

Calibration of photomultiplier tubes

L.N. Kalousis

*Physics Department, National Technical University,
157 80 Zografou, Athens, Greece*

E-mail: leonidas.kalousis@gmail.com

ABSTRACT: The purpose of the present article is to demonstrate the calibration of photomultipliers with a gaussian single photoelectron response using a numerical method based on the Discrete Fourier Transform (DFT). Conventional techniques, commonly employed in the literature, use approximate models or brute force numerical calculations of the convolution integrals that lead to the charge response function of the photomultiplier, $S_R(x)$. In this publication, we explain how a truncated gaussian model for the single photoelectron amplification can lead to rigorous results if one leans on the DFT approach. The distinct feature of this procedure is that $S_R(x)$ is calculated to all orders in the Poisson mean μ that characterizes the light intensity and no approximations are needed. This scheme was applied to the calibration of the Hamamatsu R7081 photomultiplier tube and a comparison of the DFT approach with the more standard numerical integration method is also presented. Last, toy Monte Carlo data were analyzed for different values of μ to understand the precision of the DFT method.

KEYWORDS: Analysis and statistical methods; Data analysis; Detector alignment and calibration methods (lasers, sources, particle-beams); Photon detectors for UV, visible and IR photons (vacuum) (photomultipliers, HPDs, others)

ARXIV EPRINT: [2304.07292](https://arxiv.org/abs/2304.07292)



Contents

1	Introduction	1
2	Standard theory	2
3	Data analysis	3
3.1	Single photoelectron response model	3
3.2	Experimental data and results	4
3.3	Comparison with the numerical integration method	7
4	Simulated data	8
5	Outlook	9
A	Calculation of Q_R and σ_R	9
B	Variance of $S(x)$	10
C	Mean value and variance of $g(x)$	11

1 Introduction

Most photomultipliers have a single photoelectron response function that can be parameterized with a gaussian distribution. A prime example of this case is the ten inches Hamamatsu R7081 photomultiplier tube (PMT) model that has been used in several experiments in particle physics. Double Chooz, RENO and IceCube to name only but a few [1–3]. It is known to operate at a nominal gain of 10^7 units of electron charge for about 1500 V of high voltage. It has a photocathode sensitivity that ranges between 300 and 600 nm with a maximum peak in quantum efficiency of roughly 25 % at about 400 nm. The good photoelectron resolution, the small dark noise, the low glass radioactivity levels and relatively low cost make R7081 an attractive choice for the instrumentation of various detectors.

Several methods have been employed in the literature for the calibration of PMTs with a gaussian single photoelectron (SPE) response. Perhaps, the most rigorous has been presented in a seminal paper written by Dossi et al., ref. [4]. In the aforementioned publication, the SPE response was parametrized by a combination of an exponential distribution and a truncated gaussian. In our view, the key elements of the Dossi paper can be summarized in the following points:

- i. the realization that the exponential term that models photoelectrons (PEs) that miss the first amplification stage is part of the signal (it adds to the gain) and,
- ii. the use of a properly normalized, truncated gaussian to describe the full amplification chain, avoiding thus the prediction of negative charges in the underlying formulae.

None of these points is entirely trivial. Unfortunately the final equation for the PMT charge response, $S_R(x)$, is rather complicated to be worked out analytically and one has to resort to some sort of an approximation. For example, the DarkSide collaboration solved the convolution integrals involved in the calculation of $S_R(x)$ numerically for the first two peaks, and the higher peaks were approximated by perfect gaussians [5]. That was adequate for the needs of DarkSide.

In this work, we seek to apply the Dossi model to the calibration of the R7081 PMT using the numerical method first presented in ref. [6]. The distinct feature of this procedure is that $S_R(x)$ is calculated numerically to all orders in the Poissonian mean (μ) and no approximations are needed. In section 2 we outline briefly the basic theory of gain determination and we present the Discrete Fourier Transform (DFT) technique exploited throughout the publication. In section 3 we analyze data sets of R7081 PMT showcasing the validity of our study. A comparison with the common numerical integration method is included. In section 4 we present a Monte Carlo study showing the advantages of the DFT approach. We close this publication with some general remarks concerning gain calibration.

2 Standard theory

Whenever a fixed number of photons is shot towards the photocathode of a PMT there is a certain probability that some will convert and create electrons (quantum efficiency). These PEs are then collected by the anodes and directed to the amplification chain with some certain probability (collection efficiency). The number of PEs (n) registered by the PMT is given by the well-known Poisson formula:

$$P(n; \mu) = e^{-\mu} \frac{\mu^n}{n!}. \quad (2.1)$$

The Poisson mean, μ , characterizes the light source and the quantum and collection efficiencies jointly. Now, if $S(x)$ is the probability density function (PDF) for a single PE to create a charge in the vicinity of x , the probability for n PEs to produce x is dictated by $S_n(x)$, where $S_n(x)$ is the n -times convolution of $S(x)$. The charge response of a PMT can be readily worked out:

$$\begin{aligned} S_R(x) &= \sum_{n=0}^{+\infty} P(n; \mu) S_R^{(n)}(x) \\ &= \sum_{n=0}^{+\infty} P(n; \mu) (S_n * B)(x). \end{aligned} \quad (2.2)$$

