





Embedding of Tree Tensor Networks into Shallow Quantum Circuits

Shota Sugawara¹^a, Kazuki Inomata¹^b, Tsuyoshi Okubo²^c and Syngé Todo^{1,2,3}^d

¹*Department of Physics, The University of Tokyo, Tokyo, 113-0033, Japan*

²*Institute for Physics of Intelligence, The University of Tokyo, Tokyo, 113-0033, Japan*

³*Institute for Solid State Physics, The University of Tokyo, Kashiwa, 277-8581, Japan*
{shota.sugawara, kazuki.inomata, t-okubo, wistaria}@phys.s.u-tokyo.ac.jp

Keywords: Tree Tensor Networks, Quantum Circuits, Variational Quantum Algorithms.

Abstract: Variational Quantum Algorithms (VQAs) are being highlighted as key quantum algorithms for demonstrating quantum advantage on Noisy Intermediate-Scale Quantum (NISQ) devices, which are limited to executing shallow quantum circuits because of noise. However, the barren plateau problem, where the gradient of the loss function becomes exponentially small with system size, hinders this goal. Recent studies suggest that embedding tensor networks into quantum circuits and initializing the parameters can avoid the barren plateau. Yet, embedding tensor networks into quantum circuits is generally difficult, and methods have been limited to the simplest structure, Matrix Product States (MPSs). This study proposes a method to embed Tree Tensor Networks (TTNs), characterized by their hierarchical structure, into shallow quantum circuits. TTNs are suitable for representing two-dimensional systems and systems with long-range correlations, which MPSs are inadequate for representing. Our numerical results show that embedding TTNs provides better initial quantum circuits than MPS. Additionally, our method has a practical computational complexity, making it applicable to a wide range of TTNs. This study is expected to extend the application of VQAs to two-dimensional systems and those with long-range correlations, which have been challenging to utilize.


1 INTRODUCTION


Variational Quantum Algorithms (VQAs) represent the foremost approach for achieving quantum advantage with the current generation of quantum computing technologies. Quantum computers are anticipated to outperform classical ones, with some algorithms already proving more efficient (Shor, 1994; Grover, 1996). However, the currently available Noisy Intermediate-Scale Quantum (NISQ) devices are incapable of executing most algorithms due to their limited number of qubits and susceptibility to noise (Preskill, 2018).


VQAs are hybrid methods where classical computers optimize parameters of quantum circuits' ansatz to minimize the cost function evaluated by quantum computers. VQAs require only shallow circuits, making them notable as algorithms that can be executed on NISQ devices (Cerezo et al., 2021a). VQAs come in many forms, such as Quantum Ma-


chine Learning (QML) (Biamonte et al., 2017; Mitarai et al., 2018; Schuld and Killoran, 2019), Variational Quantum Eigensolver (VQE) (Abrams and Lloyd, 1999; Peruzzo et al., 2005; Peruzzo et al., 2014), and Quantum Approximate Optimization Algorithm (QAOA) (Farhi et al., 2014; Wang et al., 2018; Zhou et al., 2020), with potential applications across diverse industries and fields.

However, a significant challenge known as the barren plateau stands in the way of realizing quantum advantage (McClean et al., 2018; Holmes et al., 2022). The barren plateau phenomenon refers to the challenge in VQAs where the gradient of the cost function decreases exponentially as the system size increases. This phenomenon occurs regardless of whether the optimization method is gradient-based (Cerezo and Coles, 2021) or gradient-free (Arasmith et al., 2021), and it has been observed even in shallow quantum circuits (Cerezo et al., 2021b). Furthermore, it has been confirmed that this phenomenon also occurs in practical tasks using real-world data (Holmes et al., 2021; Sharma et al., 2022; Ortiz Marrero et al., 2021). Avoiding the barren plateau is a critical challenge in demonstrating the

^a <https://orcid.org/0009-0005-8758-2940>

^b <https://orcid.org/0009-0008-4030-5518>

^c <https://orcid.org/0000-0003-4334-7293>

^d <https://orcid.org/0000-0001-9338-0548>

superiority of quantum algorithms using NISQ devices. To avoid the barren plateau, appropriate parameter initialization in VQAs is crucial since randomly initializing the parameters can result in the algorithm starting far from the solution or near a local minimum (Zhou et al., 2020). Although various initialization methods have been considered (Wiersema et al., 2020; Grant et al., 2019; Friedrich and Maziero, 2022), using tensor networks is natural due to their compatibility with quantum circuits.

Tensor networks are originally developed to efficiently represent quantum many-body wave functions. Any quantum circuit can be naturally regarded as a tensor network (Markov and Shi, 2008), and it is sometimes possible to simulate quantum computers with a practical amount of time using tensor networks (Liu et al., 2021). Moreover, in recent years, their utility has been recognized and applied across various fields such as machine learning (Levine et al., 2019) and language models (Gallego and Orus, 2022)

In this study, we focus particularly on Matrix Product States (MPSs) and Tree Tensor Networks (TTNs) among various structures of tensor networks. MPSs are one-dimensional arrays of tensors. Its simplest and easiest-to-use structure, along with the presence of advanced algorithms such as Density Matrix Renormalization Group (DMRG) (White, 1992) and Time-evolving block decimation (TEBD) (Vidal, 2003), has led to its application across a wide range of fields. Recently, these excellent algorithms have been applied to the field of machine learning, continuing the exploration of new possibilities (Stoudenmire and Schwab, 2016; Han et al., 2018). TTNs are tree-like structures of tensors. The procedures used in DMRG and TEBD have been applied to TTNs (Shi et al., 2006; Silvi et al., 2010). Additionally, the distance between any pair of leaf nodes scales in logarithmic order in TTNs while in linear order in MPSs. As connected correlation functions generally decay exponentially with path length within a tensor network, TTNs are better suited to capture longer-range correlations and represent two-dimensional objects than MPSs. TTNs are utilized in various fields such as chemical problems (Murg et al., 2010; Gunst et al., 2019), condensed matter physics (Nagaj et al., 2008; Tagliacozzo et al., 2009; Lin et al., 2017) and machine learning (Liu et al., 2019; Cheng et al., 2019).

