

# Pulse processing routines for neutron time-of-flight data

---

(n\_TOF Collaboration) Žugec, Petar; Weiss, C.; Guerrero, C.; Günsing, F.; Vlachoudis, V.; Sabate-Gilarte, M.; Stamatopoulos, A.; Wright, T.; Lerendegui-Marco, J.; Mingrone, F.; ...

Source / Izvornik: **Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment**, 2016, 812, 134 - 144

Journal article, Published version

Rad u časopisu, Objavljena verzija rada (izdavačev PDF)

<https://doi.org/10.1016/j.nima.2015.12.054>

Permanent link / Trajna poveznica: <https://urn.nsk.hr/urn:nbn:hr:217:390166>

Rights / Prava: [Attribution 4.0 International](#)/[Imenovanje 4.0 međunarodna](#)

Download date / Datum preuzimanja: **2024-07-17**



Repository / Repozitorij:

[Repository of the Faculty of Science - University of Zagreb](#)





ELSEVIER

Contents lists available at ScienceDirect

# Nuclear Instruments and Methods in Physics Research A

journal homepage: [www.elsevier.com/locate/nima](http://www.elsevier.com/locate/nima)

## Pulse processing routines for neutron time-of-flight data



P. Žugec<sup>a,\*</sup>, C. Weiß<sup>b</sup>, C. Guerrero<sup>c</sup>, F. Gunsing<sup>d</sup>, V. Vlachoudis<sup>b</sup>, M. Sabate-Gilarte<sup>b,c</sup>,  
A. Stamatopoulos<sup>e</sup>, T. Wright<sup>f</sup>, J. Lerendegui-Marco<sup>c</sup>, F. Mingrone<sup>g</sup>, J.A. Ryan<sup>f</sup>,  
S.G. Warren<sup>f</sup>, A. Tsinganis<sup>b,e</sup>, M. Barbagallo<sup>h</sup>

<sup>a</sup> Department of Physics, Faculty of Science, University of Zagreb, Croatia

<sup>b</sup> European Organization for Nuclear Research (CERN), Geneva, Switzerland

<sup>c</sup> Universidad de Sevilla, Spain

<sup>d</sup> Commissariat à l'Énergie Atomique (CEA) Saclay - Irfu, Gif-sur-Yvette, France

<sup>e</sup> National Technical University of Athens (NTUA), Greece

<sup>f</sup> University of Manchester, Oxford Road, Manchester, UK

<sup>g</sup> Dipartimento di Fisica, Università di Bologna, and Sezione INFN di Bologna, Italy

<sup>h</sup> Istituto Nazionale di Fisica Nucleare, Bari, Italy

On behalf of the The n\_TOF Collaboration<sup>1</sup>

### ARTICLE INFO

#### Article history:

Received 14 September 2015

Received in revised form

7 December 2015

Accepted 20 December 2015

Available online 7 January 2016

#### Keywords:

Signal analysis algorithms

Pulse recognition

Signal baseline

Pulse shape fitting

n\_TOF facility

### ABSTRACT

A pulse shape analysis framework is described, which was developed for n\_TOF-Phase3, the third phase in the operation of the n\_TOF facility at CERN. The most notable feature of this new framework is the adoption of generic pulse shape analysis routines, characterized by a minimal number of explicit assumptions about the nature of pulses. The aim of these routines is to be applicable to a wide variety of detectors, thus facilitating the introduction of the new detectors or types of detectors into the analysis framework. The operational details of the routines are suited to the specific requirements of particular detectors by adjusting the set of external input parameters. Pulse recognition, baseline calculation and the pulse shape fitting procedure are described. Special emphasis is put on their computational efficiency, since the most basic implementations of these conceptually simple methods are often computationally inefficient.

© 2016 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

After a year and a half long shutdown, the neutron time of flight facility n\_TOF [1,2] at CERN has entered a third phase of its operation, known as n\_TOF-Phase3. The new era of the n\_TOF facility is marked by the successful completion of the construction of Experimental Area 2 (EAR2) [3–5], which was recently put into operation. Experimental Area 1 (EAR1), already in function for more than a decade, operates in parallel. The in-depth description of the general features of the n\_TOF facility, such as the neutron production and the neutron transport, may be found in Refs. [4–6].

At n\_TOF a wide variety of detectors is used for measuring neutron induced reactions, including neutron capture ( $n, \gamma$ ), neutron induced fission ( $n, f$ ) and reactions of type ( $n, p$ ), ( $n, t$ ) and ( $n, \alpha$ ). Among these are solid-state detectors (such as the silicon

based neutron beam monitor [7] and CVD diamond detectors [8]), scintillation detectors (an array of BaF<sub>2</sub> scintillator crystals [9], C<sub>6</sub>D<sub>6</sub> liquid scintillators [10]) and gaseous detectors (such as MicroMegas-based detectors [11,12], a calibrated fission chamber from the Physikalisch Technische Bundesanstalt [13], a set of Parallel Plate Avalanche Counters [14]). Several other types of detectors were recently introduced and tested at n\_TOF, such as solid-state HPGe, scintillation NaI, and gaseous <sup>3</sup>He detectors.

A high-performance digital data acquisition system is used for the management and storage of the electronic detector signals. The system is based on flash analog-to-digital (FADC) units, recently upgraded to handle an amplitude resolution of 8 to 12 bits. It operates at sampling rates typically ranging from 100 MHz to 1 GHz, with a memory buffer of up to 175 MSamples, allowing for an uninterrupted recording of the detector output signals during the full time-of-flight range of approximately 100 ms (as used in EAR1). A detailed description of the previous version of this system can be found in Ref. [15].

Once stored in digital form, the electronic signals have to be accessed for offline analysis, in order to obtain the time-of-flight

\* Corresponding author. Tel.: +385 1 4605552.

E-mail address: [pzucec@phy.hr](mailto:pzucec@phy.hr) (P. Žugec).

<sup>1</sup> URL: [www.cern.ch/ntof](http://www.cern.ch/ntof)

and pulse height information for each detected pulse. The analysis procedures applied to the signals from C<sub>6</sub>D<sub>6</sub> and BaF<sub>2</sub> detectors have already been described in Refs. [15,16]. In order to efficiently and consistently accommodate analysis requirements of a wide variety of detectors used at n\_TOF, a generic type of routine was recently developed that can be applied to different types of signals. The routine is characterized by a minimal number of explicit assumptions about the nature of signals and is based on a pulse template adjustment, which we refer to as the *pulse shape fitting*. For each detector or type of detector a set of analysis parameters needs to be set externally. A number of these will be mentioned throughout this paper.

Many of the procedures adopted for the signal analysis – such as the pulse integration with the goal of extracting the energy deposited in the detectors, or the constant fraction discrimination for determining the pulses' timing properties – are all well established techniques, thus we do not find it necessary to enter their description. Consequently, we will focus on the technical aspects of the more specific methods that were found to perform very well for the wide variety of signals from different detectors, in order to provide their documentation and ensure their reproducibility. Special emphasis will be put on the computational efficiency of these procedures.

Selected examples of the signals from the detectors available at n\_TOF are shown throughout the paper. Regarding the previous works on the signal analysis procedures adapted to the specific types of detectors, the reader may consult Refs. [17–20], dealing with NaI, HPGe, silicon and organic scintillation detectors, respectively. We also refer the reader to an exhaustive comparative analysis of many different pulse shape processing methods comprehensively covered in Ref. [21], and to the references contained therein.

## 2. Pulse recognition

### 2.1. Signal derivative

The central procedure in the pulse recognition is the construction of the signal derivative  $d$ . We use the following definition:

$$d_i \equiv \sum_{j=1}^{\min[N, i, P-1-i]} (s_{i+j} - s_{i-j}) \quad (1)$$

that takes advantage of integrating the signal  $s$  at both sides of the  $i$ -th point at which the derivative is to be calculated.  $P$  is the total number of points composing the recorded signal. The points are enumerated from 0 to  $P-1$ , which is a convention used throughout this paper, unless explicitly stated otherwise. A *step-*

**Table 1**

List of recursive relations for calculating forward and backward sums  $\Sigma_i^+$  and  $\Sigma_i^-$  from Eq. (2). The signal derivative  $d_i$  may then be obtained as:  $d_i = \Sigma_i^+ - \Sigma_i^-$ . Cases are categorized based on the boundary effects (whether the integration windows defined by the step-size  $N$  reach the boundaries of the waveform, composed of total of  $P$  points), as illustrated in Fig. 1.

**Beginning of the waveform**

$i \leq N$  and  $2i \leq P-1$

$$\Sigma_i^+ = \Sigma_{i-1}^+ - s_i + s_{2i-1} + s_{2i}$$

$$\Sigma_i^- = \Sigma_{i-1}^- + s_{i-1}$$

**Middle of the waveform**

$i > N$  and  $i+N \leq P-1$

$$\Sigma_i^+ = \Sigma_{i-1}^+ - s_i + s_{i+N}$$

$$\Sigma_i^- = \Sigma_{i-1}^- + s_{i-1} - s_{i-1-N}$$

**End of the waveform**

$i+N > P-1$  and  $2i > P-1$

$$\Sigma_i^+ = \Sigma_{i-1}^+ - s_i$$

$$\Sigma_i^- = \Sigma_{i-1}^- + s_{i-1} - s_{2i-P-1} - s_{2i-P}$$

size  $N$  is the default number of points to be taken for integration. As illustrated by Fig. 1, this procedure formally resembles a convolution between the signal and a see-saw-shaped filter function of unit height, up to the boundary effects regulated by the upper summation bound from Eq. (1). Evidently, when the step-size  $N$  is adjusted so as to be wider than the period of noise in the signal (and narrower than the characteristic pulse length), the procedure acts as a low-pass filter, improving the signal-to-noise ratio in the derivative.

