h 里面存储了关于实验的全部信息,包括有效体积,漂移速度,曝光量,cut, 各个衰变元素大致的活度范围等等,是后面所有文件都需要引用的一个头文件。

Signal\_model.h 是我们定义本底模型的地方, Signal\_model\_multi.h 和 Signal\_model\_prefit.h 都是继承自这个文件,二者区别在于 Signal\_model.h 和 Signal\_model\_prefit.h 都只是针对一个run 来运行的,而 Signal\_model\_multi.h 是针对 3 个run 同时拟合设计的。另外一个区别就是, Signal\_model\_multi.h 对应于正式的拟合,选用区间 是 2000 – 3000 keV,而 Signal\_model\_prefit.h 进行的是预拟合,区间为 1000 – 3000 keV。这 几个文件全面的描写了使用的模型,而且每生成一个模型的时候,都得将一个 Dataset 作为 传入参数告诉它,这个模型才能被初始化。

StartBAT.h 顾名思义就是开始 BAT 拟合,将一个定义好的模型传入,调用这个函数后 各参数的值就会被更新为拟合值。

toyMC.h 是做 toy MC 的地方,这个程序主要是在计算半衰期下限和灵敏度的时候有用, 我们将在后面具体介绍。为了方便程序的运行,其实我们将模型对应需要拟合的能谱也放 在了模型内部。而利用 toyMC.h 定义的函数,我们可以将模型内的能谱根据模型,以及预 先设定的参数值人为的生成一张能谱来代替原有的能谱。通过这个方法,我们可以不断的 重复拟合得到半衰期的分布。

最后有 4 个 run\*.cxx 文件,是我们的主要执行文件。首先进行的是 run\_prefit.cxx,它 将对 3 个 run 进行单独的预拟合。其次是 run\_bestfit.cxx,它将对 3 个 run 进行同时的拟合, 这个拟合产生我们的最佳拟合结果。而 run\_bestfit\_sys.cxx 和 run\_exclusion\_muti.cxx 分别是 用来计算系统误差和排除灵敏度。基本上可以把它们看成是 run\_bestfit.cxx 重复 N 遍拟合, 不过每一个重复都会由 toy MC 产生一个新的能谱。



4.2.3 0vββ 拟合结果

首先,进行预拟合,我们在这个拟合中对每个 run 进行单独的拟合,每个拟合参数都有 8 个,包括 <sup>232</sup>Th 事例率,<sup>40</sup>K 事例率,<sup>60</sup>Co 事例率,<sup>228</sup>U 事例率和 4 个与分辨率相关的参 数(1173keV,1447keV,1764keV 和 2615keV 对应的分辨率),预拟合的结果见图 4-3。



图 4-3<sup>60</sup>Co 事例率计算拟合,能谱是 3 个 run 的叠加,但拟合时是分开的。

然后我们就开始进行最佳拟合,将<sup>232</sup>Th,<sup>238</sup>U,<sup>60</sup>Co,FlatBkg 和 0vββ 事例率设为变量, 同时 FlatBkg 和 0vββ 事例率的先验分布设为平的分布,即等概率分布。但是<sup>232</sup>Th,<sup>238</sup>U, <sup>60</sup>Co 这 3 者的事例率先验概率设成高斯分布,宽度是预拟合得到的不确定性。

除了上述元素外,  $^{222}Rn$  事例率被固定了下来,  $^{40}K$  和  $2\nu\beta\beta$  在这一步的拟合中都被忽略了。

一旦拟合后  $0\nu\beta\beta$  事例数知道了,我们就能算出  $0\nu\beta\beta$  衰变的衰变速率 $\Gamma_{0\nu}$ ,如下所示:

$$N = \Gamma_{0\nu} \frac{N_A}{m_a} \cdot a_I \cdot \varepsilon_{tot} \cdot m_d \cdot t \tag{4-7}$$

*N*是事例数,  $N_A$ 是阿伏伽德罗常数,  $m_a$ 是 Xe 的平均摩尔质量,  $a_I$ 是自然丰度,  $\varepsilon_{tot}$ 是 总的效率,  $m_a$ 是探测器总有效质量, t是运行时间。

每一个 run 都有各自独自的 <sup>232</sup>Th, <sup>238</sup>U, <sup>60</sup>Co 和 FlatBkg 事例率, 但它们共享同一个 0vββ 事例率值。因此, 总的拟合参数数目为:  $3 \times 4 + 1 = 13$ 。

