Fitting Single Photo-electron peak

Qing He
Princeton University

August 30, 2010

Abstract

This note describes a method of fitting PMT's single photo electron peak and related systematic errors.

1 A model of photomultiplier response

Ref. [1] presents a method for measuring the photomultiplier gain using a realistic photomultiplier response function. In this note, I present my understanding of the model and show how to use the model to fit the SPE (single photo-electron) peak for Daya Bay analysis.

The photomultiplier is treated as an instrument consisting of two independent parts:

- the photo-detector where photons are converted into electrons
- the amplifier (dynode system) which amplifies the initial charge

1.1 Photoconversion and electron collection

Suppose the light source is steady, so the average number of photons hitting the PMT is a constant. In reality, the number of photons hitting PMT is a Poisson distributed variable. Not all photons are converted into electrons, only a fraction of them (so called quantum efficiency) are collected by PMT. This process is a random binary process, the convolution of Poisson and binary processes still gives a Poisson distribution:

$$P(n;\mu) = \frac{\mu^n e^{-\mu}}{n!}, \tag{1}$$

where μ is the mean number of photo-electrons collected by the first dynode, $P(n;\mu)$ is the probability that n photo-electrons will be observed when their mean is μ.

1.2 Amplification

The response of a multiplicative dynode system to a single photo-electron in principle is a Poisson distribution, but if the amplification of the first dynode is large (>4), the response can be approximated by a Gaussian distribution:

$$G_1(x) = \frac{1}{\sigma_1\sqrt{2\pi}}\exp(-\frac{(x-Q_1)^2}{2\sigma_1^2}), \tag{2}$$

where x is the variable charge, Q_1 is the average charge at the PMT output when one electron is collected by the first dynode, σ_1 is the corresponding standard deviation of the charge distribution.

Some MC events are generated to test if the output of PMT can be approximated by a Gaussian distribution. Start from a single photo-electron at the first dynode, 8 dynodes are simulated. The amplification factor is set to 10 for the first dynode, and 7 for rest. So the gain is $10 \times 7^7 = 8.2 \times 10^6$, which is close to our PMT performance. The amplification of each dynode is simulated using Poisson distribution with μ equal to the amplification factor. The ratio of resulted number of photo-electrons to gain is shown in Fig. 1. The poor χ^2 fit

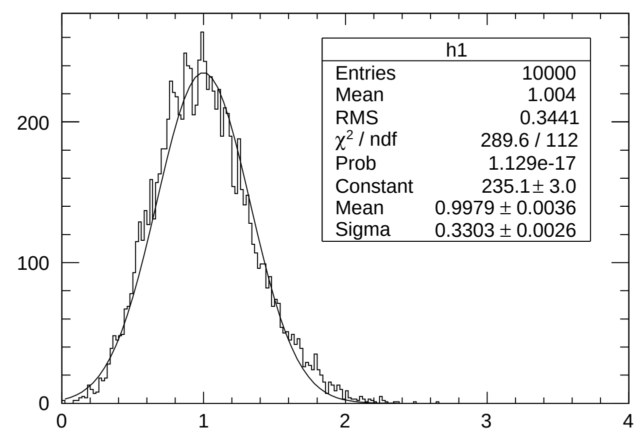

Figure 1: NPE/gain distribution.

shows that the distribution is not Gaussian. However, the mean, NPE/gain, happens to be close to 1 with small error. It can be clearly seen that the Gaussian peak is biased to the right hand side because of the high side tail. We should keep in mind that the peak position of a Poisson distribution is not same as its mean value, especially for small μ. So if we only fit the peak within a small range, it is biased to the lower side. A Gaussian fit to Fig. 1 within range 0.5∼1.5 shows that there is 4% change in the resulted mean. This becomes important in a precise calibration.

