A Three-Step Low-Power Multichannel TDC Based on Time Residual Amplifier

Florent Bouyjou, Eric Delagnes, Fabrice Couderc, Damien Thienpont, José David Gonzalez-Martinez, Frederic Dulucq, Fabrice Guilloux, and Irakli Mandjavidze

***Abstract*— This article proposes and evaluates an architecture for a low-power time-to-digital converter (TDC) with high resolution, optimized for high-rate operation (40 MSa/channel), and integration with analog front end in multichannel readout chips in 130-nm CMOS technology. The converter is based on a three-step architecture. The first step uses a counter and the following ones are based on two types of delay-line (DL) structures. A programmable time amplifier (TA) is used between the second and third steps to reach a final resolution of 24.4 ps in the standard mode of operation. In addition, this architecture uses common continuously stabilized reference blocks that control the channels against the effects of global process, voltage, and temperature (PVT) variations. We also propose a per-channel DL gain correction based on a trimmable block to correct the mismatch effect. The area of the TDC channel is only 0.051 mm2. For a 40-MSa/channel rate, the TDC average power consumption measured per channel is 2.2 mW for a 100% hit occupancy and decreases to 311 μW for the 10% occupancy specified for our main application. The demonstrated compactness and low power consumption fully match our requirements for integration into multichannel front-end chips. The experimental results demonstrate good timing performance over a broad range of operating temperatures (−35 °C and 65 °C), which conforms to our expectations. For example, the measured timing integral nonlinearity (INL) is better than ±1 LSB (±25 ps), and the overall timing precision is better than 21-ps rms.**

***Index Terms*— Calibration, delay line (DL), delay-locked loop (DLL), low power, multichannel, process, voltage, and temperature (PVT), time amplifier (TA), time-to-digital converter (TDC).**

I. Introduction

TIME-TO-DIGITAL converters (TDCs) are widely used in high-energy physics (HEP) and many other applications that require highly precise time measurements. During the last few years, fast-timing detectors have become increasingly important as the extra information they provide opens new possibilities for data filtering or particle identification.

Manuscript received 6 October 2023; revised 14 November 2023; accepted 19 November 2023. Date of publication 28 November 2023; date of current version 19 December 2023. This work was supported in part by the French National Research Agency (ANR) through P2IO Labex P2IO LabEx under Grant ANR-10 LABX-0038 and through Investissements d'Avenir under Grant ANR-11 IDEX-0003-01. *(Corresponding author: Florent Bouyjou.)*

Florent Bouyjou, Eric Delagnes, Fabrice Couderc, Fabrice Guilloux, and Irakli Mandjavidze are with IRFU, CEA, Université Paris-Saclay, 91191 Gif-sur-Yvette, France (e-mail: florent.bouyjou@cea.fr; eric.delagnes@cea.fr; fabrice.couderc@cea.fr).

Damien Thienpont, José David Gonzalez-Martinez, and Frederic Dulucq are with Ecole Polytechnique, CNRS/IN2P3, Omega, 91120 Palaiseau, France.

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TNS.2023.3335657.

Digital Object Identifier 10.1109/TNS.2023.3335657

Today, one of the main challenges of European Organization for Nuclear Research (CERN) collaborations for the high luminosity large hadron collider (HL-LHC) era [1], [2] is the large pileup with up to 200 simultaneous proton collisions per bunch crossing to record in a detector.

The most efficient way to mitigate the pileup in the most occupied regions of detectors is to use timing information, opening the way to 4-D-tracking (spatial and temporal). For this purpose, timing precision in the range of 30-ps rms is required. Outside the field of HEP, these new high-precision timing detectors have paved the way for improvements or new concepts in medical applications as time-of-flight positron emission tomography [3] or light detection and ranging instrumentation [4], [5].

Today, a wide range of TDC architectures are available for integration in application-specific integrated circuit (ASIC) and field-programmable gate arrays (FPGAs) [6]. Most of them are based on a digital delay line (DL). The most advanced DL-based TDC ASICs compatible with high-rate acquisition exhibit timing resolution in the 10-ps range, and the best ones reach a few picoseconds [7], [8], [9], using, for example, the resistive interpolation technique. Using advanced CMOS technology nodes, the design of high-performance TDC benefits from technology scaling and achieves fine resolution due to the fast transitions, small area, and low power [9], [10], [11]. In the LHC environment, where technologies must survive to high radiation levels, we also find high-performance TDCs. For example, Zhang et al. [11] proposed an architecture based on DL, very fast as not servo-controlled, whose gain and nonlinearities must be constantly calibrated and digitally compensated. This technique enables TDCs to consume as low as 200 μW for a 10% hit occupancy.

Most of these attractive architectures require the massive integration of correction logic which would occupy a large area, with the rather old technological node imposed (130 nm) by our main application, or the sending to acquisition of data containing redundancy for subsequent correction. This is also impossible because of the limited bandwidth allocated to us at the output of the front end and the philosophy adopted by the data-acquisition (DAQ) managers to minimize calibrations and corrections. However, for our target requirements (a few tens of picoseconds precision range and 40 MSa/channel), moderate power consumption has recently been reported using multistep structures based on the Vernier [12] or time amplifier (TA) methods for the last stage. Modern TA structures can

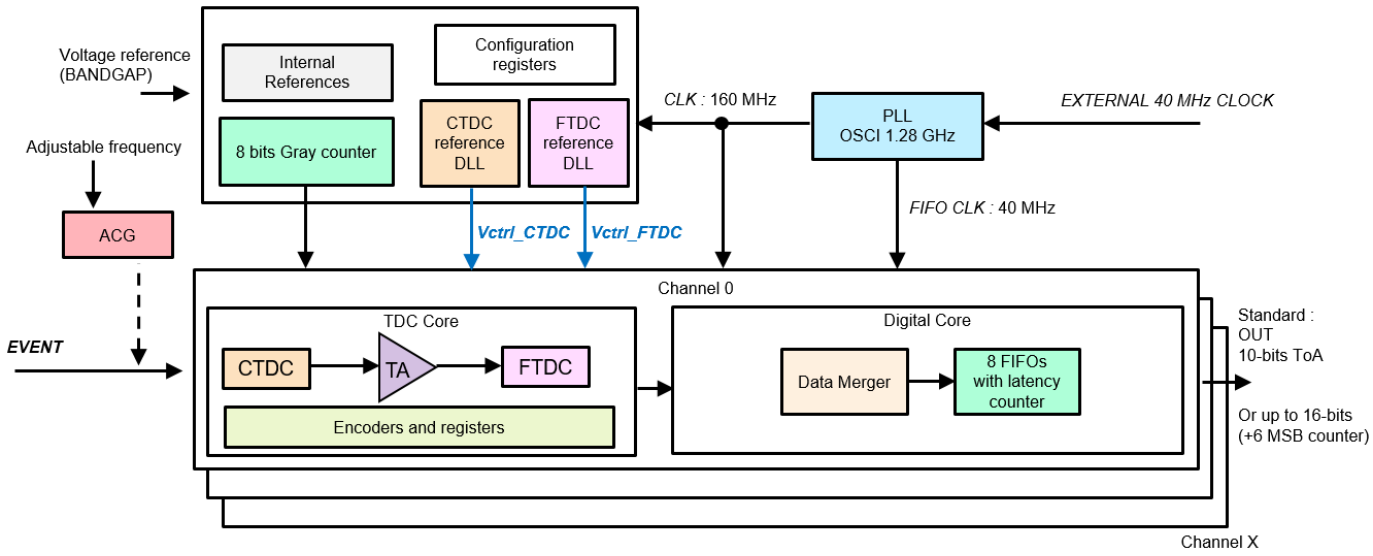

Fig. 1. Block diagram of the three-step multichannel TDC.

amplify time intervals as small as a few tens of picoseconds by an order of magnitude so that they can be measured using standard methods. This makes it a very attractive technique for fast, high-resolution TDCs [7], even if the gain of the TAs is usually set by open-loop high-speed elements, strongly depending on the process, voltage, and temperature (PVT) variations.

This article presents the design of a three-step low-power multichannel TDC, inspired by [13], based on time-residue amplification but with an extended dynamic range and improved robustness against PVT variations for HL-LHC environment. The first step captures the output of a counter operating at 160 MHz. The second step, using a DL, refines the time-stamping within the 160-MHz period, with a resolution of 200 ps. The final resolution of a few tens of picoseconds is obtained in the third step, which uses a DL fed by a programmable pulse-train TA optimized for a large dynamic range and enhanced stability.

The organization of the following sections of this article is given as follows. Section II describes the overall architecture of the three-step multichannel TDC. Section III presents the main building blocks used in the TDC in detail. The implementation and the experimental results are released in Sections IV and V. Finally, Section VI presents conclusions and outlines future work.

II. Description of the Three-Step TDC

A. Requirements

Our original goal was to design a multichannel TDC intellectual property (IP) block, implemented in the 130-nm CMOS technology from Taiwan Semiconductor Manufacturing Company (TSMC) widely used by the HEP community, compatible with integration together with low-noise analog front ends and fulfilling the Compact Muon Solenoid (CMS) high-granularity calorimeter (HGCAL) [14] requirements but also easily reusable for other applications. Its resolution must be better than 30 ps over a range of at least 25 ns (>10-bit dynamic range), which corresponds to the LHC bunch-crossing period, and its target timing precision should be 20-ps rms (after calibration). Even if optional features of the design allow the extension of its dynamic range, this study focuses on the nominal mode of operation (10-bit range with 25-ps LSB). The total conversion time must be less than 25 ns, leading to a conversion rate of up to 40 MSa/channel required for the LHC operation. Because this device will be used to read cooled silicon detectors, it is designed to operate at temperatures as low as −30 °C. However, because it will be tested at room temperature and is intended for a wide range of applications, its operating range extends to 65 °C. Finally, because this TDC will be used in a detector with several tens of millions of channels, the final users required us to minimize the calibration procedures, as well as the number of calibration parameters to be determined, stored, and used to reach its operational performance.

