

# Efficient quantum state tomography

Marcus Cramer and Martin B. Plenio

*Institut für Theoretische Physik, Albert-Einstein Allee 11, Universität Ulm, 89069 Ulm, Germany*

Steven T. Flammia and Rolando Somma

*Perimeter Institute for Theoretical Physics, Waterloo, Ontario, N2L 2Y5 Canada*

David Gross

*Institute for Theoretical Physics, Leibniz University Hannover, 30167 Hannover, Germany*

Stephen D. Bartlett

*School of Physics, The University of Sydney, Sydney, New South Wales 2006, Australia*

Olivier Landon-Cardinal and David Poulin

*Département de Physique, Université de Sherbrooke, Sherbrooke, Québec, Canada*

Yi-Kai Liu

*Institute for Quantum Information, California Institute of Technology, Pasadena, CA, USA*

Quantum state tomography [1], the ability to deduce the state of a quantum system from measured data, is the gold standard for verification and benchmarking of quantum devices. It has been realized in systems with few components [2–7], but for larger systems it becomes infeasible because the number of quantum measurements and the amount of computation required to process them grows exponentially in the system size. Here we show that we can do exponentially better than direct state tomography for a wide range of quantum states, in particular those that are well approximated by a matrix product state ansatz. We present two schemes for tomography in 1-D quantum systems and touch on generalizations. One scheme requires unitary operations on a constant number of subsystems, while the other requires only local measurements together with more elaborate post-processing. Both schemes rely only on a linear number of experimental operations and classical postprocessing that is polynomial in the system size. A further strength of the methods is that the accuracy of the reconstructed states can be rigorously certified without any *a priori* assumptions.

## I. INTRODUCTION

One of the principal features distinguishing classical from quantum many-body systems is that quantum systems require exponentially many parameters in the system size to fully specify the state, compared to only linearly many for classical systems. Put to use constructively, the exponential complexity enables the construction of information processing devices fundamentally superior to any classical device. At the same time, however, this “curse of dimensionality” makes engineering tasks—such as verifying that the quantum processing device functions as intended—a daunting challenge.

The full determination of the quantum state of a system, known as quantum state tomography [1], has already been achieved by measuring a complete set of observables whose expectation values determine the quantum state [2–7]. As it is typically formulated [8], simply to output an estimate for a generic state would take exponential time in the system size  $N$ , given that there are an exponential number of coefficients in a generic state’s description. This is but one of several inefficiencies. Most quantum states have exponentially small amplitudes in almost every basis, so to distinguish any one of those amplitudes from zero, one must take an exponential number

of samples. Assuming one were able to collect all the data from an informationally complete measurement, one is left with the intractable computational task of inverting the measured frequencies to find an estimate of the state.

However, the traditional representation of quantum states is in a sense too general. Indeed, states which occur in many practical situations are specified by a small number of parameters. An efficient description could be a practical preparation scheme which outputs the state; or, in the case of thermodynamical equilibrium states, a local Hamiltonian and a temperature. This insight is not new: researchers in many-body physics and quantum information theory have found many classes of states which are described by a number of parameters scaling polynomially in  $N$  [9–13] and which closely approximate the kind of states found in physical systems [14, 15]. However, the question of whether these restricted classes can be put to use in the context of tomography has remained largely open.

In this work, we address the above challenge. The physical system we have in mind is one where the constituents are arranged in a one-dimensional configuration (e.g., ions in a linear trap [3]). But the methods that we are presenting here can be generalized to higher-dimensional arrangements such as those realized in optical lattices [16]. It is highly plausible that in such a set-

ting, correlations between neighboring particles are much more pronounced (due to direct interaction) than correlations between distant systems (mediated e.g. by global fluctuations of control fields). An efficiently describable class of states anticipating exactly this behavior has long been studied under the names of finitely correlated states (FCS) or matrix product states (MPS) [9, 10]. Importantly, restricting to this class is not a limitation since *every* state may be written as a MPS with a suitable, albeit possibly large, matrix dimension. Since many states that are relevant for quantum information processing or quantum many-body physics have a small (independent of  $N$ ) bond dimension, our methods are directly applicable to such states; examples include, but are not limited to, ground and thermal states, the GHZ, W, cluster, and AKLT states, the latter two being universal resources states for quantum computing.

Given that standard tomography is no longer feasible in a range of recent and upcoming experiments involving large numbers of qubits, our results represent a significant advance in the ability to verify and quantitatively and efficiently benchmark systems of experimental importance.

## II. RESULTS

In the following we present two schemes for identifying systems that are well approximated by an MPS, initially focusing on pure states for simplicity. We will view each system as consisting of a linear chain of  $N$  qudits, each having dimension  $d$ . Both schemes require the measurement of linearly (in the system size  $N$ ) many local observables within finite accuracy, polynomial classical post-processing of the data and can certify the accuracy of the reconstructed state without making any technical assumptions about the state in the laboratory. The first scheme requires unitary control and local measurements while the second scheme removes the need for unitary control at the cost of more elaborate post-processing.

### A. Scheme based on unitary transformations

The key idea of the method consists of a sequential procedure to *disentangle* the left hand side of the chain from the right hand side, using a sequence of unitary operations with small interaction length independent of  $N$ . The result will be a product state and a sequence of local unitary operations from which to construct the original state.

Suppose the ideal state in the laboratory is  $\hat{\rho} = |\phi\rangle\langle\phi|$ , which we assume for clarity is a MPS of given bond dimension. This implies that the rank of reductions to one part of a bipartite (left vs. right) split of the chain is bounded by a constant  $R$ . The protocol starts by estimating, through standard state tomography, the reduced density matrix of the first  $\kappa = \lceil \log_d(R) \rceil + 1$  sites,

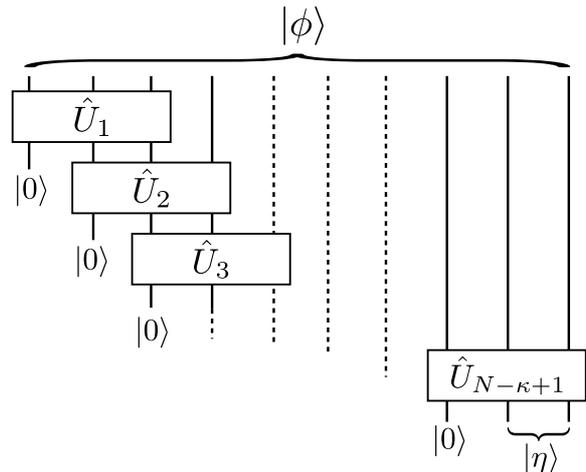


FIG. 1. Quantum circuit that transforms  $|\phi\rangle$  into  $|0\rangle^{\otimes N-\kappa+1} \otimes |\eta\rangle$  with  $\kappa = 3$ . The unitaries  $\hat{U}_i$  successively disentangle the particles and the state  $|\eta\rangle$  on the last sites acts as a boundary condition to determine the MPS description of  $|\phi\rangle$ .

