A 3.3-Gb/s SPAD-Based Quantum Random Number Generator

Pouyan Keshavarzian, *Graduate Student Member, IEEE*, Karthick Ramu, Duy Tang, Carlos Weill, Francesco Gramuglia, *Member, IEEE*, Shyue Seng Tan, Michelle Tng, Louis Lim, Elgin Quek, *Member, IEEE*, Denis Mandich, Mario Stipčević, and Edoardo Charbon, *Fellow, IEEE*

***Abstract*— Quantum random number generators (QRNGs) are a burgeoning technology used for a variety of applications, including modern security and encryption systems. Typical methods exploit an entropy source combined with an extraction or bit generation circuit in order to produce a random string. In integrated designs, there is often little modeling or analytical description of the entropy source, circuit extraction, and post-processing provided. In this work, we present a single-photon avalanche diode (SPAD)-based QRNG design, which utilizes the quantum random flip-flop (QRFF) method. Extensive modeling of detector and circuit imperfections that result in entropy degradation is performed. A new method to analytically model serial autocorrelations of the proposed bit generation method, which includes detector dead time, is proposed. Then, a Verilog-AMS model is developed in order to validate the analytical model in simulation. A novel transistor implementation of the QRFF circuit is presented, which enables compensation of the degradation in entropy inherent to the finite non-symmetric transitions of the random flip-flop. Finally, a full system containing two independent arrays of the QRFF circuit is manufactured and tested in a 55-nm bipolar–CMOS–DMOS (BCD) technology node, demonstrating bit generation statistics that are commensurate to the developed model. The full chip is able to generate 3.3 Gb/s of data when operated with an external LED. Pixelwise and spatial analysis of bias and correlation is performed. NIST STS (SP 800-22) and SP 800-90B are used to benchmark the generated bit strings.**

***Index Terms*— Entropy, hardware security, photon counting, quantum random number generator (QRNG), single-photon avalanche diodes (SPADs), Verilog-AMS.**

Manuscript received 19 July 2022; revised 15 January 2023 and 18 March 2023; accepted 30 April 2023. This article was approved by Associate Editor David Stoppa. This work was supported by the Swiss National Science Foundation under Grant 200021-169465. The work of Pouyan Keshavarzian was supported by Qrypt Inc., New York, NY, USA. *(Corresponding author: Pouyan Keshavarzian.)*

Pouyan Keshavarzian is with the Advanced Quantum Architecture (AQUA) Laboratory, École Polytechnique Fédérale de Lausanne (EPFL), 2000 Neuchâtel, Switzerland (e-mail: pouyan.keshavarzian@epfl.ch).

Karthick Ramu, Duy Tang, Carlos Weill, and Denis Mandich are with Qrypt Inc., New York, NY 10007 USA.

Francesco Gramuglia, Shyue Seng Tan, Michelle Tng, Louis Lim, and Elgin Quek are with GLOBALFOUNDRIES Singapore Pte. Ltd., Singapore 738406.

Mario Stipčević is with the Rudjer Boskovic Institute, 10000 Zagreb, Croatia.

Edoardo Charbon is with the Advanced Quantum Architecture (AQUA) Laboratory and the Center for Quantum Science and Engineering, Ecole Polytechnique Fédérale de Lausanne (EPFL), 1015 Lausanne, Switzerland.

Color versions of one or more figures in this article are available at https://doi.org/10.1109/JSSC.2023.3274692.

Digital Object Identifier 10.1109/JSSC.2023.3274692

I. Introduction

RANDOM number generators (RNGs) are well-established security primitives used in a variety of schemes ranging from key generation/distribution to, encryption, and privacy amplification [1]. With the proliferation of the Internet of Things (IoT) and connected devices, security has become a critical aspect of all system-level design. Consequently, true RNGs (TRNGs) [2], [3], which exploit some classical physical entropy source, are a mature technology available commercially as both discrete silicon devices and IP blocks inside more complex computing circuitry [4] and are able to achieve energy per bit ratios lower than pJ/bit [5]. However, due to the inherent limitations of classical entropy sources in providing sufficient randomness, i.e., limitation of bit bias and correlation, these TRNG ASICs often require complex post-processing to establish an acceptable output entropy in the generated bit stream, which, in turn, significantly reduces the output bit rate [1]. Finally, with the emergence of quantum computing, the security parameter for a generated key increases, doubling the required key length for symmetric encryption algorithms [6], [7].

Quantum RNGs (QRNGs), which exploit inherently random phenomena in nature, are promising technologies, which aim to address the challenge/trade-off between system complexity and randomness performance. Standardization of RNGs is ongoing (AIS 31 [8] and NIST SP 800 90-C [9]), while debate remains regarding requirements for and specifics of post-processing methods [10], [11], along with the validity of empirical randomness testing [12], [13]. Nevertheless, the exploitation of quantum phenomena provides advantages for the development of future RNGs, particularly for entropy-as-a-service (EaaS) [14] and quantum key distribution (QKD) applications [15], which necessitate very high bit generation rates.

Systems and methods for QRNG designs come in many flavors, including those which exploit photon timing statistics [16], [17], [18], polarization [19], quantum tunneling [20], and laser phase noise [21], to name a few. These can be broadly classified as *trusted* QRNGs, because they largely rely on the quantum nature of the entropy source as sufficient to realize a functioning generator. As an additional measure for combating environmental changes or attacks on the device itself, complex generators that are proven to be device [22], source [23],

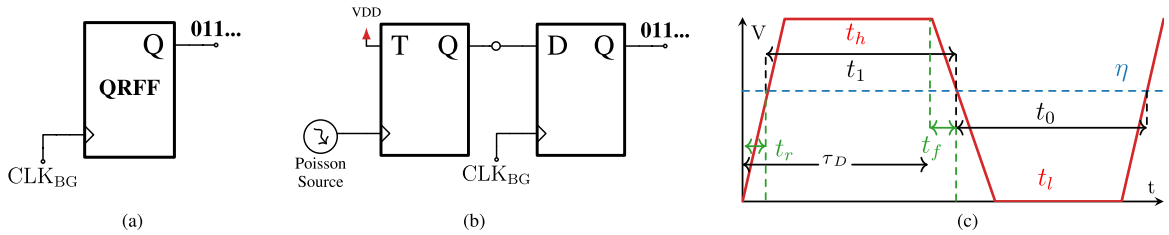


Fig. 1. Example circuit implementation of the QRFF concept presented in this work with a TFF and DFF combined with an exponentially distributed counting source. The waveform of the TFF output is shown to illustrate pertinent characteristics, such as the normalized sampling threshold (η), and rise and fall times (t_r and t_f) that contribute to bias. (a) Circuit symbol. (b) QRFF circuit implementation. Realistic waveform of TFF output (TFF_Q).

and measurement-independent [24] have been demonstrated in the literature. These device/source/measurement-independent implementations do not inherently trust that the entropy source and/or measurement device are working properly, but rather validate the quantum nature of the experiment by performing Bell tests [25]. However, these methods require bulky optical setups and, therefore, are impractical for monolithic integration. A compromise between trusted systems and those that contrive more secure bounds, using post-processing or source/device independence, are so-called self-testing QRNGs that test for generator defectiveness [26], [27]. This is performed by creating tests tailored specifically to verify the generator output string against its randomness model. Regardless of the generator design, those which provide the most pragmatic solution can be readily modeled, integrated in silicon, and scalable to produce designs with high data throughput. For these reasons, single-photon avalanche diode (SPAD)-based systems are attractive for QRNG technology development, as they are highly scalable (>1 Mpixel [28], [29]) and reproducible in silicon manufacturing.

The composition of this article is as follows. In Section II, we review some previously developed theory on the quantum random flip-flop (QRFF) circuit [30], [31] and, thereby, introduce the considerations for an integrated circuit that employs this method. Then, we augment the theory with new formulations for modeling correlations when taking into account detector dead time. From there, a Verilog-AMS model (Section III) is developed to thoroughly investigate, in simulation, how circuit imperfections manifest themselves in bias and correlations, thereby validating the analytical model of the bit generation method. In Section IV, a novel full-custom implementation of the QRFF flip-flop is proposed, which uses dynamic logic to overcome effects of finite and non-symmetric transitions present in logic circuits, on the quality of generated bit strings. This QRFF is then implemented in a 55-nm BCD process with measurements comparing the results to the analytical and simulated predictions provided. Finally, we scale the QRFF circuit to a full Gb/s QRNG design on chip. Section V presents two independent arrays that are capable of running concurrently. They are implemented on the same die with separate readout schemes and achieve a combined 3.3-Gb/s output data rate, showing the suitability of the approach in practice. These results are presented in Section VI. Several approaches for generating random bits using SPAD photon detection have been investigated, such as those which detect the presence of a photon within a gate window [32], [33], comparison of inter-arrival times [16], [34], [35], and the first detection between a pair of detectors [17], [18]. To the authors' knowledge, this is the first work where the proposed bit-generation method has been integrated on chip with SPADs. Integrated SPAD-based designs are compared in Section VII, followed by a conclusion in Section VIII.

II. Quantum Random Flip-Flop

A. Fundamental Operation

An entropy extraction/harvesting or bit generation method is required for RNG designs, regardless of the entropy source chosen. The QRFF describes a simple circuit concept that, upon the arrival of a clock strobe, generates a random bit. A symbol representation is shown in Fig. 1(a). A specific circuit realization of the QRFF concept is shown in Fig. 1(b). Here, a Poisson source, which has exponentially distributed inter-arrival times, clocks a toggle flip-flop that has its toggle input continuously held to logic 1, thereby realizing a random digital signal that can be seen as a random telegraph signal/process (RTS) with random transitions in time. In principle, as the arrival events occur randomly, the TFF output, over a sufficient integration time, is uniformly distributed $X \sim U\{0, 1\}$. Therefore, once the sampling DFF is clocked by the strobe signal CLK_{BG}, a random bit is generated. The architectural simplicity allows for accurate modeling of the bias and autocorrelation of generated bit strings, while the ability to vary internal parameters, such as the arrival rate of Poisson events, and external parameters, such as the generation rate, enables flexibility from a system point of view, which we will demonstrate in Sections II-B and II-C.

B. Model for Bias and Correlation

Evidently, no perfect source or circuit can exist, which then perfectly matches the theory of the concept of Fig. 1(a). The output of the TFF indeed has finite and non-symmetric rise/fall times. Furthermore, the sampling threshold of the signal, which distinguishes between low and high states, has some deviation from center, resulting in a RTS that resembles the waveform depicted in Fig. 1(c). On average, the time between transition edges, τ_{D}, is determined by the detection rate $\lambda_{\text{D}} = 1/\tau_{\text{D}}$. Theoretically, since the event arrivals happen at random times, the waveform can be interpreted as two equal half-periods with edge transitions controlled by the Poisson source. The rise and fall times are denoted by t_r and t_f, respectively, and represent the transition time between the "1" and "0" states until the level of the normalized sampling threshold, η, is reached.

It can be shown that the statistical bias, i.e., deviation from $P(X = 1) = 0.5$ for the high state, is described by the following equation [31]:

$$b = \frac{t_f - \eta(t_r + t_f)}{2} \cdot \lambda_D. \tag{1}$$

Some key guidelines for circuit design can be extracted from this model. First, it is clear that the bias should scale linearly in magnitude with increasing detection rate, and that it is desirable to have a fast TFF. Furthermore, it should be possible to compensate bias resulting from any non-symmetry of the rise/fall times by adjusting the sampling threshold during startup calibration.

