# Chapter 1

KamLAND-ZEN Simulation and Reconstruction

KamLAND-ZEN uses detailed simulations defined in KLG4Sim, a GEANT4-based Monte Carlo (MC) simulation software. The MC simulated events are tuned with real calibration events to carefully match real detector response. Simulated and physical events produce detector responses that are reconstructed to extract higher-level information such as energy and position. The reconstructed event information is used for data-selection and spectrum fitting. This chapter discusses the MC simulation and event reconstruction procedures used in KamLAND-ZEN 800.

1.1 Analysis Framework

1.1.1 Data Flow

Figure 1·1 outlines the data flow in KamLAND-ZEN. PMT signals are digitized in either KamFEE or MoGURA, the two DAQ systems discussed in the previous chapter, the digitized signals are stored in Kinoko Data Format (KDF). KDF files contain trigger information and timestamped, digitized PMT waveforms. KDF files also store run condition information in the header. The EventBuilder collates the waveforms of a single event and stores them in a serial file. A waveform analyzer reconstructs a hit time and charge (TQ) information for each of these waveforms. The RTQ files hold

the Raw-TQ information for each PMT. Event vertices and visible energy are derived from the RTQ files through their respective reconstruction algorithms. There are secondary reconstructions that are also applied to the RTQ files, such as muon track fitting, flasher vetos, double pulse fit, and unphysical event selections. The general vector file (GVF) is used for the main physics analyses like the one presented in this thesis.
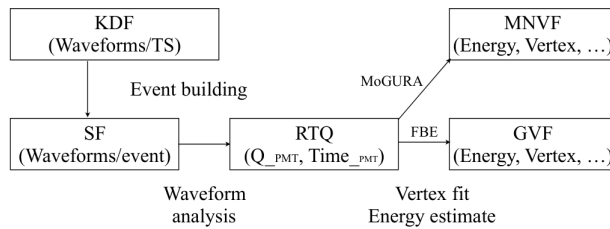
Figure 1·1: Data flow in KamLAND from raw waveforms to analysis variables such as energy, vertex, total hit PMTs, etc. [2]

GVF files contain the following information:

- **run number**

- **event number**

- **TimeStamp** based on DAQ clock time (25 nsec for KamFEE, 20 nsec for MoGDAQ)

- **unixtime** is the number of seconds from January 1st in 1970 and used for some run vetos

- **trigger type** records which trigger was used

- **event vertex and badness** event vertices and a radius from detector center is saved, along with a vertex fit quality parameter called badness

- **energy/energy17** visible energies given by the fitter, energy17 is the energy estimate only using 17-inch PMTs

- **TotalChargeID/17/OD** sum of all PMT charges of each PMT type

- **numhit/numhit17** the number of hit PMTs/17-inch PMTs in each event.

- **NsumMax** a maximum number of hit PMTs in a single DAQ cycle within each event, a "peak" nhit of the event

- **N200OD** Maximum number of simultaneous hit OD pmts within 200nsec windows

- **muon entrance and direction** muon fitter results are recorded.

Finally, MoGURA events are associated with muon events acquired in KamFEE DAQ (FBE) and stored in a Muon-Neutron Vector File (MNVF) to search for neutron capture events that occur shortly after muons.

1.2 Event Reconstruction

1.2.1 Waveform Analysis

Each digitized waveform has 128 samples with 1.5ns sample times; this corresponds to a waveform digitization window of 192 ns. The waveforms are processed and TQ reconstructed with the following procedure.

- **Smoothing** Each waveform is smoothed via a running-average first derivate.

- **Baseline adjustment** The baseline of each PMT is collected at the beginning of each run. This baseline is subtracted from each waveform.

- **Peak finding** Peaks are found with running-averaged 1st, 2nd, and 3rd derivatives.

- **Leading-edge and Trailing-edge tag** A leading-edge is stamped as 10ns before the peak voltage. The trailing edge is stamped when the waveform returns to baseline. An example of this time-stamping is shown in Figure 1·2

- **Waveform Sum calculation** The waveform is integrated from the leading-edge to the trailing-edge.
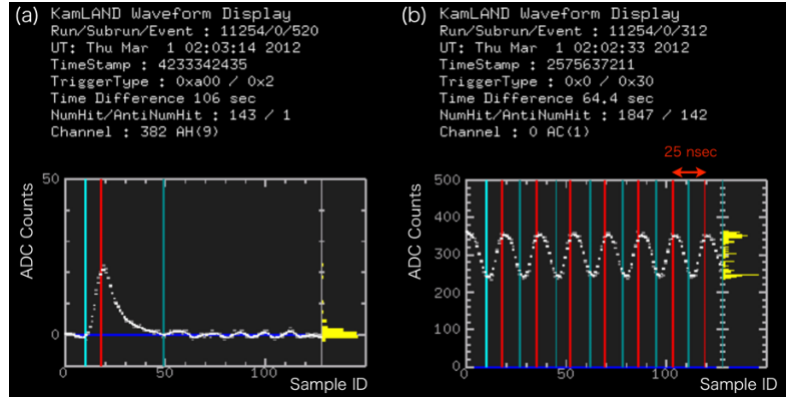
Figure 1·2: An example of waveform analysis from thesis [3]. (left) ADC counts of a real waveform after baseline subtraction. The left cyan line is the leading edge, the center red line is the peak position, and the right dark cyan line is a trailing-edge. (right) Clock calibration example on 25 nsec intervals.

