Auto-calibrating TDC for a SoC-FPGA data acquisition system

P. Carra[1,2], M. Bertazzoni[1,2], M. G. Bisogni[1,2], J. M. Cela Ruiz[3], A. Del Guerra[1,2], D. Gascon[4], S. Gomez[4], M. Morrocchi[1,2], G. Pazzi[1,2], D. Sanchez[4], I. Sarasola Martin[3], G. Sportelli[1,2], N. Belcari[1,2]

Abstract– In this work, an FPGA-based plain delay line TDC is presented, together with a theoretical model on its timing properties. The TDC features an automated calibration system implemented in the on-chip processor of a SoC-FPGA, uses a low amount of FPGA resources and is therefore suitable for applications requiring a high number of channels such as Time-of-Flight Positron Emission Tomography. We first investigated the importance of calibration and validated the theoretical model on the TDC timing properties. Finally, the device has been embodied into a two channel Positron Emission Tomography acquisition system and tested. We found the calibration essential to obtain a good time resolution (38 ps FWHM in comparison with a 78 ps FWHM obtained with the uncalibrated device). The model we developed is able to predict the TDC timing properties. They are shown to be related to the fundamental parameters of the used FPGA technology. In particular, the best achievable time resolution of this specific architecture (plain tapped delay line on FPGA) is set to about 30 ps by the sum of the setup and hold times of the registers in the FPGA. The timing resolution of the two-channel setup is about 118 ps.

Index Terms—FPGA, positron emission tomography, scintillators, time-of-flight, time-to-digital converters.

I. INTRODUCTION

HIGH resolution instruments for measuring temporal intervals find various applications both in the industry and in research. Examples are optical spectroscopy, mass spectroscopy and positron emission tomography (PET), to name a few [1]. In the case of PET, timing resolutions in the order of a few hundreds of picoseconds enable the measurement of the difference in time of flight (ToF) of the two γ-rays coming from a positron-electron annihilation. This technique improves the statistical information carried out by each pair of coincident photons and eventually allows increasing the image signal to noise ratio and its robustness against image artifacts [2]. The performance of modern SiPMs now allows to reach timing resolutions below 100 ps. In order to exploit the full performance of these photodetectors, fast time to digital converters (TDC) are required. For their massive applications in systems with many independent detectors, these devices have to be sufficiently precise, scalable to a high number of channels and use a small amount of electronical resources.

A typical TDC implementation produces a coarse timestamp, obtained with a simple clock counter, and a fine one. In the tapped delay line TDC architecture, an input trigger feeds a chain of short delay elements whose output is sampled at every clock cycle. The sampled delays are then encoded to find the fraction of the clock cycle at which the input trigger arrived, i.e., the fine timestamp.

In ASICs, these delay lines can be made of identical circuit elements, so as to provide a uniform delay throughout the chain. The delay introduced by each element is also controllable with the use of a feedback to react to modifications in the operating

P. Carra is with the Istituto Nazionale di Fisica Nucleare (INFN), section of Pisa, PI 56127, Italy and the University of Pisa, PI 56126, Italy (e-mail: pietro.carra@pi.infn.it).

M. Bertazzoni is with the Istituto Nazionale di Fisica Nucleare (INFN), section of Pisa, PI 56127, Italy and the University of Pisa, PI 56126, Italy (e-mail: matteo.bertazzoni@pi.infn.it).

M. G. Bisogni is with the Istituto Nazionale di Fisica Nucleare (INFN), section of Pisa, PI 56127, Italy and the University of Pisa, PI 56126, Italy (e-mail:giuseppina.bisogni@pi.infn.it).

J. M. Cela Ruiz is with the División de Instrumentación Científica Dpto. Tecnología. CIEMAT, Avd. Complutense 40, Ed.22.P1.16 28040 – Madrid (e-mail: josemanuel.cela@ciemat.es).

D. Gascon is with the Dept. FQA, Institut de Ciències del Cosmos (ICCUB), Universitat de Barcelona, Martí Franquès 1, E08028 Barcelona, Spain (e-mail: david.gascon@icc.ub.edu)

S. Gomez is with the Dept. FQA, Institut de Ciències del Cosmos (ICCUB), Universitat de Barcelona, Martí Franquès 1, E08028 Barcelona, Spain (e-mail: sgomez@icc.ub.edu)

A. Del Guerra is with the Istituto Nazionale di Fisica Nucleare (INFN), section of Pisa, PI 56126, Italy and the University of Pisa, PI 56126, Italy (e-mail: alberto.del.guerra@unipi.it).

M. Morrocchi is with the Istituto Nazionale di Fisica Nucleare (INFN), section of Pisa, PI 56126, Italy and the University of Pisa, PI 56126, Italy (e-mail: matteo.morrocchi@pi.infn.it).

G. Pazzi is with the Istituto Nazionale di Fisica Nucleare (INFN), section of Pisa, PI 56126, Italy and the University of Pisa, PI 56126, Italy (e-mail:giulia.pazzi@pi.infn.it).

D. Sanchez is with the Dept. FQA, Institut de Ciències del Cosmos (ICCUB), Universitat de Barcelona, Martí Franquès 1, E08028 Barcelona, Spain (e-mail: dsanchez@icc.ub.edu)

I. Sarasola Martin is with the División de Instrumentación Científica Dpto. Tecnología. CIEMAT, Avd. Complutense 40, Ed.22.P1.16 28040 – Madrid (e-mail: iciar.sarasola@ciemat.es).

G. Sportelli is with the Istituto Nazionale di Fisica Nucleare (INFN), section of Pisa, PI 56126, Italy and the University of Pisa, PI 56126, Italy (e-mail: giancarlo.sportelli@unipi.it).

