Estimation of output-channel noise for continuous-variable quantum key distribution

Oliver Thearle,* Syed M. Assad, and Thomas Symul
Centre for Quantum Computation and Communication Technology, Department of Quantum Science, Research School of Physics and Engineering, The Australian National University, Canberra, ACT 0200, Australia
(Received 21 August 2015; published 28 April 2016)

Estimation of channel parameters is important for extending the range and increasing the key rate of continuous-variable quantum key distribution protocols. We propose an estimator for the channel noise parameter based on the method-of-moments. The method-of-moments finds an estimator from the moments of the output distribution of the protocol. This estimator has the advantage of being able to use all of the states shared between Alice and Bob. Other estimators are limited to a smaller publicly revealed subset of the states. The proposed estimator has a lower variance for the high-loss channel than what has previously been proposed. We show that the method-of-moments estimator increases the key rate by up to an order of magnitude at the maximum transmission of the protocol.

DOI: 10.1103/PhysRevA.93.042343

I. INTRODUCTION

Quantum key distribution (QKD) was proposed in 1984 [1] as a solution to the key distribution problem. In this problem Alice wants to share a secret key with a remote party, Bob, but she only has public channels available to her. Alice can solve this problem by encoding the key into quantum states. Alice and Bob can then use quantum mechanics to prove their secret is unconditionally secure from an eavesdropping adversary, Eve [2]. The key can then be used later for cryptographic purposes. Continuous-variable (CV) QKD uses the quadrature modulations and measurements of phase and amplitude from a bright laser to distribute the shared secret [3–5]. The coherent-state protocol with homodyne detection [4] is commonly used to study CV QKD [6,7]. In this protocol Alice sends Bob a series of randomly displaced vacuum states through an unsecured channel. Bob then measures the received states by either switching between quadratures or measuring both simultaneously. Alice and Bob then estimate a bound on the maximum information that may have been intercepted by Eve in the channel. The optimal attack Eve can make on this protocol has been shown to be a Gaussian collective attack [8]. This attack assumes that Eve has access to all information lost in the channel. A bound on Eve's information can be found as a function channel transmission T and excess channel noise relative to the input ξ. In a practical CV QKD protocol these parameters must be estimated from the shared secret between Alice and Bob. This ensures that the correct bound is found for the final secret key. This is currently done by Alice and Bob publicly revealing a random subset of their initial shared secret [6,9]. The number of states revealed can be optimized to give an optimal key rate, as discussed in Refs. [9,10]. An alternative bound on Eve's information can be found by directly estimating the covariance matrix of the shared secret after reconciliation [11].

In this paper we build on some of the ideas presented in Ref. [9]. The authors propose a way of estimating the two parameters by modeling the protocol using a classical loss channel with additive Gaussian noise:

$$y_i = t x_i + z_i \quad i = 1,2,\ldots,N. \tag{1}$$

Here x_i is the data sent by Alice, y_i is Bob's measurement data, z_i is a Gaussian noise term with variance $\sigma^2 = 1 + T\xi$ and mean zero, and $t = \sqrt{T}$ is the amplitude transmission. This model is well understood, and a maximum-likelihood estimator (MLE) exists for both parameters t and σ^2. The authors then use these estimators to find the worst case for excess noise transmission to find the final key rate.

We propose to use the method-of-moments (MM) in conjunction with the MLE for t to find an alternative estimator for σ^2. This estimator has a lower variance than the MLE as $T \to 0$. Generally, this method produces estimators that are typically worse in terms of minimizing variance and bias than other estimation methods. However, it has the advantage of only requiring the public exchange of moments rather than sacrificing part of the key for parameter estimation. The estimator we use in this paper is a function of Alice and Bob's variance and transmission estimates. The variances can be estimated and revealed by Alice and Bob individually using the whole shared secret. Estimation of the transmission still requires some of the shared secret to be revealed. Sharing the variance allows the estimator to use more of the accessible information to decrease the variance of the MM estimator $\hat{\sigma}^2_{\mathrm{MM}}$ without sacrificing more of the shared secret. The variance of the estimator can be further improved by creating a linear combination of the MM estimator and the MLE. The resulting estimator is the optimum of the two.

This paper is organized as follows. Section II covers prior work on noise estimation and explains the coherent-state CV QKD protocol in more detail. In Sec. III we explain the method-of-moments and show how we arrive at the two estimators. The variance of the MM estimators is then found and compared in Sec. IV to other CV QKD noise estimators and shown to be asymptotically unbiased, and in Sec. V we conclude the paper with a discussion of the MM estimator and its effects on the final key rate.

II. THE PROTOCOL AND MODEL

In this paper we will consider the coherent-state protocol with homodyne detection [4]. In this protocol Alice prepares N displaced vacuum states $|q_i + ip_i\rangle$ through phase and amplitude quadrature modulation. The displacements q_i and p_i are both random variables sampled from the normal

*oliver.thearle@anu.edu.au

distribution $\mathcal{N}(0, V_A)$. These states are transmitted to Bob through an unsecure channel with transmission T and excess noise ξ. The channel is assumed to be under the control of Eve. Bob will then measure the received states using a homodyne detector switching randomly between the phase and amplitude quadratures. In reference to (1) we will take x_i as Alice's modulations and y_i as Bob's measurement outcomes. When Alice uses modulation to prepare the states for Bob, it is known as a prepare-and-measure protocol. These protocols have been shown to have an entanglement-based equivalent [12] which is used for the security analysis [13].

