I. INTRODUCTION

Device-independent (DI) quantum cryptography has caught great interests recently \[1\]–\[10\]. It aims to replace the model of the physical devices used in the cryptographic protocol with physically testable assumptions, e.g., the certification of nonlocality. Thus the devices can be treated as black boxes that produce outputs correlated with some inputs. This brings the advantage that the assumptions needed to guarantee the security of the protocol can be significantly reduced, so that the knowledge of the internal workings of the devices is not required. The protocol remains reliable even if the devices are provided by the adversary. Such a higher degree of security makes DI protocols more dependable in practical applications than traditional quantum cryptography.

In many recent proposals on quantum key distribution (QKD) protocol \[11\]–\[17\], the nonlocality is tested by observing the violation of the Clauser-Horne-Shimony-Holt (CHSH) inequality. However, being a statistic law, a certain extent of fluctuation has to be allowed. We show that the eavesdropper can make use of this property to obtain a remarkable part of the secret key by replacing some of the DI nonlocal boxes with local ones. On the contrary, the same cheating strategy does not apply to the device-dependent (DD) version of the protocol. Thus such kind of DI protocol is less secure than its DD counterpart.

II. THE DD QKD PROTOCOL

Following the notations of \[12\]–\[13\], a particular DD implementation of the QKD protocols based on the CHSH inequality can be accomplished in the following manner. Alice and Bob share a quantum channel consisting of a source that supplies many pairs of entangled particles, each of which is supposed to be in the Bell state

\[
\langle \Phi^+ \rangle = \frac{|0\rangle_A |0\rangle_B + |1\rangle_A |1\rangle_B}{\sqrt{2}},
\]

where particle \( A \) (\( B \)) goes to Alice (Bob), with \( |0\rangle \) and \( |1\rangle \) denoting the two eigenstates of the Pauli operator \( \sigma_z \).

On each of their particles, Alice chooses randomly among three measurements

\[
A_0 = \sigma_z, \\
A_1 = (\sigma_z + \sigma_x)/\sqrt{2}, \\
A_2 = (\sigma_z - \sigma_x)/\sqrt{2},
\]

while Bob chooses randomly between two measurements

\[
B_1 = \sigma_z, \\
B_2 = \sigma_x.
\]

The outcomes of the measurements are labeled as \( a_i, b_j \in \{+1, -1\} \) \( i = 0, 1, 2, j = 1, 2 \). Alice and Bob announce their inputs (i.e., their choices of the measurements) through a classical channel, which can be insecure. To check whether the states indeed have the form of Eq. (1), they gather the outcomes when Alice did not choose \( A_0 \), and compute the CHSH polynomial

\[
S = \langle a_1 b_1 \rangle + \langle a_1 b_2 \rangle + \langle a_2 b_1 \rangle - \langle a_2 b_2 \rangle,
\]
where the correlator \( \langle a_i b_j \rangle \) is defined as the probability \( P(a = b|ij) - P(a \neq b|ij) \). To generate a raw secret key, they use the outcomes when Alice chose \( A_0 \) and Bob chose \( B_1 \). When no eavesdropping present, there should always be \( a_0 = b_1 \), while the CHSH value reaches the point of maximal quantum violation of the well-known CHSH Bell inequality, i.e., the correlators satisfy \( \langle a_1 b_1 \rangle = \langle a_1 b_2 \rangle = \langle a_2 b_1 \rangle = -\langle a_2 b_2 \rangle = 1/\sqrt{2} \), so that \( S = 2\sqrt{2} \).

III. THE DI QKD PROTOCOL

In brief, the typical structure of the DI version of the above QKD protocol is as follows. Alice and Bob share many pairs of devices called nonlocal boxes, which can be supplied by either Alice or Bob, or even an untrustworthy third party, including the eavesdropper Eve. These devices can be treated as DI black boxes that take inputs and produce outputs, without the need nor the possibility to check how they work internally. Each of Alice’s (Bob’s) boxes has three (two) inputs \( A_0, A_1 \) and \( A_2 \) (\( B_1 \) and \( B_2 \)). For each input, Alice’s (Bob’s) box can produce either of the two outputs \( a_i = \pm 1 \) (\( b_j = \pm 1 \)). When no cheating present, i.e., the boxes are manufactured honestly and working properly, Alice’s and Bob’s inputs to each pair of their boxes should display the maximal nonlocality, as can be checked by the CHSH value in Eq. (3).

