Time-energy measure for quantum processes

Chi-Hang Fred Fung* and H. F. Chau
Department of Physics and Center of Theoretical and Computational Physics, University of Hong Kong, Pokfulam Road, Hong Kong
(Received 24 May 2013; published 9 July 2013)

Quantum mechanics sets limits on how fast quantum processes can run given some system energy through time-energy uncertainty relations, and they imply that time and energy are tradeoffs against each other. Thus, we propose to measure the time energy as a single unit for quantum channels. We consider a time-energy measure for quantum channels and compute lower and upper bounds of it using the channel Kraus operators. For a special class of channels (which includes the depolarizing channel), we can obtain the exact value of the time-energy measure. One consequence of our result is that erasing quantum information requires $\sqrt{(n+1)/n}$ times more time-energy resource than erasing classical information, where n is the system dimension.

DOI: 10.1103/PhysRevA.88.012307 PACS number(s): 03.67.Lx, 89.70.Eg

I. INTRODUCTION

Evolution of quantum processes (including performing quantum computation) requires physical resources, in particular, time and energy. The computation speed of a physical device is governed by physical laws and is limited by the energy of the device. Under the constraints of quantum mechanics, system evolutions are bounded by time-energy uncertainty relations (TEURs) [1]. The investigation of TEURs has a long history. The first major result of a TEUR was proved by Mandelstam and Tamm [2]. This was followed by subsequent work on isolated systems [3–10] and composite systems with entanglement [11–13]. Recently TEURs for general quantum processes have also been proved [14,15]. The general form of TEURs is an inequality that sets a lower limit on the product of the system energy (or a function of the energies) and the time it takes to evolve an initial state to a final state (e.g., an orthogonal state). Motivated by the TEURs and recognizing that time and energy are tradeoffs against each other, time energy can be regarded as a single property of a quantum process. The intuition is that the more computation or work a quantum process performs, the more time energy it requires. And it is up to the system designer (or nature) to perform it with more time but less energy, or vice versa. Thus, our goal in this paper is to investigate the time-energy requirements of quantum processes by using a time-energy measure. Chau [16] proposed a time-energy measure for unitary transformations that is based on a TEUR proved earlier [10]. In this paper we extend this measure to quantum processes. The TEUR due to Chau [10] is tight in the sense that it can be saturated by some states and Hamiltonians, and thus it serves to motivate a good definition for a time-energy measure. To see this, let us start with this TEUR. Given a time-independent Hamiltonian H of a system, the time t needed to to evolve a state $|\Phi\rangle$ under the action of H to a state whose fidelity [17] is less than or equal to ϵ satisfies the TEUR

$$t \geqslant \frac{(1-\sqrt{\epsilon})\hbar}{A\sum_j |\alpha_j|^2 |E_j|}, \tag{1}$$

where E_j are the eigenvalues of H with the corresponding normalized energy eigenvectors $|E_j\rangle$, $|\Phi\rangle = \sum_j \alpha_j |E_j\rangle$, and $A \approx 0.725$ is a universal constant. Essentially, after time t, the state transforms unitarily according to $U = e^{-iHt/\hbar}$. The same U could be implemented with either a high-energy H run for a shorter time or a low-energy H run for a longer time. Based on Eq. (1), a weighted sum of $|tE_j|$ serves as an indicator of the time-energy resource needed to perform U, and as such the following time-energy measure on unitary matrices was proposed by Chau [16]:

$$\|U\|_{\vec{\mu}} = \sum_{j=1}^{r} \mu_j |\theta_j|^{\downarrow},$$

where U has eigenvalues $\exp(-iE_jt/\hbar) \equiv \exp(\theta_j)$ and $\vec{\mu}$ is some fixed vector to be described later (see Sec. II). In essence, a large value of $\|U\|_{\vec{\mu}}$ suggests that a long time may be needed to run a Hamiltonian that implements U for a fixed energy, and vice versa.

In this paper we are interested in an analogous measure for quantum channels which include unitary transformations as special cases. We are given a quantum channel $\mathcal{F}(\rho)$ acting on system A that maps $n \times n$ density matrix ρ to another one with the same dimension. There exist unitary extensions U_{BA} in a larger Hilbert space with an ancillary system B such that $\mathcal{F}(\rho) = \mathrm{Tr}_B[U_{BA}(|0\rangle_B\langle 0| \otimes \rho_A)U_{BA}^\dagger]$. Each U_{BA} could have a different time-energy spectrum, and we want to select the one requiring the least resource for $\mathcal{F}$. We extend the resource indicator for U to quantum channel $\mathcal{F}$ by defining

$$\|\mathcal{F}\|_{\vec{\mu}} \equiv \min_U \|U\|_{\vec{\mu}} \quad \text{s.t. } \mathcal{F}(\rho) = \mathrm{Tr}_B[U_{BA}(|0\rangle_B\langle 0| \otimes \rho_A)U_{BA}^\dagger]\ \forall \rho.$$

This gives a U that consumes the least time-energy resource. Thus, $\|\mathcal{F}\|_{\vec{\mu}}$ is an indicator of the resource needed to perform $\mathcal{F}$. We formally formulate this problem in Sec. II into the "partial U problem" and the "channel problem." Then we simplify the "partial U problem" in Sec. III and solve the a special case of it in Sec. IV. The special case solution will be used to prove our major results, which are the upper bound of the time-energy resource measure $\|\mathcal{F}\|_{\vec{\mu}}$ (Sec. V), the lower bound of $\|\mathcal{F}\|_{\vec{\mu}}$ (Sec. VI), and the optimal $\|\mathcal{F}\|_{\vec{\mu}}$ for a class of quantum channels (Sec. VII). The lower and upper bounds of

*chffung@hku.hk

the time energy $\|\mathcal{F}\|_{\vec{\mu}}$ hold for any quantum channel $\mathcal{F}$ and for specific $\vec{\mu}$:

$$\|\mathcal{F}\|_{\max} \geqslant \min_{\mathbf{v}:\,\|\mathbf{v}\|\leqslant 1} \max_{1\leqslant i\leqslant n} \cos^{-1}\left\{\mathrm{Re}\left[\lambda_i\left(\sum_{j=1}^{d} v_j F_j\right)\right]\right\}, \tag{2}$$

$$\|\mathcal{F}\|_{\max} \leqslant \min_{\mathbf{v}:\,\|\mathbf{v}\|\leqslant 1} \sum_{i=1}^{n} \cos^{-1}\left\{\mathrm{Re}\left[\lambda_i\left(\sum_{j=1}^{d} v_j F_j\right)\right]\right\}, \tag{3}$$

$$\|\mathcal{F}\|_{\mathrm{sum}} \geqslant \min_{\mathbf{v}:\,\|\mathbf{v}\|=1} \max_{1\leqslant i\leqslant n} 2\cos^{-1}\left|\lambda_i\left(\sum_{j=1}^{d} v_j F_j\right)\right|, \tag{4}$$

$$\|\mathcal{F}\|_{\mathrm{sum}} \leqslant \min_{\mathbf{v}:\,\|\mathbf{v}\|\leqslant 1} \sum_{i=1}^{n} 2\cos^{-1}\left\{\mathrm{Re}\left[\lambda_i\left(\sum_{j=1}^{d} v_j F_j\right)\right]\right\}, \tag{5}$$

where $\|\mathbf{v}\| = \sqrt{\sum_{j=1}^{d} |v_j|^2}$, $F_j \in \mathbb{C}^{n\times n}, j = 1,\ldots,d$ are the Kraus operators of $\mathcal{F}$, and $\lambda_i(\cdot)$ denotes the ith eigenvalue of its argument. Here $\|\cdot\|_{\max}$ is a shorthand notation for $\|\cdot\|_{\vec{\mu}}$ with $\vec{\mu} = [1,0,0,\ldots]$ and $\|\cdot\|_{\mathrm{sum}}$ for $\|\cdot\|_{\vec{\mu}}$ with $\vec{\mu} = [1,1,1,\ldots]$. For a class of channels (which includes the depolarizing channel), we obtain the exact value for $\|\mathcal{F}\|_{\max}$ in Sec. VII. In particular, when $\mathcal{F}$ is a depolarizing channel with probability q that the input state is unchanged, its time-energy requirement is $\|\mathcal{F}\|_{\max} = \cos^{-1}\sqrt{q+(1-q)/n^2}$. Finally, in Sec. VIII, we study the time-energy resource needed to erase information in both the quantum and classical settings. We conclude that $\sqrt{(n+1)/n}$ times more resource is required in the quantum setting than in the classical setting and that the amount of time-energy resource needed for k runs of the depolarizing channel scales as $\sqrt{k}$ when the noise is small.

II. PROBLEM FORMULATION

A. Notations and assumptions

We can describe $\mathcal{F}(\rho)$ using the Kraus operators: $\mathcal{F}(\rho) = \sum_{j=1}^{d} F_j \rho F_j^\dagger$ where $F_j \in \mathbb{C}^{n\times n}$ are the Kraus operators satisfying the trace-preserving condition $\sum_{j=1}^{d} F_j^\dagger F_j = I$. Note that any channel can be described by at most n^2 Kraus operators. But here the formulation is general for any number of Kraus operators.

Denote by $\mathrm{U}(r)$ the group of $r \times r$ unitary matrices.

Decompose $U \in \mathrm{U}(r)$ into eigenvectors:

$$U = \sum_{j=1}^{r} \exp(-i\theta_j)|u_j\rangle\langle u_j|, \tag{6}$$

where $\theta_j = E_j t/\hbar$, E_j is the energy, and t is the evolution time. We call θ_j *eigenangles*. We assume that all angles are taken in the range $(-\pi,\pi]$. Define a time-energy measure for U [16]

$$\|U\|_{\vec{\mu}} = \sum_{j=1}^{r} \mu_j |\theta_j|^{\downarrow}, \tag{7}$$

where $|\theta_j|^{\downarrow}$ denotes $|\theta_j|$ ordered nonincreasingly $|\theta_1|^{\downarrow} \geqslant |\theta_2|^{\downarrow} \geqslant \cdots \geqslant |\theta_r|^{\downarrow}$. Also, $\vec{\mu} = [\mu_1,\mu_2,\ldots,\mu_r] \neq \vec{0}$ with $\mu_1 \geqslant \mu_2 \geqslant \cdots \geqslant \mu_r \geqslant 0$. Note that $\|U\|_{\vec{\mu}}$ satisfies the multiplicative triangle inequality $\|UV\|_{\vec{\mu}} \leqslant \|U\|_{\vec{\mu}} + \|V\|_{\vec{\mu}}$ [16].