Eve's optimal attack with finite-size effects has been shown to be a Gaussian attack [14]. The collective state between Alice and Bob can be assumed to be Gaussian. For the entanglement-based protocol it can be described by the covariance matrix

$$\Gamma = \begin{pmatrix} (V_A + 1)\mathbb{I}_2 & \sqrt{T\left(V_A^2 + 2V_A\right)}\sigma_z \\ \sqrt{T\left(V_A^2 + 2V_A\right)}\sigma_z & (TV_A + 1 + T\xi)\mathbb{I}_2 \end{pmatrix}, \tag{2}$$

where σ_z is the Pauli matrix

$$\sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{3}$$

In a prepare-and-measure scheme Alice and Bob want to find the covariance matrix for the equivalent entanglement-based protocol. To do this they reveal a subset of $m < N$ states for estimating the parameters t and σ. Using the channel model (1) for the protocol, we have the maximum-likelihood estimators [9]:

$$\hat{t} = \frac{\sum_{i=1}^{m} x_i y_i}{\sum_{i=1}^{m} x_i^2}, \quad \hat{\sigma}^2_{\mathrm{MLE}} = \frac{1}{m}\sum_{i=1}^{m}(y_i - \hat{t}x_i)^2. \tag{4}$$

The distributions of these estimators are

$$\hat{t} \sim \mathcal{N}\left(t, \frac{\sigma^2}{\sum_{i=1}^{m} x_i^2}\right), \qquad \frac{m\hat{\sigma}^2_{\mathrm{MLE}}}{\sigma^2} \sim \chi^2(m-1). \tag{5}$$

As described in Ref. [9], the estimates are then used to find the worst case for t and σ^2, that is, the minimum of t and the maximum of σ^2 in the confidence interval $1 - \epsilon_{\mathrm{PE}}$. The parameter ϵ_{PE} is the probability that the parameter estimation failed (typically, $\epsilon_{\mathrm{PE}} = 10^{-10}$). Using the theoretical distributions in (5), the worst-case estimators can then be written as

$$t_{\min} \approx \hat{t} - z_{\epsilon_{\mathrm{PE}}/2}\mathrm{SD}(\hat{t}), \tag{6}$$

$$\sigma^2_{\max} \approx \hat{\sigma}^2_{\mathrm{MLE}} + z_{\epsilon_{\mathrm{PE}}/2}\mathrm{SD}\left(\hat{\sigma}^2_{\mathrm{MLE}}\right). \tag{7}$$

Here SD is the standard deviation function, and $z_{\epsilon_{\mathrm{PE}}/2} = \mathrm{erf}^{-1}(1 - \epsilon_{\mathrm{PE}}/2)$, where $\mathrm{erf}(x)$ is the error function. We can rewrite (2) for the worst-case noise and transmission,

$$\Gamma_{\epsilon_{\mathrm{PE}}} = \begin{pmatrix} (V_A + 1)\mathbb{I}_2 & t_{\min}\sqrt{V_A^2 + 2V_A}\sigma_z \\ t_{\min}\sqrt{V_A^2 + 2V_A}\sigma_z & \left(t^2_{\min}V_A + \sigma^2_{\max}\right)\mathbb{I}_2 \end{pmatrix}. \tag{8}$$

Another proposed estimator from Ref. [10] uses a second modulation transmitted with the key to assist the estimation of the channel parameters. This assumes the second modulation will experience the same channel as the modulation used for the final key. For the protocol analyzed in their paper, Alice sends Bob squeezed displaced vacuum states with a squeezed quadrature variance of V_S. By setting $V_S = 1$ the protocol becomes the coherent-state protocol. The parameters they estimate are the channel transmission T and the excess noise relative to the output $V_\xi = T\xi$. Thanks to the second modulation this estimator is able to use N states for the key and parameter estimation,

$$\hat{T} = \frac{\left(\sum_{i=1}^{N} x_{\mathrm{M2},i}\, y_i\right)^2}{(N V_{\mathrm{M2}})^2}, \tag{9}$$

$$\hat{V}_\xi = \frac{1}{N}\sum_{i=1}^{N}\left(y_i - \sqrt{\hat{T}}x_{\mathrm{M2},i}\right)^2 - \hat{T}V_A - 1, \tag{10}$$

where $x_{\mathrm{M2},i}$ is the displacement of the second modulation from Alice. These estimators were shown to be asymptotically unbiased and to have the variances

$$\mathrm{Var}(\hat{T}) = \frac{4}{N}T^2\left(2 + \frac{V_N}{T V_{\mathrm{M2}}}\right), \tag{11}$$

$$\mathrm{Var}(\hat{V}_\xi) = \frac{2}{N}V_N^2 + V_A^2\mathrm{Var}(\hat{T}), \tag{12}$$

where $V_N = 1 + V_\xi + TV_A$ and V_{M2} is the variance of the second modulation. The authors then suggests using the linear combination in Eq. (13) to find the optimal estimators T^{opt} and V_ξ^{opt} at a high channel transmission,

$$\hat{\theta}_{\mathrm{opt}} = \alpha\hat{\theta}_1 + (1 - \alpha)\hat{\theta}_2, \tag{13}$$

where $\hat{\theta}_1$ and $\hat{\theta}_2$ are two different estimators for either V_ξ or T. The optimum value of α to achieve a minimum variance from two estimators with a covariance of zero is given by

$$\alpha = \frac{\mathrm{Var}(\hat{\theta}_2)}{\mathrm{Var}(\hat{\theta}_1) + \mathrm{Var}(\hat{\theta}_2)}. \tag{14}$$

This can be found by minimizing $\mathrm{Var}(\hat{\theta}_{\mathrm{opt}})$ with respect to α. A derivation of α is shown in Appendix C. The variance of $\hat{\theta}_{\mathrm{opt}}$ is then given by

$$\mathrm{Var}(\hat{\theta}_{\mathrm{opt}}) = \frac{\mathrm{Var}(\hat{\theta}_1)\mathrm{Var}(\hat{\theta}_2)}{\mathrm{Var}(\hat{\theta}_1) + \mathrm{Var}(\hat{\theta}_2)}. \tag{15}$$

By construction $\hat{\theta}_{\mathrm{opt}}$ will have a variance less than or equal to the variance of estimators $\hat{\theta}_1$ and $\hat{\theta}_2$. The linear combination will also preserve the bias properties of the two estimators. Once the channel parameters are estimated, Alice and Bob will select an appropriate reconciliation protocol and correct the remaining $n = N - m$ states for errors. In this paper we will only consider reverse reconciliation [4]. Alice and Bob then hash their raw secret key to produce an information-theoretically secure final key [6].

