Quantum mutual information redistribution in a pure tripartite state by a number partitioning algorithm

Muchun Yang, Cheng-Qian Xu, and D. L. Zhou*

Institute of Physics, Beijing National Laboratory for Condensed Matter Physics, Chinese Academy of Sciences, Beijing 100190, China and School of Physical Sciences, University of Chinese Academy of Sciences, Beijing 100049, China

(Received 3 July 2023; accepted 13 October 2023; published 2 November 2023)

Quantum information distribution in a tripartite state plays a fundamental role in quantum information processes. Here we investigate how a bipartite unitary transformation U_{AB} redistributes the quantum mutual information with the third party C in a tripartite pure state $|\psi\rangle_{ABC}$ in a $d_A \times d_B \times d_C$-dimensional Hilbert space. In particular, we focus on finding out the optimal unitary transformation U_{AB}^* that maximizes the quantum mutual entropy between party A and party C, $I(A:C) = S(\rho_A) - S(\rho_B) + S(\rho_C)$. We show that the mutual entropy $I(A:C)$ is upper bounded by $2S(\rho_C)$ derived from the Araki-Lieb inequality. This upper bound can be realized via an optimal unitary transformation for any pure state with the rank r_C of ρ_C satisfying $r_C \leqslant d_A$. For a generic pure state with $r_C > d_A$, the upper bound cannot be realized by any bipartite unitary transformation. To maximize the mutual entropy in the later case, we propose a fast numerical algorithm to produce an approximate optimal unitary transformation, where our optimization is transformed into a modified number partition problem. The validness of our algorithm is confirmed by its comparison with the results from the Adam algorithm for parameterized unitary transformations. Our approximate algorithm thus provides a practical protocol to implement redistribution of quantum mutual information for a tripartite quantum state with high dimensions.

DOI: 10.1103/PhysRevA.108.052402

I. INTRODUCTION

Quantum information distribution in a tripartite state is fundamental in quantum information processes, where the strong subadditivity inequality of von Neumann entropy gives important constrains on these information distribution. Here we investigate how a bipartite unitary transformation changes the quantum information distribution in a tripartite pure state.

To characterize the quantum information distribution, we use quantum mutual information as a basic measure. Mutual information in Shannon's theory is a fundamental quantity of information transmitting capacity [1]. Quantum mutual information is a measure of correlation with the information transmission task in a quantum state [2,3]. In the quantum on-time pad, quantum mutual information is the maximum information that can be securely sent [4,5]. Quantum mutual information also quantifies the minimal amount of noise needed to erase the correlation in a bipartite state [6,7].

In a tripartite pure state, a bipartite unitary transformation cannot change the mutual entropy between the two parties and the third party, but it changes the distribution of quantum mutual entropy among them. Thus our aim is to maximize the mutual entropy of the first party and the third one, which is shown to be equivalent to the maximization of the entropy difference of the first party relative to the second party.

The basic application of our mutual information redistribution is as follows. Assume Alice, Bob, and Charlie share an entangled state $|\psi\rangle_{ABC}$, which is used as a quantum resource for performing securely transfer classical information between Alice (and/or Bob) and Charlie. Alice and Bob can redistribute their capacities by performing an optimized bipartite unitary transformation, although their total capacity keeps invariant.

To solve the above maximization, we develop an approximate numerical algorithm, which can be transformed into a modified number partition problem. The number partitioning problem (NPP) is to partition a group numbers into a fixed number of subsets, such that the sums of each subset are as similar as possible. Finding the exact solution is difficult, which is an NP-hard problem [8,9]. There are lots of approximate algorithms to give approximate solutions [10–17]. Recently, physicists also proposed a quantum algorithm for NPP [18–20]. Our algorithm is based on the maximization of the entropy of the first party before the minimization of the second party, which are related with the concavity [21] and the majorization [22,23] properties of Von-Neumann entropy, respectively. Here the maximization of the entropy of the first party is mapped to a modified number partition problem after a disentanglement unitary transformation. The validness of our algorithm is confirmed by its comparison with the results from the Adam algorithm for parameterized unitary transformations. Our approximate algorithm thus provides a practical protocol to implement the redistribution of quantum mutual information for a tripartite quantum state.

*zhoudl72@iphy.ac.cn

II. MAXIMIZATION PROBLEM OF QUANTUM MUTUAL INFORMATION

For a bipartite quantum state ρ_{AB}, the quantum mutual information between A and B is defined as

$$I_\rho(A:B) \equiv S(\rho_A) + S(\rho_B) - S(\rho_{AB}), \tag{1}$$

where the state function S is the von Neumann entropy: for a quantum state σ, $S(\sigma) \equiv -\mathrm{Tr}[\sigma \ln \sigma]$, and the logarithm is taken on base 2. The states ρ_A and ρ_B are the reduced density matrices of the state ρ_{AB}.

For a pure tripartite state $|\Psi\rangle_{ABC}$, a direct calculation leads to the following equality on quantum mutual information:

$$I_{|\psi\rangle}(AB:C) = I_{|\psi\rangle}(A:C) + I_{|\psi\rangle}(B:C) = 2S(\rho_C), \tag{2}$$

where

$$I_{|\psi\rangle}(A:C) = S(\rho_C) + S(\rho_A) - S(\rho_B), \tag{3}$$

$$I_{|\psi\rangle}(B:C) = S(\rho_C) + S(\rho_B) - S(\rho_A). \tag{4}$$

The states ρ_A, ρ_B, and ρ_C are reduced states of $|\psi\rangle_{ABC}$. Equation (2) implies that, for a pure tripartite state, the quantum mutual information between C and $\{A, B\}$ is completely distributed into the quantum mutual information between C and A and that between C and B.

Our main task can be formulated as follows. Assume that Alice, Bob, and Charlie share a pure tripartite state $|\psi\rangle_{ABC}$, which is the main resource assisting the communications between Charlie and Alice (or Bob). By performing a unitary transformation U_{AB} between Alice and Bob, the quantum mutual information between Charlie and Alice (or Bob) can be adjustable. Our aim is to maximize the quantum mutual information

$$\max_{U_{AB}} I_{U_{AB}|\psi\rangle_{ABC}}(A:C) \equiv I_M. \tag{5}$$

Because U_{AB} is a unitary transformation on AB acting on AB, the mutual information between C and AB is invariant, i.e.,

$$I_{U_{AB}|\psi\rangle_{ABC}}(AB:C) = I_{|\psi\rangle_{ABC}}(AB:C) = 2S(\rho_C). \tag{6}$$

This implies that the unitary transformation that maximizes the mutual information between A and C must minimize the mutual information between B and C:

$$\min_{U_{AB}} I_{U_{AB}|\psi\rangle_{ABC}}(B:C) = 2S(\rho_C) - I_M. \tag{7}$$

In particular, when the dimension of the Hilbert space of party A equals to that of party B, the range of the mutual information between A and C under any unitary transformation U_{AB} is given by

$$2S(\rho_C) - I_M \leqslant I_{U_{AB}|\psi\rangle_{ABC}}(A:C) \leqslant I_M. \tag{8}$$

Following Eq. (3), the maximization in Eq. (5) can be simplified as

$$\max_{U_{AB}} \left[S\left(\rho_A^U\right) - S\left(\rho_B^U\right)\right] = I_M - S(\rho_C), \tag{9}$$

where

$$\rho_{AB}^0 = \mathrm{Tr}_C(|\psi\rangle_{ABC\,ABC}\langle\psi|), \tag{10}$$

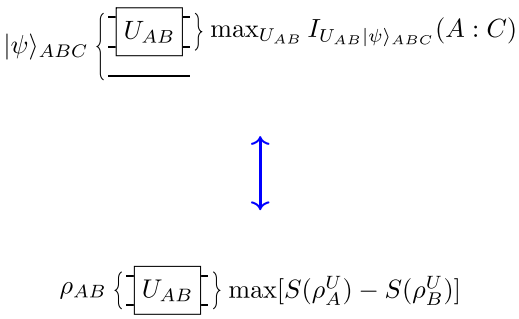

FIG. 1. Equivalent pictures of redistribution of quantum mutual information.

$$\rho_A^U = \mathrm{Tr}_B\left(U_{AB}\rho_{AB}^0 U_{AB}^\dagger\right), \tag{11}$$

$$\rho_B^U = \mathrm{Tr}_A\left(U_{AB}\rho_{AB}^0 U_{AB}^\dagger\right). \tag{12}$$

We observe that ρ_{AB}^0 determines the result of the maximization. The equivalent maximization in Eq. (5) and that in Eq. (9) are demonstrated in Fig. 1.