我们数据在 2000 – 3000 keV 范围内的最佳拟合值见图 4-4,图 4-4 是 3个 run 的叠加。 具体的各 run 对应参数见表 4-1。

	Run9	Run10	Run11
<sup>222</sup> Rn [/kg/yr](fixed)	$22.6 \pm 12.1$	$20.6 \pm 10.4$	17.5 ± 8.8
<sup>60</sup> Co [/kg/yr]	$14.2 \pm 1.0$	$13.2 \pm 1.0$	$11.6 \pm 0.6$
<sup>232</sup> Th [/kg/yr]	$65.9 \pm 2.2$	55.9 ± 1.9	68.6 ± 1.1
<sup>238</sup> U [/kg/yr]	$40.9 \pm 2.1$	37.4 ± 1.9	$42.6 \pm 1.1$
Flat Baseline [/kg/yr]	5.8 ± 3.3	$9.4 \pm 2.8$	3.1 ± 1.6
$0\nu\beta\beta$ [/kg/yr]		-0.91 ± 0.76	
Resolution [%] (fixed)	$4.9 \pm 0.1$	$4.2 \pm 0.1$	$4.3 \pm 0.1$

表 4-1 能谱最佳拟合值总结(ROI:2000-3000 keV),其中 0vββ 事例率是由拟合得到 的衰变速率换算过来的,能量分辨率是根据的 2615keV 的峰计算的。





图 4-4 左图: 2000 – 3000 keV 能谱的最佳拟合。右图: 0vββ 衰变速率的概率分布图

综上,对能谱的最佳拟合告诉我们 0vββ 衰变速率的最佳拟合值为:  $\Gamma_{0v} = (-0.24 \pm 0.21) \times 10^{-23} yr^{-1}$ 

# 4.3 系统误差

本节致力于计算系统误差,这是我们在得出 0vββ 半衰期下限之前必须考虑的因素<sup>[12]</sup>。 一般而言,我们将系统误差分为两种类型:与真实 0vββ 衰变速率成比例的误差,即系 统缩放σ<sub>sca</sub>,以及与 0vββ 衰变速率无关的附加误差,我们称之为σ<sub>add</sub>。

系统误差有几个主要的来源。研究系统最简单就能确定的来源是信号探测效率ε和有效 质量不确定性带来的缩放不确定性,因为最终的半衰期是直接与它们的不确定性成正比的。 信号探测效率的误差在上一章讲了怎么算,而有效质量的不确定性是通过求 NN 和 PAF 两 种重建算法半径方向上的距离的平均值得到的。其他误差则的主要是由更高阶段的分析导致 的误差,比如能谱拉伸,对 Rn 事例率,能量分辨率以及拟合过程本身的拟合误差。

除了由于信号探测效率和有效质量的误差引起的不确定性(即直接的缩放不确定性)之 外,我们使用蒙特卡罗方法估计了测量衰变速率的系统误差。简而言之,该过程模拟了大量 的伪实验集合,并与我们的感兴趣区域模型相匹配(式 4-1)。

计算系统误差的过程是一个多步骤过程,从本章前面的最佳拟合开始。对于我们想要计 算系统误差的每个参数,我们在最佳拟合模型中将它的值±1σ。然后我们重新拟合,以新拟 合得到的最佳拟合参数为标准产生大量的蒙特卡洛事件,同时加入 0vββ 衰变速率从 0 到 3×10<sup>-23</sup> yr<sup>-1</sup>变化的模拟信号,得到很多的"假"能谱。

利用蒙特卡洛产生"假"能谱的核心在于每一个小区间内的事例数实际上是符合泊松分布的。可以想象,如果我们有 1000 个平行宇宙,那么每个宇宙的 PandaX-II 实验得到的数据能谱都会有微小区别,这个区别就是我们在利用蒙特卡洛产生新能谱的时候引入的泊松分布。最终,我们产生 1000 个新的能谱,这就相当于我们得到了 1000 个平行宇宙的数据。

得到能谱之后再使用最佳拟合模型去拟合它并且返回新的 0vββ 衰变速率值。最终我们 可以得到 (input decay rate v.s fitted decay rate) 的分布图,我们用一次函数去拟合这张图并 且得到它的截距和斜率。图 4-5 中展示了这种分布的一个例子,测试拟合误差。蓝色的点对 应的是最佳模型拟合的衰变速率,而红线是对整体的拟合值。一次函数拟合得到的截距是  $\sigma_{add}$ ,而斜率与1的差就是 $\sigma_{sca}$ 。