N. Belcari is with the Istituto Nazionale di Fisica Nucleare (INFN), section of Pisa, PI 56127, Italy and the University of Pisa, PI 56126, Italy (e-mail: nicola.belcari@unipi.it).
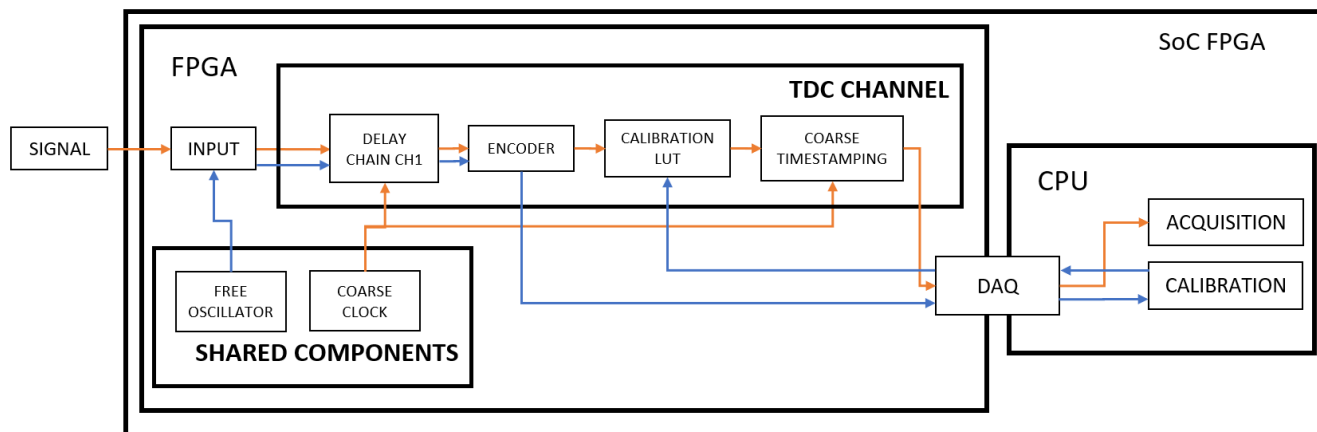
Fig. 1: Block diagram of the TDC implementation. During normal acquisitions the input signal follows the orange path, entering the delay chain, going through the encoder and the calibration look-up table (LUT), and finally being transferred to the CPU for permanent storage. During the calibration phase, the data flow follows the blue lines. The TDC uses a free oscillator as input which is encoded and then transferred directly to the CPU for the creation of the calibration LUT. The LUT is then loaded into the FPGA and will be used in the normal acquisitions.

conditions such as voltage or temperature changes.

With FPGAs, it is not possible to custom design a chain of identical circuit elements: delay chains can be implemented, though they are not fine-tunable and suffer from relatively higher differential nonlinearity (DNL). In addition, the control of the placement and routing of the used resources with the specific aim of equalizing the delays of an asynchronous path is not supported by FPGA vendor provided software. As a consequence, realizing a high-resolution, scalable and portable FPGA-based TDC architecture is still today a challenging task. However, FPGA-based data acquisition systems provide several advantages, including lower development costs and the possibility of early photon processing at the front-end to optimize the data transmission bandwidth and pixel identification efficiency [3], thus motivating the research in FPGA-based TDC architectures.

In this work, we discuss a tapped-delay-line TDC architecture implemented in FPGA. Our device focuses on addressing the following main issues: (1) the delays introduced by each element of the delay chain are not only uneven per se, but they depend on temperature and power supply voltage stability [4]. We address this issue by implementing a software auto-calibration procedure running in the on-chip processor of a SoC-FPGA, which automates the calibration and loosens the constraints on temperature and supply voltage stabilization of the FPGA. It is important to note that the presence of a processor in the device is not necessary for the implementation of a calibration procedure. It could be done, e.g., with a soft core in the FPGA, or with native HDL components, but our approach allows to save FPGA resources for implementing a higher number of channels.

(2) Automated routing of a TDC delay line has been shown to be often sub-optimal, but a hand-routed delay chain would lack in scalability and portability on different FPGA devices, thus limiting its general application. The implementation described in this paper requires no manual placing of the resources.

(3) Complex TDC architectures, while allowing to reach higher resolutions, use more FPGA resources to encode the data produced by the delay line and therefore limit the amount of TDC channels that can fit in a single device [5], [6]. The proposed architecture uses the least amount of resources while maintaining a resolution good enough not to impact the final timing performances of state-of-the-art ToF-PET systems. A different TDC implementation with similar performances and resource usage is described in [7]. The TDC used in that case is based on the Vernier method [8] and implements an automated calibration to measure the period of the two oscillators. In both cases the calibration is essential for producing a linear TDC output and for correcting for temperature variations during operations.

A theoretical model of TDCs based on plain tapped delay lines, suitable for generic FPGAs, is also described. The model allows to understand the time resolution limits of this kind of architecture considering both the delay introduced by each element of the line and the metastability issues caused by the fast propagation of the signal inside the line.

The realized TDC has been integrated into a two channel ToF-PET data acquisition system and tested, to verify the model predictions and to assess its performances.

The rest of the manuscript presents the implementation and characterization of the realized TDC and is structured as follows: firstly, the TDC implementation is outlined, together with the calibration procedure employed. Secondly, a theoretical TDC model is introduced that enumerates the sources of non-linearities encountered in FPGA-based TDCs and explains how we identify and correct these non-linearities. We then present and discuss the measurements performed and the results obtained.

II. Materials and Methods

IIa. TDC implementation

The TDC is schematized in Fig. 1 and is composed by a 400 MHz clock feeding a 32-bit counter for coarse timestamping, a

delay line, an encoder, a calibration component and a calibration look-up table (LUT). The whole design is fully pipelined, so that each channel can process inputs with a dead time of one clock period, i.e., 2.5 ns.

When the input signal enters the delay line, it is assigned a coarse timestamp (from the counter) and a fine timestamp (produced by the delay line and converted into a number by the encoder). This fine timestamp is then calibrated using the calibration LUT to eliminate any non-linearity introduced by the delay line.