The asymptotic key rate for the coherent-state protocol with reverse reconciliation is bounded by [6]

$$K \geqslant I(x : y) - S(y : E), \tag{16}$$

where $I(x : y)$ is the mutual information between Alice and Bob and $S(E : y)$ is the mutual information Eve has with Bob. Both of these terms can be calculated from the channel parameters. This bound can be rewritten to include the effects

for reconciliation efficiency β, parameter estimation $\frac{n}{N}$, and ϵ_{PE} on a finite key [9],

$$K = \frac{n}{N}[\beta I(x:y) - S_{\epsilon_{\mathrm{PE}}}(y:E)], \tag{17}$$

where $S_{\epsilon_{\mathrm{PE}}}(y:E)$ is calculated from the worst-case estimates of our channel parameters. The reconciliation efficiency β is related to the amount of information Alice and Bob must sacrifice in order to perform this step. For a given transmission and noise of a channel the choice of reconciliation protocol can be optimized to maximize β [15].

III. THE METHOD-OF-MOMENTS ESTIMATOR

The estimators in Eq. (4) are found by maximizing the log-likelihood probability function $\ln p(x_i,y_i;\sigma^2,t,V_{\mathrm{A}})$. An alternative is to use the method-of-moments [16] to find the estimators. The method-of-moments is a simple way to find an estimator, but it has no optimality properties. It performs best with a long data record, which makes it suitable to CV QKD as, typically, the data record is $>10^8$ [7]. To use the method we first find a probability distribution describing our observations in terms of the parameters we want to estimate. In the case of Bob's measurements the distribution is given by $\mathcal{N}(0,t^2V_{\mathrm{A}}+\sigma^2)$. The moments of this distribution can then be solved as a system of equations for the parameters we want to estimate. As Bob's data are normally distributed around zero, the first moment will be zero, and the second moment is given by the variance,

$$\sigma_{\mathrm{B}}^2 = t^2V_{\mathrm{A}} + \sigma^2. \tag{18}$$

All other moments for this distribution will be a function of σ_{B}^2 giving only one independent nonzero moment. This allows us to only find one estimator. We are most interested in maximizing the key rate for long-distance CV QKD. The limiting factor for protocols with a high loss channel is the excess noise [5]. For this reason we will concentrate on finding a better estimator for the output noise. The variance in Eq. (18) can be used to estimate t, but the process is made more difficult by requiring an estimate of σ^2. Starting with Eq. (18) and substituting the estimator for t and the sample variance for σ_{B}^2, we find an initial estimator for the noise relative to the output,

$$\hat{\sigma}_{\mathrm{mm}}^2 = \hat{\sigma}_{\mathrm{B}}^2 - \hat{t}^2V_{\mathrm{A}}, \tag{19}$$

where $\hat{\sigma}_{\mathrm{B}}^2$ is given by $\frac{1}{N}\sum y_i^2$. To use this estimator Alice and Bob can publicly reveal V_{A} and $\hat{\sigma}_{\mathrm{B}}^2$ without giving away any more of the shared secret to Eve [9]. We found that by treating V_{A} as an unknown parameter and using the estimate $\hat{\sigma}_{\mathrm{A}}^2 = \frac{1}{N}\sum x_i^2$ in its place the variance of the MM estimator decreased. This improvement comes from increasing the covariance between $\hat{\sigma}_{\mathrm{B}}^2$ and $\hat{t}^2\hat{\sigma}_{\mathrm{A}}^2$ and is demonstrated by the following property of variance:

$$\mathrm{Var}\big(\hat{\sigma}_{\mathrm{B}}^2 - \hat{t}^2\hat{\sigma}_{\mathrm{A}}^2\big) = \mathrm{Var}\big(\hat{\sigma}_{\mathrm{B}}^2\big) + \mathrm{Var}\big(\hat{t}^2\hat{\sigma}_{\mathrm{A}}^2\big) - 2\mathrm{Cov}\big(\hat{\sigma}_{\mathrm{B}}^2,\hat{t}^2\hat{\sigma}_{\mathrm{A}}^2\big). \tag{20}$$

Substituting $\hat{\sigma}_{\mathrm{A}}$, we arrive at our final MM estimator,

$$\hat{\sigma}_{\mathrm{MM}}^2 = \hat{\sigma}_{\mathrm{B}}^2 - \hat{t}^2\hat{\sigma}_{\mathrm{A}}^2. \tag{21}$$

The variances of $\hat{\sigma}_{\mathrm{mm}}^2$ and the estimator $\hat{\sigma}_{\mathrm{MM}}^2$ are compared in Appendix A. Using Eq. (20), we can already see the improvement in the variance of the estimator $\hat{\sigma}_{\mathrm{MM}}^2$ will have over $\hat{\sigma}_{\mathrm{MLE}}^2$ as the transmission approaches zero for a fixed value of V_{A} and m. The variance of these estimators with $t=0$ are given by

$$\mathrm{Var}(\hat{\sigma}_{\mathrm{MM}}^2) = \frac{2\sigma_{\mathrm{B}}^4}{N} \qquad \mathrm{Var}(\hat{\sigma}_{\mathrm{MLE}}^2) = \frac{2\sigma_{\mathrm{B}}^4}{m}. \tag{22}$$

When compared we find $\mathrm{Var}(\hat{\sigma}_{\mathrm{MM}}^2)$ is better by a factor of $\frac{m}{N}$.