图 4-5 从 6000 个玩具蒙特卡罗实验中拟合衰变速率的分布作为其模拟(真实)值的函数。 红色是线性拟合,我们从中提取出斜率和截距。

类似的,我们按照这个方法对能谱拉伸, Rn 事例率,能量分辨率都进行了系统误差的分析。从拟合衰变速率 v.s 真实衰变速率的分布的线性拟合中得到的系统误差*σ<sub>sca</sub>和 σ<sub>add</sub>*见 表 4-2。

	$\sigma_{add} \ (10^{-23} \ yr^{-1})$	σ <sub>sca</sub> (%)
探测效率		0.71
有效质量		14.7
能谱拉伸	0.045	0.28
Rn 事例率	0.277	0.06
能量分辨率	0.104	2.27
拟合误差	0.043	0.35

表 4-2 不同来源的系统误差贡献

最终,我们需要合成表 4-2 里面列出的不同来源的系统误差。对于每一个不同的来源, 我们取它变化1σ能够产生的最大的误差,而它引入的总误差为:

$$\sigma_{syst}(\Gamma_{0y}) = \sigma_{add} + \Gamma_{0y} \cdot \sigma_{sca} \tag{4-8}$$

最后我们假设这些列出的误差彼此是独立的,所以它们符合平方相加的规则,忽略四阶 小量后,合成方法为:

$$\sigma_{svst}^2(\Gamma_{0v}) = \sum_i \sigma_{add,i}^2 + \sigma_{sca,i}^2 \Gamma_{0v}^2$$
(4-9)

将表 4-2 中的数值代入其中,得到合成的系统误差。最后,我们对于能谱的最佳拟合是:  $\Gamma_{0\nu} = (-0.24 \pm 0.21(\text{stat.}) \pm 0.30(\text{sys.})) \times 10^{-23} yr^{-1}$ 

从这个结果可知,我们在误差范围内没有发现 0νββ 存在的证据。

# 4.4 0νββ半衰期下限

在这一节中,我将描述用于提取 0νββ 半衰期下限的方法。

4.4.1 从最佳拟合到 0νββ 半衰期下限

给定了表 4-2 中的最佳拟合参数,我们可以得到数据的后验贝叶斯概率分布:

$$P_{stat}^{post}(\Gamma_{0\nu}|E) = P_{stat}(E|\Gamma_{0\nu})\pi(\Gamma_{0\nu})$$
(4-10)



将最佳拟合的结果与系统误差相结合,可以得到 0vββ 半衰期的下限。 首先我们定义与数据相关的卡方 $\chi^2_{stat}(\Gamma_{0v}) = -2log [P_{stat}(E|\Gamma_{0v})], 然后与系统误差相$  $关的卡方<math>\chi^2_{sys}(\Gamma_{0v}) = \frac{(\Gamma_{0v} - \Gamma_{0v})^2}{\sigma^2_{syst}(\Gamma_{0v})}, 其中\sigma^2_{syst}$ 来自式 4-9。简单的将统计结果与系统误差相加的 话就是:

$$\frac{1}{\chi^2_{tot}} = \frac{1}{\chi^2_{stat}} + \frac{1}{\chi^2_{sys}}$$
(4-11)

然后再将 $\chi^2_{tot}$ 转化为 $P_{tot}(E|\Gamma_{0v}) = e^{-0.5\chi^2_{tot}}$ 。式 4-10 中的先验概率 $\pi(\Gamma_{0v})$ 确保 $\Gamma_{0v}$ 一定是 一个正数,而且大于 0 的时候是等概率的。

$$\pi(\Gamma_{0\nu}) = \begin{cases} 1, & \Gamma_{0\nu} > 0\\ 0, & \Gamma_{0\nu} \le 0 \end{cases}$$
(4-12)

通过对后验概率的积分可以得到在置信度为α<sub>c.L</sub>时的 0vββ 半衰期下限。

$$\alpha_{C.L.} = \frac{\int_{0}^{\Gamma_{0\nu}^{limit}} P_{tot}(E|\Gamma_{0\nu}) d\Gamma_{0\nu}}{\int_{0}^{inf} P_{tot}(E|\Gamma_{0\nu}) d\Gamma_{0\nu}}$$
(4-13)



图 4-6 0vββ 衰变速率概率分布,红色的点是计算值,红线是高斯拟合曲线。 上述结合过程如图 4-6 所示,最终对结合后的概率积分到 90%,得到:

 $\Gamma_{0\nu} < 0.48 \times 10^{-23} yr^{-1}$  (90% C.L.) 或者等价的写成半衰期的形式:  $T_{1/2}^{0\nu} > 1.4 \times 10^{23} yr$  (90% C.L.) 如果只考虑统计误差的话:

 $T_{1/2}^{0\nu} > 3.1 \times 10^{23} \,\mathrm{yr}$  (90% C.L.)