B. Global Architecture

Most modern high-resolution fast TDCs are based on DL, which comprises a cascade of elementary digital delay elements (DEs). Analogous to the structure of flash analog-to-digital converters (ADCs), one can imagine using a 2^N DE delay chain to encode the full range of an N-bit TDC. Of course, this would be at the cost of a large silicon area, and for a large dynamic range, the accumulated jitter in the DL would degrade precision. To avoid this problem, the proposed TDC is based on the three-step structure shown in Fig. 1. The TDC is driven by a 160-MHz CLK (6.25-ns period T_{ck}) signal generated by an on-chip 1.28-GHz phase-locked loop (PLL) block, locked on an external 40-MHz reference clock. This clock sequences a unique 8-bit Gray counter whose outputs are broadcast to all the TDC channels integrated within the chip. When an event occurs in a channel, the counter output is captured in a local 8-bit register. Only the last 2 bits are used for the nominal use case (25-ns coding range), and the six most significant bits can be used to extend the range up to 1.6 μs or to provide the redundancy required to merge

the data from the TDC with a low-resolution timestamping performed elsewhere. To refine the measurement, a coarse TDC (CTDC) based on 32-step DL (CDL) provides five more bits. The small length of the CDL limits the jitter accumulation along the delay chain, and its 195-ps step is easily achieved in the technology node we use. To extract the less significant bits, the time residue between the hit occurrence and the next step of the CDL is multiplied by 8 (optionally by 16) by a TA before being coded by a fine TDC (FTDC) by 3 (optionally by 4) bits using another DL (FDL). Then, a digital block decodes and combines the data from the Gray counter, the CTDC, and the FTDC to form the TDC data coded over up to 17 bits (up to 1.6-μs range with 12.5-ps step). However, for the nominal time-of-arrival (ToA) mode of operation in our primary application in HGCAL, only 10 bits are kept, covering 25 ns with 25-ps steps. In the HGCAL chip (HGCROC), these 10 bits are merged with the bunch-crossing identifier, counting the 40-MHz clock.

The last operations starting from the time amplification are completely asynchronous, but the entire TDC block has been designed to guarantee that the complete conversion time never exceeds 25 ns.

In practice, a central common block integrates some reference voltage generators, the PLL, the Gray-counter timebase, and the reference delay-locked loops (DLLs) used to stabilize the in-channel CDLs and FDLs. This block generates and broadcasts signals to "dependent blocks" localized in every channel. Each dependent block integrates a register to capture the Gray-counter state, CDL, TA, FDL, and associated coding logic. An integrated asynchronous clock generator (ACG) allows to optionally trigger channels, selectable by slow control, with a tunable frequency periodic signal asynchronous to the TDC clock. This block will be used for in situ testing, characterization, or calibration of TDCs using statistical methods, without the need for extra components. This block is composed of a free-running ring oscillator controlled by an internal DAC voltage reference. Great care has been taken in the design of the power supply for this clock to prevent coupling with other clocks, ensuring that its asynchronous behavior remains unimpeded.

To limit the single-event effects of highly ionizing particles, the most critical digital parts of the design, such as the configuration registers, Gray counter, the PLL, clock tree, and state machines, are triplicated.

III. Detailed Description of the Main Building Blocks

A. Counter, Event Synchronization, and Associated Output Registers

The most significant bits of the TDC are produced by a Gray counter, running at 160 MHz, which is common to all TDC channels. Its 8-bit output is broadcast to the channels. In the Gray architecture, only a single-bit transition occurs when the counter is incremented. This minimizes the power consumption and possible perturbations toward the analog front end due to the buffers broadcasting the counter output.

When an event is detected in the front-end part of a channel, the output of its discriminator (EVENT) rises to "1." A latched version of this signal (HIT) is provided by a delay flip-flop (DFF).

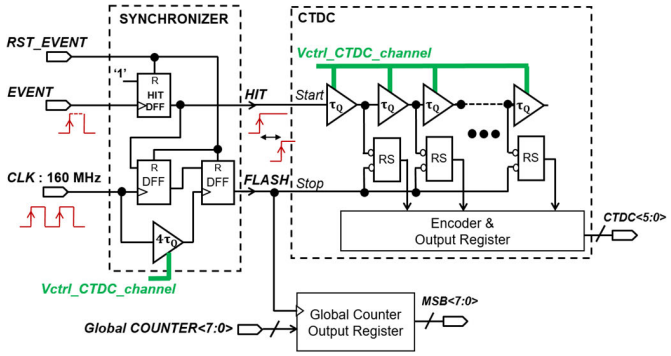

Fig. 2. Block diagram of the first two stages of the TDC and the event synchronizer.

Then, the HIT is synchronized by the 160-MHz CLK clock to produce the FLASH signal. A cascade of two DFFs, driven by CLK and CLK delayed by four elementary τ_Q delays (approximately 800 ps), reduces the risk of metastability.

As shown in Fig. 2, the FLASH signal is first used to capture the state of the common 8-bit Gray counter and stored in the output registers. This stored value is converted into binary data before being merged with other data. Only the two LSBs of this counter are used for the ToA value. Following the Nutt method [15], [16], [17], this FLASH signal is also sent as a stop signal to the second stage of the TDC, which will measure the time difference between its arrival and the arrival of the asynchronous HIT, used as a start. Thus, the time range for coding with the second stage of the TDC is between $4{\cdot}\tau_Q$ and $T_{\text{ck}} + 4 \cdot \tau_Q$. To secure the operation, the HIT DFF is released only at the end of the conversion by the RST_EVENT signal. This means that only one hit can be processed during a 25-ns period.

B. Second Stage: DL CTDC and Reference DLL

1) Introduction: One of the most popular digital methods of time-to-digital conversion for low-power multichannel applications is based on the use of DL. This type of design makes it easy to achieve a time accuracy of a few hundred picoseconds with a simple and compact structure. In addition, it can be almost completely inactive when not triggered, allowing for very low-power operation.

As shown in Fig. 2, our DL-CTDC consists of a chain of N identical DE with a propagation time τ_Q tunable by a V_{ctrl} voltage. The first delay is fed by a start signal. A memory element, which is an RS flip-flop with priority reset, is associated with each delay tap to capture the DL state when the stop signal arrives. The encoded information from the RS flip-flops is a measure of the delay between the start and the stop signals. The range of the DL is defined as the product of τ_Q by N, the number of cascaded delay cells, and its resolution is set by the minimum achievable delay and thus depends on the IC technology. The effective measurement precision is affected by the stability (with temperature, radiation, and power supply), reproducibility (of process parameters), and uniformity of the delays. Moreover, it can be shown that each

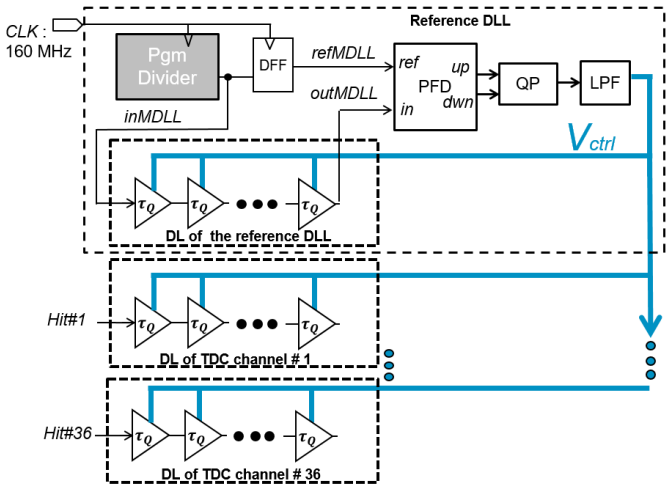

Fig. 3. Block diagram of the reference DLL that provides the DL control voltage to the TDC "dependent" channels to overcome PVT variations.

DL tap accumulates random jitter proportionally to $\sqrt{N}$ (or N in the case of a coherent noise source). Therefore, we chose to limit the number of taps. To cover the range corresponding to the $T_{\rm ck}$ duration (6.25 ns), 32 taps of 195.3 ps are used, and eight extra taps are added to provide some range margin that allows compensation for parasitic delays that will be measured as offsets. This also results in some redundancy, allowing the merging of data from the CTDC with those of the counter without any ambiguity. Consequently, the CTDC produces 6 bits, 5 of which are required.

2) DL Loop and Reference DLL: The DEs used in the DL are usually sensitive to PVT variations and irradiation. A popular solution to improve the stability of DL is to lock the total delay of a delay chain to a time reference using a DLL structure [18]. For this purpose, as shown in Fig. 3, a reference signal (refMDLL) is continuously compared with the output of the DL (outMDLL) obtained from its input (InMDLL), delayed by a reference value $T_{\rm ref}$. This comparison is achieved by a phase-frequency detector (PFD) controlling a charge pump (QP), the output of which is low-pass filtered to generate the control voltage ($V_{\rm ctrl}$), tuning the duration of the individual delays. In our design, the InMDLL signal is provided by a programmable frequency divider sequenced by the 160-MHz CLK clock. The same signal is delayed by a DFF clocked by the same clock signal to generate refMDLL such that $T_{\rm ref}$ is equal to $T_{\rm ck}$.

