

Schumacher's quantum data compression as a quantum computation

Richard Cleve

Department of Computer Science, University of Calgary, Calgary, Alberta, Canada T2N 1N4

David P. DiVincenzo

IBM Research Division, Thomas J. Watson Research Center, P.O. Box 218, Yorktown Heights, New York 10598

(Received 8 March 1996)

An explicit algorithm for performing Schumacher's noiseless compression of quantum bits is given. This algorithm is based on a combinatorial expression for a particular bijection among binary strings. The algorithm, which adheres to the rules of reversible programming, is expressed in a high-level pseudocode language. It is implemented using $O(n^3)$ two- and three-bit primitive reversible operations, where n is the length of the qubit strings to be compressed. Also, the algorithm makes use of $O(n)$ auxiliary qubits. Space-saving techniques based on those proposed by Bennett are developed which reduce this workspace to $O(\sqrt{n})$ while maintaining a running time of $O(n^3)$ (albeit with a larger constant). This latter algorithm is of interest because it has a slightly smaller time-space product, which is considered to be the relevant figure of merit for efficiency in some physical models. [S1050-2947(96)06909-0]

PACS number(s): 03.65.Bz, 07.05.Bx, 89.80.+h, 02.70.Rw

I. INTRODUCTION

There is considerable interest in the controlled generation, manipulation, and transportation of individual quantum states; applications of such resources are envisioned in new kinds of data transmission, cryptography, and computation. The quantum extension of conventional bits, called *qubits*, has been subject to considerable exploration lately. A single qubit is embodied in the state of a single two-state quantum system, such as the spin degree of freedom of an electron or other spin- $\frac{1}{2}$ particle, where the spin-up state of the particle is denoted by $|0\rangle$ and the spin-down state is denoted by $|1\rangle$. The basic laws of quantum physics dictate that a description of the entire possible state-space of the qubit is given by the wave function

$$|\Psi\rangle = \alpha|0\rangle + \beta|1\rangle, \quad (1)$$

where α and β are any two complex numbers such that $|\alpha|^2 + |\beta|^2 = 1$. This is called a "qubit" since it can assume one of two binary values, but of course it has fundamentally different properties because of the possibility of it being in a superposition of these two values. The properties with which quantum mechanics endows the qubit make possible a kind of cryptography which is fundamentally secure against eavesdropping attacks [1], and computations which apparently violate the complexity-class categorizations for ordinary Boolean computers [2].

One of the ideas of this sort that has been understood recently is the possibility of data compression for qubits [3–5]. In classical information theory, if n bits, x_1, \dots, x_n , are each sampled independently according to some probability distribution $p = (p_0, p_1)$ (on the set $\{0, 1\}$) then the string $x_1 \dots x_n$ may be compressed to a $nH_S(p)$ -bit string [where $H_S(p) = -\sum_{i=0}^1 p_i \log_2 p_i$, the Shannon entropy [6]]—and no further—in the following asymptotic sense. For any $\epsilon, \delta > 0$, for sufficiently large n , for any $\lambda(n) \geq n[H_S(p) + \delta]$, $\lambda(n) \in \{1, \dots, n\}$, there exists a

compression scheme that compresses $x_1 \dots x_n$ to $y_1 \dots y_{\lambda(n)}$, and such that $x_1 \dots x_n$ can be successfully recovered from $y_1 \dots y_{\lambda(n)}$ with probability greater than $1 - \epsilon$. Moreover, the above compression is the maximum possible in the sense that, for any $\epsilon, \delta > 0$, for sufficiently large n , for any $\lambda(n) \leq n[H_S(p) - \delta]$, for any compression scheme that maps $x_1 \dots x_n$ to $y_1 \dots y_{\lambda(n)}$, the probability that $x_1 \dots x_n$ can be successfully recovered from $y_1 \dots y_{\lambda(n)}$ is less than ϵ .

The quantum physical analogue of the above scenario involves the compression of a string of qubits, instead of bits. Note that there are a continuum of possible states for each qubit, rather than two possible values. We shall consider the "discrete" case, where a probability distribution is concentrated on some finite set of qubit states $\mathcal{S} = \{|\Psi_1\rangle, \dots, |\Psi_m\rangle\}$. Let the respective probabilities be $p = (p_1, \dots, p_m)$. In the language of quantum physics, (\mathcal{S}, p) defines an *ensemble* of states. Let $|\alpha_1\rangle \dots |\alpha_n\rangle$ be a string of n qubits, each sampled independently from (\mathcal{S}, p) . Define a *compressor* A as a unitary transformation that maps n -qubit strings to n -qubit strings. Again let $\lambda(n) \in \{1, \dots, n\}$. It is to be understood that, on input $|\alpha_1\rangle \dots |\alpha_n\rangle$, the first $\lambda(n)$ qubits that are output by the compressor $|\beta_1 \dots \beta_{\lambda(n)}\rangle$ are taken as the *compressed version* of its input, and the remaining $n - \lambda(n)$ qubits are discarded. A *decompressor* B is a unitary transformation that maps n -qubit strings to n -qubit strings. It is to be understood that the first $\lambda(n)$ qubits input to the decompressor are $|\beta_1 \dots \beta_{\lambda(n)}\rangle$, the compressed version of some sequence of n qubits, and the remaining $n - \lambda(n)$ qubits are all $|0\rangle$. An n -to- $\lambda(n)$ *quantum compression scheme* is a compressor-decompressor pair (A, B) . As in the classical case, the goal is to achieve as high a compression rate [i.e., as small a $\lambda(n)$] as possible, while permitting the original message to be recovered from its compressed version, with high probability.

Assume that the compressor knows (i.e., can be a function of) the underlying ensemble (\mathcal{S}, p) , but has no explicit

knowledge about the specific random selections made [interestingly, compressors exist that know even less than (\mathcal{S}, p) ; more about this later]. In the classical case, the compressor obtains complete information about the bits to be compressed, but complete information cannot generally be obtained from a qubit. If the possible qubit states in \mathcal{S} are not mutually orthogonal then any observation of such a qubit will only yield partial information about its state, and can irretrievably change this state. Due to this, one might expect to be able to achieve less in the quantum scenario than with classical compression schemes—in fact, the opposite is true.

Let us measure the quality of an n -to- $\lambda(n)$ compression scheme (A, B) with respect to a source distribution p in terms of its *fidelity*, defined as follows. Consider the following experiment. Let the sequence $|\alpha_1\rangle \dots |\alpha_n\rangle$ be sampled independently from (\mathcal{S}, p) . Transform $|\alpha_1\rangle \dots |\alpha_n\rangle$ according to the compressor A and let $|\beta_1 \dots \beta_{\lambda(n)}\rangle$ be the compressed version. Next, transform $|\beta_1 \dots \beta_{\lambda(n)}\rangle |0 \dots 0\rangle$ according to the decompressor B and let $|\alpha'_1 \dots \alpha'_n\rangle$ be the output. Finally, measure $|\alpha'_1 \dots \alpha'_n\rangle$ with respect to a basis containing $|\alpha_1 \dots \alpha_n\rangle$. The *fidelity* is the probability $|\langle \alpha'_1 \dots \alpha'_n | \alpha_1 \dots \alpha_n \rangle|^2$ that this measurement results in $|\alpha_1 \dots \alpha_n\rangle$.

Note that the fidelity is with respect to two sources of randomness: (a) the random choices in the original generation of $|\alpha_1 \dots \alpha_n\rangle$; and (b) the randomness that results from performing a measurement of the state $|\alpha'_1 \dots \alpha'_n\rangle$. Roughly speaking, the fidelity can be high if for “most” choices in (a), $|\alpha'_1 \dots \alpha'_n\rangle$ is “close to” $|\alpha_1 \dots \alpha_n\rangle$.

The ensemble (\mathcal{S}, p) represents a *mixed state*, which has *density matrix* ρ , defined as

$$\rho = \sum_{i=1}^m p_i |\Psi_i\rangle \langle \Psi_i|.$$

The *von Neumann entropy* corresponding to (\mathcal{S}, p) is defined in terms of the density matrix ρ as $H_{VN}(\rho) = -\text{Tr}(\rho \log_2 \rho)$. In general, $H_{VN}(\rho) \leq H_S(p)$, with equality occurring if and only if the states in \mathcal{S} are mutually orthogonal.

Roughly speaking, Schumacher's theorem [3] states that $nH_{VN}(\rho)$ is asymptotically the maximum compression at-

tainable for n qubits resulting from a source with density matrix ρ . More precisely, let (\mathcal{S}, p) be any ensemble of qubits, and ρ be the corresponding density matrix. Then, for all $\varepsilon, \delta > 0$, for sufficiently large n and $\lambda(n) \geq n[H_{VN}(\rho) + \delta]$, there exists an n -to- $\lambda(n)$ quantum compression scheme for (\mathcal{S}, p) with fidelity greater than $1 - \varepsilon$. Moreover, for all $\varepsilon, \delta > 0$, for sufficiently large n , if $\lambda(n) \leq n[H_{VN}(\rho) - \delta]$ then every n -to- $\lambda(n)$ quantum compression scheme has fidelity less than ε .

