

Simulating noisy quantum circuits with matrix product density operators

Song Cheng^{1,2}, Chenfeng Cao^{1,3}, Chao Zhang^{1,2}, Yongxiang Liu², Shi-Yao Hou^{1,4,2}, Pengxiang Xu², and Bei Zeng^{3,*}

¹Yanqi Lake Beijing Institute of Mathematical Sciences and Applications, Beijing 100407, China

²Center for Quantum Computing, Peng Cheng Laboratory, Shenzhen 518055, China

³Department of Physics, The Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong, China

⁴College of Physics and Electronic Engineering, Center for Computational Sciences, Sichuan Normal University, Chengdu 610068, China

(Received 7 July 2020; accepted 1 March 2021; published 1 April 2021)

Simulating quantum circuits with classical computers requires resources growing exponentially in terms of system size. Real quantum computer with noise, however, may be simulated polynomially with various methods considering different noise models. In this work, we simulate random quantum circuits in one dimension with matrix product density operators (MPDOs), for different noise models such as dephasing, depolarizing, and amplitude damping. We show that the method based on matrix product states (MPSs) fails to approximate the noisy output quantum states for any of the noise models considered, while the MPDO method approximates them well. Compared with the method of matrix product operators (MPOs), the MPDO method reflects a clear physical picture of noise (with inner indices taking care of the noise simulation) and quantum entanglement (with bond indices taking care of two-qubit gate simulation). Consequently, in case of weak system noise, the resource cost of the MPDO will be significantly less than that of the MPO due to a relatively small inner dimension needed for the simulation. In case of strong system noise, a relatively small bond dimension may be sufficient to simulate the noisy circuits, indicating a regime that the noise is large enough for an “easy” classical simulation, which is further supported by a comparison with the experimental results on an IBM cloud device. Moreover, we propose a more effective tensor updates scheme with optimal truncations for both the inner and the bond dimensions, performed after each layer of the circuit, which enjoys a canonical form of the MPDO for improving simulation accuracy. With truncated inner dimension to a maximum value κ and bond dimension to a maximum value χ , the cost of our simulation scales as $\sim ND\kappa^3\chi^3$, for an N -qubit circuit with depth D .

DOI: 10.1103/PhysRevResearch.3.023005

I. INTRODUCTION

The quantum computer has the potential to outperform the best possible classical computers in many tasks such as factoring large numbers. It relies on the fact that wave functions represent amplitudes that grow exponentially in terms of the system size [1]. At the heart is quantum coherence, which is fragile and easily destroyed by noise. In principle, this drawback may be overcome by the techniques of quantum error correction and fault tolerance, which, however, require tens of thousands of qubits to perform computing tasks of practical relevance [2,3]. For near-term hardware systems, precision needs to improve while system size grows, to be able to perform reasonable computing tasks before decoherence, whose performance may be measured by the so-called quantum volume [4]. It is claimed that systems with quantum

volume as large as 32 have been achieved [5], and systems of quantum volume 64 is on the way [6].

Real world quantum computers battling noise recently achieved the so-called quantum supremacy, at Google, for implementing random quantum circuits in a 53-qubit system and a circuit depth of 20 [7]. It is well known that noiseless random circuits are hard to simulate on classical computers [8–13], and simulations of (noiseless) random quantum circuits on supercomputers have been implemented for more than 40 qubits [14–17]. It still remains unclear, however, whether there are classical algorithms running on available supercomputers that may be able to simulate the behavior of the Google system, due to physical noise that results in low fidelity compared to noiseless systems. For instance, a method based on second storage has been proposed, which may simulate the system in a few days [18].

It is recently proposed in [19] that a method based on matrix product states (MPSs) (one of the one-spacial-dimensional tensor network states) can approximate the behavior of real quantum systems. The MPS has been a powerful method that faithfully represents ground states of local Hamiltonians [20–22]. The singular value decomposition (SVD) method for truncating the bond dimension for the MPS has shown great success for finding ground states of spacial-one-dimensional (1D) local Hamiltonians, both gapped and

*zengb@ust.hk

gapless. It is unclear, however, what the error model is that the MPS method represents, for simulating circuit output distribution of real quantum computers.

Since MPSs cannot represent mixed states of quantum systems, a natural idea is instead to use the matrix product operators (MPOs) [20,23,24]. The MPO method has been used to simulate quantum circuits of Shor's and Grover's algorithm with noise [25]. Very recently, the MPO method has also been used to simulate 1D random circuits [26]. For simulating two-qubit gates, the MPO tensors with a “canonical” update bring a factor of DN^2 in simulating an N -qubit random circuit with depth D in terms of complexity.

In this work, we simulate noisy 1D random quantum circuits with matrix product density operators (MPDOs), based on the MPDO construction proposed in [24]. Another well-known name of the MPDO is the MPS purification or the local purification form of the mixed-state density matrix. Recently, it is also shown that MPDOs can describe many-body thermal states and open quantum many-body system efficiently [27,28].

