Efficient factorization with a single pure qubit and $logN$ mixed qubits

S. Parker and M.B. Plenio
Optics Section, The Blackett Laboratory, Imperial College, London SW7 2BW, England

(February 1, 2008)

It is commonly assumed that Shor's quantum algorithm for the efficient factorization of a large number N requires a pure initial state. Here we demonstrate that a single pure qubit together with a collection of $log_2 N$ qubits in an arbitrary mixed state is sufficient to implement Shor's factorization algorithm efficiently.

Pacs No: 03.67.-a, 3.67.Lk

The discovery of a quantum algorithm for the efficient factorization of large numbers [1] has started a rapid development of quantum information processing [2]. Following this ground-breaking result a number of experimentally realizable proposals for the implementation of quantum computers have been made, for example, in ion trap systems [3] or Nuclear Magnetic Resonance (NMR) schemes [4]. These systems are distinguished by a low decoherence rate combined with a comparatively high gate speed and therefore promise the possibility of executing many quantum gates. While noise in these systems can be made small in principle, it nevertheless imposes limitations to the maximal size of the computation [5] and to the achievable quality (e.g. the purity) of the initial state of the quantum computer. It would therefore be interesting to see whether a quantum computation necessarily requires the preparation of an initial state of high purity, or whether some parts of the quantum computer may be left in a mixed state. Such a result would be of particular interest in NMR systems in which it is difficult to prepare physically pure quantum states of nuclear spins.

The use of mixed states in quantum algorithms has had little discussion as yet. Note, however, the work of Schulman and Vazirani [7] in which they demonstrated that, starting from a set of qubits each in a thermal state, one can obtain a certain number of pure qubits using a quantum algorithm. These were then envisaged to be used for a quantum computation, while all the other qubits which are in a mixed state are discarded. If the initial states are in a thermal mixture at high temperature, the number of mixed quantum states and quantum gates required to obtain even a single pure qubit is very high. It would greatly enhance the efficiency of this approach if it would be possible to reduce the necessary number of pure qubits as much as possible at the expense of employing some of the mixed qubits in the actual quantum computation. Recently, Knill and Laflamme [6] have investigated the power of quantum computation when only a single pure qubit together with a supply of *maximally* mixed states is available. They were able to construct a problem that such a system can solve more efficiently than the best currently *known* classical algorithm.

It would be interesting to see whether these ideas can be extended to other problems of practical relevance. In this paper we demonstrate that a *single* pure qubit together with an initial supply of $\log_2 N$ qubits in an arbitrarily mixed state is sufficient to implement Shor's algorithm for the factorization of the number N efficiently. This is the smallest number of pure states that can achieve this task. We also demonstrate that the efficiency of the modified algorithm is essentially independent of the degree of mixing of the $\log_2 N$ qubits.

We proceed by outlining the problem addressed in Shor's algorithm, followed by the formulation of Shor's algorithm introduced in [8]. Then we will describe the necessary modifications to this algorithm, that will allow it to be executed using a single pure qubit and $\log_2 N$ qubits in a maximally mixed state.

The basis of Shor's algorithm is a classical order finding method which, recast as a quantum algorithm, can be executed in polynomial time, requiring only a polynomial amount of additional classical computation to compute the factors of N. The factors of a number $N = pq$ can, with high probability, be found if the period or *order*, r, (the lowest positive integer $x \neq 0$ such that $f_a(x) = 1$) of the element a in the space of the function $f_a(x) = a^x \bmod N$, is known. Then, provided a is coprime to N (which can be checked classically in polynomial time using Euclid's algorithm), there is a high probability that $\gcd(a^{\frac{r}{2}} \pm 1, N)$ yields a factor of N, where $\gcd(\alpha, \beta)$ denotes the greatest common divisor of α and β which, again, can be determined efficiently using Euclid's algorithm [1].

