Pantoja et al. *EPJ Quantum Technology* (2024) 11:78
https://doi.org/10.1140/epjqt/s40507-024-00290-6

EPJ Quantum Technology
a SpringerOpen Journal

RESEARCH Open Access

Electromagnetic side-channel attack risk assessment on a practical quantum-key-distribution receiver based on multi-class classification

John J. Pantoja[1], Victor A. Bucheli[2] and Ross Donaldson[1*]

*Correspondence:
r.donaldson@hw.ac.uk
[1]Institute of Photonics and Quantum Sciences, Heriot-Watt University, Edinburgh, EH14 4AS, Scotland, UK
Full list of author information is available at the end of the article

Abstract

While quantum key distribution (QKD) is a theoretically secure way of growing quantum-safe encryption keys, many practical implementations are challenged due to various open attack vectors, resulting in many variations of QKD protocols. Side channels are one such vector that allows a passive or active eavesdropper to obtain QKD information leaked through practical devices. This paper assesses the feasibility and implications of extracting the raw secret key from far-field radiated emissions from the single-photon avalanche diodes used in a BB84 QKD quad-detector receiver. Enhancement of the attack was also demonstrated through the use of deep-learning model to distinguish radiated emissions due to the four polarized encoding states. To evaluate the severity of such side-channel attack, multi-class classification based on raw-data and pre-processed data is implemented and assessed. Results show that classifiers based on both raw-data and pre-processed features can discern variations of the electromagnetic emissions caused by specific orientations of the detectors within the receiver with an accuracy higher than 90%. This research proposes machine learning models as a technique to assess EM information leakage risk of QKD and highlights the feasibility of side-channel attacks in the far-field region, further emphasizing the need to utilise mechanisms to avoid electromagnetic radiation information leaks and measurement-device-independent QKD protocols.

Keywords: Quantum key distribution; Quantum communication; Single-photon avalanche diode; Single-photon detector; Information leakage; Electromagnetic security; Side channel attack

1 Introduction

Quantum key distribution (QKD) is a quantum communication protocol that allows two parties to grow a quantum-safe encryption key [1–3]. While theoretically secure, practical implementations of QKD have loopholes due to the use of real hardware and processes. A review of practical security aspects of QKD is presented in [4]. One such loophole is from side-channel attacks, which are intrusions that use fundamental characteristics of the practical system to extract secret information. Side-channels of QKD include power

consumption [5], light emission [6], degrees of freedom of photons [7, 8] or side-band modulation [9].

One relatively unexplored side channel is electromagnetic (EM) emanations of electronic components used in QKD systems. In previous studies on single-photon avalanche diode (SPAD), it was shown that SPADs in Geiger mode produce pulsed-like radio-frequency (RF) radiated emanations that can be used to determine the time of photon detection [10]. These emissions are given by breakdown current pulses in the avalanche diode at each detection of a photon [11] or even at each dark count [12]. These studies show that radiated emissions can be used to remotely detect individual SPAD triggering; however, research on the amount of data leakage and the vulnerability of a QKD receiver terminal have not been conducted.

Studies of EM side-channel attacks on QKD have shown that deep learning techniques and neural networks can be used to recover raw secret key from EM emanations of QKD senders [13] and SPADs [11]. In [11], near-field emissions of two separated SPADs (i.e. 20-cm separation) measured between 30 MHz and 300 MHz at a 2-m distance in a non-anechoic environment were used to show the feasibility of this kind of attack. Although high accuracy rates of data extraction have been achieved from emanations in the near field (i.e. higher than 99%), no key extraction has been proven from far-field emissions [13]. In addition, accuracy is the only parameter used to assess the algorithm's performance, which does not reveal possible characteristics of the attack, such as over-fitting or imbalanced classification.

Advanced QKD protocols have sought to remove issues from single-photon detector side-channel attacks [14] (measurement-device independent (MDI-QKD) protocols) or all side-channel attacks [15] (device-independent (DI-QKD) protocols). MDI-QKD protocol design [16] and implementation breakthroughs have enabled QKD to be transmitted over 1000 km using ultra-low loss fibre [17] in research environments. However, since both MDI- and DI-QKD protocols still have implementation challenges, commercial systems in the first generation of QKD networking will be non-MDI or DI protocols, meaning there is a need to evaluate and mitigate side-channel attacks from single-photon detectors. This research emphasizes how current scientific and commercial space-QKD implementations, which are based on prepare-and-measure protocols, are susceptible to measurement-device side-channel attacks.

This paper studies the feasibility of retrieving secret keys from far-field radiated emissions of a quantum receiver designed for free space and satellite QKD. Emissions from a compact quad-detector free-space polarization-based BB84 QKD quasi-receiver were measured in a semi-anechoic chamber to guarantee the repeatability of EM measurements. Radio-frequency (RF) emissions of four SPADs (different positions and orientations) were measured and analysed using antennas in both vertical and horizontal polarizations. The potential risk of raw key extraction was assessed using machine learning classification methods.

This work generated a dataset with 1200 records of 2500 samples classified into 4 labels available in a repository along with machine learning models. Additionally, a novel approach to understanding and measuring the potential risks associated with raw key extraction using machine learning classification methods is introduced. This approach allows us to measure the system's vulnerability based on evaluation metrics of machine learning models. This paper highlights the EM information leakage risk of free-space and

