Quantum Information and Computation, Vol. 17, No. 1&2 (2017) 0001–0014

SCALING AND EFFICIENT CLASSICAL SIMULATION OF THE QUANTUM FOURIER TRANSFORM

KIERAN J. WOOLFE
Center for Quantum Computation and Communication Technology, School of Physics, University of Melbourne Victoria, 3010, Australia

CHARLES D. HILL
Center for Quantum Computation and Communication Technology, School of Physics, University of Melbourne Victoria, 3010, Australia

LLOYD C.L. HOLIENBERG
Center for Quantum Computation and Communication Technology, School of Physics, University of Melbourne Victoria, 3010, Australia

Received May 25, 2016
Revised December 23, 2016

We provide numerical evidence that the quantum Fourier transform can be efficiently represented in a matrix product operator with a size growing relatively slowly with the number of qubits. Additionally, we numerically show that the tensors in the operator converge to a common tensor as the number of qubits in the transform increases. Together these results imply that the application of the quantum Fourier transform to a matrix product state with n qubits of maximum Schmidt rank χ can be simulated in $O(n\left(\log(n)\right)^2\chi^2)$ time. We perform such simulations and quantify the error involved in representing the transform as a matrix product operator and simulating the quantum Fourier transform of periodic states.

Keywords: Shor's algorithm, Matrix Product States, Tensor Networks, Entanglement
Communicated by: R Jozsa & B Terhal

1 Introduction

A central problem of quantum computing is determining the origin and nature of the speedup provided over classical computing. One approach to this problem is to study which classes of quantum computations can be simulated efficiently by classical means. Such computations must be missing a central feature of quantum computing, separating it from the classical counterpart. There have been many results in this area. These include the Gottesman-Knill theorem [1, 2], which states that circuits composed only of Clifford group gates can be efficiently simulated classically. Other results include the efficient classical simulation of match-gate circuits [3, 4], circuits which generate limited entanglement [5, 6], circuits whose graph representation has restricted topological properties [7, 8, 9] and circuits with sparse output distributions [10].

The quantum Fourier transform (QFT) is an important part of several quantum algorithms, including quantum simulation [11] and Shor's algorithm [12]. Each of these provides

an exponential speedup over the fastest known classical alternatives. In our discussion of the QFT we will focus on its role in Shor's algorithm. The QFT is the most intuitively quantum mechanical part of Shor's algorithm. That is, it contains Hadamard and controlled phase-rotation gates, neither of which have a classical analogue. The full QFT does not display any of the features found in previous studies to allow efficient classical simulation. Despite this, the approximate QFT (AQFT) is efficiently classically simulatable for input states with limited entanglement. This was first shown in [13, 14] using a tensor contraction simulation method. Together with results showing that the AQFT is sufficient for many computational tasks including Shor's algorithm [15, 16, 17, 18], this result is sufficient to show that the QFT is efficiently classically simulatable to high fidelity for a limited class of input states. It was shown in [19] using matrix product states that a terminating QFT can be efficiently simulated for any input state with limited entanglement. Additionally, in [20] a classical algorithm to obtain the results of the QFT on separable states is derived. That this algorithm is simpler than the QFT suggests that the quantum speedup of the QFT lies in the quantum parallelism of its input state rather than its innate complexity. A more general result is presented in [21, 22], where it is shown that a circuit consisting of a polynomial number of QFTs, automorphisms and quadratic phase gates can be efficiently classically simulated over a restricted set of input states on an arbitrary abelian group. These results generalize the Gottesman-Knill theorem and the restricted set of input states is much broader than in earlier work.

In this paper, we use matrix product operators to simulate the quantum Fourier transform. Our numerical results imply that for weakly entangled input states over n qubits and a maximum Schmidt rank of χ, independent of n, the resources required for this simulation scale as $O(n\left(\log(n)\right)^2\chi^2)$, a significant improvement on earlier results. The simulation tools we use are also more straightforward than those found in [13, 14] and have a wider applicability than other results concerning the classical simulatability of the QFT. In section 2 we will review matrix product states and matrix product operators. Section 3 discusses their use to simulate the QFT. In section 4 we will present numerical results showing the errors associated with such simulation to be minimal. Finally we will discuss the implications of this work and how it is related to the computational speedup of Shor's algorithm in section 7.

2 Matrix product states and operators

A matrix product states (MPS) of n qubits is a quantum states with the form [23, 24]:

$$|\psi\rangle = \sum_{i_1,i_2,\ldots,i_n} \sum_{\alpha_1,\alpha_2,\ldots,\alpha_n} \Gamma^{[1]\alpha_1}_{i_1}\lambda^{[1]}_{\alpha_1}\Gamma^{[2]\alpha_1\alpha_2}_{i_2}\ldots \Gamma^{[n-1]\alpha_{n-2}\alpha_{n-1}}_{i_{n-1}}\lambda^{[n-1]}_{\alpha_{n-1}}\Gamma^{[n]\alpha_{n-1}}_{i_n}|i_1\rangle\ldots|i_n\rangle, \quad (1)$$

where $|i_j\rangle$ is the state of the jth qubit in the system. The determination of a coefficient of this state requires the contraction of a series of one and two-dimensional tensors $\{\Gamma^{[j]\alpha_{j-1}\alpha_j}_{i_j}\}$ and one dimensional vectors $\lambda^{[j]}_{\alpha_j}$ where the i_j are set to specify the coefficient required in the computational basis. The α_i are ancillary indices and will henceforth be referred to as bonds between different parts of the system. The connectivity of the tensors reflects a one-dimensional ordering of the qubits in a state.