An interesting point is $\hat{\sigma}_{\mathrm{MM}}^2 = \hat{\sigma}_{\mathrm{MLE}}^2$ when both estimators are used on the N transmitted states, such as the case at the range limit of a protocol where we reveal almost all of the states for parameter estimation for a positive key. Using Eq. (4) on the N transmitted states, we find

$$\hat{\sigma}_{\mathrm{MLE}}^2 = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{t}x_i)^2 \tag{23}$$

$$= \frac{1}{N}\sum_{i=1}^{N}y_i^2 - \frac{1}{N}\frac{\left(\sum_{i=1}^{N}x_iy_i\right)^2}{\sum_{i=1}^{N}x_i^2} \tag{24}$$

$$= \hat{\sigma}_{\mathrm{B}}^2 - \hat{t}^2\hat{\sigma}_{\mathrm{A}}^2 \tag{25}$$

$$= \hat{\sigma}_{\mathrm{MM}}^2. \tag{26}$$

With this result we can show $\hat{\sigma}_{\mathrm{MM}}^2$ from Eq. (21) is a combination of two estimators. By ordering the exchanged states into the publicly revealed m subset and the secret $n = N - m$ subset we find

$$\hat{\sigma}_{\mathrm{MM}}^2 = \frac{1}{N}\left[\sum_{i=1}^{m}y_i^2 + \sum_{i=m+1}^{N}y_i^2 + \hat{t}^2\left(\sum_{i=1}^{m}x_i^2 + \sum_{i=m+1}^{N}x_i^2\right)\right] \tag{27}$$

$$= \frac{1}{N}\sum_{i=1}^{m}(y_i - \hat{t}x_i)^2 + \frac{1}{N}\left(\sum_{i=m+1}^{N}y_i^2 - \hat{t}^2\sum_{i=m+1}^{N}x_i^2\right) \tag{28}$$

$$= \frac{1}{N}\left(m\hat{\sigma}_{\mathrm{MLE}}^2 + n\hat{\sigma}_{\mathrm{MM}''}^2\right), \tag{29}$$

where $\hat{\sigma}_{\mathrm{MM}''}^2$ is the MM estimator applied to n states and $\hat{\sigma}_{\mathrm{MLE}}^2$ and $\hat{t}$ are applied to m states. In Appendix C we show that $\hat{\sigma}_{\mathrm{MLE}}^2$ and $\hat{\sigma}_{\mathrm{MM}''}^2$ have a covariance of zero given that $\hat{t}$ and $\hat{\sigma}_{\mathrm{MLE}}^2$ are independent [17]. This leads to the next estimator we present in this paper. As in Ref. [10], we can find an optimum linear combination of our two estimators. Using (13), we find an optimum estimate of the noise,

$$\hat{\sigma}_{\mathrm{opt}}^2 = \alpha\hat{\sigma}_{\mathrm{MLE}}^2 + (1-\alpha)\hat{\sigma}_{\mathrm{MM}''}^2. \tag{30}$$

Here α is given in Eq. (14).

IV. PERFORMANCE

For the purposes of CV QKD, it is important to consider the variance and the bias of the parameter estimators. Finding an unbiased estimator with a minimized variance will ultimately lead to an increase in the key rate and secure distance of the protocol.

For the MM estimators, the variance and mean are difficult to find due to the division required for $\hat{t}$. For this paper we use a standard method in uncertainty analysis where the variance is approximated from a first-order Taylor series expansion [16]. Given an estimator $\hat{\theta}$ that is some function of $\mathbf{J} = \{J_1(\boldsymbol{y}), J_2(\boldsymbol{y}), \ldots, J_r(\boldsymbol{y})\}$, where $J_i(\boldsymbol{y})$ is some statistic from the data vector $\boldsymbol{y}$, we find the variance is approximated by

$$\mathrm{Var}[\hat{\theta}(\mathbf{J})] \approx \frac{\partial\hat{\theta}}{\partial\mathbf{J}}\bigg|_{\mathbf{J}=\boldsymbol{\mu}}^{T} \boldsymbol{C}_J \frac{\partial\hat{\theta}}{\partial\mathbf{J}}\bigg|_{\mathbf{J}=\boldsymbol{\mu}}, \tag{31}$$

and the mean is approximated by

$$E(\hat{\theta}(\mathbf{J})) \approx \hat{\theta}(\boldsymbol{\mu}). \tag{32}$$

Here $\boldsymbol{\mu}$ is the expected value of our statistics $\mathbf{J}$, and $\boldsymbol{C}_J$ is the covariance matrix for $\mathbf{J}$. This method assumes that the statistics $\mathbf{J}$ will have a low variance and the estimator $\hat{\theta}$ will be roughly linear around $\boldsymbol{\mu}$. That is, Eqs. (31) and (32) will be the asymptotic variance and mean. To apply this method we rewrite our estimators in terms of the data statistics, $\hat{\sigma}^2_{\mathrm{B}}$, $\hat{\sigma}^2_{\mathrm{A}}$, $\hat{\sigma}_{\mathrm{A'B'}}$, and $\hat{\sigma}^2_{\mathrm{A'}}$. Here $\hat{\sigma}_{\mathrm{A'B'}}$ is the sample covariance. We use A′ and B′ to indicate the statistic was estimated from the m subset of states used for parameter estimation. The estimator $\hat{\sigma}^2_{\mathrm{MM}}$ becomes

$$\hat{\sigma}^2_{\mathrm{MM}} = \hat{\sigma}^2_{\mathrm{B}} - \left(\frac{\hat{\sigma}_{\mathrm{A'B'}}}{\hat{\sigma}^2_{\mathrm{A'}}}\right)^2 \hat{\sigma}^2_{\mathrm{A}}. \tag{33}$$

