Device-independent secret key rates via a postselected Bell inequality

Sarnava Datta,* Hermann Kampermann, and Dagmar Bruß
Institut für Theoretische Physik III, Heinrich-Heine-Universität Düsseldorf, D-40225 Düsseldorf, Germany

(Received 11 November 2021; accepted 22 February 2022; published 29 March 2022)

In device-independent quantum key distribution (DIQKD) the security is not based on any assumptions about the intrinsic properties of the devices and the quantum signals but on the violation of a Bell inequality. We introduce a DIQKD scenario in which an optimal Bell inequality is constructed from the performed measurement data rather than fixing beforehand a specific Bell inequality. Our method can be employed in a general way, for any number of measurement settings and any number of outcomes. We provide an implementable DIQKD protocol and perform finite-size security key analysis for collective attacks. We compare our approach with related procedures in the literature and analyze the robustness of our protocol. We also study the performance of our method in several Bell scenarios as well as for random measurement settings.

DOI: 10.1103/PhysRevA.105.032451

I. INTRODUCTION

Data security concerns are prevalent in the modern world. One of the most prominent domains of quantum communication is quantum key distribution (QKD), which allows to distribute a secure key between two (or more) parties, namely, Alice and Bob, where the security is only based on the laws of quantum mechanics. Since the inception of QKD [1], a variety of QKD protocols [2–12] have been introduced. However, the security of these device-dependent protocols needs complete characterization of the devices, sources, and/or the channel between the parties. In a realistic scenario, the device can be not completely characterized or could even be prepared by a malicious eavesdropper (Eve). Furthermore, hacking of existing implementations that exploit experimental imperfections was demonstrated [13–15]. To overcome these drawbacks, device-independent (DI) QKD was introduced [16], where the security does not require any assumptions about the inherent properties of the devices or the dimension of the Hilbert space of the quantum signals. The security of DI protocols is based on the observation of a loophole-free Bell inequality violation [17–29] which guarantees the quantum nature of the observed data. The length of the secret key will depend on the estimated violation of the Bell inequality.

In this article we introduce a DIQKD scenario in which the Bell inequality is not agreed upon beforehand but will be constructed from the observed probability distribution of the measurement outcomes. We follow a two-step process: From the input-output probability distribution, we construct a Bell inequality that leads to the maximum Bell violation for that particular measurement setting of Alice and Bob. Then we use this optimized Bell inequality and the corresponding violation to bound the secret key rate. Note that in [30,31], the authors introduced an alternative approach to bound the device-independent secret key rate via a Bell inequality and the corresponding violation, which is also constructed from the full measurement statistics. We will relate and compare our method with theirs in the Results section (Sec. VI). In particular, we show that our procedure is advantageous in the nonasymptotic regime.

This paper is organized as follows. We start in Sec. II by briefly reviewing classical and quantum correlations. Then we explain how to obtain the optimal Bell inequality from the observed probability distribution. We lay the framework to provide a confidence interval for the Bell expectation value in Sec. III. We provide an implementable DIQKD protocol in Sec. IV and calculate the finite-size secret key rate in Sec. V. In Sec. VI we illustrate our method with several examples.

II. GENERAL FRAMEWORK

In this section we review the concept of the classical correlation polytope in Sec. II A, and, based on this, we explain in Sec. II B how to construct Bell inequalities that are maximally violated by the measurement data.

A. Set of correlations

Consider a setup for two parties[1] (namely, Alice and Bob) connected by a quantum channel. The parties perform local measurements on a joint quantum state. Let us assume that Alice and Bob have m_a and m_b measurement settings, respectively. Alice's set of measurement settings is denoted as $X = \{1, \ldots, m_a\}$, and Bob's set of measurement settings as $Y = \{1, \ldots, m_b\}$. To estimate the probability distribution from the experimental data, we have to use the measurement device N times in succession. We assume that the devices behave independently and identically (i.i.d.) in each round, i.e., the results of the ith round are independent of the past $i-1$

*Sarnava.Datta@hhu.de

[1]Note that our method can be extended in a straightforward way to n parties.

rounds. The setting of the ith round is denoted as $x_i \in X$ for Alice and $y_i \in Y$ for Bob. Each of these measurement settings has d outcomes, which are denoted as $a_i \in A = \{1, \ldots, d\}$ for Alice and $b_i \in B = \{1, \ldots, d\}$ for Bob. We call this the $[(m_a, m_b), d]$ scenario, i.e., two parties with (m_a, m_b) measurement settings and d outcomes each. When both parties have an equal number of measurement settings, i.e., $m_a = m_b = m$, we will denote this as the $[m, d]$ scenario. The joint probability of getting outcome a when Alice is using the measurement setting $x \in X$ and b when Bob uses the measurement setting $y \in Y$ is denoted as $P(A_x^a B_y^b)$. All these joint probabilities will be collected in a probability vector

$$\mathbf{P} := \left[P\left(A_x^a B_y^b\right)\right], \tag{1}$$

where $x \in X$, $y \in Y$, $a \in A$, and $b \in B$. The associated probability space is of dimension

$$D^d_{m_a,m_b} := m_a m_b d^2. \tag{2}$$

The set of all probabilities that represent a classical or locally real theory forms a convex polytope [32–34]. We denote this polytope as $\mathcal{P}$. Any probability distribution which is not contained in $\mathcal{P}$ shows nonclassical or quantum behavior and can be witnessed by the violation of a Bell inequality [35]. As illustrated in [36], the polytope of classical correlations can be characterized by its extremal points $\mathbf{v}_p$, where $p = \{1, 2, \ldots, d^{m_a+m_b}\}$, and $\mathbf{v}_p$ has entries from the set $\{0,1\}$. The extremal points of the polytope correspond to deterministic strategies. Every classical correlation $\mathbf{P}_{cl} \in \mathcal{P}$ can be written as a convex combination of all the deterministic strategies as

$$\mathbf{P}_{cl} = \sum_{p=1}^{d^{m_a+m_b}} \lambda_p \mathbf{v}_p, \tag{3}$$

where $\lambda_p \geqslant 0$ and $\sum_{p=1}^{d^{m_a+m_b}} \lambda_p = 1$. This subsequently implies that every observed probability distribution which cannot be decomposed as shown in Eq. (3) violates at least one Bell inequality.

B. Designing Bell inequalities

Consider the $[(m_a, m_b), d]$ scenario where the parties receive the measurement data $\mathbf{P}$. In order to extract a secret key from these classical measurement data, they need to violate a Bell inequality. As shown in [36], this scenario can be translated to a linear separation problem. For illustration, see Fig. 1. Bell inequalities correspond to hyperplanes in the probability space that separate the classical correlation polytope $\mathcal{P}$ from the set of all genuine quantum correlations $\mathcal{Q} \setminus \mathcal{P}$. Such hyperplanes are specified by a normal vector $\mathbf{h} \in \mathbb{R}^{D^d_{m_a,m_b}}$ with the dimension given in Eq. (2). If $\mathbf{P} \in \mathcal{Q} \setminus \mathcal{P}$, there exists at least one hyperplane $\mathbf{h}$ that separates all the vertices $\mathbf{v}_p$ of $\mathcal{P}$ from the observed probability distribution $\mathbf{P}$. We set the objective of the linear program to find the hyperplane vector $\mathbf{h}$ corresponding to the Bell inequality which is maximally violated by the measurement data $\mathbf{P}$. This optimization problem can be formulated as

$$\begin{aligned} \max_{\mathbf{h},c} \quad & \mathbf{h}^T\mathbf{P} - c \\ \text{subject to} \quad & \mathbf{h}^T\mathbf{v}_p \leqslant c \quad \forall \quad p \in \{1, \ldots, d^{m_a+m_b}\} \end{aligned}$$

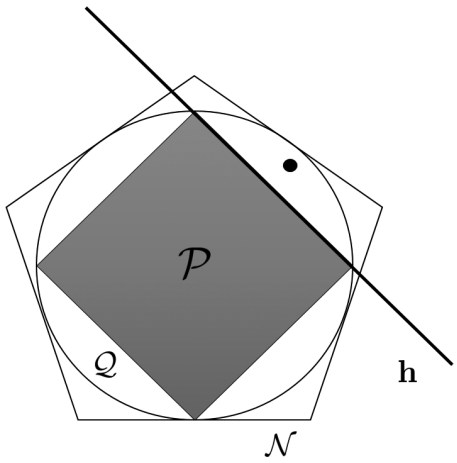

FIG. 1. A sketch for the set of correlations. All classical probabilities form a convex polytope $\mathcal{P}$, which is embedded in the set $\mathcal{Q}$ of quantum correlations, which in turn is a subset of the nonsignaling polytope $\mathcal{N}$. The Bell inequality is specified by the vector $\mathbf{h}$ defining a hyperplane which separates all vertices v_p from the observed probability distribution $\mathbf{P}$ (the black point situated outside the classical polytope $\mathcal{P}$).

$$\begin{aligned} & \mathbf{h}^T\mathbf{P} > c \\ & -1 \leqslant h_i \leqslant 1 \quad \forall \quad i \in \left\{1, \ldots, D^d_{m_a,m_b}\right\}, \end{aligned} \tag{4}$$

with the classical bound c. The additional constraint imposed on the elements of h_i of the hyperplane vector $\mathbf{h}$ keeps the maximization bounded. The chosen boundaries of h_i do not influence the result of the optimization problem besides being a global scaling factor. The hyperplane found in this manner has the form

$$\mathbf{h} = \left[h^{ab}_{A_x B_y}\right], \tag{5}$$

where $x \in X$, $y \in Y$, $a \in A$, and $b \in B$. Thus the Bell inequality found by the optimization and specified by the hyperplane vector $\mathbf{h}$ is given as

$$B[\mathbf{P}] = \sum_{a,b,x,y} h^{ab}_{A_x B_y} P\left(A_x^a B_y^b\right) \leqslant c. \tag{6}$$

Equation (6) represents the Bell inequality that is maximally violated by the observed probability distribution $\mathbf{P}$. Note that if $\mathbf{P} \in \mathcal{P}$, the optimization problem Eq. (4) is infeasible and no Bell inequality can be found.

III. STATISTICAL FLUCTUATIONS AND THEIR ESTIMATION

So far, we have concentrated on the ideal asymptotic case, that is, using the exact probabilities as entries of the observed probability distribution $\mathbf{P}$. However, in a real experiment one does not have access to probabilities but only to frequencies that are subject to statistical uncertainties and systematic errors. Since systematic errors mostly arise from specific experimental settings, we solely focus on the theoretical framework and concentrate on statistical fluctuations, as they lead to uncertainties in the observed Bell violation.

Let Alice and Bob perform N rounds of measurements. The number of instances when Alice chooses measurement $x \in X$

and Bob chooses measurement $y \in Y$ is denoted by $N_{x,y}$. In a real experiment, instead of having access to joint probabilities, we estimate them by the joint frequencies $\hat{P}(A_x^a B_y^b) = \frac{N(a,b,x,y)}{N_{x,y}}$. Here $N(a,b,x,y)$ is the number of occurrences of the corresponding input-output pair.

The Bell value $B[\hat{\mathbf{P}}]$ is a function of the joint frequencies,

$$B[\hat{\mathbf{P}}] = h^{ab}_{A_x B_y} \hat{P}(A_x^a B_y^b), \tag{7}$$

see also Eq. (6). Let $\chi(e)$ be an indicator function for a particular event e, i.e., $\chi(e) = 1$ if the event e is observed, $\chi(e) = 0$ otherwise. We introduce a random variable

$$\hat{B}_i = \sum_{a,b,x,y} h^{ab}_{A_x B_y} \frac{\chi(a_i = a, b_i = b, x_i = x, y_i = y)}{\hat{p}(x_i = x, y_i = y)},$$

where $\hat{p}(x_i = x, y_i = y) = \frac{N_{x,y}}{N}$ is the input joint frequency distribution. We get $\frac{1}{N}\sum_{i=1}^{N} \hat{B}_i = B[\hat{\mathbf{P}}]$. Defining

$$q_{\min} = \min_{a,b,x,y} \frac{h^{ab}_{A_x B_y}}{\hat{p}(x_i = x, y_i = y)},$$

$$q_{\max} = \max_{a,b,x,y} \frac{h^{ab}_{A_x B_y}}{\hat{p}(x_i = x, y_i = y)},$$

we have $q_{\min} \leqslant \hat{B}_i \leqslant q_{\max}$. We define $\gamma := q_{\max} - q_{\min}$. By using Hoeffding's inequality [37] (see Lemma 2 in Appendix A), we can bound the deviation δ of the Bell value obtained by the frequencies from the asymptotic value by a probability:

$$\Pr(B[\mathbf{P}] \geqslant B[\hat{\mathbf{P}}] - \delta) \geqslant 1 - \epsilon, \tag{8}$$

with

$$\epsilon = \exp\left(-\frac{2N\delta^2}{\gamma^2}\right). \tag{9}$$

For a given ϵ of a DIQKD protocol, one can calculate the confidence interval δ for the Bell value using Eq. (9).

IV. DIQKD MODEL AND PROTOCOL

Let us state the DIQKD protocol. We consider the i.i.d. scenario where the devices will behave independently and identically in each round. The state distributed between the parties is also the same for each round of the protocol. Alice has m measurement inputs $x \in \{1, \ldots, m\}$. Each of the inputs has d corresponding outputs $a \in \{1, \ldots, d\}$. Bob instead has $m+1$ measurement inputs $y \in \{1, \ldots, m+1\}$. Each measurement input of Bob also has d outputs $b \in \{1, \ldots, d\}$.

