

Towards ultra-fast Time-Correlated Single Photon Counting: a compact module to surpass the pile-up limit

S. Farina,^{1, a)} G. Acconcia,¹ I. Labanca,¹ M. Ghioni,¹ and I. Rech¹
*Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano,
Piazza Leonardo da Vinci 32, 20133 Milano, Italy*

(Dated: 17 January 2022)

Time-Correlated Single-Photon Counting (TCSPC) is an excellent technique used in a great variety of scientific experiments to acquire exceptionally fast and faint light signals. Above all, in Fluorescence Lifetime Imaging (FLIM) it is widely recognized as the gold standard to record sub-nanosecond transient phenomena with picosecond precision. Unfortunately, TCSPC has an intrinsic limitation: to avoid the so-called pile-up distortion, the experiments have been historically carried out limiting the acquisition rate below 5% of the excitation frequency. In 2017, we demonstrated that such a limitation can be overcome if the detector dead time is exactly matched to the excitation period, thus paving the way to unprecedented speedup of FLIM measurements. In this paper, we present the first single-channel system that implements the novel proposed methodology, to be used in modern TCSPC experimental setups. To achieve this goal, we designed a compact Detection Head including a custom Single-Photon Avalanche Diode (SPAD) externally driven by a fully-integrated Active Quenching Circuit (AQC), featuring a finely-tunable dead time and a short reset time. The output timing signal is extracted by a picosecond precision Pick-Up Circuit (PUC) and fed to a newly-developed Timing Module consisting of a mixed-architecture Fast Time to Amplitude Converter (F-TAC) followed by high performance ADCs. Data are transmitted in real-time to a PC at USB3.0 rate for specific and custom elaboration. Preliminary experimental results show that the new TCSPC system is suitable for implementing the proposed technique, achieving indeed high timing precision along with a count rate as high as 40 Mcps.

I. INTRODUCTION

Nowadays, the analysis of weak and fast optical pulses plays a key role in life sciences, especially in challenging experiments such as single-molecule analysis or sub-nanosecond Fluorescence Lifetime Imaging (FLIM)^{1,2}. Among all the available techniques, Time-Correlated Single Photon Counting (TCSPC) has gained a prominent role in FLIM experiments, thanks to its inherently high sensitivity and timing precision³.

Basically, a TCSCP fluorescence measurement consists of the periodic excitation of a sample with a pulsed laser and in the record of the arrival time of re-emitted photons. Currently-available acquisition systems can record at most one photon per laser period: in this scenario, if the impinging rate is higher than this, the histogram suffers from the so-called pile-up distortion. In this case, the reconstructed exponential waveform is altered, meaning that the predicted fluorescence lifetime significantly differs from the real one⁴. To avoid this issue, TCSPC experiments have been typically carried out limiting the acquisition rate below 5% of the excitation rate, thus guaranteeing that the average number of impinging photons in a period is well below unity. However, this limitation strongly affects the speed of a TCSPC experiment, especially when only a single channel is used.

In the last decade, a great effort has been devoted in literature to the study and design of multi-channel systems⁵⁻⁹: indeed, by exploiting N parallel independent acquisition chains, in principle it is possible to increase the counting capability of the system by a factor N . Nevertheless, large integrated arrays

proposed so far do not achieve a measurement speed proportional to the number of channels and they usually suffer from a trade-off between the high number of channels and the overall performance, especially in terms of precision and linearity. In these solutions, the single channel is still bounded to the 5% rate limitation, thus it is not exploited at its best. For this reason, new approaches are needed to significantly increase the acquisition speed of FLIM experiments.

In 2017, Cominelli *et al.*¹⁰ proposed a novel theoretical solution to overcome the historical pile-up limitation, thus maximizing the single-channel acquisition speed while keeping lifetime distortion almost at zero. In this case, by matching the detector dead time to the laser excitation period (12.5 ns), pile-up distortion is avoided and measurement speed can be increased by a factor up to 8 with respect to pile-up limited systems. It is worth noting that the proposed single-channel approach and multi-channel solutions are not mutually exclusive: multiple optimized single-channels can be placed in parallel to achieve a further increase in the overall system speed.

In this paper, we present the first compact single-channel system designed to implement the novel proposed methodology. The newly-conceived TCSPC system is composed by a Detection Head, that hosts a custom-technology Single-Photon Avalanche Diode (SPAD) with the relative integrated electronics, and a high-performance Timing Module including a Fast Time to Amplitude Converter (F-TAC) structure¹¹. In order to properly apply the new measurement technique, the integrated electronics features a finely-tunable dead time, as even a small deterministic mismatch between sensor dead time and laser period can easily cause a non-acceptable distortion¹⁰. Moreover, it is worth noting that the photons detected during the SPAD non-ideal reset transition constitute another source of distortion, since they are acquired before the actual end of the dead time¹⁰. To avoid this issue, such

^{a)}Corresponding author: serena.farina@polimi.it

photons are discarded by an appropriate masking circuit, thus guaranteeing a dead time perfectly matching the excitation period. The Detection Head is therefore capable of providing a high-precision timing output that is directly fed to the Timing Module at a frequency up to 40 Mcps by means of a coaxial cable. To properly manage such a fast signal, the Timing Module features a routing front-end connected to sixteen integrated Time to Amplitude Converters (TACs), that constitute the so-called F-TAC structure¹¹. High-performance ADCs convert the TAC analog output into digital data, that are processed and stored in histograms using a Field-Programmable Gate Array (FPGA). The reconstructed TCSPC histograms are exported to an external Personal Computer (PC) through a USB 3.0 connection, reaching a transfer rate up to 400 MB/s.

In this paper we report a detailed description of the elements composing the overall system, pointing out the adopted solutions and the achieved results. The paper is organized as follows: in Sec. II the innovative Detection Head is described, in Sec. III and IV the designed Timing Module is deeply analyzed, in Sec. V the obtained experimental results are presented, and lastly conclusions are drawn and future research developments are highlighted in Sec. VI.

II. DETECTION HEAD

The core of the newly-conceived TCSPC system consists of a compact Detection Head, designed with custom integrated circuits and Printed Circuit Boards (PCBs) to specifically implement the novel acquisition technique¹⁰. The overall module design has been mainly driven by two requirements. Firstly, the integrated circuits should be placed as close as possible to the detectors, in order to limit the length of the wire bonding and thus the parasitics. Secondly, the routing path of the timing signal should be kept of minimum length and far away from possible noise sources, to avoid signal degradation potentially leading to jitter.

As mentioned earlier, to properly implement the new TCSPC acquisition technique, the Detection Head should feature a detector dead time perfectly matched to the laser period, along with a high-precision timing output. In this scenario, the project calls for the use of a differential configuration composed by an active and a dummy SPAD, as will be clearly motivated later. The two custom-technology detectors are coupled to an external fully-integrated front-end¹² and hosted on a dedicated PCB. Even though the integrated electronics can also drive detectors with a higher excess bias, at this stage we decided to employ thin SPADs with an overvoltage in the order of 5-7 V.

A. Fully-integrated front-end

The structure of the differential front-end is illustrated in Fig. 1. Two different integrated circuits, i.e. a fast Active Quenching Circuit (AQC) and a fully-differential Pick-Up Circuit (PUC), both reported in¹², are employed to bias the detectors and to collect their output current, respectively.

