LHCb Collaboration

LHCb 2004-016 RICH
26^{th} March, 2004

Analysis of Signal Spectra of Multianode Photo Multiplier Tubes

R. Cikoski, S. Eisenhardt[1], F. Muheim
University of Edinburgh

Abstract

The signal response of Multianode Photo Multiplier Tubes (MaPMT) is studied. The pulse height spectra of MaPMT are fitted with three different methods. The first method is based on Gaussian distributions for the signal, the second method is based on Poissonian signal distributions and the third approach includes a possible production of photoelectrons at the first dynode. The simulation of MaPMT pulse height spectra is discussed with the additional feature of a possible inhomogeneity in the electron multiplication probability across the surface of the first dynode. The signal loss below a threshold cut is measured with three methods.

[1] Corresponding author. E-mail: S.Eisenhardt@ed.ac.uk

1 Introduction

Multianode Photomultiplier Tubes (MaPMT) have been evaluated by the LHCb collaboration as technology for the RICH photodetectors [1, 2, 3]. In preparation of the proposal [4] the performance of the MaPMT as a photo detector was studied. Preliminary results from the tests carried out with light emitting diodes (LED) and using a CAMAC readout were described in [5]. Tests carried out with a LHC-speed readout in a testbeam environment were reported in [6].

In reference [5] the analysis of the signal spectra recorded with MaPMTs used a relatively simple fitting algorithm based on Gaussian distributions for the signal components allowing zero to three photoelectrons. This approach was easy to implement and rather fast in computing. But it turned out that the yield of signal events was strongly underestimated by this fit in the region above the pedestal. This caused a bias in the determination of the average single photoelectron signal. As a simple fix at that time the distortion of the fit results was reduced by removing the badly described region from the fit. This allowed an approximate determination of the average single photoelectron signal. But the shape of the signal spectra was not understood and the results depended on the definition of the exclusion region. This dependency dominated the uncertainty in the determination of the signal loss caused by a threshold cut. Which is why no results on the signal loss were given in reference [5].

This note summarises our study of improvements to the description and understanding of the MaPMT signal spectra. The MaPMT data used in this study was taken originally with a CAMAC readout [5]. A comprehensive description of the used MaPMT and the used CAMAC readout can be found in [7]. The data analysis in [7] makes use of the improved algorithms discussed in this note.

Three improvements in the description of the signal spectra were made. For the signal components generated by one or two photoelectrons the Gaussian distributions were replaced by Poisson distributions. Additional signal contributions were introduced following references [8, 9]. They model an assumed photon conversion at the first dynode for photons which pass the semi-transparent photocathode without photo-conversion. These signals miss out the first multiplication stage and their charge is smaller by the average gain at the first dynode when compared to the signals of photoelectrons originating from the photocathode. Finally, an inhomogeneity of the electron multiplication capability of the dynode material would smear out any distribution of the signal produced at a particular dynode. Signal spectra taken with an MaPMT were compared to simulations with a selection of the three effects enabled. Only the first two effects, but not the inhomogeneity, finally were implemented into a fitting algorithm and tested against the data.

2 Algorithms for Fitting and Simulation

In this section the used algorithms are discussed in detail. Their implementation is documented in section 6 (Appendix A).

2.1 Signal Spectrum made from Gaussian Distributions

A simple and fast way to approximate the shape of the signal spectrum of a MaPMT channel is based on the following assumptions:

1. A major fraction of events in the spectrum are noise events, i.e. events without a charge pulse of the MaPMT as no photoelectron was produced, like when the high voltage supply to the tube is turned off. Thus the noise peak is a distinct feature of the signal spectrum and can be determined reliably by the fit. This certainly works if the fraction of noise events is 40 % or more.

2. The noise distribution is Gaussian shaped, i.e. a convolution of the electronic noise sources of the data acquisition chain, while asymmetric components like jumps or common mode are insignificant or corrected for:

$$\mathcal{G}_0(x) = \frac{\mathcal{N}_0}{\sqrt{2\pi}\sigma_0}\; e^{-\frac{1}{2}\left(\frac{(x-Q_0)^2}{\sigma_0^2}\right)} = \mathcal{N}_0 \cdot G(x, Q_0, \sigma_0) \tag{1}$$

   where $\mathcal{G}_0$ is the noise distribution, $\mathcal{N}_0$ is the number of noise events, Q_0 the pedestal position, i.e. the average pulse height x of a noise event, measured in ADC counts, and σ_0 the pedestal width, i.e. the total noise of the data acquisition system.

3. The probability to generate photoelectrons at the photocathode of an MaPMT is governed by Poisson statistics:

$$\mathcal{N}_n(\mu) = N \cdot \frac{\mu^n \; e^{-\mu}}{n!} \; , \qquad \Rightarrow \qquad \sum_{n=0}^{\infty} \mathcal{N}_n = N \tag{2}$$

where $\mathcal{N}_n$ is the number of events with a signal originating from n photoelectrons, N is the total number of events in the signal spectrum and μ is the parameter of the Poisson distribution, i.e. the average probability of producing a photoelectron. This definition explicitly includes the pedestal events with $n = 0$.

4. The gain of a MaPMT, especially after the first few dynodes, is high enough that the statistical process of the electron multiplication at the dynodes can be approximated by Gaussian statistics. Hence the final distribution of a signal of n photoelectrons is a Gaussian distribution as it is a convolution of all the signal distributions at the individual dynodes:

$$\begin{aligned} \mathcal{G}_n(x) &= \mathcal{N}_n \cdot G(x, Q_0 + Q_n, \sigma_0^2 + \sigma_n^2) \qquad \text{with} \qquad n > 0 \\ \mathcal{S}(x) &= \textstyle\sum_n \mathcal{G}_n(x) \end{aligned} \tag{3}$$

where $\mathcal{N}_n$ is the number of events of signals originating from n photoelectrons, Q_n is the average mean pulse height of a n-photoelectron spectrum after subtraction of the pedestal Q_0 and σ_n is the corresponding width of this signal contribution which gets folded in with the electronic noise σ_0. The sum $\mathcal{S}(x)$ describes the full signal spectrum.

5. Events with multiple photoelectrons ($n \geq 2$) can be regarded as the superposition of n coincident events of one photoelectron. Thus, due to the Gaussian shape of each contribution, the charge of the n signals is summed and the width is summed in quadrature:

$$Q_n = n \cdot Q_1 \qquad \text{and} \qquad \sigma_n^2 = n \cdot \sigma_1^2 \quad \Leftrightarrow \quad \sigma_n = \sqrt{n} \cdot \sigma_1 \tag{4}$$

The strength of this model is that is has a few simple assumptions which allow a computationally fast implementation. This implementation is documented in the Appendix section 6.1.

The limits of this model basically are twofold. Assumption 4 is not a good approximation for the case we want to study. The average probability μ for producing a photoelectron is usually smaller than one. In this case mostly events with $n \leq 2$ photoelectrons contribute to the signal spectrum. The statistical process of the emission of secondary electrons at dynodes of photomultipliers can be described by a Poisson distribution $\mathcal{P}(\lambda)$, like the emission of photoelectrons off the photocathode. λ is the average number of emitted electrons, or the gain g_j at dynode j. The average gain at the first dynode of a MaPMT in our application may vary in the range $4 \leq g_1 \leq 12$, depending on the configuration of the voltage divider chain and the applied high voltage. Using the phenomenological gain model described in section 2.4 we expect $g_1 \sim 7$ for our default setting with $HV = -900\,V$. The Poisson distribution can be approximated well by a Gaussian distribution for values of $\lambda \geq 10$. But for average gain values as low as $g_1 \sim 4$ at the first dynode the approximation is not good enough as the Gaussian distribution fails to describe reasonably well the shape of the signal distribution. At later dynodes the total number of electrons is large enough for the Gaussian approximation to be good. Thus the signal distribution at the first dynode is Poissonian and it gets folded with as many Gaussian distributions as there are remaining dynodes in the photomultiplier. The Poissonian shape therefore remains visible in the final signal spectrum. Hence a better description of the total signal spectrum has to account for the Poissonian nature of the electron multiplication as the first dynodes. This is what the algorithm described in section 2.2 is doing.

The second limitation of this algorithm is that additional physical processes may contribute to the total signal spectrum which are not accounted for. When the model is compared to data the data show an excess of events in the region between the pedestal and the MaPMT signal peak. Apart from incomplete signal sampling, i.e. a problem in the data acquisition setup, two other processes are worth to be considered.

There is a probability of photons passing through the semitransparent photocathodes of the MaPMT without producing a photoelectron. Instead such photons may produce photoelectrons when hitting the surface of the first dynode. Signals of such events would lack the electron multiplication at the first dynode. Hence their pulse height would fill the gap between pedestal and signal peak of the spectrum made of the signals subject to the full multiplication chain. An algorithm implementing this possibility is described in section 2.3.

In addition there is the possibility that the capability of the electron multiplication varies over the area of the dynodes, e.g. due to variations in the properties of the semiconductor layer. This would mean that the signal spectrum would effectively be a superposition of many signal spectra with varied gain and therefore its shape would be smeared out. As such a behaviour was studied in simulations which have been compared to the data. In section 2.4 the algorithm is discussed which was used for the simulation of this effect.

2.2 Signal Spectrum made from Poisson Distributions

This algorithm builds on top of the Gaussian based algorithm defined in the section 2.1. By replacing the assumption 4 with assumption 6 defined below the validity of the algorithm is extended. Its implementation is documented in the Appendix section 6.2. This is equivalent to the definition in [8, 9]:

6. The gain at the first dynode of the MaPMT might be too low to be reasonably well approximated by a Gaussian distribution. In our study this dominantly happens when only $n = 1$ or $n = 2$ photoelectrons contribute to the signal spectrum. For a higher photoelectron multiplicity the total number of electrons emitted at the first dynode exceeds 10 in average and the approximation by a Gaussian distribution is good enough. Equation 3 therefore has to be replaced by:

$$\begin{aligned} \mathcal{G}'_n(x) &= \begin{cases} \mathcal{N}_n \sum_{m=0}^{\infty} \frac{(nK_1)^m \cdot e^{-nK_1}}{m!} G(x, Q_0 + mQ_2, \sigma_0^2 + m\sigma_2^2) & n = 1, 2 \\ \mathcal{G}_n(x) \quad \text{as defined in equation 3} & n \geq 3 \end{cases} \\ \mathcal{S}'(x) &= \sum_n \mathcal{G}'_n(x) \end{aligned} \tag{5}$$

where Q_2 is the mean average pulse height, in ADC counts, of an electron emerging off the 1st dynode and impinging on the 2nd dynode[2], σ_2 is the corresponding width of the distribution in the Gaussian approximation, $K_1 = Q_1/Q_2$ is the signal gain at the 1st dynode and m is running over the number of electrons produced at the 1st dynode. The sum $\mathcal{S}'(x)$ describes the full signal spectrum.

This definition introduces two new fit parameters Q_2 and σ_2 representing the average mean pulse height and width of sub-signal distributions of electrons emerging off the first dynode, regardless of how they are produced. This sub-signal again is approximated by a Gaussian distribution. A Poisson distribution here would give an even more realistic description but also would need a further fit parameter Q_3 and would slow down the fitting time considerably by adding another level of loops.

This model gives a more realistic description of the signal spectrum than the Gaussian model presented in section 2.1. An additional strength of this model is that by fitting Q_1 and Q_2 independently it gives insight into an essential internal property of the MaPMT, namely the gain at the first dynode K_1. It also provides a platform to implement additional electron production mechanism at the first dynode, as will be seen in section 2.3.

This algorithm approximates the shape of the signal spectrum with Gaussian distributions only if more than two photoelectrons are produced. Thus the width σ_1 is determined by $\sigma_1 = \sigma_n/n$ with $n \geq 3$. This is meaningful only if there is a significant number of events with a gain $\geq 3 \cdot Q_1$ in the spectrum, i.e. the average multiplicity of photoelectrons is high enough. Otherwise the fit result for σ_1 will become arbitrary.

2.3 Adding the Photoconversion at the First Dynode

Based on the model using Poisson distributions presented in section 2.2 one can easily add the effect of a photoconversion at the first dynode. This may happen when a photon passes the semi-transparent photocathode, hits the semiconducting surface of the first dynode and causes the release of an electron via the photoeffect.

In this case the assumptions 3, 6 and 5 have to be extended to assumptions 7, 8 and 9, respectively, in the following way:

[2] This follows the logic that Q_1 is the total gain of an electron emerged off the photocathode and impinging on the 1st dynode. Hence, the symbols Q_2 and σ_2 have different meanings in the context of the Gaussian and the Poisson fit models.

7. The probabilities to generate photoelectrons at the photocathode or the first dynode of an MaPMT are governed by Poisson statistics and are independent of each other:

$$\mathcal{N}'_{n,k}(\mu,\mu_1) = N \cdot \frac{\mu^n\ e^{-\mu}}{n!}\frac{\mu_1^k\ e^{-\mu_1}}{k!}\ , \qquad \Rightarrow \qquad \sum_{n=0}^{\infty}\sum_{k=0}^{\infty}\mathcal{N}'_{n,k} = N \tag{6}$$

where $\mathcal{N}'_{n,k}$ is the number of events with a signal originating from $n+k$ photoelectrons originating from the photocathode (n) and from the first dynode (k), N is the total number of events in the signal spectrum and μ and μ_1 are the means of the Poisson distributions, i.e. the average probabilities of producing a photoelectron at the photocathode (μ) or at the first dynode (μ_1). This definition explicitly includes the pedestal events where $n+k=0$.

8. As for equation 5 it is assumed that a Poisson description is needed for the cases $n = 1, 2$ and a Gaussian distribution is sufficient for $n \geq 3$. To avoid another nested loop in the fitting routine Gaussian distributions are used for all k. So equation 5 will become:

$$\begin{aligned} \mathcal{G}''_{n,k}(x) &= \begin{cases} \mathcal{N}'_{n,k}\sum_{m=0}^{\infty}\frac{(nK_1)^m\cdot e^{-nK_1}}{m!}G(x, Q_0+(m+k)Q_2, \sigma_0^2+(m+k)\sigma_2^2) & n=1,2 \\ \mathcal{N}'_{n,k}G(x, Q_0+Q_{n,k}, \sigma_0^2+\sigma_{n,k}^2) & n\geq 3 \end{cases} \\ \mathcal{S}''(x) &= \textstyle\sum_n\sum_k \mathcal{G}''_{n,k} \end{aligned} \tag{7}$$

where the sum $\mathcal{S}''(x)$ describes the full signal spectrum.