In the case when more than one photo-electron are collected by the first dynode, the charge distribution is a convolution of n single electron cases:

$$G_n(x) = \frac{1}{\sigma_1\sqrt{2n\pi}}\exp(-\frac{(x-nQ_1)^2}{2n\sigma_1^2}). \tag{3}$$

The response of an ideal noiseless PMT is simply a convolution of Eq. 1 and Eq. 3:

$$S_{ideal}(x) = P(n;\mu) \otimes G_n(x) \tag{4}$$

$$= \sum_{n=0}^{\infty} \frac{\mu^n e^{-\mu}}{n!} \frac{1}{\sigma_1\sqrt{2n\pi}}\exp(-\frac{(x-nQ_1)^2}{2n\sigma_1^2}). \tag{5}$$

1.3 Background processes

The background processes are split into two groups in Ref. [1]:

(I) the low charge processes present in each event which are responsible for nonzero width of the signal distribution when no photo-electron was emitted from the photocathode (the pedestal);

(II) the discrete processes which can, with some probability, accompany the measured signal (such as thermoemission, noise initiated by the measured light, etc.).

The process of type I can be described by a Gaussian and those of the type II can be described by an exponential function. The background processes are parametrized as

$$B(x) = \frac{1-w}{\sigma_0\sqrt{2\pi}}\exp(-\frac{x^2}{2\sigma_0^2}) + w\theta(x)\alpha\exp(-\alpha x), \tag{6}$$

where σ_0 is the standard deviation of the type I background distribution, w is the probability that type II background present, α is the coefficient of the exponential decrease of type II background,

$$\theta(x) = \begin{cases} 0 & x < 0 \\ 1 & x \geq 0 \end{cases} \tag{7}$$

is the step function. Note that type I background is always present (100%) and the second term corresponds to the presence of both groups of background. For small σ_0 the convolution of a Gaussian with an exponential function is reduced to a pure exponential function.

1.4 The realistic response function of the PMT

The realistic PMT response is the convolution of Eq. 5 and Eq. 6:

$$\begin{aligned} S_{real}(x) &= \int S_{ideal}(x')B(x-x')dx' & (8) \\ &= \sum_{n=0}^{\infty}\frac{\mu^n e^{-\mu}}{n!}[(1-w)G_n(x-Q_0) + wI_{G_n\otimes E}(x-Q_0)], & (9) \end{aligned}$$

$$\begin{aligned} I_{G_n\otimes E}(x-Q_0) &= \int_{Q_0}^{x} G_n(x'-Q_0)\alpha\exp[-\alpha(x-x')]dx' & (10) \\ &= \frac{\alpha}{2}\exp[-\alpha(x-Q_n-\frac{\alpha\sigma_n^2}{2})]\times \\ & \left[\mathrm{erf}(\frac{|Q_0-Q_n-\sigma_n^2\alpha|}{\sigma_n\sqrt{2}}) + \mathrm{sign}(x-Q_n-\sigma_n^2\alpha)\mathrm{erf}(\frac{|x-Q_n-\sigma_n^2\alpha|}{\sigma_n\sqrt{2}})\right] & (11) \end{aligned}$$

$$Q_n = Q_0 + nQ_1, \tag{12}$$

$$\sigma_n = \sqrt{\sigma_0^2 + n\sigma_1^2}, \tag{13}$$

where Q_0 is the pedestal and $\mathrm{erf}(x)$ is the error function. Note that Ref [1] miss a "$\frac{1}{2}$" in $\exp[-\alpha(x-Q_n-\frac{\alpha\sigma_n^2}{2})]$.

2 Fit to Dry Run data

For Dry Run data, we have threshold for each PMT, which kills the $n = 0$ component in Eq. 9. Since the trigger is a random binary process, it will have no effect to the shape.

There are 7 free parameters in parametrization Eq. 9, μ, Q_0, σ_0, Q_1, σ_1, w, and α. μ is related to the intensity of the light source. Q_0, σ_0 characterize the pedestal distribution. Q_1, σ_1 characterize the SPE distribution. w, and α are parameters for background distribution. All parameters are suppose to be greater than 0, w is suppose to be $0 < w < 1$. What we after is the PMT gain, Q_1.

In the real fit, there is another overall normalization parameter.

2.1 What is Q_0 and σ_0?

The PreAdc distribution for channel 1 (Ring=1, Column=1) is shown in Fig. 2. We choose $Q_0 \approx 90.7$, $\sigma_0 \approx 1.9$ in the gain fit. The gain fit result Q_1 is not very sensitive to σ_0.

