Real-Time Digital Timing in Positron Emission Tomography

P. Guerra, *Member, IEEE*, J. E. Ortuño, G. Kontaxakis, *Senior Member, IEEE*,
M. J. Ledesma-Carbayo, *Member, IEEE*, J. J. Vaquero, *Senior Member, IEEE*, M. Desco, and
A. Santos, *Senior Member, IEEE*

Abstract—Positron emission tomography (PET) requires accurate timing of scintillation events to properly discriminate between coincident and noncoincident pairs. The traditional solution to timing is based on custom application specific integrated circuits (ASIC) designs, whose cost may not be justified in the design of experimental small animal PET scanners. The new generation of PET scanners introduces the idea of continuous sampling of the detected scintillation pulse, replacing event-triggered acquisition front-ends. This approach enables new options to the timing procedure based on digital processing of the sampled pulse signal. This work proposes a time stamping algorithm based on the optically matched filter and compares the potential performance benefits of this approach versus other FIR-based timing algorithms, some of which have been already implemented by different authors. Results show that the coincidence timing resolution may be as low as 1.5 ns without the need of expensive high-speed converters when the proper signal processing is applied.

Index Terms—Data acquisition, positron emission tomography, signal processing.

I. INTRODUCTION

THE PRECISE measurement of the elapsed time between two events is the underlying foundation for electronic collimation in positron emission tomography (PET). PET is based on the physical properties of certain radioactive isotopes known as positron emitters, which emit positrons when they undergo radioactive decay. The ejected positron rapidly annihilates on contact with electrons, producing, in the most common case, two photons with energy equal to the rest energy of the electron or positron (511 keV). Due to the conservation of the linear momentum, the gamma rays are emitted in opposite directions. These photons impinge almost simultaneously on a set of gamma detectors placed around the field of view (FOV)

and the technique relies on the detection of simultaneous events on physically opposed detectors [1].

Time-coincident photons are usually resolved by tagging every detected single event with a time stamp and searching for those events whose time difference is below a certain threshold. However, assigning a time stamp to the scintillation pulse is not a trivial task, because the pulse start point is ill-defined and also due to practical problems derived from the time scale being handled. Moreover, the system may require periodic calibrations in order to provide optimal performance [2].

The typical implementation of the PET timing module relies on a mixed-signal application-specific integrated circuit (ASIC) device designed to yield nanosecond timing accuracy [3]–[5]. However, solutions based on digital processing of the sampled signal may represent a reasonable trade-off between cost and performance in preclinical PET systems targeted to small animals [6], [7], where time-of-flight PET is not a design constraint.

This work presents and compares different algorithmic alternatives to the design of a hardware-timing module that provides an accurate estimation of the time stamp in real time, based on pulse energy samples. This paper is structured as follows. The first section summarizes the different timing methods that will be considered and the second section describes the material required for the acquisition and processing of the experimental data. Next, the result of applying the different filters to the experimental data is presented, showing that with the proposed filtering approach a coincidence timing resolution as low as 1.5 ns FWHM may be achieved for LYSO pulses at a sampling frequency of 65 MHz. The document includes an Appendix to show the feasibility of implementing the presented methods in hardware.

II. TIMING FILTERS

A. Implementation Constraints

The ultimate goal of the algorithms described in this work is to implement a hardware-timing module that computes the time stamp for every detected scintillation pulse in real time. In order to support high counting rates, a direct algorithm is preferred against model-fitting iterative approaches, such as those described in [8], which may potentially yield higher accuracy at a higher computational cost.

As part of such an implementation decision, all considered algorithms admit similar hardware implementation blocks consisting of a finite impulse response (FIR) filtering of the energy signal plus a zero-crossing detection and some sort of signal interpolation. The motivation for such a solution is twofold: on

the one hand the resulting datapath is very predictive in terms of latency and the number of clock cycles required to extract the time stamp, facts that lead to easy data pipelining, and on the other hand, FIR filters are very efficiently implemented in a field-programmable gate array (FPGA).

The discussion contained in this work focuses on the existing alternatives to the computation of the coefficients $h[n]$ of the integrated hardware FIR filter and the impact that the selected method has on the timing resolution of the timing module.

The following equations assume an energy scintillation pulse $p(t)$, taken, for example, from the last dynode of a photomultiplier, sampled at a fixed frequency f_s. The sample pulse $p[n]$ is filtered $p_{\text{filter}}[n]$, and is afterward interpolated back into the continuous as $\hat{p}_{\text{filter}}(t)$. The scintillation pulse $p(t)$ is modeled by a biexponential function $\lambda(t)$ characterized by its rise (τ_R) and fall (τ_F) times.

B. Constant Fraction Discriminator

The simplest solution to digital timing consists of implementing the digital counterpart of the constant fraction discriminator (CFD), which is by far the most popular approach to timing in nuclear applications and whose main characteristic is maintaining a constant timing edge, since it is not affected by amplitude of the input pulse $p(t)$. Its principle of operation, as shown in (1), is as follows. The analogue signal is split into two components; one of them is time delayed D and the other is attenuated by the factor CF; afterwards, both branches are remixed and a comparator detects the zero-cross point η, which determines the signal timing.

$$h_{\text{CFD}}(t) = \delta(t - D) - CF \cdot \delta(t)$$
$$\eta = \arg_t\{p(t) * h_{\text{CFD}}(t) = 0\} \qquad (1)$$

When it comes to implementing a digital version of equation (1), there is a need to design a noninteger digital delay $\delta(\text{t} - D/T_s)$, which admits several solutions depending on the chosen minimization criteria in order to approximate its ideal frequency, as thoroughly described in the literature [9]–[11].

$$h_{\text{CFD}}[n] = \delta\left[n - \frac{D}{T_s}\right] - CF \cdot \delta[n]$$
$$H_{\text{CFD}}(\omega) = \exp\left(j\omega\frac{D}{T_s}\right) - CF = \exp\left(e^{j\omega D'}\right) - CF \qquad (2)$$

The most simple and direct approximation of $H(\omega)$ comes from computing and truncating the inverse Fourier transform of the ideal frequency response $H(\omega)$ (2), which is a shifted and sampled version of the *sinc* function, as shown in (3).

$$\Im^{-1}\left[e^{j\omega D'} - CF\right] = \frac{\sin(\pi(n - D'))}{\pi(n - D')} - CF \cdot \partial[n] \qquad (3)$$

