Majorization arrow in quantum-algorithm design

J. I. Latorre
Departamento d'Estructura i Constituents de la Matèria, Universidad Barcelona, 08028 Barcelona, Spain

M. A. Martín-Delgado
Departamento de Física Teórica I, Universidad Complutense, 28040 Madrid, Spain

(Received 11 December 2001; published 9 August 2002)

We apply majorization theory to study the quantum algorithms known so far and find that there is a majorization principle underlying the way they operate. Grover's algorithm is a neat instance of this principle where majorization works step by step until the optimal target state is found. Extensions of this situation are also found in algorithms based in quantum adiabatic evolution and the family of quantum phase-estimation algorithms, including Shor's algorithm. We state that in quantum algorithms the time arrow is a majorization arrow.

DOI: 10.1103/PhysRevA.66.022305 PACS number(s): 03.67.Lx, 02.30.Lt, 89.70.+c

I. INTRODUCTION

Majorization is the natural ordering on probability distributions. One probability distribution is more uneven than another one when the former majorizes the latter. Furthermore, majorization implies an entropy decrease, thus the ordering concept introduced by majorization is more restrictive and powerful than the one associated to the Shanon's entropy. The goal of this work is to show that all known efficient quantum algorithms obey a majorization principle, in a way to be made precise later.

The classical theory of majorization was first introduced by Muirhead [1] and later developed by Hardy, Littlewood, and Pólya in their study of symmetric means [2]. Majorization was studied by economists in the beginning of the twentieth century in order to formalize the concept of unevenness in the distribution of income. In 1905, Lorenz pointed out that one distribution can be said to be more uneven than another precisely when it majorizes the other [3]. Likewise, Dalton in 1920 stated his *principle of transfers* showing that a distribution is less uneven than another if it can be obtained from the other by transferring some income from a richer to a poorer income-receiver. Moreover, majorization has found many applications in classical computer science like stochastic scheduling, optimal Huffman coding, greedy algorithms, etc.

In quantum information theory, majorization characterizes when two quantum bipartite pure states can be connected via local operations and classical communication [4,5]. This result shows that this connection is indeed possible when there exists majorization between the vectors of eigenvalues (weights) of the partial von Neumann entropies associated to each bipartite state. A further application of majorization in quantum information theory corresponds to the problem of Hamiltonian simulation [6]. There, strong restrictions based on majorization theory limit the possibility to simulate a proposed quantum evolution from a different given Hamiltonian complemented with local unitary transformations. Majorization is also present in quantum measurement theory and in the separability problem.

Majorization is often defined as a binary relation denoted by $\prec$ on vectors in $\mathbb{R}^d$. We need to fix notations by introducing some basic definitions.

Definition 1. For $\mathbf{x},\mathbf{y}\in\mathbb{R}^d$,

$$\mathbf{x}\prec\mathbf{y} \text{ iff } \begin{cases} \sum_{i=1}^{k} x_{[i]} \leqslant \sum_{i=1}^{k} y_{[i]}, & k=1,\ldots,d-1, \\ \sum_{i=1}^{d} x_{[i]} = \sum_{i=1}^{d} y_{[i]}, \end{cases} \tag{1}$$

where $[z_{[1]}\cdots z_{[d]}]\coloneqq \mathrm{sort}_{\downarrow}(\mathbf{z})$ denotes the descendingly sorted (nonincreasing) ordering of $\mathbf{z}\in\mathbb{R}^d$. An immediate consequence is that majorization is a partial order for sorted vectors in $\mathbb{R}^d$.

Definition 2. If it exists, the least element x_1 (greatest element x_g) of a partial order like majorization is defined by the condition $x_1\prec x,\forall x\in\mathbb{R}^d$ ($x\prec x_g,\forall x\in\mathbb{R}^d$).