III. UPPER BOUND OF ENTROPY DIFFERENCE AND ARAKI-LIEB INEQUALITY

Equation (9) makes reminds us of the Araki-Lieb inequality [24,25]

$$|S(\rho_A) - S(\rho_B)| \leqslant S(\rho_{AB}). \tag{13}$$

In our case, $S(\rho_{AB}) = S(\rho_C)$, which is invariant under any unitary transformation U_{AB}. Thus $S(\rho_A^U) - S(\rho_B^U) \leqslant S(\rho_C)$, which implies that the entropy difference ΔS is upper bounded by $S(\rho_C)$. Let us investigate under which condition the upper bound $S(\rho_C)$ for ΔS can be arrived at. Our main result is summarized in the following theorem.

Theorem 1. For a tripartite pure state $|\psi_{ABC}\rangle$ with Hilbert space dimension $d_A \times d_B \times d_C$, there exists an optimal unitary transformation U_{AB}^* such that $S(\rho_A^{U^*}) - S(\rho_B^{U^*}) = S(\rho_C)$ if and only if the rank of ρ^C: $r_C \leqslant d_A$.

Proof. Let us first prove the "only if" part. If the upper bound can be arrived at for a state $|\psi\rangle_{ABC}$, then there exists a unitary transformation U_{AB}^*, such that

$$S\left(\rho_A^{U^*}\right) - S\left(\rho_B^{U^*}\right) = S(\rho_C). \tag{14}$$

Let us introduce the notations

$$|\psi^*\rangle_{ABC} = U_{AB}^*|\psi\rangle_{ABC}, \tag{15}$$

$$\rho_{ABC}^* = |\psi^*\rangle_{ABC\,ABC}\langle\psi^*|. \tag{16}$$

Then the mutual information

$$I_{\rho^*}(A:C) = 2S(\rho_C), \tag{17}$$

$$I_{\rho^*}(B:C) = 0. \tag{18}$$

The later equality implies that

$$\rho_{BC}^* = \rho_B^* \otimes \rho_C = \sum_{m=1}^{r_B}\sum_{n=1}^{r_C} p_m q_n |mn\rangle\langle mn|, \tag{19}$$

where $\{|m\rangle\}$ and $\{|n\rangle\}$ are orthogonal normal bases, $p_m, q_n > 0$, $\sum_m p_m = \sum_n q_n = 1$, r_B, and r_C are the ranks of ρ_B^* and ρ_C, respectively. Thus

$$|\psi^*\rangle_{ABC} = \sum_{m=1}^{r_B}\sum_{n=1}^{r_C}\sqrt{p_m q_n}|\phi_{mn}\rangle \otimes |m\rangle \otimes |n\rangle, \tag{20}$$

where $\{|\phi_{mn}\rangle\}$ is the orthogonal normal basis of the Hilbert space of party A. Equation (20) implies that

$$r_B \leqslant d_B, \tag{21}$$

$$r_B r_C \leqslant d_A. \tag{22}$$

Because $r_B \geqslant 1$, we have

$$r_C \leqslant d_A. \tag{23}$$

Then we prove the "if" part. Our tripartite pure state can be written as

$$|\psi_{ABC}\rangle = \sum_{n=1}^{r_C}\sqrt{q_n}|\chi_n\rangle \otimes |n\rangle, \tag{24}$$

where $q_n > 0$, $\sum_n q_n = 1$, $\{|n\rangle\}$ and $\{|\chi_n\rangle\}$ are orthonormal base of Hilbert space of C and AB, respectively. Then there exist a unitary transformation such that

$$U_{AB}^*|\chi_n\rangle = |\phi_n\rangle \otimes |1\rangle, \tag{25}$$

where $\{|\phi_n\rangle\}$ are orthonormal basis of Hilbert space of A, and $|1\rangle$ is a normalized state of B. Then

$$\rho_{AB}^* = \sum_{n=1}^{r_C} q_n|\phi_n\rangle\langle\phi_n| \otimes |1\rangle\langle 1|, \tag{26}$$

which satisfies $S(\rho_A^*) - S(\rho_B^*) = S(\rho_C)$. This completes the proof of our theorem. ■

Because $r_C \leqslant d_C$, we obtain a direct corollary of the above theorem: if the dimension of Hilbert spaces A and C satisfies $d_A > d_C$, there exists an optimal unitary transformation U^* such that $S(\rho_A^*) - S(\rho_B^*) = S(\rho_C)$.

For simplicity, we focus on the case with $d_A = d_B = d$ in the following. For a general stat ρ^{AB} in a $d \times d$ Hilbert space, the rank of ρ^{AB} lies the range $[1, d^2]$. Only the states with lower rank in the range $[1, d]$ can arrive at the upper bound $S(\rho_{AB})$ from the Araki-Lieb inequality. For the states with the rank in the range $[d+1, d^2]$, the upper bound $S(\rho_{AB})$ cannot arrive at.

When the upper bound of entropy difference can be reached, Eqs. (17) and (18) implies that the optimal unitary transformation completely transforms the correlation between AB and C to that between A and C, without any correlation left between B and C. The upper bound cannot be reached by a unitary transformation originates from the limitation of the dimension of the Hilbert space.

IV. SOLVING MUTUAL INFORMATION MAXIMIZATION BY NUMBER PARTITION ALGORITHM

The eigendecomposition of ρ_{AB}^0 is

$$\rho_{AB}^0 = \sum_{m=0}^{d-1}\sum_{n=0}^{d-1} p_{mn}|\psi_{mn}\rangle\langle\psi_{mn}|. \tag{27}$$

We assume these eigenvalues are in the decreasing order, i.e., $p_{mn} \geqslant p_{m'n'}$ if $md + n \leqslant m'd + n'$. First, we apply the disentanglement unitary transformation

$$D|\psi_{mn}\rangle = |mn\rangle, \tag{28}$$

which makes ρ_{AB}^0 become a separable state

$$\rho_{AB}^D \equiv D\rho_{AB}^0 D^\dagger = \sum_{m=0}^{d-1}\sum_{n=0}^{d-1} p_{mn}|mn\rangle\langle mn|. \tag{29}$$

Then we define a unitary transformation related to a permutation operation s

$$U_s|mn\rangle = |s(mn)\rangle, \tag{30}$$

where s is an element in the permutation group S_{d^2}. Thus

$$\begin{aligned}\rho_{AB}^s \equiv U_s\rho_{AB}^D U_s^\dagger &= \sum_{m=0}^{d-1}\sum_{n=0}^{d-1} p_{mn}|s(mn)\rangle\langle s(mn)| \\ &= \sum_{m=0}^{d-1}\sum_{n=0}^{d-1} p_{s^{-1}(mn)}|mn\rangle\langle mn|,\end{aligned} \tag{31}$$

where s^{-1} is the inverse of s. The reduced states of ρ_{AB}^s are

$$\rho_A^s = \sum_{m=0}^{d-1} p_{Am}^s|m\rangle\langle m|, \tag{32}$$

$$\rho_B^s = \sum_{n=0}^{d-1} p_{Bn}^s|n\rangle\langle n|, \tag{33}$$

with

$$p_{Am}^s = \sum_{n=0}^{d-1} p_{s^{-1}(mn)}, \tag{34}$$

$$p_{Bn}^s = \sum_{m=0}^{d-1} p_{s^{-1}(mn)}. \tag{35}$$

Hence

$$S\left(\rho_A^s\right) - S\left(\rho_B^s\right) = -\sum_{m=0}^{d-1} p_{Am}^s \ln p_{Am}^s + \sum_{n=0}^{d-1} p_{Bn}^s \ln p_{Bn}^s. \tag{36}$$

Thus the maximization over the permutation operation is given by

$$\max_{s\in S_{d^2}}\left[S\left(\rho_A^s\right) - S\left(\rho_B^s\right)\right]. \tag{37}$$

In the following, we aim to show that Eq. (37) is an excellent substituent of Eq. (9) in most cases for our optimization problem.

