

# On the role of entanglement and correlations in mixed-state quantum computation \*

Animesh Datta<sup>†</sup>

Department of Physics and Astronomy, University of New Mexico, Albuquerque, New Mexico 87131-1156, USA

Guifre Vidal<sup>‡</sup>

School of Physical Sciences, The University of Queensland, QLD 4072, Australia

(Dated: March 7, 2007)

In a quantum computation with pure states, the generation of large amounts of entanglement is known to be necessary for a speed-up with respect to classical computations. However, examples of quantum computations with mixed states are known, such as the DQC1 model [E. Knill and R. Laflamme, Phys. Rev. Lett. **81**, 5672 (1998)], in which entanglement is at most marginally present, and yet a computational speed-up is believed to occur. Correlations, and not entanglement, have been identified as a necessary ingredient for mixed-state quantum computation speed-ups. Here we show that correlations, as measured through the operator Schmidt rank, are indeed present in large amounts in the DQC1 circuit. This provides evidence for the preclusion of efficient classical simulation of DQC1 by means of a whole class of classical simulation algorithms, thereby reinforcing the conjecture that DQC1 leads to a genuine quantum computational speed-up.

PACS numbers: 3.67.Lx

## I. INTRODUCTION

Quantum computation owes its popularity to the realization, more than a decade ago, that the factorization of large numbers can be solved exponentially faster by evolving quantum systems than with any known classical algorithm [1]. Since then, progress in our understanding of what makes quantum evolutions computationally more powerful than a classical computer has been scarce. A step forward, however, was achieved by identifying entanglement as a *necessary* resource for quantum computational speed-ups. Indeed, a speed-up is only possible if in a quantum computation, entanglement spreads over an adequately large number of qubits [2]. In addition, the amount of entanglement, as measured by the Schmidt rank of a certain set of bipartitions of the system, needs to grow sufficiently with the size of the computation [3]. Whenever either of these two conditions is not met, the quantum evolution can be efficiently simulated on a classical computer. These conditions (which are particular examples of subsequent, stronger classical simulation results based on tree tensor networks (TTN) [4]) are only necessary, and thus not sufficient, so that the presence of large amounts of entanglement spreading over many qubits does not guarantee a computational speed-up, as exemplified by the Gottesman-Knill theorem [5].

The above results refer exclusively to quantum computations with pure states. The scenario for mixed-state quantum computation is rather different. The intriguing *deterministic quantum computation with one quantum bit* (DQC1 or ‘the power of one qubit’) [6] involves

a highly mixed state that does not contain much entanglement [7] and yet it performs a task, the computation with fixed accuracy of the normalized trace of a unitary matrix, exponentially faster than any known classical algorithm. Thus, in the case of a mixed-state quantum computation, a large amount of entanglement does not seem to be necessary to obtain a speed-up with respect to classical computers.

A simple, unified explanation for the pure-state and mixed-state scenarios is possible [3] by noticing that the decisive ingredient in both cases is the presence of *correlations*. Indeed, let us consider the Schmidt decomposition of a vector  $|\Psi\rangle$ , given by

$$|\Psi\rangle = \sum_{i=1}^{\chi} \lambda_i |i_A\rangle \otimes |i_B\rangle, \quad (1.1)$$

where  $\langle i_A | j_A \rangle = \langle i_B | j_B \rangle = \delta_{ij}$  and  $\chi$  is the rank of the reduced density matrices  $\rho_A \equiv \text{Tr}_B[|\Psi\rangle\langle\Psi|]$  and  $\rho_B \equiv \text{Tr}_A[|\Psi\rangle\langle\Psi|]$ ; and the (operator) Schmidt decomposition of a density matrix  $\rho$  given by [8]

$$\rho = \sum_{i=1}^{\chi^\sharp} \lambda_i^\sharp O_{iA} \otimes O_{iB}, \quad (1.2)$$

where  $\text{Tr}(O_{iA}^\dagger O_{jA}) = \text{Tr}(O_{iB}^\dagger O_{jB}) = \delta_{ij}$ . The Schmidt ranks  $\chi$  and  $\chi^\sharp$  are a measure of correlations between parts  $A$  and  $B$ , with  $\chi^\sharp = \chi^2$  if  $\rho = |\Psi\rangle\langle\Psi|$ . Let the density matrix  $\rho_t$  denote the evolving state of the quantum computer during a computation. Notice that  $\rho_t$  can represent both pure and mixed states. Then, as shown in Refs. [3] and [4], the quantum computation can be efficiently simulated on a classical computer using a TTN decomposition if the Schmidt rank  $\chi^\sharp$  of  $\rho$  according to a certain set of bipartitions  $A : B$  of the qubits scales polynomially with the size of the computation. In other

---

\*Some of the results in this paper were presented at the APS March Meeting, 2007, Denver.

