Multi-Channel SPAD Chip for Silicon Photonics With Multi-Photon CoIncidence Detection

Alfonso Incoronato, *Student Member, IEEE*, Fabio Severini, *Student Member, IEEE*, Federica Villa, *Member, IEEE*, and Franco Zappa, *Senior Member, IEEE*

***Abstract*—We present the design and characterization of a microelectronic chip consisting of 32 independent single-photon counting and multi-photon time-coincidence channels, based on Single Photon Avalanche Diodes (SPADs). The chip aims at easing the assembly together with Silicon Photonics substrates and chips, for a broad spectrum of quantum applications. The chip provides not only 32 independent pulse outputs for multi-channel single-photon counting, but it features other two operation modalities. For single-photon applications, the chip provides the single-hit digital address of the channel detecting the photon among the 32 ones. For multi-photon applications, the chip provides an event-driven pulse every time more than n photons (n selectable between 2 and 4) concurrently trigger different channels. Each detection channel consists of 4 independent SPADs with different diameter (5 μm, 10 μm, 20 μm, and 50 μm), to easily match the waveguides and collection efficiency.**

***Index Terms*—Single photon avalanche diode (SPAD), photon counting, quantum sensing, silicon photonics, photonic integrated circuits.**

I. Introduction

SINGLE Photon Avalanche Diode (SPAD) detection channels and arrays are more and more required in many applications, such as Quantum Imaging [1], photonics [2], molecular studies using Fluorescence Lifetime Imaging (FLIM) [3], Förster Resonance Energy Transfer (FRET), Fluorescence Correlation Spectroscopy (FCS), to the recent exploitation of Light-Detection and Ranging (LiDAR) chips [4] for the automotive field. While some applications require dense arrays of SPADs to provide a detailed image of the specimen under observation, making pixel count of the utmost importance, many other applications requisite high performance, multi-channel (a few tens at the most) chips with on-chip pre-processing of the photon detection events. For instance, integrated photonic applications require linear arrays of SPADs to reveal single photons, so geometry and layout should be devised and tailored to assist the coupling between optics and electronic.

Several examples of high efficiency and low noise SPAD detectors monolithically integrated in CMOS technology are reported, as in [5]–[7]. Indeed, photodetection probability (PDP) can range from 20% to over 40%, also assisted by microlenses enhancement, and dark count rate (DCR) can be as low as few tens of Hz.

Depending on the application, a plethora of specifically tailored linear SPAD arrays have been developed, exhibiting various form factors (ranging from a few tens of pixels to thousands) and diverse processing capabilities. For instance, [8] can independently deliver at the output the triggering activity of each pixel, while [9] is a fast-gated sensor with photon counting operation mode. [10]–[15] are designed so to operate both in photon counting and photon timing resolution (with resolutions ranging from few hundreds to tens of ps), and [15] additionally embeds an histogramming circuit directly on chip.

In this paper we discuss the design and the experimental characterization of a novel SPAD chip based on 32 independent SPAD channels, with on-chip processing of detected photons. The chip aims at easing the assembly with Silicon Photonics substrates embedding waveguides and non-linear optics, so to develop a quantum system on-chip (QSoC) capable of enabling photonic technologies to accommodate quantum experiments into millimeter-size chips, by shoehorning complex systems, which are presently found on meter-scale optical breadboards. These very compact systems and assembly will not only reduce overall system size and cost, but will also bring improvements in terms of robustness, reproducibility, and deployment to the market.

One example of these applications is the Quantum Random Number Generator (QRNG) [16], that relies on the statistics of single-photons to provide the true random digital words needed to encrypt and decrypt data [17], [19]. Single photons can be exploited in different ways to provide random numbers, for example by measuring their interarrival time [18], by counting their number within time slots [19], by detecting which optical path they travel among many available, just to mention a few. In the last method, a single photon is generated and fed to multiple-optical paths, e.g., waveguides on a Silicon Photonic substrate, to let it travel through one of them, and finally to identify which one was randomly selected, by means of a SPAD detector at the end of each path [20]. This so-called "optical path" QRNG requires a single photon source, an integrated optical

Manuscript received July 31, 2021; revised November 1, 2021; accepted December 7, 2021. Date of publication December 9, 2021; date of current version December 31, 2021. This work was supported by the European Union's Horizon 2020 research and innovation programme, under Grant Agreement 820474 "UNIQORN" (https://quantum-uniqorn.eu/). *(Corresponding author: Alfonso Incoronato.)*

The authors are with the Dipartimento di Elettronica, Informazione e Bioingegneria of Politecnico di Milano, 20133 Milano, Italy (e-mail: alfonso.incoronato@polimi.it; fabio.severini@polimi.it; federica.villa@polimi.it; franco.zappa@polimi.it).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/JSTQE.2021.3134071.

