PMT Signal Processing

Bouke Jung (bjung@nikhef.nl)

2023-07-05 15:17:29

Abstract

This document outlines models used to describe PMT data and the routines for calibrating the PMT as seen from a Jpp perspective.

1 Introduction

The cubic kilometer neutrino telescope (KM3NeT) forms a research infrastructure, consisting of large 3D-arrays of photomultiplier tubes deployed in the deep waters of the Mediterranean Sea. It allows for the detection of neutrinos by recording the arrival times of Cherenkov photons produced by the relativistic charged particles which emerge from a neutrino interaction.
To reconstruct neutrino events, one needs to analyze the time-position correlations between photomultiplier tube hits. This can be done using a software structure called Jpp (pronounced as `yi-pee-pee`). This document presents the functionalities incorporated in two C++ classes (`JPMTSignalProcessorInterface` and `JPMTAnalogueSignalProcessor`), which simulate the response of a single photomultiplier tube to photon hits.

2 The Photomultiplier Tube

The photomultiplier tube (PMT) constitutes the principal sensor of KM3NeT. Sporting a development history of over 90 years [1], their fundamental working principle has remained the same over time. This section describes how a photomultiplier tube functions from a hardware perspective. The model used to describe a photomultiplier tube's response to an incoming light signal is also discussed.

The base design of a photomultiplier tube consists of four consecutively placed elements, encapsulated by a vacuum tube. These are the photo-cathode, the focusing electrodes, the electron multiplier and the anode. When combined, they allow for the conversion of individual photons on one end of the device into measurable currents on the other. In figure 1 sketches the overall detection principle is sketched. External light enters the vacuum tube through an input window located at the front. The light subsequently impinges on the photo-cathode surface, where an electron can be emitted into the vacuum through the photoelectric effect. A focusing electrode accelerates and focuses the electron towards the electron multiplier. This is a system consisting of several secondary-emission electrodes, called dynodes, which iteratively multiply the number of primary electrons originating from the photo-cathode, so that a measurable current at the anode is produced. A voltage divider supplies the electric fields needed to accelerate and focus the output electrons from one dynode to the next.

The material and structural design of a photomultiplier tube can be adjusted to satisfy specific needs and intents. In KM3NeT the list of specifications includes a high efficiency to detect light of the quantum level

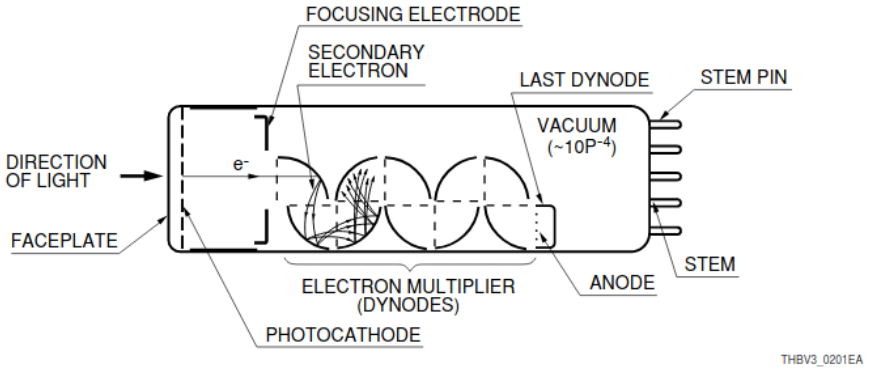


Figure 1: A schematic overview of the principal building blocks comprising a photomultiplier tube (from [2]).

with nanosecond time resolution. Taking this into account, the collaboration decided to use the R12199-02 type PMT produced by Hamamatsu.

2.1 Characteristic Variables

The quantity that determines the light sensitivity of a PMT is referred to as the quantum efficiency (QE). It is defined as the probability that an incoming photon is converted into a primary electron. Depending upon which material is used to coat the photo-cathode, the quantum efficiency of the photomultiplier tube will peak at different frequencies. For the R12199-02 type PMTs developed for KM3NeT, this coating material is Bialkali (SbKCs). Figure 2 shows the resulting quantum efficiency of the PMT as a function of the wavelength of incident light.

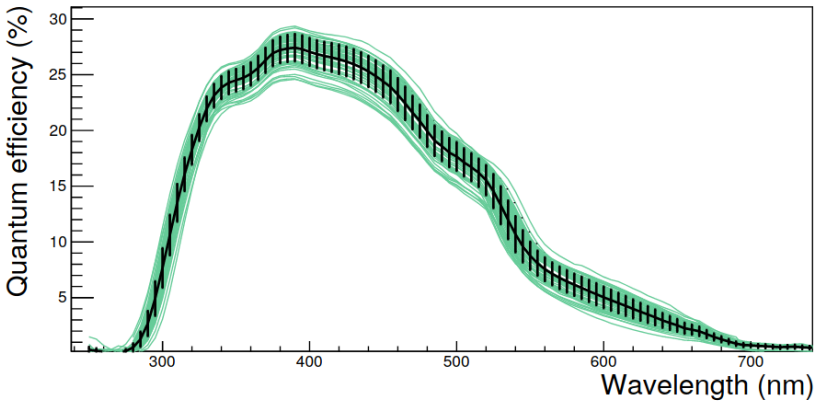


Figure 2: The photo-cathode quantum efficiency as a function of the wavelength, as measured for 56 R12199-02 type PMTs (from [3]). Curves for individual PMTs are shown in green, whilst the average is shown as a black line.

Not all photo-electrons generated at the photo-cathode or in the dynodes result in effective secondary emission in the subsequent stages of the electron multiplier. For example, certain trajectories may be

diverted such that a following dynode is missed. Effects like these are summarized in a quantity referred to as the collection efficiency, n. In principle, each inter-dynode space is characterized by its own collection efficiency, n_i. However, these can generally not be characterized individually. Instead, the collection efficiencies of all dynode stages are multiplied with their respective secondary emission coefficients, ϵ_i, to define an overall electron multiplication factor called the gain:

$$G = \prod_{i=1}^{N} n_i \epsilon_i \tag{1}$$

In general it is useful to define the gain in terms of the high voltage applied to the dynode stages. The dependence of the secondary emission coefficients on the inter-dynode voltages, V_i, is generally well-described by a power-law [4]:

$$\epsilon_i = a V_i^k, \tag{2}$$

where a is a constant associated with the dynode material and where $k < 1$ accounts for losses (e.g. due to scattering). Assuming that the voltage is distributed uniformly over the inter-dynode stages, i.e. $V_i = \frac{V}{N+1}$, and that all secondary emission coefficients are identical ($\epsilon_i = \epsilon$), this leads to the following expression for the gain:

$$G \equiv n \cdot \epsilon^N = n \cdot \frac{a^N V^{kN}}{(N+1)^{kN}} = A \cdot V^{kN}, \tag{3}$$

where the definition $A \equiv n \cdot a^N (N+1)^{-kN}$ with $n = \prod_{i=1}^{N} n_i$ has been substituted, in the last step.

