Editors' Suggestion

Quantum Hacking Against Discrete-Modulated Continuous-Variable Quantum Key Distribution Using Modified Local Oscillator Intensity Attack with Random Fluctuations

Lu Fan, Yiming Bian, Mingze Wu, Yichen Zhang,* and Song Yu

State Key Laboratory of Information Photonics and Optical Communications, School of Electronic Engineering, Beijing University of Posts and Telecommunications, Beijing 100876, China

(Received 1 April 2023; revised 3 June 2023; accepted 7 August 2023; published 29 August 2023)

The local oscillator in a practical continuous-variable quantum key distribution system fluctuates at any time during the key distribution process, which may open security loopholes for the eavesdropper to hide her eavesdropping behaviors. Based on this, we investigate a more stealthy quantum attack where the eavesdroppers simulate random fluctuations of local oscillator intensity in a practical discrete-modulated continuous-variable quantum key distribution system. Theoretical simulations show that both communicating parties will misestimate channel parameters and overestimate the secret key rate due to the modified attack model, even though they have monitored the mean local oscillator intensity and shot noise as commonly used. Specifically, the eavesdropper's manipulation of random fluctuations in LO intensity disturbs the parameter estimation in a realistic discrete-modulated continuous-variable quantum key distribution system, where the experimental parameters are always used for constraints of the semidefinite program modeling. The modified attack introduced by random fluctuations of local oscillator can only be eliminated by monitoring the local oscillator intensity in real time, which places a higher demand on the accuracy of monitoring technology. Moreover, similar quantum hacking will also occur in a practical local local oscillator system by manipulating the random fluctuations in pilot intensity, which shows the strong adaptability and the role of the proposed attack.

DOI: 10.1103/PhysRevApplied.20.024073

I. INTRODUCTION

Quantum key distribution (QKD) [1] relies on quantum mechanical properties to ensure communication security, which allows two remote parties to generate unconditionally secure keys despite the presence of potential eavesdroppers in quantum channels [2,3]. Depending on the dimension of coding, it can be divided into discrete-variable QKD (DVQKD) and continuous-variable QKD (CVQKD) [4,5]. Between them, CVQKD uses coherent detection technology to replace single-photon detection technology, which can be well compatible with classical optical communication systems [6]. In addition, the transmission distance achieved recently can support the requirement in metropolitan distances [7–9], which facilitates the large-scale application of QKD [10–16].

CVQKD based on Gaussian modulation (GM) [17–27] is widely used, where the quadratures of quantum states are modulated according to the Gaussian probability distribution. Moreover, discrete-modulated (DM) CVQKD [28], where the quadratures of quantum states are modulated discretely, has also been gaining a lot of attention. Due to the simplicity of state preparation [9,29,30] and lower complexity of classical error correction [31,32], the unconditional security of DM CVQKD protocols (e.g., two-state and three-state protocols [33,34]) has been successively proved. Recently, one of its security analysis theories is using the semidefinite program (SDP) model to give the lower bound of the secret key rate [35–39], and related experiments have also been carried out [9,31,40]. With the tremendous progress in theory and experiment of DM CVQKD in recent years, its practical system security has received increasing attention. Different from theoretical analysis, the transmission of the local oscillator in the practical DM CVQKD system would leave Eve with vulnerabilities to perform quantum hacking.

The local oscillator (LO) is required for stable coherent detection, whose nonideality within practical environments is usually ignored in the security analysis of CVQKD. Due to the imperfection of practical blocks and devices, researchers have proposed many quantum hacking models aiming at practical LO, such as local oscillator calibration attack [41], local oscillator intensity fluctuation attack [42], wavelength attacks [43,44], polarization attack [45], and so on. Particularly, the transmission of LO opens a security vulnerability, which gives eavesdropper (Eve)

*zhangyc@bupt.edu.cn

advantages to exploit the fluctuation of LO intensity [46]. We note that in the previous studies [42], Eve is assumed to reduce the intensity of LO overall using variable attenuators, which can easily be defended by commonly used monitoring technology such as monitoring the LO intensity mean and optical power or monitoring the shot noise. In other words, the above manipulation and monitoring of LO intensity is carried out on its overall mean. When Eve simulates random fluctuations in LO intensity with invariant mean value, a more serious security loophole appears because the above-mentioned monitoring techniques will fail.

In this paper, we propose a modified local oscillator intensity attack (LOIA) with random fluctuations. Eve can exploit the random fluctuating properties of LO intensity to launch a more stealthy attack by attenuating a part of LO pulses while amplifying another part of LO pulses. The security of the practical DM CVQKD system under this attack is considered, which mainly includes the effect of random fluctuations in LO intensity on the estimation of experimental parameters in the SDP security analysis model. Quantitative theoretical simulations support that Eve's manipulation of random fluctuations in LO intensity would cause insecure final keys to be shared between legitimate parties. Remarkably, the fluctuating variance of LO intensity is found to be the core character and even no secure key could be generated when the fluctuating variance is relatively large. While seriously affecting system security, the modified attack model has the advantage of evading traditional commonly used monitoring techniques. Therefore, the communication parties must employ real-time LO intensity monitoring with higher accuracy to evade such an attack. The study complements the research of quantum hacking of the practical DM CVQKD system.

This paper is organized as follows. In Sec. II, we first introduce the modified LOIA with random fluctuations, then we perform the practical security analysis of DM CVQKD under the modified attack model in Sec. III. Theoretical simulations prove the feasibility of the attack in Sec. IV. The discussion is shown in Sec. V, and finally the conclusion is given in Sec. VI.

II. THE MODIFIED LOIA WITH RANDOM FLUCTUATIONS

The practical DM CVQKD system mainly includes optical source, discrete modulation, quantum channel, heterodyne detector, and so on, as shown in Fig. 1. The coherent light generated by the laser is amplitude modulated to obtain optical pulses that meet the requirements of the certain extinction ratio and duty cycle, and then Alice divides them into two parts using the optical BS, one part includes LO pulses with strong power and the other part includes the signal pulses with weak power. Furthermore, Alice encodes the key information on the signal pulses through discrete modulation, where the prepared coherent state is selected from a finite number of constellations in phase space. Time-multiplexing technology and polarization multiplexing technology allow modulated signal pulses and LO pulses to be transmitted in the same channel without interference in the scheme. Bob demultiplexes the received pulses and then obtains the outcomes using the heterodyne detector. As can be seen from the constellation plot of two different kinds of discrete modulation (i.e., QPSK and QAM) in Fig. 2, each coherent state represents a constellation. In phase space, the regular components x and p of the coherent state are discretely distributed. For the more commonly used QAM modulation method, the mutual information decreases compared with the Shannon capacity when its constellations are uniformly distributed. Probability constellation shaping (PCS) was therefore introduced to address this limitation, with the idea of using a nonuniform distribution on a QAM lattice. The probability distribution for such a PCS MQAM is

$$p_{x,p} \sim \exp(-\nu(x^2 + p^2)). \tag{1}$$

This distribution depends on $\nu > 0$ and the spacing between the possible values of x (or p). The parameter, ν allows changing the shape of the Gaussian-like distribution on the constellation, which can be optimized to maximize the key rate.

