



Contents lists available at [SciVerse ScienceDirect](http://www.sciencedirect.com)

## Medical Engineering & Physics

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



# A novel non-linear recursive filter design for extracting high rate pulse features in nuclear medicine imaging and spectroscopy

Salar Sajedi<sup>a,b</sup>, Alireza Kamal Asl<sup>e</sup>, Mohammad R. Ay<sup>a,c,d,\*</sup>, Mohammad H. Farahani<sup>a,b</sup>, Arman Rahmim<sup>f,g</sup>

<sup>a</sup> Medical Imaging Systems Group, Research Center for Molecular and Cellular Imaging, Tehran University of Medical Sciences, Tehran, Iran

<sup>b</sup> Parto Negar Persia Co., Tehran, Iran

<sup>c</sup> Department of Medical Physics and Biomedical Engineering, Tehran University of Medical Sciences, Tehran, Iran

<sup>d</sup> Research Institute for Nuclear Medicine, Tehran University of Medical Sciences, Tehran, Iran

<sup>e</sup> Faculty of Medical Radiation Engineering, Shahid Beheshti University, Tehran, Iran

<sup>f</sup> Department of Radiology, School of Medicine, Johns Hopkins University, Baltimore, USA

<sup>g</sup> Department of Electrical and Computer Engineering, School of Engineering, Johns Hopkins University, Baltimore, USA

### ARTICLE INFO

#### Article history:

Received 22 December 2011

Received in revised form 4 June 2012

Accepted 8 August 2012

#### Keywords:

High rate spectroscopy

Nuclear medicine

Pulse processing

Non-linear recursive filter

Pulse pile-up

Time-of-flight PET

### ABSTRACT

Applications in imaging and spectroscopy rely on pulse processing methods for appropriate data generation. Often, the particular method utilized does not highly impact data quality, whereas in some scenarios, such as in the presence of high count rates or high frequency pulses, this issue merits extra consideration. In the present study, a new approach for pulse processing in nuclear medicine imaging and spectroscopy is introduced and evaluated. The new non-linear recursive filter (NLRF) performs nonlinear processing of the input signal and extracts the main pulse characteristics, having the powerful ability to recover pulses that would ordinarily result in pulse pile-up. The filter design defines sampling frequencies lower than the Nyquist frequency.

In the literature, for systems involving NaI(Tl) detectors and photomultiplier tubes (PMTs), with a signal bandwidth considered as 15 MHz, the sampling frequency should be at least 30 MHz (the Nyquist rate), whereas in the present work, a sampling rate of 3.3 MHz was shown to yield very promising results. This was obtained by exploiting the known shape feature instead of utilizing a general sampling algorithm. The simulation and experimental results show that the proposed filter enhances count rates in spectroscopy. With this filter, the system behaves almost identically as a general pulse detection system with a dead time considerably reduced to the new sampling time (300 ns). Furthermore, because of its unique feature for determining exact event times, the method could prove very useful in time-of-flight PET imaging.

© 2012 IPPEM. Published by Elsevier Ltd. All rights reserved.

## 1. Introduction

A range of pulse processing methods may be utilized for imaging and spectroscopy application; at the same time, in the presence of high count rates, traditional methods perform poorly, and new methods need to be developed. For the wide range of imaging cameras and spectrometers in the field [1], initial works involved the use of analog primitives [2], followed by more developed analog circuits [3,4]. This was followed by the use of digital signal processors [5–7], opening the doorway for the synthesis of optimum filters for radiation detection systems [8]. Signals (e.g. obtained from systems of detectors followed by photomultiplier tubes

(PMTs)) have an exponential decay time as well as associated noise resulting in imprecision in feature extraction from the original signal. Many efforts have been made to enhance signal quality (e.g. the spatial and energy resolutions achieved) by means of reducing noise [3,8], pile-up [9–18], baseline drift [19–23] and ballistic deficit [24,25]. Because signal quality deteriorates with increasing count rates, such techniques commonly seek a compromise. Some methods achieve algorithm enhancement using the off-line mode [26], which is not practical in many imaging applications because of the high throughput of the raw data. Clearly by increasing the speed and performance of digital circuits, there is a much greater opportunity for the introduction of new methods. The present study aims to introduce a novel nonlinear algorithm for extraction of features from high rate nuclear pulses; e.g. as used in nuclear medicine imaging. The proposed algorithm is applicable to real-time signal processing and data acquisition. This filter was implemented via a prototype hardware platform for a scintillation camera to test the potentials for practical performance. The implementation

\* Corresponding author at: Department of Medical Physics and Biomedical Engineering, Tehran University of Medical Sciences, Tehran, Iran. Tel.: +9821 64053255; fax: +9821 64053255.

E-mail address: [mohammadreza.ay@tums.ac.ir](mailto:mohammadreza.ay@tums.ac.ir) (M.R. Ay).

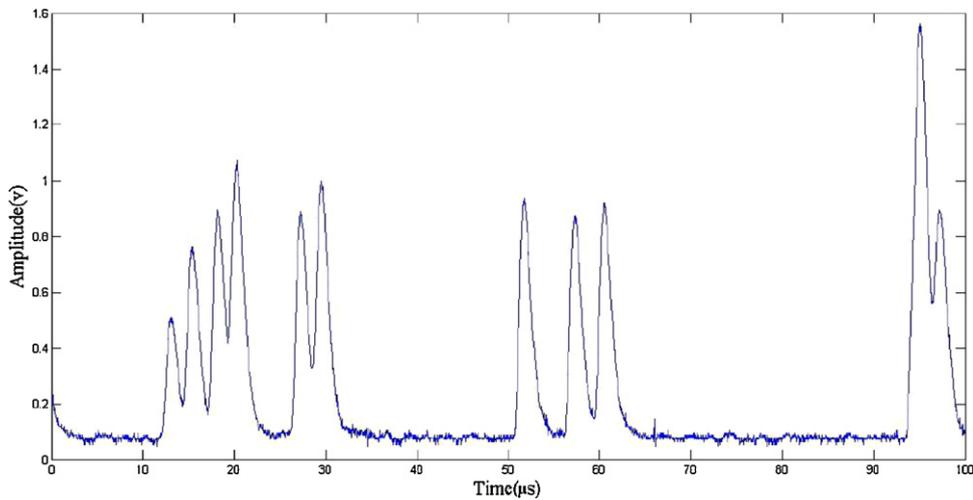


Fig. 1. A real signal shape captured from the detector of Starcam 2000 (GE Healthcare Technologies, Waukesha, WI).

algorithm is simple, requires low cost sources, and significantly reduces the required sampling rate for accurate pulse processing. By allowing the possibility to process high rate pulses, the algorithm is able to minimize pulse pile up, and as such, it is very useful for high count rate spectroscopic and imaging application, including for use in time-of-flight (TOF) PET imaging. The proposed approach was validated by simulations as well as experimental data.

## 2. Method

The NLR filter applies to signals which can be sampled from the time response of a detector. In a radiation detection system consisting of scintillation detectors and PMTs, the input signal can be considered as a series of sharp spikes whose heights represent the energy deposited in the detector. In the absence of noise, the signal recorded by a system  $s(t)$  is the convolution of the input signal  $i(t)$  and the system response function  $h(t)$  [10]:

$$s(t) = i(t) \times h(t) \quad (1)$$

The input signal  $i(t)$  to the PMT is essentially a narrow pulse whose amplitude is determined by the total collected energy in the detector; by contrast, the shape (not amplitude) of  $s(t)$  is determined by the system response function  $h(t)$ . The algorithm, exploiting the constant shape of the input signal, aims to extract, from  $s(t)$ , the desired signal features such as the signal amplitude and the relative occurrence time with respect to the sampling point. An example of a detector signal captured from a Starcam 2000 gamma camera (GE Healthcare Technologies, Waukesha, WI) is shown in Fig. 1.

