

Improved statistical fluctuation analysis for two decoy-states phase-matching quantum key distribution

Jiang-Ping Zhou(周江平), Yuan-Yuan Zhou(周媛媛), Xue-Jun Zhou(周学军), and Xuan Bao(暴轩)

Citation: Chin. Phys. B, 2023, 32 (**8**): 080306. DOI: 10.1088/1674-1056/acbdea
Journal homepage: http://cpb.iphy.ac.cn; http://iopscience.iop.org/cpb

What follows is a list of articles you may be interested in

Improving source-in-the-middle continuous-variable quantum key distribution using a heralded hybrid linear amplifier

Lei-Xin Wu(伍磊鑫), Yan-Yan Feng(冯艳艳), and Jian Zhou(周健)
Chin. Phys. B, 2023, 32 (**7**): 070310. DOI: 10.1088/1674-1056/ac9b39

Phase-matching quantum key distribution with imperfect sources

Xiao-Xu Zhang(张晓旭), Yi-Fei Lu(陆宜飞), Yang Wang(汪洋), Mu-Sheng Jiang(江木生), Hong-Wei Li(李宏伟), Chun Zhou(周淳), Yu Zhou(周雨), and Wan-Su Bao(鲍皖苏)
Chin. Phys. B, 2023, 32 (**5**): 050308. DOI: 10.1088/1674-1056/ac9b03

Effect of weak randomness flaws on security evaluation of practical quantum key distribution with distinguishable decoy states

Yu Zhou(周雨), Hong-Wei Li(李宏伟), Chun Zhou(周淳), Yang Wang(汪洋), Yi-Fei Lu(陆宜飞),Mu-Sheng Jiang(江木生), Xiao-Xu Zhang(张晓旭), and Wan-Su Bao(鲍皖苏)
Chin. Phys. B, 2023, 32 (**5**): 050305. DOI: 10.1088/1674-1056/ac8730

Security of the traditional quantum key distribution protocolswith finite-key lengths

Bao Feng(冯宝), Hai-Dong Huang(黄海东), Yu-Xiang Bian(卞宇翔), Wei Jia(贾玮), Xing-Yu Zhou(周星宇), and Qin Wang(王琴)
Chin. Phys. B, 2023, 32 (**3**): 030307. DOI: 10.1088/1674-1056/ac9b33

Performance analysis of quantum key distribution using polarized coherent-states in free-space channel

Zengte Zheng(郑增特), Ziyang Chen(陈子扬), Luyu Huang(黄露雨),Xiangyu Wang(王翔宇), and Song Yu(喻松)
Chin. Phys. B, 2023, 32 (**3**): 030306. DOI: 10.1088/1674-1056/ac7450

Improved statistical fluctuation analysis for two decoy-states phase-matching quantum key distribution

Jiang-Ping Zhou(周江平)[1], Yuan-Yuan Zhou(周媛媛)[1,†], Xue-Jun Zhou(周学军)[1], and Xuan Bao(暴轩)[2]

[1] *College of Electronic Engineering, Naval University of engineering, Wuhan* 430000, *China*
[2] *Unit* 92682, *Zhanjiang* 524000, *China*

(Received 10 January 2023; revised manuscript received 19 February 2023; accepted manuscript online 22 February 2023)

Phase-matching quantum key distribution is a promising scheme for remote quantum key distribution, breaking through the traditional linear key-rate bound. In practical applications, finite data size can cause significant system performance to deteriorate when data size is below 10^{10}. In this work, an improved statistical fluctuation analysis method is applied for the first time to two decoy-states phase-matching quantum key distribution, offering a new insight and potential solutions for improving the key generation rate and the maximum transmission distance while maintaining security. Moreover, we also compare the influence of the proposed improved statistical fluctuation analysis method on system performance with those of the Gaussian approximation and Chernoff–Hoeffding boundary methods on system performance. The simulation results show that the proposed scheme significantly improves the key generation rate and maximum transmission distance in comparison with the Chernoff–Hoeffding approach, and approach the results obtained when the Gaussian approximation is employed. At the same time, the proposed scheme retains the same security level as the Chernoff–Hoeffding method, and is even more secure than the Gaussian approximation.

Keywords: quantum key distribution, phase matching protocol, statistical fluctuation analysis, decoy state

PACS: 03.67.Dd, 03.67.Ac, 03.67.Hk, 42.50.Ex **DOI:** 10.1088/1674-1056/acbdea

1. Introduction

Quantum key distribution (QKD)[1] is one of the most successful applications in the field of quantum information, and is based on the fundamental laws of physics and enables two remote parties (Alice and Bob) to achieve unconditionally secure key sharing. Over the past three decades, quantum key distribution has become more and more mature in practical applications, and many secure and efficient QKD protocols[2–7] have been proposed to address security vulnerabilities caused by the non-ideality of physical apparatuses.[8,9] Since quantum information is always carried by transmitted photons,[10] it is generally believed that in the absence of trusted relay nodes, the key rate R is bounded by the channel transmittance η,[11,12] $R \leq O(\eta)$. Surprisingly, the twin-field QKD (TF-QKD)[13] and its variants[14–16] can overcome the rate- distance limits and increase the key rate from $O(\eta)$ to $O(\sqrt{\eta})$ and thus has attracted widespread attention. Theoretically, TF-QKD has been improved in terms of asymmetric channel, finite data size, device non-ideality and other aspects.[17–20] At the same time, breakthroughs have been made in the relevant experiments of TF-QKD. In 2020, Chen *et al*. achieved the 511-km fiber channel TF-QKD experiment,[21] and Pittaluga *et al*. used dual-band phase stabilization technology to realize TF-QKD over 600 km in optical fiber.[22] In 2021, Wang *et al*. further extended the transmission distance to 833.8 km in fiber channel by optimizing the source, channel and detectors.[23]