Compared with the MPO method, the inner indices in the MPDO method capture the classical information of the noise simulation, which also reduces the computational and memory complexity under the condition of weak noise. The MPDO model consists of two parts that are conjugated to each other. By so, the simulation of the two-qubit gates can be done in a similar way as the MPS simulation, which is taken care of by the bond indices. In case of weak system noise, a small inner dimension may be sufficient for the simulation, so the resource cost of MPDO could be significantly less than that of the MPO. In case of strong system noise, a relatively small bond dimension may be sufficient to simulate the noisy circuits, indicating a regime that the noise is large enough for an “easy” classical simulation.

Moreover, we propose a more effective canonical tensor update scheme, performed after each layer of the circuit, which would truncate the inner dimension to some maximum value κ and the bond dimension to some maximum value χ with a canonicalization of the MPDO for improving simulation accuracy. The complexity of this scheme is only proportional to DN for an N -qubit circuit with depth D . The cost of our entire simulation scales as $\sim ND\kappa^3\chi^3$.

We apply our method to simulate the random quantum circuit with different noise models, including the dephasing noise, the depolarizing noise, and the amplitude damping noise. We demonstrate that the MPDO approximates the noisy output quantum states well, while the method based on matrix product states fails to approximate the noisy output quantum states for any of the noise models considered. This indicates that the bond dimension truncation method of the MPS simulation might not represent any local noise model in real physical systems. With a further look into the deviation from the Porter-Thomas distribution for the ideal random circuit case, relatively small bond dimension for the MPDO method already grasps some “qualitative behavior” of the noisy output distribution. To test our method with a real quantum computer, we run random circuits on an IBM 16-qubit device. The comparison between experimental data and the simulation based on the MPDO method demonstrates that relatively small χ and κ can indeed simulate the noisy random circuit efficiently.

We organize our paper as follows. In Sec. II, we discuss the error models we use for our circuit simulation. In Sec. III, we discuss the MPDO method for simulating noisy quantum circuits and its complexity. In Sec. IV, we present our results based on the MPDO method, and compare with an exact noise simulation based on density matrices, and the MPS method based on bond dimension truncation, for different error models. In Sec. V, we study the effect of truncation on the bond and inner dimensions. In Sec. VI, we run several 1D random circuits on 10 qubits of the 16-qubit IBM device `ibmq_16_melbourne`, and compare with our simulation. Summary of the results and discussions on future directions will be given in Sec. VIII.

II. NOISE MODELS

Physical noise \mathcal{E} for quantum systems with a quantum state ρ are generally characterized by the Kraus representation, as given by

$$\mathcal{E}(\rho) = \sum_k E_k \rho E_k^\dagger, \quad (2.1)$$

where E_k 's are the Kraus operators and fulfill $\sum_k E_k^\dagger E_k = I$ [1].

For the quantum circuit of N qubits with single-qubit and two-qubit quantum gates, normally the fidelity of single-qubit gates are much higher than that of the two-qubit gates. We will then assume that all single-qubit gates are ideal, and model the noise only on the two-qubit gate U by

$$\rho \rightarrow \sum_k U E_k \rho E_k^\dagger U^\dagger, \quad (2.2)$$

where E_k 's acting on the same qubits of U .

Alternatively, for the noisy channel by \mathcal{E} , we can denote the channel corresponding to the two-qubit gate U by \mathcal{U} , hence rewrite Eq. (2.2) as

$$\rho \rightarrow \mathcal{U} \circ \mathcal{E}(\rho). \quad (2.3)$$

In this work, we consider the following noise models.

(i) Dephasing noise.

The dephasing noise on a single qubit can be modeled by

$$\rho \rightarrow \mathcal{E}_{DF}(\rho) = (1 - \epsilon)\rho + \epsilon Z \rho Z^\dagger, \quad (2.4)$$

where $\epsilon \in [0, 1]$, and Z is the Pauli operator.

(ii) Depolarizing noise.

The depolarizing noise on a single-qubit \mathbf{U} can be modeled by

$$\rho \rightarrow \mathcal{E}_{DP}(\rho) = (1 - \epsilon)\mathbf{U} \rho \mathbf{U}^\dagger + \epsilon \frac{I}{2}. \quad (2.5)$$

Notice that, for a quantum circuit with gate noise modeled by depolarizing noise, the density matrix of the output state under will also be given in the form of a global depolarizing noise, i.e.,

$$\rho \rightarrow (1 - \alpha)|\psi\rangle\langle\psi| + \alpha \frac{I}{M}, \quad (2.6)$$

where $|\psi\rangle$ is the corresponding noiseless output state, and $M = 2^N$ with N the number of qubits in the system.

(iii) Amplitude damping noise.

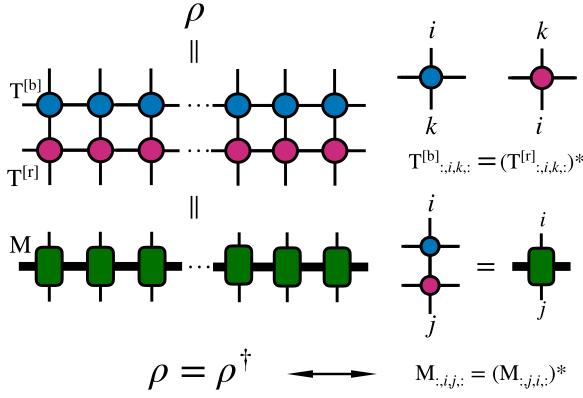


FIG. 1. The structure of the matrix product density operators (MPDO). Instead of directly using a tensor M to form a matrix product operator, we choose M to be composed by the red tensor $T^{[r]}$ and the blue tensor $T^{[b]}$, where $T^{[b]} = (T^{[r]})^\dagger$. By so, the MPDO will automatically satisfy the hermiticity of a density matrix, although T s are general fourth-order tensors. In this way, the parameters of the model can be fully used on the one hand, and the inner indices between the $T^{[b]}$ and $T^{[r]}$ have emerged on the other hand. Those inner indices can be used to describe the classic entropy of the system. At the same time, when the classic entropy is small, the model's computational cost will become significantly smaller.

