Analog analogue of a digital quantum computation

Edward Farhi*
Center for Theoretical Physics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

Sam Gutmann†
Department of Mathematics, Northeastern University, Boston, Massachusetts 02115

(Received 11 December 1996)

We solve a problem, which while not fitting into the usual paradigm, can be viewed as a quantum computation. Suppose we are given a quantum system with a Hamiltonian of the form $E|w\rangle\langle w|$ where $|w\rangle$ is an unknown (normalized) state. The problem is to produce $|w\rangle$ by adding a Hamiltonian (independent of $|w\rangle$) and evolving the system. If $|w\rangle$ is chosen uniformly at random we can (with high probability) produce $|w\rangle$ in a time proportional to $N^{1/2}/E$. If $|w\rangle$ is instead chosen from a fixed, known orthonormal basis we can also produce $|w\rangle$ in a time proportional to $N^{1/2}/E$ and we show that this time is optimally short. This restricted problem is an analog analogue to Grover's algorithm, a computation on a conventional (!) quantum computer that locates a marked item from an unsorted list of N items in a number of steps proportional to $N^{1/2}$. [S1050-2947(98)09004-0]

PACS number(s): 03.67.Lx

Although a quantum computer, beyond certain elementary gates, has not yet been constructed, a paradigm [1] for quantum computation is in place. A quantum computer is envisaged as acting on a collection of spin 1/2 particles sitting at specified sites. Each elementary operation is a unitary transformation that acts on the spins at one or two sites. A quantum computer program, or algorithm, is a definite sequence of such unitary transformations. For a given initial spin state, the output of the program is the spin state after the sequence of transformations has acted. The length of the algorithm is equal to the number of elementary unitary transformations that make up the algorithm.

This framework for quantum computation is general enough that any ordinary digital computer program can be implemented on a quantum computer. Quantum computers can go beyond ordinary computers when they act on superpositions of states and take advantage of interference effects. An example of a quantum algorithm that outperforms any classical algorithm designed to solve the same problem is the Grover algorithm [2]. There we are given a function $f(a)$ defined on the integers a from 1 to N. The function has the property that it takes the value 1 on just a single element of its domain, w, and it has the value 0 for all $a \neq w$. With only the ability to call the function f, the task is to find w. On a classical computer this requires, on average, $N/2$ calls of the function f. However, Grover showed that with a quantum computer w can be found with of order $N^{1/2}$ function calls. This remarkable speedup illustrates the power of quantum computation. (In the Appendix we explain how the Grover algorithm works.)

In this paper we consider quantum computation differently, as controlled Hamiltonian time evolution of a system, obeying the Schrödinger equation

$$i\frac{d}{dt}|\psi\rangle = H(t)|\psi\rangle, \tag{1}$$

where the Hamiltonian is designed to solve a specified problem. We illustrate this with an example. Suppose we are given a Hamiltonian in an N-dimensional vector space and we are told that the Hamiltonian has one eigenvalue $E \neq 0$ and all the others are 0. The task is to find the eigenvector $|w\rangle$ that has eigenvalue E. We consider this task accomplished when the system is known to be in the state $|w\rangle$. We now describe a solution to two versions of this problem and then discuss in what sense it is optimal.

First, suppose we are given

$$H_w = E|w\rangle\langle w| \tag{2}$$

with $|w\rangle$ chosen uniformly at random from the unit sphere in N-dimensional complex space. In some convenient fashion select a normalized vector $|s\rangle$, which of course does not depend on $|w\rangle$ since we don't yet know what $|w\rangle$ is. (We will have more to say about choosing $|s\rangle$ shortly.) Now add to H_w the ''driving'' Hamiltonian

$$H_D = E|s\rangle\langle s| \tag{3}$$

so that the full Hamiltonian is

$$H = H_w + H_D. \tag{4}$$

We now calculate the time evolution of the state $|\psi_w, t\rangle$, which at $t=0$ is $|s\rangle$, that is,

$$|\psi_w, t\rangle = e^{-iHt}|s\rangle. \tag{5}$$

It suffices to confine our attention to the two-dimensional subspace spanned by $|s\rangle$ and $|w\rangle$. The vectors $|s\rangle$ and $|w\rangle$ are (generally) not orthogonal and we call their inner product x,

$$\langle s|w\rangle = x, \tag{6}$$

*Electronic address: farhi@mitlns.mit.edu
†Electronic address:sgutm@nuhub.neu.edu

