IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 25 September 2023, accepted 18 October 2023, date of publication 25 October 2023, date of current version 1 November 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3327325

RESEARCH ARTICLE

RAVA: An Open Hardware True Random Number Generator Based on Avalanche Noise

GABRIEL GUERRER
IDOR-Pioneer Science Fellow, D'Or Institute for Research and Education (IDOR), Rio de Janeiro 22281-100, Brazil
Paradox Science Institute, Palo Alto, CA 94306, USA

e-mail: gabriel.guerrer@idor.org

This work was supported by the D'Or Institute for Research and Education (IDOR).

ABSTRACT Entropy is a crucial resource in the domains of cryptography, artificial intelligence, and science. This paper introduces RAVA, a true random number generator based on avalanche noise. RAVA is an open-source device designed to offer a transparent and customizable platform, making auditable and high-quality entropy accessible to a wider audience. The device employs a differential design, which involves comparing two similar noise sources to mitigate the impact of environmental factors. Furthermore, RAVA incorporates a dual entropy core architecture featuring two independent entropy channels that generate random bytes simultaneously. A stochastic model is theoretically derived and empirically confirmed, offering valuable insights into the entropy extraction mechanism and allowing the estimation of the minimum bias attainable. An implementation is presented as a discrete circuit with an ATmega32U4 microcontroller including a USB interface, achieving an unbiased throughput of 136.0 Kbit/s without the necessity of post-processing algorithms. The generated random bytes are evaluated for bias and serial correlation, their entropy is assessed using NIST SP 800-90B estimators, and the randomness quality is verified using the NIST 800-22R1a test suit. For comparison, the same tests are applied to a commercial device based on quantum optical phenomena, revealing similar distributions for both devices across the studied metrics.

INDEX TERMS Random number generator, entropy source, reverse-biased diode, avalanche breakdown, open-source hardware.

I. INTRODUCTION

True random number generators (TRNG) are devices that perform measurements in fundamentally unpredictable physical systems to generate random outcomes. They are used for various applications that rely on entropy as cryptography, electronic games, and scientific simulations.

The outcome of an ideal TRNG is characterized by uniform and unpredictable number sequences. If the device generates a sequence of bits, uniformity means that each bit should be 0 or 1 with 50% probability. The unpredictability condition states that the previous outcomes cannot be used to predict future measurements, meaning that the generated bits are independent of each other. However, practical TRNG implementations can exhibit imperfections due to construction limitations and natural variations in electronic component properties. These factors can introduce biases and correlations in the generated sequences, undermining their randomness. To address these issues, deterministic post-processing functions can be applied to the generated bit sequences. The output of a post-processing function is a new bit sequence, smaller than the one used as input but with an enhanced entropy content.

The review literature [1], [2], [3], [4], [5] highlights various categories of physical phenomena utilized as entropy sources in TRNGs. These include thermal noise, electronic noise (Shot, Zener, avalanche), chaos, phase jitter, radioactive decay, and quantum optical effects. Examples of recent developments can be found on [6], [7], [8], and [9], highlighting the prevalent use of phase jitter and quantum optical effects as entropy sources. These sources exhibit high speeds, enabling throughputs on the order of millions or even billions of bits per second. Moreover, their implementations can be accommodated within compact designs, such as Field Programmable Gate Arrays (FPGAs) systems, as well as

The associate editor coordinating the review of this manuscript and approving it for publication was Francesco G. Della Corte.

integrated circuits like the IDQ250C3 and QN100 chips produced by IDQuantique and Quside, respectively.

While compactness is often desirable in many applications, it can hinder direct access to the entropy source. This limitation can be advantageous in scenarios where security is of utmost importance, as it prevents unauthorized access to the circuit's core elements. However, in applications operating within a secure environment where the physical presence of a malicious third party can be excluded, direct access to the entropy source becomes valuable as it enables auditing, i.e., allowing an investigator to examine the source to verify its reliability. In contrast to phase jitter and quantum optical effects, electronic noise sources are implemented by discreet components that can be directly monitored with voltage-measuring tools. As the TRNG design presented herein aims to prioritize transparency, the scope of this paper is directed to electronic noise sources, particularly the noise found in reverse-biased Zener diodes.

A Zener diode is a semiconductor device composed of a p-n junction with a unique characteristic known as the Zener voltage, denoted as V_Z. When subjected to a reverse voltage that exceeds V_Z, it conducts current reliably while maintaining a V_Z voltage across its terminals. This property makes Zener diodes widely employed as voltage regulators in various electronic circuits.

The noise observed in reverse-biased Zener diodes has two possible mechanisms, Zener and avalanche breakdown [10]. Zener breakdown is a dominating effect in lower voltage Zener diodes ($V_Z < 6$ V). It is caused by electrons tunneling from the valence band on the p side to the conduction band on the n side. Avalanche noise is found in higher voltage Zener diodes and is caused by a cascade effect involving electric fields and free electrons, as detailed in [11]. When the current flowing through the diode remains relatively low ($<100\mu$A), the tunneling/cascade is set in a state of intermittent on-and-off switching, causing the noise.

Both breakdown mechanisms are seen in an oscilloscope as sudden voltage jumps across the diode. The inherent time unpredictability of the voltage jumps constitutes the source of entropy in reverse-biased Zener diodes. However, those processes have some memory effect, meaning that an instantaneous voltage across the component depends on the system's recent history. Consequently, voltage measurements must be conducted over sufficiently large time intervals to achieve bit independence, a matter which will be thoroughly investigated in the following sections.

Various TRNG designs based on tunneling and avalanche noise have been explored in the literature, as evidenced by [12], [13], [14], [15], and [16]. The device introduced in this paper, the RAVA circuit [17], stands out as a unique design that combines what the author considers to be the most favorable features found in the aforementioned designs: the noise source differential design [13], [16], the use of pulse counters for improving the bias [12], [13], the use of operational amplifiers to buffer the noise and raise it to a common DC level [14], the auditable design with monitoring headers [16], and the open-source design [15]. RAVA is an acronym derived from the words Random and Avalanche, symbolizing the device's foundation in utilizing avalanche noise as an entropy source.

While entropy sources based on avalanche noise have been extensively explored in recent decades, the contribution of the RAVA design resides in its singular design summarized as an open-source, fully auditable TRNG featuring two independent randomness cores operating within a differential framework. Currently, there's a shortage of open-source designs matching the trustworthiness level of the commercial TRNGs employed by government agencies and corporations. The RAVA device aims to bridge this gap by offering a solution where reliability is achieved through inherent quality, absolute transparency, and consensus within the users' community. The open-source aspect, as illustrated by the Arduino example, has the potential to extend the reach of technologies. In the case of RAVA, it may expand the access of auditable and high-quality entropy for a wider audience.

The RAVA device can find application in various domains, including:

- Personal privacy: The RAVA circuit can enhance privacy in cryptography and blockchain applications. The existence of a high-quality open-source device benefits such niche where budget restrictions may apply.
- Scientific research: There are a considerable number of studies relying on pseudo-random generators that could benefit from true randomness. The RAVA circuit can be applied in Monte Carlo simulations, random weighting for neural networks, random timing in cognitive research, random assignment of groups and conditions in double-blind studies and blind analysis, among other uses. Transparency is crucial in scientific applications, allowing researchers to fully understand, test, and monitor the used randomness source.
- Maker community: The RAVA circuit can be used to create unpredictable behavior for artificial intelligence in robotics and games. Customizability is a critical aspect in these domains. The circuit allows integration with sensors and other devices through exposed interface headers. Additionally, firmware upgrades enable users to tailor the circuit's behavior and implement new functionalities. In a system comprising multiple components, the RAVA's microcontroller can serve as the central processing unit, orchestrating its operation.
- Arts projects: The RAVA circuit can be utilized to create immersive experiences within installation artworks. By integrating the circuit, artists can generate unpredictable variations of images, colors, patterns, sounds, and music in real time. This capability adds an element of surprise, captivating the audience and fostering a sense of discovery within the artistic experience. In digital arts, randomness is applied in

diffusion models, i.e., neural networks that generate visually compelling images from textual inputs.
- Educational projects: The RAVA circuit can be employed as an educational tool as its usage incites users to delve into electronics and software programming. Additionally, toy experiments producing random bits can teach concepts related to statistics and the scientific method. The users can learn more about all the mentioned fields as they investigate the circuit, possibly guided by didactic material and online tutorials.

When evaluated alongside high-end or commercially available solutions, a limitation of the RAVA circuit lies in its throughput, rated at 136.0 Kbit/s in the current implementation. However, in contrast to the example of a web server providing cryptography services to numerous users concurrently, the mentioned applications are compatible with such throughput.

Considering the potential actions of malicious entities, the proposed applications can be categorized into two scenarios. The first encompasses environments that can be deemed safe, such as the user's home or laboratory. The second scenario involves non-critical applications, where no sensitive information is indirectly exposed in the event of an induced fault.

The general design of the RAVA device is presented in the next section, followed by the details of a specific hardware implementation, an empirical study of the noise characteristics, and the statistical analysis of the generated random bytes.

II. GENERAL DESIGN

This section highlights the key characteristics that an RAVA circuit implementation should adhere to. The RAVA device's main features are:

1) High-quality entropy: Producing unbiased and independent random bits without a post-processing algorithm.
2) Differential design: Aiming towards immunity to environmental conditions by comparing two similar and independent noise sources.
3) Dual entropy core: Incorporating two parallel and independent entropy channels that simultaneously produce random bytes. The dual design provides redundancy, a double output rate, or a unique feature for experiments employing a condition/control design.
4) Full transparency: As an open-source project providing complete access to the circuit design, firmware, and user-side software, including drivers, libraries, and utilities. At the hardware level, monitoring headers allow real-time inspection of voltages and noise sources during operation.
5) Customizability: Offering interface headers for integration with other circuits, sensors, and integrated circuit (IC) components. Users have full control over the device's operation by sending commands through a communication interface. Moreover, as an open-source project, the firmware can be updated to modify the circuit's behavior and implement new functionalities.
6) Accessibility: As a discrete circuit employing low-cost IC components and SMD resistors and capacitors of size 0805. Ensuring that the device remains affordable and can be assembled by users through manual soldering of the components to the printed circuit board.