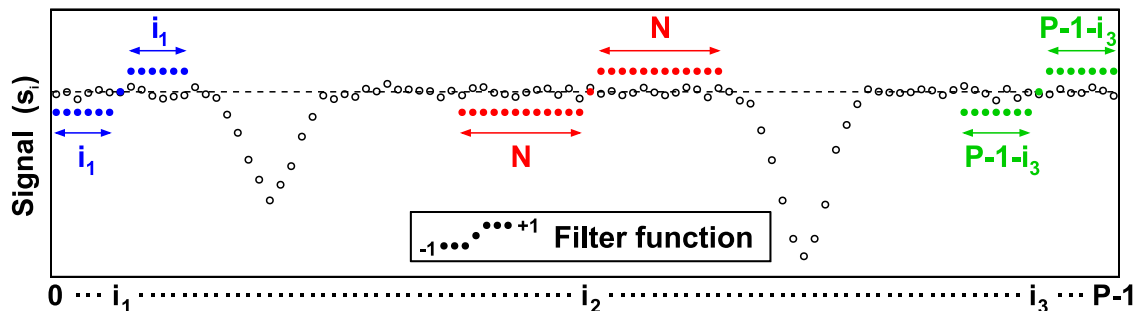
The number of operations required by the straightforward implementation of this algorithm is proportional to  $N \times P$ , making such approach computationally inefficient. Fortunately, recursive relations may be derived for calculating the consecutive  $d_i$  terms, making the entire procedure linear in the number of required operations:  $\mathcal{O}(P)$ . By defining the forward and backward sums  $\Sigma_i^+$  and  $\Sigma_i^-$ , respectively, as:

$$\Sigma_i^\pm \equiv \sum_{j=1}^{\min[N, i, P-1-i]} s_{i \pm j} \quad (2)$$

the derivative may be rewritten as:  $d_i = \Sigma_i^+ - \Sigma_i^-$ . The initial values  $\Sigma_0^+ = 0$  and  $\Sigma_0^- = 0$  follow directly from Eq. (1). The recursive relations for subsequent pairs of  $\Sigma_i^+$  and  $\Sigma_i^-$  are given in Table 1, being listed according to the boundary effects.

### 2.2. Derivative crossing thresholds

In order to recognize the presence of the pulses in the overall signal, their derivative must cross certain predefined thresholds. These thresholds need to be set in such a way as to reject most of



**Fig. 1.** Illustration of the procedure for calculating the signal derivative from Eq. (1). The filter of step-size  $N$  (red dots) is applied to the artificially constructed signal (open dots). The behavior of the filter at signal boundaries is shown as well (blue and green dots). (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)

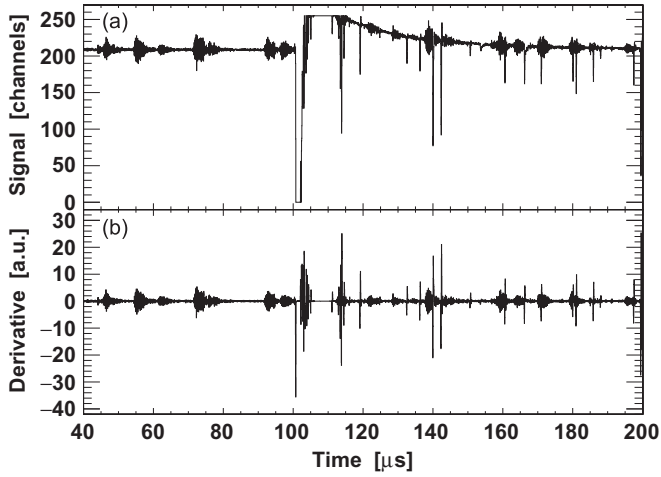


Fig. 2. Top panel (a): example of the digitized signal from MicroMegas detector. Bottom panel (b): its derivative calculated from Eq. (1).

the noise, but not to discard even the lowest pulses. Therefore, they must be adaptively brought into connection with the level of the noise characteristic of the current waveform, which is best expressed through the root mean square (RMS) of the noise. Fig. 2 shows an example of one of the most challenging signals for this task, the signal from a MicroMegas detector. Top panel (a) shows the selected fraction of an actual recorded signal, with the strongest pulse corresponding to an intense  $\gamma$ -flash caused by the proton beam hitting the spallation target, while the bottom panel (b) shows its derivative calculated from Eq. (1). This signal is heavily affected by the random beats which do not qualify as the pulses of interest to any meaningful measurement (by *beats* we consider the coherent noise resembling acoustic beats, as shown in Fig. 2 and later in Fig. 10). Several tasks are immediately evident. First, the pulses themselves must be excluded from the procedure for determining the derivative thresholds, since they can only increase the overall RMS, thus leading to a rejection of the lowest pulses. However, the pulses can not be discriminated from the noise before the thresholds have been found. Second, the beats must not be assigned to the noise RMS, since they are only sporadic and can also only lead to an unwanted increase in thresholds. Finally, in some cases one can not even rely on the assumption of a fixed number of clear presamples before the first significant pulse, such as the initial  $\gamma$ -flash pulse. This is the case in measurements with high activity samples, when their natural radioactivity causes a continual stream of pulses, independent of the external experimental conditions. Another example is the intake of waveforms for certain calibration purposes, when no external trigger is used and signals are recorded without any guarantee of clear presamples. In order to meet all these challenges, the procedure of applying the weighted fitting to the modified distribution of derivative points is used. It may be decomposed into four basic steps, described throughout this section.

*Step 1:* build the distribution (histogram) of all derivative points. As Fig. 2 shows, all the points from the derivative baseline are expected to group around the value 0, forming a peak characterized by the RMS of the noise. On the other hand, the points from the sporadic pulses and/or beats are expected to form the long tails of the distribution. Since the central peak of the distribution carries the information about the sought for RMS, it needs to be reconstructed by means of (weighted) fitting.

A technicality is related to the treatment of the central bin, corresponding to the derivative value 0. It has been observed that

in certain cases an excessive number of points is accumulated in this bin, making it reach out high above the rest of the distribution. Depending on the specific signal conditions, this feature has proven to be either beneficial or detrimental to the quality of the fitting procedure. Therefore, the content  $N_c$  of the central ( $c$ -th) bin is replaced by:

$$N_c \rightarrow \sqrt{N_c \times \frac{N_{c-1} + N_{c+1}}{2}} \quad (3)$$

i.e. by the geometrical mean between the initial content and the arithmetic mean of the neighboring bins. Since the geometric mean is biased towards the smaller of the averaged terms, this solution was selected in an attempt of finding an ideal compromise between retaining the signature of the original bin content and bringing it down towards the main fraction of the histogram. It was found that after this modification the RMS of the fitted distribution is very well adjusted to the derivative baseline in both cases: when the initial bin content would have worked either to the advantage or the detriment of the fitting procedure.

*Step 2:* adjust the histogram range. After building the initial distribution, taking into account all derivative points and adjusting the central bin, the histogram range is reduced by cutting it symmetrically around 0 until 10% of its content has been discarded. This procedure helps in localizing the relevant part of the distribution by rejecting the sporadic far-away points, thus limiting the range of the distribution from  $-d_{\max}$  to  $d_{\max}$ , which will be of central importance in defining the weights for the weighted fitting.

*Step 3:* emphasize the central peak. One must consider that even with appropriate weights, the fitting might still be heavily affected by the long tails of the distribution, increasing the final extracted RMS. In order to compensate for this effect, the central peak is better pronounced by exponentiating the entire distribution, i.e. by replacing the content  $N_i$  of the  $i$ -th histogram bin by the following value:

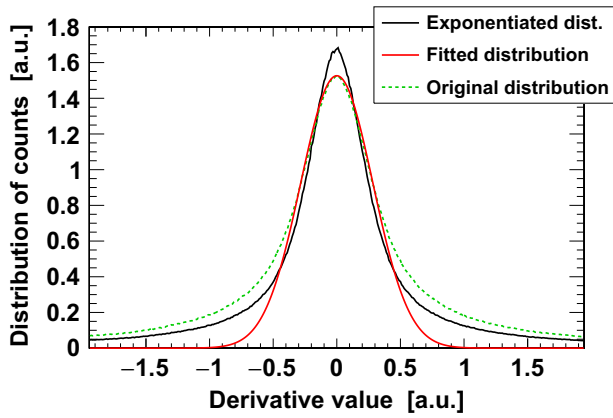
$$N_i \rightarrow e^{N_i} - 1 \quad (4)$$

This procedure affects the width of the central peak, narrowing it somewhat when there are no significant tails. The lower extracted RMS is preferred over the higher one, in order for the derivative thresholds not to reject the lowest pulses. As it will be explained later, the accidental triggering of lower thresholds by the noise will be discarded by the appropriate pulse elimination procedure. Before exponentiating the histogram content, care must be taken to rescale it appropriately – e.g. by scaling the distribution peak to unity – in order to avoid the potential numerical overflow. Furthermore, a consistent normalization is crucial in making the procedure insensitive to the length of the recorded signal (i.e. the initial height of distribution), since the exponentiation is nonlinear in the absolute number of counts  $N_i$ .

