64.9 Gbit/s Real-Time Quantum Random Number Generator with an Integrated Photonic-Electronic Detector

Axl Bomhals[(1)], Cédric Bruynsteen[(1)], Xin Yin[(1)]

[(1)] IDLab, Ghent University - imec, 9052 Ghent, Belgium, axl.bomhals@ugent.be

Abstract *We present an FPGA-based quantum random number generator using a low-power custom-designed integrated vacuum fluctuation source. An adaptive, efficient and fast Toeplitz hashing scheme is developed for FPGA, delivering a record extracting rate of 64.9 Gbit/s for secure usage in real time. ©2023 The Authors*

Introduction

True random number generators (TRNGs) play a critical role in secure encryption systems by producing numbers that are unpredictable, even to those with a comprehensive understanding of the system. These generators primarily function by extracting randomness from physical events. A specialized sub-category of TRNGs, known as quantum random number generators (QRNGs), leverages inherently random quantum mechanical phenomena. One such phenomenon, vacuum fluctuations, is utilized in this work to generate random numbers at very high speeds. The ongoing development of advanced quantum encryption schemes necessitates the creation of fast TRNGs capable of supporting the ever-increasing data rate demands, favourably with low-cost, miniaturized integrated devices.

Extensive research has been conducted in developing high-entropy sources that could potentially yield QRNGs with exceptionally high random number rates[1]–[4]. While these sources hold considerable promise, the reported speeds are typically extrapolated from offline post-processing. To render these random number generators field-ready, real-time post-processing must be implemented. Only recently have high-speed real-time QRNGs been reported[5]–[7]. A photonic integrated chip (PIC) based QRNG with subsequent paralleled post-processing in a field programmable gate array (FPGA) demonstrated a real-time output rate as high as 18.8 Gbit/s[6]. The author in[7] presents a real-time QRNG that achieves an aggregated generation rate of 50 Gbit/s using a spatially multiplexed scheme. However, the sizeable entropy source was built relying on discrete photonic and electric components (one 1x8 slicia PLC splitter, plus four discrete balanced photodiodes followed by four connectorized RF amplifiers). In addition, a lower channel rate of 12.5 Gbit/s demands a higher number of multiplexed channels and a high-power laser with $>$180 mW output power.

In this paper, we present an FPGA-based record-fast 64.9 Gbit/s real-time QRNG, which employs a quantum entropy source based on a photonic-electronic integrated dual-channel detector. Thanks to the high integration level of PIC and electric integrated circuit (EIC), the entropy source has a small footprint and works with laser optical power at 16 mW. Moreover, we provide a thorough characterization of the acquisition device, based on an advanced security proof[1] guaranteeing the true randomness of the generator, and have validated the output random numbers using the NIST and Dieharder testsuites.

Integrated Photonic-Electronic Noise Source and Entropy Framework

The maximum achievable throughput of a QRNG is primarily determined by three factors: the noise source, the device employed to capture this noise, and the security proof that provides the min-entropy. For the noise source, we make use of the intrinsic quadrature fluctuations present in the vacuum state. As shown in Fig. 1, the measurement device consists of a dual-channel integrated low-noise balanced homodyne detector. A 16 mW local oscillator (LO) laser and the vacuum fluctuations are split into two channels using 50/50 beam splitters (BSs), after which the fluctuations are amplified by the laser in the Mach-Zehnder interferometers (MZIs), and then captured by dual-channel balanced homodyne detectors. The noise is further amplified by the transimpedance amplifiers (TIAs) and subsequently converted to a digital signal by a two-channel 14-bit 3 GSamples/s analog-to-digital converter (ADC). Lastly, the digitized data is further processed on an FPGA.