The amplitude damping noise on a single-qubit gate can be modeled by

$$\rho \rightarrow \mathcal{E}_{AD}(\rho) = A_0 \rho A_0^\dagger + A_1 \rho A_1^\dagger, \quad (2.7)$$

where $A_0 = \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{1-\epsilon} \end{pmatrix}$, $A_1 = \begin{pmatrix} 0 & \sqrt{\epsilon} \\ 0 & 0 \end{pmatrix}$.

For a noisy two-qubit gate U , we could simply model the above three noises by

$$\rho \rightarrow \mathcal{U} \circ \mathcal{E}^{\otimes 2}(\rho). \quad (2.8)$$

III. MODELING NOISE SIMULATION BY MATRIX PRODUCT DENSITY OPERATORS

By applying MPOs instead of MPSs, one can extended the error model to represent mixed quantum states, which allows us to introduce typical random noise in a more direct and efficient way. As shown in Fig. 1, we use the MPO to represent the density matrix ρ of N qubits,

$$\rho = \sum_{s_1, s'_1, \dots, s_N, s'_N=1}^2 M_{L_1, R_1}^{s_1, s'_1} I^{R_1, L_2} M_{L_2, R_2}^{s_2, s'_2} \dots I^{R_{N-1}, L_N} M_{L_N, R_N}^{s_N, s'_N} \times |s_1, \dots, s_N\rangle\langle s'_1, \dots, s'_N|, \quad (3.1)$$

where I is the identity matrix connecting the left and right indices, and ρ denotes the density matrix of a mixed quantum state. s_k and s'_k are known as the “physical” indices, which expand into the $2^N \times 2^N$ density matrix ρ . L_k and R_k denote the left and right “bond” indices, which carry some information of entanglement between qubits.

However, there are at least two reasons that directly using the tensor M to form an MPO representation of the density matrix may not be the best choice. First, since the density matrix is a Hermitian matrix, we must guarantee that $M = M^\dagger$, which will make at least half of the M parameters invalid,

resulting in additional computational overhead. Second, considering the density matrix $\rho = \sum_k \lambda_k |\phi_k\rangle\langle\phi_k|$, when k is not too many, the density matrix can be regarded as the sum of the direct products of several state vectors. In this case, it is very inefficient to use tensor networks to model the direct product of state vectors instead of the sum of several states itself.

Therefore, a proper way to ensure the hermiticity of density matrix ρ is the MPDO method [24], which designs the $M_{L_k, R_k}^{s_k, s'_k}$ to be composed by a general fourth-order tensor $T_{l_k, r_k}^{s_k, a_k}$ and its conjugate copy $(T_{l'_k, r'_k}^{s'_k, a_k})^*$,

$$M_{L_k, R_k}^{s_k, s'_k} = \sum_{a_k=1}^{d_k} T_{l_k, r_k}^{s_k, a_k} \times (T_{l'_k, r'_k}^{s'_k, a_k})^*. \quad (3.2)$$

In others words, the MPDO is composed of a general MPO and its conjugation. The indices L_k and R_k of M correspond to the direct product of indices $l_k \otimes l'_k$ and $r_k \otimes r'_k$, respectively. In practice, the dimension of l_k and r_k bond indices D_k are restricted to a certain maximum value χ .

The two conjugate MPOs were connected by the “inner” indices a_k , which could carry the classical information of the system. This “inner” dimension will increase by adding statistical noise. The dimension of the inner indices d_k would be no more than $2D_k D_{k+1}$. While in some cases, we still truncate the d_k to a smaller number $\kappa < 2D_k D_{k+1}$. The memory cost of the MPDO would only be proportional to $2N\kappa\chi^2$. Note that if we use the M matrix directly, the memory cost will be $4N\chi^4$, which means that if κ reaches the upper bound, those two models cost almost the same memory. While if the system noise is small, the dimensions of the inner indices will be fixed at a smaller value, and the model will consume significantly fewer resources than using M directly.

To be more intuitive, here we give two particular examples of the MPDO. The first example is the density matrix of a pure quantum state $\rho_{\text{pure}} = |\mathbf{s}\rangle\langle\mathbf{s}|$, where $s_k \in \{0, 1\}$. The corresponding $T^{(k)}$ is $\{T_{l_k, r_k}^{s_k, a_k} : T_{0,0}^{s_k, 0} = 1, \text{others} : 0\}$. For the density matrix of a maximum mixed state $\rho_{\text{mix}} = (\frac{1}{2}|0\rangle\langle 0| + \frac{1}{2}|1\rangle\langle 1|)^{\otimes N}$, the $T^{(k)}$ can be written as $\{T_{l_k, r_k}^{s_k, a_k} : T_{0,0}^{0,0} = T_{0,0}^{1,1} = \frac{1}{\sqrt{2}}, \text{others} : 0\}$.

