Power-consumption back door in quantum key distribution

Beatriz Lopes da Costa,[1,2,3,4,*] Matías R. Bolaños,[5] Ricardo Chaves,[1,6] Claudio Narduzzi,[5] Marco Avesani,[5,7] Davide Giacomo Marangon,[5,7] Andrea Stanco,[5,7] Giuseppe Vallone,[5,7] Paolo Villoresi,[5,7] and Yasser Omar[1,2,3,4,†]

[1] *Instituto Superior Técnico, Universidade de Lisboa*

[2] *Physics of Information and Quantum Technologies Group, Centro de Física e Engenharia de Materiais Avançados (CeFEMA)*

[3] *PQI—Portuguese Quantum Institute*

[4] *LaPMET—Laboratory of Physics for Materials and Emerging Technologies, Portugal*

[5] *Dipartimento di Ingegneria dell'Informazione, Università degli Studi di Padova*

[6] *Instituto de Engenharia de Sistemas e Computadores—Investigação e Desenvolvimento (INESC-ID)*

[7] *Padua Quantum Technologies Research Center, Università degli Studi di Padova*

(Received 15 April 2025; revised 25 July 2025; accepted 5 August 2025; published 3 November 2025)

Over recent decades, quantum key distribution (QKD) has arisen as a promising solution for secure communications; however, like all cryptographic protocols, implementations of QKD can open security vulnerabilities. Until now, the study of physical vulnerabilities in QKD setups has focused primarily on the optical channel. In classical cryptoanalysis, power and electromagnetic side-channel analysis are powerful techniques that can be used to access secret information about the encryption key in symmetric-key algorithms. Such attacks have rarely been used in QKD, since they require an eavesdropper to have access to Alice's or Bob's setup; however, security proofs of QKD protocols generally assume that these setups are secure, making it crucial to understand the security measures required to ensure this protection. In this work, we propose and implement a power side-channel analysis of a QKD system by exploiting the power consumption of the electronic driver controlling the electro-optical components of the QKD transmitter. QKD modules typically require very precise electronic drivers, such as field-programmable gate arrays (FPGAs). Here, we show that the FPGA's power consumption can leak information about the QKD operation and consequently the transmitted key. Our results are consistent and show critical information leakage, having reached a maximum accuracy of 73.35% in predicting transmitted qubits at a 100-MHz repetition frequency. We also discuss possible countermeasures to prevent such an attack.

DOI: 10.1103/f92x-c3zj

I. INTRODUCTION

Quantum key distribution (QKD) leverages the features of quantum physics to offer a highly secure way for two authenticated parties to share a secret key. From satellite links reaching across intercontinental distances [1,2] to fiber networks spanning metropolitan areas [3–5], QKD is maturing as a commercial quantum technology. Alongside postquantum cryptography [6], it offers a promising way to develop cryptographic systems that are secure against attacks by classical and quantum computers [7].

Nonetheless, like any other cryptographic protocol, implementations of QKD systems may contain physical vulnerabilities that can be exploited to obtain key information. The concept of hacking a cryptographic setting via its physical flaws, known as side-channel analysis, was initially introduced in classical cryptanalysis [8]. Classical cryptographic devices are typically implemented using semiconductor transistors as logic gates, where electrons flow across the silicon substrate when a voltage is applied to or removed from the gate. Consequently, they typically exhibit power consumption and electromagnetic radiation emission that depend on the processed data. Exploiting these dependencies has evolved into the fields of power and electromagnetic side-channel analysis, respectively [9].

*Contact author: m.beatriz.costa@tecnico.ulisboa.pt

†Contact author: contact.yasser@pqi.pt

In contrast, a QKD setting typically involves optical and electronic components. Therefore, physical vulnerabilities in QKD setups may lie in the optical or the electronic band. Using weak coherent-state sources as an approximation to single photons or employing nonideal single-photon detectors are examples of physical flaws that have led to side-channel attacks in the optical band [10,11]. Regarding the electronic components in Alice's or Bob's setups, it is well known that they must be protected via classical means to guarantee the security of the entire system [12]. In reality, security proofs typically assume that Alice's and Bob's laboratories are secure locations from which no information is leaked [13,14]. Understanding which security measures must be employed to ensure that this assumption holds in practical scenarios is crucial for the safe deployment of QKD systems. To achieve this, the physical vulnerabilities of the electronic components must be characterized. Devices such as field-programmable gate arrays (FPGAs) are ubiquitous to QKD setups [15–18]; however, they are also based on transistor logic, thus opening the possibility of electronic side channels. Until now, to our knowledge, all work devising these types of side channels for QKD has focused on the electromagnetic band [19–21]. Additionally, one should note that electromagnetic side-channel attacks have also been devised, studied, and implemented for quantum random number generators (QRNGs) [22–24], which are crucial for the security of QKD.

Regarding power side channels, using power-consumption traces of cryptographic devices to access secret information about the encryption key is a standard and powerful technique in cryptoanalysis of classical devices implementing symmetric-key algorithms, such as the Data Encryption Standard (DES) [25] and the Advanced Encryption Standard (AES) [26]. In this work, we aim to devise and implement a power side-channel attack to a component of a QKD transmitter. The side-channel analysis in this work was specifically performed on the system on a chip (SoC) controlling the electro-optical components of the QKD transmitter at the University of Padua. Since side-channel vulnerabilities are highly implementation dependent, the findings presented here are specific to the QKD transmitter used.

We will start in Sec. II by briefly covering the basic operation of the QKD transmitter, focusing on its electronic components. This background is essential for comprehending the experimental setup for the power analysis, which is detailed in Sec. III. Additionally, an understanding of the working principle of the QKD transmitter is presented to shed light on the potential causes of information leakage. The analysis of the SoC's power consumption is divided into two parts: its average value, covered in Sec. IV, and its frequency spectrum, covered in Sec. V. Vulnerabilities found in the analysis of the latter are then used to propose and implement a side-channel attack on the SoC. The methodology of this attack and the results obtained from its implementation are detailed in Sec. VI, while a short discussion of possible countermeasures is presented in Sec. VII. During the entire analysis, the SoC was treated as a pseudo-black-box system, as it would be by a current eavesdropper with malicious intent, who would be expected to only have access to publicly available information about the device. Therefore, the only knowledge we assume regarding the working principle of the SoC is that made public in Ref. [15], as covered in Sec. II. We conclude in Sec. VIII by evaluating the overall performance of the side-channel attack and the limitations of this approach.

II. WORKING PRINCIPLE OF THE QKD TRANSMITTER

The QKD setup implements the three-state one-decoy BB84 protocol with polarization encoding [27]. In the transmitter, an intensity modulator (IM) [28] is used to create decoy and signal states, followed by a polarization encoder, the POGNAC [29], which encodes the polarization of each photon (Fig. 1) [16].

The three-state BB84 protocol relies on three qubit states: two that contribute to key generation (e.g., $\{|L\rangle, |R\rangle\}$) and a third that is solely used for parameter estimation (e.g., $|D\rangle$). During QKD operation, the transmitter sends these qubit states to the receiver through the quantum channel at a specific rate known as the qubit repetition frequency f_{rep}. A gain-switched distributed-feedback laser is employed as a weak coherent-pulse source.

To employ the decoy aspect of the protocol, the intensity of each weak coherent pulse must be modulated between two values, μ_1 and μ_2, one for signal states and the other for decoy states. To achieve this, a lithium niobate ($LiNbO_3$) phase modulator placed inside a Sagnac interferometer is used as a two-level IM [28]. In this configuration, voltage pulses are sent to the phase modulator at a rate f_{rep}, with a voltage chosen between two values, e.g., 0 and V, corresponding to the higher and lower intensity, respectively.