# 4.5 0vββ 灵敏度

在前面两节中,我们得到了能谱的最佳拟合值以及由最佳拟合推出的半衰期限制。这一 节是我们整个分析环节的最后一步,计算出 PandaX-II 对 0vββ 探测的统计学上的灵敏度。 值得注意的是无论是上一节提到的半衰期下限还是这一节将要计算的灵敏度,本质上都只是 以时间为单位的一个数字,两者非常相似但又有所区别。半衰期下限是根据实验实际得到的 能谱计算出来的对半衰期的限制,实际的能谱只有一张,因此这个值是直接代表了数据集告 诉我们的信息。而本节的灵敏度则是根据我们对于本底的了解,利用蒙特卡洛模拟,人为的 创造出若干符合统计涨落规律的"假"能谱,再从"假"能谱中提取信号。灵敏度代表的是 我们对于探测器的信心,表示我们探测器应该有多强的探测本领。

第 25 页 共 30 页



4.5.1 灵敏度预估 如式 4-7 所示,

$$N = \Gamma_{0\nu} \frac{N_A}{m_a} \cdot a_I \cdot \varepsilon_{tot} \cdot m_d \cdot t$$

我们希望看到的本底事例由拟合得到,在能量区间δE内的本底数目的平均值是:

$$B(\delta E) = b \cdot m_d \cdot \delta E \cdot t \tag{4-14}$$

其中 b 为本底事例率,单位为 cts/(keV kg yr)。利用式 4-7 和式 4-14,我们可以解析的 求出灵敏度的表达式。实验的灵敏度可以看成是满足大于或者等于本底波动条件的最小的N。 当本底很大的时候可以看成是高斯分布,因此我们可以用高斯标准差来描述数据的偏离情况。

$$N(\delta E) = N \cdot f(\delta E) = n_{\sigma} \cdot \sqrt{B(\delta E)}$$
(4-15)

其中 $\sigma = \sqrt{B(\delta E)}$ ,  $f(\delta E)$ 表示信号落到 $\delta E$ 区间的比例。对于一个高斯分布的信号,  $f(\delta E)$ 表达式为:

$$f(\delta E) = erf(\frac{\delta E}{\Delta E}\sqrt{ln2})$$
(4-16)

其中ΔE为能量分辨率,通过式 4-15,得到灵敏度公式:

$$T_{1/2}^{0\nu} = \frac{\ln(2)}{n_{\sigma}} \cdot \frac{N_A \cdot a_I \cdot \varepsilon_{tot}}{m_a} \cdot \sqrt{\frac{m_d \cdot t}{b \cdot \delta E}} \cdot f(\delta E)$$
(4-17)

在我们的分析中 $f(\delta E) = 1, b \cdot \delta E = 153.4/(\text{kg} \cdot \text{yr}),$ 因此 $T_{1/2}^{0\nu}(1.64\sigma) = 1.91 \times 10^{23} \text{yr}.$ 

## 4.5.2 利用 Toy MC 计算

上述的计算只是一个初步的估计,在很多涉及到探测器具体性质和响应的地方很难用简 单的式子去描述。因此为了精确计算这个数值,我们仍然是选择蒙特卡洛模拟的方法。与计 算系统误差时的步骤类似,我们利用本底模型的最佳拟合值生成"假"能谱。不断循环生成 这样的能谱并且得到半衰期的概率分布图,但是区别在于这个地方的半衰期拟合值只能是负 数。



**图 4-7 0vββ 衰变速率的概率分布,取从 0 开始积分到 90%处作为这次实验的灵敏度** 在得到概率分布后我们积分到 90%处作为灵敏度,然后不断循环<sup>[13]</sup>。最终我们能够得 到 12000 次循环后半衰期的分布,并且取它的中值作为实验灵敏度: T<sup>0v</sup><sub>1/2</sub> > 2.1 × 10<sup>23</sup> yr。



