Tight finite-key analysis for generalized high-dimensional quantum key distribution

Rong Wang,1,2 Zhen-Qiang Yin,1,2 Hang Liu,1,2 Shuang Wang,1,2 Wei Chen,1,2 Guang-Can Guo,1,2 and Zheng-Fu Han1,2

1CAS Key Laboratory of Quantum Information, CAS Center for Excellence in Quantum Information and Quantum Physics, University of Science and Technology of China, Hefei 230026, China
2State Key Laboratory of Cryptology, P. O. Box 5159, Beijing 100878, P. R. China

Due to the capability of tolerating high error rate and generating more key bits per trial, high-dimensional quantum key distribution attracts wide interest. Despite great progresses in high-dimensional quantum key distribution, there are still some gaps between theory and experiment. One of these is that the security of secret key heavily depends on the number of the emitted signals. So far, the existing security proofs are only suitable in the case with infinite or unpractically large number of the emitted signals. Here, by introducing the idea of "key classification" and developing relevant techniques based on the uncertainty relation for smooth entropies, we propose a tight finite-key analysis suitable for generalized high-dimensional quantum key distribution protocols. Benefitting from our theory, high-dimensional quantum key distribution protocols with finite resources become experimentally feasible.

INTRODUCTION

Quantum key distribution (QKD), considered as the first application in quantum information science, can provide two distant parties Alice and Bob with a string of secret key bits by the laws of quantum mechanics. Because of this amazing feature, it has been rapidly developed in both theory and experiment over last three decades [1–11]. Among all the proposed QKD protocols, most of them are based on qubit systems, such as the well-known BB84 protocol [1]. QKD protocols using qubit systems are very mature both in theory and experiment, but in some scenarios, their performances are limited due to the dimensionality. For instance, each qubit can distribute at most 1 key bit. As our requirements for protocols performance increase, more and more novel protocols have been proposed. Some of them can tolerate high error rate such as six-state protocol [12], some of them carry more than one secret key each signal [13]. Some of these QKD protocols prepare quantum states in a Hilbert space larger than 2, while others may prepare and measure quantum states in 2 or more bases. That is the reason we call them high-dimensional(HD) QKD.

Since HD-QKD has various advantages, scholars have made a lot of efforts both in its security proofs and experimental techniques [13–16]. However, the most existing security proofs [13–17,18] are only available under the assumption that we have infinite resources. In another word, the two parties Alice and Bob are required to exchange arbitrarily large quantum signals N, which cannot be achieved by practical equipments. When we remove the infinite resources assumption, that is, when we consider the finite key issue, several security proofs [17,19,20] have been proposed for some specific HD-QKD protocols. Frustratingly, the number of exchanged quantum signals N is usually too large to be realized. Thus, a more efficient method to reduce N to an acceptable level is an urgent need. Besides, the existing proof [21] for HD-QKD protocols are not general, e.g. Bob is assumed to make measurements along only two bases albeit coding states are qudit systems.

Here, we propose an efficient method to tackle finite key issues for generalized HD-QKD protocols, i.e. the dimension of Hilbert space is arbitrary and Bob’s measurement bases can be multiple. The proposed method can cover the previous proof technique [21] that is only suitable for two measurements bases. The essential of our method is introducing the idea of ”key classification”, which means classifying key bits into different types with different bit error patterns. Futhermore, applying uncertainty relation [22] for smooth entropies [23] to each type and developing relevant theoretical techniques, we derive a tight bound of secret key rate for HD-QKD in finite-key scenarios. Compared with previous methods including de Finetti theorem [24] and post-selection technique [22], our method leads to more optimistic bound. Through numerical simulations, we show that, for a variety of HD-QKD protocols, the number of exchanged quantum states N can be reduced dramatically thanks to the proposed theory.

RESULTS

Security definition. Before stating our new proof technique, let us review the security framework [5,24] that we are concerned about in this paper. A general QKD protocol is executed by two distant parties Alice and Bob. Bob receives the signals from an insecure quantum channel. Then Alice and Bob output either a pair of bit strings $S_A$ and $S_B$, or a symbol $\perp$ to indicate the abort of the protocol.

According to the definition of security, a QKD protocol has to satisfy three criteria called "correctness", "secrecy" and "robustness". Owing to the practical implementation, it is impossible to guarantee $S_A = S_B$. Then a QKD protocol is $\varepsilon_{\text{cor}}$-correct, if it is $\varepsilon_{\text{cor}}$-indistinguishable from a $S_A = S_B$ protocol.
protocol is $\varepsilon_{sec}$-secret, if
\begin{equation}
\min_{\sigma_E} \frac{1}{2} \| \rho_{AE} - U_A \otimes \sigma_E \| \leq \varepsilon_{sec},
\end{equation}

where $U_A$ is the fully mixed state of Alice’s system, $\rho_{AE}$ is the composed state of Alice and Eve, and $\| \cdot \|$ denotes the trace norm. Finally, a protocol is $\varepsilon_{rob}$-robust, if the probability that the protocol aborts is no bigger than $\varepsilon_{rob}$. In this work, we just consider the correctness and secrecy of a QKD protocol. Thereby, we say a QKD protocol is $\varepsilon_{tot}$-security, if it is both $\varepsilon_{cor}$-correct and $\varepsilon_{sec}$-secret, with $\varepsilon_{cor} + \varepsilon_{sec} \leq \varepsilon_{tot}$.

Based on this security definition, we are able to guarantee the security when we use our new technique in the HD-QKD protocols.

**Protocol definition.** In this work, we take $(d+1)$-basis QKD protocols, i.e. the generalization of the six-state protocols, as the examples to introduce our proof technique. Below, we present some assumptions of the devices first. On the one side of the insecure quantum channel, Alice controls her devices to prepare $d$-level ($d$ is a prime number in this work) quantum states which we also call qudits. We recall that there are at most $d+1$ mutually unbiased bases (MUBs) in the $d$-level Hilbert space. Then, Alice randomly chooses one of MUBs and encode the key bit into its one orthogonal eigen-vector. After Bob receiving the particle, he is able to randomly choose one MUB to do projective measurement. In order to clearly describe the protocols, we list some notations as follows.

First, we review the definition of ”overlap”. The overlap of any two measurements is defined as $c = \max[\| \sqrt{M_y} |Y\rangle \|, \| \sqrt{N_z} |Z\rangle \|]$, where $\{M_y\}$ and $\{N_z\}$ are the elements of the positive operator valued measurements (POVMs) of $Y$ basis and $Z$ basis, respectively. In this paper, we heavily rely on the fact that the overlap of any two POVMs of MUB in $d$-level Hilbert space is $1/d$.