After exiting the IM, the pulses have to be encoded with key information. To do this, they are sent through the POGNAC, a polarization encoder based on a $LiNbO_3$ phase modulator inside a Sagnac interferometer [29]. Here, the light pulses enter the Sagnac loop in the superposition state

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|H\rangle + |V\rangle), \tag{1}$$

where they pass through a polarizing beam splitter (PBS), ensuring that each light pulse is divided into two pulses that travel in opposite directions within the loop. To avoid polarization mode dispersion and birefringence effects, after the PBS, the polarization of one of the pulses is rotated into its orthogonal counterpart, so that only one

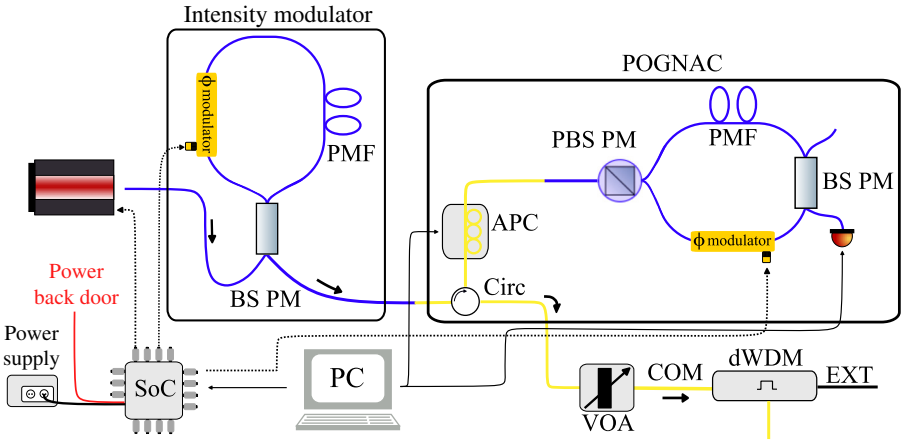

FIG. 1. Schematic representation of the QKD transmitter. The gain-switched laser emits light pulses, which are directed into an intensity modulator (IM) composed of polarization-maintaining fiber (PMF), a polarization-maintaining beam-splitter (BS PM) and a phase modulator (ϕ modulator). After exiting the IM, the pulses enter a polarization encoder (POGNAC), which employs a circulator (Circ), an automatic polarization controller (APC), beam splitters, and a phase modulator to encode the polarization of each state. Before exiting the transmitter, the pulses pass through a variable optical attenuator (VOA) and a dense wavelength-division multiplexer (dWDM). A system on a chip (SoC) is used to drive the electro-optical modulators and to gain-switch the laser.

polarization travels through the fiber. Finally, they arrive at the phase modulator with a time difference of

$$\delta t = \frac{\Delta L}{n_f\, c}, \tag{2}$$

where ΔL and n_f are, respectively, the length and the refractive index of a polarization-maintaining fiber delay line placed inside the loop. In this way, by sending synchronized voltage pulses at f_{rep} and then choosing one of the two time slots separated by δt, one is able to control the phase applied to each polarization state. By sending an electrical pulse with voltage V_ϕ [30] targeting only one of the two polarized pulses (for example, the vertically polarized one), the state $|\psi\rangle = |H\rangle + e^{i\phi}\,|V\rangle$ can be created. Thus, by choosing $\phi = 0, \pm\pi/2$, the states $|D\rangle$, $|R\rangle$, and $|L\rangle$, respectively, can be created, which is sufficient for the three-state BB84 protocol.

Overall, the optical encoder in the transmitter requires three separate electrical signals: one to gate the laser, and two for the phase modulators, one in the IM and the other in the POGNAC. To create these electrical pulses with the high temporal resolution required by fast repetition rates f_{rep}, an SoC is used [15].

The SoC comprises an FPGA and a dual-core CPU. Its architecture follows a top-down configuration, where data flows from the user or personal computer to the quantum system, passing through two different layers. First, it goes through the CPU layer, which is responsible for communication with the outside world and data-management operations. Then it reaches the FPGA, where all deterministic and high-resolution operations are carried out. Finally, from the FPGA, it goes to the chip input-output pins, where the electric pulses are emitted. This is the device whose power consumption we will exploit in a power side-channel attack.

III. EXPERIMENTAL SETUP

Our target device is the Zynq-7020 SoC mounted on a ZedBoard by Avnet. The ZedBoard receives a 12-V input power supply and comes with an inbuilt 10-mΩ, 1-W in-series resistor for current measurements, placed at the high-voltage side of the circuit. This allows us to determine the current drawn from the SoC's power supply and consequently infer its power consumption. To guarantee enough resolution for measurements of the average power consumption values, the 10-mΩ resistance was replaced with a 1-Ω, 1-W one. The 1-Ω value was chosen to ensure that the voltage across the resistor was large enough to allow resolution in the millivolt range while avoiding the need for a higher-voltage power supply ($V > 12$ V) to the ZedBoard.

The voltage difference between terminals of the 1-Ω resistor translates into the power consumption of the SoC. To characterize this, we acquire the voltage difference during a certain time interval and refer to the resulting data as a power-consumption trace.

The power traces were acquired using a SIGLENT SDS5104X oscilloscope. Two different probe setups were used to measure the voltage, according to the different requirements of the measurements. The first probe used was a KEYSIGHT N2791A Differential Probe, which directly computes the voltage difference between the two resistance terminals, as schematized in Fig. 2. This probe has a 25-MHz bandwidth, which is insufficient for an

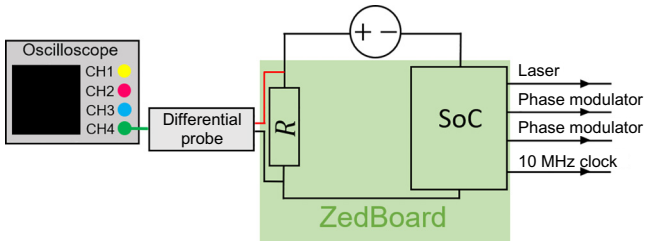

FIG. 2. Schematic of the experimental setup used for measuring the average power consumption of the SoC using a differential probe.

analysis of the frequency spectrum of the power traces, leaving this setup for the study of the average power consumption of the SoC.

For the frequency-spectrum analysis of the power traces, two SIGLENT SP2035A voltage probes with a bandwidth of 350 MHz were employed in a pseudodifferential setup, as illustrated in Fig. 3. Using the oscilloscope, the power traces were retrieved by computing the difference between the two voltage traces, each measured by one of the probes. In this case, the sampling frequency of the oscilloscope was set to 2.5 GSa/s, the maximum allowed.

The digital pulse output by the SoC to gain-switch the laser was used to trigger the acquisitions in both experimental setups. In addition, to eliminate clock drifts between the SoC and the oscilloscope, a 10-MHz clock generated by the SoC was employed as an external clock source for the oscilloscope.

A. System settings

To control the system settings, we relied on an application that allows the user to set the type of sequence to be sent by directly loading it onto the FPGA in the SoC. For all the measurements taken in this work, the qubit repetition frequency was set to 100 MHz. Regarding the key, the symbol encoded by each qubit can be chosen from H, V, and D, which are interface notations for the states $|L\rangle$, $|R\rangle$, and $|D\rangle$ in Sec. II, respectively. For the purposes of this work, only nondecoy key symbols, H and V, were considered.