In order to produce realizable fractional delay filters, some windowed $w[n]$ finite length approximation has to be used. Under

these considerations, the digital CFD is expressed in terms of a digital FIR filter $h_{\text{CFD}}[n]$ as follows.

$$h_{\text{CFD}}[n] = \frac{\sin(\pi(n - D'))}{\pi(n - D')} \cdot w[n] - CF \cdot \partial[n] \qquad (4)$$

As with the standard CFD, the time stamp of the pulse $p[n]$ is taken as the time point where the filtered signal crosses the zero level, though in the digital case, where only samples at fixed time points are available, it implies interpolating the filtered pulse $p_{\text{CFD}}[n]$ in order to estimate the point η where the estimated continuous signal $\hat{p}_{\text{CFD}}(t)$ would cross the zero level.

$$p_{\text{CFD}}[n] = p[n] * h_{\text{CFD}}[n]$$
$$\hat{p}_{\text{CFD}}(t) = \text{interpolation}_t\{p_{\text{CFD}}[n]\}$$
$$\eta = \arg_t\{\hat{p}_{\text{CFD}}(t) = 0\} \qquad (5)$$

C. Linear Interpolation

Linear interpolation may also be used for the direct computation of the pulse start point, based on the assumption that the rising edge may be linearly approximated. Despite its simplicity, this approach has proven to be accurate enough for the computation of the time stamp with LSO pulses shaped for 75 ns rise time and sampled at 40 MHz [12] and for LSO/LGSO integrated signals sampled at 45 MHz [13].

The linear interpolator, in the simplest form, takes into account two samples adjacent to the desired point. In this case, the sample m with the steepest slope $p'[m]$ in the vicinity of the rise edge is considered for the computation of the line that passes through $p[m]$ and $p[m+1]$ and the zero crossing η of the interpolated line $p_m[n]$ with the base level computed as shown in (6).

$$m = \arg\max_n\{p'[n]\}$$
$$\hat{p}_m(t) = \text{interpolation}\{p_m[n]\}$$
$$\eta = \arg_t\{\hat{p}_m(t) = 0\} \qquad (6)$$

The previous expression can also be reformulated as a pulse filter plus interpolation to compute the crossing point with the base line, as shown in (7), taking into account that $h_{\text{DIF}}[n]$ is a discrete differentiator [14].

$$m = \arg\max_n\{p[n] * h_{\text{DIF}}[n]\}$$
$$\eta = m - \frac{p[m]}{p[n] * h_{\text{DIF}}[n]|_m} \qquad (7)$$

D. Classical Matched Filter

Another possible approach to timing is based on the correlation between the sampled pulse and a reference template of the scintillation pulse. In this situation, the criterion is that the
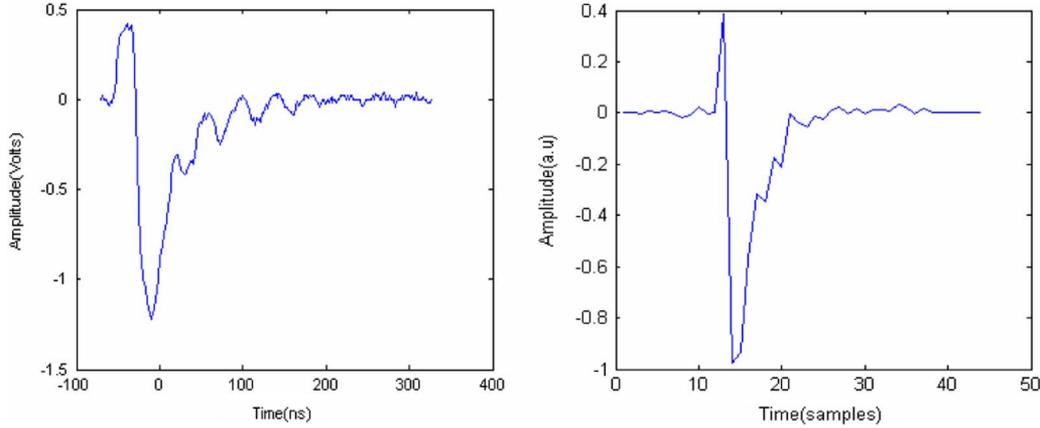
Fig. 1. Pulse output of the CFD with an analogue delay of 15 ns and a constant fraction of 0.3 and output of the digital CFD for the same pulse sampled at 60 MHz, with an equivalent delay of 15 ns and a constant fraction of 0.3.
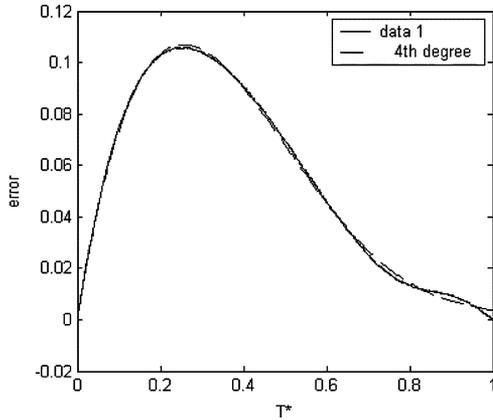

Fig. 2. Error distribution in [0,1] and the fitted 4th order polynomial.

timing of the pulse is the point where the correlation is at maximum. This can be formulated in terms of an FIR filter $h_{\mathrm{MF}}[n]$, which is derived from the reference pulse $\lambda[n]$, either based on experimental data or on nominal values. In the latter case

$$
\begin{aligned}
h_{\mathrm{MF}}[n] &= \lambda_s[n] \\
&= \frac{A}{\tau_F - \tau_R}\left(\exp\left(-\frac{n \cdot T_s}{\tau_F}\right)\right. \\
&\quad \left. - \exp\left(-\frac{n \cdot T_s}{\tau_R}\right)\right),
\end{aligned}
$$

$$
\begin{aligned}
h'_{\mathrm{MF}}[n] &= \frac{A \cdot T_s}{\tau_F - \tau_R}\left(-\frac{\exp\left(-\frac{n \cdot T_s}{\tau_F}\right)}{\tau_F}\right. \\
&\quad \left. + \frac{\exp\left(-\frac{n \cdot T_s}{\tau_R}\right)}{\tau_R}\right),
\end{aligned}
\tag{8}
$$

$$
\hat{p}_{\mathrm{MF}}(t) = \text{interpolation}\,\text{t}\{p[n] * h_{\mathrm{MF}}[n]\},
$$
$$
\eta = \arg\max_t\{\hat{p}_{\mathrm{MF}}(t)\},
$$
$$
\eta = \arg_t\{\text{interpolation}_t\{p[n] * h'_{\mathrm{MF}}[n]\} = 0\},
\tag{9}
$$

where $h'_{\mathrm{MF}}[n]$ is the first derivative of the impulse response and the time stamp η is computed as the point where the interpolated filtered pulse $\hat{p}_{\mathrm{MF}}(t)$ crosses the zero level.

