100 Gbit/s Secure and Real-Time Vacuum Fluctuation Based Quantum Random Number Generator

Axl Bomhals[1,*], Cédric Bruynsteen[1], Caro Meysmans[1], Achim Vandierendonck[1] and Xin Yin[1]

***Abstract*—True randomness is an essential requirement in numerous classical and quantum cryptographic systems and applications to guarantee secure communication, necessitating the research and design of promising quantum random number generators. As communication systems become more randomness consuming, current generators fail to deliver the required throughput mainly due to slow and often delayed post-processing. To address these performance issues, this work presents an FPGA-based quantum random number generator based on our previous work that uses a low optical power, custom-designed integrated vacuum fluctuation source. An adaptive, efficient and fast Toeplitz hashing scheme is developed with a configurable security factor, making it suitable for a wide range of applications. The high clearance and large bandwidth of the low-noise photonic-electronic detector in combination with hardware accelerated processing makes it possible for the presented dual channel generator to deliver a record generation rate of 100.3 Gbit/s for secure usage in real-time. Due to the small form factor, the design shows promise to be fully integrated into a custom-packaged device.**

***Index Terms*—Quantum, Random number generation, Optical noise, Optical receiver, Field programmable gate arrays, Digital signal processing, Real time systems, Cryptography**

I. Introduction

RANDOM numbers are an important aspect in many applications [1] ranging from entertainment such as computer games or games of chance to statistical simulations [2], but most importantly cryptography [3]. In cryptography, random numbers are used to generate keys, encoding the data and making it secure against unauthorized parties. Currently, many active encryption systems rely on pseudo-random number generators (PRNGs) to create the keys, where a seemingly random sequence is generated through an algorithm using simple arithmetic operations [4]. Due to the deterministic nature of computers, the resulting numbers are predictable if the start state and algorithm are known. This creates a vulnerability in the security of our communication systems, especially with the technological advances of machine learning and AI [5]. For this reason, there is a need for true random number generators [6] (TRNGs) that derive their numbers primarily from physical events that are fundamentally random. Quantum random number generators (QRNGs), a specialized subset of TRNGs, capitalize on inherently random quantum mechanical phenomena [7]. Some of the quantum sources from which randomness is extracted are single photons in different modes [8], laser phase noise [9] or vacuum fluctuation [10]–[12]. In this work, vacuum fluctuations are used to generate random numbers due to the source's small size, low complexity, and high achievable bit rates.

The continuous advancement of quantum encryption schemes necessitates the development of fast TRNGs that can meet the growing data rate demands, preferably through low-cost, miniaturized integrated devices. The best-known application of quantum cryptography is quantum key distribution (QKD), where unconditionally secure links are established [13]. As this link is only as strong as its weakest component, it is of vital importance that the random numbers used in these systems are truly random, making QRNGs an essential component. State-of-the-art QKD systems have pushed the need for high-speed true random number generation to its limits due to the high throughput required to implement practical systems. Looking at a recent publication [14], an estimation of the random numbers required to operate these systems can be determined based on the modulation formats and symbol-rates employed. Recently, Pi *et al.* [15] published a continuous variable QKD system with Gaussian modulation in which both channels of the IQ modulator are driven by an 8-bit digital-to-analog converter (DAC), to generate coherent states with a 1 GHz repetition rate. This system would require random numbers at a rate of 16 Gbit/s to operate in real time.

Hajomer *et al.* demonstrated continuous-variable QKD with discrete modulation, transmitting 64-QAM at 8 Gbaud [16] and recently 16-QAM at 16 Gbaud [17], requiring random number rates of 48 and 64 Gbit/s, respectively.

As these systems are expected to become more widespread due to the threat of quantum computers to traditional communication security [18] together with an increase in investments [19], their speed and consumption of random numbers are bound to further increase.

Considerable research has been dedicated to developing measurement apparatus based on high-entropy sources capable of potentially delivering RNGs with exceptionally high random number rates [12], [20]–[22]. While these sources show promise, reported speeds are often extrapolated from offline post-processing. To make these random number generators field-ready, real-time post-processing must be implemented. Only recently have real-time QRNGs with high speeds been reported [23]–[27]. For instance, a photonic integrated chip-based QRNG, with subsequent parallel post-processing in a field-programmable gate array (FPGA), demonstrated a real-time output rate of up to 18.8 Gbit/s [23]. Another study achieved an aggregated generation rate of 50 Gbit/s using a spatially multiplexed scheme [26]. However, this setup relies on an entropy measurement device built with discrete photonic

[1] IDLab, Department of Information Technology, imec-Ghent University, 9052 Ghent, Belgium

[*] axl.bomhals@ugent.be

Published in IEEE Journal of Lightwave Technology, vol. 44, no. 16, pp. 7182–7190, Aug. 15, 2026.

DOI: 10.1109/JLT.2026.3665249.

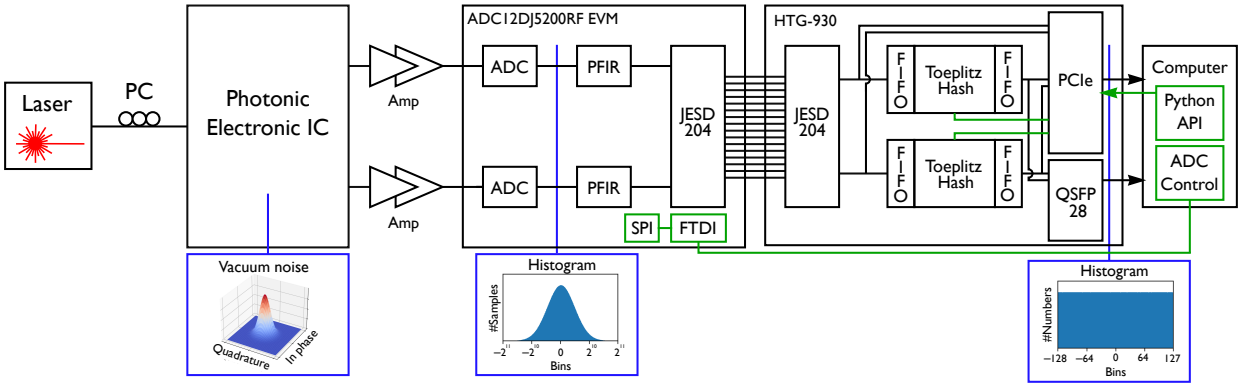

Fig. 1. A block diagram of the QRNG implementation and experimental setup containing the photonic-electronic IC, ADC evaluation board with equalization using programmable finite impulse response (PFIR) filters and a HTG-930 FPGA with the submodules and routing of the post-processing.

and electronic components, requiring a higher number of multiplexed channels and a high-power laser with >180 mW output power to achieve this rate.