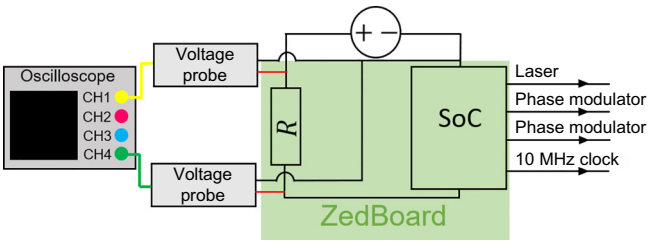

FIG. 3. Schematic of the experimental setup used for analyzing the frequency spectrum of the power consumption of the SoC using two voltage probes.

In the course of the analysis, two different key types were used. The first corresponded to fixed sequences, which consist of an infinite repetition of symbols. For example, an Only-H sequence corresponds to having the QKD transmitter sending H symbols indefinitely. Naturally, these sequences would not be used in a practical QKD setting, since to assure theoretical security, each symbol Alice sends must be chosen in a truly random manner, i.e., using a QRNG. Therefore, we use these sequences solely for the characterization of the SoC's power consumption under different settings.

The second type of key is a finite-size random key, which can be used to simulate a more realistic QKD scenario. We note that the random key strings were generated with PYTHON's pseudorandom number generator and therefore do not correspond to true randomness. Nevertheless, pseudorandomness provides the framework necessary for validating a possible side-channel attack; thus, from this point forward, we neglect its difference from true randomness.

IV. AVERAGE POWER CONSUMPTION

In this section we analyze whether the average power consumption of the FPGA depends on the values of the symbols composing the key. To do this, we consider the emission of fixed sequences with different percentages of H symbols in the key, using a randomized emission method. The fixed sequences considered were Only-H, Only-HHV, Only-HV, Only-HVV, and Only-V, chosen given their different percentages of H symbols, which are respectively 100%, 66%, 50%, 33%, and 0%. Regarding the acquisition, the time scale of the oscilloscope was set to 500 μs/div, allowing the average power consumption to be calculated during the emission of 500 000 symbols with each power-trace acquisition. We note that during acquisition, some power traces showed an abnormally low average power consumption. This appeared to affect the traces randomly. We attributed this to faults in the connection between the jack of the power cable and the ZedBoard's socket.

Figure 4 displays the time evolution of the average power consumption for the Only-V, Only-HV, and Only-H sequences, in which the power-consumption drops are visible. The average power consumption for the emission of each fixed sequence showed a slow decrease over the acquisition time, which may be explained by a temperature dependence of the 1-Ω resistor and an overall increase of the resistor's temperature and thus resistance during the measurement phase. To understand whether this decrease affected all sequences equally, we performed five linear regressions $V = \beta_0 + \beta_1 t$, one for each sequence's time evolution of the average power consumption. From these results, we concluded that the power-consumption descent was sequence independent. To mitigate its effect, we

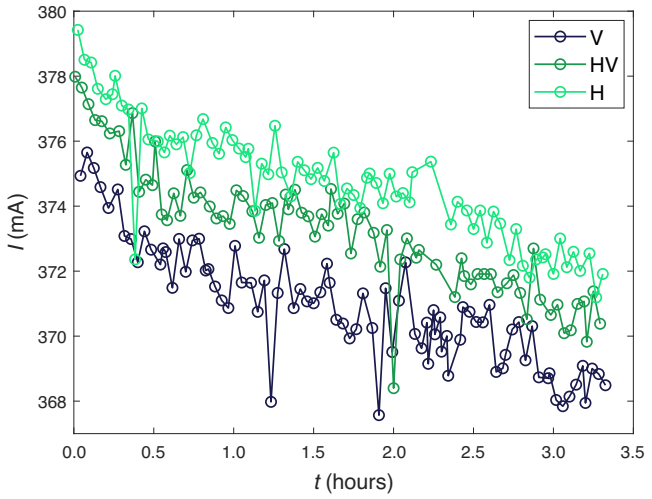

FIG. 4. Time evolution of the average power consumption of traces corresponding to different fixed sequences.

considered the average value $\overline{\beta_1} = (-1.73 \pm 0.2)$ mA/h, and shifted all data points (t, V) from all sequences to $(t, V - \overline{\beta_1} t)$.

Figure 5 shows the average power-consumption values for each fixed sequence after mitigating the global power-consumption decrease. The current-intensity distribution across the 1-Ω resistor is displayed for the different percentages of H symbols in the key, together with the corresponding power consumption. These results showed a trend: the average power consumption increased with increasing percentage of H symbols in the key.

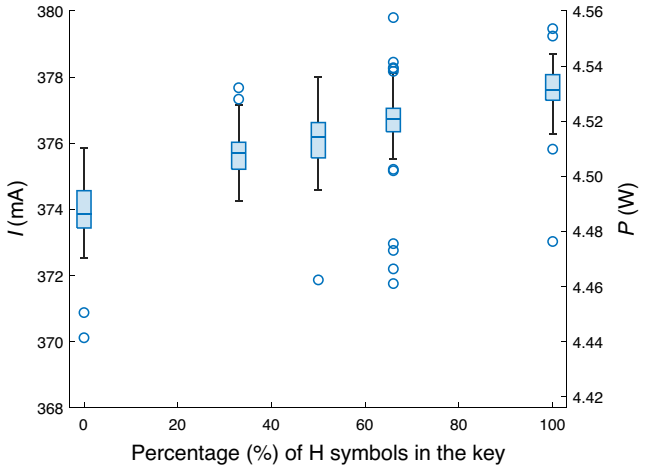

FIG. 5. Box plot displaying the average power consumption values for each H symbol percentage. Each dataset is represented by its median (the line inside the box), upper and lower quartiles (the top and bottom edges of the box), and whiskers indicating the range of nonoutlier data. Outliers are shown as circular points.

The power-consumption oscillation when sending Only-HV sequences was also analyzed for all integer qubit-repetition frequencies in the range 2–46 kHz. We used the FPGA at $f_{\mathrm{rep}} = 100$ MHz to approximately simulate a smaller qubit repetition frequency f_{delay}. To do this, one emits sequences consisting of $n = f_{\mathrm{rep}}/f_{\mathrm{delay}}$ same-valued symbols, each sequence corresponding to the emission of a single symbol in the lower-frequency case.

The time scale of the oscilloscope was set to 500 μs/div, and the delay was set to −2 ms, meaning that each power trace encoded the rise in power consumption in the first 0.5 ms and then the emission during 4.5 ms. The number of key symbols encoded in these 4.5 ms depends on the repetition frequency considered. For each frequency, ten different power traces were acquired.

To reduce noise, each trace was filtered with a low-pass filter, with its bandwidth set to the corresponding smaller repetition frequency. Additionally, the ac power supply used until this point was replaced with a dc power supply to reduce the frequency noise around 20 kHz introduced by the former. Figure 6 shows an example of a trace acquired at 10 kHz after filtering and excision of the initial sharp rise in the power consumption.

We note that during the characterization of the SoC's power consumption, it was observed that the CPU writing operation, detailed in Ref. [15], influenced the power consumption. This periodic operation caused a destabilization of the power consumption, and this can be seen in Fig. 6 at around 3.12 ms.