Digital Object Identifier 10.1109/JSTQE.2021.3134071

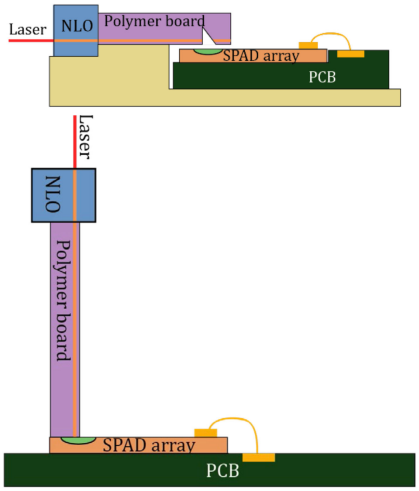

Fig. 1. Possible integration of the SPAD array with silicon photonics contained for example on a polymer board with Nonlinear Optics (NLO) to obtain an extremely compact quantum system on a chip. (Top) Planar integration with 45° V-groove; (bottom) 90° system integration for face-to-face contact between waveguides edges and SPADs. Everything can be assembled on top of a application specific Printed Circuit Board (PCB).

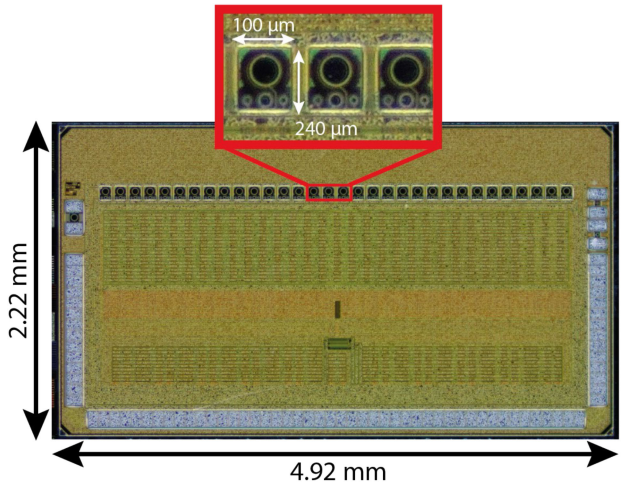

Fig. 2. Microphotograph of the 32 channel SPAD chip, with 4.92×2.22 mm^2 size. Note that the chip side above the SPADs has been left with no pad for allowing face-to-face contact with the application-specific photonics chip containing the waveguides.

circuit capable of randomly route the photon through many available paths, and a set of single-photon detectors that sense the photon and provide the address of the detection position.

II. System Integration

The primary goal of the chip is to be the assembled together with photonics substrates to provide a compact quantum system on a chip, for example an optical-path QRNG. The integrated photonics may be laid out in a polymer board, e.g., the PolyBoard in [23] and [24] designed by HHI Fraunhofer institute, consisting of polymer-embedded single-mode waveguides, coupled with a bulk nonlinear crystal (NLO). The input laser pulse at 785 nm is attenuated down to the single photon level (through polarization beam splitters, dichroic mirrors, long pass filters, and half wave plates, inserted perpendicularly to the waveguide layer into etched slots), randomly split through the waveguides as described in [25], and finally detected by the SPAD chip. A Printed Circuit Board (PCB) provides power supply and signals to the chip, as well as all the application-specific chip management and data processing.

To prevent excessive coupling losses from photon generation to photon detection and to assure signal integrity between SPAD chip and PCB, the integration among the components needs to be devised and performed thoughtfully. Fig. 1 schematically depicts two possible options. The first solution (Fig. 1 top) requires a polymer board with a 45° V-groove and a mirror to allow planar assembly. Overall, this is a space-saving approach immune to signal integrity errors, as all parts lay along the horizontal direction and the array is directly wire bonded to the PCB. However, the presence of a V-groove makes the system prone to light losses and a widening of the incoming beam.

The second solution (Fig. 1 bottom) is more favorable for the light signal, since the SPAD chip is directly facing the PolyBoard edge waveguide outputs, with no in-between mirror reflections. However, this solution occupies a larger volume, making the final system packaging more challenging.

Based on these two assembly options, we conceived the SPAD chip geometry to ease the integration with other photonic devices and to simplify alignment and mounting with the integrated optics.

III. SPAD Chip Architecture

For the chip design we selected the 160 nm BCD (Bipolar-CMOS-DMOS) microelectronic technology by STMicroelectronics, which we already tailored and tested in previous fabrication runs [21].

The produced chip shown in Fig. 2 has 4.92×2.22 mm^2 dimensions, since chip layout has been heavily influenced by considerations about system integration with other photonic substrates and components, giving constraints particularly on the pads and wire-bonding positions. In fact, the wire-bonding pad-ring cannot frame the whole chip as it is usually done (i.e., by surrounding the overall SPADs and on-chip electronics), because of the integration with the PolyBoard that would touch the chip (see Section II), thus covering and leaving inaccessible the area of the SPAD chip around the SPADs active areas. Hence, we laid out the pads only on the other three sides of the chip.

The chip was conceived to provide 32 independent detection channels with 4 selectable SPADs with different dimensions, as shown in the red-highlighted zone of Fig. 2. SPAD diameters of 5 μm, 10 μm, 20 μm, and 50 μm have been implemented to give the possibility to couple the chip to different waveguides core dimensions, and assembly procedure. Only one set of SPADs (all with the same diameter) are available at a time and they all keep the same pitch of 125 μm between identical SPADs.