Because the total delay of the DLL is locked to $T_{\rm ck}$, the average delay of each delay step is $\tau_Q = T_{\rm ck}/N$.

In practical implementation, it is difficult to use this system for every channel of a multichannel TDC without introducing any dead time due to dedicated calibration phases requiring to send signal in the DLL.

However, as shown in Fig. 3, such a servo-controlled structure can be used for a reference channel only producing the control signal $V_{\rm ctrl}$ for all the "dependent" DLs in channels, which are used for measurements. The reproducibility of the measurements then relies on the matching of the dependent DL delays with those of the reference DL such that the two types of delay lines must be strictly identical to have the same temporal response for PVT variations. The "Pgm divider" block is used to program the servo control update frequency between 1 and 40 MHz. This is not a critical parameter, and for all the measurements presented in this study, we used a 10-MHz rate.

Practically, we chose to use a reference DLL for a group of 36 dependent CDLs, such that in our prototype, two reference DLLs control 72 channels.

For a DL (reference or dependent channel), the theoretical total delay is defined as follows:

$$T_N = N \cdot \tau_Q = N \cdot K_{\rm VDL} \cdot V_{\rm ctrl} \tag{1}$$

where $K_{\rm VDL}$ is the delay-transfer function of the DL element in $\mathrm{s{\cdot}V^{-1}}$.

Usually, the voltage-to-delay transfer function of a DE is nonlinear. Nevertheless, for simplicity, in the following sections, we consider $K_{\rm VDL}$ to be constant in a reduced range of control voltage.

3) Dependent DLs and Fine Delay Trimming: Ideally, in our design, T_N should be equal to $T_{\rm ck}$. However, DLs suffer from mismatches that cause a delay spread for a fixed control voltage [19]. For this reason, each step i along the CDL of channel c exhibits a positive or negative delay error $\varepsilon_{c,i}$ [19], [20], and for the DLL, the total accumulated delay, including the errors, $T_{N\varepsilon}$, is given by the following equation:

$$T_{N\varepsilon} = N \cdot \tau_Q + \sum_{i=1}^{N} \varepsilon_{c,i}. \tag{2}$$

For the reference DLL, the closed loop forces the total DL delay to be locked to $T_N = T_{\rm ck}$, regardless of any PVT variation, such that $\sum_{i=1}^{N} \varepsilon_i = 0$. This is not ensured for dependent DLs that are not servo-controlled such that $T_{N\varepsilon}$ can be smaller or larger than $T_{\rm ck}$. This can lead to a gap or an overlap problem between the CTDC and counter ranges, which may drastically affect the overall differential nonlinearity (DNL) of the TDC to the point of generating missing codes.

If we consider that all the TDCs of a multichannel chip (here 72 channels in two groups of 36 with the same reference DLL) should have a DNL better than ± 1 LSB $= \pm 24.4$ ps, and if we want to limit the yield rejection of chips to 2% because of this criterion, this implies

$$\left[\mathrm{erf}\left(\frac{\tau_Q}{\sqrt{2}\cdot\sigma_D}\right)\right]^{72} < 0.98, \text{requiring} \sigma_D < 6.7\text{psRMS} \tag{3}$$

where σ_D quantifies the spread over the channels of the total duration $T_{N\varepsilon}$ of the DL due to mismatch.

A Monte Carlo simulation of DL with reasonably sized transistors, compatible with a compact design, leads to an estimation of σ_D =28 ps rms. This is four times larger than the requirement and even larger than the value of an LSB. Simple design optimization cannot solve this issue; therefore, we chose to adjust the total delay of each channel. From (1) and (2), we can calculate

$$T_{N\varepsilon} = N \cdot \tau_Q + \sum_{i=1}^{N} \varepsilon_{n,i} = N \cdot K_{\rm VDLs} \cdot (1 + \varepsilon_{\rm KVDL}) \cdot V_{\rm ctrl} \tag{4}$$

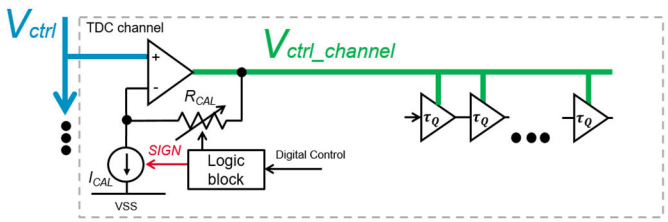

Fig. 4. Block diagram of per-channel DL gain correction.

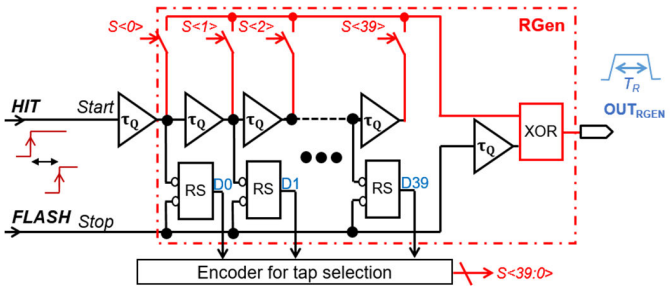

Fig. 5. Principle of the RGen.

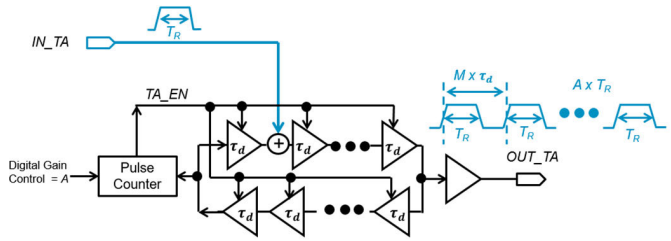

Fig. 6. Block diagram of the pulse-train time amplifier based on M delay looped cells and a pulse counter.

where K_{VDLs} is the stabilized value of K_{VDL} due to the reference DLL and $\varepsilon_{\mathrm{KVDL}}$ is the mean relative deviation from the foreseen transfer function (or gain) of the DE of the dependent DL.

To cancel the error term, for each channel, a small voltage shift V_s is added to the V_{ctrl} voltage in the form of

$$V_s = -\frac{\varepsilon_{\mathrm{KVDL}}}{1+\varepsilon_{\mathrm{KVDL}}} \cdot V_{\mathrm{ctrl}}. \tag{5}$$

This adjustment ensures that the total DL delay returns to the expected value

$$T_N = N \cdot K_{\mathrm{VDLs}} \cdot (1+\varepsilon_{\mathrm{KVDL}}) \cdot (V_{\mathrm{ctrl}} + V_s). \tag{6}$$

Fig. 4 shows the practical implementation of this correction applied to each channel. It uses an operational transconductance amplifier (OTA) with an adjustable resistor (R_{CAL}), in the feedback loop through which the current I_{CAL} flows. The "Logic block" allows programming the value of R_{CAL} and the direction of the I_{CAL} current. The corrected control voltage is given by

$$V_{\mathrm{ctrl\_channel}} = V_{\mathrm{ctrl}} + I_{\mathrm{CAL}} \cdot R_{\mathrm{CAL}} + V_{\mathrm{OFF}} \tag{7}$$

where V_{OFF} is the offset error of the OTA (less than 3-mV rms). It should be noted that this offset value will also be compensated for by the calibration system. Optionally, R_{CAL} and I_{CAL} can be set to zero so that the OTA only copies V_{ctrl}. The design of this block was optimized to reduce its sensitivity to PVT using references connected to the internal bandgap reference of the chip.

C. Third Stage: TA and FTDC

1) Introduction: This third stage improves the resolution provided by the CTDC (195 ps) by a factor of 8 (optionally 16) by finely measuring the timing of the hit signal inside a CDL step.

For this purpose, a residue generator block (RGen) provides a "residue pulse" with a T_R duration related to the time interval between the HIT signal and the next tap of the CDL reached after the HIT occurrence. This principle is shown in Fig. 5, which completes the schematic of the CTDC introduced in Fig. 2.

During the CTDC conversion, when the stop signal (synchronized version of HIT) occurs, the CDL state is memorized by the RS flip-flops. The latter's output, indicating the final tap of the DL, is promptly encoded as *S<39:0>* by the encoder for tap selection block. A multiplexer driven by *S<39:0>* selects the signal from the third tap following the last activated RS flip-flop. In parallel, following a τ_Q delay, required for encoding/multiplexing operations after the stop occurrence, the output of the RGen $\mathrm{OUT_{RGEN}}$ is set to "1" using a XOR gate. It is reset by the multiplexer output. Therefore, T_R is defined as the delay between the HIT and the next pseudo-clock edge of the CDL (τ_Q period), increased by an extra τ_Q step. Then, the duration of the "residue pulse" is amplified by a TA whose output is sent to the FTDC.

2) Concept and Design of the TA: The TA used in our design is an improvement of the pulse-train amplifier introduced in [13], which generates a train of A copies from an input pulse (IN_TA) with a T_R duration.

The design shown in Fig. 6 consists of a ring of M noninverting delay (τ_d) elements associated with a counter. Between the arrivals of events, the inputs of the DEs are forced to "0." When an event occurs, the output of the first delay is set to "1" during the input pulse duration. As long as the TA_EN signal is high, the generated pulse propagates through the ring. As the total propagation of the loop, $T_{\mathrm{loop}} = M \cdot \tau_d$, is longer than the pulse duration by design, the loop generates a train of pulses, similar to the input pulse, spaced by T_{loop}. A counter counts the number of generated pulses, and when it reaches the programmed value A, the TA_EN signal inhibits the circulation process in the TA. The total duration of the state "1" in the output train of pulses generated at the TA output, named "pulse-train duration" hereafter, corresponds to the duration of the input pulse multiplied by A.This design consumes power only during the pulse-multiplication process, which is a clear advantage for its integration into a low-power chip.

