European Conference on Optical Communication (ECOC) 2022 © Optica Publishing Group 2022

Countering Detector Manipulation Attacks in Quantum Communication through Detector Self-testing

Lijiong Shen[(1)], Christian Kurtsiefer[(1,2)]

[(1)] Centre for Quantum Technologies, National University of Singapore, 3 Science Drive 2, Singapore 117543,
[(2)] Department of Physics, National University of Singapore, 2 Science Drive 3, Singapore 117551, christian.kurtsiefer@gmail.com

Abstract *Detector manipulation attacks are the most critical vulnerabilities in practical quantum key distribution systems. We present a self-testing method of photodetectors to reveal manipulation by anything but single photon-level signals, which does not rely on specific assumptions about the detection or manipulation mechanism.*

Introduction

In practical quantum key distribution systems (QKD), imperfect physical devices open security loopholes that challenge the core premise of this technology[1]–[3]. Alternatively, A critical vulnerability of QKD systems is the detector blinding / fake state attack family on single-photon detectors[4]. This class of attacks has been experimentally demonstrated to work for detectors based on avalanche photodiodes and superconducting nanowires[5]–[7].

Countering detector manipulation attacks has been heavily investigated over the past decade[8]–[10]. However, most countermeasures have drawbacks in significantly increasing the overall cost and complexity, or reduce significantly the QKD bit rate. Here, we present a self-testing method of photodetectors to reveal manipulation by anything but single photon-level signals.

Self-testing strategy

In a typical blinding attack, the adversary first measures the photons in the quantum channel, and then replicates the result on the corresponding single-photon detector at the legitimate receiver by blinding the detector using a tailored bright illumination, and creating "fake state" using stronger light pulses. We shows three methods of detecting a detector manipulation attack using a simple light emitter (LE) controlled by the legitimate user coupled weakly to the single photon detectors.

When the LE is switched off (Fig. 1(a)), each detector will detect single photons from either the legitimate sender or background events in an unblinded scenario (labeled "N" in Fig. 1). The adversary will try to keep the "fake" state (labeled

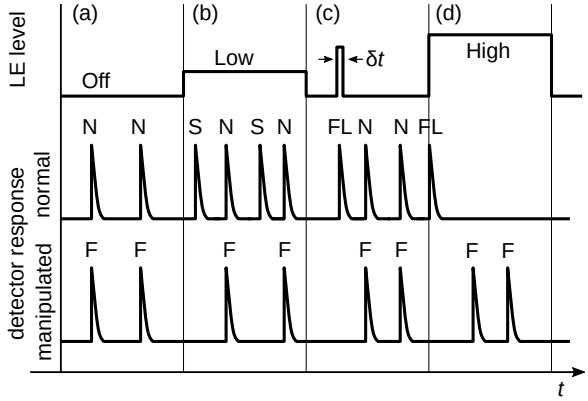

Fig. 1: Detector self testing. Top trace: light level of the light emitter LE, middle trace: normal detector response (no manipulation), lower trace: detector response under manipulation.

"F" in Fig. 1) indistinguishable from the normal detector response to make a successful attack. When the LE is switched on at a low level, the optical power coupled to the single-photon detector is much lower than the blinding power sent by the adversary. Thus, only an unblinded detector will produce additional salt events ("S" in Fig. 1(b)). The legitimate user can detect the blinding attack by monitoring photon detection statistics in time interval T when the LE is switched on.

Alternatively, the attack could also be identified by the absence of flag events ("FL" in Fig. 1(c)) when turning on the LE for a short pulse time interval δt. The optical power should be set at a level just enough to cause a detection event with almost unit probability on an unblinded single-photon detector.

Further increasing the optical power of the LE leads to the third method, where the detector would be self-blinded. A detector not manipulated by an adversary should produce a flag event immediately after switching on the LE, and stay

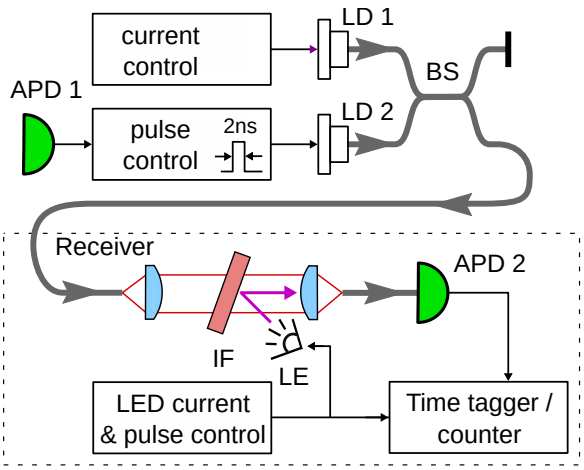

Fig. 2: Setup to demonstrate detector self-testing. Light from a cw laser diode (LD1) and a pulsed laser diode (LD2) is combined in a fiber beamsplitter (BS) to simulate different illumination scenarios. An interference filter (IF) prevents LE (a LED) light leaking to the quantum channel.

silent during the rest of the self-blinding interval (see Fig. 1(d)). Here, any positive detector manipulation will overrule the local blinding and cause a detector event, revealing the manipulation. Both the flag event and the photon detection event in the rest of self-blinding interval could then be used to reveal the adversary's existence.