In the protocol, Alice and Bob choose their inputs into the boxes randomly, and then announce the inputs through a classical channel. Whenever Alice chooses \( A_0 \) and Bob chooses \( B_1 \) for the same pair, the outputs should satisfy \( a_0 = b_1 \) so that they can use these pairs to generate a raw secret key. For other boxes, they gather the outcomes and calculate the CHSH value to detect the existence of eavesdropping. See [11, 13, 14] for other details of the protocols.

IV. STATISTICAL FLUCTUATIONS

It is worth noting that the maximal CHSH value \( S = 2\sqrt{2} \) is a statistical result, calculated from the mean values \( \langle a_i b_j \rangle \) (\( i, j = 1, 2 \)) of finite outcomes. Since any statistical property is inevitably subjected to fluctuation even in the ideal case, we cannot expect to find the actual result of an experiment exactly equal to the theoretical value.

As a simple example, let us consider the outcome of tossing a classical coin. Suppose that the coin is absolutely ideal, that the outcomes “heads” and “tails” will both occur with the probability 1/2 without any bias at all. Even so, tossing the coin twice does not necessarily result in 1 head and 1 tail. In fact this result will occur with the probability 50% only. In the rest 50% case, the ratio of heads to tails will be either 2 : 0 or 0 : 2. When tossing the coin \( n \) times, we cannot expect that the ratio will be perfectly \( n/2 : n/2 \) either. In fact, interestingly, this perfect result will occur with the probability \( \left( \frac{n}{2} \right)^2 / 2^n \) only, which drops as \( n \) increases. That is, the more samples involved in the statistics, the less we can expect to meet the theoretical value exactly. More importantly, as we mentioned, this is the result for an ideal coin. It remains valid even if we do not consider any experimental imperfection, e.g., biased coins, wind disturbance, etc. Therefore, in practice we have to accept a certain range of fluctuation deviating from the perfect result that we expect theoretically. For instance, when tossing the ideal coin for 10000 times, it is still acceptable if the ratio between heads and tails turns out to be 5500 : 4500, as this result can indeed occur with a probability not much less than that of the perfect result 5000 : 5000.

For the same reason, when checking the CHSH value in the DI QKD protocol, we can hardly expect to find \( S = 2\sqrt{2} \) even if all parties are completely honest. Even the actual value turns out to be a little lower, the secret key should still be considered secure. Otherwise the protocol will have only little probability to proceed, even when no eavesdropping present at all. Indeed, some papers suggested a precise acceptable lower bound for the CHSH value. For example, \( S \geq \frac{2}{\sqrt{5}} \) is considered acceptable in step 5 of the protocol in [17].

Note that the origin of this problem is not the channel noise studied in [11, 16], whose effect is, e.g., following the notation in [11], to transform the state Eq. (1) into the Werner state with the density matrix

\[
\rho = p \Phi^+ \langle \Phi^+ \rangle + (1 - p) \frac{I}{4},
\]

so that the expected CHSH value \( S \) will drop. Instead, the fluctuation considered here is a fundamental theoretical property of statistical quantities, which exists even when the probability \( p \) in the above equation vanishes.

V. SECURITY PROBLEMS

As \( S < 2\sqrt{2} \) has to be accepted in practice, there is more room for eavesdropping in such a DI QKD protocol, as elaborated below.

Consider that the eavesdropper Eve replaces the DI nonlocal boxes with a local box pair \( X_1 \), whose outcomes have a fixed relationship as

\[
X_1 : a_0 = a_1 = a_2 = b_1 = b_2,
\]

where \( a_0 \) is chosen beforehand by Eve to be either +1 or −1. That is, e.g., if \( a_0 = +1 \), then Alice will obtain the outcome \( a_1 = +1 \) when her input to her box is \( A_1 \), or she will obtain the outcome \( a_2 = +1 \) when her input is \( A_2 \). Meanwhile, Bob will obtain the outcome \( b_1 = +1 \) when his input to his box is \( B_1 \), etc.