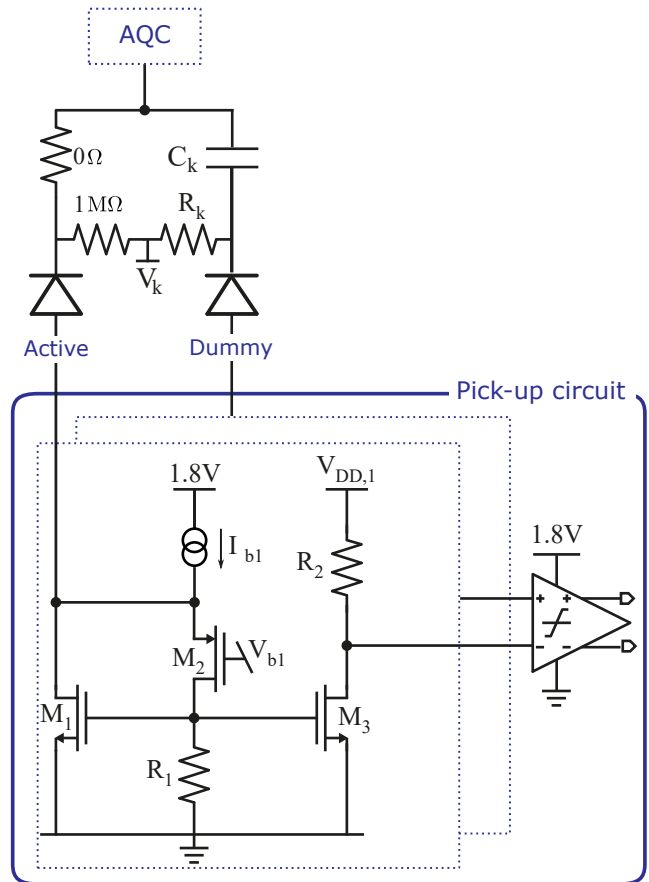


FIG. 1. Front-end structure composed by the Active Quenching Circuit and the Pick-Up Circuit. The dummy SPAD is biased through the $C_k - R_k$ network, while the active sensor features a compensation network consisting of two resistors.

The AQC is connected to the cathode terminal and it provides fast quenching pulses along with a finely-tunable dead time that was set to 12.5 ns to match the typical 80 MHz-laser period. On the other hand, the PUC is attached to the anode ensuring low-jitter timing capabilities thanks to its small input impedance.

It is worth noting that a short reset time of the detector is of utmost importance to keep the measurement distortion low and to properly implement the novel proposed technique¹⁰. However, to achieve this goal the AQC must apply a sharp voltage transient to the SPAD cathode, and this can lead to a disturbance coupled to the PUC on the anode side. When it comes to fast TCSPC measurements, photons are likely to be detected just after the reset transition and, in case of disturbances, the timing performance can be easily impaired. To avoid this issue, a differential cell has been designed: along with the active SPAD, an additional dummy SPAD biased below its breakdown voltage is used. Since the AQC drives both the dummy and the active sensor, the injected disturbances are converted into a common mode signal and eventually rejected by the output differential comparator of the PUC. To maximize the symmetry of the structure, a C-R network has been placed between the AQC and the dummy cell, thus al-

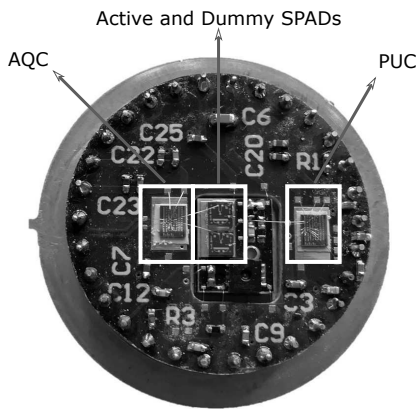


FIG. 2. Two boards compose the Detection Head: (i) a round-shaped board hosting the AQC on the left and the PUC on the right, and (ii) a rectangular sensor board housing the fully-differential SPAD structure.

lowing the exploitation of two identical SPADs, while biasing the second below the breakdown. Moreover, to compensate for the parasitics introduced by the Surface Mount Device (SMD) components constituting the C-R network, a symmetrical network with no influence on the connection impedance has been inserted also on the active branch.

B. Detection Module design

Concerning the overall module design, we decided to host the sensors and the integrated circuits on two different but adjacent PCBs, as to ensure a selective and fast cooling of the SPADs with beneficial effects on dark count rate¹³. In Fig. 2, the conceived structure is illustrated: a rectangular board of size 6 mm x 7.5 mm is placed in the center of a circular board with a diameter of 25 mm. Both boards are included in a custom 30-pin TO package with a glass window cap, to allow the sealing of the sensors in a dry nitrogen atmosphere thus reducing the risk of moisture.

The round-shaped board mounts the Active Quenching Circuit and the Pick-Up Circuit, while the rectangular black board hosts the differential SPAD structure as well as the C-R networks. Since the sensor board is intentionally made of aluminum, it provides a good thermal contact with the Thermo-Electric Cooler (TEC) underneath, thus ensuring fast thermal transients.

The desired cooling temperature is generally intended to be around -20 °C for best performance, but our system allows a much wider operating-temperature range thanks to a new control scheme for the TEC. Starting from the approach presented in¹⁴, we leveraged the TEC current as the feedback control variable, rather than its voltage as suggested in¹⁴: this refinement allows us to employ only a single temperature sensor placed on the cold side of the TEC, i.e. on the sensor board. On the contrary, in¹⁴ a further temperature sensor was needed on the TEC hot side, with a consequent increase in the complexity of the module mounting process.

Regarding the power supply generation, a dedicated PCB is needed in consideration of the high number of bias voltages required by the integrated circuits and the limited amount of space available on the previously described boards. This so-called power PCB is based on a Point-Of-Load (POL) architecture, as to attain a high efficiency and to avoid noise coupling between different loads. Indeed, for some highly-sensitive circuits, such as the PUC, a power supply ripple can easily be converted into additional timing jitter. The designed power solution exploits both switching-mode converters (SMPS) and low-dropout linear regulators (LDO): the former are very efficient, whereas the latter can provide high-precision voltages.

Moreover, most of the delivered voltages can be adjusted from a PC interface using a serial communication bus (UART). In this way, our system is suitable to host SPADs with different breakdown voltages, and to regulate the available AQC thresholds to get a fine dead time matching.

Finally, the designed power board also routes the LVDS (Low-Voltage Differential Signaling) timing output from the PUC to an external NIM connector, that is then attached to the Timing Module by means of a coaxial cable. It is worth noting that the circuit employed in the conversion between the two voltage standards also serves to reject the photons impinging during the detector reset phase. A deeper insight in the so-called photon-masking circuit is provided in the next section (Sec. II C).

C. Photon masking circuit

As deeply demonstrated in¹⁰, all the photons recorded during the sensor reset phase represent a significant source of distortion, since they are acquired before the actual end of the dead time. To avoid this issue, such photons should be first identified and then discarded from the corresponding histogram. In this scenario, many different methodologies could be adopted to manage the photon rejection, at different levels in the acquisition chain: from hardware or FPGA firmware modifications, to pure post-processing software solutions. As a first step, we opted for a pure hardware approach, i.e. we designed a circuit capable of guaranteeing a minimum time distance of 12.5 ns between two subsequent photon acquisitions.