We begin by examining the formulation of Shor's algorithm as given in [8] and use it as a basis to demonstrate the main result of this paper. First of all we introduce the transformation $U_a |x\rangle = |ax \bmod N\rangle$ where $x = 0, \cdots, N-1$. Provided a is coprime to N this is a unitary transformation and has eigenvectors

$$|\psi_j\rangle = \sum_{k=0}^{r-1} e^{\frac{-2\pi ijk}{r}} \left|a^k \bmod N\right\rangle \qquad j = 0, \cdots, r-1 \quad (1)$$

with corresponding eigenvalues $e^{\frac{2\pi ij}{r}}$. Given one of these eigenvectors we can apply U_a to it and the value of r will be encoded in the phase, $e^{\frac{2\pi ij}{r}}$. This, however, is a global phase which we cannot measure so instead we can use the "phase-kickback" technique [8] requiring the conditional unitary transformation given by

$$cU_a\,|0\rangle\,|x\rangle \;=|0\rangle\,|x\rangle\;;\;cU_a\,|1\rangle\,|x\rangle=|1\rangle\,|ax \bmod N\rangle\,. \qquad (2)$$

The effect of applying the controlled unitary transform to the state $(|0\rangle+|1\rangle)\,|\psi_j\rangle$ is

$$cU_a(|0\rangle+|1\rangle)\,|\psi_j\rangle=(|0\rangle+e^{\frac{2\pi ij}{r}}\,|1\rangle)\,|\psi_j\rangle \qquad (3)$$

'kicking' the 'global' phase shift acquired on the second qubit into a relative phase in the first qubit. We can now perform measurements on the first qubit which will allow us to estimate r, however, we cannot create the eigenstates of U_a without knowledge of r. Instead one can use the fact [8] that $\sum_{j=0}^{r-1}|\psi_j\rangle=|1\rangle$ and conditionally apply U_a to the state $|1\rangle$ (which obviously requires no knowledge of r) in the second qubit

$$cU_a(|0\rangle+|1\rangle)\,|1\rangle=\sum_{j=0}^{r-1}(|0\rangle+e^{\frac{2\pi ij}{r}}\,|1\rangle)\,|\psi_j\rangle\,. \qquad (4)$$

This state is, of course, entangled, so when we make measurements on the first qubit we will get an estimate of $e^{\frac{2\pi ij}{r}}$, with j (which corresponds to an eigenstate) selected at random.

How do we estimate this phase and the value of r accurately? The network in Fig. 1 will give us, with a sufficient probability, the best L-bit estimate of the value of $2^L j/r$ [8].

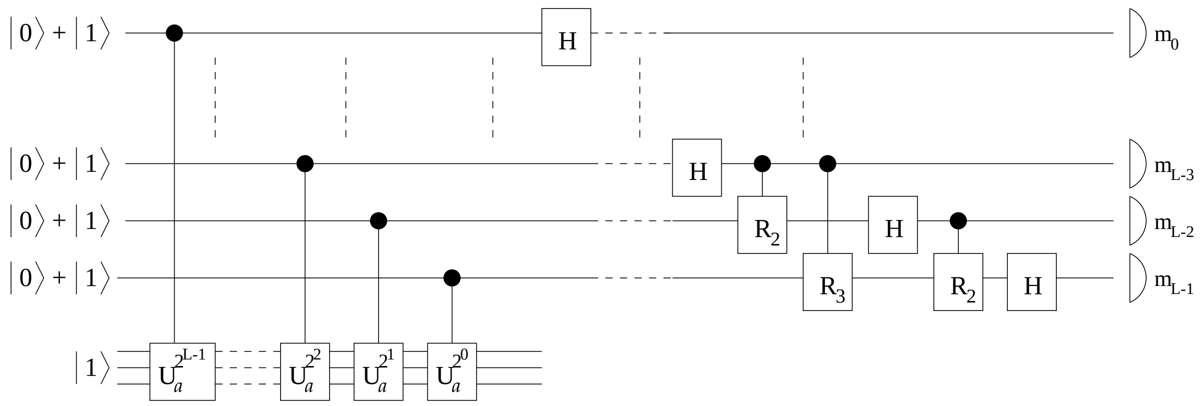

FIG. 1. An implementation of Shor's algorithm [8]. The controlled U_a operations produce phase shifts related to the order of U_a and the remaining Hadamard transformations (H) and controlled rotations $R_j=\left(\begin{smallmatrix}1&0\\0&\phi_j\end{smallmatrix}\right)$ with $\phi_j=e^{-2\pi i/2^j}$ implement the inverse Fourier transform.