可以看到这个值和上一节预估的值还是很接近的。

## 4.6 本章小结

在上一章介绍完数据质量筛选后,我在这一章具体描述了 PandaX-II 的 0vββ 分析结果 是如何产生的。

首先是对本底进行合理的分析,考虑到最佳拟合有很多的参数待定,我们选取了两种区间来分析本底。对于本底的组成,我们是通过 Geant4 模拟结果得知的。第一个是预拟合: 1000 keV – 3000 keV,这部分拟合的本底主要包含了<sup>40</sup>K、<sup>238</sup>U、<sup>60</sup>Co、<sup>232</sup>Th 和<sup>222</sup>Rn。通过这个拟合,我们确定了各同位素衰变速率的后验概率分布并且将这个值应用到第二部分的最 佳拟合当中。第二部分 2000 keV – 3000 keV 的最佳拟合中使用了 13 个参数,最终得到合理的 0vββ 信号强度拟合值。

整个拟合使用的是 BAT 工具包,对于能谱的拟合我们经过反反复复的测试。如果拟合的不好,一方面可能是我们模型还有问题,另外一方面可能是代码写错了。经过几个月的码代码以及核对矫正,我们最终得到了文章中呈现的最好的结果。

之后是关于系统误差的讨论,以及如何结合最佳拟合以及系统误差得到半衰期限制。系统误差是每个实验都需要关注讨论的东西,通过与系统误差比较我们才能知道得到的信号是 否是有实际意义。在我们最终给出的结果中也能看到系统误差是能够包括信号的,也就是我 们没有发现 0vββ。但是即便如此,我们也可以根据这些结果对 0vββ 半衰期做出限制。



# 第五章 结论

在本论文中,我报告了 PandaX-II 的 0vββ 衰变分析的主要几个步骤。我们分析了 PandaX-II 探测器从 2016 年 3 月到 2018 年 8 月间采集到的 407.1 天的数据。同时我们开发 了新的数据筛选和能量重建技术,并对探测器在高能区的表现有了更好的了解。在此分析中 我们也仔细的研究了探测器本底来源,这有助于未来的探测器的开发设计。

本论文的工作产生了第一个双相型液氙探测器对于  $0\nu\beta\beta$  的探测结果。我们没有在 PandaX-II 的数据中发现有  $0\nu\beta\beta$  事件存在的迹象,但我们进一步给出了 <sup>136</sup>Xe 的  $0\nu\beta\beta$  半衰期 限制:  $T_{1/2}^{0\nu} > 1.4 \times 10^{23}$  yr。

尽管相比于世界前沿探测器,我们的探测器还不具有优势,但是这展现了双相型液氙探测器也能用于 0vββ 探测。考虑到这个因素,很多使用吨级液氙探测器来探测暗物质的实验 组越来越注重他们探测器的充分利用<sup>[14],[15]</sup>。下一代双相液体氙探测器将被希望能实现多个 物理目标,覆盖宽能量范围,同时有低本底和良好的能量分辨率。

PandaX-4T 还在紧锣密鼓的建设当中,使用超纯水屏蔽和吨级的液氙自屏蔽将为探测器带来极低的本底,我们能够预期在未来,PandaX 探测器的灵敏度将会更好。



```
参考文献
```