(1) In every round of the protocol, the parties do the following:

(a) A state ρ_{AB} is distributed between Alice and Bob.

(b) There are two types of measurement rounds, namely, raw key generation rounds and parameter estimation rounds. According to a preshared random key T, Alice and Bob choose a random $T_i = \{0, 1\}$ such that $\Pr(T_i = 1) = \xi$. If $T_i = 0$, Alice and Bob choose the measurement input $(x = 1, y = m+1)$ to generate the raw key. Otherwise, Alice and Bob choose the measurement inputs $x \in \{1, \ldots, m\}$ and $y \in \{1, \ldots, m\}$, respectively, uniformly at random. These cases will be denoted as parameter estimation rounds.

(c) The parties record their inputs and outputs as (x_i, y_i) and (a_i, b_i). After N rounds of measurement, we denote the input bit strings as X^N and Y^N and output bit strings as A^N and B^N for Alice and Bob, respectively.

(2) Alice and Bob publicly reveal their measurement outcomes of the parameter estimation rounds. They divide the parameter estimation rounds' data randomly into three sets (Alice specifies to which set each parameter estimation round's data belongs, according to a random number generator in her possession). From the first set, Alice and Bob estimate the frequencies $\hat{\mathbf{P}}_1 = [\hat{P}(A_x^a B_y^b)]$ [see Eq. (1)]. If $\hat{\mathbf{P}}_1$ is inside the classical correlation polytope $\mathcal{P}$, the protocol aborts. Otherwise, they construct an optimal Bell inequality by solving the linear optimization in Eq. (4). Then Alice and Bob use the data from the second set to calculate the Bell value $B[\hat{\mathbf{P}}_2]$. They then bound the deviation of this estimated Bell value $B[\hat{\mathbf{P}}_2]$ from the real Bell value $B[\mathbf{P}]$ by [see Eq. (8)]

$$\Pr(B[\mathbf{P}] \geqslant B[\hat{\mathbf{P}}_2] - \delta_{\text{est}}) \geqslant 1 - \epsilon_{\text{est}}, \tag{10}$$

where $\epsilon_{\text{est}} = \exp(-\frac{2N\xi\delta_{\text{est}}^2}{3\gamma^2})$ and $\frac{N\xi}{3}$ are the number of measurement rounds used to estimate the Bell value $B[\hat{\mathbf{P}}_2]$.

The parties will use the Bell inequality and corresponding violation $B[\hat{\mathbf{P}}_2] - \delta_{\text{est}}$ as a hypothesis in the experiment. From the data of the third set, the parties calculate the Bell value $B[\hat{\mathbf{P}}_3]$. For an honest implementation, the protocol aborts if the Bell value $B[\hat{\mathbf{P}}_3]$ is smaller than $B[\hat{\mathbf{P}}_2] - \delta_{\text{est}}$.

(3) Furthermore, the parties need to estimate the Quantum bit error rate (QBER) Q to bound the error correction information. Alice and Bob publicly reveal the measurement outcomes from $N\eta$ randomly sampled key generation rounds to estimate the QBER. The QBER of the raw key can be upper bounded with high probability using the tail inequality (see Lemma 1 in Appendix A):

$$\Pr\left[Q \geqslant \hat{Q} + \gamma_{\text{est}}\left(N(1-\xi-\eta), N\eta, \hat{Q}, \epsilon_{\text{est}}^{\gamma}\right)\right] > \epsilon_{\text{est}}^{\gamma}, \tag{11}$$

where $\gamma_{\text{est}}(N(1-\xi-\eta), N\eta, \hat{Q}, \epsilon_{\text{est}}^{\gamma})$ is the positive root of the following equation:

$$\ln \binom{N(1-\xi-\eta)\hat{Q} + N(1-\xi-\eta)\gamma_{\text{est}}}{N(1-\xi-\eta)} + \ln \binom{N\eta\hat{Q}}{N\eta} = \ln \binom{(N(1-\xi)\hat{Q} + N(1-\xi-\eta)\gamma_{\text{est}}}{N(1-\xi)} + \ln \epsilon_{\text{est}}^{\gamma}. \tag{12}$$

Thus we can deduce that the QBER Q is not larger than $\hat{Q} + \gamma_{\text{est}}$ (estimated QBER + statistical correction) with very high probability of $1 - \epsilon_{\text{est}}^{\gamma}$.

(4) Alice and Bob use an one-way error correction (EC) protocol to obtain identical raw keys K_A and K_B from their bit strings A^N and B^N. During the process of error correction, Alice communicates O_{EC} (O_{EC} denotes all the classical communication in the error correction step) to Bob such that he can guess the outcomes A^N of Alice. If EC aborts, they abort the protocol. In an honest implementation, this happens with probability at most ϵ_{EC}^{c}. Otherwise, they obtain error-corrected raw keys K_A and K_B [12,38–40]. The probability that Alice and Bob do not abort but hold different raw keys $K_A \neq K_B$ is at most ϵ_{EC}. For details, see Appendix B 1.

When the real QBER Q is greater than $\hat{Q} + \gamma_{\text{est}}$ (which happens with probability $\epsilon^{\gamma}_{\text{est}}$), the hashed values of keys belonging to Alice and Bob (which is sent from Alice to Bob to check if the error correction is successful, see Appendix B 1 for details) are different with high probability [38]. This results in the abortion of the implemented error correction protocol. Thus we can upper bound the error correction abortion probability ϵ^{c}_{EC} by $\epsilon^{\gamma}_{\text{est}}$.

(5) Alice and Bob apply a privacy amplification protocol to obtain a secure final key $\tilde{K}_A = \tilde{K}_B$ of length l that is close to be uniformly random and independent of the adversary's knowledge.

V. SECRET KEY RATE

To provide a lower bound on the device-independent secret key rate, one has to estimate two terms. One is the conditional von Neumann entropy $H(A|X,E)$ and the other one is the error correction information $H(A|B)$ of the raw key [41]. To estimate the latter, one can follow the footsteps of [25,42]; the detailed derivation is shown in Appendix B. For the estimation of the conditional von Neumann entropy $H(A|X,E)$, we lower bound it by the conditional min-entropy $H_{\min}(A|X,E) = -\log_2 P_g(A|X,E)$ [see Eq. (B18)] [43], where $P_g(A|X,E)$ is Eve's guessing probability about Alice's X-measurement results conditioned on her side information E. $P_g(A|X,E)$ can be upper bounded by a function G_x of the estimated Bell violation $B[\mathbf{P}]$ [26] by solving a semidefinite program [44], i.e.,

$$P_g(A|X,E) \leqslant G_x(B[\mathbf{P}]). \tag{13}$$

In real-life experiments, one does not have access to the probabilities. Instead, one has to deal with the frequencies. In Sec. IV we discussed that the protocol will abort if the observed Bell violation $B[\hat{\mathbf{P}}_3]$ in the hypothesis testing is smaller than $B[\hat{\mathbf{P}}_2] - \delta_{\text{est}}$. We need to take into account that the observed Bell violation $B[\hat{\mathbf{P}}_3]$ is calculated from a finite number of rounds. To infer the real Bell violation of the i.i.d. implementation, we make use of Hoeffding's inequality to define a confidence interval δ_{con} and the associated error probability ϵ_{con}. We bound the probability of wrongly accepting the hypothesis with the error probability ϵ_{con} by

$$\begin{aligned}&\Pr(B[\hat{\mathbf{P}}_2] - \delta_{\text{est}} \geqslant B[\hat{\mathbf{P}}_3] + \delta_{\text{con}}) < \epsilon_{\text{con}}\\ &\Rightarrow \Pr(B[\hat{\mathbf{P}}_2] - \delta_{\text{est}} - \delta_{\text{con}} \geqslant B[\hat{\mathbf{P}}_3]) < \epsilon_{\text{con}}.\end{aligned} \tag{14}$$

Therefore given that Alice and Bob do not abort the protocol, we infer that the Bell violation of the system under consideration is higher than $B[\hat{\mathbf{P}}_2] - \delta_{\text{est}} - \delta_{\text{con}}$ (with maximum ϵ_{con} probability of error). We consider the worst possible scenario and use the Bell violation $B[\hat{\mathbf{P}}_2] - \delta_{\text{est}} - \delta_{\text{con}}$ to upper bound the guessing probability $P_g(A|X,E)$ via a semidefinite program

$$\begin{aligned}&\max_{\rho,\{A(a|x)\},\{B(b|y)\}} \quad P_g(A|X,E)\\ &\text{subject to:} \quad \text{Tr}(\rho\mathcal{G}) = B[\hat{\mathbf{P}}_2] - \delta_{\text{est}} - \delta_{\text{con}}.\end{aligned} \tag{15}$$

The guessing probability $P_g(A|X,E)$ is bounded by using the NPA hierarchy [45,46] up to level 2 in the optimization problem of Eq. (15). The optimization is performed using standard tools YALMIP [47], CVX [48–50], NCPOL2SDPA [51], and QETLAB [52]. Here we have used the SDPT3 [53] solver for solving the optimization problem of Eq. (15). One can use SEDUMI [54] or MOSEK [55] as possible alternatives. $\mathcal{G}$ is the Bell operator, defined as

$$\mathcal{G} = \sum_{a,b,x,y} h^{ab}_{A_xB_y} A(a|x)B(b|y).$$

$A(a|x)$ and $B(b|y)$ are measurement operators for Alice and Bob, respectively, and ρ is the state shared between Alice and Bob. Hence the conditional von Neumann entropy $H_{\min}(A|XYE, T=1)$ can be bounded by

$$\begin{aligned}H_{\min}(A|XYE, T=1)_\rho &= -\log_2 P(A|X,E)\\ &\geqslant -\log_2 G_x(B[\hat{\mathbf{P}}_2] - \delta_{\text{est}} - \delta_{\text{con}}).\end{aligned} \tag{16}$$

The function G is defined in Eq. (13). $T=1$ specifies the outcomes of the parameter estimation rounds which are used for the estimation of the min-entropy.

To bound the error correction information, we need to estimate the QBER Q, i.e., the probability that Alice's and Bob's measurement outcomes in the key generation rounds differ. In Sec. IV we have discussed that we can upper bound the QBER Q of the raw key with at least $1 - \epsilon^{\gamma}_{\text{est}}$ probability by $\hat{Q} + \gamma_{\text{est}}$. In Appendix B we show that we can upper bound the von Neumann entropy $H(A|B)$ [20,38]:

$$H(A|B) \leqslant f(\hat{Q} + \gamma_{\text{est}}), \tag{17}$$

where $f(x) = h(x) + x\log_2(d-1)$. Here d is the number of outcomes per measurement in the Bell scenario [56] and h is the binary entropy function.

Using the bound on the min-entropy [see Eq. (16)] and the QBER [see Eq. (17)], we derive the finite-size secret key rate of a $\epsilon^{s}_{\text{DIQKD}}$-sound, $\epsilon^{c}_{\text{DIQKD}}$-complete (see Definition 6 and Appendix B for details) DIQKD protocol for collective attacks. The statement is as follows [42]: Either the protocol in Sec. IV aborts with probability higher than $1 - (\epsilon_{\text{con}} + \epsilon^{c}_{\text{EC}})$ or an $(2\epsilon_{\text{EC}} + \epsilon_s + \epsilon_{\text{PA}})$-correct-and-secret key of length

$$\begin{aligned}l \leqslant N&(-\log_2 G_x(B[\hat{\mathbf{P}}_2] - \delta_{\text{est}} - \delta_{\text{con}})\\ &- (1-\xi-\eta) f(\hat{Q} + \gamma_{\text{est}}) + (\xi+\eta)\log_2 d)\\ &-\sqrt{N}\left(4\log_2(2\sqrt{2^{\log_2 d}}+1)\left(\sqrt{\log_2\frac{8}{\epsilon'^2_{\text{EC}}}}+\sqrt{\log_2\frac{2}{\epsilon_s^2}}\right)\right)\\ &-\log_2\left(\frac{8}{\epsilon'^2_{\text{EC}}} + \frac{2}{2-\epsilon'_{\text{EC}}}\right) - \log_2\frac{1}{\epsilon_{\text{EC}}} - 2\log_2\frac{1}{2\epsilon_{\text{PA}}},\end{aligned} \tag{18}$$

can be generated where $\epsilon^{c}_{\text{DIQKD}} \leqslant \epsilon_{\text{est}} + \epsilon^{\gamma}_{\text{est}}$ (for an honest implementation) and $\epsilon^{s}_{\text{DIQKD}} \leqslant 2\epsilon_{\text{EC}} + \epsilon_s + \epsilon_{\text{PA}}$. The expression in Eq. (18) is derived in Appendix B. Table I lists all parameters of the DIQKD protocol.

VI. RESULTS

In this section we illustrate the potential and the versatility of our method with examples. We choose $\epsilon^{c}_{\text{DIQKD}} = 10^{-2}$, $\epsilon^{s}_{\text{DIQKD}} = 10^{-5}$, $\epsilon_{\text{EC}} = 10^{-10}$ as DIQKD parameters for all the examples shown in the following section.

TABLE I. Parameters of the DIQKD protocol.