Such a local box pair can be constructed, because in a DI protocol, Alice and Bob cannot assume that the inputs \( A_i \)'s and \( B_j \)'s must correspond to the operators...
in Eqs. (2) and (3). Instead, Eve can, e.g., prepares Alice’s and Bob’s box pair in the product state $|0\rangle_A |0\rangle_B$ or $|1\rangle_A |1\rangle_B$, and let all inputs be the operator $\sigma_z$. Then we can see that Eq. (5) can be met.

Similarly, there can be other local box pairs $X_2$, $X_3$ and $X_4$, whose outcomes satisfy

$$
\begin{align*}
X_2 : & \ a_0 = a_1 = -a_2 = b_1 = b_2, \\
X_3 : & \ a_0 = a_1 = a_2 = b_1 = -b_2, \\
X_4 : & \ a_0 = -a_1 = a_2 = b_1 = -b_2, 
\end{align*}
$$

(7)

where $a_0$ is also chosen beforehand by Eve. These box pairs can be constructed using non-entangled product states too. For example, $X_2$ can be obtained by preparing Alice’s and Bob’s box pair in the three-particle product state $|01\rangle_A |0\rangle_B$, and the inputs are set as $A_0 = A_1 = \sigma_z \otimes I$, $A_2 = I \otimes \sigma_z$, and $B_1 = B_2 = \sigma_z$.

From Eq. (4) we can easily find that the CHSH value for each of $X_1$, $X_2$, $X_3$ and $X_4$ is $S = 2$ as nonlocality is absent between Alice’s and Bob’s boxes. Also, if Eve uses equal numbers of each of the four box pairs, then the correlators $\langle a_1 b_1 \rangle$, $\langle a_1 b_2 \rangle$, $\langle a_2 b_1 \rangle$, and $\langle a_2 b_2 \rangle$ in Eq. (4) all equal to 2. Therefore no bias can be found even if these correlators are checked separately. Now since all the four box pairs satisfy $a_0 = b_1$, whenever Alice and Bob chooses such a pair to generate the raw secret key, Eve will know the corresponding secret bit, while Alice’s and Bob’s keys remain consistent with each other so that they find nothing wrong. The remaining question is whether Eve can pass the CHSH value check.

Obviously, as $X_1$, $X_2$, $X_3$ and $X_4$ all have the CHSH value $S = 2 < 2\sqrt{2}$, Eve cannot pass the check if she replaced all nonlocal boxes with these four. However, as we mentioned, in the DI QKD protocol a lower CHSH value has to be allowed due to the existence of statistical fluctuations. Therefore if Eve uses the local boxes for a small portion, then she will have a non-trivial probability to pass the check, while manage to learn a part of the secret key. It turns out that the amount of the key she learned is not too small, as we will show below.

Let $S_{\text{min}}$ denote the lower bound of the CHSH value allowed in the DI QKD protocol, and $S_n = 2\sqrt{2}$ be the CHSH value expected theoretically for the boxes displaying maximal nonlocality. Suppose that Eve replaces the nonlocal boxes using the above local ones with the probability $p$. As the local boxes have the CHSH value $S = 2$, while the rest nonlocal ones that Eve has not replaced are expected to give $S_n = 2\sqrt{2}$ averagely, the final expected value $S_e$ that Alice and Bob will find in their CHSH value check is

$$
S_e = 2p + S_n (1 - p).
$$

(8)

When $S_e \geq S_{\text{min}}$ Eve can pass the check successfully. In this case

$$
p \leq \frac{S_n - S_{\text{min}}}{S_n - 2}.
$$

(9)

This sets the maximal probability that Eve can use the local boxes to cheat. Since Eve will know the secret bit once a local box pair is chosen to generate the raw key, $p$ also describes the proportion of the raw key leaked to Eve. Substituting $S_{\text{min}} = 2.5$ (as suggested in [17]) and $S_n = 2\sqrt{2}$ into Eq. (9), we have $p \leq 39.64\%$. As Eve should use each of the four local boxes $X_1$, $X_2$, $X_3$ and $X_4$ with equal probabilities, this result means that she can use each box with the probability $9.9\%$, and learn $39.6\%$ of the raw secret key, which is far from being trivial.