We have two special cases for the time-energy measure:

$$\text{Sum time energy:} \quad \|U\|_{\mathrm{sum}} \equiv \sum_{j=1}^{r} |\theta_j|, \tag{8}$$

$$\text{Max time energy:} \quad \|U\|_{\max} \equiv \max_{1\leqslant j\leqslant r} |\theta_j| = |\theta_1|^{\downarrow}. \tag{9}$$

Note that the subscript "sum" is short for $\vec{\mu} = [1,1,\ldots,1]$ and "max" for $\vec{\mu} = [1,0,\ldots,0]$.

Define $\|\mathbf{v}\| = \sqrt{\sum_{j=1}^{d} |v_j|^2}$ where $\mathbf{v} = [v_1,v_2,\ldots,v_d]$. We adopt the convention that $\cos^{-1}$ always returns an angle in the range $[0,\pi]$.

B. The "partial U problem" and the "channel problem"

We generalize the measure $\|\cdot\|_{\vec{\mu}}$ to the case where part of a unitary matrix is given. Suppose that we are given the first $n \leqslant r$ columns of U denoted by $U_{[1,n]} \in \mathbb{C}^{r\times n}$. Define the "partial U problem" for $U_{[1,n]}$:

$$\|U_{[1,n]}\|_{\vec{\mu}} \equiv \min_{V} \|V\|_{\vec{\mu}} \quad \text{s.t.} \quad V_{[1,n]} = U_{[1,n]} \quad \text{and} \quad V \in \mathrm{U}(r). \tag{10}$$

We use this as a bridge to generalize the measure to quantum channel $\mathcal{F}(\rho_A)$ acting on density matrix $\rho_A \in \mathbb{C}^{n\times n}$. For such a channel $\mathcal{F}(\rho_A) = \sum_{i=1}^{d} F_i \rho_A F_i^\dagger$ where $F_i \in \mathbb{C}^{n\times n}$, there exist unitary extensions U_{BA} in a larger Hilbert space with an ancillary system B such that $\mathcal{F}(\rho_A) = \mathrm{Tr}_B[U_{BA}(|0\rangle_B\langle 0| \otimes \rho_A)U_{BA}^\dagger]$.

Define a mapping from a sequence of Kraus operators $F_{1:d} \triangleq (F_1,F_2,\ldots,F_d)$ to an $dn \times n$ matrix as follows:

$$g(F_{1:d}) \triangleq \begin{bmatrix} F_1 \\ F_2 \\ \vdots \\ F_d \end{bmatrix}. \tag{11}$$

Because $\sum_{j=1}^{d} F_j^\dagger F_j = I$, the columns of $g(F_{1:d})$ are orthonormal and $g(F_{1:d})$ can be regarded as a submatrix of a unitary one. Thus, we can obtain $\|g(F_{1:d})\|_{\vec{\mu}}$ from problem (10).

Note that two sets of Kraus operators $\{F_1,\ldots,F_d\}$ and $\{F_1',\ldots,F_d'\}$ represent the same quantum channel if and only if $F_i' = \sum_{j=1}^{d} w_{ij} F_j$ for all i and for some unitary matrix $[w_{ij}]$ (see Ref. [18]). If more Kraus operators are desired in one set, we can supplement the other set with all-zero Kraus operators. This implies that given one Kraus representation $\{F_1,\ldots,F_d\}$, the most general form of all unitary extension implementing

$\mathcal{F}$ is

$$U_{BA} = (W_B \otimes I_A) \underbrace{\begin{bmatrix} F_1 & * & * & \cdots & * \\ F_2 & * & * & \cdots & * \\ \vdots & & & & \vdots \\ F_d & * & * & \cdots & * \\ \mathbf{0} & * & * & \cdots & * \\ \vdots & & & & \vdots \\ \mathbf{0} & * & * & \cdots & * \end{bmatrix}}_{\tilde{U}_{BA}}, \tag{12}$$

where W_B is any unitary of dimension $d' \times d'$, I_A is the identity matrix of dimension $n \times n$, $\tilde{U}_{BA}$ is any unitary of dimension $nd' \times nd'$ with the first n columns fixed as shown, and $\mathbf{0}$ is the all-zero matrix of dimension $n \times n$. Here we allow d' to be in the range $d \leqslant d' < \infty$.

We define the time-energy measure of the quantum channel $\mathcal{F}$ given a Kraus representation $\{F_1, \ldots, F_d\}$ as follows:

$$\|\mathcal{F}\|_{\vec{\mu}} \equiv \min_{W_B, d'} \|(W_B \otimes I_A)\tilde{U}_{BA}(1:n)\|_{\vec{\mu}} \quad \text{s.t. } W_B \in \mathrm{U}(d'), \quad \text{and} \quad d \leqslant d' < \infty, \tag{13}$$

where we make use of Eq. (10) in the objective function and $\tilde{U}_{BA}(1:n)$ are the first n columns of $\tilde{U}_{BA}$ given in Eq. (12). We call this the "channel problem."

Our ultimate goal is to find the minimum time-energy required to implement a quantum channel by solving the "channel problem" (13). We first consider the "partial U problem" (10).

III. SIMPLIFICATION OF THE "PARTIAL U PROBLEM"

The "partial U problem" (10) can be recast as that given U of dimension $r \times r$ of the form

$$U = [|b_1\rangle \ |b_2\rangle \ \ldots \ |b_n\rangle \ * \ * \ \cdots \ *], \tag{14}$$

where the first n columns, labeled as $|b_i\rangle, i = 1, \ldots, n$, are fixed, and our goal is to find the remaining $r - n$ columns to minimize $\|U\|_{\vec{\mu}}$ while maintaining U unitary.

It is helpful to consider U as a mapping with the requirement that it performs the following transformations:

$$|e_i\rangle \longrightarrow |b_i\rangle \quad \text{for all} \quad i = 1, \ldots, n, \tag{15}$$

where $|e_i\rangle$ is the unit vector with 1 at the ith entry and 0 everywhere else. Then the "partial U problem" (10) for $U_{[1,n]}$ is equivalent to

$$\|U_{[1,n]}\|_{\vec{\mu}} = \min_{U} \|U\|_{\vec{\mu}} \quad \text{s.t. } U|e_i\rangle = |b_i\rangle \quad \text{for all} \quad i = 1, \ldots, n, \quad \text{with} \quad U \in \mathrm{U}(r). \tag{16}$$

Note that $\{|b_i\rangle : i = 1, \ldots, n\}$ is an orthonormal set due to the trace-preserving property of quantum channels.

Lemma 1. $\|g(F_1, F_2, \ldots, F_d)\|_{\vec{\mu}} = \|g(QF_1Q^\dagger, F_2Q^\dagger, \ldots, F_dQ^\dagger)\|_{\vec{\mu}}$ for any unitary matrix Q.

Proof. Let $G_1 = g(F_1, F_2, \ldots, F_d)$ and $G_2 = g(QF_1Q^\dagger, F_2Q^\dagger, \ldots, F_dQ^\dagger)$. First note that $G_2 = \tilde{Q}G_1Q^\dagger$, where $\tilde{Q} = \left[\begin{smallmatrix} Q & 0 \\ 0 & I \end{smallmatrix}\right]$. Problem (10) for G_2 is

$$\|G_2\|_{\vec{\mu}} \equiv \min_{V} \|V\|_{\vec{\mu}} \quad \text{s.t. } V_{[1,n]} = G_2 \quad \text{and} \quad V \text{ is unitary.}$$

Pre- and postmultiplication on the constraint gives

$$\|G_2\|_{\vec{\mu}} \equiv \min_{V} \|V\|_{\vec{\mu}} \quad \text{s.t. } \tilde{Q}^\dagger V_{[1,n]} Q = \tilde{Q}^\dagger G_2 Q \quad \text{and} \quad V \text{ is unitary.}$$

Further simplification on the constraint gives

$$\|G_2\|_{\vec{\mu}} \equiv \min_{V} \|\tilde{Q}^\dagger V \tilde{Q}\|_{\vec{\mu}}, \quad \text{s.t. } (\tilde{Q}^\dagger V \tilde{Q})_{[1,n]} = G_1 \quad \text{and} \quad \tilde{Q}^\dagger V \tilde{Q} \text{ is unitary.}$$

Here we used the fact that $\|V\|_{\vec{\mu}} = \|\tilde{Q}^\dagger V \tilde{Q}\|_{\vec{\mu}}$ for any unitary $\tilde{Q}$ since the eigenvalues are preserved under the conjugation by $\tilde{Q}$. Finally, noting that minimizing over V is the same as minimizing over $\tilde{Q}^\dagger V \tilde{Q}$, the claim that $\|G_1\|_{\vec{\mu}} = \|G_2\|_{\vec{\mu}}$ is proved. ■

Remark 1. (Triangularization of F_1) According to Lemma 1, $\|g(QF_1Q^\dagger, F_2Q^\dagger, \ldots, F_dQ^\dagger)\|_{\vec{\mu}}$ is invariant to unitary Q. Thus, we may choose any Q so that the Kraus operators are in a form that we desire. In particular, we may choose Q to be the unitary matrix of the Schur decomposition of F_1. (Note that the Schur decomposition is applicable to any matrix.) This makes $QF_1Q^\dagger$ upper triangular with the eigenvalues of F_1 on the diagonal.