*Step 4:* perform the weighted fitting so as to best reconstruct the central peak. The remaining distribution is fitted to a Gaussian shape explicitly assumed to be centered at 0, by minimizing the following expression:

$$\sum_{i=i_{\min}}^{i_{\max}} W_i (N_i - A e^{-x_i^2/(2\Delta^2)})^2 \quad (5)$$

where  $x_i$  is the abscissa coordinate of the  $i$ -th bin, such that  $x_{i_{\min}} = -d_{\max}$  and  $x_{i_{\max}} = d_{\max}$ . Parameters  $A$  and  $\Delta$  are to be determined by fitting. At the end of the procedure,  $\Delta$  is identified with the RMS of the central peak, i.e. with the RMS of the noise in the derivative. The selection of a Gaussian as a prior is justified by the Central Limit Theorem, applied to a sum of random noise values from Eq. (1). Central to the fitting are the weights  $W_i$ , which have been



**Fig. 3.** Distribution of derivative points from Fig. 2, with the result of the weighted fitting designed to reconstruct the width of the central peak. The dashed line shows the true distribution of points from Fig. 2, arbitrarily scaled to a height of a fitted distribution. The exponentiated distribution was obtained starting from the original distribution scaled to unity.

selected to follow the Gaussian dependence:

$$W_i = e^{-x_i^2/(2\Lambda^2)} \quad (6)$$

with a standard deviation  $\Lambda$ . By an empirical optimization it was set to  $\Lambda = d_{\max}/4$ . These weights efficiently suppress the impact from the tails of the distribution, while giving precedence to the central peak. For the fitting a Levenberg–Marquardt algorithm was adopted, as described in Ref. [22]. Fig. 3 shows the distribution of derivative points from Fig. 2, together with the central peak reconstruction by means of the weighted fitting.

While the weighted fitting is beneficial for rejecting the long tails of the distribution, the unweighted fitting has been found more appropriate for very narrow distributions covering only a few histogram bins. Due to the low number of bins and rapidly decreasing weighting factors, the weighted fitting procedure is then sensitive only to the narrow top of the distribution, which is effectively treated as flat, yielding an outstretched fit. Therefore, the unweighted fitting to the Gaussian shape from Eq. (5) is also performed. In addition, the RMS of the distribution is calculated directly as:  $\text{RMS}^2 = \sum_{i=i_{\min}}^{i_{\max}} N_i x_i^2 / \sum_{i=i_{\min}}^{i_{\max}} N_i$ . The lowest of the three results – from the unweighted fitting, unweighted fitting and the direct calculation – is kept as the final one. The additional fitting and the direct calculation also serve as a contingency in case either of the fitting procedures fails to properly converge.

### 2.3. Pulse discrimination

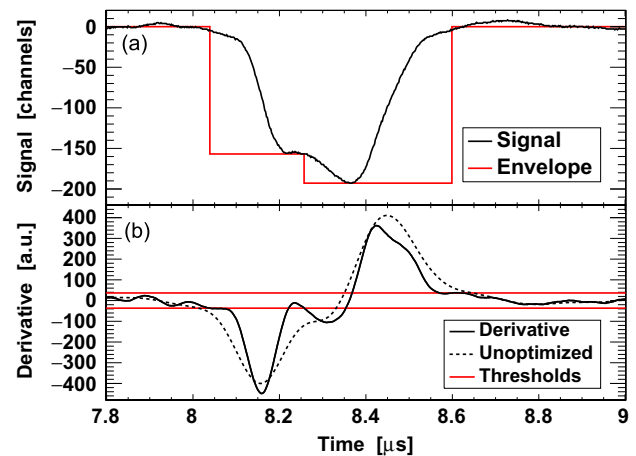
From the derivative noise RMS extracted by one of the previously described procedures, the default values for the derivative crossing thresholds have been selected as  $\pm 3.5 \times \text{RMS}$ , due to the fact that this range corresponds to 99.95% confidence interval under the assumption of normally distributed noise. Since the order of crossing these thresholds (together with some later analysis procedures) depends on the pulse polarity, all signals are treated as negative. This means that the signals are inverted, i.e. multiplied by  $-1$ , if expected to be positive from an external input parameter.

Differentiating the unipolar pulse leads to a bipolar pulse in the derivative. Therefore, the derivative of the negative unipolar pulse must, ideally, make 4 threshold crossings in this exact order: lower–lower–upper–upper. However, in case of the lowest pulses or very high pileup, the integration procedure from Eq. (1) may flatten the final derivative, not causing the second threshold crossing. Hence, the principle of 4 threshold crossings was relaxed

in order to facilitate the recognition of these pulses. Thus, crossing a single threshold suffices to trigger the pulse recognition. However, if both thresholds are crossed in the order lower–upper, a single pulse is recognized, instead of two. In summary, these are the threshold crossing possibilities that mark the presence of the pulse: lower–lower (without the subsequent upper crossing), upper–upper (without the previous lower crossing) and lower–lower–upper–upper. After initially locating the pulses between the points of the first and the last threshold crossing, their range is further extended until the derivative reaches 0 at both sides, unless there are neighboring pulses in line preventing the expansion.

The thresholds being low enough not to reject the lowest pulses will, from time to time, be accidentally triggered by the noise. These occurrences are dealt with by a set of elimination conditions, which are determined by means of the external input parameters. These conditions include the lower and upper limit for the pulse width, the lower limit for the pulse amplitude and the lower and upper limit for the area-to-amplitude ratio. The first elimination, based only on the pulse width, is performed immediately after the pulse recognition procedure. The final elimination, based on the pulse amplitudes and areas, may only be performed at a later stage, after the signal baseline has been calculated. However, it is paramount that the first stage of elimination be performed at this point, since several later procedures, such as the baseline calculation, depend on the reported pulse candidates. In case of an excessive number of falsely recognized pulses, the quality of procedures relying on the reported pulse positions may be compromised.

Fig. 4 shows an example of a demanding case of pileup, where two pulses are successfully resolved. The top panel (a) shows the actual signal, with the red envelope confining separate pulses. The bottom panel (b) shows the optimized signal derivative crossing the thresholds, triggering the pulse recognition. It also illustrates the importance of optimizing the step-size for calculating the derivative from Eq. (1), since a further increase in step-size (dashed line) would flatten the derivative at the point of the second crossing, preventing the separation of two pulses from panel (a). For visual purposes, two displayed derivatives were normalized so that their thresholds coincide.



**Fig. 4.** Pulse recognition procedure applied to the piled-up pulses. Top panel (a) shows the actual signal, with the red envelope marking the successful separation of the pulses. Bottom panel (b) shows the signal derivative crossing the appropriate thresholds and triggering the pulse recognition from panel (a). Derivative calculated with an unoptimized (too large) step-size is also shown. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)

The described pulse recognition technique was found to perform very well for signals from a wide variety of detectors in use at n\_TOF. The example from Fig. 4 confirms that with optimized parameters the procedure is able to resolve quite demanding pileups. Due to the relaxed threshold crossing conditions, it is also quite sensitive even to the lowest pulses, barely exceeding the level of the noise. Since the same sensitivity characterizing the pulse recognition procedure sporadically leads to an accidental threshold crossing due to noise, an elimination procedure has been implemented alongside it.

#### 2.4. Multiple polarities

The adopted pulse recognition procedure lends itself easily to signals that exhibit pulses of both polarities. In this case two derivative passes should be made – one over the regular derivative, one over the inverted one (multiplied by  $-1$ ). Quite often, the reported pulse candidates from two passes will overlap, since the part of a real pulse from one pass will act as a false candidate within the other pass. The pulse candidates from two passes should be analyzed independently and then submitted to the pulse elimination algorithm. It was observed that even the quite relaxed elimination conditions successfully reject the false candidates from the selection of overlapping pulses.

#### 2.5. Bipolar pulses

An additional pulse range adjustment procedure was implemented in order to accommodate bipolar pulses. Since the end of the pulse is determined by the derivative reaching 0 after the first unipolar part of the pulse, the recognition of bipolar pulses stops at the extremum of the second pole. However, once the signal baseline has been calculated, the boundary of the pulse may be shifted towards the point of the baseline crossing, keeping only the first pole of the pulse or fully covering both of them. In case of two immediate but not piled-up bipolar pulses, the first one ends at the extremum of its own second pole, where the next pulse is immediately recognized to start, due to the behavior of the derivative  $d$ . Therefore, the starting points of the pulses need to be adjusted (with respect to the calculated baseline) in accordance with the requirements of a specific signal, so that the finally determined range of the second pulse does not start prematurely, preventing also the (optional) expansion of the first pulse.

