Likelihood Model of Combined Vertex and Energy Reconstruction

Benda Xu

2014-06-19

Contents

Abstract

A complete probability model of vertex and energy of KamLAND event is derived. It is used in simutaneous event reconstruction.

1 PMT charge

The charge of PMT depends on number of photo-electron (PE). The PMTs (Hamatsu 20inch and 17inch) used in KamLAND have a good linear response.

1.1 17 inch PMT

The charge of 1 PE can be modeled with a normal distribution. Assuming different PE's are of identical independent distribution (iid), the distribution of charge given number of PE (N) is,

$$
\begin{aligned}
q = \sum_{i=1}^{N} q_i, q_i \sim \mathcal{N}(\mu_q, \sigma_q) \\
q \sim \mathcal{N}(N\mu_q, \sqrt{N}\sigma_q)
\end{aligned} . \tag{1}
$$

In KamLAND, μ_q is normalized to 1 on every PMT, and σ_q is fitted with all the PMTs on average to be 0.39 [1]. Note that q>0, and the probability of q<0 in Eqn(1) is ignored. After all, it is better treated in the following Eqn(3).

N, in turn, is described by a Poissonian of intensity λ, also called the *occupancy factor*. Noting when N=0, q=0, the probability being e$^{-\lambda}$, the cumulative distribution function(CDF) $C(q|\lambda)$ and the probability distribution function(PDF) $c(q|\lambda)$ of the resulting compound Poisson distribution of q are,

$$
\begin{aligned}
C(q|\lambda) &= e^{-\lambda}\left[1 + \sum_{N=1}^{\infty} \frac{\lambda^N}{N!}\Phi\left(\frac{q - N\mu_q}{\sqrt{N}\sigma_q}\right)\right] \\
c(q|\lambda) &= e^{-\lambda}\left[\delta(q) + \sum_{N=1}^{\infty} \frac{\lambda^N}{N!}\frac{1}{\sqrt{2\pi N}\sigma_q} e^{-\frac{(q-N)^2}{2N\sigma_q^2}}\right]
\end{aligned} \tag{2}
$$

where $\delta(q)$ is the Dirac delta function, and Φ(q) is the CDF of the standard (μ=0, σ=1) normal distribution.

In KamLAND, the *hit* of a channel refers to the charge of the channel crossing a certain threshold. A charge threshold of 0.3 for every *single* PE is imposed on the front-end electronics(FEE), and a total charge of 0.3 for *all* the PEs on offline data analysis. Formally, the *hit* in these two cases are defined as $\exists i, q_i > 0.3$ and $\sum_{i=0}^{N} q_i =: q > 0.3$ respectively. The non-hit probability of the latter is $C(0.3|\lambda) =: Z(\lambda)$, while that of the former is

$$
\begin{aligned}
\widetilde{Z}(\lambda) &:= e^{-\lambda} \sum_{N=0}^{\infty} \frac{\lambda^N}{N!}\Phi^N\left(\frac{0.3 - \mu_q}{\sigma_q}\right), \\
&= e^{-0.9652\lambda}(\mu_q = 1, \sigma_q = 0.39)
\end{aligned} \tag{3}
$$

Note that $\widetilde{Z}(\lambda_1 + \lambda_2) = \widetilde{Z}(\lambda_1)\widetilde{Z}(\lambda_2)$.

The leading orders of Φ in the summations of Z and $\widetilde{Z}$ are calculated in Table 1. In KamLAND, $\lambda = O(1)$, the common terms of $N = 0, 1$ dominates. Therefore for simplicity, we use Z and $\widetilde{Z}$ interchangeably from now on. Evaluating Z and $\widetilde{Z}$ numerically to 1×10^{-50} confirms our speculation even to $\lambda \sim 100$, as in Figure 1.

[1] The number literally comes from the source code of KatEnergyA2.cc originally by M. Batygov as 0.385861.