Applying gates and noise on the MPDO is straightforward. Considering the symmetry of the density matrix, one only needs to apply the gate to half of the density matrix, which is the MPO; the rest automatically becomes the conjugation of updated tensors. It is worth mentioning that the identity I of the depolarizing noise in (A1) can also be decomposed into two conjugate parts using $2I = \rho + \sum_{i=x,y,z} \sigma_i \rho \sigma_i^\dagger$, where σ_i are Pauli matrices. As Fig. 2(a) shown, the one-qubit gate wouldn't change any topology or dimensions of the MPDO, we could do it exactly with complexity $\sim \kappa\chi^2$.

A two-qubit gate would be represented as a fourth-order tensor U . As shown in Fig. 2(b), we first contract the gate with the corresponding qubits, which form a sixth-order tensor W . It can be separated into two new MPO tensors by the singular value decomposition. In general cases, applying a two-qubit gate would increase the dimension of bond indices between the two qubits to $\min(2D_k d_k, 2D_{k+2} d_{k+2})$. If we truncate it to χ , the approximation error is $\sim \sum_{j=\chi+1} \lambda_j^2 / \sum_{i=1} \lambda_i^2$, where λ_i are the singular values in descending order. More details

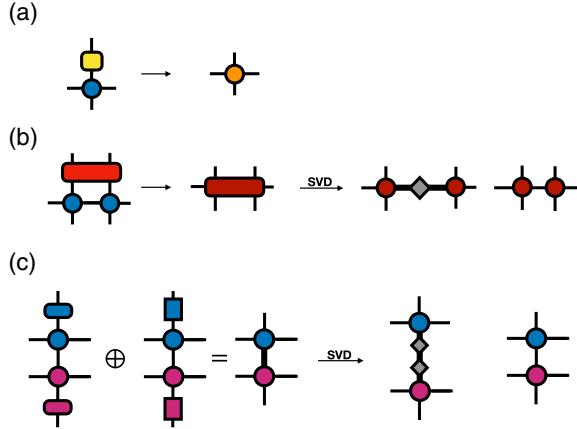


FIG. 2. (a) Applying a single-qubit gate on the corresponding tensor of the MPDO. (b) Applying a two-qubit gate on two corresponding tensors of the MPDO. The new tensors would be obtained by the singular value decomposition (SVD). The SVD would increase the bond dimension between those two tensors, while we could find the global optimal truncation of tensors based on the singular value of SVD. (c) Applying noise models on the corresponding tensor is equivalent to applying each noise operator on tensor separately and then direct sum them on the inner indices. The increased dimensions of inner indices can also be truncated using SVD.

about the strategy of this truncation are placed at the end of this section. The computation cost is $\sim \kappa^2 \chi^3$ for contraction, and $\sim \kappa^3 \chi^3$ for SVD.

Compared with the MPS method as discussed in [19], the MPDO has a clear advantage in adding noise. Unlike the state representation, the density matrix can directly express noise as the Kraus representation in (2.1), which avoids repeated Monte Carlo sampling of different noise. Let's take the amplitude damping noise as an example. What we need to do is to apply the A_0 and A_1 in (2.7) to ρ , respectively, and then add them directly. Same to the gate operators, we only need to apply noise to half of ρ to save the computational cost. Note that the summation of noise and the contraction of the conjugate tensors are not interchangeable. To avoid unnecessary cross-terms, as shown in Fig. 2(c), here we need to direct sum the different noise parts on the inner indices a_k , which writes

$$T_{l_k, r_k}^{s_k, a_k} = \begin{cases} \sum_{s'_k} A_0^{s_k, s'_k} T_{l_k, r_k}^{s'_k, a'_k} & a_k \in [0, d_k) \\ \sum_{s'_k} A_1^{s_k, s'_k} T_{l_k, r_k}^{s'_k, a'_k} & a_k \in [d_k, 2d_k), \end{cases} \quad (3.3)$$

thus,

$$M^{[\text{new}]} = \sum_{a_k} T T^\dagger = A_0 M^{[\text{old}]} A_0^\dagger + A_1 M^{[\text{old}]} A_1^\dagger. \quad (3.4)$$

The computational cost of applying noise is $\sim m\kappa \chi^2$; m is the number of terms of the noise model.

There is an interesting fact behind the structure of the MPDO. Note that the two-qubit gate introduces entanglement entropy into the system, and it only increases the dimension of the bond indices; meanwhile the single-qubit noise introduces classical statistical entropy into the system, and only increases the dimension of the inner indices. Therefore, if there is no

truncation, in the final structure of the MPDO, the dimensions of the bond indices and inner indices will be related to the quantum entanglement entropy and classical statistical entropy of the system, respectively. The exact relationship between bond/inner dimensions and entanglement/classical entropy remains an open question.