N	Number of measurement rounds in the protocol
ξ	Fraction of parameter estimation rounds for estimating the Bell violation
η	Fraction of measurement rounds for estimating the QBER
ϵ_s	Smoothing parameter
ϵ_{EC}, ϵ'_{EC}	Error probabilities of the error correction protocol
ϵ^c_{EC}	Probability of abortion of error correction protocol
δ_{est}	Width of the statistical interval for the Bell violation hypothesis test
ϵ_{est}	Error probability of the Bell violation hypothesis test
δ_{con}	Confidence interval for the Bell test
ϵ_{con}	Error probability of the Bell violation estimation
γ_{est}	Width of the statistical interval for the QBER estimation
$\epsilon^{\gamma}_{\mathrm{est}}$	Error probability of the QBER estimation
ϵ_{PA}	Error probability of the privacy amplification protocol
$\epsilon^c_{\mathrm{DIQKD}}$	Completeness parameter of the DIQKD protocol
$\epsilon^s_{\mathrm{DIQKD}}$	Soundness parameter of the DIQKD protocol

A. Scenario of m measurements each, two outcomes

We present the scenario with m measurement settings for Alice and $m+1$ for Bob (where the outcomes of only m measurement settings are used in the parameter estimation). Each of those measurement settings has two possible outcomes. Let the shared state between Alice and Bob be a maximally entangled Bell state $|\psi\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$, mixed with white noise of probability p, i.e.,

$$\rho = (1-p)|\psi\rangle\langle\psi| + p\frac{\mathbb{1}}{4}, \tag{19}$$

with $p \in [0, 1]$. Both parties use σ_z as key generation measurements, resulting in the maximal possible correlation between the outcomes of Alice and Bob.

In the case of $m = 2$, consider the measurement settings of Alice and Bob that maximally violate the Clauser-Horne-Shimony-Holt (CHSH) inequality [57], i.e.,

$$\begin{aligned} x = 1 &\Rightarrow \sigma_z, \qquad x = 2 \Rightarrow \sigma_x, \\ y = 1 &\Rightarrow \frac{\sigma_z + \sigma_x}{\sqrt{2}}, \quad y = 2 \Rightarrow \frac{\sigma_z - \sigma_x}{\sqrt{2}}. \end{aligned} \tag{20}$$

For the CHSH settings with different values of white noise p, we recover the stable hyperplane stated in Table II. The secret key rate as a function of the number of measurement rounds for different values of white noise p is shown in Fig. 2. The hyperplane in Table II is equivalent to the CHSH inequality and consequently the key rate generated by our method coincides with Ref. [26] that uses a predetermined standard CHSH inequality. Though our method finds a hyperplane equivalent to the CHSH inequality, we identify the facet with the maximal violation which is then used in the DIQKD protocol. The other facets (equivalent to CHSH inequality) may admit local hidden variable models which lead to zero key.

In Refs. [30,31], the authors introduced an approach of bounding the device-independent secret key rate (DISKR) directly by using the measurement data. In the asymptotic regime, this corresponds to using a Bell inequality that leads to the maximal DISKR for the precise setup. However, small changes in the parameters (e.g., imperfections on the measurement directions) or on the measured probability distribution may lead to different Bell inequalities corresponding to the optimal secret key rate. We compare our method with Refs. [30,31] in the finite key regime. We study two different Bell scenarios. For the [2,2] scenario, we consider the CHSH settings [see Eq. (20)] and the noisy Bell state of Eq. (19) with $p = 0$ [see graph (a) of Fig. 3] and $p = 0.02$ [see graph (b) of Fig. 3]. For the [3,2] scenario (three measurement settings

TABLE II. Optimized Bell inequality for the measurement settings in Eq. (20), performed on a Bell state. Here the entries of the hyperplane vector, see Eq. (5), are given in a tabular form. For their explicit ordering see Appendix D.

1	–1	1	–1
–1	1	–1	1
1	–1	–1	1
–1	1	1	–1

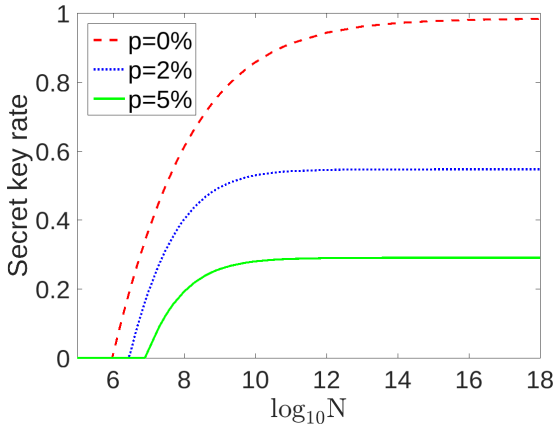

FIG. 2. Secret key rate vs logarithm of the number of rounds N using the measurement settings of Eq. (20). The state shared between two parties is the noisy Bell state [defined in Eq. (19)], where the noise is taken to be $p = 0.0$ (dashed red line), $p = 0.02$ (dotted blue line), $p = 0.05$ (green solid line).

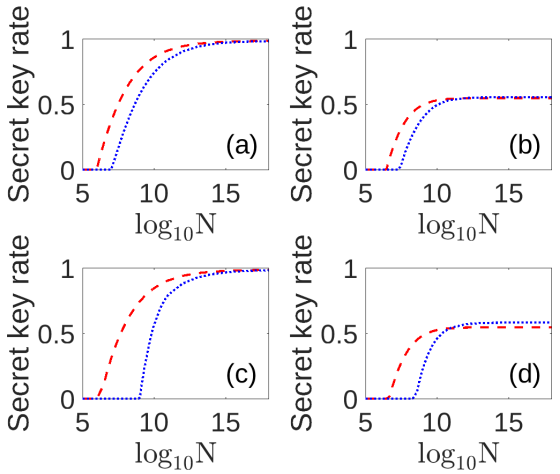

FIG. 3. Achievable secret key rate as a function of the number of measurement rounds N, comparing our method (dashed red line) and the method of Refs. [30,31] (dotted blue line) for a noisy Bell state with noise parameter p, see Eq. (19). Upper row: Measurement settings of Eq. (20) for (a) $p = 0$ and (b) $p = 0.02$. Lower row: Measurement settings of Eq. (21) for (c) $p = 0$ and (d) $p = 0.02$.

each, two outcomes per measurement), we consider the setting

$$\begin{aligned} &x = 1 \Rightarrow \sigma_z, \\ &x = 2 \Rightarrow \sin\frac{\pi}{3}\sigma_x + \cos\frac{\pi}{3}\sigma_z, \\ &x = 3 \Rightarrow \sin\frac{2\pi}{3}\sigma_x + \cos\frac{2\pi}{3}\sigma_z, \\ &y = 1 \Rightarrow \sin\frac{\pi}{6}\sigma_x + \cos\frac{\pi}{6}\sigma_z, \\ &y = 2 \Rightarrow \sigma_x, \\ &y = 3 \Rightarrow \sin\frac{5\pi}{6}\sigma_x + \cos\frac{5\pi}{6}\sigma_z, \end{aligned} \tag{21}$$

and use the noisy Bell state [Eq. (19)] with $p = 0$, $p = 0.02$ [see graphs (c) and (d) of Fig. 3]. To analyze the robustness, we incorporate fluctuations θ in the orientations in some measurement settings of Eq. (21) such that

$$\begin{aligned} &x = 1 \Rightarrow \sigma_z, \\ &x = 2 \Rightarrow \sin\left(\frac{\pi}{3} - \theta\right)\sigma_x + \cos\left(\frac{\pi}{3} - \theta\right)\sigma_z, \\ &x = 3 \Rightarrow \sin\left(\frac{2\pi}{3} + \theta\right)\sigma_x + \cos\left(\frac{2\pi}{3} + \theta\right)\sigma_z, \\ &y = 1 \Rightarrow \sin\left(\frac{\pi}{6} + \theta\right)\sigma_x + \cos\left(\frac{\pi}{6} + \theta\right)\sigma_z, \\ &y = 2 \Rightarrow \sigma_x, \\ &y = 3 \Rightarrow \sin\left(\frac{5\pi}{6} - \theta\right)\sigma_x + \cos\left(\frac{5\pi}{6} - \theta\right)\sigma_z. \end{aligned} \tag{22}$$

We use a noisy Bell state with $p = 0.02$ [see Eq. (19)] as the shared state between Alice and Bob. We use two approaches to compare the robustness of our method with Refs. [30,31]. First, we set $\theta = \frac{\pi}{60}$ [see Eq. (22)] and vary the number N of measurement rounds (see Fig. 4). Next we

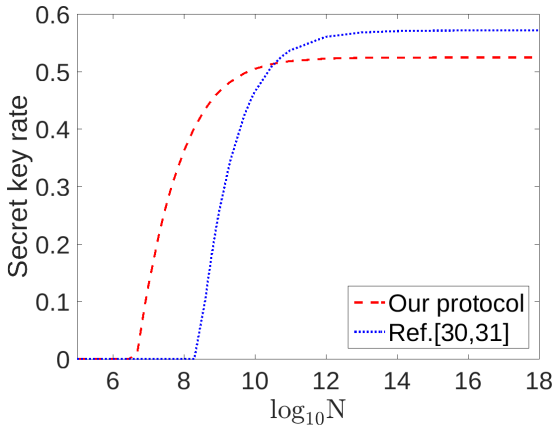

FIG. 4. Deviation of measurement direction. Secret key rate vs logarithm of the number of rounds N for our method (dashed red line) and the method of Refs. [30,31] (dotted blue line), with measurement settings of Eq. (22) where $\theta = \frac{\pi}{60}$, using a noisy Bell state with $p = 0.02$ [see Eq. (19)].

compare the methods for a range of deviations θ for $N = 10^{10}$ measurement rounds (see Fig. 5). We observe that the Bell inequality derived from our approach is stable against small fluctuations of the measurement directions or in the shared state. Our method can also generate a nonzero secret key by performing fewer measurement rounds in comparison with Refs. [30,31] (see Figs. 3–5). This is because the effect of statistical corrections in the Bell inequality violation [see Eq. (10)] is smaller in our approach. A similar behavior is also expected if the number of measurement settings per party is increased. These statistical corrections become insignificant for a high number of measurement rounds, such that the method of Refs. [30,31] yields a higher secret key in the asymptotic regime.

We point out that our method can also have advantages with respect to the CHSH scenario, when the DI secret key rate is calculated via the analytical expression from Ref. [20]: if nonoptimal measurement settings were used, we can in-

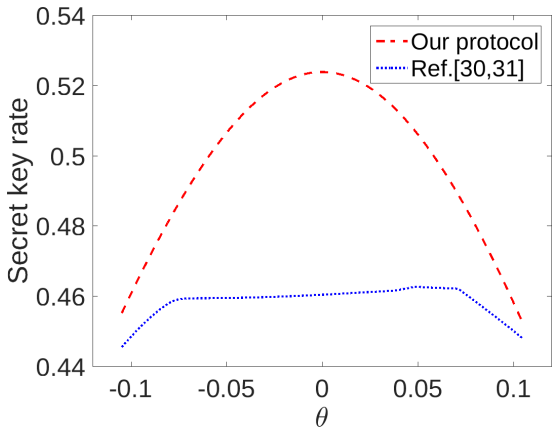

FIG. 5. Deviation of measurement direction. Secret key rate vs deviation θ of the measurement settings in Eq. (22) for our method (dashed red line) and the method of Refs. [30,31] (dotted blue line), with $N = 10^{10}$, using a noisy Bell state with $p = 0.02$ [see Eq. (19)].

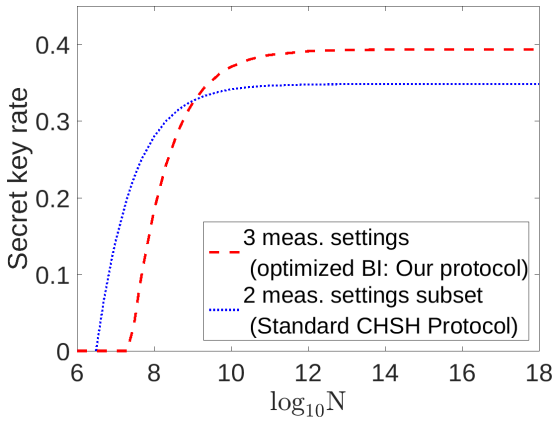

FIG. 6. Improvement for more than two measurement settings. Secret key rate vs $\log N$ for our method with measurement settings as given in Eq. (C1) (dashed red line) compared to the optimal subset of two measurement settings per party (dotted blue line). Here the secret key rate for any subset of two measurement settings per party is calculated via the analytical expression in [20] using the CHSH inequality. The shared state is a Bell state.

crease the key rate by employing additional measurement settings. As an example, we consider the observed probability distribution originating from the maximally entangled Bell state and the set of measurement settings listed explicitly in Appendix C, see Eq. (C1). With our method we can generate a higher secret key rate (for certain N) than using any subset of two measurement settings per party (and the analytical expression of [20]). See Fig. 6 for an illustration.

If the probability distribution obtained by two nonoptimal measurement settings per party does not lead to a nonzero secret key, adding another measurement setting per party and employing our strategy can be advantageous. For example, with nonoptimal measurement settings in Eq. (C2) and the maximally entangled Bell state, one cannot extract a secret key, using our method or blindly using the CHSH inequality. By adding another set of measurements for Alice and Bob, as shown in Eq. (C3), our method leads to a nonzero secret key rate.