Note that for the rest $1 - p = 60.4\%$ nonlocal boxes that Eve has not replaced, fluctuation also exists when calculating the CHSH value. Therefore Eve cannot guarantee that the actual CHSH value $S_n$ obtained in her above cheating will always satisfy $S_n > S_{\text{min}}$. But on one hand, the fluctuation can either raise or lower the actual value. Even if there is only about $1/2$ probability that Eve manages to learn $39.6\%$ of the raw secret key without being detected, it is still a serious problem to the DI QKD protocol. On the other hand, when the local boxes $X_1$, $X_2$, $X_3$ and $X_4$ take part in the calculation of the CHSH value, no fluctuation will take place as the outcomes of these boxes are deterministic. The average fluctuation range caused by the rest $60.4\%$ nonlocal boxes is surely smaller than that in the honest protocol where $100\%$ of the boxes are nonlocal. If Eve uses the local boxes with a slightly lower probability $p$, she can further raise the expected CHSH value $S_e$ so that there is even less chances that the fluctuation caused by the rest nonlocal boxes is sufficiently large to bring the actual value $S_n$ down below $S_{\text{min}}$.

VI. COMPARING WITH THE DD PROTOCOL

Intriguingly, Eve’s above cheating strategy is completely futile in the DD protocol. This is because in the DD scenario, Alice and Bob always know what are the exact measurements that correspond to their inputs to the states, as shown in section II. Eve has no chance to change the measurements, even if she can replace the states. When applying the above cheating strategy, the states she uses must be able to give a deterministic result when Alice and Bob choose them to generate the secret key. Therefore, the states have to be the eigenstates of Alice’s measurement $A_0$ and Bob’s $B_1$. As $A_0 = B_1 = \sigma_z$ are fixed in the DD protocol, the states must be either $|0\rangle_A |0\rangle_B$ or $|1\rangle_A |1\rangle_B$.

Now we calculate the CHSH values of $|0\rangle_A |0\rangle_B$ and $|1\rangle_A |1\rangle_B$ using Eq. (4). But we can no longer set all inputs as the operator $\sigma_z$ as we did in the DI case. Instead, the operators have to remain the forms in Eqs. (2) and (3). Consequently, we will find $S = \sqrt{2} < 2$. But more importantly, the correlators in $S$ will be $\langle a_1 b_1 \rangle = \langle a_2 b_2 \rangle = 1/\sqrt{2}$ while $\langle a_1 b_2 \rangle = \langle a_2 b_1 \rangle = 0$. On the contrary, when the state $|\Phi^+\rangle$ was used honestly, there should be $\langle a_1 b_1 \rangle = \langle a_2 b_2 \rangle = \langle a_2 b_1 \rangle = -\langle a_1 b_2 \rangle = 1/\sqrt{2}$. Therefore, even if Eve mixes the states $|0\rangle_A |0\rangle_B$ and $|1\rangle_A |1\rangle_B$ with $|\Phi^+\rangle$, then as long as $|0\rangle_A |0\rangle_B$ and $|1\rangle_A |1\rangle_B$ present with a non-trivial probability $p$, Alice
and Bob will be able to find a bias on the values of the correlators. That is, Eve cannot escape the detection if she wants to learn a non-trivial portion of the raw secret key.

VII. SUMMARY AND REMARKS

Thus we can see that in the DI QKD protocol based on the violation of CHSH inequality, a lower CHSH value has to be allowed due to the existence of statistical fluctuation. Then the eavesdropper can have a non-trivial probability to learn a remarkable amount of information on the raw secret key without being detected. Note once again that this fluctuation exists even in the ideal case. It is a fundamental theoretical problem of statistical properties, which is not caused by any experimental imperfection. Even if entangled states can be perfectly prepared again that this fluctuation exists even in the ideal case. It is impossible to remove it. Therefore, the raw secret key in DI QKD protocol has to be allowed due to the existence of statistical fluctuation. Then the eavesdropper can have a non-trivial probability to learn a remarkable amount of information on the raw secret key. Note once again that this fluctuation exists even in the ideal case. It is a fundamental theoretical problem of statistical properties, which is not caused by any experimental imperfection. Even if entangled states can be perfectly prepared.