Before the setup of such a CVQKD system, the communication parties will calibrate the LO intensity value in

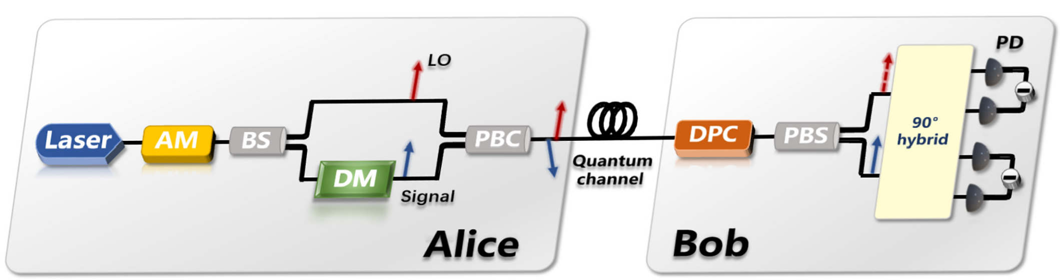

FIG. 1. The practical block diagram of a discrete-modulated CVQKD system. AM, amplitude modulation; BS, beam splitter; DM, discrete modulation; PBC, polarization beam coupler; DPC, dynamic polarization controller; PBS, polarizing beam splitter; PD, photodiode.

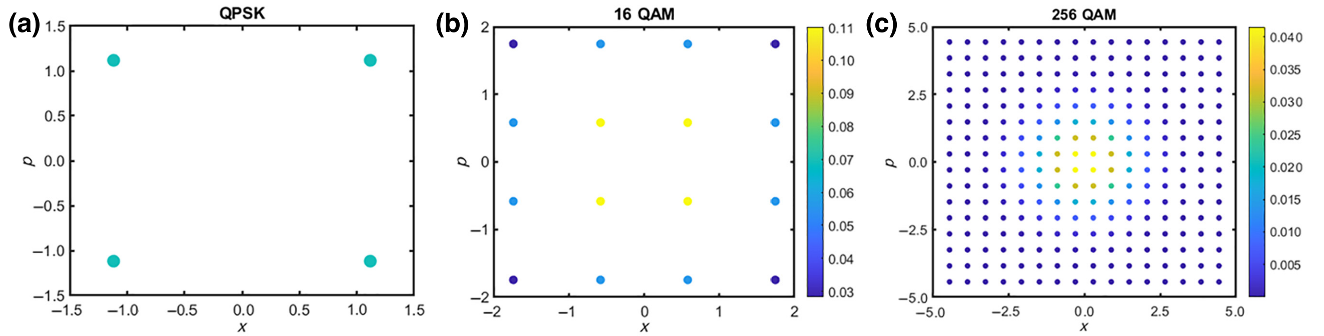

FIG. 2. The constellation plots of different discrete modulation schemes are presented. The regular components of the coherent state are no longer continuous, but discrete. The first corresponds to the constellation plot of quadrature phase-shift keying (QPSK). The latter two correspond to the constellation plots of 16 quadrature amplitude modulation (16 QAM) with $\nu = 0.085$ and 256 quadrature amplitude modulation (256 QAM) with $\nu = 0.035$ where ν is the Gaussian-like distribution parameter. Colors indicate the probabilities corresponding to each coherent state.

advance to obtain the shot-noise unit (SNU). The transmission process of LO is always ideally assumed safe in the theoretical security analysis while there are some imperfections in the practical system. During the transmission of LO pulses on the quantum channel, Eve simulates the random fluctuations in LO intensity around its mean value to cover up her attack on signal pulses that successfully steals final keys, which can be called the LOIA with random fluctuation. The modified attack LOIA with random fluctuations can be carried out in many ways where one of the implementations is presented in Fig. 3(a). Eve may actively open a security loophole: she intercepts the signal pulses and the LO pulses, replacing the imperfect channel between the communication parties as her own quantum channel. The LO pulses are transmitted on the attacked channel with attenuation factor α_{attack} and the signal pulse are transmitted on the normal fiber channel with attenuation factor α_{std}. Without changing the phase of LO, Eve can manipulate LO intensity to simulate random fluctuations by changing the attenuation factor α_{attack}. Eve then couples the fluctuating LO pulses and signal pulses together and sends them to the receiver. The attack effect is shown in Fig. 3(b). The calibrated SNU in advance is defined as $u_S = A^2 a_{\text{LO}}^2$ where LO is considered immutable during the key distribution process. Here A represents the amplification within the heterodyne detector and $I_{\text{LO}} = a_{\text{LO}}^2$ represents the intensity of LO. However, LO intensity fluctuates randomly around its mean value at any time during the key distribution process due to Eve's manipulation. Whereas the previous LOIA model considered the overall attenuation of LO, here Eve carries out the modified LOIA model where the fluctuation in LO intensity is a random variable with invariant mean to evade the LO intensity mean monitoring and shot-noise monitoring. The accurate description of LO intensity fluctuations will be directly related to the upper limit of the amount of final keys that Eve could steal. Therefore, the practical LO intensity $|a_{\text{LO}}^k|^2 = k|a_{\text{LO}}|^2$ where the attack factor k is introduced to indicate the random fluctuation of LO intensity. To simplify the analysis, we can describe the

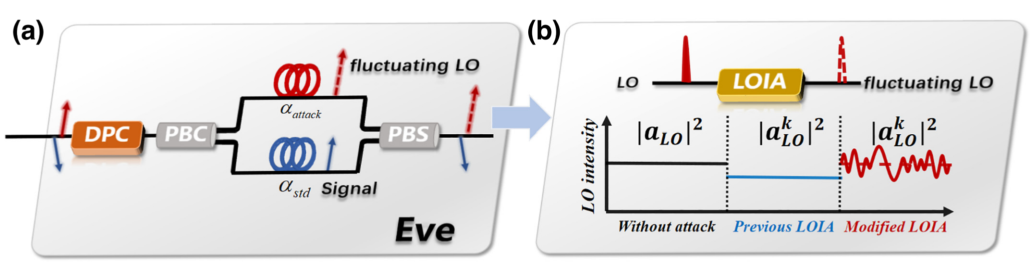

FIG. 3. (a) The implementation of the modified LOIA with random fluctuations. Due to Eve's manipulation, the LO and signal pulses pass through fibers with different attenuation factors. The LO received at the receiver end fluctuates randomly around the mean value. (b) The details of Eve's quantum hacking attack on LO. The line at the left side (black) refers to initially calibrated LO intensity; the line in the middle (blue) is the instantaneous LO intensity under the previous LOIA model where Eve attenuates the LO intensity overall; the line at the right side (red) is the instantaneous LO intensity under the modified LOIA model. Eve simulates the random fluctuations of LO intensity to launch the LOIA.

attack model through some following conditions:

(a) k is an independent random variable with the same distribution.
(b) k follows some probability distributions with a mean of $E_k = 1$ and variance V_k.
(c) k is independent of the LO intensity $I_{\rm LO}$.
(d) The probability distribution function (PDF) of k will not change during CVQKD process.