9. Events with multiple photoelectrons ($n+k \geq 2$) can be regarded as the superposition of coincident events of n-times one photoelectron produced at the photocathode and k-times one photoelectron produced at the first dynode. Thus, due to the Gaussian shape of each contribution, the mean pulse height $Q_{n,k}$ of the $n+k$ signals, measured in ADC counts above the pedestal, is summed and the width $\sigma_{n,k}$ is summed in quadrature:

$$Q_{n,k} = n\cdot Q_1 + k\cdot Q_2 \qquad \text{and} \qquad \sigma_{n,k}^2 = n\cdot\sigma_1^2 + k\cdot\sigma_2^2 \tag{8}$$

The implementation of this algorithm is documented in the Appendix section 6.3.

2.4 Inhomogeneity of First Dynode

If the electron multiplication capability of the first dynode would vary across its surface effectively a superposition of signal spectra with different gain would be measured. As the gain at the first dynode varies the width of the effective measured signal is broadened. This behaviour cannot be fitted reliably together with the determination of the average total gain and the gain at the first dynode. As if another variation of these parameters is introduced to the fit arbitrary return values occur. Hence the effect of a possible inhomogeneity of the first dynode was studied in simulation only.

Two approaches for the simulation of the MaPMT signal spectra were used. In the first approach one of the fitting algorithms described above was implemented. The Gaussian and Poisson distributions were implemented by random number generators of the proper functional shape. Their parameters were taken from actual fits or their variations. A possible inhomogeneity of the gain at the first dynode was implemented in the following way:

10. The gain on the surface of the first dynode was assumed to vary linearly. For a given distribution to describe the number of electrons produced at the first dynode, either from electron multiplication or from photoconversion, the mean value was varied on an event-by-event base like:

$$\overline{g}'_1 = \overline{g}_1 + \mathcal{R}\mathcal{I}\cdot\overline{g}_1 \tag{9}$$

where $\overline{g}_1$ is the given mean of the distribution (i.e. K_1 or μ_1, respectively), $\mathcal{R}$ is a random number in the range $[-1,+1]$ and $\mathcal{I}$ is the strength of the inhomogeneity.

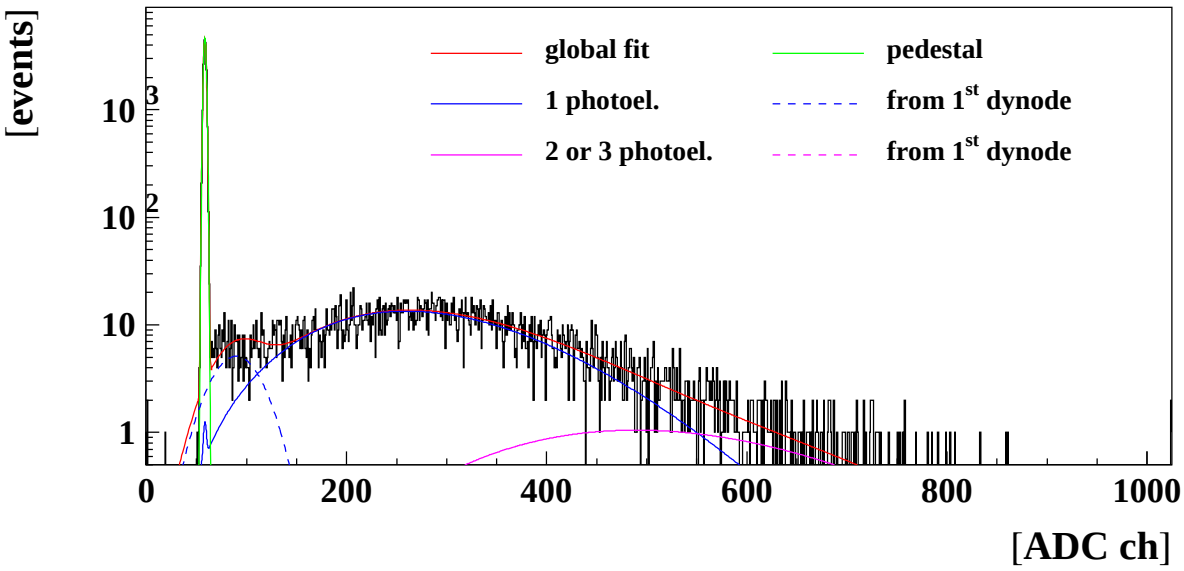

Figure 1: MaPMT signal spectrum at high gain with Poisson fit and assumed photoproduction at the first dynode, all contributions to the global fit are shown separately.

The second approach used a phenomenological gain model for the MaPMTs. In this model the gains of electrons in the whole tube G and at the individual dynodes g_j are:

$$G = \prod_{j=1}^{l} g_j = \prod_{j=1}^{l} a \cdot dV_j^k \tag{10}$$

where dV_j is the potential before dynode j, k is a material dependent constant defining the slope of the gain, a is a normalisation constant and l is the number of dynodes in the tube. In the simulation the values of the dV_j were taken from the experimental conditions of the MaPMT signal spectra, where they are determined by the applied high voltage and the resistive voltage divider chain. The values for k and a were fitted to data or variations of these results. Then the gain at the first dynode was varied as defined in assumption 10.

3 Methods for Signal Loss Determination

In this section different methods for measuring the signal loss below a threshold cut are defined and discussed. The implementation of the algorithms is documented in section 7 (Appendix B).

Consider signal spectra like in figures 1 and 2 which have been fitted with the Poisson based algorithm and the assumption of photoconversion at the first dynode as described in section 2.3. The global fit is drawn as well as the contributions which belong to the pedestal, the single and double photon signal from the photocathode and the single and double photoelectron contribution form the first dynode. The difference between the two spectra is the gain of the signal, i.e. how well the signal is separated from the pedestal. This determines how many signal events are lost if a threshold cut is applied to separate 'signal' contributions (S) form the 'noise' (N). A signal loss always is introduced as both contributions overlap. A threshold cut conveniently is defined as:

$$S > Q_0 + T \cdot \sigma_0 \tag{11}$$

where T is selecting the minimum separation of a 'signal' from the noise position Q_0 in terms of the noise width σ_0. For our studies mostly $T = 5.0$ was used as default.

There are three independent quantities in the fit parameters which can be used to determine signal loss:

- the probability of photoelectron production,
- the position and shape of the one-photoelectron distribution and

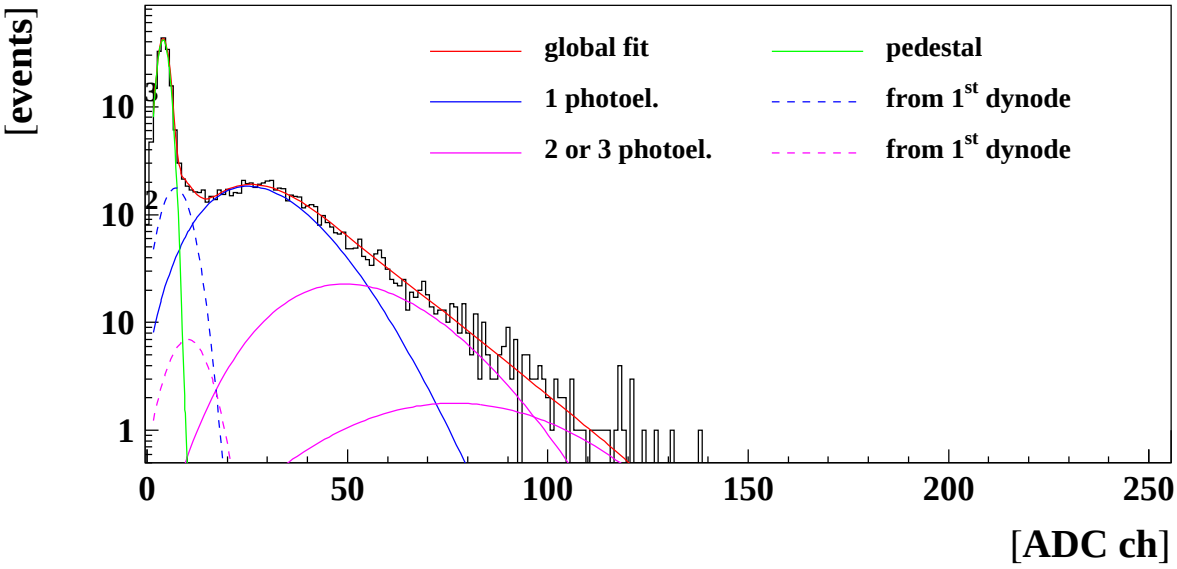

Figure 2: MaPMT signal spectrum at low gain with Poisson fit and assumed photoproduction at the first dynode, all contributions to the global fit are shown separately.

- the shape of the pedestal distribution.

In addition the first two quantities can be defined either by exclusively regarding the photoelectron production at the photocathode (PC) or by including photoelectron production at the first dynode (D) as well.

As will be seen in the following sections the methods using the three quantities above depend on different areas of the spectrum and different fit parameters. Thus, the methods are relatively independent of each other. This allows to cross-check them and possibly a choice can be made of a preferred method depending on which part of the fit seems most reliable. Table 1 gives an overview of the definitions discussed in the following sections.

ph.e. origin	ph.e. probability	1-ph.e. distribution	pedestal
PC & D	$\ell_1(\mu+\mu_1)$	$\ell_2(\int s_{1,PC}+s_{1,D})$	$\ell_3(\int p_G)$
PC	$\ell_4(\mu)$	$\ell_5(\int s_{1,PC})$	

Table 1: Overview about signal loss definitions and their dominant dependencies.

3.1 Signal Loss measured with the Photoelectron Probability

The production probability of photoelectrons $\overline{\mu}$ is used in the first method to measure the signal loss:

$$\ell\left(\overline{\mu}, T, \sigma_0\right) = \frac{\mathcal{N}_{\geq 1}(\overline{\mu}) - N_{\geq T\sigma_0}}{\mathcal{N}_{\geq 1}(\overline{\mu})} \tag{12}$$

where $\mathcal{N}_{\geq 1}(\overline{\mu})$ is the sum over the cases with emission of photoelectrons ($n = 1, 2, \ldots, \infty$) in equation 2, and $N_{\geq T\sigma_0}$ counts the number of events in the spectrum above the threshold $Q_0 + T\sigma_0$.

This method depends dominantly on the fitted production probability, $\overline{\mu}$. The signal loss also depends on the arbitrary choice of the threshold cut $T\sigma_0$. With this choice the signal loss also depends on how well the pedestal was fitted and thus how reliable the threshold is defined (see section 3.3 for a more detailed discussion on the fit of the pedestal).

The quality of this estimate of the signal loss is expected to be good if the overall number of generated photoelectrons is well described by the chosen fit method. If, however, not all contributions to the MaPMT signal are properly regarded by the fit, i.e. if the photoelectron production probability is found to low, the signal loss will be underestimated. In such a case the signal loss determined in this way even can become negative.

The actual loss definition has to state which signal contributions of the fit are regarded. For the fit algorithms used in this study either only the production of photoelectrons at the photocathode was

regarded ($\overline{\mu} = \mu$), leading to the definition ℓ_4, or the production of photoelectrons at the first dynode was included ($\overline{\mu} = \mu + \mu_1$), leading to definition ℓ_1:

$$\begin{aligned} \ell_1(\mu+\mu_1, T, \sigma_0) &= \frac{\mathcal{N}_{\geq 1}(\mu+\mu_1) - N_{\geq T\sigma_0}}{\mathcal{N}_{\geq 1}(\mu+\mu_1)} \\ \ell_4(\mu, T, \sigma_0) &= \frac{\mathcal{N}_{\geq 1}(\mu) - N_{\geq T\sigma_0}}{\mathcal{N}_{\geq 1}(\mu)} \quad . \end{aligned} \tag{13}$$

The dominant contributions to the error of these loss definitions arise from the uncertainty of the expected number of signal events $\delta\mathcal{N}_{\geq 1}(\overline{\mu})$ and the error made in the event counting $\delta N_{\geq T\sigma_0}$. The second is discussed in section 3.4. They are combined by error propagation:

$$\begin{aligned} \delta\ell_{1,4} &= \sqrt{\left(\frac{N_{\geq T\sigma_0}}{\mathcal{N}_{\geq 1}^2}\delta\mathcal{N}_{\geq 1}\right)^2 + \left(\frac{1}{\mathcal{N}_{\geq 1}}\delta N_{\geq T\sigma_0}\right)^2} \\ &\text{where} \\ \delta\mathcal{N}_{\geq 1}(\overline{\mu}) &= Ne^{-\overline{\mu}}\delta\overline{\mu} \qquad \text{and} \qquad \delta\overline{\mu} = \sqrt{\delta\mu^2 + \delta\mu_1^2} \qquad \text{or} \qquad \delta\overline{\mu} = \delta\mu \end{aligned} \tag{14}$$

depending on the chosen definition. $\delta\mu_1$ and $\delta\mu$ are the errors returned by the fit procedure for the parameters μ_1 and μ, respectively. This error estimate neglects the small contribution of the errors for the pedestal position and width which define the position of the threshold.

3.2 Signal Loss measured with the One-Photoelectron Distribution

The second handle to define signal loss is based on the integrals over the single photon contribution to the MaPMT signal:

$$\ell\left(\int s_1(x), T, \sigma_0\right) = \frac{\int_0^\infty s_1(x)dx - \int_{Q_0+T\sigma_0}^\infty s_1(x)dx}{\int_0^\infty s_1(x)dx} \tag{15}$$

where x is the pulse hight of the signal and $s_1(x)$ is the number of events assigned by the fit to the single photoelectron contribution in the bin dx. The term in the denominator counts the total number of events assigned to the single photon contribution and the second term in the nominator only starts counting above the threshold $Q_0 + T\sigma_0$. Defined in this way the signal loss is restricted to the range $[0 \ldots 1]$.

The quality of this signal loss estimate depends mainly on whether the shape and position of the single photoelectron contribution to the signal $s_1(x)$ is well represented by the fit. As before in section 3.1, the signal loss also depends on the arbitrary choice of T to define the threshold cut. And in the same way this choice depends on how well the pedestal was fitted and thus how reliable the threshold is defined. In addition this method neglects the signal loss of multi-photoelectron contributions $s_n(x)$, with $n \geq 2$. For small photoelectron multiplicities $\overline{\mu}$ such contributions to the signal loss are small and a good estimate is expected by this method. This way it provides a direct estimate of the probability of loosing a single photoelectron. If data spectra with larger photoelectron multiplicities $\overline{\mu}$ are studied signal contributions with larger n can be added.

