Dear Author

Please use this PDF proof to check the layout of your article. If you would like any changes to be made to the layout, you can leave instructions in the online proofing interface. First, return to the online proofing interface by clicking "Edit" at the top page, then insert a Comment in the relevant location. Making your changes directly in the online proofing interface is the quickest, easiest way to correct and submit your proof.

Please note that changes made to the article in the online proofing interface will be added to the article before publication, but are not reflected in this PDF proof.

If you would prefer to submit your corrections by annotating the PDF proof,
please download and submit an annotatable PDF proof by clicking the link below.

Annotate PDF

AUTHOR QUERY FORM

AIP Publishing	Journal: APL Photon. Article Number: APP24-AR-01137	Please provide your responses and any corrections by annotating this PDF and uploading it to AIP's eProof website as detailed in the Welcome email.

Dear Author,

Below are the queries associated with your article. Please answer all of these queries before sending the proof back to AIP.

Article checklist: In order to ensure greater accuracy, please check the following and make all necessary corrections before returning your proof.

1. Is the title of your article accurate and spelled correctly?
2. Please check affiliations including spelling, completeness, and correct linking to authors.
3. Did you remember to include acknowledgment of funding, if required, and is it accurate?

Location in article	Query/Remark: click on the Q link to navigate to the appropriate spot in the proof. There, insert your comments as a PDF annotation.
Q1	Please check that the author names are in the proper order and spelled correctly. Also, please ensure that each author's given and surnames have been correctly identified (given names are highlighted in red and surnames appear in blue).
Q2	Please confirm the change in page number in Refs. 1, 8, 17, and 22.
Q3	Please confirm ORCIDs are accurate. If you wish to add an ORCID for any author that does not have one, you may do so now. For more information on ORCID, see https://orcid.org/. Lijiong Shen – 0000-0002-5854-5236 Christian Kurtsiefer – 0000-0003-2190-0684
Q4	Please check and confirm the Funder(s) and Grant Reference Number(s) provided with your submission: National Research Foundation Singapore, Award/Contract Number QEP-P1, Reseach Centre of Excellence Programme Ministry of Education - Singapore, Award/Contract Number Reseach Centre of Excellence Programme Please add any additional funding sources not stated above.

Thank you for your assistance.

Countering detector manipulation attacks in quantum communication through detector self-testing

Cite as: APL Photon. 10, 000000 (2025); doi: 10.1063/5.0233772
Submitted: 17 August 2024 • Accepted: 30 December 2024 •
Published Online: 9 99 9999

Lijiong Shen[1,2] and Christian Kurtsiefer[2,3,a)]

AFFILIATIONS

[1] School of Physics, Hangzhou Normal University, Hangzhou, Zhejiang 311121, China
[2] Centre for Quantum Technologies, National University of Singapore, 3 Science Drive 2, Singapore 117543
[3] Department of Physics, National University of Singapore, 2 Science Drive 3, Singapore 117551

[a)] **Author to whom correspondence should be addressed:** christian.kurtsiefer@gmail.com

ABSTRACT

In practical quantum key distribution systems, imperfect physical devices open security loopholes that challenge the core promise of this technology. Apart from various side channels, a vulnerability of single-photon detectors to blinding attacks has been one of the biggest concerns and has been addressed both by technical means as well as advanced protocols. In this work, we present a countermeasure against such attacks based on self-testing of detectors to confirm their intended operation without relying on specific aspects of their inner working and to reveal any manipulation attempts. We experimentally demonstrate this countermeasure with a typical InGaAs avalanche photodetector, but the scheme can be easily implemented with any single photon detector.

I. INTRODUCTION

Quantum key distribution (QKD) is a communication method that uses quantum states of light as a trusted courier such that any eavesdropping attempt in this information transmission is revealed as part of the underlying quantum physics of the measurement process on the states.[1–3] While the basic protocols are secure within their set of assumptions, practical QKD systems can exhibit vulnerabilities through imperfect implementation of the original protocol scenarios, through imperfect preparation and detection devices, or through side channels that leak information out of the supposedly safe perimeter of the two communication partners.[4–6] Families of such vulnerabilities have been identified and addressed through technical measures and advanced protocols. Examples are the photon number splitting attacks where single photons were approximated by faint coherent pulses,[7,8] Trojan horse attacks,[3,9] various timing attacks,[10–12] and classes of information leakage into parasitic degrees of freedom.