In this paper, we introduce a vacuum fluctuations based QRNG with a record-fast real-time output rate of 100.3 Gbit/s using post-processing on an FPGA, a significant improvement over our previous work [27]. Our system utilizes a quantum entropy source based on a dual-channel photonic-electronic integrated detector. The high integration level of the photonic integrated circuit (PIC) and the electric integrated circuit (EIC) results in a compact footprint for the entropy measurement device, which operates with an optical power of 18 mW. Additionally, we present a comprehensive characterization of the acquisition device, incorporating an advanced device-dependent security proof [12], ensuring the true randomness of the generator. The output random numbers have also been validated using the NIST [28] and Dieharder [29] test suites.

II. System Implementation

Our practical implementation of a vacuum fluctuations based QRNG consists of two main parts: the photonic-electronic front-end that measures the vacuum fluctuations, and the post-processing that extracts the randomness. The goal of the post-processing step is to remove the effect of side-channels and to convert the raw samples to a uniform distribution of bits in real-time. In practical systems, side channels arise from the interaction of quantum effects with classical noise. It is required to properly model the underlying quantum process such that side-channel information can be determined and removed from the quantum randomness [30]. An advanced device-dependent security proof for vacuum fluctuations [12] is used to find out the achievable quantum randomness of our system. The entire QRNG processing chain is illustrated in Fig. 1.

A. Photonic-Electronic front-end

The vacuum state $|0\rangle$ is an elementary minimum-uncertainty quantum state of light and contains non-zero energy, contrary to the classical interpretation of electromagnetics. The electric field is constantly fluctuating and is considered to be omnipresent [31]. Observing the electric field quadratures of the vacuum state reveals that the amplitude and phase cannot be known simultaneously with arbitrary precision. It is represented in phase-space by a two-dimensional Gaussian distribution centered around the origin.

These fluctuations can be captured and amplified using a balanced homodyne detector (BHD) [10]. In this work, we use a dual-channel integrated low noise balanced homodyne detector [16] (Fig. 2). 50/50 beam splitters (BSs) are used to apply a local oscillator (LO) laser and vacuum fluctuations to two channels after which they are each fed into a BHD structure. This BHD contains a Mach-Zehnder interferometer (MZI) and a pair of photodiodes with a 1.1 A/W responsivity, with the MZIs used to balance the coupling of the optical signal in the photodiodes such that the common mode rejection ratio (CMRR) of the detector is maximized. The differential current of the photo-diodes is converted to a voltage by the low noise transimpedance amplifiers (TIAs) [32] and converted to digital samples using 12-bit 5.2 GSa/s analog-to-digital converters (ADCs). The digitized data is further processed on an FPGA.

Similar measurements as in work [32] were performed on the output of the TIA of the photonic-electronic integrated circuit using the Keysight N9020A MXA signal analyzer to determine the detector's performance before being included in the QRNG device. For these measurements, an optical loss of approximately 3.4 dB was measured between the laser module and the photo-diodes of the balanced detector, as the use of 20 mW LO power resulted in a photocurrent of 2.5 mA flowing through the photo-diodes.

The detector [16], shows a high shot-noise to electrical noise clearance larger than 28 dB for both channels at a photo-diode current of 3 mA, corresponding with an input laser power of 16 mW. Further increasing the optical input power results in little improvement of the clearance, but introduces more compression artifacts. For the optimal photocurrent of 2 mA the detector shows a large shot-noise limited bandwidth of 7.5 GHz in which the clearance remains above 10 dB, accommodating the usage of ADCs with large input bandwidths and sampling speeds. The tunability of the MZIs allows for a very high CMRR ranging from 70 dB at low frequencies to 20 dB at 20 GHz, reducing the effect of the excess noise present in the LO.

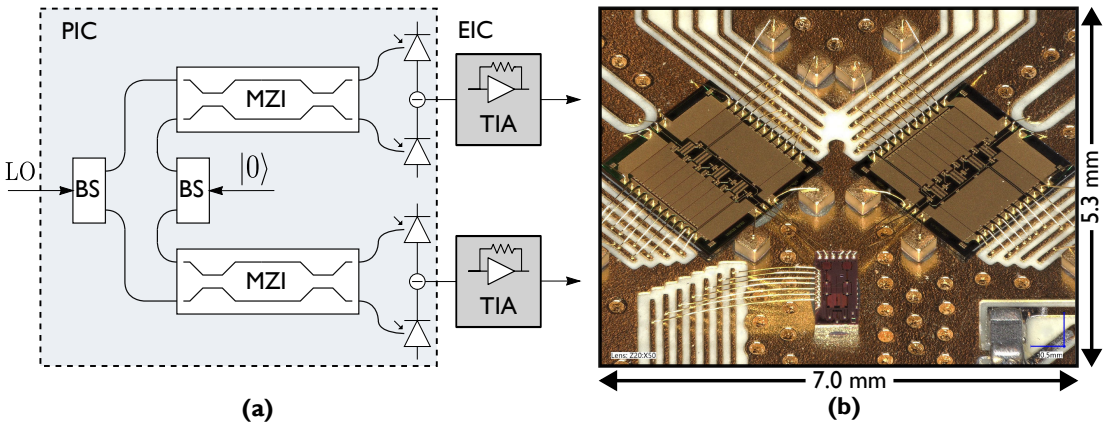

Fig. 2. The Photonic-Electronic front-end. (a) The diagram of the photonic-electronic IC containing two balanced homodyne detectors with low noise trans-impedance apmplifiers (TIAs). (b) An angled picture of the assembly of the photonic-electronic IC.

B. Security proof

The extractable entropy of the captured samples is determined based on an advanced device-dependent security framework [12], which takes into account classical side-information leakage caused by classical noise, quantum side-information leakage resulting from bandwidth limitations in the system and imperfect analog-to-digital conversion. This framework is important to theoretically define the randomness of our source. Based on it, we characterize the measurement apparatus to determine the min-entropy of the captured samples. These samples are assumed to follow a Gaussian distribution, but don't have to be identically independently distributed (IID).

The measurement apparatus comprises the optical front-end, the TIA, and the ADC. The min-entropy of each sample, $H_{\min}$, is calculated as follows:

$$H_{\min} \geq -\log\left[\Gamma(n)\mathrm{erf}\left(\frac{(\Delta x)_{\max}}{2g_*}\right)\right] \tag{1}$$

where $(\Delta x)_{\max}$ is the maximum bin-width of the N-bit ADC, $\Gamma(n)$ is a function of the effective mean photon number n of the excess noise and g_* is the optimized gain. A detailed derivation of this security proof can be found in work [12].