The delay line has been implemented using the fast carry chain logic of the adder blocks in a last generation Altera SoC-FPGA (Intel Corp., Santa Clara CA, USA) [5], [6], [9], [10]. A 128-bit adder is instantiated in the FPGA to sum a vector of '1's and a vector of '0's and the input signal is fed as the first carry in of the adder (Fig. 2). The output of the adder is registered and reset at each clock cycle. While the signal travels down the carry chain, the '1's of the output vector become '0's producing a so-called "thermometer code".
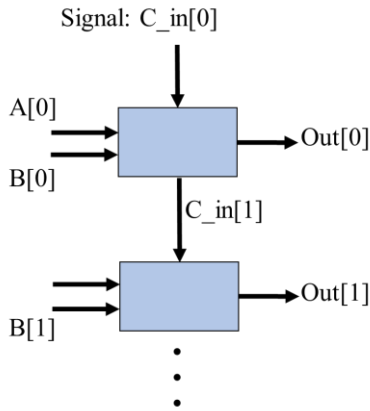

Fig. 2: Schematic of an FPGA adder.

The position of the transition point between the sequence of '0's and the sequence of '1's in the thermometer code gives the number of elements the signal has gone through before the last clock rising edge. This information represents the fine timestamp described in the introduction and is used together with the coarse counter measurement to assign the final timestamp to the input signal. Carry chains are the shortest signal path available to the user in most FPGAs and therefore they are the best suited for TDC applications. The conversion from the thermometer code to the final TDC code is performed by means of a so-called "thermometer-to-binary" encoder [11]. During registration, metastable states or inhomogeneities in signal propagation speed may generate "bubbles" in the registered output. Therefore, the encoder must be able to produce a meaningful output even when the transition point is not defined, e.g., as in the following case ('1's propagate from right to left):

$$00000 \ldots 00000 \underbrace{11001010}_{bubble} 11111 \ldots 11111).$$

With high-end FPGAs, the signal propagates faster than in lower-end devices and thus bubbles are on average bigger and more frequent. A way to obtain a reliable encoded value is to count the number of '1's in the adder output instead of finding

the transition edge [11]. A Wallace tree encoder has been used for this purpose [12]. The encoder recursively reduces the delay chain vector by summing its bits in groups as schematized in Fig. 3. The final length of the encoded value is $L = \log_2(n_{el} + 1)$, where n_{el} is the number of delay elements in the chain. In this case $n_{el} = 127$ and thus $L = 7$. The delay line is actually 128-element-long, but the last element is not fed to the encoder and is used to represent the fact that the signal has reached the end of the delay line before the rising edge of the clock, and thus must be rejected since its fine time cannot be determined.
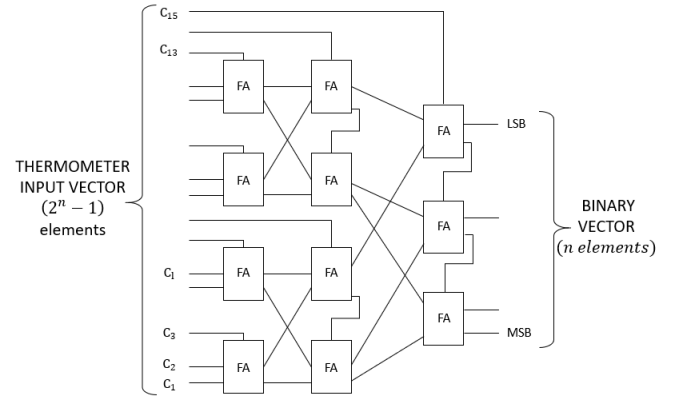

Fig. 3: Schematic of a Wallace tree encoder. A sequence of Full Adders (FA) progressively reduces the number of bits of the input vector summing its bits in pairs.

The operations of the encoder can be easily pipelined, allowing to convert the thermometer codes into fine timestamps without introducing dead time. Using this encoder severely reduces the probability of having missing TDC codes since it does not aim at locating the transition point in the thermometer code.

The encoder output is then calibrated with a LUT: a table that converts the fine timestamps coming out from the encoder (representing the number of delay elements traversed by the signal) into clock fractions, so that the coarse timestamp and the fine timestamp may be concatenated together. The LUT is also necessary to eliminate the non-linearities introduced by the various TDC components (see section IIc for more details on the source of the inhomogeneities).

The LUT is built by the calibration component using the statistical code density test. This test consists in sending random signals as inputs (generated by a free oscillator inside the FPGA) to the TDC delay line. The TDC codes produced by the delay line encoder are sent to the CPU of the SoC-FPGA where the histogram of the frequencies of each measured code is computed. The frequency of each TDC code is directly proportional to the width of the delay introduced by the corresponding element of the chain (i.e., the bin width), with the constraint that the total sum of the delays must be equal to one coarse clock period. A convenient way of expressing these widths is thus in terms of fractions of the coarse clock period.

If the delay line is sufficiently long, the TDC codes corresponding to the delay elements at the edge of the chain will never be output by the encoder. This happens because the signal is unable to travel that far before the clock rising edge. The

delay elements corresponding to these TDC codes are assigned a width of 0.

Every possible TDC code n is associated to a corresponding calibrated timestamp, according to $\sum_{i=0}^{n} w_i$, where w_i represents the width in terms of fractional clock parts of the i-th bin. In our implementation, the calibrated fine timestamp is encoded with 10 bits, i.e., its LSB is 1/1024th of the coarse clock period. Note that the real TDC code is only 7 bit long, we use 3 more bits in order not to lose precision during the LUT conversion.

Since the delay introduced by each element in the chain may vary with operating conditions such as temperature and voltage, it is important to recalibrate the TDC at regular intervals. To do this, a programmable timer in the CPU periodically activates the calibration. During the calibration, the TDC has its input redirected to the internal oscillator and thus it is not available for normal operation. However, if the input rate is sufficiently high, also acquired data can be used to calibrate the TDC channels on the fly, completely eliminating the need to suspend the acquisition during calibration.