### 3. Baseline

Three different baseline methods have been implemented, that may all be used within the same waveform, depending on the signal behavior. These are the constant baseline, the weighted moving average and the moving maximum. The use of the moving maximum is usually related only to the first part of the waveform, when the effect of the  $\gamma$ -flash upon the signal is extreme (there is also an alternative method of subtracting the baseline distortion pulse shape, designed for this region). Moving average is also related to the baseline distortion by  $\gamma$ -flash, however it is often the most appropriate method to be used throughout the entire waveform, especially if the baseline exhibits slow oscillations. Constant baseline is suitable only after the baseline has been fully restored after the initial  $\gamma$ -flash, or if the detector response to external influences is remarkably stable.

#### 3.1. Constant baseline

A constant baseline is calculated as the average of all signal points between the pulse candidates reported by the pulse

recognition procedure. In this way any need for an iterative procedure is avoided, while the baseline remains unaffected by the actual pulses.

#### 3.2. Weighted moving average

The moving average is the appropriate method for determining the baseline whenever the clear information about the baseline is, in fact, available, i.e. when the uninterrupted portions of the baseline may indeed be found within the signal. The following definition is used for the weighted moving average:

$$B_i \equiv \frac{\sum_{j=\max[0,i-N]}^{\min[i+N,P-1]} s_j w_j \left\{ 1 + \cos \left[ (j-i) \frac{\pi}{N} \right] \right\}}{\sum_{j=\max[0,i-N]}^{\min[i+N,P-1]} w_j \left\{ 1 + \cos \left[ (j-i) \frac{\pi}{N} \right] \right\}} \quad (7)$$

with  $N$  as the number of points (referred to as the *window parameter*) at each side of the  $i$ -th one to be taken for averaging the signal  $s$ , composed of the total of  $P$  points. It should be noted that the averaging window is  $2N+1$  points wide. The weighting kernel is given by the cosine (i.e. Hann [23]) window, with additional weighting factors  $w_i$  that are equal to the number of uninterrupted points within a given stretch of the baseline. Inside the reported pulse candidates, these weights should be much lower than unity ( $w_i \ll 1$ ), so as to exclude the pulses from the baseline calculation. However, a finite non-zero value is required, in order to avoid division by zero in case the averaging window is completely contained within the pulse. For the weighting factors inside the pulses we have adopted the value  $10^{-6}$ . More precisely, for  $Q$  as the total number of pulses identified inside the waveform – with  $\alpha_q$  and  $\beta_q$  denoting the first and the last index of the  $q$ -th pulse ( $q \in [1, Q]$ ), respectively – the weighting factors are defined as:

$$w_i = \begin{cases} 10^{-6} & \text{if } i \in [\alpha_q, \beta_q] \text{ for any } q \in [1, Q] \\ \alpha_{q+1} - \beta_q - 1 & \text{if } i \in (\beta_q, \alpha_{q+1}) \text{ for any } q \in [0, Q] \end{cases} \quad (8)$$

where  $\beta_0 = -1$  and  $\alpha_{Q+1} = P$ . Evidently, the window parameter  $N$ , given as an external parameter, should be large enough to connect the baseline at both sides of the widest pulse or the widest expected chain of piled-up pulses. The initial elimination of falsely recognized pulses (based on their widths) also plays a role in this procedure, since every reported pulse interrupts the baseline, affecting the weighting factors  $w$ . Still, the procedure is quite robust against this change of the weighting factors.

The form of the summation bounds from Eq. (7) properly takes into account the boundary cases, when the averaging window reaches the edges of the signal. Once again, the straightforward implementation of the algorithm for evaluating Eq. (7) is of  $\mathcal{O}(N \times P)$  computational complexity. Hence, recursive relations have been derived, which provide a linear dependence in the number of operations for calculating the baseline throughout the entire waveform:  $\mathcal{O}(P)$ . We define the following terms:

$$\begin{aligned} K_i^{(\lambda)} &\equiv \sum_{j=\max[0,i-N]}^{\min[i+N,P-1]} \lambda_j \\ C_i^{(\lambda)} &\equiv \sum_{j=\max[0,i-N]}^{\min[i+N,P-1]} \lambda_j \cos \left[ (j-i) \frac{\pi}{N} \right] \\ S_i^{(\lambda)} &\equiv \sum_{j=\max[0,i-N]}^{\min[i+N,P-1]} \lambda_j \sin \left[ (j-i) \frac{\pi}{N} \right] \end{aligned} \quad (9)$$

allowing to rewrite Eq. (7) as:

$$B_i = \frac{K_i^{(sw)} + C_i^{(sw)}}{K_i^{(w)} + C_i^{(w)}} \quad (10)$$

**Table 2**

List of recursive relations for evaluating the baseline from Eq. (7). The involved terms are defined by Eq. (9). The separate cases refer to the position of the averaging window, defined by the window parameter  $N$ , relative to the edges of the waveform composed of  $P$  points. Constants  $\kappa_c \equiv \cos \frac{\pi}{N}$  and  $\kappa_s \equiv \sin \frac{\pi}{N}$  have been introduced for the efficiency of the calculation.

**Window protrudes at the beginning of the waveform**

$$i - N \leq 0 \text{ and } i + N \leq P - 1$$

$$K_i^{(\lambda)} = K_{i-1}^{(\lambda)} + \lambda_{i+N}$$

$$C_i^{(\lambda)} = \kappa_c \times C_{i-1}^{(\lambda)} + \kappa_s \times S_{i-1}^{(\lambda)} - \lambda_{i+N}$$

$$S_i^{(\lambda)} = \kappa_c \times S_{i-1}^{(\lambda)} - \kappa_s \times C_{i-1}^{(\lambda)}$$

**Window is contained within the waveform**

$$i - N > 0 \text{ and } i + N \leq P - 1$$

$$K_i^{(\lambda)} = K_{i-1}^{(\lambda)} - \lambda_{i-1-N} + \lambda_{i+N}$$

$$C_i^{(\lambda)} = \kappa_c \times C_{i-1}^{(\lambda)} + \kappa_s \times S_{i-1}^{(\lambda)} + \kappa_c \times \lambda_{i-1-N} - \lambda_{i+N}$$

$$S_i^{(\lambda)} = \kappa_c \times S_{i-1}^{(\lambda)} - \kappa_s \times C_{i-1}^{(\lambda)} - \kappa_s \times \lambda_{i-1-N}$$

**Window protrudes at the end of the waveform**

$$i - N > 0 \text{ and } i + N > P - 1$$

$$K_i^{(\lambda)} = K_{i-1}^{(\lambda)} - \lambda_{i-1-N}$$

$$C_i^{(\lambda)} = \kappa_c \times C_{i-1}^{(\lambda)} + \kappa_s \times S_{i-1}^{(\lambda)} + \kappa_c \times \lambda_{i-1-N}$$

$$S_i^{(\lambda)} = \kappa_c \times S_{i-1}^{(\lambda)} - \kappa_s \times C_{i-1}^{(\lambda)} - \kappa_s \times \lambda_{i-1-N}$$

**Window protrudes at both ends of the waveform**

$$i - N \leq 0 \text{ and } i + N > P - 1$$

$$K_i^{(\lambda)} = K_{i-1}^{(\lambda)}$$

$$C_i^{(\lambda)} = \kappa_c \times C_{i-1}^{(\lambda)} + \kappa_s \times S_{i-1}^{(\lambda)}$$

$$S_i^{(\lambda)} = \kappa_c \times S_{i-1}^{(\lambda)} - \kappa_s \times C_{i-1}^{(\lambda)}$$

where the notation  $\lambda = sw$  implies:  $\lambda_j = s_j w_j$ . Initial values  $K_0^{(\lambda)}$ ,  $C_0^{(\lambda)}$  and  $S_0^{(\lambda)}$  are to be calculated directly from Eq. (7). The recursive relations for calculating all subsequent terms are listed in Table 2, according to the position of the averaging window, relative to the waveform boundaries. It should be noted that the efficient calculation requires the terms  $\cos \frac{\pi}{N}$  and  $\sin \frac{\pi}{N}$  to be treated as constants and calculated only once, instead of repeating the calculation at each step. Fig. 5 shows two examples of the performance of the described baseline procedure.

**3.3. Moving maximum**

The following baseline procedure is appropriate when the information about the signal baseline has been (almost)

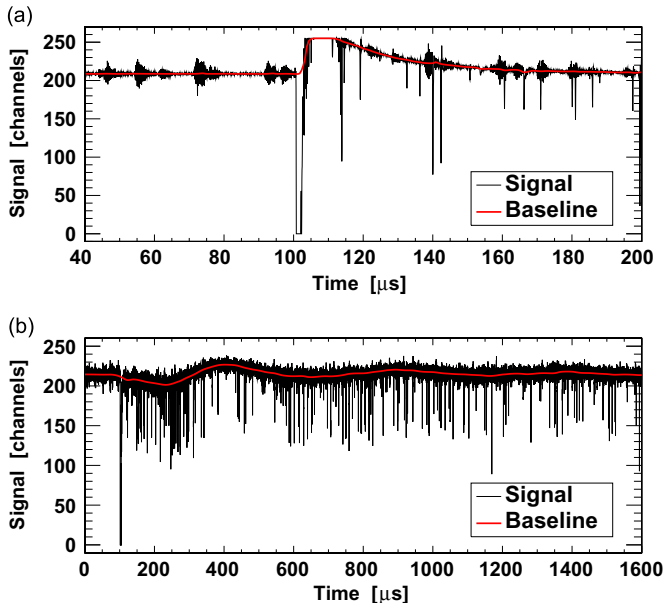


Fig. 5. Independent examples of the adaptive baseline calculated using the weighted moving average procedure from Eq. (7).

