Saturation attack on Continuous-Variable QKD systems: experimental demonstration, performance analysis and countermeasure

Rupesh Kumar,[1] Hao Qin,[1] and Romain Alléaume[1]
[1] *Institut Mines Telecom / Telecom ParisTech, CNRS LTCI, 46 rue Barrault, 75634 Paris Cedex 13, France*

Introduction –

Quantum communications with phase-sensitive detectors, such as homodyne or heterodyne detectors, constitute a promising technological platform for quantum information, exemplified by continuous variable quantum key distribution (CV-QKD). From a technology viewpoint, CV quantum communications systems meet several of the key requirements related to the uptake of future quantum optical communication networks. Firstly, CV-quantum communication systems can be entirely built with telecom-grade components, which defines a clear path towards photonics integration. Moreover, despite the detrimental effect of loss on CV signal, reverse reconciliation combined with efficient post-processing techniques allow secure key distribution over relatively large distances, in the 100 km range [1]. Third, thanks to the strong spectral selectivity associated with the coherent detection, CV quantum communication system exhibit, without any additional filtering, superior coexistence compatibility with wavelength-multiplexed classical channels, allowing deployment on DWDM networks [2].

Implementation security has become a major concern in practical QKD, where prepare and measure quantum communication systems are being used. The security of practical QKD can be jeopardized if some security assumptions are not fulfilled by Alice, respectively Bob stations. In the case of discrete-variable QKD systems, attacks on detectors have been extensively used, in particular for blinding [3] or time-shift attacks, while Trojan-horse attacks or passive side-channels such as signal distinguishability can also affect Alice's station. The practical security of CV-QKD has so far been less studied. In particular, most of the attention has been put on attacks targeting the calibration of the local oscillator intensity (and thus the shot noise) using time- or wavelength-shifting [4, 5]. These attacks can all be closed by a change of design: having a local local oscillator, which has recently been demonstrated experimentally [6]. Trojan horse attacks have been demonstrated on a CV-QKD system with low input losses but would be ruled out in realistic systems due to input losses.

We demonstrate here an attack of a different type, applicable to a large range of quantum communication protocols relying on a quantum limited coherent detection, and in particular CV-QKD: the saturation attack. CV-QKD parameter estimation is based on quadrature measurements with a coherent detector such as a homodyne detection. The linearity of the quadrature measurement happens to be a crucial assumption in security proofs, but we demonstrate experimentally that a coherent detector can be driven out of its linear regime, opening the way to a powerful attack on CV quantum communication links.

Contributions –

- We have shown that a homodyne detection can be induced to work in a region where the output voltage is not linear with the optical quadrature input. We have experimentally demonstrated two active attack strategies to induce saturation of the homodyne detection in a CV-QKD set-up: 1) By coherently displacing the signal received by Bob ; 2) By shining a powerful laser, incoherent with the local oscillator, in order to induce a DC component in the measured quadrature signal.
- We have proposed and studied theoretically a comprehensive attack strategy exploiting the saturation effect. It combines the intercept-resend attack, (intercept performed at the output Alice, and resend at the input of Bob) with a strategy to induce saturation.
- We have experimentally realized a functional "Eve", capable of actively performing a controlled displacement of the quadratures in order to induce saturation of the homodyne detection, cf Fig. 1.
- We have experimentally studied the relation between Eve parameters (gain and displacement) and Alice-Bob channel parameter estimation results. Considering two possible criteria for a successful saturation attack, we have determined the initial parameters (variance V_A and channel transmission T) for which a saturation attack can be experimentally realized, with our experimental Eve, cf Fig. 2, Right.
- We have proposed an "algorithmic" counter-measure against the saturation attack, based on Gaussian post-selection [7]. This procedure can be used to guarantee that the detector is operated in a linear regime, with a post-selected gaussian input, in a region where security proof holds.

Modeling the saturation attack –

We consider the attack scheme relying on coherent displacement, depicted on Fig.1. Alice prepares a coherent state with gaussian modulation, of variance, V_A on both quadrature (from now on we can consider only what happens to the X quadrature, by symmetry). Eve in the middle cuts down the quantum channel and places a station close to Alice in order to perform quadrature measurements with a heterodyne detection. This adds an equivalent of two units of shot noise on her quadrature measurement. She then sends her (classical) measurement information X_M to another station close to Bob, with $X_M = \frac{1}{\sqrt{2}}(X_A + X_0 + X_0')$. Eve applies an amplification g to this quadrature information X_M and prepares the corresponding coherent state, that she displaces (in order to induce saturation) before sending it to Bob $X_E = g\,X_M + X_0'' + \Delta$. Bob then performs an homodyne detection, with a detector this is assumed to be perfectly linear in the quadrature range $\{-\alpha, +\alpha\}$, and saturated beyond.
Taking into account the technical noise X_N and the electronic noise X_{ele}, and the non-linearity of the homodyne detection, the measured value $X_{B_{sat}}$ follows the following equation:

$$X_{B_{sat}} \rightarrow \begin{cases} X_{B_{sat}} = \sqrt{\eta}(X_E + X_N) + \sqrt{1-\eta}X_0''' + X_{ele} \text{ if } |X_B| < \alpha \\ X_{B_{sat}} = \pm\alpha, \text{ if } |X_B| > \alpha \end{cases} \tag{1}$$