<sup>†</sup>Electronic address: [animesh@unm.edu](mailto:animesh@unm.edu)

<sup>‡</sup>Electronic address: [vidal@physics.uq.edu.au](mailto:vidal@physics.uq.edu.au)

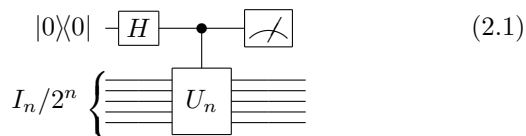
words, a necessary condition for a computational speed-up is that correlations, as measured by the Schmidt rank  $\chi^\sharp$ , grow super-polynomially in the number of qubits. In the case of pure states (where  $\chi = \sqrt{\chi^\sharp}$ ) these correlations are entirely due to entanglement, while for mixed states they may be quantum or classical.

Our endeavor in this paper is to illustrate the above line of thought by applying it to the DQC1 model of quantum computation. In particular, we elucidate whether DQC1 can be efficiently simulated with any classical algorithm, such as those in [3],[4] [and, implicitly, in [2]], that exploits limits on the amount of correlations, in the sense of a small  $\chi^\sharp$  according to certain bipartitions of the qubits. We will argue here that the state  $\rho_t$  of a quantum computer implementing the DQC1 model displays an exponentially large  $\chi^\sharp$ , in spite of it containing only a small amount of entanglement [7]. We will conclude, therefore, that none of the simulation techniques mentioned above can be used to efficiently simulate ‘the power of one qubit’.

On the one hand, our result indicates that a large amount of classical correlations are behind the (suspected) computational speed-up of DQC1. On the other hand, by showing the failure of a whole class of classical algorithms to efficiently simulate this mixed-state quantum computation, we reinforce the conjecture that DQC1 leads indeed to an exponential speed-up. We note, however, that our result does *not* rule out the possibility that this circuit could be simulated efficiently using some other classical algorithm.

## II. DQC1 AND TREE TENSOR NETWORKS (TTN)

The DQC1 model, represented in Eq. (2.1), provides an estimate of the normalized trace  $\text{Tr}(U_n)/2^n$  of a  $n$ -qubit unitary matrix  $U_n \in \text{U}(2^n)$  with fixed accuracy efficiently [6]. For discussions on the classical complexity of evaluating the normalized trace of an unitary, see [7].



$$|0\rangle\langle 0| \xrightarrow{H} \bullet \xrightarrow{\text{Measurement}} \quad (2.1)$$

$$I_n/2^n \left\{ \begin{array}{c} \text{---} \\ \text{---} \\ \text{---} \\ \text{---} \\ \text{---} \\ \text{---} \\ \text{---} \\ \text{---} \end{array} \right\} U_n \left\{ \begin{array}{c} \text{---} \\ \text{---} \\ \text{---} \\ \text{---} \\ \text{---} \\ \text{---} \\ \text{---} \\ \text{---} \end{array} \right\}$$

It transforms the highly-mixed initial state  $\rho_0 \equiv |0\rangle\langle 0| \otimes I_n/2^n$  at time  $t = 0$  into the final state  $\rho_T$  at time  $t = T$ ,

$$\rho_T = \frac{1}{2^{n+1}} \begin{pmatrix} I_n & U_n^\dagger \\ U_n & I_n \end{pmatrix}, \quad (2.2)$$

through a series of intermediate states  $\rho_t$ ,  $t \in [0, T]$ . The simulation algorithms relevant in the present discussion [2]-[4] require that  $\rho_t$  be efficiently represented with a TTN [4] (or a more restrictive structure, such as a product of  $k$ -qubit states for fixed  $k$  [2] or a matrix product

state [3]) at all times  $t \in [0, T]$ . Here we will show that the final state  $\rho_T$ , henceforth denoted simply by  $\rho$ , cannot be efficiently represented with a TTN.