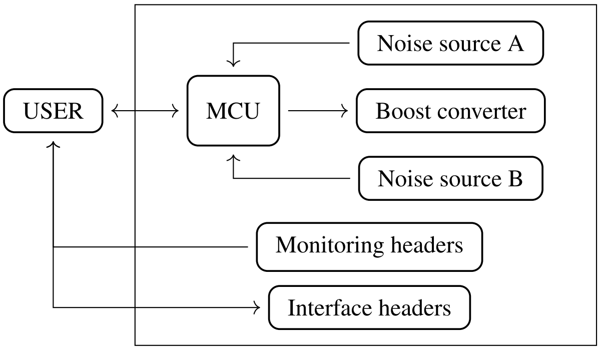


FIGURE 1. Block diagram illustrating the essential modules of the RAVA circuit. The arrows depict the direction of information flow.

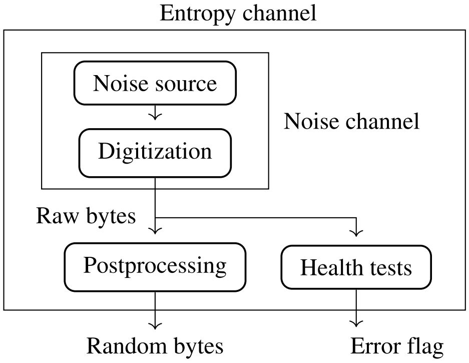


FIGURE 2. The naming convention employed for the random bytes' generation.

As shown in Fig. (1), the RAVA device consists of the following essential modules: microcontroller unit (MCU), boost converter, noise sources, and monitoring and interface headers.

The MCU serves as the central processing unit of the circuit. It encompasses a microprocessor, memory, input/output peripherals, and a communication interface. The MCU governs the circuit's operation by listening to user commands, conducting measurements and calculations, and returning data as requested. Its main task is to generate and send a certain number of random bytes once or repeatedly in a regular time interval. Optionally, it can engage in post-processing the random bytes output. The MCU executes health tests during circuit startup and over the generated byte sequences to identify errors and ensure the randomness quality. Additionally, the MCU generates a pulse width modulation (PWM) signal fed to the Boost converter – the

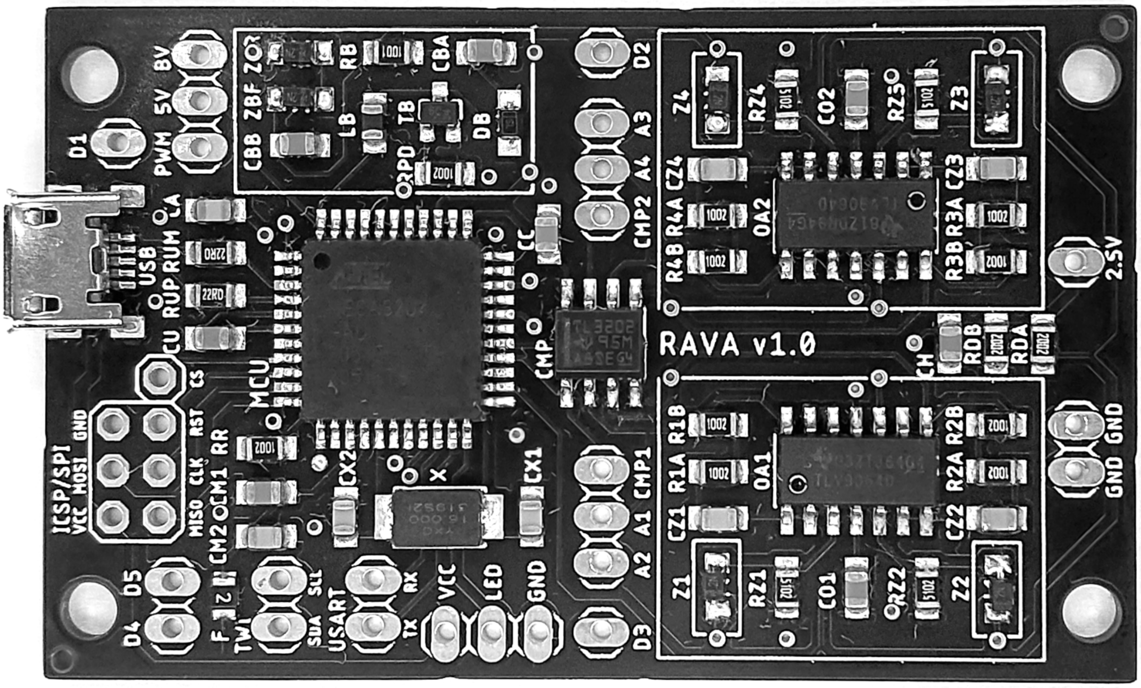


FIGURE 3. A photo of the RAVA circuit's implementation. The circuit measures 6 cm x 3,7 cm.

module responsible for increasing the USB 5V input into a higher voltage applied to the noise sources.

To describe the randomness components, the National Institute of Standards and Technology (NIST) convention is adapted, following the naming shown in Fig. (2). The NIST is a North American standards agency that provides, among other topics, recommendations on entropy sources and randomness tests for RNGs [18], [19].

The noise sources contain the fundamentally unpredictable physical processes responsible for entropy. Their output consists of digital signals characterized by rising edge pulses occurring at times that cannot be estimated using theoretical or empirical methods. The digitization follows by counting pulses in a specific interval and creating random bits associated with the counts' parity. The noise channel consists of raw bytes produced by the digitization step continuously monitored by health tests. If the raw bytes are biased, they can be post-processed into random bytes with enhanced entropy contents. When post-processing is not necessary, raw and random bytes are equivalent. It is labeled an entropy channel, combining the noise channel with the health tests and the optional post-processing.

The monitoring headers are strategically positioned ports within the circuit that grant access to crucial voltage levels. They serve multiple purposes, including diagnosing potential faults and analyzing/auditing the behavior of noise sources during circuit operation. The interface headers are ports that enable interaction with external devices, components, and sensors. They expand the circuit's functionality and the application range.

III. IMPLEMENTATION

This section describes one particular implementation of the general design previously discussed, resulting in the circuit shown in Fig. (3).

By inspecting the circuit's photo, one can identify the available headers. The voltage monitoring headers are labeled GND, 2.5V, 5V, PWM, and BV for the boost converter output. The noise source monitoring headers are labeled Ai for the four avalanche noise channels and CMPi for the two comparator outputs. The interface headers provide access to the following communication interfaces: ICSP (In-Circuit Serial Programming), TWI (Two-Wire serial Interface), SPI (Serial Peripheral Interface), and USART (Universal Synchronous and Asynchronous serial Receiver and Transmitter). Furthermore, the interface headers expose digital ports labeled as Di, which offer several peripheral features.

The upcoming subsections provide comprehensive details of the implementation. The circuit schematics are presented, showcasing the values of resistors and capacitors utilized. Details about the remaining components can be found in Table (1).

A. MICROCONTROLLER

The circuit's MCU choice is the ATmega32U4 [20] operating in a clock frequency of 16MHz. The ATmega32U4 is

TABLE 1. Components used in the RAVA circuit implementation.

Comp. name	Type	Value
C_x	Ceramic capacitor x F	±10% tolerance
CMP	Comparator 2ch	TLV3202 AIDR
F	Resettable fuse 500mA	BSMD0805-020-30V
$L_{22\mu}$	Power inductor 22μH	CB2012T220K
MCU	Microcontroller	ATMEGA32U4-AU
OA	Operational amplifier 4ch	TLV9064 IDR
Q	N-Channel MOSFET	2N7002
R_x	Resistor x Ω	±1% tolerance
S	Schottky diode	1N5819
USB	USB connector	U254-051N-4BH806
X	Crystal resonator 16MHz	X503216MSB2GI
Z_{24}	Zener diode 24V	ST MM3Z24

employed in various electronic projects, including Arduino boards. Its popularity provides several advantages, such as access to open-source libraries, extensive online documentation, and a supportive maker community.

The ATmega32U4 is an 8-bit MCU that provides: an arithmetic logic unit with 28 unique instructions, 32Kbytes of flash memory for storing the firmware, 2.5Kbytes of SRAM memory, a USB v2.0 controller that is used as the primary communication interface, four internal timer/counters with pulse width modulation (PWM), analog to digital conversion, and several communication interfaces.

The MCU wiring schematic is shown in Fig. (4) with the interface headers omitted – for more details, see [17].

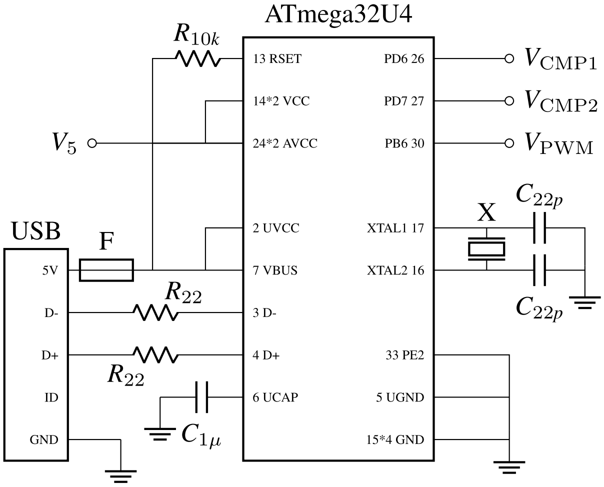


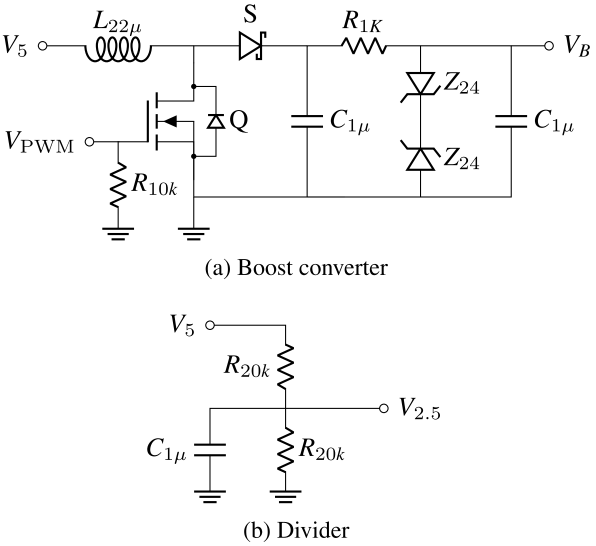
FIGURE 4. RAVA's MCU schematics including the USB connection. The communication interfaces connections are omitted.