With such an architecture, there is a risk of losing very short residue pulses during the circulation operation. This would clip the transfer function of the TA, resulting in large nonlinearities of the TDC response. For this reason, we chose to increase the residue pulse duration by one τ_Q delay during its extraction so that the duration (T_R) of the input TA pulse is comprised τ_Q and $2 \cdot \tau_Q$.

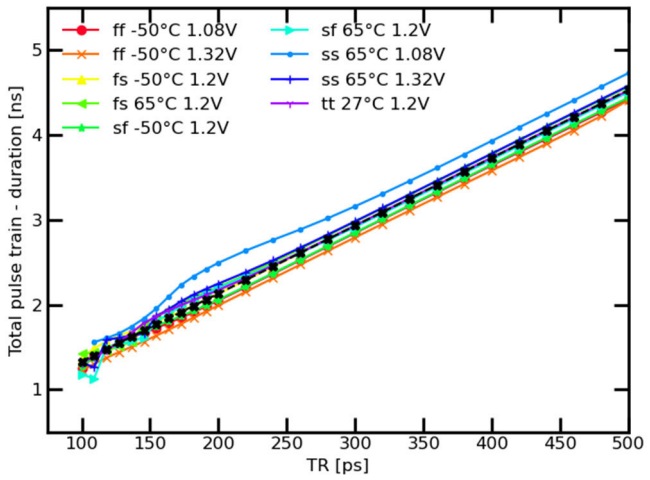

Fig. 7. Simulation of the TA transfer function for gain A = 8 and various processes (ss, tt, and ff), supply voltages, and temperatures (50 °C–65 °C).

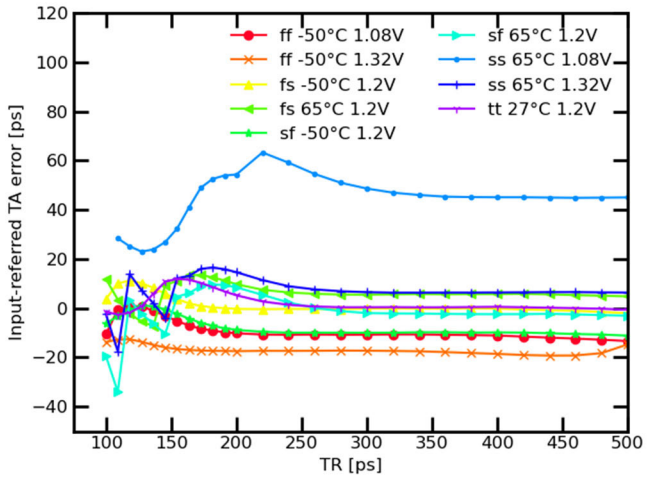

Fig. 8. Input-referred error due to the nonidealities of TA versus input pulse duration T_R for gain $A = 8$ under process (ss, tt, and ff), voltage, and temperature (−50 °C–65 °C) variations.

Notwithstanding the possible fluctuations in the pseudo-period of the pulse train, the TA gain is generally unaffected by PVT variations as long as the duration of the elementary pulses remains constant during the circulation process. For this purpose, special care was taken to ensure equal propagation and transition times for the two edges. Practically, this is easier when the elementary delays are small.

Fig. 7 shows the simulated transfer functions of the TA for various process corners (ss, tt, and ff), voltage (±10% around the nominal value), and temperature (−50 °C and 65 °C) conditions for a theoretical gain value of 8.

They exhibit a reasonable linear behavior, with an effective gain close to 8, in the useful range of T_R between 200 and 400 ps. To quantify the timing error due to TA nonlinearity or gain error, we subtracted the transfer function of an ideal TA from these transfer functions and normalized the result by the expected gain of 8. The input-referred TA errors are shown in Fig. 8. In the useful range, for a given simulation case, the error appears to be nearly constant. This constant term can be regarded as an offset that will be canceled

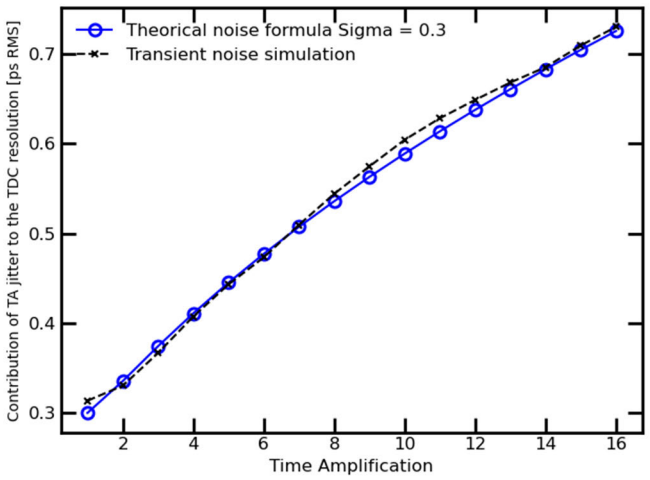

Fig. 9. Input-referred contribution of the TA jitter to the TDC timing resolution. Simulation (black crosses) and using (13) (blue circle).

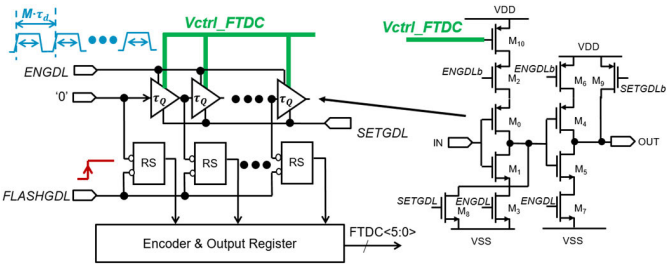

Fig. 10. (Left) architecture of the FTDC and (right) schematic of the GDEs used in FTDC.

during the calibration procedure. In the useful range, the deviation from the constant term is less than ±5 ps for all cases except the extremely slow case (ss process, 65 °C, and 1.08-V supply voltage), for which it is less than ±10 ps. Thus, for a TDC with LSB = 24.4 ps, this will result in an acceptable contribution to its integral nonlinearity (INL) of less than ±0.5 LSB (more likely ± 0.25 LSB).

The impact of the TA on the timing resolution is studied in Appendix I. This theoretical approach has been confronted with noise simulations performed using the Spectre simulator [21] to produce the plot shown in Fig. 9. With $\sigma_w = 0.3$ ps rms, the simulation results are in very good agreement with (13) taken into account a pulsewidth modulation spread of σ_w =0.3 ps rms for each circulation step. This plot shows that even for a time gain of 16, a jitter of only 0.75-ps rms is added by the TA process, when referring to the input of the TA. This is totally negligible compared to the LSB of the TDC, which is 12.2 ps in these conditions.

3) Fine DL TDC: The pulse-train duration at the TA output is measured using an FDL. It uses a gated DL (GDL) structure, made of a chain of N identical gated DE (GDE) with a propagation time τ_Q tunable by a $V_{\text{ctrl\_FTDC}}$ voltage. A pulse can propagate in the GDL only when its enable input (ENGDL) is activated, whereas the GDL state is frozen when the ENGDL signal is off. As shown in the right part of Fig. 10, this is achieved by cutting or enabling the current path in the starved and standard inverters constituting the GDEs due to nMOS

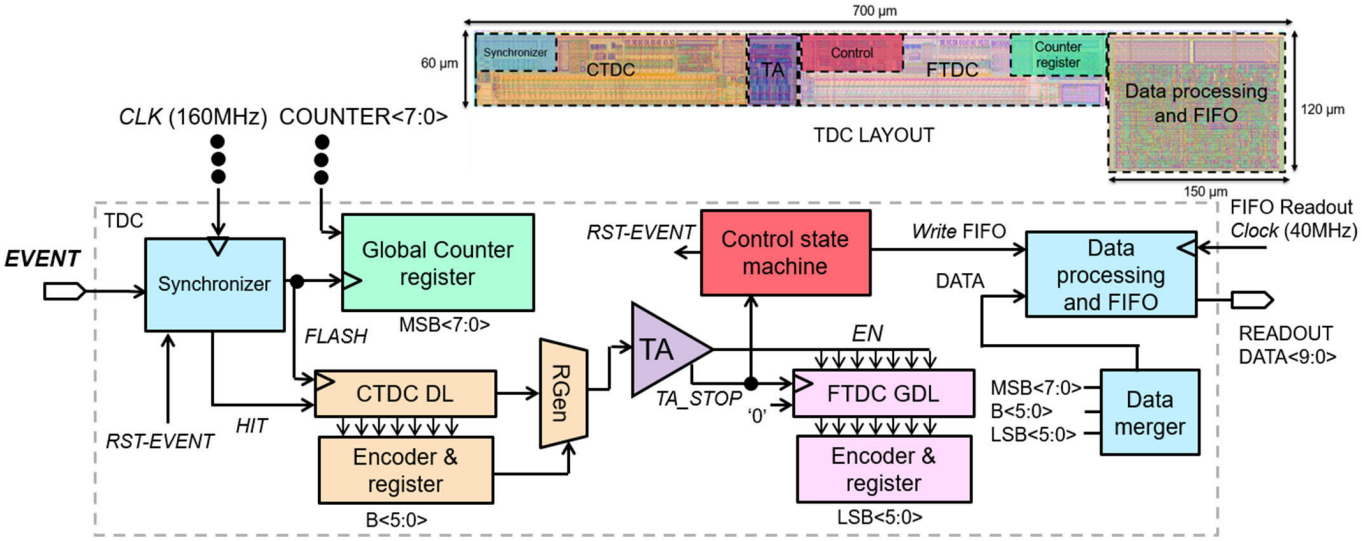

Fig. 11. Block diagram (bottom) and layout (top) of the proposed TDC channel.

and pMOS series switches (M2, M3, M6, and M7) controlled by the ENGDL input and its inverted version ENGDLb.