To compute the value of each symbol from the power traces, we considered an algorithm that segments the signal into snippets of individual symbols and extracts certain features from them, namely the average, the slope, and the curvature of the power consumption. These features are then combined using optimized weights to classify each

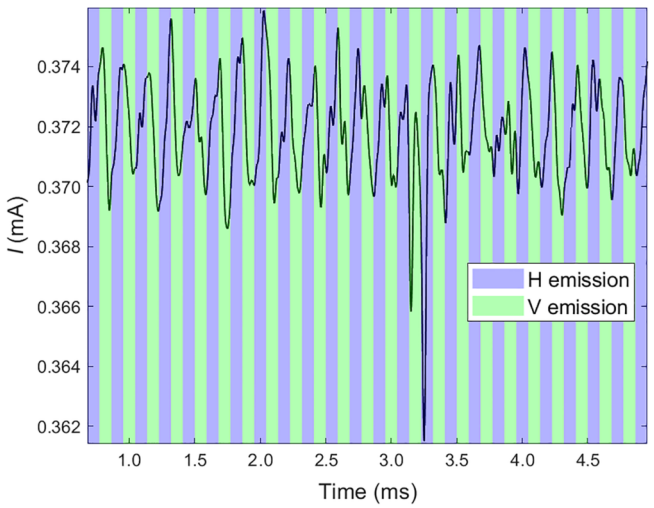

FIG. 6. Power trace acquired during the emission of an HV sequence at 10-kHz repetition frequency after filtering.

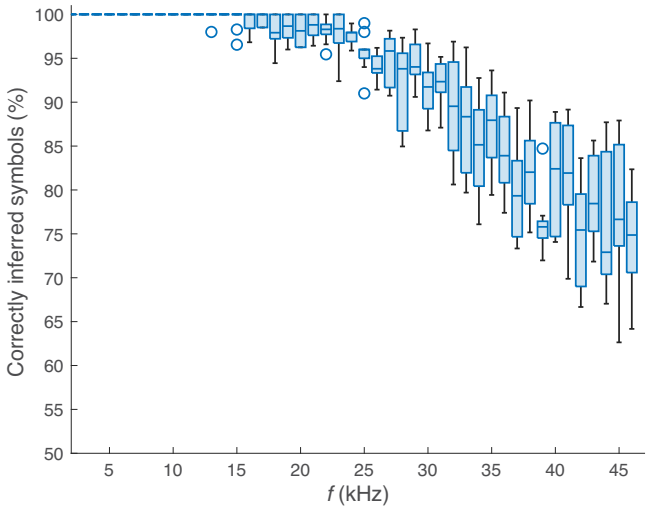

FIG. 7. Box plot displaying the percentage of correctly inferred symbols in HV sequences emitted at different repetition frequencies obtained by analyzing the oscillation in the power-consumption traces acquired during their emission.

symbol as either H or V. By additionally correcting for time shifts in the traces, we achieved an average accuracy above 73% for all repetition frequencies. Figure 7 is a box plot presenting the algorithm's accuracy when inferring individual symbols from the power traces of HV sequences at different repetition frequencies. At higher frequencies, the accuracy decreases; this can be explained by the fact that the symbol window becomes shorter as the frequency is increased, and waveform distortions caused by noise have a greater impact on the inference of each symbol. Notably, the dominant frequency in the spectrum of all acquired traces was found to be half of the qubit repetition frequency of the encoded HV sequence. This is consistent with the oscillatory behavior that the HV sequence creates in the power consumption, where the higher part of the oscillation corresponds to H-symbol emission and the lower part to V-symbol emission.

V. FREQUENCY-SPECTRUM ANALYSIS OF THE POWER CONSUMPTION

In this section, the frequency spectrum of the power traces is characterized to identify possible information leakage. Given that the focus is on the frequency domain, the bandwidth of the acquisition system is of crucial importance. Therefore, the power traces were acquired using the two SIGLENT SP2035A voltage probes with a 350-MHz bandwidth, using the setup schematized in Fig. 3. Nonetheless, data taken with the same setup but with two Tektronix P2220 voltage probes with a bandwidth of 200 MHz were also considered. Although they have a lower bandwidth, these probes can be impedance matched to the oscilloscope, which can only switch between an input impedance of 50 Ω or 1 MΩ. This, together with the fact that these probes do not have built-in signal attenuation, unlike the 350-MHz probes, reduces the noise associated with the data acquisition. Data taken with these two types of probe will be analyzed and compared. Unless stated otherwise, the measurements were taken with the 350-MHz probes.

A. Fixed sequences

Average frequency spectra of Only-H and Only-V sequences are shown in Fig. 8, averaged over the fast Fourier transforms (FFTs) of 50 power traces. Each frequency bin displays the average magnitude, together with its standard deviation. We note that the magnitudes for all the frequencies in this range that are integer multiples of 100 MHz show complete distinguishability between Only-H and Only-V.

Averaging the FFTs corresponding to emissions of the same sequence directly results in a smoothing of the noise in the spectrum of the sequence. The correlation coefficients at zero delay, computed between all possible pairs of FFTs, provide insight into how this noise affects the distinguishability of spectra. The coefficients were organized in matrix form

$$\left[\begin{array}{ccc|ccc} h_1,h_1 & \cdots & h_1,h_{50} & h_1,v_1 & \cdots & h_1,v_{50} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ h_{50},h_1 & \cdots & h_{50},h_{50} & h_{50},v_1 & \cdots & h_{50},v_{50} \\ \hline v_1,h_1 & \cdots & v_1,h_{50} & v_1,v_1 & \cdots & v_1,v_{50} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ v_{50},h_1 & \cdots & v_{50},h_{50} & v_{50},v_1 & \cdots & v_{50},v_{50} \end{array}\right], \tag{3}$$

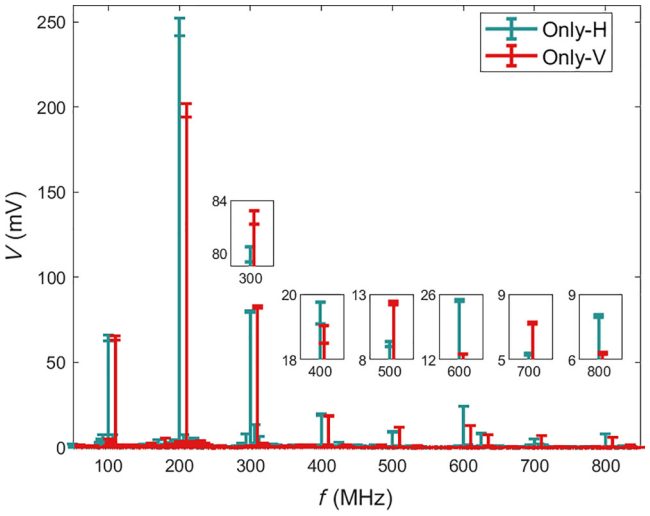

FIG. 8. Average frequency spectra of the SoC's power consumption during the emission of Only-H and Only-V sequences. Only the 50–900 MHz frequency range is displayed. The Only-V spectrum has been offset by +10 MHz for readability.