where x can be taken to be real and positive since any phase in $\langle s|w\rangle$ can ultimately be absorbed in $|s\rangle$. We will discuss how big x is shortly. Now the vectors

$$|r\rangle=\frac{1}{\sqrt{1-x^2}}(|s\rangle-x|w\rangle) \tag{7}$$

and $|w\rangle$ are orthonormal. In the $|w\rangle,|r\rangle$ basis the Hamiltonian (4) is

$$H=E\begin{bmatrix}1+x^2 & x\sqrt{1-x^2}\\ x\sqrt{1-x^2} & 1-x^2\end{bmatrix} \tag{8}$$

and

$$|s\rangle=\begin{bmatrix}x\\ \sqrt{1-x^2}\end{bmatrix}. \tag{9}$$

Now a simple calculation gives

$$|\psi_w,t\rangle=e^{-iEt}\begin{bmatrix}x\cos(Ext)-i\sin(Ext)\\ \sqrt{1-x^2}\cos(Ext)\end{bmatrix}. \tag{10}$$

Thus we see that at time t the probability of finding the state $|w\rangle$ is

$$P(t)=\sin^2(Ext)+x^2\cos^2(Ext) \tag{11}$$

and that at a time t_m given by

$$t_m=\frac{\pi}{2Ex} \tag{12}$$

the probability is one.

Next consider a restricted version of the problem, in which we are given an orthonormal basis $\{|a\rangle\}$ with $a=1,\ldots,N$, and we know that $|w\rangle$ is one of these N basis vectors. In this case it is most convenient to let

$$|s\rangle=\frac{1}{\sqrt{N}}\sum_{a=1}^{N}|a\rangle \tag{13}$$

and it then follows that $x=N^{-1/2}$. The initial state $|s\rangle$ evolves to $|w\rangle$ in a time $(\pi N^{1/2}/2E)$. This is directly analogous to Grover's problem.

Note that the eigenvalues of the Hamiltonian (8) are $E(1\pm x)$. Thus the difference in eigenvalues is $(2xE)$, which is $2E/N^{1/2}$ for $|s\rangle$ given by Eq. (13). By the time-energy uncertainty principle, the time required to evolve substantially [3], that is, from $|s\rangle$ to $|w\rangle$, must be of order $N^{1/2}/E$, which is the time we found. You might think that by increasing the energy difference, for example by using $H_D=E'|s\rangle\langle s|$ with $E'\gg E$, you could speed up the procedure for producing $|w\rangle$. However, the next result shows that this is not the case.

We now show that our procedure for producing $|w\rangle$, when $|w\rangle$ is an unknown member of a given orthonormal basis, is optimally short. The proof we give here is the analog analogue of the oracle proof [4] that can be used to show that the Grover algorithm is optimal for the problem it sets out to solve. Again we are given the Hamiltonian $H_w=E|w\rangle\langle w|$ and we wish to add some Hamiltonian $H_D(t)$ to it which drives the system to $|w\rangle$. Because $\{|w\rangle\}$ is a basis we have

$$\sum_w H_w=E\sum_w|w\rangle\langle w|=E. \tag{14}$$

Now start with some initial $|w\rangle$-independent state $|i\rangle$ and evolve it with the Hamiltonian

$$H=H_w+H_D(t). \tag{15}$$

We want to find a lower bound on the time t_0 required for the state $|i\rangle$ to evolve with the Hamiltonian $H_w+H_D(t)$ into $|w\rangle$. (The time t_0 is assumed independent of the state $|w\rangle$.) Let

$$i\frac{d}{dt}|\psi_w,t\rangle=[H_w+H_D(t)]|\psi_w,t\rangle \tag{16}$$

with

$$|\psi_w,0\rangle=|i\rangle.$$

Let $|\psi,t\rangle$ evolve with $H_D(t)$, that is,

$$i\frac{d}{dt}|\psi,t\rangle=H_D(t)|\psi,t\rangle \tag{17}$$

with

$$|\psi,0\rangle=|i\rangle.$$

At t_0 we have $|\psi_w,t_0\rangle=|w\rangle$. Therefore

$$\begin{aligned}\sum_w\||\psi_w,t_0\rangle-|\psi,t_0\rangle\|^2&=\sum_w\||w\rangle-|\psi,t_0\rangle\|^2\\ &=2N-\sum_w(\langle w|\psi,t_0\rangle+\langle\psi,t_0|w\rangle)\\ &\geqslant 2N-2\sqrt{N}\geqslant N\end{aligned} \tag{18}$$