Second, the prepared states are guaranteed to be $d$-level quantum states chosen from $d+1$ MUB $X_{j,k} \in \{X_{0,1}, X_{1,0}, \cdots, X_{1,k}, \cdots, X_{1,d-1}\}$, where the notions are analogous to [17]. On the other side of the channel, Bob controls his devices to measure quantum states in these $d+1$ basis. Thus, there exits an equivalent entanglement-based (EB) protocol according to the model described above.

Third, under the EB version of protocol, Alice prepares two entangled quantum states and sends one of them to Bob in each trial. At measurement, we assume that Bob is able to delay all the measurements in $X_{0,1}$-basis until parameter estimation completed. This assumption doesn’t affect the final key rate if the measurement statistics is same as the ones of actual devices. According to [27], the composed quantum states before measurements have the simple form

\begin{equation} \rho_{AB}^n = \sum_{j_k=0}^{d-1} \rho_{n_{00},\cdots, n_{jk},\cdots, n_{d-1,d-1}} \mu_{n_{00},\cdots, n_{jk},\cdots, n_{d-1,d-1}} \rho_{00,\cdots, 0} \cdots \rho_{d-1,d-1} \end{equation}

In this formula, the sum is taken over all $n_{00}, \cdots, n_{jk}, \cdots, n_{d-1,d-1}$ satisfying $\sum_{j_k=0}^{d-1} n_{jk} = n$ and $\mu_{n_{00},\cdots, n_{jk},\cdots, n_{d-1,d-1}}$ are some non-negative coefficients. Moreover, there exits a unitary operation $\pi$ on $H_{AB}^{n}$ which permutes the $n$ subsystems, so that $\rho_{n_{00},\cdots, n_{jk},\cdots, n_{d-1,d-1}}$ can be given by

\begin{equation} \rho_{n_{00},\cdots, n_{jk},\cdots, n_{d-1,d-1}} = \pi(\otimes_{j_k=0}^{d-1} |\Phi_{jk}\rangle \langle \Phi_{jk}|)^{\otimes n_{jk}}. \end{equation}

In this expression, the generalized Bell basis states $|\Phi_{jk}\rangle = \sum_{s=0}^{d-1} \omega^{sk} |s, s+j\rangle$ ($j_k \in \{0,1,\cdots, d-1\}$ and $\omega$ is the $d$th root of unity) [17] belong to the composed Hilbert space of Alice and Bob denoted by $H_{AB}$.

Finally, in practical optical schemes, $(d+1)$-basis QKD protocols are often realized by weak coherent light rather than single-photon source. And this doesn’t meet the assumption that Alice prepares $d$-level quantum states. Inspired by Lim et al.’s work [28], the finite-key analysis under this case can intuitively be solved by using decoy states [24 32].

We now define a family of $(d+1)$-basis QKD protocols, $\Phi[n, m, l, \varepsilon_{cor}, \text{leak}_{EC}]$, where $n$ is the block size with respect to the shifted keys in $X_{0,1}$-basis, $m$ is the number of dits used to do parameter estimation with regard to each basis, $l$ is the secret key length, $\varepsilon_{cor}$ is the required correctness, and $\text{leak}_{EC}$ is the information leakage in error correction. The protocol is asymmetric, specifically, the $n$ shifted keys used for producing final secret keys are measured in $X_{0,1}$-basis, the other $(d + 1) \times m$ dits used for parameter estimation are measured in all $d + 1$ basis. Therefore, the number of total sifted keys is defined as $N = n + (d + 1) \times m$.

The protocol is described in **Box 1**
Box 1  Protocol definition.

State Preparation: Alice and Bob repeat the first four steps the protocol for \( i = 1, \ldots, M \) until the condition in the Sifting step is met. Alice chooses a basis \( X_i \in \{X_{0,1}, X_{1,0}, \ldots, X_{1,d-1}\} \), where \( X_{j,k} \) is chosen with probability \( p_{j,k} \) respectively. Here we choose \( p_{0,1} = f(n, m) \) and \( p_{1,k} = (1 - p_{0,1})/d \) that the function \( f(n, m) \) is chosen to minimize the number \( M \) of exchanged quantum states. Then, Alice chooses a random dits \( r_i \in \{0, 1, \ldots, d-1\} \) and prepares the quantum state corresponding to \( r_i \) in a basis \( X_i \).

Distribution: Alice sends the quantum state over the insecure channel to Bob.

Measurement: Bob also chooses a basis \( \tilde{X}_i \in \{X_{0,1}, X_{1,0}, \ldots, X_{1,k}, \ldots, X_{1,d-1}\} \) with probability \( p_{j,k} \). After receiving the state, Bob measures it in the chosen basis and stores the outcome \( \tilde{r}_i \in \{0, \ldots, d-1\} \).

Sifting: Alice and Bob broadcast their basis settings over an classical authenticated channel. We define the sets \( X_{0,1} := \{i : X_i = \tilde{X}_i = X_{0,1}\} \) and \( X_{1,k} := \{i : X_i = \tilde{X}_i = X_{1,k}\} \). The protocol repeats the first four steps unless \( |X_{0,1}| \geq n + m \) and \( |X_{1,k}| \geq m \) for each \( k \in \{0, \ldots, d-1\} \).

Parameter Estimation: Alice and Bob use \( n \) random dits from \( X_{0,1} \) to form the code dit strings \( X_{0,1}^n \) and \( \tilde{X}_{0,1}^n \) respectively. Then, for \( m \) dits from \( X_{j,k} \in \{X_{0,1}, X_{1,0}, \ldots, X_{1,k}, \ldots, X_{1,d-1}\} \), they compute \( d \) types of statistical parameters \( q_{j,k}^{(t)} := \frac{1}{m} \sum_{i} \delta_{j,k}^{(t)} \) where

\[
\delta_{j,k}^{(t)} = \begin{cases} 1, & \tilde{r}_i - r_i \equiv t \pmod{d} \\ 0, & \tilde{r}_i - r_i \not\equiv t \pmod{d} \end{cases}
\]

and \( t \in \{0, 1, 2, \ldots, d-1\} \). Moreover, these parameters satisfy \( \sum_{t=0}^{d-1} q_{j,k}^{(t)} = 1 \) with the probability of no error \( q_{j,k}^{(0)} \) for each basis \( X_{j,k} \). The protocol aborts if the probability of error \( \sum_{t=1}^{d-1} q_{j,k}^{(t)} \) for each basis \( X_{j,k} \) is too high.

