Laser damage creates backdoors in quantum cryptography

Shihan Sajeed,[1] Sarah Kaiser,[1] Poompong Chaiwongkhot,[1] Mathieu Gagné,[2] Jean-Philippe Bourgoin,[1]
Carter Minshull,[1] Matthieu Legré,[3] Thomas Jennewein,[1,4] Raman Kashyap,[2] and Vadim Makarov[1]

[1]*Institute for Quantum Computing, University of Waterloo, Waterloo, ON, N2L 3G1 Canada*
[2]*Department of Engineering Physics and Department of Electrical Engineering,*
École Polytechnique de Montréal, Montréal, QC, H3C 3A7 Canada
[3]*ID Quantique SA, Chemin de la Marbrerie 3, 1227 Carouge, Geneva, Switzerland*
[4]*Quantum Information Science Program, Canadian Institute for Advanced Research, Toronto, ON, M5G 1Z8 Canada*

(Dated: April 29, 2016)

Quantum communication (QC) protocols, as opposed to classical cryptography, have theoretical proofs of being unconditionally secure [1–7]. Although the security is based on the assumed behaviour of implemented equipment, the actual behaviour often deviates from the modelled one, leading to a compromise of security as has been shown so far in case of quantum key distribution (QKD) [8–11]. However, it is widely assumed that as long as these deviations are properly characterized and security proofs are updated accordingly [3, 12], implementations are unconditionally secure. In this work we show that this is not always true. Even if a system is perfectly characterized and deviations are included into the security proofs, an adversary can still create a new deviation on-demand and make the system insecure.

Before going into details on how the adversary may do it, let's consider a few examples of deviations and their consequences. For example, a calibrated optical attenuator is required to set a precise value of outgoing mean photon number μ in the implementations of ordinary QKD [13, 14], decoy-state QKD [15], coherent-one-way QKD [16], measurement-device-independent QKD [17], continuous-variable QKD [18], digital signature [5], relativistic bit commitment [6], coin-tossing [19] and secret-sharing [7] protocols. An unexpected increase of this optical component's attenuation may cause a denial-of-service. However, a reduction in attenuation will increase μ, leading to a compromise of security via attacks that rely on measurement of multi-photon pulses [10]. Some implementations use a detector for time synchronization [6, 7, 13, 14, 16–19]. Desensitizing it may result in the denial-of-service. However, several implementations require a calibrated monitoring detector for security purposes [6, 7, 13, 14, 16, 18, 19]. A reduction in its sensitivity may lead to security vulnerabilities such as a Trojan-horse attack [20] that might leaks the key in QKD, increases the cheating probability in coin-tossing [10], leaks the program and client's data in quantum cloud computing [4] and allows forging of digital signatures [5]. Many implementations use beamsplitters and rely on their pre-characterized splitting ratio (e.g., [6, 13–16, 18, 19]). A shift in the splitting ratio may lead to either the denial-of-service or security vulnerabilities (e.g., [21] or one of the above mentioned attacks). A shift in characteristics of a phase modulator or a Faraday mirror may create imperfect qubits that will result in the denial-of-service or a breach in security [8, 9, 22]. If the dark count rate of single-photon detectors is increased it may lead to the denial-of-service [23]. Even device-independent QKD (DI-QKD) [24] assumes the absence of information-leakage channels and memory [25]. For example, lets assume, detectors in DI-QKD implementation emit light on detection [26], and to prevent this leaking of information, spectral filters and optical isolators are added to the devices. Then, unexpected deviations in characteristics of the these components become important for security. In summary, quantum communication systems rely on multiple characteristics of many components for their correct operation, and a deviation might lead to severe security consequences.

In classical communication, these 'deviations' are not too much of concern because the security-critical parts can be made physically separated from the communication channel making them isolated from an adversary in the channel. But the front-end of a quantum communication system is essentially an analog optical system connected to the channel. An eavesdropper may shoot a high-power laser from the communication channel to damage a security critical component such that its property is deviated from the modelled value [23]. Which component will yield first to laser damage and whether the newly created deviation will lead to a denial of service or a security vulnerability is not clear beforehand without in-depth experimental testing. Since security is the principal concern that necessitates QKD over classical cryptography, this issue cannot be left ignored and in-depth experimental testing for every QKD implementation must be performed. **This is what we have done in this paper. We choose two completely different and widely used practical quantum communication systems and check the consequences of laser damage on them. Unfortunately, from the security point of view, in both systems, laser damage altered the characteristics of security critical components in such a way that resulted in a compromise of security.**

Laser damage in fiber-optic quantum communication system. The first system is a commercial state-of-the-art fiber-optic system for running plug-and-play QKD and loss-tolerant quantum coin-tossing (QCT) protocols with phase-encoded qubits [13, 14]. Both were implemented using the commercial system Clavis2 from ID Quantique and require for security that Alice monitors the incoming pulse energy from Bob. Hence, a portion of the incoming energy is fed into a monitoring p-i-n photodiode $\mathrm{D_{pulse}}$ [10]. The sensitivity of $\mathrm{D_{pulse}}$ is factory-calibrated and an alarm is produced if the incoming energy is higher than a calibrated value. We tested the endurance of this countermeasure against laser damage. During normal system operation, we injected high power $1550\ \mathrm{nm}$