The computational cost associated with a TTN, in time and space, grows at most linearly in the number of qubits  $n$  and as a small power of its rank  $q$ . The rank  $q$  of a TTN is the maximum Schmidt rank  $\chi_i^\sharp$  over all bipartitions  $A_i : B_i$  of the qubits according to a given tree graph whose leaves are the qubits of our system. See [4] for details. It is not difficult to deduce that for any tree of  $n + 1$  qubits, there exists at least one edge that splits the tree in two parts  $A$  and  $B$ , with  $n_A$  and  $n_B$  qubits, where  $n_0 = \min(n_A, n_B)$  fulfills  $n/5 \leq n_0 \leq 2n/5$ . In other words, if a rank- $q$  TTN exists for the  $\rho$  in Eq. (2.2), then there is a bipartition of the  $n + 1$  qubits with  $n_0$  qubits on either  $A$  or  $B$  and such that the Schmidt rank  $\chi^\sharp \leq q$ . Theorem 1, our main technical result, shows that if  $U_n$  is chosen randomly according to the Haar measure, then the Schmidt rank of any such bipartition fulfills  $\chi^\sharp \geq O(2^{n_0})$ . Therefore for a randomly generated  $U_n$ , a TTN for  $\rho$  has rank  $q$  (and computational cost) exponential in  $n$ , and none of the techniques of [2]-[4] can simulate the outcome of the DQC1 model efficiently.

Two comments are in order here. (i) Firstly, Theorem 1 does not exclude the possibility that the quantum computation in the DQC1 model can be efficiently simulated with a TTN for particular choices of  $U_n$ . For instance, if  $U_n$  factorizes into single-qubit gates, then  $\rho$  can be seen to be efficiently represented with a TTN of rank 3, and we can not rule out an efficient simulation of the power of one qubit for that case. (ii) Secondly, the DQC1 model is defined for unitary matrices  $U_n$  that are efficiently generated in terms of some universal gate set. That is,  $U_n$  must be the product of *poly*( $n$ ) one-qubit and two-qubit gates. Such unitary matrices form a subset of measure zero in the group  $\text{U}(2^n)$  of arbitrary  $n$ -qubit unitary matrices. Strictly speaking, then, a  $U_n$  generated randomly in an efficient quantum circuit is not Haar distributed, and Theorem 1 does not apply to it. And yet, plenty of evidence suggests that, as far as our measure of correlations  $\chi^\sharp$  is concerned, such  $U_n$  behaves as if it was Haar distributed. Indeed, the results of [10] indicate that random (but efficient) quantum circuits generate pseudo-random  $n$ -qubit gates according to a measure that converges to the Haar measure. In addition, we have performed numerical simulations that unambiguously support this view. This is the content of the following section.

## III. SCHMIDT RANK IN THE DQC1 STATE: NUMERICAL STUDY

In this section, we study the Schmidt rank of an output to the DQC1 circuit, a state of the form in Eq. (2.2). We will not calculate the rank explicitly, but numerically evaluate a lower bound to it which we find is exponential, thus allowing us to conclude that there exists a splitting

of the total number of qubits for which the Schmidt rank is at least exponential.

The Schmidt rank of a pure state obtained by applying operator onto a fiducial pure state is a lower bound on the operator Schmidt rank of the operator, i.e., the Schmidt rank  $\chi$  of the vector

$$|\rho_{\phi_A\psi_B}\rangle \equiv \rho|\phi_A\rangle|\psi_B\rangle = \sum_{i=1}^{\chi^\sharp} \lambda_i^\sharp O_{iA}|\phi_A\rangle \otimes O_{iB}|\psi_B\rangle \quad (3.1)$$

is a lower bound for the Schmidt rank  $\chi^\sharp$  of  $\rho$  or  $\chi^\sharp \geq \chi$ . For the purpose of our numerics, we consider the pure state  $U_n|0\rangle^{\otimes n}$ . It will be shown in the next section that this is a correct state to consider.  $U_n$  is composed of a sequence of  $2n$  pseudo-random 2 qubit gates, applied to pairs of qubits, also chosen at random. The pseudo-random unitary is generated using the mixing algorithm presented in [9]. Note that applying  $2n$  gates means that the resulting unitary is efficiently implementable, a situation for which the DQC1 model is valid. For an even number of qubits  $n$ , we calculate the Schmidt rank for a  $n/2 : n/2$  partition of the qubits (similar results can be obtained for odd  $n$ ) by taking a minimum over all possible  $n/2 : n/2$  partitions. The resulting numbers are plotted in Fig (1).

