Tree tensor network approach to simulating Shor's algorithm

Eugene Dumitrescu
Quantum Computing Institute, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831, USA
and Bredesen Center for Interdisciplinary Research, University of Tennessee, Knoxville, Tennessee 37996, USA
(Received 15 May 2017; revised manuscript received 4 August 2017; published 20 December 2017)

Constructively simulating quantum systems furthers our understanding of qualitative and quantitative features which may be analytically intractable. In this paper, we directly simulate and explore the entanglement structure present in the paradigmatic example for exponential quantum speedups: Shor's algorithm. To perform our simulation, we construct a dynamic tree tensor network which manifestly captures two salient circuit features for modular exponentiation. These are the natural two-register bipartition and the invariance of entanglement with respect to permutations of the top-register qubits. Our construction help identify the entanglement entropy properties, which we summarize by a scaling relation. Further, the tree network is efficiently projected onto a matrix product state from which we efficiently execute the quantum Fourier transform. Future simulation of quantum information states with tensor networks exploiting circuit symmetries is discussed.

DOI: 10.1103/PhysRevA.96.062322

Introduction. Tensor networks (TNs) are graphical data structures consisting of nodal tensors, with elements related to basis amplitudes, and indexed edges which represent the physical and virtual degrees of freedom of a quantum system. In certain cases, the observables of large systems can efficiently be calculated by utilizing TN decompositions of a quantum state [1]. TNs have been especially successful in identifying the ground states of local Hamiltonians in low-dimensional condensed matter systems [2,3]. In this work, we turn our attention to the simulation of quantum information theoretic systems [4–7] by introducing a tensor network.

Besides calculating observables, tensor networks provide explicit insight into entanglement structure, which differs vastly based on dimensionality and criticality [3,8,9]. However, finding an optimal tensor network representation for a given system is not a simple task. For example, no generally efficient method exists for $d \geqslant 2$-dimensional systems [1].

While lattice geometries are natural for condensed matter systems, dimensionality and local geometry are ill-defined quantities for states generated by logical quantum circuits on abstract registers of qubits. However, quantum algorithms do have important structural and symmetric properties. Thus, motivated by the algorithmic structure and entanglement invariance with respect to permutation of qubits in the modular exponentiation step, we construct a bipartite tree tensor network (TTN) naturally suited to simulating Shor's algorithm. Our numerical analysis verifies the volume law scaling relation given by Eq. (2).

While it is impossible to efficiently simulate general quantum algorithms, tensor networks tailored to quantum algorithms increase the size and complexity of classically simulable systems; in our case, 39-qubit Shor wave functions were constructed on a laptop computer. Such simulations are of practical interest for benchmarking noisy near-term quantum experiments [10,11] and may have some bearing on long-term quantum phase estimation or hidden subgroup algorithms (i.e., the TN constructed in this work is generalizable to period-finding algorithms with inter-register entanglement generated as illustrated in Fig. 1) [12].

Shor's wave function. We first outline the logical operations comprising Shor's algorithm [13] in order to develop an intuition for the form of an appropriate tensor network. To factor a natural number $N = pq$, with p,q large prime numbers, we draw a random integer $x \in \mathbb{Z}_N$ and use Shor's algorithm [13] to find the characteristic modular periodicity r given by $x^r \mod N = 1$. Assuming that $\gcd(x,N) = 1$ (in the unlikely case that $x = p$ or q, N is trivially factored) one initializes a $2l$ (l) qubit top (bottom) register, for a total of $3l$ qubits, where $l = \log_2(N)$ is the number of bits required to represent N. The top register is initialized into the product state $|+\rangle^{\otimes 2l} = 1/(\sqrt{2})^{2l} \sum_{i=0}^{2^{2l}-1} |i\rangle$ while the bottom register is initialized as $|1\rangle$ which is the integer basis representation of the computational basis state $|0,0,\ldots,0,1\rangle$. We represent the *bottom* register in the integer basis for the remainder of this paper. The composite initial product state is thus $|\Psi_i\rangle = |+\rangle_{\text{top}}^{\otimes 2l} \otimes |1\rangle_{\text{bot}}$.