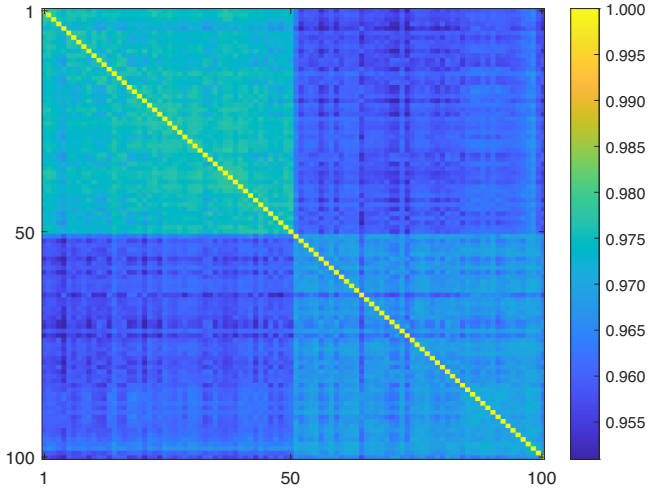

FIG. 9. Correlation matrix for the power consumption's FFTs during the emission of 200 000 symbols taken from an Only-H and an Only-V emission. Only the 50–900 MHz range was considered.

where x_i represents the FFT of the ith power trace taken during Only-X emission, $X \in \{\mathrm{H}, \mathrm{V}\}$, and x_i, x_j denotes the correlation value between x_i and x_j. The results for the 50–900 MHz frequency range are shown in Fig. 9. The top-left and bottom-right 50×50 submatrices exhibit higher values, demonstrating that it is possible to distinguish between the spectra of Only-H and Only-V sequences without noise filtering.

The correlation coefficients were also computed using the full-frequency-range spectra. In this case, all FFTs exhibited a higher degree of similarity; thus, it can be inferred that the main information leakage lies in the 50–900 MHz spectral range.

In a realistic symbol-to-symbol side-channel attack, where the transmitted key is a random string of symbols, FFTs of smaller-size sequences must be considered. The power consumption's frequency spectrum during the emission of a large quantity of symbols can only transmit information if all the symbols are either the same, as in the previous case, or differ in a known way, e.g., an HV sequence. Nonetheless, as the size of the sequence decreases, the FFT resolution will worsen, where

$$\text{FFT resolution} = \frac{f_{\text{sample}}}{\text{Record length}} \tag{4}$$

and $f_{\text{sample}} = 2.5$ GSa/s.

We also note that the average spectra obtained for an n-symbol emission extracted from a fixed sequence does not necessarily correspond to the one obtained for the same n-symbol emission during a random sequence exchange. In fact, memory effects during the exchange of fixed sequences may influence the spectra of the symbols emitted. Therefore, the next section focuses on characterizing the power consumption's spectrum during the exchange of random qubits, i.e., a realistic QKD scenario.

B. Sequences in random-sequence emission

The study of different sequences' spectra during random emission was conducted using a dataset consisting of $s = 1 \times 10^6$ randomly emitted symbols and the power consumption of the SoC during their emission. For this, five power traces were collected, each taken during the emission of an independent random raw key consisting of 200 000 symbols. After data acquisition, the five power traces were normalized. The spectra of different sequences were studied considering the following methodology:

(1) Consider sequences S_L^i, where L is the number of symbols in the sequence and i the number of possible sequences of size L, $i \in \{1, \ldots, 2^L\}$.
(2) Partition the s number of symbols in the dataset into short sequences of length L, resulting in s/L sequences with 2^L possibilities.
(3) Divide the power-consumption data accordingly.
(4) Calculate the FFTs of the occurrences of each sequence and average them to produce the sequence's FFT average F_L^i.
(5) Compute the correlation matrix of the FFT averages

$$\begin{bmatrix} F_L^1, F_L^1 & \cdots & F_L^1, F_L^{2^L} \\ \vdots & \ddots & \vdots \\ F_L^{2^L}, F_L^1 & \cdots & F_L^{2^L}, F_L^{2^L} \end{bmatrix}. \tag{5}$$

By averaging all the FFTs corresponding to the emission of the same sequence, two outcomes are achieved. First, the noise is smoothed out, increasing distinguishability. Second, since all occurrences of a given sequence were preceded and succeeded by other completely random sequences, any possible memory effect is eliminated from the spectrum.

Correlation matrices were computed for $L = \{2, 4, 6, 8\}$, and for all considered values of L, the same trend was observed: correlations between the frequency spectra of two sequences were higher the more initial symbols the two sequences shared. It was also found that this distinguishability increased when considering the phase values at each frequency. In this case, the phases of each sequence occurrence were calculated and then averaged over all occurrences of the same sequence. Figures 10 and 11 show the correlation matrices considering the average phases for $L = 8$ and $L = 4$, respectively.

To address the correlation matrices, it is useful to define the k-diagonal of a matrix A as the set of entries $a_{i,j}$ for which $j = i + k$. In the correlation matrices, the $\pm(2^n\, k)$ diagonals showcase the higher correlations between the average phases of FFTs corresponding to

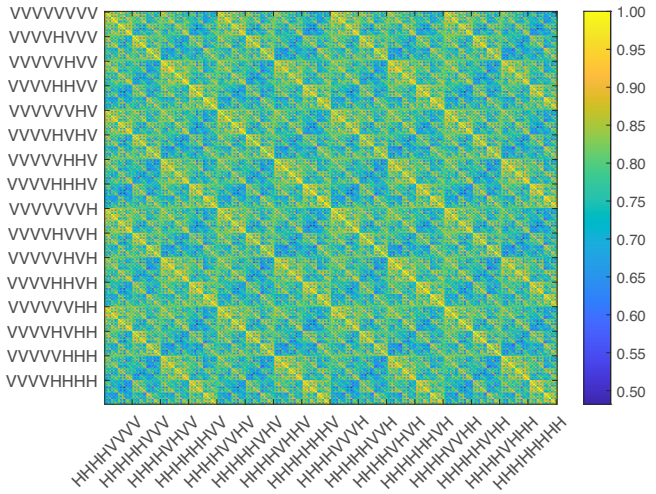

FIG. 10. Matrix displaying the zero-lag correlations between the average FFTs of all possible eight-symbol sequences. Sequences were taken from a larger dataset consisting of 1 000 000 randomly emitted symbols. The average phase value was considered at each frequency bin.

sequences that share n initial symbols, with $n < L$ and $k = 2i + 1$, $i \in \{0, \ldots, L - n - 1\}$. For example, in the $L = 4$ case, the $\pm 4, \pm 12$ diagonals showcase the correlation values between the average phases of FFTs corresponding to sequences that share the first two symbols. In fact, by applying a threshold to the correlation matrix, it is observed that the values in these diagonals are smaller than the ones in the ± 8 diagonals, which represent correlations between sequences that share the first three initial symbols.

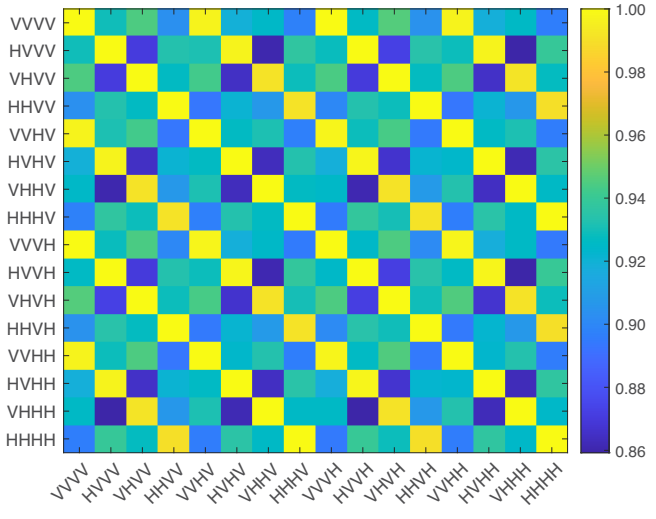

FIG. 11. Matrix displaying the zero-lag correlations between the average FFTs of all possible four-symbol sequences. Sequences were taken from a larger dataset consisting of 1 000 000 randomly emitted symbols. The average phase value was considered at each frequency bin.

