POLITECNICO
MILANO 1863

SCUOLA DI INGEGNERIA INDUSTRIALE
E DELL'INFORMAZIONE

Executive Summary of the Thesis

Hardware Implementation of a Compact, Stand-Alone Single-Channel SiPM Detection Module for Ultra-Fast and Distortion-Free TCSPC Applications

Laurea Magistrale in Electronics Engineering - Ingegneria Elettronica

Author: Simone Basile

Advisor: Prof. Ivan Rech

Co-advisors: Gennaro Fratta, Piergiorgio Daniele

Academic year: 2024-2025

1. Introduction

In recent years, a steadily growing interest has emerged toward the detection and analysis of weak optical signals, driven by the unique properties of light as a non-invasive contrast agent for investigating cellular phenomena. Within this context, Time-Correlated Single-Photon Counting (TCSPC) is a digital measurement technique that enables the reconstruction of a time-resolved intensity profile by precisely recording the arrival times of individual photons. Its method relies on the periodic excitation of a sample, followed by the detection of photons emitted in response to the excitation. The arrival times of these photons are measured relative to the excitation pulse within each repetition period. By repeating this process over a large number of excitation cycles, a histogram of photon arrival times is constructed [1]. As shown in Figure 1, after a statistically significant number of repetitions, the histogram will ideally have the same waveform as the light signal incident on the photodetector.

Despite its widespread adoption, conventional TCSPC implementations are constrained by two

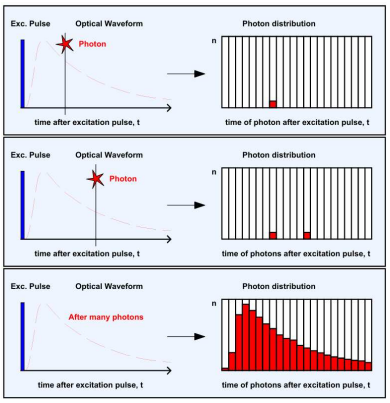

Figure 1: Photon detections over multiple periods build a histogram that reconstructs the light intensity waveform [1].

major limitations: pile-up and counting loss, which originate from detector and system dead time, and the reliance on bulky external instrumentation for signal processing. When multiple photons arrive within the same excitation cycle, the first photon triggers the system dead time, causing all subsequent photons arriving during this interval to go undetected, giving rise to the phenomenon known as counting loss. In addition, the unregistered photons introduce distor-

tions in the histogram reconstruction, commonly referred to as pile-up effects. To mitigate this effect, the average number of detected photons per excitation period must be kept well below unity (typically 0.01–0.05), significantly reducing photon efficiency, and degrading the accuracy in estimating the decay constant τ.
To overcome pile-up distortion, a novel TCSPC measurement strategy was employed [3]. The proposed approach extends the conventional acquisition scheme by also timestamping a reset signal marking the end of the dead time. This enables the construction of an auxiliary histogram encoding the time-dependent probability $\alpha(t)$ that the system is in its active state. Once the activity probability is known, the impinging photon rate $r_{\text{imp}}(t)$ can be directly retrieved from the measured data by a pointwise normalization of the recorded photon rate $r_{\text{rec}}(t)$:

$$r_{\text{imp}}(t) = \frac{r_{\text{rec}}(t)}{\alpha(t)}.$$

1.1. Objective

This work presents a compact single-channel detection module designed to overcome both limitations through the integration of a high-speed analog front-end and a custom FPGA-based processing chain. The architecture embeds a dedicated time-to-digital converter (TDC) together with the pile-up distortion correction algorithm within a fully standalone system. This integrated approach eliminates the need for external instrumentation and relaxes the low count-rate constraint typical of conventional digital TCSPC systems, enabling reliable operation even under high photon flux conditions. An example of the final assembled system is shown in Figure 2.

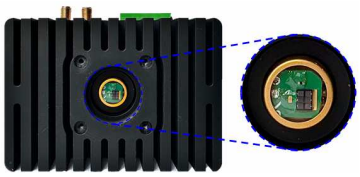

Figure 2: Stand-alone module for ultra-fast TCSPC application. On the right a close-up view of the SiPM device in use.