- K.A. Olive and Particle Data Group. Review of Particle Physics [J]. Chinese Physics C, 2014, 38(9):090001.
- [2] M Agostini, M Allardt, E Andreotti, AM Bakalyarov, M Balata, I Barabanov, M Barnabe Heider, N Barros, L Baudis, C Bauer, et al. Results on neutrino-less double decay of ge 76 from phase i of the gerda experiment [J]. Physical Review Letters, 2013, 111(12):122503.
- [3] C Alduino, F Alessandria, K Alfonso, E Andreotti, C Arnaboldi, FT Avignone III, O Azzolini, M Balata, I Bandac, TI Banks, et al. First results from cuore: A search for lepton number violation via 0vββ decay of te 130 [J]. Physical review letters, 2018, 120(13):132501.
- [4] A Gando, Y Gando, T Hachiya, A Hayashi, S Hayashida, H Ikeda, K Inoue, K Ishidoshiro, Y Karino, M Koga, et al. Search for majorana neutrinos near the inverted mass hierarchy region with kamland-zen [J]. Physical review letters, 2016, 117(8):082503.
- [5] 季向东, 刘江来, 谌勋. PandaX 实验介绍和进展[J]. 科学通报, 2016, 61(20):2264.
- [6] 符长波, 陈勋, 陈云华, 等. PandaX 实验的进展与展望[J]. 原子核物理评论, 2017, 34(3).
- [7] Yu-Cheng Wu, Xi-Qing Hao, Qian Yue, Yuan-Jing Li, Jian-Ping Cheng, Ke-Jun Kang, Yun-Hua Chen, Jin Li, Jian-Min Li, Yu-Lan Li, et al. Measurement of cosmic ray flux in china jinping underground laboratory [J]. arXiv preprint arXiv:1305.0899, 2013.
- [8] XiGuang Cao, Xun Chen, YunHua Chen, XiangYi Cui, DeQing Fang, ChangBo Fu, Karl L Giboni, HaoWei Gong, GuoDong Guo, Ming He, et al. Pandax: a liquid xenon dark matter experiment at cjpl [J]. Science China Physics, Mechanics & Astronomy, 2014, 57(8):1476-1494.
- [9] Andi Tan, Mengjiao Xiao, Xiangyi Cui, Xun Chen, Yunhua Chen, Deqing Fang, Changbo Fu, Karl Giboni, Franco Giuliani, Haowei Gong, et al. Dark matter results from first 98.7 days of data from the pandax-ii experiment [J]. Physical Review Letters, 2016, 117(12):121303.
- [10] Xiangyi Cui, Abdusalam Abdukerim, Wei Chen, Xun Chen, Yunhua Chen, Binbin Dong, Deqing Fang, Changbo Fu, Karl Giboni, Franco Giuliani, et al. Dark matter results from 54-ton-day exposure of pandax-ii experiment [J]. Physical review letters, 2017, 119(18):181302.
- [11] Allen Caldwell, Daniel Kollar, and Kevin Kroninger. Bat-the bayesian analysis toolkit [J]. Computer Physics Communications, 2009, 180(11):2197-2209.
- [12] C Alduino, K Alfonso, DR Artusa, FT Avignone III, O Azzolini, TI Banks, G Bari, JW Beeman, F Bellini, A Bersani, et al. Analysis techniques for the evaluation of the neutrino-less double decay lifetime in te 130 with the cuore-0 detector. Physical Review C, 2016, 93(4):045503.
- [13] George Casella and Roger L Berger. Statistical inference, duxbury press. Pacific Grove, CA, 2002.
- [14] XENON collaboration, G Plante, et al. The xenonnt project. Columbia University, 2016.
- [15] DN McKinsey, LZ Collaboration, et al. The lz dark matter experiment. In Journal of Physics: Conference Series, volume 718, page 042039. IOP Publishing, 2016.



# 谢辞

时间如白驹过隙,临近毕业倍感如此。回首在交大的四年生活,好似盛宴将散,意犹未 尽。在此要感谢所有帮助过我,促我成长的人。

首先要感谢的是我的导师韩柯老师。他把我从一个科研新手培养成一个能承担一定科研 任务的本科毕业生。想到大一初见到韩老师的时候,当时只是懂一点点 Root 代码。后来, 韩老师领着我入门数据分析,指导我的日常工作,给予我鼓励,带我见识到 CUORE 实验的 全貌,并且给予信任让我有机会去格兰萨索实验室学习、值班。能在本科阶段就在粒子物理 领域有这样的体验,我感到非常幸运。我时常在想要是没有见识过格兰萨索的漫天繁星,要 是没有韩老师的帮助,我的科研道路不会是如今的顺畅。韩老师兢兢业业的精神也是我学习 的榜样,让我明白一个人对自己的工作抱有极大的兴趣该是有多么的开心啊。

然后,我还要感谢的是倪凯翔学长,在做毕设期间,倪学长给予了我很大的帮助。就很 多程序上的细节我们有过反反复复的讨论,倪学长带着我一起完成了 PandaX-II 分析上的大 部分事情,让我对于一篇文章的发出要经过好几个月连续不断的努力这件事情又有了更深刻 的体会。没有倪学长,这个工作是不可能这么好的完成的。同时,倪学长对于程序编写的专 业和对科研的热爱也让我敬佩。