Before detailed numerical optimization, we explore the symmetry in our problem. Let $r \in S_d$ and $t \in S_d$, and $r \otimes t \in S_d \otimes S_d$, and $S_d \otimes S_d$ is a subgroup of S_{d^2}:

$$r \otimes t(mn) = [r(m)t(n)]. \tag{38}$$

TABLE I. Number partition of $\{p_{mn}\}$ with column sums and row sums for $d = 3$.

	p^s_{B0}	p^s_{B1}	p^s_{B2}
p^s_{A0}	$p_{s^{-1}(00)}$	$p_{s^{-1}(01)}$	$p_{s^{-1}(02)}$
p^s_{A1}	$p_{s^{-1}(10)}$	$p_{s^{-1}(11)}$	$p_{s^{-1}(12)}$
p^s_{A2}	$p_{s^{-1}(20)}$	$p_{s^{-1}(21)}$	$p_{s^{-1}(22)}$

Thus we can prove that

$$S\left(\rho_A^{r\otimes t\cdot s}\right) = S\left(\rho_A^s\right), \tag{39}$$

$$S\left(\rho_B^{r\otimes t\cdot s}\right) = S\left(\rho_B^s\right). \tag{40}$$

The proof of Eq. (39) is given as follows:

$$\begin{aligned} S\left(\rho_A^{r\otimes t\cdot s}\right) &= -\sum_{m=0}^{d-1}\sum_{n=0}^{d-1} p_{(r\otimes ts)^{-1}(mn)} \ln \sum_{n=0}^{d-1} p_{(r\otimes ts)^{-1}(mn)} \\ &= -\sum_{m=0}^{d-1}\sum_{n=0}^{d-1} p_{s^{-1}(r\otimes t)^{-1}(mn)} \ln \sum_{n=0}^{d-1} p_{s^{-1}(r\otimes t)^{-1}(mn)} \\ &= -\sum_{m=0}^{d-1}\sum_{n=0}^{d-1} p_{s^{-1}[r^{-1}(m)t^{-1}(n)]} \ln \sum_{n=0}^{d-1} p_{s^{-1}[r^{-1}(m)t^{-1}(n)]} \\ &= -\sum_{m'=0}^{d-1}\sum_{n'=0}^{d-1} p_{s^{-1}(m'n')} \ln \sum_{n'=0}^{d-1} p_{s^{-1}(m'n')} \\ &= S(\rho_A^s). \end{aligned} \tag{41}$$

Similarly, we can prove Eq. (40). Thus for any $r, t \in S_d$ and $s \in S_{d^2}$,

$$S\left(\rho_A^{r\otimes ts}\right) - S\left(\rho_B^{r\otimes ts}\right) = S\left(\rho_A^s\right) - S\left(\rho_B^s\right), \tag{42}$$

which implies that every element s in one right coset of the subgroup $S_d \otimes S_d$ will give the same value of $S(\rho_A^s) - S(\rho_B^s)$. In other words, the maximization in Eq. (37) is taken over the set of all the right cosets of the subgroup $S_d \otimes S_d$, any one element in each coset

$$\max_{s\in S_{d^2}/S_d\otimes S_d} \left[S\left(\rho_A^s\right) - S\left(\rho_B^s\right)\right]. \tag{43}$$

In particular, we realize that the maximization problem in Eq. (37) or in Eq. (56) is a type of NPP: d^2 numbers $\{p_{mn}\}$ are partitioned into a $d \times d$ lattice, every site with one element. Every partition corresponds to a permutation element. Our maximization is taken over all the ways of partitions. The case of $d = 3$ is demonstrated in Table I. To maximize $S(\rho_A^s) - S(\rho_B^s)$, we need to choose a permutation s such that the numbers in $\{p^s_{Am}\}$ are as similar as possible and the numbers in $\{p^s_{Bn}\}$ are as different as possible.

Let us consider the number of the permutations in the set S_{d^2} or in the set of $S_{d^2}/S_d \otimes S_d$, which is $(d^2)!$ or $(d^2)!/(d!d!)$. With the increasing of d, there numbers become extremely large, e.g., $(4^2)! \simeq 2.1 \times 10^{13}$ and $(4^2)!/(4!4!) \simeq 3.6 \times 10^{10}$, which prevents the numerical optimization directly by the exhaustive attack method.

A. $d = 2$ case

Let us start with the case of $d = 2$, where both A and B are one qubit. Let us take a computational basis of $\mathcal{H}_A$ as $\{|m\rangle, 0 \leqslant a \leqslant 1\}$ [the eigenvectors of Z_A with eigenvalues $(-1)^m$], and a basis of $\mathcal{H}_B$ as $\{|n\rangle, 0 \leqslant n \leqslant 1\}$ [the eigenvectors of Z_B with eigenvalues $(-1)^n$]. Then for any permutation s we construct a unitary transformation

$$U_s|mn\rangle = |s(mn)\rangle. \tag{44}$$

The number of all the unitary transformations related with permutations equals to the order of the permutation group S_4, i.e., $4! = 24$. We can show that

$$U_s Z_A U_s^\dagger = (-1)^{c_s^A} Z_A^{a_s^A} Z_B^{b_s^A}, \tag{45}$$

$$U_s Z_B U_s^\dagger = (-1)^{c_s^B} Z_A^{a_s^B} Z_B^{b_s^B}. \tag{46}$$

Note that the unitary transformations related with the subgroup $S_2 \otimes S_2$ is

$$X_A^a \otimes X_B^b, \quad a, b \in \{0, 1\}, \tag{47}$$

where X is the x component of the Pauli operator defined by

$$X|m\rangle = |1 - m\rangle, \tag{48}$$

or equivalently defined by

$$XZX^\dagger = -Z. \tag{49}$$

Then Eqs. (45) and (46) become

$$X_A^a X_B^b Z_A X_B^b X_A^a = (-1)^a Z_A, \tag{50}$$

$$X_A^a X_B^b Z_B X_B^b X_A^a = (-1)^b Z_B. \tag{51}$$

Thus the representative element in the unitary transformations corresponding to the right cosets of the subgroup $S_2 \otimes S_2$ are given by

$$U_s Z_A U_s^\dagger = Z_A^{a_s^A} Z_B^{b_s^A}, \tag{52}$$

$$U_s Z_B U_s^\dagger = Z_A^{a_s^B} Z_B^{b_s^B}, \tag{53}$$

which means the unitary transformations of $\{Z_A, Z_B\}$ can take the values of any two ordered elements in $\{Z_A, Z_B, Z_AZ_B\}$. This implies there are $P_3^2 = 6$ representative unitary transformations in the cosets. If we denote $s(mn) = (m'n')$, then these unitary transformations can be obtained by solving the following equations:

$$s^{-1}\begin{pmatrix} m' \\ n' \end{pmatrix} = \begin{pmatrix} m \\ n \end{pmatrix} \mod 2, \tag{54}$$

with

$$s^{-1} = \begin{pmatrix} a_s^A & b_s^A \\ a_s^B & b_s^B \end{pmatrix}. \tag{55}$$

The above equations have a unique solution if and only if

$$\det s^{-1} = a_s^A b_s^B - a_s^B b_s^A \neq 0. \tag{56}$$

TABLE II. Optimal number partition of $\{p_{mn}\}$ with column sums and row sums for $d = 2$, which corresponds to Eq. (66).

	$p_{B0}^{s_5}$	$p_{B1}^{s_5}$
$p_{A0}^{s_5}$	p_{00}	p_{11}
$p_{A1}^{s_5}$	p_{01}	p_{10}

Then all the permutations in $S_4/S_2 \otimes S_2$ are given by

$$s_1{}^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \tag{57}$$

$$s_2{}^{-1} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \tag{58}$$

$$s_3{}^{-1} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}, \tag{59}$$

$$s_4{}^{-1} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \tag{60}$$

$$s_5{}^{-1} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}, \tag{61}$$

$$s_6{}^{-1} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}. \tag{62}$$

Their inverse permutations are given by

$$s_i = s_i{}^{-1}, \quad 1 \leqslant i \leqslant 4, \tag{63}$$

$$s_5 = s_6{}^{-1}, \tag{64}$$

$$s_6 = s_5^{-1}. \tag{65}$$

A direct calculation gives

$$\underset{s \in S_4/S_2 \otimes S_2}{\operatorname{argmax}} \left[S\left(\rho_A^s\right) - S\left(\rho_B^s\right) \right] = s_5, \tag{66}$$

where

$$s_5^{-1} \begin{pmatrix} m \\ n \end{pmatrix} = \begin{pmatrix} n \\ m+n \mod 2 \end{pmatrix}. \tag{67}$$

Then the state

$$\begin{aligned} \rho_{AB}^{s_5} &= \sum_{m=0}^{1} \sum_{n=0}^{1} p_{s_5^{-1}(mn)} |mn\rangle\langle mn| \\ &= p_{00}|00\rangle\langle 00| + p_{11}|01\rangle\langle 01| \\ &\quad + p_{01}|10\rangle\langle 10| + p_{10}|11\rangle\langle 11|. \end{aligned} \tag{68}$$

The reduced states of $\rho_{AB}^{s_5}$ are

$$\rho_A^{s_5} = (p_{00} + p_{11})|0\rangle_A\langle 0| + (p_{01} + p_{10})|1\rangle_A\langle 1|, \tag{69}$$

$$\rho_B^{s_5} = (p_{00} + p_{01})|0\rangle_B\langle 0| + (p_{11} + p_{10})|1\rangle_B\langle 1|. \tag{70}$$

Hence the maximum value

$$\begin{aligned} S\left(\rho_A^{s_5}\right) - S\left(\rho_B^{s_5}\right) = &- (p_{00} + p_{11})\ln(\mathrm{p}_{00} + \mathrm{p}_{11}) \\ &- (p_{01} + p_{10})\ln(\mathrm{p}_{01} + \mathrm{p}_{10}) \\ &+ (p_{00} + p_{01})\ln(\mathrm{p}_{00} + \mathrm{p}_{01}) \\ &+ (p_{10} + p_{11})\ln(\mathrm{p}_{10} + \mathrm{p}_{11}). \end{aligned} \tag{71}$$

This permutation is demonstrated in Table II.

