Upper bounds on entangling rates of bipartite Hamiltonians

Sergey Bravyi

IBM Watson Research Center, Yorktown Heights, New York 10598, USA
(Received 11 October 2007; published 21 November 2007)

We discuss upper bounds on the rate at which unitary evolution governed by a nonlocal Hamiltonian can generate entanglement in a bipartite system. Given a bipartite Hamiltonian H coupling two finite dimensional particles A and B, the entangling rate is shown to be upper bounded by $c \log(d) \|H\|$, where d is the smallest dimension of the interacting particles, $\|H\|$ is the operator norm of H, and c is a constant close to 1. Under certain restrictions on the initial state we prove an analogous upper bound for the ancilla-assisted entangling rate with a constant c that does not depend upon dimensions of local ancillas. The restriction is that the initial state has at most two distinct Schmidt coefficients (each coefficient may have arbitrarily large multiplicity). Our proof is based on analysis of a mixing rate—a functional measuring how fast entropy can be produced if one mixes a time-independent state with a state evolving unitarily.

I. INTRODUCTION

Consider two remote parties Alice and Bob controlling finite-dimensional quantum systems A and B. Suppose A and B interact with each other according to a time-independent Hamiltonian H. Unless H is a sum of local Hamiltonians, a unitary evolution e^{iHt} is a nonlocal operation capable of creating entanglement between A and B. The goal of the present paper is to get an upper bound on the rate at which the entanglement between Alice and Bob can increase or decrease as a function of time.

We consider ancilla-assisted entangling, see Fig. 1. It means that Alice's laboratory consists of two subsystems: A and a, such that the Hamiltonian H acts only on the subsystem A. Similarly, Bob's laboratory is partitioned into B and b, where H acts only on the subsystem B. The subsystems a and b are the local ancillas held by Alice and Bob. In the ancilla-assisted entangling Alice and Bob start from a pure state $|\Psi\rangle$ of the composite system $aABb$ which may be already entangled.

Time evolution of the composite system $aABb$ is described by a unitary operator $U(t) = I_a \otimes e^{iH_{AB}t} \otimes I_b$. Thus the joint state of Alice and Bob always remains pure. Accordingly, the entanglement between Alice and Bob at any time t can be quantified by entanglement entropy

$$S(aA) = - \operatorname{Tr} \rho_{aA}(t) \log \rho_{aA}(t), \tag{1}$$

where $\rho_{aA}(t) = \operatorname{Tr}_{Bb} U(t) |\Psi\rangle\langle\Psi| U(t)^\dagger$. The quantity we are interested in is *entangling rate*,

$$\Gamma(\Psi, H) = \left. \frac{dS(aA)}{dt} \right|_{t=0}. \tag{2}$$

Understanding properties of the entangling rate is crucial for optimal generation of entanglement [1], computing capacities of bidirectional quantum communication channels [2], and for describing dynamics of entanglement in quantum spin lattice models [3].

Our goal is to get an upper bound on $\Gamma(\Psi, H)$ that would not explicitly depend upon dimensions of the local ancillas a and b. Upper bounds of this kind can be easily generalized via the Trotter decomposition to arbitrary multipartite Hamil-

tonians decomposable into a sum of few-party interactions. Throughout the paper we shall use notations $|a|$, $|A|$, $|B|$, and $|b|$ for the dimension of the Hilbert space describing the subsystem a, A, B, and b, respectively.

A. Previous work

The maximal entangling rate $\Gamma(H) = \max_\Psi \Gamma(\Psi, H)$ has been studied by many authors for the case when A and B are qubits. The optimal state Ψ maximizing $\Gamma(\Psi, H)$ has been found by Dür *et al.* [1] for a general Hamiltonian H assuming $|a| = |b| = 1$. These authors have also observed that for some Hamiltonians local ancillas are capable of increasing the maximal entangling rate $\Gamma(H)$. A powerful technique of getting upper bounds on $\Gamma(H)$ for arbitrarily large a and b was proposed by Childs *et al.* [4]. It was used to identify a subclass of two-qubit Hamiltonians, including the Ising interaction, for which the maximum $\Gamma(H)$ does not depend upon $|a|$ and $|b|$ (and thus can be achieved without local ancillas). An upper bound on the entangling rate of the Ising interaction has been also obtained by Cirac *et al.* [5] by showing that the time evolution with the Ising Hamiltonian for time t can be implemented by local operations and classical communication (LOCC) protocol consuming $O(|t|)$ *e*-bits of preshared entanglement.

Some progress has been also achieved for larger dimensions $|A|$ and $|B|$. The technique of [4] has been generalized by Wang and Sanders [6] to prove that $\Gamma(H) \le \beta \approx 1.9123$ for any product Hamiltonian $H = H_A \otimes H_B$ if H_A, H_B have eigenvalues ± 1. This result applies to arbitrary dimensions $|a|$ and $|b|$. An upper bound $\Gamma(H) \le \beta \|H\|$ for the ancilla-assisted case has been proved for arbitrary bipartite product Hamil-
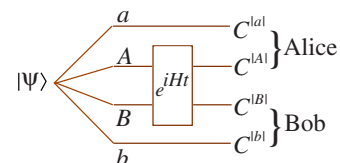

FIG. 1. (Color online) Ancilla-assisted entangling.

tonians by Childs, Leung, and Vidal [7]. Finally, it was shown by Bennett *et al.* [2] that $\Gamma(H) \le cd^4 \|H\|$, where $d = \min(|A|, |B|)$ and $c = O(1)$ does not depend upon $|a|$ and $|b|$. This result relies on a decomposition of an arbitrary bipartite Hamiltonian into a sum of product Hamiltonians. Thus for fixed d and $\|H\|$ the entangling rate has a constant upper bound independent of how large dimensions $|a|$ and $|b|$ are. The authors of [2] also proved that the supremum of $\Gamma(H)$ over all dimensions of local ancillas $|a|, |b|$ coincides with the asymptotic capacity of H to generate entanglement by any protocol in which unitary evolution with H is interspersed with LOCC. It is unknown whether the supremum over $|a|$ and $|b|$ can be actually achieved for finite dimensional ancillas.