Error Correction: For those \( n \) that pass the parameter estimation step, an information reconciliation scheme is applied. This allows Bob to obtain an estimate \( \tilde{X}_{0,1}^n \) of \( X_{0,1}^n \) by Alice sending him \( \text{leak}_{EC} \) bits of error correction data. Then, Alice computes a bit string (a hash) of length \( \log_{2} \frac{1}{\varepsilon_{cor}} \) by using a random two-universal hash function to \( \tilde{X}_{0,1}^n \). She sends the choice of function and the hash to Bob. The protocol aborts if \( \text{hash}(\tilde{X}_{0,1}^n) \neq \text{hash}(X_{0,1}^n) \).

Privacy Amplification: If the \( n \) dits pass the error correction, Alice and Bob apply a random two-universal hash function to \( X_{0,1}^n \) and \( \tilde{X}_{0,1}^n \) to extract the final secret \( l \) bits \( (l \cdot \log d \cdot 2 \text{ dits}) \)

Security analysis. We now present our main result of our paper. It says that the \( (d+1) \)-basis protocols \( \Phi[n, m, l, \varepsilon_{cor}, \text{leak}_{EC}] \) are both \( \varepsilon_{cor} \)-correct and \( \varepsilon_{sec} \)-secret, if the length of secret key is calculated according to given set of observed values. The correctness is guaranteed by the error correction step, where a hash of Alice’s shifted key is compared with the hash of its estimate of Bob. If the length of secret key \( l \) satisfies

\[
l \leq n(\log_d d - H(\xi))(1 - Q - \mu(\varepsilon)) - \text{leak}_{EC} - \log_2 \frac{2}{\varepsilon_{sec} \varepsilon_{cor}}, \tag{4}
\]

the protocols \( \Phi[n, m, l, \varepsilon_{cor}, \text{leak}_{EC}] \) are \( \varepsilon_{sec} \)-secret. In this formula, \( \xi \) is a \( d \)-level probability vector denoted by \( \xi = \{\xi_0, \xi_1, \ldots, \xi_d, \xi_{d+1}\} \), and

\[
\begin{align*}
\xi_0 &= 1 - \frac{d+1}{d}(Q + \mu(\varepsilon)) \\
\xi_1 &= \cdots = \xi_{d-1} = \frac{1}{d}(Q + \mu(\varepsilon)) \\
\xi_d &= \frac{Q}{1 - Q - \mu(\varepsilon)}
\end{align*}
\tag{5}
\]

where \( H(\cdot) \) denotes the entropy function of \( d \)-level probability vector by \( H(\xi) = \sum_{t=0}^{d} -\xi_t \log_2 \xi_t \), \( \varepsilon_{sec} = 2\sqrt{1 - (1 - \varepsilon)^2}d^{d+1} \) and \( \mu(\varepsilon) \) that accounts for statistical fluctuation is given by

\[
\mu(\varepsilon) := \sqrt{\frac{n + m m + 1}{nm m}} \ln \frac{1}{\varepsilon}. \tag{6}
\]

A sketch of the proof of equation (4) can be found in methods section and a rigorous proof including a more general version of the equation (4) can be found in Supplementary Material. When we comes to the asymptotic case of sufficient large block sizes \( n \), the statistical fluctuation term \( \mu(\varepsilon) \) can be neglected, and thus \( l \) satisfies \( l \leq n(\log_d d - H(\xi))(1 - Q) - \text{leak}_{EC} \), as obtained in previous work [13].

DISCUSSION

In this section, we analyze the behavior of our security bounds and compare our bounds with previous results by numerical simulations [17]. For this purpose, we assume that the quantum channel can be simulated as a generalization of the qubit depolarizing channel which leads to

\[
q_{j,k}^{(0)} = 1 - Q, \quad q_{j,k}^{(1)} = \cdots = q_{j,k}^{(d-1)} = \frac{Q}{d-1}, \tag{7}
\]

for each basis \( X_{j,k} \). To maximize the expected key rate, we fix \( \varepsilon_{tot} = 10^{-10} \) and assume an error correction leakage of \( \text{leak}_{EC} = \zeta n H_2(Q + \mu(\varepsilon)) \) where \( \zeta = 1.1 \) is the error correction efficiency and \( H_2(\cdot) \) denotes the binary entropy function.

In Fig. 1, Fig. 2 and Fig. 3, we compare our optimal key rates (defined as \( l/N \)) with the secret key rates in [17] of \((d+1)\)-basis QKD protocols featured by \( d = 2 \), \( d = 3 \) and \( d = 17 \) respectively. As we can see from these
The horizontal dashed lines represent the asymptotic rates for error rate $Q \in \{1\%, 2.5\%, 5\\%\}$ (from top to bottom).

Our new proof technique can give more tight finite-key bounds for other high-dimensional QKD protocols such as tomographic [33, 34] and reference-frame-independent [35] QKD protocols.

**METHODS**

**Secrecy.** Here we briefly discuss our new proof technique that is applied to establish the secrecy of the ($d+1$)-basis protocols $\Phi[n, m, l, \varepsilon_{cor}, \varepsilon_{leak} EC]$. We denote the dit strings of length $n$ by $X^n_{0,1}$ of Alice’s side and $X^n_{0,1}$ of Bob’s side respectively, which are used to extract the final key. Then, after the measurements (based on EB version), the classical-classical-quantum state of Alice, Bob and Eve is given by

$$\rho_{X^n_{0,1}X^n_{0,1}E} = (x^n_{0,1}x^n_{0,1})_{AB} \cdot (x^n_{0,1}x^n_{0,1}) \otimes \rho_{E}^{x^n_{0,1}x^n_{0,1}},$$

where $x^n_{0,1} \in X^n_{0,1}$ and $\bar{x}^{n_{0,1}} \in \bar{X}^{n}_{0,1}$ respectively, and $F_{X^n_{0,1}X^n_{0,1}}(x^n_{0,1}x^n_{0,1})$ is the probability of joint dit string $x^n_{0,1}x^n_{0,1}$. According to the Quantum Leftover Hashing lemma [8, 36], the secret key length is directly dependent on the lower bound of smooth min-entropy of $H^\varepsilon_{min}(X^n_{0,1}|E)$. Then, we use the uncertainty relation for smooth entropies [23] with two constructed POVMs to give a lower bound of $H^\varepsilon_{min}(X^n_{0,1}|E)$.