Theorem 2. For a two-qubit state ρ_{AB}, the unitary transformation $U^* = U_{s_5} D$ makes $S(\rho_A^U) - S(\rho_B^U)$ take a local maximum.

Proof. Let $U = WU^*$. Then

$$\begin{aligned} \Delta S(U) &\equiv S\left(\rho_A^U\right) - S\left(\rho_B^U\right) \\ &= S\left(\mathrm{Tr}_B\left[W \rho_{AB}^{s_5} W^\dagger\right]\right) - S\left(\mathrm{Tr}_A\left[W \rho_{AB}^{s_5} W^\dagger\right]\right), \end{aligned} \tag{72}$$

where the unitary transformation W can be parameterized as

$$\begin{aligned} W = &W_{00}(h_{00}) \prod_{j=1}^{3} W_{j0}(h_{j0}) \\ &\times \prod_{k=1}^{3} W_{0j}(h_{0j}) \prod_{m,n=1}^{3} W_{mn}(h_{mn}), \end{aligned} \tag{73}$$

with

$$W_{mn}(h_{mn}) = \exp(i h_{mn} \sigma_m \otimes \sigma_n). \tag{74}$$

Let

$$W' = \prod_{m,n=1}^{\prime 3} W_{mn}(h_{mn}), \tag{75}$$

and $U' = W'U^*$. Then we can prove

$$\Delta S(U) = \Delta S(U') = \Delta S(\{h_{mn}\}'), \tag{76}$$

where $'$ implies that it does not contain the term with $m = n = 3$.

Then we need to show that

$$\left.\frac{\partial \Delta S}{\partial h_{mn}}\right|_{\{h_{mn}=0\}} = 0, \tag{77}$$

$$H(\Delta S)|_{\{h_{mn}=0\}} \leqslant 0, \tag{78}$$

with the Hessian matrix

$$H(\Delta S)_{mn;m'n'} = \frac{\partial \Delta S}{\partial h_{mn}} h_{m'n'}. \tag{79}$$

Equations (77) and (78) show that U^* is a local maximum of $\Delta S(U)$ whose proofs are given in Appendix A. ■

There remains a question that, what is the global maximal value of entropy difference $\Delta S(U)$? Fortunately, we have not found any other unitary transformation that get larger $\Delta S(U)$. We can use the Adam optimizing algorithm [26], which is a gradient descending (GD) method, to optimize function $\Delta S(U)$. Figures 2(a) and 2(b) show that the numerical results using Adam optimization are actually the same as that using U^*, where the relative error is in the order of 10^{-8} numerically. Here the relative error is defined by $(\Delta S_{\mathrm{Adam}} - \Delta S_{\mathrm{permutation}})/\Delta S_{\mathrm{Adam}}$. But we failed to prove that U^* is the global optimal unitary transformation.

B. $d = 3$ case

In the $d = 3$ case, the optimal permutation unitary transformation that maximize the entropy difference ΔS cannot be given explicitly in a table similar as Table II in the $d = 2$ case. We observe that for different $\{p_{mn}\}$, the optimal permutation is different. This arises from a competition between increasing S_A and decreasing S_B in the case of $d = 3$, while the optimal

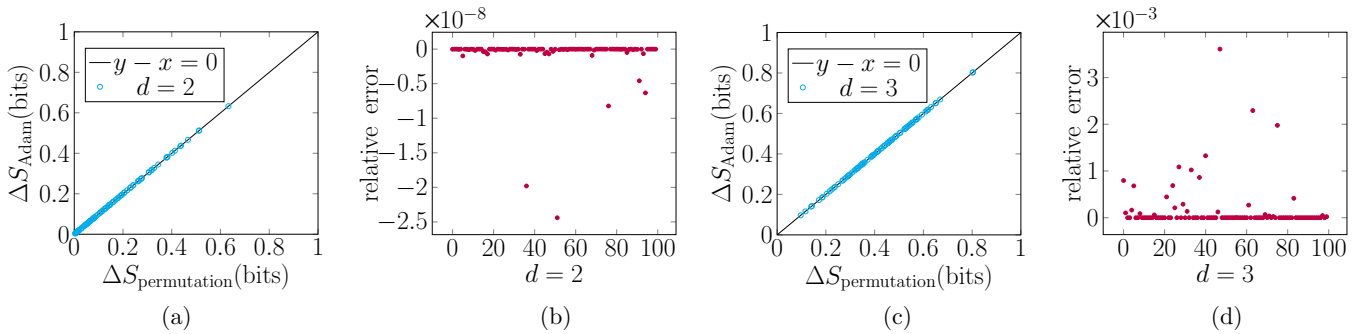

FIG. 2. Numerical results of entropy difference based on optimal permutation and Adam optimization in the two-qubit case and two-qutrit case. (a,c) The maximized entropy difference obtained by permutation $\Delta S_{\text{permutation}}$ via that obtained by Adam optimization ΔS_{Adam}, where 100 states are randomly generated in each case. (b,d) Relative error of $\Delta S_{\text{permutation}}$ with respect to ΔS_{Adam}.

permutation maximizes S_A and minimize S_B simultaneously in the $d = 2$ case. The numerical results on optimal permutations in the $d = 3$ case implies that we should give priority to maximize S_A, which is related with the number partition problem. In fact, all optimal permutations we found are the optimal solutions from the number partitioning problem.

In the Adam algorithm for the $d = 3$ case, we take the Gellman matrices as the generators of the unitary transformations whose derivatives are given in detail in Appendix B. We numerically checked that the optimal permutations are all local maximal values.

Numerical results on the comparison of optimized entropy differences between the optimal permutation and the Adam algorithm are shown in Figs. 2(c) and 2(d). The relative error can arrive at the number of 10^{-3} order, which clearly demonstrates that the optimal permutation is, in general, not a global maximum of the entropy difference ΔS. On the other hand, the entropy difference from the optimal permutation is a relatively excellent approximation of the global maximum of ΔS.

C. $d \geqslant 4$ cases

For the systems with $d \geqslant 4$, as discussed in the paragraph before Sec. IV A, it is difficult to find optimal permutation directly. Motivated by the experience in the case of $d = 3$, we take two steps to find the optimal permutation: first, find one permutation that maximizes $S(\rho_A)$; second, minimize $S(\rho_B)$ while keeping $S(\rho_A)$ invariant. The convenient way to describe the above two steps is by visualizing permutations in Table I, the first step is to make the row sums as equal as possible, which is the same aim as the number partition problem. In the second step, to keep $S(\rho_A)$ invariant, we keep the row invariant for every number. To decrease the entropy of $S(\rho_B)$, we only need to arrange the numbers in every row in the decreasing order. In the following, we will develop an approximate algorithm to find one optimized permutation that maximizes $S(\rho_A)$ in the first step.

Before we present our algorithm, we first review the greedy number partitioning (GNP) algorithm for number partitioning [8] on which our algorithm is constructed. In the number partition problem, we aim to partition n numbers into k set such that the sums of every set are as equal as possible. The GNP algorithm can be stated as follows. First, sort the numbers in the descending order and place the largest k numbers into the k set. Then process the remaining numbers sequentially, put the next number to a set that the sum of the set is currently smallest. We write the GNP in the form of pseudo code in Algorithm 1. This method is also called longest-processing-time-first scheduling. It has an approximation ratio that in the worst case, the largest sum in the greedy partition is at most $\frac{4k-1}{3k}$ times the optimal largest sum [11–14]. In addition, the minimum sum is at least $\frac{3k-1}{4k-2}$ times the optimal smallest sum [15–17].