For the fit algorithms used in this study either only the production of photoelectrons at the photocathode was regarded ($s_1 = s_{1,PC}$), leading to the definition ℓ_5, or the production of photoelectrons at the first dynode was included ($s_1 = s_{1,PC} + s_{1,D}$), leading to definition ℓ_2:

$$\begin{aligned} \ell_2(\textstyle\int s_{1,PC} + s_{1,D}, T, \sigma_0) &= \frac{\int_0^\infty s_{1,PC}(x)+s_{1,D}(x)dx - \int_{Q_0+T\sigma_0}^\infty s_{1,PC}(x)+s_{1,D}(x)dx}{\int_0^\infty s_{1,PC}(x)+s_{1,D}(x)dx} \\ \ell_5(\textstyle\int s_{1,PC}, T, \sigma_0) &= \frac{\int_0^\infty s_{1,PC}(x)dx - \int_{Q_0+T\sigma_0}^\infty s_{1,PC}(x)dx}{\int_0^\infty s_{1,PC}(x)dx} \quad . \end{aligned} \tag{16}$$

The dominant contributions to the error of these loss definitions arise from the uncertainty of the weight of the single photoelectron contribution δI_s and the error introduced by the definition of the

threshold cut $\delta I_{s\geq T\sigma_0}$. The errors are combined by error propagation:

$$\begin{aligned}
\delta\ell_{2,5} &= \sqrt{\left(\tfrac{I_{s\geq T\sigma_0}}{I_s^2}\delta I_s\right)^2 + \left(\tfrac{1}{I_s}\delta I_{s\geq T\sigma_0}\right)^2} \\
&\text{where} \\
I_s &= \textstyle\int_0^\infty s_1 dx \\
I_{s\geq T\sigma_0} &= \textstyle\int_{Q_0+T\sigma_0}^\infty s_1 dx \\
\delta I_s &= N(1-e^{-\overline{\mu}})\sqrt{\left(\tfrac{\delta\overline{\mu}}{\overline{\mu}}\right)^2 + \left(\tfrac{\delta\overline{\sigma}}{\overline{\sigma}}\right)^2} \\
\delta I_{s\geq T\sigma_0} &= \delta I_s \cdot \tfrac{I_{s\geq T\sigma_0}}{I_s} \\
&\text{and} \\
\delta\overline{\mu} = \sqrt{\delta\mu^2+\delta\mu_1^2} &\text{ or } \delta\overline{\mu} = \delta\mu \\
\delta\overline{\sigma} = \delta\sigma_2 &\text{ or } \delta\overline{\sigma} = \delta\sigma_1
\end{aligned} \tag{17}$$

depending on the chosen definition. $\delta\mu_1$, $\delta\mu$, $\delta\sigma_2$ and $\delta\sigma_1$ are the errors returned by the fit procedure for the parameters μ_1, μ, σ_2 and σ_1, respectively. Choosing $\delta\overline{\sigma} = \delta\sigma_2$ in the case of ℓ_2 neglects the minor contributions of $\delta\sigma_1$ but eases the computation. The error on the uncertainty of the fitted gain δQ_1 may be neglected if the signal contributions of the spectrum are described well by the fit. If included it were to enter as a factor $(\delta Q_1/Q_1)^2$ in the expression for $\delta I_{s\geq T\sigma_0}$. The error also neglects the even smaller contribution of the errors for the pedestal position and width which define the position of the threshold.

For a Poisson based algorithm the fit of σ_1 and thus its error becomes unreliable if the average gain is to low, as discussed in section 2.2. A way out is the calculation of σ_1 from other fit parameters. If the fit algorithm is used which also regards photoconversion at the first dynode σ_1 can be obtained by using the relation of the width and gain of a Poisson distribution in a MaPMT [8]:

$$\sigma_j^2 = Q_j^2\left(\tfrac{1}{K_j} + \tfrac{1}{K_jK_{j+1}} + \cdots + \tfrac{1}{K_jK_l}\right) \tag{18}$$

with the j looping over the amplification stages of the MaPMT as in equation 10 and a generalised definition of $K_j = Q_j/Q_{j+1}$. Some arithmetic gives the following result for σ_1 and the propagated error:

$$\begin{aligned}
\sigma_1 &= \sqrt{\tfrac{Q_1}{Q_2}(\sigma_2^2+Q_2^2)} \\
\left(\tfrac{\delta\sigma_1}{\sigma_1}\right)^2 &= \tfrac{Q_2^2}{2Q_1^2(\sigma_2^2+Q_2^2)^2}\left(\left(\left(\tfrac{\sigma_2^2}{Q_2}+Q_2^2\right)\delta Q_1\right)^2 + \left(\left(Q_1 - \tfrac{Q_1\sigma_2^2}{Q_2^2}\right)\delta Q_2\right)^2 + \left(\tfrac{2Q_1\sigma_2}{Q_2}\delta\sigma_2\right)^2\right) \quad .
\end{aligned} \tag{19}$$

3.3 Signal Loss measured with the Pedestal Distribution

The final method to measure signal loss is only depending on the fit to the pedestal and the knowledge of the shape of the pedestal distribution:

$$\ell\left(\textstyle\int p(x), T\right) = \dfrac{N_{\leq T\sigma_0} - \int_0^\infty p(x)dx}{N - \int_0^\infty p(x)dx} \tag{20}$$

where x again is the pulse height of the signal, $p(x)$ is the number of events assigned by the fit to the pedestal contribution in the bin dx, N is the total number of events in the spectrum and $N_{\leq T\sigma_0}$ counts the number of events in the fitted spectrum below the threshold $Q_0 + T\sigma_0$. Thus, the denominator estimates the number of total signal events and the nominator the number of signal events below the threshold.

This method depends on how well the actual pedestal in the spectrum gets described by the pedestal function in the fit. No knowledge is needed about the shape or size of the signal content of the spectrum. As before in sections 3.1 and 3.2, the measured signal loss also depends on the arbitrary choice of T to define the threshold cut.

In this study we assume that the pedestal is of a Gaussian shape. Thus, this method to determine the signal loss is sensitive to additional noise components as non-Gaussian tails and distortions due to common mode. In the case of tails the fit would attribute them to the signal and not to the pedestal. Hence the signal loss would be overestimated. Common mode instead could have various effects depending on its source and size. If it were a random distribution it would fold with the Gaussian pedestal to effectively broaden the pedestal distribution. This would increase the effective noise of the system and thus the signal loss as it pushes the position of the threshold cut to a higher value. If the common mode instead has a structure it generally is not clear whether a Gaussian fit would under- or overestimate the

integral of the pedestal as it depends on the actual shape and magnitude of the common mode and the starting parameters of the fit. But if the structure produces two peaks or a shoulder in the pedestal distribution it is more likely that the fit would underestimate the integral over the pedestal. In this case the signal loss would be overestimated. In summary this method is expected to give in this study an upper limit on the signal loss of all contributions to the signal.

The final loss definition assumes the Gaussian shape of the pedestal distribution $p_G(x)$:

$$\ell_3(\int p_G, T) \quad = \quad \frac{N_{\leq T\sigma_0} - \int_0^\infty p_G(x)dx}{N - \int_0^\infty p_G(x)dx} \quad . \tag{21}$$

The dominant contributions to the error of this loss definition arise from the errors made in the event counting around the threshold cut $\delta N_{\leq T\sigma_0}$ and the accuracy of the description of the pedestal distribution δI_p. The first is discussed in section 3.4. They are combined by error propagation:

$$\begin{aligned} \delta\ell_5 \quad &= \quad \sqrt{\left(\tfrac{1}{N-I_p}\delta N_{\leq T\sigma_0}\right)^2 + \left(\tfrac{|N_{\leq T\sigma_0}-N|}{(N-I_p)^2}\delta I_p\right)^2} \\ &\text{where} \\ I_p \quad &= \quad \int_0^\infty p_G(x)dx \\ \delta I_p \quad &= \quad Ne^{-\overline{\mu}}\sqrt{\left(\tfrac{\delta Q_0}{Q_0}\right)^2 + \left(\tfrac{\delta\sigma_0}{\sigma_0}\right)^2} \\ &\text{and} \\ \delta\overline{\mu} \quad &= \quad \mu + \mu_1 \quad . \end{aligned} \tag{22}$$

δQ_0 and $\delta\sigma_0$ are the errors returned by the fit procedure for the parameters Q_0 and σ_0 as defined in section 2.1 equation 1. The contribution of δQ_0 is small and may be neglected.

3.4 Error of the Event Counting

For a single measurement with N events the number of events above or below a threshold cut $S > Q_0 + T\sigma_0$ has a binomial error:

$$\begin{aligned} \delta N_{\geq T\sigma_0} \quad &= \quad \delta N_{\leq T\sigma_0} = \sqrt{N \cdot p \cdot q} \\ &\text{where} \\ p = \tfrac{N_{\geq T\sigma_0}}{N} \quad &\text{and} \quad q = 1 - p = \tfrac{N_{\leq T\sigma_0}}{N} \end{aligned} \tag{23}$$

In addition the signal spectra have a finite bin size and the cut position generally falls between the two bin centres i and $i+1$ with the event counts $N(i)$ and $N(i+1)$, respectively, cf. figure 3. The cut takes place in a region were the number of events per bin can be large. This happens especially for signal spectra with low gain and a significant contribution of signal events right above threshold, e.g. from photoelectrons generated at the first dynode. Figure 2 shows such an example spectrum. A linear interpolation of the spectrum between the two bin centres, depicted as dashed line in figure 3, improves significantly the error made in the event counting above and below the threshold cut. An upper estimate to the remaining error is:

$$\delta N_{\geq T\sigma_0} = \delta N_{\leq T\sigma_0} = \left|\frac{N(i) - N(i+1)}{4}\right| \quad , \tag{24}$$

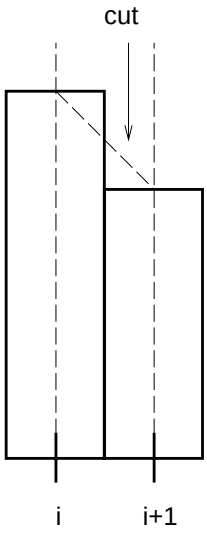

Figure 3: Possible position of threshold cut between the centres of two bins and linear interpolation of bin content.

i.e. this estimate assumes that the difference of the shape of the physical distribution and the linear interpolation is not larger than one of the triangles between the linear interpolation and the bin borders. This is a good assumption as in MaPMT spectra with small gain the bin content $N(i)$ and $N(i+1)$ are large and their associated statistical error small. In spectra with large gain the bin contents are small and the statistical error increased but the absolute error will be small.

An estimate shows that for low gain spectra the remaining error caused by the finite bin size may become of the order of the binomial error. Therefore both should be regarded and quadratically added.

4 Application to Data

4.1 Comparison of fit methods

To better understand the behaviour of the three fit algorithms (Gaussian, Poisson without and with production of photoelectrons at the first dynode as defined in the sections 2.1 to 2.3) they are applied

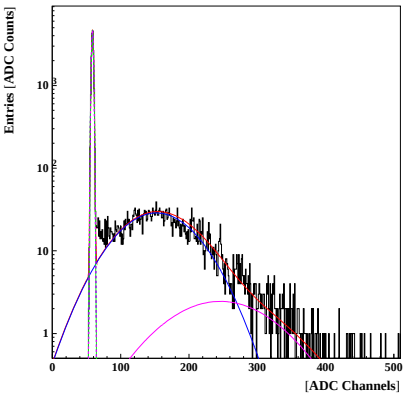

Figure 4: MaPMT spectrum of run 2041 (tube: 9C20A2, pixel: 10); with Gaussian fit.

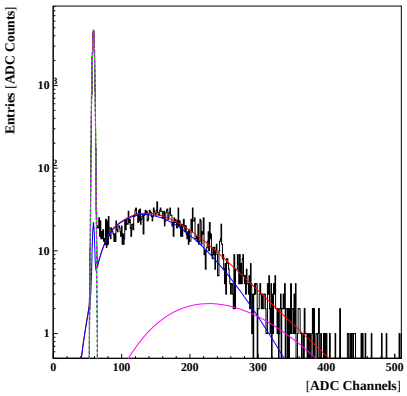

Figure 5: MaPMT spectrum of run 2041 (tube: 9C20A2, pixel: 10); with Poissonian fit.

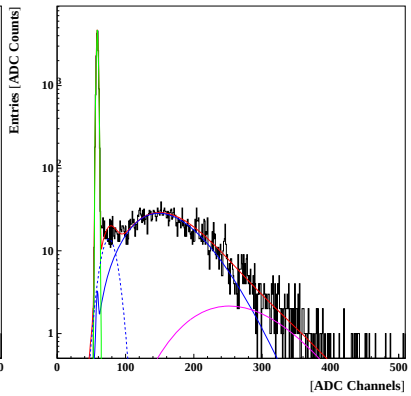

Figure 6: MaPMT spectrum of run 2041; with Poissonian fit and photoelectron conversion at the first dynode.

to the same data set and the results are compared in the following.

The data set was taken with the CAMAC readout described in [7] for a centre pixel of a MaPMT[3] operated with a default voltage divider chain at a supply voltage of $HV = -900\,V$. Figure 4 shows the spectrum of this run with a Gaussian fit superimposed. The same spectrum is shown in figures 5 and 6. There fits using the Poisson algorithm without and with first dynode effect are superimposed. The fit results displayed in the figures 4 to 6 are summarised in table 2. The errors given are those returned by the fit or determined using error propagation if applicable.

algorithm	Gauss			Poisson			Poisson & 1^{st} dynode		
μ	0.241	$\pm$	0.004	0.240	$\pm$	0.004	0.215	$\pm$	0.003
μ_1							0.024	$\pm$	0.002
Q_1	93.6	$\pm$	0.9	94.2	$\pm$	1.0	103.2	$\pm$	0.9
σ_1	52.6	$\pm$	0.8	75.2	$\pm$	19.	71.7	$\pm$	5.6
Q_2				22.8	$\pm$	1.1	16.9	$\pm$	0.7
σ_2				15.7	$\pm$	0.5	10.6	$\pm$	0.5
$K_1 = Q_1/Q_2$				4.1	$\pm$	0.2	6.1	$\pm$	0.3
$g_1 \geq Q_1^2/\sigma_1^2$	3.17	$\pm$	0.06						

Table 2: Parameters returned by the fit methods on the MaPMT spectrum of run 2041.

The production probabilities of photoelectrons found by the three algorithms, μ and $\mu + \mu_1$, match very well. As will be seen at the end of this section in average the three algorithms exhibit a small systematic difference in the measurement of $\overline{\mu}$. But this example provides a convenient common base for the following comparison of the three algorithms.

The Gaussian algorithm, displayed in figure 4, completely fails to describe the signal spectrum below the pedestal. In addition the signal is strongly underestimated by the fit in the region between the pedestal and the single photoelectron peak, between 68 and 100 ADC channels. Signal contributions which fill this region are not accounted for in the Gaussian fit model. Their presence in the data tends to pull the measured gain of the single photoelectron distribution, Q_1, to a lower value.

By using the Poisson algorithm, as shown in figure 5, the description of the region near the pedestal is significantly improved. This is due to the more physical description of the electron multiplication at the first dynode by a Poisson distribution which dominates the shape of the full pulse height spectrum. However, the data between the pedestal and the single photoelectron peak, between 68 and 130 ADC channels, are still not correctly described by the fit. The mean single photoelectron gain, Q_1, is pulled towards a low value as, again, there are no accounts for signal contributions which fill this region. The gain of the single photoelectron distribution, Q_1, is found to be the same value as for the Gaussian fit method within errors.