E. Optical Matched Filter

The last considered solution is the use of an optical matched filter that, like the matched filter, originally comes from communication theory, in particular from the design of optical receivers that are robust against uncertainty in the statistics of the observation process in photodetection. This family of filters are described in full detail in [15] and was first proposed for timing in PET in [16] from a theoretical point of view, although no actual implementation is known to date.

The filter design starts with the definition of a filtered Poisson process $i_s(t)$ for the output of the photodetector, with the addition of an independent zero-mean thermal noise process $i_{\mathrm{th}}(t)$ that corrupts the detected current signal $i(t)$.

$$
i(t) = q\sum_{i=1}^{N_t} g_k \partial(t - \tau_k) + i_{\mathrm{th}}(\tau),
\tag{10}
$$

where N_t is an inhomogeneous Poisson counting process such that N_t is the number of photoelectrons generated during the period [0,t], τ_k is the emission time of the kth electron and g_k are independent and identically distributed random variables that model the number of secondary electrons generated for each primary photoelectron with mean $E\{g\}$. Here, q is the electronic charge and the detector impulse response is assumed to be $g_k\delta(t)$, where $\delta(\bullet)$ is the Dirac delta function.

The intensity $\lambda(t)$ of the inhomogeneous Poisson process N_t is related to the incident scintillation pulse $p(t)$ and is decomposed into signal $\lambda_s(t)$ and the dark current rate from the photodetector λ_d. (Background radiation can also be lumped into this term).

$$
\lambda(t) = \lambda_s(t) + \lambda_d
$$
$$
\lambda_s(t) = \frac{A}{\tau_F - \tau_R}\left(\exp\left(-\frac{t}{\tau_F}\right) - \exp\left(-\frac{t}{\tau_R}\right)\right)
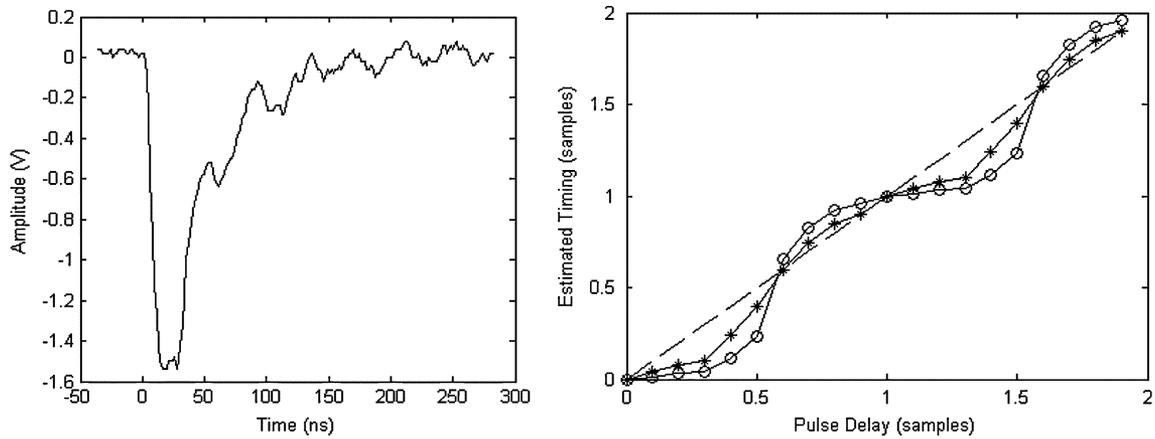\tag{11}
$$

Fig. 3. Sample of a LYSO scintillation pulse (left) and the corresponding time estimates vs. the introduced delay with the optical filter (*) and the classical matched filter (o) (right).
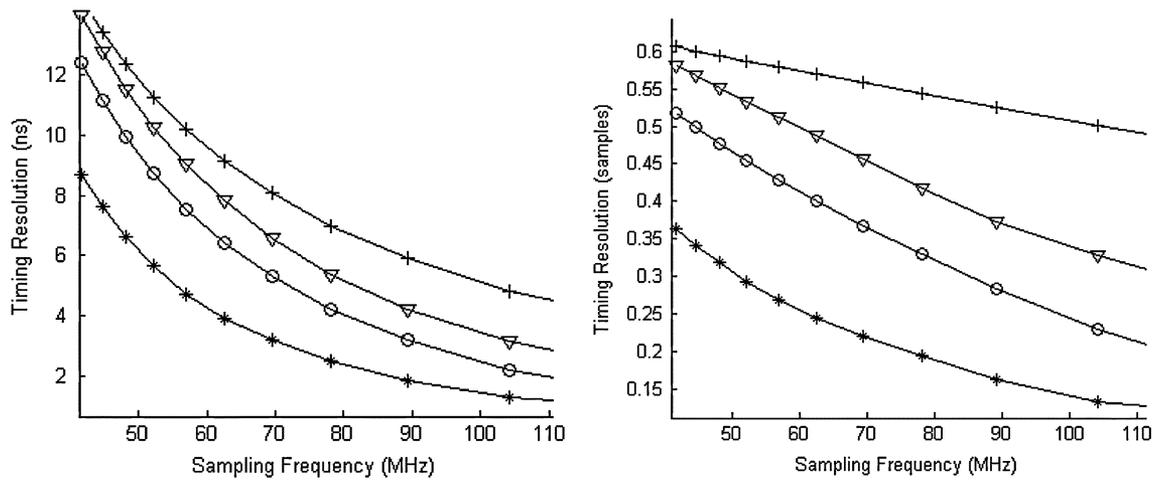

Fig. 4. Absolute (left) and normalized (right) singles timing resolution with different algorithms: linear interpolation ($+$), CFD ($\blacktriangledown$), matched filter (o) and optical filter (*) for sampling frequencies ranging from 40 to 110 MHz.
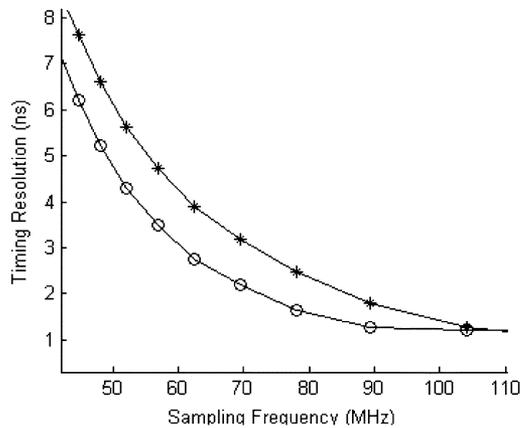

Fig. 5. Optical filter after self-calibration (o) vs. the noncorrected results (*).