B. POWER

The circuit primarily relies on V_5, the 5V power provided by the USB interface. To ensure a reliable power supply, decoupling capacitors are connected in proximity to the main IC components. They suppress high-frequency noise and provide local energy storage to mitigate voltage variations.

The boost converter module, illustrated in Fig. (5a), generates the V_B voltage necessary for producing the avalanche noise. It utilizes the V_{PWM} signal generated by the MCU to step up the V_5 input into the higher voltage level V_B. The boost circuitry follows a conventional design with an inductor, a MOSFET switching transistor, a Schottky diode, and a capacitor. Subsequently, it includes a resistor, two Zener diodes (one in forward and the other in reverse mode), and a capacitor. The resistor and Zener diodes function as a voltage regulator, ensuring that V_B remains within the desired range for generating the avalanche noise. The additional capacitor contributes to a cleaner and more stable voltage output.

The circuit includes a power divider component, as depicted in Fig. (5b), which generates $V_{2.5}$, a reference voltage of 2.5V.



(a) Boost converter

(b) Divider

FIGURE 5. RAVA's power modules schematics.

C. ENTROPY

The noise source schematics shown in Fig. (6) generate entropy through the following steps: First, the V_B voltage is applied to reverse-biased 24V Zener diodes, inducing avalanche breakdown. A 24V Zener is specifically chosen for the circuit due to its substantial noise amplitude of several hundred millivolts, which is not achievable with lower Zener voltage diodes. Next, the noise voltages are buffered using operational amplifiers (OA_1, OA_3). The purpose of the buffering stage is to prevent distortions that could be introduced in the subsequent steps. The noise voltages are then DC decoupled and raised to a common level of 2.5V using unity-gain operational amplifiers (OA_2, OA_4). These operations result in the *avalanche noise* channels V_{A1} and V_{A2} containing the original noise voltages, which have been inverted and raised to the 2.5V DC level. Finally, the analog channels V_{A1} and V_{A2} are connected to a comparator IC, which produces a digital output V_{CMP} representing which Zener produces the largest avalanche noise at a given time.

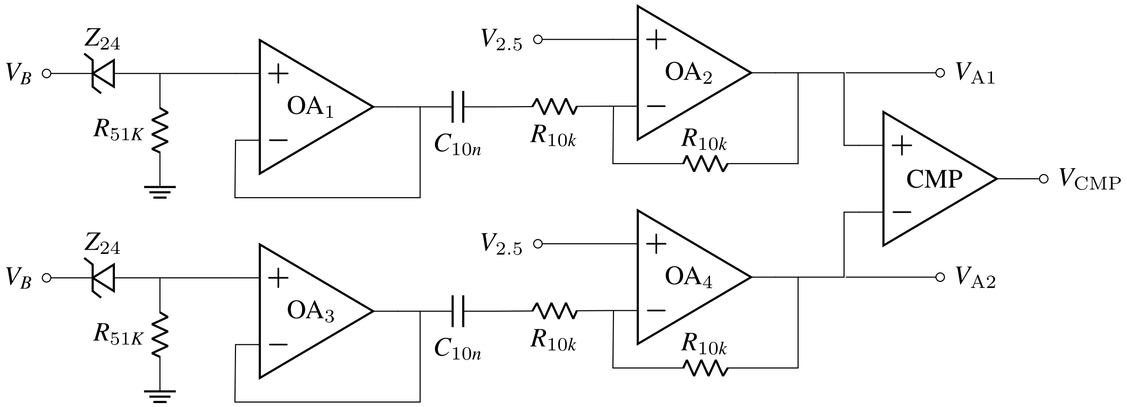


FIGURE 6. RAVA's noise source schematics.

The differential design consists in comparing the two independent avalanche noise channels, V_{A1} and V_{A2}, instead of comparing just one of them with its mean value of 2.5V. It mitigates predictable effects caused by environmental influences.

The V_{CMP} output, referred to as *differential noise*, consists of a sequence of pulses with varying lengths and unknown rising edge times at t_i. The interval between successive pulses, represented by $\Delta t_i = t_i - t_{i-1}$, depends on the avalanche breakdown occurring in the reverse-biased Zener diodes. Consequently, the Δt_i intervals are inherently unpredictable, serving as an entropy source for the circuit.

An example of the avalanche noise produced by the circuit's implementation is shown in Fig. (7), revealing the considerably large noise amplitude. The bottom part of the figure shows the mathematical simulation of the differential noise for this example. While a dual-channel oscilloscope could not simultaneously measure the actual signal, the simulation is sufficient to provide insight into how the comparator produces and sustains a digital pulse while $V_{A1} > V_{A2}$.

The noise source output V_{CMP} is wired to a timer/counter port in the MCU configured to count the measured rising edge pulses. Every *i*th-random bit is generated by evaluating the *pulse count*, labeled as n_i, after a fixed *sampling interval* denoted as t_s. The *i*-th bit results in 0 if n_i is even and 1 if odd. The Fig. (7) example reveals nine pulses in the sampling interval of $3\mu s$ that would result in a 1 bit.

Then, the steps for generating one random byte are: a) counting digital pulses in the t_s interval; b) detecting an odd pulse count and enabling the corresponding bit in the resulting byte; c) repeating a) and b) steps eight times; d) applying the generated byte to a continuous health monitoring algorithm described in the next section; e) sending the generated byte over the serial/USB interface.

The circuit contains two copies of the noise source module depicted in Fig. (6), establishing a dual entropy core architecture with two independent random byte channels. Within the dual architecture, two random bytes are generated in parallel.

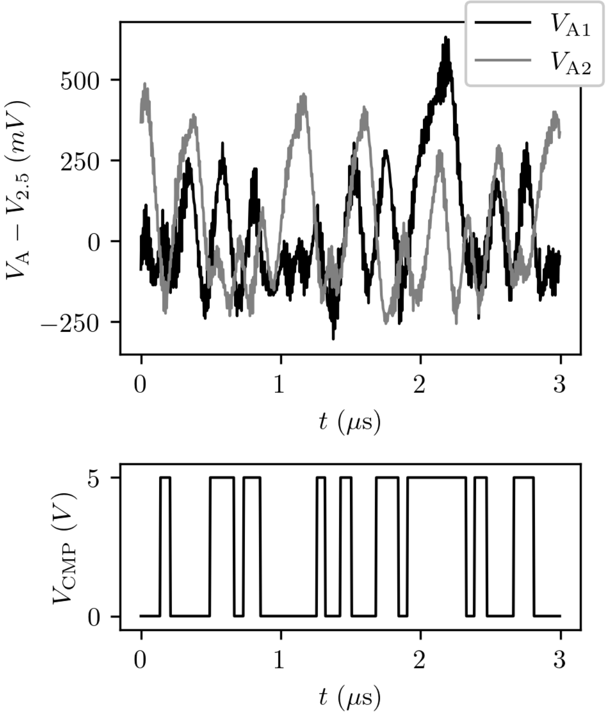


FIGURE 7. An example of the avalanche noise channels measured in a $3\mu s$ window by a dual-channel oscilloscope and the comparator's simulated response.

The bit generation in both channels is synchronized as the MCU timer/counters connected to $V_{\text{CMP}1,2}$ are sequentially zeroed and read after the same t_s delay.

D. HEALTH TESTS

The RAVA firmware implements health monitoring tests that adhere to the NIST requirements outlined in the "Recommendation for the Entropy Sources" [18] document. Upon powering up the circuit, startup tests are executed to assess the proper functioning of the noise sources. If the initial tests are successful, the circuit becomes ready to receive commands and generate random bytes; otherwise, it communicates the failure and rejects user commands. The startup tests evaluate the probability distributions of the 2-valued bits, the 256-valued bytes, and the average pulse count numbers.

In addition, continuous tests are conducted for every generated byte while the noise source is operational. The firmware implements two recommended tests: repetition count and adaptive proportion. The first detects catastrophic failures that may cause the noise source to become ''stuck'' on a single output value for a long period. The second detects a loss of entropy that might occur due to some physical failure or external factors affecting the noise source. Continuous tests' errors do not disable the randomness generation. Instead, the user is informed of the errors, allowing them to take appropriate action based on the failure rate.

IV. CHARACTERIZATION

This section discusses the implementation of a RAVA circuit using the specific layout and components outlined in section III. Once the circuit's hardware has been established, three free parameters must be defined to proceed with the random byte generation: PWM frequency f_{PWM}, PWM duty cycle d_{PWM}, and sampling interval t_s. The following subsections show the criteria to determine these values and the resulting noise characteristics. Moreover, a stochastic model is introduced to provide further insight into the system's behavior.

A. PWM CONFIGURATION

The MCU port providing the PWM capability allows the selection of various frequencies. The value chosen is $f_{\text{PWM}} = 46.9\text{kHz}$ as it enables the desired voltage outcome while keeping a relatively low frequency that minimizes interference with other circuit components.

In order to determine the duty cycle parameter, the relationship between the pulse count and the circuit's current consumption is examined. The pulse count N is a random variable with particular values $\{n_1, n_2, \ldots, n_i\}$, where a pulse count average is defined as $\bar{n} = 1/k \sum_i^k n_i$. The n values varies as a function of the sampling interval t_s, satisfying the following inequation

$$\sum_i^n \Delta t_i \leq t_s < \sum_i^{n+1} \Delta t_i. \tag{1}$$

Fig. (8) presents $\bar{n}$ and the circuit's current consumption c for different d_{PWM} values and an arbitrarily large sampling interval chosen as $t_s = 20\,\mu\text{s}$. The results reveal two different regions of current consumption. In the first region, the c increase leads to higher $\bar{n}$, implying that the power generated by the boost converter module is being converted into avalanche noise. However, as d_{PWM} exceeds 10%, a second region is observed. In this region, $\bar{n}$ reaches a plateau while the current consumption increases at a higher rate, implying the additional power being dissipated.