Phase-matching QKD(PM-QKD),[14] which employs phase encoding and phase post-compensation technique,[24] is one of the most cutting-edge protocols in the TF-QKD variants.[25] It is immune to all possible measurement terminal attacks and offers a number of advantages,[26] including good stability, strong anti-interference capability, and low bit error rate. Nevertheless, the rates and probabilities are different for finite key sizes in a real experimental setting. In order to account for this variation and ultimately derive a secure final key, a statistical fluctuation analysis method must be employed. This will allow the correct parameters to be approximated and provide confidence in the security of the system. Furthermore, if finite data size effects are taken into consideration, the performance of real PM-QKD can be further compromised, leading to a decrease in key rate and posing a challenge to practical implementation.

As is well known, several statistical fluctuation methods have been widely used in QKD systems mainly including Gaussian approximation,[2] Hoeffding inequality,[27] and Chernoff bound.[28] Among the three methods, the Gaussian approximation method is based on the assumption that the channel statistical fluctuation follows a Gaussian distribution, which provides a high key generation rate but has poor security.[29] Contrarily, the Hoeffding's inequality and Chernoff bounds yield higher levels of security, but may result in a lower key rate. In a QKD system, security is always the priority. For this reason, it is an attractive prospect to improve the existing statistical fluctuation analysis methods and further

†Corresponding author. E-mail: EPJZYY@aliyun.com

© 2023 Chinese Physical Society and IOP Publishing Ltd

to explore new methods, which can increase the key generation rate while still guaranteeing the same high level of security as the Chernoff bound. Fortunately, Zhang[30] proposed a new statistical fluctuation analysis method, which combines the high key generation rate of Gaussian approximation with the strong security offered by the Chernoff bound. The method has been validated and applied successfully to BB84-QKD and MDI-QKD protocols.[29,31]

In this paper, we adopt the new statistical fluctuation method proposed by Zhang *et al.*[30] to investigate the performance of PM-QKD protocol. Furthermore, we apply it to a two-decoy-state PM-QKD protocol that uses a weak coherent state (WCS) light source, which has proved to be a practical QKD protocol.[32] With the new statistical fluctuation method, tight bounds can be established for estimation parameters such as photon yield and quantum bit error rate (QBER). This leads to a higher key generation rate and longer maximum transmission distance, allowing the two-decoy-state PM-QKD protocol to be implemented within a reasonable time period.

The remainder of this paper is organized as follows. In Section 2, we briefly review the commonly used two-decoy-state PM-QKD. In Section 3, we introduce the new method of finite data size analysis in detail and apply it to the two-decoy-state PM-QKD. In Section 4, we carry out simulation by using the corresponding model and typical parameters, and then compare the simulation results with the results in other literature. In Section 5, we draw some conclusions from the present study.

2. Two-decoy-state PM-QKD protocol

Two decoy states, a vacuum state and a weak decoy state, have proved to be enough for practical PM-QKD.[33] The performance of practical PM-QKD system cannot be significantly improved by adding additional decoy state. The WCS is a common light source with a simple structure, low cost, and mature technology, making it ideal for practical application. The two-decoy-state PM-QKD protocol using WCS can be described as follows.

(i) Alice generates a random key κ_{A}, a random phase $\phi_{\mathrm{A}} \in [0, 2\pi)$ and a random intensity $s \in \{\mu, \nu, o\}$, and then prepares the coherent state $|\sqrt{s}\mathrm{e}^{\mathrm{i}(\phi_{\mathrm{A}}+\pi\kappa_{\mathrm{A}})}\rangle_{\mathrm{A}}$. Similarly, Bob prepares the coherent state $|\sqrt{s}\mathrm{e}^{\mathrm{i}(\phi_{\mathrm{B}}+\pi\kappa_{\mathrm{B}})}\rangle_{\mathrm{B}}$. It should be noted that we implement the two-decoy-state protocol by randomly selecting different signal intensities, where μ, ν, and o denote the intensity of the signal state, the weak decoy state, and the vacuum state, respectively.

(ii) Alice and Bob send the modulated pulses to the untrusted third party (UTP) who performs interference detection and announces the results. The successful detection means that one and only one detector clicks. After UTP announces the detection result, Alice and Bob announce the phase slice index where the random phase is located and sift the keys from the successful detections after phase post-compensation.

(iii) Alice and Bob estimate the phase error rate through the experiment data and perform post-processing such as error correction and privacy amplification to obtain the distilled keys.