Between consecutive FTDC conversions, the SETGDL signal forces all outputs of the GDEs to "1." It is released at the beginning of the FTDC conversion so that the "0" level, set at the input of the first GDE, is ready to propagate in the GDL.

This propagation is enabled by the TA output signal, connected to the ENGDL input of the GDL. After the alternations of levels "1" and "0" of the pulse train, the final position of the "0"-to-"1" transition in the GDL depends linearly on the "pulse-train duration." It is memorized and encoded when the FLASHGDL signal arrives. This signal is generated by the TA block after a small delay at the end of the multiplication process.

Similar to the CTDC, a common servo-controlled reference GDL provides control voltages to stabilize the dependent GDLs of the FTDCs embedded in each channel. For simplicity, the delay of the GDL element has the same value as that of the DL of the CTDC (τ_Q =195 ps). As explained in the previous section, the GDL must cover a time range extended to $2 \cdot A \cdot \tau_Q$. For this purpose, the GDL uses 40 GDEs corresponding to the required length for a TA gain of 16, increased by eight elements to compensate for the time offsets. Therefore, the GDL-coded output is encoded on 6 bits for four ($A = 16$) or 3 ($A = 8$) bits used. Unlike the CTDC, there is no channel-by-channel adjustment of the total duration of the FTDC GDL. Indeed, the impact of an error on the duration of these delay lines will be reduced by the gain of the TA and are therefore expected to be negligible.

At the end of the conversion, the HIT DFF, shown in Fig. 2, is reset by the RST_EVENT signal, preparing the TDC for a new event.

As the asynchronous processing time of the TDC is less than 25 ns, the TDC can potentially treat more than one event per first-in, first-out (FIFO) readout clock period. In the prototype chip, we chose to keep only the first event occurring within a 25-ns clock period, but different options are possible for future applications.

D. Channel Implementation

Fig. 11 shows the block diagram and the physical implementation of a complete TDC channel: the counter register captures the MSBs on 8 bits; the CTDC, based on a DL provides six intermediate bits (5 of which are significant); and the FTDC, using the time-residue amplification method, encodes the LSBs on 5 bits (4 or 3 of which are significant, depending on the configuration). The data from each block are combined to form the final TDC code D_{out} using

$$D_{out} = D_{\text{counter}} \cdot N \cdot A - (D_{\text{CTDC}} - 1) \cdot A - D_{\text{FTDC}} \quad (8)$$

where D_{counter}, D_{CTDC}, and D_{FTDC} are the converted values of counter, CTDC, and FTDC, respectively. Optionally, these data are truncated to 10 bits at the output of this block.

Finally, a control block incorporates the state machines that sequence the conversion, store its output data in an FIFO, and reset the TDC channel for the next conversion.

The TDC conversion time can be broken down as follows.

1) The time between the HIT signal and the stop of the CTDC, comprised between $4 \cdot \tau_Q$ and $36 \cdot \tau_Q$ corresponding to the range from 0.8 to 7 ns.
2) The time required for residue extraction and multiplication, defined by $A \cdot T_{\text{loop}}$, with $T_{\text{loop}} = 750$ ps. For a nominal TA gain of eight, this operation lasts 6 ns. This time also sets the time of use for the GDL. The CTDC value is encoded during the same period.
3) Then, less than 3 ns is required to encode the FTDC value, build the final TDC code, and write it in the FIFO.

At the end, the TDC conversion time is comprised between 9.8 and 16.8 ns, smaller than 25 ns allowed before sampling in the FIFO operating at the 40-MHz HL-LHC frequency used both for writing and readout operations. This is also true for $A = 16$. In this case, the conversion time is smaller than 23 ns.

IV. Description of the Prototype and Measurement Results

To evaluate the performance of the proposed three-step TDC structure, a 72-channel TDC prototype was designed and

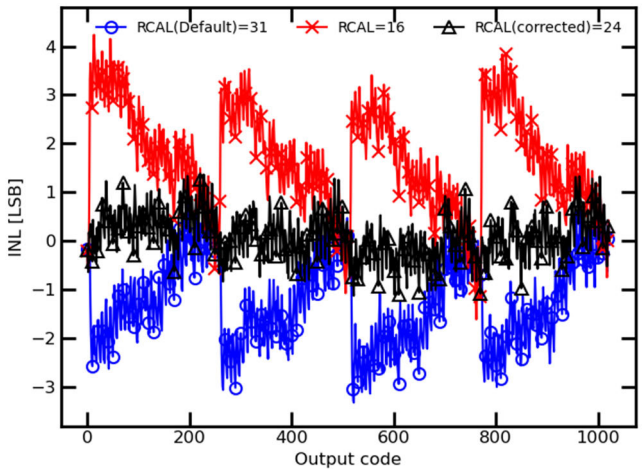

Fig. 12. INLs of a single TDC channel for $R_{CAL} = 31$ (default) (blue circle), $R_{CAL} = 16$ (red cross), and $R_{CAL} = 24$ (black triangle).

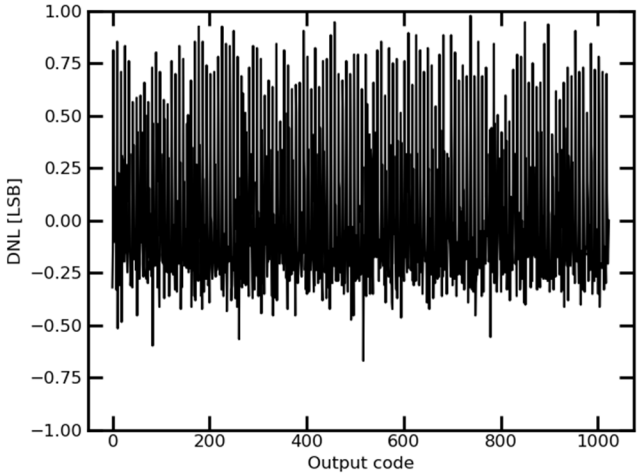

Fig. 13. DNL measured on channel 70, after internal CDL total delay trimming.

fabricated using a 1.2-V 130-nm digital CMOS process. This section describes the implementation details of the prototype and the results of the characterization measurements.

The layout of a TDC channel, occupying a 0.051-mm^2 surface, is optimized for integration within an electronic front-end channel with a rectangular small-width form factor, as shown in Fig. 11. In operation at its nominal 1.2-V supply voltage, the measured power consumption of a channel is 2.2 mW for the maximum activity corresponding to a hit every 25 ns. This power consumption is reduced to 101 μW, if there is no event and to 300 μW for 10% occupancy. Common blocks, including the reference DLLs, 8-bit Gray counter, bias references, state machines, and FIFO, consume 1.02 mW. An on-chip PLL, locked on a 40-MHz external clock, provides 160- and 40-MHz cleaned clocks used as time references for the chip. This PLL, based on a ring oscillator structure associated with a counter, will be described in a future paper. Its power consumption is 2.2 mW, and its jitter is better than 9-ps rms when measured alone with a low-jitter reference clock.

A. Measured Performance of the TDCs

Nonlinearities and time resolution are the most important parameters of TDCs. To measure the nonlinearities of the TDC channels, we used input signals asynchronous with the chip clock generated by the AGC block (shown in Fig. 1). We verified that the results were similar when external random signal sources were used. One million events are recorded to characterize a channel. The DNL is obtained using the normalized statistical density of the output digital codes provided by the TDC, and the INL is calculated by integrating the DNL. All the measurements presented in this article were performed in the nominal TDC mode, with the TA gain set to 8, resulting in an LSB of 24.4 ps.

On start-up, the control voltage V_{ctrl} provided by the reference DLLs (of the coarse and FTDCs) is directly fed to the dependent DLs located in the channels without any correction. The INL measured under these conditions on a typical TDC channel (channel 70) is plotted over the 25-ns range (circle symbols) in Fig. 12. The INL, with a peak-to-peak amplitude of four LSBs, is dominated by a pattern with nearly four identical saw teeth, each corresponding to the range of the CDL (6.25 ns). As expected, this shape reveals that the total CDL propagation time deviates slightly from its expected value. In this particular case, the CDL is actually too fast; for other channels, it could be too slow. This can be compensated channel-by-channel by using the trimming circuit introduced in Section III-B3. Its effect on the slope of the saw teeth can be seen in Fig. 12 for two values of the tuning parameter (R_{CAL}).

For the optimal tuning (triangle), referred to hereafter as hardware (HW) calibration, the INL is comprised between -1.1 and 1.4 LSB, with a standard deviation of 0.47-LSB rms (11.5-ps rms). Even for this optimal tuning, we still distinguish four zones corresponding to four 160-MHz periods with four slightly different offsets. It is possible to calibrate these four offsets and subtract them from the measurements.

The corresponding DNL, plotted in Fig. 13, is comprised between -0.7 and 0.9 LSB. As shown in the magnified plot of Fig. 14, the DNL is dominated by a modulo-8 pattern. The source of the pattern was localized in the FDL implementation, which will be improved in future versions of the chip. Even without this improvement, the measured performance, absence of missing codes, and seamless connections between the three ranges validate the TDC architecture in the first order.