Due to the statistical behavior of the electron multiplication process, the overall gain of a photomultiplier tube exhibits a spread. To good approximation, this spread can be described by regarding the secondary emissions at each dynode as individual Poisson processes, having means $\lambda_i = n_i \epsilon_i$ and standard deviations $\sigma_i \equiv \sqrt{\lambda_i} = \sqrt{n_i \epsilon_i}$. Following this line of reasoning, the gain for the overall electron multiplication can be expressed by equation 1, whilst the variance becomes [5; 6]:

$$\Sigma^2 = \sum_{i=1}^{N} \left(\frac{\sigma_i}{\lambda_i}\right)^2 \cdot \frac{\prod_{j=1}^{N} \lambda_j^2}{\prod_{k=1}^{i-1} \lambda_k} = \sum_{i=1}^{N} \frac{1}{n_i \epsilon_i} \cdot \frac{\prod_{j=1}^{N} (n_j \epsilon_j)^2}{\prod_{k=1}^{i-1} n_k \epsilon_k}. \tag{4}$$

Therefore the relative variance of the gain is given by:

$$\left(\frac{\Sigma}{G}\right)^2 = \sum_{i=1}^{N} \left(\prod_{j=1}^{i} n_j \epsilon_j\right)^{-1}, \tag{5}$$

Assuming that all secondary emission coefficients and collection efficiencies are roughly equal (i.e. $n_i \epsilon_i \approx n_1 \epsilon_1$), this can be reduced to $1/(n_1 \epsilon_1 - 1)$. Looking at equation 5, it becomes apparent that for $n_i \epsilon_i \gg 1$, any spread in the number of secondary electrons measured at the anode will be dominated by the variance in the yield of the first dynode. This can be thought of intuitively as a consequence of the limited statistics within the first multiplication stage.
Depending upon the choice of dynode material and structure, the values of the gain and the variance of the secondary electron distribution at the anode will change. Commonly used materials include AgMg, CuBe and NiAl [7]. Although these do not display substantial secondary emission by themselves, the oxidation products (e.g. MgO, BeO and Al_2O_3) forming on their surface do.

Where the dynode material influences the value of ϵ_i, its geometry holds sway over the collection efficiency, n_i. Two prevalent geometry choices are the Venetian blind and the box-and-grid structures. The first

arranges the dynodes in parallel strips, rotated slightly with respect to the longitudinal axis of the tube. This type of geometry has high collection efficiency and good gain stability. However, its time resolution is rather poor [2], because the electric fields close to the surface of each dynode are small [7]. The same holds true for the box-and-grid structure, which implements dynodes cut in a cylindrical shape which are positioned in quadrants. Since the detection of individual Cherenkov photons requires a time resolution of 1 ns, neither the Venetian blind nor the box-and-grid structure would be a good design choice for KM3NeT. Instead, the Hamamatsu R12199-02 photomultipliers used in KM3NeT, contain a linear focusing structure, where all dynodes are positioned in a row. Combined with the strong electric fields between each stage, this ensures that the transit-times of individual photo-electrons remains small, such that the required time resolutions can be achieved. A downside to this type of design is the relatively large non-uniformity of the photomultipliers response as function of incidence position and incidence angle on the photo-cathode. In the KM3NeT photomultiplier tubes, these effects lead to efficiency variations of 10% to 25% [8]. Other types of photomultipliers typically perform better in this respect, displaying efficiency variations of no more than 10% [9].

The time response of a photomultiplier tube is generally evaluated through something referred to as the transit-time distribution. This distribution maps out the time durations between the impingement of a photon on the photo-cathode and the appearance of an output current at the anode. It features a peak alongside contributions from prepulses and delayed pulses. The width of the peak is commonly referred to as the transit-time spread (TTS) and scales approximately with the number of primary photo-electrons per pulse, N_{pe} as $1/N_{pe}$. It has been estimated to be 3 ns for the photomultiplier tubes employed in KM3NeT [3]. The position of the peak in the transit time distribution relates to the voltage difference between the cathode and the anode, V, and is roughly proportional to $1/\sqrt{V}$. This follows from the acceleration of the photo-electrons across the electric fields.
The temporal behavior of output currents generated by individual photon hits is generally described using two quantities known as the current's rise-time and its decay-time. The rise-time corresponds to the time the associated analogue pulse takes to get from 10% to 90% of the maximal pulse height. Conversely, the decay-time is defined as the time to go back from 90% of the maximal pulse height to 10%. Whereas the rise-time is — to good approximation — independent of the number of photo-electrons, the decay-time generally increases with the number of generated primary photo-electrons. This is a consequence of the varying latencies of the secondary electrons and the relaxation time of the electronics.

3 Pulse shape modeling and time-over-threshold

The primary output of PMTs consist of currents, typically a few tens of ns in width, when a single photo-electron is produced. It is fed to an amplifier and subsequently passed through a threshold discriminator to cut out electronic noise. Typically, the threshold of a PMT corresponds to a 0.3 photo-electron signal. Both the leading and the trailing edge of the output of the threshold discriminator are timestamped. The least significant bit of these timestamps is 1 ns, which corresponds to a resolution of $1/\sqrt{12}$ ns[1]. Once the timestamps have been applied, the identifier of the associated PMT, the arrival time of the signal, i.e. the leading edge, and the time-over-threshold are stored for future use. The value of the leading edge depends on the height of the pulse passed to the threshold discriminator. This effect is referred to as time-slewing.

To maximize the performance of the event reconstruction and to reproduce the key features of a PMT in the simulations, the time-over-threshold values need to be related to the pulse-shapes which were originally

[1] I.e. the standard deviation of a uniform distribution with a width of 1 ns.