In Fig. 3, a conceptual electrical scheme of the so-called masking circuit is reported. The circuit is based on a dual-channel comparator (ADCMP563 from Analog Devices), followed by a high-performance flip-flop (MC100EP51 from Micrel). The first comparator receives the PUC timing output, and it sets the considered flip-flop to the high state when a photon is detected. Meanwhile, the second comparator generates an output pulse corresponding to the falling edge of the AQC output signal, that is related to detector reset period. The issued pulse is then used as an asynchronous reset to inhibit the flip-flop operation. Consequently, as long as the flip-flop is kept in the reset state, no other photon can be acquired. More precisely, the length of the reset pulse can be adjusted to perfectly match the end of the dead time by acting

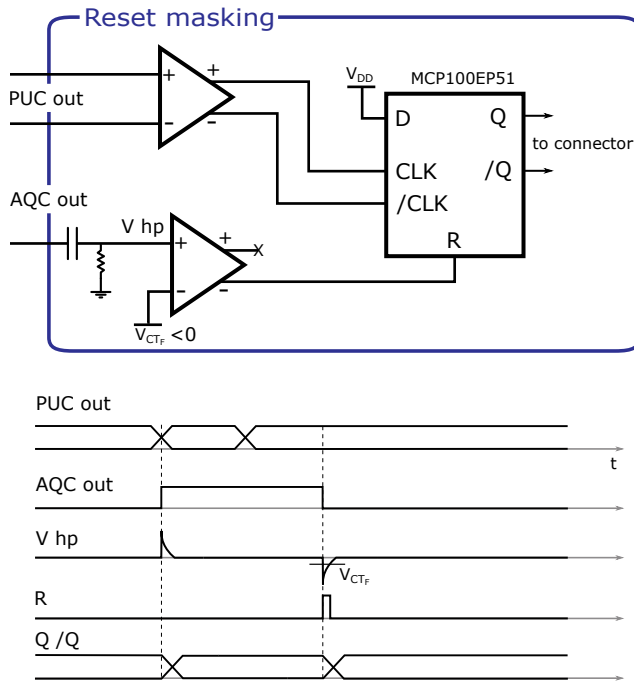


FIG. 3. Conceptual schematic of the photon masking circuit and its signal operation. The flip-flop is set in presence of a photon and kept in reset state during the detector reset interval.

on the V_{CTF} negative threshold. To this aim, a proper threshold calibration has been performed, as will be illustrated in Sec. V.

III. TIMING MODULE HARDWARE

When the novel proposed technique is practically applied, the Detection Head provides a fast timing output with an average count rate at least equal to 32 Mcps, depending on the actual duration of the sensor reset time as described in¹⁰. To properly acquire such a fast signal, we decided to develop a TCSPC module operating at a frequency as high as the laser excitation rate (typically 80 MHz). Moreover, the new module features both a single-channel and a multi-channel mode, as to provide a high flexibility to the end user. The designed TCSPC module is therefore conceived to operate either as a 16-channel module with input frequencies up to 5 MHz, or as a fast single-channel system supporting an input frequency up to 80 MHz.

The implemented structure is composed by a conversion board and a transmission board, as shown in Fig. 4. The dual-board solution was chosen to separate the time to digital conversion section of the TCSPC acquisition chain from the digital processing blocks, thus allowing an easier design reuse for future systems. Both the conversion and the transmission board feature a multi-layer stack-up, to provide several ground planes and signal layers, in order to avoid crosstalk between the analog conditioning stages and the high-rate data paths. Moreover, three differential signaling standards have

been employed (CML, LVPECL and LVDS) in the front-end routing of the timing signals, to best preserve the sharp timing edges, thus avoiding the introduction of significant jitter. Timing precision is indeed one of the most important requirements to meet for a high-quality FLIM measurement.

Depending on the selected acquisition mode, the external START signals are routed to four 4-channel integrated TACs in different ways, using sixteen commercial multiplexers (SY89543L from Micrel).

In the fast single-channel mode, the core of the conversion board is constituted by the so-called F-TAC. Indeed, a single TAC is not suitable for reaching a frequency as high as 80 Mcps and a parallel structure is consequently needed. A preliminary 8-channel F-TAC was described in¹¹ and the performance of the employed 4-channel integrated TAC are reported in¹⁵. Based on those results, we developed and tested a novel 16-channel F-TAC structure, implemented on the conversion board by combining four 4-channel integrated TACs and off-the-shelf discrete components for the routing part.

In this case, the fast START signal is first replicated sixteen times exploiting low-jitter buffers (SY58606U and SY58031U from Micrel) and it is subsequently fed to the F-TAC routing front-end, that is then connected to the four 4-channel TACs. The designed front-end is intended to sequentially route the START signal towards the sixteen available TACs, as will be more precisely explained in Sec. III A.

Conversely, in the multi-channel mode, the sixteen START signals are individually connected to the inputs of integrated TACs, by properly selecting the corresponding multiplexer outputs.

The TAC analog outputs directly drive four commercial 4-channel ADCs (AD9253 from Analog Devices), featuring 14 bits and a maximum conversion rate of 80 Msps. The chosen converter has four serial LVDS (Low-Voltage Differential Signaling) DDR (Double Data Rate) lines, that are fed to the FPGA. Along with data lines, one more LVDS signal is connected to the FPGA: the 350 MHz Data Clock Output (DCO) for source-synchronous readout of bits with SerDes (Serializer/Deserializer) interface. The ADC dual-lane configuration ensures that the data eye of the LVDS bus is large enough to preserve signal integrity, even in presence of long PCB traces used to reach the FPGA.

Concerning the FPGA, a Kintex-7 (XCK7160 from Xilinx) in an FBG676 package was chosen since it provides High Performance banks to acquire ADC signals, High Range banks to support the FT601 communication protocol and a 1 MB block-RAM, large enough to store sixteen TCSPC histograms. The FPGA is situated on the transmission board, that is connected to the conversion board using two high-speed backplane connectors (ERM8-50 and ERF8-50 from Samtec), as to preserve signal integrity.

Furthermore, with regard to FPGA data transmission, a gigabit data-link is required to properly manage the high amount of data potentially generated by a TCSPC experiment. Hence, the transmission board hosts a USB3.0 transceiver from FTDI Chip (FT601): a 32-bit parallel bus operated up to 400 MB/s connects the FPGA to the FIFO bridge chip, which is attached to the actual USB hub.

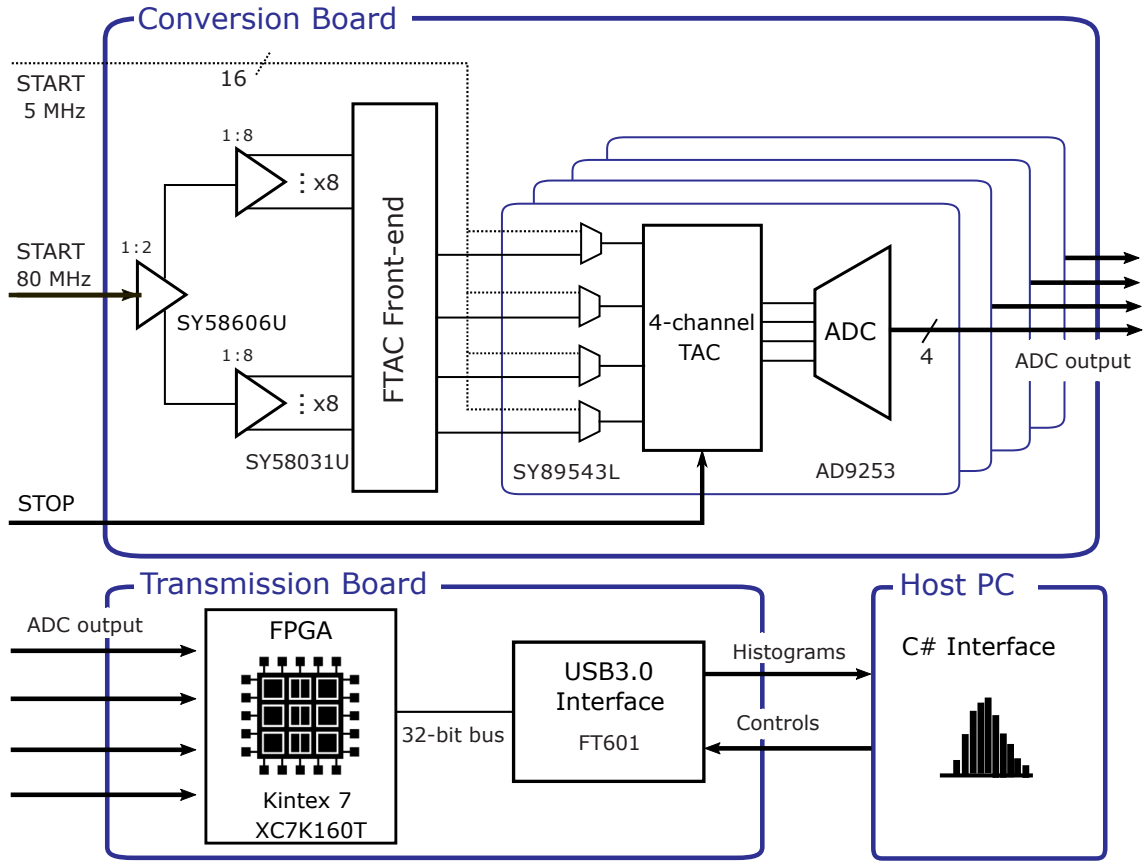


FIG. 4. Hardware structure of the Timing Module. The Conversion board receives up to sixteen inputs at 5 MHz or a single high-speed input at 80 MHz, thanks to the F-TAC routing structure. The subsequent TAC and ADC acquisition chain converts the time interval into a digital word. Finally, the Transmission Board collects the histograms to be sent to the host PC.