$\hat{\sigma} \approx \text{tr}_{\kappa+1, \dots, N}[\hat{\rho}]$ . This reduced density matrix has the eigendecomposition  $\hat{\sigma} = \sum_{r=1}^R \sigma_r |\phi_r\rangle\langle\phi_r|$  where the rank  $R \leq d^{\kappa-1}$ . Hence there exists a density matrix with one fewer qudit that has the same rank  $R$  and eigenvalues  $\sigma_r$  as  $\hat{\sigma}$ . Therefore we can disentangle the first site in  $\hat{\sigma}$  with the following unitary acting on the first  $\kappa$  sites:

$$\hat{U} = \sum_{s=0}^{d-1} \sum_{s'=0}^{d^{\kappa-1}-1} |s\rangle_1 \otimes |s'\rangle_{2, \dots, \kappa} \langle\phi_{s d^{\kappa-1} + s' + 1}|_{1, \dots, \kappa}, \quad (1)$$

where  $|\phi_1\rangle, \dots, |\phi_R\rangle$  have been extended in some arbitrary way to get a complete basis for sites  $1, \dots, \kappa$ . Applying  $\hat{U}$  produces the state  $\hat{U}|\phi\rangle = |0\rangle_1 \otimes |v\rangle_{2, \dots, N}$ , where  $|v\rangle$  is some pure state on sites  $2, \dots, N$ . Hence  $\hat{U}$  disentangles the first qudit from all the others. Now, set aside this first qudit, look at sites  $2, \dots, \kappa + 1$ , and repeat the above process as indicated on Fig. 1. In this way, one obtains a sequence of unitaries  $\hat{U}_1, \dots, \hat{U}_{N-\kappa+1}$ , where each  $\hat{U}_i$  acts on sites  $i, \dots, i + \kappa - 1$ . This sequence transforms  $|\phi\rangle$  into  $\hat{U}_{N-\kappa+1} \dots \hat{U}_1 |\phi\rangle = |0\rangle^{\otimes N-\kappa+1} \otimes |\eta\rangle$ , where  $|\eta\rangle$  is some pure state on the last  $\kappa - 1$  sites.

In summary, this scheme infers the quantum circuit used to prepare an MPS [17]. The MPS decomposition of  $|\phi\rangle$  can then be obtained readily from the  $U_i$  and  $|\eta\rangle$  [18]. If the state in the laboratory  $\hat{\rho}$  is arbitrary, then the reduced density matrices  $\hat{\sigma}$  will generally have full rank. Hence in each step we will need to truncate the  $\kappa$  qudit state  $\hat{\sigma}$  to a rank  $R$  subspace with  $R < d^{\kappa-1}$ . Then the above method will produce an MPS approximation to  $\hat{\rho}$ . The accuracy of this estimate can be certified, without any assumptions on the state, by keeping track of the effects of truncating each of the reduced states  $\hat{\sigma}$ . As shown in the Methods, errors of magnitude  $\epsilon$  due to finite measurement precision or truncation error (as measured

by the weight of the truncated space) accumulate at most linearly with the number of sites, and can be evaluated directly from the data, leading to an estimate of accuracy  $N\epsilon$ .

The present scheme requires unitary control of  $\kappa$  neighboring qudits, which is challenging to implement in many current experimental settings. We now present a second scheme that avoids unitary control and requires only local measurements.

### B. Scheme based on local measurements: Certification of estimated state

Consider again the state  $\hat{\rho}$  in the laboratory and suppose that a tomographically-complete set of local measurements on all groups of  $k$  neighbouring qubits has been performed. Further suppose that sufficient data has been taken to yield estimates  $\hat{\sigma}_i$  of the local reductions  $\hat{\rho}_i = \text{tr}_{1,\dots,i;i+k+1,\dots,N}[\hat{\rho}]$  such that

$$\|\hat{\rho}_i - \hat{\sigma}_i\|_{\text{tr}} \leq \epsilon_i. \quad (2)$$

Determining these approximate reduced density operators  $\hat{\sigma}_i$  completes the experimental work. In the remainder we describe the classical post-processing that will result in a MPS estimate  $|\psi\rangle$  to  $\hat{\rho}$  and a lower bound to the fidelity  $\langle\psi|\hat{\rho}|\psi\rangle$  [19] that does *not* require assumptions on the nature of the state  $\hat{\rho}$ .

We start with the latter, since the following easy calculation yields the fidelity bound we have in mind and hints at the MPS estimate we are after. Suppose, for the sake of the argument, that  $|\psi\rangle$  is the unique ground state (with energy zero) of a local Hamiltonian  $\hat{H} = \sum_i \hat{h}_i$ , where the  $\hat{h}_i$  is a projection operator acting only on sites  $i+1, \dots, i+k$  (as it turns out, generic MPS are of this type). Then, expanding in the eigenbasis  $\hat{H} = \sum_{n=0}^{2^N-1} E_n |E_n\rangle\langle E_n|$ ,

$$\text{tr}[\hat{H}\hat{\rho}] \geq \Delta E \sum_{n>0} \langle E_n | \hat{\rho} | E_n \rangle = \Delta E (1 - \langle\psi|\hat{\rho}|\psi\rangle), \quad (3)$$

where we denoted by  $\Delta E$  the energy gap above the ground state  $|\psi\rangle$ . Hence, we have the fidelity bound

$$\langle\psi|\hat{\rho}|\psi\rangle \geq 1 - \frac{\sum_i \text{tr}[\hat{h}_i \hat{\rho}_i]}{\Delta E} \geq 1 - \frac{1}{\Delta E} \left( \sum_i \text{tr}[\hat{h}_i \hat{\sigma}_i] + \epsilon_i \right). \quad (4)$$

In other words, the Hamiltonian acts as a *witness* for its ground state  $|\psi\rangle$ . This bound is tight: suppose the experimental estimates  $\hat{\sigma}_i$  and the reductions of the state in the laboratory  $\hat{\rho}_i$  coincide, that is  $\epsilon_i = 0$ . If, in addition, the reductions of  $|\psi\rangle$  match the  $\hat{\sigma}_i$  then, as  $|\psi\rangle$  was assumed to be the unique ground state with energy zero of  $\sum_i \hat{h}_i$ , we have  $\sum_i \text{tr}[\hat{h}_i \hat{\sigma}_i] = 0$ , i.e.,  $\langle\psi|\hat{\rho}|\psi\rangle = 1$ .