The modular exponentiation (ME) unit (see Fig. 1, yellow box) entangles each top register qubit with the entire bottom register via controlled modular multiplication operators $U \equiv U(x,N)$. The operator U is a rank-2 tensor with dimensions $2^l \times 2^l$ and matrix elements satisfying $U|b\rangle = |xb \mod N\rangle$. Powers of U^{2^i} are generated by i iterative matrix multiplications. Upon application of the last controlled operator, the state reads $|\Psi\rangle = 2^{-l} \sum_{i=0}^{2^{2l}-1} |i\rangle \otimes |x^i \bmod N\rangle$. Because $x^r \mod N = 1$ we may group together like bottom register basis vectors and write the state as

$$|\Psi\rangle = \frac{1}{\sqrt{r}} \sum_{i=0}^{r-1} \left(\sum_{j=0}^{\lceil (2^{2l}-1)/r \rceil} |jr+i\rangle \right) \otimes |x^i \bmod N\rangle. \quad (1)$$

The entanglement growth across arbitrary bipartitions is exponential in the smaller bi-partition qubit volume. The growth saturates at a critical scale which is proportional to the modular periodicity. Using the structure and symmetries present in Shor's factoring quantum circuit, we construct a unique tensor network.

Bipartite tensor network complexity. Equation (1) is by definition a bipartite Schmidt decomposition between the two registers and reveals several interesting features. We see that $|\Psi\rangle$ is r-entangled across the bipartition; that is, the Schmidt coefficient $1/\sqrt{r}$ appears r times. In the worst

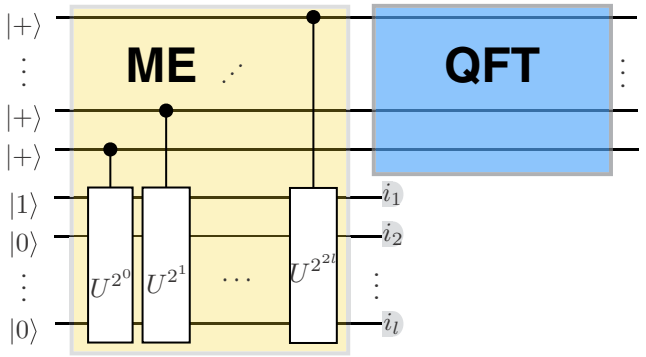

FIG. 1. Schematic of Shor's algorithm with the ME (QFT) components highlighted in yellow (blue) boxes. Equation (1) describes the state of the system one time step before the bottom register measurements preceding the ME subcircuit.

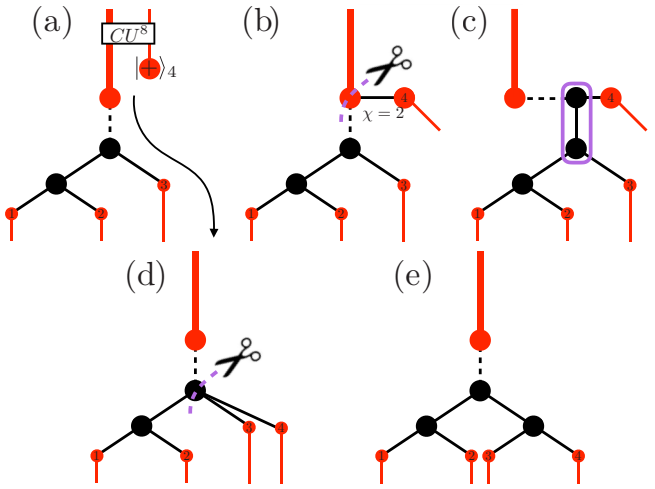

FIG. 2. Intermediate virtual networks for (b)–(d) between the application of a controlled modular multiplication gate (a) and the updated tree network (e). Red (black) lines denote physical (virtual) degrees of freedom. The thick red line at the base of the tree denotes the 2^l-dimensional qudit bottom register state. Top register qubits are represented by thin red lines at the bottom of the tree. The dashed black line denotes the bipartition defined by Eq. (1). The scissors icon and dashed purple lines (purple box) denote the bipartition chosen for tensor decompositions (contraction) generating the next tree configuration.

case $r \sim O(N)$ so the inter-register entanglement scales exponentially in the number of qubits l [14]. The equality of all Schmidt coefficients also foreshadows different correlation scalings compared to ground states of local Hamiltonians, which have exponentially (or power-law) decaying correlations. Equation (1) also demonstrates that it is natural to decompose $|\Psi\rangle$ across the inter-register bipartition, and we shall retain this feature in our tree network.

