PHYSICAL REVIEW A **87**, 060304(R) (2013)

Streamlining Shor's algorithm for potential hardware savings

Y. S. Nam and R. Blümel
Department of Physics, Wesleyan University, Middletown, Connecticut 06459-0155, USA
(Received 14 May 2013; published 26 June 2013)

We constructed a virtual quantum computer by running a complete, scaling, quantum-gate–by–quantum-gate implementation of Shor's algorithm on a 128-core classical cluster computer. In mode A [quantum period finding (PF) only, supplied with classical results for the modular exponentiation (ME) part of Shor's algorithm], factoring semiprimes up to $N = 557\,993$ with up to $n = 39$ qubits, we confirm earlier, smaller-n results concerning the performance scaling of Shor's algorithm equipped with a truncated (banded) quantum Fourier transform. Running our virtual quantum computer in mode B (full quantum implementation of ME and PF), we find that a large number of gates may be discarded in a scalable way in both the ME and PF parts of Shor's algorithm in exchange for only a small reduction in performance. We explicitly state the associated scaling laws. Implying significant savings in quantum gates, we suggest that these results are of importance for future experimental and technical large-n implementations of quantum computers.

DOI: 10.1103/PhysRevA.87.060304 PACS number(s): 03.67.Lx

Most computations of practical interest cannot be performed on classical, digital computers because they exceed the capabilities of even the largest presently existing supercomputers. A well-known example of importance in cryptanalysis is the factorization of large semiprimes $N = pq$, where p and q are prime numbers of about equal size [1]. Solution of the factorization problem would immediately break many popular encryption schemes, such as the RSA (Rivest, Shamir, and Adleman) encryption scheme [2], and would immediately reveal untold scores of government, military, and bank secrets. However, even if we build a classical supercomputer combining the resources of the entire known universe, we would still not be able to factor a relatively modest-sized semiprime with 5000 decimal digits [3]. Therefore, encryption codes that rely on the difficulty of integer factorization are considered secure—for now.

Classical digital computing, however, is not the only form of information processing. Since the early 1980s, we have known that a quantum computer [4–6], a qualitatively new form of information processor, is capable of solving problems that are strictly beyond the powers of any conceivable classical computer. In the world of integer factoring this is impressively demonstrated by Shor's algorithm [7], which is exponentially more powerful than any currently known classical factoring algorithm. In particular, at least in principle, a quantum computer running Shor's algorithm is powerful enough to break currently employed RSA-based encryption codes [8–10].

The quantum core of Shor's algorithm may be broken down into two parts: (1) modular exponentiation (ME) and (2) period finding (PF). Given an integer x, relatively prime to N, the ME part of Shor's algorithm determines the values $f(r) = x^r \mod N$ for integer exponents r. This is followed by the PF part of Shor's algorithm, which, supplemented with an efficient classical algorithm (classical postprocessing [8–10]), performs a quantum Fourier transform (QFT) [8–10] to determine the period ω of f, i.e., $f(r + \omega) = f(r)$. Given the period ω and a few additional conditions that are straightforwardly accommodated in practice [8–10], the factors of N are then determined according to $p = \gcd(x^{\omega/2} - 1, N)$ and $q = \gcd(x^{\omega/2} + 1, N)$, where gcd is the greatest common divisor, efficiently computed via Euclid's algorithm [1].

Because of its importance, several experimental groups have implemented demonstration models of quantum computers that are capable of factoring small semiprimes [11–15]. Although these experiments employ quantum circuits that implement both parts of Shor's algorithm, i.e., ME and PF, the quantum circuits are tailor-made and optimized for a single semiprime N. Thus, these quantum circuits do not scale; i.e., they are not capable of factoring any other N than the one they are designed for. The high degree of specialization in the experimental implementations of ME and PF is necessary since even with today's standards of exquisite quantum control, it is still a daunting task to coherently control more than a handful of qubits. Therefore, experimental implementations of quantum computers are currently limited to $N \leqslant 21$ [11–15].

In order to overcome the current experimental limitations on N and to be able to study, theoretically, the performance of Shor's algorithm under various conditions, we implemented a complete quantum-gate–by–quantum-gate simulation of a scaling, virtual quantum computer on a 128-core classical cluster computer. We designed our virtual quantum computer to run in two modes: (A) PF only, supplied with ME performed classically, and (B) both ME and PF executed quantum mechanically. With our currently available classical computing resources, running our virtual quantum computer in mode A, we are able to factor all semiprimes up to and including $N = 557\,993$. Running the quantum computer in mode B, we are able to factor all semiprimes up to and including $N = 57$. We view our virtual quantum computer as a convenient virtual quantum laboratory that lets us investigate the effects of various (scaling) optimizations. One of these optimizations is the use of a banded QFT and its influence on the performance of Shor's algorithm [3,16,17]. Running our virtual quantum computer in mode A, we report results that test our performance scaling laws [3,17] in the region of up to 39 qubits. Running our quantum computer in mode B, we report results on the effects of a scaling optimization of the adder and ME-QFT components [18] of ME.