In this paper we address the following basic problem of elucidating what is the role, if any, played by majorization in the way quantum algorithms operate. We find, indeed, that there is a majorization principle underlying the way quantum algorithms work that we shall now state more precisely. Let us denote by $|\Psi_m\rangle$ the pure state representing the state of the register in a quantum computer at an operating stage labeled by $m=0,1,\ldots,M-1$, where M is the total number of steps of the algorithm. We can associate naturally a set of sorted probabilities $[p_{[x]}],x=0,1,\ldots,2^n-1$ to this quantum state of n qubits in the following way: decompose the register state in the computational basis, i.e., $|\Psi_m\rangle\coloneqq\Sigma_{x=0}^{2^n-1}c_x|x\rangle$ with $\{|x\rangle\coloneqq|x_0x_1\cdots x_{n-1}\rangle\}_{x=0}^{2^n-1}$ denoting the basis states in digital or binary notation, respectively, and $x\coloneqq\Sigma_{j=0}^{n-1}x_j2^j$. The sorted vectors to which majorization theory applies are precisely $[p_{[x]}]\coloneqq[|c_{[x]}|^2]$. Thus in quantum algorithms we shall be dealing with probability densities defined in $\mathbb{R}_+^d$, with $d=2^n$. With these ingredients, our main result can be stated as follows: in the quantum algorithms known so far, the set of sorted probabilities $[p_{[x]}^m]$ associated to the quantum register at each step m are majorized by the corresponding probabilities of the next step,

$$[p^m_{[x]}] \prec [p^{m+1}_{[x]}], \quad \begin{cases} \forall \ m=0,1,\ldots,M-2, \\ x=0,1,\ldots,2^n-1. \end{cases} \tag{2}$$

This is a strong result for it means that majorization works locally in quantum algorithms, i.e., step by step, and not just globally (for the initial and final states). Our starting point is the majorization analysis of Grover's algorithm [7].

II. GROVER'S ALGORITHM

This quantum algorithm solves efficiently the problem of finding a target item in a large database. The algorithm is based on a kernel that acts symmetrically on the subspace orthogonal to the solution. This is clear from its construction

$$K := U_s U_{y_0},$$

$$U_s := 2|s\rangle\langle s| - 1, \quad U_{y_0} := 1 - 2|y_0\rangle\langle y_0|, \tag{3}$$

where $|s\rangle := 1/\sqrt{N}\Sigma_x|x\rangle$ and $|y_0\rangle$ is the searched item.

Theorem. The set of probabilities to obtain any of the N possible states in a database is majorized step by step along the evolution of Grover's algorithm when starting from a symmetric state until the maximum probability of success is reached.

Proof. To prove this result we write $[p_{[x]}]$ as the set of sorted probabilities of finding the state $|x\rangle$ when performing a measurement. We call $[p'_{[x]}]$ the set of sorted probabilities after one single application of Grover's kernel. The theorem is equivalent to prove that $[p_{[x]}] \prec [p'_{[x]}]$ until p_1, the probability of finding the correct solution, reaches its maximum value.

The hypothesis of symmetry imposes that the probabilities of finding each of the N outputs at some point during the implementation of Grover's algorithm can be ordered in the list

$$\left[p, \frac{1-p}{N-1}, \frac{1-p}{N-1}, \ldots, \frac{1-p}{N-1}\right], \tag{4}$$

where p is the one associated to the correct output. After one further action of the kernel these probabilities will be

$$\left[p', \frac{1-p'}{N-1}, \frac{1-p'}{N-1}, \ldots, \frac{1-p'}{N-1}\right]. \tag{5}$$

We first need to prove that Grover's algorithm increases the probability of success monotonically, that is $p' > p$, until it reaches a maximum and then decreases also monotonically. This part of the proof relies on the fact that the Grover algorithm can be described in a reduced two-dimensional space [8,9], which follows from the symmetry of the subspace orthogonal to $|y_0\rangle$. In this case, the dynamics can be reduced to a two-state system, $\{|y_0\rangle, |y_0^\perp\rangle\}$. Grover's kernel on this space acts as a rotation [10]

$$K = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}, \tag{6}$$

where $\cos\theta = 1 - 2/N$. Starting from the symmetric state

$$|s\rangle^t = \left(\frac{1}{\sqrt{N}} \quad \sqrt{1-\frac{1}{N}}\right), \tag{7}$$

m applications of the kernel lead to

$$K^m|s\rangle = \begin{pmatrix} \frac{1}{\sqrt{N}}\cos(m\theta) - \sqrt{1-\frac{1}{N}}\sin(m\theta) \\ \frac{1}{\sqrt{N}}\sin(m\theta) + \sqrt{1-\frac{1}{N}}\cos(m\theta) \end{pmatrix}. \tag{8}$$

The projection onto the upper component corresponds to the probability amplitude which, thus, evolves monotonically until it reaches a maximum.