Sources of correlation in any RNG must also be modeled and understood. The autocorrelation function for a binary RTS with normalized amplitudes is defined by the following equation [37]:

$$R_{XX}(\tau) = e^{(-2\cdot\lambda\cdot|\tau|)}. \tag{2}$$

The time lag interval, τ, for calculation of the autocorrelation coefficient, is controlled by the clock frequency of the sampling DFF in Fig. 1(b). Therefore, correlation coefficients, a_i, corresponding to specific bit lags, i, can be calculated with the following equation:

$$a_i = e^{(-2\cdot\lambda_D/(i\cdot f_{BG}))}. \tag{3}$$

The 1-bit lag correlation coefficient, a_1, therefore, has the highest magnitude and can be minimized by increasing the ratio λ_D/f_{BG}. Consequently, there exists an inherent trade-off between designing for acceptable bias, which increases linearly, and for correlation, which decreases exponentially, with increasing detection rate. More practically, for a generated bit string of length, n, the i-bit lag correlation coefficient can be calculated with the following equation:

$$a_i = \frac{n\left(\Sigma_{j=0}^{n-1} x_j x_{j+i}\right) - \left(\Sigma_{j=0}^{n-1} x_j^2\right)}{\left(n\Sigma_{j=0}^{n-1} x_j^2\right) - \left(\Sigma_{j=0}^{n-1} x_j\right)^2}. \tag{4}$$

C. Dead-Time Considerations

The detection statistics of an SPAD can deviate from that of a Poisson arrival process, when dead time is considered [38]. A more general approach is introduced here to model correlation of a circuit employing the QRFF shown in Fig. 1(b) when the Poisson source is replaced with a detector containing dead time. First, consider the RTS-like digital signal produced by the TFF, as shown in Fig. 2. Although the edge transitions can no longer be modeled as purely Poisson, they still occur at random times. Given a sampling period $\tau : T_{BG} = 1/f_{BG}$ chosen for evaluating autocorrelation, it can be seen that a correlated event happens when TFF_Q is in the same state at the end of the period as it was at the start. This occurs exclusively when an even number of detections has occurred during the period T_{BG}. The TFF output TFF_Q is generalized as a stochastic process, $\{X_{K(t)}\}_{t\in T} : \mathbb{R} \to [0, 1]$ with the random variable K, $k \in \mathbb{Z}_0^+$ denoting the number of edge transitions (detections) occurring in the interval T_{BG}. Specifically, given that a generated bit $x_i = 1$, then $x_{i+1} = 1$ happens if the number of detections is even, i.e., if k, in the period T_{BG}, is $k = \{0 \cup 2 \cup 4 \cup \cdots \cup k_{max}\}$. The maximum number of possible detections in the period is limited by the dead time, τ_{dead}, such that $k_{max} = T_{BG}/\tau_{dead}$. Therefore, the probability of each of $k = \{0, 1, \ldots, k_{max}\}$ detections can be calculated. The corresponding autocorrelation function is then evaluated as the probability of an even number of detections minus the probability of odd, in the interval T_{BG}

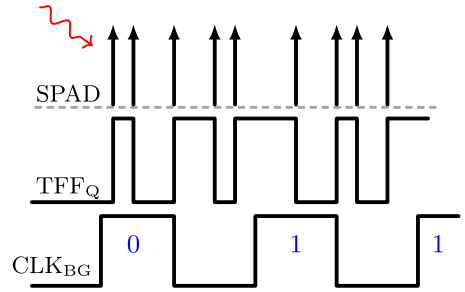


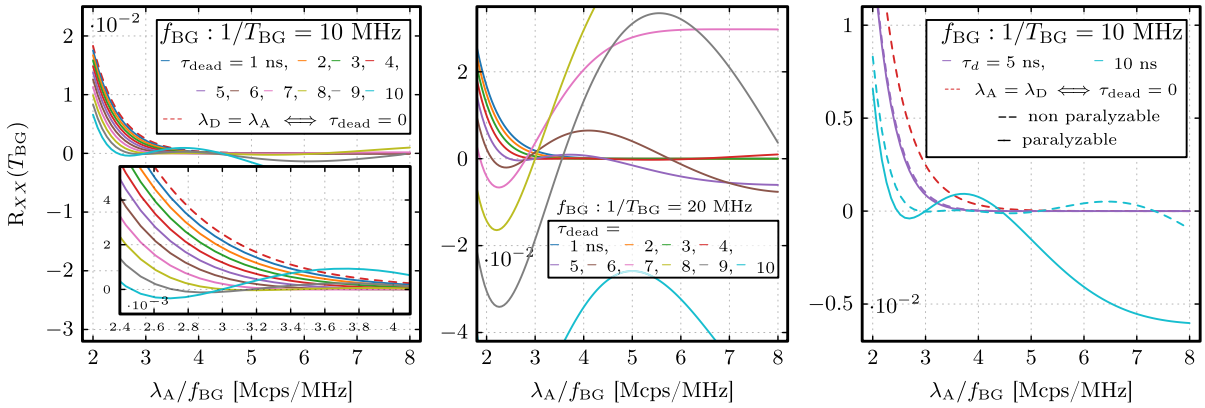
Fig. 2. Generalized timing diagram for circuit shown in Fig. 1(b). Bits are generated by sampling the TFF output (TFF_Q). Transitions of the toggle flip-flop occur at random times, but do not necessarily follow Poisson statistics. This concept is used for formulation of a generalized autocorrelation model, shown in (5).

$$R_{XX}(T_{BG}) = \mathbb{P}(K_{even}) - \mathbb{P}(K_{odd}). \tag{5}$$

In [36], new counting equations, which derived the probability of k detections in period T_{BG}, given an arrival rate of λ_A, were proposed using renewal theory. This work validated the derived equations with Monte Carlo simulations and empirical models previously proposed in the literature for both paralyzable and non-paralyzable dead times. A paralyzable dead time refers to a detector configuration that can result in the extension of the insensitive period (dead time), if a subsequent avalanche commences during the recharge phase, but before the discriminator circuit threshold is reached. In this case, the dead time is not well defined, whereas a non-paralyzable detector employs a pixel circuit that can precisely control dead time [40]. The implications of this for the SPAD-circuit interface (pixel) are discussed later on.

More recently, counting equations for non-paralyzable SPADs were also presented in [39]. The rigorous analysis in that work considered several case scenarios in order to derive the counting equations, such as the probabilities of whether the detector is in the sensitive or dead state at the beginning of the counting window. Furthermore, in their analysis, afterpulses and twilight pulses were augmented. The counting equations in [36] and [39] are used in this work to calculate autocorrelation [using (5)]; however, the equations from those works are not reprinted here for brevity. For clarity, the photon arrival rate λ_A is differentiated from photon detection rate λ_D, which are evidently not equal in the presence of dead time, as arrivals can occur, while the detector is not capable of initiating an avalanche (detection).

Using these counting equations, the autocorrelation functions defined by (5) can be evaluated for different arrival rates and sampling intervals. Fig. 3(a) plots $R_{XX}(T_{BG})$ across arrival-sampling rate ratios (λ_A/f_{BG}), with different dead times, including the ideal case, which corresponds to (2), using the paralyzable detector model. The results suggest



(a) (b) (c)

Fig. 3. Autocorrelation function of the QRFF circuit using the counting equations presented in [36], which present the paralyzable and non-paralyzable scenarios. Plots are shown at various arrival rate/sampling rate ratios (λ_D/f_{BG}). (a) $f_{BG} = 10$-MHz paralyzable dead times. The red dashed line shows the result for the ideal detector model. (b) $f_{BG} = 20$-MHz paralyzable dead times. (c) $f_{BG} = 10$-MHz comparison of dead-time models.

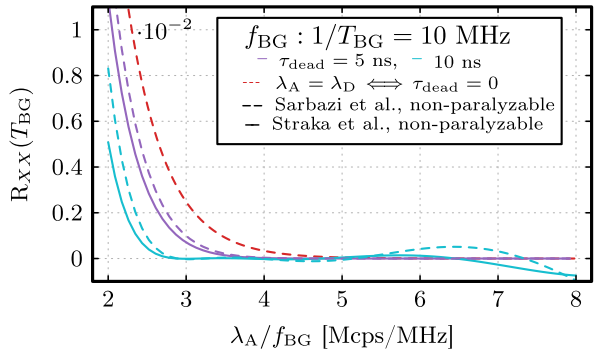


Fig. 4. Comparison of autocorrelation function with non-paralyzable dead time using the counting equations presented in [36] (dashed line) and [39] (solid line).

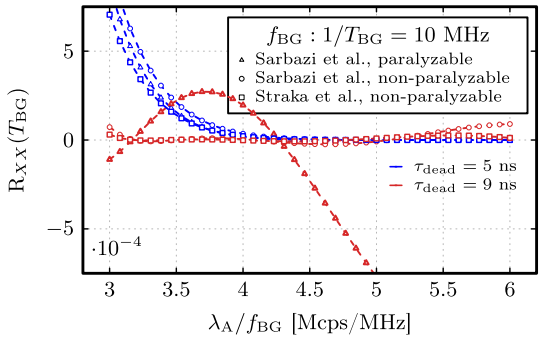


Fig. 5. Comparison of autocorrelation function with paralyzable and non-paralyzable dead times using the counting models presented in [36] and [39].

interesting phenomena. First, the serial correlation decays faster, i.e., at lower arrival rates with increased dead time. This is an advantageous characteristic, as it suggests that acceptable performance can be achieved with a lower flux requirement. Therefore, the power consumption of a system employing this method can be reduced by increasing dead time. However, at larger dead times, a natural trade-off exits caused by pileup effects, which manifests itself as high fluctuation of the correlation functions. This effect is exacerbated with increased sampling rates, which is accentuated in Fig. 3(b), where heavy oscillations are visible at low arrival rate sampling ratios (λ_A/f_{BG}), with longer dead times. Under these conditions, frequent arrivals with shorter mean inter-arrival times will have a higher probability of falling in the dead time of the detector, further extending the time between toggles in the paralyzable case. As a result, the sampling flip-flop over samples a given toggle state, resulting in higher correlation.

A comparison of the paralyzable and non-paralyzable cases is shown in Figs. 4 and 5. The results at low dead times are very similar. However, it is clear that a non-paralyzable detector is advantageous, as it allows for consistently low correlation at lower λ_A/f_{BG} ratios with longer dead times. From Fig. 5, it is shown that the non-paralyzable analysis is also representative of the paralyzable case, when the dead time is very low ($\tau_{dead} \leq 5$ ns) and the flux is not so high that pileup effects are present.

Practically, a non-paralyzable dead time is achievable when an active quenching circuit, that has a controllable hold-off time, is implemented in the SPAD pixel. During the hold-off period, the SPAD is not sensitive to incident photons. Moreover, an active recharge function must be implemented that can quickly reset the SPAD, limiting the time between when the SPAD becomes active and when the threshold of the discriminator circuit, that can register a pulse, is reached. These concepts will be expanded upon in the pixel design section.

D. Benchmarks for Performance

As noted earlier, while the security requirements of any given system and cryptographic scheme can vary, we aim to design a generator, which is capable of complying with entropy requirements for the upcoming version of the AIS-31 standard; therefore, the Shannon entropy, $H_1 = -\sum_{i=1}^{n} p_i \log p_i$, must remain $\geq$0.9998 for a sufficiently long bit string, and the min entropy, H_∞, must be $\geq$0.98 [8]. In this work, the serial correlation and bias of each individual pixel are analyzed. Moreover, cross correlation between generated bits of neighboring pixels is calculated to ensure there is no entropy degradation, i.e., spatial correlations caused by crosstalk or other phenomena. Using (6), the probability of occurrence for the most frequent n-bit symbol, generated from a string with correlation, a_1, can be

TABLE I
VERILOG-A MODEL PARAMETERS OF QRFF

Param.	Description
η	DFF sampling threshold normalized to 1 V
t_r	TFF rise time
t_f	TFF fall time
λ_{D}	Detection rate from exponential source
f_{BG}	Sampling frequency of bit gen. clock ($\mathrm{CLK_{BG}}$)
τ_{dead}	SPAD dead time.

estimated [41]

$$P_{z,\max} = \mathbf{max}\left(\left(\frac{a_1}{2}+0.5\right), 1-\left(\frac{a_1}{2}+0.5\right)\right)^n. \quad (6)$$

Equation (6) is used as a rudimentary method to estimate H_∞, while the NIST SP 800 90-B [42] test is used to characterize min entropy of the final generated bit string (Section VI). We aim to have no calculable bias or correlation higher than $|10^{-3}|$, which corresponds to $H_1(X) \simeq 0.999997$ and $H_\infty \simeq 0.9986$. The NIST Statistical Test Suite [43] (SP 800-22) is also used to validate the performance of overall bit strings generated by the final array.

III. VERILOG-AMS SIMULATION OF QRFF ANALYTICAL MODEL

A. Model Details

In order to validate these analytical equations, a simple SPICE-compatible Verilog-AMS model of the QRFF circuit was developed. An exponential source was used by taking advantage of the *$rdist_exponential* function provided by the Verilog-AMS language standard. The parameters in Table I were investigated as variables in simulation of bias, b, and correlation coefficients, a_i.

B. Simulation Results

Simulation results of bias are displayed in Fig. 6. The generation of bits is a binomial process with N trials; therefore, the variance of bias from simulation can be calculated with $\sigma^2 = 1/(4N)$. In our results, we plot $\pm\sigma$ for reference. Fig. 6(a) displays the simulated bias compared with the analytical calculation for varying t_r and t_f, given a fixed detection rate $\lambda_{\mathrm{D}} =$ 80 Mc/s and a sampling threshold, $\eta = 0.499$, placed close to the center of the waveform. As the discrepancy between the rise and fall time increases, so does bias, matching very closely to the analytical calculation. In Fig. 6(b), a similar analysis was performed but with a varying η. Here, we can see that a mismatch between t_r and t_f can be compensated for by adjusting the threshold, thereby allowing for the minimization of bias. This is a critical finding from the perspective of integrated circuit implementation, as the foundry process will always create some small, albeit present, variation, across an array, regardless of how carefully the circuit is designed. In order to confirm that the bias magnitude scales linearly with increased count rate, at a fixed sampling rate and threshold, a final simulation is performed with the results displayed in Fig. 6(c).