B. Timing Precision Measurement

To complete the characterization, we measured the resolution of the time difference between two TDC channels (after HW calibration). This was achieved by sending an asynchronous signal, provided by an external generator, directly to one channel, whereas a delayed version of the same signal was fed to the second channel. This test allows the exploration of the entire phase space between the signal and the clock, including cases that may cause metastability issues.

In our setup, eight channels from the first TDC subgroup (channels 0–35) are simultaneously pulsed with the first signal, whereas the delayed one pulses eight channels from the second TDC subgroup (channels 36–71). A typical distribution of

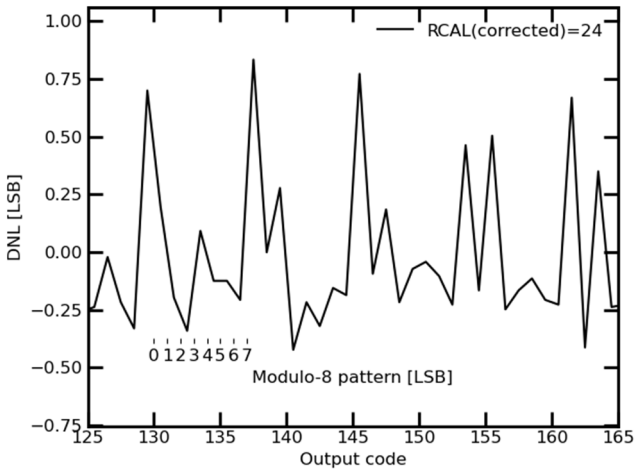

Fig. 14. Zoom on the DNL plot of Fig. 13.

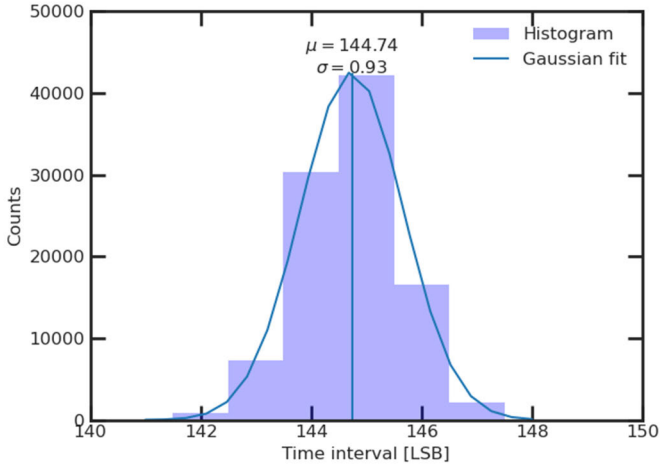

Fig. 15. Distribution of the time difference after CDL trimming (here channels 0 and 70, for a 3.5-ns delay).

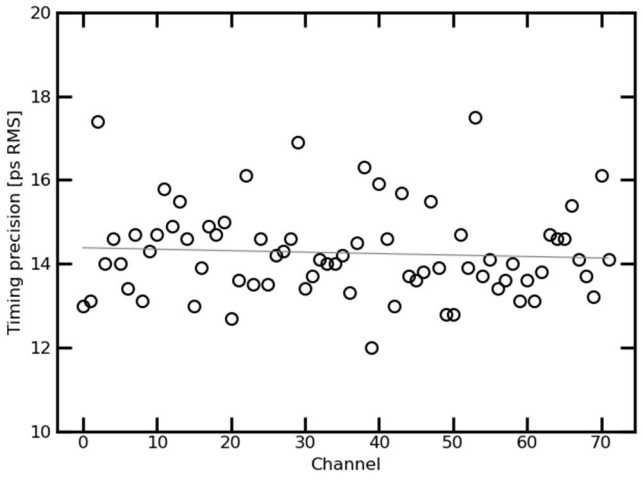

Fig. 16. Single-channel timing precision measured for all 72 channels after internal CDL calibration (3.5-ns interpulse measurement condition).

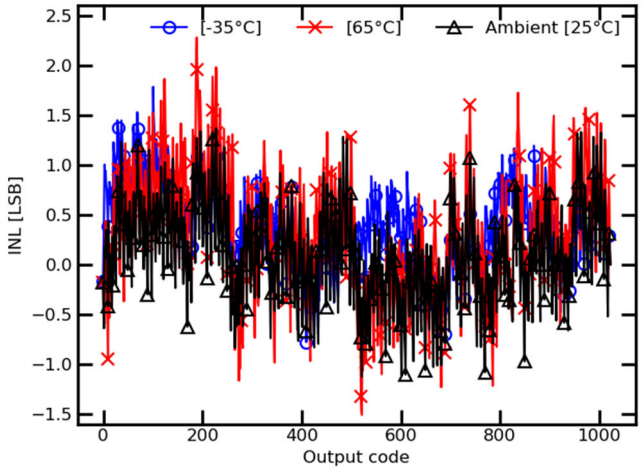

Fig. 17. INL of a TDC channel #70 after CDL delay calibration ($R_{CAL} = 24$) at 25 °C (black triangles), 65 °C (red crosses), and −35 °C (blue circles).

the measured time difference between two channels is shown in Fig. 15.

It has a Gaussian shape with a standard deviation of 0.93 LSB corresponding to 22.7-ps rms. By dividing this value by $\sqrt{2}$, we can evaluate a timing precision of 16-ps rms for a single channel. We did not notice any signs of abnormal entries in the time difference distribution, revealing a problem of metastability or range connection.

As explained in Appendix II, the timing precision of each individual channel can be extracted from simultaneous time difference measurements. These extracted values, plotted in Fig. 16 for a 3.5-ns interpulse delay, have a mean value of 14.3-ps rms with a spread of 1.1-ps rms, which is consistent with the very first estimation made previously.

C. Temperature Stability

Even if for the target applications, the TDCs will operate in an environment with a well-controlled temperature; stability with temperature remains a key point for such a circuit. For instance, we can imagine that the local temperature on the silicon can vary, depending on the chip or board activity. In this design, the variations of the key timing parameters are supposed to be controlled by the reference DLLs with respect to temperature changes. However, as the design relies on the stability of the matching between these reference elements and open-loop blocks located in the channels, it is important to verify that the TDC channel performance does not deteriorate when the temperature deviates from the value used for calibration. This was verified both for −35 °C and 65 °C temperatures by using a climatic chamber.

After HW calibration was performed at 25 °C, we performed measurements at −35 °C and +65 °C. The measured INL, shown in Fig. 17, changes only slightly and remains in the −1.5-/2.2-LSB range. The standard deviation of INL increased with temperature from 10.5-ps rms at −35 °C to 15-ps rms at 65 °C. As previously mentioned, this standard deviation was 11.5-ps rms at room temperature.

In the same conditions, we also measured the timing precision, as described in Section IV-B, for pulses delayed in steps of 3.5 ns and for the three temperatures. The results plotted in Fig. 18 reveal a noticeable performance improvement at low temperatures with a time precision of 10.1-ps rms at −35 °C.

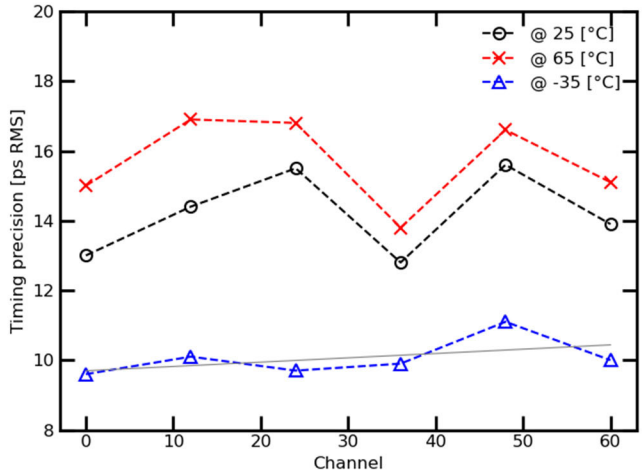

Fig. 18. Timing precision at different temperatures of six TDC channels after internal CDL trimming; 25 °C (black circle), 65 °C (red cross), and −35 °C (blue triangle) for pulses delayed by 3.5 ns.

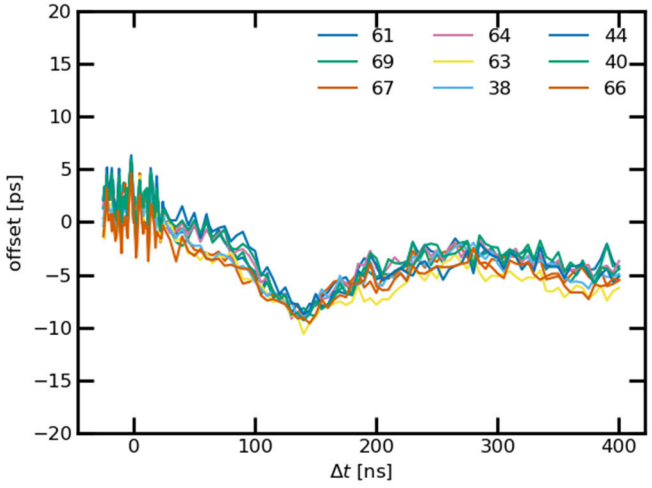

Fig. 19. Deviations between measured and expected time differences (measured in parallel with an oscilloscope) for typical channels for increasing time intervals (average of 100 kSa/channel/delay).

Unsurprisingly, the timing precision followed the evolution of INL with temperature.

This good stability of the TDC performance over a wide range of temperatures confirms the suitability of the chosen architecture for our needs.

