Quantitative analysis of Trojan-horse attacks on practical continuous-variable quantum key distribution systems

Imran Khan[1,2], Birgit Stiller[1,2,5], Nitin Jain[1,2,6], Paul Jouguet[3,4], Sébastien Kunz-Jacques[3,4], Eleni Diamanti[3,4], Christoph Marquardt[1,2] and Gerd Leuchs[1,2,7]

1 Max Planck Institute for the Science of Light, Guenther-Scharowsky-Str. 1/Bldg. 24, 91058 Erlangen Germany
2 Institute for Optics, Information and Photonics, University of Erlangen-Nuernberg, Staudtstraße 7/B2, 91058 Erlangen, Germany
3 SeQureNet, 23 avenue d'Italie, 75013 Paris, France
4 LTCI, CNRS - Telecom ParisTech, 46 rue Barrault, 75013 Paris, France
5 Centre for Ultrahigh bandwidth Devices for Optical Systems (CUDOS), School of Physics, University of Sydney, NSW 2006, Australia
6 Center for Photonic Communication and Computing, EECS Department, Northwestern University, Evanston, Illinois 60208, USA
7 Department of Physics, University of Ottawa, 25 Templeton, Ottawa, ON, Canada

Practical quantum key distribution (QKD) implementations may deviate from the assumptions made in security proofs for QKD protocols [1, 2]. Quantum hacking uses this potential gap to demonstrate possible attacks on such systems. Previously, we experimentally demonstrated a Trojan-horse attack on a laboratory continuous-variable QKD system with a success rate of 98.73 % to read out the state of Alice's modulator for binary modulation [3- 5].

In this work, we extend our analysis into the Gaussian modulation regime, studying the performance of the attack for a similar commercial system [5, 6]. We performed a first proof of principle attack, by measuring Eve's Q-function of the Trojan-horse pulse and simulating the Q-function for Bob. In this way, we calculate the correlations for different Trojan-horse pulse round trip losses. The simulation of Bob's Q-function was computed as a function of Alice's classical modulation voltages that are used to produce the coherent states required for Gaussian modulation. In Figure 1, the correlation can be estimated visually, guided by colour coding of the different phases sent by Alice. Using the correlations, we will also discuss the impact on the final key rate and analyze how well the attack performs under different conditions.

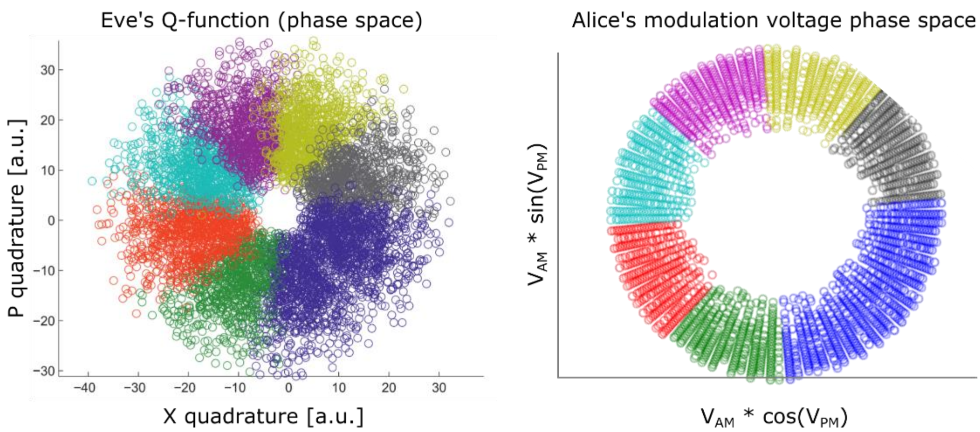

Figure 1: Eve's measured Q-function and Alice's classical modulator voltages plotted in polar coordinates; V_{AM} corresponds to the modulation voltages of the amplitude modulator, and V_{PM} to the modulation voltages of the phase modulator. The missing values in the middle of the Gaussian modulation a) can be modeled for Eve's Q-function by the finite extinction ratio of the modulator used to prepare the states in this experimental setting and b) exist in Alice's voltage phase space due to a constant voltage offset of the amplitude modulator.

[1] N. Jain, E. Anisimova, I. Khan, V. Makarov, Ch. Marquardt and G. Leuchs, New Journal of Physics **16**, 123030 (2014)
[2] N. Jain, B. Stiller, I. Khan, V. Makarov, C. Marquardt and G. Leuchs, IEEE Journal of Selected Topics in Quantum Electronics **21**, 3 (2014)
[3] I. Khan, N. Jain, B. Stiller, P. Jouguet, S. Kunz-Jacques, E. Diamanti, Ch. Marquardt and G. Leuchs, "Trojan-horse attacks on practical continuous-variable quantum key distribution systems", Conference on Quantum Cryptography (QCrypt), Paris, France, September 2014
[4] I. Khan, C. Wittmann, N. Jain, N. Killoran, N. Lütkenhaus, Ch. Marquardt and G. Leuchs, Physical Review A **88**, 010302(R) (2013)
[5] B. Stiller, I. Khan, N. Jain, P. Jouguet, S. Kunz-Jacques, E. Diamanti, C. Marquardt, and G. Leuchs, "Quantum hacking of continuous-variable quantum key distribution systems: realtime Trojan-horse attacks,", CLEO_QELS, San Jose, paper FF1A.7
[6] P. Jouguet, S. Kunz-Jacques, A. Leverrier, P. Grangier and E. Diamanti, Nature Photonics **7**, 378 (2013)