The practical SNU in the presence of the attack will be $u_S^k = ku_S$ which is different from the previous calibrated SNU. If Alice and Bob still use the calibrated SNU for parameter estimation, the system's excess noise will be misestimated and insecure secret keys will be generated. As a result, this attack model reveals a more serious loophole that cannot be avoided by commonly used monitoring techniques.

III. THE MODIFIED LOIA AGAINST A PRACTICAL DM CVQKD SYSTEM

Unlike the particularity of Gaussian modulation, the covariance of Alice and Bob for discrete modulation with finite constellations cannot be constructed, so there is no way to directly obtain a corresponding covariance matrix that conforms to entanglement-based (EB) model. Recent studies have shown that the security analysis of DM CVQKD can be performed using the SDP modeling, which represents the calculation of secret key rate as a convex optimization problem. The difference of SNU in the modified LOIA from the calibration SNU affects the normalization of the experimental parameters used for constraints in the SDP modeling.

A. The SDP modeling of DM CVQKD protocol

In the prepare-and-measure (PM) version of the DM CVQKD protocol, a set of coherent states $\{|\alpha_k\rangle\}$ are prepared, where Alice chooses one of the states $|\alpha_k\rangle$ to be sent with the probability p_k. The average state sent by Alice can be represented by the density matrix τ given by the weighted mixture of coherent states,

$$\tau := \sum_k p_k |\alpha_k\rangle \langle\alpha_k|. \tag{2}$$

We define the quadrature operators $\hat{x}$ and $\hat{p}$ in phase space as $\hat{x} := \hat{a} + \hat{a}^\dagger$ and $\hat{p} := -i(\hat{b} + \hat{b}^\dagger)$, where $\hat{a}$ and $\hat{a}^\dagger$ (respectively, $\hat{b}$ and $\hat{b}^\dagger$) are the annihilation and generation operators of Alice's register A (respectively, Bob's register B) on Fock space. In the EB version of the DM CVQKD protocol, Alice prepares the bipartite state $|\Phi_{AA'}\rangle$ and makes local POVM measurements of the reserved mode A,

$$|\Phi_{AA'}\rangle = \sum_{k=1}^{M} \sqrt{p_k}\, |\psi_k\rangle\, |\alpha_k\rangle\,, \tag{3}$$

where the $\{|\psi_k\rangle\}$ form an orthonormal basis and M stands for the number of constellations. Then another mode A' to be sent collapses to the corresponding coherent state. Alice sends A' through the quantum channel $\mathcal{N}_{A'\to B}$ and Bob obtains the received mode B. The quantum state shared by communication parties is denoted as $\rho_{AB} = (id_A \otimes \mathcal{N}_{A'\to B})(|\Phi_{AA'}\rangle\langle\Phi_{AA'}|)$, where id_A stands for the identity operator acting on mode A.

When Eve performs the collective attack, the secret key rate for reverse reconciliation can be calculated using the Devetak-Winter formula,

$$R = \beta I_{AB} - \sup_{\mathcal{N}_{A'\to B}} \chi_{BE}, \tag{4}$$

where β is the reconciliation efficiency [47]. I_{AB} is the mutual information between Alice and Bob, and χ_{BE} is the Holevo bound, which describes the mutual information between Bob and Eve. The covariance matrix between Alice and Bob is assumed to be in general form [38],

$$\gamma_{AB} = \begin{pmatrix} V\cdot I_2 & Z\cdot\sigma_z \\ Z\cdot\sigma_z & W\cdot I_2 \end{pmatrix}, \tag{5}$$

where

$$I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

The parameter V, Z, W can be given by $V := 1 + 2\,\mathrm{tr}(\rho\hat{a}^\dagger\hat{a})$, $Z := \mathrm{tr}(\rho(\hat{a}\hat{b} + \hat{a}^\dagger\hat{b}^\dagger))$, and $W := 1 + 2\mathrm{tr}(\rho\hat{b}^\dagger\hat{b})$. The Holevo bound can be given by

$$\chi_{BE} = \sum_{i=1}^{2} G\left(\frac{\lambda_i - 1}{2}\right) - G\left(\frac{\lambda_3 - 1}{2}\right), \tag{6}$$

$\lambda_{1\sim 2}$ represent the symplectic eigenvalues of γ_{AB}, and $\lambda_3 = V - [Z^2/(1+W)]$ [4]. Both V and W can be observed by Alice and Bob locally, there is left unknown $Z := \mathrm{tr}(\rho C)$ with $C = \hat{a}\hat{b} + \hat{a}^\dagger\hat{b}^\dagger$ in the covariance matrix, which needs to be bounded using some constraints to obtain the secret key rate.

In the security analysis method of SDP modeling, the objective function is the minimum value of the covariance $Z := \mathrm{tr}(\rho C)$. The first constraint merely says that ρ_{AB} is obtained by applying the channel $\mathcal{N}_{A'\to B}$ to $|\Phi_{AA'}\rangle$:

$$\mathrm{tr}_B(\rho_{AB}) = \mathrm{tr}_B((id_A \otimes \mathcal{N}_{A'\to B})(|\Phi_{AA'}\rangle\langle\Phi_{AA'}|)) = \bar{\tau}, \tag{7}$$

where we define $\bar{\tau}$ to be the complex conjugate of τ in the Fock basis. The other constraints come from observations from the PM scheme: the second moment constraint and two first moment constraints. The second moment constraint is that defining the operator $\Pi \otimes \hat{b}^\dagger\hat{b}$ where $\Pi := \sum_k |\psi_k\rangle\langle\psi_k|$ is a projector and observing that

$$\mathrm{tr}(\rho(\Pi \otimes \hat{b}^\dagger\hat{b})) = n_B, \tag{8}$$

where n_B can be measured in the protocol. An operator is introduced to define the first moment constraints,

$$\alpha_\tau := \tau^{1/2}\alpha\tau^{-1/2}. \tag{9}$$

Then the constraints are obtained:

$$\mathrm{tr}(\rho C_1) = 2c_1, \tag{10}$$

$$\mathrm{tr}(\rho C_2) = 2c_2. \tag{11}$$

The operators C_1 and C_2 are defined by

$$C_1 := \sum_k \overline{\langle\alpha_k|\,\alpha_\tau\,|\alpha_k\rangle}\,|\psi_k\rangle\langle\psi_k| \otimes \hat{b} + \text{h.c.}, \tag{12}$$

$$C_2 := \sum_k \bar{\alpha}_k\,|\psi_k\rangle\langle\psi_k| \otimes \hat{b} + \text{h.c.} \tag{13}$$

The correlation coefficients c_1 and c_2 can be estimated experimentally, and h.c. stands for Hermitian conjugate. The last constraint is that the quantum state is semidefinite. Then the SDP can be listed [38], min $\mathrm{tr}(\rho C)$,

such that

$$\times\begin{cases}\mathrm{tr}_B(\rho) = \bar{\tau},\\ \mathrm{tr}(\rho(\Pi\otimes\hat{b}^\dagger\hat{b})) = n_B,\\ \mathrm{tr}\left(\rho\sum_k \overline{\langle\alpha_k|\,\alpha_\tau\,|\alpha_k\rangle}\,|\psi_k\rangle\langle\psi_k| \otimes\hat{b} + \text{h.c.}\right) = 2c_1,\\ \mathrm{tr}(\rho\sum_k \bar{\alpha}_k\,|\psi_k\rangle\langle\psi_k|\otimes\hat{b} + \text{h.c.}) = 2c_2,\\ \rho \succeq 0.\end{cases} \tag{14}$$