Note that a final convolution with the pedestal PDF, $B(x)$, is needed to include white noise from the electronics and other sources. Of course, $S_1(x) = S(x)$ and $S_0(x) = \delta(x)$. More details on the theory of PMT calibration can be found in ref. [6]. The mean and standard deviation of $S_R(x)$, that is, Q_R and σ_R are equal to:

$$Q_R = Q_0 + \mu Q_s \quad (2.3)$$

$$\sigma_R^2 = \sigma_0^2 + \mu(\sigma_s^2 + Q_s^2), \quad (2.4)$$

where Q_0, σ_0 are the mean and standard deviation of $B(x)$ respectively and Q_s, σ_s those of $S(x)$. The calculation of these formulae can be found in the appendix A.

The essence of the DFT approach to gain determination lies in the realization that the DFT of $S_R(x)$, $\tilde{S}_R(k)$, has a very simple formula [6]:

$$\tilde{S}_R(k) = \tilde{B}(k)e^{\mu(\tilde{S}(k)-1)}, \quad (2.5)$$

where $\tilde{B}(k)$, $\tilde{S}(k)$ are the DFTs of $B(x)$ and $S(x)$ respectively. The fact that the series of eq. (2.2) can be summed formally in the Fourier inverse space, is due to the form of the Poisson factors and the simple mathematical theorem that the DFT of $S_n(x)$ is n powers of $\tilde{S}(k)$, $\tilde{S}^n(k)$. If one could invert eq. (2.5) analytically then one has a formula of $S_R(x)$ in a closed form. In practice this is an impossibility. In order to progress, in ref. [6] we proposed to perform the forward DFT and inverse DFT calculations numerically using the fftw package [7]. For this purpose a C++/ROOT based software [8] was developed that calculates $S_R(x)$ for a given number of steps. The code has been committed in a public github repository with several examples that can assist the interested reader [9]. We have analyzed all data in this article using this software.

3 Data analysis

3.1 Single photoelectron response model

The SPE response model of any given PMT can be parameterized by the general formula:

$$S(x) = w\alpha e^{-\alpha x}H(x) + (1-w)g(x). \quad (3.1)$$

A few remarks are necessary here. First, the prefactor w parameterizes, through the exponential distribution, the probability that a single PE will miss the first amplification stage. It ranges from zero to one. For $w = 0$, all PEs are amplified according to the full chain. A better discussion of this can be found in ref. [4]. $H(x)$ is the Heaviside step function. The calculation of Q_s and σ_s is not at all trivial but not that difficult to perform. It is done in the appendix B and the result is:

$$Q_s = \frac{w}{\alpha} + (1-w)Q_g \quad (3.2)$$

$$\sigma_s^2 = \frac{w}{\alpha^2} + (1-w)\sigma_g^2 + w(1-w)\left(Q_g - \frac{1}{\alpha}\right)^2, \quad (3.3)$$

where Q_g and σ_g are the mean value and standard deviation of $g(x)$. Note that when $w = 0$, we have $Q_s = Q_g$ and $\sigma_s = \sigma_g$. On the other hand, for $w = 1$ the mean value and standard deviation of $S(x)$ are those of the exponential term.

The $g(x)$ PDF model adopted by Dossi et al. is [4]:

$$g(x) = \frac{1}{g_N} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-Q)^2}{2\sigma^2}} H(x). \quad (3.4)$$

This characterizes the amplification of the full dynode system. It is a gaussian truncated at the negative values of x . The factor g_N ensures that $g(x)$ is properly normalized and equals to:

$$g_N = \frac{1}{2} \operatorname{erfc}\left(-\frac{Q}{\sqrt{2}\sigma}\right). \quad (3.5)$$

$\text{erfc}(x)$ is the complementary error function [10]. The mean value and standard deviation of $g(x)$ are deduced in the appendix C. They are equal to:

$$Q_g = Q + \kappa \quad (3.6)$$

$$\sigma_g^2 = \sigma^2 - (Q + \kappa)\kappa, \quad (3.7)$$

where κ is given by the equation:

$$\kappa = \frac{1}{g_N} \frac{1}{\sqrt{2\pi}} \sigma e^{-\frac{Q^2}{2\sigma^2}}. \quad (3.8)$$

One can see that as Q increases, κ approaches to zero due to the damping exponential factor and Q_g, σ_g approach Q and σ respectively.

3.2 Experimental data and results

A PMT model based on eq. (3.1) and (3.4) is quite difficult to derive in a closed form. Even without the final convolution with $B(x)$, it is rather cumbersome to calculate $S_n(x)$ even for the lowest value of $n = 2$. For example, for $n = 2$ one has three terms to compute while for $n = 3$ one has four different terms in $S_n(x)$! One has to resort to some sort of an approximation. For instance, in the original Dossi et al. paper a simple formula was given for $S_R(x)$ where all $S_n(x)$ distributions above $n = 1$ were modeled by symmetric gaussians. Additionally, in section 5 of that paper a brute force numerical method was presented. It is the purpose of this article to solve the Dossi model using the DFT procedure explained in section 2.