By allowing the production of photoelectrons at the first dynode as an additional contribution to the signal in the fit the overall description of the data improves considerably, as shown in figure 6. Especially in the region of the valley between the pedestal and the single photoelectron peak the data

[3] Tube: 9C20A2, pixel: 10, run no.: 2041

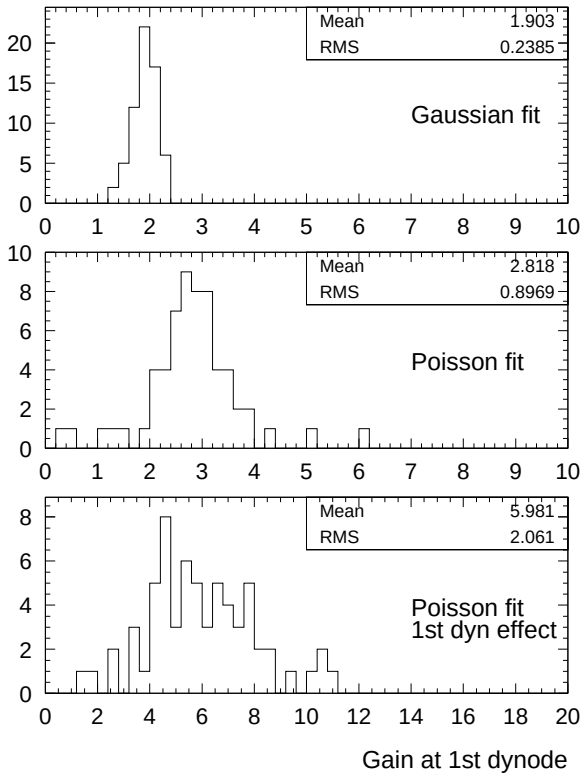

Figure 7: Derived gain at the first dynode g_1 and K_1 for the signal spectra of the 64 channels of tube 9K20C3 taken with the default divider chain at $HV = -900\,V$ and analysed with the three fit methods.

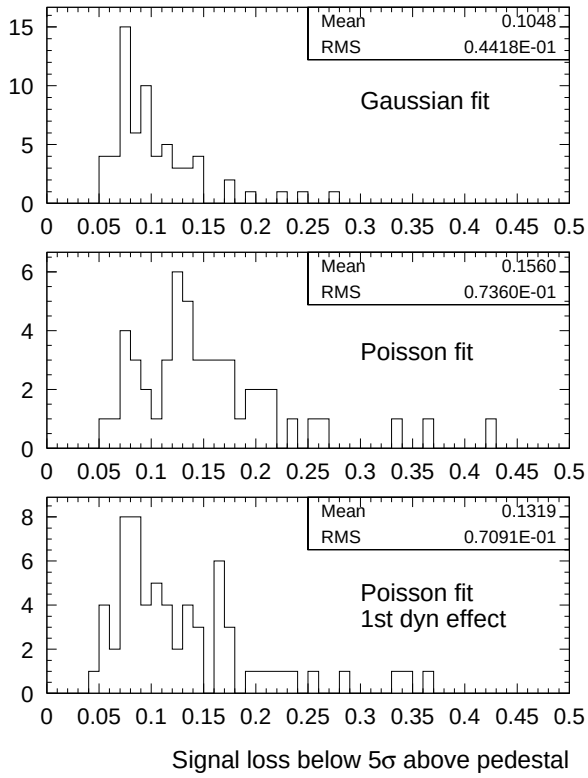

Figure 8: Derived signal loss ℓ_1 for the signal spectra of the 64 channels of tube 9K20C3 taken with the default divider chain at $HV = -900\,V$ and analysed with the three fit methods.

is now reasonably well described by the fit. The gain Q_1 is measured about 10 % larger than for the other two methods, but it appears more reliable as the shape of the pulse height spectrum is now better described. A closer look at the region between 68 and 100 ADC channels reveals that the shapes of the data and the fit still do not match perfectly. This is further studied in section 4.2 with a simulation of an additional inhomogeneity of the first dynode. The measurement of the signal width, σ_1, matches for the two Poisson algorithms. But the increased errors indicate that the determination of σ_1 becomes unreliable as for these algorithms it is determined from the width of the 3-photoelectron peak. In figures 5 and 6 this peak would be found at around 345 and 370 ADC channels, respectively, where the spectrum lacks data.

The signal contribution of electrons emerging off the first dynode can be measured in the two algorithms based on Poisson distributions. When the additional photoconversion at the first dynode is included the gain and width, Q_2 and σ_2, are measured with smaller values. In this case the signal has two overlapping components and Q_1 and Q_2 are anti-correlated. The measured gain at the first dynode, K_1, is about 50 % larger. In addition the width σ_2 tends to be smaller than for the case of only one signal contribution emerging off the photocathode.

The measured value of the gain $K_1 \approx 6$ using the Poisson fit with the production of photoelectrons at the first dynode matches reasonably well the expectation of $K_1 \approx 7$ at $HV = -900\,V$ from the gain model, described in section 2.4. A compilation of results using this model can be found in [7]. The low gain value of $K_1 \approx 4$ obtained from the Poisson fit is not compatible with expectation. From the Gaussian fits the gain is estimated by $g_1 \geq Q_1^2/\sigma_1^2$. As the Gaussian algorithm tends to overestimate the width of the single photoelectron contribution, especially if there is an additional contribution to the signal, the estimator g_1 is pulled to low values. Thus, it is a lower limit of the actual gain at the first dynode and the measured value of $g_1 \approx 3$ reflects the failure of the Gaussian fit to describe the shape of the pulse height spectrum.

These results are confirmed when all 64 pixels of a tube are studied. In Figure 7 we show the distributions found for the gains K_1 and g_1 for all pixels of a MaPMT taken at the same conditions as for figures 4 to 6. The average value of $g_1 \approx 2$ demonstrates that for the Gaussian algorithm the effect of broadening the signal contribution and pulling the gain to a lower value is even stronger than in the

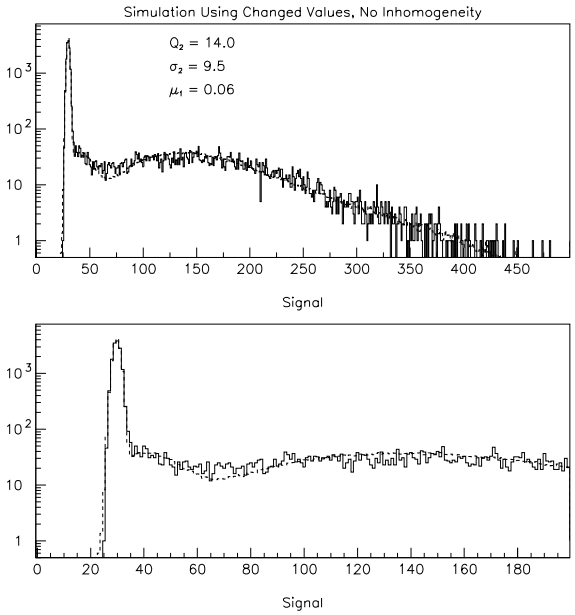

Figure 9: MaPMT spectrum of run 2169 (tube: 9K20C3, pixel: 10); simulation with Poisson distributions and photoelectron conversion at the first dynode (dashed line); top: full spectrum, bottom: zoom into the critical region.

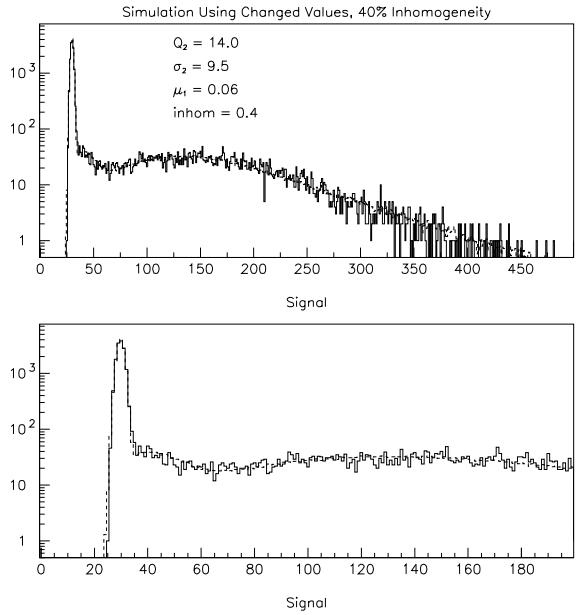

Figure 10: As figure 9 but now including an inhomogeneity of $\mathcal{I} = 0.4$ in the electron multiplication probability across the surface of the first dynode (dashed line).

example of figure 4. The gain K_1 measured for the two methods based on Poisson distributions differs in average by a factor of two ($K_1 \approx 3$ and $K_1 \approx 6$, respectively).

When studying the signal loss ℓ_1 for the same 64 signal spectra, we find a pattern which is consistent with the characteristics of the three fit methods. For the Gaussian fit method the average signal loss is measured to be $\ell_1 = 10.5\,\%$. To be added is the signal loss due to non-multiplication of electrons at the first dynode. It is $0.25\,\%$ at $K_1 = 6$. This leads to the total signal loss $\ell_1 = 10.7\,\%$ which can be compared with the other methods. For the two methods based on Poisson distributions the signal loss due to non-multiplication at the first dynode is an intrinsic part of the loss. The Poisson methods without and with the production of photoelectrons at the first dynode give an average loss $\ell_1 = 15.6\,\%$ and $\ell_1 = 13.2\,\%$, respectively. The definition of ℓ_1 is based on the photoelectron production probability μ. Therefore it is not dependent on how well the detailed shape of the signal spectrum is fitted. The Gaussian method tends to exclude the additional entries right above the pedestal and thus, it is likely to underestimate μ and ℓ_1. As visible in figure 5 the Poisson method likely overestimates the signal content in the region between the pedestal and the single photoelectron peak and hence, μ and ℓ_1 are found too high. The Poisson fit which includes the production of photoelectrons at the first dynode gives the best estimate of the signal content. Therefore it provides the most reliable estimate of the signal loss ℓ_1. A detailed discussion comparing the methods of loss determination is given in section 4.3.

4.2 Simulation results

The Poisson fit including the production of photoelectrons at the first dynode still does not match perfectly the data in the region between the pedestal and the single photoelectron peak, cf. figure 6. To test the hypothesis of an inhomogeneity in the electron multiplication probability across the surface of the first dynode simulations of the signal spectra were performed. The simulation is based on Poisson distributions including the production of photoelectrons at the first dynode. The inhomogeneity is added as described in section 2.4, equation 9. Input parameters to the simulation were the parameters returned by a fit to the data. A variety of values for the inhomogeneity was tested. Each run contained 10^6 simulated events and was normalised to the number of events in the data spectrum.

In figure 9 a MaPMT signal spectrum is shown superimposed with the simulation with no inhomogeneity ($\mathcal{I} = 0.0$). The statistical error of the simulation is negligible. The simulation closely models the fit on which parameters it was based on. A mismatch between data and simulation is visible similar to figure 6.

The inhomogeneity was increased in steps of $\Delta\mathcal{I} = 0.1$ and the result for $\mathcal{I} = 0.4$ is shown in figure 10. It represents the best description of the data. As expected the inhomogeneity smears out the contribu-

tions to the MaPMT signal. The underestimated regions get filled up and the overestimated regions are reduced. The detailed match between data and simulation is much improved and the simulation now matches the shape of the full data spectrum. At larger values of $\mathcal{I}$ the simulation tends to underestimate the number of events in the signal peak region and to overestimate the high gain tail of the spectrum as the single photoelectron distribution becomes to broad. The match between simulation and data is achieved with a large inhomogeneity of $\mathcal{I} = 0.4$. This translates into a factor 2.3 difference in the electron multiplication probability across the surface of the first dynode. In contrast in reference [8] an inhomogeneity of $\mathcal{I} \approx 0.2$ was reported to give best results. Although it is not documented there how this inhomogeneity was implemented into the simulations one can assume that the implementation is similar to the one presented in section 2.4.

4.3 Behaviour of signal loss

Method and Data sets: To compare and discuss the signal losses $\ell_1, \ldots, \ell_5$ introduced in the sections 3.1 to 3.3 they have been measured for the same series of data as for the study presented in reference [7]. The full details on the experimental conditions as well as on the used data acquisition hardware can be found there. For the discussion here it is of interest that most parameters, including the amount of light shown onto the photodetectors, were kept stable. Only the supply voltage to the voltage divider chain, i.e. the total gain of a tube, was varied. The study in reference [7] compared three different kinds of data acquisition electronics: a CAMAC based system, a system based on the APVm chip where the signals of the 12-stage MaPMTs had to be attenuated in front of the chip to match the dynamic range and a system where the MaPMT signals were put into the APVm chip directly but the MaPMT were operated at a very low total gain. Each of these three DAQ systems was operated with three settings of the voltage divider chain giving a high, medium or low intrinsic gain to the MaPMT at a given supply voltage. Thus, there are nine sets of data available for each tested MaPMT channel. The signal spectra were fitted with the algorithm based on Poisson distributions including the production of photoelectrons at the first dynode, as defined in section 2.3. Then the values for $\ell_1, \ldots, \ell_5$ were calculated based on the fit results.

While conducting the study for reference [7] it was recognised that the probability of photoelectrons $\overline{\mu}$ is a sensitive fit parameter to judge the quality of the fit. In the chosen experimental conditions the total gain of the tubes is varied and the light level applied to the photodetector is kept constant. Reliable fits are indicated if for such a series of signal spectra a constant probability of photoelectrons is measured. Figure 11 shows the photoelectron production probabilities (label $\lambda = \overline{\mu}$) for the data sets of interest here. For the data points shown the fits to the spectra were visually inspected and regarded as sufficiently reliable for a further analysis of the signal loss. In each run series the photoelectron production probability starts to rise significantly towards low HV supply values. This starts at the point where the gain in the MaPMT becomes to low for a good separation of the signal from the pedestal. Such data is excluded from the selections presented here.

General Behaviour: In figures 12 to 15 selected data sets are shown for the signal loss definitions $\ell_1 \ldots \ell_5$. Note that in these displays data points are suppressed if their symbols would cross the frame. So small values, below $\approx 2.5\,\%$, are not shown, as are negative loss values which may occur for the definitions ℓ_1 and ℓ_4.

For a comparison of options of data acquisition hardware like in reference [7] the signal loss has to be plotted against the ratio of signal-over-noise to determine the relative advantages. The signal-over-noise achieved in the particular system defines where the threshold cut is applied relative to the average pulse height. An example of such a comparison based on the signal loss definition ℓ_1 is shown in figure 12. The figure exhibits clearly the different levels of signal loss achieved by the three readout options. For the following comparison of the properties of the loss definitions it is of interest to compare what results are obtained in exactly the same signal spectra. For that the signal loss data of figure 12 have to be viewed against the setting of the supply voltage, i.e. the gain of the tube, like in figure 13.

