2020 IEEE 4th Information Technology,Networking,Electronic and Automation Control Conference (ITNEC 2020)

# Study on Statistical Analysis Method of Decoy-state Quantum Key Distribution with Finite-length Data

Wei Yu, Yuanyuan Zhou*, Xuejun Zhou, Lei Wang, Shang Chen
College of Electronic Engineering, Naval University of Engineering, Wuhan 430033, China
* Correspondence: 754093499@qq.com

**Abstract—In order to solve the statistical fluctuation problem caused by the finite data length in the practical quantum key distribution system, four commonly used statistical methods, DeMoivre-Laplace theorem, Chebyshev inequality, Chernoff boundary and Hoeffding boundary, are used to analyze. The application conditions of each method are discussed, and the effects of data length and confidence level on quantum key distribution security performance are simulated and analyzed. The simulation results show that the applicable conditions of Chernoff boundary are most consistent with the reality of the practical quantum key distribution system with finite-length data. Under the same experimental conditions, the secure key generation rate and secure transmission distance obtained by Chernoff boundary are better than those of the other three methods. When the data length and confidence level change, the stability of the security performance obtained by the Chernoff boundary is the best.**

***Keywords—quantum key distribution; finite-length data; statistical analysis; security performance***

## I. Introduction

Quantum key distribution (QKD) [1], as a combination of one-time pad (OTP) cryptography and quantum mechanics fundamentals, has become an information security technology in quantum information science.

The security of QKD system under ideal conditions has been proved [2-5]. However, there is still a big gap between the actual QKD system and the ideal condition, such as the lack of ideal single-photonic light source, the low detection efficiency of detector and other non-ideal factors, which places a serious hidden danger on the security of the actual QKD system. Fortunately, these problems have been solved by scholars' efforts. In 2003, Hwang proposed a decoy scheme [6], which not only guarantees the unconditional security of QKD system but -also improves its transmission performance [7-9]. At present, the decoy state quantum key distribution scheme has become a typical scheme of QKD theory and experimental research.

The security performance analysis of the decoy scheme essentially comes from the GLLP analysis method proposed by Gottesman et al. [10]. The basic assumption of this method is that the data exchanged between the two sides of the communication is infinite. However, the data length processed by the actual QKD system in a certain period of time is certainly limited, which will lead to the statistical fluctuation of the data, thus reducing the secure key generation rate and secure transmission distance. Therefore, it is of great significance to study the security performance limit of decoy-state QKD system under the condition of finite-length data. In 2005, Ma Xiongfeng analyzed the statistical fluctuation of finite-length data decoy-state QKD system for the first time based on DeMoivre-Laplace theorem [11]. This method requires each experiment to be independent and the same distribution and puts forward the requirement of sample size. Many subsequent statistical fluctuation analyses of QKD schemes is based on this theorem [12-14]. In reference [15], the Chernoff boundary is applied to the analysis of finite-length data QKD systems, and the problem of limiting sample size is solved. In reference [16], a statistical analysis method based on Chebyshev is proposed, which only requires the expectation of random variables to be bounded. Most of the above studies only apply a certain statistical analysis method to the statistical fluctuation analysis of QKD system, but combined with the characteristics of the practical QKD system and the applicable conditions of the above common statistical analysis methods, which method has better QKD security performance boundary is not systematic analysis.

In this paper, the performance of the decoy QKD scheme under finite-length data conditions will be studied by using the DeMoivre-Laplace theorem, Chebyshev inequality, Chernoff boundary and Hoeffding boundary. systematically combing the characteristics and applicable conditions of the above analysis methods to obtain the one that is more suitable for performance analysis of finite-length data decoy QKD scheme.

## II. Decoy-state QKD Scheme

In this paper, a typical two-decoy-state BB84 [17] protocol based on weak coherent source(WCS) is adopted. The average number of photons of the signal state light source is set as u, and the average number of photons of the two deception-state light sources are set as $v_1$, $v_2$ respectively which satisfy: $v<< u$, $v= 0$.

