We study a practical method to monitor in real-time the behavior of a Continuous Variable Quantum Key Distribution (CVQKD) system, and especially the shot noise measurement procedure, for the presence of an eavesdropper. The amount of secret key that can be extracted from the raw statistics of a CVQKD system depends strongly on a shot noise estimate since this quantity affects in particular the computation of the excess noise (i.e. noise in excess of the shot noise) added by an eavesdropper on the quantum channel. Some powerful quantum hacking attacks rely on faking the estimated value of the shot noise to hide an intercept-and-resend eavesdropping strategy [11], notably by targeting the calibration of the system [2], the saturation of the homodyne detection device [3] or the wavelength dependency of some optical components [4]. We describe a new countermeasure that consists in fairly general statistical tests build on top of an existing real-time shot-noise measurement technique [2] and show that it is able to defeat all these attacks.

I. INTRODUCTION

QKD [5] is a quantum information technology primitive which allows two distant parties, Alice and Bob, to establish a secret key out of communications through an untrusted physical channel and a public authenticated channel. Its advantage over classical cryptographic primitives is that the produced keys are secure in the information-theoretic sense even against an eavesdropper limited only by the laws of quantum mechanics, i.e. in particular with unlimited computational resources and/or advanced knowledge of classical cryptanalysis techniques.

The first QKD realizations used unique photons to carry information. In their case, the impossibility to clone photons perfectly is the source of the security. Although the requirement to send unique photons was relaxed, their measurement still requires specialized detectors, usually cooled much below 0 degree Celsius. In contrast to this technology usually named Discrete Variable (DV) QKD, Continuous Variable QKD or CVQKD has been the subject of great interest in the past few years [6] because it can be implemented with only standard components optimized for the telecommunications industry and features nice characteristics such as a coherent detector which acts as a natural filter against classical Raman noise [7] which is the main source of noise in WDM infrastructures where QKD signals coexist with intense signals at other wavelengths. Thanks to the possibility to encode more than one bit per pulse, it was shown recently that its secret key rate is expected to be competitive with DVQKD techniques assuming some reasonable clock rate improvements [8]. The practical secure distance of CVQKD is 80 km [9] taking into account finite-size effects and several other imperfections [10]. It is currently limited by the technical noise of the system which is attributed to a potential eavesdropper.

Quantum hacking against commercial DVQKD systems has been demonstrated in the past few years through powerful attacks, namely the time-shift attack [11], the phase-remapping attack [12], and the remote control of single-photon detectors using tailored bright illumination [13]. Other attacks proposed against discrete-variable QKD systems include Trojan horse [14], device calibration [15], and wavelength dependent beamsplitter [16] attacks. The latter has been adapted to CVQKD while saturation attacks [3] are CVQKD-specific. Generally, hacking of practical devices is a central topic in QKD since this is where the high theoretical security expectations of QKD meet reality.

For any QKD system, some statistics about the transmitted data are used to evaluate the amount of information that may be in possession of an attacker. For DVQKD protocols, the transmitted data are usually binary strings, and the eavesdropper information estimate only depends on the bit error rate between Alice’s and Bob’s data, and is zero for a null bit error-rate. For coherent-state CVQKD, the picture is more complex: the received data is always noisy because of a quantum noise term, the shot noise. The estimate of the eavesdropper information depends on the excess noise which is the difference between the observed noise and the shot noise, and also on the observed correlation between the emitter and the receiver. The excess noise is typically at least one order of magnitude smaller than the shot noise.

One of the main implementation difficulties of CVQKD is to measure the shot noise with the required precision [9] to be able to compute the excess noise on the signal. The historical method to estimate the shot noise is to calibrate once-and-for-all the slope of the local oscillator to shot noise linear relation on the homodyne detection, and then to measure in real time the power of the local oscillator. It was shown in [2] that this relationship is prone to changes over time, especially under the influence of an attacker. In the case of a polarization dependent scheme, as the one pictured in Fig. 1 the shot noise is usually measured at the entrance of Bob before polarizations are separated; in this setting, this shot noise estimation technique is prone to shot noise overestimation when a polarization drift causes the optical power on the detection to be smaller than expected.
Because of the shortcomings of this shot noise measurement technique, another technique was proposed in [2]. It attempts to separate the shot noise and the signal noise contributions in the total measured noise on signal by adding an active attenuation device on Bob’s signal path, and measuring the total noise for two values of the attenuation ratio, exploiting the fact that the contribution of the signal noise is proportional to the signal itself, hence should be proportional to the attenuation factor imposed to the signal.