Autocorrelation simulations with comparison to the calculation of an RTS (3) are presented in Fig. 7. At a fixed sampling

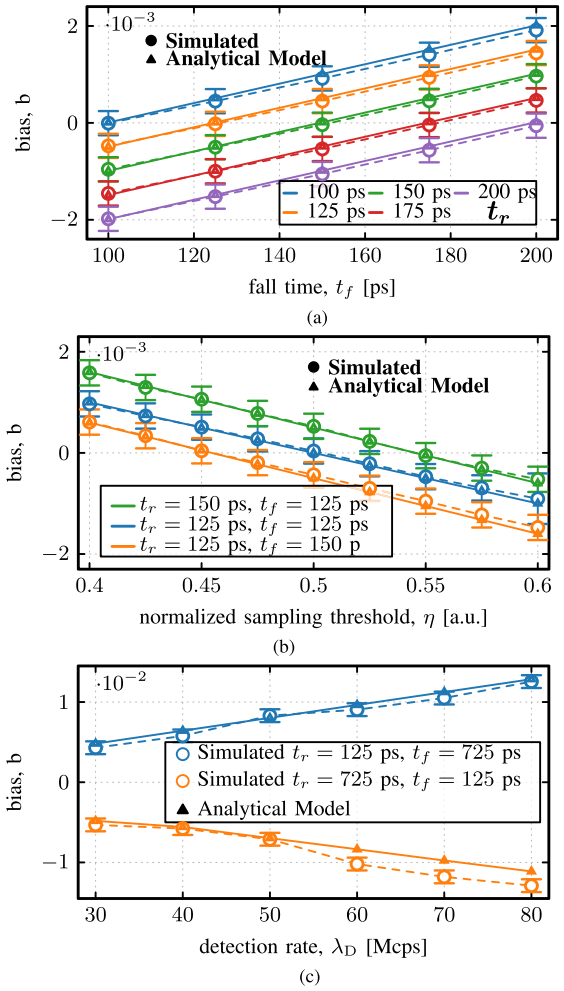


Fig. 6. Simulation of the bias from $P(X = 1) = 0.5$ with comparison to the analytical model in (1). σ for each simulation is plotted as error bars. (a) Rise/fall time (t_r and t_f) mismatch analysis: with fixed detection rate ($\lambda_{\mathrm{D}} = 80$ Mc/s), bit generation rate ($f_{\mathrm{BG}} = 25$ MHz), and normalized sampling threshold ($\eta = 0.499$). (b) Sampling threshold analysis (η): fixed detection rate ($\lambda_{\mathrm{D}} = 80$ Mc/s), and bit generation rate ($f_{\mathrm{BG}} = 25$ MHz) performed at various TFF rise/fall times. (c) Detection rate (λ_{D}) analysis with: fixed normalized sampling flip-flop threshold ($\eta = 0.475$) and fixed bit generation rate ($f_{\mathrm{BG}} = 25$ MHz). Rise/fall time discrepancy is deliberately exaggerated in order to increase bias, so the number of samples simulated can be reduced and still be statistically relevant.

rate, the 1-bit lag correlation coefficient should decrease exponentially as detection rate increases, which is indeed observed in the results of Fig. 7(a). Conversely, at a fixed detection rate, the correlation should increase exponentially for increased sample rates, as shown in Fig. 7(b). Although, as predicted, higher order coefficients remain very low. The modeling suggests that, given a constant detection rate and circuit speed parameters, the bias should remain unchanged with varied sampling rates. To demonstrate this, the simulated data for a_1 in Fig. 7(a) are plotted once more, along with the bias, in Fig. 7(c). The length of the simulation for each data point was kept constant; therefore, the total number of generated bits varies. For this reason, the σ increases with decreased sample rate.

Dead time is introduced into the simulation, and the results are plotted in Fig. 8. The comparison is performed using the counting model presented in [39]. Increased dead time

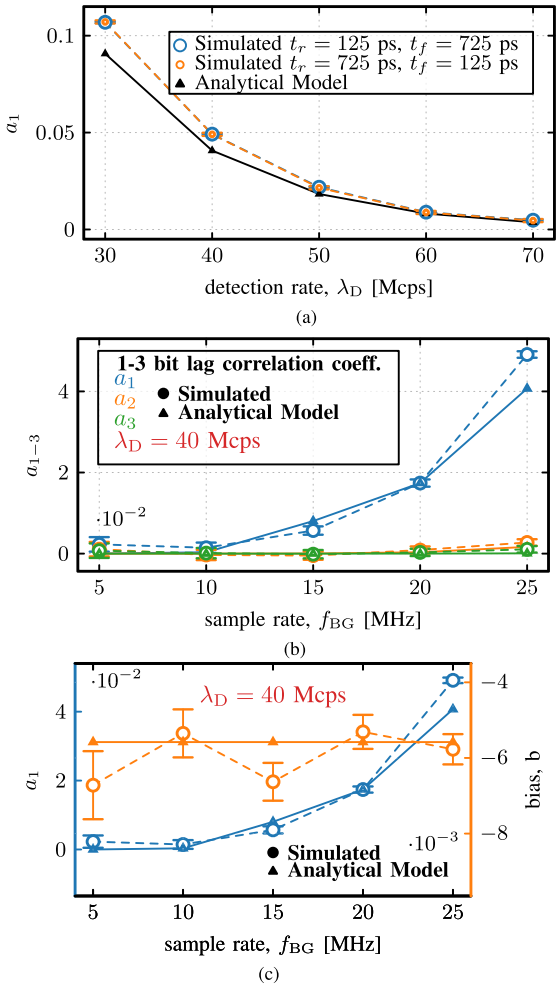


Fig. 7. Simulation using an ideal detector model to demonstrate the exponential relationship between count rate and autocorrelation. Higher order correlations are less significant as predicted by the analysis. (a) 1-bit lag autocorrelation analysis with fixed normalized sampling flip-flop threshold ($\eta = 0.475$) at $f_{BG} = 25$ MHz and swept across detection rates (λ_D). (b) Autocorrelation analyses with lags of 1–3 bits, fixed normalized sampling flip-flop threshold: $\eta = 0.475$, fixed TFF rise/fall times: $t_r = 725$ ps and $t_f = 125$ ps, and fixed detection rate: $\lambda_D = 40$ Mc/s. Higher order correlation coefficients are calculated according to (2) and (3). (c) $P = 1$ bias and 1-bit lag autocorrelation analyses with fixed normalized sampling flip-flop threshold: $\eta = 0.475$, fixed TFF rise/fall times: $t_r = 725$ ps and $t_f = 125$ ps, and fixed detection rate: $\lambda_D = 40$ Mc/s.

reduces the λ_A/f_{BG} ratio required to achieve acceptable autocorrelation, as predicted by the analytical section. Moreover, the results from simulations match closely with the proposed analytical calculations.

Some relevant system considerations can be derived from the above analysis. First, it is desirable to have a controllable, non-paralyzable dead time in the range of 5–10 ns. Under these conditions, a range of λ_A/f_{BG} ratios can be used that achieve low correlation. This aides in choosing an illumination setting that reduces activity (power consumption) but also provides margins for drift in count rates caused by environmental changes. Moreover, this is amenable to an array implementation that inevitably has some non-uniformity of count rate [44].

While, in principle, the analysis shows that higher per-pixel generation rates ($\geq$20 MHz) are achievable, the sensitivity of correlation to small changes in dead time is dramatically

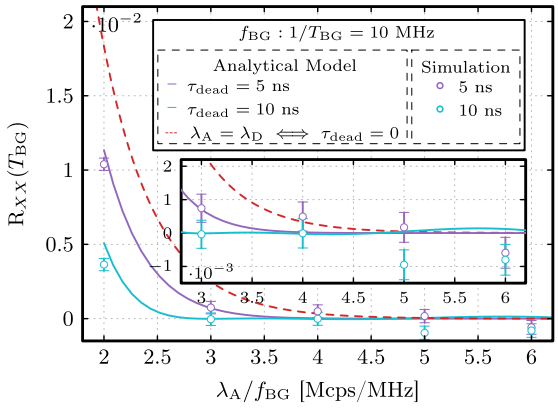


Fig. 8. Comparison between simulated and analytically calculated autocorrelation values. The analysis is performed using the non-paralyzable counting model at the τ_{dead} values of 5 and 10 ns. The ideal autocorrelation function based on a pure Poisson counting process is shown by the dashed trace.

increased. Therefore, it is desirable to keep bit generation rates lower ($\leq$10 MHz). Moreover, considering a scenario with integrated electronics, such that TFF_Q demonstrates rise and fall times on the order of 100 ps, a detection rate of 30 Mc/s would result in serial correlation and bias values $<10^{-3}$. This per-pixel generation rate is considerably higher than those demonstrated by other SPAD-based QRNG techniques [17], [17], [45]. Finally, the model could be further improved by formulating the effects of detector imperfections, in particular, those containing correlated effects, such as afterpulsing and crosstalk. Clearly, this analysis is only effective for a single QRFF; therefore, exploration of system consideration, such as PVT of the TFF, count rate/breakdown non-uniformity, metastability, comparator offset, and others, must be performed in order to have a clear view of the scalability of this circuit concept. Nevertheless, it will be shown that this model performs well in predicting the performance of individual pixels.

IV. Design of a Full-Custom CMOS QRFF

A. Pixel Design

In order to test the model presented, and take advantage of the findings from the simulation analysis, which demonstrates the ability to overcome circuit imperfections, a pixel design containing a full-custom version of the QRFF is proposed and shown in Fig. 9. Although very-high performing SPADs were recently demonstrated in the GF 55-nm BCD process [46], it is not considered a mature CMOS image sensing process, as a standard flow was used for the fabrication of this chip. Therefore, several tunable pixel functions were implemented to limit detector variability.

For the TFF, a true-single-phase clock (TSPC) logic-based circuit was implemented for enabling fast transitions, with the output buffer sized appropriately for symmetric rise/fall times. However, as previously stated, process variation will always result in some mismatch across the array. For this reason, a comparator-based sampling flip-flop is an evident choice for achieving a mean bias centered at zero, overcoming any inevitable non-symmetry. A strongARM comparator-based DFF was designed for fast latching, further enabling high-speed solution, which require serialization of many QRFFs

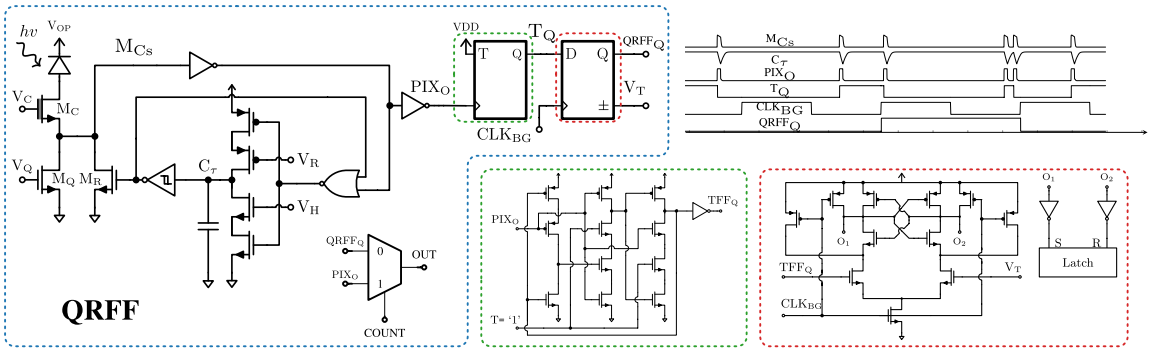


Fig. 9. Complete custom QRFF design including PQAR circuit and full custom flip-flop design for improved performance.

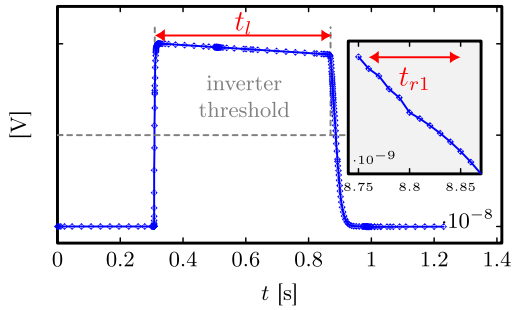


Fig. 10. Plot demonstrating the simulated pulsewidth of the SPAD anode when coupled with the pixel shown in Fig. 9. The detector is technically paralyzable in the time interval $t_l + t_{r1}$. During the hold-off interval, t_l, excess bias across the SPAD is very low (<100 mV); i.e., absorbed photons have a low probability of initiating an avalanche. t_{r1} denotes the recharge time until the inverter threshold is crossed (≃100 ps); therefore, non-paralyzable assumption is acceptable as long as flux is tuned to reduce pileup effects.