for $N\geqslant 4$. Now consider

$$\frac{d}{dt}\||\psi_w,t\rangle-|\psi,t\rangle\|^2=-2\,\mathrm{Re}\frac{d}{dt}\langle\psi_w,t|\psi,t\rangle, \tag{19}$$

which upon using Eqs. (16) and (17) gives

$$\begin{aligned}\frac{d}{dt}\||\psi_w,t\rangle-|\psi,t\rangle\|^2&=2\,\mathrm{Im}\langle\psi_w,t|H_w|\psi,t\rangle\\ &\leqslant 2|\langle\psi_w,t|H_w|\psi,t\rangle|\\ &\leqslant 2\|H_w|\psi,t\rangle\|.\end{aligned} \tag{20}$$

We now sum on w and (again) use the fact that if $\Sigma_{i=1}^N|a_i|^2=1$ then $\Sigma_{i=1}^N|a_i|\leqslant N^{1/2}$ along with Eq. (14) to obtain

$$\frac{d}{dt}\sum_w\||\psi_w,t\rangle-|\psi,t\rangle\|^2\leqslant 2EN^{1/2}. \tag{21}$$

Since $|\psi_w,0\rangle=|\psi,0\rangle$ we have

$$\sum_w \| |\psi_w,t\rangle - |\psi,t\rangle \|^2 \leq 2EN^{1/2}t. \tag{22}$$

Applying Eq. (22) at time t_0 and using Eq. (18) we have

$$t_0 \geq \frac{N^{1/2}}{2E}. \tag{23}$$

This shows that the H_D we chose in Eq. (3) allows us to produce $|w\rangle$ in a time that is within a constant factor of the best possible.

We now return to the question of finding $|w\rangle$ when it is an arbitrary normalized vector in an N-dimensional vector space, not just one of a list of N possible orthonormal vectors. In this case we assume $|w\rangle$ is picked at random from a uniform distribution on the $(2N-1)$-dimensional unit sphere. We now choose $|s\rangle$ to be some fixed unit vector. We can select $|s\rangle$ based on some criterion such as the ease of constructing the driving Hamiltonian $H_D = E|s\rangle\langle s|$. For a fixed $|s\rangle$ how big do we expect $x = |\langle s|w\rangle|$ to be? Since $|w\rangle$ is uniformly distributed on the $(2N-1)$-dimensional unit sphere we can, without loss of generality, pick $|s\rangle$ to be at the north pole. Using spherical coordinates we find

$$\mathrm{Prob}(x \leq x_0) = \frac{\int_{\cos^{-1}(x_0)}^{\pi/2} (\sin\theta)^{2N-2} d\theta}{\int_0^{\pi/2} (\sin\theta)^{2N-2} d\theta}. \tag{24}$$

This implies that for N large enough,

$$\mathrm{Prob}(x \geq \epsilon/N^{1/2}) \geq 1 - 2\epsilon \tag{25}$$

and in fact, as $N \to \infty$, $\sqrt{2N}x$ converges in distribution to the absolute value of a unit-variance Gaussian random variable.

For the case of $|w\rangle$ chosen as a random unit vector, we see that the time t_m [see Eq. (12)] for our system to evolve from $|s\rangle$ to $|w\rangle$ is typically of order $N^{1/2}/E$. In fact, by repeated measurements, we can determine t_m exactly. (Here we are assuming that we can repeatedly access the quantum system with the fixed H_w.) We run the system, starting in $|s\rangle$ with $H = H_s + H_w$, for a sequence of times of order $N^{1/2}/E$. At the end of each run we measure the operator $|s\rangle\langle s|$. Since the probability of finding the system in $|s\rangle$ is the periodic function $1-(1-x^2)\sin^2(\pi t/2t_m)$ we can determine t_m. The number of measurements required to produce a desired accuracy for t_m does not depend on N. After determining t_m if we run the system starting in $|s\rangle$ for a time t_m we are guaranteed to be in the state $|w\rangle$ [see Eq. (11)].