Based on the SDP modeling, the analytical lower bound of Z can also be obtained directly as

$$Z^* := 2c_1 - 2\left(\left(n_B - \frac{c_2^2}{\langle n\rangle}\right)w\right)^{1/2} \;[38], \tag{15}$$

where the average photo number $\langle n\rangle$ and the quantity w is defined by the protocol

$$\langle n\rangle := \sum_k^M p_k|\alpha_k|^2 \tag{16}$$

$$w := \sum_k^M p_k(\langle\alpha_k|\,a_\tau^\dagger a_\tau\,|\alpha_k\rangle - |\langle\alpha_k|\,a_\tau\,|\alpha_k\rangle|^2). \tag{17}$$

The quantities c_1, c_2, and n_B are estimated from experiments whose values are related to SNU. Therefore, the impact of our proposed attack on SNU will also affect the parameter estimation of c_1, c_2, and n_B in DM CVQKD.

B. The parameter estimation against the modified LOIA with random fluctuations

It is assumed that N symbols of the original key are used for parameter estimation. The estimation of the three parameters is expressed separately by estimators $\hat{c}_1$, $\hat{c}_2$, and $\hat{n}_B$ [48],

$$\hat{c}_1 = \frac{1}{\sqrt{u_S}}\frac{\sqrt{2}}{N}\sum_{k=1}^N a_k y_k, \tag{18}$$

$$\hat{c}_2 = \frac{1}{\sqrt{u_S}}\frac{\sqrt{2}}{2N}\sum_{k=1}^N x_k y_k, \tag{19}$$

$$\hat{n}_B + 1 = \frac{1}{\sqrt{u_S}}\frac{1}{N}\sum_{k=1}^N |y_k|^2, \tag{20}$$

where $a_{2k-1} = \mathrm{Re}(\langle\alpha_k|\,\alpha_\tau\,|\alpha_k\rangle)$, $a_{2k} = \mathrm{Im}(\langle\alpha_k|\,\alpha_\tau\,|\alpha_k\rangle)$, x_k is the transmitted symbols with variance V_A from Alice, y_k is the measurement of Bob. The three parameters c_1, c_2, and n_B can be inferred from the expected value of estimators, and there are such relationships:

$$c_1 + \frac{3}{4N}c_1 = E\left[\hat{c}_1\right] = E\left[\tfrac{1}{\sqrt{u_S}}\right]E\left[\tfrac{\sqrt{2}}{N}\textstyle\sum_{k=1}^N a_k y_k\right], \tag{21}$$

$$c_2 + \frac{3}{4N}c_2 = E\left[\hat{c}_2\right] = E\left[\tfrac{1}{\sqrt{u_S}}\right]E\left[\tfrac{\sqrt{2}}{2N}\textstyle\sum_{k=1}^N x_k y_k\right], \tag{22}$$

$$\begin{aligned}\frac{N}{N-2}(n_B+1) &= E\left[\hat{n}_B + 1\right]\\ &= E\left[\tfrac{1}{\sqrt{u_S}}\right]E\left[\tfrac{1}{N}\textstyle\sum_{k=1}^N |y_k|^2\right].\end{aligned} \tag{23}$$

Suppose Eve carries the proposed attack and randomly fluctuates LO intensity, the practical parameters c_1^k, c_2^k, and n_B^k should be

$$c_1^k + \frac{3}{4N}c_1^k = E\left[\tfrac{1}{\sqrt{ku_S}}\right]E\left[\tfrac{\sqrt{2}}{N}\textstyle\sum_{k=1}^N a_k y_k\right], \tag{24}$$

$$c_2^k + \frac{3}{4N}c_2^k = E\left[\tfrac{1}{\sqrt{ku_S}}\right]E\left[\tfrac{\sqrt{2}}{2N}\textstyle\sum_{k=1}^N x_k y_k\right], \tag{25}$$

$$\frac{N}{N-2}(n_B^k+1) = E\left[\tfrac{1}{k\sqrt{u_S}}\right]E\left[\tfrac{1}{N}\textstyle\sum_{k=1}^N |y_k|^2\right]. \tag{26}$$

In order to obtain the explicit expression for the relevant parameters, we need to evaluate the unknown quantities $E[1/\sqrt{k}]$ and $E[1/k]$. The attack factor k can be written as $k = 1 + \Delta k$ where Δk follows some probability with a mean of $E[\Delta k] = 0$ and variance $V[\Delta k] = V_k$. According to the definition of $E[1/\sqrt{k}]$,

$$E\left[\frac{1}{\sqrt{k}}\right] = E[(1+\Delta k)^{-\frac{1}{2}}] = \frac{1}{M}\sum_{i=1}^{M}(1+\Delta k_i)^{-\frac{1}{2}}. \quad (27)$$

The LO pulses with large fluctuations are usually discarded in practical environments and Δk_i refers to a small quantity. Therefore,

$$(1+\Delta k_i)^{-\frac{1}{2}} \approx 1 - \frac{1}{2}\Delta k_i + \frac{3}{8}\Delta k_i^{\,2}. \quad (28)$$

Substituting Eq. (28) into Eq. (27),

$$E\left[\frac{1}{\sqrt{k}}\right] \approx \frac{1}{M}\sum_{i=1}^{M}\left(1 - \frac{1}{2}\Delta k_i + \frac{3}{8}\Delta k_i^{\,2}\right) = 1 + \frac{3}{8}V_k. \quad (29)$$

where

$$E[\Delta k] = \frac{1}{M}\sum_{i=1}^{M}\Delta k_i, \quad (30)$$

$$V[\Delta k] = \frac{1}{M}\sum_{i=1}^{M}\Delta k_i^{\,2}. \quad (31)$$

The same reasoning leads to the fact that $E[1/k] = 1 + V_k$. After the attack, the change in SNU makes for corresponding changes in the values of experimental parameters. When the communicating parties perform the parameter estimation without the SNU being calibrated again, the estimated value c_1, c_2, and n_B will be different from the practical value c_1^k, c_2^k, and n_B^k,

$$c_1^k = E\left[\frac{1}{\sqrt{k}}\right]c_1 \approx \left(1 + \frac{3}{8}V_k\right)c_1, \quad (32)$$

$$c_2^k = E\left[\frac{1}{\sqrt{k}}\right]c_2 \approx \left(1 + \frac{3}{8}V_k\right)c_2, \quad (33)$$

$$n_B^k + 1 = E\left[\frac{1}{k}\right](n_B + 1) \approx (1 + V_k)(n_B + 1). \quad (34)$$

From the above analysis, it can be seen that Eve could exploit the random fluctuation in LO intensity to cause incorrect parameter estimation between the communicating parties, which leaves security vulnerabilities for Eve.