The relationship between the average shape of a sequence's spectrum and its initial values results in information leakage that can be exploited at high repetition rates. By comparing the spectrum of unknown sequences of size L to the known average spectra of all possible sequences of size L, it is possible to infer information about the initial values of the unknown sequence. This concept forms the basis of the strategy implemented in the next section.

VI. DISCOVERING THE TRANSMITTED QUBITS AT 100 MHz

In this section, the information leakage found in the frequency spectrum of the power consumption of the SoC is exploited to discover the transmitted qubits at a repetition rate of 100 MHz.

We propose a template side-channel attack to the SoC based on the "FFT fingerprints" of the device. Here, we assume that the eavesdropper has initial access and control of the QKD transmitter, or of an identical copy, and uses this to store power-consumption data and corresponding sequences. With the power-consumption data, the eavesdropper computes the FFT fingerprints, where for an arbitrary L, we define the FFT fingerprint as the average FFT of the power consumption during the emission of a certain sequence with L number of symbols. Note that for our attack, we considered the phase average at each frequency bin as opposed to the magnitude.

In the second stage of the attack, Eve loses control over the QKD transmitter but keeps monitoring the power consumption of the FPGA. It is in this stage that Alice and Bob use the QKD system to create shared raw keys, not knowing that Eve has had previous access and is still monitoring the power consumption of the FPGA. During qubit transmission, Eve acquires power-consumption data and extracts overlapping data segments from it corresponding to sequences with L symbols. Each segment starts δN symbols after the previous one, creating a sliding window with step size δN. For each segment, she computes the FFT and finds the FFT fingerprints that best resemble it, using zero-lag correlations. Knowing the sequence of the FFT that yields the best match, Eve uses the first δN symbols in this sequence to guess the first δN symbols of the segment.

In the proposed side-channel attack strategy, there are two parameters that Eve must choose: L, the length of the sequences for which the FFT fingerprints are calculated, and δN, the number of symbols she guesses after finding the best match for an FFT, such that $0 < \delta N \leq L$.

To demonstrate the feasibility of this side-channel attack, we implemented it to discover 15 independent random raw keys. The previously obtained set of 1 000 000 random-symbol emissions was used to calculate the FFT fingerprints. Each raw key we wished to discover consisted of 200 000 random symbols, and to quantify the

performance of our attack, we calculated the prediction accuracy as

$$\text{Prediction accuracy (\%)} = \frac{\text{No. of correctly guessed symbols}}{\text{Total no. of symbols in the key}} \times 100. \quad (6)$$

The prediction accuracy with a random-guessing method—that is, for each symbol in an infinitely sized raw key, one randomly predicts between H and V—would be 50%. Therefore, we consider the discovery of a raw key to be successful if it exceeds random guessing by 3 standard deviations of the binomial distribution [19], that is,

$$\text{Prediction accuracy (\%)} > 50.34\%. \quad (7)$$

Two side-channel attack trials were performed, one with a pair of oscilloscope probes with a bandwidth of 350 MHz, and another with a probe pair with 200-MHz bandwidth. The trials were spaced by a month and, for each, a different set of 15 independent random raw keys was considered. The FFT fingerprints were calculated using a dataset acquired with the corresponding probes. Using a strategy with $L = 8$ and $\delta N = 1$, we were able to achieve a maximum prediction accuracy of 73.35%. The prediction accuracy achieved for each of the 15 sequences in the two trials is shown in Fig. 12. The performance of the trial with the 350-MHz probes surpasses the performance of the one with the 200-MHz probes, which tells us that the bandwidth of the acquisition system is more important to the side-channel attack performance than the amount of noise associated with the measurements.

Side-channel attack strategies with $L = \{2, 4, 6\}$ and $\delta N = 1$ were also performed using the same data. In all of them, we see a decrease in the prediction accuracy compared to the $L = 8$, $\delta N = 1$ case, although all remained above the success threshold. Generally, for fixed δN, we expected the performance to decrease as L decreases, given that the FFT resolution worsens; however, although this was true for $L = \{8, 4, 2\}$, the $L = 6$ strategy performed worse than the $L = 4$ one. The former entails FFT resolution = 16.67 MHz, meaning that the FFTs computed do not have frequency bins located at the integer-multiple frequencies of the 100-MHz frequency. For $L = 4$, these frequency bins exist, which might justify the better performance of a strategy with this value. Additionally, for $L = 4$ and $L = 2$, the side-channel attack performed better for the data acquired with the 200-MHz probes, meaning that, as the spectral resolution decreases, noise becomes a greater impediment to guessing the key and the bandwidth of the system becomes less important.

To study the impact of δN, strategies with fixed L and different values of δN were considered. The results for $L = 4$ are shown in Fig. 13, where for simplicity, only the trial with the 350-MHz probes was considered. Here, we see a decrease in the side-channel attack performance as δN increases. This means that after calculating the FFT of a given snippet and finding its highest-correlated FFT fingerprint, we have a high degree of certainty on the first symbol in the snippet. This certainty decreases for the second symbol, and so forth. Therefore, the strategy with the lowest performance was the one that combined the lowest L with the highest δN allowed, this is, $L = 2$ and $\delta N = 2$.

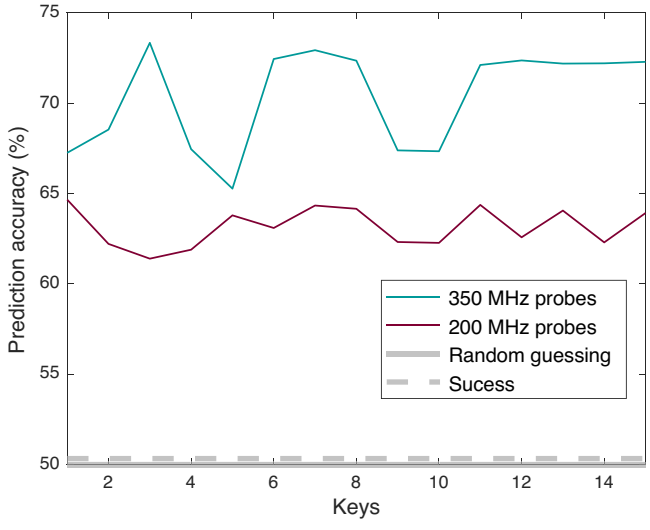

FIG. 12. Prediction accuracy achieved by the side-channel analysis for different raw keys with 200 000 symbols, using a strategy with $L = 8$ and $\delta N = 1$. Results obtained with two different sets of oscilloscope probes are shown, each set with a given bandwidth.

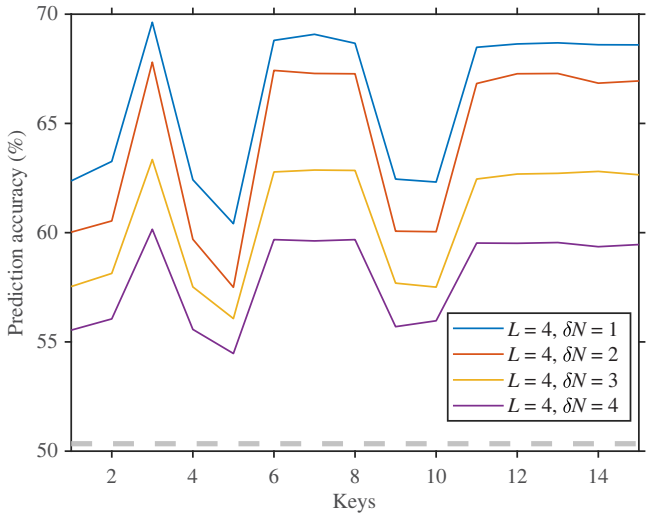

FIG. 13. Prediction accuracy achieved by the side-channel analysis for different raw keys with 200 000 symbols, using a strategy with $L = 4$ and different values of δN. Only the results from data acquired with the set of oscilloscope probes of 350-MHz bandwidth are shown.