In the canonical form of a MPS, the decomposition from a state vector to matrix product form is accomplished by a series of singular value decompositions, in which the Γ tensors are unitary and the λ vectors contain the singular values. Each bond joins two tensors and corresponds to a bipartition of the state. The rank of a bond is the Schmidt rank of this bipartition. A state with little entanglement will have low Schmidt ranks and so the tensors used to encode the state in a MPS will be small. As such, it is possible to efficiently simulate quantum states with low entanglement with a MPS [24, 5]. The MPS form is also useful because it allows the truncation of the number of singular values in each bipartition. As the singular values are ordered, it is straightforward to remove the smallest ones and then to normalize the state.

It is possible to generalise the structure of matrix product states in several ways. A simple generalisation is to keep the linear connectivity of the tensor network but to encode an operator rather than a state.

$$\mathbf{O} = \sum_{i_1,\ldots,i_n} \sum_{j_1,\ldots,j_n} \sum_{\alpha_1,\ldots,\alpha_{n-1}} O^{[1]\alpha_1}_{i_1 j_1} \gamma^{[1]}_{\alpha_1} O^{[2]\alpha_1\alpha_2}_{i_2 j_2} \ldots$$
$$O^{[n-1]\alpha_{n-2}\alpha_{n-1}}_{i_{n-1} j_{n-1}} \gamma^{[n-1]}_{\alpha_{n-1}} O^{[n]\alpha_{n-1}}_{i_n j_n} |i_1\rangle \ldots |i_n\rangle\langle j_1| \ldots \langle j_n|.$$

This is called a Matrix Product Operator (MPO) and was introduced in [25, 26]. Similarly to MPSs, the canonical form of a MPO is created with a series of singular value decompositions. The bond ranks of the MPO are then the Schmidt numbers of the operator. The maximal bond dimension of a MPO thus gives the Hartley strength of the operator [27].

The Schmidt rank of a state is a measure of the amount of entanglement in the state. However, the Schmidt number of an operator does not have as straight-forward an interpretation. An operator with a high Schmidt number may have a high amount of classical correlating power but little entangling ability. This is illustrated in the case of two-qubit unitary operations by the SWAP gate, which has the Schmidt-operator decomposition $1/2(I \otimes I + X \otimes X + Y \otimes Y + Z \otimes Z)$. The Schmidt number of four is the maximum possible for a two-qubit operator, but the gate has no entangling power.

Applying a MPO to a MPS produces a new MPS in which each tensor is the product of a tensor in the original MPS and a tensor in the MPO, each representing the same qubit:

$$|\psi'\rangle = \sum_{i_1,i_2,\ldots,i_n} \sum_{\beta_1,\beta_2,\ldots,\beta_n} \Gamma'^{[1]\beta_1}_{i_1} \lambda'^{[1]}_{\beta_1} \Gamma'^{[2]\beta_1\beta_2}_{i_2} \ldots$$
$$\Gamma'^{[n-1]\beta_{n-2}\beta_{n-1}}_{i_{n-1}} \lambda'^{[n]}_{\beta_{n-1}} \Gamma'^{[n]\beta_{n-1}}_{i_n} |i_1\rangle \ldots |i_n\rangle, \tag{2}$$

where we have relabelled the new ancillary indices to include those from both the state and the operator: c $\Gamma'^{[l]\beta_{l-1}\beta_l}_{i_l} = \Gamma'^{[l]\alpha_{l-1}\alpha_l\mu_{l-1}\mu_l}_{i_l} = \sum_{j_l} \Gamma^{[l]\alpha_{l-1}\alpha_l}_{j_l} O^{[l]\mu_{l-1}\mu_l}_{j_l i_l}$, $\lambda'^{[l]}_{\beta_l} = \lambda'^{[l]}_{\alpha_l\mu_l} = \lambda^{[l]}_{\alpha_l}\gamma^{[l]}_{\mu_l}$. This multiplication can be followed by a new singular value decomposition at each bipartition to return the state to canonical form. Before the singular value decomposition, the new state will have a bond rank of $c_j d_j$ at site j where c_j is the bond rank of the MPS and d_j is the bond rank of the MPO. This rank invites an interpretation of the Schmidt number of an operator as an upper bound for the amount by which the Schmidt rank of a state can increase upon application of the operator. If the MPO has bond rank d for that partition, the Schmidt rank will be at most multiplied by d. However, in many cases the Schmidt rank after application will be much lower than this.

3 Simulation of the QFT with a MPO

The quantum Fourier transform can be written in operator form:

$$\frac{1}{\sqrt{N}} \sum_{j,k=0}^{N-1} e^{2\pi i j k/N} |j\rangle\langle k|, \tag{3}$$

where $N = 2^n$.

This equation can be expanded to give output values at individual qubits:

$$|j_1, \ldots, j_n\rangle \rightarrow \frac{1}{2^{n/2}} \left(|0\rangle + e^{2\pi i 0.j_n}|1\rangle\right) \\ \ldots \left(|0\rangle + e^{2\pi i 0.j_1 \ldots j_n}|1\rangle\right), \tag{4}$$

where $j = j_1 2^{n-1} + j_2 2^{n-2} + \ldots + j_n 2^0$ and $0.j_l \ldots j_m = j_l/2 + \ldots + j_m/2^{m-l+1}$. This form of the equation motivates the canonical decomposition of the QFT into quantum gates, which is shown in figure 1.