The quantum Fourier transformation (QFT) is known to be efficiently simulable [15,16], suggesting that the nontrivial part of the computation occurs during the modular exponentiation step. We thus pose the following question: what are the entanglement properties of the basis state $\sum_{j=0}^{\lceil (2^{2l}-1)/r \rceil} |jr+i\rangle$? The top register qubits are clearly entangled with one another via their interaction with the bottom register. We therefore know from Eq. (1) that r sets an upper bound on the amount of entanglement. Below we elucidate the entanglement properties of the top register basis states by developing a tensor network representation whose geometry is consistent with the inter-register bipartition and, more importantly, by the permutational invariance of the top register qubits [17,18].

A first attempt at a tensor network was performed in Ref. [7], which treats the bottom register as a qudit lying at one end of a matrix product state (MPS). The ME algorithm was performed by contracting two local controlled modular multiplication gates along with a series of swap gates. In doing so, the complexity of storing the state is reduced from $O(2^{3l}) \to O(2^l r) +$ constant with the constant given by $\sum_j 2 d_l^{(j)} d_r^{(j)}$, where $d_{l(r)}^{(j)}$ are the virtual bond dimensions to the right and left of the jth top register qubit. While this approach was successful in simulating Shor's algorithm, artificially large virtual bond dimensions were generated by successive swap operations. This leads to a situation where $d_{l(r)}^{(j)} = r$ for many bonds when, as we shall see, few virtual bonds of that size are necessary.

Tree-generation algorithm. We now introduce a natural tensor network which maintains the inter-register bipartition and distributes entanglement in an unbiased manner. This network is dynamically constructed by following the ME subcircuit, as shown in Fig. 1, with intermediate virtual updates performed, as shown in Fig. 2 and discussed below, in between circuit operations. Our construction algorithm goes as follows. (i) Contract the ith controlled U^{2^i} operator with the ith single-qubit $|+\rangle$ tensor and the current bottom qudit state as shown in panel (a). (ii) Perform internal operations updating and generating virtual indices. This cascades the ith qubit from the tree root (i.e., directly connected to the bottom register) to a new bottom tree branch as illustrated in panels (b)–(e). Repeat the procedure for all qubits indexed by $i \in \mathbb{Z}_{2l}$.

The internal updates consist of the following steps: (i) After applying a controlled U^{2^i} gate, a singular value decomposition (SVD) separates the ith qubit from the root qudit [Fig. 2, panel (b)]. The ith qubit is maximally entangled with the bottom register via a $\chi = 2$-dimensional auxiliary edge with singular values $(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$. (ii) Generate the new inter-register entanglement bond by performing an SVD between the bottom register and its local complement formed by the union of the new qubit and the previous tree root as indicated by the dashed purple line in panel (b). Recall that this bond's dimensionality eventually saturates at r. (iii) The tensors encircled by the purple box in panel (c) are contracted in order to "lower" the qubit through the tree before (iv) another SVD along a bipartition, which is chosen to direct the qubit through a specific path, is performed. Step (iv) is identical to step (ii) but occurs further down the tree. Repeat steps (iii) and (iv) until each qubit settles into its final location at the bottom of the tree. An example of a final tree configuration is illustrated in Fig. 3.

Note that the choice of a binary tree is arbitrary and that the entanglement features discussed in the next section hold for an arbitrary tree data structure. Also note that the number of virtual updates cascading the ith qubit is clearly upper bounded

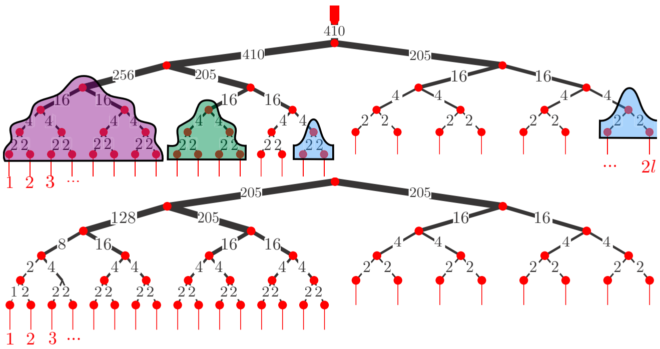

FIG. 3. Tree tensor network decomposition of $|\Psi\rangle$ from Eq. (1) for an $l = 12$ size system with $N = 3403 = 41\times 83$, $x = 346$, and cyclic order $r = 410$. Edge widths are $\log_2(d_i) + 1$ where d_i is the ith bond dimension which has been labeled. Top qubit (qudit) physical degrees of freedom appear as red edges along the bottom (top). Highlighted regions are entangled with their complement by a common parent branch whose dimension follows Eq. (2).

by final tree depth. Since the final tree depth is logarithmic in the number of qubits, that is with depth $\lceil \log_2(2l)\rceil$, the tree construction procedure is efficient. This trade-off can be compared to that in an MPS-based simulation for which at least $2l$ swaps are performed. Thus, a logarithmic number of updates to generate an unbiased and natural representation of the state is well justified and we now discuss the emergent entanglement properties.