One is no doubt the $X^n_{0,1}$, the other one is $\pi(X^n_{1,0} \cdots X^n_{1,k} \cdots X^n_{1,d-1})$
where $\sum_{k=0}^{d-1} n_k = n$. Thus, uncertainty relation is given by
\begin{equation}
H_{\min}^n (X_{0,1}^n|E) + H_{\max}^n (\pi (X_{1,0}^{\otimes n_0} \cdots X_{1,d-1}^{\otimes n_{d-1}})|B) \\
\geq \log_2 \frac{1}{e^n} = n \log_2 d.
\end{equation}

It is not obvious to apply this formula to the state $\rho_{X_{0,1}^n, E}$, to solve this problem, we introduce the idea of "classification". For any joint dit strings $x_{0,1}^n, y_{0,1}^n$, we can find that the $i$th pair dits occur $t (r_i - r_i \pmod d = t)$ type change including error ($t \neq 0$) or no error ($t = 0$) happens. According to this feature, we define new dit strings $y_{q_0,1}^n \in \mathcal{Y}_{q_0,1}^n$ that are given by
\begin{equation}
y_{q_0,1}^n = x_{0,1}^n - x_{0,1}^n \pmod d \tag{10}
\end{equation}
where the subtraction is bitwise and "t" dit occurs $n_k := n r_{q_k}^n$ times in dit strings $y_{q_0,1}^n$. Then we denote a conditional state by
\begin{equation}
\rho_{X_{0,1}^n, y_{q_0,1}^n|E} = \sum_{x_{0,1}^n} \mathbb{P}_{X_{0,1}^n} x_{0,1}^n (x_{0,1}^n \otimes \bar{x}_{0,1}^n) \mid x_{0,1}^n \bar{x}_{0,1}^n \rangle_{AB} \langle x_{0,1}^n \bar{x}_{0,1}^n | \otimes \rho_{E} \bar{x}_{0,1}^n, \tag{11}
\end{equation}
and its corresponding probability by $\mathbb{P}_{\mathcal{Y}_{q_0,1}^n} (y_{q_0,1}^n)$ which is a marginal probability distribution $\mathbb{P}_{X_{0,1}^n, y_{q_0,1}^n}(x_{0,1}^n \bar{x}_{0,1}^n)$. Therefore, $\rho_{X_{0,1}^n, y_{q_0,1}^n|E}$ can be rewritten by
\begin{equation}
\rho_{X_{0,1}^n, y_{q_0,1}^n|E} = \sum_{y_{q_0,1}^n} \mathbb{P}_{y_{q_0,1}^n} y_{q_0,1}^n \rho_{X_{0,1}^n, y_{q_0,1}^n|E} |y_{q_0,1}^n \rangle_{\mathcal{Y}_{q_0,1}^n}, \tag{12}
\end{equation}
where we have "classified" the state $\rho_{X_{0,1}^n, y_{q_0,1}^n|E}$ according to the dit string $y_{q_0,1}^n$. By uncertain relationship with appropriately constructing the POVM $\pi (X_{1,0}^{\otimes n_0} \cdots X_{1,k}^{\otimes n_k} \cdots X_{1,d-1}^{\otimes n_{d-1}})$, we can obtain an almost tight bound of smooth min-entropy of $\rho_{X_{0,1}^n, y_{q_0,1}^n|E}$. With the help of subadditivity of min-entropy [37], we can connect the two smooth min-entropies between $\rho_{X_{0,1}^n, y_{q_0,1}^n|E}$ and $\rho_{X_{0,1}^n, y_{q_0,1}^n|E}$. Finally, we obtain the lower bound of min-entropy given by
\begin{equation}
H_{\min}^n (X_{0,1}^n|E) \geq n (\log_2 d - H(X)) (1 - Q - \mu (\varepsilon)), \tag{13}
\end{equation}
where $\varepsilon = \sqrt{1 - (1 - x^2)^{d+1}}$ and $\mu (\varepsilon)$ that analogously to [21] accounts for statistical fluctuation depends on the security parameter $\tilde{\varepsilon}$.

The approach to constructing the POVM $\pi (X_{1,0}^{\otimes n_0} \cdots X_{1,k}^{\otimes n_k} \cdots X_{1,d-1}^{\otimes n_{d-1}})$ and using the subadditivity of min-entropy can be found in Supplementary Material.

Due to the Quantum Leftover Hashing lemma [38], it is possible to extract a $\Delta$-secret key of length $t$ from $X_{0,1}^n$, where
\begin{equation}
\Delta = \tilde{\varepsilon} + \frac{1}{2} \sqrt{2^{1-H_{\min}^n (X_{0,1}^n|E')} - \log_2 \frac{2}{\varepsilon_{cor}}}. \tag{14}
\end{equation}

The term $E'$ that represents all information Eve obtained can be decomposed as $E' = CE$, where $C$ is classical information revealed by Alice and Bob during the error correction step. For the revealed information $C$ is at most $\text{leak}_{EC} - \log_2 \frac{2}{\varepsilon_{cor}}$ bits, we use a chain rule for smooth entropies and then obtain
\begin{equation}
H_{\min}^n (X_{0,1}^n|E') \geq H_{\min}^n (X_{0,1}^n|E) - \text{leak}_{EC} - \log_2 \frac{2}{\varepsilon_{cor}}. \tag{15}
\end{equation}

If we choose $\varepsilon = \varepsilon_{sec}$/2, combining equation (4) and Quantum Leftover Hashing lemma, we get
\begin{equation}
\Delta \leq \tilde{\varepsilon} + \frac{1}{2} \sqrt{2^{1-H_{\min}^n (X_{0,1}^n|E')} - \log_2 \frac{2}{\varepsilon_{sec}}/2}. \tag{16}
\end{equation}

Thus, these protocols are $\varepsilon_{sec}$-secret.

**Data availability**
The data that support the findings of this study are available from the corresponding author upon reasonable request.

**ACKNOWLEDGEMENTS**
This work has been supported by the National Key Research and Development Program of China (Grant No. 2016YFA0302600), the National Natural Science Foundation of China (Grant Nos. 61822115, 61775207, 61961136004, 61702469, 61771439, 61622506, 61627820, 61675189), National Cryptography Development Fund (Grant No. MMJ20170120) and Anhui Initiative in Quantum Information Technologies.

**Competing interests**
The authors declare that they have no competing interests.

**AUTHOR CONTRIBUTIONS**
Z.-Q.Y., S.W., W.C., G.-C.G., Z.-F.H., and R.W. conceived the basic idea of the security proof. R.W. finished the details of the security proof. R.W. and H. L. designed the simulations. Z.-Q.Y. and R.W. wrote the paper.