As stated earlier, the min-entropy is a metric for the randomness of the used source, it indicates how many truly random bits can be extracted from a raw sample. Looking at Eq. 1, the argument inside the logarithm determines the min-entropy: the smaller this argument, the larger the entropy. This argument can be broken down into 3 main contribution factors:

1) Photon-number term $\Gamma(n)$, which depends on the effective mean photon number n
2) Optimized gain g_*
3) Maximum ADC bin-width $(\Delta x)_{\max}$

The effective mean photon number n and $\Gamma(n)$ is given by:

$$n = \frac{1}{2}\frac{\sigma_M^2}{\sigma_{M,c}^2}\exp\left\{\int_0^{f_N}\ln\left[1+\frac{1}{C(f)}\right]\frac{df}{f_N}\right\} - \frac{1}{2} \tag{2}$$

$$\Gamma(n) = (\sqrt{n}+\sqrt{n+1})^2 \tag{3}$$

where $C(f) = S_Q(f)/S_E(f)$ is the clearance, which is defined as the ratio of the power spectral density of the quantum signal $S_Q(f)$ and the power spectral density of the excess noise $S_E(f)$. The clearance captures the (electronic) noise performance of the analog front-end. Higher clearance means less classical contamination, reducing n and $\Gamma(n)$. More of the output signal is therefore attributed to the vacuum fluctuations, increasing the min-entropy.

The ratio of the variance of the homodyne measurement σ_M^2 and the conditional variance of the homodyne measurement $\sigma_{M,c}^2$ is defined as the temporal correlation factor $\sigma_M^2/\sigma_{M,c}^2$. This factor captures the bandwidth performance of the measurement device, which can be improved by applying digital equalization. When adjacent samples are less correlated (ratio closer to 1), n decreases, improving entropy. Adjacent samples become independent of each other, resulting in more extractable randomness from each sample.

The gain g_* determines how the Gaussian-distributed signal spans the ADC range R. The optimized gain is chosen using the following equation, which makes the Gaussian of the source spread over all the ADC bins in an optimal way.

$$\mathrm{erf}\left(\frac{(\Delta x)_{\max}}{2g_*}\right) = \frac{1}{2}\mathrm{erfc}\left(\frac{R}{g_*}\right) \tag{4}$$

If the gain is too low, the higher bins of the ADC will only trigger very rarely, thus lowering the entropy. If the gain is too high, saturation will occur, triggering the highest and lowest bins too often. This will result in part of the randomness not being captured by the ADC, lowering the entropy (Fig. 5b).

The final factor we take into account is the maximum bin size of the ADC $(\Delta x)_{max}$, based on the maximum differential nonlinearity (DNL) of an N-bit ADC. Codes with larger bin sizes will get triggered more often compared to the ideal case, resulting in a reduction in the min-entropy.

$$(\Delta x)_{\max} = 1 + \mathrm{DNL}_{\max}\ \mathrm{LSB},\ \text{with}\ 1\ \mathrm{LSB} = R/2^N \tag{5}$$

This security proof links the physical properties of the measurement system to a rigorous min-entropy bound. By optimizing these factors, we ensure that the extracted randomness remains provably secure even under realistic, non-IID conditions.

C. Real-time post-processing

Due to the digitized data containing information on both the vacuum fluctuations and side channels, an extraction process is required. The aim is to remove side channels such as the added classical noise. The post-processing consists of distilling the raw ADC samples into uniform random bit-strings in accordance with the security framework, which calculates the amount of classical and quantum information present in the samples to determine the quantum randomness that can be extracted. To evaluate the randomness of the entropy source, the min-entropy is calculated. In turn, the min-entropy determines the amount of input bits required to produce a number of truly random output bits, which is referred to as the extraction ratio $\left(e = \frac{\#\text{Output bits}}{\#\text{Input bits}}\right)$.

The first processing done on the raw samples consists of applying a 9-tap programmable finite impulse response (PFIR) filter to flatten the power spectral density (PSD) over the Nyquist bandwidth of 2.6 GHz, equalizing the channel response. This reduces the temporal correlation between samples and therefore increases the min-entropy [12].

Equalized samples are transmitted at 10.4 Gbit/s per lane over 16 lanes to the connected FPGA using the JESD204b protocol [33]. A GUI interface is used to configure the settings of the ADCs and JESD204b link at startup. The status of the link is actively monitored so that the QRNG is stopped if failure would occur, making sure that the quality of the ADC samples is not compromised. These 12-bit samples are processed in the randomness extractor.

The randomness extractor produces almost uniform bits from raw samples containing biases due to the side channel information by using a seed consisting of truly random bits [34]. As hash functions such as Trevisan's extractor [35] or the Toeplitz hash [36] satisfy these requirements, we opted to use a specialized version of this Toeplitz hash because of its lighter computational complexity [37] and viability in FPGA designs.

The extracted data can be outputted through the PCIe or QSFP28 connection. A duplicate of each raw ADC sample is sent through the PCIe connection to the connected computer, which uses this raw data to determine the min-entropy. This min-entropy calculation was repeated during the measurements each time a data packet of 1 Gbit was collected to guarantee the quality of the processed data. The frequency of this quality check can be manually set as desired, in practical situations the quality check could be set such that it happens during an unused time slot.

III. Adaptive Toeplitz hash

The Toeplitz hash consists of multiplying the raw samples by a matrix in which all diagonals, extending from the top left to the bottom right, contain the same bit. A randomly initiated Toeplitz matrix is capable of converting its input to a uniform distribution [37]. When executed on binary data, this multiplication can be easily implemented using basic logical operators, making it a suitable candidate for being implemented on FPGAs (Fig. 3a).

The dimensions of the Toeplitz matrix must be chosen so that the extraction ratio is smaller than the min-entropy per bit determined by the security proof. This ensures the randomness of the produced bit-strings.

Another aspect that determines the size of this matrix is the security factor required by the application using the processed numbers. This factor indicates how closely the extracted random sequence approximates the ideal uniform distribution. The size of the matrix is calculated using the leftover hash lemma against quantum side information [11], [38] as the upper bound on the size of the output string given the size of the input, the min-entropy per bit and the security factor:

$$m \leq kH_{\min} - \log_2 \frac{1}{2\epsilon_{\text{hash}}^2} \tag{6}$$

with m the size of the output word in bits, k the size of the input word in bits, $H_{\min}$ the min-entropy per bit and ϵ_{hash} the security factor.