We have confronted this model to an experimental characterization of the detection, and the homodyne electronics indeed shows a sharp saturation behavior. From this model, we can evaluate the different terms of the covariance matrix $Cov(X_A, X_{B_{sat}})$ and in particular T_{sat} and ξ_{sat}, the estimation of the channel transmittance and of the excess noise (both affected by the non-linearity when the value of Δ induced by Eve is close to α).

Experimental Setup –

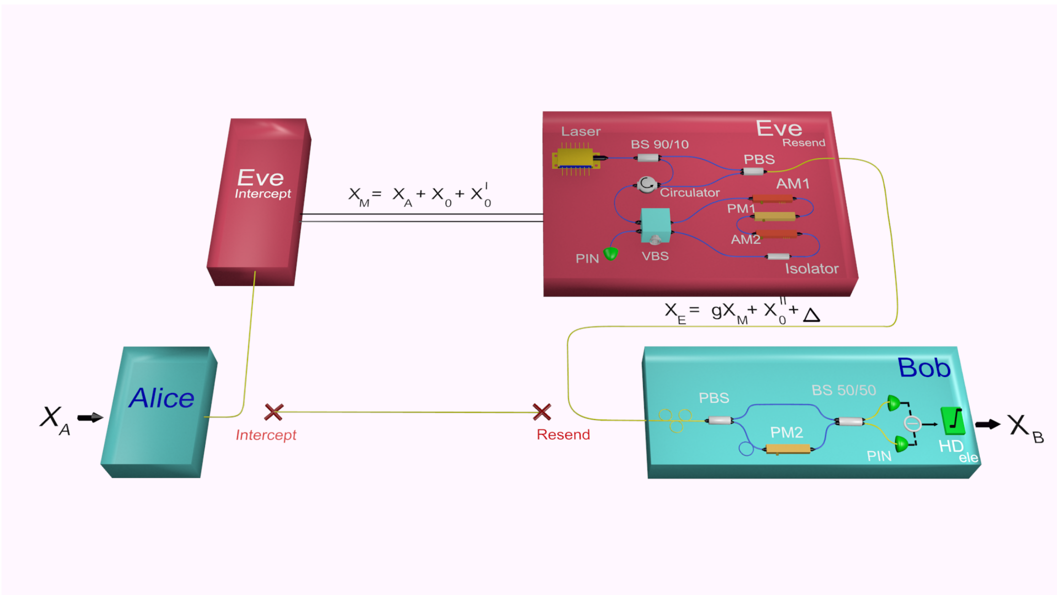

Figure 1: Experimental setup: saturation attack by coherent displacement on a CV-QKD system

In order to be able to induce a controlled displacement on Eve resent data, we have modified the "heritage" CV-QKD Alice system by introducing a Sagnac loop combined with variable beamsplitter (VBS). Displacing the signal is achieved as follows. The VBS, with splitting ratio 99.9%, splits the pulse from the circulator into two. The signal pulse, which is the less intense pulse along the clockwise direction, undergoes modulation by amplitude modulator (AM1) and phase modulator(PM1) and is further heavily attenuated by an isolator (connected in reverse to achieve an attenuation higher than 30dB). The higher intensity pulse travelling along the anti-clockwise direction, named as pump pulse, meets the signal pulse at the VBS. The interference on this strongly unbalanced beamsplitter effectively displaces the signal pulse. The amplitude modulator AM2 controls the intensity of the pump and hence the value of the displacement. A PIN diode attached to the VBS is used for monitoring. Finally, the circulator directs the displaced signal towards the polarization beam splitter (PBS) to perform polarization multiplexing of the local oscillator the displaced signal, that is then sent to output fibre channel.

