Saturation Attack on Continuous-Variable Quantum Key Distribution System

Hao QIN,[1] Rupesh KUMAR,[1] and Romain ALLEAUME[1]

[1]*Telecom ParisTech, Laboratoire Traitement et Communication de l'Information, Centre National de la Recherche Scientifique, 46 Rue Barrault, 75634 Paris Cedex 13, France*

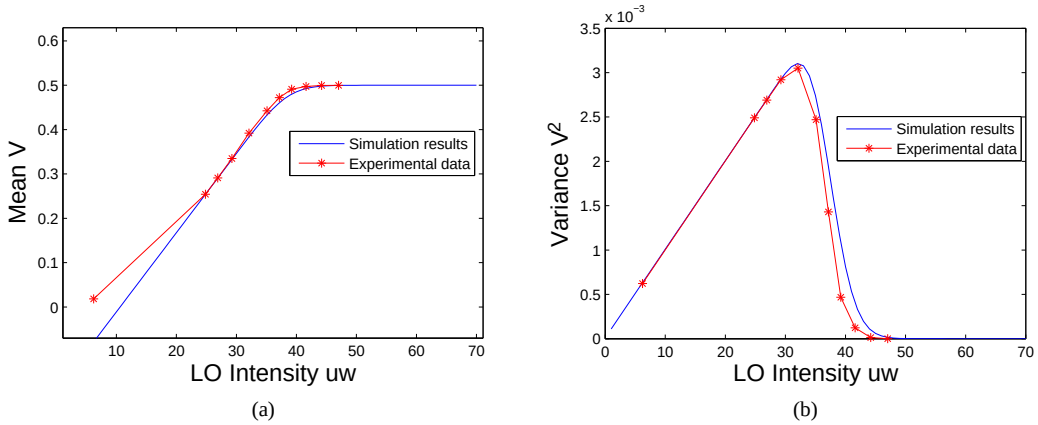
Introduction Quantum key distribution (QKD) [1] enables two remote parties Alice and Bob to share common secure keys which are unknown to a potential eavesdropper. Unconditional security of QKD is based on the fundamental laws of quantum mechanics, but in reality, securities of practical QKD systems could be jeopardized by physical implementations. In discrete-variable (DV) QKD system, due to devices imperfections, various quantum hacking strategies have been proposed and some of them are demonstrated in experiments [2–4]. Most of the practical attacks that have been demonstrated up to now are targeting the detection part of the QKD systems.

Continuous-variable (CV) QKD, as another approach, is proven secure against collective attacks and recent works have shown progress in proving its security against arbitrary attacks [5]. However, practical CV QKD systems also face the security problems linked to imperfect implementations. The validity of security proofs relies on assumptions that may be violated in practical setup, opening loopholes that may be exploited by Eve to mount attacks. For example direct [6] or indirect [7] manipulation of local oscillator (LO) intensity can fully compromise the security. This imposes to monitor LO intensity and to use filters to forbid wavelength-dependent LO intensity manipulations.

In this work, we have identified a new loophole and shown that it can be used to attack a practical CV QKD system implementing Gaussian-modulated coherent state (GMCS) protocol [8]. Instead of attacking LO, we aim at the homodyne detection located on Bob side, specifically, the electronics of the homodyne detection. We propose an attack consisting in a full intercept-resend attack [9] combined with the exploitation of the nonlinear response of homodyne detection, namely saturation attack. Under this saturation attack, we can show that Eve can manipulate the measurement results on Bob's side and get information without being discovered. Importantly, our attack is practical that can be realistically launched against existing implementations.

Saturation of homodyne detection A fundamental assumption in the security proof of CV QKD is that the response of homodyne detection is linear with respect to input field quadrature. This assumption is necessary because parameter estimation implicitly assumes the linearity of Bob quadrature measurement with respect to the value sent by Alice. However, this assumption does not hold if Bob's homodyne detection saturates. A practical homodyne detector only works normally over a limited range. Saturation typically occurs when the input field quadrature overpasses a threshold. This threshold depends on parameters of detector's electronics, such as the amplifiers linearity domains or the data acquisition card range. The important point is that since detection range cannot be infinity, saturation can always be induced by displacing the field quadratures strongly enough.

We have experimentally confirmed this prediction by observing saturation of our homodyne detection for high LO intensity. We have measured the variances and means of the homodyne output for different LO intensities. When homodyne detection is not saturated, homodyne detection outputs (mean value and variance) vary linearly with respect to LO intensity. However



(a) (b)

FIG. 1: Shot noise measurements of homodyne detection (a) Mean vs LO Intensity. (b) Shot noise variance vs LO Intensity.

when LO intensity is relatively high, the response of homodyne detection will overpass the saturation threshold. The response of homodyne detection is then saturated and the measured variance will drop quickly (Fig.1 (b)).

We have proposed a saturation model with predefined upper and lower bounds of homodyne detection response. For values between the two bounds, the response of homodyne detection behaves normally, otherwise the response is constant. This model is applied to our shot noise measurements. The simulation results match very well with our experimental data (Fig.1 (a)(b)). It indicates that our proposed saturation model is realistic and can be further used to interpret our saturation attack.