Perhaps the most critical vulnerability of QKD systems is the detector blinding/fake state attack family on single-photon detectors.[13] This attack has been experimentally demonstrated to work for detectors based on avalanche photodiodes and superconducting nanowires[14–16] and allowed to completely recover a key generated by QKD without being noticed by the error detection step in a QKD implementation.[17] The attack is based on the fact that these single photon detectors can be blinded by a macroscopic light level into not giving any response, while an even stronger light pulse or a recovery event from a blinded state could create an output signal from the blinded detector that emulates a photon detection event[13] (see Fig. 1). This vulnerability can be exploited by carrying out an undetected man-in-the-middle attack, where an eavesdropper intercepts photon states carrying the information, measures the quantum state in a basis of his/her choice, and copies the measurement results into the photon detector of the legitimate receiver with macroscopic powers of light.

Various countermeasures against the detector control attack have been suggested and implemented. One class of countermeasures addresses technical aspects of the detectors. Examples are using more than one detector or a multi-pixel detector for one measurement basis,[18–21] including a watchdog detector for the blinding light,[14,22] effectively varying the detector efficiency at random

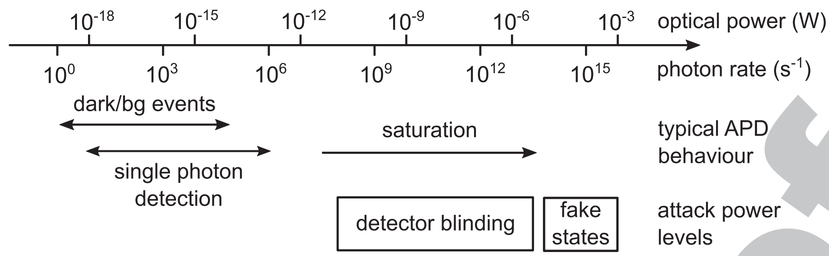

FIG. 1. Single-photon avalanche photodiode properties underlying a blinding/fake state attack. At light levels less than 10^{-12} W, these devices respond with detection events that can be used to identify single photons. At higher power levels, they saturate and can eventually brought into a blinded mode where they are not susceptible anymore to additional single photons. Very bright short pulses of light ("fake states") can lead to a detector response that is indistinguishable from the single photon response at low light levels. Photon rate/power level scaling is shown for a wavelength of 1300 nm.

timings,[23,24] and carefully monitoring the photocurrent, breakdown status, or single-photon detection efficiency of the detector[25–27] to identify a detector manipulation. However, most of these countermeasures have operational drawbacks. For example, additional single photon detectors significantly increase the overall cost and complexity, and beam splitters in the receiver for watchdog detectors introduce additional optical losses. Varying the efficiency frequently to get enough statistics to identify the blinding attack could significantly affect the QKD bit rate and changing the detector operation condition or monitoring its state increases the complexity of the electronic circuitry around the single photon detectors. Such countermeasures may also introduce additional vulnerabilities that may be exploited in an arms race style.[28]

An elegant countermeasure on the protocol level is provided by the so-called measurement-device independent quantum key distribution (MDI-QKD),[29] which further developed the idea of device-independent QKD, where a photon pair source can be made public or even controlled by an eavesdropper[30] to a scenario where the detectors receiving single photons (or approximations thereof) can be public or controlled by an eavesdropper. The scheme has been demonstrated experimentally several times by now.[31–34] It requires a pair of single photons (or weak coherent pulses) from two communication partners without a phase correlation to arrive within a coherence time on a Bell state analyzer, where single photon detection is carried out, and the result is published. This requires a matching of emission times and wavelengths of two spatially separated light sources with both communication partners.

The MDI-QKD approach counteracts any active or passive attack on single photon detectors, as their result need not to be private anymore. The communication partners can simply test if the detectors were performing single photon detection through an error detection process similar to the original QKD protocols.

In this work, we present a method of testing the proper operation of single photon detectors in a QKD scenario that does not require the synchronization of light sources such as in the MDI-QKD approach, while also not touching the specific detector mechanism. It brings the idea of self-testing of quantum systems[35–37] to single photon detectors that can remain black boxes. We use a light emitter (LE) under control of a legitimate communication partner that is weakly coupled to its single photon detector for this self-testing. When the single photon detector is under a blinding attack, it is insensitive to low-intensity light fields used for quantum key distribution. Thus, by turning on the LE at times not predictable by an eavesdropper, "salt" optical detection events are generated in the detector when it operates normal, while it does not react to the test light when blinded. Complementary, the test light intensity can be raised to blinding levels of the photodetector, which is thereby desensitized to legitimate single photons. Registration of any detector events under self-blinding then suggests the presence of fake state events.

