Preparation of Matrix Product States with Log-Depth Quantum Circuits

Malz, Daniel; Styliaris, Georgios; Wei, Zhi Yuan; Cirac, J. Ignacio

Published in:
Physical Review Letters

DOI:
10.1103/PhysRevLett.132.040404

Publication date:
2024

Document version
Publisher's PDF, also known as Version of record

Document license:
CC BY

Citation for published version (APA):
Preparation of Matrix Product States with Log-Depth Quantum Circuits

Daniel Malz,1,* Georgios Styliaris,2,3,* Zhi-Yuan Wei,2,3,* and J. Ignacio Cirac2,3

1Department of Mathematical Sciences, University of Copenhagen, Universitetsparken 5, 2100 Copenhagen, Denmark
2Max-Planck-Institut für Quantenoptik, Hans-Kopfermann-Str. 1, 85748 Garching, Germany
3Munich Center for Quantum Science and Technology (MCQST), Schellingstr. 4, 80799 München, Germany

(Received 1 August 2023; accepted 19 December 2023; published 24 January 2024)

We consider the preparation of matrix product states (MPS) on quantum devices via quantum circuits of local gates. We first prove that faithfully preparing translation-invariant normal MPS of $N$ sites requires a circuit depth $T = \Omega(\log N)$. We then introduce an algorithm based on the renormalization-group transformation to prepare normal MPS with an error $\epsilon$ in depth $T = O(\log(\frac{N}{\epsilon}))$, which is optimal. We also show that measurement and feedback leads to an exponential speedup of the algorithm to $T = O(\log \log(\frac{N}{\epsilon}))$. Measurements also allow one to prepare arbitrary translation-invariant MPS, including long-range non-normal ones, in the same depth. Finally, the algorithm naturally extends to inhomogeneous MPS.

One of the most important tasks in many-body physics and quantum information science is the preparation of useful or relevant states. This has spurred a large effort to find ways to prepare states, for example, adiabatically [1], dissipatively [2,3], or using quantum circuits. A natural class of states to consider are matrix product states (MPS), because they efficiently approximate ground states of gapped local Hamiltonians [4–6]. Moreover, many paradigmatic states can neatly be expressed as MPS, such as the cluster [7], Greenberger-Horne-Zeilinger (GHZ) [8], W [9], and Affleck-Kennedy-Lieb-Tasaki (AKLT) states [10,11].

Several ways are known to prepare MPS. Using unitary quantum circuits with strictly local gates, all MPS can be prepared using a sequential quantum circuit of depth $T \propto N$ [12]. This is provably optimal for long-range correlated states such as the GHZ state [13]. However, for so-called normal MPS [14], which have short-range correlations, shorter depths are possible. Indeed, when allowing for a small error $\epsilon$, they can be obtained by acting on a product state with a constant-depth circuit of quasilocal gates—gates whose support grows (poly-)logarithmically with system size [15,16]. However, such quasilocal gates have to be compiled into gates with strictly local support, and in the worst case such a compilation leads to circuits with a depth scaling exponentially in the support, and thus as $\text{poly}(N)$. However, since normal MPS all lie in the topologically trivial phase, one can construct adiabatic paths with a guaranteed gap [17], which means normal MPS can provably be prepared adiabatically in $T = O(\log(\frac{N}{\epsilon}))$ [18] (also see [19]).

Despite these results, it remains unclear if the scaling of the state-of-the-art algorithm [18] is optimal, or if there exist even faster algorithms to prepare normal MPS. Proving optimality requires finding a tight lower bound on the depth, or, equivalently, its complexity, which is believed to be difficult in general [13,20,21].

Here, we first resolve the question of asymptotically optimal preparation of normal translation-invariant (TI) MPS. We prove that any circuit faithfully preparing them requires a depth $T = \Omega(\log N)$, i.e., it has to scale at least logarithmically with $N$. We then introduce an algorithm that saturates this bound and prepares all normal TI-MPS in a circuit depth

$$T = O(\log(\frac{N}{\epsilon}))$$

using strictly local gates. This is asymptotically faster than the previously fastest known algorithm (adiabatic preparation [18]) and also asymptotically optimal. Moreover, the algorithm naturally extends to inhomogeneous MPS that are suitably short-range correlated.

