Short pulse attack on continuous-variable quantum key distribution system

Hao Qin,[1, 2, *] Anqi Huang,[1, 3] and Vadim Makarov[2, 1, 3]

[1] *Institute for Quantum Computing, University of Waterloo, Waterloo, ON, N2L 3G1 Canada*
[2] *Department of Physics and Astronomy, University of Waterloo, Waterloo, ON, N2L 3G1 Canada*
[3] *Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON, N2L 3G1 Canada*

As a coherent detection technique, homodyne detector (HD) is used in continuous-variable (CV) quantum key distribution (QKD) system for measurements, which is one of the main advantages over discrete-variable (DV) QKD using single photon detectors (SPDs) [1–3]. By using HD, CV QKD can be fully implemented with off-the-shelf telecom components [4–6]. The using of Local Oscillator (LO) in HD acts as single-mode filters, which enables CV QKD signals to be wavelength-multiplexed with intense classical channels over optical networks [7]. Moreover, unlike SPDs as vulnerable targets open for side channel attacks in DV QKD [8–10], CV QKD used to be believed robust against detector-based attack at early time. However, recently Ref [11, 12] has shown that an eavesdropper, Eve, can fully break the security of CV QKD taking advantage of saturation on HD's amplifier electronics. Although the concept of measurement device independent (MDI) is already introduced into CV QKD [13], there is still a large gap between practical implementation and theoretical proposal, there are even debates on whether MDI CV QKD can become practical regarding to its theoretical performances and current available technologies [14, 15]. Thus, it is worth studying detector based attacks in CV QKD to motivate the development of practical MDI CV QKD.

Here, we propose a new side channel attack on CV QKD implementing GG02 (Grosshans and Grangier, 2002) protocol [16] by exploiting HD's imperfections, such as the finite bandwidth of HD amplifiers and limited response time of HD electronics. In particular, we take advantage of the fact that HD's efficiency is dependent on the input pulse temporal mode [17] where Bob's HD can behave nonlinearly when Eve manipulates input pulse widths. In GG02 protocol, Alice modulates quadratures X and P of coherent states with a centered bivariate Gaussian modulation and sends them to Bob. Bob performs homodyne measurement on these coherent states and decodes them into continuous values as raw keys. In practice, in order to make sure Bob can correctly decode information and measure the shot noise (N_0), there is a trade off between electrical noise and bandwidth of HD. For this reason, most of CV QKD experiments [4–6] consist HD with only few MHz bandwidth to limit electrical noise, since Bob's HD must be shot noise limited. Meanwhile Alice needs to increase the pulse duration (typically 100 ns) and reduce repetition rate (1 MHz) to meet Bob's HD bandwidth requirement [5]. However, HD bandwidth will reduce the HD output efficiency significantly if it is smaller than the inverse temporal width of the signal temporal mode [17]. Such effects can be obvious when HD bandwidth is relatively small, which gives more space to a potential Eve to manipulate HD efficiency. The response time of the electronics is typically not faster than a few ns which means if the input pulse width is less than few ns, the HD efficiency also becomes very poor. In order to illustrate such effects, we perform simple experiments in which we vary width of optical pulses from 1 ns to 100 ns at 1550nm and send them to a classical optical photodiode (PD) detector with 3.5 GHz bandwidth and one of the port (PD1 or PD2) of our HD with 100 MHz bandwidth. For each measurement, we record the maximum amplitude value of PD output during the pulse duration as our measurement results (which is similar to sampling stage of CV QKD [5, 18]). As shown in Fig.1 when the pulse duration is longer than 4 ns there is no obvious degrading effect on the output, however when pulse duration becomes shorter, PD's outputs decrease in both case with 100 MHz at about 3 ns and 3.5 GHz bandwidth at about 1 ns. In order to compare the two cases, we normalize all the values by the amplitudes measurement with pulse width of 8 ns. Such observations confirm the predictions on the relation between HD efficiency and pulse width.

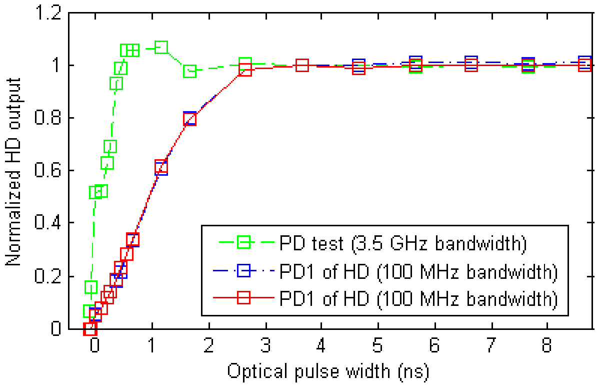

FIG. 1: Short pulse demonstration on classical PD with 3.5 GHz bandwidth and HD with 100 MHz. Each square corresponds to a measurement on the maximum amplitude value of PD (green), PD1 (blue) and PD2 (red) outputs.