Returning to the original problem, we can now check that all probabilities evolve in such a way that majorization works smoothly:

$$P \leqslant p',$$

$$\frac{(N-2)p+1}{N-1} \leqslant \frac{(N-2)p'+1}{N-1}, \tag{9}$$

$$\vdots$$

$$\frac{(N-m-1)p+m}{N-1} \leqslant \frac{(N-m-1)p'+m}{N-1}.$$

Thus $[p_{[x]}] \prec [p'_{[x]}]$ and Eq. (2) holds true. ■

Majorization works in a simple way in Grover's algorithm. Nevertheless, the proof does not hold when the initial distribution of probabilities is not symmetric in the subspace orthogonal to the solution. It is indeed easy to find numerical counterexamples to the majorization principle in absence of symmetry. We realize that this corresponds to starting with a quantum state $|s\rangle$ whose set of probabilities is the *least element* of the majorization we have introduced to study quantum algorithms. We shall see that this fact also happens in the rest of the algorithms below.

III. QUANTUM ADIABATIC EVOLUTION ALGORITHMS

Grover's algorithm can be mapped onto the evolution of the homogeneous state $|s\rangle$ into the solution $|0\rangle$ driven by a simple Hamiltonian [8]. Farhi *et al.* have proposed to use the adiabatic evolution to guarantee that the system remains in the fundamental state and reaches the target solution in the end [11]. More precisely, the idea consists of setting up a Hamiltonian of the form

$$H\left(\frac{t}{T}\right) = \left(1 - \frac{t}{T}\right)H_0 + \frac{t}{T}H_1, \tag{10}$$

such that $|s\rangle$ is the ground state of H_0 and $|0\rangle$ is the ground state of H_1. For large enough T, the evolution will be adiabatic and the system will remain in the ground state all along the flow. The adiabatic theorem dictates that T must scale as

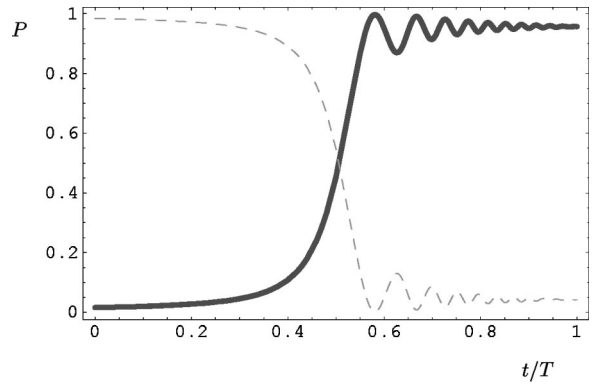

FIG. 1. Evolution of the probability of finding the target state (bold) and other states (dashed) for $n=6$.

the inverse squared of the minimum gap of the system. The question we address here is whether this evolution respects majorization.

Although the system contains n qubits, 2^n possible states, the adiabatic evolution can be computed using a subspace if sufficient symmetry is present. The simplest example is to consider the Hamiltonian

$$H\left(\frac{t}{T}\right)=-|s\rangle\langle s|\left(1-\frac{t}{T}\right)-|0\rangle\langle 0|\frac{t}{T} \tag{11}$$

and the initial state $|s\rangle$. In this particular case, the evolution can be computed using a reduced two state Hilbert space. More precisely

$$|s\rangle=\frac{1}{\sqrt{2^n}}(|0\rangle+\sqrt{2^n-1}|0^{\perp}\rangle). \tag{12}$$

Then the Hamiltonian written in the basis $\{|0\rangle,|0^{\perp}\}$ reads

$$H\left(\frac{t}{T}\right)=-\left(1-\frac{t}{T}\right)\begin{pmatrix}\frac{1}{2^n} & \sqrt{\frac{2^n-1}{2^n}}\\ \sqrt{\frac{2^n-1}{2^n}} & \frac{2^n-1}{2^n}\end{pmatrix}-\frac{t}{T}\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix}. \tag{13}$$

It is possible to verify numerically that when $T\sim 4\times 2^n$ the probability follows the graphic shown in Fig. 1. An argument similar to the previous theorem indicates that symmetry imposes majorization for the complete set of probabilities. Shorter T lead to evolutions that do not hit the solution with probability one, while a larger T smooths this evolution. Once the maximum is attained, the probabilities oscillate and majorization is obviously lost.

It is worth mentioning that a combination of H_0 and H_1 chosen as above but mixed with no time dependence leads to a Hamiltonian that rotates the ground state in the manner of the previous theorem. Then, the solution is obtained in $T=(\pi/2)2^{n/2}$ with probability 1. This is precisely the scaling law found in Grover's algorithm.