The goal is now clear: find a local gapped Hamiltonian  $\hat{H}$  such that the reductions of its ground state are

close to the  $\hat{\sigma}_i$ . *A priori* it is unclear whether such a convenient witness always exists and how it could be found efficiently. However, using formal methods, one can show that if the true state  $\hat{\rho}$  is close to a generic MPS, then such a witness Hamiltonian always exists [10]. What is more, it can be constructed from the estimate of the algorithm sketched below. Its properties, chief among them the gap, are efficiently computable. In the Methods section, we detail the efficient computation of these quantities, and we also consider how to treat states such as the GHZ for which local marginals alone are not quite sufficient for complete characterization (they violate our “generic” condition).

### C. An illustrative example

We illustrate how our certification procedure operates if our estimate for the state  $\hat{\rho}$  is a linear cluster state [20]. The cluster state is defined as the unique eigenstate (with eigenvalue +1) of stabilizers  $\hat{K}_i = \hat{Z}_{i-1} \hat{X}_i \hat{Z}_{i+1}$  for all  $i = 2, \dots, N-1$  (together with boundary terms  $\hat{X}_1 \hat{Z}_2$  and  $\hat{Z}_{N-1} \hat{X}_N$ , which we do not treat separately for simplicity). Assume that we have performed standard quantum state tomography on sets of three neighbouring spins,  $k=3$ , which is the smallest useful set because in a cluster state the rank of the reduced density matrices of contiguous blocks is upper bounded by  $R=4$ . Let us now assume that on the basis of these data, our procedure suggests that the linear cluster state is indeed our estimate. The local Hamiltonian in this case is given by  $\hat{H} = \sum_i (\mathbb{1} - \hat{K}_i)/2$ , where the  $(\mathbb{1} - \hat{K}_i)/2 = \hat{h}_i$  are projectors,  $\hat{H}$  has the cluster state as its unique ground state (with energy zero) and an energy gap  $\Delta E = 1$ . The fidelity between the cluster state  $|\psi_{CS}\rangle$  and the state  $\hat{\rho}$  is bounded by  $\langle\psi_{CS}|\hat{\rho}|\psi_{CS}\rangle = 1 - \sum_i (\text{tr}[\hat{h}_i \hat{\sigma}_i] + \epsilon_i)$ , where  $\epsilon_i$  from Eq. (3) quantifies the statistical and experimental error in the local experimental estimates, and  $\text{tr}[\hat{h}_i \hat{\sigma}_i]$  quantifies how much the laboratory state  $\hat{\rho}$  deviates from an exact cluster state.

Tomography on a cluster state can also be performed with the scheme based on unitary transformations. A cluster state is the output of a quantum circuit where each qubit is initially prepared in the state  $|+\rangle = 1/\sqrt{2}(|0\rangle + |1\rangle)$  and a controlled phase transformation  $CZ$  acts successfully on each pair of neighboring qubits. The  $CZ$  gate changes the sign of state  $|11\rangle$  and acts trivially on the other states of the computational basis. Thus, a cluster state is the output of a circuit whose structure corresponds to the one indicated on Fig. 1 with  $\kappa=2$  and the scheme based on unitary transformations directly applies. Note that the unitary scheme takes advantage of the decreased rank of a reduced density matrix on the *boundary* to save one qubit worth of local tomographic effort ( $\kappa=2$  vs.  $k=3$ ).

#### D. Scheme based on local measurements: Efficient determination of an MPS estimate

With the experimentally obtained  $\hat{\sigma}_i$ , we now turn to the task of finding an MPS  $|\psi\rangle$  such that its reductions  $\text{tr}_{1,\dots,i;i+k+1,\dots,N}[|\psi\rangle\langle\psi|]$  closely match the  $\hat{\sigma}_i$ . In other words: Let  $\hat{P}_m^{i,k}$  be all possible products of Pauli operators (enumerated by  $m$ ) that act non-trivially only on sites  $i+1, \dots, i+k$ . Then, as the  $\hat{P}_m^i$  are an orthogonal basis, the  $\hat{\sigma}_i$  may be expanded as

$$\hat{\sigma}_i = \frac{1}{2^N} \sum_m \text{tr}[\hat{\sigma}_i \hat{P}_m^i] \hat{P}_m^i, \quad (5)$$

where the expectations  $\text{tr}[\hat{\sigma}_i \hat{P}_m^i] \in \mathbb{R}$  are obtained as results of tomographic measurements. Then we need to find a matrix  $|\psi\rangle\langle\psi|$  such that for all  $m$  and  $i$  the expectations  $\text{tr}[|\psi\rangle\langle\psi| \hat{P}_m^i]$  coincide with those of the tomographic estimates,  $\text{tr}[\hat{\sigma}_i \hat{P}_m^i] = \text{tr}[\hat{\sigma}_i \hat{P}_m^i]$ .

The method of choice for such a problem is singular value thresholding (SVT) [21, 22], which has been developed very recently in the context of classical *compressive sampling* or *matrix completion* [23] and may also be applied to the quantum setting [24–28]. SVT is a recursive algorithm that provably converges to a low-rank matrix satisfying constraints of the type  $\text{tr}[|\psi\rangle\langle\psi| \hat{P}_m^i] = \text{tr}[\hat{\sigma}_i \hat{P}_m^i]$ . Unfortunately, a straightforward application of SVT requires time and memory that scale exponentially with the number of particles. However, a modification of the algorithm allows us to overcome this problem.

Given the measured values  $\text{tr}[\hat{\sigma}_i \hat{P}_m^i]$  the algorithm may then be described as follows. First set up the operator  $\hat{R} = \sum_{m,i} \text{tr}[\hat{\sigma}_i \hat{P}_m^i] \hat{P}_m^i / 2^N$  and initialize  $\hat{Y}_0$  to some arbitrary matrix (e.g., the zero matrix). Then proceed recursively by finding the eigenstate  $|y_n\rangle$  with largest eigenvalue,  $y_n$ , of  $\hat{Y}_n$  and set

$$\hat{X}_n = y_n \sum_{m,i} \frac{\langle y_n | \hat{P}_m^i | y_n \rangle}{2^N} \hat{P}_m^i, \quad (6)$$

$$\hat{Y}_{n+1} = \hat{Y}_n + \delta_n (\hat{R} - \hat{X}_n).$$

So far, this algorithm still suffers from the fact that in every step the  $2^N \times 2^N$  matrix  $\hat{Y}_n$  needs to be diagonalized. However, the  $\hat{Y}_n$  are of the form  $\sum_{m,i} a_{m,i} \hat{P}_m^i$ ,  $a_{m,i} \in \mathbb{R}$ , where the  $\hat{P}_m^i$  act non-trivially only on sites  $i+1, \dots, i+k$ , i.e., they have the form of a local ‘‘Hamiltonian’’. Hence,  $|y_n\rangle$  can be determined as the highest energy state of this Hamiltonian. For this task standard methods have been developed in condensed matter physics [13, 29] for which the number of parameters scale polynomially in the system size and converge rapidly [30, 31] to the optimal MPS approximation. The standard but exponentially inefficient SVT algorithm possesses a convergence proof while our efficient modification does not. Hence, we now present numerical examples for different target states  $|\phi\rangle$  to demonstrate the feasibility and efficiency of the proposed algorithm.