B. Scenario of 2 measurements each, d outcomes

In this section we analyze the scenario where each party has two measurement settings in the parameter estimation rounds (Bob has an additional measurement setting which will be used in key generation rounds), and each measurement has d outcomes. The state shared between Alice and Bob is a maximally entangled state of two qudits, i.e., $|\psi\rangle = \sum_{i=0}^{d-1} \frac{1}{\sqrt{d}}|ii\rangle$, which is affected by white noise with probability p, i.e.,

$$\rho = (1-p)|\psi\rangle\langle\psi| + p\frac{\mathbb{1}}{d^2}. \tag{23}$$

We consider the measurement settings from Refs. [58,59]. The measurement is carried out in three steps. In the first step Alice applies a unitary operation on her subsystem with only nonzero terms in the diagonal equal to $e^{\iota\vec{\phi}_x(j)}$, where x denotes Alice's measurement direction, i.e., $x \in \{1, 2\}$, and $j = 0, 1, 2, \ldots, d-1$. Similarly, Bob applies a unitary operation on his subsystem with only nonzero terms in the diagonal equal to $e^{\iota\vec{\varphi}_y(j)}$, where y denotes Bob's measurement direction, i.e., $y \in \{1, 2, 3\}$. These unitary operations are denoted by $U(\vec{\phi}_x)$ and $U(\vec{\varphi}_y)$ for Alice and Bob, respectively, where

$$\vec{\phi}_x \equiv [\phi_x(0), \phi_x(1), \phi_x(2), \ldots, \phi_x(d-1)],$$
$$\vec{\varphi}_y \equiv [\varphi_y(0), \varphi_y(1), \varphi_y(2), \ldots, \varphi_y(d-1)].$$

The values of these phases are chosen as

$$\begin{aligned} &\phi_1(j) = 0, \qquad \phi_2(j) = \frac{\pi}{d}j, \\ &\varphi_1(j) = \frac{\pi}{2d}j, \quad \varphi_2(j) = -\frac{\pi}{2d}j, \quad \varphi_3(j) = 0, \end{aligned} \tag{24}$$

with $j = 0, 1, 2, \ldots, d-1$. We use $\{x = 1, y = 3\}$ for the key generation rounds and $\{x \in (1, 2), \quad y \in (1, 2)\}$ for the parameter estimation rounds. The second step consists of Alice carrying out a discrete Fourier transform U_{FT} and Bob applying U_{FT}^*. The matrix elements of the Fourier transform are defined as $(U_{FT})_{jk} = \exp[(j-1)(k-1)2\pi\iota/d]$, $(U_{FT}^*)_{jk} = \exp[-(j-1)(k-1)2\pi\iota/d]$. Thus the concatenated unitaries for Alice and Bob are $V(\vec{\phi}_x) \equiv U_{FT}\, U(\vec{\phi}_x)$ and $V(\vec{\varphi}_y) \equiv U_{FT}^*\, U(\vec{\varphi}_y)$, respectively.

Finally, Alice and Bob carry out measurements in the computational basis $|i\rangle$. For $d = 3$, we find via linear optimization, see Eq. (4), the optimized Bell inequality as shown in Table III. The details of this representation of the Bell inequality are explained in Table VII of Appendix D.

TABLE III. Optimized Bell inequality for the measurement described in the text, performed on a maximally entangled state of two qutrits. Here the entries of the hyperplane vector, see Eq. (5), are given in a tabular form. For their explicit ordering see Appendix D.

1	–1	0	–1	1	0
0	1	–1	0	–1	1
–1	0	1	1	0	–1
1	0	–1	1	–1	0
–1	1	0	0	1	–1
0	–1	1	–1	0	1

The hyperplane in Table III is equivalent to the CGLMP inequality [59,60]. If the parties share the nonmaximally entangled state

$$|\Phi\rangle \equiv \frac{|00\rangle + 0.7923|11\rangle + |22\rangle}{\sqrt{2 + 0.7923^2}}, \tag{25}$$

the CGLMP inequality is maximally violated, thus resulting in a significantly higher secret key rate, as shown in Fig. 7. This trend of generating a higher secret key rate using nonmaximally entangled states is also observed for higher dimensions (i.e., $d > 3$).

Note that in this scenario with d outcomes the maximum secret key rate is $\log_2 d$. For a fair comparison, we have normalized the min-entropy [i.e., $-\log_2$ of the solution of the optimization problem of Eq. (15)] by division with $\log_2 d$ to get a rate per qubit dimension.

Comparing the DIQKD protocol with measurement settings as described around Eq. (24) for different d and the corresponding d-dimensional maximally entangled state, see

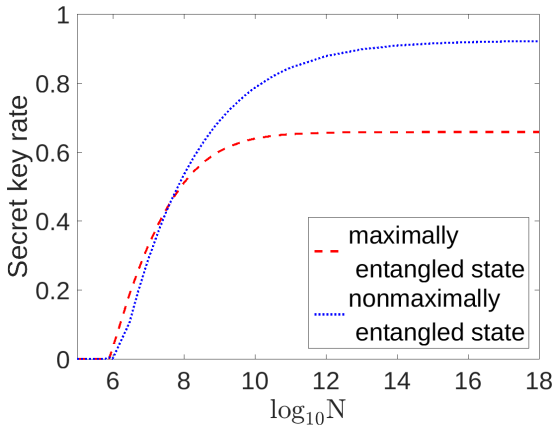

FIG. 7. Secret key rate vs $\log N$ when performing the measurement described around Eq. (24) on a maximally entangled state of two qutrits (dashed red line) and on the nonmaximally entangled state given in Eq. (25) (dotted blue line).

Eq. (23), the minimum number of measurement rounds required to have a nonzero secret key rate decreases slightly with increasing d, see Fig. 8. This follows from the fact that the minimum number of measurement rounds required to have a nonzero Bell violation decreases with increasing d. On the other side, the secret key is decreasing with increasing d (see Fig. 8) when the number of measurement rounds is sufficiently high. The nonlocality of the resultant correlation is decreasing with increasing d, which in turn results in the lower secret key.

C. Random measurement settings

In this section we analyze the case when Alice's and Bob's devices perform random measurements. We specifically focus on the fraction of events that leads to a nonzero secret key rate. First consider the $[m, 2]$ scenario, i.e., m measurement each, with two outcomes. The state shared between the parties is the noisy Bell state as in Eq. (19). We choose the raw key generation measurement operators $\{x = 1 \Rightarrow \sigma_z, y = m + 1 \Rightarrow \sigma_z\}$ in order to achieve correlated outcomes in the key measurement rounds and consequently have to exchange less error correction information. The remaining measurement operators are chosen randomly. Alice and Bob perform general unitary operators

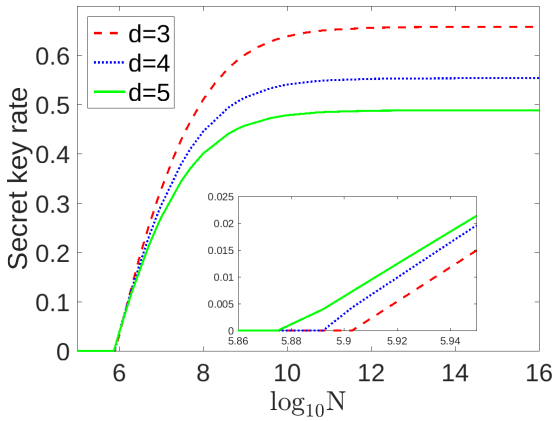

FIG. 8. Secret key rate vs $\log N$ when performing the measurement described around Eq. (24) for $d = 3$ (dashed red line), 4 (dotted blue line), and 5 (solid green line) on a maximally entangled state of two d-dimensional subsystems. The inset graph shows a zoomed-in version in the region of low number of measurement rounds, demonstrating the crossover of the curves.

TABLE IV. Approximate probability of achieving a nonzero secret key rate in the $[m, 2]$ scenario for different white noise levels p in the noisy Bell state [see Eq. (19)]. The statistics are taken over 10^5 realizations. Measurement settings of key generation rounds are fixed to be σ_z for Alice and Bob. The remaining measurement settings are performed in random orientation. For each realization, 10^{12} measurement rounds are used to compute the finite key.

	$(m_a, m_b) = 2$	$(m_a, m_b) = 3$
$p = 0\%$	∼28.6%	∼53.4%
$p = 1\%$	∼18.3%	∼46.5%
$p = 2\%$	∼10.8%	∼36.8%
$p = 3\%$	∼6.4%	∼28.2%
$p = 4\%$	∼3.9%	∼18.5%
$p = 5\%$	∼2.2%	∼11.3%

$$U(\phi, \psi, \chi) = \begin{bmatrix} e^{\iota\psi}\cos\phi & e^{\iota\chi}\sin\phi \\ -e^{-\iota\chi}\sin\phi & e^{-\iota\psi}\cos\phi \end{bmatrix} \tag{26}$$

with parameters $\psi, \chi \in [0, 2\pi]$ and $\phi \in [0, \frac{\pi}{2}]$ and then measure in the computational basis $\{|0\rangle, |1\rangle\}$. This strategy is equivalent to choosing a random measurement. In Table IV we show the fraction of events that leads to a nonzero secret key rate with random measurements. The statistics are based on 10^5 realizations. For the [2,2] scenario, the optimization in Eq. (4) will always lead to the CHSH inequality. Adding another measurement setting per party (i.e., the [3,2] scenario) significantly increases the probability of finding a hyperplane that produces a nonzero secret key rate. The first explanation of this fact is statistical. By increasing the number of settings, we increase the probability that some of them violate a Bell inequality even involving only two settings per party. Apart from that, the optimization in Eq. (4) also provides some hyperplanes for the [3,2] scenario that are independent of the hyperplanes for the [2,2] scenario. From the higher chance of Bell inequality violation, we obtain a higher chance of achieving a nonzero key. This result also reverberates the

TABLE V. Approximate probability of achieving a nonzero secret key rate in the $[2, d]$ scenario for white noise of different probability p added to the maximally entangled state of two qudits [see Eq. (23)]. All other details are as in Table IV.

	$d = 3$	$d = 4$
$p = 0\%$	∼6.4%	∼2.5%
$p = 1\%$	∼2.2%	∼0%
$p = 2\%$	∼0.3%	∼0%

results of the nonlocal volume[2] in [61–66], which increases for the pure bipartite entangled state when more measurement settings for each party are used. We observe the same phenomenon in our case, regarding the secret key rate. As the nonlocal volume shrinks by adding noise, it also reduces the probability of producing a nonzero key rate. Let us now analyze the $[2, d]$ scenario (i.e., d outcomes per measurement) with random measurement settings. The shared state is a noisy maximally entangled state of two qudits [see Eq. (23)]. We compute the approximate probability for achieving a nonzero secret key rate (see Table V). The statistics are based on 10^5 realizations. The measurements for key generation are in the computational basis. The remaining measurement settings are chosen randomly.

We observe that for $d \geqslant 3$, the probability to extract a nonzero secret key is smaller compared to the case with only two outcomes. This follows from the fact that the nonlocal volume shrinks by increasing the dimension of the maximally entangled state. This results in a smaller probability of generating nonlocal correlations and therefore a smaller chance of a Bell inequality violation [65] and smaller probability of a nonzero secret key.

VII. CONCLUSIONS

Several protocols for device-independent quantum key distribution (DIQKD) have the common feature that they rely on the violation of a predetermined Bell inequality. We propose a robust DIQKD procedure where a suitable Bell inequality is instead constructed from the measurement data. This constructed Bell inequality leads to the maximum Bell violation for the particular setup. Then we use the Bell inequality and its corresponding violation to bound the secret key rate via lower bounding the min-entropy.

We provide a finite-size key analysis of our proposed procedure. We bound the statistical fluctuations of the Bell inequality violation by Hoeffding's inequality. However, we do not claim that our choice of concentration inequality [67–69] is optimal for a finite number of measurement rounds. Note that our method could also be implemented for the estimation of global randomness in a device-independent randomness generation protocol.

We have illustrated our method with several examples for different numbers of measurement settings and different numbers of outcomes. Even though our procedure may identify a specific Bell inequality of a known type in some cases, a predefined version of this type of Bell inequalities would often lead to zero key. Our procedure identifies the one Bell inequality (out of all the equivalent ones) with maximal violation, which then leads to a nonzero secret key rate.

We have also shown cases when our method yields a higher secret key rate than using the standard CHSH inequality. In comparison to related approaches (Refs. [30,31]), we provide examples where our approach needs fewer numbers of measurement rounds to generate a nonzero secret key. Using our method, the typical number of measurement rounds to generate a nonzero key varies between 10^6 to 10^8 for the $[m, 2]$ Bell scenario and is of the order 10^6 for the $[2, d]$ Bell scenario. We further showed the performance of our method in the case of random measurement settings. Our method employs the observed measurement statistics, which can be affected by inefficient detectors. In case of no-detection events, one can follow our procedure by declaring no-detection as an additional outcome. One could also account for detector efficiencies by using the approaches of Refs. [70–72].

Finally, future work should address the use of more sophisticated methods of bounding the conditional von Neumann entropy [73–75], which could increase the secret key rate, in comparison to the bounds based on the min-entropy.

ACKNOWLEDGMENTS

The authors acknowledge support from the Federal Ministry of Education and Research (BMBF, Projects Q.Link.X and HQS). We also acknowledge support by the QuantERA project QuICHE via the German Ministry for Education and Research (BMBF Grant No. 16KIS1119K). We thank Gláucia Murta, Federico Grasselli, and Lucas Tendick for helpful discussions.