The 32 channel outputs can be independently readout in parallel, through 32 digital pulse outputs. Furthermore, other advanced processing logic is implemented on-chip for the two operational modes. In the *Single-Hit Mode*, specifically designed for optical-path QRN generation, a selectable time window is

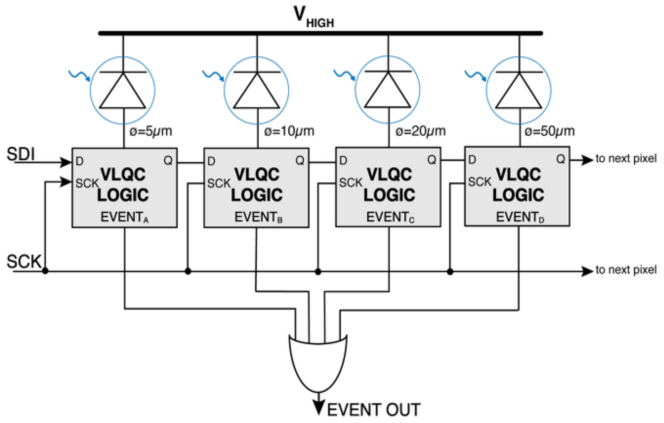

Fig. 3. Block diagram of a single channel architecture, with 4 different SPADs, individually selectable through a 1-bit serial communication.

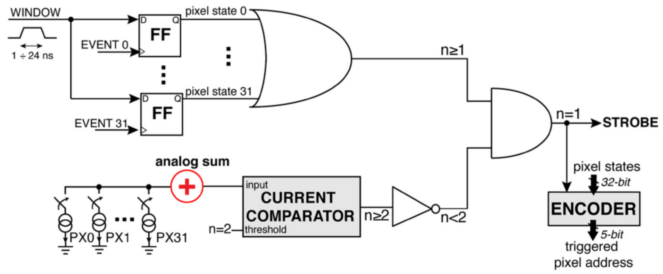

Fig. 4. Simplified block diagram of the logic for the Single-Hit Mode.

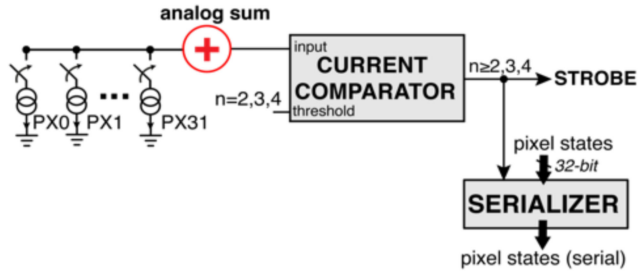

Fig. 5. Simplified diagram of the logic for the Multi-Hit Mode.

generated starting from an external signal synchronous with the laser pulse and, within that time window, a coincidence detection logic detects if one (and just one) photon has been detected across the whole channels. When this case happens, an output event is set along with the 5-bit address of the triggered SPAD. In the *Multi-Hit Mode*, the on-chip electronics detects if the number of channels detecting photons exceeds a threshold (that can be set from 2 to 4) within an asynchronous coincidence time window, configurable from 1 ns to 24 ns, with 1 ns steps. Once the coincidence happens, the SPADs' status (triggered/not triggered) is fed to a 1-bit serial output bus.

A. Channel Architecture

The architecture implemented for each channel is shown in Fig. 3. Each of the 32 channels embeds four SPADs with different diameters to make the linear array compatible with different waveguide dimensions and fiber diameters. Just one SPAD per channel must be enabled at a time, and this is made possible through a serial configuration (by means of a Serial Data In, SDI, and Serial Clock, SCK, inputs). The four SPADs have their own specifically tailored Variable Load Quenching Circuits (VLQC), necessary for promptly quenching the avalanche current and eventually to generate an EVENT digital pulse every time an avalanche is triggered. We redesigned the VLQC version presented in [26], to allow both gated and free-running operation modes, so to enable the exploitation of the array with different photon sources, either pulsed or continuous wave, respectively. While in the first mode, SPADs are sensitive to photons just in a few nanoseconds time window synchronous with an external trigger, in the latter SPADs are ideally always ready to photon detection, apart from the so-called hold-off time after each triggering. The internal logic actively quenches the SPAD as soon as the GATE signal is at low level, while in free-running the SPAD is quenched only when an avalanche occurs. In fact, since some avalanche charges can be trapped within the SPAD junction and get released at later times (afterpulsing effect), the SPAD must be kept below the breakdown voltage for a time interval known as hold-off time. For chip flexibility, this time is adjustable through an external analog voltage, from 10 ns to 200 ns, either favoring high output data rates or low afterpulsing probability, when choosing either a short or a long hold-off time, respectively.

B. Processing Logic

The processing logic is based on the photon coincidence concept and can be configured in two distinct modes, namely Single-Hit and Multi-Hit modes.