As the algorithm proceeds it uses the controlled $U_a, U_a^2, U_a^{2^2}, \cdots, U_a^{2^L}$ transformations to produce the 'kicked' phases $e^{\frac{2\pi ij}{r}}, e^{\frac{2^2\pi ij}{r}}, e^{\frac{2^3\pi ij}{r}}, \cdots, e^{\frac{2^{L-1}\pi ij}{r}}$ into the upper 'control' qubits. The remaining operations on the control qubits realise the quantum inverse Fourier transform. A measurement on each of these qubits produces a binary number $c=\sum_{i=0}^{L-1}2^i m_i$ such that with a finite probability $c/2^L$ is the best estimate of j/r for some integer j again selected at random on measurement.

The first modification to this algorithm comes when we notice that the gates within the Fourier transform are applied sequentially on the qubits. Thus instead of performing the entire transform and then making measurements on all control qubits afterwards we may apply the single qubit (Hadamard) operation to the first qubit and then measure it. The operations (controlled phase shifts) controlled by this first qubit are then replaced by single qubit operations *given the result of the measurement on the first.* This 'semi-classical' modification [10] preserves the probabilities of all measurement results.

Taking this further we need only insist on one control qubit and the remaining $\lceil\log_2 N\rceil$ qubits as we can 'recycle' the control qubit after each measurement (Fig. 2): we perform all the necessary operations of the first control qubit including measurements, followed by all the operations of the second control qubit *on the same physical qubit system* given the results of previous measurements, and so on [11].

We can, therefore, already implement Shor's algorithm with $1+\lceil\log_2 N\rceil$ pure qubits that is, one control qubit and $\lceil\log_2 N\rceil$ of the remaining qubits. We will find later that we can also replace the $\lceil\log_2 N\rceil$ pure qubits with $\lceil\log_2 N\rceil$ maximally mixed qubits and find the order r efficiently (see also [12]). To see why this is the case we first need to examine the unitary transformation U_a more closely.

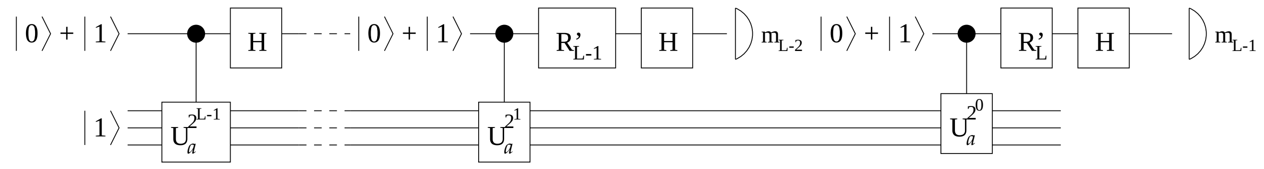

FIG. 2. An implementation of Shor's algorithm using only one control qubit which is recycled. R'_j are now combinations of the rotations R_j given the results of previous measurements: $R'_j = \left(\begin{smallmatrix}1 & 0\\ 0 & \phi'_j\end{smallmatrix}\right)$ with $\phi'_j = e^{-2\pi i \sum_{k=2}^{j} m_{j-k}/2^k}$.