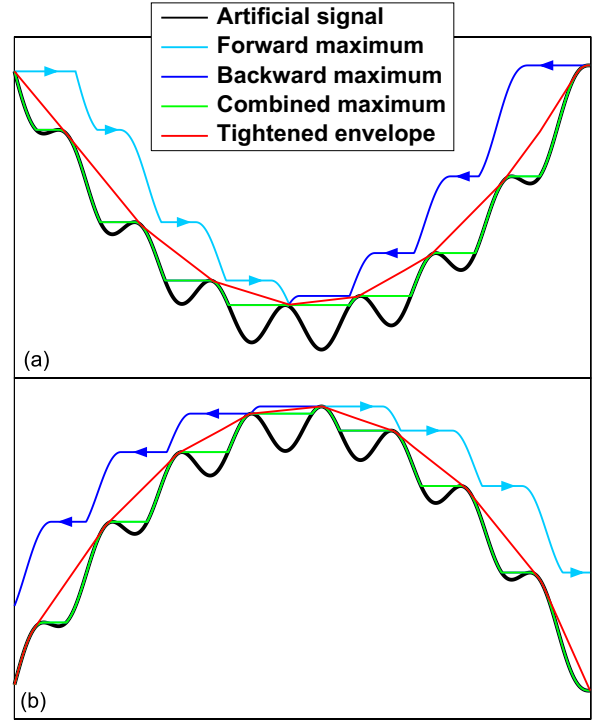


Fig. 6. Proof of concept for finding the upper signal envelope by combining the forward and backward moving maximum. The tightened envelope is also shown. The signals have been artificially constructed.

completely lost due to the sequential and persistent pileup of pulses, while the baseline itself is known not to be constant and no other *a priori* knowledge about it is available (example given later in Fig. 7). In this case the best, if not the only assumption to be made is that the baseline follows the signal envelope, defined by the dips between the pulses, especially those that manage to reach most deeply toward the true baseline. Since all signals are treated as negative, as stated before, the upper envelope needs to be found. This may be done by constructing two moving maxima – one that we refer to as the *forward maximum*, the other as the *backward maximum* – and taking the minimum of the two at each point of the signal (the advantages of this kind of competitive approach have already been explored in the past [24]). We define the forward maximum at  $i$ -th point as the maximal signal value from a *moving window* of  $N$  points before the  $i$ -th one, with backward maximum as the maximal value from the window of  $N$  points after the  $i$ -th one:

$$M_i^{(\text{forward})} = \max[s_{\max[0, i-N+1]}, \dots, s_i]$$

$$M_i^{(\text{backward})} = \max[s_i, \dots, s_{\min[i+N-1, P-1]}]$$

As before,  $P$  is the total number of points in the waveform, with  $N$  as the external input parameter. The upper envelope – following closely the upper edge of the signal, thus defining the baseline  $B$  – may simply be obtained by taking the pointwise minimum:

$$B_i \equiv \min[M_i^{(\text{forward})}, M_i^{(\text{backward})}]$$

Fig. 6 illustrates the proof of the concept on artificially constructed signals. The straightforward implementation of this procedure is again of  $\mathcal{O}(N \times P)$  computational complexity. Therefore, a very elegant and efficient algorithm was adopted from Ref. [25], that significantly speeds up the procedure, bringing it much closer to the linear dependence:  $\mathcal{O}(P)$ . A simplified version of the code from Ref. [25], excluding the calculation of the moving minimum and not requiring the *deque* data structure available in C++, is

presented in Table A1 from Appendix A. Thus obtained envelope may be additionally tightened in order to obtain a smoother and somewhat less artificial baseline. The tightening code, which is more efficient than a quadratic one, is given in Table A2 from Appendix A. Fig. 7 shows the result of this procedure on a selected portion of a real signal from a gaseous  $^3\text{He}$  detector.

### 3.4. $\gamma$ -flash removal

At neutron time-of-flight facilities the most common cause for a baseline distortion is the induction of a strong pulse by an intense  $\gamma$ -flash, which is released each time the proton beam hits the spallation target. The response of certain detectors to the  $\gamma$ -flash is remarkably consistent, which allows for a clear identification of the distorted baseline. By properly averaging a multitude of signals from an immediate vicinity of the  $\gamma$ -flash pulse, the detector response to a  $\gamma$ -flash may be recovered in form of an average baseline distortion pulse shape [26]. In effect, this pulse shape serves as a *a priori* knowledge of the baseline. In general, the baseline offset may be changed for various reasons, e.g. by simply adjusting the digitizer settings. Hence, if available, the shape of the distorted baseline is subtracted from the signal only after identifying and subtracting the primary baseline, which is – for obvious reasons – best found as the constant baseline offset. The positioning of the distorted baseline within the signal is performed relative to the  $\gamma$ -flash pulse, by fitting the externally selected portion of the pulse shape to a leading edge of the  $\gamma$ -flash pulse. The fitting routine, which is the same as for the regular pulses, is described in Section 4. Fig. 8 shows an example of the adjustment of a distorted baseline to a signal from a MicroMegas detector, clearly revealing the true pulses rising above the baseline, thus providing access to the low time-of-flight, i.e. the high-neutron-energy region.

## 4. Pulse shape analysis

After baseline subtraction, the amplitude, area, status of the pileup and timing properties such as the time of arrival are determined for each pulse. Three different methods are available for finding the amplitudes: search for the highest point, parabolic fitting to the top of the pulse and a predefined pulse shape adjustment. By *pulse shape* we refer to the template pulse of a fixed form, given by the tabulated set of points  $(t_i, p_i)$ , with  $t_i$  as the time coordinate of the  $i$ -th point and  $p_i$  as its height (i.e. the pulse shape value). The optimal pulse shape is best obtained by averaging a large number of real pulses. Several example procedures for excluding unreliable pulses from the pulse shape extraction may be found in Ref. [19].

Though the pulse shape fitting is generally the most appropriate method for pulse reconstruction, it may not always be applicable, especially if the detector exhibits pulses of strongly varying shapes. This is often the case with gaseous detectors, when the shape and length of the pulse depend on the initial point of ionization and/or the details of the particle trajectory inside the gas volume. The area under the pulse may be calculated by simple signal integration or from a pulse shape fit, if the latter option has been activated by means of the external input parameter. Finally, extraction of the timing properties relies on the digital implementation of the constant fraction discrimination, with a constant fraction factor of 30%.

### 4.1. Pulse shape fitting – the numerical procedure

Pulse shape fitting is a well established method [19–21]. However, its straightforward implementation is of  $\mathcal{O}(n^2)$

computational complexity – with  $n$  as the number of points comprising a typical pulse – whereas our adopted procedure requires only  $\mathcal{O}(n \log n)$  operations per pulse. It is important to note that any pulse shape from the following procedure is of the same sampling rate as the analyzed signal. If there is an initial mismatch between the sampling rates of the externally delivered pulse shape and the real signal, the pulse shape is first synchronized to the signal by means of linear interpolation.

Let us consider the predefined (and already synchronized) pulse shape  $p$ , consisting of  $N$  points, with the  $M$ -th one as the highest point ( $0 \leq M \leq N-1$ ). For a given pulse within the analyzed signal, the left and right fitting boundaries  $L$  and  $R$  are determined. These may correspond to the pulse boundaries coming directly from the pulse recognition procedure or may be further modified, depending on the pulse requirements. The pulse shape is shifted along the pulse, so that at each step the  $M$ -th pulse shape point is aligned with an  $i$ -th pulse point, where  $i$  is confined by the fitting boundaries:  $i \in [L, R]$ . At every position the least squares optimization is performed by minimizing the sum or residuals:

$$\mathcal{R}_i \equiv \sum_{j=\max[L, i-N_1]}^{\min[R, i+N_2]} (s_j - \alpha_i p_{j-i+N_1})^2 \quad (13)$$

where by  $N_1$  and  $N_2$  we have introduced the number of pulse shape points at each side of the  $M$ -th one:

$$\begin{aligned} N_1 &= M \\ N_2 &= N - 1 - M \end{aligned} \quad (14)$$

At each alignment position an optimal multiplicative factor  $\alpha_i$  is found from the minimization requirement:  $\partial \mathcal{R}_i / \partial \alpha_i = 0$ . Introducing the following terms:

$$\begin{aligned} S_i &\equiv \sum_{j=\max[L, i-N_1]}^{\min[R, i+N_2]} s_j^2 \\ P_i &\equiv \sum_{j=\max[L, i-N_1]}^{\min[R, i+N_2]} p_{j-i+N_1}^2 \\ C_i &\equiv \sum_{j=\max[L, i-N_1]}^{\min[R, i+N_2]} s_j p_{j-i+N_1} \end{aligned} \quad (15)$$

the optimal  $\alpha_i$  may be expressed as:

$$\alpha_i = \frac{C_i}{P_i} \quad (16)$$

The quality of the fit is evaluated at each alignment point by

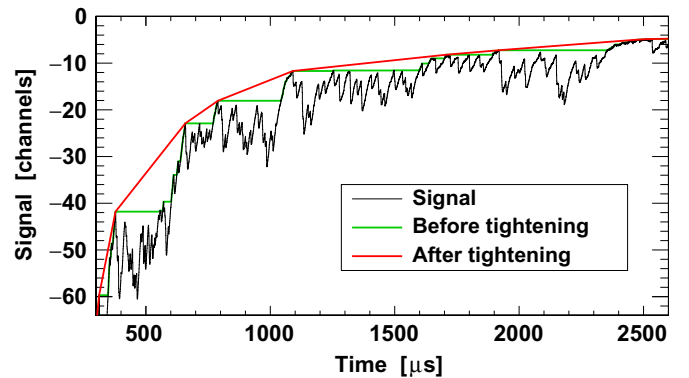
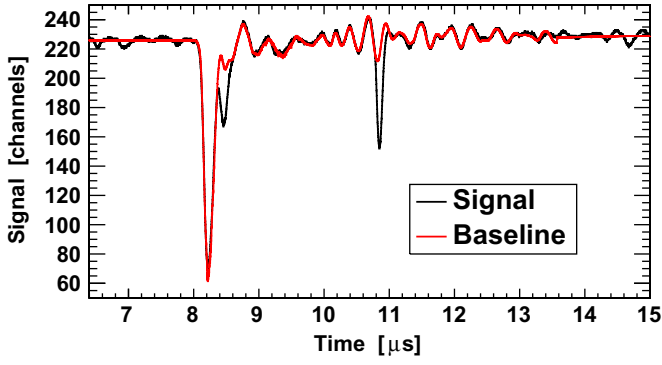


Fig. 7. Example of the signal from a gaseous  $^3\text{He}$  detector, that requires the reconstruction of the upper envelope in order to identify the baseline. The envelope is shown both before and after the tightening procedure.





**Fig. 8.** Adjustment of a distorted baseline to a signal from a MGAS detector. The horizontal adjustment is performed relative to the initial,  $\gamma$ -flash pulse. The primary (vertical) offset is identified by the constant baseline procedure.

**Table 3**

List of recursive relations for calculating the sums from Eq. (15). Different cases cover all possible combinations of summation bounds.

**Pulse shape is contained within the pulse**

$$i - N_1 > L \text{ and } i + N_2 \leq R$$

$$S_i = S_{i-1} - s_{i-1-N_1}^2 + s_{i+N_2}^2$$

$$P_i = P_{i-1}$$

**Pulse shape protrudes at the beginning of the pulse**

$$i - N_1 \leq L \text{ and } i + N_2 \leq R$$

$$S_i = S_{i-1} + s_{i+N_2}^2$$

$$P_i = P_{i-1} + p_{i-N_1}^2$$

**Pulse shape protrudes at the end of the pulse**

$$i - N_1 > L \text{ and } i + N_2 > R$$

$$S_i = S_{i-1} - s_{i-1-N_1}^2$$

$$P_i = P_{i-1} - p_{R-i+1+N_1}^2$$

**Pulse shape protrudes at both ends of the pulse**

$$i - N_1 \leq L \text{ and } i + N_2 > R$$

$$S_i = S_{i-1}$$

$$P_i = P_{i-1} + p_{L-i+N_1}^2 - p_{R-i+1+N_1}^2$$

means of a reduced  $\chi^2$ :

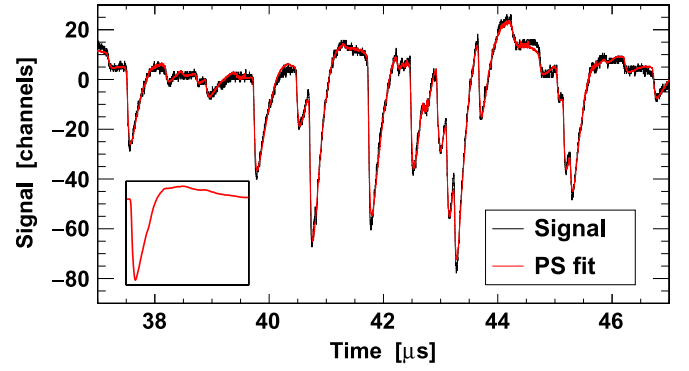
$$\chi_i^2 \propto \frac{\mathcal{R}_i}{(\min[R, i + N_2] - \max[L, i - N_1] + 1) - 2}$$

$$= \frac{S_i - C_i^2/P_i}{\min[R, i + N_2] - \max[L, i - N_1] - 1} \quad (17)$$

where the number of points taken by the fit is reduced by 2 due to 2 degrees of freedom: the horizontal and vertical alignment. A fit with a minimal reduced  $\chi^2$  is taken as the best result.

Equation (15) reveals  $\mathcal{O}(n^2)$  nature of the procedure, with  $n$  typically  $n \approx R - L$ . However, recursive relations for the terms  $S_i$  and  $P_i$  may be obtained, allowing for their calculation using only  $\mathcal{O}(n)$  operations. These relations are listed in Table 3, according to the manner in which the pulse shape and the fitted portion of the pulse are overlapped. By defining the term-wise inverted array  $\tilde{p}$  as  $\tilde{p}_i = p_{(N-1)-i}$ , it becomes evident that the final  $C_i$  term from Eq. (15) formally corresponds to a convolution of the partial signal  $s$  and a pulse shape  $p$ . In order to calculate  $C_i$  at each alignment point in a least number of operations possible, a Fast Fourier Transform algorithm – of  $\mathcal{O}(n \log n)$  computational complexity – was adopted directly from Ref. [22].

Once the best pulse shape alignment has been found by means of a minimal reduced  $\chi^2$ , the pulse shape is resampled by linear interpolation, constructing the set of  $2K$  intermediate pulse shapes  $p^{(k)}$  ( $k = \pm 1, \dots, \pm K$ ). In symbolic and self-evident notation, these



**Fig. 9.** Signal from a NaI detector characterized by a high density of piled-up pulses. The signal reconstructed by means of pulse shape fitting consists of the fitted and superimposed pulse shapes. Inset shows one of the three pulse shapes used, each adjusted to a given amplitude range.

intermediate terms may be defined as:

$$p_i^{(k)} = p_{i+k/(K+1)} \quad (18)$$

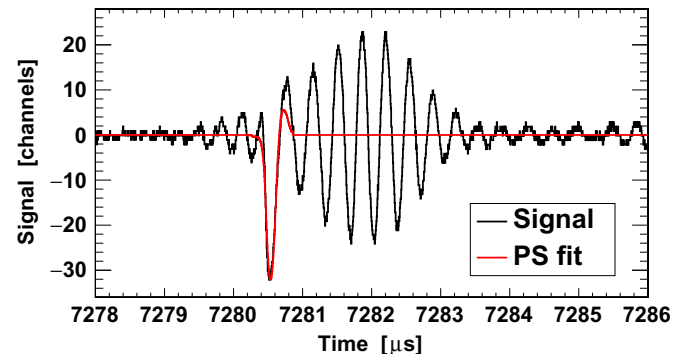
Evidently, one may treat the initial pulse shape  $p$  as the  $(2K+1)$ -th member  $p^{(0)}$ , allowing to establish the uninterrupted indexing by  $k \in [-K, K]$ . For intermediate pulse shapes the least squares adjustment by minimization of the associated Eq. (13) is performed only at the point of the best alignment of the initial pulse shape  $p^{(0)}$ , calculating the associated members from Eq. (15) by direct summation. The adjustment producing a minimal reduced  $\chi^2$  (for any  $k \in [-K, K]$ ) is kept as the final result. The value  $K=4$  has been adopted for the PSA framework described in this work.

#### 4.2. Pulse shape fitting – the saturated pulses

An important feature of the adopted pulse shape fitting routines is the exclusion of saturated points from the fitting procedure. Here, saturation is defined by the recorded signal reaching the boundaries of the data range (i.e. the minimal or maximal channel) supported by the data acquisition system (example in Fig. 2). The saturation management may be directly implemented in Eq. (13) through the introduction of appropriate weighting factors  $\theta_i$ , taking the values 0 or 1:

$$\mathcal{R}_i \equiv \sum_{j=\max[L, i-N_1]}^{\min[R, i+N_2]} \theta_j \times (s_j - \alpha_i p_{j-i+N_1})^2 \quad (19)$$

The weighting factors are given as  $\theta_i = \Theta(s_i; s_{\min}, s_{\max})$ , where we



**Fig. 10.** Example of the pulse rejection capabilities, based only on the calculated discrepancy between the signal and the adjusted pulse shape.

have introduced the following useful function:

$$\Theta(s_i; s_{\min}, s_{\max}) \equiv \begin{cases} 1 & \text{if } s_{\min} < s_i < s_{\max} \\ 0 & \text{otherwise} \end{cases} \quad (20)$$

Following the same procedure as for obtaining the expressions from Eq. (15), one arrives at the following generalized terms:

$$\begin{aligned} S_i &\equiv \sum_{j=\max[L_i-N_1]}^{\min[R_i+N_2]} \theta_j s_j^2 \\ P_i &\equiv \sum_{j=\max[L_i-N_1]}^{\min[R_i+N_2]} \theta_j p_{j-i+N_1}^2 \\ C_i &\equiv \sum_{j=\max[L_i-N_1]}^{\min[R_i+N_2]} \theta_j s_j p_{j-i+N_1} \end{aligned} \quad (21)$$

and to the corresponding expression for the reduced  $\chi^2$ :

$$\chi_i^2 \propto \frac{\mathcal{R}_i}{\sum_{j=\max[L_i-N_1]}^{\min[R_i+N_2]} \theta_j - 2} \quad (22)$$

The drawback of this generalization is immediately evident: the  $P_i$  term from Eq. (21) has become a convolution, in the same way as the  $C_i$  term, thus requiring the application of a Fast Fourier Transform, as opposed to the less computationally expensive recursive relations from Table 3 (recursive relations completely analogous to those from Table 3 may now be used only for the  $S_i$  term). Furthermore, under the assumption of properly set parameters of the data acquisition system, the saturated pulses are expected to appear only very rarely. For this reason it is advisable to keep the separate approaches – the one from Eq. (13) for unsaturated pulses and the one from Eq. (19) for saturated pulses – instead of applying the generalized and more computationally expensive procedure to both types of pulses.