The observed signal $y(t)$, which is a filtered version of the detected current signal $i(t)$, is corrupted by thermal noise that is assumed to be white and Gaussian with a spectral density of $N_o/2$.

$$y(t) = q \sum_{i=1}^{N_t} g_k h_{\text{OMF}}(t - \tau_k) + \int_o^t i_{\text{th}}(\tau) h_{\text{OMF}}(t - \tau) d\tau$$
$$= \lambda(t) + w(t) \tag{12}$$

The process $y(t)$ tends to a Gaussian process in $[0, T]$ with mean μ, and variance σ^2 equal to

$$\mu = \int_o^T \lambda_s(t) h_{\text{OMF}}(T - t) dt,$$

$$\sigma^2 = \int_o^T [\lambda_s(t) + R_n(t)] h_{\text{OMF}}^2(T - t) dt,$$

$$R_n(t) \triangleq \frac{N_0}{2} \frac{1}{q^2 E\{g^2\}} + \lambda_d, \tag{13}$$

where the term R_n incorporates the effects of dark current and thermal noise. Within this formulation, the optical matched filtering problem is obtained by the maximization of the observed signal $(y(t))$ to noise ratio over all possible filter impulse responses $(h_{\mathrm{OMF}}(t))$, which in the case of fixed $\lambda_s(t)$ and R_n is given by [15], as follows:

$$h_{\mathrm{OMF}}(t) = \frac{\lambda_s(-t)}{\lambda_s(-t) + \lambda_d + \frac{N_o}{2}} \tag{14}$$

Under these definitions, the linear estimator that provides the pulse starting point results from maximizing the interpolated filtered pulse, as shown in the following expression:

$$\begin{aligned} \hat{p}_{\mathrm{OMF}}(t) &= \operatorname*{interpolation}_{t}\{p[n] * h_{\mathrm{OMF}}[n]\}, \\ \eta &= \operatorname*{arg\,max}_{t}\{\hat{p}_{\mathrm{OMF}}(t)\}, \\ \eta &= \operatorname*{arg}_{t}\{\operatorname*{interpolation}_{t}\{p[n] * h'_{\mathrm{OMF}}[n]\} = 0\}, \end{aligned} \tag{15}$$

where $h_{\mathrm{OMF}}[n]$, and its first derivative $h'_{\mathrm{OMF}}[n]$, are functions of the underlying pulse parameters and the time stamp η is computed as the point where the interpolated filtered pulse $\hat{p}_{\mathrm{OMF}}(t)$ crosses the zero level.

F. Calibration

Time-stamp estimation algorithms show nonlinearities that are susceptible to being automatically corrected. A solution to this problem has been previously reported in [17]. This subsection details an alternative approach, which has been developed during the course of this work.

The error of the time stamp is a function of the sampling phase with periodicity equal to the sampling period T_s. Therefore, in order to develop a calibration method, it is enough to analyze the estimation error within $[0, T_s]$.

It is assumed that there is no correlation between the scintillation events and the sampling clock and therefore in the ideal case the timing error must be uniformly distributed in $[0, T_s]$. However, experimental distributions of the time stamp error show a nonuniform distribution of the error, which is mainly due to the method nonlinearities. The aim of the calibration method is to find the transformation that equalizes the histogram and to use this transformation to correct every time stamp. This is achieved based on an estimation of the error function that results from sorting N observed time stamps based on their fractional part $\mathrm{fract}(\tau)$ and using the position $k(i)$ of the time stamp i in the sorted list to eliminate the bias on the time stamp $\tau(i)$ to obtain the linearized value $\tau_k^*(i)$, as it is described in the following algorithm.

$$\begin{aligned} \breve{\tau}(i) &= \mathrm{fract}(\tau(i)) = \tau(i) - \lfloor\tau(i)\rfloor, \\ 0 &\le \breve{\tau}(i) < 1 \quad, \forall i \in [0, N-1] \\ k(i) &= \mathrm{sort}(\breve{\tau}(i)) \leftrightarrow \breve{\tau}(k(i)) \le \breve{\tau}(k(i+1)) \\ &\qquad\qquad\qquad\qquad \forall i \in [0, N-1] \end{aligned}$$

$$\begin{aligned} \breve{\tau}^*(i) &= \frac{k(i)}{N} \\ \tau^*(i) &= \lfloor\tau(i)\rfloor + \breve{\tau}^*(i) \end{aligned} \tag{16}$$

Under these definitions, the estimation error $\varepsilon(i)$ associated with each time stamp is as follows.

$$\varepsilon(i) = \frac{k(i)}{N} - \breve{\tau}(k(i)) \tag{18}$$

Afterwards, a polynomial function $f(\mathrm{x})$ of order m is computed to represent the nonlinear relationship between the fractional part $\breve{\tau}(i)$ of the time stamp provided by the algorithm and the related error $\varepsilon(i)$.

$$f(\breve{\tau}) = \sum_{j=0}^{m-1} a_j \breve{\tau}^j \tag{18}$$

Once the function is evaluated, the time stamp is updated as follows:

$$\hat{\tau}(i) = \tau(i) + f(\breve{\tau}(i)) \tag{18}$$

The polynomial coefficients are restricted to satisfy the condition that the error function is periodical, that is $f(0) = f(1)$ and therefore:

$$f(0) = f(1) \Rightarrow \sum_{j=1}^{m-1} a_j = 0. \tag{20}$$

The polynomial coefficients a_j are computed from the single list through linear least squares optimization with linear constraints, being Fig. 2 an example of such computation. This computation may be either offline, during the initial calibration steps, or online during acquisition. In the latter case, an intermediate memory buffer would be required to temporally store singles while computing polynomial coefficients in a first pass and correcting time stamps in a second pass. Nevertheless, the on-line approach is presently discarded until an adequate processing architecture that prevents an increase in the dead time of the scanner is defined.