A more refined test for the majorization principle corresponds to the Hamiltonian evolution proposed by Farhi *et al.* as a natural starting point for any adiabatic evolution [11]. Let us consider the following choice:

$$H_0=\sum_{i=1,n}(1-\sigma_x)^{(i)}. \tag{14}$$

This Hamiltonian acts as an eraser of information and has the state $|s\rangle$ as its ground state. Furthermore, it allows for a decomposition of the Hilbert space into $n+1$ symmetric subspaces. Finding the target instance $|0\rangle$ amounts to solving the dynamical evolution in this $(n+1)$-dimensional Hilbert space. Let us denote as $|k\rangle$ the symmetric space with k qubits in the state $|1\rangle$ and the rest in $|0\rangle$. The Hamiltonian becomes

$$H_0=\frac{n}{2}I-N, \tag{15}$$

where the elements of the symmetric matrix N are given by

$$\langle i|N(i,j)|j\rangle=\sqrt{j}\sqrt{n-(j-1)}\,\delta_{i+1,j}. \tag{16}$$

A numerical solution of the evolution is now easy to perform. For $T>7\times 2^n$, the system indeed evolves along the ground state and majorization holds for the set of $n+1$ probabilities, as shown in Fig. 2. Shorter evolutions perform poorly and fail to verify the majorization principle. We conclude that quantum algorithms based on adiabatic evolution naturally fulfill a majorization principle provided that the Hamiltonians and initial state are chosen with sufficient symmetry and the evolution is slow enough.

IV. QUANTUM PHASE-ESTIMATION ALGORITHMS

These represent a large family of quantum algorithms that include as particular instances the order-finding problem, Shor's algorithm [12], discrete logarithms, etc. [13]. The basic problem is: given an arbitrary unitary operator U and one

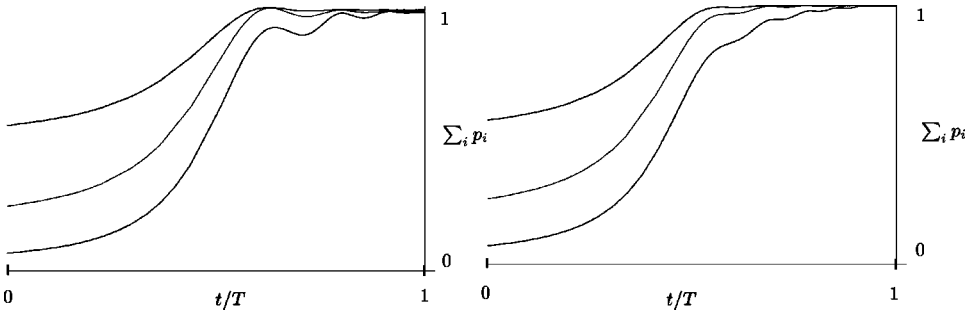

FIG. 2. Curves for p_1, p_1+p_2, and $p_1+p_2+p_3$ for $n=4$. The failure of majorization (monotonicity) for fast evolution, $T=4\times 2^n$, in the upper curves goes away for slower evolution, $T=7\times 2^n$.

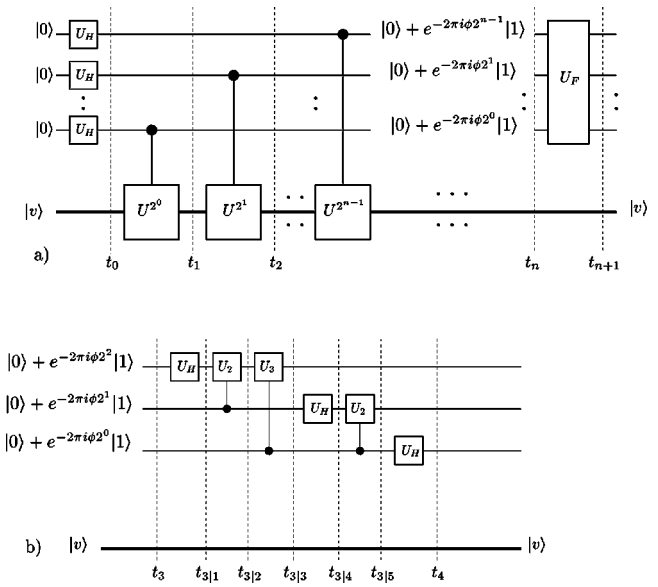

FIG. 3. (a) Quantum circuit implementing the phase-estimation algorithm constructed from Hadamard gates U_H, controlled-U gates acting as $|0\rangle\langle 0|\otimes 1+|1\rangle\langle 1|\otimes U$, and the QFT. Dashed lines represent time steps for majorization testing. (b) An example of QFT decomposition into elementary gates for $n=3$ qubits.