In figures 14 and 15 the signal loss values $\ell_1 \ldots \ell_5$ are shown for two selected, typical data sets. The filled symbols denote the losses ℓ_1, ℓ_2 and ℓ_3 which estimate the full signal loss. The open symbols mark the losses ℓ_4 and ℓ_5 which estimate the loss of signal events which have been produced by photoconversion at the photocathode only. Generally the signal loss increases when the gain is lowered and more signal events fall below the threshold cut. The gain can be lowered by reducing the supply voltage, as applied here, or by changing the ratios of the resistor chain distributing the supply voltage to the dynodes of the MaPMT.

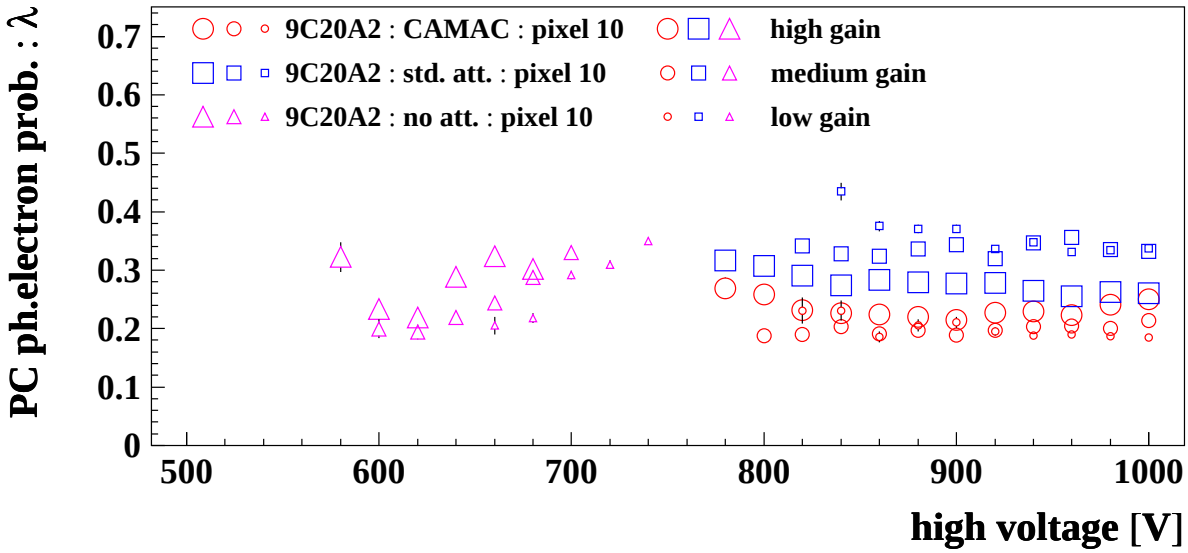

Figure 11: Control plot with the determined number of photoelectrons generated at the photo cathode of tube 9C20A2, pixel 10, for the accepted data points data sets using three different data acquisition systems with three gain options each, from [7].

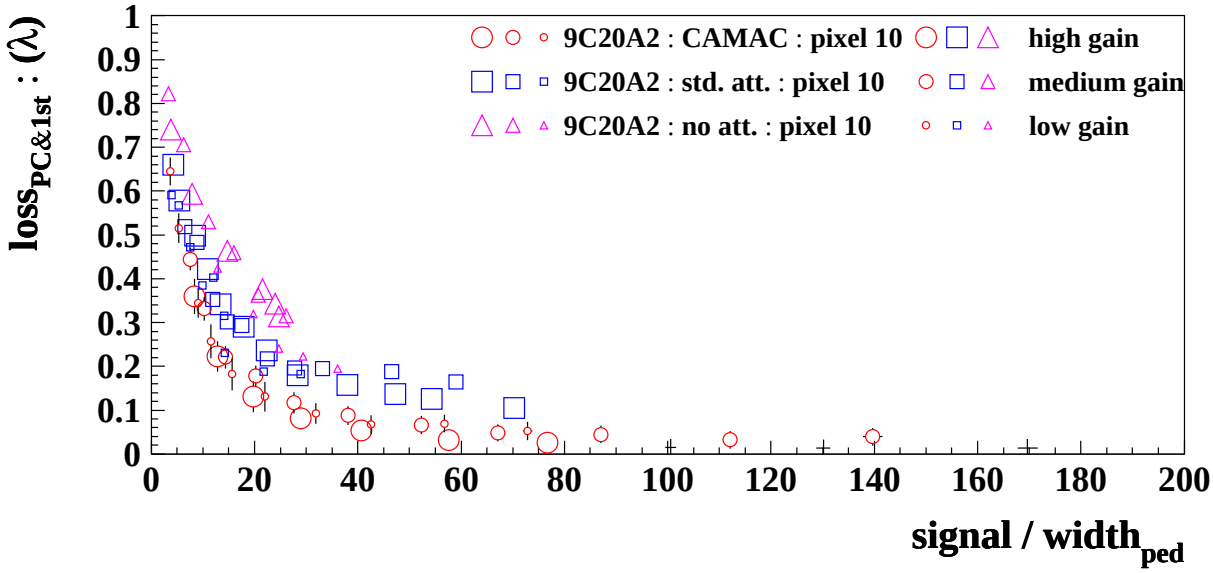

Figure 12: Signal loss ℓ_1 of the MaPMT signal due to a threshold cut $S > Q_0 + 5\sigma_0$ for the data samples of figure 11 here plotted against signal-over-noise for comparison of the readout options.

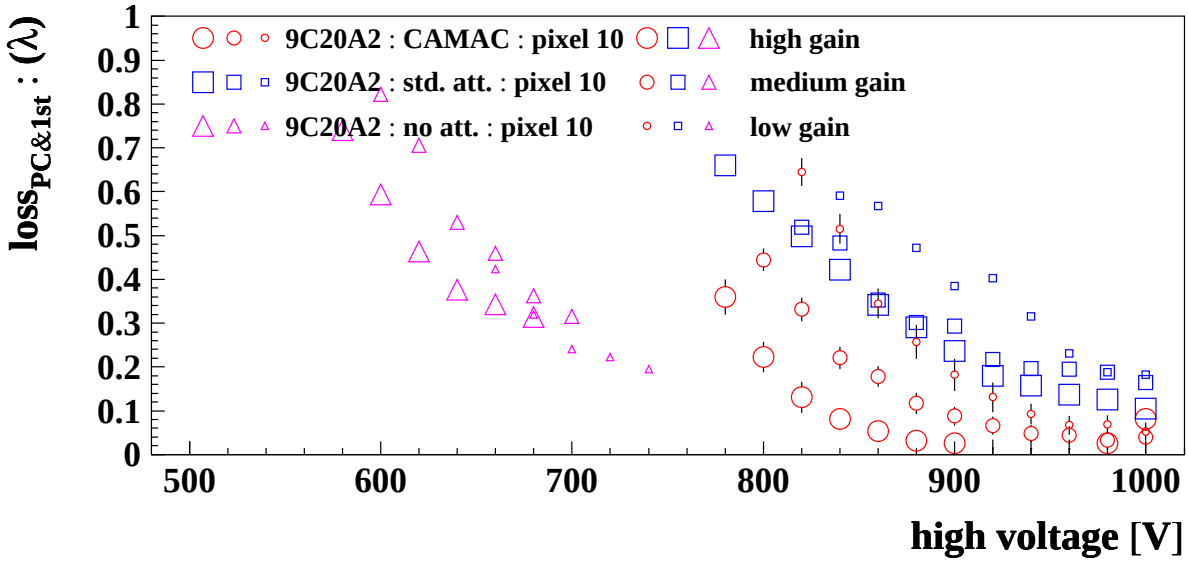

Figure 13: Same data as in figure 12 but plotted against the supply voltage (gain) of the tube.

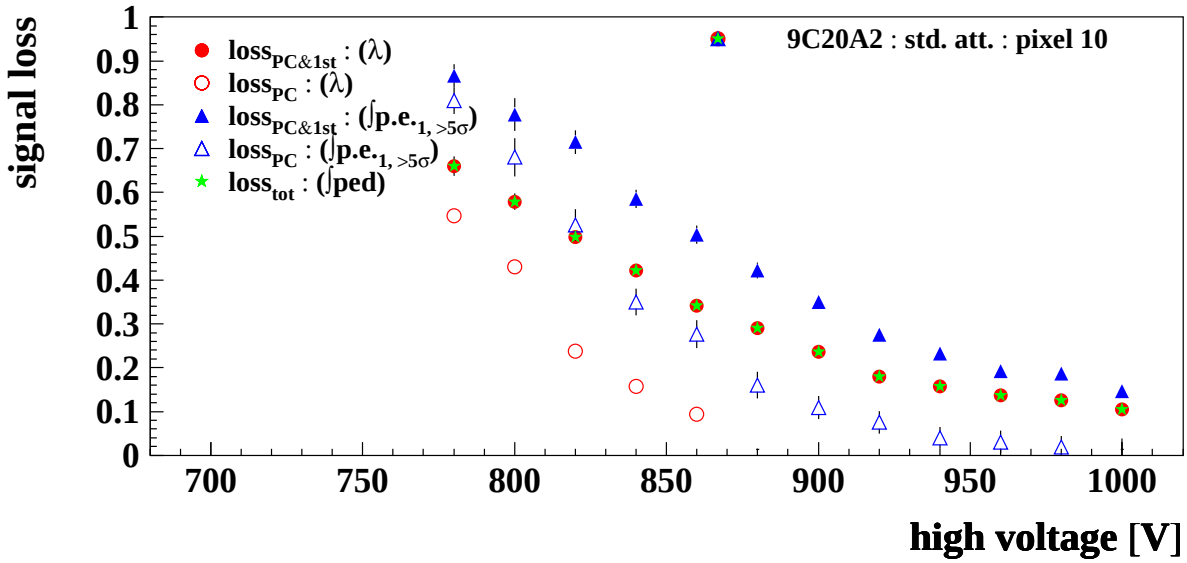

Figure 14: Signal losses $\ell_1 \ldots \ell_5$ of the MaPMT signal due to a threshold cut $S > Q_0 + 5\sigma_0$ for a selected data set of the data samples of figure 11: tube 9C20A2, pixel 10, high gain resistor chain, standard attenuation APVm readout.

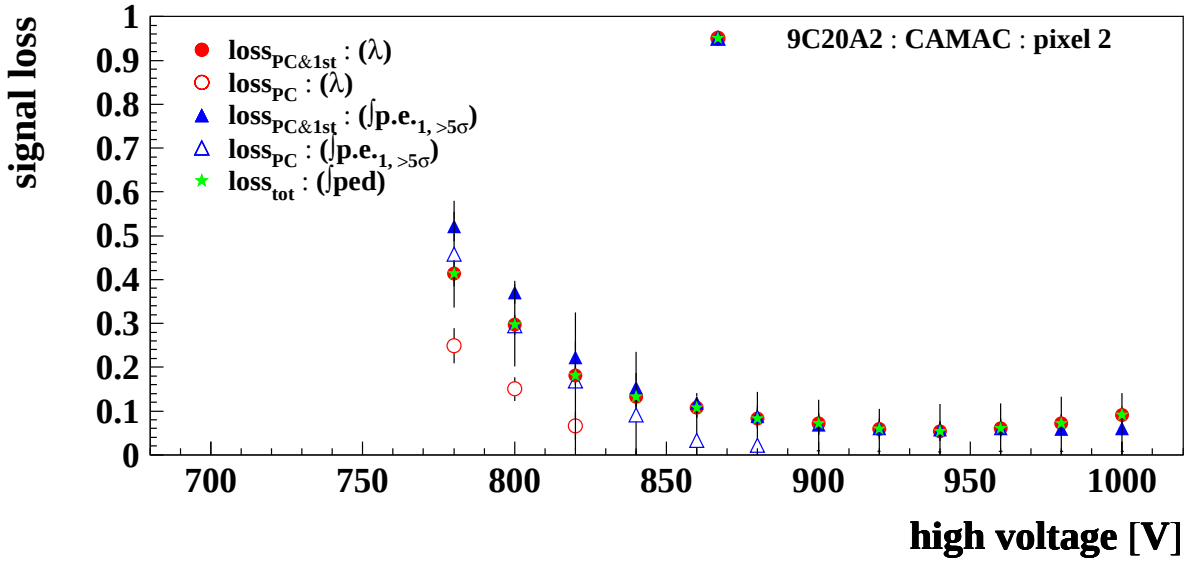

Figure 15: Signal losses $\ell_1 \ldots \ell_5$ of the MaPMT signal due to a threshold cut $S > Q_0 + 5\sigma_0$ for a selected data set of the data samples of figure 11: tube 9C20A2, pixel 2, high gain resistor chain, CAMAC readout.

For the APVm data set in figure 14 a ratio of $\mu_1/\mu \sim 0.5$ is found for the signal events which are described by photoelectrons emerging off the first dynode. These signals have a pulse height which is lower by the factor $k_1 = Q_1/Q_2$ compared to signals with electron multiplication at the first dynode. Therefore these signals are much more likely to get lost below the threshold cut. At high gain, above 900 V in figure 14, this loss dominates the total signal loss. The contributions to the signal with a larger pulse height are very small as visible in the loss found for ℓ_4 and ℓ_5 shown with the open symbols. The more the gain is lowered the more significant becomes the loss of photoelectrons emerged off the photocathode while the loss of the signal contribution with low pulse hight quickly approaches its maximum.

The definitions ℓ_1 (filled circles) and ℓ_3 (stars) measure virtually the same total signal loss and confirm each other. This is due to the good determination of the photoelectron production probability $\overline{\mu}$ in the fit and the reliable determination of the pedestal. The definition ℓ_2 (filled triangles) overestimates the total signal loss by about 30 %. This is caused by the large number of events between the pedestal and the single photoelectron peak in the data taken with the APVm readout. As these events are not caused

by photoconversion at the first dynode but rather by events sampled on the flanks of the signal pulse the shape of the signal spectrum is not described correctly by the model. In this case the fit has the tendency to find the position of the signal distributions to low and their width too wide. Thus the signal loss depending on the shape of the signal distributions gets overestimated.