Several data sets were taken with a R7081 PMT inside a light-tight box. The details of the experimental setup can be found in ref. [6]. In general, the PMT was illuminated by an optical fiber connected to a light-emitting diode (LED). The light pulses were produced at the LED by a fast pulse generator. The charge was readout by a LeCroy oscilloscope (WavePro 725Zi) triggering at the generator's second, duplicated channel. Like this, all pulses were recorded, including those with no PEs produced at the photocathode (pedestal). Figure 1 shows a few examples of the charge distributions obtained from our setup. To analyze the data and obtain the gain from the data points, we fitted the spectra with the $S_R(x)$ calculated using the DFT method. The procedure is best described in ref. [6]. First, a gaussian fit was performed on the pedestal alone to obtain approximate values for Q_0, σ_0 and μ . The model was minimized using the Minuit2 software [11].

We should point out that for all the measurements included in this publication the optical fiber was placed at the center of the photocathode, using a plastic halo-like structure, and that the fiber was always pointing vertically with respect to the surface of the PMT.¹ In this way one has to deal a single quantum efficiency (that at the center of the PMT) and we expect the assumption of eq. (2.1) to be quite valid. Note that the quantum efficiency of the PMT is expected to vary across the surface of the photocathode and a light source illuminating the whole surface of the PMT will not be accurately described by eq. (2.1). In such cases one has to take into account the variance of the quantum efficiency across the incident angle and the Poissonian factors have to be modified. Note also that any possible bias in the extraction of the gain has to be attributed to the validity of the SPE response model (which was taken as an assumption) and/or the stability of our setup.

¹More details on the experimental setup, including pictures of the halo-like structure, can be found in ref. [12].