The proposed NLR filter is based on the fact that the response of a detector follows a deterministic shape pattern [2] wherein the shape of the signal scales with amplitude and moves with time. The algorithm uses non-linear reverse functions to obtain amplitude and time of occurrence utilizing two successive samples from the input signal. Furthermore, it uses a recursive structure in which values from previous pulses are subtracted from the current pulse in order to lower the pile-up effect (see Section 2.3). Pulses from scintillation detectors have been successfully modeled including using single or double exponential decay time [9]. The proposed algorithm has been implemented for both models in the context of different detectors.

### 2.1. NLR filter algorithm

The main idea behind the design of the current filter is that two successive measurements of the pulse amplitude ( $x, y$ ) sampled

with a known time difference  $\Delta t_s$  can generate a pair of values that uniquely define (i) the maximum signal amplitude ( $a$ ) and (ii) the time  $t$  relative to pulse onset at which the first sample was obtained (which we refer to as the phase). In the other words, it can be shown that for a range of deterministic pulse shapes of the detector output, the values of the maximum signal amplitude ( $a$ ) and time of first sample ( $t_1$ ) constitute a one-to-one mapping to the values of the two successive samples ( $x, y$ ) of the input signal, as obtained with a constant known time difference of  $\Delta t_s$ . Although this mapping operation is non-linear, it can be implemented using a “lookup table” (LUT). Fig. 2 shows the definition of  $x, y, a, t_1$ , and  $\Delta t_s$  in a detector output pulse.

The above mapping is possible for a range of pulse shape models  $f(a, t)$  when the following uniqueness condition is valid:

$$\text{if } \begin{cases} x = f(a, t_1) = f(a', t'_1) \\ y = f(a, t_1 + \Delta t_s) = f(a', t'_1 + \Delta t_s) \end{cases} \xrightarrow{\text{yields}} a = a' \text{ and } t_1 = t'_1 \quad (2)$$

This is true for a wide range of function. At the same time, in the present work, without loss of generality, we considered the following pulse shape formulation:

$$f(a, t) = a, t^\alpha \cdot e^{t/\tau} \quad (3)$$

where  $\alpha$  and  $\tau$  are constants (which for our particular sample detectors had values of 4.3 and 160 ns, respectively). For this particular mode, proof for the abovementioned unique mapping is straightforward:

$$\begin{aligned} \frac{x}{y} &= \frac{t_1^\alpha}{(t_1 + \Delta t_s)^\alpha} \cdot \frac{e^{t_1/\tau}}{e^{(t_1 + \Delta t_s)/\tau}} = \frac{t_1^\alpha}{(t_1 + \Delta t_s)^\alpha} \cdot \frac{e^{-(t_1/\tau)}}{e^{-(t_1 + \Delta t_s)/\tau}} \\ &\xrightarrow{\text{yields}} \frac{t_1^\alpha}{(t_1 + \Delta t_s)^\alpha} \cdot e^{t_1/\tau} = \frac{t_1'^\alpha}{(t_1' + \Delta t_s)^\alpha} \cdot e^{t_1'/\tau} \end{aligned} \quad (4)$$

It then follows that:

$$\frac{t_1}{(t_1 + \Delta t_s)} = \frac{t_1'}{(t_1' + \Delta t_s)} \quad (5)$$

which can be simplified to:  $t_1 \Delta t_s = t_1' \Delta t_s$ . And since  $\Delta t_s \neq 0$ , one concludes that  $t_1 = t_1'$  and consequently  $a = a'$ . The same conclusion may similarly be reached for a model with two exponential decay times.

Subsequently, there exists an inverse function which can calculate values of  $a$  and  $t_1$  from  $x$  and  $y$ :

$$(a, t_1) = g(x, y) \quad (6)$$

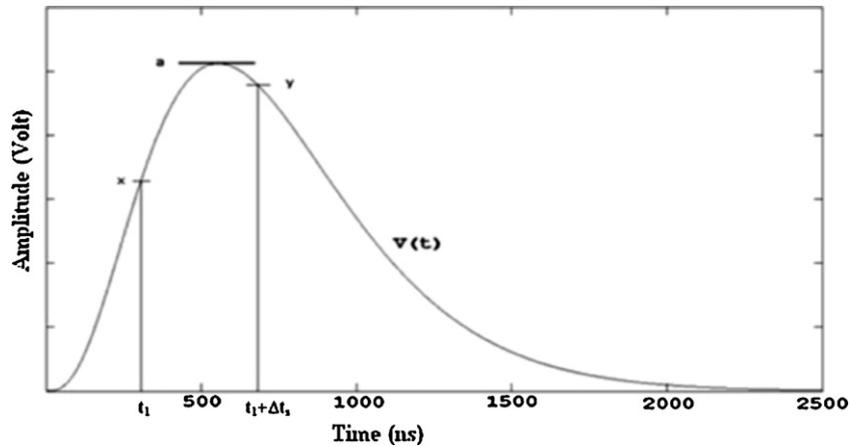


Fig. 2. An example of a detector output and digital sampling.

We separate this function into two distinct components as follows:

$$a = g_0(x, y) \tag{8}$$

$$t_1 = g_1(x, y) \tag{9}$$

As an example, for the abovementioned specific model (3), the following inverse relation may be obtained:

$$t_1 = \frac{c \Delta t_s}{1 - c} \tag{10}$$

where  $c$  is a function of  $x/y$  as follows:

$$c = e^{(\Delta t_s / \tau \alpha)} \cdot \left(\frac{x}{y}\right)^{1/\alpha} \tag{11}$$

where  $\alpha$  and  $\tau$  are the model parameters. Then  $a$  can be calculated as:

$$a = \frac{x}{t_1^\alpha \cdot e^{t_1/\tau}} \tag{12}$$

In the case of more complicated pulse models, exact inverse functions may be very difficult to obtain. Alternatively, a practical implementation of these functions can be achieved by computing look-up tables (LUTs) and storing them for subsequent use. Such LUTs would map a given  $(x, y)$  pair to its corresponding  $(a, t_1)$  pair. In implementation this can be done by utilizing memory blocks. The number of bits used to digitize the measurements and the function values determines the memory volume and detection precision. Simulations show that the precision of the function value (stored

in memory) is best kept the same as analog value digitization; e.g. 11 bit digitization by an ADC converter; 11 bit precision function value. An example of such mapping  $g_0 = a$ , according to inversion of model (3) as achieved in Eqs. (10)–(12), is depicted in Fig. 3.

As can be seen, every curve in Fig. 3 is the locus of all  $x$  and  $y$  values corresponding to the same maximum signal amplitude  $a$ . So by referring to this table for a given pair of  $x$  and  $y$  values sampled from the detector output pulse, the value of the maximum signal amplitude may be straightforwardly estimated.

For a sampling resolution of 11 bits for  $x$  and  $y$  (thus 22 bits in total) and 11 bits for the final maximum amplitude value in the LUT, the data for this table occupy  $2^{22} \times 11$  or 44 Mbits. A similar map can be constructed for the value of  $t_1$ , though for applications involving the gamma camera and SPECT, exact determination of this value is not critical; but for PET, accurate determination of  $t_1$  is very valuable to allow for smaller coincidence time-windows (minimizing the effect of randoms), and furthermore, more readily enabling TOF PET imaging [27].

### 2.2. Map construction

Determination of inverse functions for a certain pulse model readily enables the construction of the LUT for mapping. Nevertheless, if a solution for the inverse function does not exist in analytical form, the inverse map may still be constructed: this can be achieved by sweeping the full range of  $(a, t_x)$  pairs and calculating the

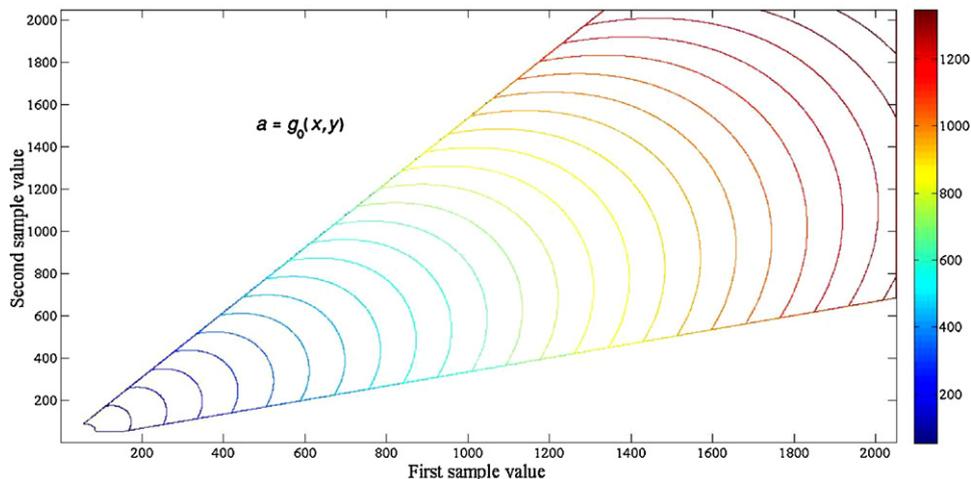


Fig. 3. Contour of amplitude, for the response pulse model (3) with a 11 bit resolution A/D converter.