By using such effects, Eve can thus manipulate Bob's HD efficiency by changing input pulse widths. If Bob's HD efficiency for certain parts of signal pulses becomes lower, then the linearity between all input quadratures

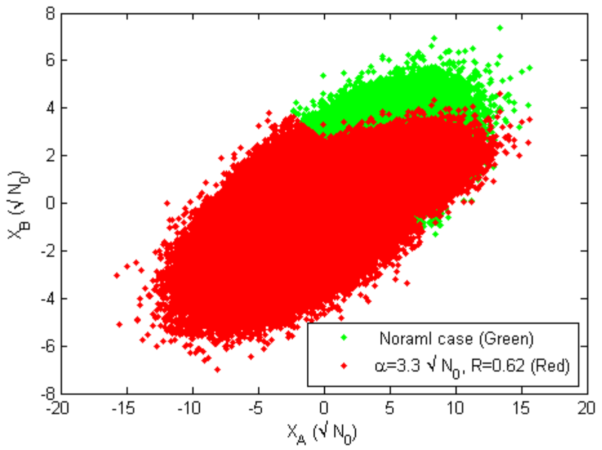

FIG. 2: Alice and Bob data distribution under the short pulse attack. Alice variance $10N_0$, Alice and Bob distance 40 km, fiber attenuation coefficient 0.2 dB/km, reconciliation efficiency 94%, Bob's HD efficiency 60%, electrical noise $0.01N_0$. Alice-Bob excess noise estimation based on green data $2.1N_0$, on red data $0.0084N_0$.

and Bob's HD outputs will not hold, which breaks the important linearity assumption in security proof of CV QKD. Assuming Alice and Bob implement GG02 CV QKD as in [5], we propose Eve's attack strategy as following: (1) Eve fully characterize Bob's HD, particularly, Eve builds the relationship between input pulse width and HD's output efficiency as a reference. As illustrated in Fig. 1, such relationship is determined by the HD's amplifier bandwidth and its electronics. (2) Eve cuts the quantum channel and measures the quadratures X and P sent by Alice with the help of a heterodyne detection [19, 20]. (3) According to her measurement results, Eve prepares corresponding signal pulses as in the intercept-resend (IR) attack [19, 20] with pulse width of 100 ns. Such"entanglement breaking" action will normally rise Alice and Bob's excess noise estimation to at least two units of shot noise ($2N_0$), which includes the vacuum noise of heterodyne detection and the vacuum noise in the new coherent states preparation. (4) Eve adjusts signal pulse widths according to following rules: 4.a) Eve sets a manipulating level ($\alpha > 0$) on her measurements. 4.b) For any Eve's measurement larger than α, she reduces the pulse width of corresponding re-prepared signal such that HD output efficiency reduces to a certain level, which relates to a efficiency reduction ratio R. The relation between HD efficiency and pulse width is determined in step (1). 4.c) For all the rest of resent signal pulses, Eve maintains their widths as Alice's pulse width (100 ns). 4.d) Eve sends all of these re-prepared signal pulses to Bob. (5) Bob performs HD measurements on Eve's resent pulses; Due to different pulse widths adjusted by Eve, Bob's HD output efficiency is not identical respect to different pulses widths (red dots in Fig. 2). (6) Alice and Bob then estimate excess noise on corrupted data which under certain conditions can lead them to underestimate the excess noise due to IR attack. Under such strategy, if the excess noise estimation can be biased below the null key threshold (collective attack [21]), then Eve's IR action won't be spotted by Alice and Bob, which fully breaks the security. We have confirmed this security break in our simulation as shown in Fig. 2, where red data corresponds to the mentioned strategy, Alice and Bob estimate excess noise as $0.0084N_0$ which is still under null key threshold ($0.091N_0$ with simulation parameters shown in Fig. 2). Overall, in our strategy Eve first performs a modified IR attack. By manipulating certain parts of resent signal pulse widths, Eve can force Bob's HD response to be non-linear, which violates the basic assumption of linear detection. Furthermore, Eve can set two target levels $\alpha_1 > 0$ and $\alpha_2 < 0$ to have more freedoms to influence Alice and Bob's data to achieve more powerful attack.

Regarding countermeasures, such attack can be prevented by MDI CV QKD [13, 22, 23]. However previous countermeasures against saturation attack may not be effective [11, 24, 25]. Since in this short pulse attack, Eve actually exploits nonlinear response of Bob's HD in the linear region that is characterized by Alice and Bob, the countermeasures of saturation attack only detect any actions that are happened beyond detection limits, which will not be enough to detect Eve's action in this new attack. On the other hand, the progress of CV QKD security proof includes additional steps such as symmetric test on Alice and Bob data [26], which may eventually cover such kind of attack. Above all, we propose a practical side channel attack targeting HD finite bandwidth and limited speed of electronics. We further propose our attack strategy and demonstrate in simulations that our attack can break the security of current GG02 CV QKD implementations.