Entanglement features. After cascading all qubits, the tree tensor network exactly encodes the wave function appearing in Eq. (1). The entanglement structure for the top register, not apparent in Eq. (1), is now revealed by the holographic dimension [19] constructed by our virtual updates. In our analysis we use the Schmidt number, given by the bond dimension of the virtual index connecting bipartitions, as the metric for shared entanglement. This is an appropriate metric because, unlike ground states of local Hamiltonians which have exponential or power-law decaying Schmidt coefficients, the Schmidt coefficients are *equal* in magnitude. Thus the Schmidt number completely describes the entanglement which can therefore be visualized as done in Fig. 3, where the drawn bonds are weighted as $\log_2(d_i) + 1$, where d_i is the local bond dimension which is also labeled.

At the bottom tree level where qubits first connect to their parent branches all qubits are *maximally* entangled to the rest of the network, with equal Schmidt coefficients $\lambda_0 = \lambda_1 = \frac{1}{\sqrt{2}}$. At the next level, all pairs of qubits (e.g., Fig. 3 blue highlighted qubits) are still maximally entangled with their complement, i.e., with degenerate Schmidt coefficients $\lambda_i = 0.5$ for $i = (0,1,2,3)$. This trend, with clusters of 2^n qubits maximally entangled to the rest of the state by 2^n identical Schmidt coefficients (e.g., Fig. 3 cluster of 4 green highlighted qubits, and so on) continues up to a critical size, at which point the entanglement rapidly saturates.

The qubit cluster size at which the entanglement scaling saturates depends on which qubits are selected and is either $l_r = \lceil \log_2(r)\rceil$ or $l_{\tilde r} = \lceil \log_2(\tilde r)\rceil$, where $\tilde r = r/2^m$ and m is the largest integer such that 2^m divides r. The critical length scale, specific to the choice of N and x determining r, can be understood by the following arguments. The r-dimensional tree root bond mediates the r-fold entanglement across the register bipartition as per Eq. (1). Further, r constrains the intra-register entanglement because all entanglement between qubits, stored in the bulk, was generated by controlled modular multiplication gates acting solely on the qudit. Descending from the tree root, bond dimensions decrease from r to either $\tilde r$, or powers of 2 less than $r,\tilde r$. An example is provided by Fig. 3(a) where we have illustrated the final tree generated for $N = 3403$. Thus the entanglement scales as

$$S = \begin{cases} 2^n, & \text{if } n < l_{r(\tilde r)},\\ r(\tilde r), & \text{otherwise},\end{cases} \tag{2}$$

where n refers to the number of qubits and the saturation dimension, r vs $\tilde r$, depends on whether qubits belong to the first l_r qubits (as seen from left to right in Fig. 3) or to the remainder of the register. Equation (2) defines a class of states whose entanglement is reminiscent of quantum error correcting codes with entanglement set by a distance d for $[[n,k,d]]$ codes [20].

We now address the seemingly strange feature of why the first l_r qubits are more entangled than the others. Note that modular exponentiation with at least l_r qubits is needed to generate the r-fold basis vectors on either side of the register bipartition. Modular exponentiation with the remaining qubits extends the orthonormal basis vectors $|jr + i\rangle$ as the Hilbert space grows, leaving the bipartite entanglement at order r. To understand why the latter qubits are less entangled than the former, consider a single orthonormal Schmidt vector $\sum_{j=0}^{\lceil (2^{2l}-1)/r\rceil} |jr + i\rangle$. The qubits are projected into such a state upon the measurement of the bottom register qudit. Since $r = 2^m\tilde r$ (if r is odd the algorithm restarts with a different x) and $i < r$, the last m bits for each $|jr + i\rangle$ are identical. Qubits 1 through m are therefore disentangled from the remaining state, and the remaining entanglement now follows the scaling law in Eq. (2) saturating at $\tilde r$. Figure 3(b) illustrates this point by replotting the tree after a qudit measurement and bond updates are performed. Note the changes in the bond dimensions along the left side of the tree.