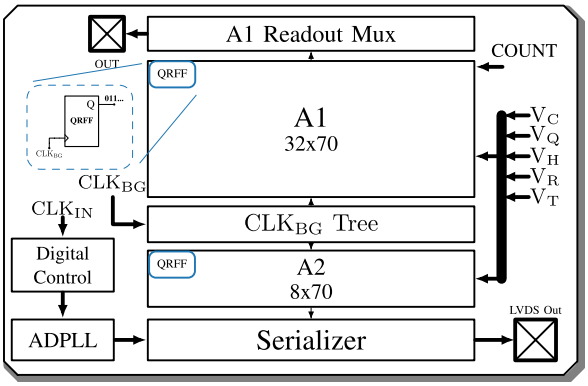


Fig. 11. FortunaSPAD full block diagram.

onto a readout bus. The sampling threshold of the DFF is controlled by a global signal V_T.

The pixel, based off of the design from [47], employs a passive-quench active-recharge (PQAR) circuit to limit afterpulsing. Despite the exclusion of an active-quenching circuit, the dead time is essentially non-paralyzable when the pixel is operated under certain conditions. The passive-quench transistor, M_Q, is designed for a high-impedance, limiting charge flow, which reduces the population of trapped carriers upon an avalanche [48], and quickly quenches the SPAD. Furthermore, when M_Q is kept off, the leakage of charge after quench (during recharge) is low. This allows for control of the hold-off time to be set using the feedback electronics. A voltage-controlled tunable delay element in the monostable feedback loop was implemented to further investigate the optimal dead time, i.e., a high count rate/afterpulsing trade-off. The hold and recharge times of the SPAD pulse are determined by the discharging and recharging time of the feedback capacitor, C_τ, which can be adjusted using the global control pins, V_H and V_R. As V_H is increased, the discharging time of C_τ decreases, thereby decreasing the length of time until M_R is turned on following an avalanche, consequently decreasing the hold time. Conversely, increasing of V_R adjusts the length of time for which M_R is on, allowing for a controllable recharge time. A simulated waveform of the anode pulse is shown in Fig. 10. After quench, the anode voltage level remains high, such that the excess bias is low, and the probability that an incident photon triggers an avalanche is also low. The recharge phase happens quickly, shortening the time when the detector is sensitive before the inverter threshold is reached, i.e., enabling essentially non-paralyzable operation. These regions are highlighted by Fig. 10.

Reduction of the SPAD bias, V_{OP}, also reduces afterpulsing. However, since the variability of breakdown voltages in this process, until this point in time, remained unexplored, it was critical to allow for a large range of excess bias values, so that all pixels in the array can be utilized. For this reason, a thick-oxide cascode transistor, M_C, was chosen, so that higher excess bias values can be used without damaging the electronics.

This complete pixel represents a realization of a QRFF, and its general functionality is described by the timing diagram in Fig. 9. Upon an avalanche detection, the SPAD becomes inactive until recharged, which is determined by the external voltage control, and the TFF is consequently toggled. With the arrival of the global bit generation clock signal, CLK_{BG}, a random bit is generated at the output, $QRFF_Q$.

V. QRNG Architecture and Characterization Setup

A sensor with 2800 total QRFF circuits, which is called the FortunaSPAD, was fabricated in the GF 55-nm BCD process with the aim of achieving multi-gigabit operation without the need for post-processing. The block diagram is shown in Fig. 11. FortunaSPAD contains two independent sub-arrays of QRFFs that can be operated simultaneously, along with readout and control circuitry. Each array has its own separate bit generation clock (CLK_{BG}). The chip micrograph and system testing infrastructure are illustrated by Fig. 12.

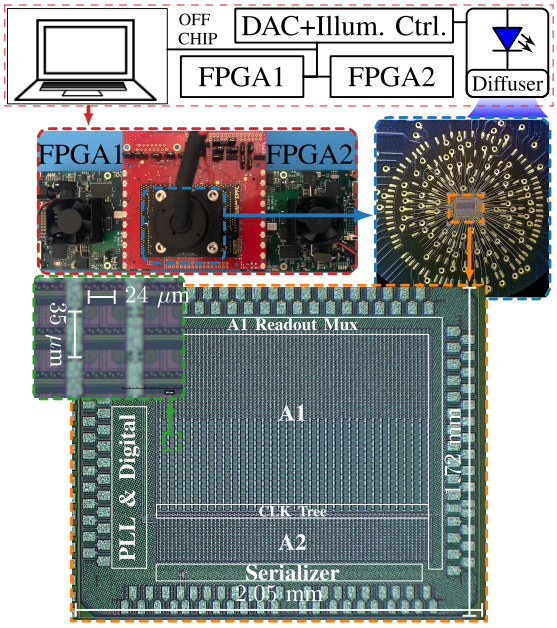


Fig. 12. FortunaSPAD micrograph with characterization setup including readout/control FPGAs. The total die area is 2.05 × 1.72 mm.

The first sub-array, denoted as $A1$, contains 70 × 32 QRFFs, which are individually read out through an output multiplexer. Furthermore, in this sub array, each individual pixel is combined with a multiplexer controlled by COUNT (Fig. 9), which can bypass the TFF/DFF circuit, allowing for monitoring of the count rate. This enables a comparison between expected results, based on the model, and measurements, along with a more quantitative method for which to decide the illumination intensity.

The second sub-array, $A2$, contains a more complex readout scheme. An on-chip digital PLL is used to operate a serializer block, which serializes 70 SPADs onto a single readout channel. Therefore, the serialization clock operates at a 70× rate compared with f_{BG}, so that data from all pixels are read out in a single f_{BG} cycle. In order to ensure that data transmitting from the chip to the FPGA is valid, the FortunaSPAD contains a control flag that, when enabled, outputs a known pattern to the FPGA. The FPGA is then able to tune the IO delays of each channel appropriately until the known pattern is received.

The two sub-arrays are read out to two separate FPGAs for firmware simplicity, although there is nothing precluding the system from using a single FPGA. The spatial readout scheme of each array is diagrammed in Fig. 13. Columnwise words are generated by $A1$ and then concatenated in the FPGA, while an entire row is serialized by $A2$. Spatial correlations of neighboring pixels are analyzed in the following sections. A motherboard containing all the required voltage generation and illumination control for the ASIC is designed, so that the entire QRNG can be operated using a USB interface. An optical tube houses the LED and a diffuser in order to provide a uniform illumination across the array while also shielding external light. The LED wavelength is 470 nm, which was chosen based on the measurements of the photon detection probability (PDP), described in the following section. The FortunaSPAD die area is 1.72 × 2.1 mm, with the horizontal and vertical pixel pitches of 24 and 35 μm, respectively.

VI. Measurement Results

A. SPAD Performance Characterization

1) Specifications: The design of the SPAD is similar to that published in [46] with the cross section shown in Fig. 14. The junction is buried deep inside the silicon using a deep p-well, buried n-well (DPW/BNW) implants. the PDP of an SPAD, which describes the probability that an incident photon will initiate an avalanche, is enhanced at longer wavelengths. While shallow junctions typically perform better in the NUV/blue region of the spectrum, proper design of the quasi-neutral region leading to the junction (PW in Fig. 14) can facilitate good PDP results at shorter wavelengths as well. Thus, generally speaking, use of a deep junction enables a larger spectrum from which to choose the illumination wavelength. The SPAD active radius is 4.4 μm, a virtual guard ring spanning 1 μm on each side, and a total radius of 6.5 μm. Pixel pitch is increased artificially to reduce crosstalk. The resulting fill factor is $\simeq$7.2%.

2) Afterpulsing: As discussed, perhaps, the most critical parameter of the SPAD is afterpulsing, as it induces correlated noise into the random bit generation circuitry. Afterpulsing occurs in SPADs when charges trapped in deep levels, during previous avalanche pulses, are then released to cause subsequent avalanches. Using the inter-arrival time histogramming technique, we estimate the afterpulsing by connecting the test pixel output to a fast 40-GS/s oscilloscope (Teledyne LeCroy WaveMaster 813 Zi-B) with an active probe and bin width of 10 ns. However, the bandwidth is limited by the maximum IO frequency ($\simeq$140 MHz). The pixel dead time was tuned to $\approx$8 ns, in order to attain accurate measurements for high-count rate applications. A low level of light was added to the measurement, to attain a count rate $\approx$1 kc/s. The results of the experiment are shown in Fig. 15. The extracted afterpulsing is $\approx$0.005%. From the histogram, it can be seen that the traps decay completely after approximately 100 ns. Both the lifetime and afterpulsing percentage are excellent results for a silicon SPAD in a deep sub-micrometer process [49], [50], which is the reasoning for the choice of this specific junction.

3) PDP: The same test pixel was used for measurement of the PDP, with results shown in Fig. 16. The data were taken using the continuous light method at 10-nm intervals up to 3-V excess bias (V_{EX}) using a setup that has been detailed in [51]. Due to the process, which was not optimized for image sensing, a clear standing wave pattern is seen across the spectrum. An LED (Cree C503B-BAN-CZ0A0452) in the blue spectrum ($\lambda = 470$ nm) is selected for the QRNG in order to avoid the efficiency troughs caused by this standing wave pattern, while maintaining a high relative detection efficiency to avoid using higher LED current.

4) DCR: The dark count rate (DCR) describes spurious SPAD avalanches in the absence of photons, degrading the signal-to-noise ratio. Measurement of DCR across all pixels in $A1$ was performed by bypassing the random flip-flop circuitry. The results are shown in Figs. 17 and 18. The DCR across all pixels remains relatively low with 95% of pixels remaining $<$10 c/s/μm^2 with only three "hot" pixels that are

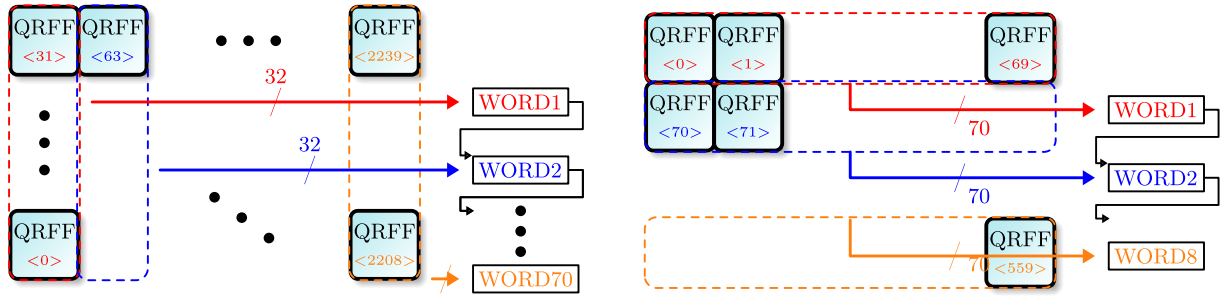


Fig. 13. Diagram demonstrating how random bits generated from individual pixels are turned into serial data. Words generated by each row/column are concatenated, and then, data from the two arrays are combined. (a) Readout scheme for $A1$. Columns are readout to generate a 32-bit word with each subsequent column concatenated. (b) Readout scheme for $A2$. Bits along one row are serialized into the 70-bit words with each subsequent row concatenated.

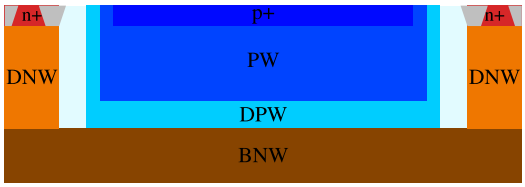


Fig. 14. Cross section of 55-nm BCD SPAD used in the FortunaSPAD. The junction is formed by the DPW–BNW interface. Deep n-well (DNW) layers are used to connect to the cathode. An additional PW is used to enhance PDP, as demonstrated and described in [46].

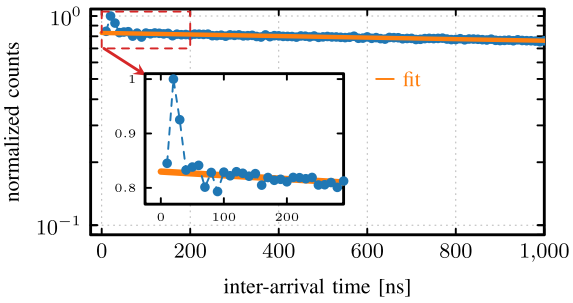


Fig. 15. Afterpulsing measurement performed at room temperature using the inter-arrival histogramming method.