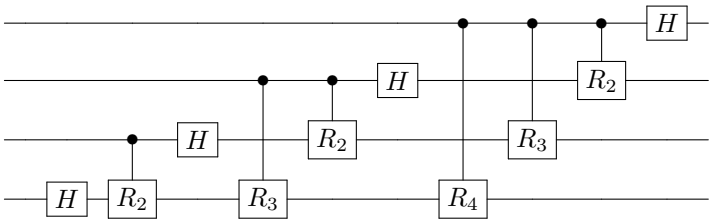

Fig. 1. The canonical decomposition of the quantum Fourier transform with four qubits.

The operator-Schmidt decomposition of the QFT has been calculated exactly [28] and the maximal Schmidt number of a n qubit transform is 2^n. Additionally, all of the singular values are equal. Simulating the QFT using a MPO representation of (4) would thus entail an exponential scaling in terms of execution resources as the number of qubits is increased.

From (4) we note that the output at the first qubit depends only upon the input value at the last qubit, the input at the second depends upon the output at the last two qubits and so on. As such, the operator displays similar classical correlations between the input and output values to those in our earlier swap gate example. These correlations are expensive to encode in a MPO. A simple re-ordering of the input or output qubit values (but not both) from equation (4) produces a more easily encoded operator:

$$|j_1, \ldots, j_n\rangle \rightarrow \frac{1}{2^{n/2}} \left(|0\rangle + e^{2\pi i 0.j_1}|1\rangle\right) \\ \ldots \left(|0\rangle + e^{2\pi i 0.j_n \ldots j_1}|1\rangle\right). \tag{5}$$

In (5) the output at the first qubit depends only upon the input at the first qubit, the output at the second qubit upon the input at the first two qubits and so on. This ordering thus requires less information to be communicated across bonds and can be encoded in a smaller MPO.

To construct a MPO encoding (5) one can construct a MPO of (4) and then apply the SWAP gates leading to the required ordering to only one side of the MPO. This has the

disadvantage that the MPO for (4) must be calculated first, which is computationally intractable for large numbers of qubits. A better approach is to apply a swap gate to the input qubits whenever one is applied to the output qubits. This makes the ordering of the input qubits the same at all times as that of the output qubits. This approach also produces the required ordering and invites an interpretation that the resulting MPO contains only interesting correlations rather than expensive swap correlations.

We constructed MPOs encoding equation (5) with a simple nearest neighbour circuit [29]. The bond ranks of the MPO encoding (5) were much lower than those required to encode (4). Figure 2 shows the size of each element in a bond in the center of a MPO representing (5) for 24 qubits. We display the probability distribution derived from the singular values $p_i = s_i^2/D$ where s_i are singular values and D is the dimension of the Hilbert space (2^{24} in this case).

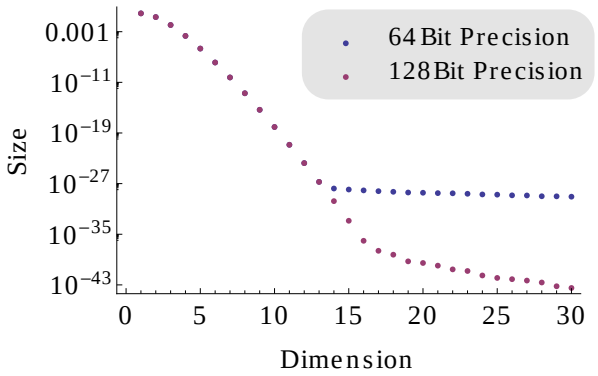

Fig. 2. The probability distribution derived from the singular values of a bipartition at the center of MPOs representing the QFT with 24 qubits at two different precisions.

The sizes of the bond elements displays a characteristic drop-off from lower to higher rank. This characteristic was present regardless of the size of the MPO (MPOs with up to 50 qubits were tested). Initially the rate of decrease is slow, but it quickly becomes exponential with the bond rank included. The exponential decrease of bond element size halted at a probability of around 10^{-40} at quadruple precision (128 bits), which corresponds to a singular value of size relative to the largest value of 10^{-20}. There were many additional bond elements of this size or slightly smaller displaying a large amount of random variation in each MPO. While their size is much larger than machine precision (these results were produced for quadruple precision numbers with $\epsilon \approx 10^{-35}$), the condition number of a singular value decomposition in this problem is very large and so instability at small element sizes is likely to result. Additionally, computing the operator with a lower numerical precision (double precision with 64 bits for example) leads to a curve with the same exponential dropoff initially, but with the dropoff halting at a larger size. It is thus likely that these elements are a result only of numerical imprecision.

The exponential dropoff of probability distribution values shown in figure 2 has the implication that the Schmidt strength of the rearranged QFT converges to a constant value as the number of qubits in the transform is increased. The Schmidt strength is the maximum entropy of the probability distribution following from the singular values of an operator U along any bipartition. It also gives the maximum entropy $E(U|\alpha\rangle|\beta\rangle)$ where $|\alpha\rangle$ and $|\beta\rangle$ are

states in two different quantum systems corresponding to any bipartition of the operator U. The states $|\alpha\rangle$ and $|\beta\rangle$ are maximally entangled with ancillary systems [27]. We compute this strength to be 0.8208. For comparison, the Schmidt strength of a CNOT or CPHASE gate is 1 and the Schmidt strength of a SWAP gate is 2.