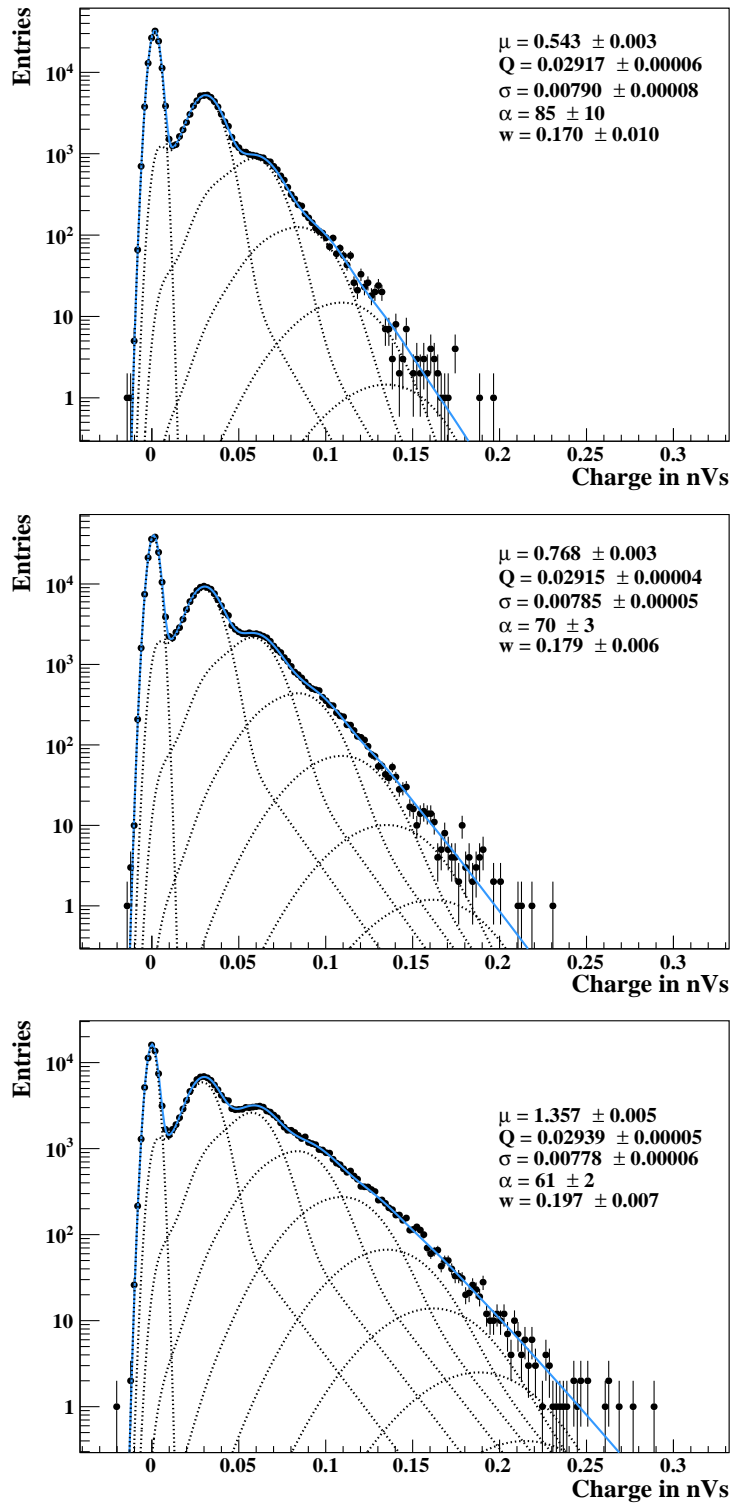


Figure 1. A few fits obtained using the Dossi et al. model through the DFT method. The data are shown in the black dots and the best fit curve is shown in azure line. The dotted lines show the contributions of the various PE peaks.

Table 1. Summary of the R7081 PMT calibration results.

μ	w	α	Q	σ	χ^2/NDOF
0.543 ± 0.003	0.170 ± 0.010	85 ± 10	0.02917 ± 0.00006	0.00790 ± 0.00008	1.52
0.632 ± 0.003	0.180 ± 0.008	71 ± 4	0.02920 ± 0.00006	0.00782 ± 0.00007	1.18
0.768 ± 0.003	0.179 ± 0.006	70 ± 3	0.02915 ± 0.00004	0.00785 ± 0.00005	1.60
0.979 ± 0.003	0.196 ± 0.006	63 ± 2	0.02923 ± 0.00005	0.00774 ± 0.00005	1.86
1.357 ± 0.005	0.197 ± 0.007	61 ± 2	0.02939 ± 0.00005	0.00778 ± 0.00006	1.68
2.014 ± 0.008	0.196 ± 0.007	61 ± 2	0.02935 ± 0.00006	0.00777 ± 0.00007	1.55

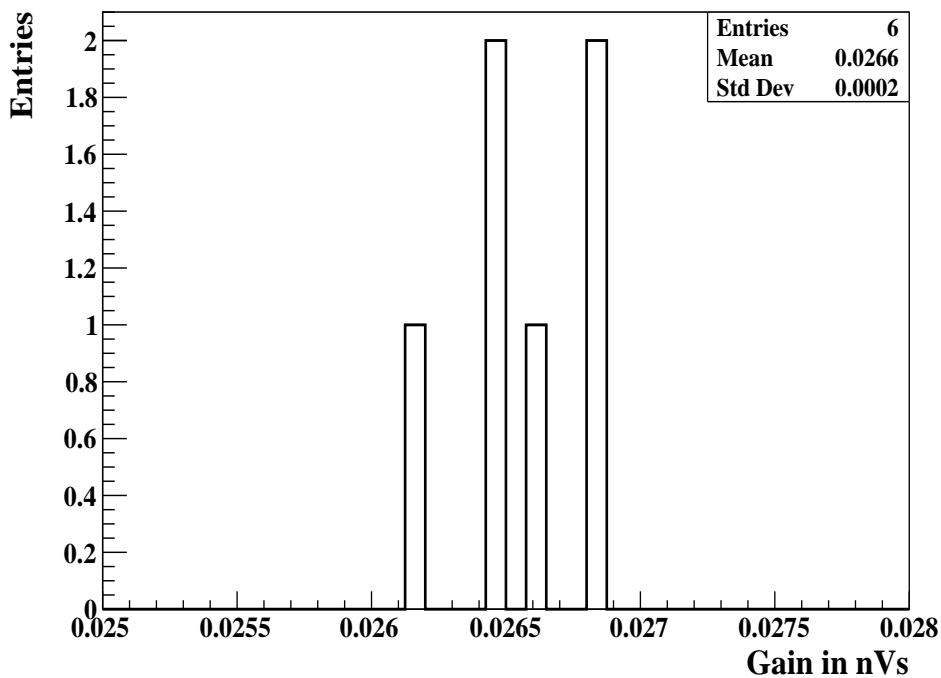
**Figure 2.** Distribution of the gain (Q_s) for the measurements compiled in table 1.

Figure 1 shows the best fit curves in azure line. One can readily see that the model follows closely the data for all examples (χ^2/NDOF was always close to one). Furthermore, more data were taken with increasing light intensity to assess the stability of gain determination. The results are gathered in table 1. Even though strong correlations exist between the various parameters, and even though there's a drift in the extraction of w and α in particular, the gain as calculated by the eq. (3.2) and (3.6) is quite stable inside the $\mu \sim 0.5\text{--}2.0$ PE plateau. Figure 2 shows the gain distribution for the results of table 1. One readily sees that Q_s is stable within a $\sim 1\%$ range. Two further remarks should be made. First, one should never identify gain with Q_g as this will overestimate the true gain. On the contrary, the gain should be calculated by the weighted average of eq. (3.2). Second, in the particular case of the R7081 PMT the truncation of the gaussian PDF at negative charge values has little impact on the final results; Q_g is very close to Q .

Table 2. Comparison of the DFT approach and the numerical integration method.