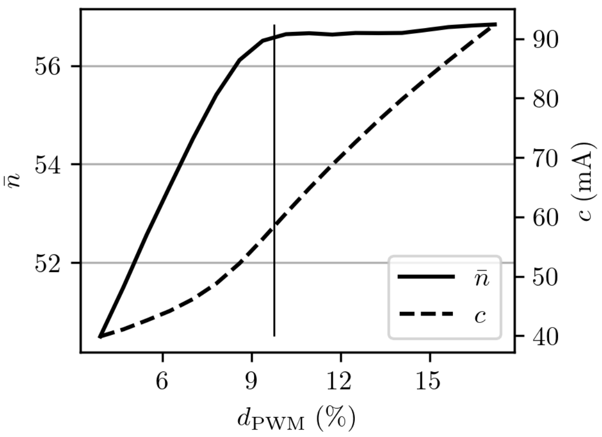


FIGURE 8. The relationship between PWM duty cycle, pulse count average, and circuit's current consumption at $f_{\text{PWM}} = 46.9\text{kHz}$ and a sampling interval of 20μs. The vertical line depicts the chosen value of $d_{\text{PWM}} = 9.8\%$.

After considering the relationship between pulse count and current consumption, a specific duty cycle value of $d_{\text{PWM}} = 9.8\%$ is chosen. This value balances achieving the maximum pulse count while maintaining a low current consumption.

At $f_{\text{PWM}} = 46.9\text{kHz}$ and $d_{\text{PWM}} = 9.8\%$ (values utilized throughout subsequent analysis), the circuit consumes 58mA, while the boost converter module yields an output voltage of $V_B = 25.5\text{V}$, accompanied by a current of 1.5mA flowing through its resistor.

B. FREQUENCY SPECTRUM

A 150MHz oscilloscope with Fast Fourier Transform (FFT) capability is used to measure the frequency spectrum of the avalanche and differential noise channels. The measurements were performed multiple times, and the average result is shown in Fig. (9). The frequency spectrum analysis reveals a white noise band in both channels up to 3.3MHz. Beyond this frequency, V_{CMP} exhibits a $1/f^2$ red noise. The spikes ranging from 20 to 100MHz in V_{A} are attributed to Radio frequencies. The results demonstrate the remarkable effectiveness of the differential design in minimizing the impact of electromagnetic interferences on the V_{CMP} channel, where the same disturbances are suppressed.

C. DIFFERENTIAL NOISE CHARACTERISTICS

Although individual Δt_i intervals vary unpredictably, it is possible to establish an average interval defined as

$$\overline{\Delta t} = \lim_{n \to \infty} \frac{1}{n} \sum_i^n \Delta t_i, \tag{2}$$

along with a mean frequency, calculated as $\bar{f} = 1/\overline{\Delta t}$. An oscilloscope is utilized for measuring those quantities in the RAVA's implementation by probing the V_{CMP} channel, leading to an average frequency of $\bar{f} = 3.2\text{MHz}$ and average interval of $\bar{\Delta t} = 313\text{ns}$.

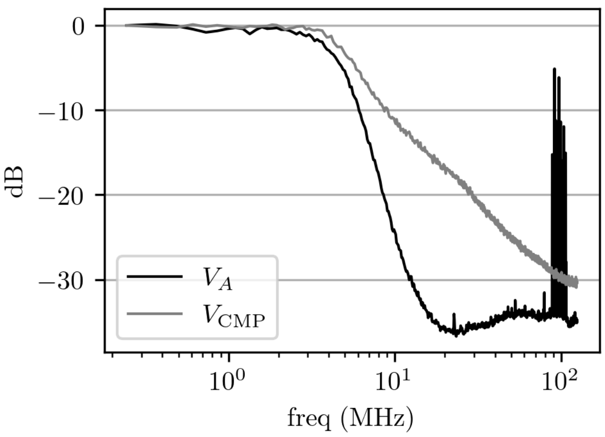


FIGURE 9. Frequency spectrum of the avalanche and differential noise channels.

Considering the differential channel, it is important to recognize a measurement limitation when connected to the timer/counter peripheral of the RAVA MCU. According to the datasheet [20], the peripheral does not accurately compute counts when the frequency between consecutive pulses is higher than the MCU clock frequency divided by 2.5. In other words, the timer/counter fails to register counts when $\Delta t < 156$ns. As a result, the RAVA circuit is anticipated to yield a lower count average than the oscilloscope. However, this discrepancy does not affect the output's entropy. Its sole impact is a reduction in the device's throughput.

Now, let us explore a different scenario where random bit generation would be achieved by periodically measuring the V_{CMP} port and assigning the bit value based on the port's digital state. In an ideal system, the avalanche channels would display similar voltage distributions, resulting in the V_{CMP} channel spending, on average, an equal amount of time in the 5V and 0V states. However, deviations in circuit component properties are natural and expected in practical implementations. These variations introduce unavoidable asymmetries, leading to an unreliable strategy contaminated by bias. In order to maximize the device's entropy, an alternative approach is employed. Rather than directly measuring the port's state, the device exploits the time uncertainty of *when* the state transitions occur. More specifically, counting the number of transitions in fixed sampling intervals, as previously discussed.

D. STOCHASTIC MODEL AND BIAS

As highlighted in [12], the pulse counting methodology's advantage is based on its intrinsic connection with the Central limit theorem (CLT) in Statistics, elucidated and deepened as follows.

The CLT states that, given certain conditions, the distribution of the sum of independent and identically distributed random variables will tend towards a normal distribution. The normal approximation holds regardless of the shape of the original distribution, provided the sum quantity is sufficiently large.

In our case, the fundamental distribution is the time associated with a single pulse count. This is represented by the random variable $\Delta_1 T$ with specific values $\{\Delta t_i\} = \{t_1 - t_0, t_2 - t_1, \ldots\}$, where t_i, as usual, denotes the time of the ith-rising edge pulse measured in the differential noise channel. Obtaining a model for the $\Delta_1 T$ distribution is challenging due to the reliance on the unique characteristics of the noise sources and the efficiency curves of the measuring components. These characteristics may vary across different instances of the same design, further complicating the task of establishing the distribution's parameters.

Let us introduce another random variable, $\Delta_2 T$, representing the time associated with two pulse counts. The specific values of $\Delta_2 T$ are given by $\{t_2 - t_0, t_4 - t_2, \ldots\}$. These values can be further expressed as $\{\Delta t_1 + \Delta t_2, \Delta t_3 + \Delta t_4, \ldots\}$. Therefore, the values of $\Delta_2 T$ are obtained by adding two values that follow the fundamental distribution. The generalization for n pulse counts leads to the insight that as n increases, ΔT_n distribution's tends to a normal curve. This result is a direct consequence of the CLT, providing a significant generalization for the ΔT_n probability distribution.

Rather than time, the variable utilized in the circuit is N, the number of pulse counts within the sampling interval t_s. The relationship between $\Delta_n T$ and N, as derived from Eq. (2), follows a linear form mediated by the constant $\overline{\Delta t}$. As a result, both variables follow the same distribution, and the N distribution also tends to a normal curve. This approximation is valid when the sampling interval t_s is sufficiently large, allowing for a substantial number of pulse counts to be accumulated.

Therefore, the RAVA's stochastic model for a compliant t_s can be summarized as $N \sim \mathcal{N}(n;\, \bar{n},\, \sigma)$, indicating that N follows a normal distribution with a mean of $\bar{n}$ and a standard deviation of σ. While the normal distribution is a continuous curve, the independent variable n assumes integer values in this case.

With knowledge on the probability distribution governing N, it is possible to estimate the theoretical bias in the conversion from pulse counts to parity, which is the step responsible for assigning the bit value. The bias, denoted as ϵ, arises from comparing the probabilities of obtaining even and odd n values. Mathematically, it is expressed as:

$$\epsilon(\bar{n}, \sigma) = \frac{1}{2}\left[\sum_n \mathcal{N}(2n;\, \bar{n},\, \sigma) - \sum_n \mathcal{N}(2n+1;\, \bar{n},\, \sigma)\right], \tag{3}$$

where $n = 0, 1, 2, \ldots$. The function $\epsilon(\bar{n}, \sigma)$ represents the minimum achievable bias by a circuit's implementation modeled as $N \sim \mathcal{N}(\bar{n},\, \sigma)$.

The numerical computation of $|\epsilon|$ is presented in Fig. (10). The results demonstrate that when $\bar{n} \geq 15$ and $\sigma \geq 1.8$, the minimum bias is below 10^{-7}. Within the depicted range,

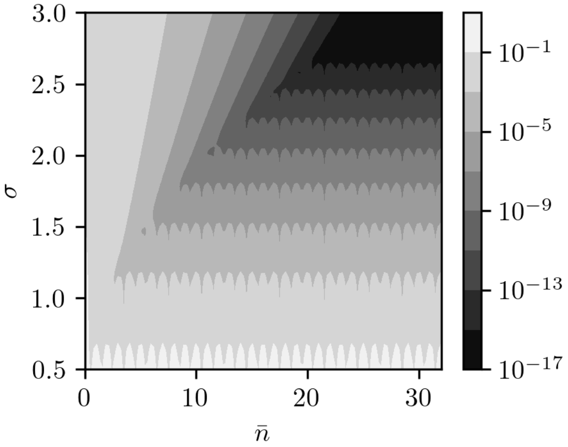


FIGURE 10. Numerical estimation of the theoretical minimum bit bias.

it takes an average of 10 million generated bits (or more for larger parameter values) to produce at least one biased bit. Beyond this range, the theoretical maximum bit entropy approaches the value of two, rendering additional post-processing algorithms unnecessary.