The final key rate[14] is given by

$$R_{\mathrm{PM}} \geq \frac{2}{M} Q_\mu \left[1 - fH\left(E_\mu^{\mathrm{Z}}\right) - H\left(E_\mu^{\mathrm{X}}\right)\right], \tag{1}$$

where M is the total number of phase slices, $2/M$ is the phase-sifting factor, Q_μ is the value of the overall gain for signal state with intensity μ, f is the error correction efficiency, $H(\cdot)$ is the Shannon function and $H(x) = -x\log_2(x) - (1-x)\log_2(1-x)$, E_μ^{Z} is the average quantum bit error rate (QBER), Q_μ and E_μ^{Z} can be derived directly from the experiment when statistical fluctuations are not considered, and E_μ^{X} is the phase error rate which is given by

$$\begin{aligned} E_\mu^X \leq\ & q_0^\mu e_0^Z + q_1^\mu e_1^Z + q_3^\mu e_3^Z + q_5^\mu e_5^Z \\ & + \left(1 - q_0^\mu - q_1^\mu - q_3^\mu - q_5^\mu\right), \end{aligned} \tag{2}$$

where e_k^Z is the error rate of k-photon state, $q_k^\mu = p_k^\mu \frac{Y_k}{Q_\mu}$ represents the contribution of k-photon state to successful detection, p_k^μ is the probability that the source with intensity μ emits k-photon states, Y_k is the yield of k-photon states.

It can be known from Eq. (2) that it is necessary to estimate the max e_k^Z and min Y_k to obtain the secure key generation rate, which can be achieved by using the decoy state method. According to the photon number channel model, the measured Q_s and E_s^Z can be expressed as follows:

$$Q_s = \sum_{k=0}^{\infty} p_k^s Y_k, \tag{3}$$

$$E_s^Z Q_s = \sum_{k=0}^{\infty} p_k^s e_k Y_k. \tag{4}$$

By replacing s with μ, ν, and o respectively, we can obtain three sets of expressions similar to the above equations, which can estimate the values of the max e_k^Z and min Y_k jointly.

3. Improved statistical fluctuation analysis for PM-QKD

In this work, the improved statistical fluctuation analysis method is extended to two-decoy-state PM-QKD, which makes the estimation of relevant parameters more accurate while ensuring safety. Before continuing analyzing the specific statistical fluctuation, some notations used throughout the paper are first defined. The capital letters N, M, and EM represent the number of pulses sent by Alice or Bob, the number of valid detections and error detections on the UTP's side, respectively. We use the subscript s to denote the fraction of each count mentioned above generated by the pulses of intensity s. The superscripts U and L denote the upper bound and

lower bound of the corresponding parameters respectively, e represents the base of the natural logarithm, and $E(\cdot)$ denotes the expected value.

According to the analysis in Section 2, the maximum value of e_k^Z and the minimum value of Y_k can be estimated under the constraints of Eqs. (3) and (4). When we consider statistical fluctuations, Q_s and $E_s^Z Q_s$ on the left side of the equations are no longer fixed values, and their confidence intervals can be determined by accepting a certain probability of error. Therefore, equations (3) and (4) will also become inequality constraints accordingly. According to the new constraints, the maximum value of e_k^Z and the minimum value of Y_k can be more reasonably estimated, thereby determining the security key generation rate. As it is difficult to analyze Q_s directly, we can analyze the M_s first according to the relationship between Q_s and M_s, *i.e.*, $Q_s = M_s/N_s$. Similarly, $E_s^Z Q_s$ can also be analyzed similarly.

The M_s can be written as the sum of a series of random variables

$$M_s = \sum_{j=1}^{N_s} \chi_j^s, \tag{5}$$

where χ_j^s is a random variable, which indicates the detection result at the UTP's end of the pulse sent when the light source intensity is s. Furthermore, 1 indicates a valid response and 0 refers to the opposite result. Then the χ_j^s can be written as

$$\chi_j^s = \begin{cases} 1, & p^{sj}, \\ 0, & 1-p^{sj}, \end{cases} \tag{6}$$

where p^{sj} is the probability with which a single pulse produces a valid detection at the UTP's end and its value remains the same no matter what j is. We should note the difference between p^{sj} and p_k^s, where the former describes the probability of the valid detection and the latter refers to the probability of k-photon state generated by the source. For the χ_j^s with different $j = 1, \ldots, N_s$, these random variables are the results obtained at different times under the same experimental conditions including light source, channel, and detector, so they are independent and identically distributions. Through the Chernoff bound with an inverse formulation, we can estimate the expected value $E(M_s)$ according to the observed value M_s. Similarly, the above analysis also applicable to $E_s M_s$.