同时我还要感谢实验室的其他老师和同学。感谢李亮老师、王少博老师、林横学长、王 秋红学长、萨拉姆学长、马文博学长等人,他们都在科研上给予过我很大的帮助与支持。

最后我想感谢我的父母,我的父亲常鼓励我说好男儿要心胸开阔,要勇敢的追求、进取。 他们对我学习物理有足够的理解与支持,让我有时间和精力投入科研。



# NEUTRINO-LESS DOUBLE BETA DECAY SEARCH WITH PANDAX 500 KILOGRAM LIQUID XENON DETECTOR

Neutrino-less double beta decay  $(0\nu\beta\beta)$  is an extremely rare nuclear transition process, during which two neutrons are converted to protons and emits two electrons without neutrinos. Once this process is found, we could claim that neutrinos are their own anti-particles, i.e., a kind of Majorana Fermions.  $0\nu\beta\beta$  is also a physical process beyond the Standard Model (SM) of particle physics, since during which the lepton number is not conserved. Another similar process "two neutrinos double beta decay  $(2\nu\beta\beta)$ " have been observed in many isotopes like <sup>76</sup>Ge, <sup>130</sup>Te and <sup>136</sup>Xe. The searches for  $0\nu\beta\beta$  must base on these isotopes in which  $2\nu\beta\beta$  can happens.

Searching for  $0\nu\beta\beta$  is an ongoing effort in many laboratories around the world. For example, Gerda and CUORE in Italy, SNO+ in Canada, EXO in the U.S. and KamLAND-Zen in Japan. Although none of them have found a signal for the existence of  $0\nu\beta\beta$ , the half-life limit of this process is getting higher and higher. If it exists, the half-life should be longer than  $10^{25} - 10^{26}$  years.

The PandaX-II detector is a large dual-phase liquid xenon time projection chamber (TPC). With the primary goal of detecting possible collisions between xenon nuclei and WIMP (weakly interacting massive particles, dark matter candidate), PandaX-II have published many important results about dark matter with the enormous data collected in the past 3 years. However, the natural isotopic abundance of xenon is 8.9%, which allows the possible detection of  $0\nu\beta\beta$  with the PandaX-II detector considering we still have 51.6 kg<sup>136</sup>Xe out of 580 kg natural xenon. In this paper, I show the  $0\nu\beta\beta$  search results with 407.1 days of data from the PandaX-II experiment. More specifically, I discuss the work I have done to improve the results.

The ultimate goal of this paper is getting a conclusion on the detection of  $0\nu\beta\beta$  with PandaX-II data, so we need to perform spectrum fitting on it. However, before that we also used some techniques to improve the data quality. Data cuts are very important for the analysis and sometimes a good cut would make the signal clear! For instance, in our analysis, the single scatter cut can remove more than half of the events. Since dark matter search is in keV region and  $0\nu\beta\beta$  search is in MeV region, I did a detail study on the performance of original dark matter cuts in MeV region and found all of them need proper modification.

The fiducial volume cut is crucial and fundamental, which is a hard restriction on the event position. We use toy MC to simulate the sensitivity at different fiducial volume and finally decide to use 80000 mm<sup>2</sup> instead of the original 72000 mm<sup>2</sup>. Another important data selection is the position cuts. We remove events with poor reconstruction performance ( $\chi^2$  is too large) and we also require the horizontal distance of the reconstructed position from PAF and NN to be small. Besides, there are some cuts we don't use anymore. The average charge cut uses to be a good way to get rid of sparking events, but in our high energy analysis, this cut turns out to be useless. The



same situation happens for many cuts relying on the top S2 signal charge. When the signal is too large, the PMT will be saturated and the event charge reconstruction is not reliable. Accordingly, the corresponding cuts are also useless.

After I list all the cuts and determine their values, I move on to the background model construction. At first, I select a region from 2000 keV to 3000 keV as our ROI (region of interest), which is reasonable considering the Q-value is 2458 keV. The main background there is  $^{232}$ Th,  $^{238}$ U,  $^{60}$ Co. With the help of Geant4, we get to know the  $\gamma$  peaks and their corresponding Compton edge. From this primary model, we get good result with half-life ~  $10^{23}$  years, but the fit result is significant influenced by  $^{60}$ Co event rate. The uncertainty of  $^{60}$ Co rate is so large! Soon we find out a way to eliminate this problem. The  $^{60}$ Co peak between 2000 keV and 3000 keV is generated by the coupling of another two peaks in 1000 keV – 2000 keV region. Only if we extend our fit region, we could get better knowledge on the event rate of  $^{60}$ Co. Therefore, the 1000 keV – 3000 keV region is fitted as a pre-fit in order to get the decay rate information of  $^{60}$ Co. The 2000 keV – 3000 keV region is fitted as a best-fit to get all the final results we present in the paper.