In order to evaluate large matrix multiplications on FPGAs at high speeds, a pipelined variant [39] of the Toeplitz hashing core was implemented (Fig. 3b). At start-up, the Toeplitz matrix is initiated by taking the least significant bit (LSB) of the 12-bit samples and applying a Von Neumann extractor [40] to this stream to remove some of the bias that might be present from the introduced classical noise. While these bits might not be truly random, they eliminate the need for a PRNG or hard-coded seed. Other options to generate this seed would be to apply the Von Neumann extractor to the sign bit or to simply XOR-reduce the 12 bits of each sample to a single bit, thereby increasing its entropy. Afterward, the last byte of every hashed bit-string is removed to refresh the Toeplitz matrix, ensuring its randomness. The whole matrix multiplication is divided into a number of smaller multiplications, where each clock cycle a part of the input bits $\left(s = \frac{k}{N_{\text{cycles}}}\right)$ is multiplied with its corresponding Toeplitz coefficients and added up to the result of the previous cycles. This is repeated until all cycles are completed and the final result is present in the register. This operation is equivalent to performing the total matrix multiplication in a single clock cycle [39].

As the amount of cycles linearly influences the input size of the hash while keeping the output size the same, we are able to control the randomness extraction ratio $\left(e = \frac{m}{k}\right)$ in real-time by adding extra control logic [27] (Fig. 4). This makes the hashing core adaptable to different min-entropy levels and security factors without requiring reprogramming of the FPGA. When a lower extraction ratio is required to guarantee the randomness of the hashed numbers, the number of cycles per output bit-string can be increased.

Inserting the extraction ratio in Eq. 6 demonstrates the effect on the security factor. The closer the ratio is to the estimated min-entropy per bit, the larger the size of the Toeplitz matrix must be to achieve a security factor of a similar magnitude. For this reason, it is preferred to maintain a lower extraction ratio than required by the security proof.

$$k(H_{\min} - e) \geq -\log_2(2\epsilon_{\text{hash}}^2) \tag{7}$$

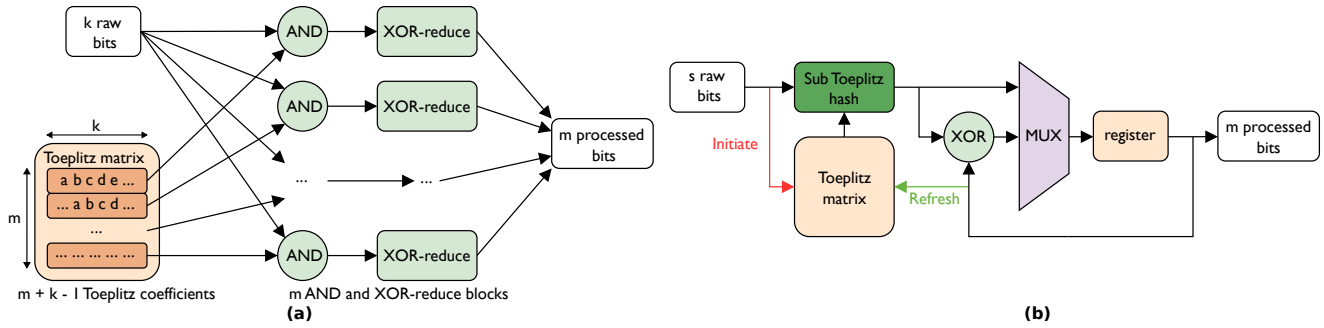

Fig. 3. Simplified graphical depiction of the Toeplitz hash, with green indicating fully logical operations, orange indication memory and purple indicating routing operations (a) The binary evaluation using basic logical operators. (b) The pipelined version of the binary Toeplitz hash, where each cycle a part of the raw bits are hashed and added to the result of the previous cycles.

The seed shift register of the adaptive Toeplitz module is sized to accommodate the minimal configurable extraction ratio used. This ratio $\left(e_{\min} = \frac{m}{s \cdot N_{\text{cycles}_{\max}}}\right)$ amounts to the maximal number of cycles the algorithm can perform before the seed shift register storing the Toeplitz coefficients runs out of space. The general configurable extraction ratio is given by:

$$e = \frac{m}{s \cdot N_{\text{cycles}}} \text{ , } N_{\text{cycles}} \in \left[\text{ceil}\left(\frac{m}{s}\right) \ \ldots \ N_{\text{cycles}_{\max}}\right] \quad (8)$$

with m the number of output bits, s the number of input bits per cycle and $N_{\text{cycles}_{\max}}$ the maximal number of cycles possible per produced bit-string. The minimum number of cycles allowed is chosen so that the extraction ratio doesn't exceed 1. The useful extraction ratios to generate true random numbers will be those lower than the min-entropy per bit. To ensure fast operation of our hash module while utilizing a large part of the resources of the FPGA, an input size s of 48 bits or 4 samples, an output size m of 1320 bits and a $N_{\text{cycles}_{\max}}$ of 48 were chosen. This results in the Toeplitz hash having a maximal size of 2304 x 1320 per individual hashing unit, with 288 iterations required to fully refresh the seed shift register using the 8 bits removed from each produced bit string. The input size consists of an integer number of samples such that information of the same sample isn't divided between different units. By placing 10 of these units in parallel for each channel, a high post-processing throughput can be guaranteed (Fig. 4).

The PCIe connection is used to upload the extraction ratio

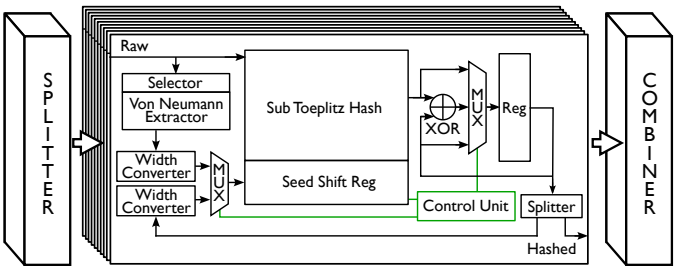

Fig. 4. Diagram of the implemented adaptive Toeplitz hash module used as randomness extractor on the FPGA.

for the Toeplitz hashing module. Currently, a Python API automatically calculates the min-entropy based on raw samples and updates the extraction ratio to its optimal value with respect to the desired security factor. Performing this quality check at regular intervals enables the detection of defects that result in a lower minimum entropy. The Toeplitz hash size is then updated so that the produced bit-strings remain truly random. This update can be scheduled at regular intervals (e.g. every hour) or during moments when there are no requests for random numbers, ensuring minimal impact on the throughput. As both channels have an independent hashing core, the extraction ratio of each core can be programmed separately in cases where applications require a different guaranteed security to maintain optimal throughput.

IV. Min-Entropy Estimation

Determining the min-entropy requires us to collect raw samples from when both the LO laser is disabled and enabled to calculate the excess $S_E(f)$ and homodyne $S_M(f)$ noise PSD. These are calculated using Welch's method [41] after which the clearance $C(f) = 10\log_{10}\left(\frac{S_M(f) - S_E(f)}{S_E(f)}\right)$ is determined.