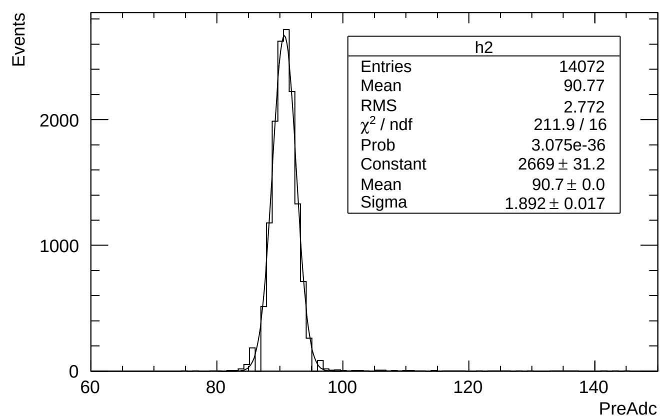

Figure 2: Pedestal distribution for Channel 1.

2.2 Dark noise fit

The dark noise is picked out by requiring Tdc value greater than 1050. The fit range is (0∼100) and the fit result is shown in Fig. 3, the n=1 and n=2 components of Eq. 9 are shown in dashed lines. The fitted gain is 19.0±0.1/PE. The mean number of photo-electrons collected by the first dynode is 0.11. χ^2 shows that the model fit the data very well.

2.3 Dry run LED fit

The model assumes the intensity of light source is stable. However, the bottom and top reflectors contribute some events which are equivalent from low intensity sources. So in principal, LED events are not suitable for gain fit unless we can separate direct light and reflect light. The effect of low intensity source is simulated in Monte Carlo samples, which are generated according the

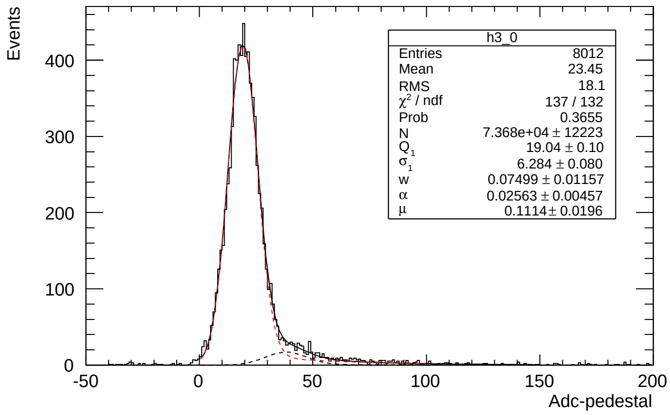

Figure 3: Fit to dark noise of Channel 1. The n=1 and n=2 components of Eq. 9 are shown in dashed lines.

above model. Results from stable source is shown in the left plot in Fig. 4. The red line shows the fit result using the same model. In the right plot in Fig. 4, 36% events are generated from low intensity source. Those events distort the distribution.

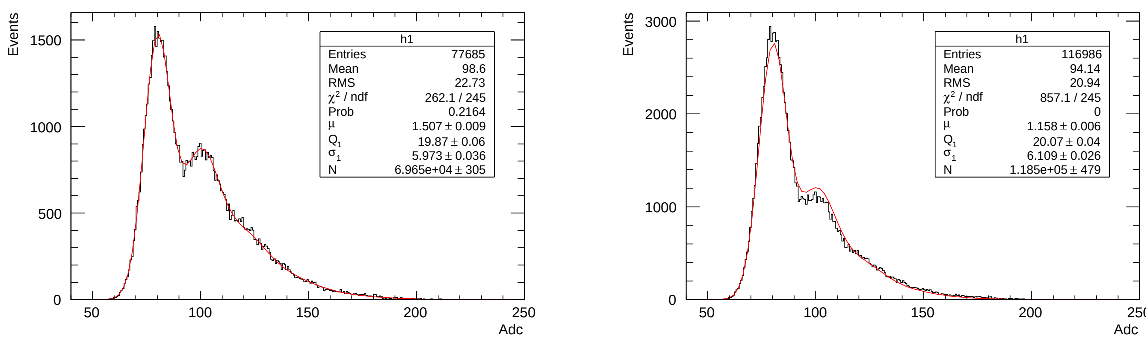

Figure 4: Monte Carlo simulation for the PMT response. Stable source (μ=1.5) is assumed in the left figure. In the right figure, 36% events are simulated from low intensity source (μ=0.5).