IV. "PARTIAL U PROBLEM" WITH ONE VECTOR

We solve the "partial U problem" (16) for the special case of $n = 1$. This case turns out to be useful in computing the upper bound of the time energy $\|\mathcal{F}\|_{\vec{\mu}}$ (Sec. V), the lower bound of $\|\mathcal{F}\|_{\vec{\mu}}$ (Sec. VI), and the optimal $\|\mathcal{F}\|_{\vec{\mu}}$ for a class of quantum channels (Sec. VII).

A. Optimal single-vector transformation: General form

We consider the optimal $U \in \mathrm{U}(r)$ $(r \geqslant 2)$ for this problem:

$$P_{\vec{\mu}}(|a\rangle, |b\rangle) \equiv \min_{U} \|U\|_{\vec{\mu}} \quad \text{s.t. } U|a\rangle = |b\rangle \quad \text{with} \quad U \in \mathrm{U}(r) \tag{17}$$

where $|a\rangle$ and $|b\rangle$ are general normalized vectors of length r. We first show that the optimal U can be achieved with two nonzero eigenangles for any $\vec{\mu}$. Then we show how to construct U given two eigenangles, find $\|U\|_{\max}$, and bound $\|U\|_{\mathrm{sum}}$.

Note that the solution is of the form

$$P_{\vec{\mu}}(|a\rangle, |b\rangle) = f_{\vec{\mu}}(\langle a|b\rangle). \tag{18}$$

This is because $\|U\|_{\vec{\mu}} = \|V^\dagger U V\|_{\vec{\mu}}$ for any unitary V, and thus $P_{\vec{\mu}}(|a\rangle, |b\rangle) = P_{\vec{\mu}}(V|a\rangle, V|b\rangle)$.

In the following, we assume $|a\rangle \neq |b\rangle$. The case $|a\rangle = |b\rangle$ is trivial since $P_{\vec{\mu}}(|a\rangle, |a\rangle) = 0$ with $U = I$.

1. Optimal U operates nontrivially on a two-dimensional subspace

Consider the constraint $U|a\rangle = |b\rangle$. We have

$$\sum_{j=1}^{r} \exp(i\theta_j)|\langle u_j|a\rangle|^2 = \langle a|b\rangle, \tag{19}$$

where $U = \sum_{j=1}^{r} \exp(i\theta_j)|u_j\rangle\langle u_j|$ is the eigendecomposition of U. Thus, the point $\langle a|b\rangle$ is a linear combination of the vertices $\exp(i\theta_j)$ on the unit circle with weights $|\langle u_j|a\rangle|^2$. We show that the optimal U for the time-energy measure $\|U\|_{\vec{\mu}}$ can always be achieved with a linear combination of two vertices.

Lemma 2. U with the minimal $\|U\|_{\vec{\mu}}$ such that Eq. (19) is satisfied can always be achieved with two nonzero eigenangles.

Proof. Suppose the number of θ_j with nonzero weights (i.e., $\langle u_j|a\rangle \neq 0$) is $m > 2$. From these m vertices, pick two adjacent vertices that are either both positive or both negative. This can always be done since $m > 2$. Without loss of generality (w.l.o.g.), denote these two vertices as $\exp(i\theta_1)$ and $\exp(i\theta_2)$, and the remaining vertices as $\exp(i\theta_3)$ to $\exp(i\theta_m)$. Since $\sum_{j=1}^{m} |\langle u_j|a\rangle|^2 = 1$, the point $\langle a|b\rangle$ lies inside the polygon defined by the vertices $\exp(i\theta_j), j = 1,\ldots,m$, according to Eq. (19). If we replace the edge connecting $\exp(i\theta_1)$ and $\exp(i\theta_2)$ by their arc on the unit circle, the resultant shape will be strictly larger and contain the original polygon. This new shape can be expressed as

$$\bigcup_{\theta_2' \in [\theta_1,\theta_2]} \text{Polygon}(\theta_2',\theta_3,\ldots,\theta_m),$$

where "Polygon" denotes the polygon defined by the vertices given in the arguments. Here we assume $\theta_1 \leqslant \theta_2$ w.l.o.g. Since the point $\langle a|b\rangle$ lies inside this shape, it must also lie inside one of the polygons, each defined with $m-1$ vertices. Therefore, the point can be obtained as a linear combination of $m-1$ vertices (see Fig. 1). Essentially, we replace θ_1 and θ_2 by some θ_2' defining the relevant polygon. It remains to verify that the time-energy measure using these $m-1$ vertices is no larger than before. Denote by $P(j), j = 1,\ldots,r$, the decreasing order of $|\theta_j|$ where $\{\theta_{m+1},\ldots,\theta_r\}$ are the eigenangles of U with zero weights (i.e., $\langle u_j|a\rangle = 0$). Denote by $P'(j), j = 1,\ldots,r$, the decreasing order of $|\theta_j'|$ where $\theta_j' = \theta_j$ for $j = 3,\ldots,m$ and $\theta_j' = 0$ for $j = 1, m+1,\ldots,r$. (θ_2' is defined above.)

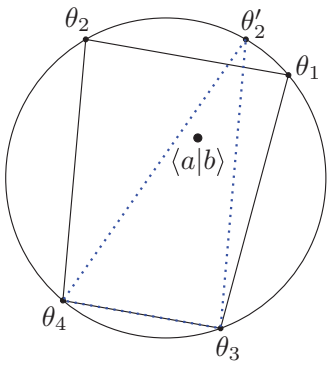

FIG. 1. (Color online) The point $\langle a|b\rangle$ is initially obtained as a linear combination of four points at eigenangles $\theta_j, j = 1,\ldots,4$ of U based on Eq. (19). A new eigenangle θ_2' can be found such that $\langle a|b\rangle$ is a linear combination of θ_2', θ_3, and θ_4.

We have

$$\begin{aligned}\|U\|_{\vec{\mu}} &= \sum_{j=1}^{r} \mu_{P(j)}|\theta_j| \\ &\geqslant \sum_{j=3}^{m} \mu_{P'(j)}|\theta_j'| + \mu_{P'(2)} \max(|\theta_2|,|\theta_1|) \\ &\quad + \mu_{P'(1)} \min(|\theta_2|,|\theta_1|) + \sum_{j=m+1}^{r} \mu_{P'(j)}|\theta_j| \\ &\geqslant \sum_{j=3}^{m} \mu_{P'(j)}|\theta_j'| + \mu_{P'(2)}|\theta_2'|,\end{aligned}$$

where the second last line is due to $\sum_{j=1}^{r} \mu_{P(j)}|\theta_j| \geqslant \sum_{j=1}^{r} \mu_{P''(j)}|\theta_j|$ for any ordering P'', and the last line is due to $\theta_2' \in [\theta_1,\theta_2]$. In summary, a new U' can be formed using these $m-1$ eigenangles $\{\theta_2',\ldots,\theta_m'\}$ with $\|U'\|_{\vec{\mu}} \leqslant \|U\|_{\vec{\mu}}$. We can repeat this argument for removing another vertex until we reach $m = 2$. This proves that the optimal U for the time-energy measure can always be achieved with a linear combination of two vertices or, in other words, two nonzero eigenangles. ■

This lemma implies that when finding an optimal U with respect to $\|U\|_{\vec{\mu}}$, it is sufficient to consider all chords (i.e., two-vertex polygons) on the unit circle passing through the desired point $\langle a|b\rangle$. Each chord defines two eigenangles, θ_1 and θ_2, which in turn define a unitary transformation from $|a\rangle$ to $|b\rangle$. This transformation $\tilde{U}$ acts on the subspace spanned by $|a\rangle$ and $|b\rangle$:

$$\begin{aligned}\tilde{U} &= \tilde{u}_1|a_i\rangle\langle a_i| + \tilde{u}_2|a_i\rangle\langle a_i^\perp| + \tilde{u}_3|a_i^\perp\rangle\langle a_i| + \tilde{u}_4|a_i^\perp\rangle\langle a_i^\perp| \\ &= \begin{bmatrix}\tilde{u}_1 & \tilde{u}_2 \\ \tilde{u}_3 & \tilde{u}_4\end{bmatrix}\end{aligned} \tag{20}$$

expressed in the basis $\{|a\rangle,|a^\perp\rangle\}$. Here

$$|a^\perp\rangle = \frac{1}{\sqrt{1-|\langle a|b\rangle|^2}}(|b\rangle - \langle a|b\rangle|a\rangle) \tag{21}$$

is a vector orthogonal to $|a\rangle$ in the plane spanned by $|a\rangle$ and $|b\rangle$. (We assume $|a\rangle \neq |b\rangle$.) The entries of $\tilde{U}$ can be found by imposing that the eigenvalues are $\exp(i\theta_1)$ and $\exp(i\theta_2)$ and $\tilde{U}|a\rangle = |b\rangle$ (see Appendix for details):

$$\tilde{U} = \begin{bmatrix}\langle a|b\rangle & -e^{i(\theta_1+\theta_2)}\sqrt{1-|\langle a|b\rangle|^2} \\ \sqrt{1-|\langle a|b\rangle|^2} & e^{i(\theta_1+\theta_2)}\langle b|a\rangle\end{bmatrix}. \tag{22}$$

The overall transformation is composed of the transformation $\tilde{U}$ in the subspace spanned by $|a\rangle$ and $|b\rangle$ and a transformation $\tilde{U}^\perp$ in the orthogonal subspace:

$$U = \tilde{U} + \tilde{U}^\perp. \tag{23}$$

We assign $\tilde{U}^\perp$ with zero eigenangles:

$$\tilde{U}^\perp = I - |a\rangle\langle a| - |a^\perp\rangle\langle a^\perp|. \tag{24}$$

This ensures that $\|U\|_{\vec{\mu}}$ is minimized. Thus, $\|U\|_{\vec{\mu}} = \|\tilde{U}\|_{\vec{\mu}} = \mu_1|\theta_i| + \mu_2|\theta_{3-i}|$ where $i = \arg\max_{j=1,2}|\theta_j|$. Note that the eigenvectors corresponding to $\tilde{U}^\perp$ in Eq. (19) have weights $|\langle u_j|a\rangle|^2 = 0$.