Here we have written $\hat{t} = \hat{\sigma}_{\mathrm{A'B'}}/\hat{\sigma}^2_{\mathrm{A'}}$. The matrix $\boldsymbol{C}_J$ can be found using the variance of the sample variance and the properties of the covariance and variance functions. The elements of $\boldsymbol{C}_J$ are given in Appendix B 1. Applying Eq. (31), the variance is given by

$$\mathrm{Var}\big(\hat{\sigma}^2_{\mathrm{MM}}\big) \approx \frac{2\sigma^4}{N} + \left(\frac{1}{m} - \frac{1}{N}\right)4t^2\sigma^2 V_{\mathrm{A}}. \tag{34}$$

The final variance in (34) was achieved by making the substitution $E[\hat{\sigma}^2_{\mathrm{A}}] = E[\hat{\sigma}^2_{\mathrm{A'}}] = V_{\mathrm{A}}$, $E[\frac{\hat{\sigma}_{\mathrm{A'B'}}}{\hat{\sigma}^2_{\mathrm{A'}}}] = t$, and $E[\hat{\sigma}^2_{\mathrm{B}}] = t^2 V_{\mathrm{A}} + \sigma^2$. For the estimator $\hat{\sigma}^2_{\mathrm{MM''}}$ we find a similar equation,

$$\mathrm{Var}\big(\hat{\sigma}^2_{\mathrm{MM''}}\big) \approx \frac{2\sigma^4}{n} + \left(\frac{1}{m} + \frac{1}{n}\right)4t^2\sigma^2 V_{\mathrm{A}}. \tag{35}$$

As $\hat{\sigma}^2_{\mathrm{MM''}}$ uses different statistics, we will have a different $\boldsymbol{C}_J$. This is given in Appendix B 2. The variance of the optimal estimator is given by [10]

$$\mathrm{Var}\big(\hat{\sigma}^2_{\mathrm{opt}}\big) = \frac{\mathrm{Var}\big(\hat{\sigma}^2_{\mathrm{MLE}}\big)\mathrm{Var}\big(\hat{\sigma}^2_{\mathrm{MM''}}\big)}{\mathrm{Var}\big(\hat{\sigma}^2_{\mathrm{MLE}}\big) + \mathrm{Var}\big(\hat{\sigma}^2_{\mathrm{MM''}}\big)}. \tag{36}$$

The standard deviation of the estimators $\hat{\sigma}^2_{\mathrm{MM}}$ and $\hat{\sigma}^2_{\mathrm{opt}}$ are plotted as a function of the channel distance in Fig. 1. Finding the expected value of our estimators using Eq. (32) shows the estimator $\hat{\sigma}^2_{\mathrm{MM}}$ is asymptotically unbiased.

We performed a series of 5000 stochastic simulations of the coherent-state protocol using $N = 10^5$. The variance of the estimators from this simulation is shown in Fig. 1 and has good agreement with Eqs. (34) and (36). In practical demonstrations N has been of the order of 10^8 to 10^9 [7].

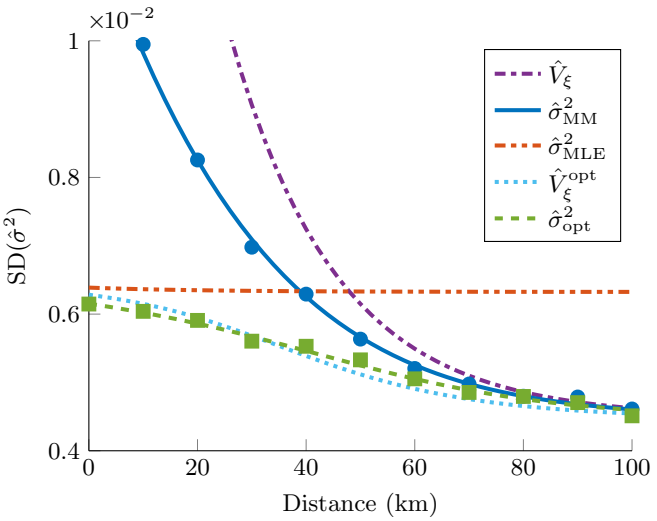

FIG. 1. Plot of the standard deviation of the different noise estimators vs distance in a fiber channel with a loss of 0.2 dB/km: $\hat{V}_\xi$ (dot-dashed line), $\hat{\sigma}^2_{\mathrm{MM}}$ (solid line), $\hat{\sigma}^2_{\mathrm{MLE}}$ (double-dot-dashed line), $\hat{V}^{\mathrm{opt}}_\xi$ (dotted line), and $\hat{\sigma}^2_{\mathrm{opt}}$ (dashed line). A stochastic simulation of the coherent-state protocol was repeated 5000 times to obtain the data points for $\hat{\sigma}^2_{\mathrm{MM}}$ (circles) and $\hat{\sigma}^2_{\mathrm{opt}}$ (squares). The parameters used were $V_A = 3$, $\xi = 0.01$, $m = 0.5\times10^5$, $N = 10^5$, and $V_{\mathrm{M2}} = 10$. The orange double-dot-dashed line is the standard deviation of the MLE with $m = N$ and represents the best estimate Alice and Bob can make of the channel noise using the MLE. The MM estimators and the double-modulation estimators approach this standard deviation as the channel losses increases.

V. DISCUSSION AND CONCLUSION

In this paper we investigated using the method-of-moments to estimate the noise in a linear channel relative to the output for a QKD protocol. The MM estimator allows for Alice and Bob to use better estimates of the variances from the complete shared secret. Using these variances allows the estimators in this paper to approach the performance of the MLE used on the entire shared secret for a high-loss channel.

To simplify our analysis we have assumed that both Alice's modulation and the channel noise are Gaussian with a mean of zero. These assumptions are necessary for finding the variance of $\hat{\sigma}^2_{\mathrm{MM}}$ and $\hat{\sigma}^2_{\mathrm{MLE}}$. The MM estimators do not require the Gaussian assumption as they are estimating the second moment. It is possible that Eve could find a non-Gaussian state that could cause Alice and Bob to underestimate Eve's influence on the channel using the method to find the key rate discussed in this paper. Which state Eve would need to do this was not investigated in this paper.