The convergence of the bond sizes is shown in figure 3 where we plot the mean difference between the size values obtained with a given number of qubits and those of the largest MPO created (44 qubits). It is clear that the values are converging towards the characteristic visible in figure 2a. Note that for reasons of speed these results were computed at double precision and so the halting of the convergence at 34 qubits represents the calculation reaching machine precision.

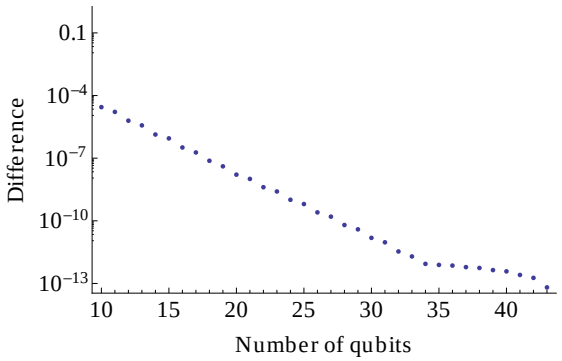

Fig. 3. The mean difference between the probability distribution of the middle bond of the QFT computed with each number of qubits and that computed with the largest number of qubits (44).

After truncation of the smaller bond elements in the MPOs encoding (5), the tensors in the middle of the transform converged to a constant tensor as the number of qubits was increased. This convergence completely specifies tensors in the middle of the transform up to phase rotations which result from the lack of uniqueness of the SVD, which can be easily corrected. The convergence is illustrated in figure 4, which shows the mean difference between the absolute values of the elements of the middle tensor of each MPO and the absolute values of the elements of middle tensor of the largest MPO computed (44 qubits). Again, these results were computed at double precision.

Two different exponential decay rates are visible in the plot. The first of these rates is the region at which we must truncate the middle tensor of the largest MPO to compare it to smaller tensors in smaller MPOs while from 20 qubits onwards, the truncation occurred at a bond size of 30 during calculation of the MPO and so the tensors are the same size.

It is clear from our results that the sizes of each bond in a MPO of the QFT decrease exponentially with increasing bond size. Truncating a bond to size t would thus create an error of $O(e^{-t})$, and $O(n)$ truncations in a n qubit transform would create errors of size $O(ne^{-t})$. To maintain a constant error as the number of qubits is increased, the bond size required would thus be $O(\log(n))$. The convergence towards a common tensor in the middle of the transform implies that it would be possible to find a standard MPO form for the QFT with a relatively small number of qubits determined by a given error tolerance. The middle tensor of this standard QFT could then be replicated a number of times to apply the transformation to any required larger number of qubits. With a n qubit QFT applied to a MPO with maximum

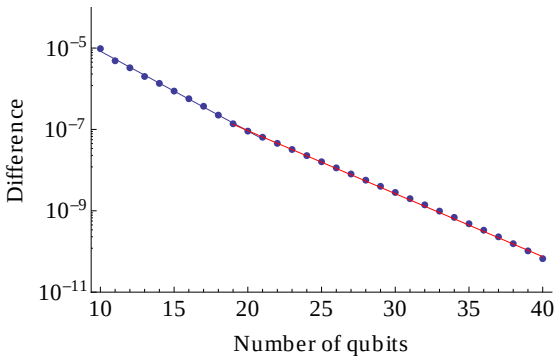

Fig. 4. The mean difference between the size of the values in a tensor in the middle of a MPO and those of a tensor in the middle of a MPO of the largest size (44 qubits). These differences are normalised by the size of the maximum value in the tensor. Two different decay rates are shown.

Schmidt rank χ, this would allow simulation of the QFT in $O(n\left(\log(n)\chi\right)^2)$ time.

4 Truncation Errors

Truncation of bonds of even small size will necessarily introduce error into the representation of an operator. In order to confirm that an efficient MPO simulation of the QFT can be run with the bond size scalings suggested by figure 2 without compromising accuracy, it is necessary to quantify this error. It is difficult to quantify the error in a large MPO because the computational cost of calculating any interesting metric will in general grow exponentially with the number of qubits. This is true of any calculation which does not take advantage of the structure of the MPO. For example, many matrix norms require the calculation of the eigenvalues or decompositions of the full matrix of an operator, or maximisation of a function defined on the full matrix. The matrix representation of the QFT is not sparse, and so the exact calculation of such quantities is intractable.

Instead, we have calculated two less rigorous but more easily computed norms. In order to perform the calculations at large system sizes, these calculations had to be performed with double precision.

Firstly, we computed the Hilbert-Schmidt inner product $\frac{1}{D}\mathrm{tr}(UV^*)$ where U is a MPO representing the QFT and whose bonds are truncated to a given size, V is the same operator but is not truncated and D is the dimension of the Hilbert space. The value obtained measures the inner product between $U|\psi\rangle$ and $V|\psi\rangle$ averaged over all states $|\psi\rangle$. The error in the result $1-\frac{1}{D}\mathrm{tr}(UV^*)$ is shown in figure 5. It can be seen from this calculation that the error drops off exponentially as the bond rank is increased, and increases sub-exponentially as the number of qubits is increased. However, this regime only extends as far as a maximum bond rank of 8, after which the observed error was zero. At these ranks, the average error is thus below the machine precision of around 10^{-16}. It is worth noting that this is only an average measure of error and so does not reflect the worse case error involved in applying a truncated MPO.