Table 1: leading orders of Φ in the summations of Z and $\widetilde{Z}$

N	$\Phi^N\left(\frac{0.3-\mu_q}{\sigma_q}\right)$	$\Phi\left(\frac{0.3-N\mu_q}{\sqrt{N}\sigma_q}\right)$
1	0.036	0.036
2	0.0013	0.0010
3	4.8×10^{-5}	3.2×10^{-5}

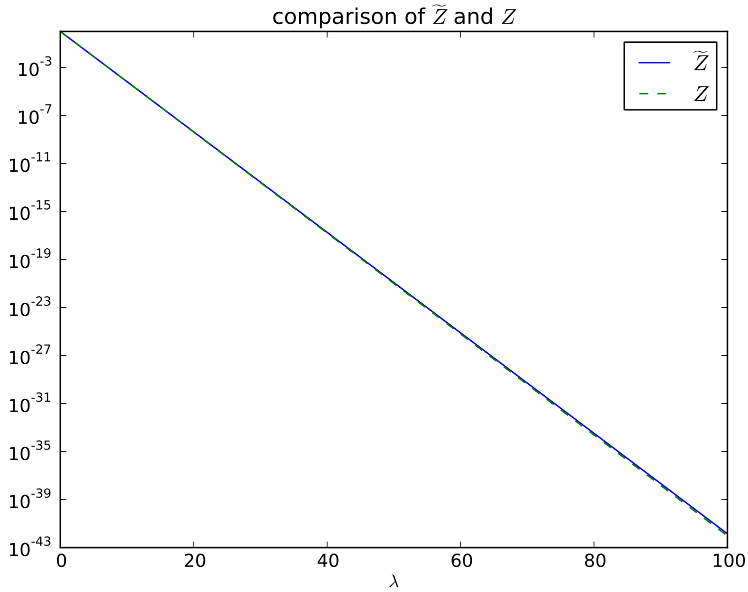

Figure 1: comparison between Z and $\widetilde{Z}$

1.2 20 inch PMT

In contrast to its 17 inch counterparts, 20 inch PMTs do not have a clear 1 PE peak, and are not eligible to be modeled with a normal distribution. The charges of 20 inch PMTs have been normalized to the surrounding 17 inch PMTs. The expected charge of 1 PE is 1. Without a well defined peak, by the maximum entropy principle, the charge distribution of 1 PE of 20 inch PMT is model by an exponential of rate 1. Similar to Eqn(1),

$$\begin{aligned} q = \sum_{i=1}^{N} q_i, q_i \sim \mathcal{E}(\mu_q), f(q_i) = \frac{e^{-\frac{q_i}{\mu_q}}}{\mu_q} \\ q \sim \text{Erlang}(N, \mu_q), f(q) = \frac{q^{N-1}e^{-\frac{q}{\mu_q}}}{\mu_q^N (N-1)!} \end{aligned}, \tag{4}$$

where $\mu_q = 1$.

In this case,

$$c_{20}(q|\lambda) := e^{-\lambda} \sum_{N=1}^{\infty} \frac{\lambda^N}{N!} \frac{q^{N-1} e^{-\frac{q}{\mu_q}}}{\mu_q^N (N-1)!}. \tag{5}$$

Exponential distribution has a large concentration near 0,

$$\int_0^{0.3} e^{-q} \mathrm{d}q = 0.259 \tag{6}$$

$$Z_{20}(\lambda) = e^{-0.741\lambda} \tag{7}$$

2 PMT timing

Each event has a light curve $\psi(\tau)$, where τ is the time offset from an arbitrary reference point (taken as $\sim$ 50 ns before rising edge in KamLAND) in the light curve. ψ, however, get distorted by the *occupancy effect* depending on λ, into $\tilde{\phi}(\tau|\lambda)$, the normalized distorted light curve (Eqn(10) of [1]),

$$\tilde{\phi}(\tau|\lambda) = \frac{\lambda\tilde{\psi}(\tau)}{1 - e^{-\lambda}} e^{-\lambda\tilde{\Psi}(\tau)}.$$

3 Variables

Two measurables are given by a hit PMT, $\mathrm{t_p}$, the time (more precisely, rising edge of the first PE), and $\mathrm{q_p}$, the charge.