III. MATERIALS AND METHODS

The experimental data have been obtained with a reengineered version [18] of the phoswich detector originally developed for the NIH ATLAS [19], whose energy dynode has been sampled at 625 MHz ($\mathrm{Ts} = 1.6$ ns) with a TDS5054B digital oscilloscope (Tektronix, Beaverton, OR, USA).

The detector consists of two 13×13 crystals layers, with the front crystal layer assembled of $1.45 \times 1.45 \times 7$ mm cerium-doped lutetium-yttrium orthosilicate (LYSO) and the back layer made of $1.45 \times 1.45 \times 8$ mm cerium-doped gadolinium orthosilicate (GSO) with a 1.45×1.45 pitch. The 20×20 mm GSO back layer is optically coupled to a Hamamatsu R8520-C12 photomultiplier tube.

TABLE I
SINGLE TIMING RESOLUTION (FWHM) IN THE FREQUENCY RANGE OF INTEREST

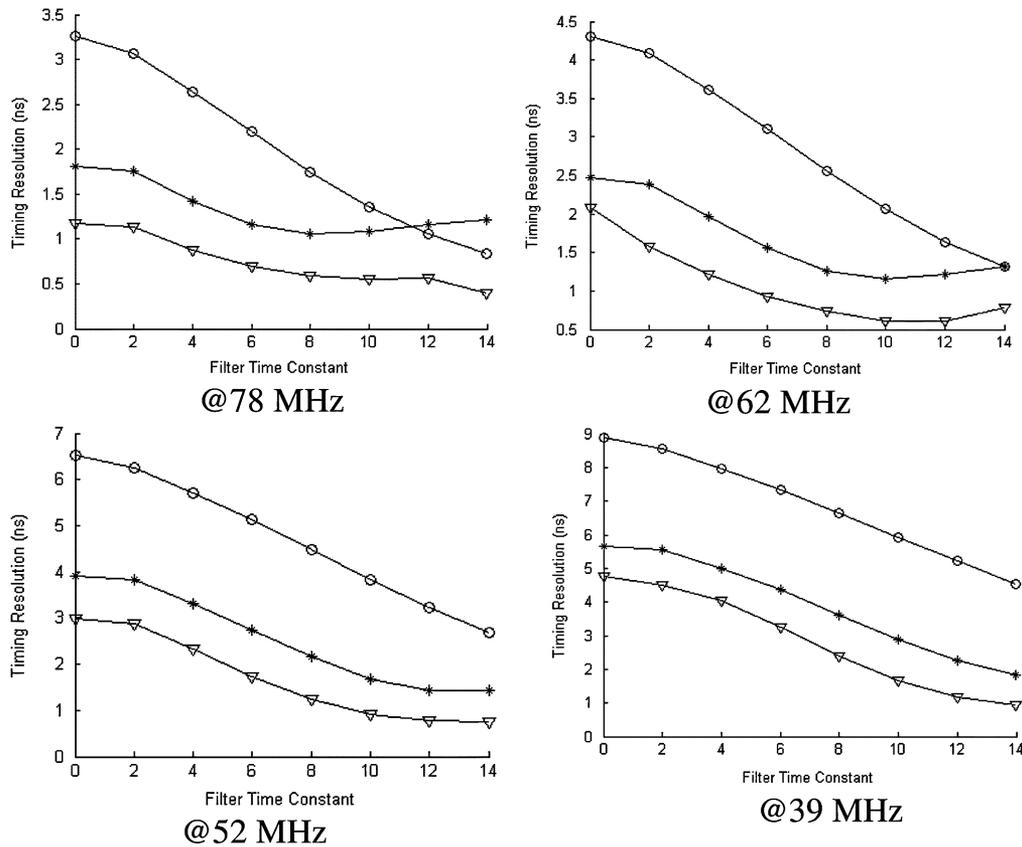| Filter Type	Resolution (ns) @ 50 MHz	Resolution (ns) @ 65 MHz	Resolution (ns) @ 80 MHz
Linear	11.80	8.71	6.77
CFD	10.82	7.32	5.11
Matched (MF)	9.33	5.97	4.02
Optical (OMF)	6.12	3.61	2.34
Optical (MF+calibr.)	4.75	2.52	1.55

Fig. 6. Self-calibrated optical filter (▼), noncalibrated optical filter (*) and matched filter (o) for sampling frequencies of 78 MHz, 62 MHz, 52 MHz and 39 MHz with shaping constants between 0 and 14 ns. (a) @78 MHz (b) @62 MHz (c) @52 MHz (d) @39 MHz.

The acquired data has been processed and analyzed with Matlab R13 (The Mathworks, Natick, MA, USA). The energy of 10 000 LYSO scintillation pulses was recorded with the TDS5054B oscilloscope and each of these pulses was subsampled by a factor between 5 and 15 with a variable offset, in order to generate a dataset of pulses at sampling frequencies ranging between 41 and 125 MHz with a known relative delay.

The digital CFD is configured for a 10 ns digital delay and a CF equal to 0.8, which is the combination of values that yielded the best results for this filter. The linear interpolation replicates the algorithm described in [12]. The correlation template was derived assuming a decay constant of 40 ns and a rise constant of 10 ns. The optical filter was computed based on the same template as before and noise estimation based on the actual data.

IV. RESULTS

The presented timing algorithms have been applied to each subsampled and delayed version of the original pulse and the error between the generated time stamp and the expected result was recorded. This procedure was performed because, although the actual start point of each pulse is unknown, the different samples of the original pulse are spaced by $Ts = 1.6$ ns, therefore the timing estimator is expected to grow linearly with the known delay. As an example, Fig. 3 (right) shows the estimated delay versus the actual delay of the matched filter and the optically matched filter for a fixed scintillation pulse shown in Fig. 3 (left) as the sampling phase between the pulse and the clock varies from 0 to 2 samples. As shown, the timing estimator fluctuates around the expected value.