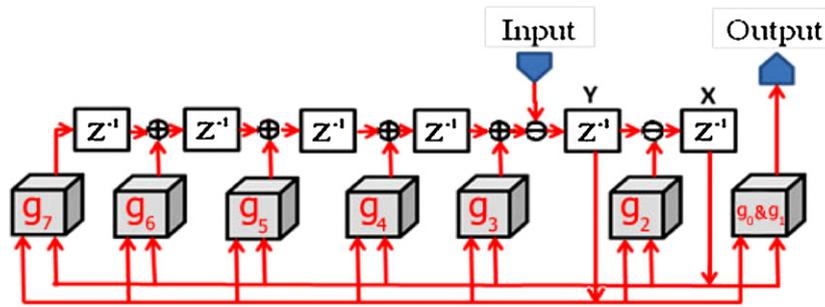


Fig. 4. Depiction of the implemented NLR filter design algorithm for nuclear pulse detection ( $Z^{-1}$  represents one cycle delay).

corresponding  $(x, y)$  pairs, and filling in the inverse mapping space. The sweeping resolution can be made small enough to minimize unfilled memory locations. Furthermore, the values for the remaining unfilled locations can be filled-in via interpolation methods.

For each pulse model there exist unfilled areas in the map indicating that such combinations of  $(x, y)$  may not be obtained for the well-defined pulse shapes. By zeroing these areas, all such pulses were rejected as noise. Furthermore, for the reason that a first sample  $x$  very near to the rising edge of the pulse may involve inaccuracies, for better performance, LUT addresses with relatively small  $x$  values were zeroed in the map (left side of contour in Fig. 3 illustrates this). In such a case, the very next sample (that was going to be the second sample) is instead taken as the first sample in the map, which coupled by the next sample, retrieves the appropriate LUT value.

Construction of the LUT typically occupies only few minutes (e.g. for 11 bit resolution) using an Intel dual core T6650 processor. In addition, once the LUT construction is finished, it does not pose an additional computational burden to subsequent look-up tasks.

### 2.3. Pileup recovery

By the abovementioned mapping procedure, the necessary pulse information is obtained. In the case of pulse pileup, a powerful and critical step that we subsequently propose and implement is to compute the corresponding pulse tail from the present signal and

to subtract it from the subsequent ones. In other words, knowing the present two samples leads to determination of the pulse shape, and as a result, it allows extrapolation of contributions to subsequent samples. Clearly, the pulse amplitude will become negligible after certain duration (depending on the extent of the pulse tail and the sampling time).

We first define sampled pulse tails as:

$$\text{Tail}_n = f(a, t_1 + \Delta t_s + n \cdot \Delta t_s), \quad n = 1, 2, \dots, N \quad (13)$$

For the specified model (3) and choosing  $\Delta t_s = 300$  ns, following the two main samples, the 6th tail was seen to result in negligible amplitude equal to 0.1% of the maximum amplitude of the pulse. Furthermore, similar to the determination of the maximum amplitude values, the pulse tails may also be determined using inverse mapping, which can similarly be performed using LUTs. The volume of each LUT is the same as the one for the maximum amplitude. The filling algorithm for tail LUTs is straightforward, and by knowing the  $(x, y)$  pair, all tails can be calculated by (13). Furthermore, we note that the tail maps were only filled in areas in which the amplitude map had values and was not zero. In fact, with this routine, we map the  $(x, y)$  space to the  $(n + 1)$ -dimensional space:  $(a, \text{Tail}_1, \text{Tail}_2, \dots, \text{Tail}_N)$ .

The suggested algorithm for implementing the NLR filter with 6 tail values ( $g_2 - g_7$ ) is described in Fig. 4. The implementation is based on transversal filter implementation. The tail functions generate filter outputs along pulse tails. Tails accumulate after a

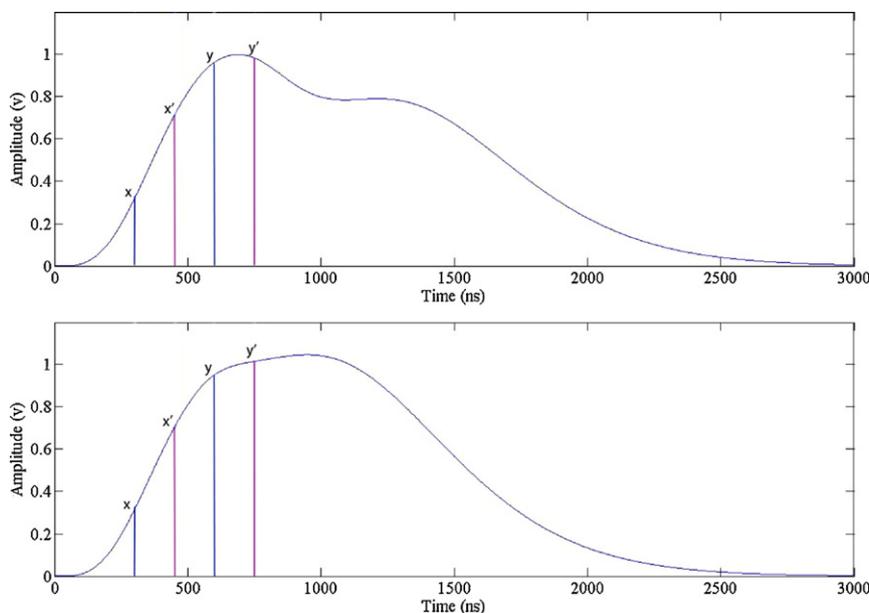


Fig. 5. Interleave sampling on (top) recoverable condition and (down) unrecoverable condition. (For interpretation of the references to color in the text, the reader is referred to the web version of this article.)

sequence of delays, according to the time difference with the samples and in accordance to the associated model, and then will be subtracted from each given sample. Thus, after detecting a pulse, tails of that pulse will be removed from the following signals, and the measured values will be processed without affecting the previous events. The analog version of this structure is elaborated in Refs. [14,34].

#### 2.4. Improvement techniques

The sampling time  $\Delta t_s$  used by the filter should be determined before applying the filter. Simulation results show that for a certain model, the filter has its best performance with sampling times about the detector time constant ( $\sim 300$  ns for NaI(Tl)). As Fig. 12 demonstrates, by changing the sampling time around this value, filter performance (FWHM) did not change radically.

By reducing the sampling times, the number of tails which will have useful information on a given pulse will increase, which requires further memory allocation. In fact, the lower limit for the sampling time is available memory volume to store the tails' LUTs. At the same time, another consideration for selecting the sampling time is the pile-up rejection ratio, as we describe next. Consider two nearly spaced pulses (what we mention below can be easily generalized for  $>2$  pulses also): if two successive samples spaced by  $\Delta t_s$  belong to the first pulse only, the pulse (and its tails), and therefore the next pulse, will consequently be recovered. It follows that if the time between two successive pulses is more than  $2\Delta t_s$ , then both pulses will be properly characterized. On the other hand, if the time between the two successive pulses is less than  $\Delta t_s$ , then, the first pulse (and its tails) may not be correctly characterized, and as such, both pulses will not be properly identified. In the intermediate range where successive pulses are separated by more than  $\Delta t_s$  but less than  $2\Delta t_s$ , success in characterization will not be definitive and will depend on the particular phase of sampling with respect to the pulses, and the corresponding likelihood of correct characterization will increase with pulse separations approaching  $2\Delta t_s$ .