Unfortunately, this security problem cannot be avoided by improving the experimental technology.

Of course, by increasing the total number $n$ of the DI boxes used in the protocol, the relative deviation from the theoretical expected value caused by the fluctuation will decrease, so that Alice and Bob can choose a higher $S_{\text{min}}$ value. Therefore it will lower the ratio $p$ between the number of Eve’s obtained bits and the entire raw secret key (even though the absolute number of bits that Eve obtained will still rise). Thus the DI QKD protocol is still secure in the limit $n \to \infty$. However, as showed in the previous section, the DD protocol is completely immune to the same cheating strategy of Eve, without requiring an infinite $n$. So we can see that quantitatively speaking, for any given finite $n$, the current DI QKD protocol is not as secure as its DD counterpart.

The work was supported in part by the NSF of China under grant No. 10975198, the NSF of Guangdong province, and the Foundation of Zhongshan University Advanced Research Center.

---

[1] J. Silman, A. Chailloux, N. Aharon, I. Kerenidis, S. Pirino, and S. Massar, Phys. Rev. Lett. 106, 220501 (2011). Fully distrustful quantum bit commitment and coin flipping

[2] H. -K. Lo, M. Curty, and B. Qi, Phys. Rev. Lett. 108, 130503 (2012). Measurement-device-independent quantum key distribution

[3] C. Branciard, E. G. Cavalcanti, S. P. Walborn, V. Scarani, and H. M. Wiseman, Phys. Rev. A 85, 010301(R) (2012). One-sided device-independent quantum key distribution: Security, feasibility, and the connection with steering

[4] K. Tamaki, H. -K. Lo, C. -H. F. Fung, and B. Qi, Phys. Rev. A 85, 042307 (2012). Phase encoding schemes for measurement-device-independent quantum key distribution with basis-dependent flaw

[5] M. Dall’Arno, E. Passaro, R. Gallego, and A. Acín, Phys. Rev. A 86, 042312 (2012). Robustness of device-independent dimension witnesses

[6] J. Barrett, R. Colbeck, and A. Kent, Phys. Rev. A 86, 062326 (2012). Unconditionally secure device-independent quantum key distribution with only two devices

[7] C. C. W. Lim, C. Portmann, M. Tomamichel, R. Renner, and N. Gisin, Phys. Rev. X 3, 031006 (2013). Device-independent quantum key distribution with local bell test

[8] H. -W. Li, et al., Phys. Rev. A 87, 020302(R) (2013). Relation between semi- and fully-device-independent protocols

[9] C. Branciard, D. Rosset, Y. -C. Liang, and N. Gisin, Phys. Rev. Lett. 110, 060405 (2013). Measurement-device-independent entanglement witnesses for all entangled quantum states

[10] S. -H. Sun, M. Gao, C. -Y. Li, and L. -M. Liang, Phys. Rev. A 87, 052329 (2013). Practical decoy-state-device-independent quantum key distribution

[11] A. Acín, S. Massar, and S. Pirino, New J. Phys. 8, 126 (2006). Efficient quantum key distribution secure against no-signalling eavesdroppers

[12] A. Acín, N. Brunner, N. Gisin, S. Massar, S. Pirino, and V. Scarani, Phys. Rev. Lett. 98, 230501 (2007). Device-independent security of quantum cryptography against collective attacks

[13] S. Pirino, A. Acín, N. Brunner, N. Gisin, S. Massar, and V. Scarani, New J. Phys. 11, 045021 (2009). Device-independent quantum key distribution secure against collective attacks

[14] M. McKague, New J. Phys. 11, 103037 (2009). Device independent quantum key distribution secure against coherent attacks with memoryless measurement devices

[15] L. Masanes, S. Pirino, and A. Acín, Nat. Commun. 2, 238 (2011). Secure device-independent quantum key distribution with causally independent measurement devices

[16] M. Lucamarini, G. Vallone, I. Gianani, P. Mataloni, and G. Di Giuseppe, Phys. Rev. A 86, 032325 (2012). Device-independent entanglement-based Bennett 1992 protocol

[17] J. Barrett, R. Colbeck, and A. Kent, Phys. Rev. Lett. 110, 010503 (2013). Memory attacks on device-independent quantum cryptography