2. Front-End Optimization

Ultra-fast TCSPC applications require a front-end architecture capable of converting optical signals into electrical pulses with minimal dead time. Starting from a previously developed front-end architecture (Figure 3), a comprehensive optimization of its parameters was carried-out. The objectives were twofold: to minimize system dead time, defined as the temporal width at the base of the SiPM-induced pulse, and to optimize the signal waveform by increasing edge steepness, thereby reducing timing jitter in both photon and reset arrival measurements.

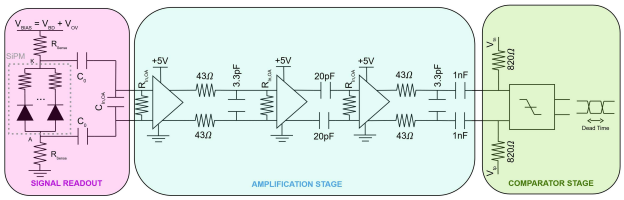

Figure 3: Front-End Block Diagram.

Following detailed simulations aimed at optimizing the differential signal readout stage, it was observed that the introduction of an RC network connected in parallel with the SiPM, between the cathode and anode terminals, was found to effectively reshape the SiPM output pulse, simultaneously narrowing its width and preserving steep edges.
Increasing R or decreasing C raises the RC branch impedance, diverting more signal current to the output and thereby increasing the signal amplitude. Higher resistance also dampens oscillations, reducing undershoot. Moreover, larger capacitance attenuates low-frequency components, lowering the average signal tail without affecting edge steepness.
The optimized values of the RC network, together with the resulting current pulse waveform, are shown in Figure 4.
Table 1 summarizes the optimal front-end parameters and the corresponding figures of merit obtained from the analysis.

3. Monostable Cicuit Design

In research contexts where minimizing dead time is critical to mitigate pile-up distortions, next-generation acquisition systems must also remain compatible with commercial electronics.

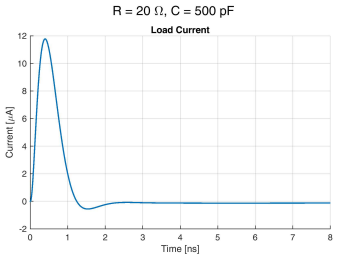

Figure 4: Optimized Current Pulse.

Table 1: Summary of the optimal parameters and the corresponding performance metrics.

Parameter	Value
Sensing Resistance [Ω]	100
Decoupling capacitance [pF]	50
RC Resistance [Ω]	20
RC Capacitance [pF]	500
FWHM [ps]	≈ 600
Signal Amplitude [uA]	12

In our case, where a ≈ 600ps signal FWHM was achieved, a monostable circuit was implemented to ensure compatibility with both the on-board FPGA and external instrumentation. The design builds upon the differential comparator employed in the Front-End described in Section 2, which can operate either in comparator or in latch mode. By exploiting this dual-mode functionality, the monostable circuit was designed to automatically enter latch mode immediately upon detection of a comparator event and to release it after a predefined time interval. Latch activation is controlled by a differential OR gate driven by the comparator outputs. One of the differential signal paths is intentionally delayed by a buffer prior to the OR gate, thereby defining the duration of the monostable output pulse. The operational behavior of the monostable circuit was thoroughly characterized on a custom PCB to validate its expected functionality. Taking into account both logic gate propagation delays and PCB trace delays, precise experimental measurements were carried out to accurately determine the resulting monostable pulse width, which can be expressed as:

$$\begin{aligned}\Delta t_{\text{pulse}} &= t_{\text{buff}} + t_{\text{OR}} + t_{\text{trace\_delay}} + t_{\text{latch}} \\ &= 210\text{ps} + 210\text{ps} + 120\text{ps} + 180\text{ps} \\ &\approx 700\text{ps}\end{aligned}$$

In addition, the introduction of an RC network between the buffer output and the OR gate input enables further fine tuning of the pulse duration by selecting an appropriate capacitance value. This ensures compatibility with both commercial time-tagging electronics and FPGA-based processing systems.