It is not known whether $\Gamma(H) = \Gamma(-H)$, i.e., whether the maximal entangling and disentangling rates of a given Hamiltonian H coincide. The results by Linden, Smolin, and Winter [8] applicable to finite unitary operators suggest that this equality might be wrong.

B. Summary of results

We start from observing that the unitary evolution with any Hamiltonian H_{AB} cannot increase (decrease) the entanglement $S(aA)$ by more than $2 \log(d)$, where $d = \min(|A|, |B|)$, see [2]. This property can be called *small total entangling*, as it says that the total increase (decrease) of entanglement throughout the unitary evolution remains bounded as the dimensions of local ancillas $|a|$ and $|b|$ go to infinity. Is there an analog of the small total entangling property for infinitely small time intervals? The following conjecture was initially proposed by Kitaev [9]. We call it small incremental entangling.

Small incremental entangling (SIE): There exists a constant $c = O(1)$ such that

$$\Gamma(\Psi, H) \le c \|H\| \log(d), \quad d \equiv \min(|A|, |B|) \quad (3)$$

for all dimensions $|a|, |b|$ and for all states $|\Psi\rangle$ of a composite system $aABb$.

It is not known whether SIE is true or false. Our first result is a proof of SIE for the special case when $|\Psi\rangle$ has at most two distinct Schmidt coefficients with respect to a partition $aA \otimes Bb$ (each Schmidt coefficient may have arbitrarily large multiplicity). Our proof yields the constant $c = 24$ although the actual value of c might be much smaller. The constraint on $|\Psi\rangle$ serves technical purposes and was introduced in order to make the problem tractable. It should be mentioned that the upper bound Eq. (3) can be confirmed by straightforward calculation if $|a| = |b| = 1$ (no local ancillas), see Sec. II. In this case the optimal pair (Ψ, H) can be found explicitly. It turns out that the optimal state $|\Psi\rangle$ has only two distinct Schmidt coefficients and thus falls into the category that we consider, see Sec. II for details. We do not know whether the optimal state has only two Schmidt coefficients in the ancilla-assisted case.

It is known that in some cases local ancillas and preshared entanglement can lead to counterintuitive effects, such as entanglement embezzling [10] or locking of classical correlations [11]. In particular, it was demonstrated by DiVin-

cenzo *et al.* [11] that sending a single qubit from Alice to Bob can increase their classical mutual information by an arbitrarily large amount in the presence of local ancillas. Thus SIE may be violated if locking effects also occur in the infinitesimal unitary transformations if one measures correlations by entanglement entropy. For future references let us give an explicit expression for the entangling rate that can be easily obtained by computing the derivative in Eq. (2).

$$\Gamma(\Psi, H) = -i \operatorname{Tr}(I_a \otimes H_{AB}[\rho_{aAB}, \log \rho_{aA} \otimes I_B]). \quad (4)$$

Our proof of SIE goes by getting an upper bound on a *mixing rate*. In order to define a mixing rate, consider a probabilistic ensemble of (mixed) states $\mathcal{E} = \{p_\alpha, \rho_\alpha\}_{\alpha=0,1}$ defined on a finite-dimensional Hilbert space (lacking any tensor product structure). Let $\rho = p_0 \rho_0 + p_1 \rho_1$ be the average state corresponding to $\mathcal{E}$. For any Hamiltonian H define a time dependent state $\rho(t) = p_0 \rho_0 + p_1 e^{iHt} \rho_1 e^{-iHt}$. Define a mixing rate as

$$\Lambda(\mathcal{E}, H) = \left. \frac{dS(\rho(t))}{dt} \right|_{t=0}. \quad (5)$$

Here $S(\rho(t))$ is the von Neumann entropy of $\rho(t)$. Basic properties of the von Neumann entropy imply that $\bar{S} \le S(\rho(t)) \le \bar{S} + h(p_0, p_1)$, where $\bar{S} = p_0 S(\rho_0) + p_1 S(\rho_1)$ is the average entropy for the ensemble $\mathcal{E}$ and $h(p_0, p_1) = -p_0 \log(p_0) - p_1 \log(p_1)$ is the Shannon entropy. This property can be called *small total mixing* as it says that the total increase (decrease) of the entropy throughout the unitary evolution goes to zero as one of the probabilities p_0 or $p_1 = 1 - p_0$ goes to zero. Is there an analog of the small total mixing property for infinitely small time intervals? A naive generalization would be as follows.

Small incremental mixing (SIM): There exists a constant $c' = O(1)$ such that

$$\Lambda(\mathcal{E}, H) \le c' \|H\| h(p_0, p_1), \quad (6)$$

where

$$h(p_0, p_1) = -p_0 \log(p_0) - p_1 \log(p_1)$$

for any probabilistic ensemble $\mathcal{E} = \{p_\alpha, \rho_\alpha\}_{\alpha=0,1}$.

Here it is meant that the constant c' is independent of the dimension of the Hilbert space. It is not known whether SIM is true or false. Although SIM might seem completely unrelated to ancilla-assisted entangling and SIE, it turns out that SIM is a stronger version of SIE. More strictly, we prove that SIM with a constant c' implies SIE with a constant $c = 4c'$, see Sec. III.

Our second result is a proof of SIM for the special case when $\rho = p_0 \rho_0 + p_1 \rho_1$ has at most two distinct eigenvalues (each eigenvalue may have arbitrarily large multiplicity), see Sec. IV. We shall refer to eigenvalues obeying this constraint as a *binary spectrum*. Our proof yields $c' = 6$, although the actual value of c' might be smaller.