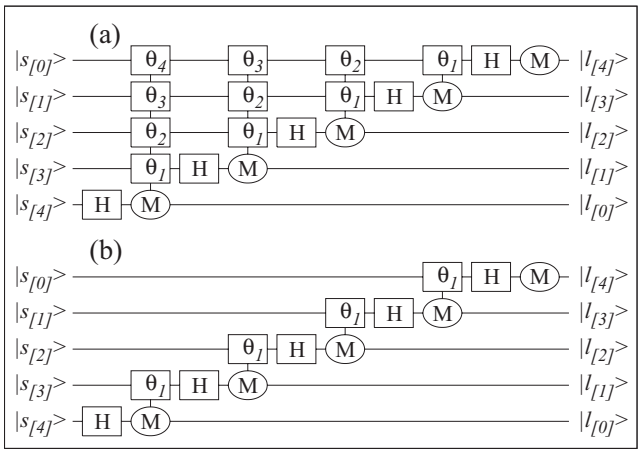

FIG. 1. Logic circuit of a five-qubit example of the Griffiths-Niu QFT [19], illustrating the concept of bandwidth, defined as the number b of off-diagonal quantum states coupled by the QFT: (a) full implementation (bandwidth $b = 4$); (b) truncated implementation (bandwidth $b = 1$). H, θ, and M denote Hadamard, single-qubit conditional rotation, and measurement gates, respectively.

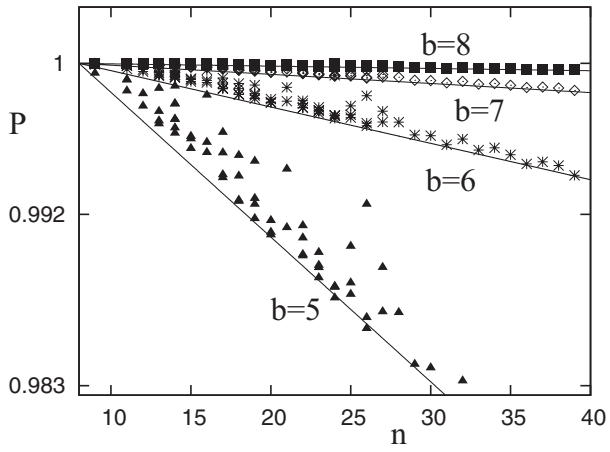

FIG. 2. Scaled probability for bandwidths $b = 5$ (triangles), 6 (asterisks), 7 (diamonds), and 8 (squares) for n ranging from 8 to 39. The data for $n = 34, \ldots, 39$ are consistent with the scaling law (3) (solid lines) predicted in [3,17].

Given the experimental challenges in achieving coherent quantum control of many qubits simultaneously, experimental implementations of Shor's algorithm are facilitated if the algorithm itself can be simplified. We distinguish two types of simplifications: performance conserving and performance changing. One such simplification, a performance-conserving optimization, is the substitution of the fully coherent QFT [8–10] in Shor's algorithm by a semiclassical version due to Griffiths and Niu [19]. A five-qubit circuit of the Griffiths-Niu QFT is shown in Fig. 1(a). Although the Griffiths-Niu version of the QFT replaces all two-qubit gates in the fully coherent QFT by (controlled) single-qubit gates and destroys phases as a result of measurement (see M gates in Fig. 1), it is exact when used in conjunction with the PF part of Shor's algorithm. Our virtual quantum computer is equipped with a banded version [3,16,17,20] of the Griffiths-Niu QFT [see Fig. 1(b)]. This means that in addition to using only single-qubit gates, as shown in Fig. 1(a), we also retain only coupling to b nearest-neighbor qubits, which results in a banded structure of the quantum circuit [see Fig. 1(b)]. Of course, denoting by n the number of qubits of our virtual quantum computer, the exact case is included for the choice $b = n - 1$. While the case $b = n - 1$ is performance conserving, $b < n - 1$ is not but leads to substantial savings in quantum gates. Therefore, for given n, our goal is to determine the optimal choice of b that corresponds to maximal savings in gates for still acceptable quantum computer performance.