In many cases, especially those with significant noise, a small dimension of bond indices is enough to ensure high fidelity. So in order to remove unnecessary parameters, we may truncate the bond indices and inner indices, simultaneously. More specifically, we will first do a local optimal approximation on inner indices by SVD, which separate tensor T into

$$T_{l_k, r_k}^{s_k, a_k} = \sum_{\mu} = U_{l_k, r_k}^{s_k, \mu} S_{\mu} V_{\mu, a_k}, \quad (3.5)$$

where S_{μ} is the singular value of T , U , and V are the unitary matrix. Then we only keep κ largest S_{μ} and corresponding orthogonal vectors in U and V . The approximate M can be written as

$$M \approx \sum_{\mu=1}^{\kappa} U' S^2 U'^\dagger. \quad (3.6)$$

We repeat this process on each of the inner indices. Then, before we truncate bond indices, we first apply QR decomposition from the left of the MPDO to the right to form a canonical form of the MPO, which writes

$$T_{l_k, r_k}^{s_k, a_k} = \sum_{\mu} Q_{l_k, \mu}^{s_k, a_k} R_{\mu, r_k}, \quad T_{l_k, r_k}^{'s_k, a_k} = Q_{l_k, \mu}^{s_k, a_k},$$

$$T_{l_{k+1}, r_{k+1}}^{'s_{k+1}, a_{k+1}} = \sum_{l_{k+1}} R_{\mu, l_{k+1}} T_{l_{k+1}, r_{k+1}}^{s_{k+1}, a_{k+1}}. \quad (3.7)$$

After applying this on all qubits, all tensors except the rightmost one are left canonicalized. In other words, they fulfill

$$\sum_{l_k, s_k, a_k} T_{l_k, r_k}^{s_k, a_k} T_{l_k, r'_k}^{s_k, a_k} = I_{r_k, r'_k}, \quad (3.8)$$

where I denotes the identity matrix. This canonicalization will ensure the SVD truncation on the rightmost tensor is globally optimal. We then use the SVD to truncate each of the bond indices from right to left:

$$\sum_{l_{k+1}} T_{l_k, l_{k+1}}^{s_k, a_k} T_{l_{k+1}, r_{k+1}}^{s_{k+1}, a_{k+1}} \approx \sum_{\mu=1}^{\chi} U_{l_k, \mu}^{s_k, a_k} S_{\mu} V_{\mu, r_{k+1}}^{s_{k+1}, a_{k+1}}. \quad (3.9)$$

Note that each time the SVD changes the right tensor from left canonicalized to right canonicalized, which results in the following SVD truncations are all globally optimal. Therefore, the most economical way is first to complete a layer of two-qubit gates and noise (see Fig. 3 for an illustration of a layer in the dotted line circuit), then to perform a canonicalization from left to right and the following SVD decomposition from right to left. Compared with canonicalizing on each qubit independently, the order of this scheme reduces the complexity of canonical truncation from a factor of N^2 to a factor of N without loss of the accuracy. The total cost of simulating an N -qubit circuit with depth D is $\sim D N \kappa^3 \chi^3$. Note that while the canonical form of the MPO cannot be used to find the global optimal truncation of inner indices, we could still use the full

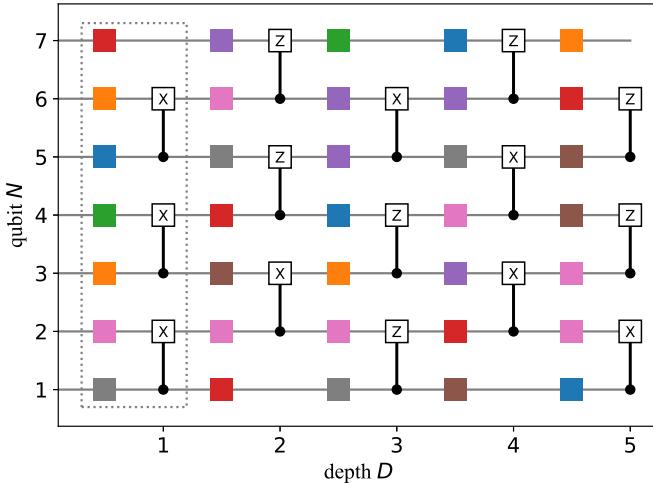


FIG. 3. Sketch of the 1D random circuit with qubits number $N = 7$ and depth $D = 5$. The gray dotted line outlines the structure of one layer. The circuit is interleaved by layers of single-qubit gates (colored boxes) and two-qubit gates (the dots connected to a blank box). The single-qubit gates are randomly sampled from the set of universal single-qubit quantum gates. The two-qubit gates are either control-NOT or control-Z with equal probability.

update method similar to that used in the higher dimensional tensor network to find its global optimal truncation, in case there are some people willing to tolerate excessive calculation costs.

IV. COMPARISON MPDO SIMULATION WITH DIFFERENT MODELS

Our numerical experiments are applied on the 1D random circuit illustrated in Fig. 3. The colored boxes represent various single-qubit gates randomly generated from $e^{i\alpha(\sigma_x \sin \theta \cos \phi + \sigma_y \sin \theta \sin \phi + \sigma_z \cos \theta)}$ with three random parameters $\alpha, \theta, \phi \in [0, 2\pi]$, which traverse the space of universal single-qubit gates. The lines and the blank boxes connecting two qubits represent either control-NOT (CNOT) or control-z gates with equal probability. It is known that such kinds of (pseudo-)random circuits with big enough depth D could yield an approximately Haar-distributed unitary, and generate entanglement efficiently [29–31]. This kind of (pseudo-)random quantum circuit has been discussed extensively for demonstrating quantum supremacy [7,13,32,33].