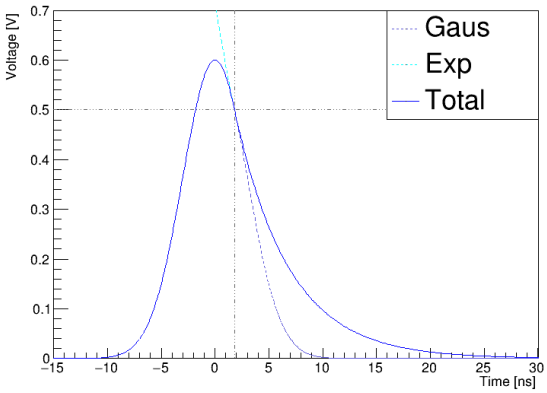


Figure 3: The model analogue pulse shape model, as described by equation 6). The match point between the Gaussian and exponential regime shown at the intersection of the black dotted lines, is found by imposing equal derivates and function values.

recorded at each PMT. In Jpp this is done by assuming that all pulses follow the same, general shape consisting of a Gaussian with an exponential tail. The resulting output voltage as a function of time and charge $V(t, q)$ can therefore be written down as:

$$V(t,q) = \begin{cases} q\tilde{R}e^{-\frac{1}{2}\left(\frac{t}{\sigma}\right)^2}, & t \leq \tilde{t}, \\ \frac{q\tilde{R}}{C}e^{-\frac{t}{\tau}}, & t > \tilde{t}. \end{cases} \tag{6}$$

In this, σ corresponds to the width of the Gaussian component of the pulse (related to the PMT rise-time as $t_r = \sigma\left(\sqrt{-2\ln(0.1)} - \sqrt{-2\ln(0.9)}\right) \approx 1.69\sigma$), τ corresponds to the decay-time associated with the exponential tail of the pulse and C and $\tilde{R}$ correspond to two normalisation constants. The first normalisation constant, C, is equivalent with the normalised pulse height evaluated at the time $\tilde{t} = \frac{\sigma^2}{\tau}$ where the Gaussian and exponential part of the pulse match both in terms of their amplitudes and derivatives, i.e. $C \equiv \frac{V(\tilde{t},q)}{q\tilde{R}} = e^{-\frac{1}{2}\left(\frac{\sigma}{\tau}\right)^2}$. The second normalisation constant is related to the electronic circuit's load resistance and carries units of resistance over time. A more detailed description of this constant can be found in appendix A.1. Figure 3 shows an example analogue pulse with an amplitude of $0.6\,\text{V}$, comprising a Gaussian centered around time $t = 0\,\text{ns}$ with a variance equal to $9.0\,\text{ns}$ and an exponential decay time of $5.0\,\text{ns}$.

From equation 6, it is possible to define the time-over-threshold as the difference between the first and last time during which the analogue pulse exceeds the voltage threshold V_0:

$$t_1 = -\sigma\sqrt{2\ln\left(\frac{q\tilde{R}}{V_0}\right)}, \tag{7}$$

$$t_2 = \begin{cases} \sigma\sqrt{2\ln\left(\frac{q\tilde{R}}{V_0}\right)}, & V(\tilde{t},q) \leq V_0, \\ \tau\ln\left(\frac{q\tilde{R}}{CV_0}\right), & V(\tilde{t},q) > V_0. \end{cases} \tag{8}$$

These times are traditionally referred to respectively as the leading and the trailing edge of the pulse. Using these definitions, the time-over-threshold ΔT can be expressed as:

$$\begin{aligned}\Delta T(q) &\equiv t_2 - t_1 \\ &= \begin{cases} 2\sigma\sqrt{\ln\left(2\frac{q}{q_0}\right)}, & q_0 < q \leq \frac{q_0}{C}, \\ \tau \ln\left(\frac{q}{Cq_0}\right) + \sigma\sqrt{-2\ln\left(\frac{q}{q_0}\right)}, & q > \frac{q_0}{C}, \\ 0, & q \leq q_0. \end{cases}\end{aligned} \tag{9}$$

In this, q_0 corresponds to the threshold-equivalent charge, set by the condition $V(0, q_0) = V_0$. Given that a nominal pulse corresponds to a time-over-threshold, $\Delta T(1)$, of 25.08 ns, it is possible to relate the rise- and decay-time to each other, such that the total number of degrees of freedom in equation 9 reduces to two (i.e. the Gaussian width, σ, and the threshold-setting, q_0) [10]. Starting from the case $q > \frac{q_0}{C\tilde{R}}$ and filling in $q = 1$ and $C = e^{-\left(\frac{\sigma}{\tau}\right)^2}$, it follows that:

$$\tau = \frac{-b + \sqrt{b^2 + 4\sigma^2 \ln(q_0)}}{2\ln(q_0)}, \tag{10}$$

where $b \equiv \sigma\sqrt{2\ln(q_0)} - \Delta T(1)$.

Sometimes an analogue signal exceeds the voltage bias inside the amplifier and discriminator. In such scenarios the top part of the pulse will be clipped. This introduces a sub-domain, where the time-over-threshold scales linearly with respect to charge as long as the pulse is in excess of the voltage bias:

$$\Delta T(q) = \tau \ln\left(\frac{q_{\mathrm{L}}}{q_0 C}\right) + \sigma\sqrt{2\left(\frac{q_{\mathrm{L}}}{q_0}\right)} + \beta(q - q_{\mathrm{L}}), \qquad q > q_{\mathrm{L}}. \tag{11}$$

In this, β corresponds to the slope associated with the linear behavior, whereas q_{L} indicates the boundary between the linear and non-linear regime. Lab studies have pointed out that $\beta \approx 7.0\,\mathrm{ns\,p.e.}^{-1}$ under normal PMT operation (i.e. with a high-voltage setting resulting in a current amplification of $3.45{\times}10^6$ [3]). Although the linear behavior holds relatively well for pulses up to a few tens of p.e., above 30 p.e. a saturation effect is observed which causes a flattening of the time-over-threshold with increasing charge [11]. This effect is modeled phenomenologically with an inverse square-root modulation defined as:

$$f(\Delta T) = \frac{\Delta T_{\max}}{\sqrt{\Delta T_{\max}^2 + \Delta T^2}}. \tag{12}$$

In this, ΔT corresponds to the time-over-threshold without saturation and $\Delta T_{\max} = 210\,\mathrm{ns}$ corresponds to the asymptotic limit for large time-over-threshold observed in the lab. Summarizing all of the above, the final expression for the time-over-threshold can be written down as a function of charge as:

$$\Delta\tilde{T} = \frac{\Delta T_{\max}}{\sqrt{\Delta T_{\max}^2 + \Delta T^2(q)}} \cdot \Delta T(q), \tag{13}$$

with:

$$\Delta T(q) = \begin{cases} 2\sigma\sqrt{-2\ln\left(\frac{q_0}{q}\right)}, & q_0 < q \leq \frac{q_0}{C}, \\ \tau\ln\left(\frac{q}{q_0 \cdot C}\right) + \sigma\sqrt{2\ln\left(\frac{q}{q_0}\right)}, & \frac{q_0}{C} < q \leq q_{\mathrm{L}}, \\ \tau\ln\left(\frac{q_{\mathrm{L}}}{q_0 \cdot C}\right) + \sigma\sqrt{2\ln\left(\frac{q_{\mathrm{L}}}{q_0}\right)} + \beta(q - q_{\mathrm{L}}), & q > q_{\mathrm{L}} \\ 0, & q \leq q_0. \end{cases} \tag{14}$$

Figure 4 shows the full functional behavior of the time-over-threshold in terms of charge, as described by equation 13.