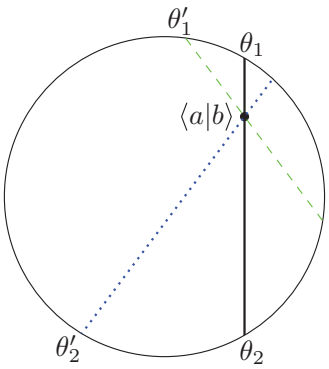

FIG. 2. (Color online) The optimal U for the max time energy consists of two nontrivial eigenangles θ_1 and θ_2, which form a vertical line passing through the point $\langle a|b\rangle$. Two other rotated lines about the point are shown. The dashed line (green) and the dotted line (blue) have max time energy θ_1' and $|\theta_2'|$, respectively, which are both greater than the optimal value θ_1.

B. Optimal max time energy

For the max time energy, we show that the optimal U of problem (17) has two nontrivial eigenangles $\theta_1,\theta_2 = \pm\cos^{-1}[\mathrm{Re}(\langle a|b\rangle)]$, where $\cos^{-1}$ always returns an angle in the range $[0,\pi]$, and U has the form of Eq. (23). The linear combination of these two eigenangles corresponds to a vertical line passing through the point $\langle a|b\rangle$ (see Fig. 2). It can easily be seen that this line gives the minimal max time energy. Consider a line obtained by rotating the vertical line about $\langle a|b\rangle$. If $\langle a|b\rangle$ is strictly inside the unit circle, then one of the two eigenangles must become larger in magnitude, giving rise to a larger max time energy of $\max(|\theta_1|,|\theta_2|)$. If $\langle a|b\rangle$ is on the unit circle, then one of the two eigenangles remains unchanged, and so the max time energy cannot become smaller. Therefore, we have

$$P_{\max}(|a\rangle,|b\rangle) = f_{\max}(\langle a|b\rangle) = \cos^{-1}[\mathrm{Re}(\langle a|b\rangle)]. \quad (25)$$

C. Sum time energy

We only derive lower and upper bounds on the sum time energy for problem (17).

Lemma 3. For each chord that passes through the point $r\exp(i\gamma)$, we associate a triangle formed by the origin and the chord. Among all such chords, the minimum angle of the triangle at the origin is 2β where $r = \cos(\beta)$.

Proof. Note that the problem is invariant to the rotation by γ. Thus, w.l.o.g. we assume $\gamma = 0$. Let the two end points of the chord be $\exp(i\zeta_1)$ and $\exp(-i\zeta_2)$, where $\zeta_1,\zeta_2 \geqslant 0$. The angle in question is $\zeta_1+\zeta_2$ and we show that $\zeta_1+\zeta_2 \geqslant 2\beta$. The point $r\exp(i0)$ is a linear combination of these two end points: $r\exp(i0) = z\exp(i\zeta_1) + (1-z)\exp(-i\zeta_2)$, where $0 \leqslant z \leqslant 1$. Thus, ζ_1, ζ_2, and z have to satisfy the constraint on the magnitude:

$$\begin{aligned} r^2 &= [z\cos(\zeta_1) + (1-z)\cos(\zeta_2)]^2 \\ &\quad + [z\sin(\zeta_1) - (1-z)\sin(\zeta_2)]^2 \\ &= [2 - 2\cos(\zeta_1+\zeta_2)](z^2 - z) + 1. \end{aligned}$$

This implies that z is a function of $\zeta_1+\zeta_2$. Solving the quadratic equation, we get

$$z = \frac{A \pm \sqrt{A^2 - 4A(1-r^2)}}{2A},$$

where $A = 2 - 2\cos(\zeta_1+\zeta_2)$. Note that if $A = 0$, then $\zeta_1 = \zeta_2 = 0$, which implies that $r = 1$ and $\beta = 0$; thus, $0 = \zeta_1 + \zeta_2 \geqslant 2\beta = 0$ as claimed. Otherwise, $A > 0$ and in this case, z has a real solution if

$$\begin{aligned} &0 \leqslant A - 4(1-r^2) \\ \Longrightarrow\ &\cos(\zeta_1+\zeta_2) \leqslant 2r^2 - 1 = 2\cos^2(\beta) - 1 = \cos(2\beta), \end{aligned}$$

where $r = \cos(\beta)$. Since cos is a decreasing function in the domain $[0,\pi]$, $\zeta_1+\zeta_2 \geqslant 2\beta$ as claimed. ■

Remark 2. Note that the minimum angle of 2β in Lemma 3 is achieved by the triangle formed by the origin and the chord perpendicular to the line connecting the origin and $r\exp(i\gamma)$.

Lemma 4. The solution to problem (17) for the sum time energy is lower bounded as follows:

$$\begin{aligned} P_{\mathrm{sum}}(|a\rangle,|b\rangle) &= f_{\mathrm{sum}}(\langle a|b\rangle) \geqslant 2\cos^{-1}|\langle a|b\rangle| \\ &\equiv f_{\mathrm{sum}}^{\mathrm{L}}(\langle a|b\rangle). \end{aligned} \quad (26)$$

Proof. The sum time energy of U is $|\theta_1| + |\theta_2|$ since U has only two nontrivial eigenangles due to Lemma 2. The chord defined by these two eigenangles, θ_1 and θ_2, passes through the point $\langle a|b\rangle = r\exp(i\gamma)$, where $r = |\langle a|b\rangle|$. We consider the triangle formed by the origin and the chord and focus on the angle at the origin. For the case $\theta_1 > 0$, $\theta_2 \leqslant 0$ and the case $\theta_1 \leqslant 0$, $\theta_2 > 0$, this angle is $\min(|\theta_1-\theta_2|, 2\pi - |\theta_1 - \theta_2|)$. According to Lemma 3, this angle is lower bounded by 2β where $\beta = \cos^{-1}(r)$. Thus,

$$\begin{aligned} |\theta_1| + |\theta_2| = |\theta_1 - \theta_2| &\geqslant \min(|\theta_1-\theta_2|, 2\pi - |\theta_1-\theta_2|) \\ &\geqslant 2\beta. \end{aligned}$$

Equality is achieved when $|\theta_1 - \theta_2| \leqslant \pi$ and the chord described by θ_1 and θ_2 is perpendicular to the line connecting the origin and $r\exp(i\gamma)$ (see Remark 2). For the case $\theta_1,\theta_2 > 0$, the case $\theta_1,\theta_2 < 0$, and the case $\theta_1 = \theta_2 = 0$, the angle is $|\theta_1 - \theta_2|$, which is lower bounded by 2β according to Lemma 3. Thus, we have

$$|\theta_1| + |\theta_2| \geqslant |\theta_1 - \theta_2| \geqslant 2\beta.$$

■

Lemma 5. The solution to problem (17) for the sum time energy is upper bounded as follows:

$$\begin{aligned} P_{\mathrm{sum}}(|a\rangle,|b\rangle) &= f_{\mathrm{sum}}(\langle a|b\rangle) \leqslant 2\cos^{-1}[\mathrm{Re}(\langle a|b\rangle)] \\ &\equiv f_{\mathrm{sum}}^{U}(\langle a|b\rangle). \end{aligned} \quad (27)$$

Proof. Any U that satisfies $U|a\rangle = |b\rangle$ [i.e., the constraint of problem (17)] serves as an upper bound to $P_{\vec{\mu}}(|a\rangle,|b\rangle)$ for any $\vec{\mu}$. Thus, for simplicity, we choose the optimal U that achieves the optimal max time energy in Eq. (25) to serve as an upper bound to $P_{\mathrm{sum}}(|a\rangle,|b\rangle)$. This U has two nontrivial eigenangles $\theta_1 = \cos^{-1}[\mathrm{Re}(\langle a|b\rangle)]$ and $\theta_2 = -\theta_1$. Thus, the sum time energy of this U is $|\theta_1| + |\theta_2| = 2\cos^{-1}[\mathrm{Re}(\langle a|b\rangle)]$, and this is an upper bound to $P_{\mathrm{sum}}(|a\rangle,|b\rangle)$. ■

V. TIME-ENERGY UPPER BOUND

In this section we consider upper bounding $\|\mathcal{F}\|_{\vec{\mu}}$ given its Kraus operators $(F_1, F_2, \ldots, F_d)$ by upper bounding $\|g(F_{1:d})\|_{\vec{\mu}}$. Any implementation U of the quantum channel $\mathcal{F}$ serves as an upper bound to $\|\mathcal{F}\|_{\vec{\mu}}$ since $\|\mathcal{F}\|_{\vec{\mu}} \leqslant \|U\|_{\vec{\mu}}$ for all U of the form of Eq. (12). We propose a simple method to construct a time-energy-efficient U that completes the partial matrix $g(F_{1:d})$, and $\|U\|_{\vec{\mu}}$ will serve as an upper bound to the "partial U problem" (10) for $\|g(F_{1:d})\|_{\vec{\mu}}$, which is an intermediate problem to the ultimate "channel problem" (13).