We start off by applying an advanced device-

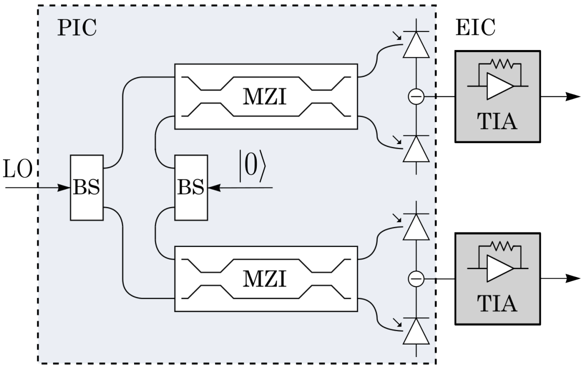

Fig. 1: Simplified schematic of integrated vacuum fluctuations source.

dependent security framework[1], 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. Based on this framework, we characterize the measurement apparatus to determine the min-entropy of the captured samples. The measurement apparatus comprises the optical front-end, the TIA, and the ADC. The min-entropy, H_{min}, is calculated as follows:

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

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. This min-entropy can be increased by reducing the temporal correlation and increasing the clearance between the excess and homodyne noise[1].

Post-Processing and FPGA Implementation

The raw samples which are being produced by the ADC follow a Gaussian distribution and contain contributions from side-channels. To obtain truly random strings from the discretized output of the ADC, post-processing is required. This post-processing step distills the raw samples into uniform random strings in accordance with the min-entropy determined by the security framework (Eq. 1), removing any bias present in the raw samples.

The key post-processing step can be implemented using a Universal Hash function. We employed Toeplitz hashing as a randomness extractor. This hashing consists of multiplying the raw samples with a Toeplitz matrix and has been proven to convert its input to a uniform distribution when performed using a random seed matrix[8]. The dimensions of the seed matrix determine the compression of the raw samples and can be chosen such that the throughput of processed bits is equal to or lower than the maximal throughput determined by the security proof. Another aspect that determines the size of this matrix is the security factor required by the application using the processed numbers. This factor gives an indication of how close the extracted random sequence is to the ideal uniform distribution. The ratio of the matrix is calculated using the leftover hash lemma against quantum side information[5]:

$$l \geq NH_{min} - log_2\frac{1}{2\epsilon_{hash}^2} \qquad (2)$$

with l the size of the output word in bits, N 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, a pipelined Toeplitz hashing core was developed (Fig. 2). At start-up, the seed matrix is initiated by taking the first bit of each sample, removing the need for a pseudo-random number generator. Afterwards it is refreshed by taking the last 8 bits of each output word, ensuring its randomness. The number of steps in the hashing pipeline is controlled by using a shift register as seed matrix that has a selectable output. The input word increases linearly with the number of steps and the output word remains constant in size. This makes the hashing core adaptable to different min-entropy levels without the need to reprogram the FPGA, and when interference occurs, one can maintain the randomness of the produced numbers by lowering extraction ratio.

The full FPGA implementation is shown in Fig. 3. It consists of a data link network, a processor network and the hashing core. The raw (equalized) samples are transported from the ADC to FIFOs on the FPGA using the JESD204b class 0 protocol. The processor network configures the settings of the ADC and JESD204b link, which is then monitored to verify the correct operation. The raw samples are sent to both the hashing core of each channel and PCIe for further processing. Before the actual hashing process,

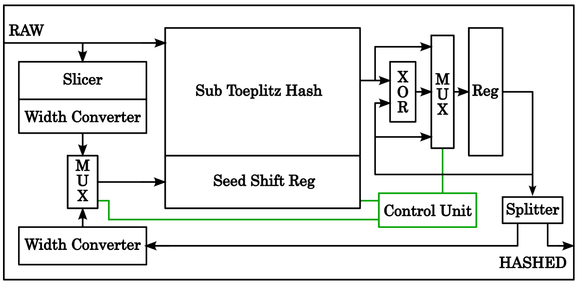

Fig. 2: Toeplitz hashing core.