IIb. Resource usage

A single delay line occupies 64 Algebraic Logic Modules (ALM, the building blocks of Altera FPGA devices), while the encoder uses up to about 300 ALMs and a LUT occupies 1270 bits of memory.

A low-cost FPGA may contain more than 100.000 ALMs; meaning that the main limit on the number of channels is posed by the input pins available in the FPGA. Our 64 channels implementation uses 128 pins out of the 288 available in the FPGA. In this implementation, each channel has its own encoder and LUT, and a shared calibration logic.

IIc. TDC model

A theoretical model of delay-line based TDCs realized in FPGAs has been developed and its predictions can help both estimating the time resolution of the TDC and assessing its limits in terms of nonlinearity and acquisition time.

A delay line is composed by a sequence of elements, each one introducing an ideally equal delay. In practice, the delays can vary greatly between the various elements of the chain, for different reasons depending on the technology used to implement them.

In the tapped delay line, every element is sampled concurrently at the rising edge of a clock signal, but in an FPGA the clock distribution network can cause inhomogeneities due to the different path lengths that the clock signal has to go through to reach every element of the chain. In other words, the exact sampling instant may not be the same for all the delay elements.

Another issue that can give rise to uneven delays between the bins is the periodic structure of the FPGA fabric: the delay elements are not equally spaced and there are periodic gaps that can increase the signal propagation time due to the interconnection between the FPGA cells.

Finally, the signal speed is influenced by temperature and voltage, being higher when temperature decreases and voltage increases. However, it has been shown that voltage variations are not significant enough to cause sensible changes in the delay structure [13].

Based on these considerations, it is possible to model the delay δ_i introduced by the i-th element of a tapped delay line implemented in an FPGA with the following formula:

$$\delta_i = d + c_i + p_i + t(T)$$

Where d is the standard delay introduced by every element, depending on the type of operation it performs on the signal and c_i, p_i and $t(T)$ are the delays due to the clock distribution network, the periodic structure of the FPGA and the temperature respectively. Note that the chain (128-element long) is short enough that the temperature can be considered uniform through it.

A first rough estimate of the best possible resolution obtainable with a tapped delay line TDC gives a value related to the bin width with the well-known quantization error formula $\sigma = \delta/\sqrt{12}$, where σ is the temporal resolution expressed as a standard deviation, δ is the average width of the bins and $\sqrt{12}$ accounts for the quantization error. However, the bin width is not the only uncertainty in these TDCs. In fact, also the metastability issues play a role in the determination of the resolution.

The input of a register must be stable for a minimum time before the clock edge (register setup time or t_{SU}) and for a minimum time after the clock edge (register hold time or t_H). If a signal transition violates a register's t_{SU} or t_H requirements, the output of the register may go into a metastable state, in which it may hover at a value between the high and low logic states or oscillate between the two states for some period of time.

In the specific TDC case, the delay line is sampled every clock cycle through a sequence of registers. When the signal generated by the photodetector is traversing the chain and the sample operation occurs, the registers immediately before (after) the signal edge may incur in metastability, due to the fact that their input signal violates their setup (hold) time.

In this case, it is not the fact that the state is metastable (i.e., not defined) that creates the problem (because we could just wait for it to resolve), but also the fact that it could resolve to the wrong logic state, which would give an incorrect measure of the number of stages that the signal has gone through.

This effect is observable only if the bin width is sufficiently short (i.e., shorter than $t_{SU} + t_H$) but in these cases, it can heavily affect the resolution of the TDC. To quantify the influence of metastability, t_{SU} and t_H must be known. In the most commonly used devices $t_H \simeq 10$ ps and $t_{SU} \simeq 20$ ps [14], [15], [16]. To get an approximate estimate of the contribution of these effects to time resolution, metastability can be viewed as an added uncertainty of $\Delta_{met} = t_{SU} + t_H$. As a direct consequence of the model presented in this work, we can identify the best resolution achievable with TDCs based on a plain delay-line implemented in FPGAs: even when the bin width is made negligibly small (e.g., 1 ps), the metastability issues remain, meaning that the resolution cannot be better than

about 30 ps FWHM. This is especially important since t_H and t_{SU} are parameters fixed by the materials and the process with which registers are built; and improving them by an order of magnitude would require changing the production techniques of these devices.

IId. DNL corrections

One of the major issues we found in developing an FPGA-based TDC is the high DNL of the delay chain; the DNL origins from all those factors that are dependent on the specific element of the line, namely the effect of the clock distribution network, c_i, and the effect of the periodic FPGA structure, p_i.

The dominating factor, and also the one that can be studied more in-depth is p_i, which is due to the Logic Array Block (LAB) structure of Altera FPGAs. Each LAB contains 10 Algebraic Logic modules (ALMs) elements and can form a 20-element-long carry chain; when the carry propagates beyond the LAB boundaries an extra delay is added causing the formation of ultra-wide periodic bins [14]. The other DNL source is c_i: the clock signal that drives the different flip flop in the adder array is not simultaneous. To mitigate the DNL effects, the chain is kept short and, therefore, the coarse clock is run as fast as possible. Given the very fast carry propagation in last generation devices a 128-element chain is needed with a 400 MHz clock to cover the whole period with a good safety margin to avoid saturating the chain.

An important feature of this TDC is that the resource placement in the FPGA does not need any manual adjustments. A uniform structure of the delay chain is essential to reduce the DNL, but while instantiating an adder and turning on speed optimization in the synthesizer usually produce a contiguous carry chain, a constant path from each adder cell to its register is more difficult to achieve, (see Fig. 4). The easiest way to ensure uniformity would be to manually place each adder cell and each register in the FPGA, but it would be extremely time consuming and is not an option when the channels grow significantly in number. Therefore, a VHDL description with which the synthesizer constantly reproduces the desired structure has been found.