---

[1] C. H. Bennett and G. Brassard, in *Proceedings of IEEE International Conference on Computers, Systems and Signal Processing* (IEEE, 1984) pp. 175–179.
Thus, we can find that

\[ F(P, Q) = \sum_{y_{i}^{(t)} \in q_{0,1}} \sqrt{\mathbb{P}_{y_{i}^{(t)} } (y_{i}^{(t)} ) Q y_{i}^{(t)} (y_{i}^{(t)} )} = \sqrt{1 - \varepsilon^2}. \]  

(A.19)
Then the purified distance \[37\] between the distributions is given by \[ P(\mathcal{P}, \mathcal{Q}) = \sqrt{1 - F^2(\mathcal{P}, \mathcal{Q})} = \varepsilon. \]

**Definition 2.** We define the \( i_{th} \) dit of string \( y^n_{q(t)} \), by \( s_i \in \{0,1,\cdots,d-1\} \). According to each dit string \( y^n_{q(t)} \), we construct a corresponding POVM denoted by \( \mathcal{X}^n|y^n_{q(t)} \), which the sub-POVM of \( i_{th} \) subsystem is \( \mathcal{X}_{1,s_i} \in \{\mathcal{X}_{1,0}, \mathcal{X}_{1,1}\cdots \mathcal{X}_{1,d-1}\} \). If we choose a comfortable permutation operation \( \pi \), \( \mathcal{X}^n|x^n_{q(t)} \) can also be written as

\[
\mathcal{X}^n|y^n_{q(t)} = \pi(\mathcal{X}^{n_0}_{1,0} \cdots \mathcal{X}^{n_t}_{1,t} \cdots \mathcal{X}^{n_{d-1}}_{1,d-1}),
\]

(A.20)

where \( n_t = n * q(t) \) and thus satisfy \( \sum_{t=0}^{d-1} n_t = n \).

**Lemma 3.** If \( \mathcal{X}^n|x^n_{q(t)} \) and \( \mathcal{X}^n|y^n_{q(t)} \) are the outcome dit strings after applying \( \mathcal{X}^n|x^n_{q(t)} \) to Alice and Bob’s quantum system respectively, then

\[
H^\varepsilon_{\min}(\mathcal{X}^n_{0,1} | E) + H^\varepsilon_{\max}(\mathcal{X}^n_{y^n_{q(t)}} | \mathcal{X}^n_{y^n_{q(t)}}) \geq n \log_2 d,
\]

(A.21)

where smooth min-entropy is for \( \rho_{X^n_{0,1},E|y^n_{q(t)}} = \text{Tr}^B(\rho_{X^n_{0,1},E|y^n_{q(t)}}) \)

**Proof.** Because the overlap of any two POVMs of MUB in \( d \)-level Hilbert space is \( 1/d \), the overlap of \( \mathcal{X}^n_{0,1} \) and \( \mathcal{X}^n_{y^n_{q(t)}} \) for any \( y^n_{q(t)} \) is \( 1/d^n \). Owing to uncertainty relationship for smooth entropies \[37\], we have

\[
H^\varepsilon_{\min}(\mathcal{X}^n_{0,1} | E) + H^\varepsilon_{\max}(\mathcal{X}^n_{y^n_{q(t)}} | \mathcal{X}^n_{y^n_{q(t)}}) \geq n \log_2 d.
\]

(A.22)

Considering the data-processing inequality for smooth max-entropy \[37\], we have

\[
H^\varepsilon_{\max}(\mathcal{X}^n_{y^n_{q(t)}} | \mathcal{X}^n_{y^n_{q(t)}}) \leq H^\varepsilon_{\max}(\mathcal{X}^n_{y^n_{q(t)}} | \mathcal{X}^n_{y^n_{q(t)}}).
\]

(A.23)

Thus, we complete the proof.

Similarly, if choosing a comfortable permutation operation \( \pi \) corresponding to \( y^n_{q(t)} \), we can rewrite \( \mathcal{X}^n|x^n_{q(t)} \) and \( \mathcal{X}^n|y^n_{q(t)} \) as

\[
\mathcal{X}^n|x^n_{q(t)} = \pi(\mathcal{X}^{n_0}_{1,0} \cdots \mathcal{X}^{n_t}_{1,t} \cdots \mathcal{X}^{n_{d-1}}_{1,d-1})
\]

\[
\mathcal{X}^n|y^n_{q(t)} = \pi(\mathcal{X}^{n_0}_{1,0} \cdots \mathcal{X}^{n_t}_{1,t} \cdots \mathcal{X}^{n_{d-1}}_{1,d-1}),
\]

(A.24)

respectively. For smooth entropies are covariant under unitary operations \[37\], we have

\[
H^\varepsilon_{\max}(\mathcal{X}^n_{y^n_{q(t)}} | \mathcal{X}^n_{y^n_{q(t)}}) = H^\varepsilon_{\max}(\mathcal{X}^n_{1,0} \cdots \mathcal{X}^n_{1,t} \cdots \mathcal{X}^n_{1,d-1} | \mathcal{X}^n_{1,0} \cdots \mathcal{X}^n_{1,t} \cdots \mathcal{X}^n_{1,d-1}).
\]

(A.25)

In the following, we focus on bounding \( H^\varepsilon_{\max}(\mathcal{X}^n_{y^n_{q(t)}} | \mathcal{X}^n_{y^n_{q(t)}}) \) by observed values. Firstly, we note that the correlation of \( \mathcal{X}^n|x^n_{q(t)} \) and \( \mathcal{X}^n|x^n_{q(t)} \) is discussed when Alice and Bob respectively output dit strings \( \mathcal{X}^n_{0,1} \) and \( \mathcal{X}^n_{1,1} \) that satisfy \( y^n_{q(t)} = x^n_{0,1} - x^n_{1,1} \) (mod \( d \)). Then we can conceive a gedankenexperiment that, if we already know the outputs are \( x^n_{0,1} \) on Alice’s side and \( x^n_{1,1} \) on Bob’s side under \( \mathcal{X}^n_{0,1} \)-basis measurement, we do the POVM \( \mathcal{X}^n|y^n_{q(t)} \) and record the output values. Finally, the output values in the gedankenexperiment help us analyze the correlation of \( \mathcal{X}^n|x^n_{q(t)} \) and \( \mathcal{X}^n|y^n_{q(t)} \) and thus bound \( H^\varepsilon_{\max}(\mathcal{X}^n_{y^n_{q(t)}} | \mathcal{X}^n_{y^n_{q(t)}}) \). In fact, we can use the actual observed values to reconstruct the values that we need in the gedankenexperiment up to a failure probability.