II. SELF-TESTING STRATEGY

In a generic QKD system, a transmitter generates photons containing quantum information in either polarization or time encoding and sends them through an optical path ("quantum channel") to a receiver. Therein, a measurement basis choice is made either through passive or active optical components, and the light arriving from the quantum channel is directed to single photon detectors. In a blinding/fake state attack, an eavesdropper measures a photon in the quantum channel and copies the result into the corresponding photon detector of the legitimate receiver using blinding and fake state light levels. For detector testing, a light emitter (LE) in the receiver is controlled by a random number generator and weakly coupled to the single photon detectors.

An unblinded single-photon detector generates events due to photons from the legitimate source or the background [labeled "N" in Fig. 2(a)]. The brightness of the legitimate source, the transmission of the quantum channel, the efficiency of the single photon detectors, and the detector dark count rate determine the average number $\bar{n}$ of the photon-detection events registered in a time interval T. An eavesdropper would choose a rate of "fake" detection events [labeled "F" in Fig. 2(a)] similar to normal QKD operation to prevent detecting the attack by monitoring photon detection statistics.

We illustrate three different examples of detector self-testing to detect detector manipulation attacks.

In the first one, the legitimate receiver switches occasionally the light emitter LE to a low light level for a test time interval T at a random timing unpredictable by an eavesdropper, while it is off for the rest of the time. In the test interval, an unblinded detector would see an increase in detector events above $\bar{n}$ due to additional salt events ["S" in Fig. 2(b)]. The legitimate receiver has complete control of the light emitter to make excess photon detection events statistically

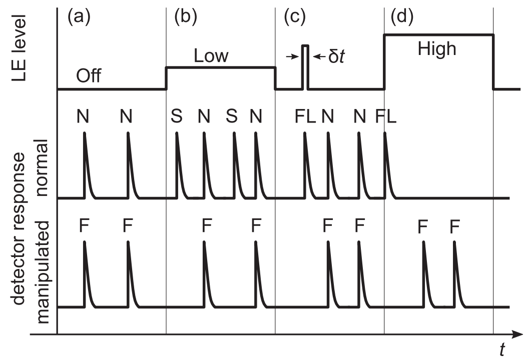

FIG. 2. Detector self-testing. Top trace: light level of the light emitter LE, middle trace: normal detector response (no manipulation), and lower trace: detector response under manipulation. Detector events are classified as normal (N), salt (S), "fake" (F), and flag (FL) signals. Segment (a) shows responses without self-testing, (b) with low LE power generating salt events, (c) with occasional test pulses at medium power, and (d) with high LE power to self-blind the detector.

detectable in the probe interval T. A single photon detector under blinding attack would be insensitive to the low light levels of LE, so only detector events generated by positive detector manipulations such as fake states would be registered [labeled "F" in Fig. 2(b)]. A statistically significant presence of salt events in a time interval T would, therefore, allow sensing a negative detector manipulation, e.g., through blinding. It should be noted that the test interval T does not need to be distributed contiguously in time.

This leads to a second self-testing example, which turns on the light emitter for a short pulse time interval δt at a random timing and with a high enough energy (a few photons) to cause a detection event with almost unit probability in an unblinded single-photon detector. A blinded detector is again insensitive to such a short optical pulse as long as the light level is way below the fake state threshold. In this situation, detecting a single flag event can witness a non-blinded detector [see Fig. 2(c)].

The third self-testing example uses the light emitter in the receiver to locally blind the detector. The typical power necessary to blind an avalanche photodetector (APD) is on the order of a few nW, which can easily be accomplished by weakly coupling even faint light sources such as LEDs. Detection events caused by single photons from the legitimate source will be suppressed by the local blinding light. In the absence of a negative detector manipulation (e.g., detector blinding), the intense light at the onset of the self-blinding period will almost deterministically create a flag event in the detector, which then remains silent during the rest of the self-blinding interval [see Fig. 2(d)]. However, any positive detector manipulation will overrule the local blinding and cause a false detection event. Both the initial flag event and any possible later event can be easily checked. This method only requires a small number of registered events in a time interval T to discover both negative and positive detector manipulation attacks.

A detector event could also be triggered when the detector recovers from a (remote) blinding exposure.[38] Local blinding will suppress such "fake" detector events, so they may not get noticed by looking for signals under local blinding. However, in such a case, the flag event will also be suppressed. Therefore, a combination of checking for detection events during self-blinding and looking for a flag event is necessary to identify such an attack.