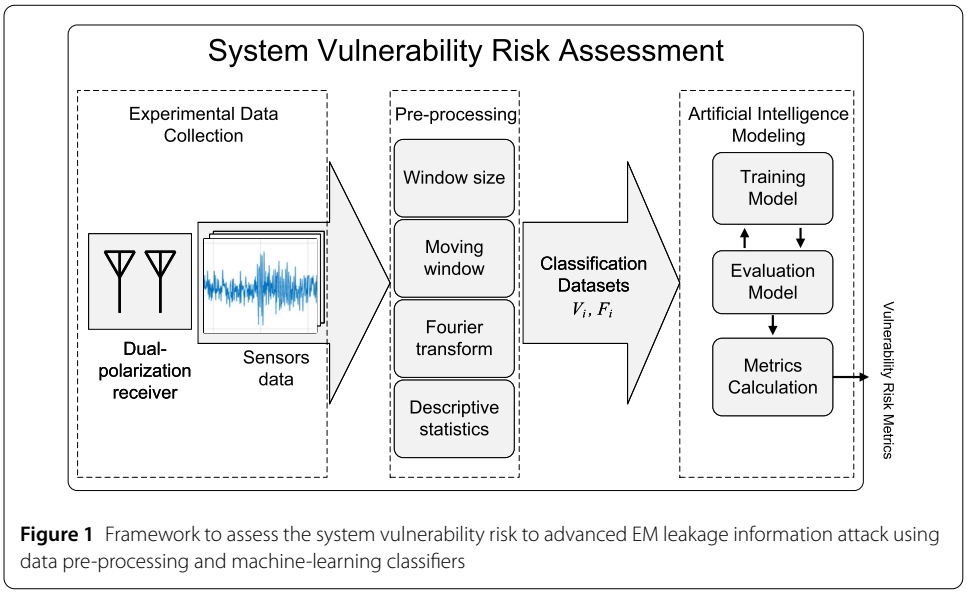

Figure 1 Framework to assess the system vulnerability risk to advanced EM leakage information attack using data pre-processing and machine-learning classifiers

satellite QKD receivers and proposes machine-learning models to assess these vulnerabilities.

The vulnerability risk evaluation approach is summarized in Fig. 1. Experimental data containing time-domain far-field radiated EM emissions were measured using a dual-polarization receiver. Measured data is then pre-processed and labelled to generate two classification datasets. The dataset V_i is based on windowing the raw data, while the dataset F_i is based on spectral and statistical properties of the captured signals. Finally, artificial intelligence models were used to determine encoding states from the classification datasets. The risk assessment is quantified using model performance metrics. Details of the experimental data collection and pre-processing procedure are presented in Sect. 2. Artificial intelligence modelling is described in Sect. 3. Results, security analysis, and conclusions are presented in Sects. 4, 5, 6 respectively.

2 Experimental setup and data collection

2.1 Quantum receiver module and attacker setup

QKD is based on encoding encryption key data on a quantum state. Polarization is typically used as the "degree of freedom" in free-space transmission due to the negligible impact of atmospheric propagation on the encoding and on the quantum bit error rate (QBER) [18]. In the decoy-state BB84 QKD protocol [19], Alice (transmitter) generates a sequence of four non-orthogonal polarization quantum states using two randomly selected bases and multiple intensity levels and transmits to Bob (receiver) to detect the quantum states, which typically contains four independent SPADs [20]. In the free-space implementation of the decoy-state BB84 protocol, the optical interface for the SPADs can be direct free-space or multimode fibre. If in free-space, the SPADs will typically be orientated to ensure a short distance between the QKD receiver and the detector. If multimode fibre is coupled, the SPADs can be positioned away from the optical system and stacked together, however, long path will result in performance degradation at high operational frequencies due to modal dispersion [21].

To demonstrate the side-channel attack, the optical system was composed of a quasi-QKD transmitter and quasi-QKD receiver, presented in Fig. 2. The transmitter and re-

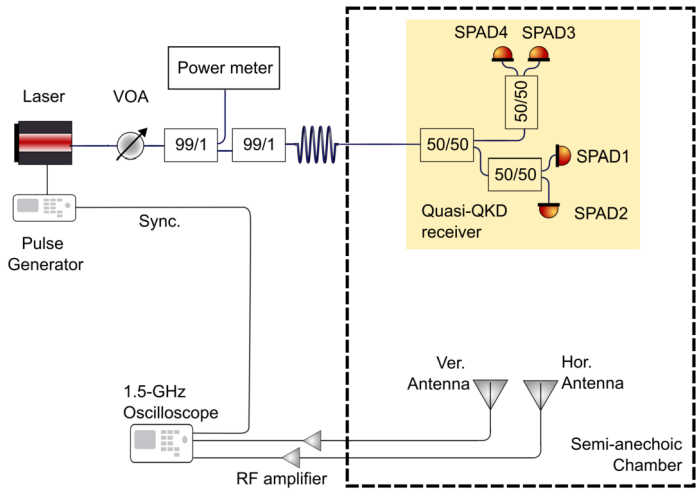

Figure 2 Diagram of experimental setup to test the quasi-quantum key distribution receiver. The transmitter created single-photon-level optical pulses from a gain-switched laser source, a variable optical attenuator (VOA), and two 99/1 fibre optic beamsplitters. The quasi-quantum key distribution receiver is composed of four single-photon avalanche diodes (SPAD)s and three 50/50 fibre-optic beamsplitters. Radiated emissions from the quasi-quantum key distribution receiver are measured using horizontal and vertical polarised antennas and a 1.5 GHz oscilloscope in a semi-anechoic chamber

ceiver were constructed of single-mode fibre components, and the "optical channel" was a single-mode fibre connection.

The quasi-QKD transmitter created single-photon-level optical pulses using a gain-switched laser with a wavelength of 940 nm. The laser driver was triggered by a pulse pattern generator that delivered 1-ns pulses at a clock frequency of 100 MHz. Optical attenuation provided by a variable optical attenuator (VOA) and two 99/1 fibre optic beamsplitters enabled adjustment of the average number of photons per pulse to 0.001. The low mean photon number meant there was a low probability of two SPADs measuring events. The transmitter was kept outside of the semi-anechoic chamber. The quasi-QKD receiver, within the semi-anechoic chamber, comprised three passive 50:50 beamsplitters and four independent silicon-based SPADs, as shown in Fig. 2. The SPADs were connected to a time-correlated single-photon counting unit with four channels that recorded time-of-arrival events with respect to a synchronisation channel from the pulse pattern generator. The time-tag events were used to correlate the information measured with other equipment. The detector layout of the quasi-QKD receiver was based on a free-space QKD receiver design, which had free-space optical interfaces to the SPADs. In our quasi-QKD receiver, the system is single-mode fibre coupled, but the SPAD layout is the same as it would be with free-space optical components. As shown in Fig. 3, physical orientations of SPADs' body are horizontal for SPADs number 1, 3, and 4, and vertical for SPAD number 2. In addition, SPADs number 3 and 4 are attached to the same face of the quantum receiver.

