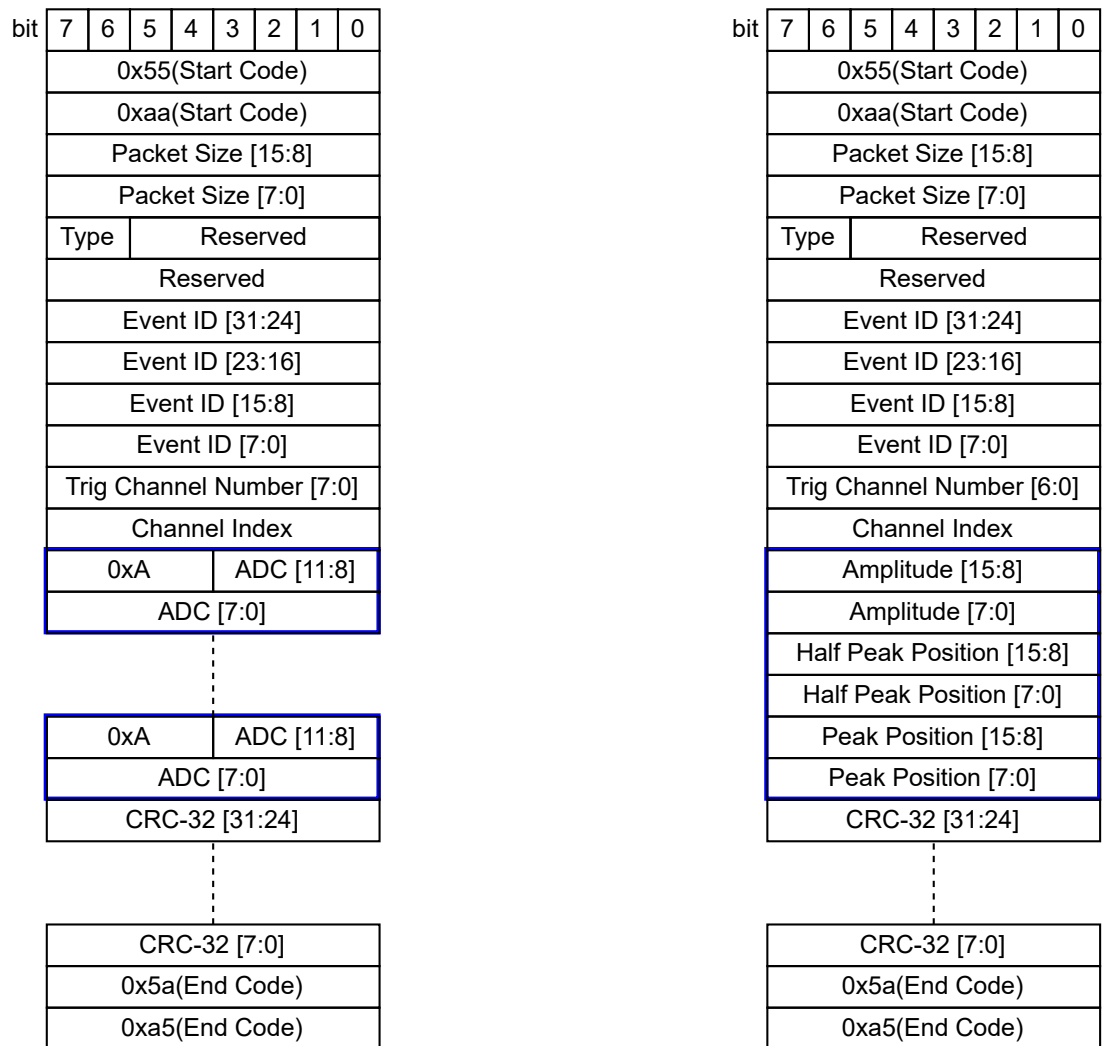


第2章 数据分析

2.1 数据处理流程

2.1.1 数据解包、信息提取

DPU 通过 SiTCP 上传的数据文件格式为 .dat，每个文件包含一次测量的所有数据帧。数据帧有两种类型，一种是波形信息，另一种是时间幅度提取信息，两种信息的数据帧格式如图 2.1 所示。