In such cases as in the present work, when the sampling rate can be increased by the hardware and the constraint is noise immunity in the algorithm, we explored the use of interleaved sampling to potentially further enhance the signal. In this approach, two filter implementations work independently and in parallel on the input signal, via interleaved samples. The ADC performs sampling at twice the frequency (6.6 MHz in this work) and the samples are passed on to the two NLR filters successively. Because the filters work on the same pulse (including abovementioned pulse pile-up recovery), they should yield similar outputs for a pulse; however, if the two outputs vary significantly, this is taken as an indication that unrecovered pile-up has occurred in the pulse. As such, the filter outputs are compared and if they are not approximately equal, they are rejected. Fig. 5 illustrates two pile-up cases one leading to recovery and the other to rejection. Two filters work in parallel on the measured values. Each filter uses samples as indicated with the same color in the figure. In the top figure, when all 4 samples belong to one pulse, both filters result in nearly similar output amplitudes. By contrast, in the bottom figure, the second pulse occurs within filter detection. In this case, blue samples belong to the first pulse and detect the peak amplitude, but red samples are affected by the arrival of the second pulse and will not detect the correct peak amplitude. By this approach, the filter has the ability to feasibly reject (and thus avoid mischaracterizing) pile-up events arising from two pulses separated by less than 450 ns (1.5 times the NLRF sampling rate). The results are further discussed in Section 4.2, showing that interleaved NLRF reduces registered counts and lowers sensitivity (by rejecting pile-up events) while achieving

improvements in the energy resolution of the spectrum especially at very high count rates.

### 3. Hardware

The implemented hardware in this work has been designed for a gamma-camera detector whose pulse profile was described in the previous section. Because of flexibility of design, data from each PMT are acquired by separate designated data acquisition cards. Then pulses are collected in the main board and are finally transmitted to the PC.

Input of this filter is generated by the analog amplifier circuit. The prototype R1847-12 PMT (Hamamatsu Photonics Co., Japan) was utilized. For the analog section, we used two-stage current-to-voltage amplification, followed by a differential ADC driver as shown in Fig. 6. The pulse height is about 200 mV, having a peak time at 688 ns and an overall extent of  $\sim 3$   $\mu$ s (in further upgrades to the hardware, one may shorten the pulse length, which we aim to implement, and accordingly modify the fitted model, to obtain higher rate acquisition and processing capabilities). The conversion resolution in this work is 11 bits. Subsequently, as described in Section 2.2, each LUT needs 44 Mbits memory; we considered 6 tail samples, and as such, we needed 7 LUTs resulting in 308 Mbits memory. In the case that application requires exact occurrence time, 1 LUT shall be added up and total memory will be 352 Mbits. Because of large volume and fast access time of SDRAMs we used a 512Mbit SDR with a 16 bit data bus chip.

For a pulse shape with a peak time of 688 ns (Fig. 11), optimal performance was obtained (see Section 4.1) using  $\Delta t_s \sim 300$  ns (i.e. sampling frequency about 3.3 MHz). At the same time, increasing this value gradually was seen to only minimally affect system performance.

Fig. 7 shows the hardware structure of an acquisition card. Main processing unit is implemented on a field programmable gate array (FPGA). The input signal from the PMT or the positioning circuit is fed through amplification to meet suitable voltage span needed for the ADC converter. The AD9224 (Analog Device, USA) with up to 40 MSPS is used with an S&H filter. As mentioned above, the selected frequency for this configuration is 3.3 MSPS (i.e. 3.3 MHz; 300 ns), nonetheless, our reason for maintaining high frequency capability is to have the ability for this hardware to work as a general acquisition card via firmware modification which is required for curve fitting and analysis in PC; and secondly, to permit implementation and comparison of conventional algorithms which require higher sampling rate, for comparison purposes.

The digital signal processing part is made by SPARTAN3 FPGA from Xilinx. This chip is chosen because of its availability and low cost. The other options in FPGA chips such as speed and volume were not critical. The processing unit consists of two parts: the data acquisition unit and the memory filling unit. The memory filling unit initializes all LUTs at the power-up and subsequently goes offline. After initialization, the data acquisition unit gathers the data and transfers them to the main processor. These units, plus the memory controller, occupy only 28% of the XCS3200-144 (Xilinx, San Jose, CA, USA) FPGA.

All LUTs explained in the text are implemented within one RAM block. Selected SDRAM has a 512 Mbit capacity which currently is the maximum memory volume in one chip. In each sampling clock cycle, the amplitude and all tail values should be read from SDRAM. In this configuration, this leads to reading  $7 \times 11$  bits in each sampling clock cycle (in the case without  $t$ -map extraction). If we execute this process in sequential form, it would confine the sampling time. Since SDRAM data bus is 16 bit, 5 reading cycles should be performed within each sampling time. The burst reading mode in SDRAM chips facilitates this procedure. For increasing

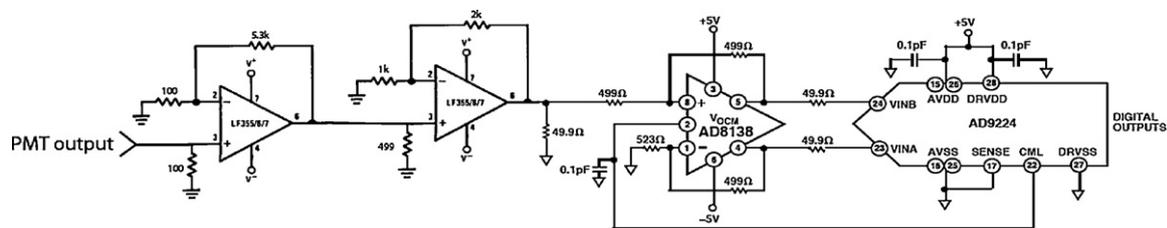


Fig. 6. The analog pulse amplification circuit.

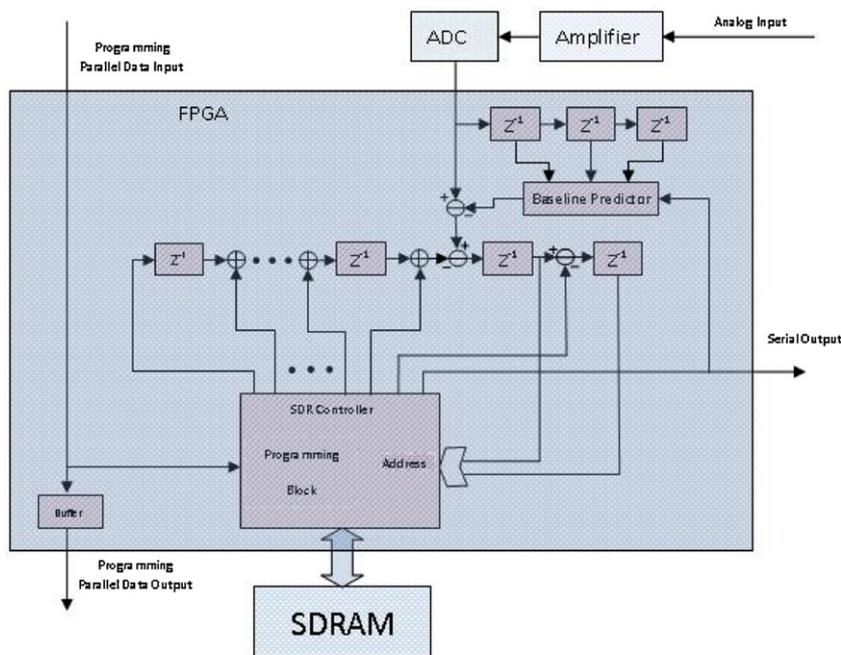


Fig. 7. Implemented hardware diagram for one data acquisition card.

sampling frequency, the data bus can be improved by paralleling the memory chips.