The connection between SIE to SIM described in Sec. III has a peculiar property that the number of distinct eigenvalues of ρ in SIM is exactly the same as the number of distinct Schmidt coefficients of the initial state $|\Psi\rangle$ in SIE. Thus a proof of SIM for the case when ρ has a binary spectrum

implies SIE for the case when the initial state $|\Psi\rangle$ has only two distinct Schmidt coefficients. In the general case we prove that the mixing rate of any ensemble $\mathcal{E}$ has an upper bound $\Lambda(\mathcal{E}, H) = O(\|H\|)$, so it does not explicitly depend on the dimension of the Hilbert space, see Sec. IV. For future reference let us give an explicit expression for the mixing rate:

$$\Lambda(\mathcal{E}, H) = -ip_1 \operatorname{Tr}(H[\rho_1, \log \rho]), \quad \rho = p_0 \rho_0 + p_1 \rho_1. \quad (7)$$

The paper is organized as follows. In Sec. II we find the maximal entangling rate and the optimal pair (Ψ, H) for the case when Alice and Bob do not use local ancillas. Section III proves that SIM implies SIE. Section IV contains the main results of the paper. It proves SIM for the case when ρ has a binary spectrum. Section V reports results of numerical maximization aimed at verifying SIM. The data obtained in numerical simulations are consistent with SIM.

II. MAXIMAL ENTANGLING RATE IN THE ABSENCE OF ANCILLAS

The maximal entangling rate can be easily found in the absence of local ancillas, i.e., when $|a| = |b| = 1$. In this case we can always write the initial state of Alice and Bob using the Schmidt decomposition $|\Psi\rangle = \sum_{j=1}^{d} \sqrt{p_j} |j_A\rangle \otimes |j_B\rangle$ with $d = \min(|A|, |B|)$. Denote $\rho_A = \sum_{j=1}^{d} p_j |j\rangle\langle j|$ the reduced density matrix of Alice. Using the general formula Eq. (4) for the entangling rate one gets

$$\Gamma(\Psi, H) = -\operatorname{Tr} \dot\rho_A \log \rho_A = -i \operatorname{Tr}(H[|\Psi\rangle\langle\Psi|, \log \rho_A \otimes I_B]). \quad (8)$$

Since the entangling rate is a linear function of H we can assume that $\|H\| = 1$. Using the fact that for any Hermitian operator X

$$\max_{H:\|H\|=1} \operatorname{Tr}(HX) = \operatorname{Tr}|X| \equiv \|X\|_1, \quad (9)$$

one can carry out the maximization over H,

$$\Gamma(\Psi) := \max_{H:\|H\|=1} \Gamma(\Psi, H) = \|[\log \rho_A \otimes I_B, |\Psi\rangle\langle\Psi|]\|_1. \quad (10)$$

For any vectors $|\Psi\rangle, |\Phi\rangle$ one has the following identity:

$$\||\Phi\rangle\langle\Psi| - |\Psi\rangle\langle\Phi|\|_1 = 2\sqrt{\langle\Psi|\Psi\rangle\langle\Phi|\Phi\rangle - |\langle\Psi|\Phi\rangle|^2}. \quad (11)$$

Substituting $|\Phi\rangle = (\log \rho_A \otimes I_B)|\Psi\rangle$ one gets

$$\Gamma(\Psi) = 2\sqrt{F(p)},$$

where

$$F(p) = \sum_{j=1}^{d} p_j \log^2 p_j - \left(\sum_{j=1}^{d} p_j \log p_j\right)^2. \quad (12)$$

We can assume that all $p_j > 0$ (otherwise replace d by $d' < d$). Then the maximum of $F(p)$ can be found by solving extremal point equations $\partial F/\partial p_j = 0$. After simple algebra one gets

$$-\log(2p_j) = S \pm \sqrt{\gamma^2 + S^2},$$

$$S = -\sum_{j=1}^{d} p_j \log p_j, \quad \gamma = \frac{1}{\ln 2}. \quad (13)$$

(Here log and ln stand for base two and natural logarithm.) The equality for the minus sign in the right-hand side (rhs) of Eq. (13) is possible only for one value of j, since it implies $p_j \geq 1/2$. Let us agree that $p_1 \geq 1/2$. Then $-\log(2p_j) = S + \sqrt{\gamma^2 + S^2}$ for all $j = 2, \ldots, d$, that is the state ρ_A must have a binary spectrum with multiplicities $1, d-1$. Introduce a variable λ such that

$$p_1 = \lambda, \quad p_2 = \cdots = p_d = \frac{1-\lambda}{d-1}, \quad \frac{1}{2} \leq \lambda \leq 1. \quad (14)$$

Let $|\Psi_\lambda\rangle$ be the state with Schmidt coefficients $p_1, \ldots, p_d$. Using Eq. (12) the maximal entangling rate with the initial state $|\Psi_\lambda\rangle$ can be written as

$$\Gamma(\Psi_\lambda) = 2\sqrt{\lambda(1-\lambda)}\log\left(\frac{\lambda(d-1)}{1-\lambda}\right). \quad (15)$$

The optimal value $\lambda = \lambda(d)$ and the optimal entangling rate $\Gamma_d = \Gamma(\Psi_{\lambda(d)})$ have to be found by maximizing Eq. (15) over $1/2 \leq \lambda \leq 1$. For example, if $d = 2$, numerical maximization yields $\lambda(2) \approx 0.9168$ and $\Gamma_2 \approx 1.9123$. It coincides with the maximal entangling rate of product two-qubit Hamiltonians found in [1,4]. It follows that under normalization condition $\|H\| = 1$ product two-qubit Hamiltonians are capable of generating entanglement with the largest rate.

One can easily infer from Eq. (15) that $\lambda(d) \approx 1/2$ and $\Gamma_d \approx \log d$ for sufficiently large d. It proves that $\Gamma(\Psi, H) = O(\|H\| \log d)$ in the absence of local ancillas. Moreover, in the limit of large d one can explicitly write down the optimal state $|\Psi_{\lambda(d)}\rangle$ and the optimal Hamiltonian H_d. Namely,

$$|\Psi_{\lambda(d)}\rangle \approx \frac{1}{\sqrt{2}}|1\rangle \otimes |1\rangle + \frac{1}{\sqrt{2}}|\Phi^+\rangle,$$

$$|\Phi^+\rangle = \frac{1}{\sqrt{d-1}}\sum_{j=2}^{d}|j\rangle \otimes |j\rangle, \quad (16)$$

$$H_d \approx -i(|\Phi^+\rangle\langle 1,1| - |1,1\rangle\langle\Phi^+|). \quad (17)$$

Thus the optimal state is a superposition of a product state and a maximally entangled state with locally orthogonal supports. The optimal Hamiltonian is a generator for a rotation in the corresponding two-dimensional subspace.