When there are mulitple hits in one PMT waveform, the total charge of the hits and the earliest hit time is returned. This simplified information is used for the vertex and energy reconstruction. The multi-pe information is used for a double-pulse fit and a muon shower tag.

1.2.2 PMT Corrections

Low Gain Problem and HV Reductions

Since 2011, we observed that the gain of some of the 17-inch PMTs gradually decreased. As the gain of the PMTs fell, this compormises signal/background ratio and

PMT waveform quality. It was also observed that the PMTs enter a low impedance state before the gain dropped. A HV current and voltage monitor allows for the real-time monitoring of this state. Usually a simple HV power cycle could recover normal PMT behavior. Since 2016, a auto HV power cycle mechanism was implemented to mitigate the low gain problem, but the root case is still unknown.

Each time the PMTs enter the low impedance state, the HV on that channel was reduced in 50-100 V increments. Over time, some of the channels had their HV reduced by 450 V. FIgure 1·3 shows the trend in low gain 17-inch PMTs.
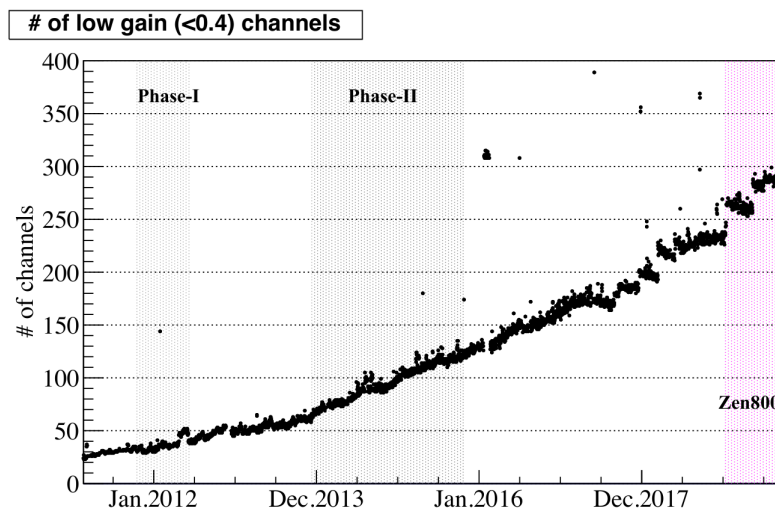
Figure 1·3: The trend in the number of low gain 17-inch PMTs, before ZEN-800. The number of low gain channels increased gradually, while the sudden increases are from HV reductions performed since 2017 [2].

Note about current low pmt gain analysis.

Bad Channel

A channels is considered bad if the PMT meets one or more of the following criteria:

- PMT pulses less than 0.6% of the time over all events

- PMT pulses below 0.48% for non-muon events

- PMT pulses less than 80% of the time for high-energy muon events

- PMT is missing a waveform more than 10% of the time

- Large discrepancy between the two ATWD hits

- High muon charge PMTs. A PMT may read much higher charge $(Q_{detected})$ than the average of its surrounding PMTs $(Q_{expected})$. A run is divided into 100 muon intervals, for each interval the criteria is defined as

$$\frac{1}{N_{interval}} \sum_{i=1}^{N_{interval}} \left(\frac{1}{N_{muon}} \sum_{j=1}^{N_{muon}} \frac{(Q_{expected} - Q_{detected})^2}{Q_{expected}} \right) > 1000 \ p.e.$$

These bad channels are excluded from event reconstruction and physics analyses.

Dark Hit

Thermal fluctuations can emit electrons off the photocathode leading to a PMT hit signal. These "dark hits" are an unavoidable hit-level background in PMT detectors, lowering the detector temperature reduces this effect. The dark hit rates are measured from run-to-run and are factored into our likelihood-maximizing reconstruction algorithms. The hit rate observed 50-100 ns before the PMT hittime rising edge is taken as the dark rate, Figure 1·4 shows the PMT hittime distribution and the dark rate window.