A. Successive construction of U

We focus on finding an upper bound to the "partial U problem" (10), which was recast as problem (16), which finds a unitary matrix U that satisfies n transformation rules: $|e_i\rangle \longrightarrow |b_i\rangle$ for all $i = 1, \ldots, n$. Here we focus on the "partial U problem" for $\|g(F_{1:d})\|_{\vec{\mu}}$. Thus, $|b_i\rangle$ is the ith column of $g(F_{1:d})$. In Sec. IV we analyzed the optimal unitary operation U_i for each single-vector transformation $|e_i\rangle \longrightarrow |b_i\rangle$, and we solved $\|U_i\|_{\max}$ and bounded $\|U_i\|_{\text{sum}}$. Motivated by this result, we propose a greedy method to construct U in which we successively construct the best unitary U_i for each $i = 1, \ldots, n$, and concatenate them. The overall U will be

$$U = U_n \cdots U_1. \tag{28}$$

We design each U_i as follows. When we consider the first transformation (i.e., $i = 1$), we seek the optimal U_1 with the minimal $\|U_1\|_{\vec{\mu}}$ such that

$$U_1|e_1\rangle = |b_1\rangle.$$

When $i = 2$, we seek the optimal U_2 such that

$$U_2 U_1 |e_2\rangle = |b_2\rangle.$$

In general, for the ith transformation, we seek the optimal U_i such that

$$U_i|a_i\rangle = |b_i\rangle, \tag{29}$$

where

$$|a_i\rangle = U_{i-1} \cdots U_1 |e_i\rangle \quad \text{for} \quad n \geqslant i \geqslant 2 \quad \text{and} \quad |a_1\rangle = |e_1\rangle. \tag{30}$$

Note that the optimal U_i with the minimal $\|U_i\|_{\vec{\mu}}$ has already been considered in problem (17). We obtained the optimal value for $\|U_i\|_{\max}$ and lower and upper bounds for $\|U_i\|_{\text{sum}}$ [see Eqs. (25)–(27)]. Thus,

$$\|U_i\|_{\vec{\mu}} = f_{\vec{\mu}}(\langle a_i|b_i\rangle), \tag{31}$$

where the RHS comes from Eq. (18).

A key feature of our construction is that we design U_i successively in a backward-looking fashion; i.e., when we design U_i, we only need to know U_j for $j < i$, and we do not use U_j for $j > i$.

In order for this successive approach to work, the action of a higher-index U_j should not affect the transformation of a lower index $i < j$, i.e.,

$$\begin{aligned} U_{i+1}U_i|a_i\rangle &= |b_i\rangle, \\ U_{i+2}U_{i+1}U_i|a_i\rangle &= |b_i\rangle, \\ &\vdots \\ U_n \cdots U_{i+2}U_{i+1}U_i|a_i\rangle &= |b_i\rangle. \end{aligned} \tag{32}$$

Only if the last equation holds for all i does the overall U transforms according to Eq. (15) as required. We show that Eq. (32) does hold.

Lemma 6. (Backward-looking design of U_i)

$$U_{i+j} \cdots U_{i+2}U_{i+1}U_i|a_i\rangle = |b_i\rangle \tag{33}$$

for $j \geqslant 1$ when Eq. (29) holds.

Proof. We need to use the essential properties that $\langle e_i|e_k\rangle = \langle b_i|b_k\rangle = \delta_{ik}$ where δ_{ik} is the Kronecker delta. We prove by induction. First, we compute $U_{i+1}U_i|a_i\rangle$. Recall from Eq. (23) that U_{i+1} performs a nontrivial transformation only in the subspace spanned by $|a_{i+1}\rangle$ and $|b_{i+1}\rangle$. We show that $U_i|a_i\rangle$ is not in this subspace. Note that

$$\begin{aligned} \langle a_{i+1}|U_i|a_i\rangle &= [\langle e_{i+1}|U_1^\dagger U_2^\dagger \cdots U_i^\dagger]U_i[U_{i-1} \cdots U_1|e_i\rangle] \\ &= 0 \end{aligned} \tag{34}$$

by Eq. (30), and also $\langle b_{i+1}|U_i|a_i\rangle = \langle b_{i+1}|b_i\rangle = 0$ by Eq. (29). Hence, $U_i|a_i\rangle$ is not in the aforementioned subspace. This shows that

$$\begin{aligned} U_{i+1}U_i|a_i\rangle &= (\tilde{U}_{i+1} + \tilde{U}_{i+1}^\perp)U_i|a_i\rangle \\ &= \tilde{U}_{i+1}^\perp U_i|a_i\rangle = U_i|a_i\rangle = |b_i\rangle \end{aligned}$$

since $\tilde{U}_{i+1}^\perp$ acts trivially on the orthogonal complement of the subspace spanned by $|a_{i+1}\rangle$ and $|b_{i+1}\rangle$ [cf. Eq. (24)].

Now consider $U_{i+j} \cdots U_{i+1}U_i|a_i\rangle$ assuming the hypothesis $U_{i+j-1} \cdots U_{i+1}U_i|a_i\rangle = |b_i\rangle$. Similar to Eq. (34),

$$\begin{aligned} &\langle a_{i+j}|U_{i+j-1} \cdots U_{i+1}U_i|a_i\rangle \\ &\quad = [\langle e_{i+j}|U_1^\dagger U_2^\dagger \cdots U_{i+j-1}^\dagger]U_{i+j-1} \cdots \\ &\qquad U_{i+1}U_i[U_{i-1} \cdots U_1|e_i\rangle] = 0 \end{aligned}$$

and

$$\langle b_{i+j}|U_{i+j-1} \cdots U_{i+1}U_i|a_i\rangle = \langle b_{i+j}|b_i\rangle = 0$$

using the hypothesis.

Therefore, $U_{i+j-1} \cdots U_{i+1}U_i|a_i\rangle$ is not in the subspace spanned by $|a_{i+j}\rangle$ and $|b_{i+j}\rangle$ (the subspace that U_{i+j} acts nontrivially), and thus

$$U_{i+j}U_{i+j-1} \cdots U_{i+1}U_i|a_i\rangle = U_{i+j-1} \cdots U_{i+1}U_i|a_i\rangle = |b_i\rangle.$$

This proves the claim. ∎

This is the key lemma that allows us to compute an upper bound in a successive manner. In essence, instead of considering the original problem of finding a unitary required to perform n simultaneous transformations, we consider the problem of finding n unitaries each required to perform one transformation. Note that we already solved the latter problem in Sec. IV A. In particular, the max time energy for a single-vector transformation is given in Eq. (25). Thus,

according to Eq. (28),

$$\|U\|_{\vec{\mu}} = \|U_n \cdots U_1\|_{\vec{\mu}} \leqslant \sum_{i=1}^{n} \|U_i\|_{\vec{\mu}}$$

$$= \sum_{i=1}^{n} f_{\vec{\mu}}(\langle a_i|b_i\rangle), \tag{35}$$

where the inequality is due to the triangle inequality of the norm $\|\cdot\|_{\vec{\mu}}$ (see Theorem 2 of Ref. [16]), and the last line is due to Eq. (31).

In light of Remark 1, we can always assume that F_1 is initially given in upper triangular form. In the following, we use this property to further deduce a simple bound on $\|U\|_{\vec{\mu}}$ and consequently $\|\mathcal{F}\|_{\vec{\mu}}$. We show that when F_1 is upper triangular (which can always be guaranteed by Remark 1), $|a_i\rangle = |e_i\rangle$ for all $1 \leqslant i \leqslant n$ in Eqs. (29) and (30). Thus, $\|U_i\|_{\vec{\mu}}$ in Eq. (35) will only depend on $\langle e_i|b_i\rangle$ which is the ith eigenvalue of F_1.

Lemma 7. (Independent design of U_i) When F_1 is upper triangular, $|a_i\rangle = |e_i\rangle$ for all $1 \leqslant i \leqslant n$ in Eqs. (29) and (30), which implies that designing according to

$$U_i|e_i\rangle = |b_i\rangle \tag{36}$$

and designing according to Eq. (29) are equivalent.

Proof. Note that since F_1 is upper triangular, $\langle b_j|e_i\rangle = 0$ for $j < i \leqslant n$. We prove by induction. We show that $|a_i\rangle = |e_i\rangle$ assuming the hypothesis that $|a_j\rangle = |e_j\rangle$ for $j < i$ is true. Note that $|a_1\rangle = |e_1\rangle$ by definition in Eq. (30). Recall from Eq. (23) that U_j performs a nontrivial transformation only in the subspace spanned by $|a_j\rangle$ and $|b_j\rangle$. We show that $|e_i\rangle$ for $i > j$ is not in this subspace: $\langle a_j|e_i\rangle = \langle e_j|e_i\rangle = 0$ where we assume that the hypothesis is true, and $\langle b_j|e_i\rangle = 0$ due to the triangular structure of F_1. This means that for $i > j$,

$$U_j|e_i\rangle = (\tilde{U}_j + \tilde{U}_j^{\perp})|e_i\rangle = \tilde{U}_j^{\perp}|e_i\rangle = |e_i\rangle$$

since $\tilde{U}_j^{\perp}$ acts trivially [cf. Eq. (24)]. This shows that

$$|a_i\rangle = U_{i-1} \cdots U_1|e_i\rangle = |e_i\rangle.$$

■

This lemma allows us to compute the upper bound easier since the upper bound in Eq. (35) now becomes

$$\|U\|_{\vec{\mu}} \leqslant \sum_{i=1}^{n} f_{\vec{\mu}}(\langle e_i|b_i\rangle) \tag{37}$$

$$= \sum_{i=1}^{n} f_{\vec{\mu}}(\lambda_i(F_1)), \tag{38}$$

where we recognize that the diagonal elements of F_1 are its eigenvalues, denoted by $\lambda_i(F_1)$. This shows that $\|U\|_{\vec{\mu}}$ only depends on the eigenvalues of one Kraus operator of the quantum channel. Since $\|\mathcal{F}\|_{\vec{\mu}} \leqslant \|U\|_{\vec{\mu}}$ for all U of the form of Eq. (12), we have

$$\|\mathcal{F}\|_{\vec{\mu}} \leqslant \min_{\mathbf{v}} \sum_{i=1}^{n} f_{\vec{\mu}}\left(\lambda_i\left(\sum_{j=1}^{d} v_j F_j\right)\right)$$

$$\text{s.t.} \sum_{j=1}^{d} |v_j|^2 \leqslant 1. \tag{39}$$

Here $\mathbf{v}$ corresponds to the first d elements of the first row of W_B in Eq. (12). This bound holds for any quantum channel $\mathcal{F}$ described by Kraus operators $F_i, i = 1, \ldots, d$.