The unitarity of the transform together with the fact that it maps a 'number' state $|x\rangle$ to a 'number' state $|ax \bmod N\rangle$ means that on repeated application of U_a periodic sequences are induced on all the numbers $x = 0, 1, \cdots, N-1$, that is, there is an $R(x)$ such that $U_a^{R(x)}\,|x\rangle = |x\rangle$. We may write the members of all possible sequences as $|ga^x \bmod N\rangle$ for some g and x. For example, for $a = 2$ and $N = 15$ on repeated application of U_a the possible sequences are

$$\begin{aligned}
g = 1: &\quad |1\rangle \rightarrow |2\rangle \rightarrow |4\rangle \rightarrow |8\rangle \rightarrow |1\rangle \\
g = 3: &\quad |3\rangle \rightarrow |6\rangle \rightarrow |12\rangle \rightarrow |9\rangle \rightarrow |3\rangle \\
g = 5: &\quad |5\rangle \rightarrow |10\rangle \rightarrow |5\rangle \\
g = 7: &\quad |7\rangle \rightarrow |14\rangle \rightarrow |13\rangle \rightarrow |11\rangle \rightarrow |7\rangle\,. \qquad (5)
\end{aligned}$$

It is the first of these sequences (with $g = 1$) whose number of members is what we previously called the 'order', r, of a modulo N and it is this period that we need to find to factorize N. However, there is a relationship between the order of the sequence with $g = 1$ and the orders of all the other sequences with $g \neq 1$. We will label each of the different sequences by d and the number of members in each sequence by r_d. U_a obeys the condition $U_a^r = I$ so it is clear that $r_d | r$, that is, the orders of all the sequences divides that of the sequence with $g = 1$. In fact we will find that nearly all of the numbers $0, 1, \cdots, N-1$ are contained within a sequence that has the *same* order as the first sequence. We can find a lower bound on the probability that for a number $g \in 0, 1, \cdots, N-1$ the state $|ga^x \mathrm{mod} N\rangle$ is contained within a sequence of order r.

Theorem 1 *Given two prime numbers p and q we define r as the lowest positive integer x such that $a^x - 1 \equiv 0 \, mod\,(pq)$ for an arbitrary integer a. Then $ga^x - g \equiv 0 \, mod\,(pq)$ with $x < r$ for at most $p + q - 1$ values of g in the interval $0 \leq a \leq pq - 1$.*

Proof: If $\gcd(g, pq) = 1$ then $g(a^x - 1) \equiv 0 \bmod pq \Rightarrow a^x - 1 \equiv 0 \bmod pq$ and therefore $x = r$. There are $(p-1)(q-1)$ positive integers less than and coprime to pq, which proves the theorem □

We can now see that the probability, P_r, of picking g at random such that the lowest x for which $ga^x \equiv g \bmod (pq)$ is r, is $P_r \geq (pq - (p+q-1))/pq = (p-1)(q-1)/pq$ which approaches unity as p and q become large.

This tells us that if we set up an algorithm that actually finds the order of a random sequence we still have a good chance that this order is in fact r.

The r eigenstates of U_a in equation 1 are orthogonal superpositions of the members of the sequence with $g = 1$. In exactly the same way we can form the remaining $N - r$ eigenstates of U_a as orthogonal superpositions of members of each of the other sequences. We write these as

$$\left|\psi_{j_d}^d\right\rangle = \sum_{k=0}^{r_d - 1} e^{\frac{-2\pi i j_d k}{r_d}} \left|g_d a^k \bmod N\right\rangle \qquad (6)$$

where d labels the sequence and $j_d = 0, \cdots, r_d - 1$ the eigenstates of U_a within the sequence d. $|g_d\rangle$ is the lowest member of the dth sequence. Each eigenstate has corresponding eigenvalue $e^{2\pi i j_d / r_d}$ so using the same phase estimation techniques allows us to estimate j_d/r_d given the state $\left|\psi_{j_d}^d\right\rangle$. Again, this requires knowledge of the sequences induced by U_a so instead we may perform the phase estimation technique on the maximally mixed state

$$\frac{\mathbf{1}}{N} = \frac{1}{N} \sum_{k=0}^{N-1} |k\rangle \langle k| = \frac{1}{N} \sum_d \sum_{j_d = 0}^{r_d - 1} \left|\psi_{j_d}^d\right\rangle \left\langle \psi_{j_d}^d\right|\,. \qquad (7)$$

Phase estimation now estimates the value of j_d/r_d for j_d and d chosen at random but as we have seen above nearly all the orders r_d are equal to r.