Quantifying our performance measure, we note that the probability of finding our quantum computer in state $|l\rangle$ after running Shor's algorithm for a specific semiprime N with maximal bandwidth $b = n - 1$ (exact case) is [3]

$$\tilde{P}_{b=n-1}(l) = \frac{\sin^2(K\pi\omega l/2^n)}{2^n K \sin^2(\pi\omega l/2^n)}, \tag{1}$$

where $K \approx 2^n/\omega$. Apparently, $\tilde{P}_{b=n-1}(l)$ exhibits ω peaks $\tilde{l}_j$ located at integer multiples of K, i.e., $\tilde{l}_j = jK$, $j = 0, \ldots, \omega - 1$. Running our quantum computer with $b < n - 1$, we find that the probabilities of all quantum states $|l\rangle$, where l is in the vicinity of $\tilde{l}_j$, i.e., states useful for factoring [8–10], respond in unison to the reduction of b [3]. Therefore, defining l_j as the closest integer to $\tilde{l}_j$, we find it useful and convenient to measure the performance of the banded Shor algorithm using the normalized performance measure [3]

$$P_b(n) = \sum_{j=0}^{\omega-1} \tilde{P}_b(l_j) \Big/ \sum_{j=0}^{\omega-1} \tilde{P}_{b=n-1}(l_j), \tag{2}$$

where $\tilde{P}_b(l)$ is the probability of collapse into the state $|l\rangle$ if the full QFT is replaced by its banded version of bandwidth b [see Fig. 1(b)]. Previously, we were able to compute $P_b(n)$ for quantum computers with up to 33 qubits [3] and found the scaling law

$$P_b(n) = \exp[-1.1 \times 2^{-2b}(n - 8)] \tag{3}$$

(solid, straight lines in Fig. 2). Testing our earlier results, we report here calculations of $P_b(n)$ that extend the range of qubits to $n = 39$. The plot symbols in Fig. 2 include our data for $P_b(n)$ in the range from $n = 34$ to 39. These data were obtained by running our virtual n-qubit quantum computer for various N, ranging from $N = 116\,939$ to $N = 557\,993$. These computations are extensive, so that in the range from $n = 34$ to $n = 39$ we can only afford to choose a single sample N for each n. We chose $N = 116\,939 = 337 \times 347$ for $n = 34$, $N = 171\,371 = 409 \times 419$ for $n = 35$, $N = 239\,117 = 487 \times 491$ for $n = 36$, $N = 265\,189 = 509 \times 521$ for $n = 37$, $N = 378\,221 = 613 \times 617$ for $n = 38$, and $N = 557\,993 = 743 \times 751$ for $n = 39$. Our chosen N values are products of consecutive primes. This is obviously not useful for cryptological applications but emulates the case $p \approx q$, known to be the most difficult case to factor [1]. For each chosen N, we determine all of its orders ω (each of these N has up to 72 different orders, which all require quantum processing) and then compute the ω-averaged P_b [3], which then appears as a plot symbol in Fig. 2. These results confirm the scaling law (3), thus confirming the conclusion in [3] that a substantial number of quantum gates can be saved by pruning

the QFT down to bandwidths around $b = 8$ (although based on a slightly different scaling law, the same conclusion was reached by the authors of [16]). This is seen clearly in Fig. 2, which shows that for $b = 8$ the performance of the quantum computer is very close to 1 for all n ranging up to $n = 39$.

We now turn to our mode-B calculations. Supplementing the QFT with quantum circuitry as described in [18], we obtain a complete quantum-gate–by–quantum-gate implementation of a virtual quantum computer that currently runs on a 128-core classical cluster computer. Our gate-by-gate implementation gives us access to each individual quantum gate and allows us to investigate the effects of gate pruning on the performance of the quantum computer. Another advantage of our implementation is that it allows us to experiment with the factoring of actual semiprimes N. We report here results for $N = 21$, although, given our classical hardware resources, $N = 33, 35, 39, 51, 55$, and 57 are within our reach [21]. Running test cases for $N = 15$ and $N = 21$ and comparing the computational results to the theoretically expected result (1), we verified that our virtual quantum computer is implemented correctly.