[*] Electronic address: `hao.qin@uwaterloo.ca`

[1] Weedbrook, C., Pirandola, S., García-Patrón, R., Cerf, N. J., Ralph, T. C., Shapiro, J. H., and Lloyd, S. *Rev. Mod. Phys.* **84**, 621–669 May (2012).

[2] Scarani, V., Bechmann-Pasquinucci, H., Cerf, N. J., Dušek, M., Lütkenhaus, N., and Peev, M. *Rev. Mod. Phys.* **81**, 1301–1350 Sep (2009).

[3] Diamanti, E. and Leverrier, A. *Entropy* **17**(9), 6072–6092 Auguest (2015).

[4] Lodewyck, J., Bloch, M., García-Patrón, R., Fossier, S., Karpov, E., Diamanti, E., Debuisschert, T., Cerf, N. J., Tualle-Brouri, R., McLaughlin, S. W., and Grangier, P. *Phys. Rev. A* **76**, 042305 Oct (2007).

[5] Jouguet, P., Kunz-Jacques, S., Leverrier, A., Grangier, P., and Diamanti, E. *Nat Photon* **7**(5), 378–381 May (2013).

[6] Fossier, S., Diamanti, E., Debuisschert, T., Villing, A.,

Tualle-Brouri, R., and Grangier, P. *New Journal of Physics* **11**(4), 045023 (2009).

[7] Kumar, R., Qin, H., and Allaume, R. *New Journal of Physics* **17**(4), 043027– (2015).

[8] Lydersen, L., Wiechers, C., Wittmann, C., Elser, D., Skaar, J., and Makarov, V. *Nat Photon* **4**(10), 686–689 October (2010).

[9] Gerhardt, I., Liu, Q., Lamas-Linares, A., Skaar, J., Kurtsiefer, C., and Makarov, V. *Nat Commun* **2**, 349– June (2011).

[10] Wiechers, C., Lydersen, L., Wittmann, C., Elser, D., Skaar, J., Marquardt, C., Makarov, V., and Leuchs, G. *New Journal of Physics* **13**(1), 013043– (2011).

[11] Qin, H., Kumar, R., and Alléaume, R. *Phys. Rev. A* **94**, 012325 Jul (2016).

[12] Qin, H., Kumar, R., and Alleaume, R. In *Proc. SPIE 9648, Electro-Optical and Infrared Systems: Technology and Applications XII; and Quantum Information Science and Technology*, volume 9648, 9648V–11, (2015).

[13] Pirandola, S., Ottaviani, C., Spedalieri, G., Weedbrook, C., Braunstein, S. L., Lloyd, S., Gehring, T., Jacobsen, C. S., and Andersen, U. L. *Nat Photon* **9**(6), 397–402 June (2015).

[14] Xu, F., Curty, M., Qi, B., Qian, L., and Lo, H.-K. *Nat Photon* **9**(12), 772–773 December (2015).

[15] Pirandola, S., Ottaviani, C., Spedalieri, G., Weedbrook, C., Braunstein, S. L., Lloyd, S., Gehring, T., Jacobsen, C. S., and Andersen, U. L. *Nat Photon* **9**(12), 773–775 December (2015).

[16] Grosshans, F., Van Assche, G., Wenger, J., Brouri, R., Cerf, N. J., and Grangier, P. *Nature* **421**(6920), 238–241 January (2003).

[17] Kumar, R., Barrios, E., MacRae, A., Cairns, E., Huntington, E., and Lvovsky, A. *Optics Communications* **285**(24), 5259–5267 November (2012).

[18] Li, H., Wang, C., Huang, P., Huang, D., Wang, T., and Zeng, G. *Opt. Express* **24**(18), 20481–20493 September (2016).

[19] Lodewyck, J., Debuisschert, T., García-Patrón, R., Tualle-Brouri, R., Cerf, N. J., and Grangier, P. *Phys. Rev. Lett.* **98**, 030503 Jan (2007).

[20] Cerf, N. J. and Grangier, P. *J. Opt. Soc. Am. B* **24**(2), 324–334 (2007).

[21] García-Patrón, R. and Cerf, N. J. *Phys. Rev. Lett.* **97**, 190503 Nov (2006).

[22] Li, Z., Zhang, Y.-C., Xu, F., Peng, X., and Guo, H. *Phys. Rev. A* **89**, 052301 May (2014).

[23] Ma, X.-C., Sun, S.-H., Jiang, M.-S., Gui, M., and Liang, L.-M. *Phys. Rev. A* **89**, 042335 Apr (2014).

[24] Zhengyu, L., Yichen, Z., Christian, W., and Hong, G. August (2016). Poster at QCrypt 2016.

[25] Huang, P., Huang, J., Wang, T., Li, H., Huang, D., and Zeng, G. *Phys. Rev. A* **95**(5), 052302 May (2017).

[26] Leverrier, A. *Phys. Rev. Lett.* **118**, 200501 May (2017).