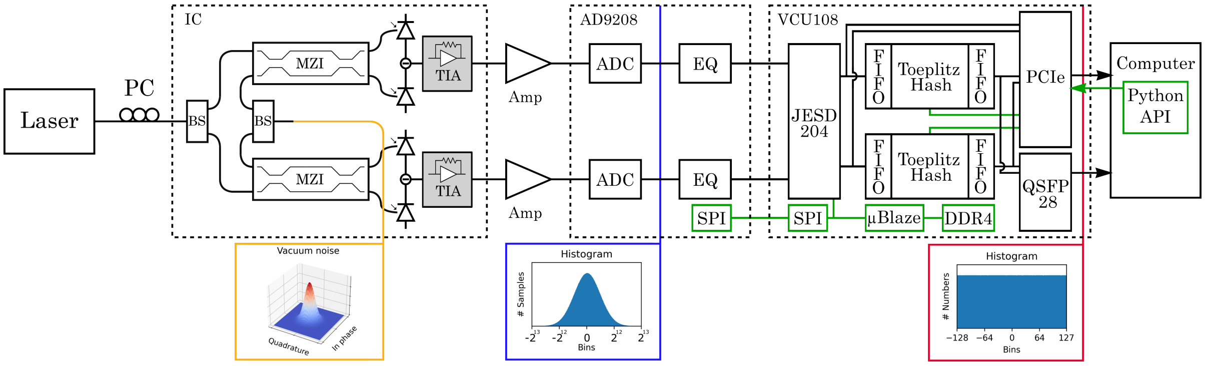

Fig. 3: Block diagram of the full implementation and experimental setup.

the raw samples are equalized using a 47-tap programmable finite impulse response (FIR) filter to achieve a flat power spectral density (PSD) to further improve the temporal correlations, increasing the min-entropy[1]. As both channels have an independent hashing core, the extraction ratio of each core may differ in case applications require a higher guaranteed security, while not unnecessarily reducing the throughput for the other applications.

The processed data can be outputted using the PCIe or QSFP28 connection. The PCIe connection is also used to upload the extraction ratio for the Toeplitz hashing module. A Python API uses the raw samples to automatically calculate the min-entropy and update the extraction ratio to its optimal value w.r.t. the desired security factor.

Experimental Results

We first characterized the measurement apparatus using different settings of the ADC range, laser power, and with or without equalization. By capturing the raw samples when the laser is turned on and off, the PSD and clearance can be determined for both channels. This measurement gives us an initial estimate of the min-entropy and the channel response. In an ideal ADC, each code has a uniform bin-width of one least significant bit (LSB). Unfortunately, due to variations in process, voltage, and temperature, the bin-width for different codes varies slightly. This effect is captured in the differential nonlinearity (DNL) of the ADC. We have measured the maximum DNL of our two-channel ADC to be 0.43 LSB, resulting in a maximum bin-width $(\Delta x)_{max} = 1.43$ LSB.

	Channel A		Channel B	
	No EQ	EQ	No EQ	EQ
Min-Entropy [bits]	11.24	11.69	11.59	11.66
Temporal Correlation	1.125	1.004	1.019	1.006
Max Throughput [Gbit/s]	33.72	35.07	34.77	34.99
Correlation Coefficient	8.15e-05			

Tab. 1: Summary of experimental results.

The gain of the amplifiers before the ADC is adjusted to reach the optimal gain g_* (Eq. 1). The 47-tap FIR filter in FPGA is used to equalize the PSD over the 1.5 GHz Bandwidth, resulting in a reduction of the temporal correlation and increasing the min-entropy of both channels as can be seen in Tab. 1.

Tab. 2 shows the measured maximum real-time QRNG rates and extraction ratios of the FPGA-based adaptive hashing scheme. The security factor of the highest output at 69 Gbit/s does not meet the requirement for the most secure usage but can still be useful for less critical applications such as simulations or gaming. The next rate at 64.94 Gbit/s with a practical security factor is suitable for standard quantum encryption schemes. The extraction ratio can be further reduced untill the requirements of the application are met at the cost of reduced throughput. A final validation was done on the produced random numbers using the Dieharder and NIST test suites, with all of these tests passing.