Unfortunately, a theoretical attack scenario against this procedure was proposed in [4]. It works by exploiting defects in typical beam splitters to introduce specially crafted signals in both the signal path and the local oscillator path on Bob’s side, which causes the attenuation-to-noise relation to become non-affine. The computations to evaluate the contributions of the shot noise and the excess noise in the total noise then yield wrong results, and with correct attack parameters this produces an underestimated excess noise and an overestimated key rate. It was remarked in [4] that using not two, but three attenuation ratios was sufficient to thwart this attack.

In this paper we present a more general shot noise / excess noise measurement technique using an active attenuation device on Bob’s side, and show how it thwarts known attacks on CVQKD systems. More importantly, we show that experimental data extracted from a real CVQKD setup are compatible with this countermeasure, i.e. real data are not incorrectly flagged as being caused by an attacker.

In Section III we present our optical setup and detail how we obtain our measurement results. We devote Section IIIA to practical excess noise measurements and study three cases: Section IIIB gives experimental results in the absence of an attacker while Section IIIC and Section IIIC respectively include simulations of the behavior of the countermeasure when faced with the wavelength attack [4] and the saturation attack [3].

II. OPTICAL SETUP, MEASUREMENTS

Figure 1 presents our implementation of the Gaussian coherent-state homodyne CVQKD protocol [17]. Alice modulates a Gaussian mixture of coherent states with a phase and an amplitude modulator. Both the strong phase reference signal (called the Local Oscillator or LO for short) and the weak information-carrying signal are sent to Bob, thanks to a polarization and time multiplexing scheme. On Bob’s side, the polarization drift that occurs on the quantum channel is actively compensated, then the LO and the signal are separated and resynchronized. A phase modulator on the LO path enables to select the quadrature of the signal that is measured on the homodyne detection through the interference of the phase-modulated LO and the signal. An amplitude modulator is added on the signal path to simulate various signal attenuation ratios. For each attenuation ratio and for large enough groups of $N$ pulses, the measured signal $Y$ is projected as a $N$-dimensional vector on the signal $X$ emitted by Alice. Thus $Y$ is written $S + (Y - S)$ where $S$ is the signal proportional to $X$ and $Y - S$ is the noise orthogonal to $X$. In particular, when the signal path attenuation ratio is set to 0, the signal is 0 up to statistical noise and, assuming there is no manipulation of the LO, the shot noise is measured on the homodyne detection. Independently of the attenuation level, the noise $Y - S$ is equal to the sum of the shot noise and the excess noise. The current let through by the homodyne detection photodiodes is amplified by a transimpedance amplifier and the raw measurement results are therefore expressed in volts.

Assuming the homodyne detection has a linear signal response, the relationship of the excess noise variance to the signal variance, as well as the relationship between the signal variance and the attenuation ratio imposed on the signal path, should be linear. Indeed, any noise added to the signal will scale with the signal. A failure of one of these two tests gives a strong indication of the presence of an attacker as we shall see in the next sections.

Of course, to prevent any adaptive behavior of an attacker that would act differently depending on the attenuation ratio applied by Bob, it is required to choose randomly the attenuation ratio applied to individual pulses. Furthermore the modulation of Alice must be independent of these attenuation choices. In our setup, these choices are made by Bob. Since secret key is produced only from the least attenuated groups, some information about the attenuation ratio selection is transmitted to Alice by Bob after the signal pulses are received. In the usual model where Alice’s and Bob’s laboratories are assumed to be out of reach of any attacker, this is sufficient to prevent any adaptive behavior of the attacker. Trojan horse attacks [14] are examples of attacks where Eve might acquire knowledge of Alice’s and/or Bob’s measurement and attenuation choices but these attacks are not considered in the present paper.