For a finite d the optimal value $\lambda(d)$ and the corresponding entangling rate Γ_d can be found numerically, see Fig. 2. For large d one has $\lambda(d) \approx (1/2)(1 + 1/\ln d)$, while entanglement entropy of $|\Psi_{\lambda(d)}\rangle$ scales as $S_d \approx (1/2)\log d$.

III. SIM IMPLIES SIE

In this section we assume that SIM is true and show that this assumption implies SIE. Consider ancilla-assisted entan-
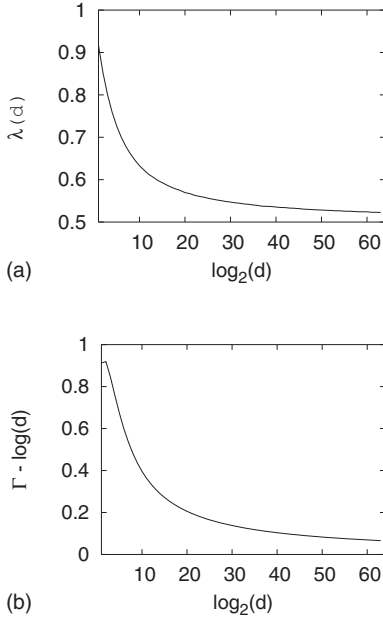
FIG. 2. Optimal λ (a) and entangling rate Γ (b) as a function of $\log d$.

gling with some fixed dimensions $|a|, |A|, |B|, |b|$ and some initial state $|\Psi\rangle$. Let us assume that $|B| \leq |A|$. Then SIE is equivalent to an upper bound

$$\Gamma(\Psi, H) \leq c \|H\| \log|B|. \qquad (18)$$

Since the rhs of this inequality does not depend on $|a|$ and $|A|$, it is enough to prove Eq. (18) in the case $|a|=1$ since we can always extend A to $A \otimes a$. Now we have a tripartite system ABb initially prepared in a pure state $|\Psi\rangle$ and evolving under a Hamiltonian H acting on A and B. Using Eq. (4) for the entangling rate one gets

$$\Gamma(\Psi, H) = -i \, \mathrm{Tr}(H[\rho_{AB}, \log \rho_A \otimes I_B])$$
$$= -i \, \mathrm{Tr}(H[\rho_{AB}, \log \tau_{AB}]), \qquad (19)$$

where ρ_A and ρ_{AB} are the reduced states of $|\Psi\rangle$, and a state τ_{AB} is defined as

$$\tau_{AB} = \rho_A \otimes \frac{I_B}{|B|}. \qquad (20)$$

In order to define a probabilistic ensemble $\mathcal{E}$ needed for a reduction to SIM we shall need the following lemma.

Lemma 1. For any mixed state ρ_{AB} there exists a mixed state μ_{AB} such that

$$\rho_A \otimes \frac{I_B}{|B|} = \frac{1}{|B|^2}\rho_{AB} + \left(1 - \frac{1}{|B|^2}\right)\mu_{AB}. \qquad (21)$$

Proof. Since the partial trace is a linear operation, it suffices to prove the lemma for the case when $\rho_{AB}=|\psi\rangle\langle\psi|$ is a pure state. The statement of the lemma is then equivalent to inequality

$$|\psi\rangle\langle\psi| \leq |B|\rho_A \otimes I_B. \qquad (22)$$

Let d be the Schmidt rank of $|\psi\rangle$. Obviously, $d \leq \min(|A|, |B|)$. Consider a state $|I\rangle = \Sigma_{j=1}^{d}|j_A\rangle \otimes |j_B\rangle$, where $\{|j_A\rangle\}$ and $\{|j_B\rangle\}$ are the local bases of A and B that diagonalize ρ_A and ρ_B. Then

$$|\psi\rangle\langle\psi| = (\rho_A^{1/2} \otimes I_B)|I\rangle\langle I|(\rho_A^{1/2} \otimes I_B).$$

Taking into account that $\langle I|I\rangle = d$, one concludes that $|I\rangle\langle I| \leq dI_{AB}$. Multiplying this inequality by $\rho_A^{1/2} \otimes I_B$ on the left and on the right we get $|\psi\rangle\langle\psi| \leq d\rho_A \otimes I_B$ which implies Eq. (22). ∎

Define an ensemble of states $\mathcal{E} = \{p_\alpha, \tau_\alpha\}_{\alpha=0,1}$ such that $\tau_0 = \mu_{AB}$, $\tau_1 = \rho_{AB}$, $p_0 = 1-|B|^{-2}$, and $p_1 = |B|^{-2}$. Here μ_{AB} is the state that appears in the decomposition Eq. (21). This ensemble has the average state $\tau = p_0\tau_0 + p_1\tau_1 = \rho_A \otimes I_B/|B|$. Let H be a Hamiltonian that appears in Eq. (19). Assuming that SIM is true one gets

$$\Lambda(\mathcal{E}, H) \leq c' h(|B|^{-2}, 1-|B|^{-2})\|H\| \leq 4c'|B|^{-2}(\log|B|)\|H\|. \qquad (23)$$

[Note that for any $x \leq 1/2$ one has $h(x, 1-x) \leq 2x|\log x|$.] On the other hand, Eq. (7) for the mixing rate leads to

$$\Lambda(\mathcal{E}, H) = -ip_1 \, \mathrm{Tr}(H[\tau_1, \log \tau]) = -i|B|^{-2}\mathrm{Tr}(H[\rho_{AB}, \log \tau_{AB}]). \qquad (24)$$

Combining Eqs. (23) and (24) one arrives to

$$-i \, \mathrm{Tr}(H[\rho_{AB}, \log \tau_{AB}]) \leq 4c'(\log|B|)\|H\|.$$

Comparing it with Eq. (19) we infer that $\Gamma(\Psi, H) \leq 4c'(\log|B|)\|H\|$. We have proved SIE with a constant $c = 4c'$, see Eq. (18).