To make this point clear, we review the Chernoff bound: for a set of n independent Bernoulli random variables $\chi_i \in \{0,1\}$, their sum is $\chi = \sum_{i=1}^{n} \chi_i$, and the average value of χ is $\bar{\chi} = E(\chi)$, the relationship between χ and $\bar{\chi}$ is described by the following two inequalities:

$$\Pr\left(\chi > \left(1+\delta^{\mathrm{L}}\right)\overline{\chi}\right) < \left[\frac{\mathrm{e}^{\delta^{\mathrm{L}}}}{\left(1+\delta^{\mathrm{L}}\right)^{1+\delta^{\mathrm{L}}}}\right]^{\overline{\chi}} = g\left(\delta^{\mathrm{L}}, \overline{\chi}\right), \tag{7}$$

$$\Pr\left(\chi < \left(1-\delta^{\mathrm{U}}\right)\overline{\chi}\right) < \left[\frac{\mathrm{e}^{-\delta^{\mathrm{U}}}}{\left(1-\delta^{\mathrm{U}}\right)^{1-\delta^{\mathrm{U}}}}\right]^{\overline{\chi}} = g\left(-\delta^{\mathrm{U}}, \overline{\chi}\right), \tag{8}$$

where $\delta^{\mathrm{L}} > 0$ and $1 > \delta^{\mathrm{U}} > 0$ will be constant when the failure probability is determined.

Assuming that the failure probability is ε, we can obtain the confidence interval of $\bar{\chi}$ according to the above inequations (7) and (8). Considering a symmetric form, we have

$$\Pr\left(\chi > \left(1+\delta^{\mathrm{L}}\right)\overline{\chi}\right) = \Pr\left(\overline{\chi} < \frac{\chi}{1+\delta^{\mathrm{L}}}\right) < g\left(\delta^{\mathrm{L}}, \overline{\chi}\right) = \frac{\varepsilon}{2}, \tag{9}$$

$$\Pr\left(\chi < \left(1-\delta^{\mathrm{U}}\right)\overline{\chi}\right) = \Pr\left(\overline{\chi} > \frac{\chi}{1-\delta^{\mathrm{U}}}\right) < g\left(\delta^{\mathrm{L}}, \overline{\chi}\right) = \frac{\varepsilon}{2}. \tag{10}$$

From the inequations (9) and (10), we can derive the following in equation:

$$\Pr\left(\frac{\chi}{1+\delta^{\mathrm{L}}} \le \overline{\chi} \le \frac{\chi}{1-\delta^{\mathrm{U}}}\right) \ge 1-\varepsilon. \tag{11}$$

In other words, the confidence interval of $\bar{\chi}$ is $\left[E^{\mathrm{L}}(\chi), E^{\mathrm{U}}(\chi)\right]$, $E^{\mathrm{L}}(\chi) = \chi/(1+\delta^{\mathrm{L}})$, and $E^{\mathrm{U}}(\chi) = \chi/(1-\delta^{\mathrm{U}})$. The δ^{L} and δ^{U} can be obtained by solving the following equations respectively:

$$\left[\frac{\mathrm{e}^{\delta^{\mathrm{L}}}}{\left(1+\delta^{\mathrm{L}}\right)^{1+\delta^{\mathrm{L}}}}\right]^{E^{\mathrm{L}}(\chi)} = g\left(\delta^{\mathrm{L}}, E^{\mathrm{L}}(\chi)\right) = \frac{\varepsilon}{2}, \tag{12}$$

$$\left[\frac{\mathrm{e}^{-\delta^{\mathrm{U}}}}{\left(1-\delta^{\mathrm{U}}\right)^{1-\delta^{\mathrm{U}}}}\right]^{E^{\mathrm{U}}(\chi)} = g\left(-\delta^{\mathrm{U}}, E^{\mathrm{U}}(\chi)\right) = \frac{\varepsilon}{2}. \tag{13}$$

If $M_s \ge -6\ln(\varepsilon/2)$, the approximate result is

$$\delta = \delta^{\mathrm{L}} = \delta^{\mathrm{U}} = \frac{3\beta + \sqrt{8\beta\chi + \beta^2}}{2(\chi-\beta)}, \tag{14}$$

where $\beta = -\ln(\varepsilon/2)$. In order to simplify the expression, we use δ to denote the identical δ^{L} and δ^{U}.

Following the above analysis, we apply the Chernoff bound with an inverse formulation to M_s and $E_s M_s$, that is, replace χ with M_s and $E_s M_s$. Thus, we can obtain the confidence interval of M_s and $E_s M_s$ as follows:

$$\begin{cases} E^{\mathrm{L}}(M_s) = \dfrac{M_s}{1+\delta_M^s}, \\ E^{\mathrm{U}}(M_s) = \dfrac{M_s}{1-\delta_M^s}, \end{cases} \tag{15}$$

where

$$\delta_M^s = \frac{3\beta + \sqrt{8\beta M_s + \beta^2}}{2(M_s-\beta)}, \tag{16}$$

$$\begin{cases} E^{\mathrm{L}}(E_s M_s) = \dfrac{E_s M_s}{1+\delta_{\mathrm{EM}}^s}, \\ E^{\mathrm{U}}(E_s M_s) = \dfrac{E_s M_s}{1-\delta_{\mathrm{EM}}^s}, \end{cases} \tag{17}$$

with

$$\delta_{\mathrm{EM}}^{s}=\frac{3\beta+\sqrt{8\beta E_{s}M_{s}+\beta^{2}}}{2\left(E_{s}M_{s}-\beta\right)}. \tag{18}$$