Fig. 5 shows the fit to the LED signal spectrum of channel 1 (Run 2317). The n in Eq. 9 runs from 1 to 6. The shape is well reproduced in the fit. However, the χ^2 is very large due to two reasons: 1) Gaussian approximation of Poisson distribution. 2) Reflection light.

2.4 Be aware of w and α

The exponential background is used to parameterize the shape more correctly, however, it also introduces more flexibilities, sometimes even results in not correct answers.

A Monte Carlo sample is generated to study the effect of w and α. In Fig. 6, the Monte Carlo sample is generated without the exponential background. A fit without the exponential background is shown in the left plot, while a fit with the exponential background is shown

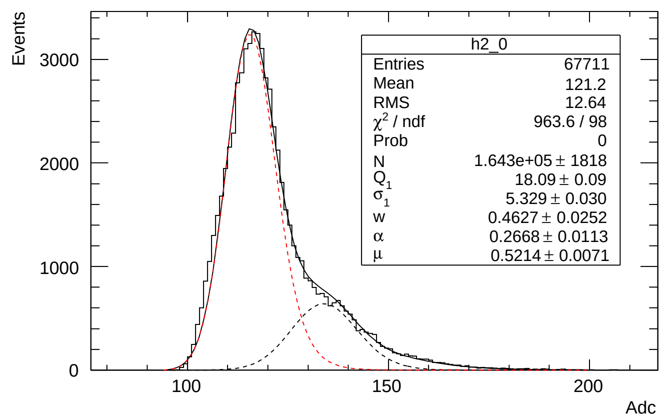

Figure 5: Fit to signal of Channel 1. The n=1 and n=2 components of Eq. 9 are shown in dashed lines.

in the right plot. The input gain is 20, the bad χ^2 in the left plot is due to using Gaussian function to fit the Poisson distribution. As shown in the right plot, if we include the exponential background to fit the sample which actually doesn't have such a background, the fit has better χ^2, but the gain is 5% smaller. This is because the exponential component make the high end tail agree better while push the peak to the left.

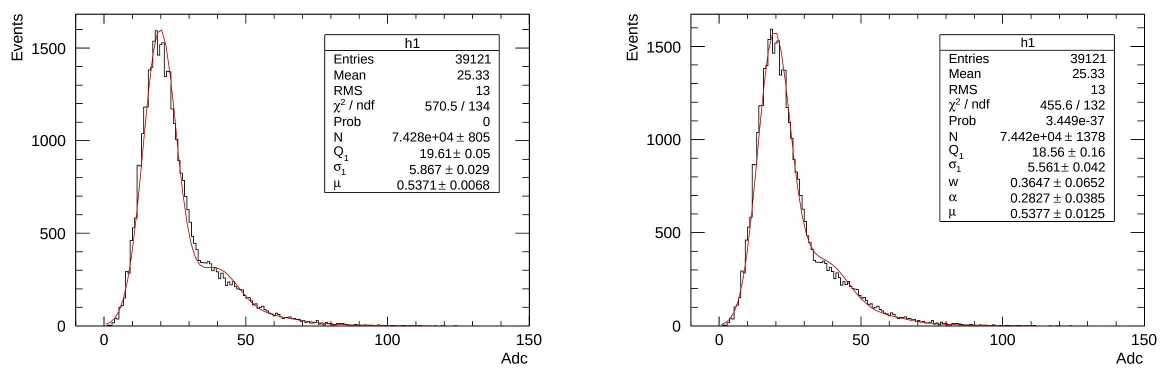

Figure 6: Fit to Monte Carlo simultion of the PMT response.

It is better if we can extract w and α beforehand, e.g. from pre-Adc fit, rather than let them float in the gain fit.

2.5 Compare dark noise SPE with low intensity LED run

Run 2314 is a low intensity LED run with external trigger. Most of hits are from single photo-electrons. We can compare the low intensity LED SPE peak with that of dark noise and see if there is any significant difference. To incerease the statistics of dark noise, we select two categories of events: 1) Hits with Tdc>1050, these are dark noise events which happen before the trigger. 2) First hits of each channel with Tdc<800, we choose first hits to avoid after pulses.