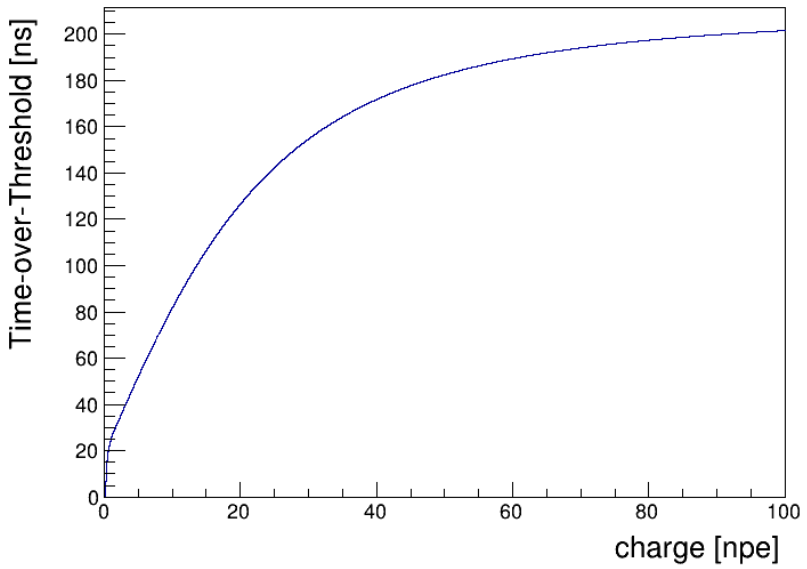


Figure 4: The time-over-threshold as function of charge, as described by equation 13.

Because equation 14 constitutes a one-to-one mapping, it is possible to define its inverse which yields the charge as function of the time-over-threshold of an analogue pulse. The definition of this inverse function is given in appendix A.2. In the following sections a description will be given of the probability density functions of the time-over-threshold and the charge, which express how frequently pulses with specific charge and time-over-threshold occur.

4 The Charge Distribution

For a given PMT it is possible to construct what is called the charge distribution by mapping out the occurrence frequencies of signals with different charges. When associated with physical light sources that give rise to the emission of single photo-electrons from the PMT's photo-cathode, the resulting histogram is called the single photo-electron distribution.
To infer the properties of light sources from data acquired with photomultiplier tubes, accurate modeling of

each of the devices' single photo-electron distributions is paramount. In an ideal scenario, the output current at the anode would perfectly reflect the Poisson statistics governing the electron multiplication at each dynode stage. However, certain processes in between the dynode stages can give rise to disturbances. In the literature one therefore encounters other functions to describe the single photo-electron response, which allow for a more tunable variance. These range from a single Gaussian approximation [12], to log-normal [13] and negative binomial distributions [14; 15]. The single Gaussian approximation is adequate in cases where signal contributions deriving from underamplification and noise effects can be neglected. However, when these need to be taken into account the negative binomial and the log-normal are more more suitable, since they accommodate an adjustable skewness. On the downside, they are computationally very difficult to manipulate.

Another solution for modeling underamplification consists of adding various distributions together. KM3NeT has opted for the latter and modeled the full single photo-electron response of the PMTs as a mixture of two Gaussian distributions:

$$f(q;1) = p \cdot \mathcal{G}(q;\mu_u,\sigma_u^2) + (1-p)\cdot\mathcal{G}(q;\mathrm{G},\Sigma^2), \tag{15}$$

where p corresponds to the occurrence probability of an underamplified hit and $\mathcal{G}(x;\mu,\sigma^2)$ to a Gaussian distribution with a mean μ and standard deviation σ. The mean and standard deviation of the nominal (underamplified) distribution are referred to as G and Σ (μ_u and σ_u), respectively. These parameters are related by the expression $\Sigma = \sigma_\mathrm{G}\sqrt{\mathrm{G}}$, where G and σ_G correspond to the gain and the gainspread of the PMT. By convoluting the single photo-electron distribution N times with itself, the multi photo-electron distribution generated by N primary electrons at the photo-cathode can be expressed as:

$$\begin{aligned} f(q;N) &\equiv f(q;N-1) \circledast f(q;1) \\ &= \sum_{k=0}^{N} \binom{N}{k} p^{N-k}(1-p)^k \cdot \mathcal{G}\left(q;\mu_k,\sigma_k^2\right), \end{aligned} \tag{16}$$

where $\mu_k \equiv (N-k)\mu_u + k\mathrm{G}$ and $\sigma_k \equiv \sqrt{(N-k)\sigma_u^2 + k\Sigma^2}$, as given by the Gaussian convolution theorem [16]. Note that equation 16 reverts back to the original single photo-electron distribution in the case where $N = 1$.

Since the underamplified and the nominal components of the single-photo-electron distribution are required to transform in a coherent way, μ_u and σ_u are defined in direct relation to the gain and gainspread in terms of two scaling factors $0 < \zeta_\mathrm{G} \leq 1$ and $0 < \zeta_\Sigma \leq 1$:

$$\mu_u = \zeta_\mathrm{G} \cdot \mathrm{G}, \tag{17}$$

$$\sigma_u = \zeta_\Sigma \cdot \Sigma. \tag{18}$$

This way, when the gain or the gainspread of the PMT is adjusted, both parts of the single photo-electron distribution (as well as the resulting multi photo-electron distribution) will be affected in the same manner. Assuming a Poissonian dynode response, where the spread in the final charge distribution is caused by statistical fluctuations in the first electron multiplication step, the scaling factor for the gain has been set to $\zeta_\mathrm{G} = \sigma_\mathrm{G}^2$, such that $\mu_u = \Sigma^2$. To ensure that $\sigma_u \propto \sqrt{\mu_u}$, the scaling factor for the gainspread has been set correspondingly to $\zeta_\Sigma = \sqrt{\zeta_\mathrm{G}} = \sigma_\mathrm{G}$.

4.1 Charge Distribution Lower Bound and Thresholdband

Due to tje application of the hardware threshold, the domain of the charge distribution is bounded towards the lower end. In the past, this lower bound was presumed to be equal to the charge of a pulse which falls just short of the threshold setting, i.e. $q > q_0$. However, this assumption does not comply with the observation that there is an excess the time-over-threshold distributions around $4\,\mathrm{ns}$. The excess has been shown in the past to correspond to genuine hits [17] and seems to coincide roughly with the RC-time of the electronics [18]. A possible explanation for its appearance would be a sub-threshold contribution of the analogue signal of the PMT. In correspondence with this, the full PMT charge distribution model in KM3NeT includes a sub-threshold component known as the thresholdband.

The thresholdband essentially constitutes an effective tuning of the lower domain-bound on the charge distribution. It is defined in Jpp using one additional parameter which is called the thresholdbandwidth, q_b. Any non-zero setting of this parameter will shift the lower bound of the charge distribution to $q_0 - q_\mathrm{b}$. The region of charge within $(q_0 - q_\mathrm{b}) < q < q_0$ is correspondingly referred to as the thresholdband.