It should be noted that the above bounds are robust in the sense that they do not change when a number of technical variations are made in the scenario. For example, the n -to- $\lambda(n)$ compression schemes that attain fidelity greater than $1 - \varepsilon$ can be restricted to being highly “oblivious” in that they depend only on knowing a basis for which the density matrix is diagonal, with nonincreasing values along the diagonal. Also [4], even if the compressor is supplied with complete information about the state of the source string $|\alpha_1 \dots \alpha_n\rangle$ that it receives, ε still bounds the fidelity attainable if $\lambda(n) \leq n[H_{VN}(\rho) - \delta]$.

The proof of Schumacher's theorem is based on the existence of a “typical subspace” Λ of the Hilbert space of n qubits, which has the property that, with high probability, a sample of $|\alpha_1 \dots \alpha_n\rangle$ has almost unit projection onto Λ . It has been shown [3,4] that the dimension of Λ is $2^{nH_{VN}(\rho)}$; thus, the operation that the compressor should perform involves “transposing” the subspace Λ into the Hilbert space of a smaller block of $nH_{VN}(\rho)$ qubits.

Bennett [7] gives a more explicit procedure for accomplishing this “transposition,” which we illustrate with an example. Suppose that $\mathcal{S} = \{|\Psi_1\rangle, |\Psi_2\rangle\}$, where $|\Psi_1\rangle = |0\rangle$ and $|\Psi_2\rangle = (1/\sqrt{2})|0\rangle + (1/\sqrt{2})|1\rangle$, and $p = (p_1, p_2)$, where $p_1 = p_2 = \frac{1}{2}$. The density matrix corresponding to (\mathcal{S}, p) is $\rho = \frac{1}{2}|0\rangle\langle 0| + \frac{1}{2}[(1/\sqrt{2})|0\rangle + (1/\sqrt{2})|1\rangle][(1/\sqrt{2})\langle 0| + (1/\sqrt{2})\langle 1|]$, or, in 2×2 matrix form,

$$\rho = \begin{pmatrix} \frac{3}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} \end{pmatrix} \quad \text{in the basis} \quad \begin{matrix} |0\rangle \\ |1\rangle \end{matrix}. \quad (2)$$

It is always possible to go to a basis in which the density matrix is diagonal:

$$\rho' = \begin{pmatrix} \lambda_{\max} & 0 \\ 0 & \lambda_{\min} \end{pmatrix} = \begin{pmatrix} \frac{3}{4} + \frac{1}{4} \tan \frac{\pi}{8} & 0 \\ 0 & \frac{1}{4} - \frac{1}{4} \tan \frac{\pi}{8} \end{pmatrix} \quad \text{in the basis} \quad \begin{matrix} |0'\rangle = \cos \frac{\pi}{8} |0\rangle + \sin \frac{\pi}{8} |1\rangle \\ |1'\rangle = -\sin \frac{\pi}{8} |0\rangle + \cos \frac{\pi}{8} |1\rangle. \end{matrix} \quad (3)$$

Both of the states $|\Psi_1\rangle$ and $|\Psi_2\rangle$ have large overlap on the basis state $|0'\rangle$ [$|\langle \Psi_i | 0'\rangle| = \cos(\pi/8)$], and small overlap on the orthogonal basis state $|1'\rangle$ [$|\langle \Psi_i | 1'\rangle| = \sin(\pi/8)$]. This observation leads to a way of compressing strings of signal states. Consider all n -qubit strings possible from the states in \mathcal{S} . These strings can all be expressed with respect to the basis

consisting of $|x_{n-1} \dots x_0\rangle = |x_{n-1}\rangle \dots |x_0\rangle$, where $x_{n-1}, \dots, x_0 \in \{|0'\rangle, |1'\rangle\}$. Each such $|x_{n-1} \dots x_0\rangle$ can be interpreted as an n -bit binary number, and, thus, can be denoted as $|x\rangle$, for $x \in \{0, \dots, 2^n - 1\}$. Now, the overlap of $|x\rangle$ with the states in \mathcal{S}^n is $|\langle x | \mathcal{S}^n \rangle| = \cos^m(\pi/8) \sin^{n-m}(\pi/8)$, where m is the number of 0's in the binary representation of

x . Because this overlap diminishes exponentially with $n - m$, basis states with large numbers of 1's are relatively unimportant for describing any string $|\alpha_1, \dots, \alpha_n\rangle$; the Hilbert space can thus be truncated to the typical subspace Λ consisting of all states $|x\rangle$ in which the binary number x contains a proportion of 1's less than $H_{VN}(\rho) < 0.601$.

Thus the "transposition" which the coder must do consists of mapping this Λ subspace for n qubits into the states spanned by less than $0.601n$ of those qubits.

We must accomplish this by a unitary transformation applied to the original states of the n qubits. In the basis $|0\rangle, \dots, |2^n - 1\rangle$, this transformation must map qubit strings with the smallest number of 1's in succession into qubit binary strings with the smallest numerical value. This is a classical combinatorial calculation, "classical" in the sense that definite binary-number states are mapped to other definite binary-number states; however, it is essential that the computation be performed quantum mechanically, since the computation must preserve the superpositions of these basis states. This means that the combinatorial computation must be performed using reversible, quantum-coherent elementary operations.

The principal object of this paper is to derive the quantum computation which is needed to do this Schumacher coding. In Sec. II we derive the analytical formula for the sorting calculation required for the coding. Section III constructs the quantum program for performing this calculation: Sec. III A illustrates a first attempt at this coding exercise; Sec. III B discusses the way in which the calculation is to be properly made reversible; and Sec. III C, which contains the essential result of the paper, gives the final quantum program for Schumacher coding. Section IV gives, in the same programming notation developed in the earlier sections, the bit-level routines needed for performing the steps in the high-level program. Appendix A discusses how these bit-level routines may be made highly space-efficient, with only a modest increase in running time (these latter routines result in a smaller time-space product, which may be desirable, due to the manner in which decoherence occurs [8,9]). Appendix B provides other ways of economizing in the bit-level implementation of these codes, by using some of the phase freedom coming from the quantum-mechanical nature of the computation.

II. COMBINATORIAL EXPRESSION FOR SCHUMACHER CODING

As Bennett [7] has described, a specific realization of the unitary transformation performing the Schumacher coding function on a set of identical qubits consists of a sorting computation in which the states $|0\rangle, \dots, |2^n - 1\rangle$ are given a lexicographical ordering according to how many 1's are in their binary expansion. So, $|0\rangle$ is mapped to itself, all the states containing exactly one 1 and $n - 1$ 0's are mapped to the states between $|1\rangle$ and $|n\rangle$, all the states with exactly two 1's and $n - 2$ 0's are mapped to the states between $|n + 1\rangle$ and $|n + n(n - 1)/2\rangle$, and, in general, all the states with exactly m 1's and $n - m$ 0's are mapped to the states between

$$\left| \sum_{i=0}^{m-1} \binom{n}{i} \right\rangle \tag{4}$$

and

$$\left| \sum_{i=0}^m \binom{n}{i} - 1 \right\rangle \tag{5}$$

inclusive. The Schumacher function does not require any particular ordering of the states within each of these blocks, except that the mapping must be 1-to-1 (i.e., a bijection); but, it turns out to be convenient to preserve lexicographical ordering within each block. Defining the index number within each block as $I[x, n, m]$, the total Schumacher function for string x (with n bits and m 1's) is

$$y = \sum_{i=0}^{m-1} \binom{n}{i} + I[x, n, m]. \tag{6}$$

The index number I obeys a recursive relationship which we now derive. Considering the possible binary-number strings representing the input state x , any string whose first 1 occurs in the $(p + 1)$ st place (i.e., whose first p bits are 0) must have a higher index number than all strings in which the first $p + 1$ places are 0. There are exactly $\binom{n-p+1}{m}$ such strings. This means that for the particular input string

$$x = \overbrace{00 \dots 00}^{p \text{ 0's}} 1 \overbrace{000 \dots 00}^{n-m-p \text{ 0's}} \overbrace{111 \dots 11}^{m-1 \text{ 1's}}, \tag{7}$$

the index number $I[x, n, m] = \binom{n-p-1}{m}$. This result permits the index number of the more complex string

$$x = \overbrace{00 \dots 00}^{p \text{ 0's}} 1 \underbrace{x'}_{n-p-1 \text{ bits}} \tag{8}$$

to be expressed recursively:

$$I[x, n, m] = \binom{n-p-1}{m} + I[x', n-p-1, m-1]. \tag{9}$$

It is probably easiest to understand Eq. (9) by writing out an example:

$$\begin{aligned}
I[00010011011,10,5] &= \binom{10-2}{5} + I[0011011,7,4] \\
&\Downarrow \\
&= \binom{7-2}{4} + I[1011,4,3] \\
&\Downarrow \\
&= \binom{4-0}{3} + I[011,3,2] \\
&\Downarrow \\
&= 0.
\end{aligned} \tag{10}$$

As this illustrates, the recursion of Eq. (9) may be iterated to produce an expression for I for a general input string x :

$$I[x, n, m] = \sum_{i=1}^{n-1} x_{n-i} \binom{n-i}{\sum_{k=i}^n x_{n-k}}. \tag{11}$$

Here the notation x_p denotes the value of the p th bit of the string x . Combining Eq. (11) with Eq. (6) yields the final expression for the Schumacher coding function:

$$y = \sum_{i=0}^{\sum_{k=0}^{n-1} x_k - 1} \binom{n}{i} + \sum_{j=1}^{n-1} x_j \binom{j}{\sum_{k=0}^j x_k}. \tag{12}$$

In this equation, binary coefficients outside their natural range [e.g., $\binom{n}{n+1}$] are understood to be zero.