4. Front-End PCB Design

The next and crucial step consists in integrating both the optimized front-end and the monostable circuit onto a single PCB, resulting in a complete and compact photon-detection board. The front-end PCB is designed to perform the following core functions:

- Acquire the current pulses generated by the SiPM upon photon detection.
- Amplify the photon-induced analog signal and convert it into a differential CML output.
- Monostabilize the differential signal to ensure robust compatibility with a wide range of downstream processing systems.
- Transmit the differential outputs to a second PCB, exclusively dedicated to FPGA-based signal processing.
- Generate and distribute all supply voltages required for the correct operation of the analog and digital components.
- Enable system control and monitoring through an on-board microcontroller, which provides USB communication with a host PC, controls the SiPM bias voltage, and regulates the detector temperature via a Peltier cell.

In parallel with the development of this board, a detailed study was conducted on bypass capacitance selection, signal and power trace routing, ground plane and return current management, as well as the implementation of stitching and shielding vias. The outcomes of this analysis were directly applied to the PCB layout to ensure optimal high-speed performance and signal integrity.

Figure 5 presents a 3D view of the whole board. The left side of the PCB hosts the buck and linear regulators used for on-board power supply generation. The microcontroller is located in the lower central region, with the Digital-to-Analog Converter (DAC) positioned to its right, while a USB 2.0 interface is placed along the top edge. The central area accommodates the socket for the sensor head containing the SiPM, followed

by the cascade of three amplification stages. The signal processing chain terminates with the monostable circuit, located in the upper-right corner of the board.

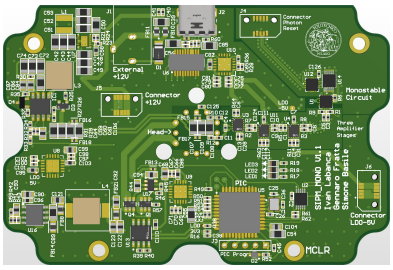

Figure 5: 3D layout view of the detection board.

5. FPGA-Based PCB Design

The primary objective of the proposed PCB is to perform all signal processing and timing operations directly on-chip. By integrating time-tagging, histogram reconstruction, and signal processing functionalities within the FPGA, the overall measurement setup is considerably simplified, removing the requirement for external instrumentation and enabling a fully integrated and portable data acquisition system. The main functions of this board are summarized as follows:

- Acquisition of the differential CML signal from the output of the monostable circuit, and its conversion into logic levels compatible with the FPGA input standards.
- Acquisition of the laser trigger signal, provided in NIM standard, and its conversion into an FPGA-compatible logic format.
- Generation of all power supply rails required for the proper operation of the FPGA and of the auxiliary components integrated on the board.
- Integration of the FPGA and all supporting components necessary for its operation, including flash memory, JTAG interface, and a quartz-based clock generator.
- Serialization of the processed data generated by the FPGA through an FTDI interface, enabling USB communication with an external computer.

The PCB interfaces with the Front-End board via three high-density board-to-board connectors located on the bottom side of the PCB. The use of three connectors was selected to improve current distribution and reinforce the mechanical robustness of the stacked assembly, obviating the need for additional structural supports. Additionally, an SMA coaxial connector is positioned on the board edge to serve as the input for the laser trigger signal in NIM standard, which provides the synchronization reference required for TCSPC measurements and the correct construction of histograms. The selection of the FPGA was primarily driven by the TDC architecture implemented in this work, which was developed and made available by the Department of Informatics, Bioengineering, Robotics, and Systems Engineering of the University of Genova [4]. Since the TDC was previously validated on an XC7K325T FPGA and the available resources were sufficient for the full processing chain, the same device was selected to ensure compatibility and minimize integration risk. A complete 3D view of the board is presented in Figures 6 and 7.

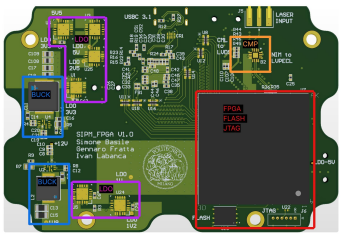

Figure 6: 3D Top view of the FPGA board.