The reduction Eq. (20) has a peculiar property that the number of distinct eigenvalues of τ_{AB} and ρ_A are the same. Since we have used identification $A = aA$, the number of distinct eigenvalues of ρ_A coincides with the number of distinct Schmidt coefficients of $|\Psi\rangle$. Thus if we can prove SIM with $\rho(0)$ having only two distinct eigenvalues, we will prove SIE for the case when $|\Psi\rangle$ has only two distinct Schmidt coefficients.

IV. UPPER BOUNDS ON THE MIXING RATE

Let us start from proving a weaker (compared to SIM) upper bound on the mixing rate.

Lemma 2. For any ensemble $\mathcal{E}$ and Hamiltonian H one has

$$\Lambda(\mathcal{E}, H) \leq 2\|H\|. \qquad (25)$$

Proof. Suppose $\mathcal{E} = \{p_\alpha, \rho_\alpha\}_{\alpha=0,1}$, where the states ρ_0, ρ_1 and a Hamiltonian H are defined on a D-dimensional Hilbert space. Consider ancilla-assisted entangling protocol shown in Fig. 1 with $a = 1$, $A = D$, $B = 2$, and $b = D$. Choose a Hamiltonian $\tilde{H}$ coupling A and B as

$$\tilde{H} = H_A \otimes |1\rangle\langle 1|_B.$$

Choose the initial state as

$$|\Psi\rangle = \sqrt{p_0}|0\rangle_B \otimes |\phi_0\rangle_{Ab} + \sqrt{p_1}|1\rangle_B \otimes |\phi_1\rangle_{Ab},$$

where $|\phi_\alpha\rangle \in \mathbb{C}^D \otimes \mathbb{C}^D$ is a purification of ρ_α. Then the state $|\Psi\rangle$ evolves in time as

$$|\Psi(t)\rangle = \sqrt{p_0}|0\rangle_B \otimes |\phi_0\rangle_{Ab} + \sqrt{p_1}|1\rangle_B \otimes (e^{iH_A t} \otimes I_b)|\phi_1\rangle_{Ab}.$$

Accordingly, $\rho_A(t) = p_0\rho_0 + p_1 e^{iHt}\rho_1 e^{-iHt}$. Therefore $\Lambda(\mathcal{E}, H) = \Gamma(\Psi, \tilde{H})$. Now we can use the upper bound on the entangling rate obtained in [6,7] for product Hamiltonians, namely $\Gamma(\Psi, \tilde{H}) \leq \beta\|\tilde{H}\| = \beta\|H\|$, where $\beta \approx 1.9123 \leq 2$. The lemma is proved. ∎

This lemma implies that the whole difficulty of proving SIM (if it is true) concerns the limiting cases $p_0 \to 0$ or $p_0 \to 1$. Note also that $\Lambda(\mathcal{E}, H) = \Lambda(\mathcal{E}^T, -H)$, where ensemble $\mathcal{E}^T$ is obtained from $\mathcal{E}$ by interchanging p_0, ρ_0 with p_1, ρ_1. Thus we can assume that $0 \leq p_1 \leq 1/2$.

In the following we shall represent an ensemble $\mathcal{E} = \{p_\alpha, \rho_\alpha\}_{\alpha=0,1}$ using the data (p, ρ, Π), where

$$p = p_1, \quad \rho = (1-p)\rho_0 + p\rho_1, \quad \Pi = p\rho^{-1/2}\rho_1\rho^{-1/2}. \quad (26)$$

One can easily check that a triple (p, ρ, Π) represents some ensemble $\mathcal{E}$ if and only if

$$0 \leq p \leq 1, \quad 0 \leq \Pi \leq I, \quad \mathrm{Tr}\,\Pi\rho = p. \quad (27)$$

Indeed, the only nontrivial statement is that $\Pi \leq I$. It can be obtained from inequality $p_1\rho_1 \leq \rho$ by multiplying it by $\rho^{-1/2}$ on the left and on the right. Using the representation Eq. (26) one can rewrite the mixing rate as

$$\Lambda(\mathcal{E}, H) = -i\,\mathrm{Tr}(H\rho^{1/2}[\Pi, \log\rho]\rho^{1/2}), \quad (28)$$

see Eq. (7). Given $\mathcal{E}$, the optimal Hamiltonian H maximizing the mixing rate can be found using Eq. (9). Thus we have

$$\Lambda(\mathcal{E}) := \max_{H:\|H\|=1} \Lambda(\mathcal{E}, H) = \|\rho^{1/2}[\Pi, \log\rho]\rho^{1/2}\|_1. \quad (29)$$

It is worth mentioning that for given ρ the upper bound Eq. (6) is true for sufficiently small p. Indeed, applying the triangle inequality for the trace norm to Eq. (29) one gets $\Lambda(\mathcal{E}) \leq 2\|\rho^{1/2}\Pi\rho^{1/2}\log\rho\|_1$. Since $\|AB\|_1 \leq \|A\|\|B\|_1$ for any operators A, B it follows that $\Lambda(\mathcal{E}) \leq 2\|\log\rho\|\|\rho^{1/2}\Pi\rho^{1/2}\|_1$. If p is smaller than the smallest eigenvalue of ρ, one has $\|\log\rho\| \leq |\log p|$ and thus $\Lambda(\mathcal{E}) \leq 2|\log p|\mathrm{Tr}(\Pi\rho) = 2p|\log p| \leq 2h(p, 1-p)$.