2.2 Radiated emissions of a quantum receiver

The radiated emissions from the quasi-QKD receiver were measured using two double-ridge antennas. Measurements were performed in a semi-anechoic dark chamber to reduce external electromagnetic interference and avoid unintended reflections. Two anten-

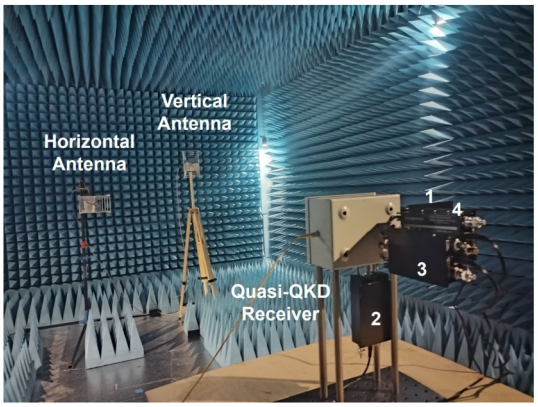

Figure 3 Experimental setup to study the polarization dependency of the radiated emissions from the quasi-quantum key distribution receiver. The SPADs are individually identified with numbers from 1-4

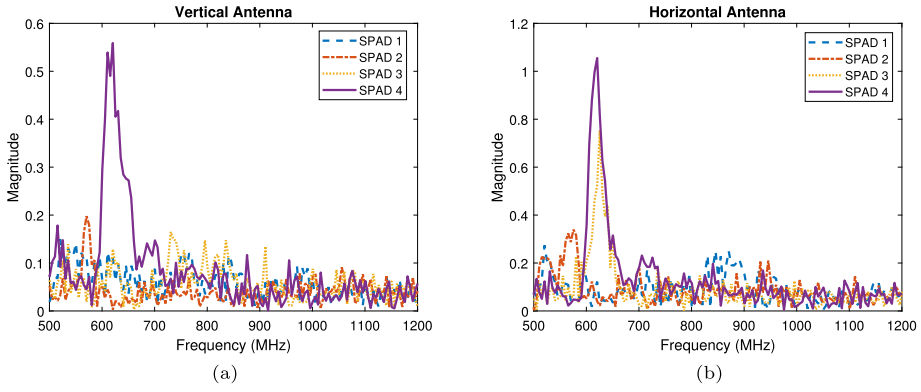

Figure 4 Spectrum of radiated emissions of a quasi-quantum receiver with 4 SPADs measured by two antennas in (a) vertical and (b) horizontal polarization

nas captured horizontal and vertical polarized emissions at a 3-m distance. The received signals were amplified and measured with a 1.5-GHz bandwidth oscilloscope. The quasi-QKD receiver was located on a non-conductive table at 80 cm height.

The spectrum of typical signals radiated by the four SPADs in the quasi-QKD receiver are shown in Fig. 4. It is shown that the received emissions from SPADs with different physical orientations have different frequency components. These differences in the spectrum for each SPAD are due to variability in the avalanche breakdown process, resonances in the assembly, and variation of paths and loads of connections such as power cable leads [22, 23]. The frequency components radiated by the quasi-QKD receiver can be used to individually identify each SPAD; however, SPADs with the same physical orientation and close position (i.e. SPADs 3 and 4 as shown in Fig. 3) produce frequency components in the same frequency range. In this case, these SPADs are indistinguishable based only on frequency value for most of the configurations. For this reason, a more advanced attack based on machine learning summarized in Fig. 1 is discussed in the next section.

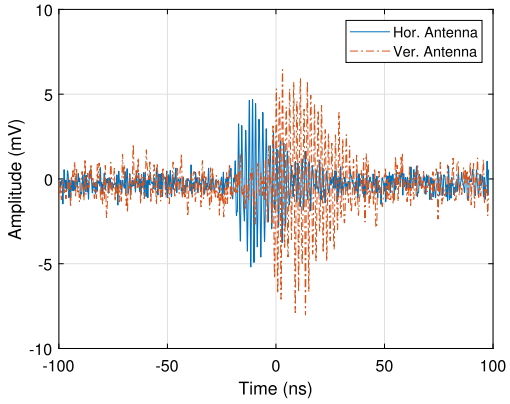

Figure 5 Example of signals measured in horizontal and vertical polarization to generate the dataset

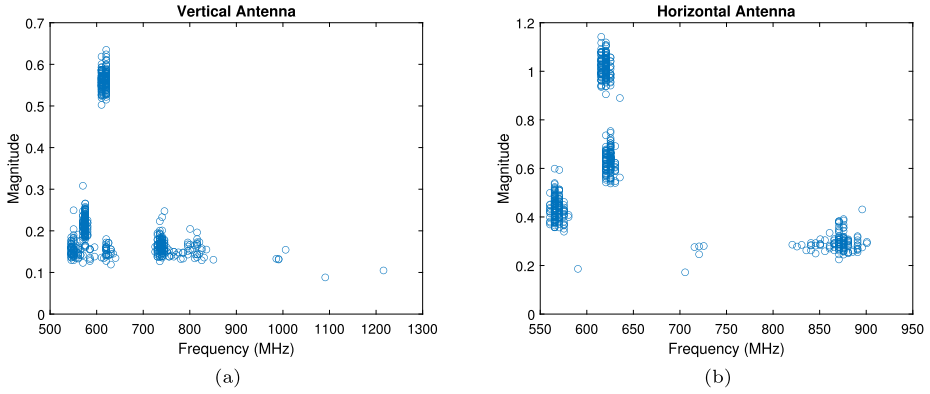

Figure 6 Main spectral component from 600 radiated-emissions records of the quasi-QKD receiver measured by (a) vertically and (b) horizontally polarized antenna