Here we notice that one of the ADC channels contains more excess noise (Fig. 5a), a lower clearance, which results in slightly worse performance of this channel. However, since both channels still have a clearance above 12 dB over the full bandwidth, this minute difference will have only a minimal effect on the min-entropy, as the factor $1 + 1/C(f)$ remains close to 1 [12]. The clearance is maximized by increasing the LO optical power while making sure that no compression occurs in the photo-diodes of the BHD, giving us the optimal optical power of 18 mW taking into account some extra losses. The next step involved adjusting the amplification g of the TIA's output to optimize the swing applied to the ADCs. Using a low gain results in only a small part of the ADC range being used, and too much gain results in overflow or strong deviation from the expected Gaussian behavior. Both cases result in a reduction of the achievable min-entropy; therefore, the optimal gain is ideally chosen (Fig. 5b). This can also be seen in the Gaussian distribution of channels A and B, which is spread

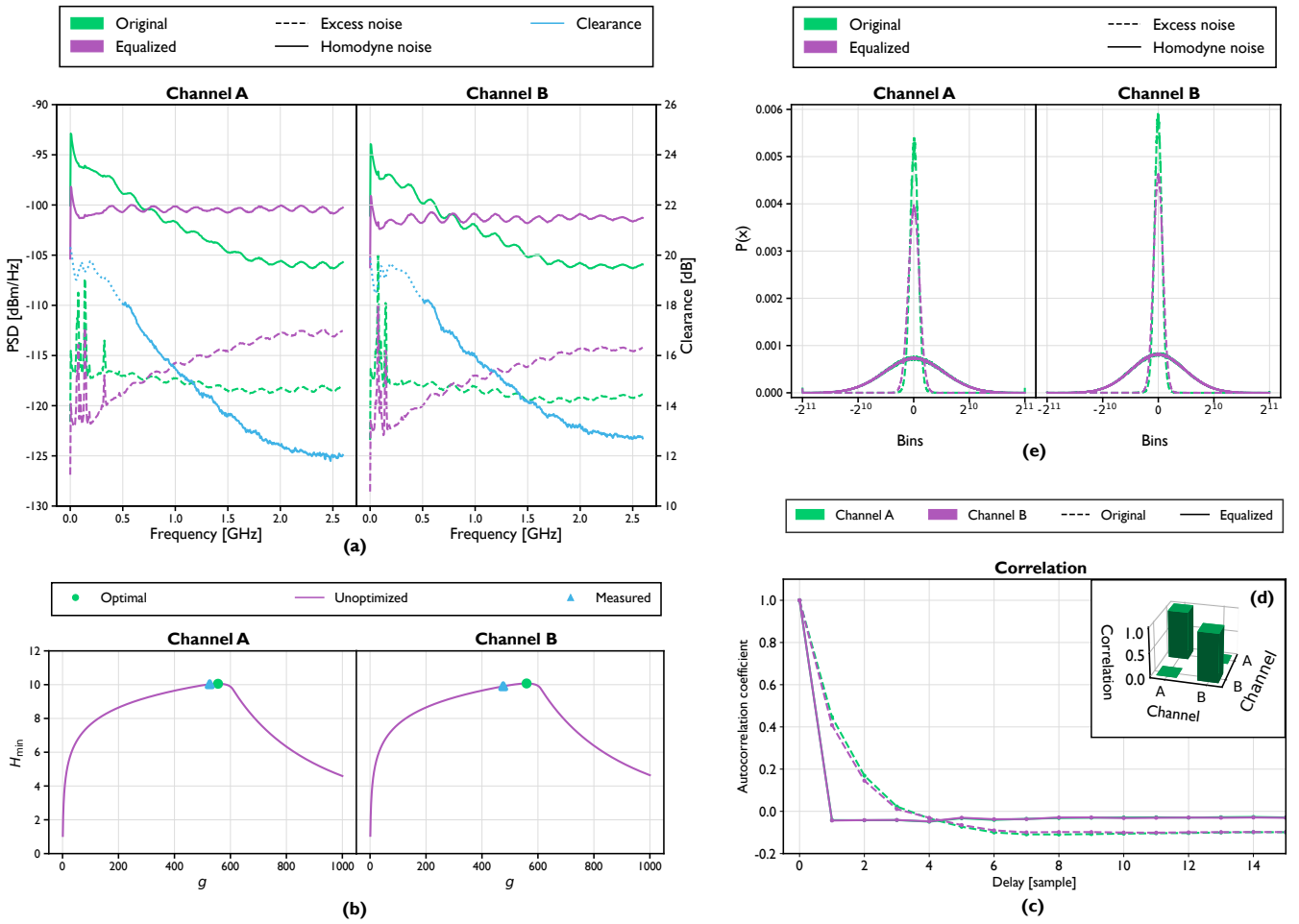

Fig. 5. The measurement results of the full system using the ADC12DJ5200RF EVM and HTG-930 FPGA containing a Xilinx XCVU13P core to capture and process the data. (a) The PSD of the excess and homodyne noise calculated using Welch's method with a segment length of 1024 samples, displayed together with the clearance. This was done for the original and equalized samples for both channels. (b) The effect of the gain between the TIA and ADC, tuned with the amplifiers, on the min-entropy for both channels. (c) The autocorrelation of both channels with a delay of up to 15 samples for the original and equalized samples. (d) The cross-correlation of channel A and channel B. (e) Density of channel A and channel B, showing the Gaussian distribution of the original and equalized samples.

over the ADC bins. The excess noise also follows a Gaussian distribution, but does not fill the full ADC range (Fig. 5e).

As the PSD of the homodyne noise should be flat [12], we can calculate the 9 taps of the PFIR filter present in the ADC evaluation board to equalize the low-pass effect of the detector and the input characteristic of the ADC, without changing the variance of the homodyne noise. Doing this for each channel helps reduce the temporal correlation $\sigma_M^2/\sigma_{M,c}^2$ and thus increases the min-entropy (Fig. 5c). This results in a significant increase in the maximal throughput achievable (Table II).

Once the maximal min-entropy is achieved, the security factor of each extraction ratio is calculated to determine the optimal setting that has a high output rate and a good security factor (Tab. I). This security factor can be further decreased in real-time to a minimum of $1.25\text{e}{-92}$ at the cost of reducing the throughput to 71.06 Gb/s. Running the generator at its optimal setting of 100.3 Gb/s with a security factor of $6.39\text{e}{-8}$ produces random bit-strings that pass both the NIST and Dieharder test suites, with a low correlation between the two channels (Fig. 5d).