For this purpose, we recall that it is sufficient to consider that the quantum states shared by Alice and Bob before any measurements have the simple form

\[
\rho_{AB} = \sum_{n, n_0,\cdots,n_{k(t)},\cdots,n_{d-1,d-1}} \mu_{n,n_0,\cdots,n_{k(t)},\cdots,n_{d-1,d-1}} \rho_{n_0,\cdots,n_{k(t)},\cdots,n_{d-1,d-1}}.
\]

(A.26)
where
\[ \rho^\pi_{ni_0,\ldots,ni_d} = \pi(\otimes_{j=0}^{d-1} |\Phi_{jk}\rangle \langle \Phi_{jk}|)^{\otimes n_j}. \]  

(A.27)

We note that the pair of qudits \(|\Phi_{t,k} - j mod d\rangle\) outcome "\(t\)" type change under measurements of \(X_{0,1}\)-basis and "\(j\)" type change under measurements of \(X_{1,0}\)-basis, and we define its corresponding expected value
\[ \lambda_{t,k - j mod d} := \frac{1}{n} \sum_{n_{i_0},\ldots,n_{i_d-1}} n_{t,k - j mod d}, \]  

(A.28)

Then, we connect the actual observed values with \(\lambda_{t,k - j mod d}\) that
\[ q^{(t)}_{0,1} = \sum_{k - j mod d} \lambda_{t,k - j mod d} \quad q^{(t)}_{1,0} = \sum_{t} \lambda_{t,k - j mod d}, \]  

(A.29)

equivalently,
\[ \lambda_{t,k - j mod d} = \frac{1}{\lambda_q}(\sum_{s} q^{(s-k)t+j mod d}) + q^{(t)}_{0,1} - 1. \]  

(A.30)

Consequently, we can define the "conditional" values as we need in the gedankenexperiment picture. The values are given by
\[ \xi_{jt} := \frac{\lambda_{t,t^2-j mod d}}{q^{(t)}_{0,1}}, \]  

(A.31)

which account for the expected probability that a pair of qudits outcome "\(j\)" type change in \(X_{1,0}\)-basis under the condition that this pair of qudits outcome "\(t\)" type change in \(X_{0,1}\)-basis.

**Definition 4.** Here, we consider a probability distribution with respect to the joint dit strings \(X_{t,0}^{n_0} \cdots X_{t,t}^{n_t} \cdots X_{t,d-1}^{n_{d-1}} X_{1,0}^{n_0} \cdots X_{1,t}^{n_t} \cdots X_{1,d-1}^{n_{d-1}}\) that is denoted by
\[ \mathbb{R}^{X_{t,0}^{n_0} \cdots X_{t,t}^{n_t} \cdots X_{t,d-1}^{n_{d-1}} X_{1,0}^{n_0} \cdots X_{1,t}^{n_t} \cdots X_{1,d-1}^{n_{d-1}}}. \]  

(A.32)

For each joint dit strings \(x_{t,0}^{n_0} \cdots x_{t,t}^{n_t} \cdots x_{t,d-1}^{n_{d-1}} \cdot x_{1,0}^{n_0} \cdots x_{1,t}^{n_t} \cdots x_{1,d-1}^{n_{d-1}}\), we can obtain a new dit string by doing subtraction bitwise, that is,
\[ \bar{x}_{1,0}^{n_0} \cdots \bar{x}_{1,t}^{n_t} \cdots \bar{x}_{1,d-1}^{n_{d-1}} \cdot \bar{x}_{1,0}^{n_0} \cdots \bar{x}_{1,t}^{n_t} \cdots \bar{x}_{1,d-1}^{n_{d-1}} \quad (mod \, d), \]  

(A.33)

where "\(j\)" dit occurs \(n_t \cdot \xi_{jt}\) times in the part \(x_{1,t}^{n_t} - \bar{x}_{1,t}^{n_t}\). Then, we define a marginal probability denoted by \(\mathbb{R}^{X_{jt}}(\xi_{jt})\) of \(\mathbb{R}^{X_{t,0}^{n_0} \cdots X_{t,t}^{n_t} \cdots X_{t,d-1}^{n_{d-1}} X_{1,0}^{n_0} \cdots X_{1,t}^{n_t} \cdots X_{1,d-1}^{n_{d-1}}(x_{1,0}^{n_0} \cdots x_{1,t}^{n_t} \cdots x_{1,d-1}^{n_{d-1}} \cdot x_{1,0}^{n_0} \cdots x_{1,1}^{n_t} \cdots x_{1,d-1}^{n_{d-1}})}\). If we exclude a small probability denoted by \(1 - (1 - \varepsilon^2)d\) event and only consider its mutually exclusive event that the error rate under all \(X_{1,t}\)-basis measurement is bounded by \(Q + \mu\), then we can find a probability distribution
\[ S_{\mathbb{R}^{X_{jt}}}(\xi_{jt}) := \begin{cases} \frac{\mathbb{R}^{X_{jt}}(\xi_{jt})}{(1 - \varepsilon^2)d}, & q^{(t)}_{1,t} \geq 1 - Q - \mu, \text{ for all } t, \\ 0, & \text{else}, \end{cases} \]  

(A.34)

where \(q^{(t)}_{1,t}\) is the no error term in the probability vector \(q_{1,t} = \{q^{(t)}_{1,t}, \cdots, q^{(t)}_{1,t}, \cdots, q^{(t)}_{1,t}\}\) with respect to each \(X_{1,t}\)-basis. Similarly, we can find that
\[ F(\mathbb{R}, S) = \sum_{\xi_{jt}} \sqrt{\mathbb{R}^{X_{jt}}(\xi_{jt}) \cdot S_{\mathbb{R}^{X_{jt}}}(\xi_{jt})} = \sqrt{(1 - \varepsilon^2)d}. \]  

(A.35)