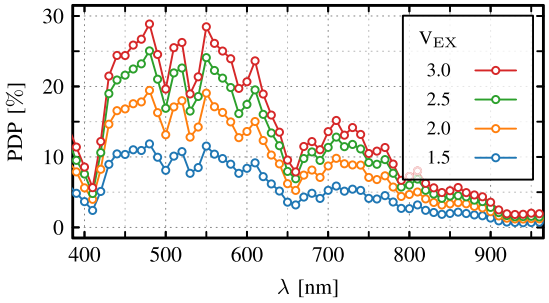


Fig. 16. PDP measured using integrated PQAR circuit at room temperature across excess bias.

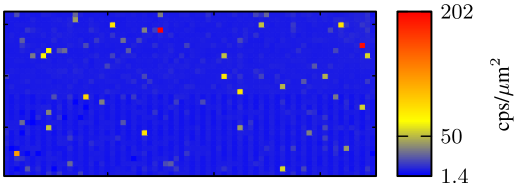


Fig. 17. Normalized DCR of each pixel in $A1$ array measured at room temperature and $V_{OP} = 34$ V.

>100 c/s/μm^2. Therefore, all QRFFs in $A1$ can be operable in the desired entropy bounds.

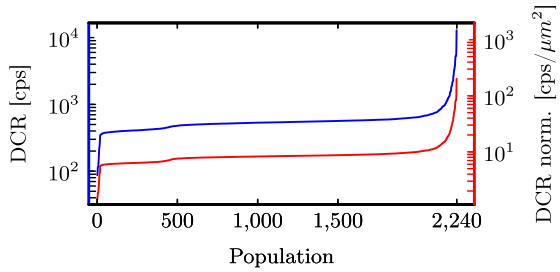


Fig. 18. DCR population for entire $A1$ array shown in c/s and normalized units at $V_{OP} = 34$ V.

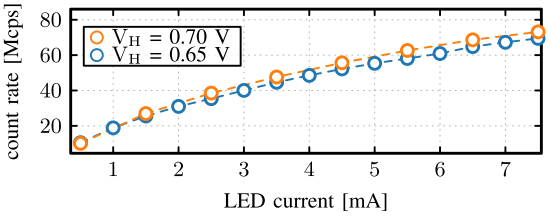


Fig. 19. Count sweep of single test pixel with swept LED current measured at room temperature.

5) Counting: As an initial validation of the model and to observe the performance capabilities of a single QRFF, the count rate is measured across swept led current, with the results shown in Fig. 19. The measurements are taken with two different control voltages for the hold time with $V_H = 0.65$ V and $V_H = 0.70$ V resulting in the dead times of $\approx$10 and $\approx$8 ns, respectively. Increasing V_H past 0.70 V, i.e., decreasing the dead time, causes the pulsewidth to shrink to a level where the count rate is not consistently measurable. Dead time can be extended to $\simeq$100 ns, although, as shown from our earlier analysis, it is advantageous to keep dead time in the 5–10-ns range to enable a variety of λ_A/f_{BG} ratios. Nevertheless, the results show counting that increases almost linearly with led current, with, perhaps, some pileup observed for $I_{LED} > 2.0$ mA at $V_H = 0.65$.

B. Comparison Between Analytical Model Values and Measured Values of a Single QRFF

The performance of a single QRFF was evaluated for comparison of bias and correlation with the expected trends described by our modeling section. Results can be seen in Fig. 20. Correlation coefficients are compared with the analytical values, since they are a function of measurable

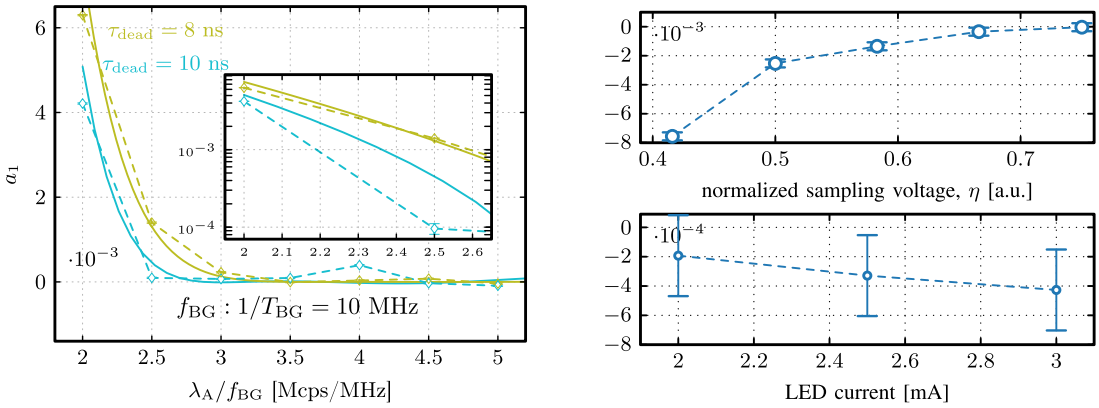


Fig. 20. Measured bias from $P(X = 1) = 0.5$ and autocorrelation results of a test pixel for comparison to expected results based on the derived analytical model. (a) Autocorrelation analysis as a function of dead time. Measurements (dashed) performed with $V_H = 0.65$ V ($\tau_{dead} \simeq 10$ ns) and $V_H = 0.7$ V ($\tau_{dead} \simeq 8$ ns) and compared with the proposed analytical model using a non-paralyzable detector. (b) Bias from $P(X = 1) = 0.5$ at $f_{BG} = 5$ MHz as a function of normalized sampling threshold and LED current. Top: performed with $I_{LED} = 2$ mA. Bottom: performed with $V_T = 0.9$ V ($\eta = 0.75$).

TABLE II
SINGLE-QRFF ENTROPY CHARACTERIZATION AT $f_{BG} = 10$ MHz AND $V_{OP} = 33.3$ V

η	I_{LED} [mA]	H_1	H_∞
0.75	2	0.999999997	0.995
0.50	2	0.999988458	0.985
0.75	3	0.999997114	0.994
0.50	3	0.999953833	0.978

qualities (λ_D and f_{BG}), for two different dead-time values. The measurements match well with the expected analytical calculations, demonstrating that very low correlation ($\leq 10^{-3}$) can be achieved with a sampling rate of $f_{BG} = 10$ MHz with $\lambda_D \simeq 25$ Mc/s, if the dead time is extended to $\simeq 10$ ns. Table II shows the entropy results for a single pixel with data sampled at $f_{BG} = 10$ MHz under two different illumination and threshold bias settings. Shannon entropy was calculated with the measured bias value, and min entropy was estimated using the SP 800-90B test suite.

The bias can be seen to scale linearly with increased illumination (count rate). This characterization is performed at lower illumination values to avoid pileup. A wider I_{LED} range is used to examine performance for the array. Moreover, we can see that the critical hypothesis regarding sampling threshold is confirmed. By adjusting the sampling threshold of the QRFF, we are able to essentially compensate bias by balancing the mismatch in the TFF output waveform. Some non-linearity is observed in the threshold correction. Furthermore, the value of threshold voltage required to balance mismatch is higher than expected. This would suggest that some of the phenomena previously mentioned, which have not been modeled, are contributing to bias. This effect's performance on the array is examined further.

C. Array Performance Characterization

1) Characterization Methodology: To demonstrate the range of operating parameter values that result in acceptable bit generation performance, the pixelwise analyses of serial correlation and bias are performed. For pixelwise characterization, a word, generated as described by Fig. 13(a) from a single column, is repeatedly generated at $f_{BG} = 5$ MHz and sorted into individual pixels. After a statistically sufficient number of bits are generated, the readout moves to the next word and repeats the process. Furthermore, spatial correlations between neighboring pixels are analyzed. This is explained in Section VI-C. Observing the results from these analyses, we can demonstrate that the FortunaSPAD generator is capable of acceptable operation across a variety of illumination and threshold voltage settings. Finally, statistical testing is done by generating a full frame of data using both arrays, in a single bit generation clock cycle, at the maximum capable speed of the readout circuitry.

2) SPAD Operating Voltage: In order to determine proper operation of the chip, the non-uniformity of breakdown voltages across the array must be understood. The V_{OP} should then be set to the minimum value of excess bias where all QRFFs are operating correctly, in order to reduce effects of afterpulsing. A method that can be used to determine this voltage is to observe the per QRFF bit bias at a constant illumination while increasing excess voltages. A visualization of the results from this test is shown in Fig. 21, where a spatial heat map of the per QRFF bias is shown. It is observed that, as the excess voltage is increased, the bit bias reaches a uniform (low) value, at a $V_{OP} = 33.3$ V, which is the operating value used for all subsequent measurements. Given the very low afterpulsing in these detectors, the FortunaSPAD could also be operated at V_{OP} values above 33.3 V, in order to have a margin for temperature drift of the SPAD breakdown voltage. Fig. 21(c) is replotted in Fig. 22 with a modified scale for properly viewing the bias distribution.

3) Bias and Correlation Analysis as a Function of Model Parameters: The root mean square (rms) and mean values of per QRFF bias, b, and serial correlation, a_1, are shown as a function of illumination intensity in Fig. 23 with $f_{BG} = 5$ MHz. From the perspective of bit bias, the rms value across the array increases with an increase in LED current, as expected, since the higher count rates scale bias proportionally. Meanwhile, it is observed that the mean bias from 2 to 3.5-mA remains constant, as a constant sampling

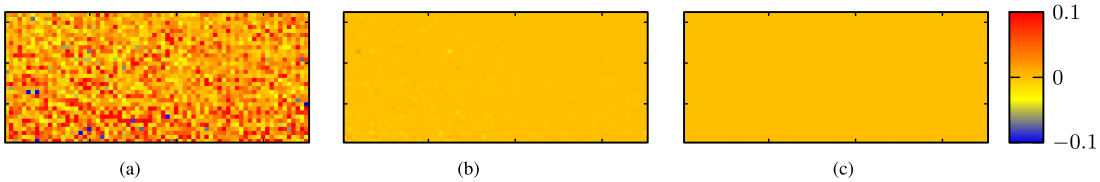


Fig. 21. Spatial bias map from $P(X = 1) = 0.5$ across excess bias with constant illumination, $I_{\mathrm{LED}} = 2$ mA, sampling rate $f_{\mathrm{BG}} = 5$ MHz, and $\eta \simeq 0.71$, i.e., $V_{\mathrm{T}} = 0.85$. These maps are shown to demonstrate the spread in breakdown voltage across the array, i.e., showing the minimum SPAD bias required to operate the chip properly. (a) $V_{\mathrm{OP}} = 32.8$ V. (b) $V_{\mathrm{OP}} = 33.1$ V. (c) $V_{\mathrm{OP}} = 33.3$ V.

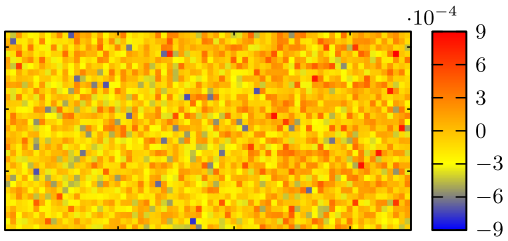


Fig. 22. Spatial map of bias from $P(X = 1) = 0.5$ shown in Fig. 21(c) with modified scale ($I_{\mathrm{LED}} = 2$ mA and $f_{\mathrm{BG}} = 5$ MHz). The results highlight how there are no patterns or spatial concentration for pixel bias across the array.

threshold (η) is maintained for all tests. A deviation, from this constant magnitude, of the mean bias between 1 and 1.5 mA, is observed. There are multiple effects, which are not taken into account by our model, such as comparator offset, metastability, and the CLK to Q transition time that can be contributing to the observed non-linearity. However, the general trend shown by the rms bias confirms that bias scales, for the majority of pixels, linearly as a function of illumination.

The sampling threshold is also swept, and in doing so, the mean bias of the entire array is shifted to zero. The results are shown in Fig. 24. Three points along the curve are also placed in a histogram to visualize the shifting of the entire array in bias, while remaining unchanged for autocorrection. At higher values for the sampling threshold, a small amount of pixels becomes stuck, as their inherent comparator offset prevents the toggling of the output. As observed for the single pixel case, the threshold voltage required to center bias across the array is relatively high $\eta \simeq 0.71$, despite TFF_Q designed with equal rise and fall times.