According to the GLLP security analysis theory, the formula for calculating the generation rate of security key [10] is

$$R \geq q\left[-f\left(E_u\right)Q_u H_2\left(E_u\right)+Q_0+Q_1\left(1-H_2\left(e_1\right)\right)\right] \quad (1)$$

Among this formula, $R$ represents quantum key generation rate, $q$ represents the base efficiency, $f(E_u)$ represents the practical error correction algorithm efficiency, $Q_u$ represents the total gain of photon state whose average photon number is

$u$, $E_u$ represents total error rate of photon state whose average photon number is $u$, $Q_0$ represents the gain of empty-photon state, $Q_1$ represents the gain of single-photon state, $e_1$ represents error rate of single-photon state, $H_2(x)$ represents binary Shannon entropy. $Q_u$, $E_u$ and $Q_0$ can be directly observed in the experiment. In order to estimate the key generation rate of the practical QKD system, the lower limit of $Q_1$ and the upper limit of $e_1$ are need to be known.

The photon number probability function of WCS is subject to Poisson distribution

$$P(X=k)=\frac{u^k}{k!}e^{-u},\ k=0,1,... \tag{2}$$

Channel transmittance is given by

$$t_{AB}=10^{-\alpha L/10} \tag{3}$$

Among of it, $\alpha$ is the channel loss, and $L$ is the communication distance.

Detector efficiency $\eta$ is given by

$$\eta=t_{AB}\eta_{Bob} \tag{4}$$

$$\eta_i=1-(1-\eta)^i \tag{5}$$

Among of it, $\eta_{Bob}$ is the detector efficiency of receiver Bob, $\eta_i$ detector efficiency of i-photon state.

The yield $Y_i$ represents the conditional probability detected by Bob in the case that the sender Alice emits i-photon state, and the calculation formula is

$$Y_i=Y_0+\eta_i-Y_0\eta_i \tag{6}$$

The gain of the i-photon state is given by

$$Q_i=Y_i\frac{u^i}{i!}e^{-u} \tag{7}$$

When the light intensity is $u$, the total gain is given by

$$Q_u=\sum_{i=0}^{\infty}Q_i=\sum_{i=0}^{\infty}Y_i\frac{u^i}{i!}e^{-u}=\mathrm{Y}_0+1-e^{-\eta u} \tag{8}$$

The total quantum bit error rate is given by

$$E_uQ_u=\sum_{i=0}^{\infty}e_iY_i\frac{u^i}{i!}e^{-u}=e_0Y_0+e_\mathrm{d}(1-e^{-\eta u}) \tag{9}$$

$e_0$ is the background error rate, and $e_d$ is the detector bit error rate.

According to formula (1), in order to calculate the lower boundary of the secure key generation rate, it is necessary to obtain the upper boundary of the gain $Q_1$ of the single-photon state and the error rate $e_1$ of the single-photon state:

$$Q_1=\frac{u^2e^{-u}}{uv-v^2}\left(Q_ve^v-\frac{v^2}{u^2}e^uQ_u-\frac{u^2-v^2}{u^2}Y_0\right) \tag{10}$$

$$e_1=\frac{\left(E_vQ_ve^v-e_0Y_0\right)ue^{-u}}{Y_1vQ_u} \tag{11}$$

## III. Statistical Analysis Method of QKD Scheme with Finite-length Data

When the number of signal pulses is finite, the upper boundary of $Q_u$ and the lower boundary of $Q_v$ and $Q_0$ need to be estimated according to the statistical fluctuation of $Q_u$, $Q_v$ and $Q_0$, so as to obtain the lower boundary of $Q_1$ and the upper boundary of $e_1$. Due to the influence of statistical fluctuation, the lower boundary value of $Q_1$ will inevitably decrease and the upper boundary value of $e_1$ will inevitably increase, so the lower boundary value of the secure key generation rate will also decrease, and the security performance of QKD scheme will correspondingly decrease. In view of this problem, the main statistical analysis methods at present include the DeMoivre-Laplace theorem, the Chebyshev inequality, the Chernoff boundary and the Hoeffding boundary. This paper will use these four analysis methods respectively to estimate the upper boundary of $Q_u$ and the lower boundary of $Q_v$ and $Q_0$.