(τ, λ), intermediate variables, depends on the event space-time $s_i := (t, \mathbf{r})$ and visible energy E (reconstruction target variables) together with PMT output ($\mathrm{t_p}$, $\mathrm{q_p}$, measured variables) and location ($\mathbf{r}_p$, detector constants),

$$\begin{aligned} \tau &= t_p - t - R_p \\ \lambda &= \frac{\lambda_0}{R_p^2} E \qquad , \\ \mathbf{R}_p &= \mathbf{r} - \mathbf{r}_p \end{aligned} \tag{8}$$

where λ_0 is the component of λ that does not depend on R_p or E, $R_p = |\mathbf{R}_p|$ is the norm of vector $\mathbf{R}_p$. The light speed is normalized to 1 for simplicity.

In KamLAND, λ_0 is modeled with light attenuation of LS including scatter and absorption, the incident angle of photon to PMT, the Kevlar rope shadow effect on PMT, and some unknown z-asymmetry effect modeled empirically. These facters do depend on $\mathbf{R}_p$. They just vary so slowly that their contribution to the gradient is ignored for simplicity.

4 Dark Charge and KamFEE Window

Dark charges are observed in PMTs that are not related to any known physics events originating from LS. They are speculated to be of electronic noise or thermal-excitated electrons, and are modeled with a constant rate, the rate coefficient δ (PE/ns) being taken as an average of over vacant time windows of each run[4, p. 63]. It is of order 10^{-5} PE/ns.

Dark charge is introduced in the probability model by taking the substitution scheme of Eqn(11-13) from [1],

$$\begin{aligned} \lambda &\to \lambda_d = \lambda + D \\ \tilde{\psi}(\tau) &\to \tilde{\psi}_d(\tau) = \frac{\lambda\tilde{\psi}(\tau) + \delta}{\lambda + D} \\ \tilde{\Psi}(\tau) &\to \tilde{\Psi}_d(\tau) = \frac{\lambda\tilde{\Psi}(\tau) + \delta(\tau - \tau_i)}{\lambda + D} \end{aligned} , \tag{9}$$

where $\tau - \tau_i = t_p - t_i$ is the PMT hit time(t_p) offset from the earliest time(t_i) at which an FBE could have been triggered, i.e. the time window permitting a dark charge to be recorded. It has two components accordingly, T_o, the first rising edge relative to the beginning of the waveform, and the *launch offset*, labeled by ξ clocks, the clocks between *launch command* (beginning of the waveform) and the digitization command from trigger. (Figure 2) [2, p. 31]

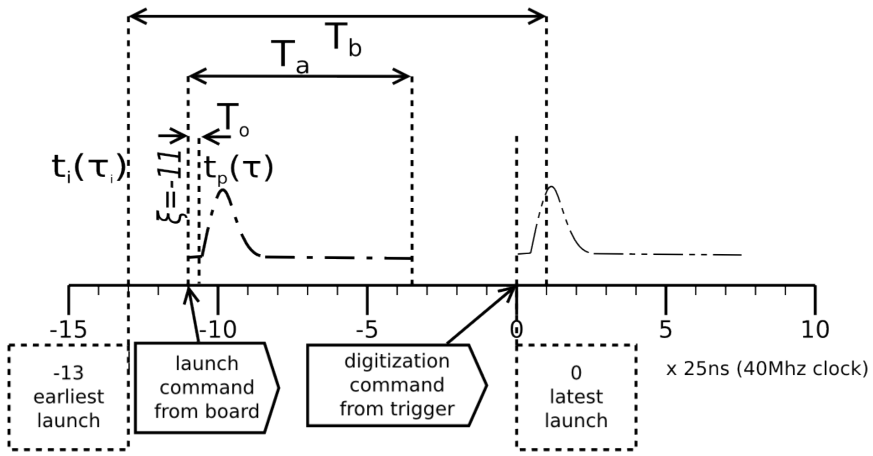

Figure 2: Event timing of KamFEE. t_i/t_p and τ_i/τ are time coordinates relative to KamFEE and light curve respectively, the difference being up to a constant offset.