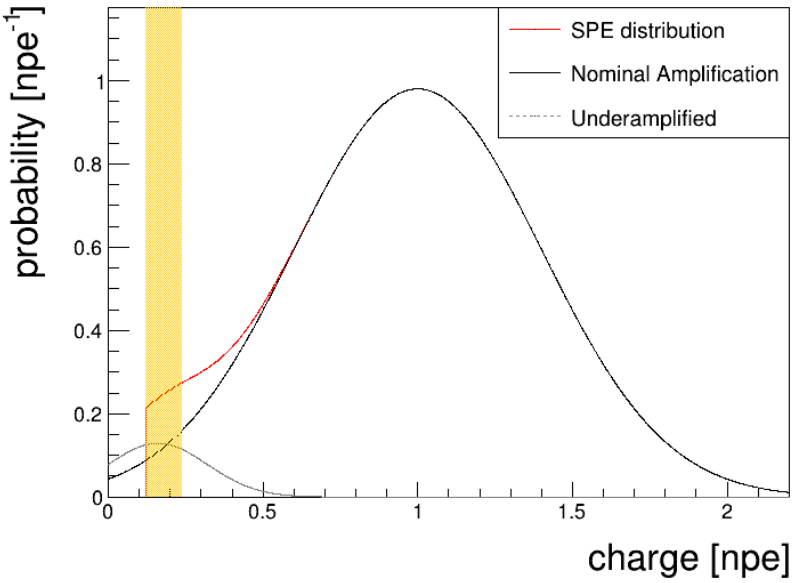


Figure 5: The single photo-electron distribution model for $p = 0.05$, $\zeta_\mathrm{G} = \sigma_\mathrm{G}^2 = 0.09$ and $\zeta_\Sigma = \sigma_\mathrm{G} = 0.3$. The individual contributions from the underamplified and nominal charge distributions are indicated with the black dotted and solid lines, respectively. The area marked in orange shows the thresholdband.

In figure 5 the modeled single photo-electron distribution is shown. Note that the presence of the lower-bound effectively constitutes a truncation of the overall distribution. To ensure proper normalization, the sum in equation 16 should therefore be divided by a factor Q, given by:

$$
\begin{aligned}
Q &= \int_{q_0-q_\mathrm{b}}^{\infty} f(q;N)dq \\
&= \frac{1}{2}\sum_{k=0}^{N}\binom{N}{k}p^{N-k}(1-p)^k \cdot \mathrm{erfc}\left(\frac{(q_0-q_\mathrm{b})-\mu_k}{\sigma_k\sqrt{2}}\right),
\end{aligned} \tag{19}
$$

where $\mathrm{erfc}(z) = 1 - \mathrm{erf}(z) = \frac{2}{\sqrt{\pi}}\int_z^\infty e^{-t^2}dt$ denotes the complementary error function.

4.2 Sampling the charge distribution

To verify the parameter configuration used to describe a certain PMT, experimental data needs to be compared directly with simulations. This requires the generation of random data-samples using specific calibration settings. In Jpp, the former can be done by drawing random samples from a PMT's charge distribution.
Since the charge distribution (c.f. equation 16) constitutes what is called a truncated mixture distribution, the generation of associated random variables requires two steps instead of one. The first step decides which of the contributions in the sum of equation 16 to draw from. Mathematically this winds down to a single sampling from a multinomial distribution with N categories, each category having an occurrence probability $p_j = \binom{N}{j}p^{N-j}(1-p)^j$. In other words, one needs to draw a single collection of random variables $(X_1, \ldots, X_N) \sim \mathcal{M}(n = 1; p_1, \ldots, p_N)$. Computationally, this can be accomplished via inversion sampling[2]. Having drawn an auxiliary random variable Y from a uniform distribution, the multi photo-electron distribution component, k, for which $X_k = 1$ is given by:

$$
k = \min\left\{j \in \{1, ..., N\} : \left(\sum_{i=1}^{j} p_j\right) \geq Y\right\}. \tag{20}
$$

The second step draws a random charge-value from this component. This is done with ROOT, using the acceptance-complement ratio method for drawing from a standard normal (see W. Hoermann and G. Derflinger (1990) [19] for more information).

5 The survival probability

As seen in section 4, introducing the concept of underamplified hits modifies the shape of the charge distribution. Similarly, it affects the survival probability. The survival probability is defined as the chance that a specific number of initial photons impinging on the photo-cathode generate a hit. Assuming that all photons would result in the emission of a primary photo-electron, this can be written down as:

$$
\epsilon(N) \equiv \frac{\int_{q_0}^{\infty} f(q;N)dq}{\int_0^{\infty} f(q;N)dq}. \tag{21}
$$

Here the influence of underamplification becomes apparent: the higher the underamplification probability p, the smaller the contribution of pulses with charge above the threshold. To simulate the detector response to Cherenkov light, it is convenient to introduce an additional average quantum efficiency $\langle\eta\rangle$ (c.f. section 2.1).

[2]This works because each of the random draws X_i has a marginal Bernoulli distribution: $X_i|X_{i-1}, \cdots, X_1 \sim \mathcal{B}(n = 1; 1 - \sum_{k=1}^{i-1} X_k, p_i/(1 - p_{i-1}))$.

The number of primary photo-electrons, n_p, generated by the photo-cathode due to the impingement of N photons, will then be distributed as a binomial. Consequently, the probability density for N photon hits, generating an observable pulse with charge q at the anode, becomes:

$$P_{\text{obs}}(q;N) = \sum_{m=1}^{N} \binom{N}{m} \cdot \langle\eta\rangle^m \cdot (1-\langle\eta\rangle)^{N-m} \cdot f(q;m), \tag{22}$$

where $f(q;m)$ corresponds to the multi photo-electron distribution from equation 16.

$$\begin{aligned} P_{\text{survival}}(N) &= \frac{\int_{q_0}^{\infty} P_{\text{obs}}(q,N)dq}{\int_0^{\infty} f(q;N)dq} \\ &= \sum_{m=1}^{N} \binom{N}{m} \cdot \langle\eta\rangle^m \cdot (1-\langle\eta\rangle)^{N-m} \cdot \frac{\int_{q_0}^{\infty} f(q;m)dq}{\int_0^{\infty} f(q;m)dq}, \\ &= \sum_{m=1}^{N} \binom{N}{m} \cdot \langle\eta\rangle^m \cdot (1-\langle\eta\rangle)^{N-m} \cdot \epsilon(m). \end{aligned} \tag{23}$$

The result obtained from equation 23 is shown in figure 6. As expected, the chance of survival decreases whenever p goes up. One of the applications where the equation comes into play, is in the calibration of the relative quantum-efficiency, $\langle\eta\rangle$. Since the function also depends on the gain and gainspread of the PMT, it is important to perform the gain-calibration first. This can be done using the time-over-threshold fitting routine implemented in Jpp (see section 6).