Obviously, the pre-fit is much more complicated than the best-fit because we are dealing with more  $\gamma$  peaks. Frist of all, we pick out all the peaks from the simulation results. Secondly, in the best-fit model, we use a constant energy resolution which is not suitable for the pre-fit. If we want to set a function for the energy resolution, the fitting process would need a very long time. Sometimes our pre-fit can take several days to be finished. Such a problem is troublesome when we want to adjust the model slightly and update the results.

With the discussion and tests day and night, we decide to smear the MC-simulated spectrum directly and four radioactive isotopes chains are considered in this energy region:  ${}^{60}$ Co,  ${}^{40}$ K,  ${}^{238}$ U and  ${}^{232}$ Th. We assume that the decay chain is in equilibrium within the material. They contribute directly to our model spectrum, and their scaling values just represent event rates. In addition,  ${}^{222}$ Rn spectrum is added with a fixed event rate come from independent analysis.  $2\nu\beta\beta$  of  ${}^{136}$ Xe is also included in this model. Since we directly smear the MC-spectrum, we could use an energy dependent function to describe the energy resolution. For energy larger than 2000 keV, we use a constant value.

Then we fit the real data spectrum with our background model using a toolkit called BAT. We also blind the  $2458 \pm 100$  keV region to prevent any bias caused by subjective factors of us. From this pre-fit, the statistical uncertainties are all known. Then we extend the <sup>60</sup>Co event rate to our best-fit region and transmit its error by multiplying the square root of the event ratio.

In the best-fit, we utilize the pre-fit information and set the <sup>60</sup>Co event rate to be float but with a Gaussian prior. And we float the <sup>232</sup>Th, <sup>238</sup>U, additional flat background and  $0\nu\beta\beta$  as parameters. <sup>222</sup>Rn is fixed. <sup>40</sup>K and  $2\nu\beta\beta$  is ignored. With this new model, we get the best-fit decay rate of  $0\nu\beta\beta$ :  $\Gamma_{0\nu} = (-0.24 \pm 0.21) \times 10^{-23} yr^{-1}$ .

An estimation of systematic uncertainties induced by fixed parameters is necessary. So I utilize pseudo-experiments, called toy MC, to investigate the effects of systematics on the final results. The ideal behind pseudo-experiment is simple. Assuming that we have 1000 parallel universes and all of them have their own PandaX-II. If we collect these 1000 results, we would expect that they distribute around the real physical results. Although we cannot get their results directly, we take the fitted results of our data as "real" model and generate pseudo-data using Poisson distribution. Therefore, we can have a comparison between pseudo-experiments and our real experiments.



Studies on systematics for experiments is usually a difficult process because multiple effects must be taken into account. It is also not trivial to investigate how the uncertainty of one parameter propagates into the systematic uncertainty of final results. Except the only scaling effect of detection uncertainty, we cannot get the contribution of any other uncertainties directly. However, by using toy MC, they can be studied straightforwardly. For instance, if I want to know how the change of resolution can affect the final results, I only need to modify the resolution a little bit and do the fitting again with the new generated pseudo-experiment. We can add some signals in the pseudo-experiment, so the difference between input signals and fitted signals is just the uncertainty propagated to final results.

Finally, we propagate to a physical Bayesian limit from the best-fit  $0\nu\beta\beta$  decay rate by combing the posterior probability distribution and systematic uncertainties. We conclude that we found no evidence of  $0\nu\beta\beta$  in PandaX-II experiment. We also redo the fitting process on pseudo-experiments with only positive  $0\nu\beta\beta$  decay rate and extract the 90% quantile value for limit calculation. After 12000 loop, the median number is taken as sensitivity.

Even though we didn't find  $0\nu\beta\beta$  in the 407.1 day's data, we have developed new techniques for data selection and energy reconstruction in high energy region. Our work yields the first  $0\nu\beta\beta$ search result from dual-phase liquid xenon detector and the half-life limit for <sup>136</sup>Xe is given to be  $T_{1/2}^{0\nu} > 1.4 \times 10^{23}$  yr. This work shows the capability of dual-phase liquid xenon detector on  $0\nu\beta\beta$  detection. Considering this, many experimental groups using tons of liquid xenon detectors to detect dark matter are paying more and more attention to the full utilization of their detectors. The next generation of dual-phase liquid xenon detectors will be expected to achieve multiple physical targets, covering a wide range of energy region while having a low background and good energy resolution.

The PandaX-4T is still under construction. Ultra-pure water shielding and tons of liquid xenon self-shielding will bring a very low background to the detector. We could expect the sensitivity of the PandaX detector to be better in the future.