Repeating this approach for the 10 000 pulses of the set, we have estimated the timing resolution at full-width-at-half maximum (FWHM) for sampling frequencies ranging between 40 and 125 MHz, as shown in Fig. 4. It seems obvious that the higher the sampling frequency, the finer the timing resolution. However, as the normalized resolution in Fig. 4 (right) shows,
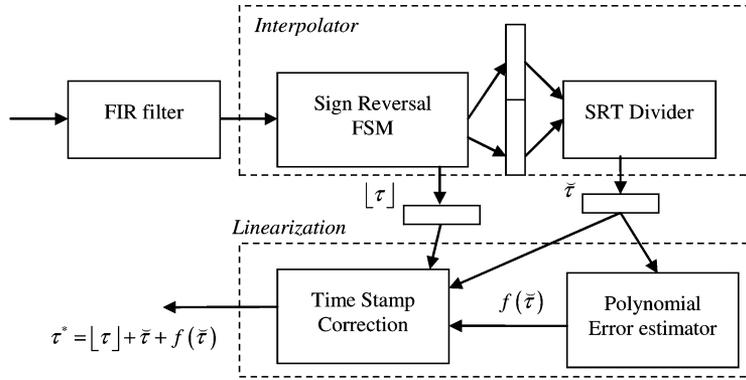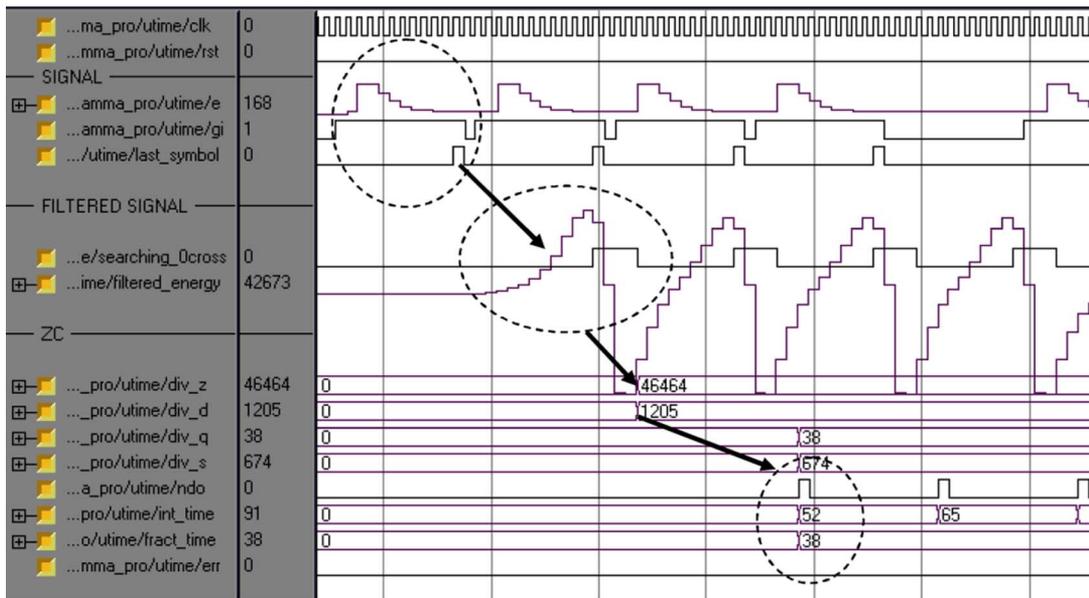
Fig. 7. Block diagram of the timing module.

Fig. 8. VHDL simulation of the timing module before polynomial correction. The input energy, the FIR filter output and the final divider output are highlighted with bubbles.
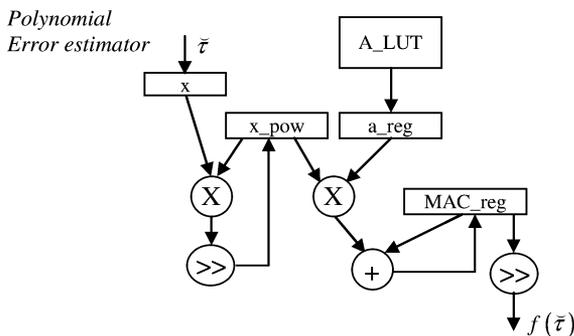

Fig. 9. Datapath for the polynomial error estimation block.

resolution improvements are exclusively due not only to the increasing sampling frequency, but also to a better performance of the estimation algorithm, which provides an increasing refinement of the sampling.

One interesting observation is that the single timing resolution improves as more knowledge about the reference pulse is included, starting from linear interpolation that assumes very

little about the reference pulse and ending with the optical filter, that includes shape and noise statistics.

From a practical point of view, we are particularly interested in the expected timing resolution for sampling frequencies over the 50–80 MHz range, where it is feasible to integrate the digital processing electronics into a medium-cost FPGA. Table I summarizes interpolated results in this range and shows that, within these sampling frequencies, the optical filter outperforms other filter designs.

As observed Fig. 3 (right), timing estimators introduce systematic errors that can be estimated and compensated for based on the acquired data, without any previous calibration. Because of this compensation, the timing resolution of the optical filter is even further reduced, as shown in Table I and Fig. 5.

The results summarized in these plots are compatible with those presented by other authors in [6], [7], [20]–[22], but nevertheless the latter represent single points of the much wider design space covered in this work.

Although it is commonly accepted that the shorter the rise time of the scintillation pulse the better the time resolution,

when it comes to timing by digital means, some smoothing is required in order to capture as many samples from the rising edge as possible. Therefore, we have also analyzed the impact of a slightly different electronics that would introduce further shaping on the energy signal pulse, whose measured rising time is 10 ns, before sampling. To emulate this electronics, a new dataset was created by filtering the original energy signals at 625 MHz with a Gaussian filter of variable time constant between 0 and 14 ns, before applying decimating data and applying again the previously described timing algorithms.

Fig. 6 shows the benefits of a slight increase in the signal rise time, that will introduce no significant change in the acquisition dead time and which additionally may provide some improvements regarding energy estimation. These results suggest that it is possible to achieve single timing resolutions below 1.0 ns (FHWM), that is, 1.41 ns in coincidence mode, when the proper smoothing is selected, even at modest sampling frequencies, when the proper filter and shaping time are selected. However, the actual figure varies significantly with the scintillation crystal, the sampling frequency, and filter type.