Attack strategy An intercept-resend adds extra excess noise and it will be noticed by Alice and Bob in their measurements. However, the saturation of homodyne detection could be taken advantage of by Eve to manipulate Bob's measurement results. We show that, by saturating on homodyne detection, Eve can further reduce the value of the excess noise and the channel transmission estimated by Alice and Bob. Since LO monitoring is performed in most of practical CV QKD setups, we should assume Eve can't saturate the homodyne detection by increasing LO intensity. However, Eve can still strongly displace the mean of quadratures to force the homodyne detection to work in a saturated region. As a matter of fact, quadrature mean value is not used in CV QKD security model and thus not monitored. Our saturation attack strategy is then simple:

- Eve implements a full intercept-resend attack [9] with the help of a heterodyne detection, she can learn information of both quadratures X and P sent by Alice.
- Eve then resends a coherent state whose quadratures consist in her measurement results combined with an appropriate displacement of the quadratures.
- Alice and Bob will estimate their key rate with a saturated homodyne detection, where excess noise is actually controlled by Eve. They will thus underestimate excess noise introduced by full intercept-resend attack and Eve's attack can remain undercover while giving her advantage over Alice and Bob.

Analysis A full intercept-resend attack will add up to two shot noise units of excess noise [9] at Alice side which will reveal the presence of Eve. However, Eve can control the displacement mean value of quadratures which she then sends to Bob. She can thus introduce saturation of the homodyne detection as much as she wants by changing the displaced value. As a consequence, Eve can reduce the two shot noise units of excess noise on the Alice-Bob channel to a arbitrary low value of excess noise estimated by Alice and Bob. This attack can of course affect the amount of information between Alice-Bob and Bob-Eve. Thus the attack will influence the key rate. But our simulation results indicate that a successful attack is possible over a large rang of distances. Under such attack, Alice and Bob may be led to believe they have positive 'secure key' rate and accept keys that are, however, totally insecure. It shows that Eve can successfully steal information without being detected.

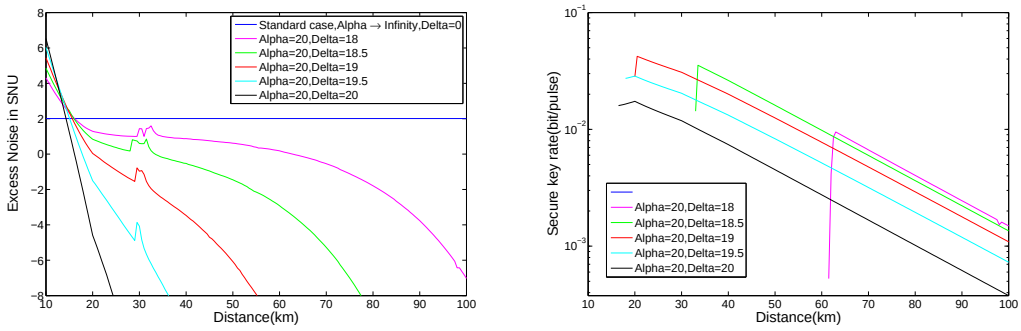
To prevent such attack, Bob should monitor the displacement value of measured quadratures to avoid the homodyne detection working in a nonlinear or saturated region. Precisely, Bob needs to make sure the mean value measured at his side is much smaller than the saturation limit. Statistical study of our counter measure is under development. We are also working on its integration into the security model, so that practical implementations could effectively protect themselves against saturation-based attacks.

Simulation results In order to simulate excess noise and key rates under our saturation attack, we follow the procedure described in Ref [11]. To achieve a high reconciliation efficiency (95%), optimal error correction codes need to work with a fixed signal to noise ratio (SNR). So Alice must optimize her modulation variance with respect to the distance in order to work at a given SNR. We have assumed that Alice variance is determined according to this procedure. We also follow the parameter estimations in Ref [12]: With the correlated variables x and y between Alice and Bob, they compute three terms of the covariance matrix: variance of x, variance of y and $\langle xy \rangle$. With an additional measurement of shot noise (variance of y where the signal port of the detection is closed), Alice and Bob can compute their covariance matrix and thus evaluate their key rates (We have assumed the collective attacks [10]). A fundamental assumption behind this parameter estimation is that the channel between Alice and Bob is linear with additive Gaussian noise:

$$y = tx + z \tag{1}$$

In which, $t = \sqrt{\eta T}$, T is channel transmission and η is efficiency of Bob. Alice modulation is Gaussian so that x is a Gaussian random variable centered on zero with variance V_A. z is the total noise which follows a centered normal distribution with unknown variance. This variance includes shot noise, excess noise and electronic noise of Bob. However, the linear model (equation (1)) doesn't hold under saturation of homodyne detection. We must replace it with the realistic saturation model that has been validated in our shot noise measurement to describe the response of homodyne detection:

$$\begin{aligned} y &= \alpha, && tx + z + \Delta \geqslant \alpha \\ y &= tx + z + \Delta, && \mid tx + z + \Delta \mid < \alpha \\ y &= -\alpha, && tx + z + \Delta \leqslant -\alpha \end{aligned} \tag{2}$$



(a) (b)

FIG. 2: (a) Excess noise in shot noise units (SNU) on Alice side with different Δ. (b) Estimated key rate (collective attack [10]) with different Δ. Alice's variance $V_A \in \{1, 100\}$, efficiency of Bob $\eta = 0.6$, excess noise of electronics $v_{\text{ele}} = 0.01$, excess noise of system $\xi_{sys} = 0.01$, reconciliation efficiency $\beta = 0.95$, attenuation coefficient $a = 0.21dB/km$.