Nonetheless, even in this case, all 15 raw keys yielded a prediction accuracy above the success threshold.

VII. COUNTERMEASURES

Side-channel attacks (SCAs) are notoriously implementation dependent. Due to their nature, the amount of information leakage depends on features such as the protocol and device architecture, among other factors. Consequently, choosing a set of suitable algorithmic countermeasures to mitigate leakage is a task that must typically be tailored for a specific combination of device and attacker model [31].

The aim of any countermeasure is to reduce the correlation between the data processed by a relevant part of the cryptographic implementation and the physical emanation exploited by the attack—in our case, power consumption. The literature on countermeasures for SCAs in classical cryptography is extensive, and they are generally classified as either hiding or masking techniques.

Masking techniques aim to eliminate leakage by applying random masks; however, they rely on the presence of algebraic manipulation of secret values during computation—which is typically absent in QKD electronic drivers [32]. This makes masking poorly suited to countering attacks such as the one proposed here.

Hiding techniques, by contrast, aim to reduce the signal-to-noise ratio of the leakage. In this way, even though correlations between the data and the power consumption may still exist, they are no longer discernible to the attacker. In fact, the nondeterministic nature of raw-key transmission in QKD makes power SCAs more susceptible to noise. Unlike block cyphers, such as AES, where due to their deterministic nature, the attacker may collect many traces under different known plaintexts and use statistical techniques to recover the key, in QKD, an attacker must rely on a single power trace. Consequently, hiding techniques, such as injecting artificial switching noise [33] or executing parallel cryptographic operations using an interfering key [34,35], might be suited to mitigating leakage.

Additionally, hiding countermeasures based on the dynamic logic reconfiguration capability of FPGAs could potentially be very effective against our attack [36–38]. This feature allows for real-time reconfiguration of parts of the FPGA's fabric, allowing changes to the logic implementation of the cryptographic algorithm that result in changes to the leakage profile. Therefore, if Alice reconfigures the FPGA logic—e.g., the circuitry generating pulses to drive the POGNAC phase modulator—each time a key is sent, the leakage pattern will vary, invalidating Eve's template.

Finally, as further discussed in the next section, the simplest countermeasure consists of physically protecting the transmitter hardware. The attack illustrated in this work requires embedding a sensing device into the chip and using short probe cables. Therefore, preventing an eavesdropper from having access to the transmitter would clearly prevent this attack. In fact, it is a necessary security assumption for any QKD protocol that Alice's and Bob's physical locations are secure and that an eavesdropper does not have physical access to them. Guaranteeing that this assumption holds is therefore enough to prevent the attack illustrated here.

VIII. CONCLUSIONS AND OUTLOOK

In this work, we have demonstrated the feasibility of a power side-channel attack to a QKD transmitter, namely by exploiting the information leakage of the SoC controlling the electro-optical components of a specific QKD transmitter.

The experimental results suggest that the average power consumption of the SoC depends on the transmitted qubit values, namely, it is larger for an H than for a V; however, we recall that this is specific of our firmware implementation, and it cannot be understood as a general feature of the SoC. Emission of HV sequences at different repetition frequencies was shown to produce an oscillation of the power consumption consistent with this symbol dependence. On average, we could correctly infer more than 70% of symbols in HV sequences for all integer repetition frequencies in the range 2–46 kHz by algorithmic analysis of their power traces.

Regarding the frequency spectrum of the SoC's power consumption, for random sequences of length L, the spectra became increasingly similar as the number of shared initial symbols between two sequences increased. This allowed for the implementation of a template side-channel attack based on the SoC's FFT fingerprints, which was implemented to predict two sets of 15 transmitted keys emitted at $f_{\mathrm{rep}} = 100$ MHz. The maximum prediction accuracy achieved for this case was 73.35%.

Notably, the side-channel attack proposed in this work requires an eavesdropper capable of measuring the power consumption of a chip integrated into Alice's transmitter. For the setup employed in this work, this requires using a sensing resistor embedded into the chip and relatively short probe cables. In a real attack scenario, if the transmitter is placed inside a box, Eve would need to dismantle or tamper with it to install this power-sensing setup, which would very likely be detected by Alice or Bob. Alternatively, Eve could attempt to measure power consumption further away from the transmitter, for instance by tapping into the power line near the ac outlet; however, this can result in noisier signals that are heavily limited by the response time of the power supply. Therefore, it would be important in future studies to understand the feasibility of power side-channel analysis performed far from the device under attack.

Nonetheless, the information leakage found, together with the success of the side-channel attack, shows that power side-channel analysis to a SoC is a feasible technique to retrieve information about the transmitted qubits. Therefore, power-consumption data from the electronics in the parties' setups must also be treated as sensitive information, making the findings in this work an important step in understanding which protective measures must be implemented for Alice's and Bob's electronics to guarantee the overall security of the QKD system.

ACKNOWLEDGMENTS

The authors are grateful for support from project QSNP—Quantum Secure Networks Partnership (GA 101114043) of the Horizon Europe Programme of the European Commission. B.C., Y.O., and R.C. thank FCT—Fundação para a Ciência e a Tecnologia (Portugal) for their support, namely through project UIDB/04540/2020 contract LA/P/0095/2020 and UIDB/50021/2020. M.R.B. acknowledges support from the European Union's Horizon Europe Framework Programme under Marie Sklodowska Curie Grant No. 101072637, Project Quantum-Safe Internet (QSI). The authors would like to thank F. Rampazzo for lending the KEYSIGHT N2791A differential probe. This work is partially supported by ICSC—Centro Nazionale di Ricerca in High Performance Computing, Big Data and Quantum Computing, funded by European Union—NextGenerationEU.

DATA AVAILABILITY

The data that support the findings of this article are not publicly available. The data are available from the authors upon reasonable request.