III. HIGH-LEVEL QUANTUM PROGRAM FOR SCHUMACHER CODING

A. First attempts

It is now our object to translate Eq. (12) into a sequence of elementary quantum-mechanical manipulations. We proceed to do this by writing out the calculation in a high-level ‘‘pseudocode’’ [10] which, when ‘‘compiled,’’ would permit the operation to be performed by a sequence of elementary spectroscopic manipulations such as two-bit XOR’s (or controlled-NOT’s), along with one-bit rotations [11]. Rather than building up the rules of this pseudocode axiomatically, we will proceed in an intuitive fashion. The principal constraint which the coded calculations must obey is that they must be done *reversibly*. Instead of going into a discourse about this, let us present the first try at coding Eq. (12) (not a perfectly successful one, in fact):

Program FIRST-TRY

quantum registers:

X : n -bit register

Y : n -bit arithmetic register (initialized to 0)

S : $[\log_2 n]$ -bit arithmetic register (initialized to 0)

if $X_0=1$ **then** $S \leftarrow S+1$

for $j=1$ **to** $n-1$ **do**

if $X_j=1$ **then** $S \leftarrow S+1$

for $m=0$ **to** $j+1$ **do**

if $X_j=1$ **and** $S=m$ **then** $Y \leftarrow Y + \binom{j}{m}$

for $i=0$ **to** $n-1$ **do**

if $i+1 \leq S$ **then** $Y \leftarrow Y + \binom{n}{i}$.

FIRST-TRY is not incorrect, but it is incomplete, in ways which we will repair by stages below. Here are some rules of this programming: All the quantum-mechanical registers are in capital letters. In FIRST-TRY, these are X (which is initialized with an input state x , or a quantum superposition of such input states), Y (which is initialized to 0, and whose final value is the output state y or their quantum superpositions), and S (a small work register, also initialized to 0). The notation X_i indicates the i th bit of X . Note that Y and S are given the data type ‘‘arithmetic,’’ indicating that ordinary integer addition and subtraction are allowed with them. Only bitwise manipulations are performed on X . (In the FINAL-SCHUMACHER program, both bitwise and arithmetic manipulations will be performed on the same registers.)

All other lower-case variables in the program always have definite values and can (and should) be implemented using classical bits. Only the quantum registers need to be explicitly treated reversibly. So, the binomial coefficients $\binom{j}{m}$ can be precomputed or evaluated by any means, reversible or not, in the implementation of the quantum computation.

In a reversible program statement, the input can always be deduced from the output. So, for example, the statement

if $X_0=1$ **then** $S \leftarrow S+1$

is reversible, because the input could be deduced by the ‘‘time-reverse’’ of this statement,

if $X_0=1$ **then** $S \leftarrow S-1$

An irreversible program statement would be

if $X_0=1$ **then** $S \leftarrow 1$

since the prior value of S cannot be deduced. As it happens, this statement would function correctly in FIRST-TRY because S actually is equal to 0 at this first executable statement of the program. However, we will enforce a rule that the only

irreversible statements permitted that involve quantum variables will be the “initialized” designations present in the declaration statements. In later programs we will introduce a “finalized” designation, which will merely serve as a reminder that certain variables will always end the program with a particular value if the program runs correctly. This designation will be an important one in constructing reversible code. It is also a reminder that physically, the finalization can serve as a useful check that no error has occurred [12]; a quantum measurement of this register at the end of the running of the program should always find the register in the finalized value.

One further comment about the program statement

if $X_0=1$ **then** $S \leftarrow S+1$.

If S were a one-bit variable, this statement would just be a quantum XOR or controlled NOT, in which the value of S is inverted conditional on the value of X_0 . In FIRST-TRY, S is a multibit register, in fact it must have about $\log_2 n$ bits. Implementation of these multibit functions in terms of primitive operations involving no more than three bits is straightforward, and is presented in Sec. IV and in Refs. [13,14]. Using quantum gates, all the three-bit primitives may be reduced to sequences of two-bit operations [11].

A few more points about FIRST-TRY are in order. Given the constraints of reversibility, FIRST-TRY is a relatively straightforward transcription of Eq. (12). The first **for** loop (indexed by j) implements the second term of Eq. (12); this is efficient because the partial sum in the binomial coefficient can be accumulated in S one term at a time, and then the completed sum can be used as the upper limit of the first term of (12), which is implemented in the second **for** loop. This inner m loop could be replaced by the single statement

if $X_j=1$ **then** $Y \leftarrow Y + \binom{j}{S}$,

but this would require a reversible calculation of the binomial function; we have chosen to make this binomial-coefficient calculation classical by writing out the loop as shown. One might also be tempted to modify the inner loop as follows:

if $X_j=1$ **then**
 for $m=0$ **to** $j+1$ **do**
 if $S=m$ **then** $Y \leftarrow Y + \binom{j}{m}$.

While moving the **if** statement from the loop is superficially more efficient, it turns out that, when these statements are reexpressed in terms of primitive operations, the **if** must be carried down to the lowest level in any case; so, we prefer a syntax in which such conditionals are explicitly shown at the lowest level.

B. Reversibility considerations

Now, what is the overall effect of FIRST-TRY, and why is it inadequate for performing the Schumacher function? Let f :

$\{0,1\}^n \rightarrow \{0,1\}^n$ denote the Schumacher function for n -bit binary strings. If the total input state is expressed as the ket

$$|X, Y, S\rangle = |x, 0, 0\rangle, \quad (13)$$

then the complete final state is

$$|X, Y, S\rangle = |x, f(x), s\rangle \quad (14)$$

(where s is the number of 1's in x). But, the correct Schumacher function must have a final state of the form

$$|X, Y, S\rangle = |0, f(x), 0\rangle. \quad (15)$$

That is, the input x should be erased and the work register S should be reset to its initial value of 0. This is possible to accomplish reversibly because the Schumacher function is bijective, so that no record of the initial state, or of the state of the work bits, needs to be retained at the end; they are completely deducible from the output. In fact, the correct operation of the Schumacher function *requires* that the output be of the form (15); if it is of the form of (14), then the final state is “entangled” with the initial state, which means that output states cannot be placed in the desired superpositions of states. Thus, the net result of the Schumacher function should be confined to the input data register only; this condition is obtainable from Eq. (15) if the final output state is swapped so that the state vector becomes

$$|X, Y, S\rangle = |f(x), 0, 0\rangle. \quad (16)$$

Thus the Schumacher function is applied, “in-place,” to the first n qubits, while the remaining $n + \log_2 n$ bits return to their original states, and may all be viewed simply as work space for the computation. We will see later that the “output” register Y can actually be removed entirely by using some clever programming. Some other workspace, not displayed explicitly in (16), appears to be necessary to do the bit-level manipulations in the Schumacher function (see Sec. IV); Appendix A shows that the size of this extra workspace does not have to exceed about $2\sqrt{n}$ bits.

These considerations have arisen previously in the context of reversible programming [15], but the rationale for constructing a function in the “fully reversible” manner as specified by the output state (16) is somewhat different than in the classical context. In traditional reversible programming the object is to avoid the small energy cost involved in irreversible erasure of any of the working bits in the computer. If such an erasure is performed, the *result* of the computation will still be correct, even though the desired goal of expending no energy is not achieved. But, in this particular quantum computation, irreversible erasure of the state of register X in Eq. (14) actually causes register Y to be in the *wrong* quantum state, in so far that, if the initial X was in a superposition of computational states, the final state of Y will be a mixed quantum state, rather than the intended, pure superposition state. Thus, the consequences of irreversibility are more serious than in conventional reversible computation.

A method for designing a calculation to arrive at the desired final states (15) or (16), as already worked out in the earlier literature [15], requires two steps: (1) zero out S and any other work space used by the program, and (2) explic-

itly implement the *inverse* of the Schumacher function Eq. (12). This can be accomplished by a program that, on input state

$$|X, Y, S\rangle = |x, y, 0\rangle, \quad (17)$$

produces the final state

$$|X, Y, S\rangle = |x \oplus f^{-1}(y), y, 0\rangle. \quad (18)$$

Note that applying such a transformation to the state

$$|X, Y, S\rangle = |x, f(x), 0\rangle \quad (19)$$

yields the required state

$$|X, Y, S\rangle = |0, f(x), 0\rangle. \quad (20)$$

Equation (18) is not implemented simply by running FIRST-TRY in reverse; indeed, the inverse function can have very different and much greater complexity than the function itself [16]. Fortunately, in this case, as we will see in a moment, the inverse Schumacher function is also relatively easy to implement.