DFT approach		Numerical integration	
Gain, Q_s	χ^2/NDOF	Gain, Q_s	χ^2/NDOF
0.0262	1.52	0.0262	2.33
0.0265	1.18	0.0263	2.03
0.0265	1.60	0.0264	2.97
0.0266	1.86	0.0264	3.31
0.0268	1.68	0.0266	2.85
0.0268	1.55	0.0267	2.96
1 sec		11 sec	

3.3 Comparison with the numerical integration method

To assess the advantages of the DFT approach towards gain determination, we analyzed the same R7081 data set using the numerical method presented in ref. [4]. Note that according to this procedure all the integrals of the convolutions in eq. (2.2) are calculated numerically. In practice, this was implemented in our analysis software using multiple for loops. Unfortunately, the execution time to compute several $S_n(x)$ distributions becomes increasingly large, and thus, is rendering this method almost inapplicable. To evade this difficulty we calculated the first PE peak² numerically and the higher PE peaks were approximated with symmetric gaussians.

Table 2 shows the results of this study. For comparison purposes the figures from the DFT analysis are also included. First columns show the gain (Q_s) obtained from the minimization procedure and second columns show the χ^2/NDOF . The last row shows the execution time that each method took for the analysis of the dataset presented in table 1. From these numbers the following conclusions can be drawn:

- i. The first thing that can be observed from table 2 is that the two methods give almost identical figures insofar as gain determination is concerned. Inside the $\mu \sim 0.5\text{--}2.0$ plateau both techniques return the same gain.
- ii. Second, the DFT approach gives a better fit, since the χ^2/NDOF is systematically smaller than that of the numerical method.
- iii. Last, the DFT method is almost ten times faster than the numerical integration.

The data that we took and analyzed in this article lie well inside the $\mu \sim 0.5\text{--}2.0$ window. Part of this decision was motivated by the fact that a pronounced pedestal always assist the fitter and enhances gain determination through strong constraints in Q_0 , σ_0 and μ . To understand how our method performs for large μ , in the next section we analyzed sets of simulated data.

²That is $(S * B)(x)$.

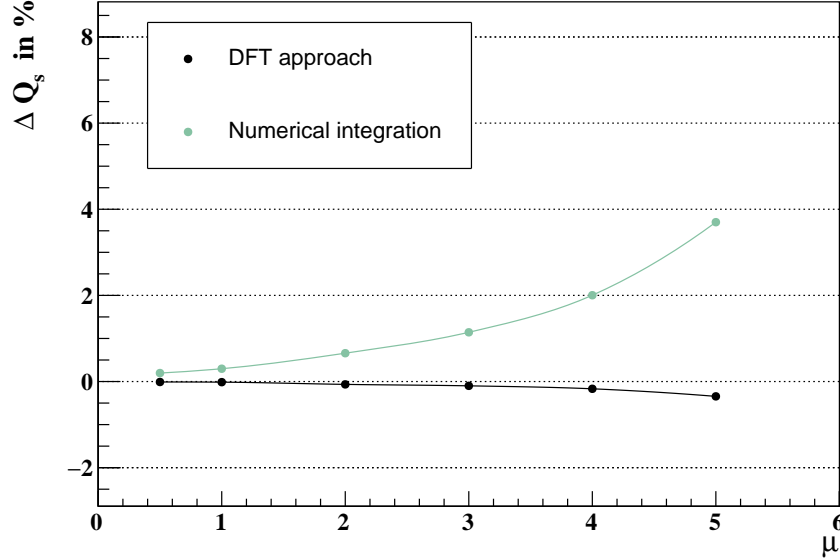


Figure 3. Comparison of the DFT (black) and numerical integration (green) methods.

4 Simulated data

Series of several toy Monte Carlo SPE spectra were generated. The algorithm that was used to produce these data sets can be summarized in the following steps:

- i. First, a number of PEs was thrown from a Poisson distribution of mean value μ .
- ii. For each PE, a charge was picked randomly from the SPE distribution of eq. (3.1) and (3.4). The total charge was calculated by summing the individual charges that each PE deposits.
- iii. A charge was thrown from the gaussian distribution of the pedestal and summed to the total charge obtained in step two.
- iv. The final, total charge was filled in a histogram.
- v. This procedure was repeated $2.5 \cdot 10^6$ times and a histogram of $2.5 \cdot 10^6$ entries was produced.

The parameters of the pedestal and SPE distributions were set to match those of the first line in table 1. For each value of μ one hundred toy spectra were generated and fitted with both the DFT and the numerical integration methods. Data were generated for μ inside the 0.5–5.0 range. For each set of one hundred toys, the distribution of the relative deviation from the true gain (Q_s):

$$\Delta Q_s = \frac{Q'_s - Q_s}{Q_s}, \quad (4.1)$$

was plotted. We note that Q'_s is the gain parameter returned by the minimizer.