KamFEE have at most 13 launch offsets of 25ns(T_c), and a waveform length (T_a) of about 192.768(=128 × 1.506)ns. The total time window (T_b) for the launch command to initiate waveform recording is 350(=25 ×

(13+1))ns. [3, p. 20] Therefore $\tau \in [\tau_i, \tau_i + T_b]$ and D, the expected number dark PE within $\mathrm{T_b}$, is δT_b, and $\xi \in \{-13, -12, \ldots, -1, 0\}$.[2] The waveform has to have a hit within first clock to get recorded, $T_o \in [0, T_c]$.

Based on the designed efficiency of the DAQ system, $\psi(\tau)$ falling outside $\mathrm{T_b}$ is ignored.

Define

$$\begin{aligned}
\Psi_d &:= (\lambda + D)\tilde{\Psi}_d \\
&= \lambda\tilde{\Psi}(\tau) + \delta(\tau - \tau_i) \\
\Psi_a^b &:= \Psi_d(b) - \Psi_d(a) \\
&= \lambda(\tilde{\Psi}(b) - \tilde{\Psi}(a)) + \delta(b - a) \\
\Psi^\tau &:= \Psi_{-\infty}^\tau = \Psi_d(\tau)
\end{aligned} \quad . \tag{10}$$

The probability density of KamFEE taking a waveform of charge q at τ is the probability of non-hit in previous $\tau - \tau_i$ multiplied by probability density of observing charge q in the remaining waveform $[\tau, \tau - T_o + T_a]$, minus the probability density of charge q concentrated away from τ by an infinitesimal η in the waveform $[\tau + \eta, \tau - T_o + T_a]$:

$$H(\tau, q|\lambda) := \tag{11}$$

$$= \frac{1}{\eta}\left[Z(\Psi^\tau)c(q|\Psi_\tau^{\tau - T_o + T_a}) - Z(\Psi^{\tau+\eta})c(q|\Psi_{\tau+\eta}^{\tau - T_o + T_a})\right] \tag{12}$$

$$= -\frac{\partial}{\partial\eta}\left[Z(\Psi^{\tau+\eta})c(q|\Psi_{\tau+\eta}^{\tau - T_o + T_a})\right], \eta \to 0 \tag{13}$$

.

Starting from requiring there to be a hit in $[\tau, \tau + \eta]$, Eqn(11) could also be obtained from

$$\begin{aligned}
&(\omega(q|\Psi_\tau^{\tau+\eta})c(q|\Psi_\tau^{\tau+\eta})) * A(q|\Psi_{\tau+\eta}^{\tau - T_o + T_a}) \\
=&A(q|\Psi_\tau^{\tau+\eta}) * A(q|\Psi_{\tau+\eta}^{\tau - T_o + T_a}) - Z(\Psi_\tau^{\tau+\eta})A(q|\Psi_{\tau+\eta}^{\tau - T_o + T_a}) \\
=&A(q|\Psi_\tau^{\tau - T_o + T_a}) - Z(\Psi_\tau^{\tau+\eta})A(q|\Psi_{\tau+\eta}^{\tau - T_o + T_a}) \\
=&c(q|\Psi_\tau^{\tau - T_o + T_a}) - Z(\Psi_\tau^{\tau+\eta})c(q|\Psi_{\tau+\eta}^{\tau - T_o + T_a}) \\
=&\frac{\eta H(\tau, q|\lambda)}{Z(\Psi^\tau)}
\end{aligned} \quad ,$$

where $*$ stands for convolution by q, $A(q|\lambda)$ is the short-hand notation of the compound distribution. $A(q|\lambda_1 + \lambda_2) = A(q|\lambda_1) * A(q|\lambda_2)$.