In *Single-hit Mode* (see Fig. 4) the array provides a flag signal named *STROBE*, every time only a single photon is detected within an adjustable coincidence time-window (from 1 to 24 ns) synchronized to an external trigger, eventually providing the address of the solely triggered channel. In simplified terms, *STROBE* signal generates from the logic AND of two flag-signals, the first signal is obtained by performing a logic OR of channels' status, latched within the programmable time window. The latter is obtained by inverting the flag-signal, stating if more photons have been detected. To discriminate the arrival of more than two photons, each channel EVENT signal generates a fixed-amplitude and fixed-duration current pulse, which are then all summed together into an analog summing node and then fed to a current comparator with an adjustable current threshold (equivalent to n = 2 photons), whose output toggles just in case the condition of more than two triggered channels is verified. At last, if *STROBE* is at high logic state (i.e., just one photon detected by the overall chip), the pre-latched channel states are encoded into the 5-bit address code of the triggered channel (from 1 to 32).

To determine the number of concurrent photons an analog approach has been preferred to a digital one, because it is intrinsically faster and with reduced circuital complexity with respect to the digital counterpart.

In *Multi-hit Mode* (see Fig. 5) the array is configured to detect if more than 2, 3 or 4 photons have been detected within a certain coincidence window, by once again exploiting the current

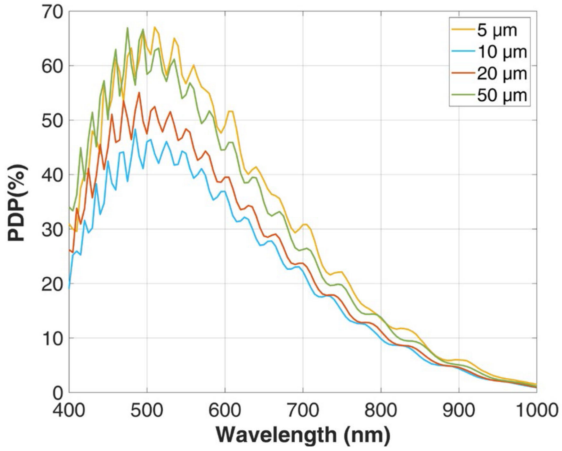

Fig. 6. Photon Detection Probability (PDP) of different size SPADs, at room temperature and at 5 V excess bias.

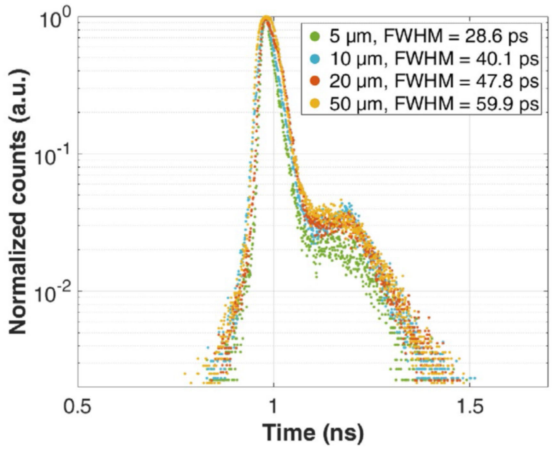

Fig. 7. Timing jitter of the different size SPADs, at room temperature and at 5 V excess bias. The tail is due to the laser and not to the SPAD intrinsic response.

comparator with adjustable threshold (equivalent to n = 2, 3, or 4 photons). Likewise single-hit mode, the signal to be compared is the sum of all current pulses generated by each triggered channel. However, in this mode, the current has a programmable duration (from 1 to 24 ns), equal to the desired coincidence time window. Thus, every time the threshold set by the user is crossed within the coincidence time-window, a *STROBE* signal is generated, and the pre-latched channel states are readout through a 1-bit serial line.

IV. Experimental Results

Both the SPAD detectors array and the on-chip electronics have been characterized and validated. At first, we assessed the performance of the four different size SPADs in each channel, then we verified the correct behavior of the whole on-chip processing. The following measurements have been performed at 5 V excess bias (with 26 V breakdown voltage SPADs).

A. Photon Detection Probability

In SPAD based systems it is very important to assess the probability to reveal a photon when it hits the active area, especially in photon starving applications. The measured Photon Detection Probability (PDP) shows a peak of 67% at 500 nm wavelength for the 5 μm and 50 μm diameter SPADs, as shown in Fig. 6. At the edge of the active area a lower detection probability is observed, thus decreasing the effective detector diameter. This effect is more evident in smaller SPADs (such as 10 μm and 20 μm) leading to a lower PDP (45% and 55% at 500 nm, respectively). The 5 μm SPAD is realized by covering a 10 μm SPAD with a tighter metal pinhole. The light is thereby concentrated in the center of the device, so no edge effects are presented, maximizing the PDP. The chip back-end stack is not optimized, featuring some interference artifacts in the visible range.

B. Timing Jitter

The SPADs timing performance have been measured in a classical Time-Correlate Single-Photon Counting (TCSPC) setup. We excited the SPAD chip with a narrow-pulsed laser at 820 nm with 15 ps Full-Width at Half-Maximum (FWHM) and acquired the photon arrival time with a SPC-630 (BECKER & HICKL) TCSPC module with 8 ps resolution. The reconstructed histogram shown in Fig. 7 the total jitter contributions from laser source, SPAD intrinsic jitter, and readout electronics jitter, resulting in 28.6 ps, 40.1 ps, 47.8 ps and 59.9 ps FWHM, respectively for the 5 μm, 10 μm, 20 μm and 50 μm SPADs, respectively. Smaller SPADs have tighter space charge region, so the avalanche builds up and the carrier drifts has a lower dispersion. The non-perfectly gaussian shape of the laser pulse leads to a noticeable tail, not due to the SPAD diffusion tail, which is much shorter.