### 4.3. Pulse shape fitting – the quality control

Multiple pulse shapes may be provided as input to the program. In this case the pulse shape adjustment is performed for each pulse shape separately and among all fits, the one with the minimal reduced  $\chi^2$  is kept. Allowing for the intake of multiple pulse shapes is not only beneficial to detectors exhibiting considerably differing pulses, but was also found specially suitable when the shape of the pulse varies slightly with its amplitude. Hence, among multiple pulse shapes that may be delivered, each may be best suited to a certain amplitude range. In addition, after each adjustment a fitted pulse shape is subtracted from the signal before proceeding to the next pulse in line. Thus, the pulse shape fitting is fully able to account and correct for pileup effects. Fig. 9 shows an example of a demanding signal from a NaI detector – exhibiting a persistent pileup of bipolar pulses – and a complete signal reconstruction by means of pulse shape fitting. Three separate pulse shapes were used, each adjusted to a given amplitude range. One is shown in an inset of Fig. 9.

An additional pulse shape fitting control was implemented in form of *discrepancy* – a quantity similar to the reduced  $\chi^2$ . Let the fitted pulse shape  $f$  be aligned with the pulse in the original signal  $s$ , so that the index-to-index correlation  $s_i \leftrightarrow f_i$  is established (we remind that the optimal pulse shape alignment is determined during the fitting procedure). For the total of  $Q$  pulses, let  $\alpha_q$  and  $\beta_q$  be the indices of the first and the last point of the  $q$ -th pulse ( $q \in [1, Q]$ ) in the signal. Similarly, let  $A_q$  and  $B_q$  be the first and the last index of the pulse shape aligned to the  $q$ -th pulse. The discrepancy  $D_q$  for the  $q$ -th pulse is calculated taking into account all

the pulse shape points around the fitted pulse – even if they are outside the fitting range – as long as the pulse shape does not intrude into any of the neighboring pulses. In addition, the fitted pulse shape point  $f_i$  is taken into account if and only if it is between the signal saturation boundaries  $s_{\min}$  and  $s_{\max}$ , even if the signal  $s_i$  itself is saturated. An explicit expression for the discrepancy  $D_q$  takes the form:

$$D_q^2 \equiv \frac{\sum_{i=\max[A_q, \beta_{q-1}+1]}^{\min[B_q, \alpha_{q+1}-1]} \Theta(f_i; s_{\min}, s_{\max}) \times (s_i - f_i)^2}{h^2 \times \sum_{i=\max[A_q, \beta_{q-1}+1]}^{\min[B_q, \alpha_{q+1}-1]} \Theta(f_i; s_{\min}, s_{\max})} \quad (23)$$

with  $\beta_0 = -1$  and  $\alpha_{Q+1} = P$  (where  $P$  is the total number of points comprising the signal  $s$ ). The  $\Theta$ -function is defined by Eq. (20) (note  $f_i$  in place of the first argument). If the discrepancy exceeds the preset threshold value, which is set as an external input parameter, the fit is rejected.

Central to the scaling of  $D_q$  is the pulse height  $h$  determined directly from the highest point of the baseline-corrected signal; not from the height of the fitted pulse shape. As opposed to  $\chi^2$ , the discrepancy has the following advantages:

- Due to the pulse height  $h$  replacing the signal baseline RMS, the high pulses – which are well discriminated from the baseline – are clearly favored by the lower discrepancy values, while the fits to the lower pulses are more susceptible to rejection.
- In case of any systematic difference between the given pulse shape and the pulses in the signal, the terms  $s_i - f_i$  from Eq. (23) scale with the pulse height  $h$ ; scaling the discrepancy by the same factor compensates for this effect, canceling the negative bias towards the higher pulses.

In addition, adopting the condition expressed through the  $\Theta(f_i; s_{\min}, s_{\max})$  term helps in rejecting the exaggerated fits to severely saturated pulses, such as the ones caused by an intense  $\gamma$ -flash. When such pulse is saturated for a longer time than a regular pulse would be, only the steep leading edge of the pulse is fitted, due to the exclusion of the saturated points. By rejecting these fits, a subtraction of the overscaled pulse tails is avoided during the pileup correction procedure.

Fig. 10 shows an example of the powerful pulse rejection capabilities, based only on the properly set discrepancy threshold. The single fitted pulse is clearly meaningful, since it significantly deviates from the envelope of the noise. Initially, each of the signal oscillations within a beat is recognized as a potential pulse. Since the shape of these false pulses is incompatible with the given pulse shape, the calculated discrepancy is large and the fit is rejected.

## 5. Conclusions

The most prominent features of the new pulse shape analysis framework developed for the n\_TOF-Phase3 have been described, including the pulse recognition, the baseline calculation and the pulse shape fitting procedures. The pulse recognition relies on the calculation of a custom derivative, as a difference between the signal integrals from both sides of a given point. A supporting procedure for defining the derivative crossing threshold was also described, which isolates the approximate root mean square of the derivative baseline, effectively rejecting the contribution from the beats and actual pulses, while avoiding the dependence on the well defined number of clear presamples.

Three different baseline calculation procedures have been adopted. The simplest one is the constant baseline, which requires a single pass through a signal, without any need for iterative techniques. One of two adaptive baseline options relies on the weighted averaging of the signal, being appropriate when clear portions of the baseline are indeed at hand. The second option is appropriate when this condition is not met – due to persistent pileup of pulses, completely concealing the baseline – and no *a priori* knowledge about the baseline is available. In this case the baseline is found as the upper signal envelope, since all regular pulses are treated as negative. In case some *a priori* knowledge of the baseline is available – coming from a consistent detector response to an intense  $\gamma$ -flash – the baseline distortion may be identified in a form of an appropriate pulse shape and may be subtracted from the signal, but only after correcting for the primary baseline offset.

The most basic implementations of previous procedures are of  $\mathcal{O}(N \times P)$  computational complexity, with  $P$  as the total number of points in a digitized signal waveform and  $N$  as a characteristic filter width of arbitrary size. Single waveforms recorded by the digital data acquisition system at n\_TOF may, at present, reach the order of magnitude of  $10^8$  points. Hence, the  $\mathcal{O}(N \times P)$  complexity constitutes a significant performance issue that had no alternative but to be resolved. Therefore, for all such procedures fast recursive algorithms were implemented, bringing the computational complexity to the  $\mathcal{O}(P)$ , or at least to the approximate  $\mathcal{O}(P)$  level. For the reasons of computational efficiency the pulse shape fitting routine was also described, though the procedure itself is well established. By the virtue of a complete *a priori* knowledge of the pulses, the pulse shape fitting procedure allows us to subtract the adjusted pulse shapes from the signal, thus correcting for pileup effects and restoring both the energy and timing resolution of the detectors which are considerably affected by pileup.

## Acknowledgments

This work was supported by the Croatian Science Foundation under Project No. 1680.

## Appendix A. Moving maximum code

Table A1 presents a computationally efficient C++ code for finding the upper envelope of a signal. The code is a simplified version of the one proposed in Ref. [25]. The array `signal` contains the signal. The external parameters `N`, `start_at` and `stop_at` define, respectively: the moving window width, the starting point and stopping point ( $0 \leq \text{start\_at} < \text{stop\_at} \leq P-1$ ) of the fraction of the waveform to be taken into account. The arrays `max`, `max_forwards` and `max_backwards` are of the same length as the array `signal` (thus establishing one-to-one correspondence between the array terms; if necessary, the code can also be adjusted so as to use only `stop_at-start_at+1` points for the array `max` and to completely avoid arrays `max_forwards` and `max_backwards`). At the end of the procedure, the baseline, i.e. the signal envelope is stored in array `max`.

Table A2 presents the code for tightening the envelope obtained using the procedure from Table A1. The main inputs to this code are an array `x` of positions of signal points and an array `max` from the previous procedure. As before, arrays `max_forwards` and `max_backwards` are only used as convenient temporary storage. The code proceeds by identifying the nodes, which define the locally steepest lines, when drawn from a previous node. The set of nodes is, in general, different when searched from the beginning or the end of the waveform. It was empirically found that the initialization `last_slope=0` from line 16 is specially favorable – in contrast to initializations to extreme values – improving the quality of the tightened baseline. The nodes are then checked for a maximum of the slope between them, within a window of a preset width. If no node is contained within this window, the next available node is used. It is to be noted from lines 29 and 37 that the same moving window width `N` was used for this procedure as for finding the initial (untightened) envelope. From the results obtained by going forwards and backwards through the waveform, a final tightened one is determined as the pointwise maximum between the two.