2.3 Dataset

The dataset contains 600 measurements per antenna of radiated signals of the quasi-QKD receiver. That is a total of 1200 measurements. Time-domain signals captured by both antennas, as shown in Fig. 5, and the 20-MHz reference clock from the pulse generator were acquired. Each record is a vector of 2500 samples of the amplitude measured by the oscilloscope that ranges between −10 and 10 mV. The records are either noise or a signal carrying the emission radiated by one of the four SPADs in the quasi-QKD receiver during each acquisition. Finally, each record was associated with one SPAD or noise with a label. Records tagged as noise were discarded and not included in the analysis with machine-learning models.

Figure 6 shows the amplitude and frequency of the highest spectral component of each measurement. It is shown that measurements are grouped in 4 data groups for the vertical antenna, which would allow for a way of classification. On the other hand, only three clusters can be identified for the horizontal antenna, which indicates that two SPADs (SPADs 3 and 4) present similar spectral components.

3 Classification modeling

3.1 Pre-processing data

Based on the dataset, two signal classification scenarios were studied. First, the raw signal was sub-sectioned and applied to a deep neural network (DNN) model. Second, features based on statistical and spectral measures were applied to the decision tree model.

For the first signal recognition approach, subsections were extracted from each record. The size of those subsections, which correspond to observation windows, were selected in a range from 937 to 1875 data points (i.e. time windows from −25 to 50 ns). Using this data, a feature vector, V_i, given by equation (1) is constructed by randomly selecting the starting position and window size, m, to store the amplitude values v_{ni}. This process samples each signal using moving windows of a specified size. Finally, each feature vector was augmented with the corresponding label, L_i, indicating the SPAD that generated the signal.

$$V_i = [v_{1i}, v_{2i}, v_{3i}. \ldots . v_{mi}, L_i] \tag{1}$$

A total of 487 records per antenna with a ratio of 60% for training and 40% for testing was used. The training dataset was built with 292 records from the original dataset, while the test dataset included 195 records to have a representative sample with 95% confidence and a margin of error of 5%. After sampling, 20000 and 5000 observation windows for training and testing were obtained, respectively. In constructing both the training and validation datasets, the choice of records for sampling was also randomized, ensuring that the validation data remained distinct from the training data. This way, the rule of using 60% of the data for training and 20% for validation is used. The data set was the input of the DNN model discussed in Sect. 3.2.

For the second recognition approach, descriptive statistical and spectral features were calculated from the 600 records measured for each antenna and used to train a decision tree model. Statistical measures considered here were mean m_i, standard deviation sd_i, and RMS value rms_i. On the other hand, considered spectral features are the frequency and magnitude of the 6 highest peaks of the fast-Fourier transform; respectively, $[f_{1i}, f_{2i}, \ldots f_{6i}]$ and $[pk_{1i}, pk_{2i}, \ldots pk_{6i}]$. Again, each feature vector is augmented with the corresponding label L_i as shown in equation (2). A total of 15 features per antenna were obtained with this approach. The performance of a decision tree classifier with this set of features is analysed in Sect. 3.2.

$$F_i = [m_i, sd_i, rms_i, f_{1i}, f_{2i}, \ldots f_{6i}, pk_{1i}, pk_{2i}, \ldots pk_{6i}, L_i] \tag{2}$$

3.2 Artificial intelligence models

Multi-class classifiers were used to analyze the far-field radiated emissions from the quasi-QKD receiver as part of a side-channel attack strategy. A DNN model with a tailored architecture was employed to capture the complex patterns and relationships in the subsections of the signals contained in V_i. The implemented model is described in Appendix 6. Other multi-class classifiers, such as gradient-boosted trees and random forest, were also implemented, but a low performance (i.e. accuracy lower than 80%) was obtained, showing that more robust techniques such as DNN are required. For the second signal recognition approach based on the feature vector F_i, a simple decision tree with 100 maximum splits produced acceptable results.

Table 1 Comparison of classifier models based on pre-processed features and raw data

Input data	Classifier	Accuracy (%)	Number of features	Feature calculation (obs/ms)	Prediction speed (obs/ms)	Overall speed (obs/ms)
Raw data: V_i	DNN	96.0	938/2500	N/A	0.6*	0.6*
Pre-processed data: F_i	Tree	99.4	30/30	2.0	6.3	1.5

*Speed for single evaluation. Batch evaluation speed is 67.5 obs/ms.

Table 2 Confusion matrix of SPAD prediction obtained with the DNN applied to V_i

		Predicted SPAD				
		1	2	3	4	Total
Actual SPAD	1	715	10	8	16	749
	2	4	1236	7	14	1261
	3	3	5	1372	28	1408
	4	2	18	83	1479	1582
	Total	724	1269	1470	1537	

Table 3 Confusion matrix of SPAD prediction obtained with the Tree Classifier applied to F_i

		Predicted SPAD				
		1	2	3	4	Total
Actual SPAD	1	114	0	0	0	114
	2	0	138	1	0	139
	3	0	0	156	0	156
	4	0	0	0	131	131
	Total	114	138	157	131	

Table 1 presents a comparison between models based on moving windows (V_i) and pre-processed signal features (F_i). The accuracy and processing speed on a 10-core Intel i5 processor are included in the table. The confusion matrices of both approaches are shown in Tables 2 and 3. The Tree classifier based on pre-processed signal features has higher accuracy and overall prediction speed than the classifier based on raw data. These results show that if proper data pre-processing is performed, basic models can be used as classifiers with superior accuracy. In terms of processing speed, results show that the processing speed can increase up to 100 times if batch evaluation instead of single evaluation is applied.