V. Discussion

To situate the performance of our QRNG, we compared it against several recent implementations, focusing on generation rate, laser power consumption, real-time capability, and the integration level of the noise source excluding the laser (Table III). These factors are critical in evaluating the practicality and scalability of QRNG systems.

Our QRNG achieves a generation rate of 100 Gbit/s with a laser power of 18 mW and supports real-time operation. Importantly, it uses a custom integrated circuit (cIC) for the noise source, which enhances its potential for on-chip integration and deployment in miniaturized systems. While Ref. [12] also achieves 100 Gbit/s using a similar cIC, this work demonstrates the feasibility of producing secure random

TABLE I
THE CONFIGURABLE EXTRACTION RATIOS, SECURITY FACTORS AND REAL-TIME THROUGHPUT PER CHANNEL BASED ON THE MAXIMAL MIN-ENTROPY FOR OPTIMAL INPUT GAIN. THE OPTIMAL SETTING WITH A HIGH THROUGHPUT AND DECENT SECURITY FACTOR IS INDICATED BY THE GREEN COLOR.

Extraction Ratio $[l/N]$	0.833	0.809	0.786	0.764	0.743	0.724	0.705	...	0.585	0.573
Real-Time Rate [Gbit/s]	51.68	50.16	48.73	47.38	46.10	44.88	43.73	...	36.29	35.53
Security Factor ϵ_{Hash}	7.18e-02	6.39e-08	5.68e-14	5.06e-20	4.50e-26	4.00e-32	3.56e-38	...	1.40e-86	1.25e-92

TABLE II
SUMMARY OF THE EXPERIMENTAL RESULTS OF EACH CHANNEL WITH (EQ) AND WITHOUT (NO EQ) EQUALIZATION. (*FOR OPTIMAL INPUT GAIN)

	Channel A		Channel B	
	No EQ	EQ	No EQ	EQ
H_{min} per sample [bits]	9.06	10.02	9.19	9.90
Maximal* H_{min} per sample [bits]	9.08	10.05	9.20	10.07
Temporal Correlation	1.438	1.003	1.357	1.004
Maximal* Throughput [Gbit/s]	47.21	52.26	47.84	52.36

numbers at high rates using vacuum fluctuations. It didn't include real-time capability, limiting its applicability in latency-sensitive environments.

Other implementations show varying trade-offs. Ref. [26], for example, achieves 50 Gbit/s in real time, requires 180 mW of laser power and relies on bulk components, which may hinder integration. Ref. [42] is highly power-efficient, uses a noise source constructed with bulk components and does not operate in real time. Our previous work [27] offers a good balance with 64.9 Gbit/s, 16 mW power, and a fully integrated noise source, but at a lower throughput than our current system.

This analysis highlights that our QRNG combines high-speed operation, moderate power consumption, real-time capability, and a fully integrated noise source. This positions it as a highly practical and scalable solution for next-generation quantum-secure systems.

VI. CONCLUSION

We have demonstrated a real-time QRNG capable of producing up to 100.3 Gbit/s of random data. This is, to our knowledge, the highest achieved rate by a real-time system. Our solution comprises an integrated photonic-electric dual-channel detector, along with high-speed data acquisition and randomness extraction, utilizing a low-power laser source. The single channel rate of 50.16 Gbit/s is a large improvement of the channel rate of the previous fastest QRNG, while using a similar level of optical power.

TABLE III
STATE-OF-THE-ART QRNGS WITH HIGH RANDOMNESS RATES (CIC: CUSTOM INTEGRATED CIRCUIT)

Ref.	Rate [Gbits/s]	Real-time?	Laser power [mW]	Integration level noise source (Laser excl.)
[23]	18.8	✓	3.36	cIC + packaged
[24]	20	✓	21.15	cIC + packaged
[25]	42	✓	∼20	IC + bulk
[26]	50	✓	180	IC + bulk
[27]	64.9	✓	16	cIC
[42]	68	✗	0.9	bulk
[12]	100	✗	20	cIC
This work	100	✓	18	cIC

An innovative adaptive Toeplitz hashing scheme was developed, which allows for dynamic changes to the min-entropy and security factor requirements. In the future, we want to expand this design by also including real-time health tests of the raw and processed data so that the quality of the random numbers is guaranteed, together with further characterization of the device such that this QRNG can be classified as a PTG.3 class random number generator following the AIS 31 standard [43]. Another future improvement we want to include is the integration of the security framework directly into the FPGA itself, as this would eliminate the need for computer software except for basic drivers to access the hashed random bit-strings and upload the desired security factor. By fully converting the core calculation of this framework to microcontroller level instructions, it is possible to do the full calculation on a soft processor implemented on the FPGA.

APPENDIX A
ADC PERFORMANCE

In an ideal ADC, each code has an equal bin-width. Unfortunately, due to variations in process, voltage, and temperature, the bin-width for different codes varies slightly. This effect is captured in the differential non-linearity (DNL) of the ADC.

The measurements of the ADC12DJ5200RF ADCs are performed based on the IEEE Standard 1241-2010 [44]. A sine wave is applied with a relatively low frequency that isn't an integer divisor of the sample frequency of the ADC, using the Keysight N9310A RF signal generator. Doing this results in each sample measuring a different phase of the sine wave, preventing only some phases from being sampled. For our characterization, a frequency of 1.234 MHz was used. As the PCIe connection doesn't support speeds high enough for a continuous stream of raw samples, an internal FIFO was used to transfer smaller packets. Each slope of the sine wave is

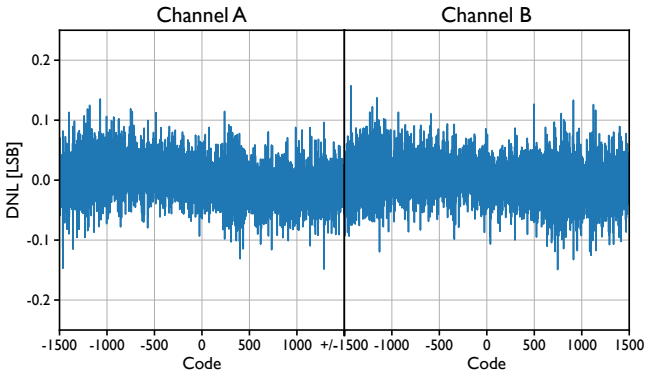

Fig. 6. DNL of Channel A and Channel B for codes ranging from -1500 to 1500.

processed individually and the results are combined so that enough data is processed to remove most of the effect of the noise on the non-linearity estimation.