III. EXPERIMENTAL RESULTS

We demonstrate our countermeasure with a single-photon detector commonly used in quantum key distribution, which is susceptible to detector manipulation attacks [see Fig. 3(a)]. Light that simulates legitimate quantum signals and provides the larger power levels required for detector manipulation is generated by combining the output of a continuous wave (cw) laser diode (LD1) with light from a pulsed laser diode (LD2) on a fiber beam splitter (BS). The 2 ns long bright fake states from LD2 can be emitted upon detection events from an auxiliary avalanche photodetector (APD1) to emulate a credible (Poissonian) event distribution. On the receiver side, the light from the quantum channel passes through an interference filter (IF) before it is focused onto the main photodetector (APD2), a passively quenched InGaAs device (S-Fifteen Instruments IRSPD1) with a maximal count rate of $5 \times 10^5\ \mathrm{s}^{-1}$ and a dark count rate of $7 \times 10^3\ \mathrm{s}^{-1}$. The light emitter (LE) for detector self-testing is a light emitting diode with a center wavelength of 940 nm (Vishay VSLY5940), which is reflected off the IF (acting as a dichroic beam splitter) onto APD2.

For the demonstration, we consider an event rate of $\approx 5 \times 10^4\ \mathrm{s}^{-1}$ at APD2, which is about an order of magnitude below the maximal detection rate to not reduce the detector efficiency significantly. Figure 3(b) shows a histogram of detection events in a time interval of $T = 200\ \mu$s generated by choosing an appropriate light level of LD1. The result with a mean photodetection number $\bar{n} \approx 10$ differs slightly from a Poisson distribution since the detector

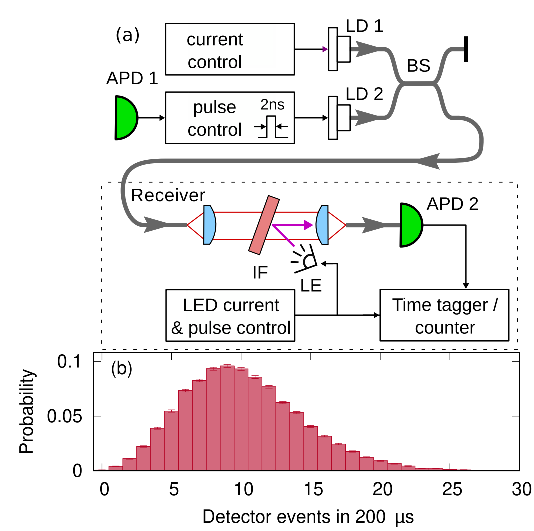

FIG. 3. (a) Setup to demonstrate detector self-testing. Light from a CW laser diode (LD1) and pulsed laser diode (LD2), both around 1310 nm, is combined in a fiber beam splitter (BS) to simulate different illumination scenarios. In addition to the single photon InGaAs detector APD2, the receiver contains an LED (940 nm) as a light emitter (LE) for local testing of APD2. An interference (IF) filter prevents leakage of LE light out of the receiver. (b) Distribution of photodetection events in a time window of $T = 200\ \mu$s under "normal" operation under illumination of the detector with a low power level from LD1.

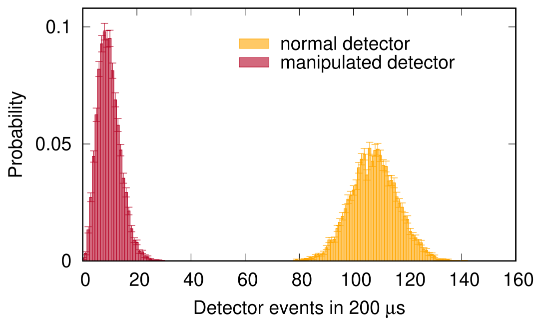

FIG. 4. Distribution of detector events in the presence of self-seeding light in a test interval of $T = 200\ \mu s$ for a normally operating and a manipulated detector. The manipulated detector shows a similar distribution as the one in Fig. 3(b), while the normally operating detector shows a distinctly higher event number. The error bars indicate Poissonian standard deviations resulting from 7432 to 7686 test runs for a normal detector and a manipulated detector, respectively.