In the rest of the section we prove SIM under the assumption that ρ has binary spectrum,

$$\rho = \sum_{j=1}^{D} \rho_j |j\rangle\langle j|, \quad \rho_j = \begin{cases} \lambda_1 & \text{if } j = 1, \ldots, m \\ \lambda_2 & \text{if } j = m+1, \ldots, D \end{cases}, \quad (30)$$

where $\lambda_1 \geq \lambda_2$ and m can be an arbitrary integer between 1 and D. Define projectors $R = \sum_{j=1}^{m} |j\rangle\langle j|$ and $R^\perp = I - R$ such that

$$\rho = \lambda_1 R + \lambda_2 R^\perp,$$

$$\rho^{1/2} = \sqrt{\lambda_1} R + \sqrt{\lambda_2} R^\perp,$$

$$\log\rho = \log\lambda_1 R + \log\lambda_2 R^\perp.$$

After some algebra one gets the following identity:

$$\rho^{1/2}[\Pi, \log\rho]\rho^{1/2} = \log\left(\frac{\lambda_1}{\lambda_2}\right)\sqrt{\lambda_1\lambda_2}[\Pi, R]. \quad (31)$$

In order to upper bound the trace norm of the commutator $[\Pi, R]$ we shall use the following.

Lemma 3. (Hölder inequality). Let R be a projector and Π be a positive semidefinite operator. Then

$$\|[\Pi, R]\|_1 \leq 2\sqrt{\mathrm{Tr}(\Pi R)\mathrm{Tr}(\Pi R^\perp)}. \quad (32)$$

Lemma 4. Let R be a projector of rank m and Π be a Hermitian operator such that $0 \leq \Pi \leq I$. Then there exists a Hermitian operator Π' such that

(i) $0 \leq \Pi' \leq \Pi$.
(ii) $[\Pi, R] = [\Pi', R]$.
(iii) $\mathrm{Tr}\,\Pi' \leq m$.

We shall postpone the proof of the two lemmas above until the end of the section. Let us apply Lemma 3 to the commutator in Eqs. (31) where Π is replaced by Π' from Lemma 4. Then we can upper bound the mixing rate in Eq. (29) as

$$\Lambda(\mathcal{E}) \leq 2\log\left(\frac{\lambda_1}{\lambda_2}\right)\sqrt{\lambda_1\lambda_2\;\mathrm{Tr}(\Pi' R)\mathrm{Tr}(\Pi' R^\perp)}. \quad (33)$$

In order to analyze the expression above introduce new variables m', p', x_1, x_2 such that

$$m' = \mathrm{Tr}\,\Pi', \quad p' = \mathrm{Tr}(\Pi'\rho), \quad x_i = \lambda_i m'. \quad (34)$$

Expressing $\mathrm{Tr}(\Pi' R)$ and $\mathrm{Tr}(\Pi' R^\perp)$ in terms of m' and p' one can rewrite Eq. (33) as

$$\Lambda(\mathcal{E}) \leq 2\log\left(\frac{x_1}{x_2}\right)\sqrt{\frac{x_1 x_2(p' - x_2)(x_1 - p')}{(x_1 - x_2)^2}} \equiv g(x_1, x_2). \quad (35)$$

Let us find constraints on the variables x_1, x_2. Noting that $\lambda_2 I \leq \rho \leq \lambda_1 I$ and taking the trace with Π' one gets $x_2 \leq p' \leq x_1$. Besides, condition (iii) of Lemma 4 implies that $m' \leq m$ and thus $x_1 \leq m\lambda_1 \leq \mathrm{Tr}\,\rho = 1$. By obvious reasons one also has $x_2 \geq 0$. Summarizing,

$$0 \leq x_2 \leq p' \leq x_1 \leq 1. \quad (36)$$

Now the problem of getting an upper bound on the mixing rate reduces to maximizing a function $g(x_1, x_2)$ in Eq. (35) under constraints Eq. (36). We prove (see Lemma 5 at the end of the section) that $\max_{x_1, x_2} g(x_1, x_2) \leq 6p'|\log p'|$ as long as $p' \leq 1/2$. Using condition (i) of Lemma 4 we get $p' = \mathrm{Tr}(\Pi'\rho) \leq \mathrm{Tr}(\Pi\rho) = p$. As was mentioned in the beginning of the section, we can assume that $p \leq 1/2$ and thus $p' \leq p \leq 1/2$. Summarizing, we get

$$\Lambda(\mathcal{E}) \leq 6p'|\log p'| \leq 6p|\log p| \leq 6h(p, 1-p),$$

where we used the fact that a function $x|\log x|$ is monotone increasing on the interval $[0, 1/2]$. Thus we have proved SIM under the assumption that ρ has a binary spectrum.

Proof of Lemma 3. Hölder inequality asserts that

$$\|XY\|_1 \le \sqrt{\mathrm{Tr}(X^\dagger X)\mathrm{Tr}(Y^\dagger Y)} \qquad (37)$$

for any operators X and Y, see [12]. In order to choose proper X and Y let us use an identity $[\Pi,R]=R^\perp \Pi R - R\Pi R^\perp$ and triangle inequality for the trace norm:

$$\|[\Pi,R]\|_1 \le \|R^\perp \Pi R\|_1 + \|R\Pi R^\perp\|_1$$
$$= 2\|R^\perp \Pi R\|_1 = 2\|(R^\perp \Pi^{1/2})(\Pi^{1/2}R)\|_1.$$

Substituting $X=R^\perp \Pi^{1/2}$ and $Y=\Pi^{1/2}R$ into Eq. (37) and noting that R, $R^\perp$ are projectors we get the inequality stated in the lemma. ∎

Proof of Lemma 4. Any Hermitian operator Π satisfying $0\le \Pi\le I$ can be written as a convex combination of projectors:

$$\Pi = \sum_\alpha p_\alpha \Pi_\alpha, \quad \Pi_\alpha^\dagger \Pi_\alpha = \Pi_\alpha.$$