All  $Z^{-1}$  units represent one cycle delay which are implemented by shift registers. The baseline predictor calculates the signal mean value by 3 successive samples in the absence of input pulse. As shown in Fig. 7, this value is subtracted from the input signal in the first stage.

SDRAM is a non-volatile memory which requires start-up memory filling. These data are the same for all acquisition cards. Since this procedure is performed once in the start-up time, the required time and particular protocol used for this procedure is not critical. The memory control unit in the FPGA multiplexes the SDRAM bus according to input data from the programming bus and releases it following programming for the normal acquisition algorithm. Table 1 summarizes parts which are used in the acquisition card.

Fig. 8 shows a data acquisition card board and its components. In spectroscopy applications, only one acquisition card is required. In imaging applications, number of required acquisition cards depends on the detector configuration and the number of

PMTs. In some applications, this card can be placed after the analog positioning circuit. The main board which contains the required number of slots for the acquisition cards should be designed for the specific application. In this work, the implemented main board

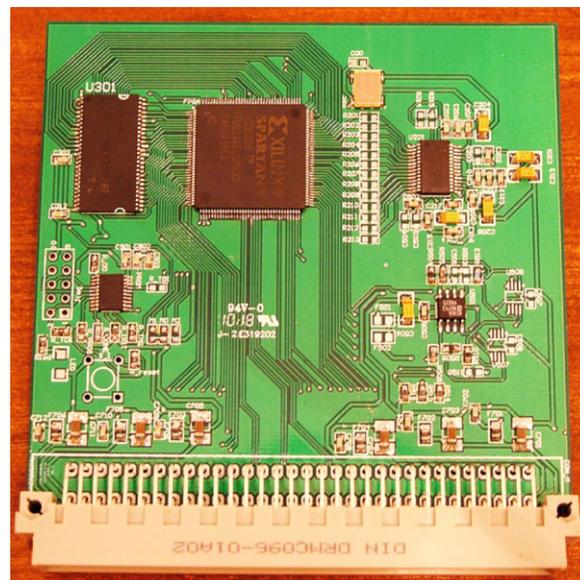


Fig. 8. Board view of the data acquisition card.

Table 1  
 Parts used in acquisition card.

Description	Part number	Specification
FPGA	XCS3200-144	Spartan3 – 144 pin – 200,000 gates
ADC	AD9224	1 channel – 12 bit – 40 Msps
Amplifier	AD8138	0.1 dB gain flatness to 40 MHz
Memory	MT48LC16M16A2TG	SDRAM – 512 Mbit – 16 bit data-bus

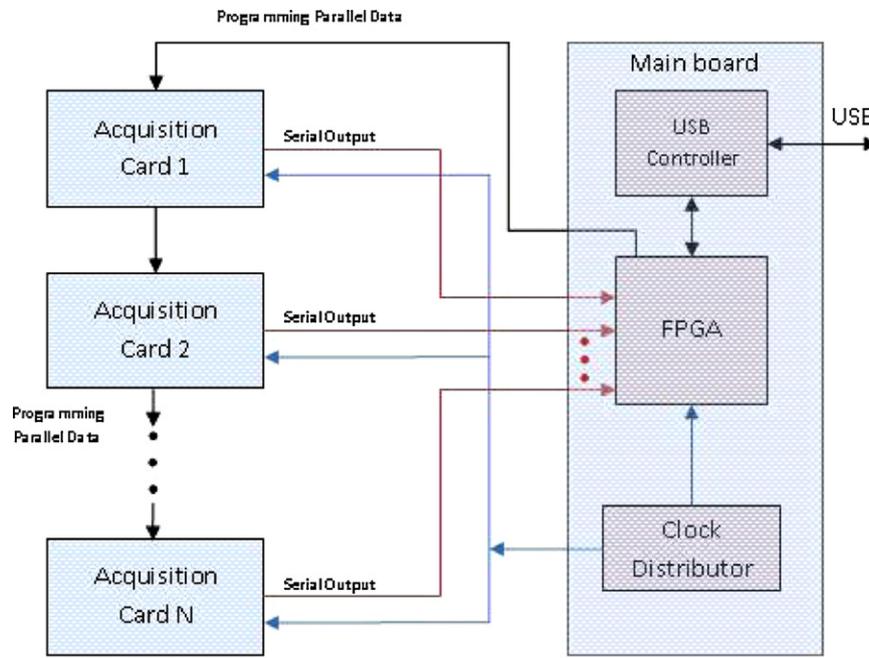


Fig. 9. Structure of the implemented main board.

contains 5 slots for imaging which also is applicable in spectroscopy application. Fig. 9 shows the design diagram of the main board.

The main board collects pulse data from each card and calculates the energy signal and position of the incident photon. These data have relatively lower rates than the raw data from detectors, and may be feasibly transferred to a PC for online processing. The suitable interface for this throughput of data is USB2.0 or high speed LAN. As shown in Fig. 9, the USB2.0 protocol is used in the implemented main board.

The clock distributor is required for synchronization of all cards and the main board. It is essential for the positioning algorithm that all samples are taken simultaneously. Preparing programming data for filling the SDRAMs is also the function of the main board. Since all data to be written are identical in all cards, the best structure is to cascade the cards. Each card will buffer the data for the next card. Consequently, by increasing the acquisition cards (i.e. up to 64) for a detector head, the structure design remains intact.

## 4. Results

### 4.1. Simulation data

The simulation procedure was performed by generating pulses with deterministic shapes. Nuclear pulses were simulated by a stream of random pulses with Poisson distribution of occurrence time and with random pulse heights. The generated stream went through the NLR filter and the obtained results were assessed.

Curve fitting and map construction, which are required before applying the filter, were performed by means of Matlab (The MathWorks Inc., Natick, MA, USA). Subsequently, C++ code, because of higher speed in execution, was used for generating detector pulses and the filter algorithm. Stream of input data were created via the Monte Carlo method. The occurrence probability of each pulse followed the Poisson distribution whose  $\lambda$  parameter was related to the count rate. Maximum amplitude of the pulse had a probability density function corresponding to the source spectrum. The source spectrum was modeled in reference to a real spectrum belonging to  $Tc^{99}$  detected by a NaI(Tl) detector.