For the CAMAC data set in figure 15 a ratio of $\mu_1/\mu \sim 0.1$ is found. The problem with events sampled on the flanks of the signal pulse is significantly reduced. As a consequence all estimates of the total signal loss, ℓ_1, ℓ_2 and ℓ_3, match well over a wide range in gain. Only towards low gain, below 840 V, the algorithm depending on the shape of the signal distributions exhibits again the tendency to overestimate the signal loss. At these gains the signal-over-noise becomes small and the valley between pedestal and single photoelectron in the signal spectra disappears. In this case the fit again shows the tendency to find the position of the signal distributions to low and their width too wide. Although the total signal loss estimates ℓ_1 and ℓ_2 match well over a wide range in gain the estimates for the signal loss of photoelectrons produced at the photocathode, ℓ_4 and ℓ_5, do not match. This means that the set of parameters found by the fit for the contribution made by photoconversion at the first dynode, μ_1 and its position and width, is not consistent with the assumption of the production of photoelectrons at the first dynode. In the case of the CAMAC data, figure 15, a small fraction of the events also is sampled incompletely as the arrival time is shifted and the integration of the charge pulse is cut off by the edge of the gate window. Any real photoconversion at the first dynode is overshadowed by this problem.

Signal Loss measured with the Photoelectron Probability: The two definitions ℓ_1 and ℓ_4 are based on the found probability of photoelectrons $\overline{\mu}$ which basically is constant over the run series. Thus, the found signal loss entirely is depending on the number of signal events found above the threshold cut, i.e. the gain at which the tube is operated.

The production of more than one photoelectron is regarded in the definitions, cf. discussion of equation 12. The average number of photoelectrons produced in the presented data is $\overline{\mu} \approx 0.3$. So there is a chance of $\approx 3.5\,\%$ producing two or more photoelectrons. This is a small effect but the chance increases rapidly to $\approx 10\,\%$ and $\approx 15\,\%$ for $\overline{\mu} = 0.5$ and $\overline{\mu} = 0.7$, respectively. Multi-photoelectron signals produce in average a larger pulse height than single photoelectron signals and thus are less likely to get lost below the threshold cut. If the fit describes the shape of the spectrum well the average number of photoelectrons will be determined correctly including the multi-photoelectron events. Therefore the definition ℓ_1 is expected to give a good estimate of the total signal loss for all light levels provided the fit describes well the shape of the spectrum as it is the case for our data.

Even if the production mechanism for the excess of the low gain signals in our data is not correctly modelled ℓ_1 remains a good estimate of the total signal loss since the overall shape of the signal spectrum is described well. Contrary to that the definition ℓ_4 gives for our data a too low estimate of the signal loss of photoelectrons produced at the photocathode. The extra component in the signal spectrum, represented by μ_1, is not regarded. Thus, the number of expected events is scaled down by $\approx \frac{\mu}{\mu+\mu_1}$. At large gain where more signal components can be separated from the pedestal this number may become smaller than the number of events measured above the threshold cut. This results in a negative loss estimate like for the data with high supply voltage in the figures 14 and 15. The signal loss definition ℓ_4 therefore is a lower limit to the signal loss of photoelectrons produced at the photocathode. It will approximate ℓ_1 if the other contributions to the signal spectrum not stemming from photoelectrons produced at the photocathode vanish or fall below the threshold cut.

Signal Loss measured with the Pedestal Distribution: The signal loss definition ℓ_3 is solely based on the Gaussian fit of the pedestal region. No knowledge of the shape of the signal spectrum is required. Even the fit to the signal region might fail dramatically to describe the shape of the spectrum. This method only requires that the pedestal of the spectrum actually has a Gaussian shape, i.e. no other noise contributions produce distortions or tails (cf. section 3.3 for a discussion of the types and effects of additional noise contributions). In general it is more likely that the total signal loss is overestimated with this method rather than underestimated, especially if an ill-shaped common mode can be avoided. The accuracy of this method can be evaluated by looking at the shape and quality of the pedestals in dedicated pedestal runs. Therefore the definition ℓ_3 is expected to give a reliable estimate of the total signal loss for all signal spectra provided the pedestal is sufficiently Gaussian shaped and constant in time.

In our data the pedestal runs to the presented data show no significant deviation from the expected Gaussian shape. The signal loss found with definition ℓ_1 and ℓ_3 matches very well: $\ell_1/\ell_3 = 1.00 \pm 0.02$. This confirms the total signal loss can be determined using the definition ℓ_1 even if the region of small gain only is described in a phenomenological way.

Signal Loss measured with the One-photoelectron Distribution: The two definitions ℓ_2 and ℓ_5 calculate the integrals over the found contributions to the signal spectrum. Therefore these definitions are highly sensitive to the found mean position and widths of the contributions. If the fit determines a slightly lower gain or a slightly larger width of a contribution to the spectrum the signal loss below the threshold cut increases significantly.

In the presented data the contribution from multi-photoelectron production is small and thus the corresponding signal loss is negligible. The shape of the one-photoelectron signal produced at the photocathode is bound by the falling tail of the signal spectrum towards the high gains. But towards low gain its position and width depends on the description of the signal contributions filling the valley between the pedestal and the single photoelectron peak. If these contributions are not well described the fit shows the tendency to pull the one-photoelectron distribution towards lower gain and broader width. This effect becomes very strong when the signal spectrum is fitted with the simpler Poisson based algorithm which omits the production of photoelectrons at the first dynode, see figure 5. By regarding this photoelectron production process in the fit the pull on the single photoelectron distribution is much reduced. Nevertheless, in our data the true shape of the low gain contribution is not well described by the model and an unquantified pull on the single photoelectron peak remains. As a consequence the corresponding signal loss ℓ_5 has to be regarded as an upper limit of the signal loss of single photoelectrons produced at the photocathode.

In our data the distribution of low gain signals has a broader shape than assumed by the model of photoelectron production at the first dynode. In this case the fit shows the tendency to pull the contribution from the first dynode towards the pedestal and in turn to increase the pull on the single photoelectron distribution produced at the photocathode. The definition ℓ_2 which regards the signal loss of both therefore overestimates the total signal loss as the loss of both contributions is assigned to high.

In our data the following typical ratios are found: $\ell_2/\ell_1 = \mathcal{O}(1.15)$ for CAMAC data and $\ell_2/\ell_1 = \mathcal{O}(1.5)$ for APVm data. This matches the experimental conditions: the APVm data were much more affected by low gain signals due to incomplete signal sampling. In addition at high gains the following typical ratios are found: $\ell_5/\ell_1 = \mathcal{O}(0.30)$ for CAMAC data and $\ell_5/\ell_1 = 0.15 \ldots 0.45$ for APVm data. At low gains this ratio is distorted to reach values of $\ell_5/\ell_1 > 1$ as ℓ_5 gets grossly overestimated. The CAMAC data is less affected by incompletely sampled signals and thus the more reliable data set. Therefore the value $\ell_5/\ell_1 = \mathcal{O}(0.30)$ gives an upper limit of the loss of single photoelectrons produced at the photocathode in our data. This indicates what level of signal loss could be achieved with a readout not affected by incompletely sampled signals.

5 Conclusion

The signal response of Multianode Photo Multiplier Tubes (MaPMT) has been studied using three methods to fit the signal spectra: one method based on Gaussian distributions, a second method based on Poissonian distributions and a third also regarding a possible production of photoelectrons at the first dynode. The three algorithms were motivated, defined and an implementation was given. The characteristics of the three definitions were discussed by the means of their application to example data spectra. The algorithm based on Poissonian distributions and regarding a production of photoelectrons at the first dynode is able to describe well the shape of the signal spectra in a phenomenological way, even though the low gain signals were not produced at the first dynode but stemmed from incomplete signal sampling.

In addition the MaPMT spectra were simulated to study a possible inhomogeneity in the electron multiplication probability across the surface of the first dynode. In the application to an example signal spectrum it was shown how conclusions could be drawn on an inhomogeneity of the first dynode in principle. But the contamination from incompletely sampled signals made the signals of photoelectrons emerged off the first dynode inaccessible.

Three definitions of signal loss were discussed. As they base on different parameters they are fairly independent of each other: one is based on the found photoelectron production probability, a second is based on the shape and position of the found signal contributions and the last is only depending on the fit of the pedestal. The first two definitions can be varied to include or exclude a contribution of low gain signals which are assumed to stem from photoelectrons emerged off the first dynode. In the application to example data sets it was shown that the total signal loss derived from the photoelectron production probability matches very well the loss derived from the fit to the pedestal. This is due to the fact that the shape of the signal spectrum and the pedestal were fitted well.

6 Appendix A

In this section the used Fortran routines for the fitting of the MaPMT signal spectra are documented. The used parameters have the following meaning:

- PAW vector *par(9)*: holds the fit parameters
  - *norm*: the overall normalisation factor; in our fits usually bound to the number of events in the spectrum $N \pm 0.1$ to ease the fit; but this can be an indicator of the quality of the fit when left free
  - *average*: average number of photoelectrons per event for the contribution from the photocathode, μ, using Poisson statistics
  - *pedestal*: pedestal position, Q_0
  - *width_ed*: pedestal width, σ_0, i.e. the electronic noise
  - *signal*: average size above pedestal of a single photoelectron signal, Q_1, with the electron generated at the photocathode
  - *width_sig*: width of the single photoelectron signal, σ_1 for electrons produced at the photocathode; this parameter only returns reliable values if the signal spectrum covers the region where Gaussian statistics is applicable for the electron multiplication, i.e. for $n \geq 3$ in the routines using the Poisson statistics
  - *fstdynavg*: average number of photoelectrons per event for the contribution from the first dynode, μ_1, using Poisson statistics
  - *firstdyn*: average size above pedestal of a single photoelectron signal, Q_2, with the electron generated at the first dynode; it is also the size of the signal fraction of secondary electrons produced at the first dynode
  - *width_dyn*: width of the single photoelectron signal, σ_2 for electrons produced at the first dynode
- PAW vector *fitpar2(6)*: holds parameters to steer the fit behaviour
  1. in the signal regions where the fit is disabled (see *fitlimits2(4)*): (*0*) let the function return the value *0.0* (default), (*1*) let the function return the value of the fitted histogram (from *hist1*)
  2. (*0*) χ^2-fit, (*1*) maximum-likelihood-fit (default); this parameter actually was used externally in the call of the fitting function
  3. minimum number of photoelectrons included in the fit of those generated at the photocathode
  4. maximum number of photoelectrons included in the fit of those generated at the photocathode
  5. minimum number of photoelectrons included in the fit of those generated at the first dynode
  6. maximum number of photoelectrons included in the fit of those generated at the first dynode
- PAW vector *fitlimits2(4)*: holds the limits to the signal range (i.e. ADC counts) where the fitting procedure is carried out; this allows for a limitation of the used range at the bottom and the top end of the signal spectrum as well as for an intermediate range which is excluded from the fit.
- PAW vector *hist1(256)* or *hist1(1024)*: holds the channel content of the spectrum to be fitted (note: the channel with *0 ADC* counts is stored in the histogram channel *1* due to the definition of the histogram in Fortran; this is the reason for the definition of *ixp1*).

In order to enable the fit to converge properly it is recommended to generate a set of starting parameters in the following way:

- *norm*: number of events in the histogram
- *average*: approximation of expected average number of photoelectrons per event originating from the photocathode depending on the amount of light
- *pedestal*: histogram channel with highest number of entries, provided events with $n = 0$ dominate
- *width_ed*: approximation of typical pedestal width

- *signal*: mean of the histogram where the pedestal has been removed (e.g. by a $5 \times \sigma$-cut above the pedestal)
- *width_sig*: RMS of the histogram where the pedestal has been removed
- *fstdynavg*: approximation of expected average number of photoelectrons per event originating from the first dynode depending on the expected ratio to *average*
- *firstdyn*: expected location above pedestal of this contribution to the total signal spectrum, scaled as $Q_2 = Q_1/K_1$, the signal size over the expected gain at the first dynode
- *width_dyn*: scaled as $\sigma_2 = \sigma_1/K_1$, which seems close enough to converge safely

6.1 Gauss based algorithm

```
      FUNCTION phe_gauss(x)

*  Fitting Gauss functions to photoelectron spectra from a MaPMT

      Double Precision norm,average,pedestal,width_ped,signal,width_sig
      Integer fp1,fp2,fp3,fp4
      Integer ixp1, ifirst, j, n
      Real fact(20), facttemp, factn, area

      data ifirst/1/
      common/pawpar/par(9)
      vector fitpar2(6)
      vector fitlimits2(4)
      vector hist1

      norm      = par(1)   ! normalisation factor
      average   = par(2)   ! average number of p.e. per event
      pedestal  = par(3)   ! position of pedestal
      width_ped = par(4)   ! width of pedestal
      signal    = par(5)   ! offset from pedestal to single p.e. signal
      width_sig = par(6)   ! width of single p.e. signal

*  Loop over number of p.e. (0 for pedestal)

      fp1 = int(fitpar2(1))
      fp2 = int(fitpar2(2))
      fp3 = int(fitpar2(3))
      fp4 = int(fitpar2(4))
      ixp1 = int(x+1)

      phe_gauss = 0.

*  Calculate vector of factorials

      if ( ifirst.eq.1 ) then
         facttemp = 1
         fact(1) = 1
         do j = 1,19
            facttemp = facttemp*REAL(j)
            fact(j+1) = facttemp
         enddo
         ifirst = 0
      endif

*  Check for valid fit region and loop over photoelectrons

      if ( x.lt.fitlimits2(1) .or.
     &     x.gt.fitlimits2(4) .or.
     &     (x.gt.fitlimits2(2) .and. x.lt.fitlimits2(3)) ) then

         if ( fp1.eq.0 ) then
            phe_gauss = 0.
         else
            phe_gauss = hist1(ixp1)
         endif

      else
```

```
         do n=fp3,fp4

            factn = fact(n+1)

*  Contribution for each peak determined using Poisson statistics
            area = norm * average**n * exp(-average) / factn

            sigma = sqrt(width_ped**2 + REAL(n)*width_sig**2)
            x0 = pedestal + REAL(n)*signal

*  Gaussian distribution for each peak
            phe_gauss = phe_gauss + area *
     &           exp(-0.5*(x-x0)**2/sigma**2)/(2.50663*sigma)

         enddo
       endif

      END
```

Comments:

- the parameter n counts over the number of photoelectrons produced at the photocathode
- the shape of the electronic noise, i.e. the pedestal, is assumed to be Gaussian
- Poisson statistics is used for the number of the photoelectrons, i.e. their generation
- the shape of the signal distribution is assumed to be Gaussian, the key feature of this routine
- the fitting parameters are defined as *Double* as fits with only *Real* parameters had an increased tendency to fail to convert; $\sqrt{2\pi}$ is given as a fixed rounded value