Step (1) above, zeroing out S , is readily performed by adding code to the end of FIRST-TRY to simply subtract away the bits which have been added to S :

```

for  $j=0$  to  $n-1$  do
    if  $X_j=1$  then  $S \leftarrow S-1$ .

```

This code, added to the end of FIRST-TRY, produces the output state (19).

Step (2) above, implementing the inverse function Eq. (18), requires a new algorithm. We have not found any way to write the inverse Schumacher coding function as a formula as in Eq. (12). Nevertheless, a straightforward algorithm can be deduced from the following two inequalities. The first is obtained by combining the information from Eqs. (4) and (5):

$$\sum_{i=0}^{m-1} \binom{n}{i} \leq y < \sum_{i=0}^m \binom{n}{i}, \quad (21)$$

where

$$m = \sum_{k=0}^{n-1} x_k \quad (22)$$

is the number of 1's in the binary string x . We will be able to write simple pseudocode to compute m (a.k.a. S). This result can then be used to compute $I[x, n, m]$ using Eq. (6). $I[x, n, m]$ satisfies an inequality which is a simple consequence of Eq. (9) and the discussion preceding it:

$$\binom{n-p-1}{m} \leq I[x, n, m] < \binom{n-p}{m}. \quad (23)$$

By finding the p which satisfies this equation, we determine that the leading p bits of x are zeros, and the next bit is a 1 (i.e., $x_j=0$, $n-1-p < j \leq n-1$, $x_{n-p-1}=1$). The index of the remaining substring can be determined from Eq. (9), and thus all the bits of x may be calculated recursively.

C. Deriving the final program

Now we will transform our procedure into reversible code. As the previous section makes clear, a necessary step for doing this will be to code the inverse of the Schumacher function. In the spirit of FIRST-TRY, we will not worry at first about the final state of the work registers as prescribed in Eq. (18); we will initially just try to code correctly the inverse function itself. We will find that reversibility will, in this case, fall out naturally from a simple modification of our first-cut program.

Program TRY-INVERSE

quantum registers:

X : n -bit register (initialized to 0)

Y : n -bit signed arithmetic register (finalized to 0)

S : $[\log_2 n]$ -bit register (initialized and finalized to 0)

for $m=0$ **to** n **do**

$$Y \leftarrow Y - \binom{n}{m}$$

if $Y \geq 0$ **then** $S \leftarrow S+1$

for $m=0$ **to** n **do**

$$\mathbf{if} \ S \geq m \ \mathbf{then} \ Y \leftarrow Y + \binom{n}{m}$$

for $p=0$ **to** $n-1$ **do**

for $i=0$ **to** $n-p$ **do**

if $S=i$ **and** $Y \geq \binom{n-p-1}{i}$

then $X_{n-p-1} \leftarrow X_{n-p-1} \oplus 1$

if $S=i$ **and** $X_{n-p-1}=1$

then $Y \leftarrow Y - \binom{n-p-1}{i}$

if $X_{n-p-1}=1$ **then** $S \leftarrow S-1$.

In this code, the m loop does the job of finding the m for which Eq. (21) is satisfied, and putting the result in the quantum register S . As a byproduct of this work, it subtracts away the first term of Eq. (12) from y , leaving in Y the value of the index $I[x, n, m]$. Actually, the m loop continues to subtract binomial coefficients from Y after it is supposed to; this is why Y is indicated to be a "signed" register, which can be handled by doing ordinary arithmetic in a register with one extra bit (see [13]). This approach has the benefit that testing that Y is non-negative only requires the examination of one bit—see the first part of Sec. IV. We might be tempted to avoid negative numbers by terminating the loop at the right moment, viz.,

for $m=0$ **to** n **do**

if $Y < \binom{n}{m}$ **then exit for-loop**

⋮

But such an **exit for-loop** statement is not reversible. There appears to be no alternative to letting the first loop go to its maximum possible upper limit, which is n , and then repairing the damage done by adding back the correct binomial coefficients in the second loop. Finally, at the end of the second loop, Y has the desired value of $I[x, n, m]$, and S has the value of m .

Then the third (p) loop of TRY-INVERSE does the iterative decomposition of the index $I[x, n, m]$. For every possible value of the leading number of zeros p [recall Eq. (7)], TRY-INVERSE checks to see if the inequality Eq. (23) is satisfied; if it is, then the program negates one bit of the X register. Then the second **if** statement decrements Y by the combinatorial coefficient in Eq. (9), so that it always contains the index of the next substring. The process continues until the index is reduced to zero. Also, S is decremented so that it always contains the current value of the number of 1's in the substring of Eq. (8). Note that, as in FIRST-TRY, an inner loop (indexed by i) is introduced to avoid the need for reversible calculation of binomial coefficients like $\binom{n-p-1}{s}$.

We now evaluate what state TRY-INVERSE has left the registers Y and S in. In fact, a very desirable thing has “accidentally” occurred. We find that, on input state

$$|X, Y, S\rangle = |0, y, 0\rangle, \quad (24)$$

TRY-INVERSE produces the final state

$$|X, Y, S\rangle = |f^{-1}(y), 0, 0\rangle. \quad (25)$$

Thus, with a final transposing of the X and Y registers, we obtain a program that implements the *inverse* of Eq. (16), so the calculation has been successfully done in-place, with the registers Y and S remaining in their initial state, having served only as “catalysts” for the calculation.

In fact, we can do even better; by a small modification of TRY-INVERSE, the Y register can be eliminated entirely. This can be done by noting that, during the course of an execution of TRY-INVERSE, the decrementing of Y sets each of its high-order bits to zero in succession, and, at the same time, the values of X are built up starting with the high-order bits and working down. Thus, the high-order bits of Y can be reused to hold the results of the final calculation. It can be shown that these high-order bits are always cleared out soon enough that they can be used for the final answer; this is done by showing that in TRY INVERSE, the same bits of X and Y are never simultaneously 1. Thus, with one small modification, TRY INVERSE can be turned into our final program for the inverse of the Schumacher coding function:

Program FINAL-SCHUMACHER-INVERSE

quantum registers:

X : n -bit signed arithmetic register

S : $\lceil \log_2 n \rceil$ -bit arithmetic register (initialized and finalized to 0)

for $m=0$ **to** n **do**

$$X \leftarrow X - \binom{n}{m}$$

if $X \geq 0$ **then** $S \leftarrow S + 1$

for $m=0$ **to** n **do**

$$\mathbf{if} \ S \geq m \ \mathbf{then} \ X \leftarrow X + \binom{n}{m}$$

for $p=0$ **to** $n-1$ **do**

for $i=0$ **to** $n-p$ **do**

$$\mathbf{if} \ S=i \ \mathbf{and} \ \mathbf{TRUNC}_{n-p-1}(X) \geq \binom{n-p-1}{i} \ \mathbf{then} \ X_{n-p-1} \leftarrow X_{n-p-1} \oplus 1$$

$$\mathbf{if} \ S=i \ \mathbf{and} \ X_{n-p-1} = 1 \ \mathbf{then} \ X \leftarrow X - \binom{n-p-1}{i}$$

if $X_{n-p-1} = 1$ **then** $S \leftarrow S - 1$

The only substantial item which has been added here is the function \mathbf{TRUNC}_j . Invocation of $\mathbf{TRUNC}_j(X)$ simply says that only the j least significant bits of the quantum register X (i.e., bit 0 to bit $j-1$) should be taken account of in the “ \geq ” comparison. This is necessary because the high-order bits are being used to store the final answer. In the final pass through the p loop, the occurrence of the zero index in $\mathbf{TRUNC}_0(X)$ indicates that the comparison should not be performed at all.

For completeness, we now record the final code for the Schumacher coding function itself. Since FINAL-SCHUMACHER-INVERSE is done “in-place,” the direct function is literally just the time-reverse:

Program FINAL-SCHUMACHER

quantum registers:

X : n -bit signed arithmetic register

S : $\lceil \log_2 n \rceil$ -bit arithmetic register (initialized and finalized by 0)

for $p=n-1$ **down to** 0 **do**

if $X_{n-p-1}=1$ **then** $S \leftarrow S+1$

for $i=n-p$ **down to** 0 **do**

if $S=i$ **and** $X_{n-p-1}=1$ **then** $X \leftarrow X + \binom{n-p-1}{i}$

if $S=i$ **and** $\text{TRUNC}_{n-p-1}(X) \geq \binom{n-p-1}{i}$ **then** $X_{n-p-1} \leftarrow X_{n-p-1} \oplus 1$

for $m=n$ **down to** 0 **do**

if $S \geq m$ **then** $X \leftarrow X - \binom{n}{m}$

for $m=n$ **down to** 0 **do**

if $X \geq 0$ **then** $S \leftarrow S-1$

$X \leftarrow X + \binom{n}{m}$.