III. SLOPE MEASUREMENTS: PRACTICAL RESULTS

A. Experimental Results Without an Attacker

We first show the noise-to-signal and signal-to-attenuation relationships on a practical experimental setup. On Bob’s side, 16 different attenuation settings are used with ratios $r_i$ satisfying $r_i/r_{i+1} \approx 0.7$. This ratio is chosen so that $r_0/r_{15}$ is approximately equal to the amplitude modulator extinction ratio (here about 25 dB) in order to exploit its attenuation range. Data are gathered according to their attenuation setting and to the quadrature measurement performed ($X$ or $P$). The noise and signal variance in each group are computed on approximately $N = 5 \times 10^8$ points in each group. The resulting signal-to-noise relationships are shown on Fig.
Figure 1: Optical scheme of a GG02 implementation using time and polarization multiplexing, and including an amplitude modulator on the signal path of the receiver.

Figure 2: Noise variance as a function of signal variance in arbitrary units, X quadrature. The noise and signal variance are computed on approximately $N = 5 \times 10^8$ points in each group. The correlation coefficient $R^2$ is 0.9974 for quadrature X. The slope is equal to $2.07 \times 10^{-3}$. Error bars corresponding to 3 standard deviations of the statistical estimators are shown for the noise only, since for the signal variance they would be too thin to see.

The correlation coefficient $R^2$ is 0.9974 for quadrature X and 0.9979 for quadrature P. In shot noise units (i.e. divided by the constant term of the affine interpolation), the absolute values of errors w.r.t the linear model never exceed $1.3 \times 10^{-4}$. This shows that a simple affine model for the noise-to-signal relationship can be observed in practice with good precision. The estimations of the shot noise derived by linear regression on both quadratures, despite being obtained from two different datasets, also agree with very good precision: on quadrature X, the estimate is 783.16 mV, while on P it is estimated to 783.19 mV, meaning that the relative error is about $3.5 \times 10^{-5}$. This error is consistent with our dataset size since it corresponds to approximately 0.55 standard deviations of the variance estimator of N normalized Gaussian values. The slope of both curves, equal respectively to $2.07 \times 10^{-3}$ and $2.12 \times 10^{-3}$, show that the excess noises observed on Bob side for small SNRs (e.g. 0.1) are compatible with key production up to 100 km approximately.

Figure 4 shows the variation of the signal variance as a function of the attenuation applied by Bob, on the same dataset. Again we have a very good linearity ($R^2 = 0.9996$). Only the X quadrature is represented, since X and P curves are visually indistinguishable. Indeed, the small error term is not random but rather caused by a systematic error in the control of the attenuation ratio, and is very similar on X and P datasets.

From these experimental results we derive the criteria that are applied to incoming data that govern whether data blocks are used to produce some secret key: When $16 \times N$ data have been acquired at various attenuation levels, the least attenuated block is used for key production if the squared correlation factor $R^2$ for the linear fit of the noise-to-signal curve is above 0.99 and the maximum error term in shot noise units (difference between the fitted line and the observed curve) does not exceed $2.10^{-4}$. Additionally the squared correlation factor of the attenuation-to-signal-variance relationship must be above 0.99. When some data is accepted, the excess noise and the shot noise used for the key rate estimation are computed with the help of the linear regression on the signal-to-noise curve. If secret key is produced only from the least attenuated group, 94% of the pulses are used for parameter estimation. This can be overcome by using other blocks for key production, or by increasing the size of the least attenuated block.
B. Simulation with Wavelength Attack

In the wavelength attack [4], the optical setup and the imperfections of beam splitters are exploited to add to the measured values a deterministic signal whose amplitude depends on Bob’s signal attenuation ratio as a degree-2 polynomial. The parameters of the attack can be chosen so that this polynomial has its minimum when the attenuation ratio is equal to 1. When the attenuation ratio is 0, a positive term is added to the shot noise variance and the sum of these two quantities is mistakenly used as the shot noise estimate, causing the excess noise to be improperly normalized. When combined with a full intercept-and-resend attack that is performed at the output of Alice, the noise added by the intercept-and-resend can be partially masked by the wavelength attack. An example of this scenario is summarized on Fig. 5.

The signal-to-attenuation relationship is unaffected by this attack, hence its linearity cannot be used to detect the presence of the attacker.