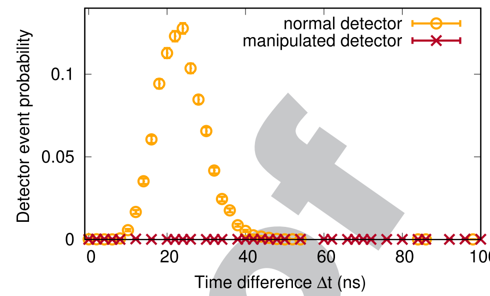

FIG. 5. Detector event probability for a 25 ns long bright pulse of the self-testing light emitter LE for a manipulated and normal detector vs the time difference Δt between detector event and a self-testing pulse edge. A non-manipulated detector reacts with an event with high probability within less than 60 ns. Optical and electrical delays shift the detector response away from $\Delta t = 0$, and the error bars indicate Poissonian standard deviations resulting from 12 542 to 12 380 test runs for the normal detector and manipulated detector, respectively.

has an after-pulse possibility of about 40%. To implement a detector manipulation with the same event characteristic, we elevate the optical output power of LD1 to 500 pW, the minimal power to completely blind detector APD2. Fake states that emulate photodetection events in APD2 are generated with optical pulses through LD2 with a peak power of 3 μW.

To demonstrate the first example of detector self-testing, we turn on the light emitter LE in the test interval T both for a normally operating and a manipulated detector. The resulting detection event distributions are shown in Fig. 4. For a normally operating detector, the observed APD2 events in the test interval increase significantly to a mean of about $\bar{n}_{T1} \approx 100$, while for a manipulated detector, the distribution is similar to the "normal" distribution with $\bar{n}_N \approx 10$ shown in Fig. 3(b). With a threshold at $n = 50$, the two distributions can be easily distinguished and a detector manipulation attempt (specifically: the presence of a blinding light level) easily identified in a single measurement interval T; in the experiment, the un-manipulated detector never showed less than 78 events, while the manipulated showed never more than 30 events.

The necessary time to detect a manipulated detector can be shortened even further with the second example of self-testing. We demonstrate this by driving the light emitter LE to emit $\delta t = 25$ ns long pulses and increasing the coupling to the detector APD2 compared to the previous example. Figure 5 shows the probability of registering a signal from APD2 as a function of the time Δt after the start of the self-testing pulse. A non-manipulated detector shows an overall detector response probability $p_s = 93.4\%$ within 60 ns (11 720 photon detection events out of 12 542 optical pulse), which is the probability for successfully identifying the detector status in a single-shot test. This number does not reach 100%, as the detector may have been in a recovery state from a previous detection event. For a manipulated detector, i.e., in the presence of both detector blinding and fake states, we find an integral detector event probability $p_f = 0.3\%$ (36 of 12 380 test pulses), which is the false-positive probability. These events were caused by fake states, not by light from the LE. A detector manipulation attack (specifically, the detector blinding) can, therefore, be identified with a few short test pulses to a very high statistical significance. For n test pulses, we classify the detector as "not manipulated" if at least n_{th} detection events are registered. The probability of a correct identification (of the non-manipulated state) is given by

$$P_s = \sum_{k=n_{th}}^{n} \frac{n!}{k!(n-k)!} p_s^k (1-p_s)^{n-k}. \quad (1)$$

Similarly, the overall false-positive probability after n test pulses is given by

$$P_f = \sum_{k=n_{th}}^{n} \frac{n!}{k!(n-k)!} p_f^k (1-p_f)^{n-k}. \quad (2)$$

For example, for the probability values p_s and p_f from the experiment above, $n = 10$, and $n_{th} = 4$, the probability of correctly identifying a non-manipulated detector is $P_s = 99.999\,95\%$, while the false positive probability P_f is only on the order of 10^{-8}. The choice n_{th} for a given n can be optimized to either increase the identification probability of a non-manipulated detector or to reduce number of false positives. The attack detection probabilities exemplified here can be reached with a sparse testing density: assuming a realistic detector dead time of $\tau_D = 1\ \mu s$ after a "true" single photon (or background) detection event and a randomized self-test pulse rate of $r_t = 2000\ s^{-1}$, the above-mentioned probability P_S of confirming a non-manipulated detector can be reached within $T = n/r_t = 5$ ms, while the detector is not available for detection of signal photons for a fraction of $\eta_t = \tau_D r_t = 0.2\%$. Such a reduction of the useful signal detection rate due to self-testing is likely lower than the uncertainties due to other environmental factors in practical systems.