E. SAMPLING INTERVAL AND THROUGHPUT

The criteria for selecting t_s is finding a value that yields sufficiently large pulse count average $\bar{n}$, satisfying the CLT requirements and enabling $N \sim \mathcal{N}(n;\ \bar{n},\ \sigma)$.

This study is implemented by varying t_s, measuring 10K pulse counts, and fitting their distribution to a normal curve. The fitting procedure aims to determine the optimal parameters, $\bar{n}$ and σ, that describe the observed N distribution. A least squares procedure is employed, resulting in a χ^2 value and an associated probability, denoted as p, which indicates the likelihood of the observed N distribution being derived from a normal curve. The fitting procedure is repeated a thousand times for each t_s, generating distributions for the normal parameters with mean values $\bar{\bar{n}}$ and $\bar{\sigma}$. The distribution of the 1K p values is then analyzed, characterized by the mean $\bar{p}$ and the standard deviation σ_p. When the normality condition for N is met, the p distribution is expected to follow a uniform distribution with $\bar{p} = 0.5$ and $\sigma_p = 1/\sqrt{12}$.

The study findings are presented in Fig. (11). The upper part depicts the obtained $\bar{p}$ along with their associated σ_p bars. The results demonstrate that as t_s increases, the $\bar{p}$ values converge towards 50%, while the standard deviation bars tend to align with the horizontal dashed lines indicating $1/\sqrt{12}$. In conclusion, as t_s increases, the N distribution tends towards normality, providing empirical evidence supporting the CLT connection discussed earlier. The lower part of Fig. (11) displays the resulting $\bar{\bar{n}}$ and $\bar{\sigma}$ values obtained for each t_s.

The sampling interval t_s is chosen as 10μs, ensuring that the N variable follows a normal distribution. While an interval of 5μs seems sufficient, selecting a larger value provides a lower bias and a safety margin for all circuit implementations to use the same value consistently. With the

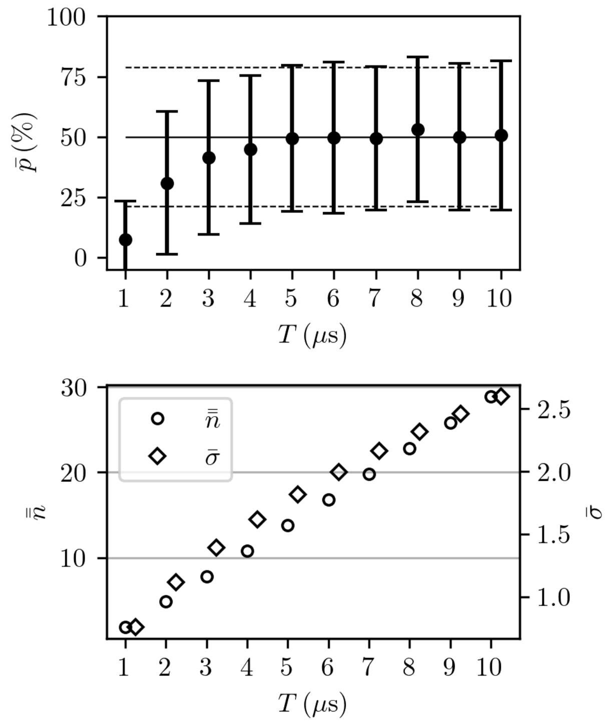


FIGURE 11. A study investigating the compatibility of the pulse count variable with a normal distribution for increasing sampling intervals.

$t_s = 10\mu$s selection, the resulting distribution is given by $N \sim \mathcal{N}(n;\ 28.9,\ 2.6)$. Based on this distribution, the theoretical minimum bias ϵ is estimated to be on the order of 10^{-15}, implying an extremely low bias and further validating the suitability of the chosen sampling interval.

For $t_s = 10\mu$, the five steps outlined in Section III-C to produce a single byte require an average time of 117.7μs to complete. This corresponds to a single channel throughput of 68.0 Kbit/s. Considering the two entropy channels combined, the overall throughput achieved by the RAVA circuit is 136.0 Kbit/s.

V. RESULTS

This section considers a RAVA circuit implemented within the layout and values described in Section III and with the parameters $f_{\mathrm{PWM}} = 46.9$kHz, $d_{\mathrm{PWM}} = 9.8\%$, and $t_s = 10\mu$. In the following subsections, statistical analyses are performed to assess the randomness of the generated bytes. For comparison, a commercial device from ID Quantique is selected as a control, and the same tests are executed on this device.

The Quantis USB device utilizes an optical quantum process as its randomness source, enabling a throughput of 4Mbits/s while consuming 73.7 mA. As described in [21], its noise source comprises a light-emitting diode, a semi-transparent mirror, and two single-photon detectors to record the which-path outcomes. As the raw bytes can exhibit a bias of up to 5%, a post-processing algorithm is employed within the device's processing unit to enhance their entropy.

The data for the first three subsections comprises a total of six files, each containing 125M random bytes – each device producing one file for each subsection. In those tests, the file's data are spit into 1K samples of 1Mbits.

A. BIAS AND SERIAL CORRELATION

The first test evaluates the bias in the bit and byte levels, as well as the serial correlation between adjacent bits. The test outcomes are presented in Fig. (12), showing the distribution of the 1K test results performed with n = 1Mbits each. The bit bias and the serial correlation distributions are normal, as informed by the Shapiro-Wilk test. The byte bias follows a χ^2-distribution with 255 degrees of freedom.

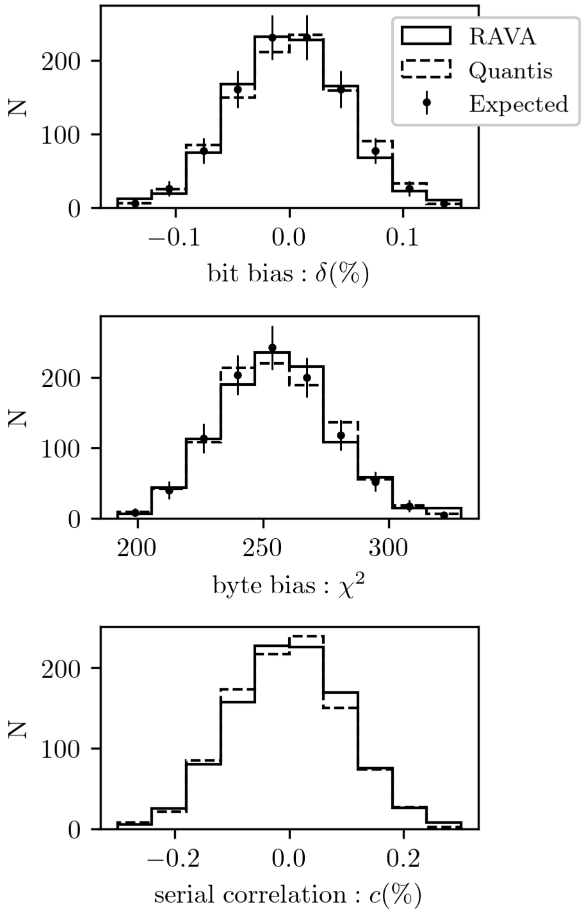


FIGURE 12. Bias and serial correlation distributions for the RAVA and Quantis circuits.

The bias at the bit level is given by

$$\delta = \frac{n_1}{n} - \frac{1}{2}, \tag{4}$$

where n_1 represents the count of 1s obtained in n = 1M random draws with a probability of 50% each. The variable n_1 follows a binomial distribution. For large values of n, the binomial distribution can be approximated by a normal distribution. With the relationship $z = 2\delta\sqrt{n}$ between the z-score and the bias, the δ distribution of 1M unbiased samples is described by a normal curve $\mathcal{N}(0, 0.05\%)$ centered at $\bar{\delta} = 0$ and with $\sigma = 1/(2\sqrt{n})$.

Black circles in the upper part of Fig. (12) represent the unbiased normal distribution expected in 1K tests, including the 95% confidence intervals. The solid and dashed lines depict how the devices' distributions align with the expected values. A least squares procedure is used to find the parameters that best describe the δ distributions. This approach determines the most suitable mean and standard deviation values, characterizing the normal distributions that match the data. The fit results, considering 95% confidence intervals, are as follows:

- RAVA circuit: $\bar{\delta} = -0.0005\% \pm 0.0032\%$, $\sigma = 0.0499\% \pm 0.0025\%$, $p = 32\%$.
- Quantis circuit: $\bar{\delta} = 0.0019\% \pm 0.0029\%$, $\sigma = 0.0502\% \pm 0.0020\%$, $p = 62\%$.

The p-values resulting from the least square procedure inform the likelihood of the observed distributions being derived from a normal curve. The results indicate that both devices generate unbiased bits, as the distributions exhibit a normal shape, with mean bias values compatible with zero and standard deviations compatible with 0.05%.

A byte consists of 8 bits, allowing for $2^8 = 256$ unique values. The bias at the byte level is evaluated by analyzing n_i, the number of bytes from the 125K-byte sample representing each unique category $i = 1, \ldots, 256$. The byte bias is assessed using the χ^2 test between the measured n_i and the expected $n_e = 125K/256$ values, calculated as

$$\chi^2 = \sum_{i=1}^{256} \frac{(n_i - n_e)^2}{n_e}, \tag{5}$$

where the test assumes an equiprobable state for each category, as expected in truly random data. The χ^2 values of unbiased samples are expected to follow a χ^2-distribution with $d = 255$ degrees of freedom.

While the bit bias δ metric quantifies the balance of 0s and 1s in a sequence of draws, the byte bias χ^2 metric goes further by incorporating the bit ordering as relevant information. To illustrate, let's consider the bit sequences 00001111 and 01010101; they yield the same δ value despite representing two different byte categories. The 256-byte categories are only equiprobable when the chance of measuring a 1 bit is the same as obtaining a 0 bit, i.e. when $\delta \rightarrow 0$. Therefore, the byte bias is a complementary test that encompasses the bit bias while also assessing the bit pattern variations over time.