1.2.3 Primary Vertex Fitter

The primary vertex fitter provides a rough estimate of a scintillating event's location. This estimate serves as the input to a more thorough, but complex secondary fitter.
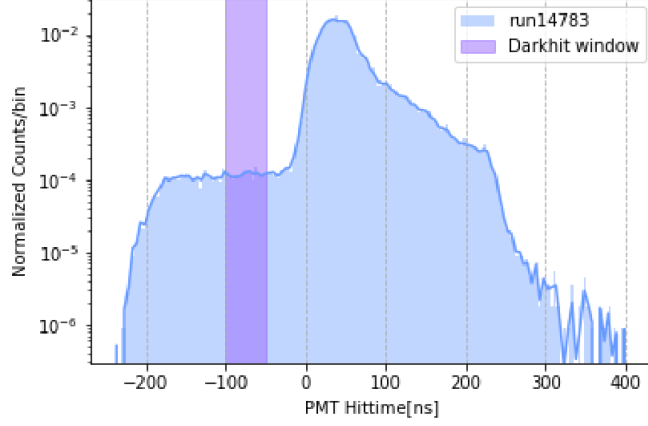
Figure 1·4: An example pmt hit time distribution from data run 14783, the 50-100 ns leading window is taken to measure the pmt dark hit rate. [1].

The fit works by constructing a hit time residual distribution:

$$T_i^{emit} = T_i - TOF_i = T_i - \frac{|R_i - r_{vertex}|}{c_e ff} \qquad (1.1)$$

Here T_i is the hit time of the ith PMT, TOF_i is the time it takes for a scintillation photon to traverse from the vertex position to the ith PMT position, R_i is the PMT position, r_{vertex} is the unknown vertex position to fit for, and c_{eff} is the speed of light in the given medium. By fitting T_i^{emit} to match the standard scintillation time profile, a primary r_{vertex} is produced by the fitter.

1.2.4 Secondary Fitter

The secondary V2 fitter uses the r_{vertex} given by the primary fitter to compute T_0 according to the equation 1.2

$$T_0 = \frac{\sum_i \left(T_i^{pmt} - TOF_i^{pmt} \right) \times Q_i}{\sum_i Q_i} - const. \qquad (1.2)$$

This T_0 is the charge weighted sum of T_emit from 1.1. This T_0 serves as the universal
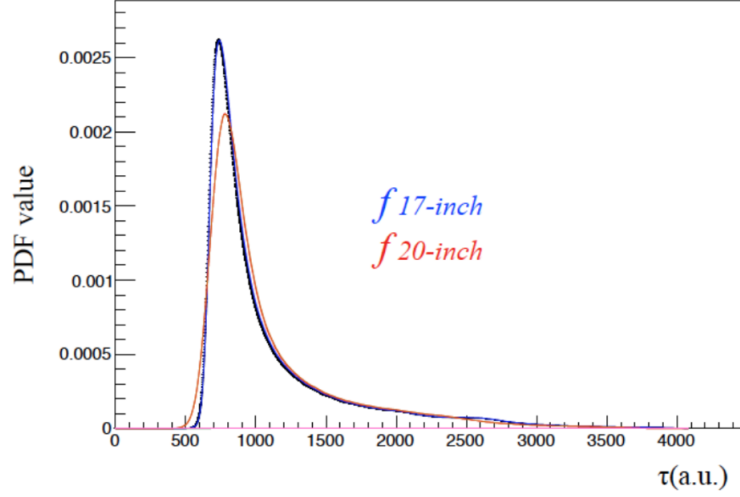
Figure 1·5: Probability density function of 17-inch and 20-inch PMT hit times calculated from calibration data. The plot is from [2] and originally from a 2005 calibration dataset.

start point of an event. From this time, each pmt hit time is

$$\tau(x, y, z, T_0) = T_i^{pomt} - TOF_i^{pmt} - T_0 \qquad (1.3)$$

Finally, these time-of-flight corrected and centered hit time distributions are used to create probability distributions for the 17 and 20 inch PMTs respectively. These PDFs are shown in Figure **??**. The likelihood function for an individual PMT is defined as:

$$\phi_i = \frac{\mu \times f_i(\tau_i) + D_i}{\mu \times C_{17/20} + D_i} \qquad (1.4)$$

Here, μ is the pulse shape determination factor, D_i is the dark hit rate for the ith PMT and $C_{17/20}$ is the normalization constant for the 17 or 20 inch PMTs. The overall log-likelihood is given by the $log(L) = \sum_i log(\phi_i)$. The log-likelihood is maximized by the Newton-Raphson method, in which the x, y, z, T_0 are adjusted to the best-fit values, giving us the V2 reconstructed vertex.

1.2.5 Energy Reconstruction

1.2.6 Muon Reconstruction

1.2.7 MoGURA Neutron Reconstruction

1.3 Event Selection

Bibliography

[1] A. Li. *The Tao and Zen of neutrinos: neutrinoless double beta decay in KamLAND-ZEN 800.* PhD thesis, Boston University, 2020.

[2] H. Ozaki. *High Sensitivity Search for Neutrinoless Double-Beta Decay in KamLAND-Zen with Double Amount of ^{136}Xe.* PhD thesis, Tohoku University, 2019.

[3] H. Yoshida. *Limit on Majorana Neutrino Mass with Neutrinoless Double Beta Decay from KamLAND-Zen.* PhD thesis, Tohoku University, 2014.