To demonstrate the third example of detector self-testing, we increased the optical power of LE on detector APD2 to a level that it could reliably blind the detector. The minimal power to blind the used InGaAs detector is only 500 pW, while the reverse bias voltage is almost unchanged under this blinding power or even two times the power with the self-blinding. Thus, the amplitude of the fake state signals caused by the intense light pulse also does not vary significantly. Figure 6 shows both a distribution of detection events in a test interval $T = 200\ \mu s$, taken 60 ns after the onset of light emission by LE. The un-manipulated detector is insensitive to single photons

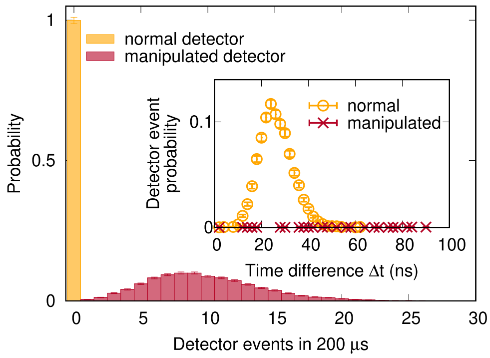

FIG. 6. Detector event distribution in a test interval $T = 200\ \mu$s in the presence of self-blinding light for a normal and manipulated detector, registered 60 ns *after* the onset of the self-blinding light. A manipulated detector still reports events due to fake states. Inset: probability of a detector event in the first 60 ns after switching on the self-blinding light. This scheme allows us to detect the presence of both blinding and fake state detector manipulations.

in this interval; we observed only 8 events in 7608 test runs (likely due to electrical noise), while a manipulated detector still reported events due to fake states present at the input; we observed 7655 of 7658 events (with the missing events compatible with statistics). The onset of the test light emission triggered a detector reaction within the first 60 ns with a probability $p_s = 97.6\%$ (7426 detector events out of 7608 test runs; see the inset of Fig. 6) for a non-manipulated detector, while the probability of an onset event was $p_f = 0.2\%$ (17 of 7658 runs) for a manipulated detector caused by fake states. A local light emitter that is able to self-blind the detector is thus able to reveal the presence of both blinding and the fake state in a detector manipulation attempt.

This countermeasure could be implemented in a QKD system based on multiple single photon detectors simply by equipping each detector with an independent light emitter. In a system based on a passive measurement base choice with a beam splitter, it can be simplified by using only one light emitter in the dark input port of the base choice beam splitter, ensuring all detectors receive roughly the same self-testing intensity.

IV. CONCLUSION

We demonstrated self-testing of single photon detectors that can reliably reveal detector manipulation attacks. The self-testing strategy relies on a light source near the detector under possible external manipulation and is able to detect both negative manipulations (i.e., suppression of single photon detections) and positive manipulations (i.e., generating detector events that are not caused by single photon detections) in a relatively short time with a high statistical significance. Contrary to efficiency variation and monitoring mechanisms to detect single photon detector manipulations, this scheme does not require a careful calibration, as manipulated and non-manipulated detector event statistics under self-testing are very different and do not strongly depend on uncertainties in the self-testing power.

The detector self-testing makes no assumption on the nature of the manipulation attack of the detector and thus also covers manipulations that are not of the known nature, such as detector blinding and fake states. It also makes no assumptions about the specific nature of the detection mechanism, as long as positive or negative detector manipulations are considered possible. Therefore, the method is applicable to all single photon detection mechanisms considered in QKD scenarios. As the self-testing can be accomplished by a relatively simple light source (as long as this is outside the control and knowledge of an adversary), this scheme can address one of the most significant hardware vulnerabilities of QKD systems in a significantly simpler way compared to device-independent or measurement-device independent approaches and may even be a suitable to retrofit existing QKD systems to make them resilient against detector manipulation attacks.

ACKNOWLEDGMENTS

This work was supported by the National Research Foundation (NRF) Singapore (Grant No. QEP-P1) and the Research Centres of Excellence programme supported by NRF Singapore and the Ministry of Education, Singapore.

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Author Contributions

Lijiong Shen: Conceptualization (equal); Data curation (lead); Formal analysis (lead); Investigation (lead); Validation (equal); Visualization (lead); Writing – original draft (lead); Writing – review & editing (equal). **Christian Kurtsiefer**: Conceptualization (equal); Formal analysis (supporting); Funding acquisition (lead); Project administration (lead); Resources (lead); Supervision (lead); Validation (equal); Writing – review & editing (equal).