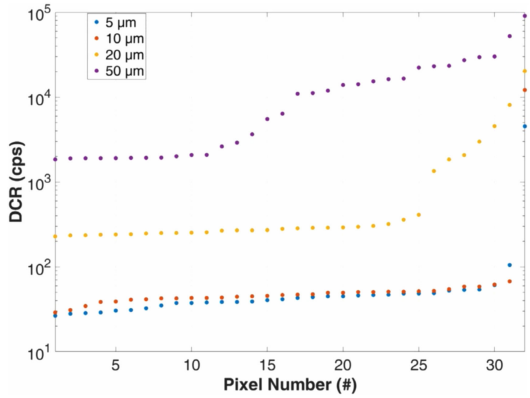

Fig. 8. DCR Cumulative Distribution Function of the 32 SPADs of different diameters, at 5 V excess bias and room temperature.

C. Dark Counts and Afterpulsing

The main SPADs noise source is related to avalanches triggered by thermal generated carriers or band-to-band tunneling [27], and is measured as Dark Counting Rate (DCR). Fig. 8 shows the Cumulative Distribution Function (CDF) of DCR at different SPAD diameter. The SPAD chips show a median of 45 cps (counts per second) for both the 5 μm and 10 μm SPADs,

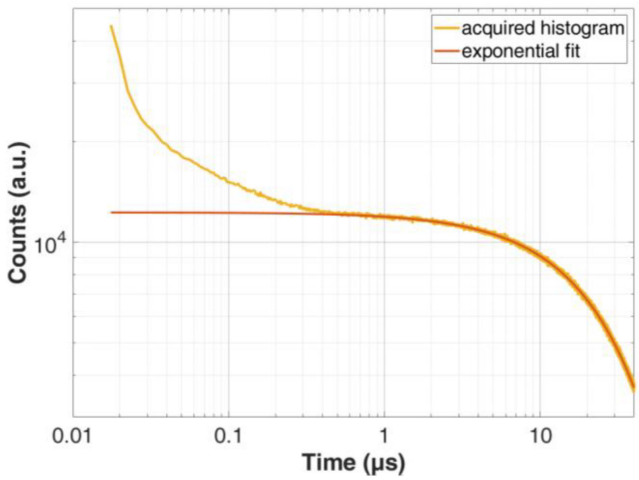

Fig. 9. Afterpulsing measurements through exponential fitting (50 μm SPAD).

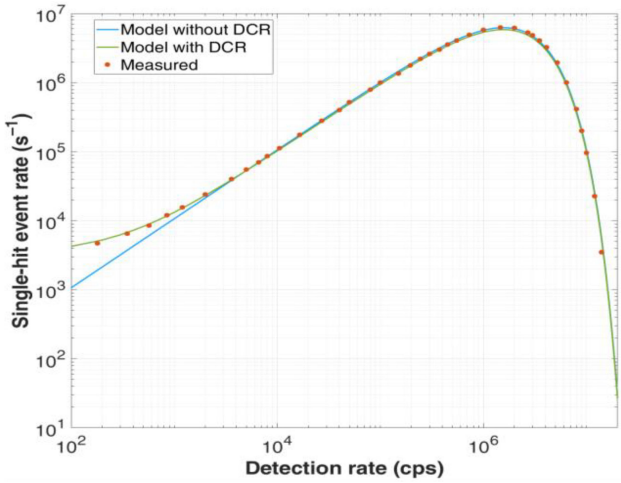

Fig. 10. Measured Single-hit event rate (*STROBE* frequency) across the 32 SPADs with 5 μm diameter and 20 ns time window, compared to the theoretical estimation computed with and without correcting for each SPAD DCR.

285 cps for the 20 μm ones and 8.6 kcps for the 50 μm ones, still at 5 V excess bias and room temperature.

The 50 μm SPADs also show a higher population of hot pixels (i.e., SPADs with a DCR much higher than the nominal one), compared to the smaller ones: the reason being the higher probability to include lattice defects and microplasmas within its active area. Identical DCR and distribution can be observed for 5 μm and 10 μm SPADs, because they have the same nominal active area diameter, while the 5 μm ones have simply a smaller metal pinhole on top.

The second noise source we analyzed is the afterpulsing effect. Due to process defects, some of the carriers flowing through the avalanche process can be trapped and released suc-cessively, thus potentially triggering a second correlated avalanche process, which does not correspond to an actual photon detection. The afterpulsing is estimated by computing the inter-arrival time of SPAD output pulses: the SPAD is illuminated by a very weak source, at around 20 kcps and all the time intervals between a first and a second output pulse are computed and stored into a histogram. Accordingly with Poisson statistics, all inter-arrival times corresponding to either revealed photons or DCR follow an exponential decay in time that can be fitted starting from the time afterpulsing can be considered negligible (after about 700 ns in this experiment, as it can be seen in Fig. 9). By subtracting such a fitting from the measurements, it is possible to compute just the spurious detections due to afterpulsing, as shown in Fig. 9, where the inter-arrival times are computed through a 2.5 ns resolution TDC implemented on an Artix-7 FPGA and across a 160 μs total range. The SPADs show 0.025% afterpulsing probability for the 5 μm and 10 μm sizes, 0.19% for the 20 μm and 0.31% for the 50 μm with the minimum achievable hold-off time of 12 ns. As already discussed for the DCR, the same performance of the two smaller diameters is due to the identical active area of both, which differ only by the metal pinhole on top, defining the actual active diameter: hence, defects and noise are similar.