To verify the DNL values reported in the data sheet of the ADC [45], the errors and DNL of the codes ranging from -1500 to 1500 were determined for both channels (Fig. 6), with raw samples having a 99.5% probability of falling in this range. The maximal DNL measured in this range aligned with the value reported in the data sheet, indicating that we can trust the characterization performed by the manufacturer. This max DNL of 0.16 LSB was used when applying the security proof. This results in a maximum bin-width $(\Delta x)_{\max} = 1.16$ LSB.

Appendix B
Programmable Finite Impulse Response filters

To equalize the spectral response and mitigate intersymbol interference (ISI), we employ digital programmable finite impulse response filters (PFIRs) integrated in the ADC (ADC12DJ5200RF). Each channel provides PFIRs with 9 configurable taps, which can be programmed via SPI. The center tap has a resolution of 18 bits and an LSB weight of 2^{-16}, while the non-center taps have a resolution of 12 bits and an LSB weight of that can be configured between 2^{-10} and 2^{-16} [45]. The filter coefficients are chosen to achieve unity gain, preserving the variance of the homodyne noise.

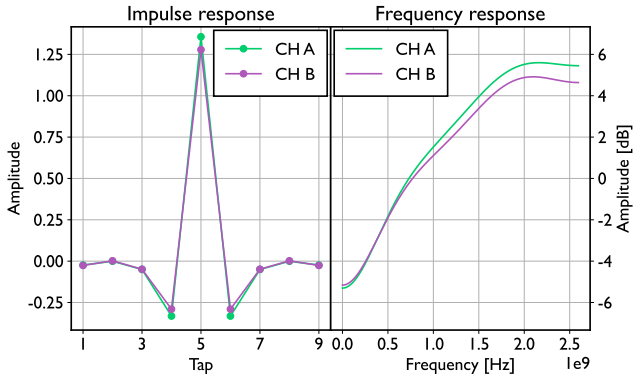

Fig. 7. Impulse (left) and frequency (right) response of the used FIR filters for equalizing the ADC channels.

Figure 7 presents the filter tap coefficients and the corresponding frequency response for both channels. The observed high-pass characteristic compensates for the low-pass behavior introduced by the TIA and ADC, ensuring a flat response across the full Nyquist bandwidth.

Appendix C
Measurement setup

The measurement setup consists of the PCB containing the photonic-electronic integrated circuit (7 mm x 5 mm) mounted on an optical table with one of the grating couplers having light from a fiber coupled in. The other end of this fiber is connected to an NKT Photonics Koheras BASIK E15 laser source, producing light with a wavelength of 1550 nm and has a tunable output power between 10 and 40 mW, with a polarization controller in between. The output of each channel of the integrated circuit is connected to two SHF low-noise amplifiers with a partially tunable gain. The output of these amplifiers is then connected to the ADC12DJ5200RF EVM containing two 12-bit 5.2 GHz ADCs with an 8 GHz analog input bandwidth, PFIR filters, and other controllable features to improve performance. An external 5.2 GHz 8 dBm clock signal is supplied to the ADCs. This evaluation board is connected to the HTG-930 FPGA using the FMC+ connector. Samples are transferred to the FPGA with the JESD204b protocol, which requires a 260 MHz clock to be supplied to the ADC evaluation board. It is also possible to fully generate the clocking on the evaluation board, but it was opted to use external clocking for finer control. The FPGA is slotted into a computer using the PCIe connector, which is also used to transfer raw and processed samples. The complete system, which includes the laser, FPGA and ADC could be integrated into a relatively compact module.

Acknowledgment

This work is partially supported by the Research Foundation–Flanders (FWO) Weave project Squeezed Quantum prOcessing with Photonics and Electronics (FWO Weave SQOPE) (3G092922), the Digital Europe project EU BeQCI (No. 101091625) and the European Quantum Flagship project Quantum Secure Networks Partnership (QSNP) (No. 101114043).

References

[1] M. M. Jacak, P. Jóźwiak, J. Niemczuk, and J. E. Jacak, "Quantum generators of random numbers," *Sci. Rep.*, vol. 11, no. 1, p. 16108, Aug. 2021.

[2] H. Bauke and S. Mertens, "Random numbers for large-scale distributed Monte Carlo simulations," *Phys. Rev. E*, vol. 75, p. 066701, Jun. 2007.

[3] K. Marton, A. Suciu, C. Sacarea, and O. Creţ, "Generation and testing of random numbers for cryptographic applications," *Proc. Romanian Acad. - Series A: Math., Phys., Tech. Sci., Inf. Sci.*, vol. 13, pp. 368–377, Oct. 2012.

[4] I. Vattulainen, K. Kankaala, J. Saarinen, and T. Ala-Nissila, "A comparative study of some pseudo random number generators," *Comput. Phys. Commun.*, vol. 86, pp. 209–226, Apr. 1993.

[5] L. Pasqualini and M. Parton, "Pseudo random number generation: A reinforcement learning approach," *Procedia Comput. Sci.*, vol. 170, pp. 1122–1127, 2020.

[6] M. Stipčević and C. Koç, *Open Problems in Mathematics and Computational Science*. Berlin, Germany: Springer, Nov. 2014, pp. 275–315.

[7] X. Ma, X. Yuan, Z. Cao, B. Qi, and Z. Zhang, "Quantum random number generation," *npj Quantum Inf.*, vol. 2, no. 1, p. 16021, Jun. 2016.

[8] T. Jennewein, U. Achleitner, G. Weihs, H. Weinfurter, and A. Zeilinger, "A fast and compact quantum random number generator," *Rev. Sci. Instrum.*, vol. 71, pp. 1675–1680, Apr. 2000.

[9] H. Guo, W. Tang, Y. Liu, and W. Wei, "Truly random number generation based on measurement of phase noise of a laser," *Phys. Rev. E*, vol. 81, p. 051137, May 2010.

[10] C. Gabriel *et al.*, "A generator for unique quantum random numbers based on vacuum states," *Nature Photon.*, vol. 4, no. 10, pp. 711–715, Oct. 2010.

[11] T. Gehring *et al.*, "Homodyne-based quantum random number generator at 2.9 Gbps secure against quantum side-information," *Nature Commun.*, vol. 12, no. 1, p. 605, Jan. 2021.

[12] C. Bruynsteen, T. Gehring, C. Lupo, J. Bauwelinck, and X. Yin, "100-Gbit/s integrated quantum random number generator based on vacuum fluctuations," *PRX Quantum*, vol. 4, p. 010330, Mar. 2023.

[13] H.-K. Lo and H. F. Chau, "Unconditional security of quantum key distribution over arbitrarily long distances," *Science*, vol. 283, no. 5410, pp. 2050–2056, 1999.

[14] Y. Zhang, Y. Bian, Z. Li, S. Yu, and H. Guo, "Continuous-variable quantum key distribution system: Past, present, and future," *Appl. Phys. Rev.*, vol. 11, p. 011318, 2024.

