Insecurity of practical quantum key distribution against long-wavelength Trojan-horse attacks

Shihan Sajeed,[1,2] Carter Minshull,[1,3] **Nitin Jain**,[4] and Vadim Makarov[3,1,2]
[1)] *Institute for Quantum Computing, University of Waterloo, Waterloo, ON, N2L 3G1 Canada*
[2)] *Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON, N2L 3G1 Canada*
[3)] *Department of Physics and Astronomy, University of Waterloo, Waterloo, ON, N2L 3G1 Canada*
[4)] *Department of Physics, Technical University of Denmark, Fysikvej, Kongens Lyngby 2800, Denmark*

Trojan-horse attacks on practical quantum key distribution (QKD) implementations have received considerable attention in the last 5 years[1–4]. In these attacks, the eavesdropper Eve directs a strong optical pulse from the quantum channel into the targeted QKD subsystem — Alice or Bob — and performs appropriate measurements on the back-reflections. These measurements can yield Eve information about the state of the modulator if it is in the attack path taken by the bright pulse and/or a back-reflection. If the attack can be carried out without alerting Alice or Bob, then the security of the QKD implementation is broken since knowing the state of the modulator is equivalent to knowing the secret bit.

While the basic ideas behind such attacks have been known for more than a decade[5,6], the first actual demonstration on 'Clavis2-Bob', the QKD receiver from ID Quantique (`www.idquantique.com`), was reported recently[1]. It was shown that information about the modulator's state can indeed be gleaned successfully even with back-reflected pulses containing just a few photons. Nonetheless, the overall attack failed, owing to the side effect of increased afterpulsing in the single photon detectors (SPDs) of Bob. This afterpulsing dramatically elevates the noise response of the SPDs, thereby alerting Alice and Bob.

Here we report that a Trojan-horse attack is likely to stay inconspicuous if the attacker uses bright Trojan-horse pulses at a wavelength > 1900 nm. This is primarily because the afterpulsing probability due to such bright pulses is significantly lower than that observed in the previous study[1], where bright pulses at the normal communication wavelengths (around 1550 nm) were used. Figure 1 shows the two afterpulsing profiles, experimentally measured by synchronizing a single Trojan-horse pulse to the first in a sequence of detection gates of Bob, and recording the times at which clicks occurred in the onward gates.

The benefit of reduced afterpulsing at λ_l unfortunately comes at the expense of a much higher attenuation of the Trojan-horse pulse inside Bob. Additionally, the degree of modulation received at λ_l differs from that at λ_s substantially. We quantify the increased optical attenuation and the sub-optimal modulator response by means of further experimental measurements. Taking all these factors into account as well as devising a new attack path through Bob, we evaluate the attack performances in the two wavelength regimes. By means of a numerical simulation, we conclude that a Trojan-horse attack at λ_l is likely to breach the security of the QKD system. We note that a full-fledged apparatus, though hard to build, should be mostly implementable with commercial off-the-shelf components. The attack can be mitigated by using a wavelength filter at the input of the QKD device.

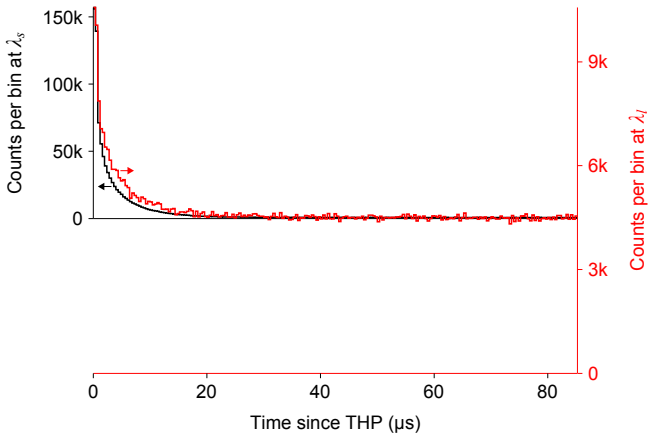

FIG. 1. Afterpulse profiles measured at $\lambda_s = 1536$ nm and $\lambda_l = 1924$ nm. For easier visual comparison, the histograms are rescaled so that their peak counts and dark count rates match in the plot.

[1] N. Jain, E. Anisimova, I. Khan, V. Makarov, C. Marquardt, and G. Leuchs, "Trojan-horse attacks threaten the security of practical quantum cryptography," New J. Phys. **16**, 123030 (2014).
[2] M. Lucamarini, I. Choi, M. B. Ward, J. F. Dynes, Z. L. Yuan, and A. J. Shields, "Practical security bounds against the Trojan-horse attack in quantum key distribution," Phys. Rev. X **5**, 031030 (2015).
[3] N. Jain, B. Stiller, I. Khan, V. Makarov, C. Marquardt, and G. Leuchs, "Risk analysis of Trojan-horse attacks on practical quantum key distribution systems," IEEE J. Sel. Top. Quantum Electron. **21**, 6600710 (2015).
[4] H.-X. Ma, W.-S. Bao, H.-W. Li, and C. Chou, "Quantum hacking of two-way continuous-variable quantum key distribution using trojan-horse attack," Chinese Physics B **25**, 080309 (2016).
[5] 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).
[6] N. Gisin, S. Fasel, B. Kraus, H. Zbinden, and G. Ribordy, "Trojan-horse attacks on quantum-key-distribution systems," Phys. Rev. A **73**, 022320 (2006).