Rudolph et al. have successfully avoided the barren plateau problem by utilizing tensor networks (Rudolph et al., 2023b). They proposed a method of first optimizing tensor networks, then mapping the optimized tensor networks to quantum circuits, and finally executing VQAs. While numeri-

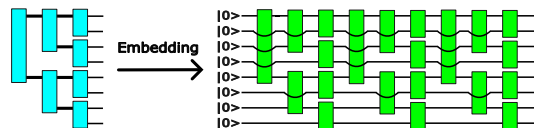


Figure 1: A schematic diagram illustrating the embedding of a TTN into a shallow quantum circuit composed solely of two-qubit gates. The aim is to ensure that the quantum state output by the quantum circuit closely approximates the quantum state represented by the TTN. The diagram depicts a case with three layers, and the approximation accuracy improves as the number of quantum circuit layers increases.

cal results have shown that this method can indeed avoid the barren plateau, there is a general challenge in mapping tensor network states into shallow quantum circuits. The case where the tensor network structure is an MPS has already been well-studied (Ran, 2020; Rudolph et al., 2023a; Shirakawa et al., 2024; Malz et al., 2024). However, effective methods for mapping tensor networks other than MPSs to shallow quantum circuits have not yet been devised.

In this paper, we propose a method to embed TTNs into shallow quantum circuits composed of only two-qubit gates as shown in Figure 1. The primary obstacle to embedding TTNs into shallow quantum circuits has been the complexity of contractions arising from the intricate structure of TTNs. In general, the contraction of tensor networks requires an exponentially large memory footprint relative to the number of qubits if performed naively. Furthermore, determining the optimal contraction order for tensor networks is an NP-complete problem (Chi-Chung et al., 1997). Through the innovative design of a contraction method that balances minimal approximation error and computational efficiency, we successfully extend the embedding method of MPSs (Rudolph et al., 2023a) to TTNs. Additionally, by applying the proposed method to practical problems, we successfully prepare better initial quantum circuits for VQAs than those provided by MPSs with a practical computational complexity. This study is expected to extend the application of VQAs to two-dimensional systems and those with long-range correlations, which have previously been challenging to utilize.

2 BACKGROUND

2.1 Tensor Networks

Tensor networks are powerful mathematical frameworks that efficiently represent and manipulate high-dimensional data by decomposing tensors into interconnected lower-dimensional components (Bridgeman and Chubb, 2017; Orús, 2014). They have

been widely applied in various fields, including quantum many-body physics for approximating ground states of complex systems (White, 1992; Vidal, 2003; Xiang, 2023) and machine learning for real data (Stoudenmire and Schwab, 2016; Han et al., 2018). Recently, increasing attention has been directed towards their compatibility with quantum computing, as their structure aligns well with quantum circuits and facilitates hybrid quantum-classical algorithms (Markov and Shi, 2008).

The widespread application of tensor networks across various fields can be attributed to their simple and comprehensible notation. Tensor network notation provides a clear and intuitive way to represent complex tensor contractions and operations. In tensor network notation, a rank- r tensor is depicted as a geometric shape with r legs. The geometric shape and the direction of the legs are determined by the properties of the tensor and its indices. When representing quantum states, the direction of the legs often indicates whether the vectors are in the Hilbert space for kets or its dual space. When two tensors share a single leg, the leg is referred to as the bond, and the dimension of that leg is referred to as the bond dimension. By adjusting the bond dimension, the expressive power of the tensor network can be controlled. Reducing the bond dimension to decrease memory requirements is referred to as truncation. One of the most commonly used operations is contraction, which combines multiple tensors into a single tensor. In tensor network diagrams, contraction corresponds to connecting tensors with lines. The contraction of A 's x -th index and B 's y -th index is defined as

$$C_{i_1, \dots, i_{x-1}, i_{x+1}, \dots, i_r, j_1, j_{y-1}, j_{y+1}, \dots, j_s} = \sum_{\alpha} A_{i_1, \dots, i_{x-1}, \alpha, i_{x+1}, \dots, i_r} B_{j_1, \dots, j_{y-1}, \alpha, j_{y+1}, \dots, j_s}, \quad (1)$$

where C is the result of the contraction. Operations such as the inner product of vectors, matrix-vector multiplications, and matrix-matrix multiplications are specific examples of contractions. A contraction network forms a tensor network, allowing for the consideration of various networks depending on the objective.

2.2 MPSs and TTNs

Let $|\psi\rangle$ be a tensor network in the form of either an MPS or a TTN. Number the tensors in $|\psi\rangle$ from left to right for MPS, and in breadth-first search (BFS) order from the root node for TTN, denoting the i -th tensor as $A^{(i)}$ ($i = 1, \dots, N$). By adding a leg with bond dimension one, a two-legged tensor can be converted into a three-legged tensor, thus all tensors in $|\psi\rangle$ can

be considered three-legged. $|\psi\rangle$ is in canonical form if there exists a node $A^{(i)}$, called the canonical center, such that for all other nodes $A^{(i')}$ the following holds

$$\sum_{l,m} A_{l,m,n}^{(i')} A_{l,m,n'}^{(i)*} = I_{n,n'}, \quad (2)$$

where the leg denoted by index n is the unique leg of $A^{(i')}$ pointing towards $A^{(i)}$. A tensor that satisfies this equation is referred to as an isometric tensor. Any $|\psi\rangle$ can be transformed into this canonical form and the position of the canonical center can be freely moved without changing the quantum state.

The canonical form offers numerous advantages. In this paper, the key benefit is that at the canonical center, the local Singular Value Decomposition (SVD) matches the global SVD, allowing for precise bond dimension reduction through truncation while appropriately moving the canonical center. Additionally, in the canonical form, each tensor is an isometry, facilitating easy conversion to unitary form and embedding into quantum circuits. Unless otherwise specified, this paper assumes that any MPS and TTN are converted to the canonical form with $A^{(0)}$ as the canonical center.

2.3 Embedding of MPSs

Although embedding general tensor networks into shallow quantum circuits is challenging, several workable methods have been proposed for MPSs. In this subsection, we overview the technique for seamlessly embedding MPSs into shallow quantum circuits.

Ran introduced a systematic decomposition method for an MPS into several layers of two-qubit gates with a linear next-neighbor topology (Ran, 2020). First, the MPS $|\psi^{(k)}\rangle$ with bond dimension χ is truncated to a bond dimension two MPS $|\psi_{\chi=2}^{(k)}\rangle$.