MPS conversion and interpretation. We now briefly comment on the conversion of a tree network to an MPS network, which is useful in simulating Shor's algorithm in its entirety, due to an efficient representation of the QFT component for MPS systems [7,15,16]. A computational basis state measurement $\mathcal{M}_i$ [illustrated by the nodes indexed $(i_1,\ldots,i_l)$ in Fig. 1] projects the qudit register onto a basis state $|i\rangle$ from Eq. (1). We can probabilistically simulate this measurement by contracting the bottom register tensor with basis states $|i'\rangle$ to find nonzero matrix elements and then choose one randomly, with each measurement outcome being equally probable.

A series of virtual updates now reduce the tree network into the MPS form. The first (leftmost) qubit in Fig. 3 is already in the MPS form because a single tensor connects a physical and virtual degree of freedom. We proceed by selecting the next qubit and contracting its parent bonds until it connects to the virtual degree of freedom to the right of qubit 1. After contracting the parent bonds of qubit 3, a decomposition is performed to its right in order to generate a virtual MPS bond not connected to qubit 4. This procedure is iteratively repeated for all remaining qubits, at which point the resulting network

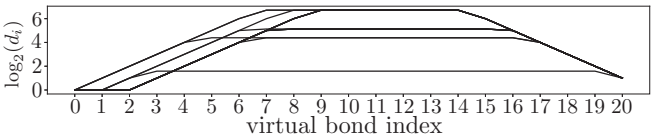

FIG. 4. Entanglement as quantified by MPS auxiliary bond dimensions for an $N = 1763, l = 12$ system. The virtual bond dimensions and entanglement grow exponentially between adjacent bipartitions for the first few sites from both the left and right end points and are saturated in the bulk of the system by the characteristic modular periodicity $\tilde{r}$.

is an MPS. Again, the number of updates is bounded by the number of qubits and the logarithmic tree depth.

Inspecting the MPS virtual Schmidt coefficients provides a complementary perspective to that provided by Fig. 3. In Fig. 4 we plot the virtual bond dimensions across the MPS network for 24 simulations involving the same $N = 1763$ with x randomly chosen. Many distinct x values share the same order r_x, so their entanglement spectrums are superimposed. The MPS spectrum in Fig. 4 verifies the entanglement scaling described by Eq. (2); namely, (i) the first m qubits are disentangled, (ii) entanglement grows exponentially up to a critical length scale, and (ii) the entanglement saturates at the scale $\tilde{r} \equiv r/2^m$.

Discussion and conclusion. Previous works applying TNs to quantum algorithms have either abstractly studied the complexity of TNs as they relate to quantum algorithms [5,6] or have directly simulated circuits using conventional TNs [7]. Going beyond this work, we have explicitly generated a tree tensor network using properties of Shor's algorithm. Our construction proceeded by contracting the ME controlled-U gates into the TN and performing a logarithmic number of virtual updates. A prevalent visible entanglement feature was the fully broadened spectrum of the Schmidt coefficients at all virtual cuts [14]. We also saw that the Schmidt rank across (global) system bipartitions scaled exponentially in the minimum number of qubits enclosed until saturation at critical threshold $\log_2(\tilde{r})$, where $\tilde{r} \propto r$. The volume entanglement scaling violates the entanglement area law and is like that recently observed in restricted Boltzmann neural networks [21].

Our construction is useful because it provides a simple and intuitive inspection of the quantum features for wave functions which are involved in the famous exponential quantum speedup. In doing so, our calculation reaffirms the difficulty in classically simulating Shor's algorithm. Further, due to the broad Schmidt spectrum, typical TN truncation techniques do not apply to our system. It is therefore interesting to consider alternative approximations. For example, a first approximation could be performed by simply eliminating a random set ratio of the bottom register basis vectors to enforce polynomial scaling. This would deform the modular multiplication operators and generate a state similar to Eq. (1) except for coefficients having support on the remaining basis vectors. However, we anticipate that the coherence of the modular periodicity r states would be destroyed by such methods.

Recently, quantum computing devices have been successfully scaled to intermediate system sizes in an attempt to tackle problems at the edge of classically intractability [10,11]. In order to validate the progress of quantum hardware, it is important to extend the reach of classical algorithms (as long as possible), especially when validation cannot be performed by conventional methodologies, e.g., with exponentially scaling state tomography. Our TN construction thus provides a powerful methodology for the verification of constantly growing quantum computations. For example, unitary noise can be directly simulated by biasing the unitary gates involved in the circuit away from their ideal description or by stochastically sampling additionally noisy unitary operations to model dephasing channels [22].