Finally, at the system power-on, two microcontrollers from Microchip (PIC18F67J11 for the conversion board and PIC16F15344 for the transmission board) manage the FPGA programming from a quad-SPI FLASH memory, initiate the ADC calibration procedure and set the operating mode of the module.

A. Fast-TAC front-end

In Fig. 5, a block scheme of the F-TAC front-end is illustrated. As previously mentioned, each START pulse needs to be sequentially routed to a free Time to Amplitude converter, to achieve an acquisition frequency at least equal to the sensor detection rate. Moreover, the architecture is conceived to support multi-hit events: even though the maximum average photon rate is strictly determined by the parallel TAC structure, the minimum time distance between two consecutive START signals is limited only by the front-end. Consequently, the system is capable of recording also fast adjacent photons, with a minimum distance in the order of 100 ps.

The designed circuit is constituted by a circular shift register, adopting sixteen Current-Mode Logic (CML) flip-flops (SY55852U from Micrel). Each flip-flop output directly con-

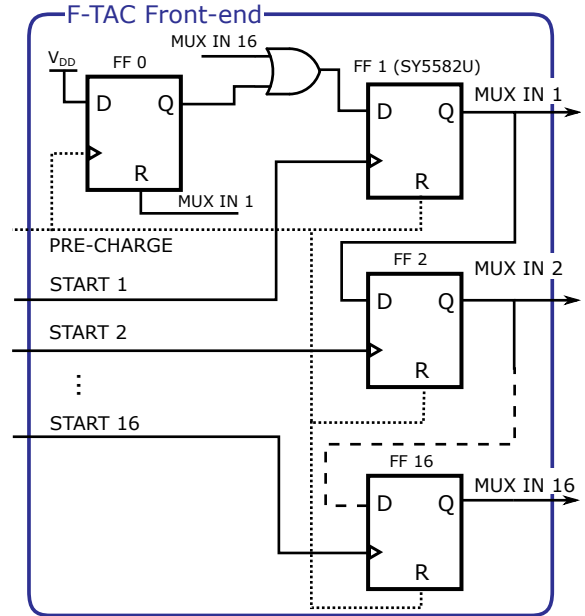


FIG. 5. Schematic of the F-TAC front-end, composed by sixteen flip-flops organized as a single circular shift register.

nects to the subsequent flip-flop input, thus implementing an overall chain. The circular register is pre-charged with a single high value ('1'), that is propagated through the chain on the clock rising edge. For this reason, each replicated START signal has been routed to the corresponding clock input of the flip-flops: when a new external START appears, the shift register is activated and a new TAC conversion is initiated. In addition, as the chosen flip-flops do not feature a set input, another flip-flop and an OR logic gate are inserted to properly initialize the shift register. The conversion-board microcontroller handles the start-up of the front-end chain when the F-TAC operating mode is selected.

Finally, it is worth noting that an excessive time skew between the replicated START signals can cause the circular chain failure: indeed, the pre-charged single high value could duplicate or disappear, thus causing a circuit malfunctioning. To avoid this issue, particular care has been taken during PCB layout to minimize and equalize flip-flop interconnection paths. Experimental results confirmed the proper operation of the circular shift register, with frequencies up to 80 MHz as required.

IV. TIMING MODULE FIRMWARE

Typically, in a TCSPC system two approaches can be used to deal with data management: on-board channel histograms construction and time-tagging operating mode. The former implies that the histograms are directly built within the on-board FPGA and downloaded to the host PC, while in the second case each arrival time and address of the fired SPAD is acquired and immediately sent to the host PC. The Timing Module hardware is suitable for both approaches, but, for the sake of simplicity, we chose to implement the histogram mode at this stage.

In Fig. 6, the block scheme of the Timing Module firmware is illustrated. Even though a first implementation of a similar structure was reported in¹⁶, substantial improvements are introduced in this work. First of all, a newer FPGA from Xilinx 7-series was chosen, as to leverage the powerful Vivado Design Suite for a faster design and debug. Even more, the available logic and I/O resources are increased and a lower power consumption is expected, that is of utmost importance when supporting a higher number of TCSPC channels. Finally, particular effort was devoted to develop a more accurate algorithm to dynamically calibrate the ADC data eye and to implement a USB3.0 communication channel with the host PC.

The firmware is generally intended to accomplish three major tasks: (i) properly acquire the ADC output data, (ii) manage TACs operation and (iii) transfer histograms to the host PC.

Since the ADC operates in free-running mode, the FPGA uses a synchronization signal (STROBE) to acquire the correct ADC channel output. More precisely, each TAC issues an asynchronous STROBE signal just after a valid conversion, indicating that the conversion value is ready to be sampled by the subsequent block. It is worth mentioning that

the faster sampling frequency of the newly-adopted ADC ensures a smaller uncertainty on the actual sampling instant after STROBE assertion with respect to¹⁶. Therefore, the timing jitter introduced due to TAC output drift during hold operation is definitely reduced.

The FPGA not only performs data sampling and processing, but it also sends a reset signal to the TACs and it manages the update and subtraction of the dithering contribution. The exploitation of the dithering technique relies on a 10-bit DAC integrated along with the TAC on the same chip, thus allowing us to minimize the impact of the commercial-ADC differential non linearity. A detailed analysis of the implemented methodology is described elsewhere¹⁶⁻¹⁸.

The ADC output, after dithering compensation, corresponds to the histogram memory cell whose value has to be updated. Dual-port Block-RAMs store the sixteen reconstructed histograms, that are individually constituted by 16384 bins, each one having a 4-byte depth. Measurement data are written into port-A and they can be simultaneously read and sent to the PC using port-B. All the collected histograms, along with the STROBE and the STOP frequencies, are transferred to the PC through the FT601 USB3.0 transceiver.

Finally, during firmware design particular attention was paid to issues potentially deriving from clock domain crossing¹⁹. Indeed, the general FPGA management, the ADC data acquisition and the data transmission with FT601 are all based on different input clocks, thus requiring a careful exchange of data among the three clock domains. For this reason the block-RAM and FIFO built-in resources of the FPGA are extensively exploited.

A. ADC data acquisition and calibration

As previously described, the ADC outputs are acquired and deserialized inside the Kintex 7 FPGA, using a Xilinx Intellectual Property (IP) that manages the SerDes instantiation and usage.

In Fig. 7, the structure of the implemented digital architecture for a single channel is illustrated. First of all, the input DCO is fed to an internal dedicated Phased-Locked Loop (PLL), to generate both a DCO replica and an aligned Frame Clock Output (FCO) signal at 80 MHz. The FCO is used to identify each ADC conversion inside the serial stream of bits. Even though the ADC itself is providing a FCO clock, we decided to regenerate this signal inside the FPGA, as to achieve better alignment and lower jitter between DCO and FCO.

Since the chosen transmission protocol is based on two lanes carrying respectively the MSB and LSB bits, two SerDes are employed in parallel, each one providing an 8-bit output. The 8-bit results are then concatenated and two meaningless bits are discarded, thus obtaining the desired 14-bit conversion value.

For a proper data acquisition and alignment, two main operations are required: (i) data eye dynamical calibration and (ii) bitslip. Both of them rely on a fixed output pattern, that is issued by the ADC when prompted at system power-up.