Next, the isometric tensors in $|\psi_{\chi=2}^{(k)}\rangle$ are converted into unitary tensors. The resulting set of unitary tensors, $L[U]^{(k)}$, is referred to as a layer and can be embedded into a quantum circuit composed of two-qubit gates. Additionally, since

$$|\psi_{\chi=2}^{(k)}\rangle = L[U]^{(k)} |0\rangle \quad (3)$$

and

$$L[U]^{(k)\dagger} |\psi_{\chi=2}^{(k)}\rangle = |0\rangle \quad (4)$$

are hold, $L[U]^{(k)\dagger}$ can be considered a disentangler, transforming the quantum state into a product state. Finally, a new MPS $|\psi^{(k+1)}\rangle = L[U]^{(k)\dagger} |\psi^{(k)}\rangle$ can be obtained. Since $L[U]^{(k)\dagger}$ acts as a disentangler, $|\psi^{(k+1)}\rangle$ should have reduced entanglement

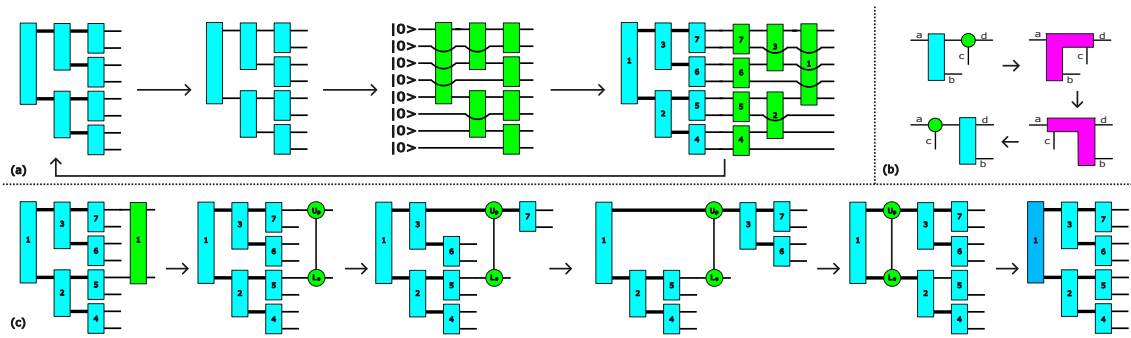


Figure 2: A schematic diagram illustrating the systematic decomposition for a TTN. (a) The TTN is truncated to a bond dimension of two using SVD. This simplified TTN is then embedded into a quantum circuit by converting tensors into unitaries via Gram-Schmidt orthogonalization. These unitaries act as disentglers and are applied to the original TTN. By repeating these steps, a multi-layer quantum circuit is formed. (b) A diagram of a penetration algorithm. It contracts two connected tensors, reorders axes, and separates them using SVD, making it seem like their positions have swapped. (c) A schematic diagram explaining the transformation from a complex tensor network into a TTN. We first use SVD to split it into an upper and a lower tensor. Then, we apply the penetration algorithm iteratively until the upper tensor is connected with the corresponding tensor. This process is repeated for the lower tensor as well. Finally, we contract the upper and lower tensors with the corresponding tensor to form a new tensor in the TTN. This process is performed sequentially from the highest-numbered component of the disentangler.

compared to $|\psi^{(k)}\rangle$. By repeating this process starting from the original MPS $|\psi^{(0)}\rangle$, a quantum circuit $\prod_{k=1}^K L[U^{(k)}] |0\rangle$ with multiple layers can be generated.

Optimization-based methods are employed as an alternative approach to embedding tensor networks into quantum circuits. This method sequentially optimizes the unitary operators within the quantum circuit to maximize the magnitude of the inner product between the quantum circuit and the tensor network. Evenbly and Vidal proposed an iterative optimization technique that utilizes the calculation of environment tensors and SVD (Evenbly and Vidal, 2009). Similarly, Shirakawa et al. employed this iterative optimization method to embed quantum states into quantum circuits (Shirakawa et al., 2024). An environment tensor is calculated to update a unitary by removing the unitary from the circuit and contracting the remaining tensor network. We perform the SVD of the environment tensor and utilize the fact that the resulting unitary matrix serves as the optimal operator to increase the fidelity between the original tensor network and the constructed quantum circuit. For a more detailed explanation of the optimization algorithm, please refer to (Rudolph et al., 2023a; Shirakawa et al., 2024) or the chapter on optimization algorithms for TTNs described later.

Rudolph et al. investigated the optimal sequence for combining these two methods to achieve the highest accuracy (Rudolph et al., 2023a). The study concludes that the best results are obtained by adding new layers through a systematic decomposition step and optimizing all layers with each addition. This

method has achieved exceptional accuracy in embedding a wide range of MPSs in fields such as physics, machine learning, and random systems. Therefore, it can be considered one of the best current techniques for embedding MPSs into shallow circuits. However, this method does not support embedding tensor networks other than MPSs. Consequently, embedding tensor networks such as TTNs and other non-MPS structures remains an unresolved issue.

3 PROPOSED METHOD

We propose a method for embedding TTNs into shallow quantum circuits. Any shape of TTN can be converted into a binary tree using SVD (Hikihara et al., 2023), so we assume a binary tree for simplicity. However, this method can be generalized to embed TTNs of any shape. The fundamental approach of the proposed method is similar to that of MPSs (Rudolph et al., 2023a). By repeatedly adding layers using the systematic decomposition algorithm and optimizing the entire circuit, we generate highly accurate embedded quantum circuits. This section first introduces the systematic decomposition algorithm for TTNs. Due to the increased structural complexity of TTNs compared to MPSs, the contraction methods become non-trivial. We propose a method that balances approximation error and computational cost. Next, we describe the optimization algorithm for TTNs and integrate it with the systematic decomposition. Finally, we discuss the computational cost and demonstrate that embedding TTNs is feasible within

practical computational limits.

3.1 Systematic Decomposition

Our systematic decomposition algorithm for TTNs is an extension of the algorithm for MPSs introduced in (Ran, 2020). We denote the original TTNs as $|\psi_0\rangle$, the k -th layer of the quantum circuit as $L[U]^{(k)}$, the number of layers in resulting quantum circuit as K , and the resulting quantum circuit as $\prod_{k=1}^K L[U]^{(k)} |0\rangle$. Algorithm 1 details the systematic decomposition process for TTNs, as depicted in Figure 2(a).