Taking the statistical fluctuation analysis of the total detection rate $Q_u$ of the signal state as an example, the detection event of each signal pulse is denoted as $X_i$。 If it is detected by detector of Bob is denoted as $X_i$=1, then

$$\Pr\left(X_i=1\right)=\sum_{k=1}^{\infty}Y_i\frac{u^k}{k!}e^{-u}=Q_u,\ i=1,\ 2,..,n \tag{12}$$

It is supposed that $X=\sum_{i=1}^{n}X_i, i=1,2,...,n$, so

$$E[X]=nQ_u \tag{13}$$

$$D[X]=nQ_u\left(1\text{-}Q_u\right) \tag{14}$$

$$\sigma[X]=\sqrt{nQ_u\left(1\text{-}Q_u\right)} \tag{15}$$

### *A. Statistical Fluctuation Analysis Based on DeMoivor-Laplace Theorem*

**Theorem 1** (DeMoivor-Laplace theorem) If $X_1$, $X_2$, …$X_n$ are the same distribution random variables which subject to Bernoulli distribution Pr($X_i$=1)=$p$ and are independent of each other, where i=1, 2, …$n$, then for any finite interval ($a$, $b$]

$$\lim_{n\to\infty}\Pr\{a<\frac{X_n-np}{\sqrt{np\left(1-p\right)}}\le b\}=\int_a^b\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}dt \tag{16}$$

The theorem shows that the normal distribution is the limit distribution of the binomial distribution. The above formula can be used to calculate the probability of binomial distribution when the number of experiments is enough.

The upper boundary of $Q_u$ is calculated from theorem 1.

$$Q_u^{\ U}=Q_u\left(1+\mathrm{u}_a\sqrt{\frac{1\text{-}Q_u}{nQ_u}}\right) \tag{17}$$

Where $u_a$ is a multiple of the standard deviation. It is given by according to the same theorem

$$Q_u^{\ U} = Q_u\left(1+\mathrm{u}_a\sqrt{\frac{1\text{-}Q_u}{N_uQ_u}}\right) \tag{18}$$

$$Q_v^{\ L} = Q_v\left(1+\mathrm{u}_a\sqrt{\frac{1\text{-}Q_u}{N_vQ_u}}\right) \tag{19}$$

$$Q_v^{\ L} = Q_v\left(1+\mathrm{u}_a\sqrt{\frac{1\text{-}Q_u}{N_vQ_u}}\right) \tag{20}$$

$$Q_0^{\ L} = Y_0^{\ L}e^{-u}\left(1-\mathrm{u}_a\sqrt{\frac{1\text{-}Q_0}{N_0Q_0}}\right) \tag{21}$$

### *B. Statistical Fluctuation Analysis Based on Chebyshev Inequality*

**Theorem 2** (Chebyshev inequality) For any random variable whose expectation is bounded

$$\Pr\{\left|X-E[X]\right|\geq c\}\leq\frac{D[X]}{c^2} \tag{22}$$

That's true for all $c$>0.

The deviation $\varepsilon$ of $Q_u$ is calculated by theorem 2.

$$\varepsilon=\sqrt{\frac{D[X]}{\theta}}\cdot\frac{1}{E[X]}=\sqrt{\frac{nQ_u(1-Q_u)}{\theta}}\cdot\frac{1}{nQ_u}=\sqrt{\frac{1-Q_u}{\theta nQ_u}} \tag{23}$$

where the confidence level is 1-$\theta$.