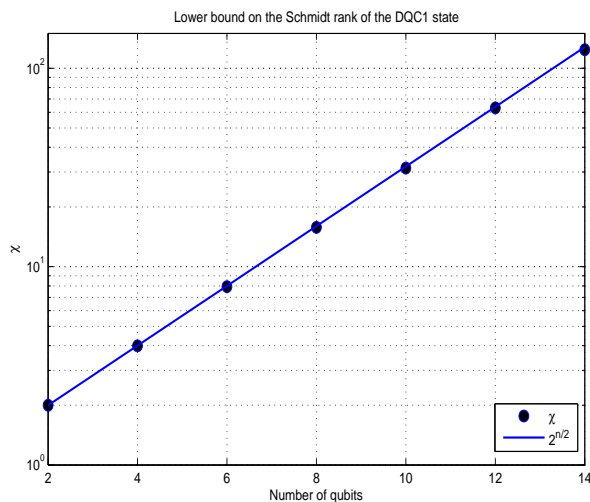


FIG. 1: A lower bound on the Schmidt rank of the DQC1 state for an equipartition. The dots are for even numbers of qubits, and the fit is the line  $2^{n/2}$ .  $\chi$  is calculated for a pure state obtained by applying  $2n$  random 2-qubit gates on the state  $|0\rangle^{\otimes n}$ . This is evidence that the Schmidt rank of a DQC1 state due to an efficiently generated unitary goes at least exponentially.

The above numerical results induces us to conclude that the final state in the DQC1 circuit has exponential Schmidt rank for an efficiently generated pseudo-random unitary. In the following section, we prove this analytically for any unitary chosen randomly according to the Haar measure from the group  $\mathbb{U}(n)$ .

#### IV. SCHMIDT RANK IN THE DQC1 STATE: ANALYTICAL STUDY

Our objective in this section is to analyze the Schmidt rank  $\chi^\sharp$  of  $\rho$  in Eq. (2.2), where  $U_n$  is Haar-distributed. Consider any bipartition  $A : B$  of the  $n+1$  qubits, where  $A$  and  $B$  contain  $n_A$  and  $n_B$  qubits, with the minimum  $n_0$  of those restricted by  $n/5 \leq n_0 \leq 2n/5$ . Without loss of generality we can assume that the top qubit lies in  $A$ . Actually, we can also assume that  $A$  contains the top  $n_A$  qubits. Indeed, suppose  $A$  does not have the  $n_A$  top qubits. Then we can use a permutation  $P_n$  on all the  $n$  qubits to bring the  $n_A$  qubits of  $A$  to the top  $n_A$  positions. This will certainly modify  $\rho$ , but since

$$\begin{pmatrix} P_n & 0 \\ 0 & P_n \end{pmatrix} \begin{pmatrix} I_n & U_n^\dagger \\ U_n & I_n \end{pmatrix} \begin{pmatrix} P_n^T & 0 \\ 0 & P_n^T \end{pmatrix} = \begin{pmatrix} I_n & V_n^\dagger \\ V_n & I_n \end{pmatrix} \quad (4.1)$$

where  $V_n = P_n U_n P_n^T$  is another Haar-distributed unitary, we obtain that the new density matrix is of the same form as  $\rho$ . Finally, in order to ease the notation, we will assume that  $n_A = n_0$  (identical results can be derived for  $n_B = n_0$ ). Thus  $n/5 \leq n_A \leq 2n/5$ .

We note that

$$\begin{pmatrix} I_n & U_n^\dagger \\ U_n & I_n \end{pmatrix} = \mathbb{I}_2 \otimes \mathbb{I}_n + \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \otimes U_n^\dagger + \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \otimes U_n, \quad (4.2)$$

so that if we multiply  $\rho$  by the product state

$$|\phi_{\vec{\alpha}}\rangle \equiv |t, i, j\rangle \equiv |t, i_A\rangle |j_B\rangle, \quad (4.3)$$

where  $\vec{\alpha} \equiv (t, i, j)$ ,  $t = 0, 1$ ;  $i = 1, \dots, d_A$ ;  $j = 1, \dots, d_B$ , we obtain  $|\psi_{\vec{\alpha}}\rangle \equiv \rho|\phi_{\vec{\alpha}}\rangle$  where

$$|\psi_{\vec{\alpha}}\rangle = \begin{cases} \frac{1}{2^{n+1}} (|0, i, j\rangle + |1\rangle \otimes U_n |i, j\rangle) & \text{if } t = 0 \\ \frac{1}{2^{n+1}} (|1, i, j\rangle + |0\rangle \otimes U_n^\dagger |i, j\rangle) & \text{if } t = 1 \end{cases} \quad (4.4)$$