However, having a short and uniform delay chain is not sufficient to mitigate the effects of DNL and some sort of calibration is necessary. In this TDC, the calibration is done as described in section IIa. The results of the calibration can serve as a look up table to convert the TDC encoded value from the number of delay elements the signal has gone through to a time measure in picoseconds.
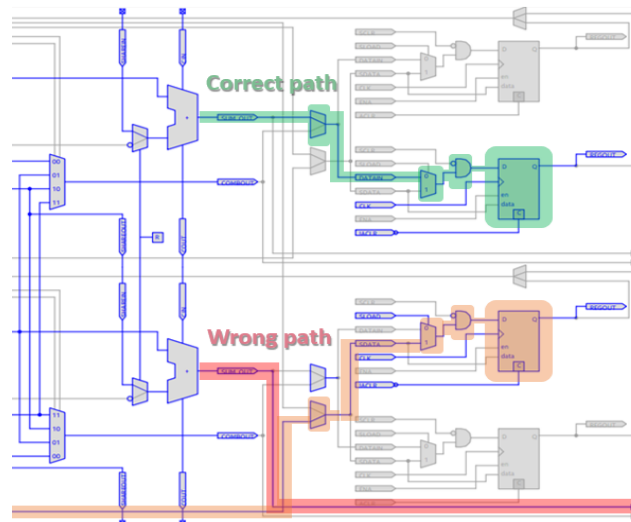

Fig. 4: View of an adaptive logic module (ALM) from the Chip Planner in Quartus. In the upper cell the correct path from the adder to the register is highlighted in green. In the lower one the register is occupied by the output of another adder coming from outside the ALM.

IIe. TDC characterization

We have implemented a calibration procedure in order to correct the effects of possible variation of the device operating conditions like temperature and voltage. To study the magnitude of these effects, we calibrated the TDC at power-up and then made a series of measurements of a known time gap at regular intervals.

At the very beginning of operations, the device is relatively cold, then for some time the temperature grows until it reaches a stable value. When the temperature varies, the accuracy of the measurement decreases. The effects of temperature on delay-line based TDCs have also been reported in [7], where a 15 °C temperature variation had a significant impact on the period of the delay line oscillators. To assess the importance of these effects in our implementation, 100 measurements of a known time interval have been taken every minute calibrating the TDC only at the beginning of the acquisition. These measurements have been compared with those obtained calibrating the TDC every 10 s. For each measurement, 250,000 events have been registered and the gaussian that best fits their time distribution has been used to estimate the time delay.

Once we determined a proper calibration procedure, we characterized the TDC performances, measuring its resolution, DNL and linearity.

The TDC resolution has been measured sending a clock synchronized with the TDC clock, but with a different phase as input signal (Fig. 5). The phase between the two clocks can be set in the FPGA firmware with a precision of 2 ps.

Controlling the phase precisely means that the delay is well known. Thus, it can be used to measure the accuracy of the TDC. This also allows to assess whether the bubble elimination performed by the Wallace tree encoder significantly affects the accuracy.
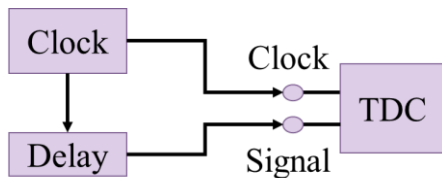
Fig. 5: Schematic of the FPGA setup for the TDC resolution measurement.

The DNL is a measure of the homogeneity between the delay introduced by each element of the delay line. A lower DNL means a higher homogeneity and thus a higher linearity of the whole line. The DNL can be formally defined as:

$$DNL_D = \left| \frac{(V_{D+1} - V_D)}{V_{IDEAL}} - 1 \right|$$

where V_D is the calibrated timestamp corresponding to the TDC output code D and V_{IDEAL} is the ideal delay introduced by a single element, which is given by the length of the time interval measured by the full delay line (in this case 2.5 ns, one clock period) divided by the number of active delay elements (the ones with a bin width higher than 0).

The linearity of the TDC can be assessed comparing the calibrated TDC codes with their nominal value in picoseconds, obtained considering the delay introduced by each element of the chain equal to the ideal one. Note that no measurements are needed for the calculation of these two parameters, only the TDC codes obtained from the calibration.

IIf. Experimental validation

To validate the system and to check the feasibility of its integration in a real PET application a two channel ToF-PET data acquisition system has been developed.

The experimental set-up consists in two NUV-SiPMs (Near-UV SiPMs) manufactured by AdvanSiD, (Trento, Italy) connected to two (3mm x 3mm x 5mm) LYSO scintillators from Saint Gobain that are used to measure the difference in the arrival times of the annihilation γ-rays emitted by a ^{68}Ge radioactive source. The SiPMs signals are read with two FlexTOT ASICs developed by the Institute of cosmos sciences (University of Barcelona) and by the Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas of Madrid [17].

The FlexTOT outputs are sent to two TDC channels.

III. RESULTS

Fig. 6 reports the measurements done in order to assess the effect of the operating conditions previously described. The great variability of measured values shows that the accuracy of the TDC decreases greatly already in the first minutes. These oscillations have an amplitude of about 50 ps and can be attributed to changes in the device operating conditions.
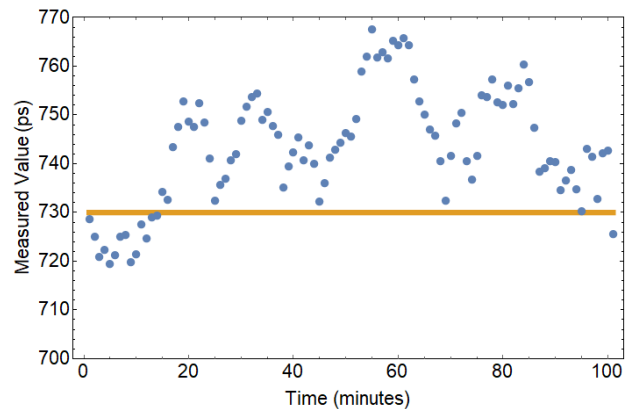

Fig. 6: Series of measurements taken without calibration. Expected (real) value is 729.9 ps (yellow line).