One example of LED fit and dark noise fit is shown in Fig. 7. The gain differences of all channels (excluding bad channels) are shown in Fig. 8. The average gain from dark noise fit is a little bit higher (<1%) than that from LED fit.

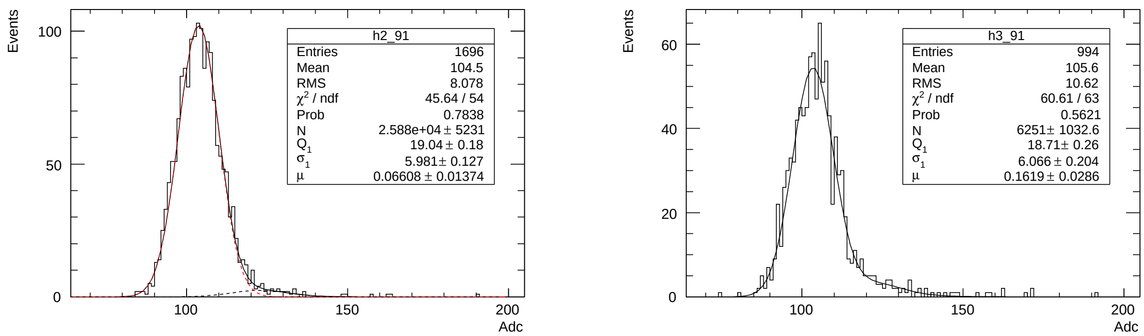

Figure 7: Left: Fit to LED signal of channel 92. Right: Fit to Dark noise of channel 92.

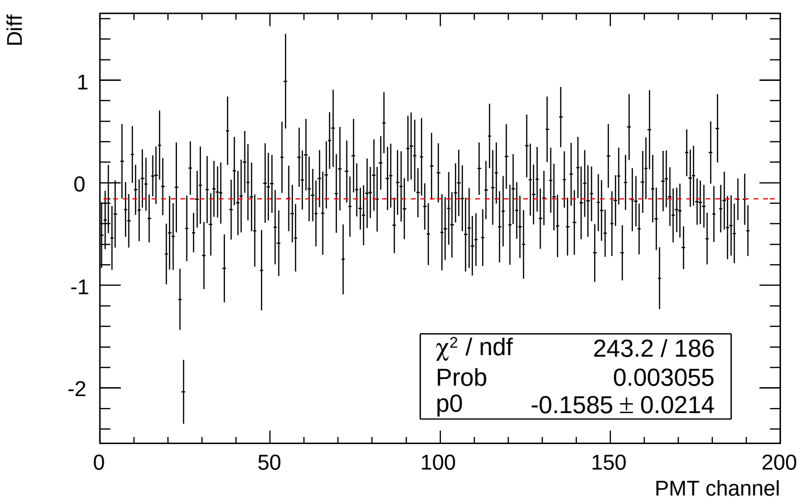

Figure 8: The differences of the gains from low intensity LED fit and dark noise fit.

2.6 Systematic error

- Systematics due to Q_0

  The gain, Q_1, is strongly correlated with Q_0 in the dark noise SPE fit. $Q_1 + Q_0 \approx Const.$ in the dark noise fit since the $n \geq 2$ components are small.

  Basically, a precise calibration of the pedestal is also crucial for a precise gain calibration using this method.

- Systematics due to σ_0

  σ_1 is strongly correlated with σ_0 in the dark noise SPE fit. However, Q_1 is not sensitive to σ_0.

- Gaussian approximation of the amplification process

  The peak position of Poisson distribution is about 4% lower than the mean, assume the amplification factor for the first dynode is 10. Even though we can make a correction for this difference, it will contribute some systematic uncertainty in the correction. The approximation also affect χ^2 for large statistics sample.

- Fit range

  Fits with different ranges, (0∼80), (0∼100), (0∼150) are performed to check the systematic error. Only 0.5% change in Q_1 is seen.

- The model itself

  Comparing with the results from Crystal Ball function, which is shown below, the difference is about 1.2%.