The second measurement of error we computed is the amount of error associated with Fourier transforming a periodic state. A periodic state with L qubits and period r takes the form $\sum_{n=0}^{\left(2^L/r\right)-1}|k_0+nr\rangle$. These states are produced by the modular exponentiation stage of Shor's algorithm. Applying the QFT to a periodic state produces a state which is strongly

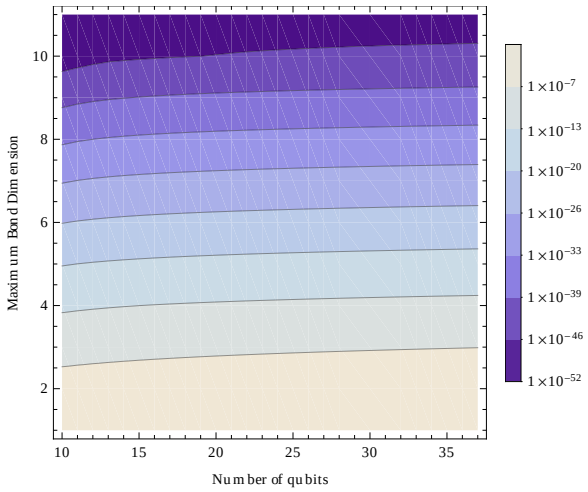

Fig. 5. The error in the trace inner product between two MPOs of the QFT, one with truncation and one without any truncation.

peaked around the values $\left|(i/r)\,2^L\right\rangle$ for $i < r$ and so measuring the Fourier Transform of a periodic state reveals the period. This is the basis of Shor's algorithm.

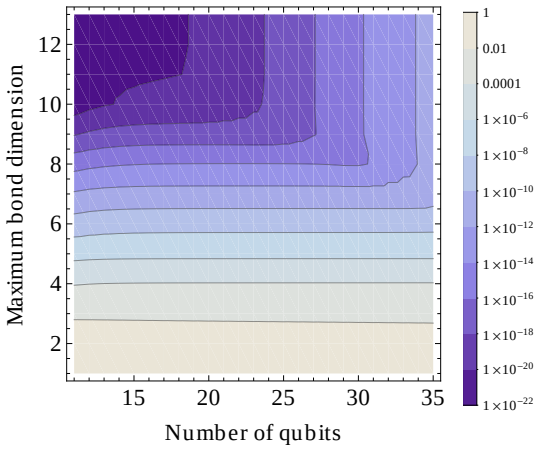

Fig. 6. The difference between the probability of measuring a peak after simulating a QFT with a MPO on a periodic state and the analytic probability. Shown for a period of 9.

We prepared periodic states with a range of periods and numbers of qubits between 10 and 28 and computed the deviation of the sizes of the peaks from the analytic values. These results are shown in figure 6 for period 9. The results were similar for other periods for other values tested ($2 \leq r \leq 15$). As with the trace inner product, the error seems to decrease exponentially as the bond rank is increased at low bond ranks. At higher bond ranks, the error appears to increase quickly as the number of qubits is increased. It is difficult to obtain data with larger numbers of qubits due to the exponential scaling of the calculation of the analytic size of the peaks.

The results in figure 2 indicate that many extra bond elements appear at double precision with sizes relative to the largest element of 10^{-13} or less. We would expect that errors observed after simulating the QFT of periodic inputs would be at less than or equal to these levels. This is the case for the range of qubits tested.

5 Fragility of the scaling

It is difficult to apply precise phase shifts in a physical quantum computer. As such, we consider the effect of small errors in the controlled phase gates. Doing so also allows us to determine how robust the scaling of the size of the MPO with increasing numbers of qubits is to a particular kind of noise in the quantum circuit. We prepared MPOs of the QFT, but when applying a controlled phase of θ between qubits a and b, we instead applied a controlled phase of $\theta+\delta r$, where δ is the size of the random error and r is a uniformly distributed random number between 0 and 1. While operational noise in a physical quantum computer is generally modelled by a normally distributed random phase, our choice of a uniform distribution allows us to relate these results with those reported in section 3.

We compared the MPOs with errors to those without by fixing the number of qubits and subtracting the singular value vector in the middle bond with errors from the singular value vector of the MPO without errors. Each bond vector with errors deviated from the error-free vector by a roughly constant amount along the length of the vector. Because the singular values of the QFT exponentially decay, this leads to characterstics similar to those shown in figure 2. We show a set of probability distributions derived from four MPOs with different δ values in figure 7.

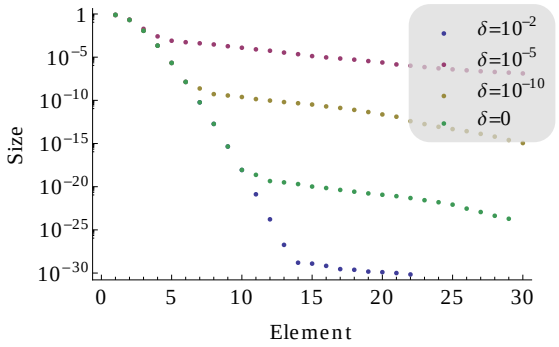

Fig. 7. The probability distribution derived from the singular values of the middle bonds of MPOs representing the QFT with 15 qubits computed with random phases errors of size δ. These MPOs were produced at double precision.

It is no coincidence that the curves in figures 2 and 7 look similar. Because controlled phase rotations constitute most of the gates in the QFT (not including SWAP, which is a permutation of a two qubit tensor), introducing a small random error of order δ to the angle of the phase rotation has a similar effect to limiting the precision of the calculation of the MPO to δ. As such, by varying the value of δ we are doing a similar operation to computing the MPO of the QFT at arbitrary precision.