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

REFERENCES

[1] C. H. Bennett and G. Brassard, "Quantum cryptography: Public key distribution and coin tossing," in *Proceedings of IEEE International Conference on Computers, Systems and Signal Processing* (IEEE, 1984), pp. 175–179.
[2] A. K. Ekert, "Quantum cryptography based on Bell's theorem," Phys. Rev. Lett. **67**, 661–663 (1991).
[3] N. Gisin, G. Ribordy, W. Tittel, and H. Zbinden, "Quantum cryptography," Rev. Mod. Phys. **74**, 145–195 (2002).
[4] V. Scarani, H. Bechmann-Pasquinucci, N. J. Cerf, M. Dušek, N. Lütkenhaus, and M. Peev, "The security of practical quantum key distribution," Rev. Mod. Phys. **81**, 1301–1350 (2009).
[5] V. Scarani and C. Kurtsiefer, "The black paper of quantum cryptography: Real implementation problems," Theor. Comput. Sci. **560**, 27–32 (2014).
[6] F. Xu, X. Ma, Q. Zhang, H. K. Lo, and J. W. Pan, "Secure quantum key distribution with realistic devices," Rev. Mod. Phys. **92**, 025002 (2020).
[7] G. Brassard, N. Lütkenhaus, T. Mor, and B. C. Sanders, "Limitations on practical quantum cryptography," Phys. Rev. Lett. **85**, 1330–1333 (2000).