6.2 Poisson based algorithm

```
      FUNCTION phe_poiss(x)

*  Fitting Poisson functions to photoelectron spectra from a MaPMT

      Double Precision norm,average,pedestal,width_ped,signal,width_sig
      Double Precision firstdyn, width_dyn, fstdynavg, K1
      Integer fp1,fp2,fp3,fp4
      Integer ixp1, ifirst, j, n, m
      Double Precision fact(40), facttemp, faktn, faktm
      Double Precision arean, aream, areatot

      data ifirst/1/
      common/pawpar/par(9)
      vector fitpar2(6)
      vector fitlimits2(4)
      vector hist1

      norm      = par(1)   ! normalisation factor
      average   = par(2)   ! average number of p.e. per event
      pedestal  = par(3)   ! position of pedestal
      width_ped = par(4)   ! width of pedestal
      signal    = par(5)   ! offset from pedestal to single p.e. signal
      width_sig = par(6)   ! width of single p.e. signal
      fstdynavg = par(7)   ! average number of p.e. created by 1st dynode
      firstdyn  = par(8)   ! position of 1st dynode p.e. signal
      width_dyn = par(9)   ! width of 1st dynode signal

*  Loop over number of p.e. (0 for pedestal)

      fp1 = int(fitpar2(1))
      fp2 = int(fitpar2(2))
      fp3 = int(fitpar2(3))
      fp4 = int(fitpar2(4))
      ixp1 = int(x+1)
      K1 = signal/firstdyn

      phe_poiss = 0.

*  Calculate vector of factorials

      if ( ifirst.eq.1 ) then
         facttemp = 1
         fact(1) = 1
         do j = 1,39
            facttemp = facttemp*REAL(j)
            fact(j+1) = facttemp
         enddo
         ifirst = 0
      endif

*  Check for valid fit region and loop over photoelectrons

      if ( x.lt.fitlimits2(1) .or.
     &     x.gt.fitlimits2(4) .or.
     &     (x.gt.fitlimits2(2) .and. x.lt.fitlimits2(3)) ) then

         if ( fp1.eq.0 ) then
```

```
         phe_poiss = 0.
      else
         phe_poiss = hist1(ixp1)
      endif

   else

      do n=fp3,fp4

         factn = fact(n+1)

         arean = norm * average**n * exp(-average) / factn

         if (n.eq.0 .or. n.gt.2) then

            sigma = sqrt(width_ped**2 + REAL(n)*(width_sig**2))
            x0 = pedestal + REAL(n)*signal

            phe_poiss = phe_poiss + arean *
&              exp(-0.5*(x-x0)**2/sigma**2)/(2.50663*sigma)

         else

            mhigh = n*10 + 10

            do m=0,mhigh

               factm = fact(m+1)

               aream = (REAL(n)*K1)**m * exp(-REAL(n)*K1) / factm
               areatot = arean * aream

               sigma = sqrt(width_ped**2 + REAL(m)*(width_dyn**2))
               x0 = pedestal + REAL(m)*firstdyn

               phe_poiss = phe_poiss + areatot *
&                 exp(-0.5*(x-x0)**2/sigma**2)/(2.50663*sigma)

            enddo
         endif
      enddo
   endif

   END
```

Comments: see also comments for routine *phe_gauss.f* in Section 6.1

- with respect to the routine *phe_gauss.f* this algorithm replaces the assumption of a Gaussian shape of the signal distribution with a Poisson distribution for the case of $n = 1$ or $n = 2$ photoelectrons
- for the cases $n = 0$ (pedestal) and $n \geq 3$ photoelectrons still a Gaussian distribution of the signal is assumed; the approximation for the latter case seems good enough
- the main loop actually sums over the electrons emitted from the first dynode; so n runs over the electrons emitted from the photocathode and m runs over the electrons emitted from the first dynode while *arean* and *aream* provide the corresponding weights
- *fstdynavg* has to be held at 0 external to the fit at it has no meaning in this model
- to actually cover the shape of the Poisson distribution well enough the summation is carried out to $m_{max} = n * 10 + 10$

- K_1 is the average signal gain at the first dynode: $K_1 = \frac{Q_1}{Q_2}$

6.3 Poisson based algorithm including photoconversion at the first dynode

```
      FUNCTION phe_poiss_dyn(x)

*  Fitting Poisson functions to photoelectron spectra from a MaPMT
*  and allowing for photoconversion at the first dynode

      Double Precision norm,average,pedestal,width_ped,signal,width_sig
      Double Precision firstdyn, width_dyn, fstdynavg, K1
      Integer fp1,fp2,fp3,fp4,fp5,fp6
      Integer ixp1, ifirst, j, n, k, m
      Double Precision fact(40), facttemp, faktn, faktk, faktm
      Double Precision area, arean, areak, aream, atreatot

      data ifirst/1/
      common/pawpar/par(9)
      vector fitpar2(6)
      vector fitlimits2(4)
      vector hist1

      norm      = par(1)   ! normalisation factor
      average   = par(2)   ! average number of p.e. per event
      pedestal  = par(3)   ! position of pedestal
      width_ped = par(4)   ! width of pedestal
      signal    = par(5)   ! offset from pedestal to single p.e. signal
      width_sig = par(6)   ! width of single p.e. signal
      fstdynavg = par(7)   ! average number of p.e. created by 1st dynode
      firstdyn  = par(8)   ! position of 1st dynode p.e. signal
      width_dyn = par(9)   ! width of 1st dynode signal

*  Loop over number of p.e. (0 for pedestal)

      fp1 = int(fitpar2(1))
      fp2 = int(fitpar2(2))
      fp3 = int(fitpar2(3))
      fp4 = int(fitpar2(4))
      fp5 = int(fitpar2(5))
      fp6 = int(fitpar2(6))
      ixp1 = int(x+1)
      K1  = signal / firstdyn

      phe_poiss_dyn = 0.

*  Calculate vector of factorials

      if ( ifirst.eq.1 ) then
         facttemp = 1
         fact(1) = 1
         do j = 1,39
            facttemp = facttemp*REAL(j)
            fact(j+1) = facttemp
         enddo
         ifirst = 0
      endif

*  Check for valid fit region and loop over photoelectrons

      if ( x.lt.fitlimits2(1) .or.
     &     x.gt.fitlimits2(4) .or.
```

```
&     (x.gt.fitlimits2(2) .and. x.lt.fitlimits2(3)) ) then

    if ( fp1.eq.0 ) then
       phe_poiss_dyn = 0.
    else
       phe_poiss_dyn = hist1(ixp1)
    endif

 else

    do n = fp3,fp4
       do k = fp5,fp6

          nphe = n + k
          factn = fact(n+1)
          factk = fact(k+1)
          arean = average**n * exp(-average) / factn
          areak = fstdynavg**k * exp(-fstdynavg) / factk
          area  = norm * arean * areak

          if (n.eq.0 .or. n.gt.2) then

             sigma = sqrt(width_ped**2 + REAL(n)*(width_sig**2) +
&                          REAL(k)*(width_dyn**2))
             x0 = pedestal + REAL(n)*signal + REAL(k)*firstdyn

             phe_poiss_dyn = phe_poiss_dyn + area *
&                 exp(-0.5*(x-x0)**2/sigma**2)/(2.50663*sigma)

          else

             mhigh = n*10 + 10

             do m = 0, mhigh

                factm = fact(m+1)
                aream = (REAL(n)*K1)**m * exp(-REAL(n)*K1) / factm
                areatot = area * aream

                sigma = sqrt(width_ped**2 +
&                             (REAL(m)+REAL(k))*(width_dyn**2))
                x0 = pedestal + (REAL(m) + REAL(k))*firstdyn

                phe_poiss_dyn = phe_poiss_dyn + areatot *
&                    exp(-0.5*(x-x0)**2/sigma**2)/(2.50663*sigma)

             enddo
          endif
       enddo
    enddo
 endif

 END
```

Comments: see also comments for routine *phe_gauss.f* in Section 6.1 and routine *phe_poiss.f* in Section 6.2

- with respect to the routine *phe_poiss.f* again the summation is done for the electrons emitted from the first dynode, but now an additional contribution for a photon conversion at the first dynode is

permitted; n counts over electrons produced from photon conversion at the photocathode, k counts over electrons produced from photon conversion at the first dynode and m counts over all electrons emitted from the first dynode

- in the case of $n = 0$ photoelectrons the signal shape of electrons produced at the first dynode by photon conversion is still approximated by a Gauss distribution; a description by Poisson could be allowed by swapping in the condition $((n + k) = 0$ *.or.* $(n + k) > 2)$ for $(n = 0$ *.or.* $n > 2)$

7 Appendix B

In this section the used Fortran routine for the calculation of the signal loss from the fits to the MaPMT signal spectra is documented. The used parameters have the following meaning:

- PAW vector *lossthresh(1)*: defining the position of the cut separating 'signal' (S) from 'noise' (N) in terms of the pedestal position Q_0 and width σ_0 like: $S > Q_0 + T\sigma_0$; our default was $T = 5.0$
- PAW vector *nhistoch(1)*: number of channels of the fitted histogram
- PAW vector *hits(1)*: number of events in the fitted histogram
- PAW vector *hist1(256)* or *hist1(1024)*: holds the channel content of the fitted spectrum as defined in Section 6
- PAW vector *dofitflag(1)*: used fit option: (2 or 12) Gauss-fit, (3 or 13) Poisson-fit, (4 or 14) Poisson-fit with photo conversion at the first dynode; the options 12, 13 and 14 are used in the external routines to signal just the drawing of the function in contrast to options 2, 3 and 4 which actually called the fitting routine
- PAW vector *par(9)*: holds the fit parameters as defined in Section 6, here with the variable names: *norm*, *npe*, *ped*, *sped*, *phe*, *sphe*, *dnpe*, *dphe*, *dsphe*
- PAW vector *par_sig(9)*: holds the errors on the fit parameters
- PAW vector *fitpar2(6)*: holds parameters to steer the fit behaviour as defined in Section 6
- PAW vector *fitlimits2(4)*: holds the limits to the signal range as defined in Section 6
- PAW vector *loss(5)*: holds the returned signal loss for definitions 1, ..., 5
- PAW vector *loss_sig(5)*: holds the error of the returned signal loss for definitions 1, ..., 5
- *del_xxxx*: error of variable *xxxx*
- *i_xxxx*: integral over variable *xxxx*
- *xx_ped*: pedestal events
- *xx_phe1*: events with 1 photoelectron from the photocathode
- *xx_phe1dyn*: events with 1 photoelectron from the photocathode OR 1 photoelectron from the first dynode
- *xx_ev_cut*: number of events in the histogram below the threshold cut, i.e. counted noise events (N)
- *xx_cut_ev*: number of events in the histogram above the threshold cut, i.e. counted signal events (S)
- *corr*: correction to event count from linear interpolation
- *cut*: cut value T separating 'signal' (S) from 'noise' (N); our default was $T = 5.0$
- *del_sphe_sphe_2*: error of signal width (σ_1) calculated from Q1, Q2 and σ_2