IV. BIT-LEVEL QUANTUM PROGRAM FOR SCHUMACHER CODING

In this section, we explain how the statements in programs FINAL-SCHUMACHER and FINAL-SCHUMACHER-INVERSE can be implemented by a gate array with fundamental bit-level operations. These fundamental operations are essentially Toffoli gates [17]. The Toffoli gate that negates bit B iff bits C and D are both 1 (and does not change the values of C and D) is denoted as

$$B \leftarrow B \oplus (C \wedge D).$$

In [11] it is shown that such an operation can be simulated in terms of eight one-bit operations and six XOR operations (which are of the form $B \leftarrow B \oplus C$). For convenience, we expand our repertoire of allowable basic operations to include

$$B \leftarrow B \oplus 1,$$

$$B \leftarrow B \oplus C,$$

$$B \leftarrow B \oplus \bar{C},$$

$$B \leftarrow B \oplus (\bar{C} \wedge D),$$

$$B \leftarrow B \oplus (\bar{C} \wedge \bar{D}),$$

$$B \leftarrow B \oplus (C \vee D).$$

As with Toffoli gates, each of these gates can be simulated by at most eight one-bit operations and six XOR operations.

In many cases a quantum phase freedom can be used to simulate these in fewer one- and two-bit gates (see Appendix B).

The first step to converting the programs into gate-arrays is to “unravel” the **for** loops. Since the ranges of these loops are all fixed prior to any computation, this is straightforward. Next, we note that (once the **for** loops have been unraveled) there are essentially five types of program statements:

1. $X \leftarrow X+k$
2. **if** B **then** $X \leftarrow X+k$
3. **if** $Y>l$ **then** $X \leftarrow X+k$
4. **if** $Y=l$ **and** B **then** $X \leftarrow X+k$
5. **if** $Y=l$ **and** $Z>k$ **then** $B \leftarrow B \oplus 1$

(where B is a bit, X, Y, Z are signed arithmetic registers, and k, l are signed integers).

Also, there are *a priori* upper bounds on the ranges of the arithmetic registers (and thus on the number of bits required to specify them). An arithmetic register whose range of values is known to be an integer within $[0, 2^n)$ can be naturally represented by n bits and arithmetic operations on it can be simulated by reversibly performing them modulo 2^n . Also, a *signed* arithmetic register whose range of values is known to lie within $[-2^n, +2^n)$ can be naturally represented in “two’s complement” form by $n+1$ bits, and it is well known that arithmetic operations on such a two’s complement integer can be simulated by interpreting it as an integer in the range $[0, 2^{n+1})$ and performing arithmetic modulo 2^{n+1} (see, for example, [18]).

A. Addition and conditional addition

In view of the above discussion, to simulate

if B then $X \leftarrow X + k$,

it suffices to perform

$$X \leftarrow (X + B \cdot k) \bmod 2^n$$

(in other words, to add k to X modulo 2^n iff $B = 1$). In the case where X is an n -bit signed register, it suffices to substitute $n + 1$ for n above. The following program performs this using n auxiliary bits C_0, C_1, \dots, C_{n-1} (which are assumed to have initial value 0, and are reset to 0 by the end of the computation).

A note about this program: the code introduces two modified assignment symbols “ \leftarrow ” and “ \leftarrow_R .” For the present purposes these can be thought of as identical to the ordinary “ \leftarrow ” assignment; however, they signal a freedom in how the quantum phase may be handled in these assignments, as discussed in Appendix B.

Program CONDITIONAL-ADD-k

quantum registers:

X : n -bit signed arithmetic register

B : bit register

C_0, C_1, \dots, C_{n-1} : bit registers (initialized and finalized to 0)

for $i = 1$ to $n - 1$ do

$$C_i \leftarrow C_i \oplus \text{MAJ}(k_{i-1}, X_{i-1}, C_{i-1})$$

for $i = n - 1$ down to 1 do

$$X_i \leftarrow X_i \oplus (k_i \wedge B)$$

$$X_i \leftarrow X_i \oplus (C_i \wedge B)$$

$$C_i \leftarrow_R C_i \oplus \text{MAJ}(k_{i-1}, X_{i-1}, C_{i-1})$$

$$X_0 \leftarrow X_0 \oplus k_0$$

where

$$\text{MAJ}(l, S, T) = \begin{cases} S \wedge T & \text{if } l = 0 \\ S \vee T & \text{if } l = 1. \end{cases}$$

The number of basic operations performed by the above program is bounded above by $4n + O(1)$. In particular, if the **for** loops of this program are unraveled then the program corresponds to a gate array consisting of $2n + 1$ bits and $4n + O(1)$ gates. (A more space-efficient $[n + O(\sqrt{n})]$ -bit program is described in Appendix A.)

The unconditional addition statement

$$X \leftarrow X + k$$

can be easily simulated by replacing $(k_i \wedge B)$ and $(C_i \wedge B)$ in the above program with k_i and C_i (respectively).

One final note about CONDITIONAL-ADD: it involves only the addition of a quantum register with an ordinary, classical number. It is possible to write a similar program which adds two quantum registers, as has been illustrated in [13]; how-

ever, this more complex routine is never needed for the implementation of the Schumacher function. Actually, it is generally possible to implement a full quantum adder as a sequence of calls to CONDITIONAL-ADD.

B. Equality and inequality testing

In order to simulate the remaining types of statements, it suffices to simulate *equality test* statements of the form

$$B \leftarrow B \oplus (X = k)$$

(which negate B iff $X = k$), and *inequality test* statement of the form

$$B \leftarrow B \oplus (X > k)$$

(which negate B iff $X > k$).

With implementations of the above tests, the statement

if $Y > l$ then $X \leftarrow X + k$

is then easily simulated by the sequence

$$B \leftarrow B \oplus (Y > l)$$

if B then $X \leftarrow X + k$

$$B \leftarrow_R B \oplus (Y > l)$$

where B is a bit register distinct from the bits of X and Y , and whose initial value is 0 (note that B must be reset to 0 after the addition is performed). Also, the compound conditional

if $Y = l$ and B then $X \leftarrow X + k$

is simulated by the sequence

$$C \leftarrow C \oplus (Y = l)$$

$$D \leftarrow D \oplus (C \wedge B)$$

if D then $X \leftarrow X + k$

$$D \leftarrow_R D \oplus (C \wedge B)$$

$$C \leftarrow_R C \oplus (Y = l)$$

where C and D are bit registers distinct from the bits of X , Y , and B , and whose initial (and final) values are 0. Again, the meaning and usefulness of the phase-modified assignments is discussed in Appendix B.

The following program simulates an equality test. It uses n auxiliary bit registers C_0, C_1, \dots, C_{n-1} . The auxiliary registers are initialized to 0, and have final value 0.

Program TEST-EQUALITY-TO-k

quantum registers:

X : n -bit signed arithmetic register

B : bit register

C_0, C_1, \dots, C_{n-1} : bit registers (initialized and finalized to 0)

$$C_{n-1} \leftarrow C_{n-1} \oplus (X_{n-1} = k_{n-1})$$

for $i=n-2$ **down to** 0 **do**

$$C_i \leftarrow C_i \oplus (C_{i+1} \wedge (X_i = k_i))$$

$$B \leftarrow B \oplus C_0$$

for $i=0$ **to** $n-2$ **do**

$$C_i \leftarrow {}_R C_i \oplus (C_{i+1} \wedge (X_i = k_i))$$

$$C_{n-1} \leftarrow {}_R C_{n-1} \oplus (X_{n-1} = k_{n-1})$$

where

$$(S=l) = \begin{cases} \bar{S} & \text{if } l=0 \\ S & \text{if } l=1. \end{cases}$$

The number of basic operations performed by the above program is bounded above by $2n + O(1)$. (The above program is very similar to the so-called \wedge_n -gate construction in [11]).

Finally, the following program simulates an inequality test. It uses n auxiliary bit registers C_0, C_1, \dots, C_{n-1} . The auxiliary registers are initialized to 0, and have final value 0.

Program TEST-GREATER-THAN-k
quantum registers:

X : n -bit signed arithmetic register

B : bit register

C_0, C_1, \dots, C_{n-1} : bit registers (initialized and finalized to 0)

$$C_{n-1} \leftarrow C_{n-1} \oplus (X_{n-1} = k_{n-1})$$

$$B \leftarrow B \oplus (X_{n-1} < k_{n-1})$$

for $i=n-2$ **down to** 0 **do**

$$C_i \leftarrow C_i \oplus (C_{i+1} \wedge (X_i = k_i))$$

$$B \leftarrow B \oplus C_{i+1} \wedge (X_i > k_i)$$

for $i=0$ **down to** $n-2$ **do**

$$C_i \leftarrow {}_R C_i \oplus (C_{i+1} \wedge (X_i = k_i))$$

$$C_{n-1} \leftarrow {}_R C_{n-1} \oplus (X_{n-1} = k_{n-1})$$

where $(S=l)$ is as in the previous subsection,

$$(S>l) = \begin{cases} S & \text{if } l=0 \\ 0 & \text{if } l=1, \end{cases}$$

and

$$(S<l) = \begin{cases} 0 & \text{if } l=0 \\ \bar{S} & \text{if } l=1. \end{cases}$$

The number of basic operations performed by the above program is bounded above by $3n + O(1)$. Once again, we employ phase-modified assignments \leftarrow , \leftarrow_R , and $\leftarrow<$, which are explained in Appendix B.