D. Single-Hit Characterization

We characterized the single-hit mode, where the generated *STROBE* pulse is synchronous with an event corresponding to just one photon detected within the time window. We measured the single-hit event rate (*STROBE* signal frequency) at different photon rates and with all 32 SPADs activated, when the chip is uniformly illuminated. We may expect a low single-hit event rate both at low and very high photon rates, because of the low probability to detect one photon within the synchronous window at low photon rates, while the high probability to detect more than one photon at very high photon rates. To better understand this trend, we computed the probability of single-hit events, given the window duration w, the detection rate r (i.e., the average rate of photons detected by each SPAD), and the number N of SPADs:

$$p \;=\; N \cdot r \cdot w \cdot (1 - r \cdot w)^{N-1} \tag{1}$$

With a periodic window and a given duty-cycle d, the single-hit *STROBE* rate s can be estimated by:

$$s \;=\; \frac{p}{w} \;\cdot d \tag{2}$$

Fig. 10 shows the perfect match at high detection rates, between measured and theoretical computed single-hit event rate. By correcting (1) for the individual rates of each SPAD (differing cause of different DCR), the estimated rate (green line), perfectly matches the measure one, also at low detection rates. As expected, the trend peaks at intermediate rates, about 2 Mcps for measurement performed with 20 ns windows at 16.6 MHz (0.33 duty-cycle), employing all 32 5 μm SPADs.

Then, we evaluated the spread of single-hit events among all channels, by acquiring the address value of the triggered channel SPAD at every *STROBE* event and by building a histogram for each channel. Fig. 11 shows the standard deviation of single-hit events among the 32 channels over the average rate at different photon rates. At very low light ($<$ 1000 cps) the relative variation is very high due to hot pixels (showing 4 kcps vs 45 cps) detecting more single-hit events than other channels. At higher photon rates, events are more equally distributed, achieving a relative variation of about 2%. At different time window durations, measurements do not significantly change.

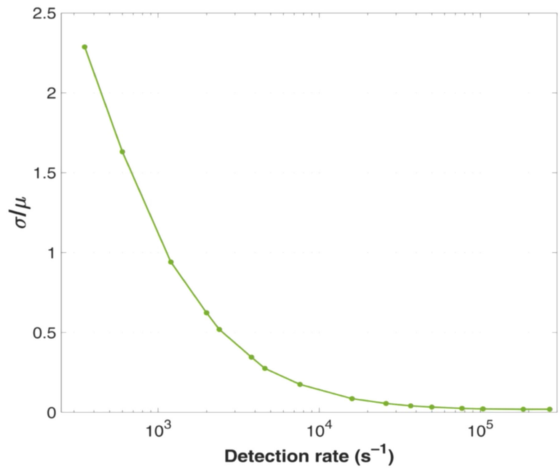

Fig. 11. Relative variation (standard deviation/average) of single-hit events among 32 channels (5 μm SPADs and 5 ns window) at different photon rates.

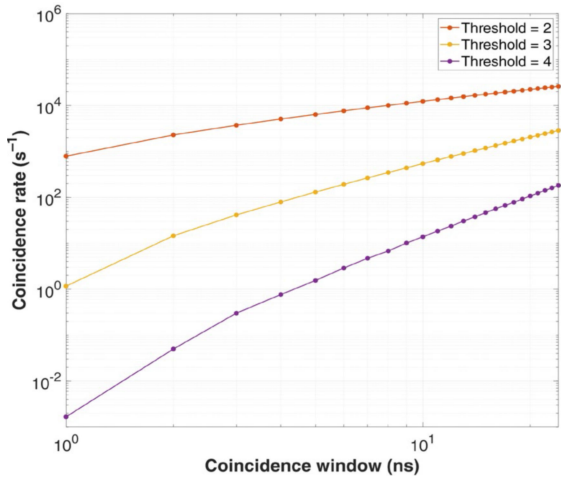

Fig. 12. Coincidence rate at different coincidence time windows duration and photon thresholds.

E. Multi-Hit Logic Characterization

We also tested the multi-hit logic, by exploiting the same setup. One of the available photon thresholds (2, 3 or 4) is set and coincidences are acquired at different windows durations (1 ÷ 24 ns). Fig. 12 shows the coincidence rate at different coincidence window duration and coincidence threshold, with 300 kcps average photon detection rate in each SPAD. As expected, the higher the threshold the lower the coincidence rate, while the larger the window the higher the probability to detect more photons within it. The three acquisitions at different thresholds asymptotically tend to polynomial curves, whose order increases with the number of photon threshold.