An accuracy of 96.0% was obtained with the DNN model. As shown in Table 2, this high accuracy means that there is a low loss (4.0%) of the raw key retrieved from the EM emanations. In particular, errors in classification are given when predicting SPADs 3 and 4. Errors in classification for SPAD 4 can be explained by the close physical location of SPADs 3 and 4, which leads to similar radiated emissions, and to the partial shielding provided by the metal body of the QKD receiver.

4 Results

The system vulnerability against information leakage can be quantified with the accuracy of the multi-class classifiers to show the general risk of the attack. However, other metrics can be used to understand the severity of the risk. In this section, we show vulnerability risk metrics based on other parameters of the multi-class classifiers' performances to get more

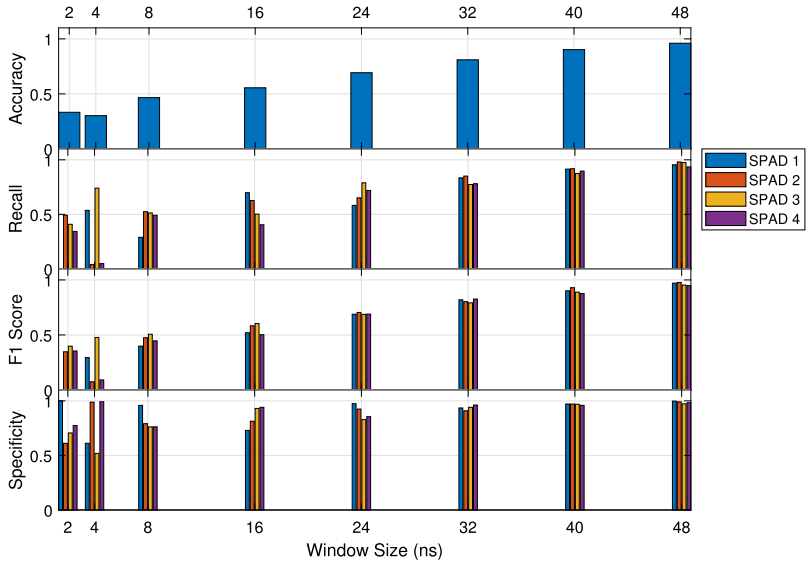

Figure 7 Vulnerability risk metrics of DNN model as a function of the moving window size

insights into the effects. In addition to Accuracy, we have considered Recall, F1 Score, and Specificity as assessment metrics. Recall quantifies the classification performance for each individual SPAD, while Specificity measures the proportion of other SPADs which are correctly identified by the model. F1 Score, on the other hand, shows a balance between precision and recall.

Figure 7 shows the selected vulnerability risk metrics for the DNN model for different window sizes. As the time window is higher, the vulnerability increases since more information from the radiated emission is collected by the eavesdropper. A window over 40 ns is required to exceed 90% accuracy. For shorter windows, the parameters quickly degrade the performance. These results show that this attack, as implemented in this example, is unsuitable for high key rates and or fast clock frequencies in which short acquisition times would be required. 90% accuracy is obtained for a key rate of 25 Mb/s. For higher key rates, accuracy, recall, and F1 score are degraded. For illustration, a key rate of 125 Mb/s, where an observation window size of 8 ns is available, these three metrics are below 52%.

Figure 8 shows the vulnerability risk metrics as a function of the number of training records used in the DNN model. In general, the figure shows that the accuracy depends on the number of training records. All cases show an accuracy higher than 90%. However, if the other metrics are considered, a degradation of the model performance in predicting SPAD 1 and 4 is shown as the number of records is reduced. Since the performance is reduced with 2 SPADs, probably due to confusion in the prediction, the best parameter to see the trend is F1 Score and Recall. More than 100 training records are required to have at least 80% recall and F1 Score in all the SPADs.

5 Security analysis

Experimental results show that the data from the raw secret key received by the quasi-QKD receiver can be extracted from its far-field radiated emissions with an accuracy of 99.6%. This highlights a possible vulnerability risk in unprotected SPAD-based QKD implementations where the raw secret key can be used for single-trace attacks without re-

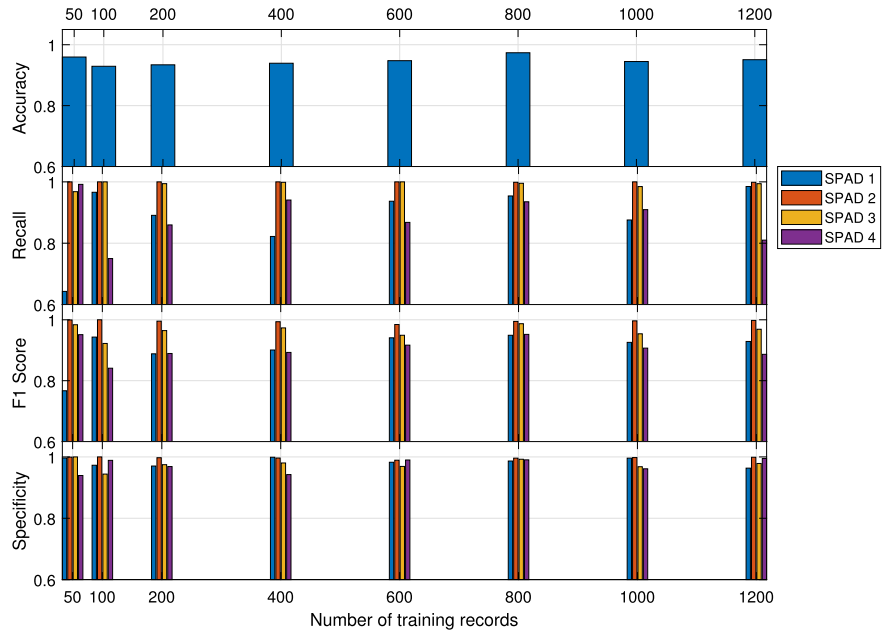

Figure 8 Vulnerability risk metrics of DNN model as a function of number of training records per antenna

vealing the eavesdropper since the quantum channel is not altered. From a security perspective, this profiled attack [13] assumes the following conditions:

- The attacker has access to the physical perimeter of the QKD receiver and is able to measure the radiated emissions. In these measurements, 3-m standard distance and general-purpose instruments were used; however, longer distances can be achieved with specialized receivers as discussed in [24].
- The attacker has a copy of the victim's device [13] or particular knowledge about the victim's set-up to characterize the expected radiated emissions in a controlled environment and train and test the classifier model or a dataset with enough amount of labelled signals.
- The attacker can access the classical public channel to perform and/or monitor the public discussion, including basis set reconciliation to perform key sifting, QBER estimation, error correction, and privacy amplification. As a result, the attacker is able to obtain the shared secret key.