Suppose we can find the operator promised in the lemma for every projector in the sum, that is we can find Π_α' such that (i) $0\le \Pi_\alpha'\le \Pi_\alpha$, (ii) $[\Pi_\alpha',R]=[\Pi_\alpha,R]$, and (iii) $\mathrm{Tr}\,\Pi_\alpha'\le m$. Then we can choose the desired operator Π' as $\Pi' = \sum_\alpha p_\alpha \Pi_\alpha'$. Thus it suffices to prove the lemma for the case when Π is a projector. Let D be the dimension of the Hilbert space. Consider a direct sum decomposition

$$\mathbb{C}^D = \mathcal{H} \oplus \mathcal{H}^\perp,$$

where $\mathcal{H}$ is the range of R, so that $\dim(\mathcal{H})=m$. Then we can write

$$R = \begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix}, \quad \Pi = \begin{pmatrix} A & C \\ C^\dagger & B \end{pmatrix}. \qquad (38)$$

Here $0\le A\le I$ and $0\le B\le I$ are Hermitian operators on $\mathcal{H}$ and $\mathcal{H}^\perp$ and C is some operator $C:\mathcal{H}^\perp \to \mathcal{H}$. The requirement that Π is a projector implies

$$A(I-A) = CC^\dagger, \quad B(I-B) = C^\dagger C, \quad C(I-B) = AC. \qquad (39)$$

Consider a decomposition $A=A'\oplus A''$, $B=B'\oplus B''$ where $0<A'<I$ and $0<B'<I$, while A'' and B'' have only eigenvalues 0, 1. It follows from Eq. (39) that

$$A''CC^\dagger = CC^\dagger A'' = 0, \quad B''C^\dagger C = C^\dagger C B'' = 0.$$

Thus $C^\dagger|\psi\rangle=0$ for any $|\psi\rangle$ from the range of A'' and $C|\phi\rangle=0$ for any $|\phi\rangle$ from the range of B''. Accordingly, the projector Π has the following block structure:

$$\Pi = \begin{pmatrix} A'' & 0 & 0 & 0 \\ 0 & A' & C' & 0 \\ 0 & (C')^\dagger & B' & 0 \\ 0 & 0 & 0 & B'' \end{pmatrix}. \qquad (40)$$

Let Π' be the central block in Π. Clearly Π' is a projector and Π' satisfies conditions (i) and (ii) of the lemma. It remains to check that $\mathrm{Tr}\,\Pi'\le m$. Let m' be the dimension of the block A'. By definition, $m'\le m$. Since Π' is a projector,

the operators A',B',C' obey the same constraint as Eq. (39), that is

$$A'(I-A') = C'(C')^\dagger,$$

$$B'(I-B') = (C')^\dagger C,$$

$$C'(I-B') = AC'. \qquad (41)$$

Since $0<A'<I$ and $0<B'<I$ we conclude that $C'(C')^\dagger$ and $(C')^\dagger C'$ are nonsingular matrices and thus the dimensions of the blocks A' and B' both equal m'. Also from Eq. (41) we infer that $A'=C'(I-B')(C')^{-1}$, that is the spectrum of A' coincides with the spectrum of $I-B'$ including multiplicities. Accordingly, $\mathrm{Tr}\,\Pi'=\mathrm{Tr}\,A'+\mathrm{Tr}\,B'=m'\le m$. Therefore Π' satisfies all three conditions of the lemma. ∎

Lemma 5. Let $0\le q\le 1/2$ be a real number. Consider a function

$$g(x_1,x_2) = 2\log\left(\frac{x_1}{x_2}\right)\sqrt{\frac{x_1 x_2(q-x_2)(x_1-q)}{(x_1-x_2)^2}}.$$

Suppose $0\le x_2\le q$ and $q\le x_1\le 1$. Then $g(x_1,x_2)\le 6q|\log q|$.

Proof. Let us consider three cases.

Case 1. $2q\le x_1\le 1$, $q/4\le x_2\le q$.

Then $\log(x_1/x_2)\le\log(4/q)\le 3|\log q|$. Using an upper bound $x_2(q-x_2)\le q^2/4$ we arrive at

$$g(x_1,x_2) \le 3q|\log q|\sqrt{\frac{x_1(x_1-q)}{(x_1-x_2)^2}} \le 3q|\log q|\sqrt{\frac{x_1}{(x_1-q)}}$$

$$\le 3q|\log q|\sqrt{\frac{2q}{2q-q}} \le 3\sqrt{2}q|\log q|.$$

The last inequality follows from monotonicity of a function $x_1/(x_1-q)$ on the interval (q,∞).

Case 2. $q\le x_1\le 2q$, $q/4\le x_2\le q$.

First note that $\log(x_1/x_2)\le\log(8)=3$. Take into account that

$$\frac{(q-x_2)(x_1-q)}{(x_1-x_2)^2} \le \frac{1}{2}.$$

Therefore

$$g(x_1,x_2) \le 3\sqrt{2}\sqrt{x_1 x_2} \le 3\sqrt{2}\sqrt{(2q)q} \le 6q \le 6q|\log q|.$$

Case 3. $q\le x_1\le 1$, $0\le x_2\le q/4$.

Introduce new variable y such that $x_2=qy$, that is $0\le y\le 1/4$. Then

$$g(x_1,x_2) \le 2q\sqrt{y(1-y)}|\log qy|k(x_1),$$

where

$$k(x_1) = \sqrt{\frac{x_1(x_1-q)}{(x_1-q/4)^2}}.$$

One can check that $k(x_1)\le 2/\sqrt{3}$ for all $q\le x_1\le 1$. Using inequality

$$\sqrt{y(1-y)}|\log qy| \le \sqrt{y}|\log qy| = \sqrt{y}|\log q| + 2\sqrt{y}|\log\sqrt{y}|$$

and noticing that $t|\log t|$ is monotone increasing for $0 \le t \le 1/2$ we conclude that