V. DISCUSSION AND CONCLUSIONS

We can finally put all the above results together to evaluate the total cost, in time and space, to perform Schumacher coding. It is easy to see that the two **if** statements inside the i loop of FINAL-SCHUMACHER are the most expensive part of

the procedure. The first **if** statement requires one call to CONDITIONAL-ADD. Although X is an n -bit register, the addition only affects the $n-p-1$ low-order bits of X . Thus, the addition can be performed on $\text{TRUNC}_{n-p-1}(X)$ rather than X , which amounts to a total running time of

$$\sum_{p=0}^{n-1} \sum_{i=0}^{n-p} 4(n-p-1) + O(1) = \frac{2}{3}n^3 + O(n^2). \quad (26)$$

The expensive part of the second **if** statement is its two calls to TEST-GREATER-THAN, performed on an $(n-p-1)$ -bit quantum register (because of the action of TRUNC). The time involved for this is

$$\sum_{p=0}^{n-1} \sum_{i=0}^{n-p} 2 \times 3(n-p-1) + O(1) = n^3 + O(n^2). \quad (27)$$

Thus, the total time required (i.e., number of bit-level primitive steps) is $\frac{5}{3}n^3 + O(n^2)$. The total number of qubits used is n , to hold the input-output string X ; plus $\lceil \log_2 n \rceil$, to hold S ; plus $n + O(1)$ to implement the conditional additions and inequality tests (the same work registers that store carries and so forth may be reused throughout the execution of the program). Thus, the total number of qubits is $2n + \lceil \log_2 n \rceil + O(1)$.

If the space-efficient routines CONDITIONAL-ADD' and TEST-GREATER-THAN' introduced in Appendix A are used instead, the execution time is increased to $\frac{8}{3}n^3 + O(n^{2.5})$, but the total number of qubits is reduced to $n + 2\sqrt{n} + O(\log_2 n)$. If the relevant figure of merit for the tractability of the quantum computation is the product of time and space, as it is in certain physical models [8,9], then the space-efficient procedures we have introduced would be preferred.

A final note about these operation counts: they are all in terms of the primitive operations listed at the beginning of Sec. IV, which includes both two- and three-bit primitives. It is known [11,19] that all three-bit operations can be simulated in quantum logic by a sequence of two-bit primitives. Most of the three-bit operations can be simulated using seven operations (three quantum XOR's and four one-bit gates); see Appendix B. So, in terms of these primitive operations the total time to do the Schumacher function would be roughly $7 \times \frac{8}{3}n^3 < 19n^3$. Computing the exact prefactor would require a considerable amount of detailed calculation, and would have to take into account that fact that many one-bit gates in the network could be merged together and executed in one step (see [11]). All of this work could easily be done if an actual physical implementation of Schumacher compression were ever undertaken.

To conclude, we believe that the pseudocode in which our results are presented is the most concise and economical form in which to present a quantum computation like Schumacher coding. The bit level primitives for addition and comparison which we have presented are similar to ones which have been presented elsewhere [13], but have a few features which may make them superior in the development of other quantum programs. The Schumacher coding can be done in $O(n^3)$ steps, with $O(\sqrt{n})$ auxiliary workspace. We cannot exclude the possibility that a lower polynomial-order

algorithm may be found, but we are not presently aware of what form this would take. The techniques in [15] enable further shrinkage of the auxiliary work space, but with a larger penalty in the running time. We think that further useful shrinkage of the auxiliary work space is unlikely; in the present scheme, only a vanishingly small fraction of quantum bits are used as workspace for large block size n .

ACKNOWLEDGMENTS

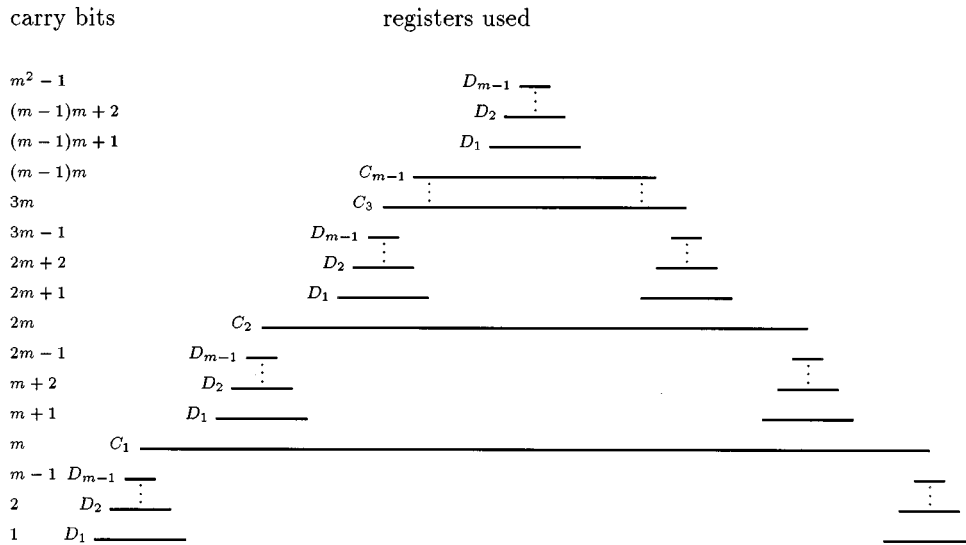
We are grateful to C. H. Bennett, R. Jozsa, B. Schumacher, and J. Smolin for useful discussions. Thanks to the authors of Ref. [13] for a preview of their work prior to publication. R. C. is supported in part by NSERC of Canada.

**APPENDIX A: IMPROVEMENTS
IN THE WORKSPACE EFFICIENCY
OF THE BIT-LEVEL IMPLEMENTATIONS**

The bit-level implementations proposed in Secs. IV A and IV B require n auxiliary bit registers. By applying techniques that were introduced in [15], we derive the following alternate programs that employ only $O(\sqrt{n})$ auxiliary bit registers while maintaining the same asymptotic operation complexity. (The space-reduction techniques in [15] can also be used

to reduce the auxiliary space further, but this causes an increase in the running time, as well as in the space-time product.)

Assume that $n = m^2$. The program **CONDITIONAL-ADD'- k** that follows employs $2m - 1$ auxiliary bit registers rather than the n auxiliary bit registers that **CONDITIONAL-ADD- k** employs. In **CONDITIONAL-ADD- k** , registers C_0, \dots, C_{n-1} are used to store information about carry propagation. In **CONDITIONAL-ADD'- k** , this is accomplished by registers C_1, \dots, C_{m-1} and D_1, \dots, D_{m-1} instead. The idea is to reset some of the registers to 0 at various checkpoints during the course of the computation. This is illustrated by the diagram below, where the horizontal direction represents time, and the placement of the lines indicate the time intervals during which the registers are active, containing the various carry bits. Registers D_1, \dots, D_{m-1}, C_1 are first set to the first m carry bits. Then D_1, \dots, D_{m-1} are reset to 0. Registers D_1, \dots, D_{m-1}, C_2 can then be used to store the $(m + 1)$ st to $(2m)$ th carry bits and then D_1, \dots, D_{m-1} are reset to 0 again—since C_1 stores the m th carry bit, this can be accomplished without recomputing the first m carry bits. The process is repeated with the remaining carry bits, and then applied in reverse to reset the carry bits to 0, as illustrated here:



The detailed program follows. D_0 is used for convenience to store the value of C_i at the beginning of each iteration of the for-loop with respect to i .

Program **CONDITIONAL-ADD'- k**

quantum registers:

X : n -bit arithmetic register

B : bit register

C_1, C_2, \dots, C_{m-1} : bit registers (initialized and finalized to 0)