(a) 包含波形信息的数据帧（一个数据帧包括一个电子学通道的 1024 个波形数据点）

(b) 包含时间幅度提取信息的数据帧（一个数据帧包括一个电子学通道波形的幅度、半峰值时间、峰值时间提取信息）

图 2.1 数据帧格式

将上述感兴趣的数据统一存入名为 Packet 的结构体中，其结构如下：

```

1 struct Packet
2 {
3     uint8_t u8Type = 0;
4     uint32_t u32EventID = 0;
5     uint8_t u8TrigChanum = 0;
6     uint8_t u8ChannelIndex = 0;
7     uint16_t au16ADC[Config::ADCPOINTS] = {};
8     uint16_t u16Amplitude = 0;
9     uint16_t u16HalfTime = 0;
10    uint16_t u16PeakTime = 0;
11 };

```

如果数据帧为波形信息，先判断波形数据是否饱和，如果饱和则将饱和部分的数据点进行拟合。其中波形前沿使用线性拟合，波形后沿使用指数拟合，拟合函数如下：

$$f(x) = \begin{cases} a \cdot x + b, & x \leq x_0, \\ c \cdot \exp(-d \cdot x) + e, & x > x_0. \end{cases} \quad (2.1)$$

然后，对波形数据进行信息提取，包括波形的幅度、半峰值时间、峰值时间等信息。最后，将波形提取信息或时间幅度提取信息存入名为 TreeEntry 的结构体中，再将其以以下结构写入 ROOT 文件中。

```

1 tTree = new TTree("tTree", "Tree for data
    ↪ extraction");
2
3 tTree->Branch("EventID", &sTreeEntry.iEventID,
    ↪ Form("EventID/I"));
4 tTree->Branch("TriggerChannelNumber",
    ↪ &sTreeEntry.iTrigChanum,
    ↪ Form("TriggerChannelNumber/I"));
5 tTree->Branch("ChannelIndex",
    ↪ sTreeEntry.aiChannelIndex,
    ↪ Form("ChannelIndex[%d]/I", Config::CHANNELNUM));
6 tTree->Branch("ADC", sTreeEntry.viADC.data()); // Do
    ↪ not store ADC data in the tree.
7 //tTree->Branch("ADC", &sTreeEntry.viADC); // If ADC
    ↪ data is usefull, use this line.

```

```

8  tTree->Branch("EventIndex", &sTreeEntry.iEventIndex,
   ↪  Form("EventIndex/I"));
9  tTree->Branch("MeanOfBaseline",
   ↪  sTreeEntry.aiMeanOfBaseline,
   ↪  Form("MeanOfBaseline[%d]/I", Config::CHANNELNUM));
10 tTree->Branch("SubMeanOfBaseline",
   ↪  sTreeEntry.aiSubMeanOfBaseline,
   ↪  Form("SubMeanOfBaseline[%d]/I",
   ↪  Config::CHANNELNUM));
11 tTree->Branch("SigmaOfBaseline",
   ↪  sTreeEntry.aiSigmaOfBaseline,
   ↪  Form("SigmaOfBaseline[%d]/I",
   ↪  Config::CHANNELNUM));
12 tTree->Branch("Amplitude", sTreeEntry.aiAmplitude,
   ↪  Form("Amplitude[%d]/I", Config::CHANNELNUM));
13 tTree->Branch("HalfTime", sTreeEntry.aiHalfTime,
   ↪  Form("HalfTime[%d]/I", Config::CHANNELNUM));
14 tTree->Branch("PeakTime", sTreeEntry.aiPeakTime,
   ↪  Form("PeakTime[%d]/I", Config::CHANNELNUM));
15 tTree->Branch("Peak2Peak", sTreeEntry.aiPeak2Peak,
   ↪  Form("Peak2Peak[%d]/I", Config::CHANNELNUM));
16 tTree->Branch("SumOfAmplitude",
   ↪  &sTreeEntry.iSumOfAmplitude,
   ↪  Form("SumOfAmplitude/I"));

```

2.1.2 反符合剔除

将 2.1.1 节中保存的 ROOT 文件中的数据进行反符合剔除，剔除方法为：依次遍历每个事例，再对每个事例中的每个通道的通道号进行判断，如果通道号对应的 Map 为反符合通道，则将该事例整体剔除。剔除后的数据保存在新的 ROOT 文件中。

2.1.3 特征提取

1. 特征

对 2.1.2 剔除后的数据或 2.1.1 的数据进行特征提取，提取的特征包括：

- 事例总幅度

- X 维击中数
- Y 维击中数
- X&Y 维击中数矢量和
- X&Y 维击中数差
- 事例各通道最大峰值时间和最小峰值时间之差
- X 维最远击中位置
- Y 维最远击中位置
- 事例最大能量沉积的相对时间位置
- θ 入射角
- ϕ 入射角
- X 维 Pearson 相关系数
- Y 维 Pearson 相关系数
- X 维径迹长度
- Y 维径迹长度
- 单位距离能量沉积
- 是否击中边缘通道