In a situation where the added noise is non-Gaussian, Alice and Bob should not use the Gaussian approximation for estimating the variance of their estimators. Instead, they should use the general formula for estimating the distribution of the estimators; for example, when N is large, $\mathrm{Var}(\hat{\sigma}^2_{\mathrm{B}})$ in Appendix B 1 should be replaced by $\mu_4/N - \mu_2^2/N$, where μ_k is the kth moment of Bob's measurements [18].

When making a comparison with other estimation methods in Fig. 1, we find that the method-of-moments-based estimators are comparable in performance to the double-modulation

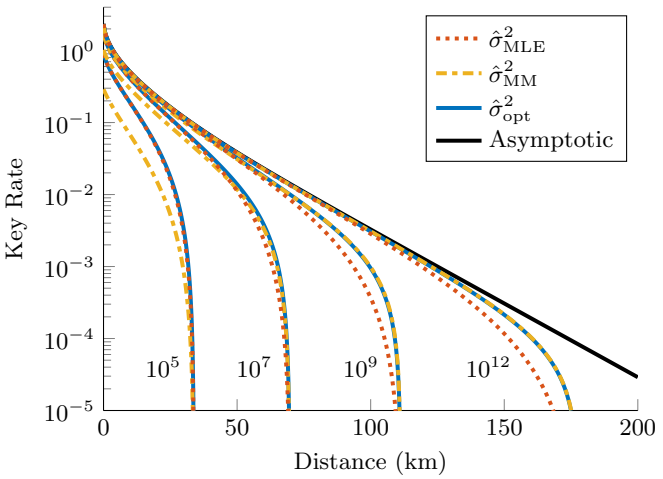

FIG. 2. Plot of the key rate with finite key effects related to parameter estimation. The values V_{A} and m have been optimized with $\xi = 0.01$ and $\beta = 0.95$ to maximize the key rate using $\hat{\sigma}^2_{\mathrm{MLE}}$ (dotted line), $\hat{\sigma}^2_{\mathrm{MM}}$ (dot-dashed line), and $\hat{\sigma}^2_{\mathrm{opt}}$ (solid line) to estimate the excess noise for (from left to right) $N = 10^5$, $N = 10^7$, $N = 10^9$, and $N = 10^{12}$. The asymptotic key rate with V_{A} optimized is also plotted (black solid line). As expected, the maximum distance increases with the size of N. We see that the optimal estimator outperforms the MLE and MM estimator. As with Fig. 1, we find the MM estimator is worse than the MLE at low channel loss but is better for a lossy channel.

method in Ref. [10] without requiring extra modulations and have an improved performance over the MLE for high-loss channels. We can see the result of the improvement in Fig. 2, where the MM-estimator-based key rate is higher than when the MLE is used and the optimum estimator always produces the best key rate. It is interesting to note that the estimators discussed in this paper will never increase the maximum distance for a QKD protocol. The reason for this is shown in Fig. 3, where we see at the maximum transmission distance the optimal m is N and the MLE and MM estimators are equal.

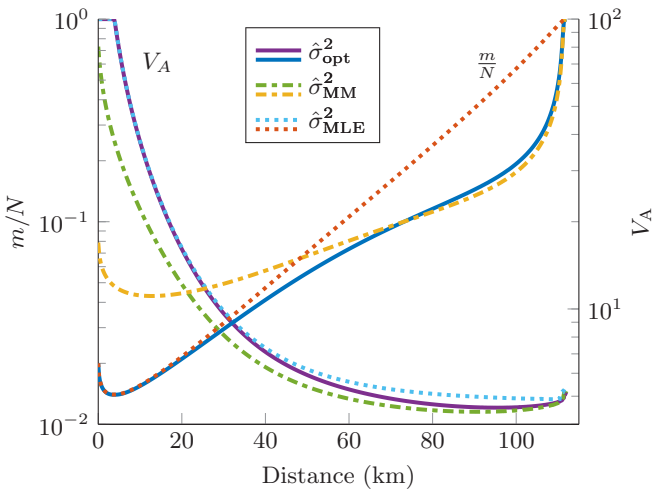

FIG. 3. The optimized values of $\frac{m}{N}$ and V_{A} for the key rates in Fig. 2, where $N = 10^9$ using $\hat{\sigma}^2_{\mathrm{MLE}}$ (dotted line), $\hat{\sigma}^2_{\mathrm{MM}}$ (dot-dashed line), and $\hat{\sigma}^2_{\mathrm{opt}}$ (solid line) to estimate the excess noise. We see that beyond 40 km more states are able to be used in the final key when the optimal or the MM estimator is used.

After reconciliation, Alice and Bob are able to estimate the key-rate bound again but this time with all N measurements. Doing this will give an improved key rate. For high-loss channels this improvement will be mostly due to the improved estimate of the covariance. The MM estimators could be used to determine a rough key rate before the protocol commits to performing the reconciliation step.

With the simplicity of the method-of-moments, this estimator can also be modified to be used with other CV QKD protocols such as the four-state protocol [5] or to include more protocol parameters [19].

ACKNOWLEDGMENTS

We wish to thank to A. Lance and M. James for useful discussions leading to this work. This research is supported by the Australian Research Council (ARC) under the Centre for Quantum Computation and Communication Technology (CE110001027).