In the case of a full intercept and resend, the signal that must be added to cancel the excess noise induces a rather strong nonlinearity which results in a squared correlation factor of 0.94, and is therefore easily detected. In a more general scenario, the attack can be used to tamper slightly with the excess noise evaluation of the receiver. For a difference to linearity of $2 \times 10^{-4}$ shot noise units as we observe normally in our setup, it is possible to use the wavelength attack to hide approximately $1 \times 10^{-3}$ units of excess noise on the receiver side for each unit of signal. This slope is below the technical noise of our system which is about $2 \times 10^{-3}$ units of excess noise for each unit of signal as seen on Fig. 2 and Fig. 3. Therefore adopting the data acceptance criteria laid out at the end of section IIIA the measured slope of the noise-to-signal relationship can be trusted up to $1 \times 10^{-3}$ units of noise per unit of signal variance. With this conservative parameter estimation procedure, results of [9], which were established at a SNR of 0.08 for long distances, remain valid: a noise slope of $2 \times 10^{-3}$ due to $1 \times 10^{-3}$ (where the first term corresponds to the technical noise of our system and the second term is the margin of error possibly caused by the wavelength attack) corresponds to an estimated excess noise on Bob side of $0.8 \times 3.10^{-3} = 2.4 \times 10^{-4}$ shot noise units which gives a positive key rate at 80.5 Km.
C. Simulation With Saturation Attack

As the wavelength attack, the saturation attack is aimed at hiding a full or partial intercept and resend by tampering with the excess noise measurement of the receiver. It assumes that the homodyne detection, when measuring a quadrature \( x \) of some state, outputs not \( x \) but \( \text{Min} (\text{Max}(x, -\alpha), \alpha) \) for some \( \alpha > 0 \). Indeed, only a finite range of quadrature values can be adequately measured, either because of limitations of the analog electronics of the detection, or because of the range of the analog-to-digital conversion. To exploit this, the states received by Bob are shifted by an offset which causes the detection to be near the saturation point. When the offset is increased, both the signal and the noise variance tend to 0; indeed, for a sufficiently large value of the offset, the detection is stuck to its saturation level with overwhelming probability. If one uses an attenuation device on the signal path to shut off the signal and measure the shot noise, this measure is unaltered because the saturated signal is blocked when the shot noise is measured. As a result, a simple excess noise measurement where only the difference between the signal noise and the shot noise, this measure is unaltered because the saturated signal is blocked when the shot noise is measured. As a result, a simple excess noise measurement where only the difference between the signal noise and the shot noise is computed may be defeated by this attack. However, simulations show that a more complete look at the signal and noise for several signal attenuations ratios easily reveals the saturation phenomenon.

All simulations are done with a value of \( \alpha \) equal to 4 times the shot noise standard deviation, so that the shot noise variance measurement is not significantly affected by the saturation limit of the detection. The offset imposed by the attacker is also equal to 4 times the shot noise standard deviation. A full intercept and resend attack is assumed, so that the excess noise is 2 units of shot noise for an unattenuated signal.

For a high variance on the receiver side, we obtain Figure 6. In this case the obtained curve is clearly not a line (the squared correlation coefficient is 0.386). Furthermore the apparent noise-vs-signal slope is higher than in the non-saturated case: in this regime, the attack does not improve the perceived excess noise slope.

On the contrary, for low variance on the receiver side such as the results shown in Figure 7 (here a signal variance of 1 unit of shot noise), the noise variance falls before the signal variance falls too, and the resulting curve is below the normal unsaturated line, hence a direct comparison of the points with attenuation ratios of 0 and 1 would yield an excess noise slope lower than the real, unsaturated value. However the full curve observed is again very far from being a line (the squared correlation coefficient is now 0.665), which enables us to reject the acquired data without even considering the signal-to-attenuation response curve.

The limit case is represented by Figure 8. In that case, the noise and signal variance values, despite being severely distorted compared to the unsaturated case, still lie on the unsaturated line: the squared correlation coefficient is 0.998, with a slope very close to the unsaturated line (0.39 vs 0.40). Thus a direct look at this relationship may not be sufficient to reject the data. However, the signal-to-attenuation plot is clearly abnormal as shown on Figure 9 when the attenuation ratio is close enough to 1, the signal variance falls instead of increasing. Even without resorting to the signal-to-attenuation linearity check, it can be noted that the attack fails to lower significantly the excess noise slope, and therefore fails to hide the intercept-and-resend added noise in that case.

Figure 6: Noise/Signal curve similar to experimental curves of Fig. 2 and Fig. 3 with and without saturation, for a modulation variance on the receiver side \( V_B = 20 \).

Figure 7: Noise/Signal curve with and without saturation, for a line transmission of 1 and a modulation variance \( V_B = 1 \).