We perform four different models to simulate this circuit.

(i) A simulator based on the state vector [15,34] for exact noiseless simulation.

(ii) A simulator based on the density matrix [35] for exact noise simulation. The noise would apply to each two-qubit gate.

(iii) An MPS simulator based on approximating a pure state by SVD method, as discussed in Ref. [19].

(iv) An MPDO simulator based on approximating a density matrix by a conjugated tensor network structure and SVD, as discussed in Sec. III. The noise would apply to each two-qubit gate.

TABLE I. Noise rates in different noise models.

Fidelity with ρ_0	MPS bond dim	Dephasing noise rate	Depolarizing noise rate	Amplitude damping noise rate
0.102	2	0.0231	0.0302	0.0454
0.183	3	0.0167	0.0220	0.0332
0.378	4	9.47×10^{-3}	0.0125	0.0188
0.450	5	7.75×10^{-3}	0.0102	0.0155
0.559	6	5.63×10^{-3}	7.45×10^{-3}	0.0113
0.644	7	4.25×10^{-3}	5.64×10^{-3}	8.51×10^{-3}
0.745	9	2.84×10^{-3}	3.76×10^{-3}	5.69×10^{-3}
0.847	12	1.59×10^{-3}	2.12×10^{-3}	3.20×10^{-3}
0.931	15	6.88×10^{-4}	9.14×10^{-4}	1.38×10^{-3}
0.999	28	9.94×10^{-6}	1.33×10^{-5}	2.01×10^{-5}

Various noise models have been considered, including the dephasing, the depolarizing, and the amplitude damping noise model, as discussed in Sec. II.

A. Comparison based on fidelity

We consider the random circuits with 10 qubits and depth $D = 24$. For clarity, we define the following notations.

(i) ρ_0 denotes the output density matrix of the exact noiseless simulator, which corresponds to a pure state representing the exact output the noiseless random circuit.

(ii) ρ_e denotes the output density matrix of the exact noise simulator, which gives the exact result of simulating the circuit with given noise models by state vector technique.

(iii) ρ_s denotes the output density matrix of the MPS simulator, which corresponds to a pure state subject to different truncation up to some maximum bond dimension χ .

(iv) ρ_d denotes the output density matrix of the MPDO simulator, which is subject to different truncation up to some maximum bond dimension χ and maximum inner dimension κ .

The fidelity between two quantum states are given by

$$\mathcal{F}(\rho, \sigma) \equiv \text{Tr} \sqrt{\sqrt{\rho} \sigma \sqrt{\rho}}. \quad (4.1)$$

For comparing the MPS simulator with others, we define a parameter r as

$$r = \mathcal{F}(\rho_0, \rho_s), \quad (4.2)$$

which corresponds to a certain bond dimension truncation χ . For each χ and noise model, we could find a corresponding error rate ϵ in ρ_e that satisfied

$$\mathcal{F}(\rho_0, \rho_e) = r. \quad (4.3)$$

We summarize the values of χ and corresponding ϵ in Table I.

We then use the MPDO simulator to simulate the noisy random quantum circuit for different error models with error rate ϵ as given in Table I. We set max bond dimension $\chi = 32$, and max inner dimension $\kappa = 48$.

For each r and each noise model, we calculate the fidelity of $\mathcal{F}(\rho_e, \rho_s)$, which demonstrates how the MPS method approximates the exact result of the noisy output density matrix, given the same fidelity of $r = \mathcal{F}(\rho_0, \rho_s) = \mathcal{F}(\rho_0, \rho_e)$. We also calculate the fidelity of $\mathcal{F}(\rho_e, \rho_d)$, which demonstrates

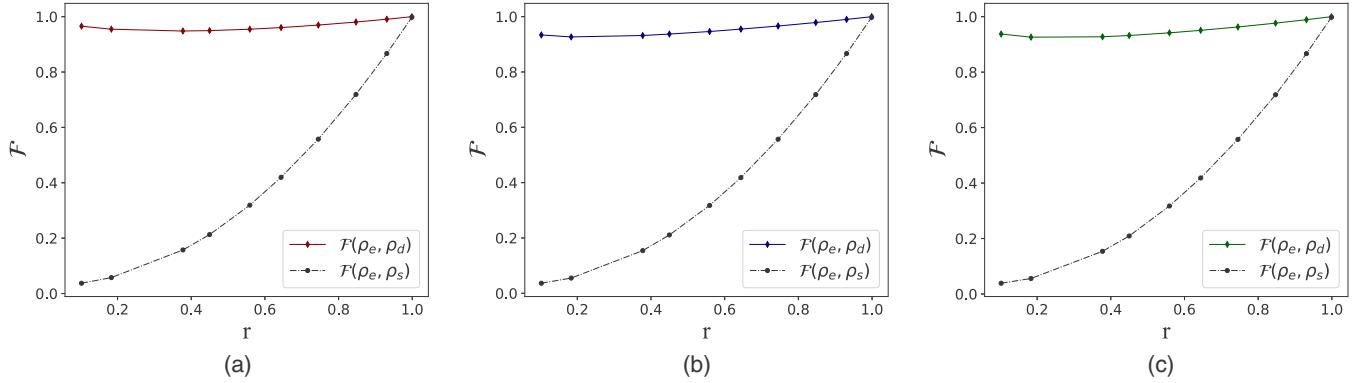


FIG. 4. Fidelity comparison for ρ_0 , ρ_e , ρ_d , and ρ_s , where $r = \mathcal{F}(\rho_0, \rho_s) = \mathcal{F}(\rho_0, \rho_e)$. (a) The dephasing noise; (b) the depolarizing noise; (c) the amplitude damping noise.

how the MPDO method approximates the exact result of the noisy output density matrix. Our results are shown in Fig. 4.