Figure 3 shows the results of this exercise. The x axis shows μ and the y axis shows the mean value of the ΔQ_s distribution for the one hundred toys. The black dots depict the results from the DFT approach and the green dots those of the numerical integration respectively. One sees that DFT provides excellent results inside the $\mu = 0.5$ – 5.0 range and a deviation of $\sim 0.3\%$ is observed only at the high value of $\mu = 5$. In contrast, the numerical integration method starts to deviate significantly from the true gain and for $\mu = 5$ the deviation reaches the value of $\sim 4\%$.

5 Outlook

In this article, we presented the calibration of PMTs with gaussian SPE charge response. The analysis was based on a truncated gaussian PDF for the SPE response and we relied on the DFT method to solve the $S_R(x)$ model numerically. We have showed that despite the complications involved in the multidimensional fit of the data, one can extract the gain with good precision. In particular, for the R7081 PMT the gain can be determined with $\sim 1\%$ accuracy or better within the $\mu \sim 0.5\text{--}2.0$ plateau. Attention was paid to emphasize the fact that the gain should not be confused with the mean value of the SPE response $g(x)$ and, instead, a weighted average between the exponential and gaussian terms should be preferred. The same procedure can be applied to other PMTs which share the same characteristics in charge response with the R7081 PMT. The analysis software used in these studies exists in a public github repository and can be used by other investigators *mutatis mutandis*.

A comparison of the DFT and numerical integration methods was also attempted. In particular we showed, using R7081 data, that within the 0.5–2.0 PE range both techniques provide consistent results, but DFT gives better χ^2/NDOF and runs much faster. In this respect, we should point out that DFT is more appropriate for the analysis of large data samples. Several analysis were performed with toy Monte Carlo data for various values of μ . Again, it was demonstrated that the DFT approach outperforms the numerical integration having an accuracy of better than 0.5 % inside the $\mu \sim 0.5\text{--}5.0$ window.

The following remarks are necessary. In this publication we have only tried to demonstrate the calibration of the R7081 PMT model in just the simplest case. That is, when a single PMT was placed inside a dark, light-tight box and with the optical fiber positioned at the center of the photocathode, and always pointing vertically towards the surface of the PMT. We have demonstrated that in this simple example, the gain remains remarkably stable (within $\sim 1\%$ or better) inside the $\mu \sim 0.5\text{--}2.0$ plateau. We should note that in large (monolithic) detectors equipped with a sizable number of PMTs this simplistic picture ceases to apply. In particular, in those circumstances we can expect the extraction of the gain to depend on the geometry of the detector and the position of the event. We can only expect that the accuracy achieved in this article will not be attainable in such cases. The question of the *in situ* gain calibration of similar, complicated instruments lies beyond the scope of this article and was not treated here. More details on the energy and spatial resolution of large-volume liquid scintillator detectors can be found elsewhere [13].

A Calculation of Q_R and σ_R

In order to compute Q_R and σ_R we first write down the formulae:

$$\sum_{n=0}^{+\infty} P(n; \mu) = 1, \quad (\text{A.1})$$

$$\sum_{n=0}^{+\infty} nP(n; \mu) = \mu, \quad (\text{A.2})$$

$$\sum_{n=0}^{+\infty} n^2 P(n; \mu) = \mu(\mu + 1). \quad (\text{A.3})$$

The first equation stems from probability conservation and it is very easy to derive. The other two can be proved by a shifting of the summing parameter n . We also note the formulae which are deduced from the properties of the convolution:

$$\langle (S_n * B)(x) \rangle = Q_0 + nQ_s, \quad (\text{A.4})$$

$$\text{Var}[(S_n * B)(x)] = \sigma_0^2 + n\sigma_s^2. \quad (\text{A.5})$$

We simplify our notation by setting $S_R^{(n)}(x) = (S_n * B)(x)$. Using these equations, the mean value Q_R becomes:

$$\begin{aligned} Q_R &= \langle S_R(x) \rangle \\ &= \sum_{n=0}^{+\infty} P(n; \mu) \int_{-\infty}^{+\infty} x S_R^{(n)}(x) dx \\ &= \sum_{n=0}^{+\infty} P(n; \mu) (Q_0 + nQ_s) \\ &= Q_0 + \mu Q_s. \end{aligned} \quad (\text{A.6})$$

On the other hand, to calculate the variance σ_R^2 we first find the integral:

$$\begin{aligned} \int_{-\infty}^{+\infty} x^2 S_R(x) dx &= \sum_{n=0}^{+\infty} P(n; \mu) \int_{-\infty}^{+\infty} x^2 S_R^{(n)}(x) dx \\ &= \sum_{n=0}^{+\infty} P(n; \mu) (\sigma_0^2 + n\sigma_s^2 + (Q_0 + nQ_s)^2) \\ &= \sum_{n=0}^{+\infty} P(n; \mu) (\sigma_0^2 + n\sigma_s^2 + Q_0^2 + n^2 Q_s^2 + 2nQ_0 Q_s) \\ &= \sigma_0^2 + \mu\sigma_s^2 + Q_0^2 + \mu(\mu + 1)Q_s^2 + 2\mu Q_0 Q_s \\ &= \sigma_0^2 + \mu\sigma_s^2 + \mu Q_s^2 + (Q_0 + \mu Q_s)^2. \end{aligned} \quad (\text{A.7})$$