Input: TTN $|\psi_0\rangle$, Maximum layers K
Output: Quantum Circuit $\prod_{k=1}^K L[U]^{(k)} |0\rangle$
 $|\psi^{(1)}\rangle \leftarrow |\psi_0\rangle$;
for $k = 1$ **to** K **do**
 Truncate $|\psi^{(k)}\rangle$ to $|\psi_{\chi=2}^{(k)}\rangle$ via SVD;
 Convert $|\psi_{\chi=2}^{(k)}\rangle$ to $L[U]^{(k)}$;
 $|\psi^{(k+1)}\rangle \leftarrow L[U]^{(k)\dagger} |\psi^{(k)}\rangle$;
end

Algorithm 1: Systematic decomposition.

In this algorithm, a copy of the original TTN is truncated to a lower dimension using SVD. Truncating bond dimension χ TTNs to bond dimension two TTNs can be done accurately, similar to MPSs, by shifting the canonical center so that the local SVD matches the global SVD. This truncated TTN is then transformed into a single layer of two-qubit gates by converting the isometric tensors in the layer into unitary tensors using the Gram-Schmidt orthogonalization process. The inverse of this layer is applied to the original TTN, resulting in a partially disentangled state with potentially reduced entanglement and bond dimensions. This process can be iteratively repeated to generate multiple layers, which are indexed in reverse order to form a circuit that approximates the target TTN. Notably, the final layer of the disentangling circuit is used as the initial layer of the quantum circuit for approximation.

While this algorithm appears to function effectively at first glance, the computation of $|\psi^{(k+1)}\rangle \leftarrow L[U]^{(k)\dagger} |\psi^{(k)}\rangle$ is exceedingly challenging from the perspective of tensor networks. The structure of $|\psi^{(k+1)}\rangle$ is TTN, whereas tensor network $L[U]^{(k)\dagger} |\psi^{(k)}\rangle$ has a more complex structure, necessitating its transformation into the shape of TTN. As previously mentioned, naively transforming a general tensor network can require exponentially large memory relative to the number of qubits. Additionally, although limiting the bond dimension during the trans-

formation can facilitate the process, the sequence of transformations can lead to substantial approximation errors.

To address the issue of tensor network transformations arising from the complexity of TTN structures, we employ a penetration algorithm as a submodule. As illustrated in Figure 2(b), the penetration algorithm operates by contracting two tensors connected by a single edge along the connected axis, appropriately reordering the axes, and then separating them using SVD. This makes it appear as if the positions of the two tensors have swapped. Additionally, by adjusting the number of singular values retained during the SVD, we can balance approximation accuracy and computational cost.

Input: $L[U]^{(k)\dagger} |\psi^{(k)}\rangle$
Output: TTN $|\psi^{(k+1)}\rangle$
for $i = N - 1$ **to** 1 **do**
 Split $L[U]_i^{(k)\dagger}$ into U_p and L_o via SVD;
 while U_p is not connected to $A^{(i)}$ **do**
 $A \leftarrow U_p$'s left tensor;
 Make U_p penetrate A ;
 end
 while L_o is not connected to $A^{(i)}$ **do**
 $A \leftarrow L_o$'s left tensor;
 Make L_o penetrate A ;
 end
 $A^{(i)} \leftarrow \text{Contract } A^{(i)}, U_p, \text{ and } L_o$;
end

Algorithm 2: Transformation of tensor networks using the penetration algorithm.

Algorithm 2 details the transformation of tensor networks using the penetration algorithm, as depicted in Figure 2(c). We assign numbers to $|\psi^{(k)}\rangle$ in breadth-first search (BFS) order from the root. We denote i -th tensor of $|\psi^{(k)}\rangle$ as $A^{(i)}$. We also assign numbers to $L[U]^{(k)\dagger}$ based on its origin in the TTN, denoting the tensor with number i as $L[U]_i^{(k)\dagger}$. For each tensor $L[U]_i^{(k)\dagger}$ in $L[U]^{(k)\dagger}$, first use SVD to split it into upper tensor U_p and lower tensor L_o to reduce the computational complexity in the penetration algorithm. Then apply the penetration algorithm iteratively until the upper tensor is connected with $A^{(i)}$. Repeat this for the lower tensor too. Finally, contract the upper and lower tensors with $A^{(i)}$ to form a new TTN's i -th tensor. Perform this process sequentially from the highest-numbered tensor in $L[U]^{(k)\dagger}$. This algorithm allows contraction with each tensor in $L[U]^{(k)\dagger}$ without changing the TTN structure of $|\psi^{(k)}\rangle$. Consequently, despite minor approximation errors during penetration, $L[U]^{(k)\dagger} |\psi^{(k)}\rangle$ can be trans-

formed into a TTN structure.

3.2 Integration with Optimization

As demonstrated in (Rudolph et al., 2023a), the systematic decomposition algorithm achieves the highest embedding accuracy when appropriately combined with the optimization algorithm. In this study, we also integrate systematic decomposition with the optimization algorithm. The fundamental concept of the optimization algorithm for quantum circuits embedded with TTNs is analogous to that for quantum circuits embedded with MPS. The primary difference lies in the ansatz of the quantum circuits; however, the optimization algorithm can be executed in a similar manner.

```

Input: Quantum Circuit  $\prod_{k=1}^K L[U]^{(k)} |0\rangle$ ,
        number of sweeps  $T$ , learning rate
         $r \in [0, 1]$ 
Output: Optimized Quantum Circuit
         $\prod_{k=1}^K L[U]^{(k)} |0\rangle$ 
for  $t = 1$  to  $T$  do
    for  $k = 1$  to  $K$  do
        for  $i = 1$  to  $N - 1$  do
             $U_{old} \leftarrow L[U]_i^{(k)}$ ;
            Calculate environment tensor  $E$ ;
            SVD  $E = USV^\dagger$ ;
             $U_{new} \leftarrow UV^\dagger$ ;
             $L[U]_i^{(k)} \leftarrow U_{old}(U_{old}^\dagger U_{new})^r$ ;
        end
    end
end
    
```

Algorithm 2: Optimization.