In the middle part of Fig. (12), the unbiased χ^2 distribution is represented by black circles, while the solid and dashed lines depict the devices' distributions. A least squares procedure is employed to determine the actual degrees of freedom from the data, resulting as follows:

- RAVA circuit: $d = 254.9 \pm 1.5$, $p = 23\%$.
- Quantis circuit: $d = 254.6 \pm 1.7$, $p = 11\%$.

These values demonstrate that both devices generate unbiased bytes, as the distributions follow a χ^2-distribution aligning with the expected d value of 255.

The serial correlation measures the degree to which a bit in a sequence depends on the previous bit. It is computed as

$$c = \frac{N\sum_{i=1}^{N} b_i b_{i+1} - (\sum_{i=1}^{N} b_i)^2}{N\sum_{i=1}^{N} b_i^2 - (\sum_{i=1}^{N} b_i)^2}, \tag{6}$$

where i ranges from 1 to N, representing the index of the bit in the sequence, and b_i denotes the bit value (0 or 1) of the ith-bit. The serial correlation ranges from -1 to 1 and tends to zero when applied to truly random and independent samples.

In Fig. (12), the lower part show the devices' c distributions. A least squares procedure is employed to determine the normal parameters, resulting as follows:

- RAVA circuit: $\bar{c} = 0.0023\% \pm 0.0047\%, \quad \sigma = 0.0971\% \pm 0.0034\%, \quad p = 88\%$.
- Quantis circuit: $\bar{c} = -0.0017\% \pm 0.0045\%, \quad \sigma = 0.0958\% \pm 0.0032\%, \quad p = 91\%$.

The results indicate that both devices generate independent bits, as the extracted serial correlation values are compatible with zero.

The correlation test based on Eq. (6) is also applied to analyze the correlation between the bits simultaneously generated by the two entropy channels within the RAVA circuit. An additional file of 125Mbytes parallelly produced by the second channel is utilized for this analysis. The normal parameters resulting from the least squares procedure are

- RAVA cores: $\bar{c} = 0.0031\%\pm0.0053\%, \quad \sigma = 0.1036\%\pm 0.0038\%, \quad p = 80\%$.

These values demonstrate that the RAVA's entropy channels produce parallel bits that are independent.

B. ENTROPY ESTIMATION

This test evaluates the devices' entropy based on the guidelines outlined in the NIST document "Recommendation for Random Number Generation Using Deterministic Random Bit Generators" [18]. The chosen metric is the min-entropy, which represents the uncertainty in predicting a byte value and is calculated as

$$h = -\log_2(\max p_i), \tag{7}$$

where $\max p_i$ represents the most probable category among the 256 unique byte values. If a device generates random bytes with h, it implies that the probability of observing any particular byte value is no greater than 2^{-h}. It's worth highlighting that $\max p_i$ arises from the interplay between the theoretically minimum bias discussed in Section IV-D and inherent statistical variations linked to the evaluated sample size, 125 KBytes in this test.

The NIST entropy estimation involves two distinct procedures, one considering the input as an independent and identically distributed (IID) sample and a more conservative approach considering the input as generated by a non-IID source. In the IID procedure, an estimate is obtained by finding $\max p_i$, constructing a 99% confidence interval for this value, and applying the upper p value into Eq. (7) to calculate the min-entropy. In addition, the IID procedure also includes permutation and chi-square tests to evaluate the IID assumption of the input data. The non-IID procedure applies ten different estimators to the input dataset, and the minimum of all the estimates is taken as the entropy assessment of the entropy source.

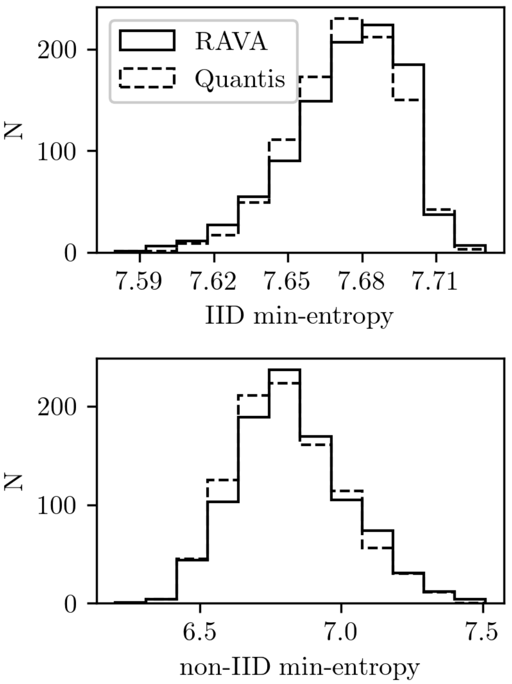


FIGURE 13. NIST min-entropy distributions. The distributions reflect the 1K min-entropy values obtained for each 1Mbits sample.

The results presented in Fig. (13) are obtained using the NIST-provided software. The min-entropy distributions are obtained from 1K tests with 125 KBytes each. The mean and standard deviation of the distributions resulting from the IID tests are as follows:

- RAVA circuit: $\bar{h} = 7.673, \quad \sigma = 0.023$.
- Quantis circuit: $\bar{h} = 7.673, \quad \sigma = 0.021$.

For the non-IID tests, the corresponding statistics are:

- RAVA circuit: $\bar{h} = 6.80, \quad \sigma = 0.25$.
- Quantis circuit: $\bar{h} = 6.79, \quad \sigma = 0.23$.

The IID assumption fail rate is 3.4% in the RAVA circuit and 4.7% in the Quantis circuit. While the distributions are not normal, both devices follow the same IID and non-IID distributions, as confirmed by the nonparametric Mann-Whitney U-test. The results indicate that both devices exhibit similar distributions of min-entropy according to the NIST methodology.

For completeness, min-entropy is also calculated using the standard definition of Eq. (7), which does not incorporate the upper bound of the p_i confidence interval as in the NIST IID metric. The mean and standard deviation of the resulting entropy distribution are as follows:

- RAVA circuit: $\bar{h} = 7.823, \quad \sigma = 0.024$.
- Quantis circuit: $\bar{h} = 7.823, \quad \sigma = 0.022$.

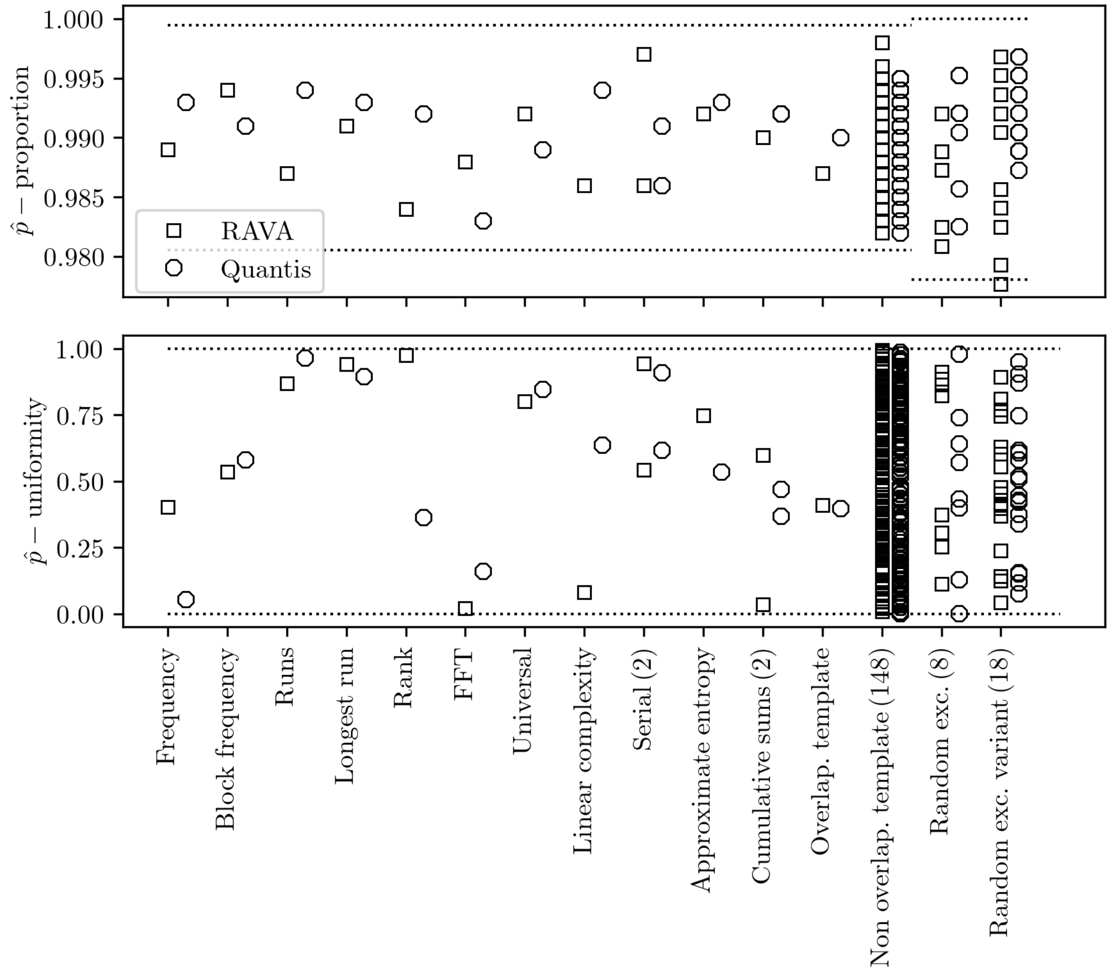


FIGURE 14. Results of the NIST randomness test suite. Some tests are repeated with several variations shown as the parenthesis number after the test name. The dotted lines represent the confidence intervals. The last two tests have additional criteria that lead to a reduction in the total number of tests performed and a consequent adjustment in the confidence interval for the$\hat{p}$proportion.