4) Spatial Correlations: The analysis performed so far is only valid for serial data generated from the per-pixel bases. However, spatial correlations must also be evaluated to make sure the entropy is not degraded due to crosstalk. To evaluate spatial correlations, data from $A1$ are used. We view this as a representative of $A2$ as well, since the pixel construction and pitch are identical. The cross correlation between adjacent pixels is measured by generating two full columns of data in a $f_{\mathrm{BG}} = 5$-MHz cycle at a time. Then, horizontally and vertically adjacent cross correlations are evaluated. The results are shown in Figs. 25 and 26. No evident cross correlation between pixels is present, and the magnitude is lower compared with those of serial correlations.

D. A2 Performance

For per-pixel characterization of the serialized array, a strobe signal is also implemented inside the FortunaSPAD, which is synchronized to the first QRFF output in the array. This enables spatial analysis to make sure there are no malfunctioning circuits/detectors and no particular "hot" spots in the array. As previously described, the serialization rate of this array is 70× that of the bit generation rate; therefore, f_{BG} is limited to 2 MHz. The calculated bias and correlation coefficient of all QRFFs in $A2$ at $f_{\mathrm{BG}} = 2$ MHz and $\mathrm{I}_{\mathrm{LED}} = 2$ mA are shown in Fig. 27. All QRFFs in the serialized array achieve a bias and serial correlation coefficient within the benchmark of 10^{-3}. The max calculated bias and correlation are 4.09×10^{-4} and 4.41×10^{-4}, respectively, with rms values across the array of 1.69×10^{-4} and 1.32×10^{-4}, respectively.

E. Entropy Evaluation and Overall Bit Generation Rate

Seeing as there is no evident spatial correlation present, and the mean value of bias and correlation across the array is always lower than an individual pixel, we perform entropy estimations using the worst performing pixel on chip as an estimate for entropy before testing is performed. Under an illumination setting of 2 mA, with a normalized threshold of $\eta = 0.71$ ($V_{\mathrm{T}} = 0.85$ V), all 2800 pixels achieve the correlation and bias benchmarks of 10^{-3} previously set ($H_1 \geq 0.999997$ and $H_\infty \geq 0.9986$). When per-pixel characterization is performed, for both arrays, within a normalized threshold range 0.65–0.8 ($V_{\mathrm{T}} = 0.76$–0.92 V), and an illumination range of 1.5–3 mA, the poorest performing pixel results to $b = -2.33 \times 10^{-3}$. This translates to an estimated $H_1 \simeq 0.99998$. Similarly, the worst pixelwise serial autocorrelation is $a_1 = 4.26 \times 10^{-3}$. This results to an estimated $H_\infty \simeq 0.994$. Therefore, in principle, all pixels on a single die are capable of generating 5 MHz × 2800 pixels = 14 Gb/s of high entropy data across a wide range of operating parameters. However, the limitations of the readout circuitry speed and IOs result in a combined achievable data rate of 3.3 Gb/s. Data generated from both arrays are combined and evaluated by more comprehensive statistical tests. To ensure that no spatial cross correlations affect the results of the generated bit strings, a full frame of data is read in a single $\mathrm{CLK}_{\mathrm{BG}}$ cycle for statistical testing. This translates to $f_{\mathrm{BG}} \simeq 0.8$ MHz for $A1$ and $f_{\mathrm{BG}} = 2$ MHz for $A2$. The calculated bias and correlation for the total generated string from a full frame of data are $b = -4.9 \times 10^{-5}$ and $a_1 = 4.1 \times 10^{-5}$. NIST SP 800-90B is used in the following section to test the min entropy of the generated data.

F. NIST SP 800-22 and 800-90B

The ability to achieve erroneous results from the NIST Statistical Test Suite when incorrect parameters are chosen is

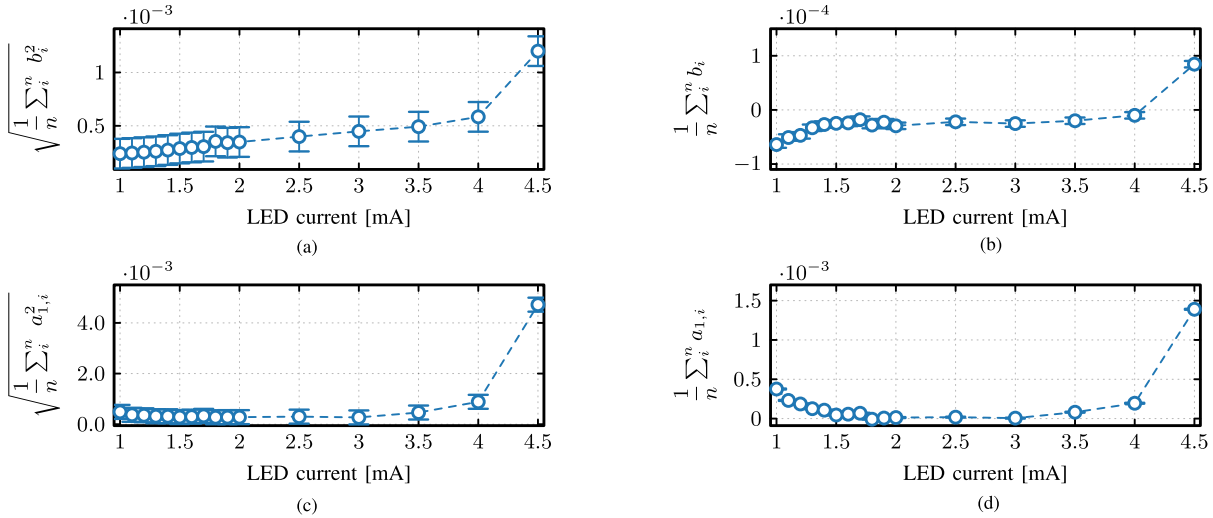


Fig. 23. Bias and correlation analysis across all QRFFs in $A1$ ($n = 2240$) as a function of LED current at a $f_{BG} = 5$ MHz. (a) RMS bias. (b) Mean bias. (c) RMS 1-bit lag autocorrelation coefficient. (d) Mean 1-bit lag autocorrelation coefficient.

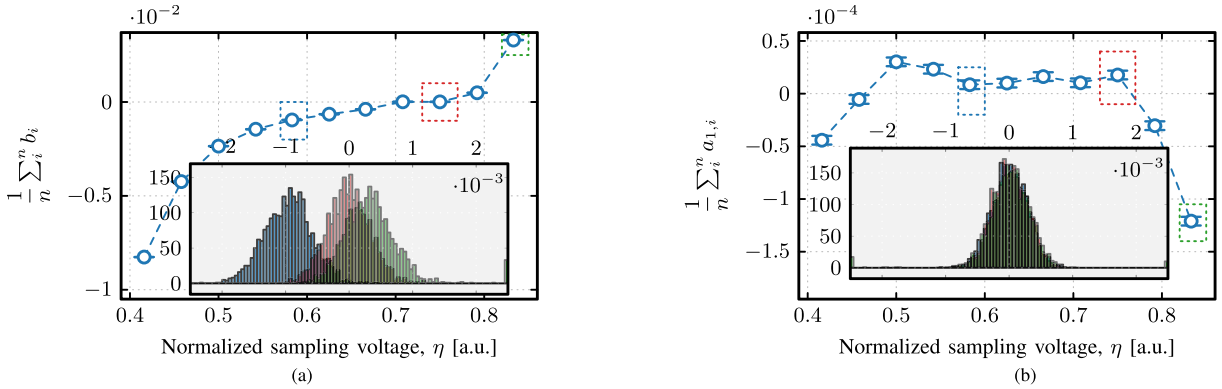


Fig. 24. Per QRFF analysis across $A1$ array at a 5-MHz bit generation rate with swept sampling threshold voltage at $I_{LED} = 2.5$-mA illumination. (a) Mean bias. (b) Mean 1-bit lag autocorrelation coefficient.

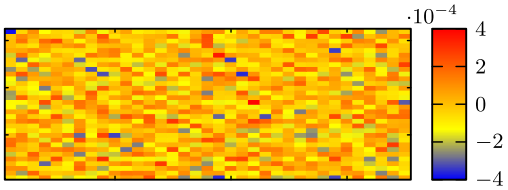


Fig. 25. Spatial map showing the calculated cross correlation value of horizontally adjacent pixels of generated bits. Columns are paired together (in 2) as a method to estimate any spatial correlations.

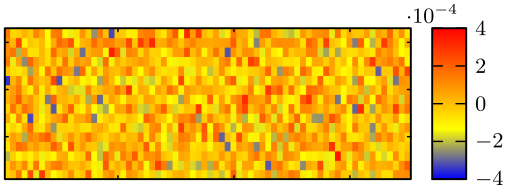


Fig. 26. Spatial map showing the calculated cross correlation value of vertically adjacent pixels of generated bits. Rows are paired together (in 2) as a method to estimate any spatial correlations.

TABLE III

SAMPLE SUMMARY OF NIST RESULTS. DATA GENERATED AT 3.3-GB/S OVERALL RATE WITH PARAMETERS: $\eta \simeq 0.71$ AND $I_{LED} = 2$ MA

Test	Min. pass rate	p-value	Pass rate
Frequency	996	0.8831	998/1000
Block frequency	996	0.0278	1000/1000
Cumulative sums	996	0.1855	997/1000
Runs	996	0.4521	999/1000
Longest run	996	0.4885	998/1000
Rank	996	0.9723	998/1000
FFT	996	0.1364	999/1000
Non overlapping template	996	0.8429	997/1000
Overlapping template	996	0.6454	998/1000
Universal	996	0.7830	1000/1000
Approximate entropy	996	0.5769	1000/1000
Random excursions	616	0.3258	618/619
Random excursions variant	616	0.5457	616/619
Serial	996	0.9737	1000/1000
Linear complexity	996	0.5523	999/1000

well documented [12], [13], [52]. Therefore, we choose strict parameters for NIST testing with 1 Gb of data generated with a full frame for both arrays, as explained above, split into 1000 bit strings using a significance level (α) of 0.001. The results for the NIST test are outlined in Table III with all tests passing. The same file was also used for testing using the 800-90B test suite. The results pass all tests (chi squared, longest repeated, and permutation), thereby confirming the random data are independent and identically distributed (i.i.d.). The estimated min entropy using MCV is $H_\infty \simeq 0.9954$.

VII. DISCUSSION AND COMPARISON

A summary of relevant integrated SPAD-based QRNGs, which include the bit generation/extraction method on chip,

TABLE IV
PUBLISHED INTEGRATED SPAD-BASED QRNGS WITH BIT GENERATION/EXTRACTION ON CHIP

Ref. & Year	Physical Principle	Circuit/Extraction Implementation	Array Size	Bitrate (per pixel)	Further Post Processing	Evaluation Method
[53] Burri, 2013.	SPAD triggering prob.	Frame readout	512×128 ×2^{α}	5 Gbps (≃ 0.04 Mbps)	Von Neumann filter	NIST STS DIEHARD
[45] Tisa, 2015.	SPAD triggering prob.	LFSR counter	32x32	200 Mbps (≃ 0.2 Mbps)	Whitening algorithm	DIEHARDER TestU01
[17] Massari, 2016.	First detected photon	Inter-arrival arbiter	16×16	128 Mbps (0.5 Mbps)	none	NIST STS
[32] Acerbi, 2018.	SPAD triggering prob.	Frame readout	1	≃ 0.1 Mbps	none	NIST STS
[18] Xu, 2018.	First detected photon	Inter-arrival arbiter	16×16	18 Mbps (≃ 0.07 Mbps)	none	NIST STS
[54] Regazzoni, 2021	Photon detection	Vector matrix multiplication†	128×128†	400 Mbps (≃ 0.024 Mbps)	none	NIST STS Diehard
[33] Massari, 2022	Photon timing statistics	Time-difference of photon arrival$^{\beta}$	16×8×2	400 kbps (≃ 1.57 kbps)	Linear corrector Elias' Integer addition	NIST SP 800-90B AIS 31
This work$^{\zeta}$	Photon timing statistics	QRFF	40×70‡	3.3 Gbps (≃ 1.2 Mbps)$^{\bowtie}$ (10 Mbps)*	none	NIST STS NIST SP 800-90B

† a fixed matrix and reconfigurable matrix are included on chip
‡ two independent arrays (70 x 32 & 70 x 8) with different readout architectures
$\bowtie$ calculated per pixel throughput with both arrays readout, limited by readout and IO speeds
* single pixel capable of $H_1 \geq 0.99997$ and $H_\infty \geq 0.98$.
$^{\alpha}$ two of the same die were used in parallel.
$^{\beta}$ a correlator circuit is used to validate that photons arrived from the emitter.
$^{\zeta}$ **Bias and correlation modelled with considerations for detector and circuit. Validated by simulation and measurement.**

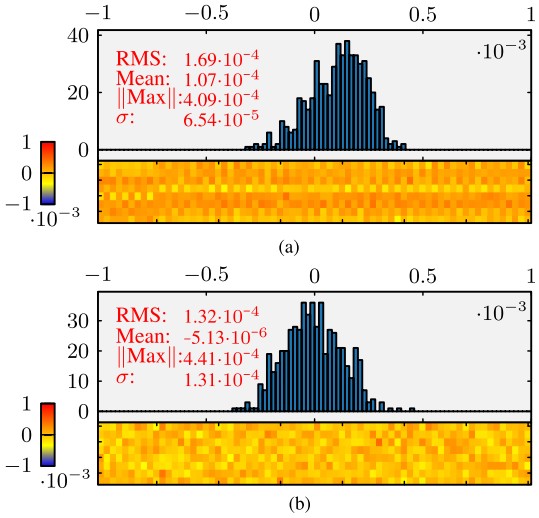


Fig. 27. $A2$ spatial maps at $f_{BG} = 2$ MHz, $I_{LED} = 2$ mA, and $\eta \simeq 0.71$, i.e., $V_T = 0.85$. (a) b. (b) a_1.

is shown in Table IV. It can be seen that for an SPAD array-based solution with bit generation on chip, we demonstrate the highest per-pixel generation rate reported. Most prior works rely either on the quantum nature of the entropy source or an arbitrarily chosen post-processing method for justification of the bit generation quality. However, in our work, we systematically model the degradation of entropy and validate it through simulation. As a result, we were able to propose a circuit innovation, which was capable of overcoming this, without the expense of a reduced generator speed, an outcome that would inevitably be the case if post-processing was employed.