Where in the fourth line we made use of the identities provided in the beginning of this section. The variance now becomes:

$$\begin{aligned} \sigma_R^2 &= \text{Var}[S_R(x)] \\ &= \int_{-\infty}^{+\infty} x^2 S_R(x) dx - Q_R^2 \\ &= \sigma_0^2 + \mu(\sigma_s^2 + Q_s^2). \end{aligned} \quad (\text{A.8})$$

B Variance of $S(x)$

The mean value of $S(x)$ is straightforward to work out and its proof will not be presented here. We only remark that the mean of the sum of two terms equals the sum of the two individual means.

$$Q_s = \frac{w}{\alpha} + (1 - w)Q_g \quad (\text{B.1})$$

The variance on the other hand is quite complicated and it will be treated in great detail. We first remind the reader that:

$$\sigma_s^2 = \int_{-\infty}^{+\infty} x^2 S(x) dx - Q_s^2, \quad (\text{B.2})$$

and we proceed to calculate each term in eq. (B.2) separately.

$$\begin{aligned} Q_s &= \left(\frac{w}{\alpha} + (1-w)Q_g \right)^2 \\ &= \frac{w^2}{\alpha^2} + 2 \frac{w(1-w)}{\alpha} Q_g + (1-w)^2 Q_g^2. \end{aligned} \quad (\text{B.3})$$

$$\begin{aligned} \int_{-\infty}^{+\infty} x^2 S(x) dx &= w \int_0^{+\infty} x^2 \alpha e^{-\alpha x} dx + (1-w) \int_{-\infty}^{+\infty} x^2 g(x) dx \\ &= \frac{2w}{\alpha^2} + (1-w)(\sigma_g^2 + Q_g^2) \end{aligned} \quad (\text{B.4})$$

Plugging eq. (B.3) and (B.4) into σ_R we have:

$$\begin{aligned} \sigma_s^2 &= \frac{2w}{\alpha^2} + (1-w)(\sigma_g^2 + Q_g^2) - \frac{w^2}{\alpha^2} - 2 \frac{w(1-w)}{\alpha} Q_g - (1-w)^2 Q_g^2 \\ &= \frac{w}{\alpha^2} + (1-w)\sigma_g^2 + \frac{w(1-w)}{\alpha^2} + w(1-w)Q_g^2 - 2 \frac{w(1-w)}{\alpha} Q_g \\ &= \frac{w}{\alpha^2} + (1-w)\sigma_g^2 + w(1-w) \left(\frac{1}{\alpha^2} + Q_g^2 - 2 \frac{Q_g}{\alpha} \right) \\ &= \frac{w}{\alpha^2} + (1-w)\sigma_g^2 + w(1-w) \left(Q_g - \frac{1}{\alpha} \right)^2. \end{aligned} \quad (\text{B.5})$$

C Mean value and variance of $g(x)$

Mean value. The mean value of $g(x)$ is equal to:

$$\begin{aligned} Q_g &= \int_{-\infty}^{+\infty} x g(x) dx \\ &= \frac{1}{g_N} \frac{1}{\sqrt{2\pi}\sigma} \int_0^{+\infty} x e^{-\frac{(x-Q)^2}{2\sigma^2}} dx. \end{aligned} \quad (\text{C.1})$$

The integral of eq. (C.1) can be calculated using the substitution of variables:

$$u = \frac{x-Q}{\sqrt{2}\sigma}. \quad (\text{C.2})$$

It turns out to be:

$$\begin{aligned}
I &= \int_0^{+\infty} x e^{-\frac{(x-Q)^2}{2\sigma^2}} dx \\
&= \int_{-\frac{Q}{\sqrt{2}\sigma}}^{+\infty} (\sqrt{2}\sigma u + Q) e^{-u^2} \sqrt{2}\sigma du \\
&= 2\sigma^2 \int_{-\frac{Q}{\sqrt{2}\sigma}}^{+\infty} u e^{-u^2} dx + \sqrt{2}\sigma Q \int_{-\frac{Q}{\sqrt{2}\sigma}}^{+\infty} e^{-u^2} dx \\
&= \sigma^2 e^{-\frac{Q^2}{2\sigma^2}} + \sqrt{2\pi}\sigma Q \frac{1}{2} \left(1 - \operatorname{erf}\left(-\frac{Q}{\sqrt{2}\sigma}\right) \right) \\
&= \sigma^2 e^{-\frac{Q^2}{2\sigma^2}} + \sqrt{2\pi}\sigma Q \frac{1}{2} \operatorname{erfc}\left(-\frac{Q}{\sqrt{2}\sigma}\right) \\
&= \sigma^2 e^{-\frac{Q^2}{2\sigma^2}} + \sqrt{2\pi}\sigma Q g_N.
\end{aligned} \tag{C.3}$$