This justifies our choice of the pure state used in the numerical calculations in the previous section.

Let us consider now the reduced density matrix

$$\begin{aligned} \sigma_{\vec{\alpha}}^B &\equiv \text{Tr}_A[|\psi_{\vec{\alpha}}\rangle\langle\psi_{\vec{\alpha}}|] \\ &= \frac{1}{2^{n+1}} (|j\rangle\langle j| + \text{Tr}_A[U_n |i, j\rangle\langle i, j| U_n^\dagger]) \end{aligned} \quad (4.5)$$

for  $t = 0$  (for  $t = 1$ ,  $U_n$  and  $U_n^\dagger$  need to be exchanged). For a unitary matrix  $U_n$  randomly chosen according to the Haar measure on  $\mathbb{U}(n)$ ,  $U_n |i, j\rangle$  is a random pure state on  $A \otimes B$ . Here, and henceforth  $A$  is the space of the first  $n_A$  qubits without the top qubit. It follows from [12] that the operator

$$Q = \text{Tr}_A[U_n |i, j\rangle\langle i, j| U_n^\dagger] \quad (4.6)$$

has rank  $d_A$ . Therefore the rank of  $\sigma_{\vec{\alpha}}^B$  (equivalently, the Schmidt rank  $\chi$  of  $|\psi_{\vec{\alpha}}\rangle$ ) is at least  $2^{n_0}$ . From Eq. (3.1) we conclude that the Schmidt rank of  $\rho$  fulfills  $\chi^\sharp \geq 2^{n_0} \geq 2^{n/5}$ . We can now collate these results into

**Theorem 1** Let  $U_n$  be an  $n$ -qubit unitary transformation chosen randomly according to the Haar measure on  $U(2^n)$ , and let  $A : B$  denote a bipartition of  $n + 1$  qubits into  $n_A$  and  $n_B$  qubits, where  $n_0 \equiv \min(n_A, n_B)$ . Then  $n/5 \leq n_0 \leq 2n/5$  and the Schmidt decomposition of  $\rho$  in Eq. (2.2) according to bipartition  $A : B$  fulfills  $\chi^\sharp \geq 2^{n/5}$ .

We have seen that we cannot efficiently simulate DQC1 with an algorithm that relies on having a TTN for  $\rho$  with low rank  $q$ . It is however possible that the evolution of all the exponential number of Schmidt coefficients are irrelevant in the classical simulation of the DQC1 state. In other words, the DQC1 state could be well approximated by another state with much lower Schmidt rank. That this is not the case is the content of Appendix A.

## V. DISCUSSIONS

In this paper, we have proved that the DQC1 state evaluating the normalized trace of a unitary given by Eq (2.2) has an exponential number of Schmidt coefficients all of which are approximately equal. The class of algorithms dealt with in this paper will therefore, not simulate efficiently a DQC1 circuit with a unitary randomly chosen according to the Haar measure from  $\mathbb{U}(2^n)$ .

In quantum computation, we are interested in efficiently generated unitaries. Procedures are known for generating pseudo-random unitaries efficiently [9]. It is for this class of operations that the DQC1 circuit is efficient [6] and analyzed [7]. The measure over ensembles of pseudo-random unitaries efficiently generated by a quantum circuit are known to converge uniformly to Haar measure over the unitary group, although the rate of convergence decreases exponentially in this case. However, we do not require uniform convergence. A weaker form of convergence to the Haar measure, with respect to some test function  $f$  is sufficient. For the Schmidt rank  $f = \chi^\sharp$ , and a convergence to Haar measure for this ‘weak topology’ is possible [10]. This suggests that the final state in the DQC1 circuit cannot be efficiently simulated classically by the class of algorithms discussed in this paper. This is corroborated by the numerical simulations shown in Fig 1.