Extraction Ratio $[l/N]$	0.830	0.782	0.738	0.699
Real-Time Rate [Gbit/s]	69.00	64.94	61.33	58.11
Security Factor ϵ_{Hash}	2.81e-01	2.66e-08	2.51e-15	2.38e-22

Tab. 2: Extraction ratio settings and generation rates.

Ref.	# Channels	Rate / Channel	Laser Power
[5]	1	2.90 Gbit/s	1.60 mW
[9]	7	440 Mbit/s	1.00 mW
[6]	1	18.8 Gbit/s	3.36 mW
[7]	4	12.5 Gbit/s	180 mW
This work	2	32.5 Gbit/s	16.0 mW

Tab. 3: Real-time vacuum fluctuation QRNGs comparison.

Conclusion

We have demonstrated a record-fast real-time QRNG with an integrated photonic-electric dual-channel detector, generating 64.9 Gbit/s of random numbers using a 16 mW laser source. Its per channel rate of 32.5 Gbit/s is more than twice as fast as the per channel rate of the previous fastest QRNG, while using ten times less optical power of[7] (Tab. 3). An innovative adaptive Toeplitz hashing scheme allows for dynamic changes in min-entropy and security factor requirements.

Acknowledgements: This work is partially supported by FWO Weave SQOPE, EU BeQCI (101091625) and QSNP (101114043).

References

[1] 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. 010 330, 1 Mar. 2023. DOI: 10.1103/PRXQuantum.4.010330.

[2] 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”, *Scientific Reports*, vol. 11, no. 1, p. 13 182, Jun. 2021, ISSN: 2045-2322. DOI: 10.1038/s41598-021-92668-0.

[3] K. Kim, S. Bittner, Y. Zeng, *et al.*, “Massively parallel ultrafast random bit generation with a chip-scale laser”, *Science*, vol. 371, no. 6532, pp. 948–952, 2021. DOI: 10.1126/science.abc2666.

[4] H. Wu, J. Xiong, B. Han, *et al.*, “Ultra-high speed random bit generation based on rayleigh feedback assisted ytterbium-doped random fiber laser”, *Science China Technological Sciences*, vol. 64, no. 6, pp. 1295–1301, Jun. 2021, ISSN: 1869-1900. DOI: 10.1007/s11431-020-1806-7.

[5] T. Gehring, C. Lupo, A. Kordts, *et al.*, “Homodyne-based quantum random number generator at 2.9 gbps secure against quantum side-information”, *Nature Communications*, vol. 12, no. 1, p. 605, Jan. 2021, ISSN: 2041-1723. DOI: 10.1038/s41467-020-20813-w.

[6] B. Bai, J. Huang, G.-R. Qiao, *et al.*, “18.8 gbps real-time quantum random number generator with a photonic integrated chip”, *Applied Physics Letters*, vol. 118, no. 26, p. 264 001, 2021. DOI: 10.1063/5.0056027.

[7] T. Ken, K. Kentaro, and F. Fumio, “Real-time 50-gbit/s spatially multiplexed quantum random number generator based on vacuum fluctuation”, in *2023 OFC Conference*, 2023.

[8] M. Hayashi and T. Tsurumaru, “More efficient privacy amplification with less random seeds via dual universal hash function”, *IEEE Transactions on Information Theory*, vol. 62, no. 4, pp. 2213–2232, 2016. DOI: 10.1109/TIT.2016.2526018.

[9] B. Haylock, D. Peace, F. Lenzini, C. Weedbrook, and M. Lobino, “Multiplexed Quantum Random Number Generation”, *Quantum*, vol. 3, p. 141, May 2019, ISSN: 2521-327X. DOI: 10.22331/q-2019-05-13-141.