We show the mean error in the middle bond vector that resulted from varying δ in figure 8. These errors were computed by forming 100 MPOs for each value of δ and averaging over the

deviations from the error free MPO. Also shown is a linear fit between the logarithm of the error and the logarithm of δ. This is a log-log plot with a clear linear relationship, and so the error varies polynomially with δ and the gradient of the fit gives the power of this dependence. In this case the power was found to be 1.031 ± 0.005, an almost linear relationship between the random errors introduced in the controlled phases and the resulting errors in the singular values.

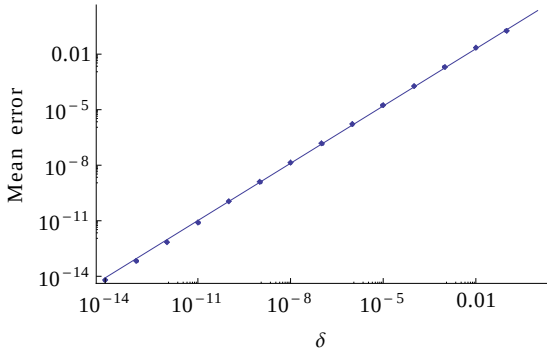

Fig. 8. The mean differences between the singular value vectors of the middle bonds of the QFT with 15 qubits and random phase errors of size δ, and of the QFT with no phase errors. These values are normalised by the largest singular value in the bond.

The middle bond vectors found in each of the random MPOs generated for each value of δ displayed a similar dependence on element number to those displayed in figure 7. Each small additional singular value in the vectors generated with errors will create a random error in the final coefficients of the QFT operator. As the deviations from the more accurate set of singular values result from a random phase, we do not expect that the random errors created by each singular value will add to produce a much larger error. Despite there being a large number of additional singular values created by the noise in each case, we thus expect that if the size of the mean error in the singular value vector is ϵ, the error in the coefficients will be $O(\epsilon)$. A phase imprecision of δ will then bring about an overall error in coefficients of the resulting QFT operator of roughly the same size. Conversely, if the coefficients of the QFT are required to a precision of ϵ, the MPO only needs to be computed with numbers of approximately the same precision.

6 Reasons for the efficient representation

The fact that ordering the input values of the qubits differently to the output values can lead to a dramatic reduction in MPO complexity raises the question of whether a different ordering to that considered in (5) may be optimal. We tested this by constructing MPOs with all possible input qubit orderings for QFTs with up to twelve qubits. In every case the ordering in (5) was optimal. For the reasons described above (the output at the nth qubit depends only upon the input at the first n qubits), we expect this to be the case for larger numbers of qubits as well.

It would seem natural to explain the exponential decrease of bond element size shown in figure 2 with the small effect of the rotation gates correlating far away qubits. That is, the full QFT introduces correlations across every qubit pairing. However, these correlations take the

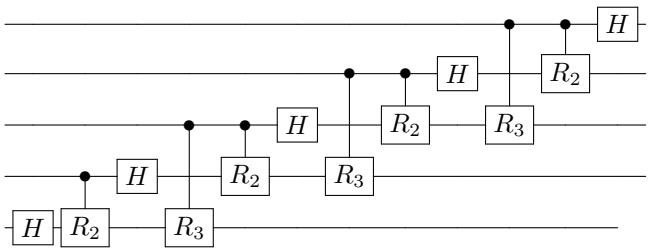

Fig. 9. The AQFT for five qubits with a maximum of three controlled phase gates.

form of controlled phase rotations and the size of the rotations decreases inverse-exponentially with the one-dimensional distance between qubits. As such, we should be able to neglect long range correlations. We would expect this to cause the sizes of the tensors at the qubits within the MPO to be almost entirely unaffected by the number of far-away qubits. This is the idea behind the AQFT [15], where the number of controlled phase gates conditioned upon each qubit, henceforth the bandwidth, is set at a fixed value irrespective of the number of qubits in the transform.

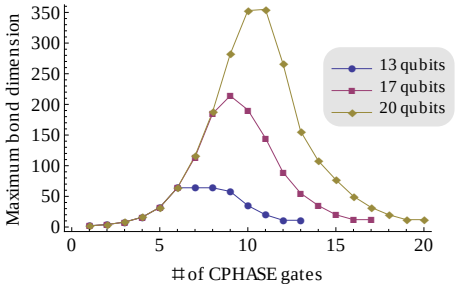

Fig. 10. The maximum bond rank in MPOs corresponding to AQFTs with different numbers of qubits. Each AQFT was constructed with a different maximum number of controlled phase gates conditioned on each qubit.

However, we constructed MPOs using a nearest neighbour quantum circuit of the AQFT and found that the bond ranks produced were larger than those produced for the full transform. The maximum bond ranks for a series of AQFTs after truncation are shown in figure 10. Maximum bond ranks in an AQFT increased by a factor of 2 per additional controlled phase rotation included. This increase levelled off in the middle of the transform but still quickly became computationally intractable. Furthermore, the trace inner product between an operator truncated at any bond rank and a series of operators with reduced bandwidths was a maximum for the full transform and decreased monotonically as the bandwidth decreased. It thus seems that the low required bond ranks observed in (5) are a feature of the full QFT.