APPENDIX A: DEFINITIONS

We start with the definition of some quantities that will help us to derive the key rates for the DIQKD protocol.

Definition 1 (Min and max-entropy [76,77]). Let $\rho_{AB} \in \mathcal{P}(\mathcal{H}_A \otimes \mathcal{H}_B)$ and $\sigma_B \in \mathcal{P}(\mathcal{H}_B)$. $\mathcal{P}(\mathcal{H}_B)$ is the set of positive-semidefinite operators on the Hilbert space $\mathcal{H}_B$. The min-entropy of ρ_{AB} conditioned on σ_B is

$$H_{\min}(\rho_{AB}|\sigma_B) := -\log_2 \lambda, \tag{A1}$$

where λ is the minimum real number such that $\lambda.(\mathbb{I} \otimes \sigma_B) - \rho_{AB} \geqslant 0$. The max-entropy of ρ_{AB} conditioned on σ_B is

$$H_{\max}(\rho_{AB}|\sigma_B) := \log_2 \mathrm{Tr}\big((\mathbb{I} \otimes \sigma_B)\rho_{AB}^0\big), \tag{A2}$$

where ρ_{AB}^0 denotes the projector onto the support of ρ_{AB}.

Definition 2 (Smoothed min and max-entropy [76,78]). For a quantum state ρ_{AB} and $\epsilon \geqslant 0$, the smooth min-entropy of system A conditioned on B is defined as

$$H_{\min}^{\epsilon}(A|B) := \max_{\tilde{\rho}_{AB} \in \mathcal{B}^{\epsilon}(\rho_{AB})} H_{\min}(A|B)_{\tilde{\rho}_{AB}}, \tag{A3}$$

and the smooth max-entropy of system A conditioned on B is defined as

$$H_{\max}^{\epsilon}(A|B) := \min_{\tilde{\rho}_{AB} \in \mathcal{B}^{\epsilon}(\rho_{AB})} H_{\max}(A|B)_{\tilde{\rho}_{AB}}. \tag{A4}$$

$\mathcal{B}^{\epsilon}$ is an ϵ-ball of subnormalized operators around the state ρ_{AB} defined in terms of the purified distance.

[2]The nonlocal volume is a statistical measure of nonlocality introduced in [61]. It is defined as the probability that the correlations, generated from randomly chosen projective measurements made on a given state $|\psi\rangle$, violate any Bell inequality (a witness of nonlocality) by any extent. Generally, the nonlocal volume for a given state $|\psi\rangle$ is obtained by $\int d\Omega f(|\psi\rangle, \Omega)$, where one integrates over the measurement parameters Ω [62]. $f(|\psi\rangle, \Omega)$ is an indicator function that takes the value 1 if the resultant correlations, generated from the state and measurements, are nonlocal. Otherwise, it will take the value 0.

Now we focus on the security parameters of quantum key distribution. The security of quantum key distribution can be split into two conditions.

Definition 3 (Correctness [5,25,42]). A DIQKD protocol is ϵ_{corr}-correct if the final key $\tilde{K}_A$ of Alice differs from the final key $\tilde{K}_B$ of Bob with probability less than ϵ_{corr}, i.e.,

$$\Pr(\tilde{K}_A \neq \tilde{K}_B) \leqslant \epsilon_{\text{corr}}. \tag{A5}$$

Definition 4 (Secrecy [5,25,42]). For any $\epsilon_{\text{sec}} \geqslant 0$, a DIQKD protocol is ϵ_{sec} with respect to the adversary E if the joint state satisfies

$$p(\Omega)\tfrac{1}{2}\|\rho_{\tilde{K}_A E|\Omega} - \tau_{\tilde{K}_A} \otimes \rho_E\|_1 \leqslant \epsilon_{\text{sec}}, \tag{A6}$$

where $\tau_{\tilde{K}_A}$ is the maximally mixed state on $\tilde{K}_A$ of the protocol. Here $p(\Omega)$ is the probability of not aborting the protocol.

If a protocol is ϵ_{corr}-correct and ϵ_{sec}-secret, then it is $\epsilon^s_{\text{DIQKD}}$-correct and secret for any $\epsilon^s_{\text{DIQKD}} \geqslant \epsilon_{\text{corr}} + \epsilon_{\text{sec}}$. The correctness (see Def. 3) of the final key is ensured by the error correction step. During error correction, Alice sends a sufficient amount of information to Bob so that he can correct his raw key. If Alice and Bob do not abort in this step, then the probability that they end up with different raw keys is guaranteed to be very small (below ϵ_{EC}). For the secrecy of the protocol (see Def. 4) one needs to estimate how far the final state describing Alice's key and the eavesdropper's system is from the ideal one.

Definition 5 (Secret key rate [25,42]). If a protocol generates a correct and secret key of length l after n rounds, the secret key rate is defined as

$$r = \frac{l}{n}. \tag{A7}$$

Any useful DIQKD protocol should not abort almost all the time. This is apprehended by the concept of completeness.

Definition 6 (Security [25,42]). A DIQKD protocol is $(\epsilon^s_{\text{DIQKD}}, \epsilon^c_{\text{DIQKD}}, l)$-secure if

(1) (Soundness) For any implementation of the protocol, either it aborts with probability greater than $1 - \epsilon^s_{\text{DIQKD}}$ or an $\epsilon^s_{\text{DIQKD}}$-correct and secret key of length l is obtained.

(2) (Completeness) There exists an honest implementation of the protocol such that the probability of not aborting, $p(\Omega)$, is greater than $1 - \epsilon^c_{\text{DIQKD}}$.

In the privacy amplification step, Alice and Bob want to turn their equal string of bits, which may be partially known to an eavesdropper, into a shorter completely secure string of bits. For this step, a 2-universal family of hash functions is needed.

Definition 7 (2-universal hash function). A family of hash functions $\mathcal{F} = \{f : \{0,1\}^n \to \{0,1\}^\ell\}$ is called 2-universal if for every two strings $x, x' \in \{0,1\}^n$ with $x \neq x'$ then

$$\Pr_{f\in\mathcal{F}}(f(x) = f(x')) = \frac{1}{2^\ell}, \tag{A8}$$

where f is chosen uniformly at random in $\mathcal{F}$. The property of 2-universality ensures a good distribution of the outputs. For $\ell \leqslant n$ there always exists a 2-universal family of hash functions [79].

Now we will state the quantum leftover hashing lemma [77,80]. It quantifies the secrecy of a protocol as a function of a conditional entropy of the state before privacy amplification and the length of the final key.

Theorem 1 (Leftover hashing lemma with smooth min-entropy [25,42,80]). Let $\rho_{A^n E}$ be a classical quantum state. Let $\mathcal{F}$ be a 2-universal family of hash functions, from $\{0,1\}^n$ to $\{0,1\}^l$, that maps the classical n-bit string A^n into K_A. Then

$$\|\rho_{K_A F E} - \tau_{K_A} \otimes \rho_{FE}\| \leqslant 2^{-\frac{1}{2}(H^\epsilon_{\min}(A^n|E)_\rho - l)} + 2\epsilon,$$

where F is a classical register that stores the hash function f.

With the leftover hash lemma and the definition of secrecy (see Def. 4), we express the length of a secure key as a function of the entropy of Alice's raw key conditioned on Eve's information before privacy amplification.

Theorem 2 (Key length [25,42]). Let $P(\Omega)$ be the probability that the DIQKD protocol does not abort for a particular implementation. If the length of the key generated after privacy amplification is given by

$$l \leqslant H^{\epsilon_s/P(\Omega)}_{\min}(A^n|E)_{\rho_{|\Omega}} - 2\log\frac{1}{2\epsilon_{\text{PA}}},$$

then the DIQKD protocol is $\epsilon_{\text{PA}} + \epsilon_s$ secret.

In this paper we have considered the II D scenario (collective attacks). In the assumption of collective attacks, the distributed state and the behavior of Alice's and Bob's devices are the same in every round of the protocol. Eve can carry out arbitrary operations in her quantum side information. This assumption implies that after n rounds of the protocol, the state shared by Alice, Bob, and Eve is $\rho_{A^n B^n E} = \rho^{\otimes n}_{ABE}$. The quantum asymptotic equipartition property [76,81] allows us to bound the conditional smooth min-entropy of state $\rho^{\otimes}_{AE}$ by the conditional von Neumann entropy of the state ρ_{AE}.

Theorem 3 (Asymptotic equipartition property [81]). Let $\rho = \rho^{\otimes n}_{AE}$ be an IID state. Then for $n \geqslant \frac{8}{5}\log\frac{2}{\epsilon^2}$,

$$H^\epsilon_{\min}(A^n|E^n)_{\rho^{\otimes n}_{AE}} > nH(A|E)_{\rho_{AE}} - \sqrt{n}\delta(\epsilon,\chi),$$

and similarly,

$$H^\epsilon_{\max}(A^n|E^n)_{\rho^{\otimes n}_{AE}} < nH(A|E)_{\rho_{AE}} + \sqrt{n}\delta(\epsilon,\chi),$$

where $\delta(\epsilon,\chi) = 4\log(\chi)\sqrt{\log\frac{2}{\epsilon^2}}$ and $\chi = \sqrt{2^{-H_{\min}(A|E)_{\rho_{AE}}}} + \sqrt{2^{H_{\max}(A|E)_{\rho_{AE}}}} + 1$.

Lemma 1 [82,83]. Let $\mathcal{X}_{n+k}$ be a random binary string of $n+k$ bits, $\mathcal{X}_k$ be a random sample (without replacement) of m entries from the string $\mathcal{X}_{n+k}$, and $\mathcal{X}_n$ be the remaining bit string. Λ_k and Λ_n are the frequencies of bit value 1 in string $\mathcal{X}_k$ and $\mathcal{X}_n$, respectively. For any $\varepsilon_1 > 0$, it holds the upper tail inequality:

$$\Pr[\Lambda_n \geqslant \Lambda_k + \gamma_1(n,k,\Lambda_k,\varepsilon_1)] > \varepsilon_1, \tag{A9}$$

where $\gamma_1(a,b,c,d)$ is the positive root of

$$\ln\binom{bc}{b} + \ln\binom{ac + a\gamma_1(a,b,c,d)}{a} = \ln\binom{(a+b)c + a\gamma_1(a,b,c,d)}{a+b} + \ln d.$$

For $\varepsilon_2 > 0$, we have the lower tail inequality:

$$\Pr[\Lambda_n \leqslant \Lambda_k - \gamma_2(n,k,\Lambda_k,\varepsilon_2)] > \varepsilon_2, \tag{A10}$$

where $\gamma_2(a,b,c,d)$ is the positive root of

$$\ln\binom{bc}{b} + \ln\binom{ac - a\gamma_2(a,b,c,d)}{a} = \ln\binom{(a+b)c - a\gamma_2(a,b,c,d)}{a+b} + \ln d.$$

Lemma 2 [37]. Let $X_1, X_2, \ldots, X_n$ be independent random variables strictly bounded by the intervals $[a_i, b_i]$, i.e., $a_i \leqslant X_i \leqslant b_i$. We define

$$\bar{X} = \frac{1}{n}(X_1 + X_2 + \cdots + X_n).$$

Then Hoeffding's inequality reads

$$\Pr(\bar{X} - E[\bar{X}] \geqslant t) \leqslant \exp\left(-\frac{2n^2t^2}{\sum_{i=1}^n (b_i - a_i)^2}\right).$$

Let $c_i := b_i - a_i$ and $c_i \leqslant C \; \forall \; i$. Then Hoeffding's inequality reads

$$\Pr(\bar{X} - E[\bar{X}] \geqslant t) \leqslant \exp\left(-\frac{2n^2t^2}{nC^2}\right) = \exp\left(-\frac{2nt^2}{C^2}\right).$$

APPENDIX B: SECRET KEY ANALYSIS

Theorem 4 (Completeness). The DIQKD protocol stated in Sec. IV is $\epsilon_{\text{est}} + \epsilon_{\text{est}}^{\gamma}$ complete.

Proof. The protocol can abort in two instances. Either it will abort if the error correction failed or if the estimated Bell violation $B[\hat{\mathbf{P}}_3]$ is not high enough. The probability that the error correction fails can only happen if the real QBER Q is larger than $\hat{Q} + \gamma_{\text{est}}$, which happens with probability $\epsilon_{\text{est}}^{\gamma}$, see Sec. IV for details. The protocol also aborts if the estimated Bell violation $B[\hat{\mathbf{P}}_3]$ is smaller $B[\hat{\mathbf{P}}_2] - \delta_{\text{est}})$, see Sec. IV for details. Thus, considering an honest implementation consisting of IID rounds, we can bound the probability of abortion of the protocol:

$$\begin{aligned}
\text{p(abort)} &= \text{p((EC aborts) or (Bell test fails))} \\
&\leqslant \text{p(EC aborts)} + \text{p(Bell test fails)} \\
&\leqslant \text{p(QBER test fails)} + \text{p(Bell test fails)} \\
&= p(Q > \hat{Q} + \gamma_{\text{est}}) + p(B[\hat{\mathbf{P}}_3] < B[\hat{\mathbf{P}}_2] - \delta_{\text{est}}) \\
&= \epsilon_{\text{est}}^{\gamma} + \epsilon_{\text{est}}, \qquad \text{(B1)}
\end{aligned}$$

where ϵ_{est} is defined in Eq. (10), and $\epsilon_{\text{est}}^{\gamma}$ is defined in Eq. (11). Thus we get $\epsilon_{\text{DIQKD}}^{c} \leqslant \epsilon_{\text{est}} + \epsilon_{\text{est}}^{\gamma}$.