To obtain the final key generation rate, we need to further study Q_s and E_sQ_s. Based on the relationship between M_s and Q_s, we have

$$\begin{cases} E^{\mathrm{L}}\left(Q_{s}\right)=E^{\mathrm{L}}\left(\dfrac{M_{s}}{N_{s}}\right)=\dfrac{E^{\mathrm{L}}\left(M_{s}\right)}{N_{s}} \\ \qquad\quad =\dfrac{M_{s}}{\left(1+\delta_{M}^{s}\right)N_{s}}=\dfrac{Q_{s}}{1+\delta_{M}^{s}}, \\ E^{\mathrm{U}}\left(Q_{s}\right)=E^{\mathrm{U}}\left(\dfrac{M_{s}}{N_{s}}\right)=\dfrac{E^{\mathrm{U}}\left(M_{s}\right)}{N_{s}} \\ \qquad\quad =\dfrac{M_{s}}{\left(1-\delta_{M}^{s}\right)N_{s}}=\dfrac{Q_{s}}{1-\delta_{M}^{s}}. \end{cases} \tag{19}$$

Similarly, for the E_sQ_s, we have

$$\begin{cases} E^{\mathrm{L}}\left(E_{s}Q_{s}\right)=\dfrac{E_{s}Q_{s}}{1+\delta_{\mathrm{EM}}^{s}}, \\ E^{\mathrm{U}}\left(E_{s}Q_{s}\right)=\dfrac{E_{s}Q_{s}}{1-\delta_{\mathrm{EM}}^{s}}. \end{cases} \tag{20}$$

When considering statistical fluctuations, the left side of Eqs. (3) and (4) should be expected values instead of experiment values. In practice, the range of the corresponding expected value can only be inferred from the measured value. Therefore, equations (3) and (4) will become inequalities, and the value range is determined by Eq. (19). At the same time, according to the end of Section 2, substituting the intensity of signal state and decoy state into s respectively, our objective becomes the estimation of the min Y_k, the results are as follows:

$$\begin{cases} \dfrac{Q_{\mu}}{1+\delta_{M}^{\mu}}\leq\displaystyle\sum_{k=0}^{\infty}\frac{\mathrm{e}^{-\mu}\mu^{k}}{k!}Y_{k}\leq\frac{Q_{\mu}}{1-\delta_{M}^{\mu}}, \\ \dfrac{Q_{\nu}}{1+\delta_{M}^{\nu}}\leq\displaystyle\sum_{k=0}^{\infty}\frac{\mathrm{e}^{-\nu}\nu^{k}}{k!}Y_{k}\leq\frac{Q_{\nu}}{1-\delta_{M}^{\nu}}, \\ \dfrac{Q_{o}}{1+\delta_{M}^{o}}\leq\displaystyle\sum_{k=0}^{\infty}\frac{\mathrm{e}^{-o}o^{k}}{k!}Y_{k}\leq\frac{Q_{o}}{1-\delta_{M}^{o}}, \end{cases} \tag{21}$$

and evaluation of max e_k^{Z}, and the results are as follows:

$$\begin{cases} \dfrac{E_{\mu}Q_{\mu}}{1+\delta_{\mathrm{EM}}^{\mu}}\leq\displaystyle\sum_{k=0}^{\infty}\frac{\mathrm{e}^{-\mu}\mu^{k}}{k!}e_{k}Y_{k}\leq\frac{E_{\mu}Q_{\mu}}{1-\delta_{\mathrm{EM}}^{\mu}}, \\ \dfrac{E_{\nu}Q_{\nu}}{1+\delta_{\mathrm{EM}}^{\nu}}\leq\displaystyle\sum_{k=0}^{\infty}\frac{\mathrm{e}^{-\nu}\nu^{k}}{k!}e_{k}Y_{k}\leq\frac{E_{\nu}Q_{\nu}}{1-\delta_{\mathrm{EM}}^{\nu}}, \\ \dfrac{E_{o}Q_{o}}{1+\delta_{\mathrm{EM}}^{o}}\leq\displaystyle\sum_{k=0}^{\infty}\frac{\mathrm{e}^{-o}o^{k}}{k!}e_{k}Y_{k}\leq\frac{E_{o}Q_{o}}{1-\delta_{\mathrm{EM}}^{o}}. \end{cases} \tag{22}$$

In order to find the secure key rate, we investigate the lower bound Y_1, Y_3, Y_5 and upper bound e_1^{Z}, e_3^{Z}, e_5^{Z} separately according to Eqs. (1) and (2). These problems can be solved by using linear programming with the linear constraints given by Eqs. (21) and (22) respectively.

In fact, the main difference among Gaussian analysis, Chernoff–Hoeffding bound, and the improved statistical analysis method proposed in this paper lies in the way of constructing the probability model and estimating the confidence intervals of the corresponding expected values in Eqs. (15) and (17) according to the observed values. According to the central limit theorem, it is assumed in Gaussian analysis that the observed values follow a normal distribution, but this assumption is usually not valid unless the data are infinitely long, which does not match the real system. The utilizing of Gaussian analysis is equivalent to artificially increasing the system constraints. Eve can take advantage of the difference in information between the real system and the constrained system to perform attacks without being detected, thereby leading to security risks. In the Chernoff–Hoeffding bound the Hoeffding inequality is used to estimate the confidence interval, which is based on independent random variables and is in line with the actual system, without introducing additional information. However, this method ignores the fact that the effective detection of pulses in the process of quantum key distribution is a Bernoulli event, which causes some information to be lost. Thus, though this method can ensure security, the estimation of the confidence interval is loose. The improved method proposed in this paper is also in line with the real system and utilizes the information that the successful detection of pulses in the process of quantum key distribution is a Bernoulli event, thus allowing the proposed method to meet the security standards of the Chernoff–Hoeffding bound while further improving the accuracy of parameter estimation.