7.1 Algorithm for the calculation of signal loss

```
REAL FUNCTION calcloss()

Integer i,j,Ihistch
Real n_cut_ev,         del_n_cut_ev
Real n_ev_cut,         del_n_ev_cut
Real del_n_ev_binom
Real i_ped,            del_i_ped
Real i_phe1,           del_i_phe1
Real i_phe1_cut_ev,    del_i_phe1_cut_ev
Real i_phe1dyn,        del_i_phe1dyn
Real i_phe1dyn_cut_ev, del_i_phe1dyn_cut_ev
Real loss_N,           del_loss_N
Real loss_N_dyn,       del_loss_N_dyn
Real loss_P,           del_loss_P
Real cut,corr,del_sphe_sphe_2
Real norm,npe,ped,sped,phe,sphe,dnpe,dphe,dsphe
Real del_norm,del_npe,del_ped,del_sped,del_phe,del_sphe
Real del_dnpe,del_dphe,del_dsphe

vector lossthresh
vector nhistoch
vector hits
vector hist1
vector dofitflag
vector par
vector par_sig
vector fitlimits2
vector fitpar2
vector loss
vector loss_sig

Ihistch = int(nhistoch(1))

n_cut_ev = 0.0
n_ev_cut = 0.0
i_ped = 0.0
i_phe1 = 0.0
i_phe1_cut_ev = 0.0
i_phe1dyn = 0.0
i_phe1dyn_cut_ev = 0.0

norm  = par(1)
npe   = par(2)
ped   = par(3)
sped  = par(4)
phe   = par(5)
sphe  = par(6)
dnpe  = par(7)
dphe  = par(8)
dsphe = par(9)
del_norm  = par_sig(1)
del_npe   = par_sig(2)
del_ped   = par_sig(3)
del_sped  = par_sig(4)
del_phe   = par_sig(5)
del_sphe  = par_sig(6)
del_dnpe  = par_sig(7)
```

```
      del_dphe  = par_sig(8)
      del_dsphe = par_sig(9)

C     # of events above and below cut

      j = 1
      cut = ped+lossthresh(1)*sped
      do i=1,Ihistch
        if ( i-1.gt.cut ) then
           n_cut_ev = n_cut_ev + real(hist1(i))
        else
           n_ev_cut = n_ev_cut + real(hist1(i))
           j = i
        endif
      enddo

      corr = ( real(hist1(j)) + real(hist1(j+1)) ) / 2 *
     &     ( cut+real(1) - (real(j)+0.5) )

      n_cut_ev = n_cut_ev - corr
      n_ev_cut = n_ev_cut + corr

C     # of events above and below cut : 'upper' error estimate

      del_n_cut_ev = abs( ( real(hist1(j)) - real(hist1(j+1)) ) / 4 )
      del_n_ev_cut = abs( ( real(hist1(j)) - real(hist1(j+1)) ) / 4 )

C     binomial error

      del_n_ev_binom = sqrt( n_cut_ev * n_ev_cut / (n_cut_ev+n_ev_cut) )

C     combined error

      del_n_cut_ev = sqrt( del_n_cut_ev**2 + del_n_ev_binom**2 )
      del_n_ev_cut = sqrt( del_n_ev_cut**2 + del_n_ev_binom**2 )

C     integral over pedestal

      fitlimits2(1) = 0.
      fitlimits2(2) = 0.
      fitlimits2(3) = 0.
      fitlimits2(4) = 1025.

      fitpar2(3) = 0
      fitpar2(4) = 0
      fitpar2(5) = 0
      fitpar2(6) = 0
      do i=1,Ihistch
         if ( dofitflag(1).eq.1 .or. dofitflag(1).eq.11 ) then
            i_ped = i_ped + photoel_dr(real(i))
         elseif ( dofitflag(1).eq.2 .or. dofitflag(1).eq.12 ) then
            i_ped = i_ped + phe_gauss_dr(real(i))
         elseif ( dofitflag(1).eq.3 .or. dofitflag(1).eq.13 ) then
            i_ped = i_ped + phe_poiss_dr(real(i))
         elseif ( dofitflag(1).eq.4 .or. dofitflag(1).eq.14 ) then
            i_ped = i_ped + phe_poiss_dyn_dr(real(i))
         endif
*         write (6,*) i,i_ped
      enddo
```

```
C     integral over pedestal : error estimate (ped?,width)

      if ( ped.ne.0. .and. sped.ne.0. ) then
*         del_i_ped = norm * exp( -1.*(npe+dnpe) ) *
*     &        sqrt( (del_ped/ped)**2 + (del_sped/sped)**2 )
         del_i_ped = norm * exp( -1.*(npe+dnpe) ) *
     &        sqrt( (del_sped/sped)**2 )
      else
         del_i_ped = 99999.
      endif

C     integral over 1st photoelectron from PC

      fitpar2(3) = 1
      fitpar2(4) = 1
      fitpar2(5) = 0
      fitpar2(6) = 0
      do i=1,Ihistch
         if ( dofitflag(1).eq.1 .or. dofitflag(1).eq.11 ) then
            i_phe1 = i_phe1 + photoel_dr(real(i))
         elseif ( dofitflag(1).eq.2 .or. dofitflag(1).eq.12 ) then
            i_phe1 = i_phe1 + phe_gauss_dr(real(i))
         elseif ( dofitflag(1).eq.3 .or. dofitflag(1).eq.13 ) then
            i_phe1 = i_phe1 + phe_poiss_dr(real(i))
         elseif ( dofitflag(1).eq.4 .or. dofitflag(1).eq.14 ) then
            i_phe1 = i_phe1 + phe_poiss_dyn_dr(real(i))
         endif
      enddo
      do i=1,Ihistch
         if ( i-1.ge.(ped+lossthresh(1)*sped) ) then
      if ( dofitflag(1).eq.1 .or. dofitflag(1).eq.11 ) then
      i_phe1_cut_ev = i_phe1_cut_ev + photoel_dr(real(i))
      elseif ( dofitflag(1).eq.2 .or. dofitflag(1).eq.12 ) then
      i_phe1_cut_ev = i_phe1_cut_ev + phe_gauss_dr(real(i))
      elseif ( dofitflag(1).eq.3 .or. dofitflag(1).eq.13 ) then
      i_phe1_cut_ev = i_phe1_cut_ev + phe_poiss_dr(real(i))
      elseif ( dofitflag(1).eq.4 .or. dofitflag(1).eq.14 ) then
      i_phe1_cut_ev = i_phe1_cut_ev + phe_poiss_dyn_dr(real(i))
      endif
         endif
      enddo

C     integral over 1st photoelectron from PC : error estimate (l,phe?,width)

      del_sphe_sphe_2 = dphe**2 / ( 2*phe**2 * (dsphe**2+dphe**2)**2 ) *
     &     ( ( ( dsphe**2/dphe + dphe**2 ) * del_phe )**2
     &     + ( ( phe - phe*dsphe**2/dphe**2 ) * del_dphe )**2
     &     + ( ( 2 * phe*dsphe/dphe ) * del_dsphe )**2 )

      if ( npe.ne.0. .and. phe.ne.0. ) then
*         del_i_phe1 = norm * (1.0 - exp( -1.*npe ) ) *
*     &        sqrt( (del_npe/npe)**2 + (del_phe/phe)**2 +
*     &              del_sphe_sphe_2 )
         del_i_phe1 = norm * (1.0 - exp( -1.*npe ) ) *
     &        sqrt( (del_npe/npe)**2 +
     &              del_sphe_sphe_2 )
      else
         del_i_phe1 = 99999.
```

```
      endif

      if ( i_phe1.ne.0. ) then
         del_i_phe1_cut_ev = del_i_phe1 * i_phe1_cut_ev / i_phe1
      else
         del_i_phe1_cut_ev = 99999.
      endif

C     integral over 1st photoelectron from PC + 1st dynode

      i_phe1dyn = i_phe1
      i_phe1dyn_cut_ev = i_phe1_cut_ev

      fitpar2(3) = 0
      fitpar2(4) = 0
      fitpar2(5) = 1
      fitpar2(6) = 1
      do i=1,Ihistch
         if ( dofitflag(1).eq.1 .or. dofitflag(1).eq.11 ) then
            i_phe1dyn = i_phe1dyn + photoel_dr(real(i))
         elseif ( dofitflag(1).eq.2 .or. dofitflag(1).eq.12 ) then
            i_phe1dyn = i_phe1dyn + phe_gauss_dr(real(i))
         elseif ( dofitflag(1).eq.3 .or. dofitflag(1).eq.13 ) then
            i_phe1dyn = i_phe1dyn + phe_poiss_dr(real(i))
         elseif ( dofitflag(1).eq.4 .or. dofitflag(1).eq.14 ) then
            i_phe1dyn = i_phe1dyn + phe_poiss_dyn_dr(real(i))
         endif
      enddo
      do i=1,Ihistch
         if ( i-1.ge.(ped+lossthresh(1)*sped) ) then
      if ( dofitflag(1).eq.1 .or. dofitflag(1).eq.11 ) then
      i_phe1dyn_cut_ev = i_phe1dyn_cut_ev + photoel_dr(real(i))
      elseif ( dofitflag(1).eq.2 .or. dofitflag(1).eq.12 ) then
      i_phe1dyn_cut_ev = i_phe1dyn_cut_ev + phe_gauss_dr(real(i))
      elseif ( dofitflag(1).eq.3 .or. dofitflag(1).eq.13 ) then
      i_phe1dyn_cut_ev = i_phe1dyn_cut_ev + phe_poiss_dr(real(i))
      elseif ( dofitflag(1).eq.4 .or. dofitflag(1).eq.14 ) then
      i_phe1dyn_cut_ev = i_phe1dyn_cut_ev + phe_poiss_dyn_dr(real(i))
      endif
         endif
      enddo

C     integral over 1st photoelectron from PC + 1st dynode : error estimate (1,phe?,width)

      if ( npe.ne.0. .and. phe.ne.0. ) then
*         del_i_phe1dyn = norm * ( 1.0 - exp( -1.*dnpe ) ) *
*     &        sqrt( (del_dnpe/dnpe)**2 + (del_dphe/dphe)**2 +
*     &              (del_dsphe/dsphe)**2 )
         del_i_phe1dyn = norm * ( 1.0 - exp( -1.*dnpe ) ) *
     &        sqrt( (del_dnpe/dnpe)**2 +
     &              (del_dsphe/dsphe)**2 )
      else
         del_i_phe1dyn = 99999.
      endif

      if ( i_phe1.ne.0. ) then
         del_i_phe1dyn_cut_ev = del_i_phe1dyn *
     &                          i_phe1dyn_cut_ev / i_phe1dyn
      else
```

```
         del_i_phe1dyn_cut_ev = 99999.
       endif

C      loss calculation : intermediate parameters

       loss_N = real(hits(1)) * (1. - exp( -1.*npe) )
       del_loss_N = real(hits(1)) * exp( -1.*npe) * del_npe

       loss_N_dyn = real(hits(1)) * (1. - exp( -1.*(npe+dnpe) ) )
       del_loss_N_dyn = real(hits(1)) * exp( -1.*(npe+dnpe) ) *
      &     sqrt( del_npe**2 + del_dnpe**2 )

       loss_P = n_ev_cut - i_ped
       del_loss_P = sqrt( del_n_ev_cut**2 + del_i_ped**2 )

C      loss calculation : final parameters

       if ( loss_N_dyn.ne.0. ) then
          loss(1) = ( loss_N_dyn - n_cut_ev ) / loss_N_dyn
          loss_sig(1) = sqrt(
      &               (n_cut_ev*del_loss_N_dyn/loss_N_dyn**2)**2 +
      &               (del_n_cut_ev/loss_N_dyn)**2 )
       else
          loss(1) = 1.0
          loss_sig(1) = 2.0
       endif

       if ( i_phe1dyn.ne.0. ) then
          loss(2) = ( i_phe1dyn - i_phe1dyn_cut_ev ) / i_phe1dyn
          loss_sig(2) = sqrt(
      &               (i_phe1dyn_cut_ev*del_i_phe1dyn/i_phe1dyn**2)**2 +
      &               (del_i_phe1dyn_cut_ev/i_phe1dyn)**2 )
       else
          loss(2) = 1.0
          loss_sig(2) = 2.0
       endif

       if ( ( n_cut_ev+n_ev_cut-i_ped ).ne.0. ) then
          loss(3) = ( n_ev_cut - i_ped ) / ( n_cut_ev+n_ev_cut-i_ped )
          loss_sig(3) = sqrt(
      &               ( 1*del_n_ev_cut / (n_cut_ev+n_ev_cut-i_ped) )**2 +
      &               ( n_cut_ev*del_i_ped /
      &                 ( n_cut_ev+n_ev_cut-i_ped )**2 )**2 )
       else
          loss(3) = 1.0
          loss_sig(3) = 2.0
       endif

       if ( loss_N.ne.0. ) then
          loss(4) = ( loss_N - n_cut_ev ) / loss_N
          loss_sig(4) = sqrt( (n_cut_ev*del_loss_N/loss_N**2)**2 +
      &                       (del_n_cut_ev/loss_N)**2 )
       else
          loss(4) = 1.0
          loss_sig(4) = 2.0
       endif

       if ( i_phe1.ne.0. ) then
          loss(5) = ( i_phe1 - i_phe1_cut_ev ) / i_phe1
```

```
         loss_sig(5) = sqrt( (i_phe1_cut_ev*del_i_phe1/i_phe1**2)**2 +
     &                       (del_i_phe1_cut_ev/i_phe1)**2 )
       else
         loss(5) = 1.0
         loss_sig(5) = 2.0
       endif

* protection still against .NAN. entries
       if ( loss(1).lt.-9999 .or. loss(1).gt.9999 ) then
         loss(1) = -9999.
       endif
       if ( loss(2).lt.-9999 .or. loss(2).gt.9999 ) then
         loss(2) = -9999.
       endif
       if ( loss(3).lt.-9999 .or. loss(3).gt.9999 ) then
         loss(3) = -9999.
       endif
       if ( loss(4).lt.-9999 .or. loss(4).gt.9999 ) then
         loss(4) = -9999.
       endif
       if ( loss(5).lt.-9999 .or. loss(5).gt.9999 ) then
         loss(5) = -9999.
       endif
       if ( loss_sig(1).lt.-9999 .or. loss_sig(1).gt.9999 ) then
         loss_sig(1) = -9999.
       endif
       if ( loss_sig(2).lt.-9999 .or. loss_sig(2).gt.9999 ) then
         loss_sig(2) = -9999.
       endif
       if ( loss_sig(3).lt.-9999 .or. loss_sig(3).gt.9999 ) then
         loss_sig(3) = -9999.
       endif
       if ( loss_sig(4).lt.-9999 .or. loss_sig(4).gt.9999 ) then
         loss_sig(4) = -9999.
       endif
       if ( loss_sig(5).lt.-9999 .or. loss_sig(5).gt.9999 ) then
         loss_sig(5) = -9999.
       endif

       write (6,*) '>>> calcloss:'
       write (6,*) 'fitpar(9),K1:'
       write (6,*) norm,npe,ped,sped,phe,sphe
       write (6,*) dnpe,dphe,dsphe,phe/dphe
       write (6,*) del_norm,del_npe,del_ped,del_sped,del_phe,del_sphe
       write (6,*) del_dnpe,del_dphe,del_dsphe
       write (6,*) 'CUT = ped+lossthresh(1)*sped    )           :',
     &     ped+lossthresh(1)*sped
       write (6,*) 'n_ev_cut,n_cut_ev,loss_N,loss_N_dyn,loss_P:'
       write (6,*) 'N....    ',n_ev_cut,n_cut_ev,loss_N,loss_N_dyn,loss_P
       write (6,*) 'ERR      ',del_n_ev_cut,del_n_cut_ev,del_loss_N
     &                        ,del_loss_N_dyn,del_loss_P
       write (6,*) 'i_ped,i_phe1,i_phe1_cut_ev:'
       write (6,*) 'INT:     ',i_ped,i_phe1,i_phe1_cut_ev
       write (6,*) 'ERR      ',del_i_ped,del_i_phe1,del_i_phe1_cut_ev
       write (6,*) 'i_phe1dyn,i_phe1dyn_cut_ev:'
       write (6,*) 'INT_DYN: ',i_phe1dyn,i_phe1dyn_cut_ev
       write (6,*) 'ERR      ',del_i_phe1dyn,del_i_phe1dyn_cut_ev
       write (6,*) 'loss(1),loss(2),loss(3),loss(4),loss(5):'
       write (6,*) 'LOSS: ',loss(1),loss(2),loss(3),loss(4),loss(5)
```

```
      write (6,*) 'ERR : ',loss_sig(1),loss_sig(2),loss_sig(3)
     &                    ,loss_sig(4),loss_sig(5)
      write (6,*) 'hits:',hits

      calcloss = 1.

      END
```

Comments:

- structures like *i-1* are due to the definition of the histogram *hist1* in Fortran where the channel with *0 ADC* counts is stored in the histogram channel *1*
- the calculation of the integrals calls almost exactly the same routine as it had been used for the fit before (this is taken care of by choosing the same value for *dofitflag(1)* in the external call as for the fit); the only difference of *fit_function_dr.f* wrt. *fit_function.f* is that the parameter vector *par(9)* is handed over as PAW vector instead of in a Fortran COMMON block, and that *fit_function_dr.f* only is looped over once
- the definition of the loss variables is discussed in the main text
- the calculation of the errors is discussed in the main text
- the calculation of *corr* and its error is discussed in the main text
- the protection against extreme numerical results seemed necessary for rate pathological cases

References

[1] "A Large Hadron Collider Beauty Experiment for Precision Measurements of CP Violation and Rare Decays"; LHCb Technical Proposal, LHCC 98/04, LHCC/P4 (Feb. 1998).

[2] S. Amato et al., "LHCb RICH Technical Design Report", CERN/LHCC/2000-037, LHCb TDR 3, Sep 2000.

[3] The LHCb RICH detector web page: "http://lhcb.cern.ch/rich/".

[4] F. Muheim et. al., "Proposal for Multi-Anode Photo Multiplier Tubes as photo detectors for the LHCb RICH", LHCb 2000-065 RICH.

[5] S. Eisenhardt et. al., "Study of Multi Anode Photo Multiplier Tubes at Low Gains", LHCb 2000-088 RICH.

[6] E. Albrecht et. al., "Performance of a cluster of multi-anode photomultipliers equipped with lenses for use in a prototype RICH detector", Nucl. Inst. Meth. **A 488** (2002) 110; LHCb 2000-083 RICH, LHCb 2001-091 RICH.

[7] S. Eisenhardt et. al., "Performance of Multianode Photo Multiplier Tubes at Low Gain", LHCb 2003-043 RICH.

[8] I.E. Chirikov-Zorin, I. Fedorko, A. Menzione, M. Pikna, I. Sykora, S. Tokar, "Method of precise analysis of the metal package photomultiplier single photoelectron spectra", Nucl. Inst. Meth. **A 456** (2001) 310

[9] I.E. Chirikov-Zorin, I. Fedorko, A. Menzione, I. Sykora, S. Tokar, "Precise analysis of the metal package photomultiplier spectra", Nucl. Inst. Meth. **A 461** (2001) 587