B. Max time-energy upper bound

For the max time energy, we substitute $f_{\max}$ in Eq. (25) for $f_{\vec{\mu}}$ in Eq. (39) to get

$$\|\mathcal{F}\|_{\max} \leqslant \min_{\mathbf{v}:\,\|\mathbf{v}\|\leqslant 1} \sum_{i=1}^{n} \cos^{-1}\left\{\mathrm{Re}\left[\lambda_i\left(\sum_{j=1}^{d} v_j F_j\right)\right]\right\}. \tag{40}$$

C. Sum time-energy upper bound

For the sum time energy, we substitute f_{sum}^{U} in Eq. (27) for $f_{\vec{\mu}}$ in Eq. (39) to get

$$\|\mathcal{F}\|_{\text{sum}} \leqslant \min_{\mathbf{v}:\,\|\mathbf{v}\|\leqslant 1} \sum_{i=1}^{n} 2\cos^{-1}\left\{\mathrm{Re}\left[\lambda_i\left(\sum_{j=1}^{d} v_j F_j\right)\right]\right\}. \tag{41}$$

D. Special case: Diagonal F_1

Here, we consider the special class of channels for which F_1 is diagonal, which will be useful for showing the optimal time energy for a class of channels in Sec. VII. Following the construction of U in Eq. (28), we argue not only that we can independently design U_i (Lemma 7), but also that they act on orthogonal subspaces (meaning that they commute). According to Lemma 7, U_i transforms $|e_i\rangle$ to $|b_i\rangle$, and is the solution to problem (17) where $|a\rangle = |e_i\rangle$ and $|b\rangle = |b_i\rangle$. Thus, its nontrivial part $\tilde{U}_i$ [see Eq. (23)] acts on the subspace spanned by $\{|e_i\rangle, |b_i\rangle\}$.

The fact that F_1 is diagonal gives rise to the following property: $\langle e_i|b_j\rangle = 0$ if $i \neq j$ where $1 \leqslant i, j \leqslant n$. Also, we already have $\langle e_i|e_j\rangle = 0$ if $i \neq j$ by definition and $\langle b_i|b_j\rangle = 0$ if $i \neq j$ by the trace-preserving property of quantum channels. Thus, the subspaces spanned by $\{|e_i\rangle, |b_i\rangle\}$ for $i = 1, \ldots, n$ are orthogonal to each other. This means that $\tilde{U}_i$ acting on these subspaces [see Eq. (20)] are orthogonal to each other. Then we may bypass the construction of U_i in Eq. (23) and directly form

$$U = \sum_{i=1}^{n} \tilde{U}_i + P_{U^{\perp}},$$

where $P_{U^{\perp}}$ is the projection onto the subspace orthogonal to the summation term. Thus, the set of nonzero eigenangles of U is composed of the eigenangles of $\tilde{U}_i$ for all $i = 1, \ldots, n$. This means that

$$\|U\|_{\max} = \max_{1\leqslant i\leqslant n} \|\tilde{U}_i\|_{\max}, \tag{42}$$

$$\|U\|_{\text{sum}} = \sum_{i=1}^{n} \|\tilde{U}_i\|_{\text{sum}}. \tag{43}$$

Thus, for this class of channels, by using Eqs. (25), (27), (31) and $\|U_i\|_{\vec{\mu}} = \|\tilde{U}_i\|_{\vec{\mu}}$, we have

$$\|\mathcal{F}\|_{\max} \leqslant \|U\|_{\max} = \max_{1\leqslant i\leqslant n} \cos^{-1}\{\mathrm{Re}[\lambda_i(F_1)]\}, \tag{44}$$

$$\|\mathcal{F}\|_{\mathrm{sum}} \leqslant \|U\|_{\mathrm{sum}} \leqslant \sum_{i=1}^{n} 2\cos^{-1}\{\mathrm{Re}[\lambda_i(F_1)]\}. \tag{45}$$

VI. TIME-ENERGY LOWER BOUND

A. General form of lower bound

We lower bound $\|\mathcal{F}\|_{\vec{\mu}}$ of the "channel problem" (13). We first consider lower bounding $\|g(F_{1:d})\|_{\vec{\mu}}$ for a fixed set of Kraus operators $(F_1, F_2, \ldots, F_d)$. Note that $\|g(F_{1:d})\|_{\vec{\mu}}$ is obtained from the "partial U problem" (16), and we propose a modified problem whose solution lower bounds this problem. The modified problem is formed by removing all except one transformation constraints in problem (16) as follows:

$$\begin{aligned} P_{\vec{\mu}}(|e_i\rangle,|b_i\rangle) = \min_{U}\ & \|U\|_{\vec{\mu}} \\ \text{s.t. }\ & U|e_i\rangle = |b_i\rangle \quad \text{with} \\ & U \in \mathrm{U}(r), \end{aligned}$$

which is defined for $i = 1, \ldots, n$. Here $|b_i\rangle$ is the ith column of $g(QF_1Q^\dagger, F_2Q^\dagger, \ldots, F_dQ^\dagger)$ where $QF_1Q^\dagger$ is an upper triangular matrix corresponding to the Schur decomposition of F_1 (see Remark 1). Note that this problem is the single-vector problem (17) which we analyzed in Sec. IV. Certainly, the feasible set of this problem contains that of problem (16), and so with the help of Lemma 1, we have $\|g(F_1, F_2, \ldots, F_d)\|_{\vec{\mu}} = \|g(QF_1Q^\dagger, F_2Q^\dagger, \ldots, F_dQ^\dagger)\|_{\vec{\mu}} \geqslant P_{\vec{\mu}}(|e_i\rangle,|b_i\rangle)$ for all $i = 1, \ldots, n$. Thus,

$$\begin{aligned} \|g(F_1, F_2, \ldots, F_d)\|_{\vec{\mu}} &\geqslant \max_{1\leqslant i\leqslant n} P_{\vec{\mu}}(|e_i\rangle,|b_i\rangle) \\ &= \max_{1\leqslant i\leqslant n} f_{\vec{\mu}}(\langle e_i|b_i\rangle) \\ &= \max_{1\leqslant i\leqslant n} f_{\vec{\mu}}(\lambda_i(F_1)), \end{aligned}$$

where we used Eq. (18) in the second line and the third line is due to Remark 1 with $\lambda_i(F_1)$ being the ith eigenvalue of F_1. Note that the last inequality holds when the LHS is replaced by $\|g(F_1, F_2, \ldots, F_d, \mathbf{0}, \ldots, \mathbf{0})\|_{\vec{\mu}}$ for any number of extra all-zero Kraus operators inserted. Combining with the "channel problem" (13), we have

$$\begin{aligned} \|\mathcal{F}\|_{\vec{\mu}} \geqslant \min_{\mathbf{v}} \max_{1\leqslant i\leqslant n} & f_{\vec{\mu}}\left(\lambda_i\left(\sum_{j=1}^{d} v_j F_j\right)\right) \\ \text{s.t. } & \sum_{j=1}^{d} |v_j|^2 \leqslant 1. \end{aligned} \tag{46}$$

Here $\mathbf{v}$ corresponds to the first d elements of the first row of W_B in Eq. (12). This bound holds for any quantum channel $\mathcal{F}$ described by Kraus operators $F_i, i = 1, \ldots, d$.

B. Max time-energy lower bound

For the max time energy, we substitute $f_{\max}$ in Eq. (25) for $f_{\vec{\mu}}$ in Eq. (46) to get

$$\|\mathcal{F}\|_{\max} \geqslant \min_{\mathbf{v}:\,\|\mathbf{v}\|\leqslant 1} \max_{1\leqslant i\leqslant n} \cos^{-1}\left\{\mathrm{Re}\left[\lambda_i\left(\sum_{j=1}^{d} v_j F_j\right)\right]\right\}. \tag{47}$$

C. Sum time-energy lower bound

For the sum time energy, we substitute f_{sum} for $f_{\vec{\mu}}$ in Eq. (46) to get

$$\|\mathcal{F}\|_{\mathrm{sum}} \geqslant \min_{\mathbf{v}:\,\|\mathbf{v}\|\leqslant 1} \max_{1\leqslant i\leqslant n} f_{\mathrm{sum}}\left(\lambda_i\left(\sum_{j=1}^{d} v_j F_j\right)\right), \tag{48}$$

and it is a simple argument to argue that the lower bound of $\|\mathcal{F}\|_{\mathrm{sum}}$ can be given in terms of $f_{\mathrm{sum}}^{\mathrm{L}}$ defined in Eq. (26). First, note that

$$\begin{aligned} P(\mathbf{v}) &\equiv \max_i f_{\mathrm{sum}}\left(\lambda_i\left(\sum_{j=1}^{d} v_j F_j\right)\right) \\ &\geqslant \max_i f_{\mathrm{sum}}^{\mathrm{L}}\left(\lambda_i\left(\sum_{j=1}^{d} v_j F_j\right)\right) \equiv Q(\mathbf{v}) \end{aligned}$$

for all $\mathbf{v}$. And we have

$$\min_{\mathbf{v}} P(\mathbf{v}) = P(\hat{\mathbf{v}}) \geqslant Q(\hat{\mathbf{v}}) \geqslant \min_{\mathbf{v}} Q(\mathbf{v}),$$

where $\hat{\mathbf{v}} = \arg\min_{\mathbf{v}} P(\mathbf{v})$. Therefore, we have

$$\begin{aligned} \|\mathcal{F}\|_{\mathrm{sum}} &\geqslant \min_{\mathbf{v}:\,\|\mathbf{v}\|\leqslant 1} Q(\mathbf{v}) \\ &= \min_{\mathbf{v}:\,\|\mathbf{v}\|\leqslant 1} \max_{1\leqslant i\leqslant n} 2\cos^{-1}\left|\lambda_i\left(\sum_{j=1}^{d} v_j F_j\right)\right|. \end{aligned}$$

Note that $Q(\mathbf{v}) \leqslant Q(\kappa\mathbf{v})$ for $0 \leqslant \kappa \leqslant 1$, and thus we have

$$\|\mathcal{F}\|_{\mathrm{sum}} \geqslant \min_{\mathbf{v}:\,\|\mathbf{v}\|=1} \max_{1\leqslant i\leqslant n} 2\cos^{-1}\left|\lambda_i\left(\sum_{j=1}^{d} v_j F_j\right)\right|. \tag{49}$$

VII. OPTIMAL TIME ENERGY FOR A CLASS OF CHANNELS

Definition. Define a class $\mathcal{C}(n)$ of quantum channels acting on $n \times n$ density matrices where each channel is described by Kraus operators $\{F_j \in \mathbb{C}^{n\times n} : j = 1, \ldots, d\}$ of the form

$$\begin{aligned} F_1 &= \sqrt{p}I \quad \text{where} \quad 0 \leqslant p \leqslant 1, \\ \mathrm{Tr}(F_j) &= 0, \quad j = 2, \ldots, d. \end{aligned} \tag{50}$$

The number d of Kraus operators of each channel can be different.