For comparison, the same time delay has been measured calibrating the TDC every 10 s as described in section II. Results are shown in Fig. 7.

The calibration has a significant impact on the TDC performance, reducing the oscillation amplitude to about 20 ps.
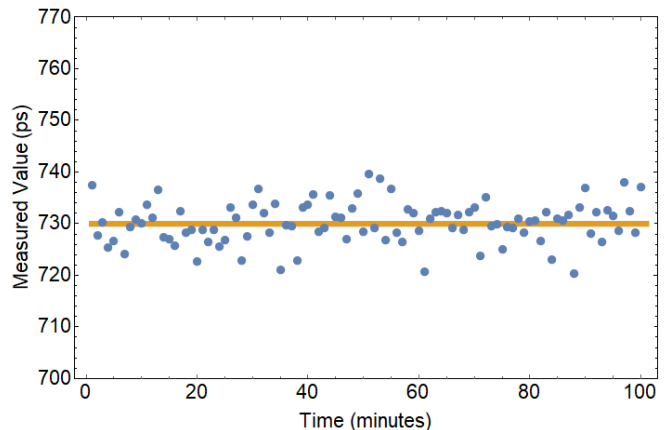

Fig. 7: Series of measurements taken with calibration. Expected (real) value is 729.9 ps (yellow line).

The resolution of the TDC has been measured with the method previously described in section IId for choosing the TDC input signals and calibrating the TDC every 10 s. The delay between the two clocks has been set from 0 ps to 2500 ps in multiples of 182.482 ps (14 total measurements); for each delay, 10,000 events have been acquired. Given that the delay is known, both accuracy and resolution can be evaluated. The difference between the real value (μ) and the measured value (m, obtained as the position of the peak of the Gaussian that best fits the time distributions) can be used to assess accuracy. The plot of ($m - \mu$) versus μ is shown in Fig. 8. The error stays approximately constant and anyway remains below 15 ps.
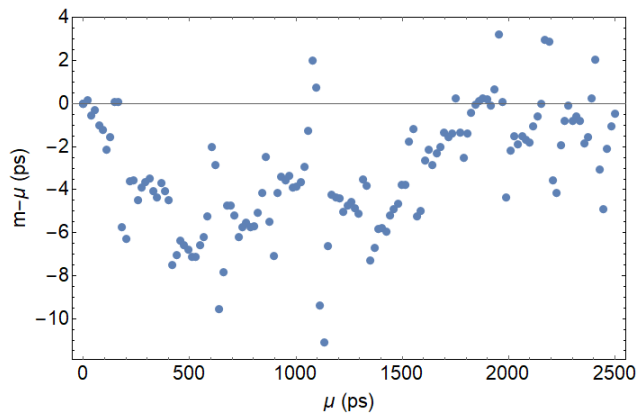
Fig. 8: Accuracy of the TDC.

The FWHM of the Gaussian that fits the distribution of the measured values is used to evaluate the TDC resolution. A histogram of the time distribution measured is presented in Fig. 9a. The average FWHM value is 38.2 ps, the maximum value is 43 ps.
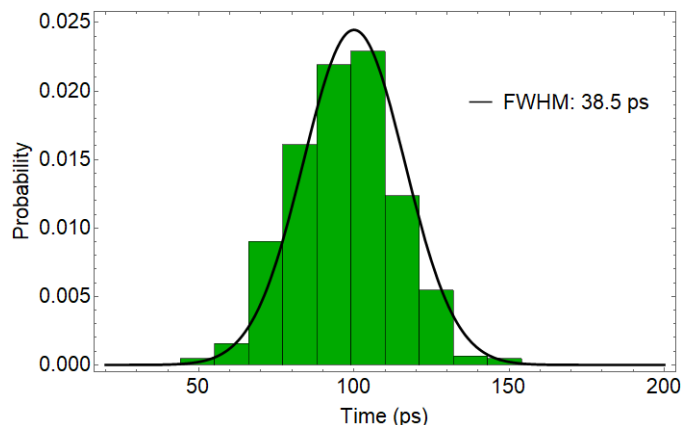

Fig. 9a: Histogram of one of the time distributions acquired to measure resolution. Calibrated TDC.
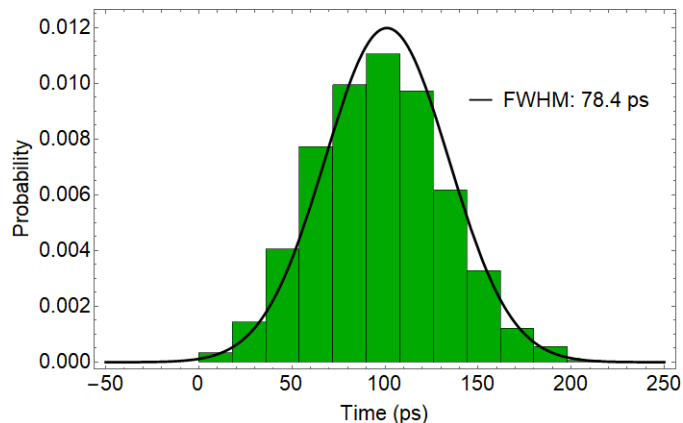

Fig. 9b: Histogram of one of the time distributions acquired to measure resolution. Uncalibrated TDC.

By contrast, in Fig. 9b we also report the resolution of the TDC on a long acquisition (1 hour) with the calibration procedure run only once at the beginning of the measurement. The time resolution is significantly worse (78.4 ps) as a

consequence of the oscillations in the bin width during the acquisition.

Once properly calibrated, the TDC can be characterized in terms of its DNL, linearity and resolution. Fig. 10 and Fig. 11 present the DNL and linearity of one TDC channel that is representative of the others. The DNL values do not vary significantly between one channel and the other.

The number of delay elements with a non-zero width is 99, thus $V_{IDEAL} = 2500/99\ ps = 25.3\ ps$.