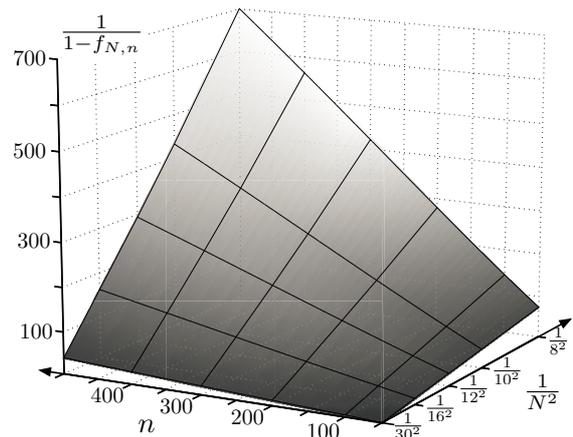


FIG. 2. Performance of the MPS-SVT algorithm for the ground state (the target state  $|\phi\rangle$ ) of the critical Ising model. We chose  $k=2$ , i.e., only nearest neighbour reductions are used to reconstruct the state. Plot illustrates the scaling of the error in the fidelity  $1 - f_{N,n} = 1 - |\langle\phi|y_n\rangle|^2 \sim N^2/n$  with the number of spins  $N$  and the iterations  $n$  of the algorithm.

These numerical examples suggest convergence of our algorithm to a MPS that closely matches the experimentally obtained reductions  $\hat{\sigma}_i$ . To arrive at the fidelity bound we follow the steps of Section II B: (i) Obtain estimates  $\hat{\sigma}_i$  of the reductions to  $k$  adjacent spins  $\hat{\rho}_i$  of the state in the laboratory such that  $\|\hat{\sigma}_i - \hat{\rho}_i\|_{\text{tr}} \leq \epsilon_i$ , (ii) compute the expectations  $p_{m_i} = \text{tr}[\hat{\sigma}_i \hat{P}_m^i]$ , which are the input to the MPS-SVT algorithm, (iii) obtain a MPS estimate  $|y_n\rangle$ , the reductions of which closely match the  $\hat{\sigma}_i$  by utilizing the MPS-SVT algorithm. As  $|y_n\rangle$  is an MPS, one can then efficiently obtain a *parent Hamiltonian* [10] and a lower bound,  $\Delta$ , on the energy gap above the ground state (see Methods for details). Putting all this together, the above programme returns a state  $|y_n\rangle$ , its parent Hamiltonian  $\hat{H} = \sum_i \hat{h}_i$ , and a number  $\Delta$  such that

$$\langle y_n | \hat{\rho}_i | y_n \rangle \geq 1 - \frac{\sum_i (\epsilon_i + \text{tr}[\hat{h}_i \hat{\sigma}_i])}{\Delta}. \quad (7)$$

#### E. Example: Strongly interacting quantum systems

We start with ground states of nearest-neighbor Hamiltonians on a chain, i.e., the  $|\phi\rangle = |gs\rangle$  are completely determined by all the reductions to two adjacent spins and we expect the above algorithm not only to produce states that match the reduced density matrices of the ground states but, in fact, states that are themselves close to the ground states. Among ground states of one-dimensional nearest-neighbor Hamiltonians, those at a critical point are the most challenging to approximate by MPS as they violate the entanglement area law [32]. We demonstrate

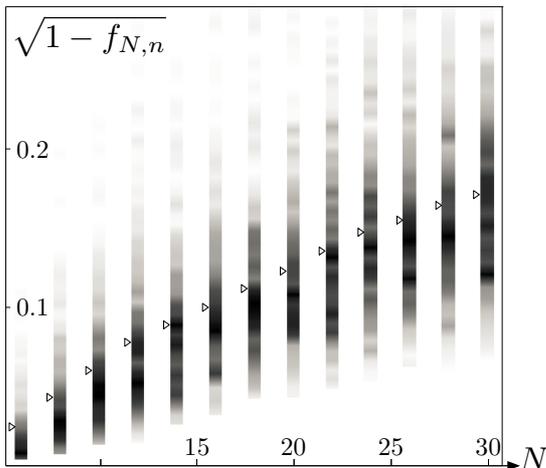


FIG. 3. Fidelity  $f_{N,n} = |\langle \phi | y_n \rangle|^2$ , for a fixed value  $n = 5$ , as a function of the number of spins  $N$  for the ground state (the target state  $|\phi\rangle$ ) of the random nearest-neighbour Hamiltonians of Eq. 8. As in Fig. 2, we chose  $k = 2$  to reconstruct the state. The plot shows densities (arrows indicate the mean) obtained from 1000 random realizations. Similar to the Ising model, the scaling for fixed  $n$  of  $1 - f_{N,n}$  is better than  $\sim N^2$ .

the effectiveness of our algorithm for such an example: the critical Ising model in a transverse field on a chain of length  $N$  with open boundary conditions. The ground state of this Hamiltonian is unique and hence it is completely characterized by its reductions to  $k = 2$  neighbouring sites. In other words, if we find an MPS that has the same reductions, the fidelity will be one. We proceed as follows. (i) We obtain the reductions  $\hat{\rho}_i$  to two neighbouring sites of the true ground state, (ii) from these reductions we obtain the expectations  $p_{m_i} = \text{tr}[\hat{\rho}_i \hat{P}_{m_i}]$ , which are the input to the MPS-SVT algorithm. In Fig. 2 we show the fidelity  $f_{N,n} = |\langle y_n | \phi \rangle|^2$  of the true ground state  $|\phi\rangle$  and the  $n$ 'th iterate of the above algorithm as a function of  $n$  and the length  $N$  of the chain. It shows that for fixed system size  $1 - f_{N,n}$  decreases as  $\sim 1/n$  while for fixed  $n$  it increases slower than  $N^2$ . This provides an indication that our algorithm is polynomial in the system size.