6 The time-over-threshold Distribution

In this section, the procedure to estimate the gain and the gainspread of a photomultiplier tube from the measured time-over-threshold distributions is presented. Relating the time-over-threshold distribution to the

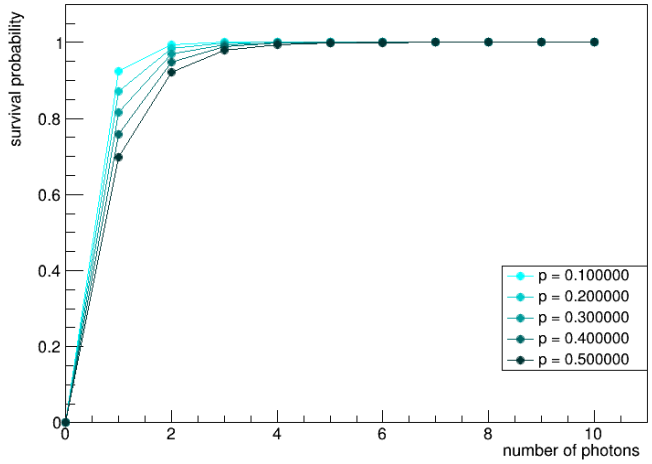


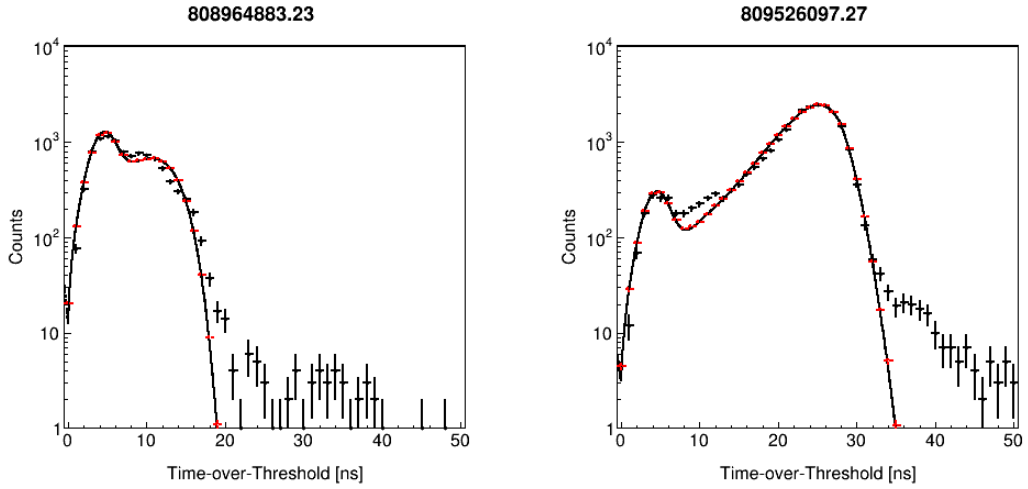
Figure 6: The survival probability of as a function of the number of photons impinging on the photo-cathode, for different settings of the underamplification probability, p.

charge distribution requires a change of variables. For this purpose, one can use the fact that for a function $g : \mathbb{R} \to \mathbb{R}$ with inverse g^{-1}, the probability density function of $Y = g(X)$, where X is a random variable distributed as $f_X(x)$, is given by:

$$f_Y(y) = f_X(g^{-1}(y)) \left| \frac{d}{dy}(g^{-1}(y)) \right|. \tag{24}$$

Associating Y with the time-over-threshold, ΔT, and X with the charge, q, this yields:

$$f(\Delta T; N) = f(q(\Delta T); N) \cdot \frac{\partial q(\Delta T)}{\partial \Delta T}. \tag{25}$$



(a) Low Gain (b) Nominal Gain

Figure 7: The time-over-threshold distributions retrieved for a low gain photomultiplier tube (a) and a photomultiplier tube with nominal gain (b) from the ORCA data corresponding to run number 6018. Black dots correspond to the original data and black solid lines to the result of a fit derived according to the model in equation 26, with $\mu_n = 4.5\,\text{ns}$, $\sigma_n = 1.5\,\text{ns}$, $q_0 = 0.22\,\text{p.e.}$, $q_\text{b} = 0.12\,\text{p.e.}$, $p = 0.05$, $\zeta_\text{G} = \sigma_\text{G}^2$ and $\zeta_\sigma = \sigma_\text{G}$. The red dots display simulation data generated with the model input obtained from the fit.

As previously discussed in section 4.1, not all time-over-threshold measurements can be accounted for using the analytical pulseshape description given in section 3. Compared to the time-over-threshold values which can be derived from 6, an excess of pulses with time-over-thresholds around $4.5\,\text{ns}$ is observed in nearly all KM3NeT PMTs. To account for the feature, an additional Gaussian distribution is included in equation 25 which scales according to the number of pulses with charge inside the thresholdband (c.f. section 4.1):

$$\begin{aligned} f(\Delta T; N) &= f_\text{a}(\Delta T) + f_b(\Delta T) \\ &= W_\text{a} \cdot f(q(\Delta T); N) \cdot \frac{\partial q(\Delta T)}{\partial \Delta T} + W_b \cdot \mathcal{G}(\Delta T; \mu_n, \sigma_n), \end{aligned} \tag{26}$$

where $\mu_n \approx 4.5\,\text{ns}$ and $\sigma_n \approx 1.5\,\text{ns}$ denote the mean and standard deviation of the observed excess. The normalization constants W_a and W_b are defined as $W_b = \int_{q_0 - q_b}^{q_0} f(q; N) dq$ and $W_\text{a} = 1 - W_\text{b} = \int_{q_0}^{\infty} f(q; N) dq$.