Algorithm 1. Greedy Number Partitioning (GNP)

Input: n numbers and the number of partitions k
Output: k sets, with an indefinite number of numbers in each set
sort n numbers in descending order;
create k empty sets;
for *num* in the first k largest numbers **do**
  append *num* to one empty set;
end
for *num* in the remained $n - k$ numbers **do**
  sum each set, and append *num* to the set that has the smallest sum;
end
for each set **do**
  sort set in descending order;
end
return k sets

Note that, in the GNP algorithm, the size of every set is not required to be the same. For our problem, however, we need to partition d^2 numbers, $\{p_{mn}\}$ into d rows, every row has exactly d numbers. The relative number partition problem can be stated as follows. To partition $k_A k_B$ numbers into k_A sets with each set having k_B numbers, such that the sum of numbers in every set as equal as possible. In our algorithm, we recurrently call function GNP and we name it the recurrent greedy number partitioning (RGNP). The procedure of the RGNP is given as follows. In the first partitioning iteration, by applying GNP to $k_A k_B$ numbers, we get k_0 sets having k_B numbers, k_0^+ sets having more than k_B numbers, k_0^- sets having less than k_B numbers. We cut all the smaller numbers from

any set having more than k_B numbers until the set remains k_B numbers. Keep the sets having exact k_B numbers remained. Then we mix all the cut numbers with sets that have less than k_B numbers, and using GNP to partition them into k_0^- sets. Then we get k_1 sets having k_B numbers, with k_1^+ sets having more than k_B numbers, k_1^- sets having less than k_B numbers. Then repeat the procedure above until the jth iteration, $k_j^+ = 0$, the iteration is finished. We write RGNP in the form of the pseudocode in Algorithm 2. In Appendix C, we give an example for the RGNP algorithm.

Algorithm 2. Recurrent greedy number partitioning (RGNP)

Input: $k_A k_B$ eigenvalues of initial state
Output: k_A sets with each set containing k_B numbers in descending order
$partition_{done}$ = empty set;
$set_{todo} = k_A k_B$ eigenvalues of initial state;
$k,\ k_+ = 0$;
$k_- = k_B$;
$i = 0$;
while $k_+ \neq 0$ **do**
  i **to** $i+1$;
  $partition^i = \mathrm{GNP}(set_{todo}, k_-)$;
  set_{todo} = empty set;
  $k,\ k_+,\ k_- = 0$;
  for set in $partition^i$ **do**
    if the number of elements in $set \leq k_B$ **then**
      append numbers in set into set_{todo};
      $k_- \leftarrow k_- + 1$;
    end
    if the number of elements in $set \geq k_B$ **then**
      remain the first largest k_B numbers, cut other numbers;
      append cut numbers into set_{todo}, append the set containing remained numbers into $partition_{done}$;
      $k_+ \leftarrow k_+ + 1$;
    end
    if the number of elements in $set = k_B$ **then**
      append set into $partition_{done}$;
      $k \leftarrow k + 1$;
    end
  end
end
return $partition_{done}$

The time complexity of greedy number partitioning is $O(d^2 \ln d^2)$ [8], where d is the dimension of one subsystem. As for the Gradient Descending algorithm, there is not a general expression $O(\cdot)$ of time complexity. However, we know that we need to optimize $d^2 d^2 - 1$, which is $O(d^2 d^2)$ independent parameters using GD, which is much slower than greedy number partitioning.

We use this method to maximize the entropy difference for $d = 3, 4, 5, 6$, and 8 cases. We generate 100 random states for each case, and compare with Adam algorithm. Here we set the GD algorithm convergence condition is $\Delta S < 1 \times 10^{-8}$, which we find this condition is sufficient to converge. The results are shown in Fig. 3. The relative error is shown in Table. III. We find that the maximized entropy differences agrees well in all the cases, and the average relative error becomes smaller with the increasing of d. In particular, for a larger system, e.g., the systems with $d = 8$, the maximized entropy difference from RGNP is larger than those from Adam algorithm. In other words, when the system has a larger dimension, our RGNP is better than the Adam in most cases, which may be attributed to too many local maximums for a larger system such that the Adam finds only one of local maximums in most cases. The numerical results show that our RGNP is an excellent approximate algorithm to present a protocol to maximize the entropy difference by an analytical unitary transformation.

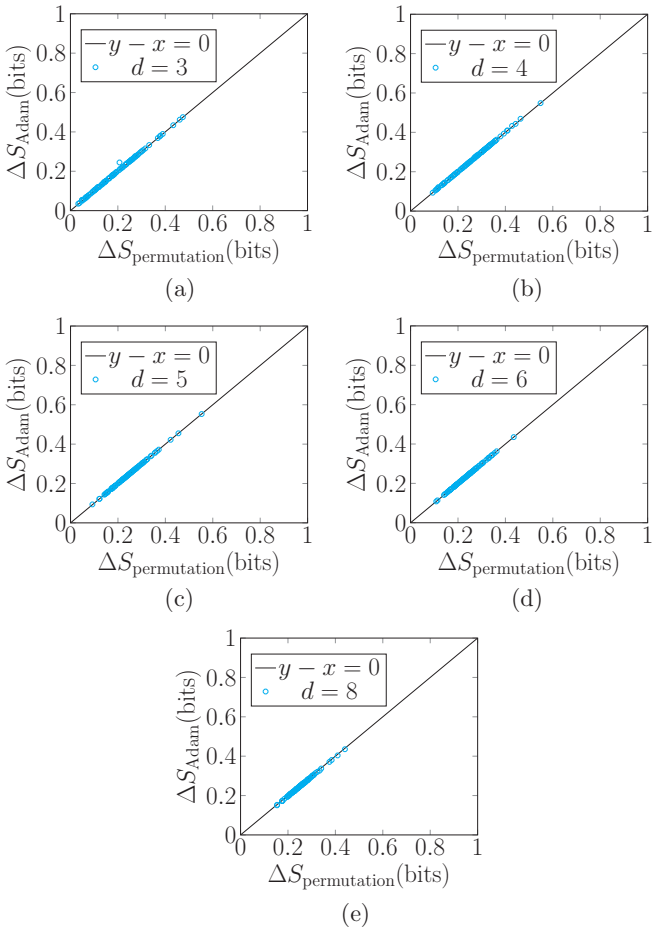

FIG. 3. Numerical results of entropy difference based on greedy permutation and Adam optimization for $d = 3, 4, 5, 6$, and 8 cases corresponding to (a)–(e). The maximized entropy difference obtained by permutation $\Delta S_{\mathrm{permutation}}$ via that obtained by Adam optimization ΔS_{Adam}, where 100 states are randomly generated in each case.

V. SUMMARY AND DISCUSSION

We introduce a mutual entropy redistribution protocol via a bipartite unitary transformation in a tripartite pure state. We show that if the dimension of Hilbert space of Alice is not less than that of Charlie, then the mutual entropy between Charlie and the other two can be completely redistributed into that

between Charlie and Alice. Otherwise, the maximization of the mutual entropy between Alice and Charlie via a bipartite unitary transformation becomes complex, especially when the Hilbert space dimension of Alice (and Bob) is large.

Furthermore we develop an approximate algorithm for the above complex problem, which is based on the greedy number partitioning algorithm and combined with applications of the two basic properties of von Neumann entropy, majorization, and concavity. Our numerical experiments show that, in a small system, this algorithm gives a nearly same results as the optimal result using gradient descending algorithm. In a large system, our algorithm gives better results with a faster speed. In practice, this algorithm helps us get the unitary for redistributing mutual information with a relative small time complexity.

However, we do not understand completely why our approximate RGNP algorithm is so successful for our maximization problem. In the RGNP algorithm, we first disentangle the initial state ρ_{AB}^0, then we perform an optimal permutation related unitary transformation to complete the maximization, where the final optimal state keeps disentangled. In our RGNP algorithm we disentangle A and B, which is due to the monogamy of entanglement [27], i.e., the entanglement between A and B decreases some degrees to maximize the mutual entropy between A and C.

In Fig. 4 we demonstrate entanglement monogamy in redistributing the mutual information. We find that in Fig. 4(a), $r_C = 2 \leqslant d_A = 2$, by theorem 1, I_{AC} can be optimized to $2S(\rho_C)$, and I_{BC} can be optimized to 0. We see that N_{AB}, N_{BC}, and CKW three-tangle are all 0 after the optimization. Here I_{AC} is maximized to up bound, thus A and C cannot share entanglement with B, the entanglement between B and other party is 0. And the three-way entanglement CKW three-tangle is also 0.