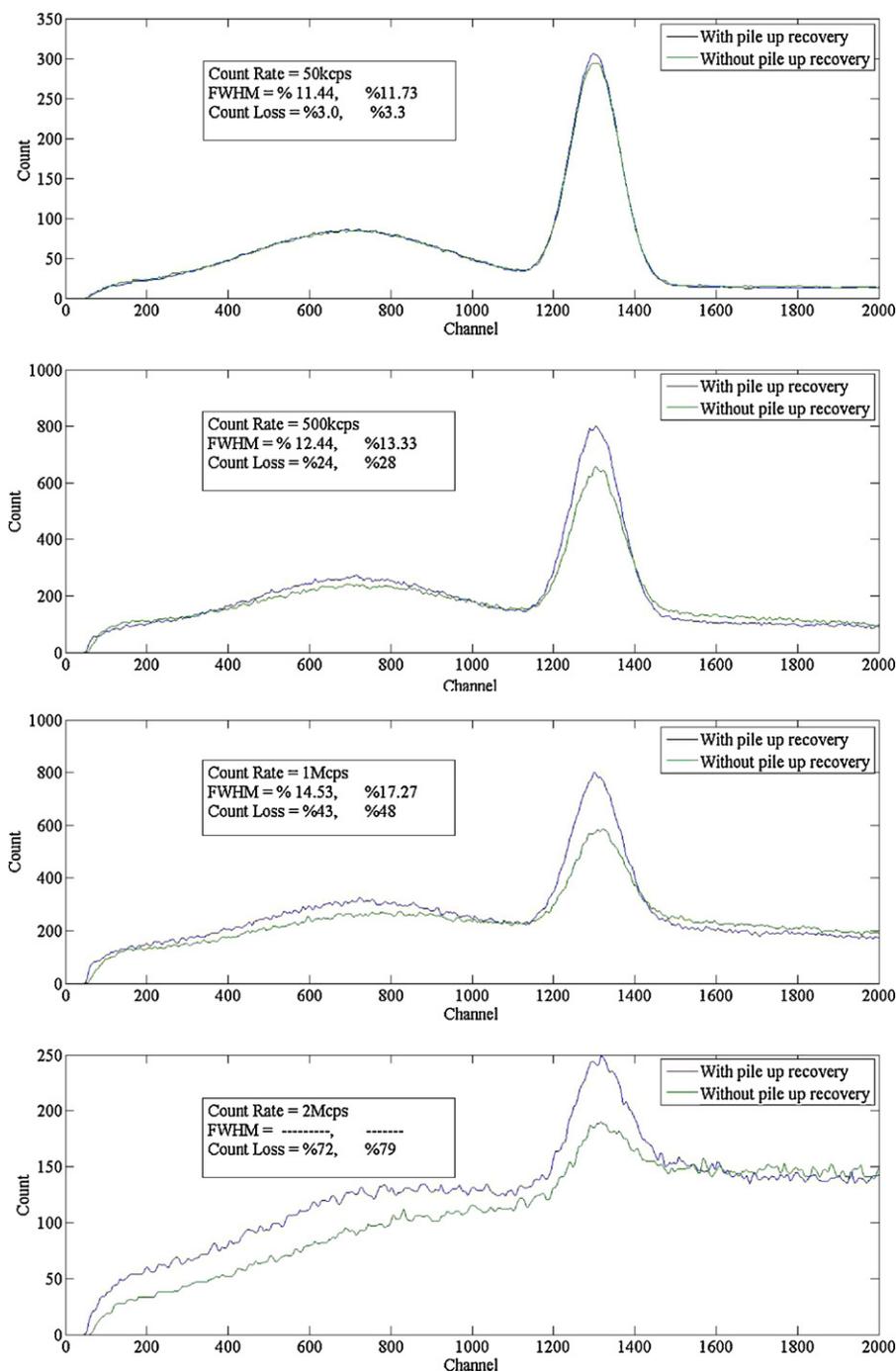
The pulse shape model was fitted by least squares in reference to the actual detector pulse shape for the NaI(Tl) crystal. This pulse was sampled from the PMT output passing the preamplifier. The sampling rate in the acquisition was  $\Delta t_s = 300$  ns and the resulting regressed pulse shape model was:

$$f(a, t) = at^{4.3}e^{t/160 \text{ ns}} \quad (14)$$

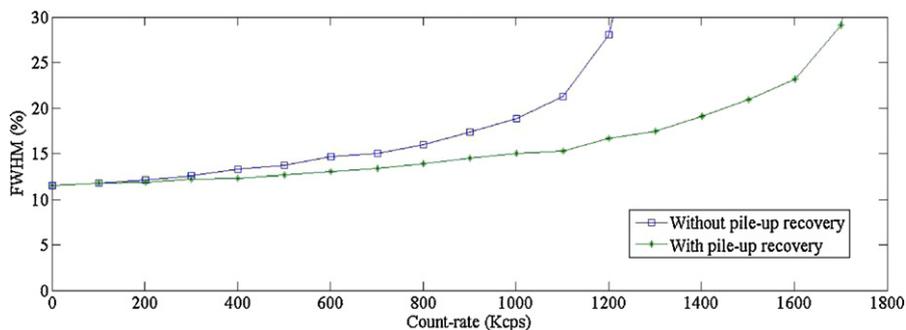
which has a peak at  $t_{\text{peak}} = 688$  ns. The energy spectra of the acquired data for different count-rates are depicted in Fig. 10. In low count rate the pileup effect is not dominant and pileup recovery does not highly affect spectrum quality whereas in high count rate the improvement in signal quality is clear. Without pileup recovery FWHM of photo-peak grows 47% from low count rates up to 1 Mcps. This growth with pileup recovery reaches 27% which shows recovery effect on high count rates. This fact should be noticed that without pileup recovery the filter still counts pileup effects in some extend (with possibly wrong values), therefore, number of registered counts in low count rates and even high count rates does not change as it would be in pileup rejection algorithms.

Fig. 11 illustrates the resulting energy FWHM values for the photo peaks as obtained by varying the count rates, without and with inclusion of pile-up recovery and tail storing.

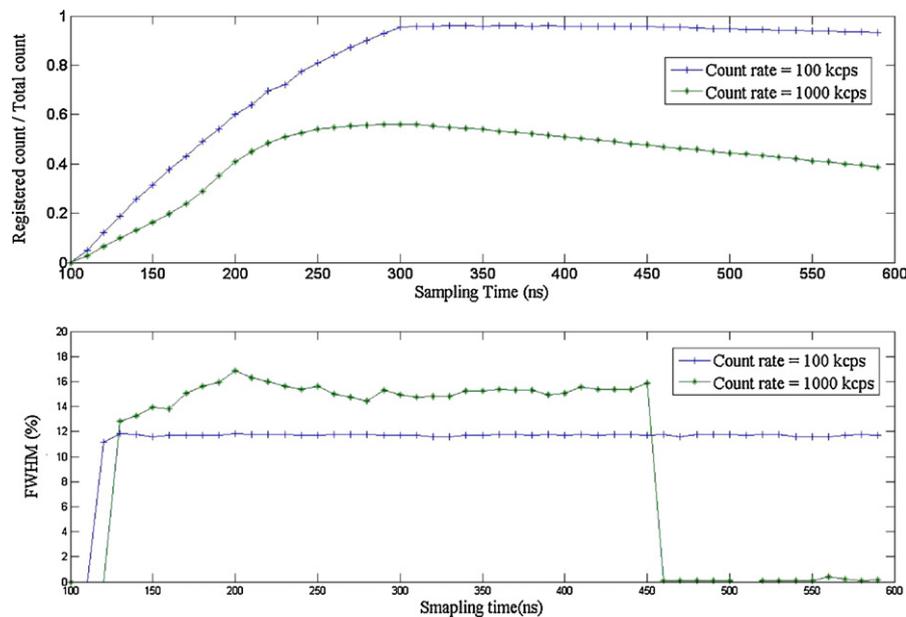
The sampling time, used by the NLR filter, plays an important role in the results. Different sampling times were set for the filter to evaluate the resulting count loss and photo peak FWHM values. As seen in Fig. 12, the count loss ratio has an optimum value of  $\Delta t_s = 300$  ns for the particular model studied; at the same time, it is observed that the FWHM of photo peak only changed minimally with small variation in the sampling times. Because in map construction, we zeroed map values when first samples are lower than a noise threshold (<50 in ADC channel), in low sampling times (<100 ns) two first samples are both under threshold. Subsequently, filter results no registered counts. When sampling period is less than 300 ns the noise threshold still rejects some pulses and lowers number of registered counts.



**Fig. 10.** Simulated spectra of source for different count rates with and without pile up recovery. The left values for FWHM and count loss correspond to filter not using pileup recovery and right values correspond to filter with pileup recovery. Count loss is determined by all registered counts (without using a window) per all generated counts.



**Fig. 11.** FWHM of the photo peak as a function of count-rate with and without pile-up recovery.



**Fig. 12.** (Down) Photo peak FWHM (top) registered count to total count ratio, versus different sampling times. The evaluation performed for 2 count rates representing high and low count rates.

#### 4.2. Experimental data

The proposed technique was additionally tested by means of prototype hardware implementation. The energy signals from NaI(Tl) detectors belonging to a gamma camera (large detector) were captured in order to obtain the TC<sup>99</sup> spectrum.

The test condition was set up for different count rates from low (150 kcps) up to high count rates (2 Mcps). To achieve high count rates we used collimator with large hole so that obtained FWHM, even in low count rates was about 14%. The detection configuration held intact for different count rates in order to evaluate the effect of count rate changes only.