Analysis of experimental results, perspectives –

For a given input variance V_A, we have studied numerically and experimentally under which conditions (gain g and displacement Δ induced by Eve) an attack can be launched. The criteria for a successful attack corresponds to the situation where Alice and Bob parameter estimation, biased by the attack, lead them to a positive key rate, although an intercept-resend (and thus entanglement-breaking) attack has been performed by Eve. One can demonstrate that provided the variance of $X_{B_{sat}}$ is large enough, increasing the displacement value close to α always leads the estimated excess noise to fall below the null key threshold, as it can be seen on Fig.2 Left. From the curve Fig.2 Right, we can see that this opens the possibility of an attack: a positive key rate can be obtained, for $V_A = 5$ for well controlled value of g and Δ, that are within experimental reach. Large displacement values however couple with the phase noise, leading to impose severe requirements regarding the precision with which Δ should be set. This required precision Δ is around $10^{-3}\ \sqrt{N_0}$, which represents a challenge with our experimental system. However, a second saturation attack, relying on a laser incoherent with the modulated signal can lift this issue. Finally we plan to demonstrate the efficiency of the countermeasure by implementing it experimentally, against an active attacker.

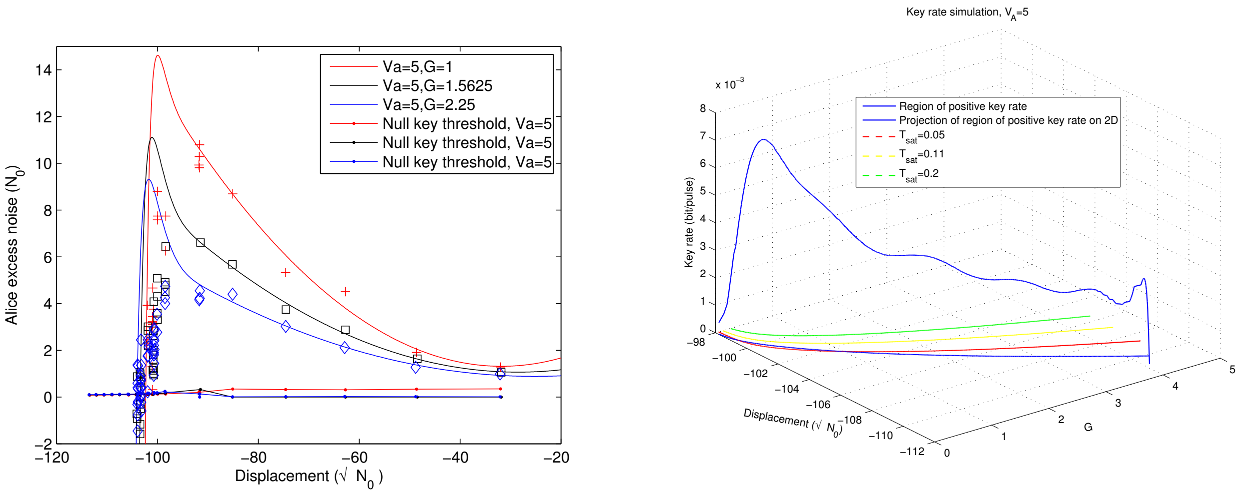

Figure 2: Left: Measured excess noise versus induced displacement Δ (experimental data and theoretical prediction) ; Right: Evaluated key rate versus Δ and g (simulation), positive key rate indicates that an effective attack can be launched when Eve carefully sets Δ and g. If Alice and Bob monitor the value of the estimated transmission and check that $T_{sat} = T$, the attack is only possible for some values of T (for example $T = 0.05$ is a possible value here).

[1] P. Jouguet, S. Kunz-Jacques, A. Leverrier, P. Grangier, E. Diamanti, Experimental demonstration of long-distance continuous-variable quantum key distribution Nature Photonics 7, 378 (2013).

[2] R. Kumar, H. Qin, R. Alléaume, Coexistence of continuous variable QKD with intense DWDM classical channels, New J. Phys. 17, 043027 (2015).

[3] I. Gerhardt, Q. Liu, A. Lamas-Linares, J. Skaar, C. Kurtsiefer, and V. Makarov, Full-field implementation of a perfect eavesdropper on a quantum cryptography system, Nat. Commun. 2, 349 (2011).

[4] P. Jouguet, S. Kunz-Jacques, E Diamanti, Preventing calibration attacks on the local oscillator in continuous-variable quantum key distribution Phys. Rev. A 87, 062313 (2013).

[5] X.-C. Ma, S.-H. Sun, M.-S. Jiang, L.-M. Liang, Wavelength attack on practical continuous-variable quantum-key-distribution system with a heterodyne protocol Phys. Rev. A 87, 052309 (2013).

[6] B. Qi, P. Lougovski, R. Pooser, W. Grice, M. Bobrek, Generating the local oscillator "locally" in continuous-variable quantum key distribution based on coherent detection arXiv:1503.00662 (2015).

[7] J. Fiurasek, N. J. Cerf, Virtual noiseless amplification and Gaussian post-selection in continuous-variable quantum key distribution, Phys. Rev. A 86, 060302(R) (2012)