In Fig. 4(b), N_{AB} is 0 after the optimization, but $N_{BC} > 0$. Here $r_C = 4 > d_A = 2$, by theorem 1, I_{AC} cannot be optimized to $2S(\rho_C)$, and it always holds that $I_{BC} > 0$. Thus B and C can have entanglement. To increase the entanglement between A and C, the entanglement between A and B must be 0 since the entanglement between A and B will prevent us from increasing the entanglement between A and C. On the other hand, it is the entanglement between B and C that decreases some degrees to maximize the mutual entropy between A and C. This is consistent with the entanglement monogamy.

In Fig. 4(c), $N_{AB} > 0$ after the optimization, but $N_{BC} = 0$. Here $r_C = 2 \leqslant d_A = 4$, by theorem 1, I_{AC} can be optimized to $2S(\rho_C)$. The reason for $N_{AB} > 0$ is that the Hilbert dimension of A is sufficiently large. After maximizing I_{AC}, A still has capacity to entangle with B. This is also consistent with the entanglement monogamy. The in-depth theoretical analysis of our RGNP algorithm is left as an open problem to be investigated in future.

Our results can be intuitively summarized as follows. If the dimension of the Hilbert space of A is sufficient large, we can always find a unitary transformation that transforms all the entanglement between two parties (A, B) and C into the entanglement between A and C, and in the same time B is disentangled with C. It is worth pointing out that the entanglement between A and B can exist after optimization in this case since our aim is to maximize $I(A, C)$ [or equivalently to minimize $I(B, C)$], which does not make any direct constraint on $I(A, B)$.

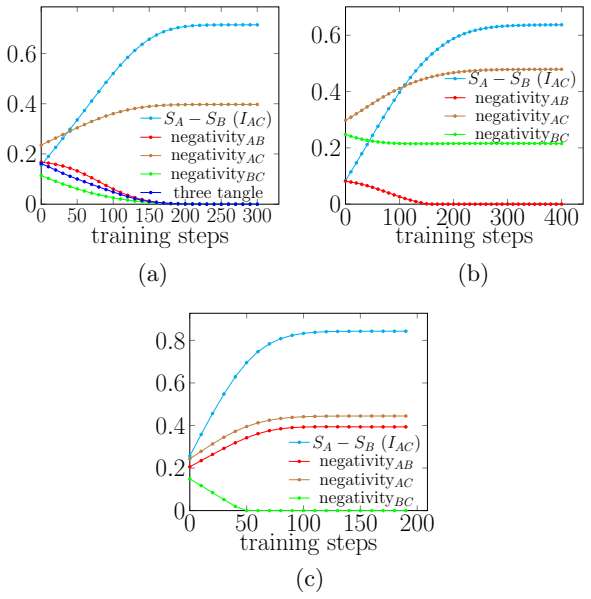

FIG. 4. (a) Numerical results of I_{AC} (in graph we use $S_A - S_B$ to represent I_{AC}, note that $I_{AC} = S_A - S_B + S_C$), CKW three-tangle, negativity between A and B (N_{AB}), negativity between A and C (N_{AC}) and negativity between B and C (N_{BC}) during the Adam optimization for a random $2 \otimes 2 \otimes 2$ pure state ρ_{ABC}, where $r_C = 2 \leqslant d_A = 2$. (b) Numerical results of I_{AC}, N_{AB}, N_{AC}, and N_{BC} during the Adam optimization for a random $2 \otimes 2 \otimes 4$ pure state ρ_{ABC}, where $r_C = 4 > d_A = 2$. (c) Numerical results of I_{AC}, N_{AB}, N_{AC}, and N_{BC} during the Adam optimization for a random $4 \otimes 2 \otimes 2$ pure state ρ_{ABC}, where $r_C = 2 \leqslant d_A = 4$. Note that the three-tangle is not computed in (b) and (c) because numerically computable three-tangle is only defined for three-qubit states [27]. The units of above quantities are bits.

ACKNOWLEDGMENTS

This work is supported by National Key Research and Development Program of China (Grants No. 2021YFA0718302 and No. 2021YFA1402104), National Natural Science Foundation of China (Grant No. 12075310), and the Strategic Priority Research Program of Chinese Academy of Sciences (Grant No. XDB28000000).

APPENDIX A: PROOF OF THEOREM 2

In this Appendix, we aim to prove Theorem 2 in the main text. Specifically, we are going to prove Eqs. (77) and (78) by evaluating second derivatives of the function with respect to parameters h_{mn}

$$\begin{aligned}\Delta S(U) &\equiv S\left(\rho_A^U\right) - S\left(\rho_B^U\right) \\ &= S\left(\mathrm{Tr}_B\left[W\rho_{AB}^{s_5}W^\dagger\right]\right) - S\left(\mathrm{Tr}_A\left[W\rho_{AB}^{s_5}W^\dagger\right]\right),\end{aligned} \tag{A1}$$

where S is von Neumann entropy, $U = WU^*$ is the unitary transformation, U^* is the optimal transformation, and unitary transformation W can be parameterized as

$$W = W_{00}(h_{00})\prod_{j=1}^{3}W_{j0}(h_{j0})\prod_{k=1}^{3}W_{0j}(h_{0j})\prod_{m,n=1}^{3}W_{mn}(h_{mn}), \tag{A2}$$

with

$$W_{mn}(h_{mn}) = \exp(ih_{mn}\sigma_m \otimes \sigma_n). \tag{A3}$$

Proof. We can take $\frac{\partial\Delta S}{\partial h_{31}}|_{h_{31}=0}$ and $\frac{\partial^2\Delta S}{\partial h_{31}^2}|_{h_{31}=0}$, for example,

$$\begin{aligned}W_{31}(h_{31}) &= \exp(ih_{31}\sigma_3 \otimes \sigma_1) \\ &= \begin{pmatrix}\cos h_{31} & i\sin h_{31} & 0 & 0\\ i\sin h_{31} & \cos h_{31} & 0 & 0\\ 0 & 0 & \cos h_{31} & -i\sin h_{31}\\ 0 & 0 & -i\sin h_{31} & \cos h_{31}\end{pmatrix}.\end{aligned} \tag{A4}$$

The density matrix $\rho_{AB}^{W_{31}V^*}$ is

$$\rho_{AB}^{W_{31}V^*}(h_{31}) = W_{31}\times\rho_{AB}^{V^*}\times W_{31}^\dagger, \tag{A5}$$

and the corresponding reduced density matrix $\rho_A^{W_{31}V^*}$ and $\rho_B^{W_{31}V^*}$ are

$$\begin{aligned}\rho_A^{W_{31}V^*}(h_{31}) &= \begin{pmatrix}p_1+p_4 & 0\\ 0 & p_2+p_3\end{pmatrix},\\ \rho_B^{W_{31}V^*}(h_{31}) &= \begin{pmatrix}\frac{1}{2}+\left(p_1+p_2-\frac{1}{2}\right)\cos(2h_{31}) & -i(-1+2p_1+2p_3)\cos(h_{31})\sin(h_{31})\\ i(-1+2p_1+2p_3)\cos(h_{31})\sin(h_{31}) & \frac{1}{2}-\left(p_1+p_2-\frac{1}{2}\right)\cos(2h_{31})\end{pmatrix}.\end{aligned} \tag{A6}$$

Then we evaluate $\Delta S[W_{31}(h_{31})U^*] = S[\rho_A^{W_{31}V^*}(h_{31})] - S(\rho_B^{W_{31}V^*}(h_{31}))$. Finally, we can obtain

$$\begin{aligned}\left.\frac{\partial\Delta S}{\partial h_{31}}\right|_{h_{31}=0} &= \lim_{h_{31}\to 0}\left[\frac{\Delta S[W_{31}(h_{31})U^*] - \Delta S[W_{31}(0)U^*]}{h_{31}}\right],\\ \left.\frac{\partial^2\Delta S}{\partial h_{31}^2}\right|_{h_{31}=0} &= \lim_{h_{31}\to 0}\left[\frac{\frac{\partial\Delta S}{\partial h_{31}}(h_{31}) - \frac{\partial\Delta S}{\partial h_{31}}(0)}{h_{31}}\right].\end{aligned} \tag{A7}$$