2. 数据筛选

对特征提取后的数据进行筛选，筛选条件为：X/Y 维击中数为 0、X&Y 维径迹长度小于 2（这里径迹长度的算法有问题，待修正）和击中位置在膜窗外的事例。筛选后的数据保存在新的 ROOT 文件中。

2.1.4 TMVA 训练

使用 BDT(Bosted Decision Tree) 方法进行训练，训练的特征包括：

- 事例总幅度
- X 维击中数
- Y 维击中数
- 事例各通道最大峰值时间和最小峰值时间之差
- X 维最远击中位置
- Y 维最远击中位置
- 事例最大能量沉积的相对时间位置
- θ 入射角
- X 维 Pearson 相关系数
- Y 维 Pearson 相关系数
- X 维径迹长度
- Y 维径迹长度

- 单位距离能量沉积

2.1.5 TMVA 测试

用 TMVA 训练好的 BDT 方法对测试数据进行测试，测试数据的特征提取方法与训练数据相同。测试结果保存在 ROOT 文件中。

2.2 数据分析结果

2.2.1 使用前期阳极条两两合并的数据进行训练

表 2.1 阳极条两两合并的数据训练结果

源类型	^{90}Sr	本底
工作气体	C_4H_{10}	C_4H_{10}
高压 [V]	370	370
时间戳	20240804231807	20240807112240
测试时间	2min	9h
原始计数	150293	299061
去除反符合计数	147597	227822
去除单维事例计数	147597	227812
去除短径迹、膜窗外计数	142647	41313
BDT 筛选后计数 (率)	85637(714 cps)	4339(8.0 cpm)
发射率 [s^{-1}]	—	—
效率 [%]	57	98.5
阈值	0.06	0.06
类型	WV 训练数据	WV 训练数据

2.2.2 使用太原标准源中心阳极条测试数据进行训练

表 2.2 太原标准 ^{90}Sr 源中心阳极条测试数据训练结果

源类型	^{90}Sr	本底	^{90}Sr	本底
工作气体	CO_2	CO_2	CO_2	CO_2
高压 [V]	540	540	540	540
时间戳	20240906172844	20240923194604 20240924214355	20240906181811	20240925200257
	20240906180546		20240906185136	
	20240906180839		20240906185441	
	20240906181151		20240906185732	
	20240906181440		20240906190011	
测试时间	10 min	24 h	10 min	12 h
原始计数	592221	449884	591632	225801
去除反符合计数	587572	412595	586741	207047
去除单维事例计数	586408	321045	585652	160728
去除短径迹、膜窗外计数	568140	146392	567579	74081
BDT 筛选后计数 (率)	397135(661.9 cps)	1544(1.1 cpm)	398677(664.5 cps)	888(1.2 cpm)
发射率 [s^{-1}]	1190	—	1190	—
效率 [%]	55.6	99.6	55.8	99.6
阈值	0.12	0.12	0.12	0.12
类型	TQ 训练数据	WV 训练数据	TQ 测试数据	WV 测试数据

表 2.3 太原标准 ^{36}Cl 源中心阳极条测试数据训练结果

源类型	^{36}Cl	本底	^{36}Cl	本底
工作气体	CO_2	CO_2	CO_2	CO_2
高压 [V]	540	540	540	540
时间戳	20240906182240	20240923194604 20240924214355	20240906183738	20240925200257
	20240906182634		20240906184015	
	20240906182953		20240906184257	
	20240906183223		20240906184537	
	20240906183458		20240906184819	
测试时间	10 min	24 h	10 min	12 h
原始计数	283263	449884	284321	225801
去除反符合计数	280745	412595	281692	207047
去除单维事例计数	279594	321045	280569	160728
去除短径迹、膜窗外计数	268472	146392	269790	74081
BDT 筛选后计数 (率)	172816(288.0 cps)	1617(1.1 cpm)	174151(290.3 cps)	920(1.3 cpm)
发射率 [s^{-1}]	502	—	502	—
效率 [%]	57.4	99.6	57.8	99.6
阈值	0.12	0.12	0.12	0.12
类型	TQ 训练数据	WV 训练数据	TQ 测试数据	WV 测试数据