V. Conclusion

We have presented the design and the characterization of a chip integrating 32 single-photon counting and multi-photon time-coincidence channels, based on SPADs. More specifically, the 32 channel embeds four different diameter SPADs (5 μm, 10, μm, 20 μm, 50 μm) and outputs can be either independently readout in parallel or fed to more advanced on-chip processing, with two operational modes.

In the first mode, *Single-Hit Mode*, the array is configured to detect if one (and just one) photon has been detected across the whole channels within a synchronous time window, along with the 5-bit address of the triggered SPAD. The maximum sin gle-hit rate is limited by the single detector hold-off time, that is 12 ns minimum, achieving about 85 MHz and 425 Mbps of random address. In the second mode, *Multi-Hit Mode*, the array is configured to discriminate if the triggered channels exceed a threshold (that can be set from 2 to 4) within an asynchronous coincidence time window, and the pixel states are readout serially.

Through an in-house designed system, we experimentally characterized the chip both in terms of SPAD performance and processing logic behavior. The devices characterization proofed state-of-the art performance of the BCD SPADs. PDP peaking at about 60% in the visible range, low noise, afterpulsing probability, and timing jitter.

Single-hit mode was characterized by measuring the single-hit signal rate, over different time windows, resulting in a fitting with the theoretical model, proofing absence of significant mismatches among channels.

In Multi-hit mode the number of multi-photon coincidences decreased with the equivalent photon threshold as expected.

Based on these results and on the optimized geometry, the chip can be considered ready to be assembled within a quantum system on chip.

References

[1] F. Madonini, F. Severini, F. Zappa, and F. Villa, "Single photon avalanche diode arrays for quantum imaging and microscopy," *Adv. Quantum Technol.*, vol. 4, no. 7, 2021, Art. no. 2100005.

[2] F. Ceccarelli *et al.*, "Recent advances and future perspectives of single-photon avalanche diodes for quantum photonics applications," *Adv. Quantum Technol.*, vol. 4, no. 2, 2020, Art. no. 2000102.

[3] R. Datta, T. M. Heaster, J. T. Sharick, A. A. Gillette, and M. C. Skala, "Fluorescence lifetime imaging microscopy: Fundamentals and advances in instrumentation, analysis, and applications," *J. Biomed. Opt.*, vol. 25, no. 07, 2020, Art. no. 1.

[4] F. Villa, F. Severini, F. Madonini, and F. Zappa, "SPADs and sipms arrays for long-range high-speed light detection and ranging (LiDAR)," *Sensors*, vol. 21, 2021, Art. no. 3839.

[5] S. Pellegrini *et al.*, "Industrialised SPAD in 40 nm technology," in *Proc. IEEE Int. Electron Devices Meeting*, Jan. 2017, pp. 16.5.1–16.5.4.

[6] J. A. Richardson, E. A. Webster, L. A. Grant, and R. K. Henderson, "Scaleable single-photon avalanche diode structures in nanometer CMOS technology," *IEEE Trans. Electron Devices*, vol. 58, no. 7, pp. 2028–2035, Jul. 2011.

[7] J. A. Richardson, L. A. Grant, and R. K. Henderson, "Low dark count single-photon avalanche diode structure compatible with standard nanometer scale CMOS technology," *IEEE Photon. Technol. Lett.*, vol. 21, no. 14, pp. 1020–1022, Jul. 2009.

[8] S. Burri, C. Bruschini, and E. Charbon, "LinoSPAD: A compact linear SPAD camera system with 64 FPGA-based TDC modules for versatile 50 ps resolution time-resolved imaging," *Instruments*, vol. 1, no. 1, 2017, Art. no. 6.

[9] Y. Maruyama, J. Blacksberg, and E. Charbon, "A 1024 × 8, 700-PS time-gated SPAD line sensor for planetary surface exploration with laser Raman spectroscopy and LIBS," *IEEE J. Solid-State Circuits*, vol. 49, no. 1, pp. 179–189, Jan. 2014.

[10] J. N. Nissinen, P. Keranen, A.-K. Lansman, J. Holma, and J. Kostamovaara, "A 2×(4)×128multitime-gated SPAD line detector for pulsed Raman spectroscopy," *IEEE Sensors J.*, vol. 15, no. 3, pp. 1358–1365, Mar. 2015.