D_0, D_1, \dots, D_{m-1} : bit registers (initialized and finalized to 0)

for $i = 0$ **to** $m - 2$ **do**

if $i > 0$ **then** $D_0 \leftarrow D_0 \oplus C_i$

for $j = 1$ **to** $m - 1$ **do**

$D_j \leftarrow \in D_j \oplus \text{MAJ}(k_{im+j-1}, X_{im+j-1}, D_{j-1})$

$C_{i+1} \leftarrow \in C_{i+1} \oplus \text{MAJ}(k_{im+m-1}, X_{im+m-1}, D_{m-1})$

for $j=m-1$ **down to** 1 **do**

$$D_j \leftarrow \leftarrow_R D_j \oplus \text{MAJ}(k_{im+j-1}, X_{im+j-1}, D_{j-1})$$

if $i>0$ **then** $D_0 \leftarrow D_0 \oplus C_i$

$$D_0 \leftarrow D_0 \oplus C_{m-1}$$

for $j=1$ **to** $m-1$ **do**

$$D_j \leftarrow \leftarrow_R D_j \oplus \text{MAJ}(k_{(m-1)m+j-1}, X_{(m-1)m+j-1}, D_{j-1})$$

for $j=m-1$ **down to** 1 **do**

$$X_{(m-1)m+j} \leftarrow X_{(m-1)m+j} \oplus (k_{(m-1)m+j} \wedge B)$$

$$X_{(m-1)m+j} \leftarrow X_{(m-1)m+j} \oplus (D_j \wedge B)$$

$$D_j \leftarrow \leftarrow_R D_j$$

$$\oplus \text{MAJ}(k_{(m-1)m+j-1}, X_{(m-1)m+j-1}, D_{j-1})$$

$$D_0 \leftarrow D_0 \oplus C_{m-1}$$

for $i=m-2$ **down to** 0 **do**

if $i>0$ **then** $D_0 \leftarrow D_0 \oplus C_i$

for $j=1$ **to** $m-1$ **do**

$$D_j \leftarrow \leftarrow_R D_j \oplus \text{MAJ}(k_{im+j-1}, X_{im+j-1}, D_{j-1})$$

$$X_{im+m} \leftarrow X_{im+m} \oplus (k_{im+m} \wedge B)$$

$$X_{im+m} \leftarrow X_{im+m} \oplus (C_{i+1} \wedge B)$$

$$C_{i+1} \leftarrow \leftarrow_R C_{i+1} \oplus \text{MAJ}(k_{im+m-1}, X_{im+m-1}, D_{m-1})$$

for $j=m-1$ **down to** 1 **do**

$$X_{im+j} \leftarrow X_{im+j} \oplus (k_{im+j} \wedge B)$$

$$X_{im+j} \leftarrow X_{im+j} \oplus (D_j \wedge B)$$

$$D_j \leftarrow \leftarrow_R D_j \oplus \text{MAJ}(k_{im+j-1}, X_{im+j-1}, D_{j-1})$$

if $i>0$ **then** $D_0 \leftarrow D_0 \oplus C_i$

$$X_0 \leftarrow X_0 \oplus k_0.$$

The above program uses $n + O(\sqrt{n})$ registers in total and runs in $6n + O(\sqrt{n})$ steps [compared to $2n + O(\log n)$ registers in total and $4n + O(1)$ steps for CONDITIONAL-ADD- k].

There also exist more space-efficient versions of TEST-EQUALITY-TO- k and TEST-GREATER-THAN- k . For the former case, the program is as follows (where again $n=m^2$).

Program TEST-EQUALITY-TO'- k

quantum registers:

X : n -bit arithmetic register

B : bit register

C_0, C_1, \dots, C_{m-1} : bit registers (initialized and finalized to 0)

D_1, D_2, \dots, D_m : bit registers (initialized and finalized to 0)

for $i=m-1$ **down to** 0 **do**

if $i=m-1$ **then** $D_m \leftarrow D_m \oplus 1$ **else**

$$D_m \leftarrow D_m \oplus C_{i+1}$$

for $j=m-1$ **down to** 1 **do**

$$D_j \leftarrow \leftarrow_R D_j \oplus (D_{j+1} \wedge (X_{im+j} = k_{im+j}))$$

$$C_i \leftarrow \leftarrow_R C_i \oplus (D_1 \wedge (X_{im} = k_{im}))$$

for $j=1$ **to** $m-1$ **do**

$$D_j \leftarrow \leftarrow_R D_j \oplus (D_{j+1} \wedge (X_{im+j} = k_{im+j}))$$

if $i=m-1$ **then** $D_m \leftarrow D_m \oplus 1$ **else**

$$D_m \leftarrow D_m \oplus C_{i+1}$$

$$B \leftarrow B \oplus C_0$$

for $i=0$ **to** $m-1$ **do**

if $i=m-1$ **then** $D_m \leftarrow D_m \oplus 1$ **else**

$$D_m \leftarrow D_m \oplus C_{i+1}$$

for $j=m-1$ **down to** 1 **do**

$$D_j \leftarrow \leftarrow_R D_j \oplus (D_{j+1} \wedge (X_{im+j} = k_{im+j}))$$

$$C_i \leftarrow \leftarrow_R C_i \oplus (D_1 \wedge (X_{im} = k_{im}))$$

for $j=1$ **to** $m-1$ **do**

$$D_j \leftarrow \leftarrow_R D_j \oplus (D_{j+1} \wedge (X_{im+j} = k_{im+j}))$$

if $i=m-1$ **then** $D_m \leftarrow D_m \oplus 1$ **else**

$$D_m \leftarrow D_m \oplus C_{i+1}$$

The above program uses $n + O(\sqrt{n})$ registers in total and runs in $4n + O(\sqrt{n})$ steps [compared to $2n + O(1)$ registers in total and $2n + O(1)$ steps for TEST-EQUALITY-TO- k].

The program TEST-GREATER-THAN'- k is similar to the above attaining $5n + O(\sqrt{n})$ time and $n + O(\sqrt{n})$ space [vs. $3n + O(1)$ time and $2n + O(1)$ space].

APPENDIX B: PHASE FREEDOM IN IMPLEMENTATION OF REVERSIBLE ROUTINES

Here we will explain ways in which the quantum phase can be treated in the essentially classical reversible routines which we have been discussing throughout this paper. In the language of quantum logic gates, the bit-level logic statements used in the programs here are represented by unitary matrices applied to the quantum wave function of all the registers. These unitary matrices have a special restriction which make them "classical," which is that the matrix elements are only 0 or 1; this means that every definite computational state $|x\rangle$ is taken to another definite computational state $|f(x)\rangle$, and not to a superposition of states. To give an example, the Toffoli gate, the three-bit implementation of the AND gate in reversible logic, involves the following unitary matrix:

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}. \quad (\text{B1})$$

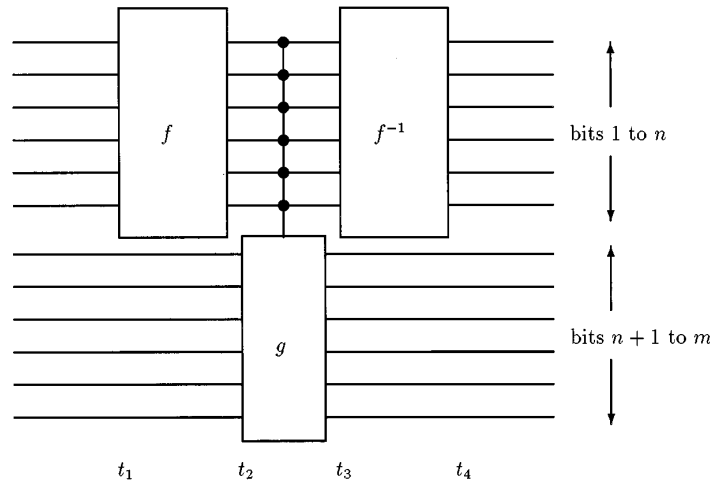
Here we consider the question of whether elementary operations with modified phases (i.e., in which the matrix elements are unimodular complex numbers $e^{i\theta}$, rather than being 1) could be used in the implementation of the Schumacher function. We are motivated to investigate this because we found in our previous study [11] that the implementation of a modified Toffoli gate with a single non-zero phase

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}, \tag{B2}$$

requires fewer resources in the following sense: we showed that the zero-phase Toffoli gate [Eq. (B1)] can be implemented with six two-bit XOR gate and eight one-bit gates,

while the Toffoli gate with modified phases [Eq. (B2)] requires only three XOR's and four one-bit gates. We will establish here that the less-costly gate can in fact be used for most of the Toffoli gates, and related three-bit operations, that are used in the implementation of the Schumacher compression function.

Note that it is necessary that the complete Schumacher calculation be carried out with all the quantum phases equal to zero, in order that the superposition states discussed in Sec. I maintain the correct phase relation to one another. Thus the question becomes: how can the effect of the non-zero phase in Eq. (B2), if it is introduced in one Toffoli gate, be undone at some later step of the calculation? The answer (which we will establish shortly) is the obvious one: many of the reversible routines which we have introduced (although not the high-level Schumacher program itself) have a palindromic character, so that a Toffoli gate on three bits is exactly "undone" at a later stage of the computation, roughly as far from the end of the subroutine as the original gate is from the beginning. It turns out that the effect of the -1 phase factor can be precisely undone at the second occurrence of the gate, too.



We will now establish the desired basic result using the setup of the figure, that the Boolean function f can be implemented with any arbitrary phase factors, so long as they also appear in f^{-1} , no matter what the intervening Boolean function g is, so long as g does not modify the values of the bits on which f and f^{-1} act. By applying this results repeatedly to the subroutines which we have introduced, starting at the innermost level, we deduce all the three-bit primitives which can be implemented with nonzero phase. Assignments in which these nonzero phases are permitted have been identified by the special assignment symbol

$$\leftarrow \in. \tag{B3}$$

These statements are always paired with others, denoted by

$$\leftarrow \in_R \tag{B4}$$

in which the reverse phases are implemented. [For the phases in Eq. (B2), the operation is self-inverse.] In one case, the pairing is between statements in different, palindromically arranged calls of the same subroutine; for these we have used a distinct symbol

$$\leftarrow \leftarrow. \tag{B5}$$

After establishing the basic result, we will review a few of the details of this implementation in the Schumacher function.