It is important to note that the synchronization signal was not used to classify the radiated emissions, whereas, in other attacks, it has been needed [13, 25]. The synchronization signal and other auxiliary signals from the classical public channel may be used in scenarios with a high signal-to-noise ratio or with multiple reflections.

Security proofs have been studied using practical light sources in QKD transmitters. An example is the method proposed in [26], where the security of QKD protocols is guaranteed when the fidelity between the side-channel-affected states and the ideal emitted states is known. However, security proofs considering side-channel attacks on the QKD receiver terminal have not been sufficiently established. For this reason, countermeasures to reduce the vulnerability of EM attacks are advised [27].

Countermeasures for the side-channel attack discussed here include attenuating the EM fields using shielding barriers and distance, as detailed in [24]. Also, a specific design

of low-EM-emissions receivers dedicated to QKD is advisable. As explained in Sect. 2.2 and shown in Fig. 6, radiated emissions are easily distinguishable if SPADs have different physical orientations and longer separations. Therefore, limiting the distance between the SPADs and circuits would be beneficial, for example, integrating all circuitry onto one board. In this case, the risk of information leakage is reduced; however, advanced processing may still be used to retrieve the raw key from the circuit's near-field emissions [13]. More sophisticated countermeasures applicable for critical infrastructure include jamming and reducing transients in the detectors [28]. One alternative is the use of superconducting nano-wire single photon detectors (SNSPD) [29, 30] that operate at lower voltages than SPADs [31] and, as a consequence, produce lower-amplitude radiated emissions. The use of MDI-QKD or DI-QKD protocols would also mitigate the EM single-photon detector side-channel attack, as they were designed to close the loophole to detector side-channel issues.

6 Conclusions

While QKD may be theoretically secure (within the bounds of the security analysis), real-world implementation results in loopholes that need to be addressed for the practical implementation to also be secure. One common group of loopholes is side channels, which a passive or active eavesdropper can use to gain partial or all secret key information in QKD.

EM side channels are relatively unexplored. Within this paper, we analyze in more detail an EM side-channel attack based on far-field radiated emissions from the single-photon detectors in a quad-detector quasi-QKD receiver, which does not have shielding. Results show that the raw secret key received by a quasi-QKD receiver can be extracted from its far-field radiated emissions with an accuracy of 99.6%. Experimental results show that variations in the physical location and orientation of single-photon detectors in the quantum receiver produce variations in the waveform and frequency components of their radiated emissions which can be used to identify the activation of individual detectors. Here, it was shown that SPADs' radiated emissions can be used to recover the raw key in polarization-based QKD protocol.

It is worth noting that the EM emissions used in this work to retrieve the secret raw key were measured in the far-field region of the quasi-QKD receiver. Emissions radiated in the far field region propagate through the media, decreasing its amplitude with the inverse of the distance. Consequently, this attack can be done at distances longer than the 3-m measurement distance used here using an RF receiver with higher sensitivity, as discussed in [24]. Signals from an unshielded and unprotected QKD receiver could be detected at a distance of up to 50 m with a basic receiver [12], which highlights the need for shielding to prevent significant emissions from being detected.

As shown in this paper, machine-learning classifiers can take advantage of any information leak. Radiated emissions used in this study to retrieve the raw secret key do not discriminate statistical descriptors. Emissions have noise and overlapped amplitudes and frequency components due to different encoding states. To guarantee secure communication systems, the identification of side-channel attacks and the inclusion of countermeasures are necessary from the design stage. It is important to mention that the tests conducted here were performed with CE-marked and metal-enclosed components, which means that normal electromagnetic compatibility (EMC) testing does not avoid information leakage

through EM channels. Our demonstration shows that artificial intelligence tools can be used by attackers to enhance eavesdropping capabilities in actual scenarios.

The risk assessment framework proposed here may be used to assess side-channel attacks on quantum communication systems and quantify the risk using vulnerability metrics for pre-compliance testing and for preparation for security certification. This paper introduced the use of machine-learning metrics as vulnerability risk metrics. It was shown that using machine-learning metrics to assess vulnerabilities allows the identification of the prediction and generalization capabilities that a model can achieve from information leakage. Specific metrics, such as recall that determine vulnerabilities in individual detectors that could not be identified with the model accuracy give important information for implementing countermeasures to reduce the effectiveness of the attack.

These results highlight the need for further measurement of side channels and standardization of EM side channel test procedures for first-generation QKD systems and a move to MDI-QKD or DI-QKD protocols for future generations to overcome EM side channel attacks.

Appendix: Deep neural network model

A batch norm classifier was employed as multi-class classifier. This model is a neural network designed for classification tasks. It begins with a fully connected layer with 1000 units, followed by batch normalization to stabilize training. The network then applies activation functions (Ramp and Sigmoid) before processing through another fully connected layer with 4 units. The final output is passed through a Softmax layer, which converts the predictions into probabilities for four distinct classes. The architecture, shown in Table 4, comprises the following layers:

- Linear Layer: A fully connected layer with 1000 units
- Batch normalization layer: A layer that normalizes the activation to improve training stability.
- Element-wise layer: A layer that applies the Sigmoid activation function element-wise.
- Linear layer: A fully connected layer with 4 units, which serves as the final layer before the output.
- Softmax layer: A layer that applies the Softmax function to the output, converting the network's predictions into probabilities for each class.