eigenvector $|v\rangle$, estimate the phase ϕ of the corresponding eigenvalue $U|v\rangle := e^{-2\pi i\phi}|v\rangle, \phi\in[0,1)$, with n bits of accuracy. The efficient quantum solution of this problem can be encoded in the quantum circuit shown in Fig. 3, and we shall always refer to this circuit when performing the majorization analysis stepwise. The algorithm clearly has two parts: (i) application of Hadamard gates U_H and controlled-U^j gates, $j=0,1,\ldots,n-1$; and (ii) application of the quantum Fourier transform (QFT) U_F.

Part (i). The whole quantum register is made up of first and second registers. The initialization stage is such that the quantum computer is in the state $|\Psi_{\text{in}}\rangle := |00\cdots 0\rangle|v\rangle$, where the first register has been prepared at the state $|0\rangle$ for short, and the second holds the eigenvector of U. In what follows, we denote by $[p^m_{[x]}]$ the sorted probabilities distributions of the first register, at time steps $m=0,1,\ldots,n+1$ that we show in Fig. 3 as time slices.

Clearly, the probability distribution of $|\Psi_{\text{in}}\rangle$ is the greatest element of the majorization. However, an application of the Hadamard gates yields a lowest element as in Grover's algorithm. Thus our starting point for majorization is $|\Psi_0\rangle := (U_H^{\otimes n}\otimes 1)|\Psi\rangle_{\text{in}} = 2^{-n/2}\Sigma_{x=0}^{2^n-1}|x\rangle|v\rangle$. Then, $[p^0_{[x]}] = [2^{-n}], \forall x$.

Next, a series of controlled-U^{2^j} gates encompassing time steps from t_1 to t_n (Fig. 3) are applied. The outcome of these steps is the factorized state

$$|\Psi_n\rangle = 2^{-n/2}[|0\rangle + e^{-2\pi i 2^{n-1}\phi}|1\rangle]\cdots[|0\rangle + e^{-2\pi i 2^0\phi}|1\rangle] = 2^{-n/2}\sum_{x=0}^{2^n-1} e^{-2\pi i x\phi}|x\rangle|v\rangle. \quad (17)$$

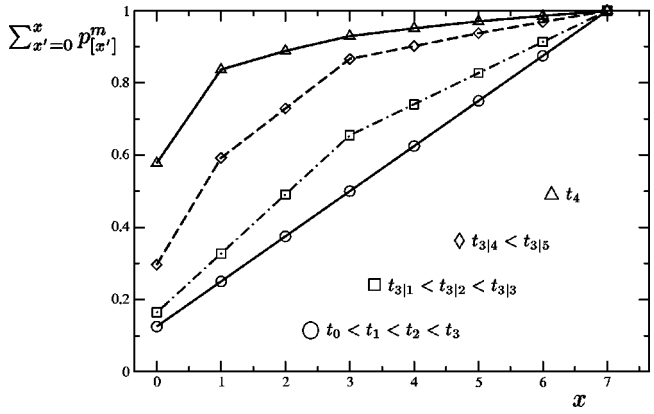

FIG. 4. Lorenz diagram (partial probability sums) for the quantum phase-estimation algorithm with $\phi=0.2$ and $n=3$ qubits as in Fig. 3. It shows how majorization works along the time arrow $\bigcirc \rightarrow \square \rightarrow \Diamond \rightarrow \triangle$.

As the action of these gates only introduces phases locally in the computational states, then we obtain again the uniform distributions $[p^m_{[x]}] = [2^{-n}], \forall x, m=0,1,\ldots,n$.