Note that in Shor's original algorithm the $\lceil \log_2 N \rceil$ qubits encode a phase change into the control qubits which is quantum mechanically correlated to eigenstates of U_a our modification encodes a phase change which is *classically* correlated to the eigenstates. This includes not only the group of eigenstates consisting of superpositions of elements in the first sequence (see Eq. 5) but groups of eigenstates consisting of superpositions of elements in each of the other sequences. However by theorem 1 most of these sequences have the same order and will encode the value $r_d = r$ into the control qubits. This makes it intuitively clear that the algorithm is still efficient. Note however that although the $\lceil \log_2 N \rceil$ mixed qubits are only classically correlated to the pure qubit, entanglement still exists in the system: one can partition the system into two halves one containing some mixed qubits and the other containing the remaining mixed qubits and the pure qubit. Then it can be checked, that this bipartite system can have negative partial transpose and is therefore entangled [13].

In the following we will prove strictly that this modified version of Shor's algorithm is indeed still efficient for order finding. Shor's algorithm requires $O(\log\log r)$ repetitions for it to have a high chance of finding the order whereas the mixed state Shor's algorithm uses exactly the same resources as Shor's original algorithm but requires

$$O\left(\frac{pq}{(p-1)(q-1)}\log\log r\right) \tag{8}$$

repetitions for it to have a high chance of finding the order which, in the limit $p, q \to \infty$, is equally as efficient as Shor's algorithm. For simplicity we will prove this efficiency result for a mixed state algorithm with L control qubits. For the reasons outlined above the result will be identical using a single pure control qubit. The proof follows very closely that of Shor [1].

Pick an L such the $N^2 < t = 2^L < 2N^2$. The initial state of our system with all the control qubits grouped into the first state is

$$\rho_{ini} = \frac{1}{Nt}\sum_{a=0}^{t-1}\sum_{b=0}^{t-1}|a\rangle\langle b| \otimes \sum_{d}\sum_{j_d=0}^{r_d-1}\left|\psi_j^d\right\rangle\left\langle\psi_j^d\right|. \tag{9}$$

Application of the controlled $U_a, U_a^2, \cdots, U_a^{2^{L-1}}$ gates and the inverse Fourier transform yields the state

$$\rho_2 = \frac{1}{Nt^2}\sum_{d}\sum_{j_d=0}^{r_d-1}\sum_{a,b,k,l=0}^{t-1} e^{2\pi i a\left(\frac{j_d}{r_d}-\frac{k}{t}\right)} e^{-2\pi i b\left(\frac{j_d}{r_d}-\frac{l}{t}\right)}|k\rangle\langle l| \otimes \left|\psi_{j_d}^d\right\rangle\left\langle\psi_{j_d}^d\right|. \tag{10}$$

We now make a measurement on the first state. The probability that the result c is obtained is

$$P(c) = \frac{1}{Nt^2}\sum_{d}\sum_{j_d=0}^{r_d-1}|S|^2\,,\;\; S = \sum_{a=0}^{t-1} e^{2\pi i a\left(\frac{j_d}{r_d}-\frac{c}{t}\right)}. \tag{11}$$

S is just an arithmetic progression and $|S|^2$ can easily be bounded by

$$|S|^2 > \frac{4t^2}{\pi^2} \qquad \text{for} \qquad \left|\frac{j_d}{r_d}-\frac{c}{t}\right| < \frac{1}{2t}. \tag{12}$$

Because $t > N^2$ this is a sufficient condition that given c/t there is only one fraction j_d/r_d with $r_d < N$ such that the above condition is obeyed. For a given measurement result c there are at least $(p-1)(q-1)/r$ corresponding values of r_d with $r_d = r$ by theorem 1. So the probability that c/t is the best estimate of a fraction with denominator r is

$$P'(c) > \frac{1}{Nt^2}\sum_{d}\sum_{j_d=1}^{r_d-1}|S|^2 > \frac{4(p-1)(q-1)}{N\pi^2 r}. \tag{13}$$

We now require that the numerator, j_d, is coprime to r otherwise cancellation of common factors will occur in j_d/r. There are $\phi(r)$ values of j_d which are less than and coprime to r, where ϕ is Euler's totient function [9]. Thus the probability that we can calculate r is $P > 4(p-1)(q-1)\phi(r)/Nr\pi^2$. Using a theorem by Hardy and Wright (theorem 328) [9] that $\phi(r)/r > \delta/\log\log r$ for some constant δ we find that the number of times that we need run the algorithm to have a high chance of finding the period, r □is given by Eq. (8).