$$\sqrt{y}|\log qy| \le \frac{1}{2}|\log(q/4)| \le \frac{3}{2}|\log q|$$

for any $0 \le y \le 1/4$. Thus

$$g(x_1, x_2) \le 2\sqrt{3}q|\log q|.$$

Combining all three cases we get

$$g(x_1, x_2) \le 6q|\log q|.$$

∎

V. NUMERICAL MAXIMIZATION OF THE MIXING RATE

This section describes numerical simulations aimed at verifying SIM. Let us start from expression Eq. (28) for the mixing rate and the constraints Eq. (27). It is convenient to represent the Hamiltonian as $H = 2K - I$. Note that $\|H\| \le 1$ if and only if $0 \le K \le I$. Then $\Lambda(\mathcal{E}, H) = F(K, \Pi)$ where

$$F(K, \Pi) = -2i\,\text{Tr}(K\rho^{1/2}[\Pi, \log\rho]\rho^{1/2})$$
$$= 2i\,\text{Tr}(\Pi\rho^{1/2}[K, \log\rho]\rho^{1/2}). \tag{42}$$

For a fixed average state ρ SIM is equivalent to an upper bound

$$\max_{K, \Pi} F(K, \Pi) \le c'h(p, 1-p), \tag{43}$$

where the maximization is subject to

$$0 \le K \le I, \quad 0 \le \Pi \le I, \quad \text{Tr}(\Pi\rho) = p. \tag{44}$$

We found the maximum of $F(K, \Pi)$ numerically for the average states

$$\rho = \rho^{(D)} = Z^{-1}\sum_{j=1}^{D} j^{-1}|j\rangle\langle j|, \quad Z = \sum_{j=1}^{D} j^{-1}, \tag{45}$$

with the dimension $D = 4, 8, 16, 32$ and for several values of p between 0 and 1/2, see Fig. 3 (as was mentioned in Sec. IV, it is enough to consider p between the smallest eigenvalue of ρ and 1/2).

The motivation for this particular choice of the average state comes from the fact that the states $\rho^{(D)}$ can "embezzle"
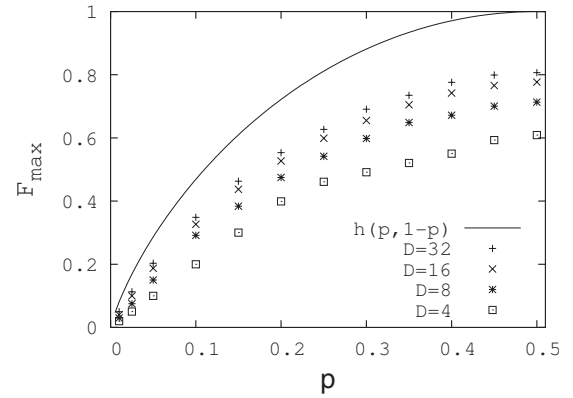

FIG. 3. Maximal mixing rate $F_{\max} = \max F(K, \Pi)$ (vertical axis) computed for a fixed average state $\rho = \rho^{(D)}$, see Eq. (45), and fixed p (horizontal axis). The maximization was performed for $D = 4, 8, 16, 32$. The upper curve shows the Shannon entropy $h(p, 1-p) = -p\log p - (1-p)\log(1-p)$.

any other mixed state for sufficiently large D, see [10]. More strictly, it was proved in [10] that for any state ρ there exists an isometry V such that $\rho^{(D)} \approx V^\dagger(\rho^{(D)} \otimes \rho)V$ in the sense that fidelity between the two states goes to 1 as D goes to infinity. In particular this is true for the optimal state ρ corresponding to the optimal ensemble $\mathcal{E} = (p, \rho, \Pi)$ maximizing the mixing rate in Eq. (29) for a fixed p. On the other hand, the state $\rho^{(D)} \otimes \rho$ is at least as good as ρ as far as the mixing rate is concerned. Indeed, one can easily verify that the mixing rate for the ensemble (p, ρ, Π) is the same as the mixing rate for the ensemble $(p, \tilde{\rho}, \tilde{\Pi})$, where $\tilde{\rho} = \rho^{(D)} \otimes \rho$ and $\tilde{\Pi} = I \otimes \Pi$. Thus a verification of SIM for the family of states $\rho^{(D)}$ is a good test for general validity of SIM.

The data obtained in the numerical maximization of $F(K, \Pi)$ are presented in Fig. 3. They are consistent with the conjecture Eq. (43) with a constant $c' = 1$. The numerical algorithm that we used is based on reformulation of the maximization problem as a semidefinite program. The details of the algorithm can be found in [13].

[1] W. Dür, G. Vidal, J. I. Cirac, N. Linden, and S. Popescu, Phys. Rev. Lett. **87**, 137901 (2001).

[2] C. H. Bennett, A. W. Harrow, D. W. Leung, and J. A. Smolin, IEEE Trans. Inf. Theory **49**, 1895 (2003).

[3] S. Bravyi, M. B. Hastings, and F. Verstraete, Phys. Rev. Lett. **97**, 050401 (2006).

[4] A. M. Childs, D. W. Leung, F. Verstraete, and G. Vidal, Quan-

tum Inf. Comput. **3**, 97 (2003).

[5] J. I. Cirac, W. Dür, B. Kraus, and M. Lewenstein, Phys. Rev. Lett. **86**, 544 (2001).

[6] X. Wang and B. C. Sanders, Phys. Rev. A **68**, 014301 (2003).

[7] A. Childs, D. Leung, and G. Vidal, IEEE Trans. Inf. Theory **50**, 1189 (2004).

[8] N. Linden, J. Smolin, and A. Winter, e-print arXiv:quant-ph/0511217.

[9] A. Kitaev (private communication).

[10] W. van Dam and P. Hayden, Phys. Rev. A **67**, 060302(R) (2003).

[11] D. DiVincenzo, M. Horodecki, D. Leung, J. Smolin, and B. Terhal, Phys. Rev. Lett. **92**, 067902 (2004).

[12] R. Bhatia, *Graduate Textx in Mathematics* (Springer-Verlag, New York, 1997).

[13] S. Bravyi, e-print arXiv:0704.0964.