[15] Y. Pi *et al.*, "Sub-Mbps key-rate continuous-variable quantum key distribution with local local oscillator over 100-km fiber," *Opt. Lett.*, vol. 48, no. 7, pp. 1766–1769, Apr. 2023.

[16] A. A. E. Hajomer *et al.*, "Continuous-variable quantum key distribution at 10 GBaud using an integrated photonic-electronic receiver," *Optica*, vol. 11, no. 9, pp. 1197–1204, Sep. 2024.

[17] A. A. E. Hajomer *et al.*, "Chip-based 16 GBaud continuous-variable quantum key distribution," *Optica*, vol. 13, no. 6, pp. 1035–1042, Jun. 2026.

[18] M. Mosca and M. Piani, "2023 quantum threat timeline report," 2023. [Online]. Available: https://globalriskinstitute.org/publication/2023-quantum-threat-timeline-report/

[19] "Quantum key distribution market size and forecast to 2030," Nov. 2023. [Online]. Available: https://www.coherentmarketinsights.com/market-insight/quantum-key-distribution-market-5971

[20] F. Monet, J.-S. Boisvert, and R. Kashyap, "A simple high-speed random number generator with minimal post-processing using a random Raman fiber laser," *Sci. Rep.*, vol. 11, no. 1, p. 13182, Jun. 2021.

[21] K. Kim *et al.*, "Massively parallel ultrafast random bit generation with a chip-scale laser," *Science*, vol. 371, no. 6532, pp. 948–952, 2021.

[22] H. Wu *et al.*, "Ultra-high speed random bit generation based on Rayleigh feedback assisted ytterbium-doped random fiber laser," *Sci. China Technol. Sci.*, vol. 64, no. 6, pp. 1295–1301, Jun. 2021.

[23] B. Bai *et al.*, "18.8 Gbps real-time quantum random number generator with a photonic integrated chip," *Appl. Phys. Lett.*, vol. 118, no. 26, p. 264001, 2021.

[24] T. Bertapelle *et al.*, "High-speed source-device-independent quantum random number generator on a chip," *Optica Quantum*, vol. 3, no. 1, pp. 111–118, Feb. 2025.

[25] J. Zhao *et al.*, "Over 40 Gbps tri-type quantum random number generator," in *Proc. SPIE 13391, Quantum Comput., Commun., Simul. V*, Mar. 2025, p. 133910Y.

[26] K. Tanizawa, K. Kato, and F. Futami, "Real-time 50-Gbit/s spatially multiplexed quantum random number generator based on vacuum fluctuation," presented at the Opt. Fiber Commun. Conf., San Diego, CA, USA, Mar. 5–9, 2023, Paper Th4A.8.

[27] A. Bomhals, C. Bruynsteen, and X. Yin, "64.9 Gbit/s real-time quantum random number generator with an integrated photonic-electronic detector," *IET Conf. Proc.*, vol. 2023, pp. 1174–1177, 2023.

[28] L. E. Bassham III *et al.*, "A statistical test suite for random and pseudorandom number generators for cryptographic applications," Nat. Inst. Standards Technol., Gaithersburg, MD, USA, Tech. Rep. NIST Special Pub. (SP) 800-22, Rev. 1a, Apr. 2010.

[29] G. B. Robert, E. Dirk, and B. David, "Dieharder: A random number test suite," 2020. [Online]. Available: https://webhome.phy.duke.edu/~rgb/General/dieharder.php

[30] 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*, vol. 87, p. 062327, Jun. 2013.

[31] M. Fox, *Quantum Optics: An Introduction*. London, U.K.: Oxford Univ. Press, Apr. 2006.

[32] C. Bruynsteen, M. Vanhoecke, J. Bauwelinck, and X. Yin, "Integrated balanced homodyne photonic–electronic detector for beyond 20 GHz shot-noise-limited measurements," *Optica*, vol. 8, no. 9, pp. 1146–1152, Sep. 2021.

[33] "JEDEC Standard JESD204B," Arlington, Tech. Rep., Jul. 2011. [Online]. Available: https://www.jedec.org

[34] S. P. Vadhan, "Pseudorandomness," *Found. Trends Theor. Comput. Sci.*, vol. 7, no. 1–3, pp. 1–336, 2012.

[35] L. Trevisan, "Extractors and pseudorandom generators," *J. ACM*, vol. 48, no. 4, pp. 860–879, Jul. 2001.

[36] H. Krawczyk, "New Hash Functions for Message Authentication," in *Proc. Adv. Cryptol.* Berlin, Heidelberg: Springer, 1995, pp. 301–310.

[37] M. Hayashi and T. Tsurumaru, "More efficient privacy amplification with less random seeds via dual universal hash function," *IEEE Trans. Inf. Theory*, vol. 62, no. 4, pp. 2213–2232, Apr. 2016.

[38] M. Tomamichel, C. Schaffner, A. Smith, and R. Renner, "Leftover hashing against quantum side information," *IEEE Trans. Inf. Theory*, vol. 57, no. 8, pp. 5524–5535, Aug. 2011.

[39] X. Zhang, Y.-Q. Nie, H. Liang, and J. Zhang, "FPGA implementation of Toeplitz hashing extractor for real time post-processing of raw random numbers," in *Proc. IEEE-NPSS Real Time Conf.*, 2016, pp. 1–5.

[40] C. Foreman, R. Yeung, A. Edgington, and F. J. Curchod, "Cryptomite: A versatile and user-friendly library of randomness extractors," *Quantum*, vol. 9, p. 1584, Jan. 2025.

[41] P. Welch, "The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms," *IEEE Trans. Audio Electroacoust.*, vol. 15, no. 2, pp. 70–73, Jun. 1967.

[42] Y.-Q. Nie, L. Huang, Y. Liu, F. Payne, J. Zhang, and J.-W. Pan, "The generation of 68 Gbps quantum random number by measuring laser phase fluctuations," *Rev. Sci. Instrum.*, vol. 86, no. 6, p. 063105, Jun. 2015.

[43] M. Peter and W. Schindler, "A proposal for functionality classes for random number generators," Bundesamt für Sicherheit in der Informationstechnik (BSI), Bonn, Germany, Tech. Rep. AIS 31 Version 3, Sep. 2024. [Online]. Available: https://www.bsi.bund.de

[44] "IEEE Standard for Terminology and Test Methods for Analog-to-Digital Converters," *IEEE Standard 1241-2010 (Revision of IEEE Standard 1241-2000)*, pp. 1–139, Jan. 2011.

[45] "ADC12DJ5200RF data sheet," 2025. [Online]. Available: https://www.ti.com/product/ADC12DJ5200RF