[11] A. Kufcsák *et al.*, "Time-resolved spectroscopy at 19,000 lines per second using a CMOS SPAD line array enables advanced biophotonics applications," *Opt. Exp.*, vol. 25, no. 10, 2017, Art. no. 11103.
[12] F. Villa, R. Lussana, D. Tamborini, A. Tosi, and F. Zappa, "High-fill-factor 60×1 SPAD array with 60 subnanosecond integrated tdcs," *IEEE Photon. Technol. Lett.*, vol. 27, no. 12, pp. 1261–1264, Mar. 2015.
[13] I. Cusini, K. Pasquinelli, E. Conca, and F. Villa, "Time-gated 128×1 and 8×8 SPAD cameras for 2D photon-counting and 3D time-of-flight maps," *Quantum Opt. Photon Counting 2021*, pp. 27–32, 2021.
[14] C. Niclass, M. Soga, H. Matsubara, S. Kato, and M. Kagami, "A 100-M range 10-frame/s 340×96-pixel time-of-flight depth sensor in 0.18μm CMOS," *IEEE J. Solid-State Circuits*, vol. 48, no. 2, pp. 559–572, Feb. 2013.
[15] H. Seo *et al.*, "A 36-channel SPAD-integrated scanning lidar sensor with multi-event histogramming TDC and embedded interference filter," in *Proc. IEEE Symp. VLSI Circuits*, Aug. 2020, pp. 1–2.
[16] M. Herrero-Collantes and J. C. Garcia-Escartin, "Quantum random number generators," *Rev. Modern Phys.*, vol. 89, no. 1, 2017, Art. no. 015004.
[17] J. Rarity, P. Owens, and P. Tapster, "Quantum random-number generation and key sharing," *J. Modern Opt.*, vol. 41, no. 12, pp. 2435–2444, 1994.
[18] Q. Yan, B. Zhao, Z. Hua, Q. Liao, and H. Yang, "High-speed quantum-random number generation by continuous measurement of arrival time of photons," *Rev. Sci. Instruments*, vol. 86, no. 7, 2015, Art. no. 073113.
[19] S. Tisa, F. Villa, A. Giudice, G. Simmerle, and F. Zappa, "High-speed quantum random number generation using CMOS photon counting detectors," *IEEE J. Sel. Topics Quantum Electron.*, vol. 21, no. 3, pp. 23–29, May/Jun. 2015.
[20] M. Gräfe *et al.*, "On-chip generation of high-order single-photon W-states," *Nat. Photon.*, vol. 8, no. 10, pp. 791–795, 2014.
[21] M. Sanzaro *et al.*, "Single-photon avalanche diodes in a 0.16 μm BCD technology with sharp timing response and red-enhanced sensitivity," *IEEE J. Sel. Topics Quantum Electron.*, vol. 24, no. 2, Mar./Apr. 2018, Art. no. 3801209.
[22] I. R. Berchera and I. P. Degiovanni, "Quantum imaging with sub-Poissonian light: Challenges and perspectives in optical metrology," *Metrologia*, vol. 56, no. 2, 2019, Art. no. 024001.
[23] M. Kleinert *et al.*, "A platform approach towards hybrid photonic integration and assembly for communications, sensing, and quantum technologies based on a polymer waveguide technology," in *Proc. IEEE CPMT Symp. Jpn.*, Nov. 2019, pp. 25–30.
[24] M. Kleinert *et al.*, "Photonic integrated devices and functions on hybrid polymer platform," *Physics and Simulation of Optoelectronic Devices XXV*, pp. 220–230, 2017.
[25] A. Sarkar and C. M. Chandrashekar, "Multi-bit quantum random number generation from a single qubit quantum walk," *Sci. Rep.*, vol. 9, no. 1, 2019.
[26] D. Portaluppi, E. Conca, and F. Villa, "32 × 32 CMOS SPAD imager for gated imaging, photon timing, and photon coincidence," *IEEE J. Sel. Topics Quantum Electron.*, vol. 24, no. 2, Mar/Apr. 2018, Art no. 3800706.
[27] A. Migdall, *Single-Photon Generation and Detection: Experimental Methods in the Physical Sciences*, Amsterdam: Elsevier, 2013.

Alfonso Incoronato (Student Member, IEEE) received the B.Sc. degree in biomedical engineering in 2016 and the M.Sc. degree in electronics engineering in 2019, from Politecnico di Milano, Milan, Italy, where he has been working toward the Ph.D. degree in information technology (electronics), since November 2019. His research interests include the microelectronic design of SPADs and SiPM arrays for automotive LiDAR and quantum communication.

Fabio Severini (Student Member, IEEE) received the B.Sc. and M.Sc. degrees in electronics engineering from Politecnico di Milano, Milan, Italy, in 2016 and 2018, respectively. He is currently working toward the Ph.D. degree in information technology with the Di-partimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano. His research interests include the design and development of CMOS circuits containing SPAD detectors, targeting quantum imaging and quantum communication applications.

Federica Villa (Member, IEEE) was born in Milan, in 1986. Since 2015, she has been an Assistant Professor of electronics engineering with Politecnico di Milano, Milan, Italy. Her academic research interests mainly include the design of microelectronic integrated sensors and circuits in CMOS and bipolar-CMOS-DMOS (BCD) technologies for the development of single photon detectors, such as single photon avalanche diodes (SPADs) and silicon photomultipliers (SiPMs), photon-counting array imagers with single-photon sensitivity, and high-resolution and linearity time-to-digital converters (TDCs).

Franco Zappa (Senior Member, IEEE) was born in Milano, Italy, in 1965. Since 2011, he has been a Full Professor of Electronics with Politecnico di Milano, Milan, Italy. His research interests include microelectronic circuitry for single-photon detectors (SPAD) and CMOS SPAD imagers, for high-sensitivity time-resolved measurements, 2D imaging, and 3D depth ranging. In 2004, he cofounded Micro Photon Devices, producing SPAD modules and cameras for single photon-counting and photon-timing, and in 2020 pioNIRS, developing near-infrared spectroscopy instrumentation for non-invasive and continuative monitoring of biological tissues and food.