V. CONCLUSION

This work analyzes the expected resolution of the time stamp when computed by digital means for a LYSO pulse with a nominal rise time of 10 ns and a decay time of 40 ns. Different algorithms with a similar hardware implementation were considered in the analysis. In all cases, the time stamp was computed after filtering the input energy signal and finding its zero crossing through interpolation. The motivation to a solution of this type versus other more elaborated solutions was that FIR filtering and dividing are common operations in signal processing that are easily and efficiently implemented on programmable hardware, enabling direct real-time computation of the time stamp. As an example case, the implementation of an optical matched filter is described in the Appendix, including HDL simulations that show the real time feasibility of methods. For instance, the implemented hardware block is able to compute a new time stamp every 13 clock cycles, i.e., around 5 Mcps for a 65 MHz sampling clock.

Some of the presented methods, namely the matched filter and the optical filter, require a template of the expected input signal, in which case the reference is derived from the two-exponentials model, which is suitable for describing pulses from most scintillators. This model-based approach has the advantage of not requiring previous preprocessing or calibration in order to compute the adequate filter response at the cost of potential performance degradation when the template does not match the actual scintillation pulse shape.

Filter performance has been analyzed for a wide range of sampling frequencies and it has been shown that, within the sampling rate of interest (around 40–100 MHz), results are promising, as most of them provide enough resolution for the application of a 10 ns timing window, and in the particular case of the optically matched filter it seems feasible to apply a 5 ns timing window.

However, we must be aware that the presented results are applicable to a single detector and neglect fixed point effects as well as timing variations due to component and clock jitter among several detectors in the ring. Some of these uncertainties may be compensated by calibration and some not, and the impact of these on timing resolution is open for further investigation.

APPENDIX
HARDWARE IMPLEMENTATION

To show the feasibility of the presented methods for real-time digital timing, a timing block that combines the timing algorithm described in (15) and the correction methods of (20) was described with VHDL and synthesized with ISE 8.3 (Xilinx Inc., San José, CA, USA) into a Virtex-II FPGA from Xilinx.

The timing module assumes that polynomial coefficients are computed beforehand during calibration.

The implemented timing module comprises the following subunits, as shown in Fig. 7:

- An FIR filter, which is implemented as a 16-tap distributed filter and generated with the core generator Coregen from Xilinx,
- an interpolation unit, consisting of a zero-crossing (ZC) detector and a nonrestoring fixed point Sweeney, Robertson, Tocher (SRT) divider, and
- a polynomial linearization block.

The FIR filter processes the scintillation pulse energy at the sampling frequency and, as previously explained, its output has an abrupt transition from positive to negative values that is related to pulse timing. Fig. 8 shows the waveforms of a VHDL simulation, where the sequence of events is highlighted. A finite state machine (FSM) monitors for the sign transition of the filtered energy signal and performs a fixed point division to interpolate the zero crossing point, based on the values of the last positive and first the negative values of the filtered energy signal $p_{OMF}[n]$. The result of this division provides the fractional part $\widetilde{\tau}$ of the time stamp, whose value is corrected by the linearization block in order to compensate for the method's nonlinearities.

The timing block is pipelined in such a way that a new time stamp may be computed every 13 clock cycles, being the limiting factor the area/performance trade-off adopted at the SRT divider. Figure shows an VHDL simulation of the first two blocks, where the input energy, the filtered energy and the division output are highlighted with dashed circles.

The main complexity of the linearization block lies in the polynomial error estimation block, which can with a look-up table or with the datapath is shown in Fig. 9. For its development, (18) is unrolled into two branches, one to iteratively compute the powers of $\widetilde{\tau}$ and another to compute the summation as a sequence of multiplication and accumulations, as follows.

$$
\begin{aligned}
y_0 &= 1 \\
y_j &= x^j = x \cdot y_{j-1} \\
f_0 &= a_0 \\
f^j &= f^{j-1} + a_j \cdot y_j
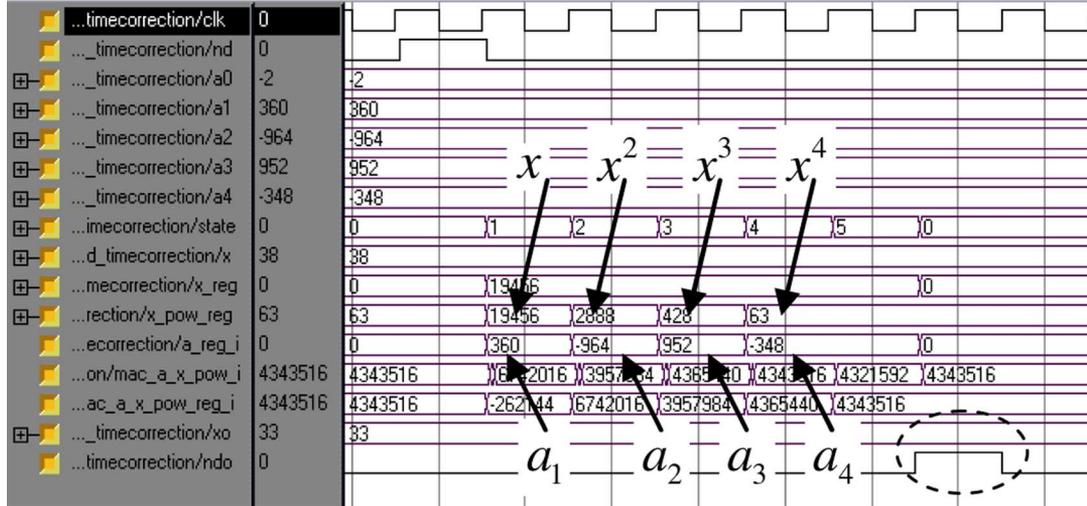\end{aligned}
\tag{21}
$$

Fig. 10.   VHDL simulation of the polynomial correction module.