From Fig. 4, it is clearly shown that when noise gradually became significant, the MPS simulator gradually failed to simulate all three noise models considered, even if it gives the fidelity with the exact noiseless state with ρ_e . This indicates that the MPS truncated approximation is not simulating physical noise in real systems. On the other hand, the MPDO method approximates ρ_e well, which can, in fact, simulate any physical noise as given by the MPDO model construction.

B. Deviation from the Porter-Thomas distribution

For general random circuits, when $D \sim N$, the output states would reach to the Porter-Thomas distribution. However, in the real world, a certain physical foundation would cause certain types of noises, which leads to the deviation of output states from the Porter-Thomas distribution. This deviation should be correctly captured by a proper noisy simulator.

In this section, we consider random circuits with 15 qubits and depth $D = 24$. We focus on analyzing how the Porter-

Thomas distribution changes due to the effect of noise, and compare the different method of simulation.

For a density matrix ρ , consider a random variable $p_i = \langle x_i | \rho | x_i \rangle$ for x_i is the i th bit string from $\{0, 1\}^N$, thus $|x_i\rangle$ is one of the computational basis.

If there is no noise, the output pure state $|\psi\rangle$ is resulted from the random circuit $U|\psi\rangle = U|0\rangle$. Then for a pure state, the probability of getting a certain base $|x_i\rangle$ is

$$p_i = |c_i|^2 = |\langle x_i | U | 0 \rangle|^2. \quad (4.4)$$

For a random circuit with sufficiently large depth, the distribution of $\{p = p_i(x)\}$ is known to follow the Porter-Thomas distribution,

$$Pr(p) = (M - 1)(1 - p)^{M-2} = M(e^{-Mp} + O(1)), \quad (4.5)$$

with expectation $1/M$ [32], where $M = 2^N$.

For random circuits with 15 qubits and depth 24, we calculate the cumulated p distribution for different noise models with the following.

- (i) The Porter-Thomas distribution (corresponding to a exact noiseless simulator).
- (ii) Distribution given by an exact noise simulator.
- (iii) Distribution given by an MPS simulator.

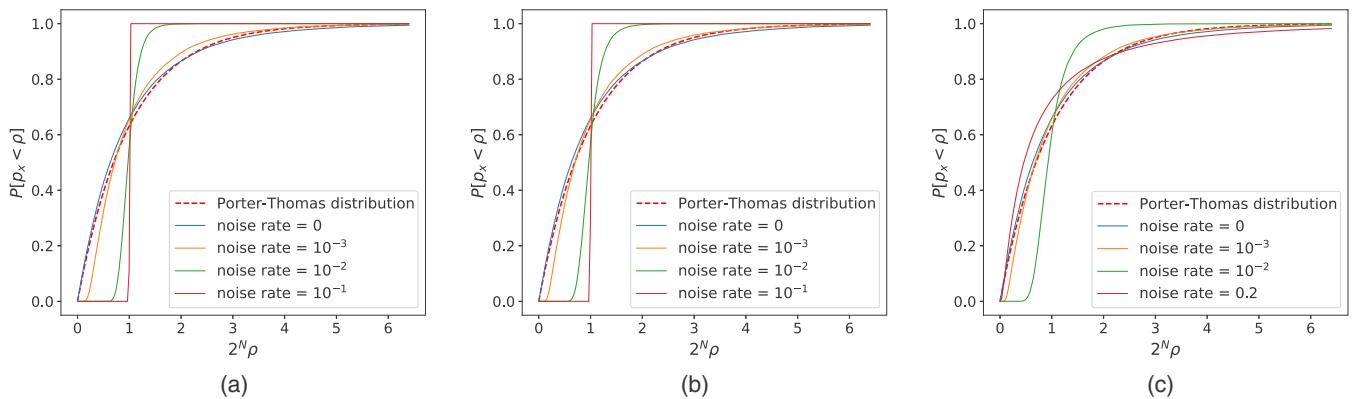


FIG. 5. Exact noisy simulation of cumulated p distributions for 1D random circuit with 15 qubits and 24 layers, under three different error models with different noise rate. (a) The dephasing noise; (b) the depolarizing noise; (c) the amplitude damping noise. As the system noise increases, the cumulated p distribution gradually deviates from the Porter-Thomas distribution.