3 Fit with Crystal Ball function

The SPE spectrum is also fitted with a Crystal Ball function to extract the SPE peak value. The Crystal Ball function is defined as,

$$CB(x;\mu,\sigma,n,\alpha) = \begin{cases} \exp(-\frac{(x-\mu)^2}{2\sigma^2}) & \text{if } \frac{x-\mu}{\sigma} > -\alpha \\ A\cdot(B-\frac{x-\mu}{\sigma})^{-n} & \text{if } \frac{x-\mu}{\sigma} \leq -\alpha \end{cases} \tag{14}$$

$$A = \left(\frac{n}{|\alpha|}\right)^n \exp\left(-\frac{\alpha^2}{2}\right), \tag{15}$$

$$B = \frac{n}{|\alpha|} - |\alpha|. \tag{16}$$

The fit result is shown in Fig. 9.

4 Fit with Gaussian

We also fit the SPE spectrum with a Gaussian with different fit range, r. The fit range is from $mean-r$ to $mean+r$, and the results are shown in Fig. 10. The results from above model and Crystal Ball function are also shown in Fig. 10 for comparison.

5 Conclusion

Ref. [1] provides a good model to extract PMT gain. Fit to Dry Run data and its associated systematic error has been studied in this document.

Some important points for using this model to fit Dry run data:

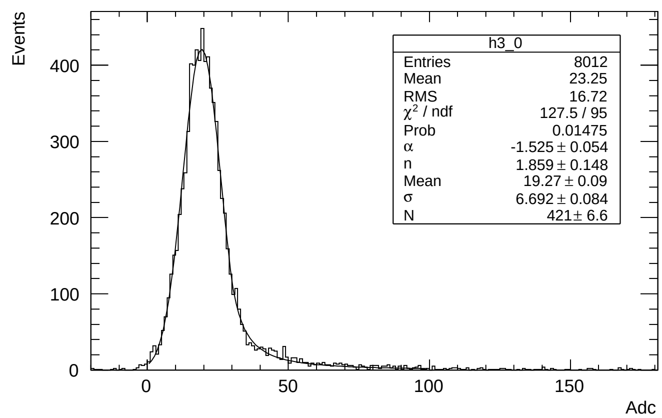

Figure 9: Fit to Dark noise of Channel 1.

- The model assumes stable light intensity, Dry run data usually doesn't satisfy this requirement due to reflections.
- The model use a Gaussian approximation for a Poisson distribution, the gain from Gaussian approximation is about 4% lower than the real gain.
- Beaware of the exponential background, sometimes it has a big impact on the fit result.
- Dark noise is good for gain calibration.

Crystal Ball and Gaussian functions are also tested. The fit using a Gaussian function within a limited range is not a good method. The results depend on the defined range and usually give high gain due to the high end tail. Crystal Ball function can produce a better result comparing with Gaussian. However, it is purely base on the shape of the spectrum rather than from the underlying physics.

6 Acknowledgement

Thank Weili and Liangjian for helpful discussions.

References

[1] Nuclear Instruments & Methods in Physics Research A 339(1994) 468-476

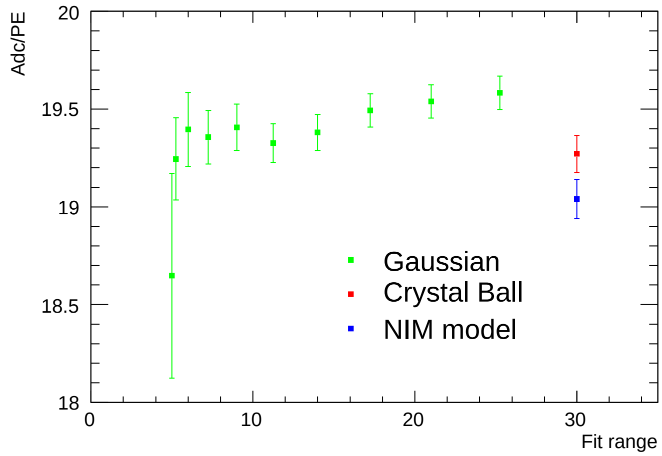

Figure 10: Fit the SPE spectrum with a Gaussian with different fit range. The results from above model and Crystal Ball function are also shown for comparison.