The Ising model is solvable and, in order to show that we are not considering a special case that is particularly favourable, we also consider one-dimensional random Hamiltonians of the form

$$\hat{H} = \sum_{i=1}^{N-1} \hat{r}_i^{(i)} \hat{r}_{i+1}^{(i)}, \quad (8)$$

where the  $\hat{r}_i^{(i)}$ ,  $\hat{r}_{i+1}^{(i)}$  act on spin  $i$  and  $i + 1$ , respectively, and are hermitian matrices with entries that have real and imaginary part picked from a uniform distribution over  $[-1, 1]$ . For each Hamiltonian, as before, we first

determine the ground state  $|gs\rangle$  (our target state  $|\phi\rangle$ ) and its reductions and then computed the fidelity  $f_{N,n}$  after  $n$  iterations of the MPS-SVT algorithm, see Fig. 3.

#### F. Example: W-state preparation in ion traps

Our method is of interest for many situations in which standard tomography will not be feasible. This is the case for the verification of state preparation in experiments with too many particles. An example is the recent ion trap experiment [3] for the preparation of W-states,  $|\phi\rangle = (|10\cdots 0\rangle + |010\cdots 0\rangle + \cdots + |0\cdots 01\rangle)/\sqrt{N}$ , that were limited to 8 qubits principally because the classical postprocessing of data became prohibitive for longer chains. Here we demonstrate the efficiency of our approach (we are not limited to few ions and demonstrate convergence for up to 20 ions – even higher number of ions are easily accessible due to the MPS alteration of the SVT method) by illustrating how one would postprocess experimentally obtained reduced density matrices to guarantee the generation of  $|\phi\rangle$  or a state very close to it. Note that, as in the above spin chain examples, one need only take tomographic data on pairs ( $k = 2$ ) of neighbouring qubits. We mimic experimental noise by adding Gaussian distributed random numbers with zero mean to the  $p_{m_i}$ . After initializing the algorithm with  $\hat{Y}_0 = \hat{R}$ , where we obtain  $\hat{R}$  from the MPS representation of  $|\phi\rangle$ , we use  $x_n := \sum_{m_i} |p_{m_i} - \langle y_n | \hat{P}_m^{i,k} | y_n \rangle|$  as a figure of merit for convergence, i.e., after a given number of iterations, we pick the  $|y_n\rangle$  with minimum  $x_n$ . The result of such a procedure is shown in Fig. 4.

### III. DISCUSSION

We have presented two schemes that efficiently produce an MPS description as a tomographic estimate of a quantum state, along with a tight fidelity bound. We emphasize that no assumptions are necessary for our scheme; if no such MPS description exists, this will be evident from the local tomographic data and our schemes will herald a failure. However, the enormous successes of MPS for describing both one-dimensional quantum systems found in nature as well as a host of states relevant to quantum information ensures that our methods will be very useful in practice.

So far we presented the method for pure states and one-dimensional systems. Various generalizations are possible: Our first scheme using unitary control can also handle mixed states corresponding to small ensembles of pure MPS. Suppose we are presented with a state  $\hat{\rho}$  that is a mixture of  $M$  pure states, each of which is an MPS having bond dimension  $D$ . Then the reduction of  $\hat{\rho}$  to any subsystem has rank at most  $MD$ . We can proceed as before, performing unitary operations on blocks of  $\kappa = \lceil \log_d(MD) \rceil + 1$  sites, in order to disentangle individual sites from the rest of the chain. At the end of

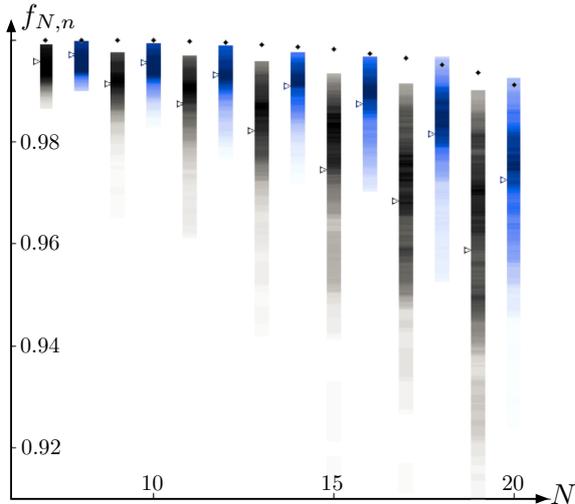


FIG. 4. Performance of the MPS-SVT algorithm for W-states,  $|\phi\rangle = (|10\dots 0\rangle + |010\dots 0\rangle + \dots + |0\dots 01\rangle)/\sqrt{N}$ . We are not limited to few ions and demonstrate convergence for up to 20 ions – even higher number of ions are easily accessible due to the MPS alteration of the SVT method, demonstrating the efficiency of our approach. We mimic experimental noise by adding Gaussian distributed random numbers with zero mean to the local expectations  $p_{m_i} = \text{tr}[\hat{\sigma}_i \hat{P}_{m_i}]$  and show results for  $n = 4000$  MPS-SVT iterations. Plot shows  $f_{N,n} = |\langle \phi | y_n \rangle|^2$  as a function of the number of ions,  $N$ , for no noise (dots) and Gaussian noise (densities obtained from 100 realizations for each  $N$ , arrows indicate mean) with a standard deviation of 0.005 (blue; only even values of  $N$  plotted for clarity) and 0.01 (black, odd  $N$ ).

the chain, we will find a mixed state  $\eta$  of rank  $M$  on the last  $\kappa - 1$  sites. We decompose  $\eta$  as a mixture of  $M$  pure states. This yields a representation of  $\hat{\rho}$  as an ensemble of  $M$  pure MPS, each with bond dimension at most  $dMD$ .

Our second scheme using local measurements can also be extended to handle mixed states. While the  $k$ -particle reduced density matrices do not uniquely determine the mixed state  $\hat{\rho}$ , reconstructions of better and better quality can be obtained by increasing  $k$ . As an example, suppose  $\hat{\rho}$  is the Gibbs state corresponding to a  $k$ -local Hamiltonian  $\hat{H}$ , i.e. the state minimizing the free energy

$$\text{tr}[\hat{\rho}\hat{H}] - TS(\hat{\rho}). \quad (9)$$

The first term is, as before, determined by the reduced density matrices. The entropy of the total state however can only be learned exactly from the complete density matrix. Yet, for essentially all reasonable physical systems, the entropy density  $\lim_{k \rightarrow \infty} S(\text{tr}_{k+1, \dots}(\hat{\rho}))/k$  in the thermal state of a Hamiltonian exists [33] (In particular, finitely correlated states are precisely those states whose entropy density is exactly equal to  $S(\text{tr}_{k+1, \dots}(\hat{\rho})) - S(\text{tr}_{k, \dots}(\hat{\rho}))$  for some finite value of  $k$ ). As a consequence, the total entropy of the state can be estimated efficiently from knowledge of the reduced density matrices. Hence, our second scheme using local measurements may be ex-

tended to mixed states by considering purifications and can then, e.g., also handle thermal states of local Hamiltonians, under the physically reasonable assumption that the entropy density exists [33]. In addition, it can be generalized to all mixed finitely correlated states (FCS) [9], though it is not always possible to certify the resulting estimates.