The saturation limit of the homodyne detection response, α is intrinsic to the detector. The value of α should be chosen large enough when the system is designed. However, no matter how large the working range of homodyne detector is, α cannot go to infinity. For a fixed unknown α, Eve can always displace field quadratures to saturate the homodyne detection. We define Δ as a displacement value which can be introduced by Eve.

In Fig.2 (a), when the homodyne detection is not saturated ($\alpha \to$ Infinity, $\Delta = 0$), the total estimated excess noise coincides with the estimation in the linear model under full intercept-resend attack, which is 2.01 in shot noise units including 0.01 system noise. Under such an excess noise, the presence of Eve can be spotted by Alice and Bob, and the secure key rate is null. Therefore we don't show this curve in Fig.2 (b).

Under realistic experimental conditions concerning the homodyne detection (α large but not infinite), Eve can perform a intercept-resend attack in combination with saturation of homodyne detection: she resends displaced quadratures (she manipulates Δ) to saturate homodyne detection. When Δ value is close to α, we can see in Fig.2 (a) that the estimated excess noise is significantly reduced at long distance. If the excess noise becomes negative, Eve can always make extra noise to realize a reasonable noise (We set it to 0.01 to calculate key rates in Fig.2 (b)). However at short distance (typically below 20 km), the estimated excess noise is significantly bigger than zero whatever the value of Δ is. In this case, our proposed saturation attack cannot fool Alice and Bob efficiently. Nevertheless, in Fig.2 (b), for longer distance and when Δ becomes close to α, positive key rate can be obtained and therefore an attack can be mounted. For example when $\Delta = 20$, positive key rate can be observed from 17 km to 100 km. In conclusion, Eve can successfully use a saturation attack to fully break the security of the system (no secure key exist, but Alice and Bob instead accept a key that Eve can fully recover) for distances above 17 km. Finding more efficient attack strategy is moreover possible but remains the subject of future studies.

Our saturation attack is achievable with current technology and impacts the security of a practical CV QKD system. It highlights the importance of exploring the assumptions in security proofs when implementing QKD protocol on practical setups. Suitable counter measures are necessary for practical CV QKD to fix the loopholes that attackers can exploit.

[1] V. Scarani, H. Bechmann-Pasquinucci, N. J. Cerf, M. Dušek, N. Lutkenhaus, and M. Peev,*Rev. Mod. Phys.* **81**, 1301 (2009).
[2] F. -H. Xu, B. Qi, and H. -K. Lo, *New J. Phys.*, **12**, 113026 (2010).
[3] L. Lydersen, C. Wiechers, C. Wittmann, D. Elser, J. Skaar, and V. Makarov, *Nature Photonics.* **4**, 686-689 (2010).
[4] I. Gerhardt, Q. Liu, A. Lamas-Linares, J. Skaar, C. Kurtsiefer, and V. Makarov, *Nature Comm.* **2**, 349 (2011).
[5] A. Leverrier, R. García-Patrón, R. Renner, and N. J. Cerf, *Phys. Rev. Lett.* **110**, 030502 (2013).
[6] H. Häseler, T. Moroder and N. Lütkenhaus, *Phy. Rev. A.* **77**, 032303 (2008).
[7] J. -Z. Huang, C. Weedbrook, Z. -Q. Yin, S. Wang, H. -W. Li, W. Chen, G. -C. Guo,and Z. -F. Han, arXiv:1302.0090v1 (2013).
[8] F. Grosshans, G. van Assche, J. Wenger, R. Brouri, N. J. Cerf, and P. Grangier, *Nature*. **421**, 238 (2003).
[9] J. Lodewyck, T. Debuisschert, R. Garcia-Patron, R. Tualle-Brouri, N J. Cerf, and P. Grangier, *Phys. Rev. Lett.* **98**, 030503 (2007).
[10] J. Lodewyck, M. Bloch, R. Garcia-Patron, S. Fossier, E. Karpov, E. Diamanti, T. Debuisschert, N J. Cerf, R. Tualle-Brouri, S. McLaughlin and P. Grangier, *Phys. Rev. A.* **76**, 042305 (2007).
[11] P. Jouguet, S. Kunz-Jacques, and A. Leverrier, *Phy. Rev. A.* **84**, 062317 (2011).
[12] P. Jouguet, S. Kunz-Jacques, E. Diamanti, and A. Leverrier, *Phy. Rev. A.* **86**, 032309 (2012).