If one has additionally access to measurements and feedback, it is known that MPS can be prepared exactly in a depth $T = O(\log N)$ by expressing them in terms of the multiscale entanglement renormalization ansatz (MERA) [22]. Including measurements also yields a speedup for our algorithm, and allows us to extend it to non-normal MPS, such that all TI-MPS can provably be prepared in depth...
Our goal is to devise an algorithm that approximates the target $N$-site MPS $|\phi_N\rangle$ by $|\tilde{\phi}_N\rangle$ with error $\epsilon = \epsilon(\phi_N, \tilde{\phi}_N)$, where

$$
\epsilon(\phi, \psi) = 1 - |\langle\phi|\psi\rangle|
$$

and $|\tilde{\phi}_N\rangle$ is prepared using a local quantum circuit. Our first result is that it is impossible to approximate well normal TI-MPS in depth $o(\log N)$. Subsequently, we provide an explicit algorithm with the asymptotically optimal depth $O(\log(N/e))$.

Lower bound.—Given (i) $\{|\phi_N\rangle\}$ a sequence of normalized TI-MPS on $N$ sites, generated by a normal tensor $A$, with finite correlation length $\xi > 0$ [cf. Eq. (3)], and (ii) $\{|\psi_N\rangle\}$ a sequence obtained from depth-$T$ local quantum circuits applied to product states, we are interested in determining how fast $T$ has to grow in order to approximate the MPS well, as measured by the error $\epsilon = \epsilon(\phi_N, \psi_N)$. We prove here that no quantum circuit with depth $T = o(\log N)$ can faithfully approximate this class.
Theorem 1.—If $T = \Theta(\log N)$ there is some $N_0$ such that
for all $N > N_0$ we have $\epsilon > 1/2$.

The proof can be found in the Appendix. To establish
this result, we use the fact that $|\psi_N\rangle$ have a strictly finite
light cone, whereas in a normal TI-MPS $|\phi_N\rangle$ correlation
functions decay only exponentially. This leads to a
mismatch in the expectation value of correlators outside
the light cone, which gives a lower bound on the error
between the two states. We additionally use the fact that
sufficiently distant parts of the system are statistically
independent, such that the error accumulates with increas-
ing system size $N$, unless the circuit depth grows suf-
ficiently quickly.

The algorithm.—We now present the key steps for our
algorithm. We will (i) approximate $|\phi_N\rangle$ by $|\tilde{\phi}_N\rangle$, then
(ii) show that $|\tilde{\phi}_N\rangle$ can be efficiently prepared, and
(iii) prove that the approximation error decays sufficiently
fast with $N$. We begin with the case of normal TI-MPS and
return to the general case later.

Approximation through the fixed-point state.—To make
the approximation, we follow the steps of the RG trans-
formation [24]. After blocking $q$ sites, we perform a polar
decomposition on the blocked tensor $B$, interpreting it
as a map from the $D^2$-dimensional virtual space to the
d$^q$-dimensional physical space [28]. This way we can
write $B = VP$ where $V$ is an isometry with $V^\dagger V = \mathbb{1}_D$;
and $P > 0$ is positive definite. Thus,

$$ E_B = \frac{BP^*}{P} \xrightarrow{q \to \infty} \frac{P^*}{P} = \rho $$

(8)

The approximation consists of replacing $P$ by its fixed-
point version $P_\infty$ in the tensor $B$, while keeping the
isometry $V$ intact. Graphically,

$$
\begin{array}{c}
\text{B} \\
\xrightarrow{P^*} \\
\text{V} \\
\xrightarrow{q \to \infty} \\
\text{P} \\
\xrightarrow{P} \\
\text{P}_{\infty} \\
\end{array}
\approx
\begin{array}{c}
\text{B} \\
\xrightarrow{P^*} \\
\text{V} \\
\xrightarrow{q \to \infty} \\
\text{P} \\
\xrightarrow{P} \\
\text{P}_{\infty} \\
\end{array}
\xrightarrow{\rho}.
\end{array}
$$

(9)

Later we will assign meaning to the approximation sign
in Eq. (9) by bounding the global error between the MPS $|\phi_N\rangle$ and $|\tilde{\phi}_N\rangle$ resulting from the two tensors, $B$ and $\tilde{B}$. To obtain
a vanishing error in the thermodynamic limit we will
need $q \propto \log N$, which we assume for now and justify
subsequently.