IV. SIMULATIONS AND RESULTS

To observe the performance of the practical DM CVQKD system under the proposed attack more clearly, the quantum channel between Alice and Bob is modeled as a phase-insensitive Gaussian channel characterized by transmittance T_c and excess noise ξ_c. Against such a Gaussian channel, the practical parameters c_1^k, c_2^k, and n_B^k can be described as the following form [38]:

$$c_1^k = \sqrt{T_c}\,\mathrm{tr}(\bar{\tau}^{-1/2}a\bar{\tau}^{1/2}a^\dagger), \quad (35)$$

$$c_2^k = \sqrt{T_c}\,\langle n\rangle, \quad (36)$$

$$n_B^k + 1 = T_c\langle n\rangle + T_c\xi_c/2 + 1. \quad (37)$$

When the communicating parties perform incorrect parameter estimation, the equivalent estimated channel parameters T_e and ξ_e will be

$$T_e = \frac{1}{(1+\frac{3}{8}V_k)^2}T_c, \quad (38)$$

$$\xi_e = \frac{(1+\frac{3}{8}V_k)^2}{1+V_k}\xi_c - \left(1 - \frac{(1+\frac{3}{8}V_k)^2}{1+V_k}\right)V_A - \left(1 - \frac{1}{1+V_k}\right)\frac{2}{T_e}. \quad (39)$$

It can be easily seen that the estimated excess noise ξ_e of Gaussian channel is lower than the practical excess noise ξ_c on small random fluctuations of LO intensity, which means that there is a certain probability that the key rate is overestimated, resulting in the security of final secret keys not being guaranteed. The numerical simulations under the Gaussian channel assumption will be given next.

Figure 4 compares the estimated key rate K_e and the practical key rate K_p of QPSK-modulated CVQKD protocol under the modified LOIA with random fluctuation.

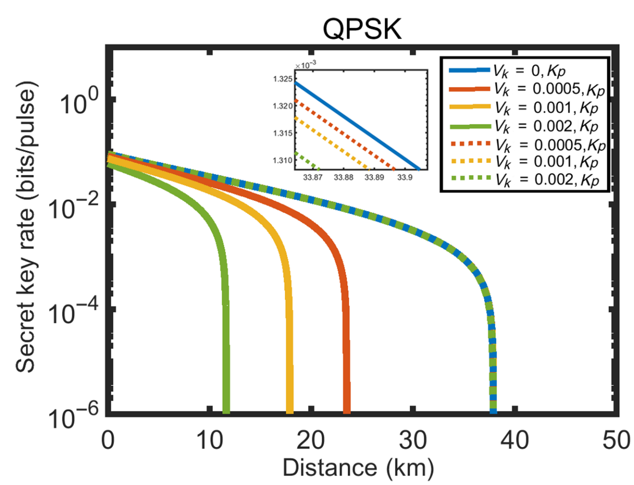

FIG. 4. The estimated and practical secret key rate versus transmission distance for the QPSK-modulated protocol against the attack with different fluctuation variances. The modulation variance $V_A = 0.456$, the estimated channel excess noise $\xi_e = 0.007$ and $\beta = 0.95$.

The modulation variance $V_A = 0.456$, the estimated channel excess noise $\xi_e = 0.007$ [9] and $\beta = 0.95$ [49]. The solid lines are the secret key rates estimated by Alice and Bob without monitoring the instantaneous LO intensity while the dashed lines reveal the practical secret key rates. Curves of different colors correspond to different degrees of fluctuation in LO intensity with the fluctuation variance $V_k = 0, 5 \times 10^{-4}, 10^{-3}, 2 \times 10^{-3}$. It can be seen that the practical maximum transmission distance will decrease dramatically if the variances is 2×10^{-3}. However, the estimated key rate obtained by both parties through parameter estimation is still close to the key rate without attack. Therefore, this attack will significantly cause both parties to overestimate the secret key rate.

Figure 5 clearly shows the amount of overestimated secret key rate by communication parties due to Eve's modified LOIA with the fluctuation variance $V_k = 2 \times 10^{-3}$ and the estimated channel excess noise $\xi_e = 0.01$. The practical maximum transmission distance is reduced to less than 10 km, while the maximum transmission distance estimated by both communicating parties can still reach more than 20 km. The secret keys generated over a communication distance of more than 10 km are insecure. It is worth noting that it can be seen from Eq. (38) that the communicating parties underestimate the transmittance due to the LO intensity random fluctuations. Therefore, the estimated secret key rate is slightly lower than the ideal one without attack as shown in the simulation plots.

Figure 6 shows the minimum fluctuation variance when Eve could steal all the keys without being detected. Two different solid lines correspond to the case with the estimated channel excess noise $\xi_e = 0.007$ and $\xi_e = 0.01$, respectively. The minimum fluctuation variance is maintained at a low value (lower than 10^{-2}), which indicates that even a small fluctuation around initial calibrated LO

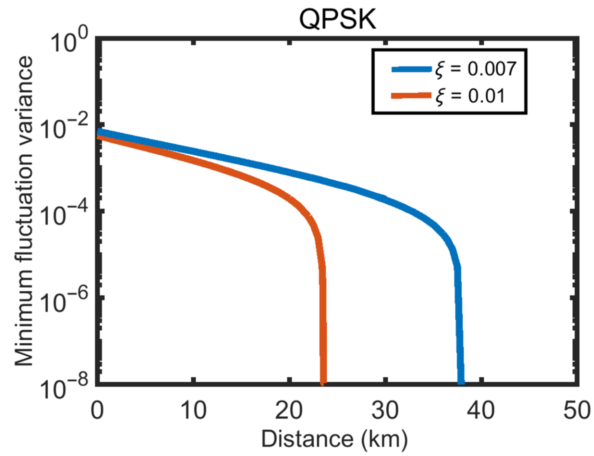

FIG. 6. The minimal fluctuation variance of LO intensity for the QPSK-modulated protocol where Eve could acquire all keys without being detected. The modulation variance $V_A = 0.456$ and the $\beta = 0.95$.

intensity will severely affect the security. Once the fluctuation variance of LO intensity exceeds the minimum value, no secret key can be generated. It can also be seen that as the distance increases or the excess noise increases, the minimum fluctuation variance is getting smaller and the attack effect is better.