Acknowledgments. E.D. would like to thank R. Bennink for careful reading of the manuscript. Research sponsored by the Intelligence Community Postdoctoral Research Fellowship and the Laboratory Directed Research and Development Program of Oak Ridge National Laboratory, managed by UT-Battelle, LLC, for the US Department of Energy.

This manuscript has been authored by UT-Battelle, LLC, under Contract No. DE-AC0500OR22725 with the US Department of Energy. The United States Government retains and the publisher, by accepting the article for publication, acknowledges that the United States Government retains a non-exclusive, paid-up, irrevocable, world-wide license to publish or reproduce the published form of this manuscript, or allow others to do so, for the United States Government purposes. The Department of Energy will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan.

[1] R. Orus, A practical introduction to tensor networks: Matrix product states and projected entangled pair states, Ann. Phys. **349**, 117 (2014).

[2] S. R. White, Density Matrix Formulation for Quantum Renormalization Groups, Phys. Rev. Lett. **69**, 2863 (1992).

[3] U. Schollwoeck, The density-matrix renormalization group in the age of matrix product states, Ann. Phys. **326**, 96 (2011).

[4] T. H. Johnson, J. D. Biamonte, S. R. Clark, and D. Jaksch, Solving search problems by strongly simulating quantum circuits, Sci. Rep. **3**, 1235 (2013).

[5] Y.-Y. Shi, L.-M. Duan, and G. Vidal, Classical simulation of quantum many-body systems with a tree tensor network, Phys. Rev. A **74**, 022320 (2006).

[6] I. L. Markov and Y. Shi, Simulating quantum computation by contracting tensor networks, SIAM J. Comput. **38**, 963 (2008).

[7] D. S. Wang, C. D. Hill, and L. C. L. Hollenberg, Simulations of Shor's algorithm using matrix product states, arXiv:1501.07644.

[8] M. B. Hastings, An area law for one-dimensional quantum systems, J. Stat. Mech. (2007) P08024.

[9] B. Pirvu, G. Vidal, F. Verstraete, and L. Tagliacozzo, Matrix product states for critical spin chains: Finite-size versus finite-entanglement scaling, Phys. Rev. B **86**, 075117 (2012).

[10] S. Boixo, S. V. Isakov, V. N. Smelyanskiy, R. Babbush, N. Ding, Z. Jiang, M. J. Bremner, J. M. Martinis, and H. Neven, Characterizing quantum supremacy in near-term devices, arXiv:1608.00263.

[11] A. Kandala, A. Mezzacapo, K. Temme, M. Takita, M. Brink, J. M. Chow, and J. M. Gambetta, Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets, Nature **549**, 242 (2017).

[12] M. A. Nielsen and I. L. Chung, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, UK, 2010).

[13] P. W. Shor, Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer, SIAM J. Comput. **26**, 1484 (1997).

[14] R. Orus and J. I. Latorre, Universality of entanglement and quantum-computation complexity, Phys. Rev. A **69**, 052308 (2004).

[15] D. Aharonov, Z. Landau, and J. Makowsky, The quantum FFT can be classically simulated, arXiv:quant-ph/0611156.

[16] N. Yoran and A. J. Short, Efficient classical simulation of the approximate quantum Fourier transform, Phys. Rev. A **76**, 042321 (2007).

[17] R. Van Meter and K. M. Itoh, Fast quantum modular exponentiation, Phys. Rev. A **71**, 052320 (2005).

[18] A. Pavlidis and D. Gizopoulos, Fast quantum modular exponentiation architecture for Shor's factorization algorithm, Quantum Inf. Comput. **14**, 0649 (2014).

[19] R. Orus, Advances on tensor network theory: Symmetries, fermions, entanglement, and holography, Eur. Phys. J. B **87**, 280 (2014).

[20] D. Gottesman, Stabilizer codes and quantum error correction, arXiv:quant-ph/9705052.

[21] D.-L. Deng, X. Li, and S. Das Sarma, Quantum Entanglement in Neural Network States, Phys. Rev. X **7**, 021021 (2017).

[22] A. S. Darmawan and D. Poulin, Tensor-Network Simulations of the Surface Code under Realistic Noise, Phys. Rev. Lett. **119**, 040502 (2017).