For the **soundness**, we have to evaluate the correctness and secrecy, defined in Def. 3 and Def. 4, respectively. In case of correctness, if we have an error correction protocol that does not abort, then Alice (Bob) will have the raw key K_A (K_B) after the protocol. The string K_B differs from K_A with probability less than ϵ_{EC}, and as the final keys $\tilde{K}_A$ and $\tilde{K}_B$ are equal if the raw keys are equal, it follows that

$$P(\tilde{K}_A \neq \tilde{K}_B) \leqslant P(K_A \neq K_B) \leqslant \epsilon_{\text{EC}}.$$

For secrecy, let us recall that Ω is defined as the event when the protocol does not abort. This happens when the error correction protocol does not abort and achieved the required Bell violation according to Alice's and Bob's outputs (and inputs). Now define the event $\hat{\Omega}$ as the event Ω (protocol not aborting) and the error correction being successful, i.e., $K_A = K_B$. Thus

$$\begin{aligned}
\|\rho_{\tilde{K}_A E_{|\Omega}} - \tau_{\tilde{K}_A} \otimes \rho_E\|_1 &\leqslant \|\rho_{\tilde{K}_A E_{|\Omega}} - \rho_{\tilde{K}_A E_{|\hat{\Omega}}}\|_1 \\
&+ \|\rho_{\tilde{K}_A E_{|\hat{\Omega}}} - \tau_{\tilde{K}_A} \otimes \rho_E\|_1 \\
&\leqslant \epsilon_{\text{EC}} + \|\rho_{\tilde{K}_A E_{|\hat{\Omega}}} - \tau_{\tilde{K}_A} \otimes \rho_E\|_1. \qquad \text{(B2)}
\end{aligned}$$

The first inequality follows from the triangular inequality of the trace distance [84]. $\rho_{\tilde{K}_A E_{|\Omega}}$ is the joint classical quantum state of Alice and Eve if the protocol does not abort. $\rho_{\tilde{K}_A E_{|\hat{\Omega}}}$ is the joint classical quantum state of Alice and Eve if the protocol does not abort and the error correction is successful. When error correction succeeds, the probability of $K_A = K_B$ is higher than $(1 - \epsilon_{\text{EC}})$. Conversely, the probability $K_A \neq K_B$ is less than ϵ_{EC}. Thus the second inequality of Eq. (B2) comes from

$$\begin{aligned}
\|\rho_{\tilde{K}_A E_{|\Omega}} - \rho_{\tilde{K}_A E_{|\hat{\Omega}}}\|_1 &\leqslant (1 - \epsilon_{\text{EC}})\|\rho_{\tilde{K}_A E_{|\hat{\Omega}}} - \rho_{\tilde{K}_A E_{|\hat{\Omega}}}\|_1 \\
&+ \epsilon_{\text{EC}}\|\rho_{\tilde{K}_A E_{|\hat{\Omega}}} - \rho_{\tilde{K}_A E_{|\hat{\Omega}^c}}\|_1 \leqslant \epsilon_{\text{EC}}, \qquad \text{(B3)}
\end{aligned}$$

where $\hat{\Omega}^c$ is defined as the event when the protocol does not abort but error correction is not successful, i.e., $K_A \neq K_B$.

Now we proceed to evaluate the term $\|\rho_{\tilde{K}_A E_{|\hat{\Omega}}} - \tau_{\tilde{K}_A} \otimes \rho_E\|_1$ of Eq. (B2). We will follow the path shown in [25,42]. Given that the protocol did not abort, the maximal length of a secure key is determined by the smooth min-entropy of Alice's raw key conditioned on all information available to the eavesdropper (see the leftover hashing lemma in Theorem. 1). In our protocol (see Sec. IV), it is given by $H_{\min}^{\epsilon_s}(A^N|X^N Y^N T^N E O_{\text{EC}})_{\rho_{|\hat{\Omega}}}$. Here we recall that O_{EC} is the information exchanged by Alice and Bob during the error correction protocol. X^N and Y^N are the input bit strings (measurement settings) for Alice and Bob, respectively. A^N is the output bit string of Alice. T^N is the shared random key that determines whether the round is a test or a key generation round. $\hat{\Omega}$ is the event that the protocol does not abort and error correction succeeds.

In order to bypass the conditioned state of $H_{\min}^{\epsilon_s}(A^N|X^N Y^N T^N E O_{\text{EC}})_{\rho_{|\hat{\Omega}}}$, we can start from the definition of secrecy (see Def. 4). Then we have to bound the term

$$p(\Omega)\|\rho_{\tilde{K}_A F E_{|\Omega}} - \tau_{\tilde{K}_A} \otimes \rho_{FE}\|_1 = \|\rho_{\tilde{K}_A F E \wedge \Omega} - \tau_{\tilde{K}_A} \otimes \rho_{FE \wedge \Omega}\|_1, \qquad \text{(B4)}$$

where $\rho_{\tilde{K}_A F E \wedge \Omega} = p(\Omega)\rho_{\tilde{K}_A F E_{|\Omega}}$ is a subnormalized state. Here we recall that F is the classical register that stores the hash function f (see Def. 7).

Now using the leftover hashing lemma in Theorem 1, we can generate an $(\epsilon_s + \epsilon_{\text{PA}})$-secret key of length [42]

$$l \leqslant H_{\min}^{\epsilon_s}(A^N|E)_{\rho \wedge \Omega} - 2\log\frac{1}{2\epsilon_{\text{PA}}}. \qquad \text{(B5)}$$

In Ref. [85] it is proved that

$$H_{\min}^{\epsilon_s}(A^N|E)_{\rho\wedge\Omega} \geqslant H_{\min}^{\epsilon_s}(A^N|E)_\rho. \tag{B6}$$

Thus we proceed to estimate the quantity $H_{\min}^{\epsilon_s}(A^N|X^NY^NT^NEO_{\mathrm{EC}})_\rho$ in order to bound the achievable secret key of length l.

Using the chain rule relation for the smooth min-entropy conditioned on classical information [76], we can write

$$H_{\min}^{\epsilon_s}(A^N|X^NY^NT^NEO_{\mathrm{EC}})_\rho = H_{\min}^{\epsilon_s}(A^N|X^NY^NT^NE)_\rho - \mathrm{leak}_{\mathrm{EC}}. \tag{B7}$$

Thus, in order to bound $H_{\min}^{\epsilon_s}(A^N|X^NY^NT^NEO_{\mathrm{EC}})_\rho$, we have to lower bound $H_{\min}^{\epsilon_s}(A^N|X^NY^NT^NE)_\rho$ and upper bound $\mathrm{leak}_{\mathrm{EC}}$ (the leakage due to the error correction).

1. Estimation of $\mathrm{leak}_{\mathrm{EC}}$

Alice and Bob perform an EC procedure so that Bob can compute a guess of Alice's raw key A^N. In order to verify if EC is successful, Alice chooses a two-universal hash function (uniformly at random) from the family of hash functions and computes a hash of length $\log(\frac{1}{\epsilon_{\mathrm{EC}}})$ from her raw keys A^N. Then she sends the chosen hash function and the hashed value of her bits to Bob via a public channel. We denote all the classical communication (information leaked during EC, hash function, and the hashed value for verification) by O_{EC}. Bob computes the hash function on his key. If the hashed values are equal, then Alice's and Bob's keys are the same with high probability. If the hashed values are different, the parties will abort the protocol. During this whole process the amount of information about the key exposing to the adversary Eve is termed as $\mathrm{leak}_{\mathrm{EC}}$. In Ref. [77] the $\mathrm{leak}_{\mathrm{EC}}$ is bounded by

$$\mathrm{leak}_{\mathrm{EC}} \leqslant H_0^{\epsilon'_{\mathrm{EC}}}(A^N|B^NX^NY^NT^N) + \log\frac{1}{\epsilon_{\mathrm{EC}}}, \tag{B8}$$

where $\epsilon^c_{\mathrm{EC}} = \epsilon_{\mathrm{EC}} + \epsilon'_{\mathrm{EC}}$ (see Table I). H_0 is the Rényi entropy introduced in Ref. [77]. In Ref. [76], it is denoted as $\bar{H}_0^{\uparrow}$. If Alice and Bob do not abort, then their resultant bit string is identical ($K_A = K_B$) with at least $1-\epsilon_{\mathrm{EC}}$ probability. We can bound the entropy $H_0^{\epsilon'_{\mathrm{EC}}}(A^N|B^NX^NY^NT^N)$ in the following way:

$$\begin{aligned}
&H_0^{\epsilon'_{\mathrm{EC}}}(A^N|B^NX^NY^NT^N)\\
&\leqslant H_{\max}^{\frac{\epsilon'_{\mathrm{EC}}}{2}}(A^N|B^NX^NY^NT^N) + \log\left(\frac{8}{\epsilon'^2_{\mathrm{EC}}} + \frac{2}{2-\epsilon'_{\mathrm{EC}}}\right)\\
&\leqslant NH(A|BXYT) + 4\sqrt{N}\log(2\sqrt{2^{\log_2 d}}+1)\sqrt{\log\frac{8}{{\epsilon'_{\mathrm{EC}}}^2}}\\
&\quad + \log\left(\frac{8}{\epsilon'^2_{\mathrm{EC}}} + \frac{2}{2-\epsilon'_{\mathrm{EC}}}\right).
\end{aligned} \tag{B9}$$

For the definition of $H_0^\epsilon(A|B)$, see Ref. [77]. The first inequality of Eq. (B9) comes from Ref. [80] and Eq. (B11) of Ref. [42]. The last inequality comes from the asymptotic equipartition property (see Theorem 3), where we used the relations

$$\begin{aligned}
\delta(\epsilon,\chi) &= 4\log(\chi)\sqrt{\log\frac{2}{\epsilon^2}}\\
&\leqslant 4\log(2\sqrt{2^{\log_2 d}}+1)\sqrt{\log\left(\frac{2}{\epsilon^2}\right)}.
\end{aligned} \tag{B10}$$

Here we have used $\chi \leqslant 2\sqrt{2^{\log_2 d}}+1$, which comes from

$$\begin{aligned}
\chi &= \sqrt{2^{-H_{\min}(A|E)_{\rho_{AE}}}} + \sqrt{2^{H_{\max}(A|E)_{\rho_{AE}}}} + 1\\
&\leqslant 2\sqrt{2^{H_{\max}(A|XYTE)_\rho}} + 1\\
&\leqslant 2\sqrt{2^{\log_2 d}} + 1.
\end{aligned} \tag{B11}$$

The first inequality of Eq. (B11) follows from the fact that A is a classical register and therefore has positive conditional min-entropy, which implies $-H_{\min}(A|XYTE) \leqslant H_{\min}(A|XYTE) \leqslant H_{\max}(A|XYTE)$. For the second inequality of Eq. (B11), we use $H_{\max}(A|XYTE) \leqslant \log_2 d$.

Therefore, from Eqs. (B8) and (B9), we can bound the leakage in the following way:

$$\begin{aligned}
\mathrm{leak}_{\mathrm{EC}} &\leqslant NH(A|BXYT)\\
&\quad + \sqrt{n}(4\log(2\sqrt{2^{\log_2 d}}+1))\sqrt{\log\frac{8}{\epsilon'^2_{\mathrm{EC}}}}\\
&\quad + \log\left(\frac{8}{\epsilon'^2_{\mathrm{EC}}} + \frac{2}{2-\epsilon'_{\mathrm{EC}}}\right) + \log\frac{1}{\epsilon_{\mathrm{EC}}}.
\end{aligned} \tag{B12}$$

Now we bound the single-round von Neumann entropy $H(A|BXYT)$ as

$$\begin{aligned}
H(A|BXYT) &= p(T=0)H(A|BXYT=0)\\
&\quad + p(T=1)H(A|BXYT=1)\\
&\leqslant (1-\xi)H(A|BXYT=0) + \xi\log_2 d\\
&\leqslant (1-\xi-\eta)H(A|BXYT=0) + (\xi+\eta)\log_2 d.
\end{aligned} \tag{B13}$$

See Table I for the details of ξ, η, and γ_{est}. For the first equality, we have used that for the conditional von Neumann entropy it holds that $H(A|BX)_\rho = \sum_x p(X=x)H(A|BX=x)$. We divide the measurement rounds into key generation (specified by $T=0$) and parameter estimation (specified by $T=1$), for details see Sec. IV. The first inequality comes from the fact that parameter estimation round's measurements were publicly communicated to estimate the Bell inequality and the corresponding violation. η rounds of the raw key generation measurement were communicated through a public channel to estimate the QBER, which leads to the last inequality.

Now our goal is to estimate $H(A|BXYT=0)$. For dichotomic observables and uniform marginals, $H(A|B)$ can be expressed as $h(Q)$ [20], where h is the binary entropy function, $h(p) := -p\log_2 p - (1-p)\log_2(1-p)$. Similarly for the $[(m_a,m_b),d]$ Bell scenario, $H(A|B)$ can be expressed as a function of the QBER, $H(A|B) = -Q\log_2 Q - (1-Q)\log_2(1-Q) + Q\log_2(d-1)$ [56].