Relating to our experiments with a reduced bandwidth b in the PF part of Shor's algorithm and making use of our access to each individual quantum gate of our virtual quantum computer, we introduced the bandwidth b_{ME} in the adder and ME-QFT parts [18] of Shor's algorithm by removing all gates causing single-qubit phase shifts of less than $\exp(i\pi/2^{b_{\mathrm{ME}}})$. For fixed b_{ME} we compute $P_b(b_{\mathrm{ME}})$ as in (2) and define $\Gamma_b(b_{\mathrm{ME}}) = 1 - P_b(b_{\mathrm{ME}})$. We chose $N = 21$ with $n = 10$ to investigate the effects of this pruning operation. Figure 3 shows $\Gamma_b = 1 - P_b$ for our test case $N = 21$ with b ranging from 1 to 7. We find that Γ_b depends only weakly on b_{ME} and decays exponentially for increasing b. This implies that satisfactory performance of Shor's algorithm can be achieved with relatively small b (b_{ME}), resulting in tremendous savings in quantum gates. Quantitatively, we find that

$$\Gamma_b \approx 2^{-2b}, \tag{4}$$

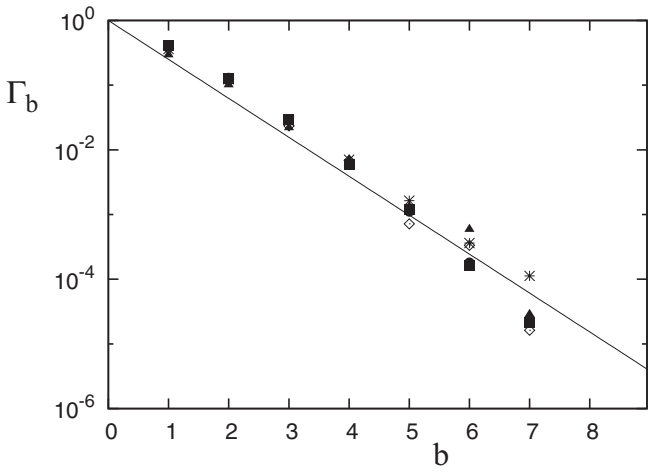

FIG. 3. Complement $\Gamma_b(b_{\mathrm{ME}}) = 1 - P_b(b_{\mathrm{ME}})$ of quantum computer performance as a function of b for five different values of b_{ME}. Triangles: $b_{\mathrm{ME}} = 1$; asterisks: $b_{\mathrm{ME}} = 2$; diamonds: $b_{\mathrm{ME}} = 3$; squares: $b_{\mathrm{ME}} = 4$; circles: $b_{\mathrm{ME}} = 5$. The solid line is the scaling function (4).

which is reminiscent of the b scaling in mode A [see (3)]. Because exponentially many quantum operations need to be simulated, our quantum computer currently runs at an accuracy of $\approx 10^{-5}$. The point corresponding to $b = 8$ in Fig. 3 is not shown because it is at the limit of our accuracy and therefore unreliable.

Multiprocessor simulations of Shor's algorithm on classical hardware are not new (see, e.g., [22–25]), and applications range from proofs of principle of the suitability of the multiprocessor architecture [22,23] and the creation of convenient parallel-computing environments [25] to massively parallel implementations on state-of-the-art supercomputer facilities [24]. While we cannot compete with supercomputer implementations of Shor's algorithm as far as raw computing power is concerned, the results reported in this paper push the envelope in different ways. Our mode-A calculations allow us to confirm the scaling behavior of Shor's algorithm equipped with a banded QFT for up to $n = 39$ qubits, which is substantially larger (the execution time essentially doubles with each unit increase in n) than the number of qubits used in earlier investigations of this type [3,16,17]. The results of our calculations confirm our analytical model of performance scaling [3,17], which predicts that the scaling (3) persists for quantum computers with several thousand qubits, relevant for factoring semiprimes of practical interest. However, our main result is the realization that the ME part of Shor's algorithm may be banded in the same way as the PF part, resulting in the scaling law (4). This proves that the adder and ME-QFT components of ME may be considerably streamlined, resulting in substantial savings in quantum gates in exchange for only a negligible reduction in performance. Detailed circuit diagrams illustrating the ME pruning operation defined above will be published elsewhere [21].

We also mention the investigations by García-Mata *et al.* [26,27]. These investigations are related to our mode-A calculations since in [26,27] the ME part of Shor's algorithm is represented by the product of unitary matrices computed classically. However, the focus in [26,27] is not on the effects of bandedness but on the influence of noise on the performance of Shor's algorithm. This motivates the question of whether in the presence of noise and decoherence a larger bandwidth b (b_{ME}) may be required than predicted by our noise-free model, possibly erasing the benefits that a small b (b_{ME}) entails. For the following simple reason this is highly unlikely. The gates pruned now are (classically controlled) single-qubit rotation gates with exponentially small rotation angles. Because the rotation angles are so small, these gates are easily drowned out by noise, and instead of performing their function, they would merely act as "antennas" to pick up noise and channel it into the quantum circuit. Therefore, in the presence of noise, it may actually be beneficial to prune even more gates, i.e., to work with an even smaller bandwidth than indicated by the noise-free model in order to avoid this "antenna effect." Our preliminary calculations confirm these conclusions and will be reported elsewhere [21].