Higher-dimensional systems are more challenging, because the most straightforward generalization of MPS, known as projected entangled-pair states (PEPS) [34], cannot be computed as efficiently. However, the certification method using a frustration-free parent Hamiltonian remains efficient in the case of qubits with nearest-neighbor couplings [35]. Combinations of our techniques can be used to reconstruct other classes of states, such as tree tensor networks [36] and multiscale entanglement renormalization ansatz (MERA) states [37], for which efficient heuristics for minimizing local Hamiltonians are available.

## ACKNOWLEDGMENTS

The authors acknowledge discussion with F.G.S.L. Brandão at early stages of this project. The work at Ulm University has been supported by the EU Integrated Project QAP, the EU STREP's CORNER and HIP and an Alexander von Humboldt Professorship. SDB acknowledges the support of the Australian Research Council and the Perimeter Institute. STF and RS were supported by the Perimeter Institute for Theoretical Physics. Research at Perimeter is supported by the Government of Canada through Industry Canada and by the Province of Ontario through the Ministry of Research & Innovation. DG is glad to acknowledge support from the EU (CORNER). OLC and DP are partially funded by NSERC and FQRNT. YKL is funded by the US ARO/NSA.

## IV. METHODS

We start by recalling the MPS representation of a state  $|\psi\rangle$  with open boundary conditions (generalisations to periodic boundary conditions are entirely straightforward).

$$|\psi\rangle = \sum_{s_1=0}^{d_1-1} \dots \sum_{s_N=0}^{d_N-1} M_1[s_1] \dots M_N[s_N] |s_1 \dots s_N\rangle, \quad (10)$$

where the  $M_i[s]$  are  $D_i \times D_{i+1}$  matrices with  $D_1 = D_{N+1} = 1$ . We denote the *bond dimension* by  $D = \max D_i$ .

### A. Direct tomography

This method proceeds by disentangling all the qudits of the chain sequentially. Thus, it yields a valid MPS

description if every unitary exactly disentangles one qudit. Put another way, while it is crucial to obtain a good estimate of the  $d^{\kappa-1}$ -dimensional subspace on which  $\hat{\sigma}$  is supported, it is *not necessary* to identify the eigenvectors of  $\hat{\sigma}$  exactly: *any* set of orthonormal vectors generating the subspace is sufficient for our tomography procedure and leads to an MPS description in another gauge [10]. This property will be central to our error analysis.

To understand the effect of errors and imperfections in our tomography procedure, consider the very first step of the recursive procedure. Tomography is performed on the first  $\kappa$  sites to ideally find a state with non-maximal support, and unitary  $\hat{U}_1$  is applied to rotate that state into the subspace  $\mathcal{H}_1^{\text{cutoff}} = |0\rangle \otimes (\mathbb{C}^d)^{\otimes n-1}$ . In any experimental setting, the resulting state  $\hat{U}_1|\phi\rangle$  would surely not lie entirely in that subspace. This can be either because the state of the system is not exactly an MPS with bond dimension  $D$ , but merely close to one or because our estimate of the density matrix  $\hat{\sigma}$  on sites 1 to  $\kappa$  is slightly wrong due to measurements that are, in practice, noisy and restricted to finite precision. Indeed, we can expect that the reduced density matrix on the first  $\kappa$  sites will actually be full-rank, though most of its probability mass will lie on a subspace of dimension at most  $D$ . So, each time we apply a disentangling operation  $\hat{U}_i$ , we also want to *truncate* the reduced state to the desired subspace. Similarly, a faulty estimate of  $\hat{\sigma}$  will result in a small probability mass that lies outside the estimated support.

Given an estimated disentangling unitary  $\hat{U}_1$ , any state  $|\phi\rangle$  can be expressed as

$$\hat{U}_1|\phi\rangle = \frac{|0\rangle \otimes |\eta_1\rangle + |e_1\rangle}{\sqrt{1 + \langle e_1|e_1\rangle}} \quad (11)$$

where  $|e_1\rangle$  is some sub-normalized error vector supported on the subspace orthogonal to  $\mathcal{H}_1^{\text{cutoff}}$ . The unitary  $\hat{U}_1$  is chosen to minimize the norm  $\epsilon_1 = \langle e_1|e_1\rangle$  of this error vector. It is possible to directly estimate the magnitude of the error  $\epsilon_1$  by measuring the first qudit in the standard basis; the error is equal to the population of the non-zero states.

In subsequent steps of the recursion, we are given a state of the form

$$\hat{U}_i \dots \hat{U}_1|\phi\rangle = \frac{|0\rangle^{\otimes i} \otimes |\eta_i\rangle + |e_i^{\text{cm}}\rangle}{\sqrt{1 + \langle e_i^{\text{cm}}|e_i^{\text{cm}}\rangle}} \quad (12)$$

where  $|e_i^{\text{cm}}\rangle$  is the accumulated error vector that lies in the subspace orthogonal to  $\mathcal{H}_i^{\text{cutoff}} = |0\rangle^{\otimes i} \otimes (\mathbb{C}^d)^{\otimes N-i}$ . As a first step, we can truncate this error vector by measuring the first  $i$  particles in the standard basis and post-select on the all-zero outcome. This occurs with a probability roughly  $1 - \epsilon_i^{\text{cm}}$ , and leaves the system in the state  $|0\rangle^{\otimes i} \otimes |\eta_i\rangle$ . We then repeat the steps leading to Eq. 11 with the post-selected state  $|\eta_i\rangle$ . The resulting state will

be

$$\hat{U}_{i+1}\hat{U}_i \dots \hat{U}_1|\phi\rangle = \frac{|0\rangle^{\otimes i} \otimes \hat{U}_{i+1}|\eta_i\rangle + \hat{U}_{i+1}|e_i^{\text{cm}}\rangle}{\sqrt{1 + \langle e_i^{\text{cm}}|e_i^{\text{cm}}\rangle}} \quad (13)$$

$$= \frac{|0\rangle^{\otimes i+1} \otimes |\eta_{i+1}\rangle + |e_{i+1}^{\text{cm}}\rangle}{\sqrt{1 + \langle e_{i+1}^{\text{cm}}|e_{i+1}^{\text{cm}}\rangle}} \quad (14)$$

where

$$|e_{i+1}^{\text{cm}}\rangle = \frac{|e_i\rangle + U_{i+1}|e_i^{\text{cm}}\rangle}{\sqrt{1 + \langle e_i^{\text{cm}}|e_i^{\text{cm}}\rangle}} \quad (15)$$

and therefore  $\epsilon_{i+1}^{\text{cm}} \leq \epsilon_i^{\text{cm}} + \epsilon_{i+1}$ .