VIII. CONCLUSION

We have demonstrated a full multi-Gb/s integrated SPAD-based QRNG system when using external illumination based on the QRFF method. The QRFF is an architecturally simple but feature-rich, scalable, model-testable bit generation method. By analyzing the degradation of entropy caused by circuit limitations, we were able to propose and validate a simple circuit innovation, namely, the addition of a tunable sampling threshold, in order to essentially eliminate bias from a single QRFF. This opens the door for more complex QRNG systems based on our circuit technique, that can continually monitor and correct for changes in operation caused by, for example, changes in environmental settings. A method for estimating the serial autocorrelation of pixels when detector dead time is considered was proposed and tested. Further additions to that analysis method, that take into account circuit and detector perfections, could help improve modeling at low and high flux conditions. The ability to precisely control the generator bias and correlation is interesting for certain applications, such as stochastic computing [55].

To the best of authors' knowledge, the total throughput of 3.3 Gb/s is the highest reported for a single-die SPAD-based system that also integrates its bit generation circuitry. Statistical testing was used to validate the performance of the QRNG and to estimate min entropy, resulting in a value of $H_\infty \simeq 0.9954$.

ACKNOWLEDGMENT

The authors would like to thank the members of the Beryllium Project Team and the members of the Advanced Quantum Architecture (AQUA) Laboratory, Neuchâtel, Switzerland,

who have contributed valuable input in discussions during development.

REFERENCES

[1] M. Alioto, "Trends in hardware security: From basics to ASICs," *IEEE Solid StateCircuits Mag.*, vol. 11, no. 3, pp. 56–74, Jul. 2019.

[2] S. K. Satpathy et al., "An all-digital unified physically unclonable function and true random number generator featuring self-calibrating hierarchical von Neumann extraction in 14-nm tri-gate CMOS," *IEEE J. Solid-State Circuits*, vol. 54, no. 4, pp. 1074–1085, Apr. 2019.

[3] S.-G. Bae, Y. Kim, Y. Park, and C. Kim, "3-Gb/s high-speed true random number generator using common-mode operating comparator and sampling uncertainty of d flip-flop," *IEEE J. Solid-State Circuits*, vol. 52, no. 2, pp. 605–610, Feb. 2017.

[4] V. von Kaenel and T. Takayanagi, "Dual true random number generators for cryptographic applications embedded on a 200 million device dual CPU SoC," in *Proc. IEEE Custom Integr. Circuits Conf.*, Jan. 2007, pp. 269–272.

[5] S. Larimian, M. R. Mahmoodi, and D. B. Strukov, "Lightweight integrated design of PUF and TRNG security primitives based on eFlash memory in 55-nm CMOS," *IEEE Trans. Electron Devices*, vol. 67, no. 4, pp. 1586–1592, Apr. 2020.

[6] S. Suhail, R. Hussain, A. Khan, and C. S. Hong, "On the role of hash-based signatures in quantum-safe Internet of Things: Current solutions and future directions," *IEEE Internet Things J.*, vol. 8, no. 1, pp. 1–17, Jan. 2021.

[7] C. Cheng, R. Lu, A. Petzoldt, and T. Takagi, "Securing the Internet of Things in a quantum world," *IEEE Commun. Mag.*, vol. 55, no. 2, pp. 116–120, Feb. 2017.

[8] W. Killmann and W. Schindler, "A proposal for: Functionality classes for random number generators," DRAFT, Bundesamt für Sicherheit in der Informationstechnik (BSI), Bonn, Germany, Tech. Rep., Version V2.35 DRAFT, 2022.

[9] E. Barker, J. Kelsey, K. McKay, A. Roginsky, and M. S. Turan, "Recommendation for random bit generator (RBG) constructions," Special Publication, 3rd Draft, NIST, Gaithersburg, MD, USA, Tech. Rep. 800 90C, 2022.

[10] X. Ma, F. Xu, H. Xu, X. Tan, B. Qi, and H.-K. Lo, "Postprocessing for quantum random-number generators: Entropy evaluation and randomness extraction," *Phys. Rev. A, Gen. Phys.*, vol. 87, no. 6, Jun. 2013, Art. no. 062327.

[11] V. Rožic and I. Verbauwhede, "Hardware-efficient post-processing architectures for true random number generators," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 66, no. 7, pp. 1242–1246, Jul. 2019.

[12] D. Hurley-Smith and J. Hernandez-Castro, "Quantum leap and crash: Searching and finding bias in quantum random number generators," *ACM Trans. Privacy Secur.*, vol. 23, no. 3, pp. 1–25, Aug. 2020.

[13] M. Sys at al, "On the interpretation of results from the NIST statistical test suite," *Romanian J. Inf. Sci. Technol.*, vol. 18, no. 1, pp. 18–32, 2015.

[14] A. Vassilev and R. Staples, "Entropy as a service: Unlocking cryptography's full potential," *Computer*, vol. 49, no. 9, pp. 98–102, Sep. 2016.

[15] P. Sharma, A. Agrawal, V. Bhatia, S. Prakash, and A. K. Mishra, "Quantum key distribution secured optical networks: A survey," *IEEE Open J. Commun. Soc.*, vol. 2, pp. 2049–2083, 2021.

[16] A. Tontini, L. Gasparini, N. Massari, and R. Passerone, "SPAD-based quantum random number generator with an N^{th}-order rank algorithm on FPGA," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 66, no. 12, pp. 2067–2071, Dec. 2019.

[17] N. Massari et al., "16.3 A 16×16 pixels SPAD-based 128-Mb/s quantum random number generator with −74dB light rejection ratio and −6.7ppm/°C bias sensitivity on temperature," in *IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, San Francisco, CA, USA, Jan. 2016, pp. 292–293.

[18] H. Xu, D. Perenzoni, A. Tomasi, and N. Massari, "A 16 × 16pixel post-processing free quantum random number generator based on SPADs," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 65, no. 5, pp. 627–631, May 2018.

[19] T. Jennewein, U. Achleitner, G. Weihs, H. Weinfurter, and A. Zeilinger, "A fast and compact quantum random number generator," *Rev. Sci. Instrum.*, vol. 71, no. 4, pp. 1675–1680, Apr. 2000.

[20] H. Zhou, J. Li, W. Zhang, and G.-L. Long, "Quantum random-number generator based on tunneling effects in a Si diode," *Phys. Rev. Appl.*, vol. 11, no. 3, Mar. 2019, Art. no. 034060.

[21] B. Qi, Y. M. Chi, H.-K. Lo, and L. Qian, "High-speed quantum random number generation by measuring phase noise of a single-mode laser," *Opt. Lett.*, vol. 35, no. 3, pp. 312–314, 2010.

[22] Y. Liu et al., "Device-independent quantum random-number generation," *Nature*, vol. 562, no. 7728, pp. 548–551, 2018.

[23] Z. Cao, H. Zhou, X. Yuan, and X. Ma, "Source-independent quantum random number generation," *Phys. Rev. X*, vol. 6, no. 1, Feb. 2016, Art. no. 011020.

[24] Z. Cao, H. Zhou, and X. Ma, "Loss-tolerant measurement-device-independent quantum random number generation," *New J. Phys.*, vol. 17, no. 12, Dec. 2015, Art. no. 125011.

[25] Y. Liu et al., "High-speed device-independent quantum random number generation without a detection loophole," *Phys. Rev. Lett.*, vol. 120, no. 1, Jan. 2018, Art. no. 010503.

[26] D. Rusca et al., "Self-testing quantum random-number generator based on an energy bound," *Phys. Rev. A, Gen. Phys.*, vol. 100, no. 6, Dec. 2019, Art. no. 062338.

[27] T. Lunghi et al., "Self-testing quantum random number generator," *Phys. Rev. Lett.*, vol. 114, no. 15, Apr. 2015, Art. no. 150501.

[28] K. Morimoto et al., "Megapixel time-gated SPAD image sensor for 2D and 3D imaging applications," *Optica*, vol. 7, no. 4, pp. 346–354, Apr. 2020.

[29] K. Morimoto et al., "3.2 megapixel 3D-stacked charge focusing SPAD for low-light imaging and depth sensing," in *IEDM Tech. Dig.*, Dec. 2021, p. 20.

[30] M. Stipčević, "Quantum random flip-flop and its applications in random frequency synthesis and true random number generation," *Rev. Sci. Instrum.*, vol. 87, no. 3, Mar. 2016, Art. no. 035113.

[31] M. Stipcevic, I. M. Antolovic, C. Bruschini, and E. Charbon, "Scalable quantum random number generator for cryptography based on the random flip-flop approach," 2021, *arXiv:2102.12204*.

[32] F. Acerbi, Z. Bisadi, G. Fontana, N. Zorzi, C. Piemonte, and L. Pavesi, "A robust quantum random number generator based on an integrated emitter-photodetector structure," *IEEE J. Sel. Topics Quantum Electron.*, vol. 24, no. 6, pp. 1–7, Nov. 2018.

[33] N. Massari et al., "A monolithic SPAD-based random number generator for cryptographic application," in *Proc. IEEE 48th Eur. Solid State Circuits Conf. (ESSCIRC)*, Sep. 2022, pp. 73–76.

[34] 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. Instrum.*, vol. 86, no. 7, Jul. 2015, Art. no. 073113.

[35] A. Dervić, N. Tadić, H. Mahmoudi, B. Goll, M. Hofbauer, and H. Zimmermann, "Single-pixel postprocessing-free 5 mbps quantum random number generator using a single-photon avalanche diode detector and a T/(T-t) pulse-shaped laser driver," *Opt. Eng.*, vol. 59, no. 12, Dec. 2020, Art. no. 127105.

[36] E. Sarbazi, M. Safari, and H. Haas, "Statistical modeling of single-photon avalanche diode receivers for optical wireless communications," *IEEE Trans. Commun.*, vol. 66, no. 9, pp. 4043–4058, Sep. 2018.

[37] S. M. Ross, *Introduction to Probability Models*. New York, NY, USA: Academic, 2014.

[38] E. Sarbazi, M. Safari, and H. Haas, "The impact of long dead time on the photocount distribution of SPAD receivers," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2018, pp. 1–6.

[39] I. Straka, J. Grygar, J. Hlousek, and M. Jezek, "Counting statistics of actively quenched SPADs under continuous illumination," *J. Lightw. Technol.*, vol. 38, no. 17, pp. 4765–4771, Sep. 1, 2020.

[40] A. Eisele et al., "185 mhz count rate 139 db dynamic range single-photon avalanche diode with active quenching circuit in 130 nm CMOS technology," in *Proc. Int. Image Sensor Workshop*, 2011, pp. 278–280.

[41] D. Johnston, "Random number generators-principles and practices," in *Random Number Generators-Principles and Practices*. Berlin, Germany: de Gruyter Press, 2018.

[42] M. S. Turan, E. Barker, J. Kelsey, K. McKay, M. Baish, and M. Boyle, "Recommendation for the entropy sources used for random bit generation," *NIST Special Publication*, vol. 800, no. 90B, p. 102, Jan. 2018.