$$Q_u^{\ U} = Q_u\cdot\left(1+\varepsilon_{Q_u}\right),\varepsilon_{Q_u}=\sqrt{\frac{1-Q_u}{\theta N_uQ_u}} \tag{24}$$

$$Q_v^{\ L} = Q_u\cdot\left(1-\varepsilon_{Q_u}\right),\varepsilon_{Q_v}=\sqrt{\frac{1-Q_v}{\theta N_vQ_v}} \tag{25}$$

$$Y_0^{\ L} = Y_0\cdot\left(1-\varepsilon_{Q_u}\right),\varepsilon_{Y_0}=\sqrt{\frac{1-Y_0}{\theta N_0Y_0}} \tag{26}$$

$$Q_0^{\ L} = Y_0^{\ L}e^{-u}\left(1-e^{-u}\right),\varepsilon_{Q_0}=\sqrt{\frac{1-Q_0}{\theta N_0Q_0}} \tag{27}$$

### *C. Statistical Fluctuation Analysis Based on Chernoff Boundary*

**Theorem 3** (chernoff boundary) If $X_1$, $X_2$, …$X_n$ are the same distribution random variables which subject to Bernoulli distribution Pr($X_i$=1)=$p$ and are independent of each other, where i=1, 2, …$n$. It is supposed that $c=\sum_{i=1}^{n}E\left[X_i\right]$ ,so for any $\delta>0$

$$\Pr\{\sum_{i=1}^{n}X_i\geq\left(1+\delta\right)\cdot c\}\leq\left[\frac{e^{\delta}}{\left(1+\delta\right)^{1+\delta}}\right]^c<e^{-c\cdot\frac{\delta^2}{2}} \tag{28}$$

$$\Pr\{\sum_{i=1}^{n}X_i\geq\left(1\text{-}\delta\right)\cdot c\}\leq\left[\frac{e^{\text{-}\delta}}{\left(1\text{-}\delta\right)^{1\text{-}\delta}}\right]^c<e^{-c\cdot\frac{\delta^2}{2}} \tag{29}$$

The deviation $\varepsilon$ of $Q_u$ is calculated by theorem 3.

$\varepsilon=\sqrt{\frac{2\left(\ln 2-\ln\theta\right)}{nQ_u}}$, where the confidence level is 1-$\theta$.

$$Q_u^{\ U} = Q_u\left(1+\varepsilon_{Q_u}\right),\varepsilon_{Q_u}=\sqrt{\frac{2\left(\ln 2-\ln\theta\right)}{N_uQ_u}} \tag{30}$$

$$Q_v^{\ L} = Q_u\cdot\left(1-\varepsilon_{Q_u}\right),\varepsilon_{Q_v}=\sqrt{\frac{2\left(\ln 2-\ln\theta\right)}{N_vQ_v}} \tag{31}$$

$$Y_0^{\ L} = Y_0\cdot\left(1-\varepsilon_{Q_u}\right),\varepsilon_{Y_0}=\sqrt{\frac{2\left(\ln 2-\ln\theta\right)}{N_0Y_0}} \tag{32}$$

$$Q_0^{\ L} = Y_0^{\ L}\cdot e^{-u}\cdot\left(1-\varepsilon_{Q_0}\right),\varepsilon_{Q_0}=\sqrt{\frac{2\left(\ln 2-\ln\theta\right)}{N_0Q_0}} \tag{33}$$

### *D. Statistical Fluctuation Analysis Based on Hoeffding Boundary*

**Theorem 4** (Hoeffding boundary)If $X_1$, $X_2$, …$X_n$ are the same distribution random variables which subject to Bernoulli distribution Pr($X_i$=1)=$p$ and are independent of each other, where i=1, 2, …$n$, so for any $a>0$

$$\Pr\{\frac{1}{n}\sum_{i=1}^{n}X_i\geq p+a\}\leq e^{-2na^2} \tag{34}$$

$$\Pr\{\frac{1}{n}\sum_{i=1}^{n}X_i\leq p\text{-}a\}\leq e^{-2na^2} \tag{35}$$

The deviation $\varepsilon$ of $Q_u$ is calculated by equation (12) and theorem 4.