[2] The KamFEE later changed this scheme and shifted ξ 4 clocks earlier, in order to use the waveform before events to estimate the dark charge rate.

Writing Eqn(11) explicitly,

$$H(\tau,q|\lambda)= \tag{14}$$

$$=-\frac{\partial}{\partial\eta}\left[Z(\Psi^{\tau+\eta})c(q|\Psi_{\tau+\eta}^{\tau-T_o+T_a})\right],\eta\to 0 \tag{15}$$

$$=\psi_d\left[Z(\Psi^{\tau})\partial_\lambda c(q|\Psi_{\tau}^{\tau-T_o+T_a})-Z'(\Psi^{\tau})c(q|\Psi_{\tau}^{\tau-T_o+T_a})\right] \tag{16}$$

.

Having $Z(a+x)Z(b)=Z(a+b+x)=Z(a)Z(b+x)$, differentiating by x and letting $x=0$, $Z'(a)Z(b)=Z(a)Z'(b)$ is obtained. Integrate q out,

$$\begin{aligned}
&\int_{0.3}^{\infty}\mathrm{d}qH(\tau,q|\lambda)=\\
=&-\psi_d\left[Z(\Psi^{\tau})Z'(\Psi_{\tau}^{\tau-T_o+T_a})+Z'(\Psi^{\tau})(1-Z(\Psi_{\tau}^{\tau-T_o+T_a}))\right]\\
=&-\psi_d Z'(\Psi^{\tau})\\
&(Z'(\Psi^{\tau})Z(\Psi_{\tau}^{\tau-T_o+T_a})=Z(\Psi^{\tau})Z'(\Psi_{\tau}^{\tau-T_o+T_a}))\\
\approx&\psi_d e^{-\Psi_d}\text{ (ignoring threshold effect)}\\
=&(1-e^{-(\lambda+D)})\tilde{\phi}_d(\tau|\lambda)=:\phi_d(\tau|\lambda)
\end{aligned} \tag{17}$$

in which $\tilde{\phi}_d$ is from Eqn(14) of [1].

The probability model including dark charge is,

$$Pr(\tau,q|\lambda)=\begin{cases} Z(\lambda+D) & \text{, non-hit}\\ H(\tau,q|\lambda)\mathrm{d}\tau\mathrm{d}q & \text{, hit}\end{cases} \tag{18}$$

Note from Eqn(11) that,

$$\begin{aligned}
&\int_{-\infty}^{+\infty}\mathrm{d}\tau\int_{0.3}^{+\infty}\mathrm{d}qH(\tau,q|\lambda)\\
=&-\int_{-\infty}^{+\infty}\mathrm{d}\tau\frac{\partial}{\partial\eta}\left[Z(\Psi_d(\tau+\eta))(1-Z(\Psi_{\tau+\eta}^{\tau-T_o+T_a})\right]\\
=&-\int_{-\infty}^{+\infty}\mathrm{d}\tau\frac{\partial}{\partial\eta}\left[Z(\Psi_d(\tau+\eta))-Z(\Psi_d(\tau-T_o+T_a))\right]\\
=&-\int_{-\infty}^{+\infty}\mathrm{d}\tau\frac{\mathrm{d}Z(\Psi_d(\tau))}{\mathrm{d}\tau}\\
=&Z(\Psi_d(-\infty))-Z(\Psi_d(+\infty))\\
=&1-Z(\lambda+D)
\end{aligned} \tag{19}$$

Thus Eqn(18) is properly normalized. It is even more obvious from Eqn(17).

Eqn(18) can be rewritten in the likelihood form,

$$L(\tau,\lambda)=\begin{cases} Z(\lambda+D) & =: L_n(\lambda)\\ H(\tau,\lambda) & \end{cases}, \tag{20}$$

q, being a measured quantity, is ignored in the notation. $\mathrm{L_n}$ represents the likelihood of non-hit, respectively. H holds the form as Eqn(14).