The maximum absolute DNL value is 0.58; this value has been obtained excluding the first and final codes, since they have a DNL of -1 because the width of the corresponding delay element is 0 and therefore $V_{D+1} = V_D$.
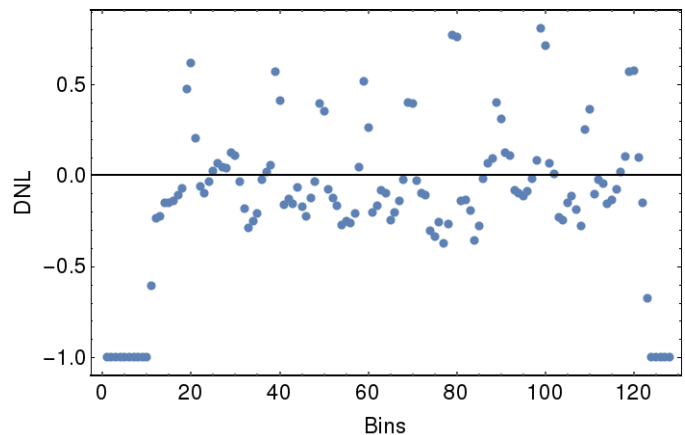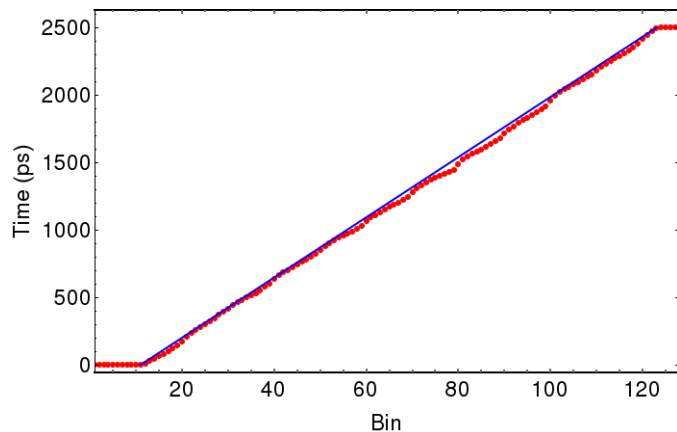

Figure 10: DNL of one TDC channel

Fig. 11: Linearity of one TDC channel

Finally, the coincidence time resolution (CTR) of the system has been measured as described in section IIe. Fig. 12 reports the histogram of the timing distribution of the coincidences, with three different source positions. The average FWHM is 116 ps, similar results have been obtained in [18] (123 ps FWHM, same experimental setup) and in [19], but in this case the TDC is a very low-cost and simple to use device.
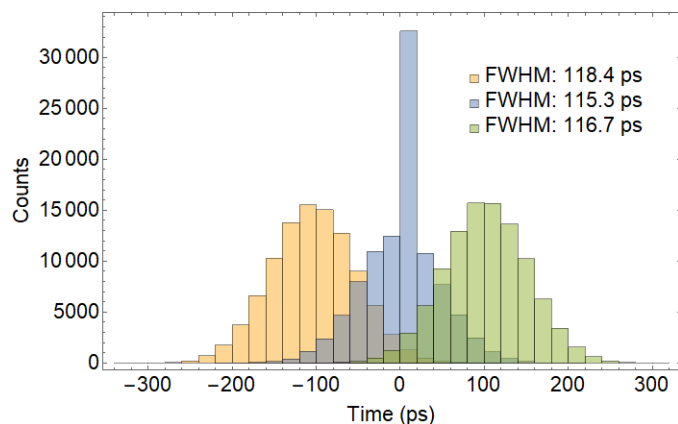
Figure 12: Histogram of the timing distributions for different source positions; from left to right: -1.5 cm, 0 cm and 1.5 cm from the center of the field of view of the instrument.

IV. DISCUSSION

As shown in Fig. 6 and Fig. 7, calibrating the TDC regularly is fundamental in order to get the best possible performances. In particular, the accuracy of the measurement can be severely impacted already after the first minutes of data acquisition. This, in turn, may also degrade temporal resolution if the measurement is sufficiently long.

An interesting aspect regarding accuracy is also shown in Fig. 8, which reports the relative error of the TDC measure with respect to the true value. All the data points are slightly negative. This can be due to the fact that the setup time of the registers that sample the delay line is higher than the hold time. Thus, metastability effects cause an underestimation of the measured values. Note that, however, the absolute error made is always < 15 ps. This means that the bubble elimination performed by the encoder does not introduce significant biases in the measurement.

The TDC also exhibits good linearity, as confirmed by Fig. 10 and Fig. 11. As expected from the use of the Wallace tree encoder, there are no missing codes.

Finally, Fig. 12 shows the obtained CTR of 116 ps. This value is compatible with what is predicted by timing models, both of LYSO scintillators and of the TDC. In fact, the TDC CTR is the resolution of one channel, 38 ps, multiplied by $\sqrt{2}$: 53 ps. This means that the rest of the detection system is responsible for $\sqrt{116^2 - 53^2}$ ps $= 103$ ps, which corresponds to some of the best time resolutions obtained in literature with L(Y)SO crystals of comparable dimensions (3x3x5 mm^3) [19], [20], [21].

V. CONCLUSIONS

The TDC we implemented and described in this work is a suitable instrument to be used both in a fully-fledged ToF-PET acquisition system and as a device for evaluating experimental setups performance given its high resolution, portability and simplicity of use.

Thanks to the relatively simple architecture (plain delay line) and encoding procedure, and to the software implementation of the calibration, the device uses a very limited amount of resources, making it possible to implement a high number of channels in a single FPGA. These design choices still allow to reach a resolution that does not degrade the time resolution of state-of-the-art ToF-PET systems. Also, since a plain delay line was used, the encoding is straightforward and can be pipelined without using a significant amount of resources, allowing for a dead time as low as 2.5 ns.