APPENDIX A: VARIANCE OF $\hat{\sigma}^2_{\mathrm{mm}}$

Using the same method as described in Sec. IV, we find the variance for $\hat{\sigma}^2_{\mathrm{mm}}$ is given by

$$\mathrm{Var}(\hat{\sigma}^2_{\mathrm{mm}}) \approx \frac{2\sigma^4}{N} + \frac{2t^4V_{\mathrm{A}}^2}{N} + \left(\frac{1}{m} - \frac{1}{N}\right)4t^2\sigma^2V_{\mathrm{A}}. \quad \text{(A1)}$$

Here the covariance $\boldsymbol{C}_J$ can be found using the information in Appendix B 1. With Eq. (A1) we find

$$\mathrm{Var}(\hat{\sigma}^2_{\mathrm{mm}}) = \frac{2t^4V_{\mathrm{A}}^2}{N} + \mathrm{Var}(\hat{\sigma}^2_{\mathrm{MM}}). \quad \text{(A2)}$$

This agrees with our claim that $\mathrm{Var}(\hat{\sigma}^2_{\mathrm{mm}}) > \mathrm{Var}(\hat{\sigma}^2_{\mathrm{MM}})$.

APPENDIX B: ELEMENTS OF $\boldsymbol{C}_J$

1. $\boldsymbol{C}_J$ for $\hat{\sigma}^2_{\mathrm{MM}}$

The diagonal terms for the covariance matrix $\boldsymbol{C}_J$ for the estimator $\hat{\sigma}_{\mathrm{MM}}$ are given by

$$\mathrm{Var}(\hat{\sigma}^2_{\mathrm{A}}) = \frac{2\sigma^4_{\mathrm{A}}}{N}, \quad \mathrm{Var}(\hat{\sigma}^2_{\mathrm{A}'}) = \frac{2\sigma^4_{\mathrm{A}'}}{m},$$

$$\mathrm{Var}(\hat{\sigma}^2_{\mathrm{B}}) = \frac{2\sigma^4_{\mathrm{B}}}{N} \quad \mathrm{Var}(\hat{\sigma}_{\mathrm{A}'\mathrm{B}'}) = \frac{1}{m}\left(2t^2\sigma^4_{\mathrm{A}'} + \sigma^2\sigma^2_{\mathrm{A}'}\right).$$

The off-diagonal terms are given by

$$\mathrm{Cov}(\hat{\sigma}^2_{\mathrm{A}},\hat{\sigma}^2_{\mathrm{B}}) = 2t^2\frac{\sigma^4_{\mathrm{A}}}{N}, \quad \mathrm{Cov}(\hat{\sigma}^2_{\mathrm{A}},\hat{\sigma}^2_{\mathrm{A}'}) = 2\frac{\sigma^4_{\mathrm{A}'}}{N},$$

$$\mathrm{Cov}(\hat{\sigma}^2_{\mathrm{A}},\hat{\sigma}_{\mathrm{A}'\mathrm{B}'}) = 2t\frac{\sigma^4_{\mathrm{A}'}}{N}, \quad \mathrm{Cov}(\hat{\sigma}^2_{\mathrm{A}'},\hat{\sigma}_{\mathrm{A}'\mathrm{B}'}) = 2t\frac{\sigma^4_{\mathrm{A}'}}{m},$$

$$\mathrm{Cov}(\hat{\sigma}^2_{\mathrm{B}},\hat{\sigma}^2_{\mathrm{A}'}) = 2t^2\frac{\sigma^4_{\mathrm{A}'}}{N},$$

$$\mathrm{Cov}(\hat{\sigma}^2_{\mathrm{B}},\hat{\sigma}_{\mathrm{A}'\mathrm{B}'}) = 2t\frac{t^2\sigma^4_{\mathrm{A}'} + \sigma^2\sigma^2_{\mathrm{A}'}}{N}.$$

2. C_J for $\hat{\sigma}^2_{\mathrm{MM}''}$

The diagonal terms for the covariance matrix $\boldsymbol{C}_J$ for the estimator $\hat{\sigma}_{\mathrm{MM}''}$ are given by

$$\mathrm{Var}\big(\hat{\sigma}^2_{\mathrm{A}''}\big) = \frac{2\sigma^4_{\mathrm{A}''}}{n}, \qquad \mathrm{Var}\big(\hat{\sigma}^2_{\mathrm{A}'}\big) = \frac{2\sigma^4_{\mathrm{A}'}}{m},$$
$$\mathrm{Var}\big(\hat{\sigma}^2_{\mathrm{B}''}\big) = \frac{2\sigma^4_{\mathrm{B}''}}{n}, \qquad \mathrm{Var}\big(\hat{\sigma}_{\mathrm{A}'\mathrm{B}'}\big) = \frac{1}{m}\big(2t^2\sigma^4_{\mathrm{A}'} + \sigma^2\sigma^2_{\mathrm{A}'}\big).$$

The off-diagonal terms are given by

$$\mathrm{Cov}\big(\hat{\sigma}^2_{\mathrm{A}''},\hat{\sigma}^2_{\mathrm{B}''}\big) = 2t^2\frac{\hat{\sigma}^4_{\mathrm{A}''}}{n}, \quad \mathrm{Cov}\big(\hat{\sigma}^2_{\mathrm{A}''},\hat{\sigma}^2_{\mathrm{A}'}\big) = 0,$$
$$\mathrm{Cov}\big(\hat{\sigma}^2_{\mathrm{A}''},\hat{\sigma}_{\mathrm{A}'\mathrm{B}'}\big) = 0, \quad \mathrm{Cov}\big(\hat{\sigma}^2_{\mathrm{A}'}, \quad \hat{\sigma}_{\mathrm{A}'\mathrm{B}'}\big) = 2t\frac{\sigma^4_{\mathrm{A}'}}{m},$$
$$\mathrm{Cov}\big(\hat{\sigma}^2_{\mathrm{B}''},\hat{\sigma}^2_{\mathrm{A}'}\big) = 0, \qquad \mathrm{Cov}\big(\hat{\sigma}^2_{\mathrm{B}''},\hat{\sigma}^2_{\mathrm{A}'\mathrm{B}'}\big) = 0.$$

Here we use A'' and B'' to indicate the statistic was calculated using the n subset of states used for generating the final key.