The computations reported in this paper are expensive. They took three months to execute on a 128-core cluster-computer and are thus at the limit of computer power that even a university computing center can provide for a single research group. Therefore, we need streamlined versions of Shor's

algorithm not only for efficient practical implementations of Shor's algorithm but also for quantum simulations performed on classical computers in order to be able to explore the high-n regions of practical importance for meaningful quantum computations.

In summary, we presented some recent results on extending our mode-A calculations to 39 qubits, thereby testing and confirming our scaling laws for finite-bandwidth quantum computer performance [3,17]. These calculations involved quantum factorization of actual semiprimes up to and including $N = 557\,993$. Running our quantum computer in mode B, we showed that quantum adders and ME-QFTs may be banded without significant loss in factoring performance. We are sure that this will be of considerable interest for technological implementations of Shor's algorithm.

The authors are grateful for a generous allotment of computer resources by the Wesleyan Scientific Computing Center. R.B. acknowledges financial support by Grant No. 216687 of the Research Council of Norway.

[1] N. Koblitz, *A Course in Number Theory and Cryptography* (Springer, New York, 1994).
[2] R. Rivest, A. Shamir, and L. Adleman, Commun. ACM **21**, 120 (1978).
[3] Y. S. Nam and R. Blümel, Phys. Rev. A **87**, 032333 (2013).
[4] D. Deutsch and A. Ekert, Phys. World **11**, 47 (1998).
[5] R. Feynman, Int. J. Theor. Phys. **21**, 467 (1982).
[6] R. Feynman, Found. Phys. **16**, 507 (1986).
[7] P. W. Shor, in *Proceedings of the 35th Annual Symposium on the Foundations of Computer Science*, edited by S. Goldwasser (IEEE Press, Los Alamitos, CA, 1994), pp. 124–134.
[8] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2000).
[9] N. D. Mermin, *Quantum Computer Science* (Cambridge University Press, Cambridge, 2007).
[10] R. Blümel, *Foundations of Quantum Mechanics: From Photons to Quantum Computers* (Jones and Bartlett, Sudbury, MA, 2010).
[11] L. M. K. Vandersypen, M. Steffen, G. Breyta, C. S. Yannoni, M. H. Sherwood, and I. L. Chuang, Nature (London) **414**, 883 (2001).
[12] C.-Y. Lu, D. E. Browne, T. Yang, and J.-W. Pan, Phys. Rev. Lett. **99**, 250504 (2007).
[13] B. P. Lanyon, T. J. Weinhold, N. K. Langford, M. Barbieri, D. F. V. James, A. Gilchrist, and A. G. White, Phys. Rev. Lett. **99**, 250505 (2007).
[14] A. Politi, J. C. F. Matthews, and J. L. O'Brien, Science **325**, 1221 (2009).
[15] E. Martín-López, A. Laing, T. Lawson, R. Alvarez, X.-Q. Zhou, and J. L. O'Brien, Nat. Photonics **6**, 773 (2012).
[16] A. G. Fowler and L. C. L. Hollenberg, Phys. Rev. A **70**, 032329 (2004).
[17] Y. S. Nam and R. Blümel, Phys. Rev. A **86**, 044303 (2012).
[18] S. Beauregard, Quantum Inf. Comput. **3**, 175 (2003).
[19] R. B. Griffiths and C.-S. Niu, Phys. Rev. Lett. **76**, 3228 (1996).
[20] D. Coppersmith, arXiv:quant-ph/0201067.
[21] Y. S. Nam and R. Blümel (unpublished).
[22] K. M. Obenland and A. M. Despain, arXiv:quant-ph/9804039.
[23] J. Niwa, K. Matsumoto, and H. Imai, Phys. Rev. A **66**, 062317 (2002).
[24] K. De Raedt, K. Michielsen, H. De Raedt, B. Trieu, G. Arnold, M. Richter, Th. Lippert, H. Watanabe, and N. Ito, Comput. Phys. Commun. **176**, 121 (2007).
[25] F. Tabakin and B. Juliá-Díaz, Comput. Phys. Commun. **180**, 948 (2009).
[26] I. García-Mata, K. M. Frahm, and D. L. Shepelyansky, Phys. Rev. A **75**, 052311 (2007).
[27] I. García-Mata, K. M. Frahm, and D. L. Shepelyansky, Phys. Rev. A **78**, 062323 (2008).