It is also interesting to note that Theorems 1 and 2 can be generalized for any fixed polarization  $\tau$ , ( $0 < \tau \leq 1$ ) of the initial state  $\tau|0\rangle\langle 0| + (1 - \tau)\mathbb{I}/2$  of the top qubit of the circuit in Eq (2.1), implying that the algorithms of [2]-[4] are unable to efficiently simulate the power of even the  *tiniest*  fraction of a qubit.

## Acknowledgements

AD acknowledges the US Army Research Office for support via Contract No. W911NF-4-1-0242 and a Visiting Fellowship from the University of Queensland, where this work was initiated. GV thanks support from the

Australian Research Council through a Federation Fellowship.

## APPENDIX A: DISTRIBUTION OF THE SCHMIDT COEFFICIENTS

In this Appendix we explore the robustness of the statement of Theorem 1. To this end, we consider the Schmidt rank  $\tilde{\chi}^\sharp$  for a density matrix  $\tilde{\rho}$  that approximates  $\rho$  according to a fidelity  $F(O_1, O_2)$  defined in terms of the natural inner product on the space of linear operators,

$$F(O_1, O_2) \equiv \text{Tr}(O_1^\dagger O_2) / \sqrt{\text{Tr}(O_1^\dagger O_1)} \sqrt{\text{Tr}(O_2^\dagger O_2)},$$

where  $F = 1$  if and only if  $O_1 = O_2$  and  $F = |\langle \psi_1 | \psi_2 \rangle|^2$  for projectors  $O_i = P_{\psi_i}$  on pure states  $|\psi_i\rangle$ . We will show that if  $\tilde{\rho}$  is close to  $\rho$ , then  $\tilde{\chi}^\sharp$  for a bipartition as in Theorem 1 is also exponential. To prove this, we will require a few lemmas which we now present.

**Lemma 1** Let  $|\Psi\rangle$  be a bipartite vector with  $\chi$  terms in its Schmidt decomposition,

$$|\Psi\rangle = N_\Psi \sum_{i=1}^{\chi} \lambda_i |i_A\rangle |i_B\rangle, \quad \lambda_i \geq \lambda_{i+1} \geq 0, \quad \sum_{i=1}^{\chi} \lambda_i^2 = 1,$$

where  $N_\Psi \equiv \sqrt{\langle \Psi | \Psi \rangle}$ , and let  $|\Phi\rangle$  be a bipartite vector with norm  $N_\Phi$  and Schmidt rank  $\chi'$ , where  $\chi' \leq \chi$ . Then,

$$\max_{|\Phi\rangle} |\langle \Psi | \Phi \rangle| = N_\Psi N_\Phi \sqrt{\sum_{i=1}^{\chi'} \lambda_i^2}. \quad (\text{A1})$$

*Proof:* Let  $\mu_i$  denote the Schmidt coefficients of  $|\Phi\rangle$ . It follows from Lemma 1 in [11] that  $\max_{|\Phi\rangle} |\langle \Psi | \Phi \rangle| = N_\Psi N_\Phi \sum_{i=1}^{\chi'} \lambda_i \mu_i$ , and the maximization over  $\mu_i$  is done next. A straightforward application of the method of Lagrange multipliers provides us with  $\mu_i = c \lambda_i$ ,  $i = 1, 2, \dots, \chi'$  for some constant  $c$ . Since  $\sum_{i=1}^{\chi'} \mu_i^2 = 1 = c^2 \sum_{i=1}^{\chi'} \lambda_i^2$ ,  $c = 1 / \sum_{i=1}^{\chi'} \lambda_i^2$ . Thus,

$$\max_{|\Phi\rangle} |\langle \Psi | \Phi \rangle| = c N_\Psi N_\Phi \sum_{i=1}^{\chi'} \lambda_i^2$$

and the result follows.  $\square$

We will also use two basic results related to majorization theory. Recall that, by definition, a decreasingly ordered probability distribution  $\vec{p} = (p_1, p_2, \dots, p_d)$ , where  $p_\alpha \geq p_{\alpha+1} \geq 0$ ,  $\sum_\alpha p_\alpha = 1$ , is *majorized* by another such probability distribution  $\vec{q}$ , denoted  $\vec{p} \prec \vec{q}$ , if  $\vec{q}$  is more ordered or concentrated than  $\vec{p}$  (equivalently,  $\vec{p}$  is flatter or more mixed than  $\vec{q}$ ) in the sense that the following inequalities are fulfilled:

$$\sum_{\alpha=1}^k p_\alpha \leq \sum_{\alpha=1}^k q_\alpha \quad \forall k = 1, \dots, d \quad (\text{A2})$$

with equality for  $k = d$ . The following result can be found in Exercise II.1.15 of [13]:

**Lemma 2** Let  $\rho_{\vec{x}}$  and  $\rho_{\vec{y}}$  be density matrices with eigenvalues given by probability distributions  $\vec{x}$  and  $\vec{y}$ . Let  $\sigma(M)$  denote the decreasingly ordered eigenvalues of hermitian operator  $M$ . Then

$$\sigma(\rho_{\vec{x}} + \rho_{\vec{y}}) \prec \vec{x} + \vec{y}.$$

The next result follows by direct inspection.

**Lemma 3** Let coefficients  $\delta_i$ ,  $1 \leq i \leq d$ , be such that  $-\delta \leq \delta_i \leq \delta$  for some positive  $\delta \leq 1$  and  $\sum_i \delta_i = 1$ , and consider the probability distribution  $\vec{p}(\{\delta_i\})$ ,

$$\vec{p}(\{\delta_i\}) \equiv \left( \frac{1}{2} + \frac{1 + \delta_1}{2d}, \frac{1 + \delta_2}{2d}, \dots, \frac{1 + \delta_d}{2d} \right).$$

Then

$$\vec{p}(\{\delta_i\}) \prec \vec{p}(\{\delta_i^*\}),$$

where

$$\delta_i^* \equiv \begin{cases} \delta & i \leq d/2 \\ -\delta & i > d/2 \end{cases}$$

and we assume  $d$  to be even.

Finally, we need a result from [12]:

**Lemma 4** With probability very close to 1,

$$\begin{aligned} & \Pr \left[ (1 - \delta) \frac{\Upsilon}{d_A} \leq Q \leq (1 + \delta) \frac{\Upsilon}{d_A} \right] \\ & \geq 1 - \left( \frac{10 d_A}{\delta} \right)^{2d_A} 2^{-(d_B \delta^2 / 14 \ln 2)} \\ & \geq 1 - O \left( \frac{1}{\exp(\delta^2 \exp(n))} \right), \end{aligned} \quad (\text{A3})$$

where  $d_A = 2^{n_A} = 2^{n_0}$  and  $d_B = 2^{n_B} = 2^{n - n_0 + 1}$ , and the operator  $Q$  defined in Eq. (4.6) is within a ball of radius  $\delta$  of a (unnormalized) projector  $\Upsilon/d_A$  of rank  $d_A$  [provided  $d_B$  is a large multiple of  $d_A \log d_A / \delta^2$  [12], which is satisfied for large  $n$ , given that  $n/5 \leq n_0 \leq 2n/5$ ].

Our second theorem uses the fact that the Schmidt decomposition of  $\rho$  does not only have exponentially many coefficients, but that these are roughly of the same size.

**Theorem 2** Let  $\rho$ ,  $U_n$ , and  $A:B$  be defined as in Theorem 1. If  $F(\rho, \tilde{\rho}) \geq 1 - \epsilon$ , then with probability  $p(\delta, n) = 1 - O(\exp(-\delta^2 \exp(n)))$ , the Schmidt rank for  $\tilde{\rho}$  according to bipartition  $A:B$  satisfies  $\tilde{\chi}^\# \geq (1 - 4\epsilon - \delta)2^{n/5}$ .

*Proof:* For any product vector of Eq. (4.3) we have

$$\begin{aligned} |\langle tij | \rho \tilde{\rho} | tij \rangle| & \leq N_{\tilde{\alpha}} \tilde{N}_{\tilde{\alpha}} \sqrt{\sum_{k=1}^{\tilde{\chi}^\#} (\lambda_k^{ij})^2} \\ & \leq N_{\tilde{\alpha}} \tilde{N}_{\tilde{\alpha}} g(\tilde{\chi}^\# / d_A), \end{aligned} \quad (\text{A4})$$