Preparing the approximate state.—The approximate
state $|\tilde{\phi}_N\rangle$ can be prepared by acting on the fixed-point
state with a product of unitaries of support $q$ (for simplicity
$D = d$ in the illustration):

$$
\begin{array}{c}
\text{U} \\
\xrightarrow{0} \\
\text{L} \\
\xrightarrow{C} \\
\text{R} \\
\xrightarrow{C} \\
\text{R}_{\infty} \\
\end{array}
\approx
\begin{array}{c}
\text{U} \\
\xrightarrow{0} \\
\text{L} \\
\xrightarrow{C} \\
\text{R} \\
\xrightarrow{C} \\
\text{R}_{\infty} \\
\end{array}
\xrightarrow{\tilde{U}}.
\end{array}
$$

(10)

where in the last step we set $A_i^q = A_i^q \otimes \mathbb{1}_D$ and contracted
$P^*$ with the rightmost $A_i^q$ to obtain $C$. As in sequential
preparation of MPS [12] and in the left-canonical
form [31], we can now iteratively apply singular value
decompositions, starting from the tensor on the left and
moving right, but stopping before the last tensor [32]. This
defines a new set of tensors that describe the same isometry
$V$ but now each tensor is a local isometry (arrows indicate
isometry direction, $q = 4$ in illustration):

$$
V = V_q \ldots V_1 = \frac{F_4 \ldots F_1}{2^{14}} \frac{G_4 \ldots G_1}{2^{14}}, \quad V_1 : \mathbb{C}^{D'_{ij}} \to \mathbb{C}^{D'_{ij+1}},
$$

(14)

with every $V_i$ an isometry $V_i^\dagger V_i = 1_{D'_i}$ satisfying $D'_i \leq D^2$
($D'_{i+1} = 1$). Importantly, $\tilde{C} = V_1$ is automatically also an
isometry, as
Since this sequential circuit comprises \( q \) sites, its depth is \( O(q) \). This scaling is unchanged if we additionally take into account that the inputs of the unitary in Eq. (13) are separated by \( O(q) \) sites, which requires one to implement SWAP gates.

The tree-RG circuit.—Blocking two neighboring sites followed by a polar decomposition is the basis for the real-space RG transformation and halves the correlation length \([24,35,36]\). Instead of directly blocking \( q \) sites, we can repeatedly apply this transformation \( k \sim \log_{2} q \) times to the same effect [illustration below, and in Fig. 1(b)]. This generates a treelike circuit with \( k \) layers, in which each layer but the lowest consists of isometries from dimension \( D^{2} \) to \( D^{q} \) (below \( A_{k} \) is obtained from \( A \) by blocking \( k \) sites, and \( q = 8 \)):

\[
\begin{align*}
V^{(q)} A_1 A_2 A_3 A_4 &= V^{(q)} D_1 D_2 D_3 D_4 = V^{(q/2)} V^{(q/4)} = \cdots \ .
\end{align*}
\] (16)

In Eq. (16), the lowest layer is again the part that is replaced by the fixed-point state in our algorithm, i.e., a product of \( \ket{\omega} \) [cf. Eq. (12)]. In this scheme, the lowest isometry \( V^{(k)} \) acts across a distance \( q \). Though not strictly local, this can be done in a depth \( O(q) \) utilizing SWAP gates. Subsequent isometries act over distances \( q/2, q/4 \) and so forth, leading to an overall circuit depth \( T = O(q) \).

Approximation error.—So far, we constructed efficient circuits for preparing \( \ket{\phi_{N}} \), having assumed that we block \( q \) sites. The scaling of \( q \) is a consequence of the following Lemma, which is adapted from Ref. [16].