Figures 7 and 8 show the performance under the 256-QAM-modulated protocol against the modified LOIA with random fluctuations. The modulation variance $V_A = 6.332$, $\nu = 0.039$ [31], and $\beta = 0.95$. Similar to QPSK, the fluctuation variance of LO intensity plays a role in the attack model. The greater the fluctuation variance, the more underestimation of practical excess noise on both the communication sides, and the more insecure keys

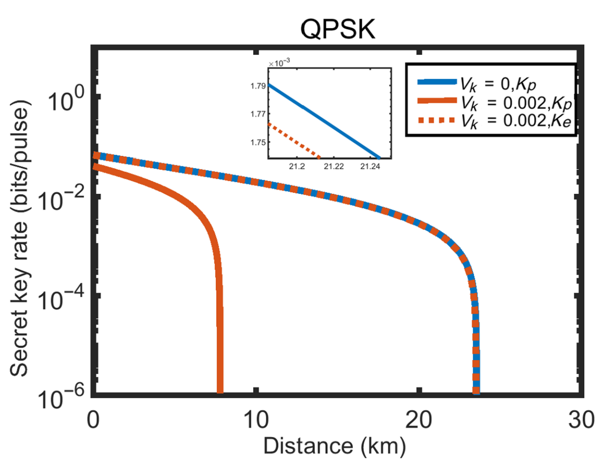

FIG. 5. The estimated and practical secret key rate for the QPSK-modulated protocol against the attack with the fluctuation variance $V_k = 2 \times 10^{-3}$. The modulation variance $V_A = 0.456$, the estimated channel excess noise $\xi_e = 0.01$ and $\beta = 0.95$.

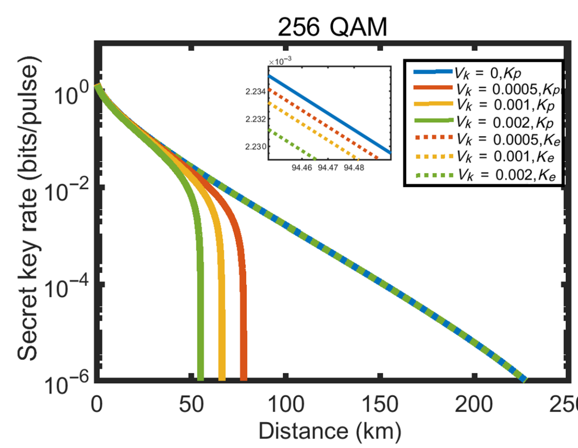

FIG. 7. The estimated and practical secret key rate versus transmission distance for the 256-QAM-modulated protocol against the attack with different fluctuation variances. The modulation variance $V_A = 6.332$, $\nu = 0.039$, the estimated channel excess noise $\xi_e = 0.029$ and $\beta = 0.95$.

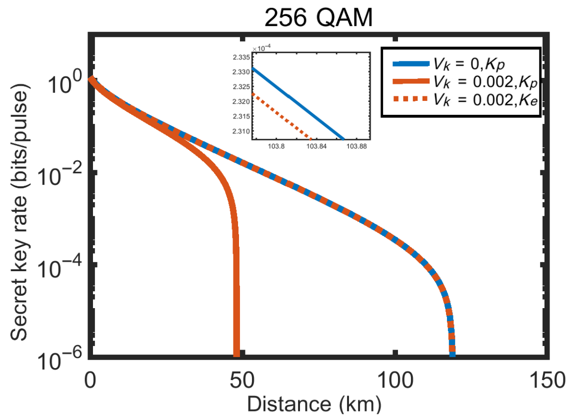

FIG. 8. The estimated and practical secret key rate for the 256-QAM-modulated protocol against the attack with the fluctuation variance $V_k = 2 \times 10^{-3}$. The modulation variance $V_A = 6.332$, $\nu = 0.039$, the estimated channel excess noise $\xi_e = 0.05$ and $\beta = 0.95$.

generated. This demonstrates the permeability of the modified LOIA with random fluctuations under DM CVQKD protocols.

Figure 9 shows the minimum fluctuation variance when Eve could steal all the keys without being detected. We consider the case with the estimated channel excess noise $\xi_e = 0.029$ and $\xi_e = 0.05$, respectively. Eve could mask the attack to steal the secret keys by simply making the LO intensity fluctuate randomly with a very small fluctuation variance.

As can be seen from the above, within the DM CVQKD system under the modified LOIA with random fluctuation, both Alice and Bob overestimate certain areas of the secret key rate so that Eve could hide her attacks on the

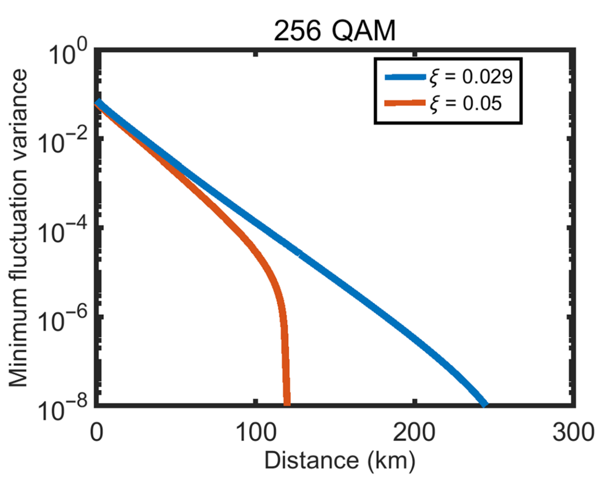

FIG. 9. The minimal fluctuation variance of LO intensity for the 256-QAM-modulated protocol where Eve could acquire all keys without being detected. The modulation variance $V_A = 6.332$, $\nu = 0.039$, and $\beta = 0.95$.

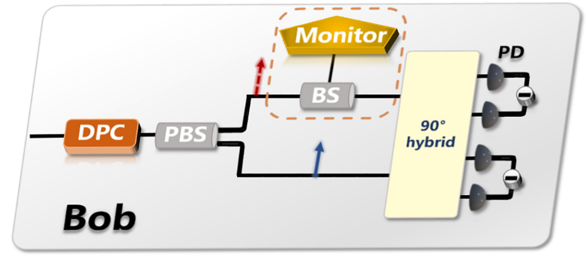

FIG. 10. The practical setup of the real-time monitoring of the LO intensity. It consists of monitoring the intensity of each LO pulse by splitting a small part of the LO pulses and calibrating the instantaneous intensity. The dotted line diagram depicts the schematic diagram of the LO monitoring module.

signal pulse. The effect of the attack is closely related to the fluctuation variance of LO intensity. The greater the fluctuation variance, the more the communication parties overestimate the secret key rate. As long as Eve chooses the appropriate fluctuation variance, she can obtain all the key information without being detected. The attack effect is more pronounced at high channel noise and long range.

V. DISCUSSION

In the practical environment, it has been analyzed that the LO intensity level during key distribution process in the presence of Eve's manipulation cannot be determined directly using the initial calibrated LO intensity. Moreover, in the modified LOIA model with random fluctuations, the statistical mean of LO intensity is consistent with the calibrated LO intensity value. Traditional commonly used defense countermeasures, including monitoring the mean intensity and optical power of LO or monitoring shot noise will fail. We must carefully monitor the instantaneous fluctuations in LO intensity and the corresponding countermeasure is to design modules that allow the real-time monitoring of LO intensity. Figure 10 gives a schematic setup of the countermeasure. It involves the LO intensity monitoring module, which splits a small portion with an asymmetric splitter and monitors the instantaneous value of LO intensity. Then communication parties use the monitored LO intensity to scale the environment measurements. If the intensity of each LO pulse is calibrated to obtain the outcomes during the key distribution, Eve's attack on LO can be avoided, making the practical DM CVQKD system more robust. However, the countermeasure requires a higher degree of accuracy for the real-time monitoring technology. Moreover, we think that other schemes are also effective, such as the local local oscillator (LLO) CVQKD system, the measurement-device-independent (MDI) QKD protocol and so on. But the introduced security vulnerabilities and the difficulty of experimental implementation should also be carefully considered.