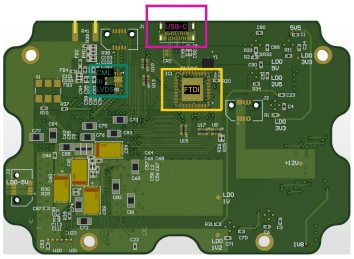

Figure 7: 3D Bottom view of the FPGA board.

6. 1GHz Digital Shift-Register

The intrinsic dead time of the TDC employed in this work is approximately 4ns, which is significantly larger than the sub-nanosecond dead time achieved by the front-end. Consequently, the TDC constitutes the dominant bottleneck of the overall system, ultimately limiting the maximum achievable throughput. Considering that

the Xilinx Kintex-7 XC7K325T FPGA can operate at a maximum clock frequency of approximately 710 MHz, corresponding to a clock period of about 1.4 ns, the design objective was to reduce the effective TDC dead time by a factor of four. To achieve this goal, a four-channel digital shift-register architecture was implemented on the FPGA. The primary challenge associated with the shift-register is the efficient and deterministic routing of each incoming photon event to a distinct TDC channel immediately upon detection, while avoiding the introduction of additional dead time. For this reason, an asynchronous architecture was chosen; however, this approach introduces the risk of metastability, which requires careful design to ensure reliable operation. To mitigate the probability of metastability, a finite state machine (FSM) based on Gray coding was adopted in place of a conventional shift-register implementation. In a Gray code sequence, only a single bit transitions between consecutive states, thereby minimizing the probability of timing violations. The implemented design is specifically developed to timestamp both the arrival of the photon signal and the corresponding reset event, which represents the end of the SiPM current. Consequently, by knowing both the photon arrival and reset times, it's possible to measure the dead time of the system, which is a prerequisite for the effective application of the pile-up correction algorithm [3]. Following design synthesis and implementation, the final stage consisted of in-depth debugging to verify its correct functionality. A measurement was carried out by simulating four consecutive photon pulses with a period of 1 ns and a 50% duty cycle. The results, shown in Figure 8, confirmed the correct operation of the shift-register even at a frequency of 1 GHz, which exceeds the maximum frequency of the BUFG (710 MHz). The proposed architecture is therefore capable of reliably detecting consecutive photon pulses with a minimum period of 1 ns, provided that each pulse exhibits a FWHM of at least 500 ps. Implementing the entire processing chain directly on a commercial FPGA platform eliminates the need for external instrumentation.

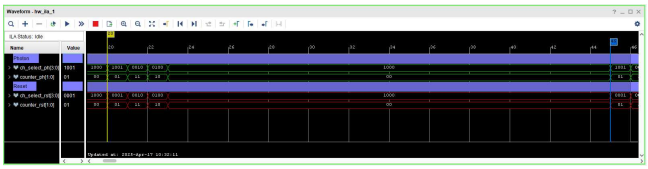

Figure 8: Between the two markers, exactly four pulses are captured. Both the photon and reset counters increment by the number of input pulses, transitioning from "00" and returning to "00". The shift register advances its output pattern accordingly.

7. 3GHz Discrete Shift-Register

To overcome the limitations of the FPGA-based implementation, represented by a maximum switching frequency of 1 GHz and photon pulses no shorter than 500 ps FWHM, a discrete-component four-channel shift register was designed and implemented in the research laboratory, on which I focused exclusively on its characterization and validation. This module is capable of operating at switching frequencies up to 3 GHz and reliably processing pulses as short as 150 ps FWHM. The overall system comprises two identical four-channel shift-register modules: one dedicated to photon detection and the other to the associated reset signal. The Gray counter is implemented as a fully asynchronous sequential circuit composed of two cascaded rising-edge–triggered flip-flops, both clocked directly by the photon pulse. In this configuration, the output Q_1 of the first flip-flop serves as the input to the second, while the output Q_2 of the second flip-flop is fed back, after inversion, to the input of the first. A block diagram of the implemented architecture is shown in Figure 9.