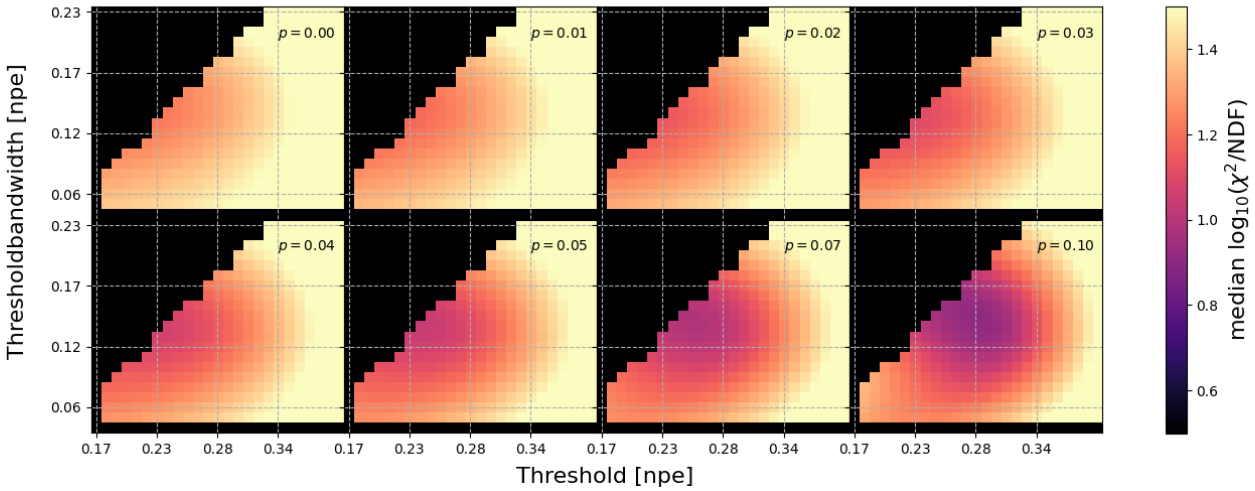


Figure 8: A gridscan over the parameter space relevant for fitting time-over-threshold distributions. The color scale displays the computed logarithmic median reduced chi-square over the fits of all time-over-threshold distributions which have been extracted from single coincident hit data (L1-data) in the 4-string ORCA run, 6018. The black areas indicate regions without data, deriving from the constraint that $q_\mathrm{b} \leq q_0$. The images are arranged according to increasing underamplification probability, p, in top-to-bottom, left-to-right order.

By varying over the gain and gainspread, implicitly contained through $f(q(\Delta T); N)$, it becomes possible to use equation 26 to perform gain calibration with time-over-threshold data. Two time-over-threshold distributions are shown in figure 7. The original data points are indicated with black dots. A black curve presents the result of a fit obtained by varying equation 26 over the gain and gainspread. Additionally, the results of a simulation acquired using the derived fit parameters are shown in red.

Besides the gain and the gainspread, a number of other parameters are of influence on the outcome of the fit derived using equation 26. In an attempt to minimize the degrees of freedom, all of these parameters are fixed to a set value before the start of the fitting procedure. In figure 7, the chosen parameter values were $\mu_n = 4.5\,\mathrm{ns}$, $\sigma_n = 1.5\,\mathrm{ns}$, $q_0 = 0.33\,\mathrm{p.e.}$, $q_\mathrm{b} = 0.1\,\mathrm{p.e.}$, $p = 0.05$. To understand which parameter settings yield optimal results, a gridscan was set up to iterate over different values of the thresholdband variables, q_0 and q_b, and of the underamplification probability, p, and calculate the reduced chi-square of the corresponding time-over-threshold fits. The results are shown in figure 8. Based on the reduced chi-square values, one might judge that a high setting of the underamplication probability greater than 0.05, yields the most reliable fit results. However, underamplification probabilities beyond the few percent level cannot be considered physical. Hence a trade-off has to be made between the setting of p and the width of the thresholdband, such that the effect of underamplification is not overestimated; Version 15 of Jpp uses a default parameter configuration where $p = 0.05$, $q_0 = 0.24\,\mathrm{p.e.}$ and $q_\mathrm{b} = 0.12\,\mathrm{p.e.}$, which comes with a median reduced chi-square of approximately $10^{1.1}$. The largest contribution to the chi-square for this setting is made by a second excess of pulses with time-over-threshold around 10 ns, which can be seen in both example distributions in figure 7. An additional, secondary thresholdband could be implemented in order to take this subpopulation into account. This has not yet been done.

p	q_0 [p.e.]	q_b [p.e.]	$\langle\chi^2/\mathrm{NDF}\rangle$
0.05	0.24	0.12	$10^{1.1}$

Table 1: The default parameter settings for the underamplification probability p, the threshold-equivalent charge q_0 and the thresholdbandwidth q_b, chosen to model the time-over-threshold distribution and the resulting median reduced chi-square of the fit results obtained for the ORCA data corresponding to run 6018.

7 Conclusion

The previous sections have presented an overview of the current photomultiplier tube response model implemented in Jpp, specifically the functions found within the classes `JPMTSignalProcessorInterface` and `JPMTAnalogueSignalProcessor`. Starting with the underlying working principles of the photomultiplier tube, the concepts of PMT gains, gainspreads, quantum efficiencies and risetimes were introduced in section 2. These variables reappeared during the subsequent discussions of the modeling of the analogue pulse shape, charge distributions and the time-over-threshold. Section 3 showed how the photomultiplier pulses can be described using a Gaussian with an exponential tail. Such a definition leads to a direct, analytical formula for the time-over-threshold as function of charge. Together with the definition of the charge distribution, introduced in section 4, this formula allows for the derivation of the time-over-threshold probability density function (c.f. section 6), which can be used to fit the gain and gainspread from in-situ data, as explained in section 6.

An excess around 4.5 ns (and 10 ns) in the time-over-threshold distributions has been observed. As shown in section 6, this can be taken into account by a sub-threshold contribution of the signal. An additional contribution to the lower end of the charge distribution from underamplification was introduced in order to account for the fraction of pulses observed within the thresholdband for photomultiplier tubes with nominal gain, as opposed to those with nominal gain. To reduce the number of degrees of freedom, all additional parameters needed for the model were fixed during the gain-fitting procedure. The optimal parameter settings were evaluated using a gridscan, which lead us to a default configuration where the probability of an underamplified hit is $p = 0.05$, the charge-equivalent threshold setting is $q_0 = 0.24$ p.e. and the width of the thresholdband is $q_\mathrm{b} = 0.12$ p.e. (c.f. table 1).

While implementing the new concepts of the thresholdband and the phenomenon of underamplification, care was taken to make all changes complient with already present functionalities. For example, the computation for the photon hit survival probability had to be modified in order to account for the dependence of the fraction of pulses with charge below the threshold on the chance of having underamplified hits (see section 5). Additionally, all modifications were carried out such that turning the newly added features off (i.e. setting $q_\mathrm{b} = 0.0$ p.e. and $p = 0.0$), results in exactly the same computational behavior as before their implementation. All in all then, it can be said that the present PMT model implemented in Jpp allows for reproduction and fitting of photomultiplier tube time-over-threshold data, including the noisy feature seen around 4.5 ns, in a phenomenological and backwards-compatible way. For the future, the implementation of an additional thresholdband which models the additional peak around 10 ns might be considered.