We evaluate these expressions above by MATHEMATICA 12.0 and we find that all first-order derivatives are equal to 0 at $W = 0$ (these expressions are not shown below because they are too long) while the diagonal second-order derivatives are less than or equal to 0:

$$\frac{\partial^2\Delta S}{\partial h_{mn}^2} = 0,\ mn = 00, 01, 02, 03, 10, 20, 30, 33, \tag{A8}$$

$$\frac{\partial^2\Delta S}{\partial h_{11}^2} = \frac{\partial^2\Delta S}{\partial h_{12}^2} = \frac{\partial^2\Delta S}{\partial h_{21}^2} = \frac{\partial^2\Delta S}{\partial h_{22}^2} = \frac{2}{\ln 2}\left[(p_1+p_2-p_3-p_4)\ln\left(\frac{p_3+p_4}{p_1+p_2}\right) + (p_1+p_4-p_2-p_3)\ln\left(\frac{p_1+p_4}{p_2+p_3}\right)\right] \leqslant 0, \tag{A9}$$

$$\frac{\partial^2\Delta S}{\partial h_{13}^2} = \frac{\partial^2\Delta S}{\partial h_{23}^2} = 8(p_1-p_2)(p_3-p_4)\frac{\ln(1-p_x)-\ln(1+p_x)}{p_x\ln 2} \leqslant 0, \tag{A10}$$

$$\frac{\partial^2 \Delta S}{\partial h_{31}^2} = \frac{\partial^2 \Delta S}{\partial h_{32}^2} = 8(p_2 - p_3)(p_1 - p_4)\frac{\ln(1-p_y) - \ln(1+p_y)}{p_y \ln 2} \leqslant 0, \tag{A11}$$

where

$$\begin{aligned} p_x &= |-1+2p_2+2p_3|, \\ p_y &= |-1+2p_1+2p_2|. \end{aligned} \tag{A12}$$

The sign of Eq. (A9) is determined by

$$\begin{aligned} &(p_1+p_2-p_3-p_4)\ln\left(\frac{p_3+p_4}{p_1+p_2}\right) + (p_1+p_4-p_2-p_3)\ln(\frac{p_1+p_4}{p_2+p_3}) \\ &\leqslant (p_1+p_2-p_3-p_4)\ln\left(\frac{p_3+p_4}{p_1+p_2}\right) + (p_1+p_2-p_3-p_4)\ln\left(\frac{p_1+p_2}{p_3+p_4}\right) \\ &= 0. \end{aligned} \tag{A13}$$

Assuming $p_1 \neq p_2 \neq p_3 \neq p_4$, all the diagonal second derivatives in Eqs. (A9), (A10), and (A11) are less than 0.

As for those off-diagonal derivatives terms, we can repeat the procedure above by setting

$$W_{mn,m'n'} = \exp(ih_{mn}\sigma_m \otimes \sigma_n + ih_{m'n'}\sigma_{m'} \otimes \sigma_{n'}). \tag{A14}$$

By evaluating using MATHEMATICA, we find all off-diagonal terms $\frac{\partial^2 \Delta S}{\partial h_{mn} h_{m'n'}}$ are equal to 0. So the Hessian matrix $H(\Delta S)_{mn;m'n'}$ is negative-definite and $\Delta S(U^*)$ is at its local maximal point.

During the computation we find that parameters in Eq. (A8) have no effect on the reduced density matrix, i.e., for m, n in Eq. (A8), $W_{mn}\rho_A(V)W_{mn}^\dagger = \rho_A(V)$ and $W_{mn}\rho_B(V)W_{mn}^\dagger = \rho_B(V)$. So we can remove such parameters from the Hessian matrix and only maintain those "useful" parameters that have nonzero second derivatives. ■

APPENDIX B: DERIVATIVES OF TWO-QUTRIT SYSTEM

In this Appendix, we aim to evaluate the derivatives of entropy difference function of two-qutrit system with respect to parameters of parameterized unitary transformation. In $d = 3$ cases we cannot give a theorem like theorem 2 because there is no consistent result for different $d = 3$ states. Let $d = 3$ state after being applied the disentanglement unitary transformation and permutation unitary transformation be

$$\rho_{\text{qutrit}} = s_1|00\rangle\langle 00| + s_2|01\rangle\langle 01| + s_3|02\rangle\langle 02| + \ldots + s_9|22\rangle\langle 22|, \tag{B1}$$

where s_i are eigenvalues after permutation. There are in total three kinds of forms for two-qutrit states ρ_{qutrit}. The first form is, taking $\frac{\partial^2 \Delta S}{\partial h_{12}^2}$, for example,

$$\frac{\partial^2 \Delta S}{\partial h_{12}^2} = \frac{2}{\ln 2}\left[(s_1+s_2-s_4-s_5)\ln\left(\frac{s_1+s_2+s_3}{s_4+s_5+s_6}\right) + (-s_1+s_2-s_4+s_5)\ln\left(\frac{s_1+s_4+s_7}{s_2+s_5+s_8}\right)\right]. \tag{B2}$$

The second form is, taking $\frac{\partial^2 \Delta S}{\partial h_{13}^2}$, for example,

$$\frac{\partial^2 \Delta S}{\partial h_{13}^2} = \text{Part1} \times \text{Part2}, \tag{B3}$$

where

$$\begin{aligned} \text{Part1} &= s_3s_4 + s_3s_5 - 4s_4s_5 - (s_4+s_5)s_6 + s_2(-s_3+4s_4+s_6) + s_1(-4s_2-s_3+4s_5+s_6) \\ &= 4(s_1-s_4)(s_2-s_5) + (s_6-s_3)(s_4+s_5) + (s_1+s_2)(s_3-s_6), \end{aligned} \tag{B4}$$

$$\text{Part2} = -\frac{2}{\ln 2}\, g(s_1+s_2+s_3,\ s_4+s_5+s_6) > 0, \tag{B5}$$

where

$$g(x,y) = \frac{\ln(x+y-|x-y|) - \ln(x+y+|x-y|)}{|x-y|}.$$

The third form is, taking $\frac{\partial^2 \Delta S}{\partial h_{18}^2}$, for example,

$$\frac{\partial^2 \Delta S}{\partial h_{18}^2} = \text{Part1} \times \text{Part2}, \tag{B6}$$

where

$$\text{Part1} = (s_1 + s_2 - s_4 - s_5)(s_3 - s_6) < 0, \tag{B7}$$

$$\text{Part 2} = -\frac{6}{\ln 2}\, g(s_1 + s_2 + s_3,\ s_4 + s_5 + s_6) > 0. \tag{B8}$$

Note that we cannot determine the sign of Eqs. (B2) and (B4) since different $\{s_i\}$ will result in different signs. However, by plugging the eigenvalues of the ρ_{qutrit} applied by optimal permutations into these derivative expressions, we find that these optimal permutations actually make $S(\rho_A) - S(\rho_B)$ take a local maximum.

APPENDIX C: EXAMPLES OF RECURRENT GREEDY NUMBER PARTITIONING

In this Appendix, we give an example for the recurrent greedy number partitioning algorithm. For a $d = 6$ state, the eigenvalues are

$p = \{0.184233, 0.172701, 0.167875, 0.130484, 0.007168, 0.006866, 0.005525, 0.00415, 0.002577, 0.002313, 0.101274, 0.009832, 0.008887, 0.008416, 0.007561, 0.006997, 0.006116, 0.004571, 0.003275, 0.000357, 0.000128, 0.043433, 0.011384, 0.011262, 0.010695, 0.010573, 0.010166, 0.010124, 0.009745, 0.008469, 0.008223, 0.007007, 0.006151, 0.005061, 0.003894, 0.002506\}$.

We first partition p into six sets

$$\begin{aligned}
s_1 &= \{0.184233\}, \\
s_2 &= \{0.172701\}, \\
s_3 &= \{0.167875\}, \\
s_4 &= \{0.130484, 0.007168, 0.006866, 0.005525, 0.00415, 0.002577, |0.002313\}, \\
s_5 &= \{0.101274, 0.009832, 0.008887, 0.008416, 0.007561, 0.006997, |0.006116, 0.004571, 0.003275, 0.000357, 0.000128\}, \\
s_6 &= \{0.043433, 0.011384, 0.011262, 0.010695, 0.010573, 0.010166, | \\
&\quad 0.010124, 0.009745, 0.008469, 0.008223, 0.007007, 0.006151, 0.005061, 0.003894, 0.002506\}.
\end{aligned} \tag{C1}$$

So we get three sets having more than six numbers. We denote the number of sets having six numbers in the ith iteration is k_i, the number of sets having more than six numbers is k_i^+, the number of sets having less than six numbers is k_i^-. We cut numbers after the "|" and put these cut numbers with s_1, s_2, and s_3, and get a new set p^1 to partition, where $k_1^- = 3$:

$p^1 = \{0.184233, 0.172701, 0.167875, 0.002313, 0.006116, 0.004571, 0.003275, 0.000357, 0.000128, 0.010124, 0.009745, 0.008469, 0.008223, 0.007007, 0.006151, 0.005061, 0.003894, 0.002506\}$.