The network is configured with an input size defined by the input vector size and outputs are decoded using NetDecoder[Class", 1., 2., 3., 4.], which means the output is classified into one of four possible classes.

Table 4 Deep neural network architecture

Parameter	Type	Size
Input	vector	1000
Linear layer	vector	1000
Batch normalization	vector	1000
Ramp layer	vector	1000
Element-wise layer	vector	1000
Linear layer	vector	4
Softmax layer	vector	4
Output	class	4

Abbreviations
EMC, Electromagnetic compatibility; DI-QKD, Device-independent quantum key distribution; DNN, Deep Neural Network; MDI-QKD, Measurement-device independent quantum key distribution; OBS, Observations; QBER, Quantum bit error rate; QKD, Quantum key distribution; RF, Radio-frequency; SNSPD, Superconducting nano-wire single photon detectors; SPAD, Single-photon avalanche diode; VOA, Variable optical attenuator.

Acknowledgements
The authors thank M. Natrella, J. Kirrane, and M. Stonehouse from Arqit Quantum for the technical discussions during the project's development and A. Tello and C. Simmons from Heriot-Watt University for the suggestions on the QKD receiver implementation and discussions on the QKD protocol. The authors thank Dimitrios Anagnostou from Heriot-Watt University for the support during the testing in the microwave laboratory.

Author contributions
John Pantoja performed the experimental work, data analysis, formulated the feature extraction method, and wrote the first version of the manuscript. Ross Donaldson conceived the idea for this research and reviewed the manuscript. Victor Bucheli performed the deep-learning classification modeling and reviewed the manuscript.

Funding
This project is funded by Innovate UK (App No. 10005967), Engineering and Physical Sciences Research Council (EP/T001011/1), and the Royal Academy of Engineering Early Career Fellowship scheme (Grant No. RF/201718/1746).

Data Availability
Data underlying the results presented are available in Heriot-Watt University Archive at [32]

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
All authors reviewed and approved the

Competing interests
The authors declare no competing interests.

Author details
[1]Institute of Photonics and Quantum Sciences, Heriot-Watt University, Edinburgh, EH14 4AS, Scotland, UK. [2]Escuela de Ingeniería de Sistemas y Computación, Universidad del Valle, Ciudad Universitaria Meléndez, Cali, Colombia.

Received: 15 August 2024 Accepted: 4 November 2024 Published online: 18 November 2024

References

1. Scarani V, Bechmann-Pasquinucci H, Cerf NJ, Dusek M, Lütkenhaus N, Peev M. The security of practical quantum key distribution. Rev Mod Phys. 2009;81:1301–50. https://doi.org/10.1103/RevModPhys.81.1301.
2. Lo H-K, Curty M, Tamaki K. Secure quantum key distribution. Nat Photonics. 2014;8:595–604. https://doi.org/10.1038/nphoton.2014.149.
3. Inamori H, Lütkenhaus N, Mayers D. Unconditional security of practical quantum key distribution. Eur Phys J D. 2007;41:599–627. https://doi.org/10.1140/epjd/e2007-00010-4.
4. Xu F, Ma X, Zhang Q, Lo H-K, Pan J-W. Secure quantum key distribution with realistic devices. Rev Mod Phys. 2020;92:25002. https://doi.org/10.1103/RevModPhys.92.025002.
5. Kim S, Jin S, Lee Y, Park B, Kim H, Hong S. Single trace side channel analysis on quantum key distribution. In: 2018 international conference on Information and Communication Technology Convergence (ICTC). 2018. p. 736–9. https://doi.org/10.1109/ICTC.2018.8539703.
6. Meda A, Degiovanni IP, Tosi A, Yuan Z, Brida G, Genovese M. Quantifying backflash radiation to prevent zero-error attacks in quantum key distribution. Light: Sci Appl. 2017;6:16261–16261. https://doi.org/10.1038/lsa.2016.261.
7. Nauerth S, Fürst M, Schmitt-Manderbach T, Weier H, Weinfurter H. Information leakage via side channels in freespace bb84 quantum cryptography. New J Phys. 2009;**11**. https://doi.org/10.1088/1367-2630/11/6/065001.
8. Sharma T, Biswas A, Chandravanshi P, Prabhakar S, Singh RP. Vulnerability in free space qkd due to detection coupling mismatch. IEEE J Quantum Electron. 2023;59(6):1–7. https://doi.org/10.1109/JQE.2023.3318585.
9. Jain N, Derkach I, Chin H-M, Filip R, Andersen UL, Usenko VC, Gehring T. Modulation leakage vulnerability in continuous-variable quantum key distribution. Quantum Sci Technol. 2021;6:045001. https://doi.org/10.1088/2058-9565/ac0d4c.
10. Pantoja J, Tello A, Anagnostou D, Kirrane J, Stonehouse M, Koehler-Sidki A, Natrella M, Donaldson R. Radiofrequency emanations of a single-photon detector. In: IET conference proceedings. 2023. p. 55–594.
11. Durak K, Jam NC, Karamzadeh S. Attack to quantum cryptosystems through rf fingerprints from photon detectors. IEEE J Sel Top Quantum Electron. 2022;**28**. https://doi.org/10.1109/JSTQE.2021.3089638.
12. Pantoja JJ, Tello A, Anagnostou DE, Kirrane J, Stonehouse M, Koehler-Sidki A, Natrella M, Donaldson R. Investigation on the electromagnetic radiated emissions of a single-photon avalanche diode. In: Padgett MJ, Fedrizzi A, Holynski M, Politi A, editors. Quantum technology: driving commercialisation of an enabling science IV, vol. 12795. SPIE; 2023. p. 1279509. https://doi.org/10.1117/12.2689901.