For our specific protocol (see Sec. IV), we bound $H(A|BXYT=0)$ by a function of $\hat{Q}_1 + \gamma_{\mathrm{est}}$ (observed QBER

TABLE VI. Bell inequality table for the [2,2] scenario.

$h^{11}_{A_1B_1}$	$h^{12}_{A_1B_1}$	$h^{11}_{A_1B_2}$	$h^{12}_{A_1B_2}$
$h^{21}_{A_1B_1}$	$h^{22}_{A_1B_1}$	$h^{21}_{A_1B_2}$	$h^{22}_{A_1B_2}$
$h^{11}_{A_2B_1}$	$h^{12}_{A_2B_1}$	$h^{11}_{A_2B_2}$	$h^{12}_{A_2B_2}$
$h^{21}_{A_2B_1}$	$h^{22}_{A_2B_1}$	$h^{21}_{A_2B_2}$	$h^{22}_{A_2B_2}$

+ estimated statistical error), see Sec. V for details:

$$H(A|BXY, T=0) \leqslant f(\hat{Q}+\gamma_{\text{est}}), \tag{B14}$$

where $f(x) = h(x) + x\log_2(d-1)$ (d is the number of outcomes per measurement in the Bell scenario) and h is the binary entropy function. From Eqs. (B13) and (B14), it then follows that

$$H(A|BXYT) \leqslant (1-\xi-\eta)f(\hat{Q}+\gamma_{\text{est}}) + (\xi+\eta)\log_2 d. \tag{B15}$$

The leakage due to error correction is given by [from Eqs. (B12) and (B15)]

$$\begin{aligned}\text{leak}_{\text{EC}} \leqslant\ & N[(1-\xi-\eta)f(\hat{Q}+\gamma_{\text{est}}) + (\xi+\eta)\log_2 d] \\ & + \sqrt{N}\left(4\log(2\sqrt{2^{\log_2 d}}+1)\right)\sqrt{\log\frac{8}{\epsilon'^2_{\text{EC}}}} \\ & + \log\left(\frac{8}{\epsilon'^2_{\text{EC}}} + \frac{2}{2-\epsilon'_{\text{EC}}}\right) + \log\frac{1}{\epsilon_{\text{EC}}}.\end{aligned} \tag{B16}$$

2. Estimation of min-entropy $H^{\epsilon_s}_{\min}(A^N|X^NY^NT^NE)_\rho$

Finally, we lower bound $H^{\epsilon_s}_{\min}(A^N|X^NY^NT^NE)_\rho$. We use the asymptotic equipartition property (see Theorem 3) to lower bound the min-entropy of N rounds by the von Neumann entropy of single rounds:

$$\begin{aligned}&H^{\epsilon_s}_{\min}(A^N|X^NY^NT^NE)_\rho \\ &\geqslant NH(A|XYTE)_\rho - 4\sqrt{N}\log(2\sqrt{2^{\log_2 d}}+1)\sqrt{\log\frac{2}{\epsilon_s^2}}.\end{aligned} \tag{B17}$$

Since Alice's actions (and her device's) are independent of Bob's choice of input, adding information about Y (Bob's input) does not increase (or decrease) the conditional von Neumann entropy $H(A|X,E)_\rho$. Since $H(A|X,E)_\rho$ and $H(A|XYE,T=1)_\rho$ are equivalent in our setup, we will use both terms interchangeably. In the general scenario, the conditional von Neumann entropy is hard to calculate analytically. But the conditional von Neumann entropy can be lower bounded by the conditional min-entropy as

$$H(A|XYT,E)_\rho \geqslant H_{\min}(A|XYT,E)_\rho. \tag{B18}$$

The advantage of looking at the conditional min-entropy is that we can express it as $H_{\min}(A|XYE,T=1)_\rho = -\log_2 P_g(A|X,E)$ [43], where $P_g(A|X,E)$ is Eve's guessing probability about Alice's X-measurement results A conditioned on her side information E. $P_g(A|X,E)$ can be upper bounded by a function G_x of the expected Bell violation $B[\mathbf{P}]$ [26] by solving a semidefinite program [44], i.e., $P_g(A|X,E) \leqslant G_x(B[\mathbf{P}])$. For our specific protocol (see Sec. IV), we will lower bound the min-entropy (via upper bounding the guessing probability $P_g(A|X,E)$ using the Bell inequality B and corresponding Bell value $B[\hat{\mathbf{P}}_2] - \delta_{\text{est}} - \delta_{\text{con}}$ (explained in Sec. V):

$$H_{\min}(A|XYE,T=1)_\rho \geqslant -\log_2 G_x(B[\hat{\mathbf{P}}_2] - \delta_{\text{est}} - \delta_{\text{con}}). \tag{B19}$$

Finally, putting Eqs. (B16) and (B19) together, we have either the protocol mentioned in Sec. IV aborts with probability higher than $1-(\epsilon_{\text{con}}+\epsilon^c_{\text{EC}})$ or a $(2\epsilon_{\text{EC}}+\epsilon_s+\epsilon_{\text{PA}})$-correct and secret key can be generated of length l. The length l is bounded by

$$\begin{aligned}l \leqslant\ & N[-\log_2 G_x(B[\hat{\mathbf{P}}_2] - \delta_{\text{est}} - \delta_{\text{con}}) \\ & - (1-\xi-\eta)f(\hat{Q}+\gamma_{\text{est}}) - (\xi+\eta)\log_2 d] \\ & - \sqrt{N}\left(4\log(2\sqrt{2^{\log_2 d}}+1)\left(\sqrt{\log\frac{8}{\epsilon'^2_{\text{EC}}}} + \sqrt{\log\frac{2}{\epsilon_s^2}}\right)\right) \\ & - \log\left(\frac{8}{\epsilon'^2_{\text{EC}}} + \frac{2}{2-\epsilon'_{\text{EC}}}\right) - \log\frac{1}{\epsilon_{\text{EC}}} - 2\log\frac{1}{2\epsilon_{\text{PA}}}.\end{aligned} \tag{B20}$$

TABLE VII. Bell inequality table for the [2,3] scenario.

$h^{11}_{A_1B_1}$	$h^{12}_{A_1B_1}$	$h^{13}_{A_1B_1}$	$h^{11}_{A_1B_2}$	$h^{12}_{A_1B_2}$	$h^{13}_{A_1B_2}$
$h^{21}_{A_1B_1}$	$h^{22}_{A_1B_1}$	$h^{23}_{A_1B_1}$	$h^{21}_{A_1B_2}$	$h^{22}_{A_1B_2}$	$h^{23}_{A_1B_2}$
$h^{31}_{A_1B_1}$	$h^{32}_{A_1B_1}$	$h^{33}_{A_1B_1}$	$h^{31}_{A_1B_2}$	$h^{32}_{A_1B_2}$	$h^{33}_{A_1B_2}$
$h^{11}_{A_2B_1}$	$h^{12}_{A_2B_1}$	$h^{13}_{A_2B_1}$	$h^{11}_{A_2B_2}$	$h^{12}_{A_2B_2}$	$h^{13}_{A_2B_2}$
$h^{21}_{A_2B_1}$	$h^{22}_{A_2B_1}$	$h^{23}_{A_2B_1}$	$h^{21}_{A_2B_2}$	$h^{22}_{A_2B_2}$	$h^{23}_{A_2B_2}$
$h^{31}_{A_2B_1}$	$h^{32}_{A_2B_1}$	$h^{33}_{A_2B_1}$	$h^{31}_{A_2B_2}$	$h^{32}_{A_2B_2}$	$h^{33}_{A_2B_2}$

TABLE VIII. Bell inequality table for the $[m, d]$ scenario.

$h^{11}_{A_1B_1}$	$\cdots$	$h^{1d}_{A_1B_1}$		$\cdots$		$h^{11}_{A_1B_m}$	$\cdots$	$h^{1d}_{A_1B_m}$
$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
$h^{d1}_{A_1B_1}$	$\cdots$	$h^{dd}_{A_1B_1}$		$\cdots$		$h^{d1}_{A_1B_m}$	$\cdots$	$h^{dd}_{A_1B_m}$
	$\cdots$			$\cdots$			$\cdots$	
$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
	$\cdots$			$\cdots$			$\cdots$	
$h^{11}_{A_mB_1}$	$\cdots$	$h^{1d}_{A_mB_1}$		$\cdots$		$h^{11}_{A_mB_m}$	$\cdots$	$h^{1d}_{A_mB_m}$
$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
$h^{d1}_{A_mB_1}$	$\cdots$	$h^{dd}_{A_mB_1}$		$\cdots$		$h^{d1}_{A_mB_m}$	$\cdots$	$h^{dd}_{A_mB_m}$

APPENDIX C: MEASUREMENT SETTINGS

Here we list the explicit measurement settings employed in Sec. VI A.

$$\begin{aligned} &x=1 \Rightarrow \sigma_z, && y=1 \Rightarrow \begin{bmatrix} 0.7064 & -0.6632+0.2473i \\ -0.6632-0.2473i & -0.7064 \end{bmatrix}, \\ &x=2 \Rightarrow \begin{bmatrix} -0.1817 & 0.1307+0.9746i \\ 0.1307-0.9746i & 0.1817 \end{bmatrix}, && y=2 \Rightarrow \begin{bmatrix} -0.6882 & -0.2128-0.6936i \\ -0.2128+0.6936i & 0.6882 \end{bmatrix}, \\ &x=3 \Rightarrow \begin{bmatrix} -0.7746 & 0.6186-0.1315i \\ 0.6186+0.1315i & 0.7746 \end{bmatrix}, && y=3 \Rightarrow \begin{bmatrix} 0.4046 & -0.1960+0.8932i \\ -0.1960-0.8932i & -0.4046 \end{bmatrix}. \end{aligned} \tag{C1}$$

Using the following set of measurement settings for Alice and Bob in Eq. (C1), one can generate a higher secret key rate employing our method than using any subset of two measurement settings per party using the standard CHSH inequality.

$$\begin{aligned} &x=1 \Rightarrow \sigma_z, && y=1 \Rightarrow \begin{bmatrix} -0.4091 & -0.5937+0.6930i \\ -0.5937-0.6930i & 0.4091 \end{bmatrix}, \\ &x=2 \Rightarrow \begin{bmatrix} 0.7019 & 0.5167-0.4903i \\ 0.5167+0.4903i & -0.7019 \end{bmatrix}, && y=2 \Rightarrow \begin{bmatrix} -0.6133 & -0.2514+0.7488i \\ -0.2514-0.7488i & 0.6133 \end{bmatrix}. \end{aligned} \tag{C2}$$

Using the following measurement settings in Eq. (C2) and the state in Eq. (19) with no white noise, one cannot extract a secret key using our method or blindly using the CHSH inequality.

$$x=3 \Rightarrow \begin{bmatrix} -0.1457 & -0.9777+0.1513i \\ -0.9777-0.1513i & 0.1457 \end{bmatrix}, \quad y=3 \Rightarrow \begin{bmatrix} -0.9020 & -0.3795-0.2056i \\ -0.3795+0.2056i & 0.9020 \end{bmatrix}. \tag{C3}$$

However, by adding another set of measurements for Alice and Bob mentioned in Eq. (C3), it is possible to achieve a nonzero secret key rate using our method.

APPENDIX D: TABULAR REPRESENTATION OF BELL INEQUALITY

Here we introduce an alternative representation of the hyperplane vector [see Eq. (5)]. We rearrange the entries (coefficients of the Bell inequality) in a tabular construction. For the [2,2] Bell scenario, it is represented in Table VI.

This representation is used in Table II. Similarly, we reorder the elements of the hyperplane vector for the [2,3] Bell scenario in the following way (see Table VII):

This tabular representation is used to describe the Bell inequality in Table III. For the generalized $[m, d]$ scenario, the reordered hyperplane vector is represented in Table VIII.

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

[2] A. K. Ekert, Quantum Cryptography Based on Bell's Theorem, Phys. Rev. Lett. **67**, 661 (1991).

[3] C. H. Bennett, Quantum Cryptography Using any Two Nonorthogonal States, Phys. Rev. Lett. **68**, 3121 (1992).

[4] D. Bruß, Optimal Eavesdropping in Quantum Cryptography with Six States, Phys. Rev. Lett. **81**, 3018 (1998).

[5] R. Renner, Security of quantum key distribution, Int. J. Quantum Inf. **06**, 1 (2008).

[6] H.-K. Lo, X. Ma, and K. Chen, Decoy State Quantum Key Distribution, Phys. Rev. Lett. **94**, 230504 (2005).

[7] D. Gottesman, H.-K. Lo, N. Lutkenhaus, and J. Preskill, Security of quantum key distribution with imperfect devices, in *International Symposium on Information Theory, ISIT 2004* (IEEE, New York, 2004), p. 136.

[8] P. W. Shor and J. Preskill, Simple Proof of Security of the BB84 Quantum Key Distribution Protocol, Phys. Rev. Lett. **85**, 441 (2000).

[9] V. Scarani, H. Bechmann-Pasquinucci, N. J. Cerf, M. Dušek, N. Lütkenhaus, and M. Peev, The security of practical quantum key distribution, Rev. Mod. Phys. **81**, 1301 (2009).

[10] X. Ma, B. Qi, Y. Zhao, and H.-K. Lo, Practical decoy state for quantum key distribution, Phys. Rev. A **72**, 012326 (2005).