[43] L. E. Bassham III et al., "Sp 800-22 REV. 1a. A statistical test suite for random and pseudorandom number generators for cryptographic applications," Nat. Inst. Standards Technol., Gaithersburg, MD, USA, Tech. Rep. SP 800-22, Version V1a, 2010.

[44] I. M. Antolovic, S. Burri, C. Bruschini, R. Hoebe, and E. Charbon, "Nonuniformity analysis of a 65-kpixel CMOS SPAD imager," *IEEE Trans. Electron Devices*, vol. 63, no. 1, pp. 57–64, Jan. 2016.

[45] 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 2015.

[46] F. Gramuglia et al., "Engineering breakdown probability profile for PDP and DCR optimization in a SPAD fabricated in a standard 55 nm BCD process," *IEEE J. Sel. Topics Quantum Electron.*, vol. 28, no. 2, pp. 1–10, Mar. 2022.

[47] P. Keshavarzian et al., "Low-noise high-dynamic-range single-photon avalanche diodes with integrated PQAR circuit in a standard 55nm BCD process," in *Proc. SPIE*, vol. 12089, May 2022, pp. 73–82.

[48] D. Bronzi, S. Tisa, F. Villa, S. Bellisai, A. Tosi, and F. Zappa, "Fast sensing and quenching of CMOS SPADs for minimal afterpulsing effects," *IEEE Photon. Technol. Lett.*, vol. 25, no. 8, pp. 776–779, Apr. 2013.

[49] S. Pellegrini et al., "Industrialised SPAD in 40 nm technology," in *IEDM Tech. Dig.*, Dec. 2017, p. 16.

[50] M. Sanzaro, P. Gattari, F. Villa, A. Tosi, G. Croce, and F. Zappa, "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, pp. 1–9, Mar. 2018.

[51] F. Gramuglia, M.-L. Wu, C. Bruschini, M.-J. Lee, and E. Charbon, "A low-noise CMOS SPAD pixel with 12.1 Ps SPTR and 3 Ns dead time," *IEEE J. Sel. Topics Quantum Electron.*, vol. 28, no. 2, Mar. 2022, Art. no. 3800809.

[52] A. L. Rukhin, "Statistical testing of randomness: New and old procedures," in *Randomness Through Computation*, 1st ed., H. Zenil, Ed. Singapore: World Scientific, 2011, pp. 160–174.

[53] S. Burri, D. Stucki, Y. Maruyama, C. Bruschini, E. Charbon, and F. Regazzoni, "Jailbreak imagers: Transforming a single-photonimage sensor into a true random numbergenerator," in *Proc. Int. Image Sensors Works. (IISW)*, Snowbird, UT, USA, Jun. 2013, pp. 1–4.

[54] F. Regazzoni, E. Amri, S. Burri, D. Rusca, H. Zbinden, and E. Charbon, "A high speed integrated quantum random number generator with on-chip real-time randomness extraction," 2021, *arXiv:2102.06238*.

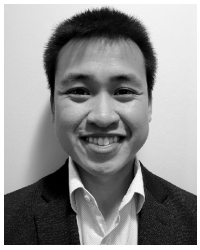
[55] A. Alaghi, W. Qian, and J. P. Hayes, "The promise and challenge of stochastic computing," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 37, no. 8, pp. 1515–1531, Aug. 2018.

Pouyan Keshavarzian (Graduate Student Member, IEEE) received the B.Sc. and M.Sc. degrees in electrical engineering from the University of Calgary, Calgary, AB, Canada, in 2015 and 2019, respectively. He is currently pursuing the Ph.D. degree with the Advanced Quantum Architecture (AQUA) Laboratory, École Polytechnique Fédérale de Lausanne, Neuchâtel, Switzerland.

From 2013 to 2016, he had various roles in hardware engineering with Garmin, Cochrane, AB, Canada, where he was involved in the design of low-power wireless sensors for sport and health monitoring applications. For his M.Sc. research, he focused on developing microwave backscattering circuits and systems for radar applications. His current research interests include the statistical modeling and development of circuits and detectors for single-photon avalanche diode (SPAD)-based quantum random number generators.

Karthick Ramu received the B.Eng. degree in electronics and communication engineering from the University of Madras, Chennai, India, and the M.S. degree in electrical and computer engineering from the Illinois Institute of Technology, Chicago, IL, USA.

In 2020, he joined Qrypt Inc., New York, NY, USA, where he is currently a Senior FPGA Engineer and also the FPGA Lead for Qrypt's PCIe-based quantum random number generator (RNG) cards. His research interests include high-speed and low-latency FPGA designs and hardware implementation of cryptography.



Duy Tang received the B.S. degree in computer engineering from the University of Maryland, College Park, MD, USA, in 2017.

From 2017 to 2021, he had various roles in FPGA engineering designing applications for wireless communication, high-performance computing, and high-speed packet processing. Since 2021, he has been with Qrypt Inc., New York, NY, USA, where he is part of the embedded systems and the FPGA Team. His scientific interests include high-speed FPGA designs and quantum random number generators.

Carlos Weill is a senior engineer with over 30 years of experience leading the development of a variety of hardware projects. He has brought to market both wired and wireless communication products, including ADSL modems, that were manufactured in the hundreds of thousands. Since 2019, he has been with Qrypt Inc., New York, NY, USA, where he has focusing on the development and integration of quantum random number generators for use in Qrypt's entropy-as-a-service solution.

Francesco Gramuglia (Member, IEEE) received the B.S. degree in biomedical engineering (major in electronics) and the M.S. degree in electronics engineering from the Politecnico di Milano, Milan, Italy, in 2013 and 2016, respectively, and the Ph.D. degree in microelectronics from the Advanced Quantum Architecture (AQUA) Laboratory, École Polytechnique Fédérale de Lausanne (EPFL), Neuchâtel, Switzerland, in 2022.

From 2015 to 2016, he was a trainee with the STORMLab, Vanderbilt University, Nashville, TN, USA, where he collaborated on the development of robotic endoscope platforms. His research interests include the design of deep-submicrometer Si-single-photon avalanche diode (SPAD) and SPAD-based sensors in standard CMOS and bipolar–CMOS–DMOS (BCD) technologies, and system development for use in several applications based on single-photon detection, e.g., time-of-flight positron emission tomography (TOFPET).

Dr. Gramuglia received the Best Application Prize and Overall Winner at Surgical Robot Challenge during the Hamlyn Symposium on Medical Robotics at Imperial College, London, in 2016. In 2021, he received the First Place Nuclear Science Symposium (NSS) Student Paper Award.

Shyue Seng (Jason) Tan received the B.Eng. (Hons.) and Ph.D. degrees in electrical and electronics engineering from Nanyang Technological University, Singapore, in 2001 and 2004, respectively.

He is currently a PMTS (Deputy Director) working with the Technology Development Department, GLOBALFOUNDRIES Singapore Pte. Ltd., Singapore, where he is leading a group of engineers working on the CMOS logic, HV, non-volatile memory (NVM), HBT device design, and development. He has authored or coauthored more than 30 journals and conference papers that include one invited paper. He holds 123 U.S. patents and more are in the progress of filing/searching. His research interests include semiconductor device physics and reliability physics; and biosensor, CIS, and single-photon avalanche diode (SPAD) development.

Dr. Tan served as a Peer Reviewer for International Journal Papers, such as the IEEE TRANSACTIONS OF ELECTRON DEVICES (T-ED), *Electron Device Letters* (EDL), *Journal of Electrochemical Society* (JES), *Applied Physics Letter* (APL), and *Journal of Applied Physics* (JAP).

Michelle Tng received the B.Eng. degree in electrical and electronic engineering from Nanyang Technological University, Singapore, in 2012.

Since 2012, she has been with GLOBALFOUNDRIES Singapore Pte. Ltd., Singapore, where she is a part of the Global TCAD Team. Her research interests include semiconductor device modeling and single-photon avalanche diodes.

Louis Lim received the B.E. degree in electrical and electronics from Nanyang Technological University, Singapore, in 1995.

He is currently with GLOBALFOUNDRIES Singapore Pte. Ltd., Singapore, where he works on logic, non-volatile memory (NVM) technology, next generation single-photon avalanche diode (SPAD) integration, and development of 2-D and 3-D image sensor technologies. As the development manager, he has more than 20 years of experience in semiconductor wafer fabrication. His work focus on developing and optimizing the performance of logic, NVM devices, and single-photon avalanche diode (SPAD) sensor; improve the process yield and reliability; and bring in new prototype to qualification and production. He has authored/coauthored more than five U.S. patent.

Elgin Quek (Member, IEEE) received the B.Eng. degree (Hons.) in electrical engineering from the National University of Singapore, Singapore, and the M.S. degree in electrical engineering from Stanford University, Stanford, CA, USA.

From 1988 to 2009, he was with Chartered Semiconductor, Singapore, where he worked on process integration, yield enhancement, device engineering, and SPICE modeling for CMOS and floating gate memories. Since 2009, he has been with GLOBALFOUNDRIES Singapore Pte. Ltd., Singapore, where he is currently a GF Senior Fellow in technology development responsible for device design for CMOS-based logic, SRAM, non-volatile memory, display driver, and sensor technologies. He has coauthored more than 40 technical papers and holds more than 150 U.S. patents.

Denis Mandich received the B.A. and M.Sc. degrees in theoretical physics from Rutgers University, New Brunswick, NJ, USA.

He has authored several cryptographic hardware and software patents.

Prof. Mandich is a Founding Member of the Quantum Economic Development Consortium (QED-C), the Mid-Atlantic Quantum Alliance (MQA), the ANSI Accredited Standards Committee X9, the ITU Telecommunications Standardization Sector (ITU-T), Cloud Security Alliance, and Forbes Technology Council. He is a 20-year veteran of the U.S. Intelligence Community, where he worked on classified technologies essential to national security.

Mario Stipčević started as a nuclear and particle fields scientist in 1991 working on CERN experiments NOMAD and ATLAS. In 1994, he defended Ph.D. thesis in experimental high energy physics on ATLAS at L'Universite de Savoie, Chambery, France. In continuation, he worked on CERN's NOMAD, NOMAD-STAR, and OPERA experiments as the leader of the Zagreb group. He is currently a Senior Scientific Associate with the Rudjer Boskovic Institute (RBI), Zagreb, Croatia. Since 2004, his interests turned to wards quantum information. Since 2014, he has been the Head of the Photonics and Quantum Optics Research Unit, Centre of Excellence for Advanced Materials and Sensing Devices, RBI. He has authored over 100 scientific articles in CC journals cited over 3500 times, 12 invited conference talks, 17 popular articles in the field of electronics, and three granted patents. His research interests include quantum communication, information, entanglement and optics, bioinspired random pulse computer, holography, and neutrino physics.

Dr. Stipčević has been an Editorial Board Member of Nature's Scientific Reports since 2017. He obtained a Fulbright Scholar at the University of California at Santa Barbara (UCSB) in 2011, followed by one-year sabbatical leave at UCSB and Duke University, working on a high-speed quantum cryptography.

Edoardo Charbon (Fellow, IEEE) received the Diploma degree from ETH Zürich, Zürich, Switzerland, in 1988, the M.S. degree from the University of California at San Diego, La Jolla, CA, USA, in 1991, and the Ph.D. degree from the University of California at Berkeley, Berkeley, CA, USA, in 1995, all in electrical engineering and EECS.

He has consulted with numerous organizations, including Bosch, X-Fab, Texas Instruments, Maxim, Sony, Agilent, and the Carlyle Group. He was with Cadence Design Systems from 1995 to 2000, where he was an Architect of the company's initiative on information hiding for intellectual property protection. In 2000, he joined Canesta Inc., as a Chief Architect, where he led the development of wireless 3-D CMOS image sensors. From 2008 to 2016, he was the Chair of VLSI design with the Delft University of Technology, Delft, The Netherlands. He has been the driving force behind the creation of deep-submicrometer CMOS single-photon avalanche diode (SPAD) technology, which has been mass-produced since 2015 and is present in telemeters, proximity sensors, and medical diagnostics tools. Since 2002, he has been a Faculty Member of Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland. He has authored or coauthored over 400 papers and two books, and he holds 24 patents. His interests span from 3-D vision, LiDAR, FLIM, FCS, and NIROT to super-resolution microscopy, time-resolved Raman spectroscopy, and cryo-CMOS circuits and systems for the control of qubit arrays in quantum computers.

Dr. Charbon is a fellow of the Kavli Institute of Nanoscience Delft. He is a Distinguished Visiting Scholar of the W. M. Keck Institute for Space at Caltech and a Distinguished Lecturer of the IEEE Photonics Society.