While the low required bond rank of the QFT cannot be attributed entirely to decreasing phase rotations, the size of these rotations and the rate of their decrease are important. We found that transforms with the same structure as the QFT but with phase rotations decreasing as $\exp(2\pi i/k^n)$ instead of $\exp\left(2\pi i/2^k\right)$, where k is the qubit distance and n an integer, did not display the characteristic dropoff of bond size. Rather, the required bond ranks appeared to increase with increasing numbers of qubits, presumably until the phase

rotations become smaller than machine precision. Using a rotation with some randomness in the form of $\exp\left(2\pi i/2^{k+\delta}\right)$ or $\exp\left(2\pi i/(2+\delta)^k\right)$, with δ a small random number, also removed the exponential dropoff. Rotations of the form $\exp\left(2\pi i/n^k\right)$ for $n \geq 2$ still lead to Fourier transforms, although not over Z_{2^m}, and were found to still lead to an exponential dropoff in bond element size. The rate of this dropoff increased as n increased.

As such, while the small bond rank required to accurately represent the QFT with a MPO is not due solely to the decreasing size of the phase rotations used, it is related to them. It is likely the exponential dropoff of bond size is the result of a symmetry in the structure of the QFT. In order to obtain a low bond rank it is necessary to have phase rotations which decrease at least exponentially with the distance between qubits and to have the same phase rotation for each conditioned gate at a given linear qubit distance.

7 Discussion

While we have not proven that the QFT can be efficiently represented as a MPO, our numerical results are strongly suggestive of this. If appears that the numerical error associated with the very small amount of truncation required for a tractable representation is very close to zero over a range of numbers of qubits. Additionally, the differences between a matrix in the middle of each operator and an adjacent matrix decreases as the system size increases.

Together, these results suggest that a MPO representing a QFT for an arbitrary number of qubits can be created from the MPO representation of a QFT of a smaller size. With an appropriate bond rank, this would allow the QFT to be performed on a MPO with maximum Schmidt rank χ with computational cost $O(n\left(\log(n)\right)^2\chi^2)$. It could similarly be performed on weakly correlated mixed states. Our method allows the QFT to be efficiently simulated in a straightforward fashion in any case in which the qubits are ordered linearly.

Application of the QFT to a MPS of n qubits with this method increases the bond rank by at most a factor scaling as $O(\log(n))$. Denoting this factor by d, the application of m QFTs increases the bond ranks by a maximum factor of d^m. As such, the application of a constant number of QFTs can be efficiently simulated with a large number of qubits. These QFTs can be interspersed by quantum circuits that do not increase the Schmidt rank.

Our results strengthen earlier work. In [14] the AQFT is show to be classically simulatable in polynomial time, although an explicit scaling is not derived. Our method of simulation uses the full QFT and has a more advantageous scaling with respect to the number of qubits of $O(n\left(\log(n)\right)^2)$.

In [14] a condition is also derived for when two efficiently simulatable quantum circuits composed may be efficiently simulated. From this condition it follows that any circuit composed of a constant number of AQFTs and log-depth limited interaction range circuits can be efficiently classically simulated. We provide a different perspective on the composability criteria. That is, our method makes explicit the scaling of the cost of the QFT with the Schmidt rank of the bipartitions in the input state. We have shown the difficulty of simulating the QFT to be mostly determined by the complexity of the state being transformed. A log-depth limited interaction circuit will produce an input state with small Schmidt ranks across each bond partition, and so the previous result follows from our results.

That the QFT can be represented to very high fidelity with a MPO with limited bond ranks implies that the QFT can produce only a limited amount of entanglement. This conclusion

was originally shown in [30], however our methods are more straightforward.

We take a different approach to simulating the QFT to that presented in [21, 22], which uses an algebraic extension of stabilizer techniques to efficiently simulate a range of circuits including a polynomial number of QFTs. These results also allow the QFT to be simulated on highly entangled states and on different abelian groups. Our results take a tensor network approach and only allow a constant number of QFTs to be simulated efficiently on states with entanglement that grows polynomially with the number of qubits. Additionally, our results allow a classification of the difficulty of simulating the QFT on arbitrary input states to be calculated.

With respect to the question of where the quantum speedup in Shor's algorithm originates, our results provide further evidence that it originates in the highly entangled state generated by modular exponentiation. Periodic states are generated by modular exponentiation, and a state of period r will have bond ranks in a MPS of r. As the maximum period of a modular exponentiation process factoring a number N scales as $O(N)$ [18], states with very high Schmidt numbers are generated. These states are very difficult to represent in a MPS and thus are very difficult to Fourier transform. This conclusion is similar to that reached in other works such as [14, 20]. It is additionally worth noting that while our method makes very clear the connection between the Schmidt rank of the input state and the difficulty in Fourier transforming it, the same conclusion can be drawn about the computational speedup from the results of [14].

Acknowledgments

This research was conducted by the Australian Research Council Centre of Excellence for Quantum Computation and Communication Technology (project number CE110001027).