Without loss of generality, when the measured value Q is 0.1, the confidence intervals of the expected value $E\left(Q\right)$ estimated by the three different methods are shown in Table 1.

Table 1. Bound of $E\left(Q\right)$ estimated by three different methods with $Q=0.1$.

Bound of $E\left(Q\right)$	Gaussian	Chernoff–Hoeffding	Our improved
Upper bound	0.10001118	0.10002510	0.10001227
Lower bound	0.09998881	0.09998496	0.09998772

From Table 1, it can be seen that the tightest bound of $E\left(Q\right)$ is obtained by using Gaussian analysis, then the bounds of $E\left(Q\right)$ become looser when the proposed method in this work is used, and the bounds of $E\left(Q\right)$ provided by the Chernoff–Hoeffding bound are the loosest. Since the accuracy of parameter estimation is positively correlated with the tightness of the bounds, this accuracy ranking of estimations of parameters made by these three statistical fluctuation analysis methods is consistent with the previous order of bound tightness.

4. Numerical simulation

In this section, we conduct some numerical simulations for two-decoy-state PM-QKD by utilizing Gaussian approximation in Ref. [2], Chernoff–Hoeffding bound in Ref. [28], and the improved statistical fluctuation analysis method, respectively. We analyze in detail the effects of these three different statistical fluctuation analysis methods on the key generation rate and the maximum transmission distance, under various data sizes. For simplicity, we use GS, C–H, IM to represent Gaussian approximation, the Chernoff–Hoeffding bound, and the improved statistical fluctuation analysis method respectively. The experimental parameters used in this paper are listed in Table 2,[33] which are typical values or practical values. These parameters are $e_{\rm d}$ (misalignment error), $p_{\rm d}$ (dark count rate), f (error correction efficiency), $\eta_{\rm d}$ (detector efficiency), M (number of phase slices), n_α (multiples of standard deviations determined by the failure probability, $n_\alpha = 5$ means that the failure probability used later is 5.73×10^{-7}), and α (transmission fiber loss constant).

Table 2. Parameters used in simulation.

$e_{\rm d}$	$p_{\rm d}$	f	$\eta_{\rm d}$	M	n_α	α
1.5%		1.15	14.5%	16	5	0.2

We compare the performance of IM with those of the GS and C–H with different data sizes in Fig. 1. When $N = 10^{16}$, the performances of the three methods are very close as shown in Fig. 1 where the three dashed lines are barely distinguishable. From the inset, we can see that the GS outperforms the IM, while the IM performs better than the C–H. When $N = 10^{10}$, the performance ranking of different statistical fluctuation methods remains the same, though the disparities among them are more distinct. Obviously, GS provides a maximum transmission distance 5 km, greater than that of the IM, while IM yields a maximum transmission distance 24 km greater than C–H. In addition, the difference in the key generation rate is more obvious when the distance exceeds 300 km. This is mainly because the M_s and E_sM_s decrease with distance increasing. According to Eq. (13), the smaller M_s and E_sM_s result in larger δ^s_M and $\delta^s_{\rm EM}$, creating a larger confidence interval of M_s and E_sM_s, which consequently produces a more obvious statistical fluctuation effect. It is clear that the system with GS has the highest key generation rate and the longest transmission distance regardless of the number of data size. However, its assumption renders it vulnerable to coherent attacks. On the other hand, although C–H is secure enough, its performance is the weakest in the three methods especially for small data size. The IM can significantly improve the system performance and approach the GS when the data size is small. In reality, the smaller data size is still very important when the system cost is taken into account. There are two ways to increase the data size: one is to prolong the experiment time, and the other is to increase the pulse frequency of the light source. The extension of experimental time will lead the system latency to increase and the system stability to decrease, which is generally undesirable. Increasing the pulse frequency means that the operating frequency of the source will rise, which requires better equipment such as the source and detector, thus leading to a higher overall cost. In conclusion, the statistical fluctuation analysis method we studied here is of great significance in improving the system performance in the case of small data size.

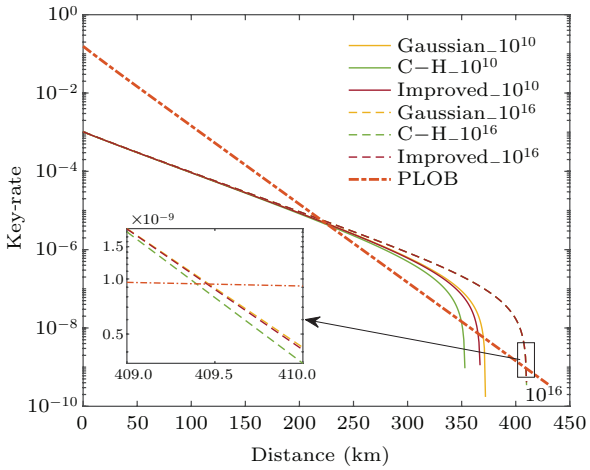


Fig. 1. Comparison among key generation rates generated by twodecoy-state PM-QKD with three different statistical fluctuation analysis methods, with yellow, purple, and red denoting curves obtained by using Gaussian approximation, Chernoff–Hoeffding bound, and the improved statistical fluctuation analysis method, respectively, with solid line and dashed lines representing the case when $N = 10^{10}$ and $N = 10^{16}$, respectively, and with inset showing magnified part within frame.