Experiment

Figure 2 illustrates the experimental setup for the countermeasure demonstration. A low-cost LED acts as the LE for self-testing. We consider an normal event rate of 50 000 s^{-1} on our InGaAs detector (APD2), which is about an order of magnitude below the maximal detection rate, and does not significantly reduce the detector efficiency. For the first method of detector self-testing, we choose a time interval of 200 μs, which contains a mean photon number of 10 in normal operation. We switch on the LED for more than 7000 trials when the detector is non-manipulated as well as when it is blinded. The mean photon number significantly increases to 100 for a non-manipulated detector with a minimum measured photon detection events of 79. On the other hand, the mean photon number of the blinded detector remains at 10 with a maximum photon detection events of 29.

For the second method, a function generator drives the LED to emit 25 ns long pulses. In more than 10000 trials, a non-manipulated detector has 93.4% probability to register output photon detection events within 60 ns, while for a blinded detector, the probability drops down to 0.3%. Therefore, a few short test pulses can identify the attack with very high statistical significance.

For the last self-binding method, the non-manipulated detector has a probability of 97.6% in more than 7000 trials producing flag events 60 ns after the onset of LED light emission and a probability of 99.9% of staying silent in the rest of the time interval T (set at 200 μs). The blinded detector only has 0.2% accidental flag events in the first 60 ns, and has a close to unit probabilityto produce at least one detection event in the remaining of the time interval T.

Conclusion

We experimentally demonstrated that the self-testing concept is able to reliably reveal detector manipulation attacks on a typical InGaAs avalanche photodetector within a very short time. This self-testing concept neither relies on specific assumptions on the detection or manipulation mechanism, nor requires the technologically more complex protocol of a measurement-device independent quantum key distribution[11]. This scheme could also be easily implemented with any single-photon detector and retrofit to existing QKD systems.

Acknowledgements

We thank S-Fifteen Instruments Pte. Ltd. providing us the InGaAs APD for testing.

References

[1] 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.*, vol. 81, no. 3, pp. 1301–1350, 2009, ISSN: 00346861. DOI: 10.1103/RevModPhys.81.1301.

[2] V. Scarani and C. Kurtsiefer, "The black paper of quantum cryptography: Real implementation problems", *Theor. Comput. Sci.*, vol. 560, no. P1, pp. 27–32, Dec. 2014, ISSN: 03043975. DOI: 10.1016/j.tcs.2014.09.015.

[3] F. Xu, X. Ma, Q. Zhang, H. K. Lo, and J. W. Pan, "Secure quantum key distribution with realistic devices", *Rev. Mod. Phys.*, vol. 92, no. 2, p. 025 002, 2020. DOI: 10.1103/REVMODPHYS.92.025002.

[4] V. Makarov, "Controlling passively quenched single photon detectors by bright light", *New J. Phys.*, vol. 11, no. 6, p. 065 003, Jun. 2009, ISSN: 13672630. DOI: 10.1088/1367-2630/11/6/065003.

[5] L. Lydersen, C. Wiechers, C. Wittmann, D. Elser, J. Skaar, and V. Makarov, "Hacking commercial quantum cryptography systems by tailored bright illumination", vol. 4, no. 10, pp. 686–689, Oct. 2010, ISSN: 17494885. DOI: https://doi.org/10.1038/nphoton.2010.214.

[6] L. Lydersen, M. K. Akhlaghi, A. H. Majedi, J. Skaar, and V. Makarov, "Controlling a superconducting nanowire single-photon detector using tailored bright illumination", *New J. Phys.*, vol. 13, no. 11, p. 113 042, Nov. 2011, ISSN: 1367-2630. DOI: 10.1088/1367-2630/13/11/113042.

[7] G. Goltsman, M. Elezov, R. Ozhegov, and V. Makarov, "Countermeasure against bright-light attack on superconducting nanowire single-photon detector in quantum key distribution", *Opt. Express, Vol. 27, Issue 21, pp. 30979-30988*, vol. 27, no. 21, pp. 30 979–30 988, Oct. 2019, ISSN: 1094-4087. DOI: 10.1364/OE.27.030979.

[8] 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*, vol. 21, no. 3, p. 2667, Feb. 2013, ISSN: 1094-4087. DOI: 10.1364/oe.21.002667.

[9] 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.*, vol. 21, no. 3, pp. 192–196, May 2015, ISSN: 15584542. DOI: 10.1109/JSTQE.2015.2389528.

[10] 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*, vol. 6, no. 9, p. 1178, Sep. 2019, ISSN: 2334-2536. DOI: 10.1364/optica.6.001178.

[11] H.-K. Lo, M. Curty, and B. Qi, "Measurement-device-independent quantum key distribution", *Phys. Rev. Lett.*, vol. 108, p. 130 503, 13 Mar. 2012. DOI: 10.1103/PhysRevLett.108.130503.