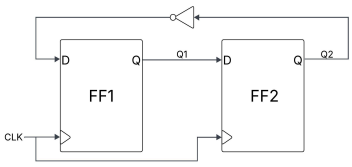

Figure 9: Gray-code block diagram.

The circuit operates as follows: at each rising edge of the photon signal, and after the associated clock-to-output propagation delay, the first

flip-flop (FF1) updates its output Q_1 with the value present at its input D_1, while the second flip-flop (FF2) simultaneously latches the previous value of Q_1 onto Q_2. After the inverter delay, the input D_1 is refreshed with the inverted value of Q_2, thereby generating the canonical two-bit Gray-code progression. The resulting Gray-coded outputs are then fed into a bank of four OR-gate decoding stages implementing the channel-selection conditions.

To assess the practical advantages of this architecture, comparative measurements were performed against an external time-tagger instrumentation. The experiment intentionally employed input pulse conditions beyond those recommended in its datasheet, which specifies a minimum FWHM of 450ps and a dead time below 680ps. Two input pulses, each 200ps wide and separated by 500ps, were applied first directly to the time-tagger and then through the implemented shift-register. When applied directly to the time-tagger, the system failed to detect all events and exhibited pronounced timing inaccuracies. Figure 10 illustrates this effect, showing the broadened, multi-peaked distribution produced by the commercial instrumentation, highlighting its inability to resolve closely spaced photon events precisely. In contrast, the shift-register architecture successfully resolves the two events, producing two narrow and well-separated peaks, spaced by approximately 500ps, exactly matching the temporal separation between them.

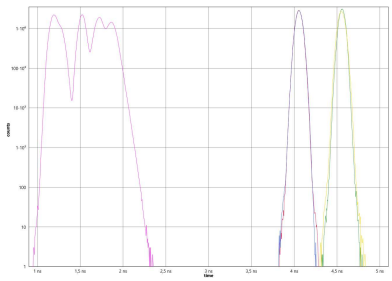

Figure 10: Time-tagger VS Shift-Register.

8. Conclusions

In conclusion, the resulting system combines modularity, a dead time as low as 1ns, and compatibility with the pile-up correction strategy, making it particularly well suited for high photon-flux applications such as Fluorescence Lifetime Imaging Microscopy (FLIM) for real-time visualization of biological dynamics [2] and intraoperative tumor margin delineation [5].

Future developments may first focus on completing the FPGA firmware for full processing, including pile-up correction, and subsequently on extending the architecture to a multi-channel configuration, thereby broadening its applicability. Additionally, the integration of the 3 GHz shift-register architecture directly onto the FPGA-based PCB could replace the current digital shift-register implementation, further increasing the maximum achievable photon detection rate and overall system throughput.

References

[1] Wolfgang Becker. *The bh TCSPC Handbook*. Becker & Hickl GmbH, Berlin, Germany, 9 edition, 2021.

[2] Fratta Gennaro, Gandolfi Valerio, Daniele Piergiorgio, Simoni Federico, Labanca Ivan, farina Andrea, Acconcia Giulia, Gola Alberto, D'Andrea Cosimo, and Ivan Rech. Up to 200% of the excitation frequency in compressive flim: A constraint-free pile-up correction methodology for advanced biomedical imaging. *Optics Express*, 33(12):26444–26460, 2025.

[3] Ivan Rech, Angela Bovolenta, Alessandro Cominelli, and Giulia Acconcia. Toward constraintless time-correlated single-photon counting measurements: A new method to remove pile-up distortion. *IEEE Journal of Selected Topics in Quantum Electronics*, 30(1), January 2024.

[4] Alessandro Rossetta. *The BrightEyes-TTM: an open-source time-tagging module for fluorescence lifetime imaging microscopy applications*. PhD thesis, University of Genova, 12 2021.

[5] Jakob Unger, Christoph Hebisch, Jennifer E. Phipps, João L. Lagarto, Hanna Kim, Morgan A. Darrow, Richard J. Bold, and Laura Marcu. Real-time diagnosis and visualization of tumor margins in excised breast specimens using fluorescence lifetime imaging and machine learning. *Biomedical Optics Express*, 11(3):1216–1230, 2020.