C. RANDOMNESS TESTS

To assess randomness, NIST developed a comprehensive test suite consisting of 15 statistical tests, as described in [19]. Each test takes a sample of n_b bits as input and produces a p-value that evaluates the null hypothesis of randomness. In this context, if the p-value exceeds the significance level of $\alpha = 1\%$, it indicates that the sample is considered random with a 99% confidence level. The tests are repeated n_t times, resulting in sequences of p-values which are evaluated based on two metrics: proportion and uniformity. The proportion metric measures the percentage of p-values that surpass the significance threshold, while the uniformity metric evaluates the distribution of the p-values across the test suite.

The proportion metric for a given test is calculated as follows: it starts by computing the number of tests yielding a $p-$value above $\alpha = 1\%$, a quantity denoted as n_r; the proportion of tests that conform to the randomness hypothesis is then obtained as $\hat{p} = n_r/n_t$; a 99.73% confidence interval for the proportion is computed as $\hat{p} \pm 3\sqrt{\alpha(1-\alpha)/n_t}$. The proportion results are depicted in the upper part of Fig. (14). It can be observed that the majority of test proportions for both devices fall within the confidence interval, indicating compliance with the randomness hypothesis.

The uniformity metric evaluates if the p-values distribution in a given test is uniform as expected in a random scenario. The uniformity is determined by partitioning the p-values in 10 intervals and obtaining a χ^2 value that compares the partitions' occupation with the expected $n_t/10$ value. A $\hat{p}$-value is obtained by applying the χ^2 value to the cumulative distribution function with 9 degrees of freedom. The samples are considered uniformly distributed if $\hat{p} \geq 0.01\%$, as stated in the NIST documentation. The uniformity results are pictured in the lower part of Fig. (14). It can be observed that the majority of tests for both devices surpass the threshold, indicating conformity to the randomness hypothesis.

D. ENVIRONMENTAL INFLUENCES

This subsection is dedicated to the impact of environmental factors on the RAVA circuit's operation and the role of the differential design.

The frequency spectrum analysis presented in Section IV-B provided insights into the effects of electromagnetic radiation. It revealed that the circuit captures Radio frequencies ranging from 20-100 MHz, causing low amplitude interference in the avalanche noise channels. Remarkably, the same perturbations do not affect the differential noise channel.

This empirical finding can be understood as follows. Let $\Delta V(t)$ represent the time-varying voltage induced by radiation. The operation within the comparator IC can then be described as:

$$V_{A1}(t) + \Delta V(t) >_? V_{A2}(t) + \Delta V(t). \tag{8}$$

Since the ΔV factor originates from a source distant enough to equally impact both avalanche channels, it is subtracted during the comparison step that produces the V_{CMP} output. Such an outcome arises as a direct consequence of the differential design, showcasing its ability to isolate the avalanche breakdown as the exclusive source of entropy in the system. This property extends to various environmental influences such as sound, vibration, luminosity, electric and magnetic fields, and temperature.

Next, the effects of temperature are explored in greater detail. Regarding the device's operating range, a review of the components specifications results in an overlap from −40°C to 85°C, establishing the circuit's operation within the so-called industrial temperature range.

An empirical study is developed to investigate the impact of temperature on pulse count and bias measurements. Three basic measurements are repeated within an interval of 20 minutes, resulting in 76 data points containing:

- The temperature T measured by a DS18B20 digital temperature sensor coupled to the RAVA circuit.
- The average pulse count $\bar{n}$ resulting from 10K measurements.
- An amount of 125K random bytes are generated, resulting in a bit bias δ value, and a byte bias χ^2 value – respectively obtained by the use of equations (4)) and (5).

Over the experimental course, the circuit experiences temperature variations as it transitions between environments, moving from a freezer at −8°C, to ambient temperature, and an oven at 90°C.

The first graph in Fig. (15) depicts the temperature variations over the 20-minute duration. The circuit was initially exposed to each environment for 3 minutes and then alternated between the oven and the freezer to maximize the temperature gradient. The device's operation under those conditions provides an indication of its compliance with the industrial temperature range.

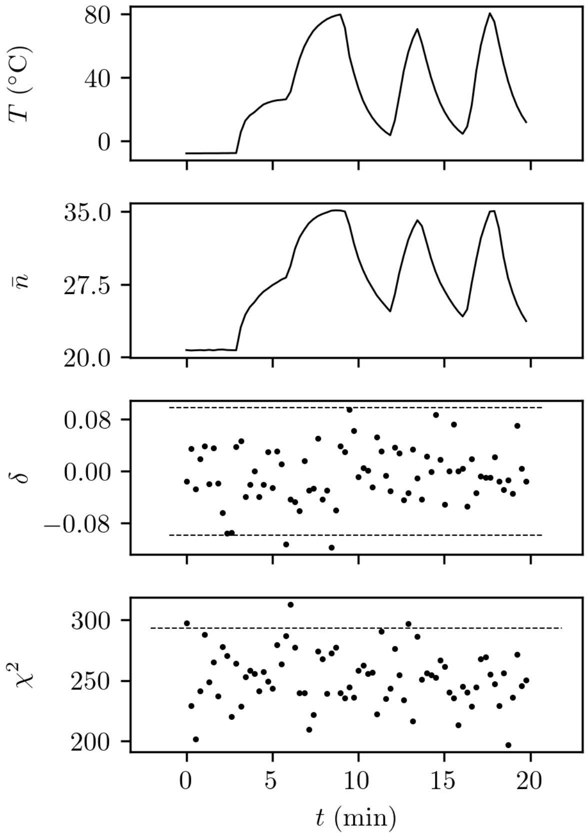


FIGURE 15. Temperature variation study including pulse count average, bit bias, and byte bias measurements. The dashed lines represent the 95% confidence intervals.

The second graph unveils the influence of temperature in the Zener diodes. It shows how higher temperatures increase the avalanche breakdown events, resulting in a higher average pulse count. The similarity with the temperature plot indicates a fast response and a linear behavior of the diodes when exposed to temperature gradients.

The third and fourth graphs reveal the bias outcomes following the temperature variations. The quantity of values falling outside the 95% confidence intervals aligns with the 3.8 expected by chance. This result implies that the device can operate in different temperatures without any discernible impact on its entropy output. Moreover, Pearson's correlation tests were conducted between the bias and temperature, as well as between the bias and temperature gradient, revealing the variables' independence in all tests.

The results indicate that although Zener diodes are sensitive to temperature variations, no conditions such as extreme temperatures or fast variations are able to bias the devices's outcomes.

VI. DISCUSSION AND CONCLUSION

This paper introduces RAVA, an open-source TNRG that employs reverse-biased Zener diodes as its entropy source. The manuscript presents the general architecture and a specific implementation that realizes the concept. The criteria for determining the three essential parameters governing the circuit's operation are outlined. The noise source is thoroughly characterized, and a stochastic model is introduced to describe the probability distribution of the main variable, the pulse count. The paper concludes with the results of statistical tests assessing the randomness quality.

The statistical tests applied to the RAVA's random bytes output are also applied to a Quantis device from ID Quantique, which extracts its entropy from a quantum optical process. The results show that both devices produce unbiased and independent bit sequences that pass in the NIST randomness test suite. The results reveal similar distributions for the two devices in all the studied metrics (bias, serial correlation, NIST's min-entropy), showing that given two random byte sequences produced by each circuit, no metric was found that could distinguish the data sources.

The physical phenomenon associated with the RAVA circuit's entropy is the avalanche breakdown of reverse-biased diodes and the time unpredictability of those events. While those processes can be initially seen in a macro-oriented view as the effect of electromagnetic fields applied to many particles, at its most fundamental level, the physical modeling of such process reaches the quantum nature of the charge carriers, inheriting a fundamental indeterminacy of *when* a particle will trigger an avalanche event.

While the textbook interpretation of quantum mechanics assumes nature as intrinsically random, deterministic interpretations are also viable, as seen in [22]. Independent of the broader metaphysical discussion, it seems enough to assume that the entropy source in the RAVA circuit is associated with a fundamental unpredictability/uncomputability of its underlying physical system, an empirical fact shared among different quantum interpretations.

By employing high V_Z Zener diodes rated at 24V, the avalanche noise amplitude reaches several hundred millivolts, as illustrated in Fig. (7). Such amplitude level is achieved without any amplification methodology, rendering the noise less susceptible to electromagnetic radiation and signal injection attacks. As the avalanche noise dominates other noise sources, it is possible to conclude that the device's entropy is predominantly derived from the avalanche process.

While the discovery of avalanche noise in reverse-biased Zener diodes dates back to the 1970s, it is important to emphasize that its choice as a noise source in the RAVA device was deliberate and motivated by its qualities. Specifically, Zener diodes enable two fundamental characteristics of the circuit: auditability and the implementation of an analog differential design. The use of Zeners allows for isolating the noise source within a discrete component, providing physical access for direct monitoring and even replacement in the event of fault detection. In contrast, the unpredictable physical events on FPGA chips, light sensors, and most modern designs occur deep within the intricate layers of the electronic components comprising the system. In such instances, the randomness machine operates as a black box system, preventing users from scrutinizing the intermediate processes and obstructing the establishment of a prior degree of belief in the digital output's quality. Furthermore, the differential design implemented with Zener diodes operates at the analog level, affording it the advantages of swift response times and more precise responses compared to what could be achieved after digital conversion.

The random bit generation initiates by feeding two independent avalanche noise channels into a comparator IC. The comparator produces a digital output, referred to as differential noise, which indicates the largest input at a given time. This differential design has the capability of mitigating environmental influences that equally affect both avalanche channels as detailed in Section V-D.

The circuit implementation includes an ATmega32u4 microcontroller with timer/counter peripherals connected to the differential noise channels. The random bit generation proceeds by counting the rising edge pulses received during a sampling interval and deriving the bit value based in the pulse count parity. In Section IV-D, it is demonstrated that for sufficiently large sampling intervals, the pulse count distribution adheres to a normal curve. This result, which underpins the noise source's stochastic model, is theoretically derived from the Central limit theorem in Statistics. To empirically validate the stochastic model, Section IV-E obtains pulse count distributions for increasing sampling intervals while fitting a normal curve to the data. As anticipated, with the increase in sampling interval, the pulse count distributions become progressively more aligned with a normal pattern, reinforcing the validity of the stochastic model.