5 Gradient

The gradient of Ψ_{d} calls for special attention. As $\tau - \tau_i = t_p - t_i$, it does not depend on τ if $\mathrm{t_p}$ and $\mathrm{t_i}$ are all observed constants. Thus from Eqn(10),

$$\frac{\partial \Psi_d}{\partial(\tau,\lambda)} = \left(\lambda\tilde{\psi}, \tilde{\Psi}\right). \tag{21}$$

Similarly, $\tau - \tau_i$ in quantities like $\tilde{\phi}_d, \tilde{\Psi}_d$, should be treated as constant.

From Eqn(3),

$$lnZ = -0.9652\lambda \tag{22}$$

$$(lnZ)' = -0.9652 \tag{23}$$

,

For 20 inch PMTs,

$$lnZ_{20} = -0.741\lambda \tag{24}$$

$$(lnZ_{20})' = -0.741 \tag{25}$$

, The following calculates 17 inch only. For 20 inch, replacing 0.9652 with 0.741, and $c(q|\lambda)$ with $c_{20}(q|\lambda)$ will do.

The derivative of (τ, λ) over $(t, \mathbf{r}, E)$ is,

$$\frac{\partial(\tau,\lambda)}{\partial(t,\mathbf{r},E)} = \begin{pmatrix} -1 & -\hat{\mathbf{R}}_p & 0 \\ 0 & -\frac{2\lambda}{R_p}\hat{\mathbf{R}}_p & \frac{\lambda}{E} \end{pmatrix}, \tag{26}$$

a 2×5 matrix, where $\hat{\mathbf{R}}_p := \frac{\mathbf{R}_p}{R_p}$ is the unit vector of $\mathbf{R}_p$.

Chain together,

$$\begin{aligned} \frac{\partial lnL_n}{\partial(t,\mathbf{r},E)} &= (0, (lnZ)')\frac{\partial(\tau,\lambda)}{\partial(t,\mathbf{r},E)} \\ &= 0.9652\left(0, \frac{2\lambda}{R_p}\hat{\mathbf{R}}_p, -\frac{\lambda}{E}\right) \end{aligned}. \tag{27}$$

H is a bit complicated. Its matrix form is kept, from Eqn(14)

$$\begin{aligned}
H =&\psi_d \begin{vmatrix} Z & Z' \\ C & C' \end{vmatrix}, Z := Z(\Psi^{\tau}), C := c(\Psi_{\tau}^{\tau - T_o + T_a}) \\
\frac{\partial H}{\partial \tau} =&\psi' \begin{vmatrix} Z & Z' \\ C & C' \end{vmatrix} \\
&+\psi_d \left\{ \begin{vmatrix} Z' & Z'' \\ C & C' \end{vmatrix} \psi + \begin{vmatrix} Z & Z' \\ C' & C'' \end{vmatrix} (\psi(\tau - T_o + T_a) - \psi) \right\} \\
\frac{\partial H}{\partial \lambda} =&\tilde{\psi} \begin{vmatrix} Z & Z' \\ C & C' \end{vmatrix} \\
&+\psi_d \left\{ \begin{vmatrix} Z' & Z'' \\ C & C' \end{vmatrix} \tilde{\Psi} + \begin{vmatrix} Z & Z' \\ C' & C'' \end{vmatrix} (\tilde{\Psi}(\tau - T_o + T_a) - \tilde{\Psi}) \right\} \\
&\frac{\partial lnH}{\partial(\tau, \lambda)} = \frac{1}{H} \frac{\partial H}{\partial(\tau, \lambda)} \\
&\frac{\partial lnH}{\partial(t, \mathbf{r}, E)} = \frac{\partial lnH}{\partial(\tau, \lambda)} \frac{\partial(\tau, \lambda)}{\partial(t, \mathbf{r}, E)}
\end{aligned}, \tag{28}$$

which can be calculated by referring to Eqn(26).