Figure 8: Noise/Signal curve with and without saturation, for a line transmission of 1 and a modulation variance \( V_B = 5 \).

One area not explored in these graphs is the variation of the offset used by the attacker to cause the detection saturation. Its effect is to scale the saturated
noise-to-signal curve: with a higher offset, the saturation effects appear for a smaller attenuation ratio. It does not change the general shape of the curve and in particular the squared correlation factor of the linear fit can only be improved by reducing the offset to reduce the overall saturation effect.

IV. CONCLUSION

We proposed some specific criteria to decide whether data should be accepted for post-processing in an implementation of Gaussian coherent CVQKD. These criteria depend on the normal behavior of an optical system implementing the protocol and define acceptable deviations from the ideal apparatus. It was shown that such a data filtering can defeat known attacks on coherent states CVQKD, except in the limit where the magnitude of the signature of the attack is comparable to the technical noise of the system. The parameter estimation can be modified in a conservative way to get around this limitation. Checking the linearity of the signal-to-attenuation response is an approximate way to test whether the homodyne detection output is proportional to a quadrature of the incoming electromagnetic field, which is an implicit hypothesis in all CVQKD security proofs. Because of the general nature of this test, we believe it applies to a wide range of potential attacks aimed at biasing the parameter estimation process and considerably improves the practical security of CVQKD.

V. ACKNOWLEDGEMENTS

This research was supported by the French National Research Agency, through the HIPERCOM (2011-CHRI-006) project, by the DIRECCTE Ile-de-France through the QVPN (FEDER-41402) project, and by the European Union through the Q-CERT (FP7-PEOPLE-2009-IAPP) project.

[1] J. Lodewyck, T. Debuisschert, R. García-Patrón, R. Tualle-Brouri, N. Cerf, and P. Grangier, Phys. Rev. Lett. 98, 030503 (2007).
[2] P. Jouguet, S. Kunz-Jacques, and E. Diamanti, Phys. Rev. A 87, 062313 (2013).
[3] H. Qin, R. Kumar, and R. Alléaume, Proc. SPIE 8899, 88990N (2013).
[4] J.-Z. Huang, S. Kunz-Jacques, P. Jouguet, C. Weedbrook, Z.-Q. Yin, S. Wang, W. Chen, G.-C. Guo, and Z.-F. Han, Phys. Rev. A 89, 032304 (2014).
[5] V. Scarani, H. Bechmann-Pasquinucci, N. J. Cerf, M. Dušek, N. Lütkenhaus, and M. Peev, Rev. Mod. Phys. 81, 1301 (2009).
[6] C. Weedbrook, S. Pirandola, R. García-Patrón, N. J. Cerf, T. Ralph, J. Shapiro, and S. Lloyd, Rev. Mod. Phys. 84, 621 (2012).
[7] B. Qi, W. Zhu, L. Qian, and H. Lo, New J. Phys. 12, 103042 (2010).
[8] P. Jouguet, D. Elkouss, and S. Kunz-Jacques, Arxiv preprint arXiv:1405.1050 [quant-ph] (2014).
[9] P. Jouguet, S. Kunz-Jacques, A. Leverrier, P. Grangier, and E. Diamanti, Nature Photonics 7, 378 (2013).
[10] P. Jouguet, S. Kunz-Jacques, E. Diamanti, and A. Leverrier, Phys. Rev. A 86, 032309 (2012).
[11] Y. Zhao, C.-H. F. Fung, B. Qi, C. Chen, and H.-K. Lo, Phys. Rev. A 78, 042333 (2008).
[12] F. Xu, B. Qi, and H.-K. Lo, New J. Phys. 12, 113026 (2010).
[13] L. Lydersen, C. Wiechers, C. Wittmann, D. Elser, J. Skaar, and V. Makarov, Nature Photonics 4, 686 (2010).
[14] N. Gisin, S. Fasel, B. Kraus, H. Zbinden, and G. Ribordy, Phys. Rev. A 73, 022320 (2006).
[15] N. Jain, C. Wittmann, L. Lydersen, C. Wiechers, D. Elser, C. Marquardt, V. Makarov, and G. Leuchs, Phys. Rev. Lett. 107, 110501 (2011).
[16] 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, et al., Phys. Rev. A 84, 062308 (2011).
[17] F. Grosshans and P. Grangier, Phys. Rev. Lett. 88, 057902 (2002).