Dynamical eye calibration is needed since crosstalk and

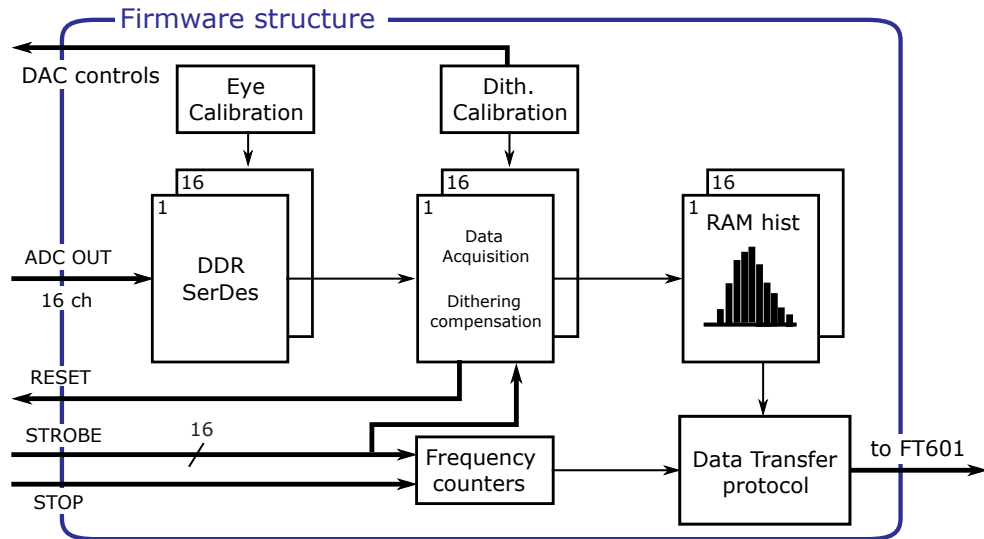


FIG. 6. Schematic representation of the firmware developed for the Timing Module. The firmware is intended to perform three main tasks: (i) ADC calibration and data management, (ii) dithering calibration and compensation, and (iii) data transfer to the FT601 chip.

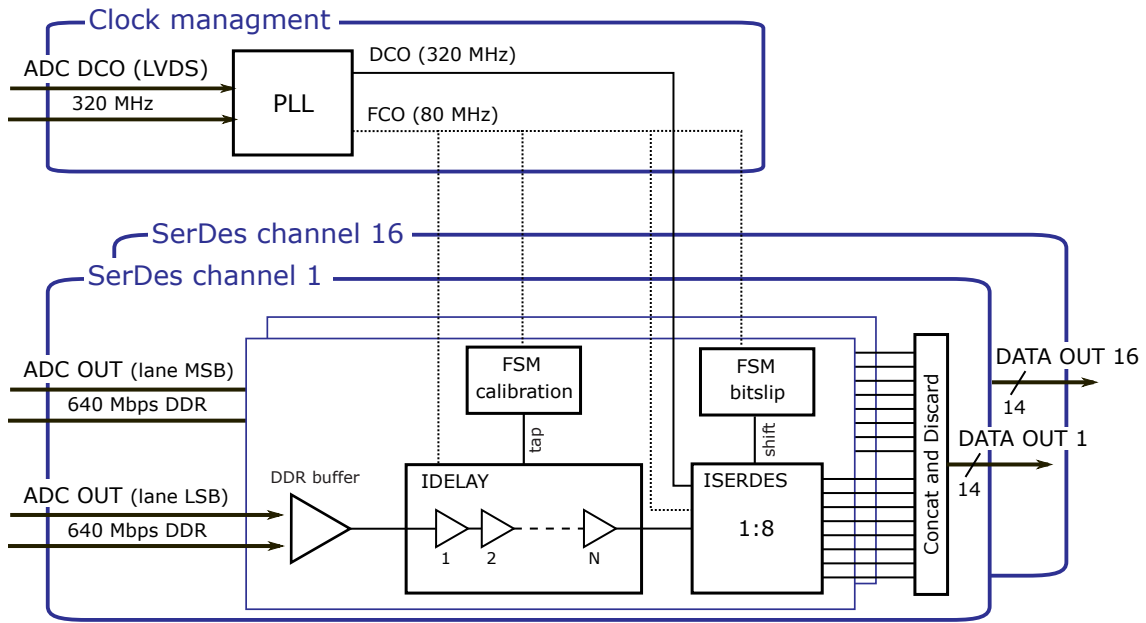


FIG. 7. Block scheme of the ADC data acquisition and calibration. The input clock is managed through a PLL. Instead, the input data is properly delayed and sampled by SerDes.

electrical disturbances, acting on the PCB traces, can easily induce a misalignment between the input DCO and the data path. As a consequence, the FPGA may sample the data during a transition, thus resulting into an incorrect value.

To properly realign input signals, data coming from the ADC are delayed within the FPGA, exploiting an embedded primitive structure, i.e. the IDELAY²⁰. The employed variable delayer is constituted by 32 taps, each one of 78 ps duration, that can be progressively applied on the input data. In order to choose the proper time shift, we developed an algorithm that sequentially scans the available taps and it memorizes, for

each step, whether the input data is stable or not. Based on this information, the data eye transitions are extracted and the proper tap corresponding to the center of the eye is selected.

It is worth noting that the IDELAY primitive is affected by an amount of data-dependent jitter equal to ± 5 ps. As a consequence, for high tap values applied on small data eyes the sampling time uncertainty is remarkable. To avoid this issue, the DCO clock can be shifted with 60° steps thanks to a proper ADC functionality, thus achieving lower tap values and more robust data sampling.

Once the calibration of the data eye has been performed,

the SerDes output should perfectly match the ADC fixed pattern. Nevertheless, since the FCO is not directly issued from the ADC, one of the eight possible bit-shifted versions of the pattern can appear at the output. The bitslip functionality is natively integrated into Xilinx SerDes and it is intended to iteratively shift the output bits, until the proper pattern version is recognized.

B. Digital data management

Even though the first version of the Timing Module is mainly intended for on-board histogram reconstruction and download, we decided to design a flexible hardware and firmware structure, as to allow future time-tagging implementations. Since the bandwidth of a simple USB2.0 connection is not sufficient for time-tagging operations, a faster link has been employed, i.e. the USB3.0.

The selected chip (FT601 from FTDI Chip) represents a FIFO to USB3.0 bridge featuring a source-synchronous interface when communicating with the FPGA. More precisely, the operating clock is sent from the FT601 source to the FPGA along with data, and it is meant to clock both input and output data travelling on the 32-bit bus. Obviously, the proprietary communication protocol ensures bidirectional transmission with the FPGA as master of the bus, that is indeed allowed to initiate reading and writing operations.

In Fig. 8, the implemented firmware structure is represented, both from the data transmission path and the clocking point of view. The firmware is generally expected to send TCSPC data to the host PC, while reading commands coming from the user interface. Three main blocks compose the transmission path: two FIFOs to exchange data with the other sections of the firmware, a Finite State Machine (FSM) to implement the communication protocol, and an input-output interface to properly capture and send data.

In case of histogram reconstruction, data coming from the sixteen B-RAMs are downloaded into a 32-bit wide FIFO, to create a data stream matching the bus width. During the writing phase, data are routed to the output interface through a multiplexer and consequently sent to the slave chip. Instead, during the reading phase incoming data are sampled and saved into an input FIFO, that is then accessible from the so-called system manager, i.e. the managing unit of the overall firmware.

Moreover, to correctly implement the FT601 proprietary communication protocol, a FSM is required. The latter constitutes the core of the transmission firmware and it is intended to start a writing or reading phase on the basis of the available space in the FT601 internal FIFOs. Since in a TCSPC system many information need to be transferred from the board to the host PC, the coded FSM gives priority to the writing operation with respect to the reading one. Lastly, the FSM controls the multiplexer output selection, thus allowing us to transfer either the protocol signals coming from the FSM itself or the histogram values on the bus.