V. Time Interval Measurements

To go further, we measured the time-difference precision as a function of the delay value between input pulses. For this purpose, we used two signals provided by a GFT1004 digital delay signal generator [22], with programmable variable delays ranging from 0 to 400 ns. The delays between these two signals were precisely measured using an oscilloscope after averaging. For each delay explored, the channels of the TDC were pulsed, as previously described, and the average time difference and timing precision were extracted with a statistic of 100 kSa/channel. The deviations between the measured and expected values are plotted as a function of delay in Fig. 19 for a representative set of channels.

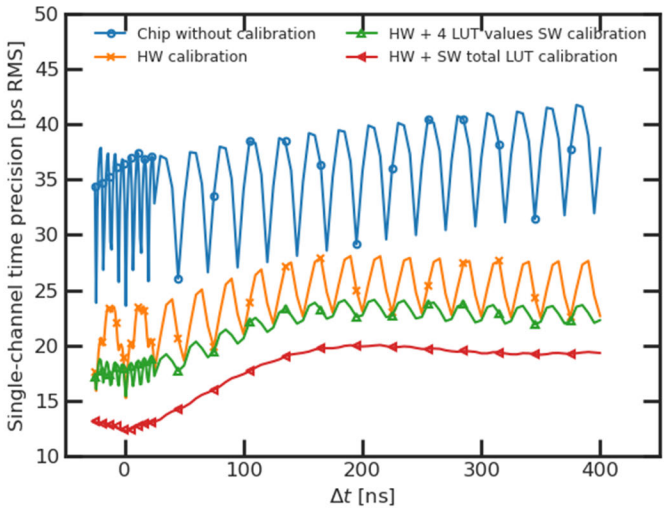

Fig. 20. Single-channel time precision for delays between −25 and 400 ns, without and with three different types of INL calibrations. The data are average values over 72 channels of a chip.

For the entire explored delay range, the measured time difference values followed the same trend, with less than one LSB deviation from the expected values. This pattern is common to all channels and is reproducible.

Using the same setup, we extracted the single-channel time precision for each delay in various configurations and plotted it in Fig. 20.

Without any calibration, we noticed a large pattern with a periodicity corresponding to the 25-ns clock period with a minimum value of 24-ps rms for zero delay and a maximum value of 40-ps rms for a long delay.

As described previously, "HW calibration" entails adjusting on-chip the total duration of the CDL of each channel. This technique leads to a significant decrease in both the floor and the amplitude of the periodic pattern of the single-channel time-precision characteristics. In this mode of operation, with no further correction, the single-channel time precision was between 17- and 27-ps rms, which corresponds to approximately one LSB rms. At the end of Section IV-A, we have provided evidence that there are still four zones of 6.25 ns (which is the CTDC range) with different offsets in the INL characteristic after HW calibration. We extracted these four offsets and corrected, off-chip, the TDC data. This operation, called "4 lookup table (LUT) Software (SW) correction," requires only a few storage and computing resources. This drastically reduces the 25-ns periodic pattern of the single-channel time-precision characteristic. Finally, it is possible to compensate for the complete INL characteristics. We call this operation "SW total LUT correction." In this case, the time precision was improved in the 12–20-ps rms range.

For standard usage, we do not intend to use this complete correction for setups with a large number of channels because it requires substantial resources for implementation. The "4 LUT SW correction" appears to be a good compromise because it requires only few computing or HW resource, which can be very fast as it requires only statistics of only a few 10 kSa/channel.

TABLE I
PERFORMANCE COMPARISON OF THE PROPOSED TDC WITH PRIOR ARTS

Parameters	TIM_20[4]	TCSI_22[5]	JINST_23[7]	TIM_22[9]	JSSC_20[10]	TNS_21[11]	This Work
CMOS [nm]	180	65	65	28	65	65	130
Architecture	Counter Middle and Fine DLLs	URFC+Vernier	2-stage interpolation (TDL + passive interpolation)	DEM Vernier	Stochastic	Untuned delay-line	2-stage interpolation (DLL + TA)
Supply voltage [V]	1.8	1.2	1.2	1	1.2	1.2	1.2
Resolution [ps]	50	27.63	3.4	8.5	0.36	5.6	22
Input range [μs]	13.1	2.5	204	0.0174	0.00005	0.0116	1.6
Conversion rate [MSa/channel]	333	0.999	320	15	100	40	40
DNL	0.47	1.679	2.75	0.3	0.77	0.5	1
INL	0.71	2.7	2.97	2.3	0.75	1	1.4
Channels	5	1	64	1	1	1	72
Power [mW]	87.6	51.4	1300 @1 MS/s	0.2	6.2	2.52	162.4
Core area [mm²]	2.25	0.432	NC	0.006	0.068	0.079	3.774
Calibration and corrections	None	None	Internal	None	Internal	External	Internal + External (option)
Temperature range [°C]	NC	NC	25–50	NC	NC	23–77	–35–65

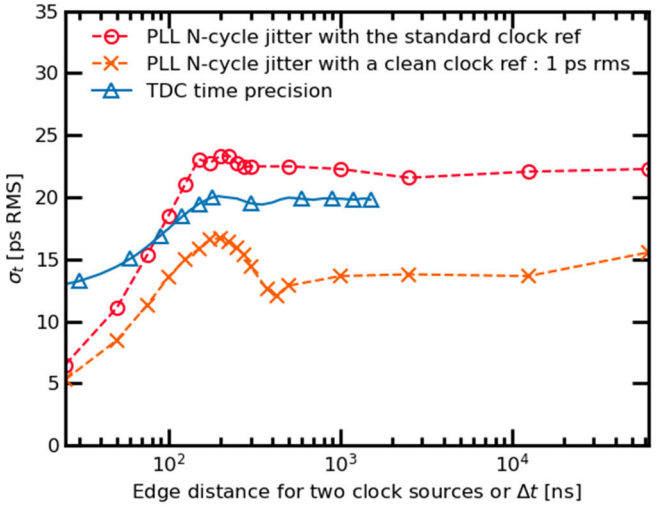

Fig. 21. N-cycle jitter of the embedded PLL as a function of edge distance measured with two clock sources. σ_t is the N-cycle jitter divided by $\sqrt{2}$ (corresponding to the contribution of one edge). The TDC time precision from Fig. 20 is overlaid for comparison.

Nevertheless, the fully corrected single-channel timing precision, which can be considered as a performance limit, is worth studying for better understanding of the design.

A dedicated pin on the test chip provides access to the clock generated by the on-chip PLL. Thus, we measured the N-cycle jitter of this clock, defined as the standard deviation of the duration of N consecutive clock periods for increasing numbers of periods (thus increasing the time difference Δt between the clock edges considered up to 60 μs). Considering that each of the two edges contributes equally to this N-cycle jitter, we can deduce the single-edge contribution (σ_t) by dividing the measured N-cycle jitter value by $\sqrt{2}$ to compare it with the single-channel time precision measured previously. This is plotted in Fig. 21 for two different configurations. The first configuration is used for the previous measurements: the reference clock of the PLL is generated by the PLL of the FPGA, which is also used for DAQ. The second configuration uses a low (1-ps rms) jitter clock as the PLL reference. Note that this configuration was foreseen for PLL-debugging only and, as such, does not allow for the acquisition of TDC data. The behavior of the N-cycle jitter is very similar to that of the time precision plotted for reference in the same figure, showing an increase for Δt up to 200 ns before reaching a plateau. N-cycle clock jitter appears to be a good candidate for explaining the moderate degradation of timing precision for large delays. We observed similar behavior for the low-jitter reference clock, even though the obtained N-cycle jitter is lower. This indicates that the PLL does not clean the input clock efficiently in the low-frequency range and probably adds itself some low-frequency jitter.

The values of σ_t (5-ps rms) obtained for a small N are dominated by noise added by our imperfect test setup of the integrated PLL.

Considering these results, we expect an improvement in timing precision for large-delay measurements when operating in the final setup with a clean clock. We also conclude that the enhancement of the PLL is a key point in order to significantly improve the timing precision of a future chip based on the same architecture.

VI. CONCLUSION

A low-power multichannel TDC based on a three-step structure using DLLs associated with the time-residue amplification technique was implemented and successfully tested. In its nominal mode, it can perform a time conversion in less than 25 ns for a 10-bit range with an LSB of 24.4 ps. It operates with a 1.2-V supply and consumes between 101 μW and 2.2 mW per channel, depending on the hit rate. Special care was taken in the design to ensure stability with variations in PVT. This includes the use of servo-control by a reference DLL and the possibility of trimming each channel to compensate for mismatches. It should be noted that the on-chip programmable ACG allows autonomous on-detector calibration of the chip. Once trimmed using a very simple procedure, all TDC channels exhibit excellent INL (between −1.1 and 1.4 LSB) and DNL (better than 1 LSB). The time precision for asynchronous signals is better than 25-ps rms

without any correction and can be reduced to 22 ps after a very simple offline correction. As described, the performance can be further improved using more advanced calibrations and a better clock reference.

Meanwhile, our circuit gives off quite respectable performance compared to prior works (Table I), often using more advanced technologies. In this table, we consider the six additional bits of the counter gray allowing to cover 1.6 μs and the power consumption of the common blocks shared or controlling the 72 channels of a chip. Moreover, the performance is stable when the temperature is varied. A time precision as good as 10-ps rms was measured at a temperature of -35 °C. This work validates our TDC architecture, which is currently being integrated, with minor improvements, into three different ASICs.