Then the purified distance between the distributions is given by \(P(\mathbb{R}, S) = \sqrt{1 - F^2(\mathbb{R}, S)} = \sqrt{1 - (1 - \varepsilon^2)d}\).
Lemma 5. We define a probability vector

\[ \xi^{(n)}_\mathbf{0} = \{\xi^{(n)}_0, \xi^{(n)}_1, \ldots, \xi^{(n)}_{d-1}\}, \]

where \(\xi^{(n)}_0 = \frac{d(1 - Q - \mu) + q^{(0)}_0 - 1}{d \cdot q^{(0)}_0} \) and \(\xi^{(n)}_1 = \cdots = \xi^{(n)}_{d-1} = \frac{1 - \xi^{(n)}_0}{d - 1}\). (A.36)

and let \(\bar{c} = \sqrt{1 - (1 - \varepsilon^2)^d}\). Then

\[ H_{\bar{c}}^d(\mathbf{X}_{1,0}^{n_0} \cdots \mathbf{X}_{1,t}^{n_t} \cdots \mathbf{X}_{1,d-1}^{n_{d-1}}) \leq \Omega_1 \cdot \cdots \cdot \cdots \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \·
It remains to prove equation (A.41). We figure that $Ω_0 ≥ N/2 ≥ \text{Max}\{Ω_1, \cdots , Ω_{d-1}\}$, thus

$$Ω_0^{\omega_0-Ω_0} ≥ (\text{Max}\{Ω_1, \cdots , Ω_{d-1}\})^{\sum_{j=1}^{d-1} Ω_j-ω_j} = \prod_{j=1}^{d-1} (\text{Max}\{Ω_1, \cdots , Ω_{d-1}\})^{Ω_j-ω_j} ≥ \prod_{j=1}^{d-1} Ω_j^{Ω_j-ω_j}.$$  \hspace{1cm} (A.43)

Equivalently, $\prod_{j=0}^{d-1} Ω_j^{Ω_j} ≥ \prod_{j=0}^{d-1} Ω_j^{Ω_j}$ which completes the proof of equation (A.41), and consequently we complete the full proof of Lemma 6. Moreover, if $Ω_0/N ≥ ξ^{(u)}_{0|0}$, $H(Ω)$ reaches maximum when $Ω_1/N = \cdots = Ω_{d-1}/N = \frac{1-ξ^{(u)}_{0|0}}{d-1}$, that is, $H(Ω) ≤ H(ξ^{(u)})$. Besides, it is not hard to note that

$$d^{n_t} = (1 + \cdots + 1)^{n_t} = \sum_{ω_0|1, \cdots , ω_{d-1}|n_t} n_t! \omega_0|1 \cdots \omega_{d-1}|n_t.$$ \hspace{1cm} (A.44)

**Lemma 7.** For any normalized density $ρ = \sum_{i=1}^d p_i ρ_i$ with the constraint $\sum_{i=1}^d p_i = 1$, if there exists a unnormalized density $\tilde{ρ} = \sum_{i=1}^d \tilde{p}_i \tilde{ρ}_i$ that satisfy $P(ρ_i, \tilde{ρ}_i) ≤ ε$ and $\text{Max}_i P(ρ_i, \tilde{ρ}_i) ≤ \tilde{ε}$ where $P(\cdot, \cdot)$ denotes purified distance \[37\], then

$$P(ρ, \tilde{ρ}) ≤ \sqrt{1 - (1 - ε^2)(1 - \tilde{ε}^2)}.$$ \hspace{1cm} (A.45)

**Proof.** Because of the constraint $\sum_{i=1}^d p_i = 1$, we find that $ρ_i$ is normalized. According to the definition of the purified distance, we have

$$P(ρ_i, \tilde{ρ}_i) = \sqrt{1 - F^2(ρ_i, \tilde{ρ}_i)} = \sqrt{1 - F^2(ρ_i, \tilde{ρ}_i)} ≤ \tilde{ε},$$ \hspace{1cm} (A.46)

where $F(\cdot, \cdot)$ denotes purified fidelity. Owing to the strong concavity of the fidelity, we find that

$$F(ρ, \tilde{ρ}) ≥ \sum_{i=1}^d \sqrt{p_i \tilde{p}_i} F(ρ_i, \tilde{ρ}_i) \geq \sum_{i=1}^d \sqrt{p_i \tilde{p}_i} \sqrt{1 - \tilde{ε}^2} = \sqrt{1 - \tilde{ε}^2} \sum_{i=1}^d \sqrt{p_i \tilde{p}_i} = \sqrt{1 - \tilde{ε}^2} \sqrt{1 - \tilde{ε}^2}.$$ \hspace{1cm} (A.47)

Thus, we have

$$P(ρ, \tilde{ρ}) = \sqrt{1 - F^2(ρ_i, \tilde{ρ}_i)} ≤ \sqrt{1 - (1 - \tilde{ε}^2)(1 - \tilde{ε}^2)},$$ \hspace{1cm} (A.48)

which completes the proof.

**Lemma 8.** For a normalized density $ρ_{AB} = \sum_{i=1}^d p_i ρ_i$ with the constraint $\sum_{i=1}^d p_i = 1$ and a unnormalized density $\tilde{ρ}_{AB} = \sum_{i=1}^d \tilde{p}_i \tilde{ρ}_i$ that satisfy $P(ρ_i, \tilde{ρ}_i) ≤ ε$ and $H_{\text{min}}(A|B)_{\tilde{ρ}_i} = H_{\text{min}}^\varepsilon(A|B)_{\tilde{ρ}_i}$ for each index $i$, we have

$$2^{-H_{\text{min}}^\varepsilon(A|B)_{\tilde{ρ}_i}} ≤ \sum_i \tilde{p}_i 2^{-H_{\text{min}}(A|B)_{\tilde{ρ}_i}}.$$ \hspace{1cm} (A.49)

**Proof.** $H_{\text{min}}(A|B)_{\tilde{ρ}_i} = H_{\text{min}}^\varepsilon(A|B)_{\tilde{ρ}_i}$ suggest that Max$_i P(ρ_i, \tilde{ρ}_i) ≤ \tilde{ε}$. Combining Lemma 7, we have $P(ρ_{AB}, \tilde{ρ}_{AB}) ≤ \sqrt{1 - (1 - \tilde{ε}^2)(1 - \tilde{ε}^2)}$. Thus, we obtain

$$H_{\text{min}}^\varepsilon(A|B)_{\tilde{ρ}} ≥ H_{\text{min}}(A|B)_{\tilde{ρ}}.$$ \hspace{1cm} (A.50)

Owing to the subadditivity of min-entropy, we have

$$2^{-H_{\text{min}}(A|B)_{\tilde{ρ}}} ≤ \sum_i \tilde{p}_i 2^{-H_{\text{min}}(A|B)_{\tilde{ρ}_i}}.$$ \hspace{1cm} (A.51)
Consequently,
\[ 2^{-H_{\min}^{1-(1-\varepsilon/2)(1-\varepsilon)}}(A|B)_\rho \leq 2^{-H_{\min}(A|B)_\rho} \leq \sum_i p_i 2^{-H_{\min}(A|B)_{\rho_i}} = \sum_i p_i 2^{-H_{\min}^x(A|B)_{\rho_i}}. \] (A.52)

which completes the proof.