Regarding the input-output interface, the design was mainly driven by the tight timing specifications of the FT601 bus,

that can easily impair a fast and simple FPGA timing closure. While the input data can be easily sampled with a flip-flop interface thanks to a wide data valid window (6.5 ns), the writing phase represents the main challenge: indeed, the FPGA must guarantee a hold time as long as 4.8 ns. Considering a bus clock of 100 MHz, this implies that the new output data should be written on the bus in correspondence of the falling edge of the clock. To avoid setup timing issues between the FSM clocked on the rising edge and the output interface clocked on the falling edge, a simple FF interface is not sufficient. For this reason, the Output Double Data Rate (ODDR) primitives from Xilinx²⁰ have been employed along with the particular connection illustrated in Fig. 8. This new strategy allows us to provide the output data from the FSM on the rising edge of the clock and to change the actual output value on the bus at the subsequent falling edge.

It is finally worth mentioning that two different clocks have been employed in the design, both at 100 MHz. The FT601 CLK is issued from the integrated chip and managed through a PLL to remove clock insertion delay and obtain a perfect synchronization with the FT601 source interface. However, this clock is only active during transmission and therefore a second general clock (SYSTEM CLK) is needed to interface with the other parts of the TCSPC firmware.

V. EXPERIMENTAL RESULTS

The described modules have been manufactured and extensively tested. In particular, at this initial stage we carried out an experimental characterization of each module separately, to better assess their individual performance before exploiting them together.

A. Detection Module

The preliminary tests performed on the Detection Module were mainly focused on three aspects: reset time measurement, timing performance and masking circuit calibration.

Regarding the former, since reset photons are discarded by the masking circuit, the photon collection efficiency is reduced, with a consequent decrease in the average acquisition rate of the system. Nevertheless, Cominelli *et al.*¹⁰ demonstrated that a detector reset time shorter than few nanoseconds allows to achieve a speed improvement of almost an order of magnitude with respect to classic pile-up limited systems. Considering for example a 1 ns fluorescence decay and a reset time of 4 ns, the system can provide a remarkable gain of a factor 8 in the measurement speed. Accordingly, a faster voltage transition during the reset phase can naturally lead to even better results. Therefore, it is of utmost importance to measure the actual reset transition applied by the AQC on the detector cathode, as to verify whether the Detection Module is able to meet the desired requirements.

Due to the fragile wire bonding connection between the SPAD terminals and the AQC, a direct probing of the cathode voltage during the reset phase is unfortunately precluded

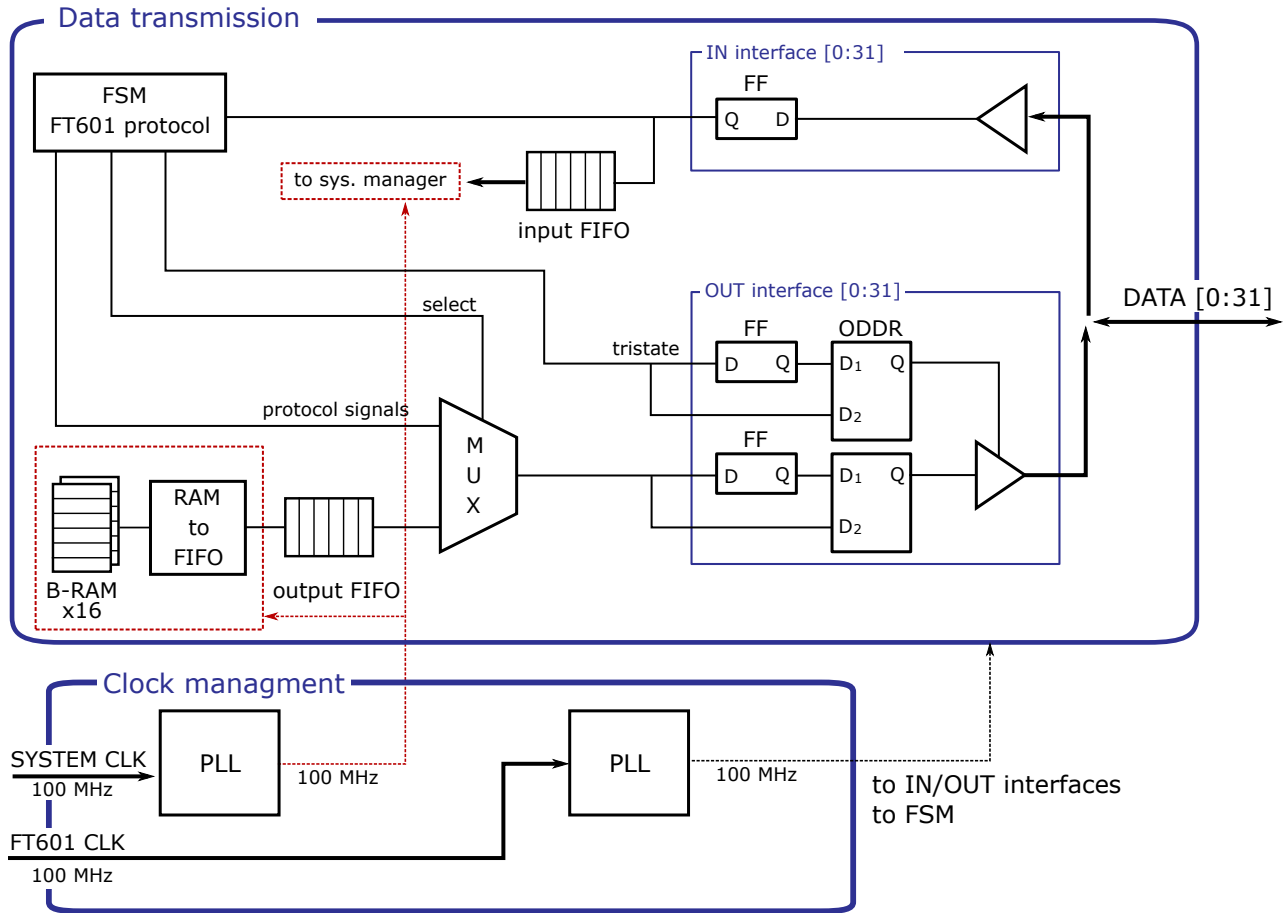


FIG. 8. Structure of the implemented firmware for data transmission through FT601. Three main elements compose the structure: a FSM to manage the operation protocol, an input/output interface to correctly acquire and send data, and two FIFOs to exchange data with the other sections of the firmware. Two PLLs are employed for clock management. The SYSTEM CLK is used to interface with the other parts of the FPGA firmware, while the FT601 CLK is directly coming from the bridge chip.

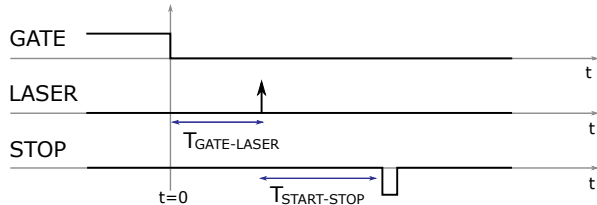


FIG. 9. Signal timeline employed for reset characterization.

and a different method must be adopted. Consequently, we resorted to the indirect procedure proposed in¹², that is based on a gated photon measurement.