$\varepsilon=\sqrt{\left(\ln\frac{\theta}{2}\right)/\left(-2n\right)}$, where the confidence level is 1-$\theta$.

$$Q_u^{\ U} = Q_u+\varepsilon_{Q_u},\varepsilon_{Q_u}=\sqrt{\frac{\ln\frac{\theta}{2}}{\text{-}2N_u}} \tag{36}$$

$$Q_v^{\ L} = Q_v-\varepsilon_{Q_v},\varepsilon_{Q_v}=\sqrt{\frac{\ln\frac{\theta}{2}}{\text{-}2N_v}} \tag{37}$$

$$Y_0^{\ L} = Y_0\text{-}\varepsilon_{Y_0},\varepsilon_{Y_0}=\sqrt{\frac{\ln\frac{\theta}{2}}{\text{-}2N_0}} \tag{38}$$

$$Q_0{}^L = Y_0{}^L \cdot \mathrm{e}^{-u} \text{-} \varepsilon_{Q_0}, \varepsilon_{Q_0} = \sqrt{\frac{\ln\frac{\theta}{2}}{-2N_0}} \tag{39}$$

By analyzing the applicable conditions of the above four methods, it can be found that the DeMoivor-Laplace theorem requires that all random variables obey Bernoulli distribution and are independent of each other and uniformly distributed. Moreover, the larger the sample size is, the more accurate the statistical analysis will be. Chebyshev inequality only requires the expectation of the random variable to be bounded. The Chernoff boundary requires that the random variables obey Bernoulli distribution and are independent of each other. The applicable conditions of Hoeffding bound are the same as that of Chernoff boundary. Compared with its applicable conditions, Chebyshev inequality is the most relaxed. The application conditions of the Chernoff and Hoeffding boundary are the same, and the random variables are required to obey Bernoulli distribution and be independent of each other. On the basis of the former, the DeMoivor-Laplace theorem also puts forward a higher requirement on the sample size.

## IV. Simulation and Analysis

The above statistical analysis method is applied to the two-decoy-state BB84 protocol, and the experimental parameters used in the simulation are from the GYS experiment, as shown in Table 1. The following simulation selects the optimal signal state strength according to the transmission distance.

TABLE 1 EXPERIMENTAL PARAMETERS

| $a$（dB/km） | $e_d$（%） | $Pd$ | $\eta_{Bob}$ | $N$ | $N_0$ | $N_u$ | $N_v$ |
|---|---|---|---|---|---|---|---|
| 0.21 | 3.3 | $1.7 \times 10^{-6}$ | 0.045 | $6\times 10^9$ | $3.98\times 10^9$ | $1.76\times 10^9$ | $2.52\times 10^8$ |