[11] H.-K. Lo, M. Curty, and K. Tamaki, Secure quantum key distribution, Nat. Photon. **8**, 595 (2014).

[12] M. Tomamichel, C. Lim, N. Gisin, and R. Renner, Tight finite-key analysis for quantum cryptography, Nat. Commun. **3**, 634 (2012).

[13] L. Lydersen, C. Wiechers, C. Wittmann, D. Elser, J. Skaar, and V. Makarov, Hacking commercial quantum cryptography systems by tailored bright illumination, Nat. Photon. **4**, 686 (2010).

[14] I. Gerhardt, Q. Liu, A. Lamas-Linares, J. Skaar, C. Kurtsiefer, and V. Makarov, Full-field implementation of a perfect eavesdropper on a quantum cryptography system, Nat. Commun. **2**, 349 (2011).

[15] Y. Zhao, C. H. F. Fung, B. Qi, C. Chen, and H. K. Lo, Quantum hacking: Experimental demonstration of time-shift attack against practical quantum-key-distribution systems, Phys. Rev. A **78**, 042333 (2008).

[16] D. Mayers and A. Yao, Quantum cryptography with imperfect apparatus, in *Proceedings of the 39th Annual Symposium on Foundations of Computer Science* (IEEE, Piscataway, NJ, 1998), pp. 503–509.

[17] J. Barrett, L. Hardy, and A. Kent, No Signaling and Quantum Key Distribution, Phys. Rev. Lett. **95**, 010503 (2005).

[18] A. Acín, N. Brunner, N. Gisin, S. Massar, S. Pironio, and V. Scarani, Device-Independent Security of Quantum Cryptography against Collective Attacks, Phys. Rev. Lett. **98**, 230501 (2007).

[19] A. Acin, N. Gisin, and L. Masanes, From Bell's Theorem to Secure Quantum Key Distribution, Phys. Rev. Lett. **97**, 120405 (2006).

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

[21] E. Hänggi and R. Renner, Device-independent quantum key distribution with commuting measurements, arXiv:1009.1833.

[22] E. Hänggi, R. Renner, and S. Wolf, Efficient device-independent quantum key distribution, in *Annual International Conference on the Theory and Applications of Cryptographic Techniques* (Springer, New York, 2010), pp. 216–234.

[23] L. Masanes, R. Renner, M. Christandl, A. Winter, and J. Barrett, Full security of quantum key distribution from no-signaling constraints, IEEE Trans. Inf. Theory **60**, 4973 (2014).

[24] R. Arnon-Friedman, F. Dupuis, O. Fawzi, R. Renner, and T. Vidick, Practical device-independent quantum cryptography via entropy accumulation, Nat. Commun. **9**, 459 (2018).

[25] R. Arnon-Friedman, R. Renner, and T. Vidick, Simple and tight device-independent security proofs, SIAM J. Comput. **48**, 181 (2019).

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

[27] J. Ribeiro, G. Murta, and S. Wehner, Fully device-independent conference key agreement, Phys. Rev. A **97**, 022307 (2018).

[28] T. Holz, H. Kampermann, and D. Bruß, A genuine multipartite Bell inequality for device-independent conference key agreement, Phys. Rev. Res. **2**, 023251 (2020).

[29] U. Vazirani and T. Vidick, Fully Device-Independent Quantum Key Distribution, Phys. Rev. Lett. **113**, 140501 (2014).

[30] O. Nieto-Silleras, S. Pironio, and J. Silman, Using complete measurement statistics for optimal device-independent randomness evaluation, New J. Phys. **16**, 013035 (2014).

[31] J.-D. Bancal, L. Sheridan, and V. Scarani, More randomness from the same data, New J. Phys. **16**, 033011 (2014).

[32] I. Pitowski, *Quantum Probability–Quantum Logic*, Lecture Notes in Physics (Springer Nature Switzerland AG, Cham, Switzerland, 1989).

[33] A. Fine, Hidden Variables, Joint Probability, and the Bell Inequalities, Phys. Rev. Lett. **48**, 291 (1982).

[34] I. Pitowsky, Correlation polytopes: Their geometry and complexity, Math. Program. **50**, 395 (1991).

[35] J. S. Bell, On the Einstein Podolsky Rosen paradox, Phys. Phys. Fiz. **1**, 195 (1964).

[36] J. Szangolies, H. Kampermann, and D. Bruß, Device-Independent Bounds on Detection Efficiency, Phys. Rev. Lett. **118**, 260401 (2017).

[37] W. Hoeffding, Probability inequalities for sums of bounded random variables, in *The Collected Works of Wassily Hoeffding*, edited by N. I. Fisher and P. K. Sen, Springer Series in Statistics (Springer Nature Switzerland AG, Cham, Switzerland, 1994), pp. 409–426.

[38] F. Grasselli, *Quantum Cryptography: From Key Distribution to Conference Key Agreement, Quantum Science and Technology Series* (Springer Nature Switzerland AG, Cham, Switzerland, 2020).

[39] N. J. Beaudry, Assumptions in quantum cryptography, arXiv:1505.02792.

[40] V. Scarani and R. Renner, Security bounds for quantum cryptography with finite resources, in *Workshop on Quantum Computation, Communication, and Cryptography* (Springer, 2008), pp. 83–95.

[41] I. Devetak and A. Winter, Distillation of secret key and entanglement from quantum states, Proc. R. Soc. A **461**, 207 (2005).

[42] G. Murta, S. B. van Dam, J. Ribeiro, R. Hanson, and S. Wehner, Towards a realization of device-independent quantum key distribution, Quantum Sci. Technol. **4**, 035011 (2019).

[43] R. Konig, R. Renner, and C. Schaffner, The operational meaning of min- and max-entropy, IEEE Trans. Inf. Theory **55**, 4337 (2009).

[44] N. Johnston, GETLAB: A Matlab toolbox for quantum entanglement, version 0.9, getlab.com, 2016.

[45] M. Navascués, S. Pironio, and A. Acín, Bounding the Set of Quantum Correlations, Phys. Rev. Lett. **98**, 010401 (2007).

[46] M. Navascués, S. Pironio, and A. Acín, A convergent hierarchy of semidefinite programs characterizing the set of quantum correlations, New J. Phys. **10**, 073013 (2008).

[47] J. Löfberg, YALMIP: A toolbox for modeling and optimization in Matlab, in *Proceedings of the IEEE CACSD Conference* (IEEE, New York, 2004).

[48] M. Grant and S. Boyd, CVX: Matlab software for disciplined convex programming, version 2.1. 2014, http://cvxr.com/cvx.

[49] S. Boyd, S. P. Boyd, and L. Vandenberghe, *Convex Optimization* (Cambridge University Press, Cambridge, England, 2004).

[50] M. Grant and S. Boyd, Graph implementations for nonsmooth convex programs, in *Recent Advances in Learning and Control*, edited by V. Blondel, S. Boyd, and H. Kimura, Lecture Notes in Control and Information Sciences (Springer-Verlag, Berlin, 2008), pp. 95–110, http://stanford.edu/~boyd/graph_dcp.html.

[51] P. Wittek, Algorithm 950: Ncpol2sdpa—Sparse semidefinite programming relaxations for polynomial optimization problems of noncommuting variables, ACM Trans. Math. Softw. **41**, 1 (2015).

[52] N. Johnston, QETLAB: A Matlab toolbox for quantum entanglement, version 0.9, http://qetlab.com, Jan. 2016.

[53] K.-C. Toh, M. J. Todd, and R. H. Tütüncü, SDPT3—A Matlab software package for semidefinite programming, Version 1.3, Optim. Method. Softw. **11**, 545 (1999).

[54] J. F. Sturm, Using SeDuMi 1.02, A Matlab toolbox for optimization over symmetric cones, Optim. Method. Softw. **11**, 625 (1999).

[55] M. ApS, MOSEK optimization toolbox for MATLAB, *User's Guide and Reference Manual, Version*, 4 (2019).

[56] K. Brádler, M. Mirhosseini, R. Fickler, A. Broadbent, and R. Boyd, Finite-key security analysis for multilevel quantum key distribution, New J. Phys. **18**, 073030 (2016).

[57] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, Proposed Experiment to Test Local Hidden-Variable Theories, Phys. Rev. Lett. **23**, 880 (1969).

[58] A. Acin, T. Durt, N. Gisin, and J. I. Latorre, Quantum nonlocality in two three-level systems, Phys. Rev. A **65**, 052325 (2002).

[59] D. Collins, N. Gisin, N. Linden, S. Massar, and S. Popescu, Bell Inequalities for Arbitrarily High-Dimensional Systems, Phys. Rev. Lett. **88**, 040404 (2002).

[60] D. Collins and N. Gisin, A relevant two qubit Bell inequality inequivalent to the ChSh inequality, J. Phys. A: Math. Gen. **37**, 1775 (2004).

[61] E. A. Fonseca and F. Parisio, Measure of nonlocality which is maximal for maximally entangled qutrits, Phys. Rev. A **92**, 030101(R) (2015).

[62] V. Lipinska, F. J. Curchod, A. Máttar, and A. Acín, Towards an equivalence between maximal entanglement and maximal quantum nonlocality, New J. Phys. **20**, 063043 (2018).

[63] A. de Rosier, J. Gruca, F. Parisio, T. Vértesi, and W. Laskowski, Multipartite nonlocality and random measurements, Phys. Rev. A **96**, 012101 (2017).

[64] A. de Rosier, J. Gruca, F. Parisio, T. Vértesi, and W. Laskowski, Strength and typicality of nonlocality in multisetting and multipartite Bell scenarios, Phys. Rev. A **101**, 012116 (2020).

[65] A. Fonseca, A. de Rosier, T. Vértesi, W. Laskowski, and F. Parisio, Survey on the Bell nonlocality of a pair of entangled qudits, Phys. Rev. A **98**, 042105 (2018).

[66] A. Barasiński and M. Nowotarski, Volume of violation of Bell-type inequalities as a measure of nonlocality, Phys. Rev. A **98**, 022132 (2018).

[67] P. Massart, *Concentration Inequalities and Model Selection*, Lecture Notes in Mathematics Vol. 6. (Springer, Berlin, 2007).

[68] S. Boucheron, G. Lugosi, and P. Massart, *Concentration Inequalities: A Nonasymptotic Theory of Independence* (Oxford University Press, Oxford, 2013).

[69] F. Chung and L. Lu, Concentration inequalities and martingale inequalities: A survey, Internet Math. **3**, 79 (2006).

[70] F. Xu, Y.-Z. Zhang, Q. Zhang, and J.-W. Pan, Device-independent quantum key distribution with random post selection, arXiv:2110.02701.

[71] W.-Z. Liu, Y.-Z. Zhang, Y.-Z. Zhen, M.-H. Li, Y. Liu, J. Fan, F. Xu, Q. Zhang, and J.-W. Pan, High-speed device-independent quantum key distribution against collective attacks, arXiv:2110.01480.

[72] E. Y.-Z. Tan, C. C.-W. Lim, and R. Renner, Advantage Distillation for Device-Independent Quantum Key Distribution, Phys. Rev. Lett. **124** (2), 020502 (2020).

[73] R. Schwonnek, K. T. Goh, I. W. Primaatmaja, E. Y.-Z. Tan, R. Wolf, V. Scarani, and C. C.-W. Lim, Device-independent quantum key distribution with random key basis, Nat. Commun. **12**, 2880 (2021).

[74] P. Brown, H. Fawzi, and O. Fawzi, Computing conditional entropies for quantum correlations, Nat. Commun. **12**, 1 (2021).

[75] P. Brown, H. Fawzi, and O. Fawzi, Device-independent lower bounds on the conditional von neumann entropy, arXiv:2106.13692.

[76] M. Tomamichel, *Quantum Information Processing with Finite Resources: Mathematical Foundations* (Springer, New York, 2015), Vol. 5.

[77] R. Renner and S. Wolf, Simple and tight bounds for information reconciliation and privacy amplification, in *Advances in Cryptology–ASIACRYPT 2005*, edited by B. Roy, Lecture Notes in Computer Science Vol. 3788 (Springer, Berlin, Heidelberg, 2005).

[78] A. Vitanov, F. Dupuis, M. Tomamichel, and R. Renner, Chain rules for smooth min-and max-entropies, IEEE Trans. Inf. Theory **59**, 2603 (2013).

[79] J. L. Carter and M. N. Wegman, Universal classes of hash functions, J. Comput. Syst. Sci. **18**, 143 (1979).

[80] M. Tomamichel, C. Schaffner, A. Smith, and R. Renner, Leftover hashing against quantum side information, IEEE Trans. Inf. Theory **57**, 5524 (2011).

[81] M. Tomamichel, R. Colbeck, and R. Renner, A fully quantum asymptotic equipartition property, IEEE Trans. Inf. Theory **55**, 5840 (2009).

[82] F. Grasselli, H. Kampermann, and D. Bruß, Conference key agreement with single-photon interference, New J. Phys. **21**, 123002 (2019).

[83] H.-L. Yin and Z.-B. Chen, Finite-key analysis for twin-field quantum key distribution with composable security, Sci. Rep. **9**, 17113 (2019).

[84] M. A. Nielsen and I. Chuang, *Quantum Computation and Quantum Information*, 1st ed (Cambridge University Press, Cambridge, 2011).

[85] M. Tomamichel and A. Leverrier, A largely self-contained and complete security proof for quantum key distribution, Quantum **1**, 14 (2017).