Table A1

Simplified version of the code from Ref. [25], adopted for the calculation of the moving maximum. Code input consists of the array `signal` and the integer parameters `N`, `start_at` and `stop_at`. Arrays `max`, `max_forwards` and `max_backwards` are to be initialized in advance, having the same number of points as the array `signal`. At the end of the procedure, array `max` holds the signal envelope as the final result.

	Forward maximum	Backward maximum
1		
2		
3		
4	<code>U_first=0;</code>	<code>U_first=0;</code>
5	<code>U_last=0;</code>	<code>U_last=0;</code>
6	<code>U[U_first]=start_at;</code>	<code>U[U_first]=stop_at;</code>
7	<code>for (i=start_at; i&lt;= stop_at; i++) {</code>	<code>for (i=stop_at; i&gt;= start_at; i--) {</code>
8	<code>if (U[U_first]==i-N)</code>	<code>if (U[U_first]== i + N)</code>
9	<code>U_first++;</code>	<code>U_first++;</code>
10	<code>while (U_last&gt;= U_first &amp;&amp;</code>	<code>while (U_last&gt;= U_first &amp;&amp;</code>
11	<code>signal[i]&gt;= signal[U[U_last]])</code>	<code>signal[i]&gt;= signal[U[U_last]])</code>
12	<code>U_last--;</code>	<code>U_last--;</code>
13	<code>U[+ + U_last] = i;</code>	<code>U[+ + U_last] = i;</code>
14	<code>max_forwards[i] = signal[U[U_first]];</code>	<code>max_backwards[i] = signal[U[U_first]];</code>
15	<code>}</code>	<code>}</code>
16	<code>for (i = start_at; i&lt;= stop_at; i++) max[i]=MIN(max_forwards[i],max_backwards[i]);</code>	

Table A2

Code for tightening the signal envelope calculated by the code from Table A1. The final result is again stored in the array **max**, i.e. its contents are overwritten.

Forward tightening	Backward tightening
1	<b>#define MAX(a,b) (a &gt; b?a : b)</b>
2	<b>int ij,last,node,NODES;</b>
3	<b>int index[(const int)(stop_at–start_at+1)];</b>
4	<b>double slope,last_slope,past_slope,A,B;</b>
5	<b>index[0]=start_at;</b>
6	<b>index[0]=stop_at;</b>
7	<b>NODES=1;</b>
8	<b>NODES=1;</b>
9	<b>last_slope=–1.e300;</b>
10	<b>last_slope=1.e300;</b>
11	<b>past_slope=–1.e300;</b>
12	<b>past_slope=1.e300;</b>
13	<b>last=start_at;</b>
14	<b>last=stop_at;</b>
15	<b>for (i=start_at–1; i &gt;= start_at; i––) {</b>
16	<b>  slope = (max[i]–max[last])/(x[i]–x[last]);</b>
17	<b>  for (i=stop_at+1; i &lt;= stop_at; i++) {</b>
18	<b>    slope = (max[i]–max[last])/(x[i]–x[last]);</b>
19	<b>    if (last_slope &gt; past_slope &amp;&amp;</b>
20	<b>      last_slope &gt; slope) {</b>
21	<b>      index[NODES++] = i–1;</b>
22	<b>      last=i–1;</b>
23	<b>      last_slope=0;</b>
24	<b>      slope=(max[i]–max[last])/(x[i]–x[last]);</b>
25	<b>    }</b>
26	<b>    past_slope = last_slope;</b>
27	<b>    last_slope = slope;</b>
28	<b>  }</b>
29	<b>index[NODES++] = stop_at;</b>
30	<b>last=0;</b>
31	<b>for (i=1; i &lt; NODES; i++) {</b>
32	<b>  if (i == last + 1) {</b>
33	<b>    A=(max[index[i]]–max[index[last]])/</b>
34	<b>      (x[index[i]]–x[index[last]]);</b>
35	<b>    node=i;</b>
36	<b>  } else if (index[i]–index[last] &lt;= N) {</b>
37	<b>    slope=(max[index[i]]–max[index[last]])/</b>
38	<b>      (x[index[i]]–x[index[last]]);</b>
39	<b>    if (slope &gt;= A) {</b>
40	<b>      A=slope;</b>
41	<b>      node=i;</b>
42	<b>    }</b>
43	<b>  if (index[i]–index[last] &gt;= N    i == NODES–1) {</b>
44	<b>    B=max[index[last]]–A*x[index[last]];</b>
45	<b>    for (j=index[last]; j &lt;= index[node]; j++)</b>
	<b>      max_forwards[j]=A*x[j]+B;</b>
	<b>    last=node;</b>
	<b>    i=last;</b>
	<b>  }</b>
	<b>  for (i=start_at; i &lt;= stop_at; i++) max[i]=MAX(max_forwards[i],max_backwards[i]);</b>

## References

- [1] C. Rubbia, S. Andriamonje, D. Bouvet-Bensimon, et al., A high resolution spallation driven facility at the CERN-PS to measure neutron cross sections in the interval from 1 eV to 250 MeV, CERN/LHC/98-02, 1998.
- [2] C. Rubbia, S. Andriamonje, D. Bouvet-Bensimon, et al., A high resolution spallation driven facility at the CERN-PS to measure neutron cross sections in the interval from 1 eV to 250 MeV, CERN/LHC/98-02-Add. 1, 1998.
- [3] E. Chiaveri, Proposal for n\_TOF experimental area 2 (EAR-2), CERN-INTC-2012-029/INTC-O-015, 2012.
- [4] C. Weiß, E. Chiaveri, S. Girod, et al., Nuclear Instruments and Methods in Physics Research A 799 (2015) 90.
- [5] S. Barros, I. Bergström, V. Vlachoudis, C. Weiß, Journal of Instrumentation 10 (2015) P09003.
- [6] C. Guerrero, A. Tsinganis, E. Berthoumieux, et al., The European Physical Journal A 49 (2013) 27.
- [7] S. Marrone, P.F. Mastinu, U. Abbondanno, et al., Nuclear Instruments and Methods in Physics Research A 517 (2004) 389.
- [8] C. Weiß, E. Griesmayer, C. Guerrero, et al., Nuclear Instruments and Methods in Physics Research A 732 (2013) 190.
- [9] C. Guerrero, U. Abbondanno, G. Aerts, et al., Nuclear Instruments and Methods in Physics Research A 608 (2009) 424.
- [10] R. Plag, M. Heil, F. Käppeler, et al., Nuclear Instruments and Methods in Physics Research A 496 (2003) 425.
- [11] Y. Giomataris, Ph. Rebourgeard, J.P. Robert, G. Charpak, Nuclear Instruments and Methods in Physics Research A 376 (1996) 29.
- [12] S. Andriamonje, M. Calviani, Y. Kadi, et al., Journal of Korean Physical Society 59 (2011) 1597.
- [13] D.B. Gayther, Metrologia 27 (1990) 221.
- [14] C. Paradela, L. Tassan-Got, L. Audouin, et al., Physical Review C 82 (2010) 034601.
- [15] U. Abbondanno, G. Aerts, F. Álvarez, et al., Nuclear Instruments and Methods in Physics Research A 538 (2005) 692.
- [16] E. Berthoumieux, Preliminary Report on BaF<sub>2</sub> Total Absorption Calorimeter Test Measurement, Rap. Tech. CEA-Saclay/DAPNIA/SPhN, 2004.
- [17] W. Xiao, A.T. Farsoni, H. Yang, D.M. Hamby, Nuclear Instruments and Methods in Physics Research A 769 (2015) 5.
- [18] R.J. Cooper, D.C. Radford, K. Lagergren, et al., Nuclear Instruments and Methods in Physics Research A 629 (2011) 303.
- [19] S.N. Liddick, I.G. Darby, R.K. Grzywacz, Nuclear Instruments and Methods in Physics Research A 669 (2012) 70.
- [20] C. Guerrero, D. Cano-Ott, M. Fernández-Ordóñez, et al., Nuclear Instruments and Methods in Physics Research A 597 (2008) 212.
- [21] J. Kamleitner, S. Coda, S. Gnesin, Ph. Marmillod, Nuclear Instrumentation and Methods in Physics Research A 736 (2014) 88.
- [22] W.H. Press, S.A. Teukolsky, W.T. Vetterling, B.P. Flannery, Numerical Recipes in C: the Art of Scientific Computing, Second ed., Cambridge University Press, Cambridge, 1992.
- [23] F.J. Harris, Proceedings of the IEEE 66 (1) (1978) 51.
- [24] M. Niedziwiecki, W.A. Sethares, IEEE Transactions on Signal Processing 43 (1) (1995) 1.
- [25] D. Lemire, J. Nord. Comput. 13 (4), 2006, 328.
- [26] A. Tsinganis, E. Berthoumieux, C. Guerrero, et al., Nuclear Data Sheets 119 (2014) 58.