In addition, for the three methods, we provide the effect of the total number of pulses on key rate and maximum transmission distance. The results are shown in Figs. 2 and 3, respectively.

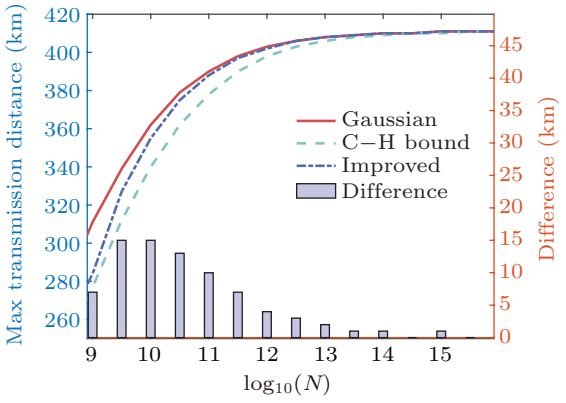


Fig. 2. Comparison among maximum safe transmission distances of two-decoy-state PM-QKD using different statistical fluctuation analysis methods for different data sizes, with curves describing maximum safe transmission distance, and bars depicting difference in maximum safe transmission distance between the improved statistical fluctuation analysis method and Chernoff–Hoeffding bound.

In Fig. 2, the curves depict the changes in the maximum transmission distance with data size for the three methods, and the histogram represents the discrepancy between CH and IM in the maximum transmission distance. It is clear that the maximum transmission distance gradually increases with the increase of the data size, but at a diminishing rate and eventually approaching 411 km. In accordance with the conclusion in Fig. 1, GS has the highest maximum transmission distance, followed by IM and C–H. Owing to the lack of security of GS, we further analyze the performance improvement provided by IM relative to C–H through the histogram. As can be seen from the histogram, the maximum gap between C–H and IM is 15 km when the data size is $N = 10^{10}$ and the minimum difference is 0 km when the data size is $N = 10^{14.5}$. Furthermore, the difference will be less than 3 km when the data size is greater than $N = 10^{13}$. This further demonstrates that the improved statistical fluctuation analysis method can significantly improve the system performance when the data size is small.

In Fig. 3, we compare the key generation rates of the systems by using the three different methods at 100 km. By combining Fig. 2 with Fig. 3, it can be seen that the key generation rate and the maximum transmission distance have similar trends. In addition, when the data size is small, the performance of IM approaches that of GS, which demonstrates a great improvement compared with that of C–H.

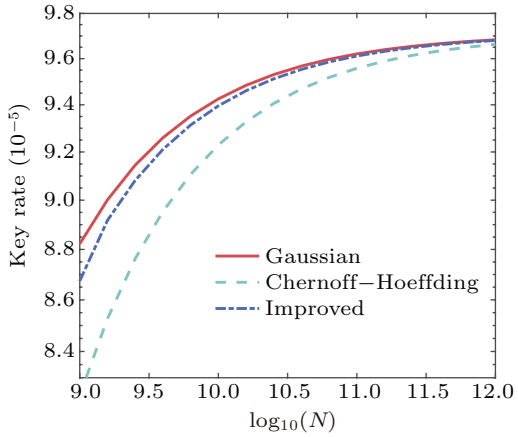


Fig. 3. Comparison among key rates of two-decoy-state PM-QKD using different statistical fluctuation analysis methods for different data sizes at 100 km, with solid line, dashed line, and dash-dotted line representing Gaussian approximation, Chernoff–Hoeffding bound, and improved statistical fluctuation analysis method studied in this paper respectively.

5. Conclusions

In this work, a new QKD scheme is proposed by combining the improved statistical fluctuation analysis method with the two decoy-states PM-QKD. We analyzed the performance of the scheme and compared it with those of the schemes in which the Gaussian analysis and the Chernoff–Hoeffding bound are used respectively under the limitation of finite data size. The simulation results show that the PM-QKD considering statistical fluctuation can still break through the PLOB bound. When the data size is larger than 10^{12}, the key generation rates of the three statistical fluctuation analysis methods are almost the same at 100 km. Equally, when the data length is larger than 10^{13}, the maximum transmission distances of the three statistical fluctuation analysis methods are also essentially the same. This indicates that when the data size is relatively large, the three different statistical fluctuation analysis methods have no remarkable influence on the system. On the contrary, when the data size is 10^{11}, the improved statistical fluctuation analysis method can obtain the approximate performance of Gaussian approximation and maintain the same security level of the Chernoff–Hoeffding bound, compared with the Chernoff–Hoeffding bound, the key generation rate is increased by 10% at 100 km, and the maximum secure transmission distance is increased by 13 km. The improved statistical fluctuation analysis method is of great significance since a smaller data size plays an important role in practice.