We have thus found that one pure qubit and a supply of maximally mixed qubits is sufficient to implement Shor's algorithm, requiring no more resources in terms of quantum operations or physical systems than the algorithm operating on pure quantum states. This implies that the algorithm presented here is a 'true' quantum algorithm, achieving an exponential speedup using only polynomial resources. This may be suprising as the degree of mixing of the state of the computer is high. However, the mixing decrease as the algorithm proceeds but never below a mixture of N/r_d eigenstates where r_d is the *measured* period. Furthermore, it should be noted that despite this strong degree of mixing the quantum computer actually evolves into an entangled state. It is this entanglement that appears to be responsible for the computational speedup.

Maximally mixed states are intuitively a less 'costly' resource than pure states but, in fact, we do not need to require *maximally* mixed states: we could equally well use any random state (mixed or pure) on which to perform the controlled U_a operations. The *average* efficiency over all these states would then be as we have shown in this paper. In particular thermal states of nuclear spins (e.g. in NMR), where the occupation of the ground state is only slightly greater than that of the first excited state, would change the efficiency of this algorithm by only a small amount leaving it an efficient algorithm. This ability of highly mixed states to support efficient quantum computation points towards the possibility of the implementation of true quantum computation for example in NMR systems.

The authors would like to thank Kevin Buzzard for valuable advice on number theory and S. Bose, R. Jozsa, R. Laflamme, M. Nielsen and C. Zalka for helpful comments. This work is supported by the EPSRC, the Leverhulme Trust, and the EU project EQUIP.

[1] P. W. Shor, SIAM J. Computing **26** 1484 (1997).

[2] A. Ekert and R. Jozsa, Rev. Mod. Phys **68**, 733 (1996); V. Vedral and M.B. Plenio, Prog. Quant. Elect. **22**, 1 (1998).

[3] J.I. Cirac and P. Zoller, Phys. Rev. Lett. **74**, 4091 (1995).

[4] D.G. Cory et. al., Proc. Natl. Acad. Sci. **94**, 1634 (1997); N.A. Gershenfeld and I.L. Chuang, Science **275**, 350 (1997).

[5] D.P. DiVincenzo, Science **270**, 255 (1995), M.B. Plenio and P.L. Knight, Phys. Rev. A **53**, 2986 (1996); M.B. Plenio and P.L. Knight, Proc. Roy. Soc. A **453**, 2017 (1997).

[6] E. Knill and R. Laflamme, Phys. Rev. Lett. **81**, 5672 (1998).

[7] L.J. Schulman and U. Vazirani, quant-ph/9804060.

[8] R. Cleve et. al., Complexity **4**, 33 (1998); R. Cleve et. al., Proc. Roy. Soc. A **454**, 339 (1998).

[9] G.H. Hardy and E.M. Wright, *An Introduction to the Theory of Numbers*, 5th ed., (Clarendon Press, Oxford 1984).

[10] R. B. Griffiths and C.-S. Niu, Phys. Rev. Lett. **76**, 3228 (1996)

[11] M. Mosca and A. Ekert, *The hidden subgroup problem and eigenvalue estimation on a quantum computer*, Proceedings of the 1st NASA International Conference on Quantum Computing and Quantum Communication, Palm Springs, USA, Lecture Notes in Computer Science 1509 (1999), 174-188.

[12] M. Mosca, page 54, PhD thesis, University of Oxford, 1999.

[13] A. Peres, Phys. Rev. Lett. **77** 1413 (1996); M. Horodecki *et al*, Phys. Lett. A **223**, 1 (1996).