The strategy outlined above will not succeed if x is smaller than of order $N^{-1/2}$. If we view $|w\rangle$ as being uniformly distributed on the unit sphere, then as a consequence of Eq. (24), for any $\delta > 0$,

$$\mathrm{Prob}\left(x \leq \frac{1}{N^{1/2+\delta}} \right) \leq \frac{2}{N^{\delta}}, \tag{26}$$

which is very small for N large enough. (In the unlikely event that x is too small, it is natural to try again with a new $|s\rangle$.) It would be interesting to find out if there exists an algorithm that outperforms the one described above for this continuous version of the problm.

E.F. would like to thank the theory group at Università di Roma 1 for their hospitality and discussions as well as the INFN for partial support. This work was supported in part by the Department of Energy under Cooperative Agreement No. DE-FC02-94ER40818.

APPENDIX: GROVER'S ALGORITHM

We are given a function $f(a)$ with $a = 1, \ldots, N$ such that $f(w) = 1$ and $f(a) = 0$ for $a \neq w$. We assume that the function $f(a)$ can be calculated using ordinary (reversible) computer code. The goal is to find w. Classically this requires, on average, $N/2$ evaluations of the function f.

We now explain how the Grover algorithm solves this problem; see also [5]. The quantum computer acts on a vector space that has an orthonormal basis $|a\rangle$ with $a = 1, \ldots, N$. It is possible to write a quantum computer algorithm that implements the unitary transformation

$$U_f|a\rangle = (-1)^{f(a)}|a\rangle. \tag{A1}$$

Equivalently we can write

$$U_f = 1 - 2|w\rangle\langle w|. \tag{A2}$$

The quantum computer algorithm that implements U_f requires two evaluations of the function f because it is necessary to erase certain work bits, which we have suppressed.

Now consider the vector

$$|s\rangle = \frac{1}{N^{1/2}} \sum_a |a\rangle. \tag{A3}$$

It is also possible to write quantum computer code that implements the unitary operator

$$U_s = 2|s\rangle\langle s| - 1. \tag{A4}$$

The number of two bit operations required to implement U_s grows more slowly than N to any positive power.

The Grover algorithm consists of letting the operator $U_s U_f$ act k times on the vector $|s\rangle$. To see what happens we can restrict our attention to the two-dimensional subspace spanned by $|s\rangle$ and $|w\rangle$. Let

$$|r\rangle = \frac{1}{\sqrt{N-1}} \sum_{a \neq w} |a\rangle \tag{A5}$$

so that $|w\rangle$ and $|r\rangle$ form an orthonormal basis for the relevant subspace. In the $|w\rangle, |r\rangle$ basis the operator $U_s U_f$ takes the form

$$U_s U_f = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}, \tag{A6}$$

where $\cos\theta = 1 - 2/N$. This implies that

$$(U_s U_f)^k = \begin{bmatrix} \cos(k\theta) & -\sin(k\theta) \\ \sin(k\theta) & \cos(k\theta) \end{bmatrix}. \tag{A7}$$

Now for N large $\theta \sim 2N^{-1/2}$ so each application of $U_s U_f$ is a rotation by an angle $\sim 2N^{-1/2}$. In the $|w\rangle, |r\rangle$ basis, the initial state $|s\rangle$ is

$$|s\rangle = \begin{bmatrix} N^{-1/2} \\ \left(1 - \frac{1}{N}\right)^{1/2} \end{bmatrix}, \tag{A8}$$

which is very close to $|r\rangle$. However, after k steps where $k\theta \approx \pi/2$ the algorithm has rotated the initial state to lie (almost) along $|w\rangle$. This requires $k \sim \pi N^{1/2}/4$ steps. Each step actually requires two evaluations of f so the number of evaluations of f required to find w grows like $N^{1/2}$.

[1] For a summary and references see A. Barenco, C. H. Bennett, R. Cleve, D. P. DiVincenzo, N. Margolus, P. Shor, T. Sleator, J. A. Smolin, and H. Weinfurter, Phys. Rev. A **52**, 3457 (1995).
[2] L. K. Grover, Phys. Rev. Lett. **78**, 325 (1997).
[3] N. Margolus and L. B. Levitin, in *Phys. Comp. 96*, edited by T. Toffoli, M. Biafore, and J. Leão (New England Complex Systems Institute, Cambridge, 1996), pp. 208–211.
[4] C. H. Bennett, E. Bernstein, G. Brassard, and U. V. Vazirani, e-print quant-ph/9701001.
[5] M. Boyer, G. Brassard, P. Høyer, and A. Tapp, in *Phys. Comp. 96* (Ref. [3]), pp. 36–43.