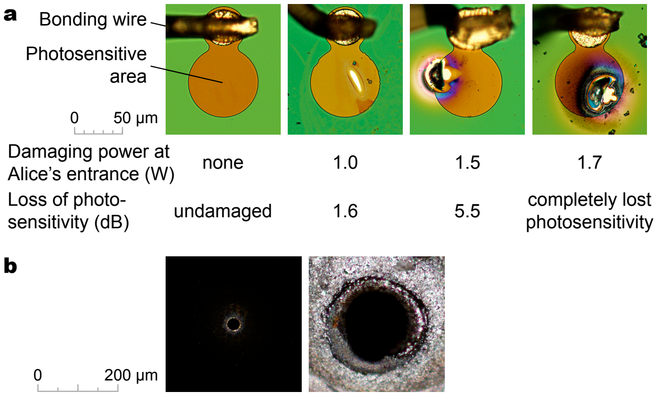
a
Bonding wire
Photosensitive area
0
50 μm
Damaging power at Alice's entrance (W)
none
1.0
1.5
1.7
Loss of photo-sensitivity (dB)
undamaged
1.6
5.5
completely lost photosensitivity
b
0
200 μm

FIG. 1. **After-effect of laser damage.** **a**, Monitoring photodiode $\mathrm{D}_{\mathrm{pulse}}$ before and after damage. The last two samples have holes melted through their photosensitive area. **b**, Spatial filter before and after damage.

laser light into Alice's entrance. The damage made the diode either less sensitive to incoming light (by 1–6 dB) or completely insensitive (after ≥ 1.7 W). The physical damage is shown in Fig. 1a. We repeated the experiment with 6 photodiode samples and found the sensitivity reduced every time. No other optical component was damaged. As modeled in [10], in the QKD protocol, Eve can eavesdrop partial or full key using today's best technology if the sensitivity of $\mathrm{D}_{\mathrm{pulse}}$ drops by more than 5.6 dB. In the QCT implementation, a sensitivity reduction by 2.6 dB increases Bob's cheating probability above a classical level, removing any quantum advantage. Laser damage thus compromises both the QKD and QCT implementations [27]. Furthermore, we performed laser damage on this system during its QKD operation with factory-supplied software, and it did not even interrupt the key generation.

Laser damage in free-space quantum communication system. The second system is a free-space QKD system with polarization-encoded qubits for long distance satellite communication [15]. It has been shown in [11] that in free-space QKD implementations, by tilting the beam going towards Bob by an angle (ϕ, θ), Eve can have control over Bob's basis choice and steal the key unnoticed. Hence, it is necessary for security to use a spatial filter or 'pinhole' to limit the tilt angle. We tested the endurance of this countermeasure against laser damage. From a distance of 26.1 m, we shot an 810 nm laser beam delivering 3.6 W c.w. power at the pinhole inside Bob's setup. The intensity there was sufficient to melt the material and enlarge the hole diameter from 25 μm to ≈ 150 μm. The physical damage is shown in Fig. 1b. Although Bob was up and running during the damage, none of his other components were affected. With the created pinhole opening, it was again possible for Eve to perform the beam-tilting attack under realistic conditions of a channel loss in 1–15 dB range with quantum bit error ratio (QBER) $< 6.6\%$. Thus laser damage completely neutralizes the spatial filter countermeasure, and makes this free-space QKD system insecure [27].

At present, no quantum communication system has countermeasures specifically designed to stop the laser-damage attack, neither do they have a mechanism to check all possible deviations in device characteristics from the modelled values. Since, before testing a system it is not clear whether this attack will result in denial-of-service only, or in a useful attack, we assume no prior knowledge of the outcome. We make a conservative assumption that at least 50 distinct quantum communication systems exist on which this attack could be attempted. The Bayesian probability that at least 20% of the untested systems are vulnerable to this attack is 98.6% (98.9%), assuming a uniform (Jeffreys) prior [28]. This high risk is in stark contrast with very low expected theoretical security risk, which for QKD is that the key is secure except with a vanishingly small probability [2, 3, 12].

We have thus practically introduced the laser damage as a new eavesdropping tool that alters parameters of a well-characterized system. Any alteration of system characteristics might compromise the security either directly by leading to an attack as we have demonstrated, or indirectly by shifting some parameter in the security proof so it would no longer apply. Currently available security proofs do not accommodate this. We expect that thorough testing against optical attacks including laser damage will become an obligatory part of security assurance for future quantum communications. **More details of this work can be found in [27].**

[1] C. H. Bennett and G. Brassard, in *Proceedings of IEEE International Conference on Computers, Systems, and Signal Processing* (IEEE Press, New York, Bangalore, India, 1984) pp. 175–179.

[2] H.-K. Lo and H. F. Chau, Science **283**, 2050 (1999).

[3] D. Gottesman, H.-K. Lo, N. Lütkenhaus, and J. Preskill, Quant. Inf. Comp. **4**, 325 (2004).