In Fig. 9, the employed signal timeline is reported. When the GATE input is deasserted, the AQC induces a reset phase on the SPAD. After a predefined time interval, i.e. $T_{GATE-LASER}$, a laser is activated, emitting a light pulse that is recorded by the sensor. Finally, an external STOP signal is generated at a fixed time distance from the GATE activation. The $T_{START-STOP}$ delay is measured using a Becker&Hickl SPC-130 commercial module, operated there-

fore in reversed START-STOP mode. In particular, the photon recorded due to the laser pulse represents the START event and the $T_{START-STOP}$ interval can be modified by acting on the laser time position. The experiment was performed by progressively decreasing the $T_{GATE-LASER}$ delay, thus scanning the whole reset transition. In this way, the system is punctually stimulated by the laser and the time response to the GATE step can be retrieved by interpolating the acquired single Instrument Response Functions (IRF). The achieved result is illustrated in Fig. 10, where the optical transient is obtained by plotting the total number of acquired photons in each IRF with respect to the $T_{GATE-LASER}$ delay. The measured reset interval is as short as 1.5 ns, with a rising edge of only 750 ps, thus allowing us to correctly implement the new proposed technique.

Further measurements were performed to investigate the Detection Head response in correspondence of the reset transition. Indeed, a nanosecond reset transient at the SPAD cathode can easily couple at the anode side and consequently impair the expected timing performance due to induced voltage oscillations. To verify whether the obtained results are still acceptable, we carried out a standard timing measurement ex-

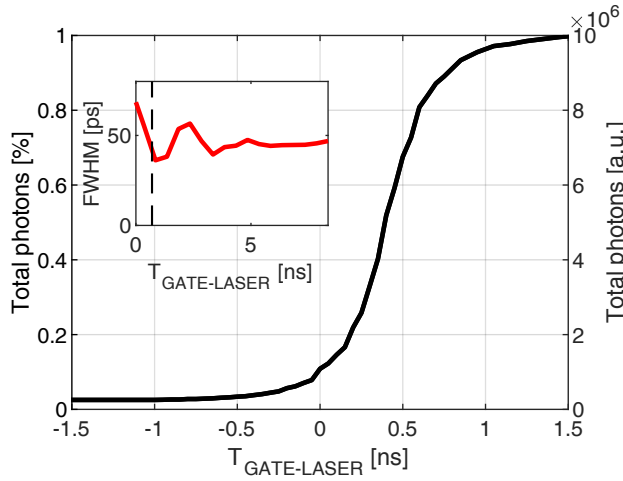


FIG. 10. Optical reset transient, defined as the total number of photons in percentage acquired at different $T_{GATE-LASER}$ delays. The initial reset point is arbitrarily chosen to be 10% of the stationary value, while the reset end point is collocated at 90% of the final stationary value. In the inset, the achieved FWHM is reported with respect to the same $T_{GATE-LASER}$ delay, where the dashed line represents the end of the reset transition.

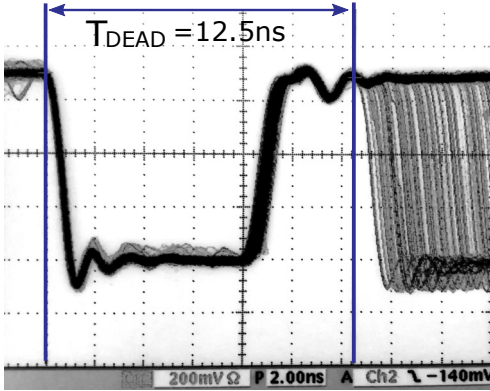


FIG. 11. NIM timing signal, observed with the oscilloscope at the output of the Detection Module. The V_{CT_F} threshold calibration is performed so that the minimum distance between two subsequent pulses is perfectly matched to the typical laser period, i.e. 12.5 ns.

exploiting the TCSPC module Becker&Hickl SPC-130. Since the Full Width at Half Maximum (FWHM) of events recorded just after the reset transition is always below 55 ps (Fig. 10), the designed module can be considered suitable for a typical FLIM experiment.

It is worth noting that all the tests concerning the reset characterization were performed using a bench power supply instead of the actual power board, as to correctly observe reset photons without introducing any masking operation.

After tests, the power board was connected and we performed a fine-tuning of the V_{CT_F} threshold for the photon masking circuit. In Fig. 11, the achieved result is reported, where an oscilloscope was used to observe the NIM output signal of the overall Detection Module. In this case, a LED

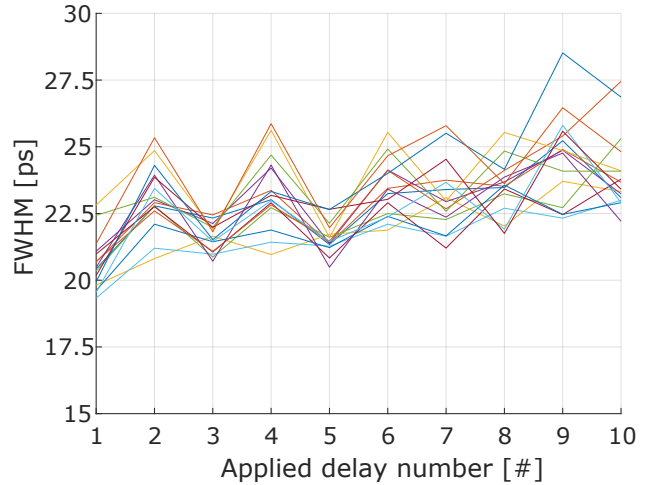


FIG. 12. Measured timing jitter for all sixteen channels in F-TAC mode with a FSR of 12.5 ns and ten different START-STOP delays. The average FWHM is equal to 20-25 ps. The timing performance worsens at higher START-STOP delays due to the internal TAC architecture, as already shown in¹⁶.

(Light Emitting Diode) was employed to generate uncorrelated photons and to quantify the minimum time distance between two consecutive photon pulses. After calibration, the dead time interval is properly matched to the laser period, with a final value of 12.5 ns. At this first stage, the threshold regulation was manually performed to better investigate the masking circuit behaviour, but future works envision the development of a real-time calibration procedure in order to comply with potential laser period drifts.

B. Timing Module

The TCSPC module was specifically designed to acquire the high-precision timing signal provided by the Detection Head. To this aim, some requirements concerning the acquisition speed and the introduced timing jitter and distortion have to be fulfilled.

Firstly, the overall F-TAC chain should be able to process incoming signals with an average acquisition rate of at least 40 Msp. From a conceptual standpoint, the maximum rate is mainly limited by three factors: the TAC settling time after STOP arrival, the time interval employed by the FPGA to sample the STROBE and issue a reset signal, and, finally, the TAC reset time. In our tests, the TAC contributions have proven to be equal to nearly 100 ns with a FSR of 12.5 ns, while the FPGA contribution is generally equal to 5-6 ADC clock cycles, i.e. 60-72.5 ns. Therefore, the overall maximum frequency is equal to almost 5 Msp for a single channel and 80 Msp for the overall sixteen-channel chain, thus complying with the system specifications.

A second experimental test was carried out to characterize the jitter performance of the module, that is of utmost importance when dealing with lifetime imaging. The START and