Acknowledgement

We are grateful to Mr. D Z Lan and Mr. Y M Zhu for their enlightening discussion and friendly help in mathematical simulation.

References

[1] Bennett C H and Brassard G 1984 *Proc. IEEE Int. Conf. Comput. Syst. Signal Process.* 1984 pp. 175–179
[2] Ma X F, Qi B, Zhao Y and Lo H K 2005 *Phys. Rev. A* **72** 012326
[3] Xu F, Ma X, Zhang Q, Lo H K and Pan J W 2020 *Rev. Mod. Phys.* **92** 25002
[4] Tan Y G and Liu Q 2016 *Chin. Phys. Lett.* **33** 090303
[5] Li X, Yuan H W, Zhang C M and Wang Q 2020 *Chin. Phys. B* **29** 070303
[6] Zhang S J, Xiao C, Zhou C, Wang X, Yao J S, Zhang H L and Bao W S 2020 *Chin. Phys. B* **29** 020301
[7] Zhang Y Y, Bao W S, Zhou C, Li H W, Wang Y and Jiang M S 2017 *Chin. Phys. Lett.* **34** 040301
[8] Jain N, Stiller B, Khan I, Elser D, Marquardt C and Leuchs G 2016 *Contemp. Phys.* **57** 366
[9] Gottesman D, Hoi-Kwonglo L O, Lütkenhaus N and Preskill J 2004 *Quantum Inf. Comput.* **4** 325
[10] Chen G, Wang L, Li W, Zhao Y, Zhao S and Gruska J 2020 *Quantum Inf. Process.* **19** 416
[11] Takeoka M, Guha S and Wilde M M 2014 *Nat. Commun.* **5** 5235
[12] Pirandola S, Laurenza R, Ottaviani C and Banchi L 2017 *Nat. Commun.* **8** 15043
[13] Lucamarini M, Yuan Z L, Dynes J F and Shields A J 2018 *Nature* **557** 400
[14] Ma X F, Zeng P and Zhou H Y 2018 *Phys. Rev. X* **8** 031043
[15] Wang X B, Yu Z W and Hu X L 2018 *Phys. Rev. A* **98** 1
[16] Mao Y, Zeng P and Chen T Y 2021 *Adv. Quantum Technol.* **12** 1012
[17] Cui C H, Yin Z Q, Wang R, Chen W, Wang S, Guo G C and Han Z F 2019 *Phys. Rev. Appl.* **11** 034053
[18] Zhou F, Qu W, Wang J, Dou T, Li Z, Yang S, Sun Z, Miao G and Ma H 2020 *Eur. Phys. J. D* **74** 185
[19] Currás-Lorenzo G, Navarrete Á, Azuma K, Kato G, Curty M and Razavi M 2021 *Npj Quantum Inf.* **7** 22
[20] Wang W and Lo H K 2020 *New J. Phys.* **22** 013020

[21] Chen J P, Zhang C, Liu Y, Jiang C, Zhang W J, Han Z Y, Ma S Z, Hu X L, Li Y H, Liu H, Zhou F, Jiang H F, Chen T Y, Li H, You L X, Wang Z, Wang X B, Zhang Q and Pan J W 2021 *Nat. Photon.* **15** 570
[22] Pittaluga M, Minder M, Lucamarini M, Sanzaro M, Woodward R I, Li M J, Yuan Z and Shields A J 2021 *Nat. Photon.* **15** 530
[23] Wang S, Yin Z Q, He D Y, Chen W, Wang R Q, Ye P, Zhou Y, Fan-Yuan G J, Wang F X, Chen W, Zhu Y G, Morozov P V, Divochiy A V, Zhou Z, Guo G C and Han Z F 2022 *Nat. Photon.* **16** 154
[24] Yu B, Mao Q, Zhu X, Yu Y and Zhao S 2021 *Phys. Lett. A* **418** 127702
[25] Zhou J P, Zhou Y Y, Gu R W and Zhou X J 2022 *Int. J. Quantum Inf.* **20** 2250005
[26] Namiki R and Hirano T 2006 *Phys. Rev. A* **74** 032302
[27] Lim C C W, Curty M, Walenta N, Xu F and Zbinden H 2014 *Phys. Rev. A* **89** 022307
[28] Curty M, Xu F, Cui W, Lim C C W, Tamaki K and Lo H K 2014 *Nat. Commun.* **5** 3732
[29] Mao C C, Zhou X Y, zhu J R, Zhang C H, Zhang C M and Wang Q 2018 *Opt. Express* **26** 13289
[30] Zhang Z, Zhao Q, Razavi M and Ma X 2017 *Phys. Rev. A* **95** 012333
[31] Ding H J, Mao C C, Zhang C M and Wang Q 2018 *Quantum Inf. Process.* **17** 332
[32] Yu W, Zhou Y Y and Zhou X J 2021 *Chin. J. Quantum Electron.* **38** 37
[33] Zhou J P, Zhou Y Y, Zhou X J and Nie N 2021 *J. Univ. Electron. Sci. Technol. China* **50** 650