The proposed algorithm applying pile-up recovery and pile-up rejection (by interleaved sampling) was compared with a digital signal processing algorithm presented in [15–17], namely pulse clipping (PC) followed by integration. For this algorithm, the signal is delayed and attenuated. The resulting signal is then subtracted from the original signal. This technique thus clips the pulse tail, and by shortening the pulse length, it degrades the pile-up defect. Delay time for the input pulse had an optimal value at 750 ns and an attenuation factor of 0.30 (in [15]), it was mentioned that the delay should be more than the pulse rise time and the attenuation factor should return the tail to baseline after delay time). After clipping, the energy of the incident photon is obtained by integrating the resulting pulse. Integration starts when pulse height reaches a threshold (channel 80 of ADC) and terminates when pulse drops to that threshold again. This integration time interval is used to check for the presence of pile-up, and for intervals greater than one pulse length (~2.5 μs), the signal is rejected as having two or more pile-up signals. The sampling frequency for applying this filter is 40 Msps (ADC maximum frequency) whereas for the proposed NLRF it is only 3.3 Msps and for the proposed interleaved NLRF it is 6.6 Msps. Furthermore, we note that the proposed NLRF included no pile-up rejection (possibly leading to mischaracterization of pileup) while interleaved NLRF included pile-up rejection for pulses arriving within 450 ns of one another (as discussed previously).

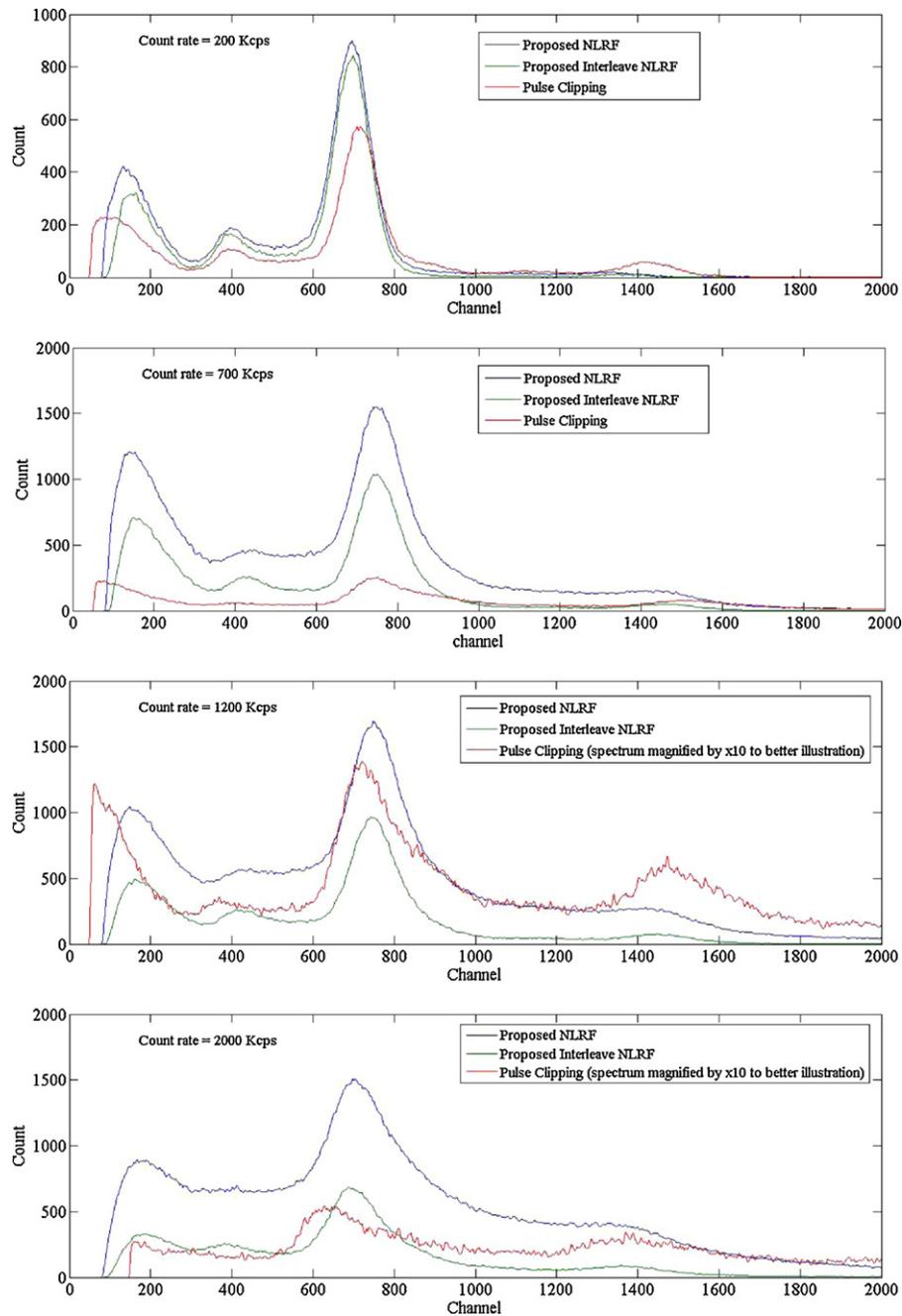
At low count rates (<50 kcps), the NLRF energy spectrum results are identical to those obtained without the use of NLRF, and the photo peak has FWHM energy resolution of 14.5%. Fig. 13 depicts the energy spectra obtained at higher count rates of 200 Kcps,

700 Kcps, 1200 Kcps and 2000 Kcps, as obtained using the proposed NLRF, its interleaved variant and the reference pulse clipping method. At rates beyond 1 Mcps, pulse clipping was seen not to generate acceptable spectral shapes from the input signal while the NLRF algorithm retained the spectral shape at higher count rates. Table 2 shows FWHM and integrated count comparisons between the 3 methods. Overall, both NLR filter methods clearly outperformed the reference PC method. Furthermore, it was seen that the proposed NLR filter depicted poorer energy FWHM compared to its interleaved variant, attributed to the latter's ability to reject more invalid (piled up) counts from the spectrum. However, this resulted in the integrated counts to drop more than half, at higher count rates, and as such, depending on task (whether sensitivity or energy resolution are more critical at a particular range of count-rates), the NLF filter vs. its interleaved variant could be more appropriately selected.

#### 5. Discussion

In our validation step, the feasibility and practicality of the proposed method for signal processing were studied. The proposed method was seen to enable signal recovery at relatively small sampling frequencies (by means of incorporating knowledge of the determinant pulse model). Furthermore, by proposing and implementing a pile up recovery scheme, enhanced pulse information was obtained. Proposed algorithm compared with a reference method to evaluate performance. Although there is no way to be sure it is best implementation of this algorithm but the results are consistent with original results presented in [15]. As such, the proposed approach has many significant potential applications to various imaging systems.

A considerable feature, which with current hardware platform, can only be simulated, is the time of flight measurement capability of the proposed algorithm, which we plan to study experimentally in future work. In PET, and some methods of spectroscopy, there is a need for exact determination of the interaction time for a photon or particle. The proposed NLR filter has the ability to output the relative occurrence time of pulse to the sampling time in system ( $t$ ) which improves the stamped time resolution to values much higher than the sampling frequency. A TOF PET device has two



**Fig. 13.** Spectra of  $Tc^{99m}$  as obtained from the NaI(Tl) detector for different count rates. In all figures, counts were recorded within constant 800 ms time period.

features: first, it needs the accurate measurement of occurrence time of interaction, which poses a considerable challenge; and second, because of the absence of a collimator, one encounters relatively high count rates and enhanced pile-up effects in imaging [28]. TOF PET scanners use fast detectors and apparatus to enable the

TOF modality, which result in improved image qualities given the additional information [27]. The most popular analog techniques to improve time resolution include the leading edge (LE), constant fraction discriminator (CFD) and amplitude and rise time compensated (ARC) methods [29]. Digital version of these methods has also

**Table 2**  
 Resolution and number of registered counts in different count rates.

Count rate(kcps)	Pulse clipping FWHM (%) counts <sup>a</sup>		Standard NLRF FWHM (%) counts		Interleaved NLRF FWHM (%) counts	
200	15.1	50,000	15.7	79,000	15.4	71,000
700	24.7	37,000	21.5	203,000	18.0	111,000
1200	29.9	22,000	25.4	253,000	19.7	109,000
2000	–	20,000	–	335,000	27.6	96,000

<sup>a</sup> Number of registered counts is equal to area under the spectrum curve within a fixed energy window (covering FWHM of photo-peak).

been implemented [30–33]. Significant degradations in detected time resolution stem from dead time in crystals and propagation delays in PMT. The exact improvement in the time resolution by the proposed NLR filter should be evaluated by further analysis involving fast detectors (such as LSO) and PMTs within the coincidence detection system. In order to work with pulses with higher frequencies (shorter pulse lengths) in fast detectors, the sampling rate and filter speed should be increased. From an implementation point of view, the bottleneck for increasing speed is the SDRAM's access time, which should be assessed by means of DDR, DDR2 and DDR3 architectures.