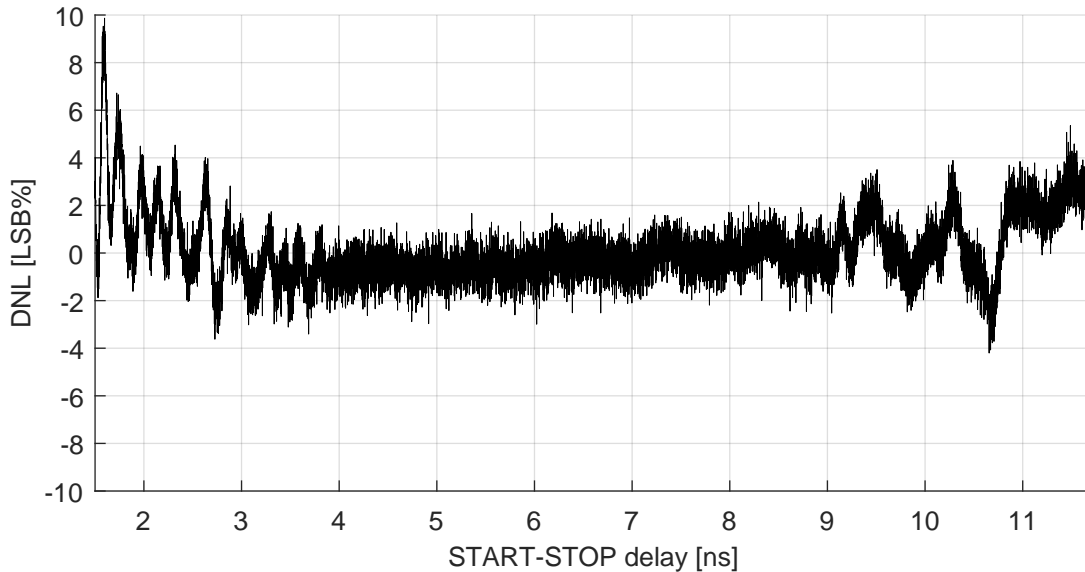


FIG. 13. Measured Differential Non-Linearity (DNL) of the Timing Module for every single channel in F-TAC mode, when connected to the Detection Head. The central portion features a DNL of $\pm 2\%$ LSB, while the oscillations are generated by START and STOP interferences. After ferrite bead insertion, the amplitude of the oscillations was significantly reduced from the previous $\pm 8 - 10\%$ LSB to the reported $\pm 3\%$ LSB.

STOP signals were obtained by splitting up the output of a pulse generator, while the STOP was also delayed by means of an adjustable passive delay line. For each channel we performed ten measurements at different START-STOP delays, to evaluate the obtained jitter in terms of FWHM along the whole Full Scale Range (FSR). The chosen FSR size was 12.5 ns, as it corresponds to the maximum measurable delay value with a 80 MHz-laser; moreover, a larger FSR would worsen the timing performance as already proven in¹⁶. In Fig. 12 the results of the characterization for every single channel in FTAC mode are reported: the module features an average FWHM in the order of 20-25 ps that makes it suitable for correctly converting the input timing signal without adding a significant jitter contribution.

Finally, we performed a Differential Non-Linearity (DNL) test to verify whether the applied dithering technique is effective in removing the ADC non linearity. In this case, two uncorrelated START and STOP pulses were fed to the Timing Module in the F-TAC mode, thus measuring a uniform delay distribution. In particular, the STOP signal was obtained from an external pulse generator, while the START signal was directly connected to the output of the Detection Head for a more realistic characterization. The achieved results are similar for all the F-TAC channels, so a single channel is reported in Fig. 13. The central portion of the graph is characterized by an excellent DNL ($\pm 2\%$ LSB peak to peak and 0.85% LSB rms) comparable with the best available TCSPC systems^{21,22}, whereas we observed high frequency oscillations for low and high delay values.

Such oscillations have been deeply investigated, as to better understand their origin and to improve the current sys-

tem performance. More precisely, it was demonstrated that the START signal activation is coupled to the first part of the DNL, while the second fluctuation is correlated to the STOP signal deassertion. The source of such non-idealities was identified both into a not perfectly matched termination for the STOP line and high frequency transmissions in the power delivery network of the Conversion Board. As a consequence, proper termination and power decoupling ferrites have been inserted in the PCB, thus reducing the second DNL fluctuation down to $\pm 3\%$ of the LSB as shown in Fig. 13. Moreover, future developments will include further refinements in the current PCB layout.

VI. CONCLUSION

In this paper we presented the first single-channel system capable of implementing a new high-speed TCSPC acquisition technique. The system is specifically designed to target FLIM measurements, that are still intrinsically slow due to the classical pile-up limitation. To this aim, two separate modules have been designed, manufactured and extensively tested, from the hardware and firmware point of views. As the obtained results are fully satisfactory, the system is ready to be tested in a real on-field fluorescence application. Finally, we envision to extend the novel technique to a multi-channel system, to further speed up TCSPC measurements while keeping high performance.

ACKNOWLEDGMENTS

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 777222.

DATA AVAILABILITY

Data available on request from the authors.

- ¹K. Suhling, L. M. Hirvonen, J. A. Levitt, P.-H. Chung, C. Tregidgo, A. Le Marois, D. A. Rusakov, K. Zheng, S. Ameer-Beg, S. Poland, *et al.*, *Medical Photonics* **27**, 3 (2015).
- ²R. Datta, T. M. Heaster, J. T. Sharick, A. A. Gillette, and M. C. Skala, *International Society for Optics and Photonics* **25**, 071203 (2020).
- ³E. Gratton, S. Breusegem, J. D. Sutin, Q. Ruan, and N. P. Barry, *International Society for Optics and Photonics* **8**, 381 (2003).
- ⁴W. Becker, *Advanced Time-Correlated Single Photon Counting Techniques* (Springer, 2005).
- ⁵C. Veerappan, J. Richardson, R. Walker, D.-U. Li, M. W. Fishburn, Y. Maruyama, D. Stoppa, F. Borghetti, M. Gersbach, R. K. Henderson, *et al.*, in *2011 IEEE International Solid-State Circuits Conference* (2011) p. 312.
- ⁶F. Villa, R. Lussana, D. Bronzi, S. Tisa, A. Tosi, F. Zappa, A. Dalla Mora, D. Contini, D. Durini, S. Weyers, and W. Brockherde, *IEEE Journal of Selected Topics in Quantum Electronics* **20**, 364 (2014).
- ⁷L. Parmesan, N. Dutton, N. J. Calder, N. Krstajic, A. J. Holmes, L. A. Grant, and R. K. Henderson, in *International Image Sensor Workshop, Vaals, Netherlands, Memory*, Vol. 900 (2015) p. M5.
- ⁸T. Al Abbas, N. A. Dutton, O. Almer, N. Finlayson, F. M. D. Rocca, and R. Henderson, *IEEE Sensors Journal* **18**, 3163 (2018).
- ⁹R. K. Henderson, N. Johnston, H. Chen, D. D.-U. Li, G. Hungerford, R. Hirsch, D. McLoskey, P. Yip, and D. J. Birch, in *ESSCIRC 2018-IEEE 44th European Solid State Circuits Conference (ESSCIRC)* (2018) p. 54.
- ¹⁰A. Cominelli, G. Acconcia, P. Peronio, M. Ghioni, and I. Rech, *Review of Scientific Instruments* **88**, 123701 (2017).
- ¹¹P. Peronio, G. Acconcia, I. Rech, and M. Ghioni, *Review of Scientific Instruments* **86**, 113101 (2015).
- ¹²G. Acconcia, A. Cominelli, M. Ghioni, and I. Rech, *Optics express* **26**, 15398–15410 (2018).
- ¹³M. Ghioni, A. Gulinatti, I. Rech, F. Zappa, and S. Cova, *IEEE Journal of selected topics in quantum electronics* **13**, 852–862 (2007).
- ¹⁴P. Peronio, I. Labanca, M. Ghioni, and I. Rech, *Review of Scientific Instruments* **88**, 116102 (2017).
- ¹⁵M. Crotti, I. Rech, and M. Ghioni, *Review of Scientific Instruments* **47**, 699 (2011).
- ¹⁶S. Antonioli, L. Miari, A. Cuccato, M. Crotti, I. Rech, and M. Ghioni, *Review of Scientific Instruments* **84**, 064705 (2013).
- ¹⁷I. De Lotto and G. E. Paglia, *IEEE transactions on instrumentation and measurement* **2**, 170 (1986).
- ¹⁸D. Resnati, I. Rech, and A. Geraci, *Review of Scientific Instruments* **79**, 064706 (2008).
- ¹⁹H. Kaeslin, *Top-Down Digital VLSI Design* (Morgan Kaufmann, 2014).
- ²⁰Xilinx, (2011), 7 Series FPGAs SelectIO Resources, UG471.
- ²¹Becker&Hickl, <https://www.becker-hickl.com/products/spc-160/>.
- ²²PicoQuant, <https://www.picoquant.com/products/category/tcspc-and-time-tagging-modules/hydraharp-400-multichannel-picosecond-event-timer-tcspc-module>.