where

$$g(x) \equiv \sqrt{\frac{1 + (1 + \delta)x}{2}} \quad (\text{A5})$$

and  $N_{\tilde{\alpha}} \equiv \sqrt{\langle tij | \rho^2 | tij \rangle}$ ,  $\tilde{N}_{\tilde{\alpha}} \equiv \sqrt{\langle tij | \tilde{\rho}^2 | tij \rangle}$ . The first inequality in (A4) follows from Lemma 1, whereas the second one follows from the fact that the spectrum  $\vec{p}$  of

$$\rho_B \equiv (N_{\tilde{\alpha}})^{-2} \text{Tr}_A[\rho | tij \rangle \langle tij | \rho] = \frac{1}{2}(|j\rangle \langle j| + Q),$$

where  $Q$  has all its  $d_A$  non-zero eigenvalues  $q_i$  in the interval  $2^{-n_0}(1 - \delta) \leq q_i \leq 2^{-n_0}(1 + \delta)$ , is majorized by  $\vec{p}(\{\delta_i^*\})$ , as follows from Lemmas 2 and 3. Then,

$$\begin{aligned} 1 - \epsilon & \leq \frac{\text{Tr} \rho \tilde{\rho}}{\sqrt{\text{Tr} \rho^2} \sqrt{\text{Tr} \tilde{\rho}^2}} \\ & = \frac{\sum_{\tilde{\alpha}} \langle \tilde{\alpha} | \rho \tilde{\rho} | \tilde{\alpha} \rangle}{\sqrt{\sum_{\tilde{\alpha}'} \langle \tilde{\alpha}' | \rho^2 | \tilde{\alpha}' \rangle} \sqrt{\sum_{\tilde{\alpha}''} \langle \tilde{\alpha}'' | \tilde{\rho}^2 | \tilde{\alpha}'' \rangle}} \\ & \leq g(\tilde{\chi}^\# / d_A) \frac{\sum_{\tilde{\alpha}} N_{\tilde{\alpha}} \tilde{N}_{\tilde{\alpha}}}{\sqrt{\sum_{\tilde{\alpha}'} (N_{\tilde{\alpha}'})^2} \sqrt{\sum_{\tilde{\alpha}''} (\tilde{N}_{\tilde{\alpha}'})^2}} \\ & \leq g(\tilde{\chi}^\# / d_A), \end{aligned}$$

where in the last step we have used the Cauchy-Schwarz inequality,  $|\langle x | y \rangle| \leq \sqrt{\langle x | x \rangle} \sqrt{\langle y | y \rangle}$ . The result of the theorem follows from  $g(\tilde{\chi}^\# / 2^{n_0}) \geq 1 - \epsilon$ .  $\square$

[1] P. Shor, Proceedings of the 35th Annual Symposium on Foundations of Computer Science, Santa Fe, NM, 20 to 22 November 1994, S. Goldwasser, Ed. (IEEE Computer Science, Los Alamitos, CA, 1994) p. 124.  
[2] R. Jozsa and N. Linden, Proc. Roy. Soc. Lond. A **459**, 2011 (2003).  
[3] G. Vidal, Phys. Rev. Lett. **91**, 147902 (2003).  
[4] Y.-Y. Shi, L.-M. Duan, and G. Vidal, Phys. Rev. A **74**, 022320 (2006). M. Van den Nest, W. Dür, G. Vidal, H. J. Briegel, quant-ph/0608060.

[5] M. Nielsen and I. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, England, 2000).  
[6] E. Knill and R. Laflamme, Phys. Rev. Lett. **81**, 5672 (1998).  
[7] A. Datta, S. T. Flammia, and C. M. Caves, Phys. Rev. A **72**, 042316 (2005).  
[8] M. Zwolak and G. Vidal, Phys. Rev. Lett. **93**, 207205 (2004).  
[9] J. Emerson, Y. S. Weinstein, M. Saraceno, S. Lloyd, and

- D. G. Cory, *Science* **302**, 2098 (2003).
- [10] J. Emerson, E. Livine, S. Lloyd, *Phys. Rev. A* **72**, 060302 (2005).
- [11] G. Vidal, D. Jonathan, and M. A. Nielsen, *Phys. Rev. A* **62**, 012304 (2000).
- [12] P. Hayden, D. W. Leung, and A. Winter, *Commun. Math. Phys.* **265**, 95 (2006).
- [13] Rajendra Bhatia, *Matrix Analysis* (Springer-Verlag, New York, 1997).