Algorithm 2 details the optimization process for TTNs, as depicted in Figure 3. The environment tensor E is obtained by contracting all tensors except for the one of interest. In this algorithm, the tensor of interest is $L[U]_i^{(k)}$ and E becomes a four-dimensional tensor with two legs on the left and two on the right. We compute the SVD of E and utilize the fact that the product UV^\dagger is the unitary matrix that maximizes the magnitude of the inner product between the original TTN and the generated quantum circuit (Shirakawa et al., 2024). Given the strength of this local update, we introduce a learning rate r , which modifies the unitary update rule via $U_{old}(U_{old}^\dagger U_{new})^r$. Replacing $L[U]_i^{(k)}$ with the unitary operator calculated in this manner for all operators constitutes one step, and repeating this process for T steps completes the optimization algorithm.

The integration of the systematic decomposition and optimization algorithms involves a method where a new layer is added using the systematic decomposition algorithm, followed by optimizing the entire quantum circuit with the optimization algorithm. This process is repeated iteratively. Rudolph et al. refer to this method as $\text{Iter}[D_i, O_{all}]$, confirming that it achieves the highest accuracy regardless of the type of MPSs (Rudolph et al., 2023a). When creating the $k + 1$ -th layer using the systematic decomposition algorithm, the layers up to k are first absorbed into the original TTN using the penetration algorithm before executing the systematic decomposition algorithm. It should be noted that since we are using the penetration algorithm, the $|\psi\rangle$ that has absorbed $L[U]^{j\dagger}$ retains its TTN structure. The integrated algorithm is presented as Algorithm 3.

```

Input: TTN  $\psi_0$ , Maximum layers  $K$ 
Output: Quantum Circuit  $\prod_{k=1}^K L[U]^{(k)} |0\rangle$ 
 $|\psi\rangle \leftarrow |\psi_0\rangle$ ;
for  $k = 1$  to  $K$  do
    Truncate  $|\psi\rangle$  to  $|\psi_{\chi=2}\rangle$  via SVD;
    Convert  $|\psi_{\chi=2}\rangle$  to  $L[U]^{(k)}$ ;
    Optimize  $\prod_{k'=1}^k L[U]^{(k')}$ ;
    for  $j = 1$  to  $k$  do
        Absorb  $L[U]^{j\dagger}$  into  $|\psi\rangle$ ;
    end
end
    
```

Algorithm 3: Proposed method.

3.3 Computational Complexity

For a TTN with a maximal bond dimension χ , the memory requirements scale $O(N\chi^3)$. However, the computational complexity of transforming the TTN into its canonical form scale $O(N\chi^4)$. Most of the algorithms related to TTNs require transformation into canonical form. Therefore, it is reasonable to assume that we set χ such that computations of $O(N\chi^4)$ can be performed efficiently.

In Algorithm 3, truncating $|\psi\rangle$ to $|\psi_{\chi=2}\rangle$ requires $O(N\chi^3)$. Optimizing the entire circuit is significantly more challenging compared to MPS due to the complexity of the TTN structure. Generally, exponential memory is required with respect to N , but by using an appropriate contraction order and caching, the computation can be performed in $O(N\chi^3 4^K)$, which is linear in N . Although the computational complexity increases exponentially with the number of layers, this algorithm is designed for embedding into shallow quantum circuits, and it operates efficiently for $K \approx 7$, as used in our experiments. Furthermore, when embedding TTNs into a larger number of layers, it is

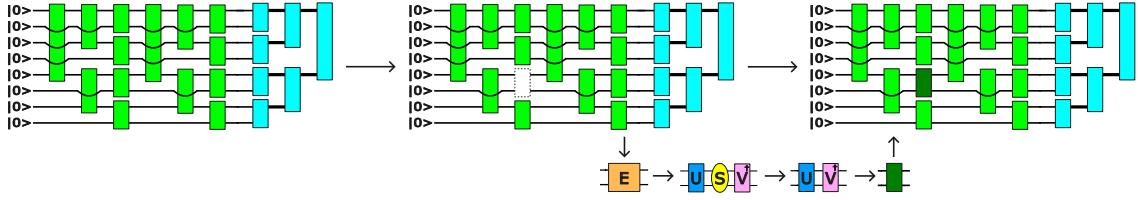


Figure 3: The environment tensor E is obtained by contracting all tensors except the one of interest. We compute its SVD $E = USV^\dagger$ and calculate UV^\dagger to find the unitary operator closest to the environment tensor. The generated unitary operator is positioned at the location of the removed unitary operator.

possible to achieve $O(N \log N \chi^4 K T)$, where T is the number of sweeps, linear computational complexity with respect to the number of layers, bond dimension by introducing approximations that prevent the bond dimension from exceeding χ during the contraction of environment tensors, as used in the embedding of MPS (Rudolph et al., 2023a).

The computational complexity of absorbing K layers into the original TTN is $O(N \log N \chi^4 K)$. In the penetration algorithm, the SVD of a matrix with row dimension χ^2 and column dimension χ has a complexity of $O(\chi^4)$. Given that the depth of the TTN is $O(\log N)$, the penetration operation is performed $O(\log N)$ times per tensor. Since each layer contains $O(N)$ tensors, the complexity of absorbing one layer is $O(N \log N \chi^4)$, leading to a total complexity of $O(N \log N \chi^4 K)$ for K layers. In summary, the computational complexity to generate one layer is $O(\max(N \log N \chi^4 K, N \chi^3 4^K))$, and it is $O(\max(N \log N \chi^4 K^2, N \chi^3 4^K))$ for K layers. It scales with the number of qubits. Additionally, the bond dimension is constrained to χ^4 , making it suitable for embedding TTNs with large bond dimensions. Furthermore, by allowing approximations in optimization, the complexity can be reduced to $O(N \log N \chi^4 K^2 T)$ for K layers.

4 EXPERIMENTS

We conducted experiments using two distinct state vectors from the fields of machine learning and physics. The first state vector represents a uniform superposition over the binary data samples in the 4×4 bars and stripes (BAS) dataset (MacKay, 2003), which has become a canonical benchmark for generative modeling tasks in quantum machine learning. The second state vector represents the ground state of the J_1 - J_2 Heisenberg model, a model that characterizes competing interactions in quantum spin systems. The Hamiltonian for this model is given by the following equation,

$$H = J_1 \sum_{\langle i,j \rangle} \mathbf{S}_i \cdot \mathbf{S}_j + J_2 \sum_{\langle\langle i,j \rangle\rangle} \mathbf{S}_i \cdot \mathbf{S}_j, \quad (5)$$

where J_1 (J_2) represents the nearest (next-nearest) neighbor interactions. By varying the ratio of the first and second nearest-neighbor interactions, the J_1 - J_2 Heisenberg model generates complex quantum many-body phenomena and has been widely studied. In this paper, we utilized $J_2/J_1 = 0.5$. Both state vectors are two-dimensional systems with long-range correlations, making them more suitably represented by TTNs rather than MPSs.