[1] S.-K. Liao, *et al.*, Satellite-relayed intercontinental quantum network, Phys. Rev. Lett. **120**, 030501 (2018).
[2] J. Yin, *et al.*, Satellite-based entanglement distribution over 1200 km, Science **356**, 1140 (2017).
[3] M. Peev, *et al.*, The SECOQC quantum key distribution network in Vienna, New J. Phys. **11**, 075001 (2009).
[4] J. F. Dynes, A. Wonfor, W. W. Tam, A. W. Sharpe, R. Takahashi, M. Lucamarini, A. Plews, Z. L. Yuan, A. R. Dixon, J. Cho, Y. Tanizawa, J. P. Elbers, H. Greißer, I. H. White, R. V. Penty, and A. J. Shields, Cambridge quantum network, npj Quantum Inf. **5**, 101 (2019).
[5] M. Sasaki, *et al.*, Field test of quantum key distribution in the Tokyo QKD network, Opt. Express **19**, 10387 (2011).
[6] D. J. Bernstein and T. Lange, Post-quantum cryptography, Nature **549**, 188 (2017).
[7] P. W. Shor, in *Proceedings 35th Annual Symposium on Foundations of Computer Science* (IEEE Computer Society, Los Alamitos, CA, USA, 1994), p. 124.
[8] P. Kocher, J. Jaffe, and B. Jun, in *Advances in Cryptology—CRYPTO' 99*, edited by M. Wiener (Springer Berlin Heidelberg, Berlin, Heidelberg, 1999), p. 388.
[9] F.-X. Standaert, in *Secure Integrated Circuits and Systems* (Springer, Boston, MA, 2010), p. 27.
[10] A. Huang, S.-H. Sun, Z. Liu, and V. Makarov, Quantum key distribution with distinguishable decoy states, Phys. Rev. A **98**, 12330 (2018).
[11] V. Makarov, A. Anisimov, and J. Skaar, Effects of detector efficiency mismatch on security of quantum cryptosystems, Phys. Rev. A **74**, 22313 (2006).
[12] N. Gisin, S. Fasel, B. Kraus, H. Zbinden, and G. Ribordy, Trojan-horse attacks on quantum-key-distribution systems, Phys. Rev. A **73**, 22320 (2006).
[13] R. Wolf, *Quantum Key Distribution* (Springer International Publishing, Cham, 2021), Vol. 988.
[14] V. Zapatero, T. van Leent, R. Arnon-Friedman, W.-Z. Liu, Q. Zhang, H. Weinfurter, and M. Curty, Advances in device-independent quantum key distribution, npj Quantum Inf. **9**, 10 (2023).
[15] A. Stanco, F. B. L. Santagiustina, L. Calderaro, M. Avesani, T. Bertapelle, D. Dequal, G. Vallone, and P. Villoresi, Versatile and concurrent FPGA-based architecture for practical quantum communication systems, IEEE Trans. Quantum Eng. **3**, 1 (2022).
[16] M. Avesani, G. Foletto, M. Padovan, L. Calderaro, C. Agnesi, E. Bazzani, F. Berra, T. Bertapelle, F. Picciariello, F. B. Santagiustina, D. Scalcon, A. Scriminich, A. Stanco, F. Vedovato, G. Vallone, and P. Villoresi, Deployment-ready quantum key distribution over a classical network infrastructure in Padua, J. Lightwave Technol. **40**, 1658 (2022).
[17] K. Wei, W. Li, H. Tan, Y. Li, H. Min, W.-J. Zhang, H. Li, L. You, Z. Wang, X. Jiang, T.-Y. Chen, S.-K. Liao, C.-Z. Peng, F. Xu, and J.-W. Pan, High-speed measurement-device-independent quantum key distribution with integrated silicon photonics, Phys. Rev. X **10**, 031030 (2020).
[18] H.-F. Zhang, J. Wang, K. Cui, C.-L. Luo, S.-Z. Lin, L. Zhou, H. Liang, T.-Y. Chen, K. Chen, and J.-W. Pan, A real-time QKD system based on FPGA, J. Lightwave Technol. **30**, 3226 (2012).
[19] A. Baliuka, M. Stöcker, M. Auer, P. Freiwang, H. Weinfurter, and L. Knips, Deep-learning-based radio-frequency side-channel attack on quantum key distribution, Phys. Rev. Appl. **20**, 054040 (2023).
[20] K. Durak, N. C. Jam, and S. Karamzadeh, Attack to quantum cryptosystems through RF fingerprints from photon detectors, IEEE J. Sel. Top. Quantum Electron. **28**, 1 (2022).
[21] J. J. Pantoja, V. A. Bucheli, and R. Donaldson, Electromagnetic side-channel attack risk assessment on a practical quantum-key-distribution receiver based on multi-class classification, EPJ Quantum Technol. **11**, 78 (2024).
[22] P. Smith, D. Marangon, M. Lucamarini, Z. Yuan, and A. Shields, Out-of-band electromagnetic injection attack on a quantum random number generator, Phys. Rev. Appl. **15**, 044044 (2021).
[23] A. Kuznetsov, O. Nariezhnii, I. Stelnyk, T. Kokhanovska, O. Smirnov, and T. Kuznetsova, in *2019 10th IEEE International Conference on Intelligent Data Acquisition and*

Advanced Computing Systems: Technology and Applications (IDAACS) (IEEE, Metz, France, 2019), Vol. 2, p. 713.

[24] Y. H. Li, Y. Y. Fei, W. L. Wang, X. D. Meng, H. Wang, Q. H. Duan, Y. Han, and Z. Ma, Effect of external magnetic fields on practical quantum random number generator, EPJ Quantum Technol. **10**, 49 (2023).

[25] T. Messerges, E. Dabbish, and R. Sloan, Examining smart-card security under the threat of power analysis attacks, IEEE Trans. Comput. **51**, 541 (2002).

[26] C. O'Flynn and Z. David Chen, in *2015 IEEE 28th Canadian Conference on Electrical and Computer Engineering (CCECE)* (IEEE, Halifax, NS, Canada, 2015), p. 750.

[27] F. Grünenfelder, A. Boaron, D. Rusca, A. Martin, and H. Zbinden, Simple and high-speed polarization-based QKD, Appl. Phys. Lett. **112**, 051108 (2018).

[28] G. L. Roberts, M. Pittaluga, M. Minder, M. Lucamarini, J. F. Dynes, Z. L. Yuan, and A. J. Shields, Patterning-effect mitigating intensity modulator for secure decoy-state quantum key distribution, Opt. Lett. **43**, 5110 (2018).

[29] C. Agnesi, M. Avesani, A. Stanco, P. Villoresi, and G. Vallone, All-fiber self-compensating polarization encoder for quantum key distribution, Opt. Lett. **44**, 2398 (2019).

[30] Here, V_ϕ corresponds to the voltage required for the modulator to perform a phase shift of ϕ.

[31] T. Güneysu and A. Moradi, in *Cryptographic Hardware and Embedded Systems—CHES 2011*, edited by B. Preneel and T. Takagi (Springer Berlin Heidelberg, Berlin, Heidelberg, 2011), p. 33.

[32] S. Nikova, C. Rechberger, and V. Rijmen, in *Information and Communications Security*, edited by P. Ning, S. Qing, and N. Li (Springer Berlin Heidelberg, Berlin, Heidelberg, 2006), p. 529.

[33] C. Tokunaga and D. Blaauw, Securing encryption systems with a switched capacitor current equalizer, IEEE J. Solid-State Circuits **45**, 23 (2010).

[34] N. Kamoun, L. Bossuet, and A. Ghazel, in *2009 3rd International Conference on Signals, Circuits and Systems (SCS)* (IEEE, Medenine, Tunisia, 2009), p. 1.

[35] E. Tena-Sánchez, F. E. Potestad-OrdÓñez, V. Zúñiga-González, and A. J. Acosta, Low-cost full correlated-power-noise generator to counteract side-channel attacks, Appl. Sci. **15**, 3064 (2025).

[36] S. R. Bommana, S. Veeramachaneni, S. Ershad, and M. B. Srinivas, Mitigating side channel attacks on FPGA through deep learning and dynamic partial reconfiguration, Sci. Rep. **15**, 13745 (2025).

[37] P. Socha, J. Brejník, S. Jeřábek, M. Novotný, and N. Mentens, in *2019 22nd Euromicro Conference on Digital System Design (DSD)* (IEEE, Kallithea, Greece, 2019), p. 277.

[38] P. Sasdrich, A. Moradi, O. Mischke, and T. Güneysu, in *2015 IEEE International Symposium on Hardware Oriented Security and Trust (HOST)* (IEEE, Washington, DC, USA, 2015), p. 130.