Lemma 1.—Given a sequence of TI-MPS generated from a normal tensor, and for all \( \gamma < 1/2 \),

\[
\epsilon(\phi_{N}, \phi_{N}) = O\left(\frac{N}{q} e^{-q\gamma/\xi}\right) .
\] (17)

The proof can be found in [37]. Using Lemma 1, it follows that \( q = O(\log(\log(N/e))) \). In particular, blocking \( q = \left[2(1+\eta)\ln N\right] \log N \) sites gives \( \epsilon = O(N^{-\eta}) \) for any \( \eta > 0 \).

We also numerically illustrate the exponential decay of Eq. (17) in [37] for preparing the 1D AKLT state [10,11] and an MPS family with tunable correlation length [44], which demonstrates that the circuit is also efficient in practice.

Inhomogeneous short-range correlated MPS.—Our results can be straightforwardly extended to MPS that have a finite correlation length, but are not TI. The setting here is that we are given a sequence of MPS \( \{\ket{\phi_{N}}\} \) with bond dimension at most \( D \). We define such a sequence to have finite correlation length if, after blocking \( q = O(\log N) \) times, the resulting states can be approximated up to quasilocal isometries by a state consisting of nearest-neighbor entangled pairs \( \ket{\Omega} = \otimes_{i=1}^{N/q} \ket{\omega_{i}}_{R,L,i} \), with an error \( \epsilon(\Omega, \phi_{N}) \to 0 \) as \( N \to \infty \). Here,

\[
\ket{\phi_{pos}} = \cdots \cdots \ ,
\] (18)

arises after blocking \( q \) sites and keeping the positive part of the decomposition of \( \ket{\phi_{N}} \). If the finite correlation assumption is satisfied, then the preparation scheme consists of preparing \( \ket{\Omega} \) and implementing the isometry, decomposed with either of the two methods. The resulting total depth is again \( O(\log(N/e)) \) with error \( \epsilon(\Omega, \phi_{N}) \), as in the TI case.

In the Supplemental Material [37] we numerically show that this protocol can prepare inhomogeneous random MPS [45–48] efficiently. For that, we use a simple extension of the Evenbly-Vidal algorithm [49,50], which efficiently variationally finds \( \ket{\Omega} \). This illustrates that our finite correlation length assumption, as defined earlier, is satisfied in a practical setting.

Preparations using measurements.—Measurements and subsequent conditional unitaries can make state preparation much faster [22,26,51–55]. Here, we elaborate how such measurements could be used in our algorithm.

Tree-RG circuit with measurements.—Local measurements and conditional local unitaries are the standard framework to perform quantum teleportation [56,57], which can be used to reduce the depth of the tree-RG circuit. Isometries appearing in Eq. (16) act on a constant number of sites, which, although spatially separated, can be teleported at neighboring registers with a constant overhead. This can be achieved by creating nearest-neighbor entangled pairs, then performing simultaneous measurements, and correcting (without postselection) based on the measurement outcomes [58] (this process is also detailed in Ref. [22]).

Therefore every isometry in Eq. (16) takes constant time using measurement. Crucially, however, the tree-RG circuit requires only \( O(\log(\log(N/e))) \) layers (in contrast to Ref. [22]). Since the fixed-point state can be prepared in constant time as before, this gives a preparation algorithm for short-range correlated MPS with depth \( O(\log(\log(N/e))) \).

Long-range MPS using measurements.—Another consequence of including measurements is that the creation of GHZ-like states \( \ket{\chi_{M}} = \sum_{i=1}^{b} a_{i} \ket{i}^{\otimes M} \) becomes possible in only constant depth [16,26,59]. These states are closely related to the fixed points of TI-MPS [33], which, up to an isometry, take the form [25]
\[ |\Omega\rangle = \sum_{j=1}^{b} a_j^{(N)} \otimes |\omega_j\rangle_{R_iL_i}, \]  

(19)

The normal case corresponds to \( b = 1 \) for which \( |\Omega\rangle = |\Omega\rangle \) while, in general, \( b \) is upper bounded by the number of blocks in the canonical form and \( a_j^{(N)} \) may depend on \( N \) [25].