## 6. Conclusion

A new filter with recursive algorithm for extracting nuclear pulse properties has been introduced by using LUTs for implementing nonlinear functions. This filter developed pile-up and noise features simultaneously in high count rate and almost had no effect in low count rate. The piled-up pulses have been recovered after applying filter instead of rejecting them. The simulation results showed desirable output for this filter at a very low sampling rate of 3.3 MHz.

For hardware implementation we used FPGA as the central signal processor, SDRAM for implementing LUTs, A/D for digitization and USB2.0 for PC interface. The filter was tested by gathering data and drawing spectrum of  $Tc^{99}$ , demonstrating applicability of the filter and its performance at high count rates.

The NLR filter was tested for a spectroscopic application evaluated on NaI(Tl) detectors belonging to a gamma camera. Relative to conventional pile-up rejection systems, the proposed technique showed significant enhancements in terms of recovered count-rates and energy resolution at increasing high count rates.

## Competing interests

None declared.

## Funding

None.

## Ethical approval

Not required.

## References

- [1] Bushberg JT, Seibert JA, Leidholdt EM, Boone JM. The essential physics of medical imaging. 2nd ed. Lippincott Williams & Wilkins; 2002.
- [2] Knoll G. Radiation detection and measurement. 3rd ed. John Wiley & Sons. Inc; 1999.
- [3] Dunn DR, Kaifer RC, McQuaid JH. A high-rate spectroscopy system for use with Ge(Li) coaxial detectors. IEEE Trans Nucl Sci 1972;19:461–8.
- [4] Goulding FS, Landis DA, Madden NW. Design philosophy for high-resolution rate and throughput spectroscopy systems. IEEE Trans Nucl Sci 1983;30(1):301–10.
- [5] Georgiev A, Gast W. Digital pulse processing in high resolution, high throughput GAMMA-RAY spectroscopy. IEEE Trans Nucl Sci 1993;40:770–9.
- [6] Reguigui N, Morel J, Kraiem HB, Mahjoub A. Characterization of three digital signal processor systems used in gamma ray spectrometry. Appl Radiat Isot 2002;56:93–7.
- [7] Germano G, Hoffman EJ. An investigation of methods of pileup rejection for 2-D array detectors employed in high resolution PET. IEEE Trans Med Imaging 1991;10:223–7.
- [8] Pullia A, Gritti G, Ripamonti G. Stochastic digital filtering of nuclear radiation signals: a new approach. In: IEEE. 1997.
- [9] Yang H, Wehe DK, Bartels DM. Spectroscopy of high rate events during active interrogation. Nucl Instrum Methods Phys Res Sect A 2009;598:779–87.
- [10] Spanoudaki VC, McElroy DP, Ziegler SI. An analog signal processing ASIC for a small animal LSO-APD PET tomograph. Nucl Instrum Methods Phys Res Sect A 2006;564:451–62.
- [11] Kafae M, Saramad S. Pile-up correction by genetic algorithm and artificial neural network. Nucl Instrum Methods Phys Res Sect A 2009;607:652–8.
- [12] Odell DMC, Bushart BS, Harpring LJ, Moore FS, Riley TN. Zero dead time spectroscopy without full charge collection. Nucl Instrum Methods Phys Res Sect A 1999;422:363–7.
- [13] Jordanov V, Knoll GF. Digital pulse processor using a moving average technique. IEEE Trans Nucl Sci 1993;40:764–9.
- [14] Wong W-H, Li H. A scintillation detector signal processing technique with active pileup prevention for extending scintillation count rates. IEEE Trans Nucl Sci 1998;45:838–42.
- [15] Bolic M, Drndarevic V. Processing architecture for high count rate spectroscopy with NaI(Tl) detector. In: IEEE international instrumentation and measurement technology conference. 2008.
- [16] Bolic M, Drndarevic V, Gueaieb W. Pileup correction algorithm for very-high-count-rate gamma-ray spectrometry with NaI(Tl) detectors. IEEE Trans Instrum Meas 2010;59:122–30.
- [17] Bolic M, Drndarevic V. Digital gamma-ray spectroscopy based on FPGA technology. Nucl Instrum Methods Phys Res Sect A 2002;482:761–6.
- [18] Raad MW, Deriche M, Noras J, Shafiq M. A novel approach for pileup detection in gamma-ray spectroscopy using deconvolution. IOP Meas Sci Technol 2008;19(6), article: 065601.
- [19] Riboldi S, Geraci A, Gatti E, Ripamonti G. A new digital auto-tracking pole-zero compensation technique for high-resolution spectroscopy. Nucl Instrum Methods Phys Res Sect A 2002;482:475–90.
- [20] Cova S, Colombo S, Freitas I. Automated regulation of critical parameters and related design aspects of spectroscopy amplifiers with time invariant filters. IEEE Trans Nucl Sci 1982;29:609–13.
- [21] Morgado AMLS, Simoes JB, Correia CM. Baseline restoration using current conveyors. IEEE Trans Nucl Sci 1995;1:326–9.
- [22] Pullia A. Spectroscopic technique for optimal P-Z setting in gamma-ray detection. IEEE Trans Nucl Sci 2001;48(4):1234–8.
- [23] Xiangyang W, Yixiang W. Optimum digital filters for baseline restoration. IEEE Trans Nucl Sci 2006;53(6):3865–9.
- [24] Imperiale C, Imperiale A. On nuclear spectroscopy pulses digital shaping and processing. Measurement 2001;30:49–73.
- [25] Pullia A, Geraci G, Ripamonti G. Quasi-optimum gamma and X spectroscopy based on real-time digital techniques. Nucl Instrum Methods Phys Res Sect A 2000;439:378–84.
- [26] Yuan Z, Jin Y. High performance data acquisition methods based on waveform digitization. In: IEEE nuclear science symposium. 2004.
- [27] Rahmim A, Zaidi H. PET versus SPECT: strengths, limitations and challenges. Nucl Med Commun 2008;29:193–207.
- [28] Jian Y, Mulnix T, Carson RE. Count-rate dependent resolution degradation from pulse pile-up on the HRRT. In: IEEE nuclear science symposium. 2010. p. 2765–8.
- [29] Monzo JM, Esteve R, Lerche CW, Ferrando N, Toledo J, Aliaga RJ, et al. Digital signal processing techniques to improve time resolution in positron emission tomography. In: IEEE-NPSS. 2010. p. 1–6.
- [30] Fallu-Labruyere A, Tan H, Hennig W, Warburton WK. Time resolution studies using digital constant fraction discrimination. Nucl Instrum Methods Phys Res Sect A 2007;579:247–51.
- [31] Visser G, Cherry S, Clajus M, Yiping S, Tumer TO. Development of low power high speed readout electronics for high resolution PET with LSO and APD arrays. vol. 4. In: IEEE nuclear science symposium, medical imaging conference. 2001. p. 1988–92.
- [32] Nakhostin M, Kikuchi Y, Ishii K, Matsuyama S, Yamazaki H. Development of a digital front-end electronics for the CdTe PET systems. Nucl Instrum Methods Phys Res Sect A 2010;614:308–12.
- [33] Joly B, Montarou G, Lecoq J, Bohner G, Crouau M, Brossard M, et al. An optimal filter based algorithm for PET detectors with digital sampling front-end. IEEE Trans Nucl Sci 2010;57:63–70.
- [34] Tanaka E, Ohmura T, Yamashita T. A new method for preventing pulse pileup in scintillation detectors. Phys Med Biol 2002;47:327–39.