One application of the stochastic model is determining the theoretical minimum bias. This involves subtracting the probability of obtaining an even pulse count from the probability of obtaining an odd pulse count. The numerical result, which relies on the normal distribution parameters, is visually represented in Fig. (10). Furthermore, by linking the bias with the normal parameters obtained for increasing sampling times, as depicted in Fig. (11), it is demonstrated that the RAVA device can achieve a reliable entropy level without the need for post-processing algorithms.

By providing the physical reasoning of the unpredictability factor behind the entropy source, implementing startup and continuous randomness health tests shown in section III-D, and estimating the IID assumption fail rate and min-entropy measures shown in section V-B, the RAVA device fulfills the key NIST compliance requirements [18]. Moreover, by presenting a stochastic model that provides entropy bounds, the RAVA circuit also conforms to more stringent standards as the BSI's AIS 31 [23], and ITU-T's X.1702 [24]. Meeting industry standards and possibly attaining official certifications may further enhance RAVA's trustworthiness.

The RAVA implementation here presented achieves a throughput of 136.0 Kbit/s. While other devices employing different noise sources can achieve throughputs in the millions or even billions of bits per second, the RAVA device remains well-suited for a variety of applications, as discussed in the Introduction Section. Notably, it finds valuable use in personal privacy, scientific research, and projects within education, arts and the maker community.

If a given application requires a higher throughput, it can be initially achieved by reducing the sampling interval. For instance, with a sampling interval of $t_s = 5\mu$s, it is possible to attain 204.8 Kbit/s. Further improvements require upgrading the circuit implementation. Two key approaches for hardware-level improvement include using lower V_Z Zeners and employing a microcontroller with a higher clock rate. Lower V_Z Zeners can generate avalanche noise at higher frequencies with the tradeoff of a smaller noise amplitude. A microcontroller at a higher clock rate can detect more pulse counts within its timer/counter peripheral. Additionally, it enables faster processing and transmission of random bytes, further improving the device's output rate. With the mentioned upgrades, it should be possible to achieve a throughput in the order of 500 Kbit/s.

An application of the RAVA circuit must evaluate the throughput compatibility and address security concerns. As outlined in the Introduction Section, while exposing the randomness source has the advantage of transparency and auditing, it may facilitate malicious actors to compromise the integrity of the circuit's output. Consequently, users must determine whether their application operates in a safe environment where the physical presence of malicious third parties can be excluded or if the application is non-critical, implying that no sensitive information is indirectly exposed in the event of an induced fault.

The RAVA device, accessible as an open-source project at [17], emphasizes transparency and customizability. Transparency is fostered by providing monitoring headers used for auditing the noise sources during circuit operation. Customizability is achieved by offering interface headers that facilitate interaction with external devices. Furthermore, all the relevant software can be downloaded and adapted as needed.

Unlike the commercial scenario, where companies may omit some details of their intellectual property, the RAVA device provides users unrestricted access to explore the device at any level they desire. The journey begins with open circuit schematics and board designs, allowing users to delve into the rationale of the noise source and investigate the wiring connections between all components. For real-time verification of the noise source's random behavior, users can plug an oscilloscope into the monitoring headers of a powered circuit. On the software front, users can study the firmware to understand how the microcontroller generates and sends the random bytes. If desired, the users can upload the approved firmware to their devices. The driver, which establishes the link between the device and the user's computer, can also be examined. To ensure the entropy quality, users can generate substantial amounts of random bytes and subject them to comprehensive analysis using standard test suites. Lastly, an internet forum may serve as a platform for users to communicate their findings, fostering a community of knowledge-sharing and validation.

The RAVA implementation showcased is not intended to be a final version but a first step in a project with the additional goal of answering the broader question: What is the most reliable reverse-biased diode RNG design that can be achieved and benefit from community-based development under the open-source philosophy? By being tested and improved by its users, the RAVA device has the potential to become a standard device in scientific projects and other use cases that require a transparent and trusted randomness device compatible with the provided throughput and security considerations.

ACKNOWLEDGMENT

The author thanks Jorge Moll Neto, Jan Walleczek, and Nikolaus von Stillfried for their helpful comments on earlier drafts of this manuscript.

REFERENCES

[1] L. Gong, J. Zhang, H. Liu, L. Sang, and Y. Wang, "True random number generators using electrical noise," *IEEE Access*, vol. 7, pp. 125796–125805, 2019.

[2] M. Bakiri, C. Guyeux, J.-F. Couchot, and A. K. Oudjida, "Survey on hardware implementation of random number generators on FPGA: Theory and experimental analyses," *Comput. Sci. Rev.*, vol. 27, pp. 135–153, Feb. 2018.

[3] M. Stipčević and Ç. K. Koç, "True random number generators," in *Open Problems in Mathematics and Computational Science*. Cham, Switzerland: Springer, 2014, pp. 275–315.

[4] M. Herrero-Collantes and J. C. Garcia-Escartin, "Quantum random number generators," *Rev. Mod. Phys.*, vol. 89, no. 1, 2017, Art. no. 015004.

[5] V. Mannalath, S. Mishra, and A. Pathak, "A comprehensive review of quantum random number generators: Concepts, classification and the origin of randomness," 2022, *arXiv:2203.00261*.

[6] B. K. Park, H. Park, Y.-S. Kim, J.-S. Kang, Y. Yeom, C. Ye, S. Moon, and S.-W. Han, "Practical true random number generator using CMOS image sensor dark noise," *IEEE Access*, vol. 7, pp. 91407–91413, 2019.

[7] P. Keshavarzian, K. Ramu, D. Tang, C. Weill, F. Gramuglia, S. S. Tan, M. Tng, L. Lim, E. Quek, D. Mandich, M. Stipcevic, and E. Charbon, "A 3.3-Gb/s SPAD-based quantum random number generator," *IEEE J. Solid-State Circuits*, vol. 58, no. 9, pp. 2632–2647, Sep. 2023.

[8] N. N. Anandakumar, S. K. Sanadhya, and M. S. Hashmi, "FPGA-based true random number generation using programmable delays in oscillator-rings," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 67, no. 3, pp. 570–574, Mar. 2020.

[9] F. Frustaci, F. Spagnolo, S. Perri, and P. Corsonello, "A high-speed FPGA-based true random number generator using metastability with clock managers," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 70, no. 2, pp. 756–760, Feb. 2023.

[10] P. I. Somlo, "Zener-diode noise generators," *Electron. Lett.*, vol. 11, no. 14, p. 290, 1975.

[11] J. Krieger, "The noise of avalanche breakdown diodes," Vishay Intertechnol., Malvern, PA, USA, Vishay Tech. Note 85966, 2017.

[12] M. Stipčević, "Fast nondeterministic random bit generator based on weakly correlated physical events," *Rev. Sci. Instrum.*, vol. 75, no. 11, pp. 4442–4449, Nov. 2004.

[13] W. Killmann and W. Schindler, "A design for a physical RNG with robust entropy estimators," in *Proc. 10th Int. Workshop Cryptograph. Hardw. Embedded Syst. (CHES)*, Washington, DC, USA. Berlin, Germany: Springer, Aug. 2008, pp. 146–163.

[14] G. Vazzana. (2012). *Random Sequence Generator Based on Avalanche Noise*. [Online]. Available: http://holdenc.altervista.org/avalanche/

[15] P. Campbell. (2014). *One RNG: Open Hardware Random Number Generator*. [Online]. Available: https://onerng.info

[16] B. Lampert, R. S. Wahby, S. Leonard, and P. Levis, "Robust, low-cost, auditable random number generation for embedded system security," in *Proc. 14th ACM Conf. Embedded Netw. Sensor Syst. CD-ROM*, Nov. 2016, pp. 16–27.

[17] G. Guerrer. (2023). *RAVA Open-Source Repository*. [Online]. Available: https://github.com/gabrielguerrer/rng_rava

[18] M. Sonmez, E. Barker, J. Kelsey, K. McKay, M. Baish, and M. Boyle, "SP 800-90b: Recommendation for the entropy sources used for random bit generation," NIST, Gaithersburg, MD, USA, Tech. Rep. SP 800-90b, 2018.

[19] L. E. Bassham, A. L. Rukhin, J. Soto, J. R. Nechvatal, M. E. Smid, E. B. Barker, S. D. Leigh, M. Levenson, M. Vangel, D. L. Banks, N. A. Heckert, J. F. Dray, and S. Vo, "SP 800-22 rev. 1a: A statistical test suite for random and pseudorandom number generators for cryptographic applications," NIST, Gaithersburg, MD, USA, Tech. Rep. SP 900-22 rev. 1a, 2010.

[20] *Datasheet: ATmega32U4*, Microchip, Chandler, AZ, USA, 2016.

[21] *White Paper: Random Number Generation Using Quantum Physics*, ID Quantique, Geneva, Switzerland, 2010.

[22] J. Walleczek, G. Grössing, P. Pylkkänen, and B. Hiley, "Emergent quantum mechanics: David Bohm centennial perspectives," *Entropy*, vol. 21, no. 2, p. 113, Jan. 2019.

[23] W. Killmann and W. Schindler, "AIS 31: A proposal for functionality classes for random number generators," BSI, Bonn, Germany, Tech. Rep. AIS 31, 2011.

[24] *Quantum Noise Random Number Generator Architecture*, document Rec. x.1702, International Telecommunication Union (ITU), 2019.

GABRIEL GUERRER received the B.S. degree in physics from Universidade Federal do Parana (UFPR), Curitiba, Brazil, in 2004, and the M.S. and Ph.D. degrees from Centro Brasileiro de Pesquisas Fisicas (CBPF), Rio de Janeiro, Brazil, in 2007 and 2009, respectively.

From 2005 to 2009, he studied high-energy physics, collaborating with the LHCb experiment with the CERN Laboratory. Since 2018, he has been a Postdoctoral Researcher with the D'Or Institute for Research and Education (IDOR), Rio de Janeiro. His current research interests include information theory, philosophy of mind, anomalistic psychology, metascience, and developing open-source tools and experiments for studying how we perceive and relate to randomness.