Thus, we see that errors accumulate *linearly* with the number of particles; if we denote  $|\psi\rangle = \hat{U}_1^\dagger \dots \hat{U}_{N-\kappa+1}^\dagger |0\rangle^{N-\kappa+1} |\eta_{N-\kappa+1}\rangle$  the estimated MPS, we have

$$\|\phi\rangle - |\psi\rangle\| = \left\| |e_{N-\kappa+1}^{\text{cm}}\rangle \right\| \leq \sum_{i=1}^{N-\kappa+1} \| |e_i\rangle \| \leq N\epsilon \quad (16)$$

where  $\epsilon = \max_i \| |e_i\rangle \|$ . The overall error is at most the sum of the individual errors on each step. In addition, each of the  $\epsilon_i$  is revealed during the tomographic procedure because they correspond to the post-selection success probability. This provides a direct method to *certify* the inferred state.

## B. Parent Hamiltonians

Let  $|\psi\rangle$  be as in (10) and such that  $\sum_s M_i[s]M_i[s]^\dagger = \mathbb{1}$  for all  $i = 1, \dots, N$ . This can always be achieved by subsuming qudits at the beginning and end of the chain into qudits with higher dimension and successive singular value decompositions [9, 10]. Now let  $N, k \in \mathbb{N}$  such that  $N/k \in \mathbb{N}$ , and assume that the MPS is injective [10] such that for all  $j = k, \dots, N - 2k$  the set

$$\left\{ M_{j+1}[s_1] \dots M_{j+k}[s_k] \mid s_i = 1, \dots, d_i \right\} \quad (17)$$

spans  $\mathbb{C}^{D_{j+1} \times D_{j+k+1}}$ . Then  $|\psi\rangle$  is the unique ground state of

$$\hat{H} = \sum_{n=0}^{N/k-2} \hat{P}_n, \quad (18)$$

where  $\hat{P}_n$  is the projector onto the subspace orthogonal to the range of the mapping  $\Gamma_n : \mathbb{C}^{D_{nk+2k+1} \times D_{nk+1}} \rightarrow \mathbb{C}^{d_{nk+1} \dots d_{nk+2k}}$ ,

$$\Gamma_n(X) = \sum_{\substack{s_{nk+1}, \\ \dots \\ s_{nk+2k}}} \text{tr} \left[ X M_{nk+1}[s_{nk+1}] \dots M_{nk+2k}[s_{nk+2k}] \right] \\ \times |s_{nk+1} \dots s_{nk+2k}\rangle. \quad (19)$$

To get an efficiently computable lower bound the energy gap, we use

$$\Delta E = \max \left\{ \lambda \mid \hat{H}(\hat{H} - \lambda) \geq 0 \right\}. \quad (20)$$

We find

$$\hat{H}^2 = \hat{H} + \sum_{\substack{n,m \\ n \neq m}} \hat{P}_n \hat{P}_m \geq \hat{H} + \sum_{\substack{n,m \\ 1 \leq |n-m|/k \leq 2k}} \frac{\hat{P}_n \hat{P}_m + \hat{P}_m \hat{P}_n}{2}, \quad (21)$$

where we omitted non-negative summands. Now consider the following quantity, which will bound the individual terms in the previous equation,

$$\begin{aligned} \gamma_{n,m} &= \min \left\{ \lambda \mid \hat{P}_n \hat{P}_m + \hat{P}_m \hat{P}_n + \lambda(\hat{P}_n + \hat{P}_m) \geq 0 \right\} \\ &= \min \left\{ \lambda \mid (\hat{P}_n + \hat{P}_m)^2 \geq (1 - \lambda)(\hat{P}_n + \hat{P}_m) \right\} \\ &= 1 - \max \left\{ \lambda \mid (\hat{P}_n + \hat{P}_m)^2 \geq \lambda(\hat{P}_n + \hat{P}_m) \right\}, \end{aligned} \quad (22)$$

where the maximum is given by the smallest non-zero eigenvalue of  $\hat{P}_n + \hat{P}_m$ . Hence

$$\hat{H}^2 \geq \hat{H} - \sum_{\substack{n,m \\ 1 \leq |n-m|/k \leq 2k}} \gamma_{n,m} \hat{P}_n \geq (1 - \gamma) \hat{H}, \quad (23)$$

and therefore we have the lower bound  $\Delta E \geq 1 - \gamma$ , where

$$\gamma = \max_{0 \leq n \leq N} \sum_{\substack{m \\ 1 \leq |n-m|/k \leq 2k}} \gamma_{n,m}. \quad (24)$$

Injectivity may fail to hold in certain singular cases. The simplest example given by the family of GHZ-type states  $\frac{1}{\sqrt{2}}(|0, \dots, 0\rangle + e^{i\phi}|1, \dots, 1\rangle)$ . Since any local reduced density matrix is independent of  $\phi$ , it is *impossible* to distinguish the members of that family based on local information alone. Indeed, in this example, the ground state space of  $\hat{H}$  will be two-dimensional, spanned by  $|0, \dots, 0\rangle$  and  $|1, \dots, 1\rangle$ . One may check that the gap analysis above continues to hold in the degenerate case (unlike the original in [9]), now certifying the overlap between  $\hat{\varrho}$  and the ground-state space. This fact alone implies an exponential reduction in the number of unknown parameters. It is easy to see that the small remaining uncertainty about  $\hat{\varrho}$  can be lifted in our example by measuring the ‘‘string operator’’  $\sigma_x \otimes \dots \otimes \sigma_x$ , which has expectation value  $\cos \phi$ . Using e.g. the results of [38], the method of this particular example immediately generalizes to any MPS with a two-fold degeneracy, such as the W state. Higher degeneracies may be treated by straight-forward, but more tedious, methods.