Appendix A
Pulse-Train Width Modulation Due to Recirculation in the Pulse Replicator

While the pulses recirculate in the TA during the multiplication process, the jitter of the elementary digital delays contributes to the degradation of the information about the width of the original pulse [23], [24]. The problem is not as simple as the jitter accumulation during edge propagation but is related to the dispersion of the width of the pulses of the train. To evaluate this effect, we must consider the jitter of both the rising and falling edges. For this, we hypothesize that each circulation i modulates the pulsewidth a standard deviation σ_{wi} and that these modulations are uncorrelated between two consecutive revolutions (hypothesis of high-frequency jitter). The variance of the width of the second pulse (generated after two circulations) is given by

$$\sigma^2_{\text{w2nd pulse}} = \sigma^2_{w1} + \sigma^2_{w2} \tag{9}$$

because the pulse is modulated for a first time by σ_{w1} before to enter in the second loop where it is modulated by σ_{w2}.

The duration of the train pulse after two circulations is the sum of the duration of the two pulses generated during each circulation. Its variance can be expressed by

$$\sigma^2_{\text{train},2} = (2 \cdot \sigma_{w1})^2 + \sigma^2_{w2} \tag{10}$$

as the two terms related to the first circulation are fully correlated. Using, the same analysis, iteratively, we can calculate the variance of the total width of the output pulse train after n circulations

$$\sigma^2_{\text{train},n} = \sum_{i=1}^{n} (i \cdot \sigma_{wi})^2 = \sum_{i=1}^{n} (i \cdot \sigma_w)^2. \tag{11}$$

Considering that the magnitude of the pulsewidth modulation has a constant value, σ_w, for all the circulation steps, this gives a standard deviation of the total width of the amplified pulse of

$$\sigma_{\text{train},n} = \sigma_w \cdot \sqrt{\frac{n \cdot (n+1) \cdot (2 \cdot n + 1)}{6}}. \tag{12}$$

By dividing it by the TA gain, we obtain the input-referred standard deviation on the TDC time measurement due to the TA jitter

$$\sigma_{in,n} = \sigma_w \cdot \sqrt{\frac{(n+1) \cdot (2 \cdot n + 1)}{6 \cdot n}}. \tag{13}$$

Appendix B
Δt Resolution

The single-channel timing resolution σ_i for each channel i is extracted by measuring the standard deviation σ_{ij} of distribution of the time difference between channels i and j. Assuming that the time measurements are not correlated in-between channels, we have for $i \neq j$

$$\sigma^2_{ij} = \sigma^2_i + \sigma^2_j. \tag{14}$$

For N channels, this gives a set of $(N1) \cdot N/2$ independent equations for N unknown σ_i, which can be solved for $N \geq 3$ (and is overconstrained for $N > 3$). For convenience, we introduce the variable $\sigma_t^2 \equiv \sum_i \sigma_i^2$. We then consider a subset of $N + 1$ equations with $N + 1$ unknowns σ_i^2 for $i \in [1, N]$ and σ_t^2 formed by summing over $j \neq i$ (14). Thus, this subset of equations is given by

$$\begin{cases} \sigma_t^2 = \displaystyle\sum_j \sigma_j^2 \\ (N-2) \cdot \sigma_i^2 + \sigma_t^2 = \displaystyle\sum_{j \neq i} \sigma_{ij}^2 \quad \text{for} i \in [1, N]. \end{cases} \tag{15}$$

This system admits as a solution

$$\begin{aligned} \sigma_t^2 &= \frac{1}{2 \cdot N - 2} \cdot \sum_{i,j \neq i} \sigma_{ij}^2 \\ \sigma_i^2 &= \frac{1}{N-2} \cdot \left[\left(\sum_{j \neq i} \sigma_{ij}^2 \right) - \sigma_t^2 \right]. \end{aligned} \tag{16}$$

Acknowledgment

The authors would like to thank the CMS HGCAL—Phase II Calorimeter Endcaps (CE) community for their helpful discussions and suggestions regarding this work. They also would like to thank the Laboratoire Leprince-Ringuet (LLR) and the CERN engineers for software and firmware implementation.

References

[1] *A MIP Timing Detector in the CMS Experiment Phase 2 Upgrade*, Tech. Proposal, CMS Collaboration, Cessy, France, 2017.

[2] H. Imam, "A high granularity timing detector for the ATLAS detector phase-II upgrade," *IEEE Trans. Nucl. Sci.*, vol. 69, no. 4, pp. 677–686, Apr. 2022.

[3] P. Lecoq et al., "Roadmap toward the 10 ps time-of-flight PET challenge," *Phys. Med. Biol.*, vol. 65, no. 21, Oct. 2020, Art. no. 21RM01.

[4] A. Hejazi et al., "A low-power multichannel time-to-digital converter using all-digital nested delay-locked loops with 50-ps resolution and high throughput for LiDAR sensors," *IEEE Trans. Instrum. Meas.*, vol. 69, no. 11, pp. 9262–9271, Nov. 2020.

[5] W. Zhang, R. Ma, X. Wang, H. Zheng, and Z. Zhu, "A high linearity TDC with a united-reference fractional counter for LiDAR," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 69, no. 2, pp. 564–572, Feb. 2022.

[6] J. Szyduczyński, D. Kościelnik, and M. Miśkowicz, "Time-to-digital conversion techniques: A survey of recent developments," *Measurement*, vol. 214, Jun. 2023, Art. no. 112762.

[7] S. Altruda, J. Christiansen, M. Horstmann, L. Perktold, D. Porret, and J. Prinzie, "PicoTDC: A flexible 64 channel TDC with picosecond resolution," *J. Instrum.*, vol. 18, no. 7, Jul. 2023, Art. no. P07012.

[8] Z. Gao, C. Yang, J. Xu, and K. Nie, "A dynamic range enhanced readout technique with a two-step TDC for high speed linear CMOS image sensors," *Sensors*, vol. 15, no. 11, pp. 28224–28243, Nov. 2015.

[9] V.-N. Nguyen, X. T. Pham, and J.-W. Lee, "An 8.5 ps resolution, cyclic Vernier TDC using a stage-gated ring oscillator and DWA-based dynamic element matching in 28 nm CMOS," *IEEE Trans. Instrum. Meas.*, vol. 71, pp. 1–12, 2022.

[10] H. Chung, M. Hyun, and J. Kim, "A 360-fs-time-resolution 7-bit stochastic time-to-digital converter with linearity calibration using dual time offset arbiters in 65-nm CMOS," *IEEE J. Solid-State Circuits*, vol. 56, no. 3, pp. 940–949, Mar. 2021.

[11] W. Zhang et al., "A low-power time-to-digital converter for the CMS endcap timing layer (ETL) upgrade," *IEEE Trans. Nucl. Sci.*, vol. 68, no. 8, pp. 1984–1992, Aug. 2021.

[12] P. Lu, Y. Wu, and P. Andreani, "A 2.2-ps two-dimensional gated-Vernier time-to-digital converter with digital calibration," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 63, no. 11, pp. 1019–1023, Nov. 2016.

[13] K. Kim, Y.-H. Kim, W. Yu, and S. Cho, "A 7 bit, 3.75 ps resolution two-step time-to-digital converter in 65 nm CMOS using pulse-train time amplifier," *IEEE J. Solid-State Circuits*, vol. 48, no. 4, pp. 1009–1017, Apr. 2013.

[14] D. Contardo and A. Ball, "The phase-2 upgrade of the CMS endcap calorimeter," CMS Collaboration, Cessy, France, Tech. Rep. CERN-LHCC-2017-023, CMS-TDR-019, 2017.

[15] P. Keränen, "High precision time-to-digital converters for applications requiring a wide measurement range," Ph.D. dissertation, Fac. Inf. Technol. Elect. Eng., Elect. Eng., Univ. Oulu, Oulu, Finland, 2016.

[16] F. Bouyjou, O. Gevin, O. Limousin, and E. Delagnes, "A 32-channel 13-b ADC for space applications," *IEEE Trans. Nucl. Sci.*, vol. 64, no. 4, pp. 1071–1079, Apr. 2017.

[17] E. Delagnes, D. Breton, F. Lugiez, and R. Rahmanifard, "A low power multi-channel single ramp ADC with up to 3.2 GHz virtual clock," *IEEE Trans. Nucl. Sci.*, vol. 54, no. 5, pp. 1735–1742, Oct. 2007.

[18] 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.

[19] B. I. Abdulrazzaq, I. A. Halin, S. Kawahito, R. M. Sidek, S. Shafie, and N. A. M. Yunus, "A review on high-resolution CMOS delay lines: Towards sub-picosecond jitter performance," *SpringerPlus*, vol. 5, Apr. 2016, Art. no. 434.

[20] S. Henzler, "Theory of TDC operation," in *Time-to-Digital Converters* (Springer Series in Advanced Microelectronics), vol. 29. Dordrecht, The Netherlands: Springer, 2010.

[21] *Virtuoso Spectre Circuit Simulator Reference*, Cadence, San Jose, CA, USA, 2010.

[22] Greenfield Technology (GFTY). *GFT1004, Digital Signal Generator, Specifications*. [Online]. Available: https://www.greenfieldtechnology.com/wp-content/uploads/2021/11/GFT1004-data-sheet_august-2022.pdf

[23] G. Li, Y. M. Tousi, A. Hassibi, and E. Afshari, "Delay-line-based analog-to-digital converters," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 56, no. 6, pp. 464–468, Jun. 2009.

[24] M. J. Figueiredo and R. L. Aguiar, "Noise and jitter in CMOS digitally controlled delay lines," in *Proc. 13th IEEE Int. Conf. Electron., Circuits Syst.*, Dec. 2006, pp. 1356–1359.