Note that this class contains the depolarizing channel, the bit-flip channel, and the phase-flip channel.

A. Optimal max time energy

We show that the lower bound of $\|\mathcal{F}\|_{\max}$ in Eq. (47) is achievable for any $\mathcal{F} \in \mathcal{C}(n)$ with $2 \leqslant n < \infty$. The RHS of Eq. (47) can be written as

$$\cos^{-1}\left\{\max_{\mathbf{v}:\,\|\mathbf{v}\|\leqslant 1}\ \min_{1\leqslant i\leqslant n} \mathrm{Re}\left[\lambda_i\left(\sum_{j=1}^{d} v_j F_j\right)\right]\right\} \tag{51}$$

since $\cos^{-1}$ is a decreasing function in the range $[0,\pi]$. Consider part of this term:

$$\begin{aligned} P(\mathbf{v}) &\equiv \min_{1\leqslant i\leqslant n} \mathrm{Re}\left[\lambda_i\left(\sum_{j=1}^{d} v_j F_j\right)\right] \\ &\leqslant \frac{1}{n}\sum_{i=1}^{n} \mathrm{Re}\left[\lambda_i\left(\sum_{j=1}^{d} v_j F_j\right)\right] \equiv Q(\mathbf{v}). \end{aligned}$$

Let the eigenvalues of $\sum_{j=2}^{d} v_j F_j$ be $\{\sigma_1,\ldots,\sigma_n\}$. Note that $\sum_{i=1}^{n}\sigma_i = 0$. Thus,

$$\lambda_i\left(\sum_{j=1}^{d} v_j F_j\right) = v_1\sqrt{p} + \sigma_i$$

and

$$Q(\mathbf{v}) = \frac{1}{n}\sum_{i=1}^{n} \mathrm{Re}(v_1\sqrt{p} + \sigma_i) = \mathrm{Re}(v_1\sqrt{p}) \leqslant \sqrt{p}.$$

Considering the maximization in Eq. (51),

$$\max_{\mathbf{v}} P(\mathbf{v}) \leqslant Q(\hat{\mathbf{v}}) \leqslant \max_{\mathbf{v}} Q(\mathbf{v}),$$

where $\hat{\mathbf{v}}$ is the optimal value of the maximization of P. Thus,

$$\|\mathcal{F}\|_{\max} \geqslant \cos^{-1}[\max_{\mathbf{v}} P(\mathbf{v})] \geqslant \cos^{-1}(\sqrt{p})$$

since $\cos^{-1}$ is a decreasing function in the range $[0,\pi]$. Note that the RHS coincides with the upper bound from Eq. (44) where $\lambda_i(F_1) = \sqrt{p}$. Therefore,

$$\|\mathcal{F}\|_{\max} = \cos^{-1}(\sqrt{p}) \tag{52}$$

for any $\mathcal{F} \in \mathcal{C}(n)$ with $2 \leqslant n < \infty$.

VIII. SOME INTERESTING CONSEQUENCES

A. Comparison of quantum and classical noisy channels

The quantum noisy channel or the quantum depolarizing channel acting on $n \times n$ density matrices is defined as

$$\mathcal{F}_Q(\rho) \triangleq q\rho + (1-q)\frac{I}{n},$$

where complete positivity requires that $-1/(n^2-1) \leqslant q \leqslant 1$ [19].

The Weyl operators are the n-dimensional generalization of the Pauli operators and are defined as

$$S_{jk} = \sum_{s=0}^{n-1} \omega^{sk}|s+j\rangle\langle s| \in \mathbb{C}^{n\times n},$$

where $j,k = 0,\ldots,n-1$ and ω is the nth root of unity. These operators have the following properties:

(1) (Identity) $S_{00} = I$.

(2) (Traceless) $\mathrm{Tr}(S_{jk}) = 0$ for $(j,k) \neq (0,0)$.

(3) (Complete erasure) $n^{-2}\sum_{j,k=0}^{n-1} S_{jk}\rho S_{jk}^{\dagger} = n^{-1}I$ for any density matrix ρ.

(4) (Trace preserving) $n^{-2}\sum_{j,k=0}^{n-1} S_{jk}^{\dagger}S_{jk} = I$.

(5) (Complete erasure) $n^{-1}\sum_{j=0}^{n-1} S_{j0}\rho S_{j0}^{\dagger} = n^{-1}I$ for any diagonal density matrix ρ.

(6) (Trace preserving) $n^{-1}\sum_{j}^{n-1} S_{j0}^{\dagger}S_{j0} = I$.

We may express the quantum noisy channel using the Weyl's operators:

$$\mathcal{F}_Q(\rho) = qI\rho I + (1-q)\frac{1}{n^2}\sum_{j,k=0}^{n-1} S_{jk}\rho S_{jk}^{\dagger}. \tag{53}$$

The advantage of doing so is that we can now see that $\mathcal{F}_Q(\rho)$ is in the class $\mathcal{C}(n)$ (defined in Definition 1), and we can apply Eq. (52) to get the max time energy of it.

In a similar manner, we define the classical noisy channel, which adds classical noise (i.e., classical states are being remapped):

$$\mathcal{F}_C(\rho) \triangleq qI\rho I + (1-q)\frac{1}{n}\sum_{j=0}^{n-1} S_{j0}\rho S_{j0}^{\dagger}, \tag{54}$$

where $-1/(n-1) \leqslant q \leqslant 1$ for complete positivity. When ρ is diagonal (i.e., a mixture of classical states),

$$\mathcal{E}_C(\rho) = q\rho + (1-q)\frac{I}{n}.$$

We now verify that we have a fair comparison between the quantum and classical noisy channels, by checking that the same amount of noise is added to the input states of both channels. To quantify this, we use the trace distance to measure the difference between the input state and the output state, and we take the input state to be a pure state. For the quantum depolarizing channel, consider any pure input state $|\Phi\rangle\langle\Phi|$ and the trace distance is

$$\begin{aligned} \frac{1}{2}\||\Phi\rangle\langle\Phi| - \mathcal{F}_Q(|\Phi\rangle\langle\Phi|)\|_{\mathrm{tr}} &= \frac{1}{2}(1-q)\left\||\Phi\rangle\langle\Phi| - \frac{I}{n}\right\|_{\mathrm{tr}} \\ &= \frac{(1-q)(n-1)}{n} \equiv \delta, \end{aligned} \tag{55}$$

where $\|A\|_{\mathrm{tr}}$ denotes the trace norm of A and is equal to the sum of all singular values of A. For the classical noisy channel, we only consider classical states and so the input state is $|j\rangle, j = 0,\ldots,n-1$, and in this case, the trace distance is

$$\begin{aligned} \frac{1}{2}\||j\rangle\langle j| - \mathcal{F}_C(|j\rangle\langle j|)\|_{\mathrm{tr}} &= \frac{1}{2}(1-q)\left\||j\rangle\langle j| - \frac{I}{n}\right\|_{\mathrm{tr}} \\ &= \delta. \end{aligned}$$

This shows that both $\mathcal{F}_Q$ and $\mathcal{F}_C$ add the same amount of noise when both are characterized by the same parameter q. Note that the two channels have different valid ranges of q, but this does not affect our discussion since we will focus on $q \approx 1$.

It can be easily seen that both $\mathcal{F}_Q$ and $\mathcal{F}_C$ written with the Weyl's operators are in the class $\mathcal{C}(n)$. In this case, $\|\mathcal{F}_Q\|_{\max}$ and $\|\mathcal{F}_C\|_{\max}$ depend only on the respective scaling factors of

the identity Kraus operators [cf. Eq. (52)]. Note that $S_{00} = I$, and thus from Eqs. (53) and (54), the identity Kraus operators [cf. Eq. (50)] are

$$F_{Q1} = \sqrt{q + \frac{1-q}{n^2}}\, I = \sqrt{1 - \delta\left(\frac{n+1}{n}\right)}\, I,$$

$$F_{C1} = \sqrt{q + \frac{1-q}{n}}\, I = \sqrt{1-\delta}\, I,$$

where we have used Eq. (55). Note that when $q = 1$, $F_{Q1} = F_{C1} = I$; when $q = -1/(n^2-1)$, $F_{Q1} = 0$, and when $q = -1/(n-1)$, $F_{C1} = 0$. Using these in Eq. (52), we have

$$\|\mathcal{F}_Q\|_{\max} = \cos^{-1}\sqrt{1 - \delta\left(\frac{n+1}{n}\right)}, \tag{56}$$

$$\|\mathcal{F}_C\|_{\max} = \cos^{-1}\sqrt{1-\delta}, \tag{57}$$

where in both cases the distance between the input and output states is δ. When $\delta \approx 0$, using the approximation $\cos^{-1}\sqrt{x} \approx \sqrt{1-x}$ for $x \approx 1$, it can be shown that

$$\|\mathcal{F}_Q\|_{\max} = \sqrt{\frac{n+1}{n}}\|\mathcal{F}_C\|_{\max}. \tag{58}$$

This shows that it takes $\sqrt{(n+1)/n}$ times more time-energy resource for a quantum process to erase information of the input state by the same distance δ compared to a classical process. For two-level systems ($n = 2$), this is 1.22 times larger.