Importantly, the different \( |\omega_j\rangle \) are orthogonal [25], which suggests a preparation procedure for \( |\Omega\rangle \). First create \( |\chi_{N/q}\rangle \), which can be done in constant depth with measurements (following, e.g., Ref [16]). Subsequently, apply in parallel the isometries \( W: |j\rangle \mapsto |\omega_j\rangle_{R_iL_i} \) such that \( |\Omega\rangle = W^{\otimes N/d}|\chi_{N/q}\rangle \), which also takes constant depth.

In [37] we show how to explicitly obtain a state of the form Eq. (19) that approximates well the target \( |\phi_N\rangle \) up to local isometries by blocking \( q \propto \log(N/e) \) sites. As a result, following the same steps as in the tree-RG circuit with measurements we have a scheme that approximates all TI-MPS (short- or long-range correlated) with depth \( T = O(\log \log N/e) \) [cf. Eq. (2)]. If instead measurements are only used for the preparation of \( |\chi_{N/q}\rangle \), the depth is \( O(\log(N/e)) \).

Our construction generalizes to inhomogeneous long-range correlated MPS exactly as in the short-range case.

Connection to MERA.—The circuit in the tree-RG scheme can be interpreted as a finite-range MERA with \( O(\log \log N) \) layers, namely a shallow tensor tree acting on the fixed-point state. Specifically, the isometries \( V^{(i)} \) [cf. Eq. (16)] are identified with the isometries in finite-range MERA, and all disentanglers are the identity, save for the first layer, which is identified with the single layer of unitaries that prepare the fixed-point state. Hence, within the approximation error \( \epsilon \),

normal TI-MPS \( \subset \) finite-range MERA  

with \( O(\log \log N) \) layers.  

(20)

Discussion and outlook.—Our results also imply that MPS in the same phase can be transformed into each other using a log-depth circuit, in contrast to the well-known quasilocal evolution corresponding to polylogarithmic depth circuit [17,60–62]. It would be interesting to explore whether our results could be exploited for applications other than state preparation. Specifically, a number of protocols [50,63–67] implicitly or explicitly depend on the ability to prepare (or disentangle) MPS using a sequential circuit. It may be possible to replace the sequential circuit with ours to reduce the circuit depth in these protocols. Another direction would be to extend our lower-bound proof and the preparation algorithm to prepare certain higher dimensional tensor network states [34].

We thank Yujie Liu, Miguel Frías Pérez, and Rahul Trivedi for insightful discussions. D. M. acknowledges support from the Novo Nordisk Fonden under Grants No. NNF22OC0071934 and No. NNF20OC0059939. G. S. is supported by the Alexander von Humboldt Foundation. The research is part of the Munich Quantum Valley, which is supported by the Bavarian State Government with funds from the Hightech Agenda Bayern Plus. We acknowledge funding from the German Federal Ministry of Education and Research (BMBF) through EQUAHUMO (Grant No. 13N16066) within the funding program quantum technologies—from basic research to market. The numerical calculations were performed using the ITensor Library [68].

Appendix: Proof of Theorem 1.—Before we present the proof, let us introduce the following lemma, which we will use to distinguish the states based on the mismatch between states with strictly finite correlation length and states with exponentially decaying correlations.

Lemma 2 (exponentially decaying correlations).—Let \( \{\phi_N\} \) be a sequence of TI-MPS generated by an injective tensor \( A \) with finite correlation length \( \xi > 0 \). Then, we can always find two local operators \( O_i, O_s \) acting on spins 1 and \( s \) with \( ||O_i|| = ||O_s|| = 1 \) such that for any integer \( s > 1 \) and sufficiently large \( N \),

\[ \langle \phi_N | O_i | \phi_N \rangle = \langle \phi_N | O_s | \phi_N \rangle = 0, \]  

(A1a)

\[ \langle \phi_N | O_i | \phi_N \rangle \geq c e^{-(s-1)/\xi}, \]  

(A1b)

where \( c > 0 \) is independent of \( N, s \).