13. Baliuka A, Stöcker M, Auer M, Freiwang P, Weinfurter H, Knips L. Deep-learning-based radio-frequency side-channel attack on quantum key distribution. Phys Rev Appl. 2023;20:54040. https://doi.org/10.1103/PhysRevApplied.20.054040.
14. Lo H-K, Curty M, Qi B. Measurement-device-independent quantum key distribution. Phys Rev Lett. 2012;108:130503. https://doi.org/10.1103/PhysRevLett.108.130503.
15. Zapatero V, Leent T, Arnon-Friedman R, Liu W-Z, Zhang Q, Weinfurter H, Curty M. Advances in device-independent quantum key distribution. npj Quantum Inf. 2023;9:10. https://doi.org/10.1038/s41534-023-00684-x.
16. Lucamarini M, Yuan ZL, Dynes JF, Shields AJ. Overcoming the rate–distance limit of quantum key distribution without quantum repeaters. Nature. 2018;557:400–3. https://doi.org/10.1038/s41586-018-0066-6.
17. Liu Y, Zhang W-J, Jiang C, Chen J-P, Zhang C, Pan W-X, Ma D, Dong H, Xiong J-M, Zhang C-J, Li H, Wang R-C, Wu J, Chen T-Y, You L, Wang X-B, Zhang Q, Pan J-W. Experimental twin-field quantum key distribution over 1000 km fiber distance. Phys Rev Lett. 2023;130:210801. https://doi.org/10.1103/PhysRevLett.130.210801.
18. Liao S-K, Cai W-Q, Liu W-Y, Zhang L, Li Y, Ren J-G, Yin J, Shen Q, Cao Y, Li Z-P, Li F-Z, Chen X-W, Sun L-H, Jia J-J, Wu J-C, Jiang X-J, Wang J-F, Huang Y-M, Wang Q, Zhou Y-L, Deng L, Xi T, Ma L, Hu T, Zhang Q, Chen Y-A, Liu N-L, Wang X-B, Zhu Z-C, Lu C-Y, Shu R, Peng C-Z, Wang J-Y, Pan J-W. Satellite-to-ground quantum key distribution. Nature. 2017;549:43–7. https://doi.org/10.1038/nature23655.
19. Lo H-K, Ma X, Chen K. Decoy state quantum key distribution. Phys Rev Lett. 2005;94:230504. https://doi.org/10.1103/PhysRevLett.94.230504.
20. Noblet Y, Donaldson R. Bb84 quantum key distribution transmitter utilising broadband sources and a narrow spectral filter. Opt Express. 2023;31:15145–55. https://doi.org/10.1364/OE.487424.
21. Lee A, Castillo AT, Whitehill C, Donaldson R. Quantum bit error rate timing jitter dependency on multi-mode fibers. Opt Express. 2023;31:6076–87. https://doi.org/10.1364/OE.477156.
22. Pantoja JJ, Vega F, Román F. Susceptibility frequency regions of deterministic loads connected to circuits with arbitrary characteristics. In: 2016 IEEE global electromagnetic compatibility conference (GEMCCON). 2016. p. 1–4. https://doi.org/10.1109/GEMCCON.2016.7797311.
23. Pantoja JJ, Peña N, Rachidi F, Vega F, Román F. Characterization, modeling, and statistical analysis of the electromagnetic response of inert improvised explosive devices. IEEE Trans Electromagn Compat. 2014;56(2):393–403. https://doi.org/10.1109/TEMC.2013.2284964.
24. ITU-T: K.84: Test methods and guide against information leaks through unintentional electromagnetic emissions. 2011.
25. Smith PR, Marangon DG, Lucamarini M, Yuan ZL, Shields AJ. Out-of-band electromagnetic injection attack on a quantum random number generator. Phys Rev Appl. 2021;15:044044. https://doi.org/10.1103/PhysRevApplied.15.044044.
26. Pereira M, Kato G, Mizutani A, Curty M, Tamaki K. Quantum key distribution with correlated sources. Sci Adv. 2020;6:4487. https://doi.org/10.1126/sciadv.aaz4487.
27. Li Y-H, Fei Y-Y, Wang W-L, Meng X-D, Wang H, Duan Q-H, Han Y, Ma Z. Effect of external magnetic fields on practical quantum random number generator. EPJ Quantum Technol. 2023;10:49. https://doi.org/10.1140/epjqt/s40507-023-00206-w.
28. Martin M, Sunmola F, Lauder D. Unintentional compromising electromagnetic emanations from it equipment: a concept map of domain knowledge. Proc Comput Sci. 2022;200:1432–41. https://doi.org/10.1016/j.procs.2022.01.344. 3rd International Conference on Industry 4.0 and Smart Manufacturing.
29. Dolphin JA, Paraïso TK, Du H, Woodward RI, Marangon DG, Shields AJ. A hybrid integrated quantum key distribution transceiver chip. npj Quantum Inf. 2023;9:84. https://doi.org/10.1038/s41534-023-00751-3.
30. Grünenfelder F, Boaron A, Resta GV, Perrenoud M, Rusca D, Barreiro C, Houlmann R, Sax R, Stasi L, El-Khoury S, Hänggi E, Bosshard N, Bussières F, Zbinden H. Fast single-photon detectors and real-time key distillation enable high secret-key-rate quantum key distribution systems. Nat Photonics. 2023;17:422–6. https://doi.org/10.1038/s41566-023-01168-2.
31. Verma VB, Lita AE, Vissers MR, Marsili F, Pappas DP, Mirin RP, Nam SW. Superconducting nanowire single photon detectors fabricated from an amorphous mo0.75ge0.25 thin film. Appl Phys Lett. 2014;105:022602. https://doi.org/10.1063/1.4890277.
32. Pantoja J, Donaldson R. Radiated-emissions records of the quasi-QKD receiver measure - main spectral component. https://doi.org/10.17861/3db4374a-d394-42ce-b4ba-3e4612220817

Publisher's Note