In a previous LOIA model, Eve hides the excess noise introduced by her eavesdropping behavior by overall attenuating the LO intensity. In this paper, we focus on the random fluctuating property of LO intensity and propose a modified LOIA model with random fluctuations with invariant mean value. The manipulation of LO intensity avoids detection by LO intensity and power monitoring technology or shot-noise monitoring technology. The theoretical parameter estimation model constructed in a practical DM CVQKD system suggests that the underestimation of excess noise leads to key leakage without being detected by both communicating parties. From the simulation results, it is found that the fluctuation variance V_k affects the excess noise and key rate estimated by both communicating parties greatly. If the fluctuating variance is relatively large, Eve can secretly steal all the keys, thus seriously compromising system security.

Similarly, not only the transmitted local oscillator (TLO) CVQKD system, the modified attack model also poses a serious challenge to the security of the pilot in the LLO CVQKD system. Traditionally, the communication parties use the mean intensity of the pilot to quantify the trusted phase noise, and then calculate the system secret key rate. Eve could exploit the loophole of the random fluctuations in pilot intensity to reduce the trusted phase noise [50,51], thereby hiding her attack on signal pulses.

VI. CONCLUSION

In conclusion, we propose a modified LO intensity attack with random fluctuations that can successfully mask Eve's eavesdropping behavior and evade commonly used monitoring technologies. To analyze its attack effect, modified LO intensity attack is carried out to observe the parameter estimation of semidefinite program modeling against a practical discrete-modulated CVQKD system. The theoretical simulation results under the Gaussian channel show that even very small random fluctuations in modified LO intensity attack will cause no secure keys to be generated. The stealthy quantum hacking can also be launched in the practical LLO CVQKD system with random fluctuations in pilot intensity. Therefore, in a realistic CVQKD system, the LO intensity should be well monitored and stabilized. The modified LO intensity attack model with random fluctuations in this paper places greater demands on current real-time intensity monitoring techniques. The related study will be of great significance to the practical security analysis of CVQKD.

ACKNOWLEDGMENTS

This research was supported by the National Natural Science Foundation of China (62001044), the Basic Research Program of China (JCKY2021210B059), the Equipment Advance Research Field Foundation (315067206), and the Fund of State Key Laboratory of Information Photonics and Optical Communications (IPOC2021ZT02).

[1] C. H. Bennett and G. Brassard, Proceedings of the IEEE International Conference on Computers, Systems and Signal Processing (1984).

[2] S. Pirandola, U. L. Andersen, L. Banchi, M. Berta, D. Bunandar, R. Colbeck, D. Englund, T. Gehring, C. Lupo, and C. Ottaviani, *et al.*, Advances in quantum cryptography, Adv. Opt. Photonics **12**, 1012 (2020).

[3] F. Xu, X. Ma, Q. Zhang, H. Lo, and J. Pan, Secure quantum key distribution with realistic devices, Rev. Mod. Phys. **92**, 025002 (2020).

[4] C. Weedbrook, S. Pirandola, R. García-Patrón, N. J. Cerf, T. C. Ralph, J. H. Shapiro, and S. Lloyd, Gaussian quantum information, Rev. Mod. Phys. **84**, 621 (2012).

[5] E. Diamanti and A. Leverrier, Distributing secret keys with quantum continuous variables: Principle, security and implementations, Entropy **17**, 6072 (2015).

[6] H. Guo, Z. Li, S. Yu, and Y. Zhang, Toward practical quantum key distribution using telecom components, Fundam. Res. **1**, 96 (2021).

[7] F. Karinou, H. H. Brunner, C.-H. F. Fung, L. C. Comandar, S. Bettelli, D. Hillerkuss, M. Kuschnerov, S. Mikroulis, D. Wang, and C. Xie, *et al.*, Toward the integration of CV quantum key distribution in deployed optical networks, IEEE Photonics Technol. Lett. **30**, 650 (2018).

[8] Y. Zhang, Z. Li, Z. Chen, C. Weedbrook, Y. Zhao, X. Wang, Y. Huang, C. Xu, X. Zhang, and Z. Wang, *et al.*, Continuous-variable QKD over 50 km commercial fiber, Quantum Sci. Technol. **4**, 035006 (2019).

[9] H. Wang, Y. Li, Y. Pi, Y. Pan, Y. Shao, L. Ma, Y. Zhang, J. Yang, T. Zhang, and W. Huang, *et al.*, Sub-GBPS key rate four-state continuous-variable quantum key distribution within metropolitan area, Commun. Phys. **5**, 1 (2022).

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

[11] G. Zhang, J. Y. Haw, H. Cai, F. Xu, S. Assad, J. F. Fitzsimons, X. Zhou, Y. Zhang, S. Yu, and J. Wu, *et al.*, An integrated silicon photonic chip platform for continuous-variable quantum key distribution, Nat. Photonics **13**, 839 (2019).

[12] T. A. Eriksson, T. Hirano, B. J. Puttnam, G. Rademacher, R. S. Luís, M. Fujiwara, R. Namiki, Y. Awaji, M. Takeoka, and N. Wada, *et al.*, Wavelength division multiplexing of continuous variable quantum key distribution and 18.3 tbit/s data channels, Commun. Phys. **2**, 1 (2019).

[13] Y. Zhang, Z. Chen, S. Pirandola, X. Wang, C. Zhou, B. Chu, Y. Zhao, B. Xu, S. Yu, and H. Guo, Long-Distance Continuous-Variable Quantum Key Distribution over 202.81 km of Fiber, Phys. Rev. Lett. **125**, 010502 (2020).

[14] P. Huang, T. Wang, R. Chen, P. Wang, Y. Zhou, and G. Zeng, Experimental continuous-variable quantum key distribution using a thermal source, New J. Phys **23**, 113028 (2021).

[15] M. Zhang, P. Huang, P. Wang, S. Wei, and G. Zeng, Experimental free-space continuous-variable quantum key distribution with thermal source, Opt. Lett. **48**, 1184 (2023).

[16] Y. Bian, Y. Zhang, C. Zhou, S. Yu, Z. Li, and H. Guo, High-rate point-to-multipoint quantum key distribution using coherent states, preprint ArXiv:2302.02391 (2023).

[17] F. Grosshans and P. Grangier, Continuous Variable Quantum Cryptography using Coherent States, Phys. Rev. Lett. **88**, 057902 (2002).

[18] C. Weedbrook, A. M. Lance, W. P. Bowen, T. Symul, T. C. Ralph, and P. K. Lam, Quantum Cryptography without Switching, Phys. Rev. Lett. **93**, 170504 (2004).

[19] M. Navascués, F. Grosshans, and A. Acin, Optimality of Gaussian Attacks in Continuous-Variable Quantum Cryptography, Phys. Rev. Lett. **97**, 190502 (2006).

[20] R. García-Patrón and N. J. Cerf, Unconditional Optimality of Gaussian Attacks Against Continuous-Variable Quantum Key Distribution, Phys. Rev. Lett. **97**, 190503 (2006).