[4] S. Barz, E. Kashefi, A. Broadbent, J. F. Fitzsimons, A. Zeilinger, and P. Walther, Science **335**, 303 (2012).

[5] R. J. Collins, R. J. Donaldson, V. Dunjko, P. Wallden, P. J. Clarke, E. Andersson, J. Jeffers, and G. S. Buller, Phys. Rev. Lett. **113**, 040502 (2014).

[6] T. Lunghi, J. Kaniewski, F. Bussières, R. Houlmann, M. Tomamichel, A. Kent, N. Gisin, S. Wehner, and H. Zbinden, Phys. Rev. Lett. **111**, 180504 (2013).

[7] W. P. Grice, P. G. Evans, B. Lawrie, M. Legré, P. Lougovski, W. Ray, B. P. Williams, B. Qi, and A. M. Smith, Opt. Express **23**, 7300 (2015).

[8] F. Xu, B. Qi, and H.-K. Lo, New J. Phys. **12**, 113026 (2010).

[9] S.-H. Sun, M.-S. Jiang, and L.-M. Liang, Phys. Rev. A **83**, 062331 (2011).

[10] S. Sajeed, I. Radchenko, S. Kaiser, J.-P. Bourgoin, A. Pappa, L. Monat, M. Legré, and V. Makarov, Phys. Rev. A **91**, 032326 (2015).

[11] S. Sajeed, P. Chaiwongkhot, J.-P. Bourgoin, T. Jennewein, N. Lütkenhaus, and V. Makarov, Phys. Rev. A **91**, 062301 (2015).

[12] C.-H. F. Fung, K. Tamaki, B. Qi, H.-K. Lo, and X. Ma, Quant. Inf. Comp. **9**, 131 (2009).

[13] D. Stucki, N. Gisin, O. Guinnard, G. Ribordy, and H. Zbinden, New J. Phys. **4**, 41 (2002).

[14] Clavis2 specification sheet, `http://www.idquantique.com/images/stories/PDF/clavis2-quantum-key-distribution/clavis2-specs.pdf`, visited 20 March 2016.

[15] J.-P. Bourgoin, N. Gigov, B. L. Higgins, Z. Yan, E. Meyer-Scott, A. K. Khandani, N. Lütkenhaus, and T. Jennewein, Phys. Rev.

A **92**, 052339 (2015).

[16] N. Walenta, A. Burg, D. Caselunghe, J. Constantin, N. Gisin, O. Guinnard, R. Houlmann, P. Junod, B. Korzh, N. Kulesza, M. Legré, C. W. Lim, T. Lunghi, L. Monat, C. Portmann, M. Soucarros, R. T. Thew, P. Trinkler, G. Trolliet, F. Vannel, and H. Zbinden, New J. Phys. **16**, 013047 (2014).

[17] 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, Phys. Rev. Lett. **113**, 190501 (2014).

[18] P. Jouguet, S. Kunz-Jacques, A. Leverrier, P. Grangier, and E. Diamanti, Nat. Photonics **7**, 378 (2013).

[19] A. Pappa, P. Jouguet, T. Lawson, A. Chailloux, M. Legré, P. Trinkler, I. Kerenidis, and E. Diamanti, Nat. Commun. **5**, 3717 (2014).

[20] A. Vakhitov, V. Makarov, and D. R. Hjelme, J. Mod. Opt. **48**, 2023 (2001).

[21] H.-W. Li, S. Wang, J.-Z. Huang, W. Chen, Z.-Q. Yin, F.-Y. Li, Z. Zhou, D. Liu, Y. Zhang, G.-C. Guo, W.-S. Bao, and Z.-F. Han, Phys. Rev. A **84**, 062308 (2011).

[22] F. Xu, K. Wei, S. Sajeed, S. Kaiser, S. Sun, Z. Tang, L. Qian, V. Makarov, and H.-K. Lo, Phys. Rev. A **92**, 032305 (2015).

[23] A. N. Bugge, S. Sauge, A. M. M. Ghazali, J. Skaar, L. Lydersen, and V. Makarov, Phys. Rev. Lett. **112**, 070503 (2014).

[24] A. Acín, N. Gisin, and L. Masanes, Phys. Rev. Lett. **97**, 120405 (2006).

[25] J. Barrett, R. Colbeck, and A. Kent, Phys. Rev. Lett. **110**, 010503 (2013).

[26] P. V. P. Pinheiro *et al.,* poster presented at *QCrypt 2015, Tokyo, Japan, September 28 – October 2, 2015;* P. V. P. Pinheiro *et al.,* manuscript in preparation.

[27] V. Makarov, J.-P. Bourgoin, P. Chaiwongkhot, M. Gagne, T. Jennewein, S. Kaiser, R. Kashyap, M. Legré, C. Minshull, and S. Sajeed, arXiv:1510.03148 [quant-ph].

[28] A. Gelman, J. B. Carlin, H. S. Stern, and D. B. Rubin, *Bayesian Data Analysis*, 2nd ed. (Chapman and Hall, 2004).