Let us first write down what the set of operations in the figure is supposed to do. Beginning with the basis state

$$|x\rangle = |\overbrace{x_1 x_2 \dots x_n \dots x_m}^{m \text{ bits}}\rangle \quad (\text{B6})$$

at time t_1 , it becomes at t_2 , after the operation of f ,

$$|x'\rangle = |\overbrace{f(x_1 x_2 \dots x_n)}^{n \text{ bits}} x_{n+1} \dots x_m\rangle \quad (\text{B7})$$

Then at time t_3 the state is

$$\exp(i\theta_g(x')) |\overbrace{f(x_1 x_2 \dots x_n)}^{n \text{ bits}} \underbrace{g(x')}_{m-n \text{ bits}}\rangle. \quad (\text{B8})$$

$g(x')$ depends on the state of the entire m -bit register x' , but only modifies the last $m-n$ bits, as indicated. Note that we allow for the possibility that g itself is a modified Boolean function with nonzero phases. This is necessary because we will apply this result in a nested fashion in the Schumacher subroutines. Finally at time t_4 the state is

$$\exp[i\theta_g(x)] |x_1 x_2 \dots x_n g(x')\rangle. \quad (\text{B9})$$

That is, the first n bits are restored to their original state, and bits $n+1$ through m remain in the state $g(x)$.

Now, the question is, will the state Eq. (B9) still result if the function f is modified to introduce nonzero phases $\theta_f(x_1 \dots x_n)$? If we establish that this is true for all Boolean inputs $|x\rangle$, this will suffice to prove that these networks have the same action on any arbitrary quantum states (this follows directly from the linear superposition principle of quantum mechanics). We follow the time evolution as before with the modified f . At time t_2 the state is

$$\exp[i\theta_f(x_1 \dots x_n)] |f(x_1 x_2 \dots x_n) x_{n+1} \dots x_m\rangle. \quad (\text{B10})$$

Then at time t_3 :

$$\exp\{i[\theta_g(x') + \theta_f(x_1 \dots x_n)]\} |f(x_1 x_2 \dots x_n) g(x')\rangle \quad (\text{B11})$$

and finally at t_4 :

$$\exp\{i[\theta_f(x_1 \dots x_n) + \theta_g(x') + \theta_{f^{-1}}(f(x_1 \dots x_n))]\} |x_1 x_2 \dots x_n g(x')\rangle. \quad (\text{B12})$$

The final term in the phase factor can be simplified. Recall that the unitary transformation corresponding to f^{-1} is the transpose of the complex conjugate of the unitary transformation corresponding to f . (This follows directly from the definition of unitarity.) Therefore, to get $\theta_{f^{-1}}$ from θ_f , we flip the sign (this is the complex conjugation), and we make the argument of the θ function the *output* values of the bits

rather than the *input* values (this is the transpose). Here we use the fact that g does not modify the first n bits—their output values are the same as the original inputs $x_1 x_2 \dots x_n$. Rendering this in mathematical language:

$$\theta_{f^{-1}}(f(x_1 \dots x_n)) = -\theta_f(x_1 \dots x_n). \quad (\text{B13})$$

Thus, the two θ_f terms in the phase in Eq. (B12) cancel out, and Eq. (B12) becomes identical to Eq. (B9), which is the desired result QED.

Finally, we briefly review the application of this result to the programs introduced in the text. The first appearance is in CONDITIONAL-ADD- k , where the role of f is played in the innermost part of the program by the assignment statement

$$C_{n-1} \leftarrow C_{n-1} \oplus \text{MAJ}(k_{n-2}, X_{n-2}, C_{n-2}).$$

This is a three-bit operation of the Toffoli type (or a trivial modification of it) involving the bits C_{n-1} , C_{n-2} , and X_{n-2} . f^{-1} occurs a short distance down,

$$C_{n-1} \leftarrow {}_R C_{n-1} \oplus \text{MAJ}(k_{n-2}, X_{n-2}, C_{n-2}).$$

The role of g is played by the two statements

$$X_{n-1} \leftarrow X_{n-1} \oplus (k_{n-1} \wedge B),$$

$$X_{n-1} \leftarrow X_{n-1} \oplus (C_{n-1} \wedge B).$$

Obviously, only X_{n-1} is modified by g , so the condition that g modify only bits not touched by f is satisfied; so, we are allowed to introduce a phase-modified f as indicated by the \leftarrow and \leftarrow_R assignments. Moving away from the innermost part of the program, we see that all the above is nested inside a larger g in which C_{n-1} , X_{n-2} , and X_{n-1} are modified, surrounded by a $f-f^{-1}$ pair involving the bits C_{n-2} , X_{n-3} , and C_{n-3} ; working outward in succession this way, we conclude that all the C_i assignments may be replaced with phase-modified \leftarrow and \leftarrow_R assignments.

In Sec. IV B we exhibit a pair of statements,

$$B \leftarrow B \oplus (Y > I),$$

$$B \leftarrow {}_R B \oplus (Y > I),$$

playing the role of f and f^{-1} . These are not primitive three-bit operations as in the earlier examples, but they are themselves implemented with bit-level programs (TEST-GREATER-THAN- k). For this f^{-1} , the \leftarrow_R assignment requires that the bit-level routine be run in the time-reversed order. This can be done since classically, this Boolean function is its own inverse. In the time-inverted TEST-GREATER-THAN- k , the \leftarrow 's and \leftarrow_R 's should be interchanged. The B assignment involving the symbol \leftarrow in this routine is special, in that it is paired with the same statement in the time-reversed call to TEST-GREATER-THAN- k . This special symbol is a reminder that this statement should be implemented with the phases corresponding to the \leftarrow assignment in the first call to the program, and with those corresponding to the \leftarrow_R assignment in the second call.

We have not indicated phase-modifying assignments for any of the two-bit gate level operations in these programs. We take as given that these two-bit gates could be implemented with zero phases. But if this were not the case, then many of these paired assignments may be phase-modified in exactly the way we have shown for the three-bit primitives.

- [1] C. H. Bennett, G. Brassard, and A. K. Ekert, *Sci. Am.* **267** (4), 50 (1992).
- [2] A. Ekert and R. Jozsa, *Rev. Mod. Phys.* **68**, 733 (1996).
- [3] B. Schumacher, *Phys. Rev. A* **51**, 2738 (1995).
- [4] R. Jozsa and B. Schumacher, *J. Mod. Opt.* **41**, 2343 (1994).
- [5] There has been further recent work on qubit string compression, which we do not touch upon in the present paper. See Howard Barnum, Christopher A. Fuchs, Richard Jozsa, and Benjamin Schumacher, Report No. quant-ph/9603014 (1996).
- [6] C. E. Shannon, *Bell Syst. Tech. J.* **27**, 379 (1948); **27**, 623 (1948).
- [7] C. H. Bennett, *Phys. Today* **48** (10), 24 (1995).
- [8] W. G. Unruh, *Phys. Rev. A* **51**, 992 (1995).
- [9] G. M. Palma, K.-A. Suominen, and A. Ekert, *Proc. R. Soc. London Ser. A* **452**, 567 (1996).
- [10] We adopt a notation which is common in textbooks on algorithm theory; see A. Aho, J. E. Hopcroft, and J. D. Ullman, *The Design and Analysis of Computer Algorithms* (Addison-Wesley, New York, 1974); G. Brassard and P. Bratley, *Algorithmics: Theory and Practice* (Prentice-Hall, Englewood Cliffs, NJ, 1988); T. H. Cormen, C. E. Leiserson, and R. L. Rivest, *Introduction to Algorithms* (McGraw-Hill, New York, 1990).
- [11] A. Barenco, C. H. Bennett, R. Cleve, D. P. DiVincenzo, N. Margolus, P. Shor, T. Sleator, J. Smolin, and H. Weinfurter, *Phys. Rev. A* **52**, 3457 (1995).
- [12] I. Chuang and Y. Yamamoto, *Phys. Rev. A* **52**, 3489 (1995).
- [13] V. Vedral, A. Barenco, and A. Ekert, Report No. quant-ph/9511018 (1995).
- [14] See also D. Beckman, A. N. Chari, S. Devabhaktuni, and J. Preskill, Report No. quant-ph/9602016 (1996).
- [15] C. H. Bennett, *SIAM J. Comput.* **18**, 766 (1989); *IBM J. Res. Dev.* **17**, 525 (1973).
- [16] See, e.g., H. F. Chau and H.-K. Lo, Report No. quant-ph/9506012 (1995).
- [17] T. Toffoli, in *Automata, Languages and Programming*, edited by J. W. de Bakker and J. van Leeuwen (Springer, New York, 1980), p. 632; Technical Memo No. MIT/LCS/TM-151 (unpublished).
- [18] I. Koren, *Computer Arithmetic Algorithms* (Prentice Hall, Englewood Cliffs, NJ, 1993).
- [19] D. P. DiVincenzo, *Phys. Rev. A* **51**, 1015 (1995); D. Deutsch, A. Barenco, and A. Ekert, *Proc. R. Soc. London Ser. A* **449**, 669 (1995); S. Lloyd, *Phys. Rev. Lett.* **75**, 346 (1995).