[8]N. Lütkenhaus, "Security against individual attacks for realistic quantum key distribution," Phys. Rev. A **61**, 052304 (2000).
[9]A. Vakhitov, V. Makarov, and D. R. Hjelme, "Large pulse attack as a method of conventional optical eavesdropping in quantum cryptography," J. Mod. Opt. **48**, 2023–2038 (2001).
[10]B. Qi, C. H. F. Fung, H. K. Lo, and X. Ma, "Time-shift attack in practical quantum cryptosystems," Quantum Inf. Comput. **7**, 73–82 (2007).
[11]A. Lamas-Linares and C. Kurtsiefer, "Breaking a quantum key distribution system through a timing side channel," Opt. Express **15**, 9388 (2007).
[12]Y. Zhao, C. H. F. Fung, B. Qi, C. Chen, and H. K. Lo, "Quantum hacking: Experimental demonstration of time-shift attack against practical quantum-key-distribution systems," Phys. Rev. A **78**, 042333 (2008).
[13]V. Makarov, "Controlling passively quenched single photon detectors by bright light," New J. Phys. **11**, 065003 (2009).
[14]L. Lydersen, C. Wiechers, C. Wittmann, D. Elser, J. Skaar, and V. Makarov, "Hacking commercial quantum cryptography systems by tailored bright illumination," Nat. Photonics **4**, 686–689 (2010).
[15]L. Lydersen, M. K. Akhlaghi, A. Hamed Majedi, J. Skaar, and V. Makarov, "Controlling a superconducting nanowire single-photon detector using tailored bright illumination," New J. Phys. **13**, 113042 (2011).
[16]M. Elezov, R. Ozhegov, G. Goltsman, and V. Makarov, "Countermeasure against bright-light attack on superconducting nanowire single-photon detector in quantum key distribution," Opt. Express **27**(21), 30979–30988 (2019).
[17]I. Gerhardt, Q. Liu, A. A. Lamas-Linares, J. Skaar, C. Kurtsiefer, and V. Makarov, "Full-field implementation of a perfect eavesdropper on a quantum cryptography system," Nat. Commun. **2**, 349 (2011).
[18]T. Honjo, M. Fujiwara, K. Shimizu, K. Tamaki, S. Miki, T. Yamashita, H. Terai, Z. Wang, and M. Sasaki, "Countermeasure against tailored bright illumination attack for DPS-QKD," Opt. Express **21**, 2667 (2013).
[19]T. Ferreira da Silva, G. C. Do Amaral, G. B. Xavier, G. P. Temporao, and J. P. von der Weid, "Safeguarding quantum key distribution through detection randomization," IEEE J. Sel. Top. Quantum Electron. **21**, 159–167 (2015).
[20]J. Wang, H. Wang, X. Qin, Z. Wei, and Z. Zhang, "The countermeasures against the blinding attack in quantum key distribution," Eur. Phys. J. D **70**, 5 (2016).
[21]G. Gras, D. Rusca, H. Zbinden, and F. Bussières, "Countermeasure against quantum hacking using detection statistics," Phys. Rev. Appl. **15**, 034052 (2021).
[22]A. R. Dixon, J. F. Dynes, M. Lucamarini, B. Fröhlich, A. W. Sharpe, A. Plews, W. Tam, Z. L. Yuan, Y. Tanizawa, H. Sato, S. Kawamura, M. Fujiwara, M. Sasaki, and A. J. Shields, "Quantum key distribution with hacking countermeasures and long term field trial," Sci. Rep. **7**, 1978 (2017).
[23]C. C. W. Lim, N. Walenta, M. Legré, N. Gisin, and H. Zbinden, "Random variation of detector efficiency: A countermeasure against detector blinding attacks for quantum key distribution," IEEE J. Sel. Top. Quantum Electron. **21**, 192–196 (2015).
[24]Y.-J. Qian, D.-Y. He, S. Wang, W. Chen, Z.-Q. Yin, G.-C. Guo, and Z.-F. Han, "Robust countermeasure against detector control attack in a practical quantum key distribution system," Optica **6**, 1178 (2019).
[25]Z. Yuan, J. F. Dynes, and A. J. Shields, "Avoiding the blinding attack in QKD," Nat. Photonics **4**, 800–801 (2010).
[26]Z. L. Yuan, J. F. Dynes, and A. J. Shields, "Resilience of gated avalanche photodiodes against bright illumination attacks in quantum cryptography," Appl. Phys. Lett. **98**, 231104 (2011).
[27]Ø. Marøy, V. Makarov, and J. Skaar, "Secure detection in quantum key distribution by real-time calibration of receiver," Quantum Sci. Technol. **2**, 044013 (2017).
[28]A. Huang, S. Sajeed, P. Chaiwongkhot, M. Soucarros, M. Legré, and V. Makarov, "Testing random-detector-efficiency countermeasure in a commercial system reveals a breakable unrealistic assumption," IEEE J. Quantum Electron. **52**, 8000211 (2016).
[29]H.-K. Lo, M. Curty, and B. Qi, "Measurement-device-independent quantum key distribution," Phys. Rev. Lett. **108**, 130503 (2012).
[30]S. Pironio, A. Acín, N. Brunner, N. Gisin, S. Massar, and V. Scarani, "Device-independent quantum key distribution secure against collective attacks," New J. Phys. **11**, 045021 (2009).
[31]Y. Liu, T.-Y. Chen, L.-J. Wang, H. Liang, G.-L. Shentu, J. Wang, K. Cui, H.-L. Yin, N.-L. Liu, L. Li, X. Ma, J. S. Pelc, M. M. Fejer, C.-Z. Peng, Q. Zhang, and J.-W. Pan, "Experimental measurement-device-independent quantum key distribution," Phys. Rev. Lett. **111**, 130502 (2013).
[32]Y.-L. Tang, H.-L. Yin, S.-J. Chen, Y. Liu, W.-J. Zhang, X. Jiang, L. Zhang, J. Wang, L.-X. You, J.-Y. Guan, D.-X. Yang, Z. Wang, H. Liang, Z. Zhang, N. Zhou, X. Ma, T.-Y. Chen, Q. Zhang, and J.-W. Pan, "Measurement-device-independent quantum key distribution over 200 km," Phys. Rev. Lett. **113**, 190501 (2014).
[33]H.-L. Yin, T.-Y. Chen, Z.-W. Yu, H. Liu, L.-X. You, Y.-H. Zhou, S.-J. Chen, Y. Mao, M.-Q. Huang, W.-J. Zhang, H. Chen, M. J. Li, D. Nolan, F. Zhou, X. Jiang, Z. Wang, Q. Zhang, X.-B. Wang, and J.-W. Pan, "Measurement-device-independent quantum key distribution over a 404 km optical fiber," Phys. Rev. Lett. **117**, 190501 (2016).
[34]H. Liu, W. Wang, K. Wei, X.-T. Fang, L. Li, N.-L. Liu, H. Liang, S.-J. Zhang, W. Zhang, H. Li, L. You, Z. Wang, H.-K. Lo, T.-Y. Chen, F. Xu, and J.-W. Pan, "Experimental demonstration of high-rate measurement-device-independent quantum key distribution over asymmetric channels," Phys. Rev. Lett. **122**, 160501 (2019).
[35]D. Mayers and A. C.-C. Yao, "Quantum cryptography with imperfect apparatus," in *Proceedings 39th Annual Symposium on Foundations of Computer Science (Cat. No. 98CB36280)* (IEEE, 1998), pp. 503–509.
[36]W. van Dam, F. Magniez, M. Mosca, and M. Santha, "Self-testing of universal and fault-tolerant sets of quantum gates," in *Proceedings of the Thirty-Second Annual ACM Symposium on Theory of Computing, STOC'00* (Association for Computing Machinery, New York, NY, 2000), p. 688696.
[37]I. Šupić and J. Bowles, "Self-testing of quantum systems: A review," Quantum **4**, 337 (2020).
[38]M. G. Tanner, V. Makarov, and R. H. Hadfield, "Optimised quantum hacking of superconducting nanowire single-photon detectors," Opt. Express **22**(6), 6734–6748 (2014).