Now, all ingredients are ready, we prove our result as follows

**Theorem 9.** For classical-quantum state \( \rho_{X_0^n,Y_0^n}\) where \( \rho_{X_0^n,Y_0^n} = \sum_{y_0^n} \mathbb{P}_{Y_0^n}(y_0^n) \rho_{X_0^n,Y_0^n} \), its smooth min-entropy is bounded by
\[ H_{\min}^{\varepsilon}(X_0^n|E) \geq n \log_2 d - H(\xi)(1 - Q - \mu), \] (A.53)

Proof. For \( \rho_{X_0^n,Y_0^n} = \sum_{y_0^n} \mathbb{P}_{Y_0^n}(y_0^n) \rho_{X_0^n,Y_0^n} \) where \( \rho_{X_0^n,Y_0^n} = \mathbb{P}_{Y_0^n}(y_0^n) = \mathbb{P}_{X_0^n}(x_0^n) \) and \( P(\mathbb{P}, Q) = \varepsilon \), owing to Lemma 8, we have
\[ H_{\min}^{X_0^n|E}(X_0^n|E) \rho_{X_0^n,Y_0^n} \geq -\log_2 \sum_{y_0^n} \mathbb{P}_{Y_0^n}(y_0^n) 2^{-H_{\min}^{X_0^n|E}(X_0^n|E) \rho_{X_0^n,Y_0^n}}. \] (A.54)

According to Lemma 3, for each \( \rho_{X_0^n,Y_0^n} \), its smooth min-entropy satisfies
\[ H_{\min}^{X_0^n|E}(X_0^n|E) \geq n \log_2 d - H_{\max}^{X_0^n|E}(X_0^n|E) \rho_{X_0^n,Y_0^n} \geq 0 \] (A.55)

Combining Lemma 5 and the probability distribution \( \mathbb{Q}_{Y_0^n}(y_0^n) = 0 \) when \( q_0^n 1 \leq 1 - Q - \mu \), we obtain
\[ H_{\max}^{X_0^n|E}(X_0^n|E) \rho_{X_0^n,Y_0^n} \leq -\log_2 \sum_{y_0^n} \mathbb{P}_{Y_0^n}(y_0^n) 2^{-H_{\min}^{X_0^n|E}(X_0^n|E) \rho_{X_0^n,Y_0^n}} \leq n[H(\xi)(1 - Q - \mu)] \geq n[H(\xi)(1 - Q - \mu)] \leq n[H(\xi)(1 - Q - \mu)] \]
(A.56)

where we use the fact that \( H(\xi)(1 - Q - \mu) \leq H_{\max}^{X_0^n|E}(X_0^n|E) \rho_{X_0^n,Y_0^n} \leq d \). Combining equation (A.54), (A.55), (A.56) and the fact that \( \sqrt{1 - (1 - \varepsilon/2)}(1 - \varepsilon/2) = \sqrt{1 - (1 - \varepsilon/2)}(1 - \varepsilon/2) = \varepsilon \), we finally obtain that
\[ H_{\min}^{X_0^n|E}(X_0^n|E) \rho_{X_0^n,Y_0^n} \geq -\log_2 \sum_{y_0^n} \mathbb{P}_{Y_0^n}(y_0^n) 2^{-H_{\min}^{X_0^n|E}(X_0^n|E) \rho_{X_0^n,Y_0^n}} \]
(A.57)

which completes the proof.

**Theorem 10.** The \((d+1)\)-basis protocols \( \Phi[n, m, l, \varepsilon, \text{cor}, \text{leak}_{EC}] \) using \( d \)-level quantum states is \( \varepsilon_{\text{sec}} \)-secret for some \( \varepsilon_{\text{sec}} \) > 0 if \( l \) satisfies
\[ l \leq \max_{\varepsilon, \varepsilon} n(\log_2 d - H(\xi)(1 - Q - \mu(\varepsilon))) - 2d \log_2 \frac{1}{2\varepsilon^2} - \text{leak}_{EC} - 2 \log_2 \frac{1}{\varepsilon_{\text{cor}}}, \]
(A.58)

where
\[ H(\xi) = -\frac{1 - d+1}{1 - Q - \mu(\varepsilon)} \log_2 \frac{1 - d+1}{1 - Q - \mu(\varepsilon)} - (d - 1) \frac{1}{d(d-1)} \log_2 \frac{1}{1 - Q - \mu(\varepsilon)}. \]
(A.59)
and we optimize over $\varepsilon > 0$ and $\varepsilon' > 0$ with constraints

$$\tilde{\varepsilon} + \varepsilon' \leq \varepsilon_{sec}, \quad \tilde{\varepsilon} = \sqrt{1 - (1 - \varepsilon^2)^{d+1}}$$

and $\mu(\varepsilon) := \sqrt{\frac{n + m}{m} \frac{m + 1}{m} \ln \frac{1}{\varepsilon}}.$ \hspace{1cm} (A.60)

**Proof.** Due to the Quantum Leftover Hashing lemma \[3, 36\], it is possible to extract a $\Delta$-secret key of length $l$ from $X_{0,1}^n$, where

$$\Delta = \tilde{\varepsilon} + \frac{1}{2} \sqrt{2^{l - H_{\min}^2(X_{0,1}^n|E')}}.$$ \hspace{1cm} (A.61)

The term $E'$ that represents all information Eve obtained can be decomposed as $E' = CE$, where $C$ is classical information revealed by Alice and Bob during the error correction step. For the revealed information $C$ is at most $\text{leak}_{EC} - \log_2 \frac{2}{\varepsilon_{cor}}$ bits, we use a chain rule for smooth entropies and then obtain

$$H_{\min}^2(X_{0,1}^n|E') \geq H_{\min}^2(X_{0,1}^n|E) - \text{leak}_{EC} - \log_2 \frac{2}{\varepsilon_{cor}}.$$ \hspace{1cm} (A.62)

With the lower bound of smooth min-entropy of $\rho_{X_{0,1}^nE}$, we consequently get

$$\Delta \leq \tilde{\varepsilon} + \frac{1}{2} \sqrt{2^{l - H_{\min}^2(X_{0,1}^n|E')}} \leq \tilde{\varepsilon} + \varepsilon' \leq \varepsilon_{sec}.$$ \hspace{1cm} (A.63)

Thus, these protocols are $\varepsilon_{sec}$-secret.