The state vector of the BAS was manually prepared, while the ground state of the J_1 - J_2 Heisenberg model was generated using a numerical solver package $\mathcal{H}\Phi$ (Kawamura et al., 2017), which is designed for a wide range of quantum lattice models. MPSs and TTNs were constructed by iteratively applying SVD to the state vector from the edges.

The conversion from MPSs and TTNs to quantum circuits employed the $\text{Iter}[D_i, O_{all}]$ method, which is also utilized in the proposed method. Additionally, to compare with the proposed method, we also used D_{all} , O_{all} , and $\text{Iter}[I_i, O_{all}]$ from (Rudolph et al., 2023a). The D_{all} method generates circuits solely through systematic decomposition. The O_{all} method optimizes circuits starting from an initial state composed only of identity gates. The $\text{Iter}[I_i, O_{all}]$ method sequentially adds identity layers, optimizing the entire circuit at each step. The number of sweeps in the optimization process was set to 1000. Experiments were conducted using various learning rates ranging from 0.5 to 0.7, and the rate that demonstrated the highest convergence accuracy was selected.

To measure the accuracy of embedding into quantum circuits, we utilized infidelity as the evaluation metric. The infidelity I_f between two quantum states, $|\Psi\rangle$ and $|\Phi\rangle$, is expressed by

$$I_f = 1 - |\langle \Phi | \Psi \rangle|, \quad (6)$$

and a smaller infidelity indicates that the two quantum states are closer. In this study, infidelity quantifies the success of our transformations.

Figure 4 illustrates the infidelity between the original state vector and the quantum circuits embedded with either TTN or MPS. Despite the greater difficulty in embedding TTNs into shallow quantum circuits due to their hierarchical structure, the infidelities

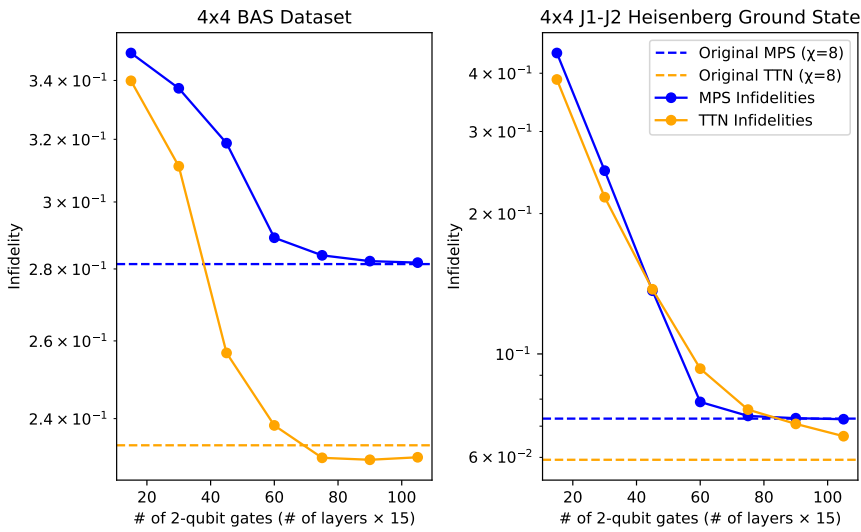


Figure 4: Infidelties between original state vectors and generated quantum circuits. The bond dimensions of tensor networks are eight and the number of optimization steps is 1000. The learning rate for the BAS Dataset and 0.6 for the J_1 - J_2 Heisenberg model. MPS is represented by the blue line, while TTN is depicted by the orange line. Additionally, the infidelity between the tensor network and the state vector is indicated by the dotted line.

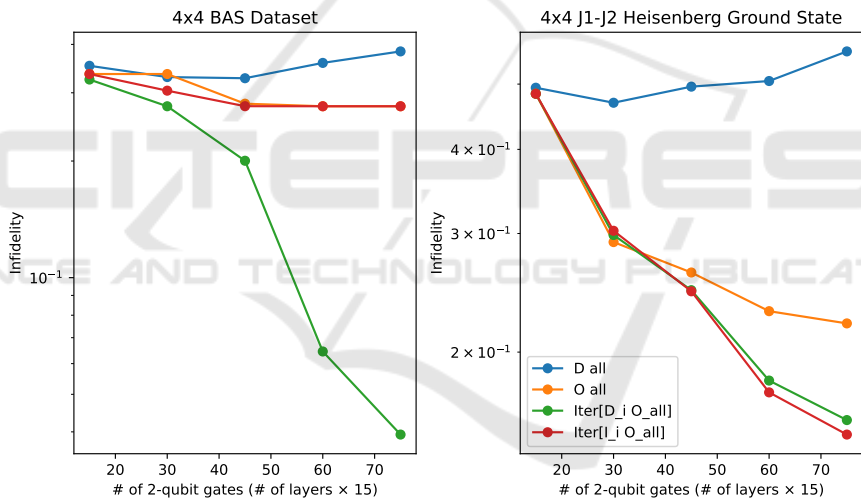


Figure 5: Infidelties between original TTNs and generated quantum circuits. The bond dimensions of TTNs are 16 and the number of optimization steps is 1000. The learning rate for optimization is set to 0.6.

between original state vectors and generated quantum circuits from TTNs are sufficiently low, indicating the successful embedding of the TTNs into shallow quantum circuits. Notably, in the BAS dataset, the TTN’s superior representational capacity enables more accurate embeddings than MPS. The result of the J_1 - J_2 Heisenberg model reveals an intriguing pattern: TTN outperforms in both shallow and deep quantum circuits, while MPS excels in the intermediate region. The superior performance of TTN in shallow circuits is due to the minimal loss from the penetration algorithm, enabling high-precision systematic decomposition. In deep circuits, TTN’s higher representational

capacity leads to better convergence accuracy, as indicated by the dotted lines. Thus, the J_1 - J_2 model results enhance our understanding of the differences between MPS and TTN embeddings.