For a proper fixed point implementation of (21), the FPGA must also deal with data scaling and with the architectural properties of the device, particularly the presence of 18×18 embedded multipliers. For the discussion, we consider that the fixed point representation of the different powers x^i of x is (1.17), that is 18 bits and a scale factor of 2^{-N} with $N = 17$. Moreover, the fixed point representation for the polynomial coefficients a_j is (A,0), with an implicit scale factor of 2^{-M}. The output values f are represented with P bits and an implicit scale factor of 2^{-M}.

$$a'_i|_{(A,0)} = a_i \cdot 2^M$$
$$y_1|_{(1.N)} = x|_{(1.N)} = x \cdot 2^N$$
$$y_j|_{(1.2N)} = (x|_{(1.N)} \cdot y_{j-1}|_{1.N})$$
$$y_j|_{(1.N)} = y_j|_{1.2N} \gg N$$
$$f_0|_{(P-N.N)} = a'_0|_{A,0} \cdot 2^N$$
$$f_j|_{(P-N.N)} = f_{j-1}|_{(P-N.N)} + a_j|_{(A,0)} \cdot y_j|_{(1.N)} \quad (22)$$

Figure shows the simulation of the polynomial error estimation block, taking as input the time stamp value provided by the interpolator block and, for the polynomial $(-348 \cdot x^4 + 952 \cdot x^3 - 964 \cdot x^2 + 360 \cdot x - 2) \cdot 2^{-9}$, values that were obtained from experimental data based on the linearization method described. In this example, the fractional part of the time stamp is estimated as $38_{(1.8)}$, which is equivalent to $19456_{(1.17)}$, and the module produces after six clock cycles $f = 33$, which represents the value 0.0645 in the fixed point simulation, whereas the actual value of f in the floating point simulation would be 0.0647.

Finally, the analysis results show the feasibility of the proposed timing module, requiring as little as 2 *MULT1* 8×18, 1590 *flip-flops* and 1419 look-up tables *LUTs*, values that represent less than 50% of the resources in the smallest available Virtex-II and barely 2% in the biggest device.

REFERENCES

[1] G. Kontaxakis, "Positron emission tomography," in *Encyclopedia of Medical Devices and Instrumentation*. New York: Wiley, 2006, pp. 406–418.
[2] C. J. Thompson and A. L. Goertzen, "A method for determination of the timing stability of PET scanners," *IEEE Trans. Med. Imag.*, vol. 24, no. 8, pp. 1053–1057, 2005.
[3] E. Raisanen-Ruotsalainen, T. Rahkonen, and J. Kostamovaara, "An integrated time-to-digital converter with 30-ps single-shot precision," *IEEE J. Solid-State Circuits*, vol. 35, no. 10, pp. 1507–1510, 2000.
[4] P. Dudek, S. Szczepanski, and J. V. Hatfield, "A high-resolution CMOS time-to-digital converter utilizing a Vernier delay line," *IEEE J. Solid-State Circuits*, vol. 35, no. 2, pp. 240–247, Feb. 2000, 2000.
[5] B. K. Swann *et al.*, "A 100-ps time-resolution CMOS time-to-digital converter for positron emission tomography imaging applications," *IEEE J. Solid-State Circuits*, vol. 39, pp. 1839–1852, 2004.
[6] A. Mann *et al.*, A Sampling ADC Data Acquisition System for Positron Emission Tomography. pp. 296–300.
[7] G. Hegyesi *et al.*, "Ethernet based distributed data acquisition system for a small animal pet," *IEEE Trans. Nucl. Sci.*, vol. 53, no. 4, pp. 2112–2117, 2006.
[8] A. Bousselham and C. Bohm, "Sampling pulses for optimal timing," *IEEE Trans. Nucl. Sci.*, vol. 54, no. 2, pp. 320–326, 2007.
[9] P. Soo-Chang and W. Peng-Hua, "Closed-form design of all-pass fractional delay filters," *IEEE Signal Process. Lett.*, vol. 11, no. 10, pp. 788–791, 2004.
[10] D. Tian-Bo, "Noniterative WLS design of allpass variable fractional-delay digital filters," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 53, no. 2, pp. 358–, 2006.
[11] V. Valimaki and M. Karjalainen, Implementation of Fractional Delay Waveguide Models Using Allpass Filters. pp. 1524–1527, 1995.
[12] M. Streun *et al.*, "Coincidence detection by digital processing of free-running sampled pulses," *Nucl. Instrum. Methods Phys. Res. A*, vol. A487, no. 3, pp. 530–534, 2002.
[13] R. Fontaine *et al.*, "Real time digital signal processing implementation for an APD-based PET scanner with phoswich detectors," *IEEE Trans. Nucl. Sci.*, vol. 53, no. 3, pp. 784–788, 2006.
[14] M. Z. Komodromos, S. F. Russell, and P. T. P. Tang, "Design of FIR Hilbert transformers and differentiators in the complex domain," *IEEE Trans. Circuits Syst. I, Fund. Theory Appl.*, vol. 45, no. 1, pp. 64–67, 1998.
[15] E. Geraniotis and H. Poor, "Robust matched filters for optical receivers," *IEEE Trans. Commun.*, vol. 35, no. 12, pp. 1289–1296, 1987.

[16] A. O. Hero III, "Timing estimation for a filtered Poisson process in Gaussian noise," *IEEE Trans. Inf. Theory*, vol. 37, no. 1, pp. 92–106, 1991.

[17] M. Streun *et al.*, "Timemark correction for the ClearPET™ scanners.," in *Proc. IEEE Nuclear Science Symp.*, 2005, pp. 2057–2060.

[18] Y. Wang *et al.*, "Performance evaluation of the GE healthcare eXplore VISTA dual-ring small-animal PET scanner," *J. Nucl. Med.*, vol. 47, no. 11, pp. 1891–1900, 2006.

[19] J. Seidel, J. J. Vaquero, and M. V. Green, "Resolution uniformity and sensitivity of the NIH ATLAS small animal PET scanner: Comparison to simulated LSO scanners without depth-of-interaction capability," *IEEE Trans. Nucl. Sci.*, vol. 50, no. 5, pp. 1347–1350, 2003.

[20] J.-D. Leroux *et al.*, "Time discrimination techniques using artificial neural networks for positron emission tomography.," in *Proc. IEEE Nuclear Science Symp.*, 2004, pp. 2301–2305.

[21] E. Albuquerque *et al.*, "Performance simulation studies of the clear-PEM DAQ/trigger system.," *Proc. 14th IEEE-NPSS/RT*, pp. 151–155, 2005.

[22] R. Abbiati *et al.*, "Application of a digital technique for timing of events from scintillation detectors," *IEEE Trans. Nucl. Sci.*, vol. 53, no. 6, pp. 3850–3854, 2006.