Proof.—Consider the connected correlation function

\[ \Delta = \langle \phi_N | O_i | \phi_N \rangle - \langle \phi_N | O_s | \phi_N \rangle \langle \phi_N | O_s^* | \phi_N \rangle, \]  

(A2)

where \( O_i \) and \( O_s \) are two (potentially different) operators placed at sites 1 and \( s \). We have

\[ \Delta = \frac{1}{c_N} \left[ \text{Tr}(E_1^{N-s-1} E_1^{-1} E_s^{N-1} E_s^{-1}) \right. \]  

\[ \left. - \text{Tr}(E_1^{N-s} E_1^{-1} E_s^{-1}) \text{Tr}(E_1^{N-1} E_s^{-1}) \right], \]  

(A3)

where the normalization \( c_N = \sqrt{\text{Tr}(E_1^N)} \), and

\[ E_Q = \sum_{i,j=1}^{d} \langle i | Q | j \rangle (A^i)^* \otimes A^j, \]  

(A4)

where \( d \) is the physical dimension of the MPS. Given the spectrum of \( E_1 \) we can always take \( N \) sufficiently large so that we can approximate with an arbitrarily small error, \( E_1^{N-1} = E_1^{N-s-1} = |R_1\rangle \langle L_1| + O(e^{-N/\xi}) \), where \( \xi = -1/\ln(|\lambda_2|) \) and \( 1 = A_1 > |\lambda_2| > \ldots \) are the eigenvalues of \( E_1 \). Note that in the main text we use the gauge in which \( |R_1\rangle = |\rho\rangle \) and \( |L_1\rangle = |\uparrow\rangle \). Here, \( (L_1|R_1) = 1 \), so that \( c_N \approx 1 \) and
\[ \Delta \approx \sum_{i=2}^{D^4} \lambda_i^{4-1} \langle L_i | E_{O_i} | R_i \rangle \langle L_i | E_{O'} | R_i \rangle, \]  
(A5)

where \( D \) is the bond dimension and we have written \( E_i^{4-1} = \sum_i \lambda_i^{4-1} | R_i \rangle \langle L_i |. \) Since the tensor \( A \) is injective \([14]\), we can always choose \( O \) (and \( O' \)), such that the corresponding transfer matrix \( E_O = |A\rangle\langle B| \) for arbitrary \( A \), \( B \) (up to a normalization constant). In particular, we can use this to impose that

\[ \langle L_i | E_{O'} | R_i \rangle = 0, \quad \forall \ i \]  
(A6a)

\[ \langle L_i | E_{O} | R_i \rangle = \langle L_i | E_{O'} | R_i \rangle = 0, \quad \forall \ i \neq j \quad (A6b) \]

\[ \langle L_i | E_{O} | R_i \rangle = c' > 0. \]  
(A6c)

The first line ensures (A1a), while the second and third ensure (A1b) for sufficiently large \( N \), with \( c = c'/2 \), where \( 1/2 \) is an arbitrary constant chosen for concreteness. 

Now, we can prove Theorem 1. Let \(|\{\phi_N\}\rangle\rangle\) be a sequence of TI normalized MPS on \( N \) sites generated by a normal tensor \( A \) and \(|\{\psi_N\}\rangle\rangle\) a sequence of states obtained by applying a depth-\( T \) local quantum circuit to a product state and define the error \( \varepsilon = 1 - |\langle \psi_N | \phi_N \rangle| \).

**Theorem 1 (restated).—**If \( T = O(\log N) \) there is some \( N_0 \) such that for all \( N > N_0 \) we have \( \varepsilon > 1/2 \).

**Proof.**—Let us assume that \( T = O(\log N) \) and \( T > 2\xi \), since we can always add layers of identity operators to increase the depth of the circuit. We approximate \(|\{\phi_N\}\rangle\rangle\) through \(|\{\bar{\phi}_N\}\rangle\rangle\) [cf. Eq. (10)] obtained by blocking \( q_N = [2(1 + \eta)]^\xi \log N \) with \( \eta > 0 \) and use Lemma 1 to bound the error as

\[ \varepsilon = 1 - |\langle \bar{\phi}_N | \phi_N \rangle| < c_0 N^{-\eta} \]  
(A7)

for some constant \( c_0 \) independent of system size. We take \( N \) such that we have a large number of blocks, all of the same size, \( q_N \), except for the last one, which may be larger. This is always possible, as \( q_N = O(\log N) \). We also take \( N \) large enough to ensure \( q_N > T \).