If we substitute I into eq. (C.1) we then have:

$$Q_g = Q + \frac{1}{g_N} \frac{\sigma}{\sqrt{2\pi}} e^{-\frac{Q^2}{2\sigma^2}}. \tag{C.4}$$

This equation can be simplified by setting:

$$\kappa = \frac{1}{g_N} \frac{\sigma}{\sqrt{2\pi}} e^{-\frac{Q^2}{2\sigma^2}}, \tag{C.5}$$

so that Q_g is equal to:

$$Q_g = Q + \kappa. \tag{C.6}$$

Variance. To calculate the variance of $g(x)$ we proceed to find the integral:

$$\int_{-\infty}^{+\infty} x^2 g(x) dx = \frac{1}{g_N} \frac{1}{\sqrt{2\pi}\sigma} \int_0^{+\infty} x^2 e^{-\frac{(x-Q)^2}{2\sigma^2}} dx. \tag{C.7}$$

$$\begin{aligned}
I &= \int_0^{+\infty} x^2 e^{-\frac{(x-Q)^2}{2\sigma^2}} dx \\
&= \int_{-\frac{Q}{\sqrt{2}\sigma}}^{+\infty} (\sqrt{2}\sigma u + Q)^2 e^{-u^2} \sqrt{2}\sigma du \\
&= \sqrt{2\pi}\sigma g_N (\sigma^2 + Q^2) + \sigma^2 Q e^{-\frac{Q^2}{2\sigma^2}}
\end{aligned} \tag{C.8}$$

$$\begin{aligned}
\int_0^{+\infty} x^2 g(x) dx &= \frac{1}{g_N} \frac{1}{\sqrt{2\pi}\sigma} I \\
&= \sigma^2 + Q^2 + \frac{1}{g_N} \frac{\sigma Q}{\sqrt{2\pi}} e^{-\frac{Q^2}{2\sigma^2}} \\
&= \sigma^2 + Q^2 + \kappa Q
\end{aligned} \tag{C.9}$$

Finally, the variance σ_g^2 is given by the formula:

$$\begin{aligned}\sigma_g^2 &= \int_{-\infty}^{+\infty} x^2 g(x) dx - Q_g^2 \\ &= \sigma^2 + Q^2 + \kappa Q - (Q + \kappa)^2 \\ &= \sigma^2 - (Q + \kappa)\kappa.\end{aligned}\tag{C.10}$$

Acknowledgments

The data used in this publication were taken in the laboratory of M. Dracos and we wish to thank him for allowing us to use them for the purposes of this communication.

References

- [1] DOUBLE CHOOZ collaboration, *The Double Chooz antineutrino detectors*, *Eur. Phys. J. C* **82** (2022) 804 [[arXiv:2201.13285](https://arxiv.org/abs/2201.13285)].
- [2] RENO collaboration, *RENO: An Experiment for Neutrino Oscillation Parameter θ_{13} Using Reactor Neutrinos at Yonggwang*, [arXiv:1003.1391](https://arxiv.org/abs/1003.1391).
- [3] ICECUBE collaboration, *Calibration and Characterization of the IceCube Photomultiplier Tube*, *Nucl. Instrum. Meth. A* **618** (2010) 139 [[arXiv:1002.2442](https://arxiv.org/abs/1002.2442)].
- [4] R. Dossi, A. Ianni, G. Ranucci and O.Y. Smirnov, *Methods for precise photoelectron counting with photomultipliers*, *Nucl. Instrum. Meth. A* **451** (2000) 623.
- [5] DARKSIDE collaboration, *Light Yield in DarkSide-10: A Prototype Two-Phase Argon TPC for Dark Matter Searches*, *Astropart. Phys.* **49** (2013) 44 [[arXiv:1204.6218](https://arxiv.org/abs/1204.6218)].
- [6] L.N. Kalousis, J.P.A.M. de André, E. Baussan and M. Dracos, *A fast numerical method for photomultiplier tube calibration*, *2020 JINST* **15** P03023 [[arXiv:1911.06220](https://arxiv.org/abs/1911.06220)].
- [7] <http://www.fft.w.org>.
- [8] R. Brun and F. Rademakers, *ROOT: An object oriented data analysis framework*, *Nucl. Instrum. Meth. A* **389** (1997) 81.
- [9] <https://github.com/lkalousis/PMTCalib>.
- [10] M. Abramowitz and I.A. Stegun, *Handbook of Mathematical Functions: with Formulas, Graphs, and Mathematical Tables*, Dover Publications, 0009-Revised edition (June 1, 1965).
- [11] F. James and M. Winkler, *C++ MINUIT User's Guide*, <https://root.cern.ch/root/html/doc/guides/minuit2/Minuit2.html>.
- [12] L.N. Kalousis, *Calibration of the Double Chooz detector and cosmic background studies*, Ph.D. Thesis, University of Strasbourg (2012).
- [13] O.Y. Smirnov, *Energy and spatial resolution of a large-volume liquid-scintillator detector*, *Instrum. Exp. Tech.* **46** (2003) 327 [[arXiv:1811.02321](https://arxiv.org/abs/1811.02321)].