- 
- [1] Vogel, K. and Risken, H. Determination of quasiprobability distributions in terms of probability distributions for the rotated quadrature phase. *Phys. Rev. A* **40**, 2847–2849 (1989).
- [2] Smithey, D.T., Beck, M., Raymer, M.G. and Faridani, A. Measurement of the Wigner distribution and the density matrix of a light mode using optical homodyne tomography: Application to squeezed states and the vacuum. *Phys. Rev. Lett.* **70**, 1244–1247 (1993).
- [3] Häffner, H., Hänsel, W., Roos, C.F., Benhelm, J., Chekhal-kar, D., Chwalla, M., Körber, T., Rapol, U.D., Riebe, M., Schmidt, P.O., Becher, C., Gühne, O., Dür, W. and Blatt, R. Scalable multiparticle entanglement of trapped ions. *Nature* **438**, 643–646 (2005).
- [4] Leibfried, D., Knill, E., Seidelin, S., Britton, J., Blakestad, R.B., Chiaverini, J., Hume, D.B., Itano, W.M., Jost, J.D., Langer, C., Ozeri, R., Reichle, R. and Wineland, D.J. Creation of a six-atom ‘Schrödinger cat’ state. *Nature* **438**, 639–643 (2005).
- [5] James, D.F.V., Kwiat, P.G., Munro, W.J. and White, A.G. Measurement of qubits. *Phys. Rev. A* **64**, 052312 (2001).
- [6] Dunn, T. J. and Walmsley, I. A. and Mukamel, S. Experimental Determination of the Quantum-Mechanical State of a Molecular Vibrational Mode Using Fluorescence Tomography, *Phys. Rev. Lett.* **74**, 884 (1995).
- [7] Lvovsky, A.I. and Raymer, M.G. Continuous-variable optical quantum-state tomography. *Rev. Mod. Phys.* **81**, 299–332 (2009).
- [8] Paris, M. and Řeháček, J., eds., *Quantum State Estimation*, no. 649 in *Lect. Notes Phys.* (Springer, Heidelberg, 2004).
- [9] Fannes, M., Nachtergaele, B. and Werner, R.F. Finitely correlated states on quantum spin chains. *Comm. Math. Phys.* **144**, 443–490 (1992).
- [10] Perez-Garcia, D., Verstraete, F., Wolf, M.M and Cirac, J.I. Matrix product state representations. *Quant. Inf. Comp.* **7**, 401–430 (2007).
- [11] Vidal, G. Entanglement Renormalization. *Phys. Rev. Lett.* **99** 220405 (2007).
- [12] Rizzi, M., Montangero, S. and Vidal, G. Simulation of time evolution with multiscale entanglement renormalization ansatz. *Phys. Rev. A* **77**, 052328 (2008).
- [13] Östlund, S. and Rommer, S. Thermodynamic Limit of Density Matrix Renormalization. *Phys. Rev. Lett.* **75**, 3537–3540 (1995).
- [14] Hastings, M.B. Solving gapped Hamiltonians locally. *Phys. Rev. B* **73**, 085115 (2006).
- [15] Hastings, M.B. An area law for one-dimensional quantum systems. *J. Stat. Mech.* P08024 (2007).
- [16] Bloch, I., Dalibard, J. and Zwerger, W., Many-Body Physics with Ultracold Gases, *Rev. Mod. Phys.* **80**, 885 (2008).
- [17] Schoen, C., Hammerer, K., Wolf, M.M., Cirac, J. I. and Solano, E. Sequential Generation of Entangled Multi-qubit States. *Phys. Rev. A* **75**, 032311 (2007).

- [18] Vidal, G. Efficient Classical Simulation of Slightly Entangled Quantum Computations. *Phys. Rev. Lett.* **91**, 147902 (2003).
- [19] Gilchrist, A., Langford, A. K., Nielsen, M. A. Distance measures to compare real and ideal quantum processes, *Phys. Rev. A* **71**, 062310 (2005).
- [20] Briegel, H. J. and Raussendorf, R. Persistent Entanglement in arrays of Interacting Particles. *Phys. Rev. Lett.* **86**, 910–913 (2001).
- [21] Candes, E.J. and Recht, B. Exact Matrix Completion via Convex Optimization. arXiv:0805.4471 [cs.IT].
- [22] Cai, J-F., Candes, E.J. and Shen, Z. A Singular Value Thresholding Algorithm for Matrix Completion. arXiv:0810.3286 [math.OA].
- [23] Candes, E.J. and Wakin, M.B. An introduction to compressive sampling. *IEEE Sig. Proc. Mag.* **25**, 21–30 (2008).
- [24] Kosut, R.L. Quantum Process Tomography via  $l_1$ -norm Minimization. arXiv:0812.4323 [quant-ph].
- [25] Gross, D., Liu, Y-K., Flammia, S.T., Becker, S. and Eisert, J. Quantum state tomography via compressed sensing. arXiv:0909.3304 [quant-ph].
- [26] Gross, D. Recovering low-rank matrices from few coefficients in any basis. arXiv:0910.1879 [cs.IT].
- [27] Shabani, A., Kosut, R.L. and Rabitz, H. Compressed Quantum Process Tomography. arXiv:0910.5498 [quant-ph].
- [28] Shabani, A., Mohseni, M., Lloyd, S., Kosut, R.L. and Rabitz, H. Efficient estimation of nearly sparse many-body quantum Hamiltonians. arXiv:1002.1330 [quant-ph].
- [29] Schollwöck, U. The density-matrix renormalization group. *Rev. Mod. Phys.* **77**, 259–315 (2005).
- [30] Schuch, N. and Cirac, J.I. Matrix Product State and mean field solutions for one-dimensional systems can be found efficiently. *Phys. Rev. A* **82**, 012314 (2010).
- [31] Aharonov, D., Arad, I. and Irani, S. An Efficient Algorithm for approximating 1D Ground States. *Phys. Rev. A* **82**, 012315 (2010).
- [32] Eisert, J., Cramer, M., and Plenio, M.B. Area laws for the entanglement entropy - a review, *Rev. Mod. Phys.* **82**, 277 (2010).
- [33] Bratteli, O. and Robinson, D.W. *Operator Algebras and Quantum Statistical Mechanics*, Texts and Monographs in Physics Vol. 2 (Springer, New York) 1979.
- [34] Verstraete, F. and Cirac, J.I. Renormalization algorithms for Quantum-Many Body Systems in two and higher dimensions. arXiv:cond-mat/0407066.
- [35] de Beaudrap, N., Ohliger, M., Osborne, T.J. and Eisert, J. Solving frustration-free spin systems. arXiv:1005.3781 (2010).
- [36] Tagliacozzo, L., Evenbly, G. and Vidal, G. Simulation of two-dimensional quantum systems using a tree tensor network that exploits the entropic area law. *Phys. Rev. B* **80**, 235127 (2009).
- [37] Evenbly, G. and Vidal, G. Algorithms for entanglement renormalization. *Phys. Rev. B* **79**, 144108 (2009).
- [38] Walgate, J., Short, A., Hardy, L. and Vedral, V. Local Distinguishability of Multipartite Orthogonal Quantum States. *Phys. Rev. Lett* **85**, 4972–4975 (2000).