We thus have

\[ d(\phi_N, \psi_N) \geq d(\psi_N, \bar{\phi}_N) - \sqrt{2c_0 N^{-\eta/2}}, \]  
(A8)

where \( d(\rho, \sigma) = ||\rho - \sigma||_1/2 \) is the trace distance \([58]\), as well as an upper bound on trace distance from fidelity combined with Eq. (A7)

\[ d(\phi_N, \bar{\phi}_N) \leq \sqrt{1 - |\langle \phi_N | \bar{\phi}_N \rangle|^2} \leq \sqrt{2c_0 N^{-\eta/2}}. \]  
(A9)

In the following, we will find a lower bound to the first term in Eq. (A8) to make the difference larger than \( 1/2 \). We will also drop the subscript \( N \) to simplify notation.

To obtain a bound on the distance of \(|\psi_N\rangle\rangle\) and \(|\bar{\phi}_N\rangle\rangle\), let us consider instead a suitable subsystem. To that end, we divide the chain into \([N/(2q)]\) blocks of size \( 2q \) each, with the last block potentially smaller than \( 2q \). We then trace over all \( 2q \) spins at the sites contained in the intervals \([4mq + 1, 2(m + 1)q] \), with \( m = 0, 1, \ldots \) in both states \(|\bar{\phi}\rangle\rangle\) and \(|\psi\rangle\rangle\). In case the last block we constructed is smaller than \( 2q \), we trace it as well. If we perform such an operation on \(|\phi\rangle\rangle\), we obtain a product state

\[ \rho = \rho_0^k, \]  
(A10)

which follows from the definition of \(|\bar{\phi}_N\rangle\rangle\) and the fact that it is invariant under translation by \( q \) sites. We have

\[ k = [N/4q]. \]  
(A11)

Analogously, applying the same trace to \(|\psi\rangle\rangle\) we also obtain a product state, because \( q > T \),

\[ \sigma = \sigma_1 \otimes \ldots \otimes \sigma_k. \]  
(A12)

Using the fact that the trace distance is contractive under tracing, and bounding it in terms of the Uhlmann fidelity, we have \([58]\)

\[ d(\bar{\phi}, \psi) \geq d(\rho, \sigma) \geq 1 - F(\rho, \sigma), \]  
(A13)

with the Uhlmann fidelity between two density matrices \( \rho \) and \( \sigma \) defined as \( F(\rho, \sigma) = \text{Tr} \sqrt{\sqrt{\rho} \sigma \sqrt{\rho}} \).

Given that \( \rho \) and \( \sigma \) are product states, we have

\[ F(\rho, \sigma) = \prod_{i=1}^k F(\rho_0, \sigma_i) \leq (1 - \delta)^{k/2}, \]  
(A14)

where \( \delta = \min d(\rho_0, \sigma_i)^2 \), and where we have used another bound between the fidelity and the trace distance \([58]\).

Next we will lower bound \( \delta \) using Lemma 2. However, we have to be a bit careful since this lemma applies to \( \phi \) instead of \( \bar{\phi} \). Fortunately, we can use Eq. (A9) to replace one with the other. For the sake of concreteness, we will bound \( d(\rho_0, \sigma) \) but the same analysis applies to every \( \sigma_i \).

Let us take \( s = 2T + 1 \), and the operators \( O, O' \) from Lemma 2 to define

\[ a_Q = \langle \phi | Q | \phi \rangle, \]  
(A15)

\[ a_Q = \text{Tr}(\rho_0 Q) = \langle \bar{\phi} | Q | \bar{\phi} \rangle, \]  
(A16)

\[ b_Q = \text{Tr}(\sigma_1 Q) = \langle \psi | Q | \psi \rangle, \]  
(A17)

where \( Q \in \{ O_1, O'_1, O_1 O'_1 \} \). Given that \( ||O|| = ||O'|| = 1 \), we can bound
\[
\begin{align*}
\mu(\sigma_1) & \geq \max_{O^C=\sigma_1} |\hat{a}_O| - |\hat{b}_O|.
\end{align*}
\] (A18)