[21] S. Pirandola, S. L. Braunstein, and S. Lloyd, Characterization of Collective Gaussian Attacks and Security of Coherent-State Quantum Cryptography, Phys. Rev. Lett. **101**, 200504 (2008).

[22] Z. Li, Y. Zhang, F. Xu, X. Peng, and H. Guo, Continuous-variable measurement-device-independent quantum key distribution, Phys. Rev. A **89**, 052301 (2014).

[23] Y. Zhang, Z. Li, S. Yu, W. Gu, X. Peng, and H. Guo, Continuous-variable measurement-device-independent quantum key distribution using squeezed states, Phys. Rev. A **90**, 052325 (2014).

[24] S. Pirandola, C. Ottaviani, G. Spedalieri, C. Weedbrook, S. L. Braunstein, S. Lloyd, T. Gehring, C. S. Jacobsen, and U. L. Andersen, High-rate measurement-device-independent quantum cryptography, Nat. Photonics **9**, 397 (2015).

[25] Y. Tian, P. Wang, J. Liu, S. Du, W. Liu, Z. Lu, X. Wang, and Y. Li, Experimental demonstration of continuous-variable measurement-device-independent quantum key distribution over optical fiber, Optica **9**, 492 (2022).

[26] N. Jain, H.-M. Chin, H. Mani, C. Lupo, D. S. Nikolic, A. Kordts, S. Pirandola, T. B. Pedersen, M. Kolb, and B. Ömer, *et al.*, Practical continuous-variable quantum key distribution with composable security, Nat. Commun. **13**, 4740 (2022).

[27] Y. Pi, H. Wang, Y. Pan, Y. Shao, Y. Li, J. Yang, Y. Zhang, W. Huang, and B. Xu, Sub-MBPS key-rate continuous-variable quantum key distribution with local local oscillator over 100-km fiber, Opt. Lett. **48**, 1766 (2023).

[28] A. Leverrier and P. Grangier, Unconditional Security Proof of Long-Distance Continuous-Variable Quantum Key Distribution with Discrete Modulation, Phys. Rev. Lett. **102**, 180504 (2009).

[29] Z. Li, Y. Zhang, and H. Guo, User-defined quantum key distribution, preprint ArXiv:1805.04249 (2018).

[30] T. Matsuura, K. Maeda, T. Sasaki, and M. Koashi, Finite-size security of continuous-variable quantum key distribution with digital signal processing, Nat. Commun. **12**, 252 (2021).

[31] Y. Pan, H. Wang, Y. Shao, Y. Pi, Y. Li, B. Liu, W. Huang, and B. Xu, Experimental demonstration of high-rate discrete-modulated continuous-variable quantum key distribution system, Opt. Lett. **47**, 3307 (2022).

[32] F. Roumestan, A. Ghazisaeidi, J. Renaudier, L. T. Vidarte, A. Leverrier, E. Diamanti, and P. Grangier, Experimental demonstration of discrete modulation formats for continuous variable quantum key distribution, preprint ArXiv:2207.11702 (2022).

[33] Y. Zhao, M. Heid, J. Rigas, and N. Lütkenhaus, Asymptotic security of binary modulated continuous-variable quantum key distribution under collective attacks, Phys. Rev. A **79**, 012307 (2009).

[34] K. Brádler and C. Weedbrook, Security proof of continuous-variable quantum key distribution using three coherent states, Phys. Rev. A **97**, 022310 (2018).

[35] S. Ghorai, P. Grangier, E. Diamanti, and A. Leverrier, Asymptotic Security of Continuous-Variable Quantum Key Distribution with a Discrete Modulation, Phys. Rev. X **9**, 021059 (2019).

[36] J. Lin, T. Upadhyaya, and N. Lütkenhaus, Asymptotic Security Analysis of Discrete-Modulated Continuous-Variable Quantum Key Distribution, Phys. Rev. X **9**, 041064 (2019).

[37] J. Lin and N. Lütkenhaus, Trusted Detector Noise Analysis for Discrete Modulation Schemes of Continuous-Variable Quantum Key Distribution, Phys. Rev. Appl. **14**, 064030 (2020).

[38] A. Denys, P. Brown, and A. Leverrier, Explicit asymptotic secret key rate of continuous-variable quantum key distribution with an arbitrary modulation, Quantum **5**, 540 (2021).

[39] W. Liu, C. Li, Y. Xie, C. Weng, J. Gu, X. Cao, Y. Lu, B. Li, H. Yin, and Z. Chen, Homodyne Detection Quadrature Phase Shift Keying Continuous-Variable Quantum Key Distribution with High Excess Noise Tolerance, PRX Quantum **2**, 040334 (2021).

[40] D. Pereira, M. Almeida, M. Facão, A. N. Pinto, and N. A. Silva, Probabilistic shaped 128-APSK CV-QKD transmission system over optical fibres, Opt. Lett. **47**, 3948 (2022).

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

[42] X. Ma, S. Sun, M. Jiang, and L. Liang, Local oscillator fluctuation opens a loophole for Eve in practical continuous-variable quantum-key-distribution systems, Phys. Rev. A **88**, 022339 (2013).

[43] J. Huang, C. Weedbrook, Z. Yin, S. Wang, H. Li, W. Chen, G. Guo, and Z. Han, Quantum hacking of a continuous-variable quantum-key-distribution system using a wavelength attack, Phys. Rev. A **87**, 062329 (2013).

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

[45] Y. Zhao, Y. Zhang, Y. Huang, B. Xu, S. Yu, and H. Guo, Polarization attack on continuous-variable quantum key distribution, J. Phys. B. **52**, 015501 (2018).

[46] X. Ma, S. Sun, M. Jiang, M. Gui, Y. Zhou, and L. Liang, Enhancement of the security of a practical continuous-variable quantum-key-distribution system by manipulating the intensity of the local oscillator, Phys. Rev. A **89**, 032310 (2014).

[47] G. Van Assche, J. Cardinal, and N. J. Cerf, Reconciliation of a quantum-distributed Gaussian key, IEEE Trans. Inf. **50**, 394 (2004).

[48] F. Roumestan, Ph.D. thesis, Sorbonne Université (2022).

[49] L. Ma, J. Yang, T. Zhang, Y. Shao, J. Liu, Y. Luo, H. Wang, W. Huang, F. Fan, and C. Zhou, *et al.*, Practical continuous-variable quantum key distribution with feasible optimization parameters, Sci. China Inf. Sci. **66**, 1 (2023).

[50] Y. Shao, H. Wang, Y. Pi, W. Huang, Y. Li, J. Liu, J. Yang, Y. Zhang, and B. Xu, Phase noise model for continuous-variable quantum key distribution using a local local oscillator, Phys. Rev. A **104**, 032608 (2021).

[51] Y. Shao, Y. Li, H. Wang, Y. Pan, Y. Pi, Y. Zhang, W. Huang, and B. Xu, Phase-reference-intensity attack on continuous-variable quantum key distribution with a local local oscillator, Phys. Rev. A **105**, 032601 (2022).