B. Cascade of depolarizing channels with small noise

As discussed in the last subsection, the quantum depolarizing channel

$$\mathcal{F}_Q(\rho) \triangleq q\rho + (1-q)\frac{I}{n}$$

is in the class $\mathcal{C}(n)$ (defined in Definition 1), and we can apply Eq. (52) to get

$$\|\mathcal{F}_Q\|_{\max} = \cos^{-1}\sqrt{q + \frac{1-q}{n^2}}. \tag{59}$$

Suppose that we run this channel k times. We can either run (1) a unitary implementation of $\mathcal{F}_Q$ k times (with the ancilla reset before the start of a new run) or (2) a unitary implementation of

$$\mathcal{F}_Q^{(k)}(\rho) \triangleq \underbrace{\mathcal{F}_Q \circ \mathcal{F}_Q \circ \cdots \circ \mathcal{F}_Q}_{k \text{ times}}(\rho) = q^k\rho + (1-q^k)\frac{I}{n}.$$

For case (1), since each run is executed independently of each other, the total time energy is $k\|\mathcal{F}_Q\|_{\max}$.

For case (2), applying Eq. (52) gives

$$\left\|\mathcal{F}_Q^{(k)}\right\|_{\max} = \cos^{-1}\sqrt{q^k + \frac{1-q^k}{n^2}}.$$

Using Taylor series expansion, we approximate $\cos^{-1}\sqrt{x} \approx \sqrt{1-x}$ for $x \approx 1$ and

$$q^k + \frac{1-q^k}{n^2} \approx 1 - k\left(1 - \frac{1}{n^2}\right)(1-q)$$

for $q \approx 1$. Thus, we have

$$\left\|\mathcal{F}_Q^{(k)}\right\|_{\max} \approx \sqrt{k}\sqrt{1 - \left(q + \frac{1-q}{n^2}\right)},$$

which implies that

$$\left\|\mathcal{F}_Q^{(k)}\right\|_{\max} \approx \sqrt{k}\,\|\mathcal{F}_Q\|_{\max}. \tag{60}$$

This means that considering all k channels together saves time-energy resource by a factor of $\sqrt{k}$ compared to separately running the channels when the noise is small ($q \approx 1$).

IX. CONCLUDING REMARKS

In this paper we extend the time-energy measure proposed by Chau [16] to general quantum processes. This measure is a good indicator of the time-energy tradeoff of a quantum process. Essentially, a large time-energy value suggests that the quantum process takes a longer time or more energy to run. Here, we prove lower and upper bounds for the sum time energy and max time energy. We also prove the optimal max time energy for a class of channels, which includes the depolarizing channel. A consequence of this result is that erasing information takes more time-energy resource in the quantum setting than in the classical setting.

A related concept about erasure and energy is Landauer's principle [20], which puts lower limits on the energy dissipated to the environment in erasing (qu)bits. There is a difference between the erasure considered here and the erasure of Landauer's principle. First, we erase information by making the initial pure state more mixed, whereas Landauer's principle concerns resetting a possibly mixed state to a standard pure state. Second, Landauer's principle concerns erasure in the thermodynamic setting where temperature plays a key role. Third, tradeoff between time and energy is implicated in our approach.

For future investigation, it is instructive to obtain the time energy for various quantum processes such as some standard gates or algorithms, to consider this time-energy measure in the thermodynamic setting, and to explore deeper operational meaning about this measure.

ACKNOWLEDGMENTS

We thank H.-K. Lo and X. Ma for enlightening discussion. This work is supported in part by RGC under Grant 700712P of the HKSAR Government.

APPENDIX: DERIVATION OF THE MATRIX $\tilde{U}$

Here we derive Eq. (22). We are given that the eigenvalues of $\tilde{U}$ are $\exp(i\theta_1)$ and $\exp(i\theta_2)$. The constraints are that (1) the chord connecting $\exp(i\theta_1)$ and $\exp(i\theta_2)$ intersects $\langle a|b\rangle$ and (2) $\tilde{U}|a\rangle = |b\rangle$.

$\tilde{U}$ has the following decomposition:

$$\tilde{U} = \exp(i\theta_1)|\tilde{u}_1\rangle\langle\tilde{u}_1| + \exp(i\theta_2)|\tilde{u}_2\rangle\langle\tilde{u}_2|, \tag{A1}$$

where $|\tilde{u}_j\rangle$ is the eigenvector corresponding to $\exp(i\theta_j)$. We assume that $|\tilde{u}_j\rangle, j = 1,2$ take the following forms:

$$\tilde{u}_1 = \begin{bmatrix} \sqrt{z} \\ e^{ix}\sqrt{1-z} \end{bmatrix}, \quad \tilde{u}_2 = \begin{bmatrix} -e^{-ix}\sqrt{1-z} \\ \sqrt{z} \end{bmatrix}, \tag{A2}$$

which are expressed in the basis $\{|a\rangle, |a^{\perp}\rangle\}$. Note that $\langle\tilde{u}_1|\tilde{u}_2\rangle = 0$. Constraint (2) implies that

$$|b\rangle = \tilde{U}|a\rangle = [ze^{i\theta_1} + (1-z)e^{i\theta_2}]|a\rangle + [\sqrt{z(1-z)}e^{ix}(e^{i\theta_1} - e^{i\theta_2})]|a^{\perp}\rangle.$$

Comparing this with

$$|b\rangle = \langle a|b\rangle|a\rangle + \sqrt{1 - |\langle a|b\rangle|^2}|a^{\perp}\rangle$$

obtained from Eq. (21), we require that

$$re^{i\gamma} = ze^{i\theta_1} + (1-z)e^{i\theta_2}, \tag{A3}$$

$$\sqrt{1-r^2} = \sqrt{z(1-z)}e^{ix}(e^{i\theta_1} - e^{i\theta_2}), \tag{A4}$$

where $re^{i\gamma} = \langle a|b\rangle$ expressed in the polar form. We need to verify that both equations hold for some $0 \leqslant z \leqslant 1$ and $x \in \mathbb{R}$. Constraint (1) means that there exists some z that Eq. (A3) holds. We take this z as fixed and find x so that Eq. (A4) holds, which can occur only if $e^{ix}(e^{i\theta_1} - e^{i\theta_2}) \in \mathbb{R}_+$. We make an ansatz for x by setting

$$e^{ix} = i(-1)^s e^{-i\frac{\theta_1+\theta_2}{2}}, \tag{A5}$$

giving

$$e^{ix}(e^{i\theta_1} - e^{i\theta_2}) = 2(-1)^{s+1}\sin\frac{\theta_1 - \theta_2}{2}, \tag{A6}$$

where $s = 0,1$ is chosen so that this term is non-negative. We square both sides of Eq. (A4) and compare both sides. For the LHS, r^2 can be obtained from Eq. (A3) as follows:

$$r^2 = |ze^{i\theta_1} + (1-z)e^{i\theta_2}|^2 = z^2 + (1-z)^2 + 2z(1-z)\cos(\theta_1 - \theta_2). \tag{A7}$$

Squaring the RHS of Eq. (A4) gives

$$4z(1-z)\sin^2\frac{\theta_1 - \theta_2}{2} = 2z(1-z)[1 - \cos(\theta_1 - \theta_2)],$$

which can be checked to be equal to $1 - r^2$ using Eq. (A7). Therefore, Eqs. (A3) and (A4) hold. Finally, Eqs. (A1)–(A6) together give Eq. (22).

[1] S. Lloyd, Nature (London) **406**, 1047 (2000).
[2] L. Mandelstam and I. Tamm, J. Phys. USSR **9**, 249 (1945).
[3] K. Bhattacharyya, J. Phys. A **16**, 2993 (1983).
[4] J. Anandan and Y. Aharonov, Phys. Rev. Lett. **65**, 1697 (1990).
[5] A. Uhlmann, Phys. Lett. A **161**, 329 (1992).
[6] L. Vaidman, Am. J. Phys. **60**, 182 (1992).
[7] P. Pfeifer, Phys. Rev. Lett. **70**, 3365 (1993).
[8] N. Margolus and L. B. Levitin, in *Proceedings of Fourth Workshop on Physics and Computation*, edited by T. Toffoli, M. Biafore, and J. Leão (New England Complex Systems Institute, Boston, 1996), pp. 208–211.
[9] N. Margolus and L. B. Levitin, Physica D **120**, 188 (1998).
[10] H. F. Chau, Phys. Rev. A **81**, 062133 (2010).
[11] V. Giovannetti, S. Lloyd, and L. Maccone, Phys. Rev. A **67**, 052109 (2003).
[12] V. Giovannetti, S. Lloyd, and L. Maccone, Europhys. Lett. **62**, 615 (2003).
[13] C. Zander, A. R. Plastino, A. Plastino, and M. Casas, J. Phys. A **40**, 2861 (2007).
[14] M. M. Taddei, B. M. Escher, L. Davidovich, and R. L. de Matos Filho, Phys. Rev. Lett. **110**, 050402 (2013).
[15] A. del Campo, I. L. Egusquiza, M. B. Plenio, and S. F. Huelga, Phys. Rev. Lett. **110**, 050403 (2013).
[16] H. F. Chau, Quant. Inf. Comput. **11**, 0721 (2011).
[17] We adopt the fidelity definition $F(\rho,\sigma) = (\mathrm{Tr}\sqrt{\rho^{1/2}\sigma\rho^{1/2}})^2$ for two quantum states ρ and σ.
[18] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2000).
[19] C. King, IEEE Trans. Inf. Theory **49**, 221 (2003).
[20] R. Landauer, IBM J. Res. Dev. **5**, 183 (1961).