Figure 5 highlights the importance of the systematic decomposition algorithm for embedding TTNs into quantum circuits. As demonstrated in previous research on MPS (Rudolph et al., 2023a), methods such as $\text{Iter}[D_i, O_{all}]$ and $\text{Iter}[I_i, O_{all}]$, which sequentially add layers and optimize all circuits, achieve high accuracy. Furthermore, as suggested in previous studies, it was confirmed that $\text{Iter}[I_i, O_{all}]$ struggles with the BAS dataset even in TTN embeddings, where

the singular values at the bond become discontinuous. This indicates that $\text{Iter}[D_i, O_{all}]$ is the most effective embedding method, achieving high accuracy in multiple fields, including machine learning and physics, thus affirming the critical importance of the systematic decomposition algorithm. This result also verifies the effective performance of the proposed systematic decomposition method.

5 CONCLUSION

To avoid the barren plateau issue in VQAs, there is increasing interest in using tensor networks to initialize quantum circuits. However, embedding tensor networks into shallow quantum circuits is generally difficult, and prior research has been limited to embedding MPSs. In this study, we propose a method for embedding TTNs, which have a more complex structure than MPSs and can efficiently represent two-dimensional systems and systems with long-range correlations, into shallow quantum circuits composed solely of two-qubit gates. We applied our proposed method to various types of TTNs and confirmed that it prepares quantum circuit parameters with better accuracy than embedding MPSs. Additionally, the computational complexity is $O(\max(N \log N \chi^4 K^2, N \chi^3 4^K))$, or $O(N \log N \chi^4 K^2 T)$ with approximation, making it applicable to practical problems. This study will serve as an important bridge for implementing hybrid algorithms combining tree tensor networks and quantum computing.

ACKNOWLEDGEMENTS

This work was supported by the Center of Innovation for Sustainable Quantum AI, JST Grant Number JP-MJPF2221, and by Japan Society for the Promotion of Science KAKENHI, Grant Numbers 22K18682 and 23H03818. We acknowledge the use of Copilot (Microsoft, <https://copilot.microsoft.com/>) for the translation and proofreading of certain sentences within our paper.

REFERENCES

- Abrams, D. S. and Lloyd, S. (1999). Quantum algorithm providing exponential speed increase for finding eigenvalues and eigenvectors. *Physical Review Letters*, 83(24):5162.
- Arrasmith, A., Cerezo, M., Czarnik, P., Cincio, L., and Coles, P. J. (2021). Effect of barren plateaus on gradient-free optimization. *Quantum*, 5:558.
- Biamonte, J., Wittek, P., Pancotti, N., Rebentrost, P., Wiebe, N., and Lloyd, S. (2017). Quantum machine learning. *Nature*, 549(7671):195–202.
- Bridgeman, J. C. and Chubb, C. T. (2017). Hand-waving and interpretive dance: an introductory course on tensor networks. *Journal of Physics A: Mathematical and Theoretical*, 50(22):223001.
- Cerezo, M., Arrasmith, A., Babbush, R., Benjamin, S. C., Endo, S., Fujii, K., McClean, J. R., Mitarai, K., Yuan, X., Cincio, L., et al. (2021a). Variational quantum algorithms. *Nature Reviews Physics*, 3(9):625–644.
- Cerezo, M. and Coles, P. J. (2021). Higher order derivatives of quantum neural networks with barren plateaus. *Quantum Science and Technology*, 6(3):035006.
- Cerezo, M., Sone, A., Volkoff, T., Cincio, L., and Coles, P. J. (2021b). Cost function dependent barren plateaus in shallow parametrized quantum circuits. *Nature Communications*, 12(1):1791.
- Cheng, S., Wang, L., Xiang, T., and Zhang, P. (2019). Tree tensor networks for generative modeling. *Physical Review B*, 99(15):155131.
- Chi-Chung, L., Sadayappan, P., and Wenger, R. (1997). On optimizing a class of multi-dimensional loops with reduction for parallel execution. *Parallel Processing Letters*, 7(02):157–168.
- Evenbly, G. and Vidal, G. (2009). Algorithms for entanglement renormalization. *Physical Review B*, 79(14):144108.
- Farhi, E., Goldstone, J., and Gutmann, S. (2014). A quantum approximate optimization algorithm. *arXiv preprint arXiv:1411.4028*.
- Friedrich, L. and Maziero, J. (2022). Avoiding barren plateaus with classical deep neural networks. *Physical Review A*, 106(4):042433.
- Gallego, A. J. and Orus, R. (2022). Language design as information renormalization. *SN Computer Science*, 3(2):140.
- Grant, E., Wossnig, L., Ostaszewski, M., and Benedetti, M. (2019). An initialization strategy for addressing barren plateaus in parametrized quantum circuits. *Quantum*, 3:214.
- Grover, L. K. (1996). A fast quantum mechanical algorithm for database search. In *Proceedings of the twentieth annual ACM symposium on Theory of computing*, pages 212–219.
- Gunst, K., Verstraete, F., and Van Neck, D. (2019). Three-legged tree tensor networks with SU(2) and molecular point group symmetry. *Journal of Chemical Theory and Computation*, 15(5):2996–3007.
- Han, Z.-Y., Wang, J., Fan, H., Wang, L., and Zhang, P. (2018). Unsupervised generative modeling using matrix product states. *Physical Review X*, 8(3):031012.
- Hikihara, T., Ueda, H., Okunishi, K., Harada, K., and Nishino, T. (2023). Automatic structural optimization of tree tensor networks. *Physical Review Research*, 5:013031.