In fact, the TDC alone reaches a resolution of 38 ps FWHM, while the experimental setup as a whole has a CTR of 116 ps. These values make the importance of a continuous calibration evident, since without it, the obtained TDC time resolution is 78 ps. More complicated architectures such as those outlined in [7] may allow for resolutions that are not limited by the dimension of the minimum delay element. However, they are characterized by a higher dead-time and a more complex calibration procedure. In fact, the method described in [7] needs the TDC to be connected to an external pulse generator, while our calibration is done with internally generated pulses as described in section IIa.

The theoretical model developed is able to predict the TDC timing properties, relating them to fundamental parameters regarding the FPGA technology in which it has been implemented. In particular the lower bound to the time resolution of this specific architecture (plain tapped delay line on FPGA) is set by the sum of the setup and hold times of the registers in the FPGA, to about 30 ps in state-of-the-art devices. These limits are still unmet by the current PET detectors, but with future generations of fast scintillators and photodetectors, new and better techniques for time-to-digital conversion that conjugate low resource utilization and high resolution will soon be needed.

VI. ACKNOWLEDGMENT

REFERENCES

[1] A. Del Guerra, N. Belcari and M. Bisogni, "Positron emission tomography: its 65 years," *La Rivista Del Nuovo Cimento,* vol. 39, no. 4, pp. 155-223, Apr. 2016.

[2] L. Eriksson and M. Conti, "Randoms and TOF gain revisited," *Physics in Medicine & Biology,* vol. 60, no. 4, pp. 1613-1623, Jan. 2015.

[3] G. Sportelli *et al.,* "The TRIMAGE PET data acquisition system: initial results," *IEEE Transactions on Radiation and Plasma Medical Sciences,* vol. 1, no. 2, pp. 168-177, Mar. 2017.

[4] S. S. Junnarkar, P. O'Connor and R. Fontaine, "FPGA based self-calibrating 40 picosecond resolution, wide range time to digital converter," in *Proc. IEEE Nuclear Science Symposium Conference Record,* Dresden, Germany, 2008, pp. 3434-3439.

[5] A. Aloisio *et al.*, "High-precision Time-to-Digital Converter in a FPGA device," in *Proc. IEEE Nuclear Science Symposium Conference Record*, Orlando, FL, USA, 2009, pp. 290-294.

[6] J. Wu, "Several key issues on implementing delay line based TDCs using FPGAs," *IEEE Transactions on Nuclear Science*, vol. 57, no. 3, pp. 1543-1548, Jun. 2010.

[7] Junnarkar, Sachin S., *et al*. "FPGA-based self-calibrating time-to-digital converter for time-of-flight experiments." In *IEEE Transactions on Nuclear Science*, vol. 56 no. 4, pp. 2374-2379, Aug 2009.

[8] D. Piotr, *et al.* "A high-resolution CMOS time-to-digital converter utilizing a Vernier delay line." In *IEEE Journal of Solid-State Circuits*, vol. 35 no. 2, pp. 240-247, Feb. 2000.

[9]   L. Zhao *et al.*, "The design of a 16-channel 15 ps TDC implemented in a 65 nm FPGA," *IEEE Transactions on Nuclear Science*, vol. 60, no. 5, pp. 3532-3536, Oct. 2013.

[10]   J. Wu, S. Zonghan, "The 10-ps wave union TDC: Improving FPGA TDC resolution beyond its cell delay," in *Proc. IEEE Nuclear Science Symposium Conference Record,* Dresden, Germany, 2008, pp. 3440-3446.

[11]   P. Pereira, J. R. Fernandes, and M. M. Silva, "Wallace tree encoding in folding and interpolation ADCs," in *Proc. IEEE International Symposium on Circuits and Systems Circuits and Systems*, Phoenix-Scottsdale, AZ, USA, 2002, pp. I-I.

[12]   F. Kaess, R. Kanan, B. Hochet, and M. Declercq, "New encoding scheme for high-speed flash ADC's," in *Proc. IEEE International Symposium on Circuits and Systems Circuits and Systems*, Hong Kong, Hong Kong, 1997, pp. 5-8.

[13]   R. Szplet, J. Kalisz, and R. Szymanowski, "Interpolating time counter with 100 ps resolution on a single fpga device," *IEEE transactions on instrumentation and measurement*, vol. 49, no. 4, pp. 879–883, Aug. 2000.

[14]   Altera, "Cyclone FPGA family datasheet" [Online].
Available:
https://www.altera.com/en_US/pdfs/literature/ds/ds_cyc.pdf
Accessed: Jun. 1, 2018.

[15]   Altera, "Cyclone device handbook" [Online].
Available:
https://www.altera.com/en_US/pdfs/literature/hb/cyc/cyc_c5v1.pdf
Accessed: Jun. 1, 2018.

[16]   Altera, "Altera arria 10 device datasheet" [Online].
Available:
http://www.altera.com/en_US/pdfs/literature/hb/arria_10/a10_data sheet.pdf
Accessed: Jun. 1, 2018.

[17]   A. Comerma *et al.*, "FlexToT-Current mode ASIC for readout of common cathode SiPM arrays," in *Proc. IEEE Nuclear Science Symposium and Medical Imaging Conference,* Seoul, South Korea, 2013, pp. 1-2.

[18]   I. Sarasola *et al.*, "A comparative study of the time performance between NINO and FlexToT ASICs," *Journal of Instrumentation*, vol. 12, no. 04, p. P04016, Apr. 2017.

[19]   S. Gundacker et al., "Time of flight positron emission tomography towards 100ps resolution with 1 (y) so: an experimental and theoretical analysis," *Journal of Instrumentation*, vol. 8, no. 07, p. P07014, Jul. 2013.

[20]   R. Lecomte, "Novel detector technology for clinical pet," *European journal of nuclear medicine and molecular imaging*, vol. 36, no. 1, pp. 69–85, Mar. 2009.

[21]   M. V. Nemallapudi *et al.,* "Sub-100 ps coincidence time resolution for positron emission tomography with LSO: Ce codoped with Ca." *Physics in Medicine & Biology*, vol. 60, no. 12, pp. 4635-4649, May 2015.