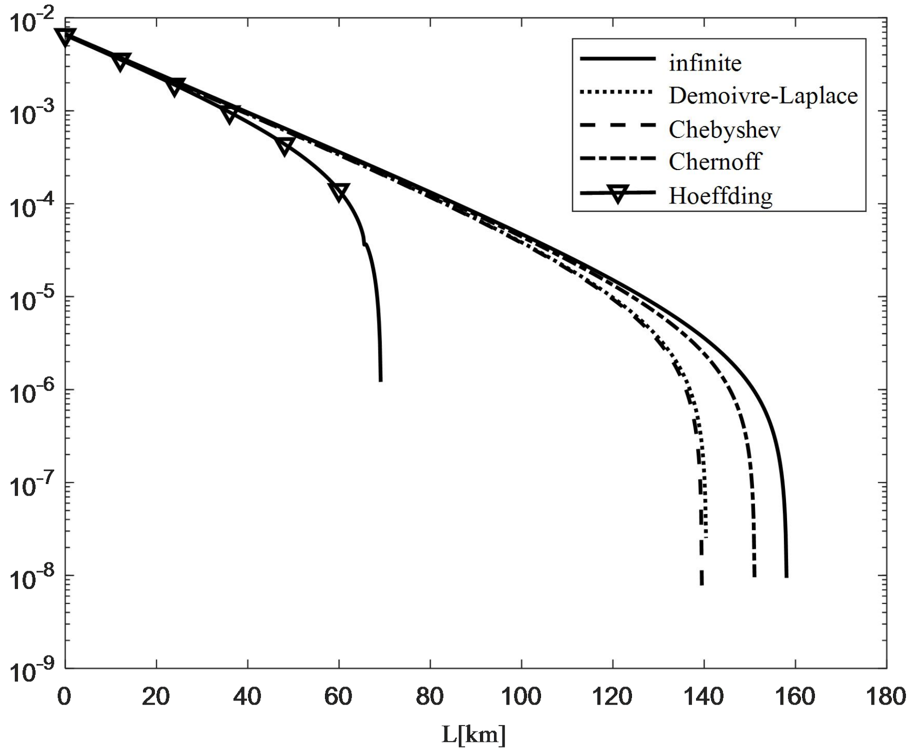
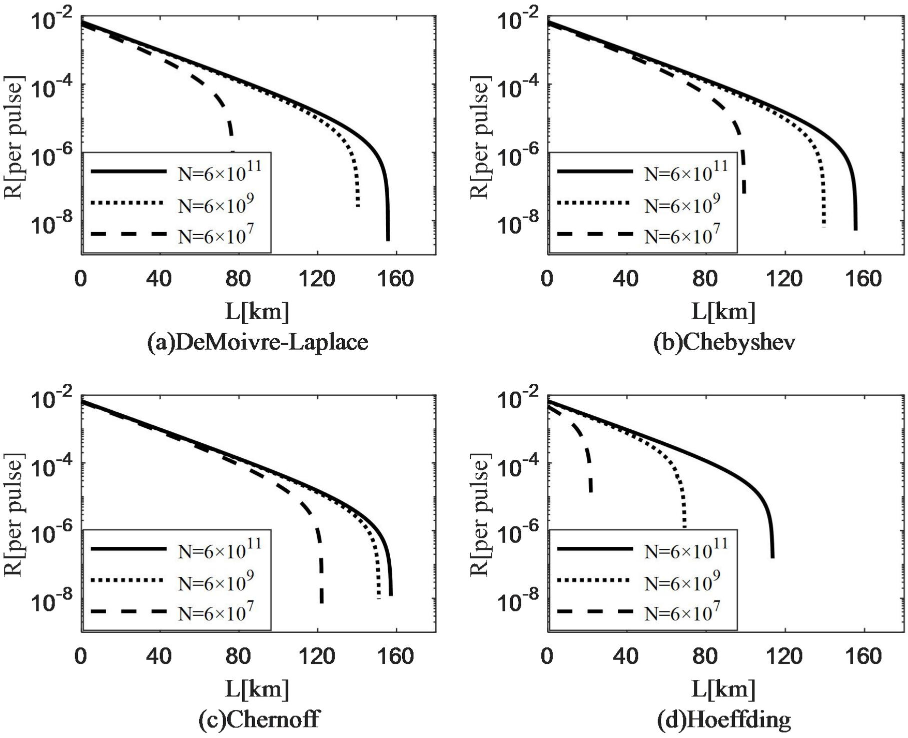


Fig. 1. Comparison of QKD security performance based on different statistical analysis methods

Figure 1 simulates the QKD security performance curve based on four statistical analysis methods under the same conditions. The confidence level is set at 1-8.7 × $10^{-3}$ and the data length is set at 6 × $10^9$. It can be seen from the figure that: compared with the infinite data QKD system, the secure key generation rate and the secure transmission distance of all the finite data QKD systems obtained by different statistical analysis methods are reduced. In the QKD system with finite-length data, the security performance obtained by using the Chernoff boundary is the best, and the result obtained by adopting the DeMoivor-Laplace theorem is not much different from that obtained by the Chebyshev inequality, the security performance of the former is slightly higher than that of the latter. While the security performance obtained by using the Hoeffding boundary is significantly lower than that of the other three methods.