6 The Ignored

If dark charge and E is ignored, from Eqn(17) the likelihood function used in v2 is recovered,

$$\begin{aligned}
\frac{\partial ln\phi}{\partial(t, \mathbf{r})} &= \frac{\partial ln\phi}{\partial(\tau, \lambda)} \frac{\partial(\tau, \lambda)}{\partial(t, \mathbf{r})} \\
&= -\left(\frac{\psi' - \lambda\psi^2}{\psi}, \hat{\mathbf{R}}_p \left[\frac{\psi' - \lambda\psi^2}{\psi} + \frac{2}{R_p}(1 - \lambda\Psi) \right] \right)
\end{aligned}. \tag{29}$$

Notice that the gradient used in v2 (and Eqn(17) in [1]) is recovered if $\frac{2}{R_p}(1 - \lambda\Psi)$ is further ignored,

$$\frac{\partial ln\phi}{\partial s_i} \sim -\frac{\psi' - \lambda\psi^2}{\psi} \left(1, \hat{\mathbf{R}}_p\right) = \frac{1}{\phi} \frac{\partial\phi(\tau, \lambda)}{\partial\tau} \frac{\partial\tau}{\partial s_i}, \tag{30}$$

Let's take a closer look at $\epsilon := \frac{2}{R_p}(1 - \lambda\Psi)$.

First, $\Psi \in [0, 1]$. Taking Ψ on average ~ 0.5, and λ less than 2 for low energy events, $\epsilon > 0$. ϵ therefore contributes in Newton method as a term (according to Eqn (18) in [1]) of,

$$-B^{-1}\nabla lnL \sim B^{-1}\epsilon\hat{\mathbf{R}}_p \tag{31}$$

In a well-posed maximum likehood problem, lnL is concave and $B < 0$. Therefore $B^{-1}\epsilon\hat{\mathbf{R}}_p$ is antiparallel to $\hat{\mathbf{R}}_p$: towards the PMT and **away from** center. Ignoring ϵ creates a fitter bias **towards** the center.

Second, $\hat{\mathbf{R}}_p$ cancels out each other for opposite PMTs. The more away from center the event is, the more unbalanced the PMT hits, the more this bias manifests.

Third, ϵ is of order $\frac{1}{R_p}$ and $R_p > 250cm$ for a 6m fiducial volume. In v2, multiplying SFactor, $R_p > 200$. The bias being small, the fitter shows a reasonable accuracy with this term ignored.

7 Likelihood

From Eqn(20),

$$L = \prod_{p \in \text{hit}} H(\lambda_p, \tau_p) \prod_{p \in \text{non-hit}} L_n(\lambda_p) \tag{32}$$

$$lnL = \sum_{p \in \text{hit}} lnH(\lambda_p, \tau_p) + \sum_{p \in \text{non-hit}} lnL_n(\lambda_p) \tag{33}$$

$$\frac{\partial lnL}{\partial(t, \mathbf{r}, E)} = \sum_{p \in \text{hit}} \frac{\partial lnH(\lambda_p, \tau_p)}{\partial(t, \mathbf{r}, E)} + \sum_{p \in \text{non-hit}} \frac{\partial lnL_n(\lambda_p)}{\partial(t, \mathbf{r}, E)} \tag{34}$$

where λ_p is understood as $\frac{\lambda_0}{R_p^2}E$, as in Eqn(8). Combinded with Eqn(27,28), the gradient of lnL could be obtained.

References

[1] Mikhail Batygov and Jason A. Detwiler. *occupancy-correction-v3.pdf*. Tech. rep. University of Tennessee, 2005–2013.

[2] Thomas Michael ODonnell. "Precision Measurement of Neutrino Oscillation Parameters with KamLAND". PhD thesis. University of California, Berkeley, 2011.

[3] Nikolai R. Tolich. "Experimental Study of Terrestrial Electron Anti-Neutrinos with KamLAND". PhD thesis. Stanford University, 2005.

[4] Hiroko Watanabe. "Comprehensive Study of Anti-neutrino Signals at KamLAND". PhD thesis. Tohoku University, 2012.