- Holmes, Z., Arrasmith, A., Yan, B., Coles, P. J., Albrecht, A., and Sornborger, A. T. (2021). Barren plateaus preclude learning scramblers. *Physical Review Letters*, 126(19):190501.
- Holmes, Z., Sharma, K., Cerezo, M., and Coles, P. J. (2022). Connecting ansatz expressibility to gradient magnitudes and barren plateaus. *PRX Quantum*, 3(1):010313.
- Kawamura, M., Yoshimi, K., Misawa, T., Yamaji, Y., Todo, S., and Kawashima, N. (2017). Quantum lattice model solver $H\Phi$. *Computer Physics Communications*, 217:180–192.
- Levine, Y., Sharir, O., Cohen, N., and Shashua, A. (2019). Quantum entanglement in deep learning architectures. *Physical Review Letters*, 122(6):065301.
- Lin, Y.-P., Kao, Y.-J., Chen, P., and Lin, Y.-C. (2017). Griffiths singularities in the random quantum Ising antiferromagnet: A tree tensor network renormalization group study. *Physical Review B*, 96(6):064427.
- Liu, D., Ran, S.-J., Wittek, P., Peng, C., García, R. B., Su, G., and Lewenstein, M. (2019). Machine learning by unitary tensor network of hierarchical tree structure. *New Journal of Physics*, 21(7):073059.
- Liu, Y., Liu, X., Li, F., Fu, H., Yang, Y., Song, J., Zhao, P., Wang, Z., Peng, D., Chen, H., et al. (2021). Closing the “quantum supremacy” gap: achieving real-time simulation of a random quantum circuit using a new sunway supercomputer. In *Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis*, pages 1–12.
- MacKay, D. J. (2003). *Information Theory, inference and learning algorithms*. Cambridge university press.
- Malz, D., Styliaris, G., Wei, Z.-Y., and Cirac, J. I. (2024). Preparation of matrix product states with log-depth quantum circuits. *Physical Review Letters*, 132(4):040404.
- Markov, I. L. and Shi, Y. (2008). Simulating quantum computation by contracting tensor networks. *SIAM Journal on Computing*, 38(3):963–981.
- McClellan, J. R., Boixo, S., Smelyanskiy, V. N., Babbush, R., and Neven, H. (2018). Barren plateaus in quantum neural network training landscapes. *Nature Communications*, 9(1):4812.
- Mitarai, K., Negoro, M., Kitagawa, M., and Fujii, K. (2018). Quantum circuit learning. *Physical Review A*, 98(3):032309.
- Murg, V., Verstraete, F., Legeza, Ö., and Noack, R. M. (2010). Simulating strongly correlated quantum systems with tree tensor networks. *Physical Review B*, 82(20):205105.
- Nagaj, D., Farhi, E., Goldstone, J., Shor, P., and Sylvester, I. (2008). Quantum transverse-field Ising model on an infinite tree from matrix product states. *Physical Review B*, 77(21):214431.
- Ortiz Marrero, C., Kieferová, M., and Wiebe, N. (2021). Entanglement-induced barren plateaus. *PRX Quantum*, 2(4):040316.
- Orús, R. (2014). A practical introduction to tensor networks: Matrix product states and projected entangled pair states. *Annals of physics*, 349:117–158.
- Peruzzo, A., McClean, J., Shadbolt, P., Yung, M.-H., Zhou, X.-Q., Love, P. J., Aspuru-Guzik, A., and O’Brien, J. L. (2005). Simulated quantum computation of molecular energies. *Science*, 309(5741):1704–1707.
- Peruzzo, A., McClean, J., Shadbolt, P., Yung, M.-H., Zhou, X.-Q., Love, P. J., Aspuru-Guzik, A., and O’Brien, J. L. (2014). A variational eigenvalue solver on a photonic quantum processor. *Nature Communications*, 5(1):4213.
- Preskill, J. (2018). Quantum computing in the NISQ era and beyond. *Quantum*, 2:79.
- Ran, S.-J. (2020). Encoding of matrix product states into quantum circuits of one- and two-qubit gates. *Physical Review A*, 101(3):032310.
- Rudolph, M. S., Chen, J., Miller, J., Acharya, A., and Perdomo-Ortiz, A. (2023a). Decomposition of matrix product states into shallow quantum circuits. *Quantum Science and Technology*, 9(1):015012.
- Rudolph, M. S., Miller, J., Motlagh, D., Chen, J., Acharya, A., and Perdomo-Ortiz, A. (2023b). Synergistic pre-training of parametrized quantum circuits via tensor networks. *Nature Communications*, 14(1):8367.
- Schuld, M. and Killoran, N. (2019). Quantum machine learning in feature Hilbert spaces. *Physical Review Letters*, 122(4):040504.
- Sharma, K., Cerezo, M., Cincio, L., and Coles, P. J. (2022). Trainability of dissipative perceptron-based quantum neural networks. *Physical Review Letters*, 128(18):180505.
- Shi, Y.-Y., Duan, L.-M., and Vidal, G. (2006). Classical simulation of quantum many-body systems with a tree tensor network. *Physical Review A*, 74(2):022320.
- Shirakawa, T., Ueda, H., and Yunoki, S. (2024). Automatic quantum circuit encoding of a given arbitrary quantum state. *Physical Review Research*, 6(4):043008.
- Shor, P. W. (1994). Algorithms for quantum computation: discrete logarithms and factoring. In *Proceedings 35th annual symposium on foundations of computer science*, pages 124–134. Ieee.
- Silvi, P., Giovannetti, V., Montangero, S., Rizzi, M., Cirac, J. I., and Fazio, R. (2010). Homogeneous binary trees as ground states of quantum critical Hamiltonians. *Physical Review A*, 81(6):062335.
- Stoudenmire, E. and Schwab, D. J. (2016). Supervised learning with tensor networks. *Advances in neural information processing systems*, 29.
- Tagliacozzo, L., Evenbly, G., and Vidal, G. (2009). Simulation of two-dimensional quantum systems using a tree tensor network that exploits the entropic area law. *Physical Review B*, 80(23):235127.
- Vidal, G. (2003). Efficient classical simulation of slightly entangled quantum computations. *Physical Review Letters*, 91(14):147902.
- Wang, Z., Hadfield, S., Jiang, Z., and Rieffel, E. G. (2018). Quantum approximate optimization algorithm for maxcut: A fermionic view. *Physical Review A*, 97(2):022304.
- White, S. R. (1992). Density matrix formulation for quantum renormalization groups. *Physical Review Letters*, 69(19):2863.

- Wiersema, R., Zhou, C., de Sereville, Y., Carrasquilla, J. F., Kim, Y. B., and Yuen, H. (2020). Exploring entanglement and optimization within the Hamiltonian variational ansatz. *PRX quantum*, 1(2):020319.
- Xiang, T. (2023). *Density Matrix and Tensor Network Renormalization*. Cambridge University Press.
- Zhou, L., Wang, S.-T., Choi, S., Pichler, H., and Lukin, M. D. (2020). Quantum approximate optimization algorithm: Performance, mechanism, and implementation on near-term devices. *Physical Review X*, 10(2):021067.