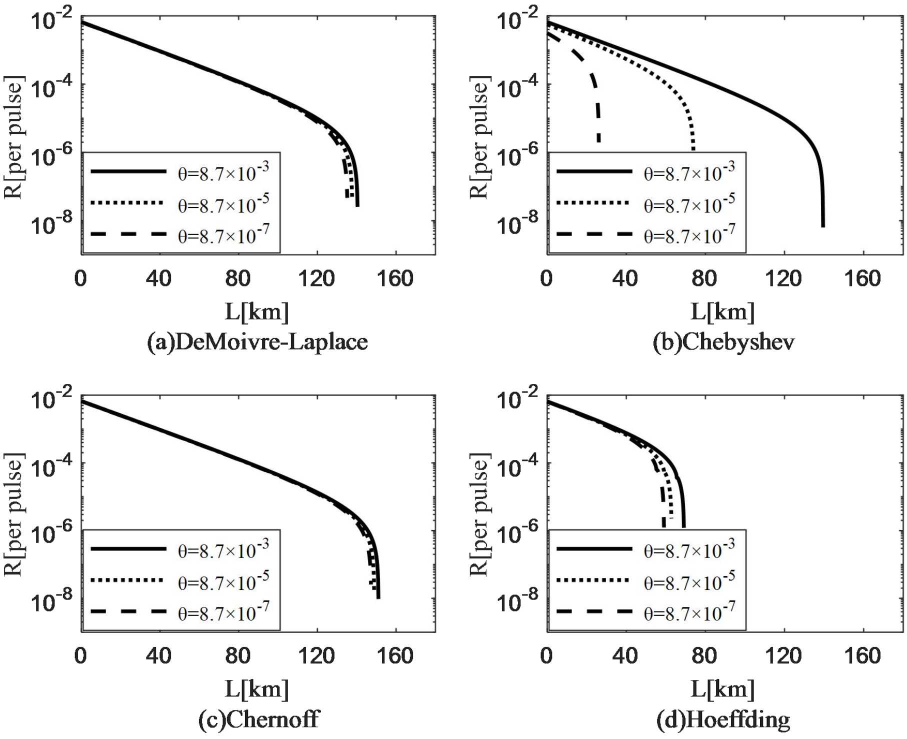
(a)DeMoivre-Laplace

(b)Chebyshev

(c)Chernoff

(d)Hoeffding

Fig. 2. Comparison of QKD security performance based on different statistical analysis methods under different data length conditions



(a)DeMoivre-Laplace

(b)Chebyshev

(c)Chernoff

(d)Hoeffding

Fig. 3. Comparison of QKD security performance based on different statistical analysis methods under different confidence level conditions

Figure 2 simulates the QKD security performance curve based on four statistical analysis methods under different data length conditions. The confidence level is set at 1-8.7 × $10^{-3}$, with three data lengths of N=6 × $10^{11}$, N=6 × $10^{9}$ and N=6 × $10^{7}$, respectively. It can be seen from the figure that with the decrease of data length, the security performance of QKD system is gradually reduced; when the data length is reduced from N=6 × $10^{11}$ to N=6 × $10^{7}$, the safe transmission distance of the DeMoivre-Laplace theorem, the Chebyshev inequality, the Chernoff boundary and the Hoeffding boundary is reduced by 80km, 65km, 35km and 95km. The ability of keeping the original security under the change of data length from high to low is the Chernoff boundary, the Chebyshev inequality, DeMoivre-Laplace theorem, the Hoeffding boundary. Therefore, the performance of the Chernoff boundary is the best.

Figure 3 simulates the QKD security performance curves based on four statistical analysis methods under different confidence level. The data length here is set as n = 6 × $10^{9}$, with three confidence level levels of 1-8.7 × $10^{-3}$, 1-8.7 × $10^{-5}$ and 1-8.7 × $10^{-7}$, respectively. It can be seen from the figure that with the increase of confidence level, the security performance of QKD system decreases gradually. When the confidence level is increased from 1-8.7 × $10^{-3}$ to 1-8.7 × $10^{-7}$, the secure transmission distance of the DeMoivre-Laplace theorem, the Chebyshev inequality, the Chernoff boundary and the Hoeffding boundary successively decreases by 8km, 110km, 5km and 10km. The ability of keeping the original security under the change of confidence level from high to low is the Chernoff boundary, the DeMoivre-Laplace theorem, the Hoeffding boundary and the Chebyshev inequality. The Chernoff boundary still performs the best.