According to Lemma 2, \(a_{\sigma_1} = a_{\sigma_2} = 0\), and
\[
\begin{align*}
a_{\sigma_1} \geq a_{\sigma_2} & = \mu - c_1 e^{-2T/\xi}.
\end{align*}
\] (A19)

In order to use Eq. (A18), we need to connect \(\hat{a}_O\) to \(a_{\sigma_1}\). Using Eq. (A9), we choose sufficiently large \(N\) to obtain \(\hat{d}(\phi, \hat{\phi}) \leq \mu/3\). This immediately implies that \(|\hat{a}_O|, |\hat{b}_C| < \mu/3\) and \(\hat{a}_{\sigma_1} > 2\mu/3\). Moreover, since \(\psi\) is created from a product state by a depth-\(T\) circuit, every connected correlation for operators at a distance larger than \(2T\) vanishes. Since \(s = 2T + 1\) we therefore have \(b_{\sigma_1} = b_{\sigma_2} = b_{\sigma_2}\). We now show that for any choice of \(b_{\sigma_1}\), \(b_{\sigma_2}\), the distance \(d(\rho_0, \sigma_1)\) is bounded below by a constant. Thus, we minimize Eq. (A18) with respect to \(x_1 = b_{\sigma_1}\) and \(x_2 = b_{\sigma_2}\), imposing \(\mu \leq 1/2\):
\[
\begin{align*}
d(\rho_0, \sigma_1) & \geq \min_{x_1, x_2} \left[ \max \left( x_1 - \frac{\mu}{3}, x_2 - \frac{\mu}{3}, \frac{2\mu}{3} - x_1 x_2 \right) \right]
\end{align*}
\]
\[
\begin{align*}
> \min_x \left[ \max (|x - \mu/3|, |2\mu/3 - x_2|) \right] > \mu/3.
\end{align*}
\] (A20)

From this it immediately follows that \(\delta > \mu^2/9\).

Putting Eqs. (A20) and (A14) into Eq. (A8), we arrive at
\[
\begin{align*}
\hat{d}(\phi, \psi) & > 1 - (1 - \mu^2/9)^{3/2} - \sqrt{2c_0 N^{-\eta}}.
\end{align*}
\] (A21)

The last term is negligible for large \(N\). If then \(k > 2/\delta\), one has that \((1 - \delta)^{3/2} < 1/e\) and thus \(\hat{d}(\phi, \psi) > \sqrt{3}/4\), which implies \(\epsilon > 1/2\).

So we need to find \(N\) for which \(k > 2/\delta\). Using Eq. (A11) and the definition of \(q\), we have \(k > N/(10\xi \ln N)\), where we chose \(\eta < 1/4\). Using \(\delta > \mu^2/9\) and the definition of \(\mu\) [Eq. (A19)], we find \((18e^{4T/\xi})/c_1 > 2/\delta\). Putting this together, we need to find \(N\) such that
\[
\begin{align*}
k > \frac{N}{5q} > \frac{N\eta}{10\xi \ln N} > \frac{18}{c_1} e^{4T/\xi} > \frac{2}{\delta}.
\end{align*}
\] (A22)

Since \(T = o(\log(N))\), we can always find an \(N_0\) such that this is fulfilled for all \(N > N_0\). \(\Box\)

*These authors contributed equally to this work.


[28] To do a polar decomposition we need q large enough that $d^q \geq D^2$. Later, we will have q scaling with the system size, so here we assume that B is injective [14]. This is always true for normal tensors after blocking finite (independent of $N$) sites [29].


[30] The subsequent derivation remains valid also for non-injective tensors B. In that case $P^{-1}$ is understood as pseudoinverse.


[32] The method can easily be generalized to absorb $P^{-1}$ into any of the tensors. With C in the bulk, the sequential circuit is obtained by repeated singular value decomposition starting both left and right and stopping at C, as in the mixed-canonical form [31].

[33] Here, we consider TI-MPS with constant D. Note that this does not include all states that are TI and have area law entanglement (e.g., W state [14,34]).


[40] P. Rakya and Z. Zimborás, Approaching the theoretical limit in quantum gate decomposition, Quantum 6, 710 (2022).