Then we partition p^1 into three sets:

$$\begin{aligned}
s_1 &= \{0.184233, 0.007007, 0.005061, 0.003894, 0.000128\}, \\
s_2 &= \{0.172701, 0.009745, 0.008223, 0.006116, 0.003275, 0.000357\}, \\
s_3 &= \{0.167875, 0.010124, 0.008469, 0.006151, 0.004571, 0.002506, 0.002313\}.
\end{aligned} \tag{C2}$$

We find that s_2 has six numbers, s_1 has five numbers, and s_3 has seven numbers, so we move the last number in s_3 to s_1 and finish the recurrent greedy number partitioning. The final result is

$$\begin{aligned}
s_1 &= \{0.184233, 0.007007, 0.005061, 0.003894, 0.000128, 0.002313\}, \\
s_2 &= \{0.172701, 0.009745, 0.008223, 0.006116, 0.003275, 0.000357\}, \\
s_3 &= \{0.167875, 0.010124, 0.008469, 0.006151, 0.004571, 0.002506\}, \\
s_4 &= \{0.130484, 0.007168, 0.006866, 0.005525, 0.00415, 0.002577\}, \\
s_5 &= \{0.101274, 0.009832, 0.008887, 0.008416, 0.007561, 0.006997\}, \\
s_6 &= \{0.043433, 0.011384, 0.011262, 0.010695, 0.010573, 0.010166\}.
\end{aligned} \tag{C3}$$

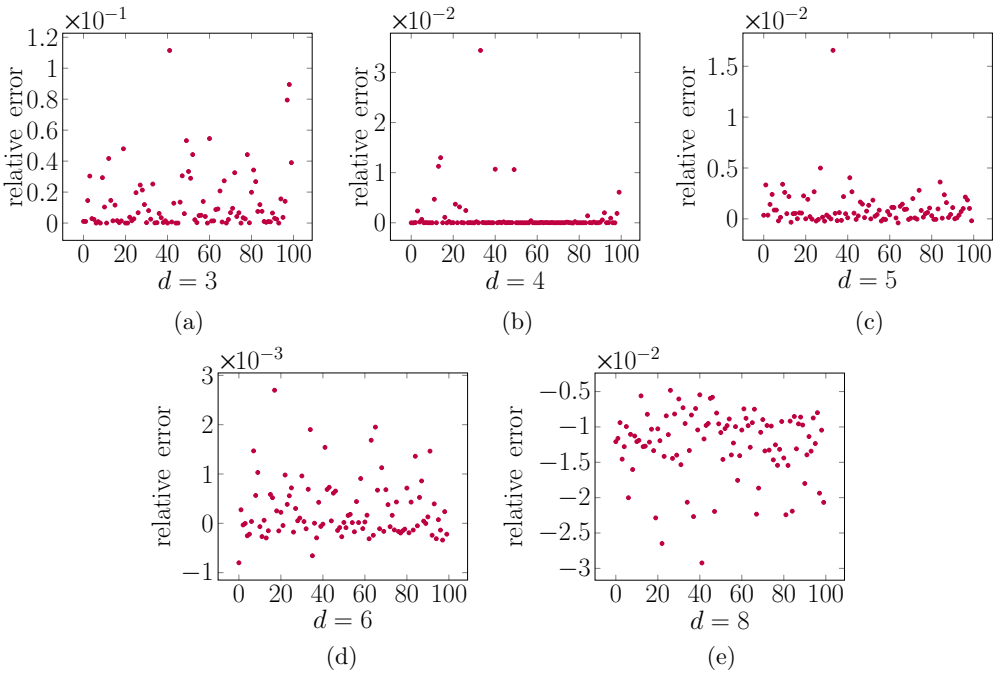

FIG. 5. Relative errors of the RGNP algorithm for $d = 3, 4, 5, 6$, and 8 cases.

APPENDIX D: RELATIVE ERROR OF RGNP ALGORITHM

In this Appendix, we plot the relative errors of the RGNP algorithm. Figure 5 gives the relative errors of the RGNP algorithm for $d = 3, 4, 5, 6$, and 8 cases.

Table III writes the average of relative errors of 100 states for each case in Fig. 5.

TABLE III. Relative error of greedy permutation for $d = 3, 4, 5, 6$, and 8 cases. Relative error is defined by $(\Delta S_{\text{Adam}} - \Delta S_{\text{permutation}})/\Delta S_{\text{Adam}}$. Minus relative error means RGNP algorithm is better than Adam optimization algorithm.

$d = 3$	$d = 4$	$d = 5$	$d = 6$	$d = 8$
1.1378%	0.2503%	0.1017%	0.0266%	−1.2422%

[1] C. E. Shannon, Bell Syst. Tech. J. **27**, 379 (1948).

[2] C. H. Bennett, P. W. Shor, J. A. Smolin, and A. V. Thapliyal, Phys. Rev. Lett. **83**, 3081 (1999).

[3] D.-L. Z. Bei Zeng, X. Chen, and X.-G. Wen, *Quantum Information Meets Quantum Matter* (Springer, New York, 2019).

[4] B. Schumacher and M. D. Westmoreland, Phys. Rev. A **74**, 042305 (2006).

[5] C.-Q. Xu and D. L. Zhou, Phys. Rev. A **104**, 022610 (2021).

[6] B. Groisman, S. Popescu, and A. Winter, Phys. Rev. A **72**, 032317 (2005).

[7] M. Berta, F. G. S. L. Brandão, C. Majenz, and M. M. Wilde, Phys. Rev. Lett. **121**, 040504 (2018).

[8] R. E. Korf, in *Proceedings of the 14th International Joint Conference on Artificial Intelligence - Volume 1, IJCAI'95, Montreal, Quebec, Canada* (Morgan Kaufmann Publishers Inc., Burlington, MA, 1995), pp. 266–272.

[9] E. Schreiber, R. Korf, and M. Moffitt, J. ACM. **65**, 1 (2018).

[10] I. P. Gent and T. Walsh, Comput. Intell **14**, 430 (1998).

[11] R. L. Graham, SIAM J. Appl. Math. **17**, 416 (1969).

[12] X. Xiao, in *Proceedings of the 5th International Conference on Frontiers of Manufacturing Science and Measuring Technology (FMSMT 2017)* (Atlantis, Amsterdam, 2017), pp. 486–489.

[13] E. G. Coffman and R. Sethi, in *Proceedings of the ACM SIGMETRICS Conference on Computer Performance Modeling Measurement and Evaluation, SIGMETRICS '76* (Association for Computing Machinery, New York, 1976), pp. 306–310.

[14] B. Chen, Oper. Res. Lett. **14**, 139 (1993).

[15] B. L. Deuermeyer, D. K. Friesen, and M. A. Langston, SIAM J. Discrete Math. **3**, 190 (1982).

[16] J. Csirik, H. Kellerer, and G. Woeginger, Oper. Res. Lett. **11**, 281 (1992).

[17] B. Y. Wu, Theor. Comput. Sci. **349**, 407 (2005).

[18] T. Graß, D. Raventós, B. Juliá-Díaz, C. Gogolin, and M. Lewenstein, Nat. Commun. **7**, 11524 (2016).
[19] G. Anikeeva, O. Marković, V. Borish, J. A. Hines, S. V. Rajagopal, E. S. Cooper, A. Periwal, A. Safavi-Naeini, E. J. Davis, and M. Schleier-Smith, PRX Quantum **2**, 020319 (2021).
[20] N. A. Sinitsyn and B. Yan, Phys. Rev. A **108**, 022412 (2023).
[21] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information: 10th Anniversary Edition* (Cambridge University Press, Cambridge, England, 2010).
[22] Y. Li and P. Busch, J. Math. Anal. Appl. **408**, 384 (2013).
[23] Z. An, C. Cao, C.-Q. Xu, and D. L. Zhou, arXiv:2107.03542.
[24] H. Araki and E. H. Lieb, Commun. Math. Phys. **18**, 160 (1970).
[25] E. A. Carlen and E. H. Lieb, Lett. Math. Phys. **101**, 1 (2012).
[26] D. P. Kingma and J. Ba, arXiv:1412.6980.
[27] V. Coffman, J. Kundu, and W. K. Wootters, Phys. Rev. A **61**, 052306 (2000).