APPENDIX C: THE OPTIMAL ESTIMATOR

An optimal estimator can be found from a linear combination of two estimators, $\hat{\theta}_1$ and $\hat{\theta}$, with $\mathrm{Cov}(\hat{\theta}_1,\hat{\theta}_2) = 0$. The optimal estimator is given by

$$\hat{\theta}_{\mathrm{opt}} = \alpha\hat{\theta}_1 + (1-\alpha)\hat{\theta}_2, \tag{C1}$$

with a variance of

$$\mathrm{Var}(\hat{\theta}_{\mathrm{opt}}) = \alpha^2\mathrm{Var}(\hat{\theta}_1) + (1-\alpha)^2\mathrm{Var}(\hat{\theta}_2), \tag{C2}$$

which is a convex function of α. The optimal value of α can be found by minimizing $\mathrm{Var}(\hat{\theta}_{\mathrm{opt}})$.

$$0 = \frac{d}{d\alpha}\mathrm{Var}(\hat{\theta}_{\mathrm{opt}}), \tag{C3}$$

$$0 = 2\alpha\mathrm{Var}(\hat{\theta}_1) - 2\mathrm{Var}(\hat{\theta}_2) + 2\alpha\mathrm{Var}(\hat{\theta}_2), \tag{C4}$$

$$\alpha = \frac{\mathrm{Var}(\hat{\theta}_2)}{\mathrm{Var}(\hat{\theta}_1) + \mathrm{Var}(\hat{\theta}_2)}. \tag{C5}$$

1. Covariance of $\hat{\sigma}^2_{\mathrm{MLE}}$ and $\hat{\sigma}^2_{\mathrm{MM}''}$

We can show that $\mathrm{Cov}(\hat{\sigma}^2_{\mathrm{MM}''},\hat{\sigma}_{\mathrm{MLE}}) = 0$ given that $\mathrm{Cov}(\hat{\sigma}^2_{\mathrm{B}''},\hat{\sigma}^2_{\mathrm{MLE}}) = 0$, $\mathrm{Cov}(\hat{\sigma}^2_{\mathrm{A}''},\hat{\sigma}^2_{\mathrm{MLE}}) = 0$, and $\mathrm{Cov}(\hat{t},\hat{\sigma}^2_{\mathrm{MLE}}) = 0$ [17],

$$\mathrm{Cov}\big(\hat{\sigma}^2_{\mathrm{MM}''},\hat{\sigma}_{\mathrm{MLE}}\big) = \mathrm{Cov}\big(\hat{\sigma}^2_{\mathrm{B}''} - \hat{t}^2\hat{\sigma}^2_{\mathrm{A}''},\hat{\sigma}^2_{\mathrm{MLE}}\big) \tag{C6}$$
$$= \mathrm{Cov}\big(\hat{\sigma}^2_{\mathrm{B}''},\hat{\sigma}^2_{\mathrm{MLE}}\big) - \mathrm{Cov}\big(t^2\hat{\sigma}^2_{\mathrm{A}''},\hat{\sigma}^2_{\mathrm{MLE}}\big) \tag{C7}$$
$$= 0. \tag{C8}$$

[1] *Proceedings of IEEE International Conference on Computers Systems and Signal Processing*, edited by C. H. Bennett and G. Brassard (IEEE, New York, 1984).
[2] V. Scarani, H. Bechmann-Pasquinucci, N. J. Cerf, M. Dušek, N. Lütkenhaus, and M. Peev, Rev. Mod. Phys. **81**, 1301 (2009).
[3] T. C. Ralph, Phys. Rev. A **61**, 010303 (1999).
[4] F. Grosshans and P. Grangier, Phys. Rev. Lett. **88**, 057902 (2002).
[5] A. Leverrier and P. Grangier, Phys. Rev. Lett. **102**, 180504 (2009).
[6] J. Lodewyck, M. Bloch, R. García-Patrón, S. Fossier, E. Karpov, E. Diamanti, T. Debuisschert, N. J. Cerf, R. Tualle-Brouri, S. W. McLaughlin, and P. Grangier, Phys. Rev. A **76**, 042305 (2007).
[7] P. Jouguet, S. Kunz-Jacques, A. Leverrier, G. Philippe, and D. Eleni, Nat. Photonics **7**, 378 (2013).
[8] R. Renner and J. I. Cirac, Phys. Rev. Lett. **102**, 110504 (2009).
[9] A. Leverrier, F. Grosshans, and P. Grangier, Phys. Rev. A **81**, 062343 (2010).
[10] L. Ruppert, V. C. Usenko, and R. Filip, Phys. Rev. A **90**, 062310 (2014).
[11] A. Leverrier, Phys. Rev. Lett. **114**, 070501 (2015).
[12] F. Grosshans, N. J. Cerf, W. Jérôme, R. Tualle-Brouri, and P. Grangier, Quantum Inf. Comput. **3**, 535 (2003).
[13] R. García-Patrón and N. J. Cerf, Phys. Rev. Lett. **97**, 190503 (2006).
[14] A. Leverrier and P. Grangier, Phys. Rev. A **81**, 062314 (2010).
[15] P. Jouguet, S. Kunz-Jacques, and A. Leverrier, Phys. Rev. A **84**, 062317 (2011).
[16] S. M. Kay, *Estimation Theory*, Fundamentals of Statiscal Signal Processing Vol. 1 (Prentice Hall, Englewood Cliffs, NJ, 1993).
[17] J. K. Patel and C. B. Read, *Handbook of the Normal Distribution*, 2nd ed., Statistics: Textbooks and Monographs Vol. 150 (Marcel Dekker, New York, 1996).
[18] M. H. Kutnereter, C. J. Nachtsheim, J. Neter, and W. Li, *Applied Linear Statistical Models*, 5th ed. (McGraw-Hill Irwin Press, New York, 2005).
[19] P. Jouguet, S. Kunz-Jacques, E. Diamanti, and A. Leverrier, Phys. Rev. A **86**, 032309 (2012).