References

1. D. Gottesman, The Heisenberg representation of quantum computers, arXiv:quant-ph/9807006.
2. M. Van den Nest, Classical simulation of quantum computation, the Gottesman-Knill theorem, and slightly beyond, Quantum Information & Computation 10 (3–4) (2010) 0258–0271.
3. L. G. Valiant, Quantum computers that can be simulated classically in polynomial time, in: Proceedings of the Thirty-third Annual ACM Symposium on Theory of Computing, STOC '01, ACM, New York, NY, USA, 2001, pp. 114–123. doi:10.1145/380752.380785.
   URL http://doi.acm.org/10.1145/380752.380785
4. R. Jozsa, A. Miyake, Matchgates and classical simulation of quantum circuits, Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences 464 (2100) (2008) 3089–3106. doi:10.1098/rspa.2008.0189.
5. G. Vidal, Efficient Simulation of One-Dimensional Quantum Many-Body Systems, Physical Review Letters 93 (4) (2004) 040502. doi:10.1103/PhysRevLett.93.040502.
6. R. Jozsa, N. Linden, On the role of entanglement in quantum-computational speed-up, Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences 459 (2036) (2003) 2011–2032. doi:10.1098/rspa.2002.1097.
7. I. L. Markov, Y. Shi, Simulating quantum computation by contracting tensor networks, SIAM J. Comput. 38 (3) (2008) 963–981. doi:10.1137/050644756.
   URL http://dx.doi.org/10.1137/050644756
8. N. Yoran, A. Short, Classical simulation of limited-width cluster-state quantum computation, Physical Review Letters 96 (2006) 170503. doi:10.1103/PhysRevLett.96.170503.
9. R. Jozsa, On the simulation of quantum circuits, arXiv:quant-ph/0603163arXiv:quant-ph/

0603163.
10. M. Schwarz, M. Van den Nest, Simulating Quantum Circuits with Sparse Output Distributions, arXiv:1310.6749arXiv:arXiv:1310.6749v1.
11. S. Lloyd, Universal quantum simulators, Science 273 (5278) (1996) 1073–1078.
12. P. W. Shor, Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer, SIAM Journal on Computing 26 (1997) 1484–1509.
13. D. Aharonov, Z. Landau, J. Makowsky, The quantum FFT can be classically simulated, quant-ph/0611156arXiv:0611156v2.
14. N. Yoran, A. J. Short, Efficient classical simulation of the approximate quantum fourier transform, Physical Review A 76 (2007) 042321. doi:10.1103/PhysRevA.76.042321.
URL http://link.aps.org/doi/10.1103/PhysRevA.76.042321
15. D. Coppersmith, An approximate Fourier transform useful in quantum factoring, eprint arXiv:quant-ph/0201067arXiv:0201067v1.
16. A. Barenco, A. Ekert, K.-A. Suominen, P. Törmä, Approximate quantum fourier transform and decoherence, Physical Review A 54 (1996) 139–146. doi:10.1103/PhysRevA.54.139.
URL http://link.aps.org/doi/10.1103/PhysRevA.54.139
17. A. Fowler, L. Hollenberg, Scalability of Shors algorithm with a limited set of rotation gates, Physical Review A 70 (3) (2004) 032329. doi:10.1103/PhysRevA.70.032329.
18. Y. S. Nam, R. Blümel, Scaling laws for Shor's algorithm with a banded quantum Fourier transform, Physical Review A 87 (3) (2013) 032333. doi:10.1103/PhysRevA.87.032333.
19. D. E. Browne, Efficient classical simulation of the quantum Fourier transform, New Journal of Physics 9 (5) (2007) 146–146. doi:10.1088/1367-2630/9/5/146.
20. A. A. Abbott, De-quantisation of the quantum fourier transform, Applied Mathematics and Computation 219 (1) (2012) 3 – 13. doi:http://dx.doi.org/10.1016/j.amc.2011.06.057.
URL http://www.sciencedirect.com/science/article/pii/S0096300311009027
21. M. Van Den Nest, Efficient classical simulations of quantum fourier transforms and normalizer circuits over abelian groups, Quantum Info. Comput. 13 (11-12) (2013) 1007–1037.
22. J. Bermejo-Vega, M. V. den Nest, Classical simulations of Abelian-group normalizer circuits with intermediate measurements, Quantum Information & Computation 14 (3) (2014) 181–216.
23. M. Fannes, B. Nachtergaele, R. Werner, Finitely correlated states on quantum spin chains, Communications in Mathematical Physics 144 (3) (1992) 443–490.
24. G. Vidal, Efficient Classical Simulation of Slightly Entangled Quantum Computations, Physical Review Letters 91 (14) (2003) 147902. doi:10.1103/PhysRevLett.91.147902.
25. F. Verstraete, J. Garcia-Ripoll, J. Cirac, Matrix product density operators: simulation of finite-temperature and dissipative systems, Physical Review Letters 93 (2004) 207204.
26. M. Zwolak, G. Vidal, Mixed-State Dynamics in One-Dimensional Quantum Lattice Systems: A Time-Dependent Superoperator Renormalization Algorithm, Physical Review Letters 93 (2004) 207205. doi:10.1103/PhysRevLett.93.207205.
27. M. Nielsen, C. Dawson, J. Dodd, A. Gilchrist, D. Mortimer, T. Osborne, M. Bremner, A. Harrow, A. Hines, Quantum dynamics as a physical resource, Physical Review A 67 (5) (2003) 052301. doi:10.1103/PhysRevA.67.052301.
28. J. Tyson, Operator-Schmidt decomposition of the quantum Fourier transform on Bbb CN1 Bbb CN2, Journal of Physics A: Mathematical and General 36 (24) (2003) 6813.
URL http://stacks.iop.org/0305-4470/36/i=24/a=317
29. A. Fowler, S. Devitt, L. Hollenberg, Implementation of Shor's algorithm on a linear nearest neighbour qubit array, Quantum Information & Computation 4 (4) (2004) 237–251.
30. N. Yoran, Efficiently contractable quantum circuits cannot produce much entanglement, arXiv:0802.1156arXiv:arXiv:0802.1156v1.