Part (ii). Although the local phases in $|\Psi_n\rangle$ do not play any role in majorization, so far, they become relevant when combined with the application of the QFT on the first register, due to interference of quantum amplitudes. The state after time step t_{n+1} (Fig. 3) is

$$|\Psi_{n+1}\rangle := (U_F\otimes 1)|\Psi_n\rangle = 2^{-n}\sum_{x,y=0}^{2^n-1} e^{-2\pi i x(\phi - y/2^n)}|y\rangle|v\rangle. \quad (18)$$

Now, $p^{n+1}_{[y]} := |2^{-n}\Sigma_{x=0}^{2^n-1} e^{-2\pi i x(\phi - y/2^n)}|^2$ majorizes the least element distribution at step $m=n$. Interestingly enough, there is a stronger majorization working stepwise when the QFT is applied by means of its canonical decomposition in terms of n Hadamard and $n(n-1)/2$ controlled-phase gates [14]. For concreteness, we show such decomposition in Fig. 3(b) for $n=3$ qubits and with the corresponding time slices (majorization checkpoints). The proof of this result relies on the recursive application of the following inequalities

$$\left|\frac{1}{\sqrt{2}}(1\pm e^{2\pi i\alpha_\pm(y,\phi)})\right|^2 \geqslant 1,$$

$$\{(\alpha_+\in[0,\tfrac{1}{4}],[\tfrac{3}{4},1];\ \ \alpha_-\in[\tfrac{1}{4},\tfrac{3}{4}]), \quad (19)$$

where, at each step, $\alpha_\pm$ depends on y,ϕ in a computable way [15]. To illustrate this fact, we show in Fig. 4 a numerical plot for $n=3$ qubits in the form of a Lorenz diagram: partial probability sums vs x, for each time step. Therefore, as a consequence of our analysis we find that the majorization principle is working locally in algorithms like order-finding $a^r=1 \mod N$, where the unitary operator is given by $U|x\rangle := |ax \mod N\rangle$ and $\phi=1/r$; Shor's algorithm, where order-finding is used combined with controlled-U gates implementing the modular exponentiation; Chuang's

algorithm for quantum clock synchronization, where $U := U_{\mathrm{cnot}} U_{\mathrm{TQP}} U_{\mathrm{cnot}}$ and U_{TQP} is the so-called Ticking Qubit Protocol [16]; etc.

V. CONCLUSIONS

Efficient quantum algorithms are scarce as compared with their classical counterparts, suggesting that we are missing the basic principles for quantum algorithm design [17]. In this paper, we have produced evidence for the general idea that there is a majorization principle acting step by step during the time evolution in efficient quantum algorithms. We may say that majorization is a sort of driving force for such algorithms. Learning to tame majorization may be useful for devising quantum-algorithm design. When majorization is not at work, the quantum algorithm is neither efficient nor successful.

ACKNOWLEDGMENTS

We are grateful to G. Vidal for introducing us to majorization theory ideas and for his advice. We acknowledge financial support from the projects: AEN99-0766, 1999SGR-00097, IST-1999-11053, and PB98-0685.

[1] R.F. Muirhead, Proc. Edinburgh Math. Soc. **21**, 144 (1903).
[2] G. H. Hardy, J. E. Littlewood, and G. Pólya, *Inequalities* (Cambridge University Press, Cambridge, England, 1978).
[3] A. W. Marshall and I. Olkin, *Inequalities: Theory of Majorization and its Applications* (Academic, New York, 1979).
[4] M.A. Nielsen, Phys. Rev. Lett. **83**, 436 (1999).
[5] M.A. Nielsen and G. Vidal, Quantum Inf. Comput. **1**, 76 (2001).
[6] G. Vidal and J.I. Cirac, e-print quant-ph/0108076.
[7] L.K. Grover, Phys. Rev. Lett. **79**, 325 (1997).
[8] E. Farhi and S. Gutmann, Phys. Rev. A **57**, 2403 (1998).
[9] A. Galindo and M.A. Martín-Delgado, Phys. Rev. A **62**, 062303 (2000).
[10] R. Jozsa, e-print quant-ph/9901021.
[11] E. Farhi, J. Goldstone, S. Gutmann, and M. Sipser, e-print quant-ph/0001106.
[12] P. W. Shor, in *Proceedings of the 35th Annual Symposium on the Foundations of Computer Science* (IEEE Computer Society Press, Los Alamitos, CA, 1994), p. 124, e-print quant-ph/9508027.
[13] R. Cleve, A. Ekert, C. Macchiavello, and M. Mosca, Proc. R. Soc. London, Ser. A **454**, 339 (1998).
[14] D. Coppersmith, IBM Research Report No. RC 19642, 1994 (unpublished).
[15] For $n=2$ qubits, the proof is simple but lengthy. Then, the extension to an arbitrary number of qubits follows.
[16] I.L. Chuang, Phys. Rev. Lett. **85**, 2006 (2000).
[17] A. Galindo and M.A. Martín-Delgado, Rev. Mod. Phys. **74**, 347 (2002).