References

[1] *B.K. Lubsandorzhiev. On the history of photomultiplier tube invention.* Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 567(1):236 – 238, 2006. Proceedings of the 4th International Conference on New Developments in Photodetection.*

[2] *KK Hamamatsu Photonics. Photomultiplier tube handbook.* Electron Tube Division,*, 2006.*

[3] *S. Aiello, S.E. Akrame, F. Ameli, et al. Characterisation of the hamamatsu photomultipliers for the KM3net neutrino telescope.* Journal of Instrumentation*, 13(05):P05035–P05035, may 2018.*

[4] *AG Wright.* The photomultiplier handbook*. Oxford University Press, 2017.*

[5] *W. Shockley and J. R. Pierce. A theory of noise for electron multipliers.* Proceedings of the Institute of Radio Engineers*, 26(3):321–332, March 1938.*

[6] *Glenn F Knoll. Radiation detection and measurement(book).* New York, John Wiley and Sons, Inc., 1979. 831 p*, 1979.*

[7] *Philips Photonics. Photomultiplier tubes, principles and applications.* Philips Export BV*, 1994.*

[8] *Thijs Juan van Eeden. Measuring the Angular Acceptance of the KM3NeT Digital Optical Module. Master's thesis, Technische Universiteit Delft, the Netherlands, 2019.*

[9] *L Baudis, S D'Amato, G Kessler, A Kish, and J Wulf. Measurements of the position-dependent photo-detection sensitivity of the hamamatsu r11410 and r8520 photomultiplier tubes.* arXiv preprint arXiv:1509.04055*, 2015.*

[10] *Bouke Jung. KM3NeT Computing and Software ELOG, message ID 444.*

[11] *Jonas Reubelt. Hardware studies, in-situ prototype calibration and data analysis of the novel multi-pmt digital optical module for the km3net neutrino telescope. 2018.*

[12] *E.H. Bellamy, G. Bellettini, J. Budagov, et al. Absolute calibration and monitoring of a spectrometric channel using a photomultiplier.* Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 339(3):468 – 476, 1994.*

[13] *David J. Kissick, Ryan D. Muir, and Garth J. Simpson. Statistical treatment of photon/electron counting: Extending the linear dynamic range from the dark count rate to saturation.* Analytical Chemistry*, 82(24):10129–10134, 2010. PMID: 21114249.*

[14] *J.R. Prescott. A statistical model for photomultiplier single-electron statistics.* Nuclear Instruments and Methods*, 39(1):173 – 179, 1966.*

[15] *Pavel Degtiarenko. Precision analysis of the photomultiplier response to ultra low signals.* Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 872:1–15, Nov 2017.*

[16] *Paul Bromiley. Products and convolutions of gaussian probability density functions.* Tina-Vision Memo*, 3(4):1, 2003.*

[17] *Jannik Hofestaedt. KM3NeT Analysis ELOG, message ID 220.*

[18] *Deepak Gajanana. private communication.*

[19] *W Hoermann and G Derflinger. The acr method for generating normal random variables.* Operations-Research-Spektrum*, 12(3):181–185, 1990.*

A Appendix

A.1 Derivation of the analogue pulse voltage as a function of time and charge

The shape of PMT analogue pulses are usually defined in terms of an output voltage like in equation 6. What is actually measured at the anode of a PMT, however, is a time-varying current $I(t)$. The voltage is retrieved from the current by multiplying with the load-resistance R of read-out electronics:

$$V(t) = I(t) \cdot R \tag{27}$$

In equation 6, the voltage is expressed not only as a function of the time t, but also in terms of the integrated charge q of the analogue pulse. As it turns out, this choice of definition is intimately related to the meaning of the normalisation constant $\tilde{R}$. Assuming the same pulse-shape model used in section 3, the current of the analogue pulse as a function of time can be written down as as:

$$I(t) = \begin{cases} I_0 e^{-\frac{1}{2}\left(\frac{t}{\sigma}\right)^2}, & t \leq \tilde{t} \\ \frac{I_0}{C} e^{-\frac{t}{\tau}}, & t > \tilde{t} \end{cases}, \tag{28}$$

where I_0 is the amplitude of the pulse and σ, τ and $\tilde{t}$ are respectively defined as the width of the Gaussian component of the pulse, the decay-constant associated with the exponential tail of the pulse and the time at which the Gaussian and exponential part of the pulse match, like before. To retrieve the total electric charge carried by the pulse, the current needs to be integrated over time:

$$\begin{aligned} q &= \int_{-\infty}^{+\infty} I(t)dt && (29)\\ &= \int_{-\infty}^{\tilde{t}} I(t)dt + \int_{\tilde{t}}^{+\infty} I(t)dt && (30)\\ &= \int_{-\infty}^{\tilde{t}} I_0 e^{-\frac{1}{2}\left(\frac{t}{\sigma}\right)^2} dt + \int_{\tilde{t}}^{+\infty} \frac{I_0}{C} e^{-\frac{t}{\tau}} dt && (31)\\ &= \frac{1}{2}\sqrt{2\pi}\sigma I_0 \left(1 + erf\left(\frac{\tilde{t}}{\sqrt{2}\sigma}\right)\right) + \frac{I_0 \tau}{C} e^{-\frac{\tilde{t}}{\tau}}. && (32) \end{aligned}$$

It follows that:

$$I_0 = q\left[\frac{1}{2}\sqrt{2\pi}\sigma \left(1 + erf\left(\frac{\tilde{t}}{\sqrt{2}\sigma}\right)\right) + \frac{\tau}{C} e^{-\frac{\tilde{t}}{\tau}}\right]^{-1}. \tag{33}$$

Coming back to the original definition of the voltage in terms of the current in equation 27, this means that:

$$\begin{aligned} V(t,q) &= I(t) \cdot R && (34)\\ &= \begin{cases} q\tilde{R} e^{-\frac{1}{2}\left(\frac{t}{\sigma}\right)^2}, & t \leq \tilde{t} \\ \frac{q\tilde{R}}{C} e^{-\frac{t}{\tau}}, & t > \tilde{t} \end{cases}, && (35) \end{aligned}$$

Thus the normalisation constant $\tilde{R} = RI_0/q$ is directly related to the load resistance of the electronics board and has units of resistance over time.

A.2 Charge as function of time-over-threshold

As mentioned at the end of section 3, the function for the time-over-threshold as a function of the charge is bijective. This means that it is possible to derive the inverse transformation which yields the charge of an analogue pulse in terms of its time-over-threshold. Using equation 13 to convert between saturated and non-saturated time-over-threshold, the inverse can be defined as:

$$q(\Delta T) = \begin{cases} q_0 \cdot \exp\left[\frac{1}{8}\left(\frac{\Delta T}{\sigma}\right)^2\right], & 0 < \Delta T \leq \Delta T\left(\frac{q_0}{C}\right), \\ q_0 \cdot \exp\left[\frac{z^2}{2}\right], & \Delta T\left(\frac{q_0}{C}\right) < \Delta T \leq \Delta T\left(q_{\mathrm{L}}\right), \\ q_{\mathrm{L}} + \frac{1}{\beta}(\Delta T - \Delta T\left(q_{\mathrm{L}}\right)), & \Delta T > \Delta T\left(q_{\mathrm{L}}\right), \end{cases} \tag{36}$$

In this, $z \equiv \frac{-\sigma+\sqrt{\sigma^2+2\tau(\Delta T+\tau\ln(C))}}{\tau}$.