## V. CONCLUSIONS

The signal pulse and data length are always limited in practical QKD system, therefore, the statistical analysis method of statistical fluctuation problem has a great influence on the security performance of finite-length data QKD system. In order to solve this problem, a decoy-state BB84 protocol based on WCS, four statistical analysis methods, the DeMoivre-Laplace theorem, the Chebyshev inequality, the Chernoff boundary and the Hoeffding boundary, are compared. Through the theoretical analysis, it can be concluded that the applicable conditions of the Chernoff boundary meet the practical QKD system. The simulation results show that the QKD security performance obtained by Chernoff boundary is the best under the same conditions, and when the data length decreases and the confidence level increases, the QKD security performance obtained by Chernoff boundary is more stable than the other three methods.

## REFERENCES

[1] Lo, Hoi-Kwong, Curty, Marcos, Tamaki, Kiyoshi. Secure quantum key distribution[J]. Nature Photonics, 8(8):595-604.

[2] Leverrier, Anthony, García-Patrón, Raúl, Renner, Renato. Security of Continuous-Variable Quantum Key Distribution Against General Attacks[J]. Physical Review Letters, 2013, 110(3):030502.

[3] Moroder T, Curty M, Lim C C W, et al. Security of Distributed-Phase-Reference Quantum Key Distribution[J]. Physical Review Letters, 2012, 109(26):260501.

[4] Tomamichel M, Renner R, Uncertainty Relation for Smooth Entropies[J]. Physical Review Letters, 2011, 106(11):110506.

[5] Woodhead, Erik. Quantum cloning bound and application to quantum key distribution[J]. Physical Review A, 2013, 88(1):012331.

[6] Hwang, Won-Young. Quantum Key Distribution with High Loss: Toward Global Secure Communication[J]. Physical Review Letters, 91(5):057901.

[7] Peng C Z, Liang H, Wang J, et al. Decoy-state quantum key distribution with polarized photons over 200 km[J]. 2010, 18(8):8587-8594.

[8] Y. L. Tang, H. L. Yin, S. J. Chen, et al. Measurement-device-independent quantum key distribution over 200 km[J]. Phys. Rev. Lett. 2014, 113: 190501.

[9] H. L. Yin, T. Y. Chen, Z. W. Yu, et al. Measurement-device-independent quantum key distribution over 404 km optical fiber[J]. Phys. Rev. Lett., 2016, 117(19): 190501.

[10] Gottesman, Daniel, Lo, Hoi-Kwong, Lütkenhaus, Norbert, et al. Security of quantum key distribution with imperfect devices[J].

[11] X. F. Ma, B. Qi, Y. Zhao, et al. Practical decoy state for quantum key distribution[J]. Phys. Rev. A, 2005, 72: 012326.

[12] Xuan Wen, Qiong Li, Hongjuan Wang, et al. A Bayesian based finite-size effect analysis of QKD[M]. Springer International Publishing, 2017.

[13] Zhu J R, Li J, Zhang C M, et al. Parameter optimization in biased decoy-state quantum key distribution with both source errors and statistical fluctuations[J]. 2017, 16(10):238.

[14] Song, Ting-Ting, Qin, Su-Juan, Wen, Qiao-Yan, et al. Finite-key security analyses on passive decoy-state QKD protocols with different unstable sources[J]. 2018, Scientific Reports, 5:15276.

[15] Zhichao Wei, Ming Gao, Zhi Ma. Statistical fluctuation analysis method of quantum key distribution based on Chernoff boundary [J]. Journal of information engineering university, 2014, 15(4):399-404.

[16] Haodong Jiang, Ming Gao, Zhi Ma. Research on statistical fluctuation analysis method in decoy-state quantum key distribution protocol [J]. Journal of information engineering university, 2016(6):694-697, 4 pages.

[17] Lo H K, Ma X, Chen K. Decoy state quantum key distribution [J]. Phys. Rev. Lett. 2005, 94(23): 230504.