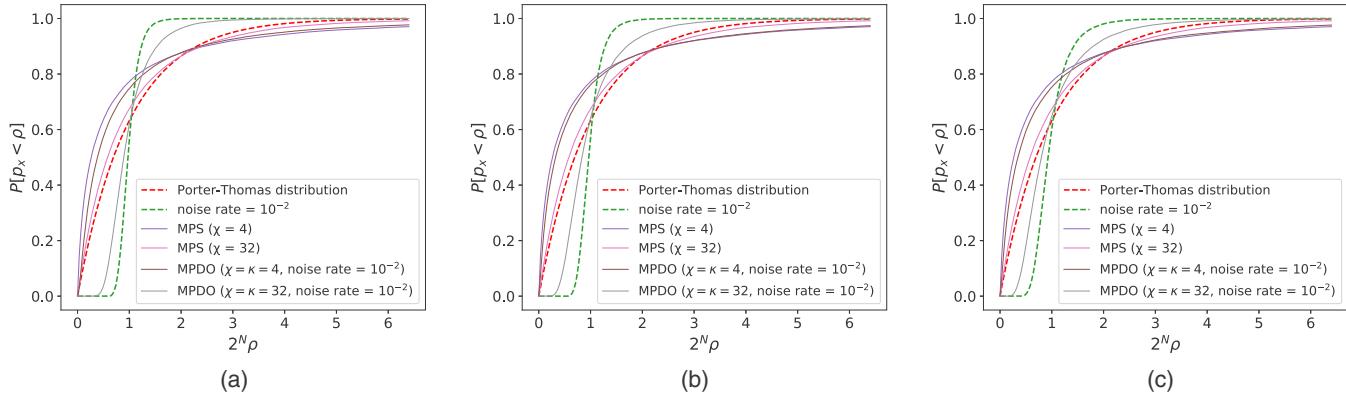


FIG. 6. Approximate results of the cumulated p distribution for different noise simulators. The system is a 1D random circuit with 15 qubits and 24 layers, under three different error models. (a) The dephasing noise; (b) the depolarizing noise; (c) the amplitude damping noise. The green dashed line refers to the result of exact simulation. As a signal that MPS cannot faithfully be simulated with noise, when χ increase, the distribution of MPS simulation will only get closer and closer to the Porter-Thomas distribution, while MPDO can gradually approach the correct cumulated p distribution.

(iv) Distribution given by an MPDO simulator discussed in Sec. III.

In Fig. 5, we compared the Porter-Thomas distribution with exact simulations of different cumulated p distributions under three different types of noise to show the derivation of the output from the Porter-Thomas distribution. Notice that the behavior of amplitude damping in Fig. 5(c) is different compared to the other noise models. This is expected, because when the noise rate became large, the density matrix through the dephasing channel or the depolarizing channel would reach a maximum mixed state, while the amplitude damping channel makes the density matrix a pure state.

Results of approximated simulation are summarized in Fig. 6, which clearly shows that the MPDO distribution can approach its corresponding exact noisy distribution with the increase of the bond dimension χ and the inner dimension κ .

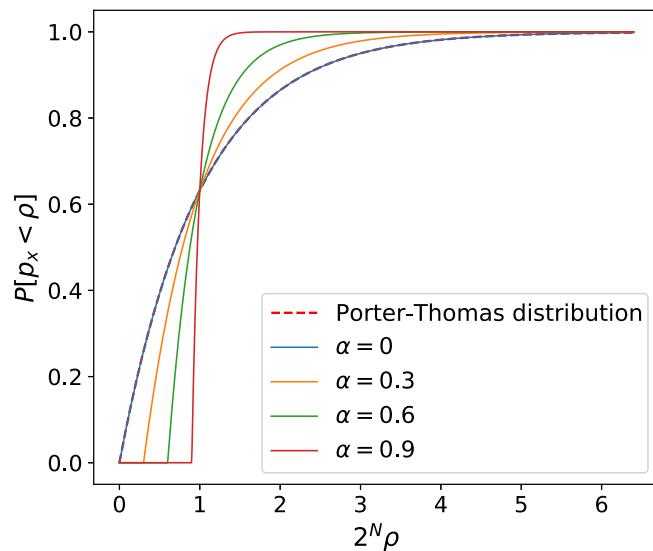


FIG. 7. Analysis result of the deviation from Porter-Thomas distribution under the depolarizing noise

The MPS method does not approximate the actual output distribution of ρ_e . Notice that a relatively small bond and inner dimensions $\chi = \kappa = 32$ for the MPDO simulation already grasp some qualitative behavior of the cumulated p distribution.

For the depolarizing noise, we can also study the deviation from Porter-Thomas distribution analytically. As discussed in Sec. II, the output density matrix of the random circuit will be in the form of certain global depolarizing noise, as given in (2.6). In this case, the distribution of $\{p = p_i\}$ will be given by

$$P((1 - \alpha)p_1 + \alpha p_2 < p) = \begin{cases} 0, & \text{if } 0 \leq p \leq \frac{\alpha}{M}, \\ 1 - \left(1 - \frac{p - \alpha/M}{1 - \alpha}\right)^{M-1}, & \text{if } \frac{\alpha}{M} < p < 1 - \alpha + \frac{\alpha}{M}, \\ 1, & \text{if } 1 - \alpha + \frac{\alpha}{M} \leq p \leq 1. \end{cases} \quad (4.6)$$

The detailed derivation of Eq. (4.6) is given in Appendix.

For different values of α , we plot the accumulated p distribution in Fig. 7. For large α , the distribution approaches the jump function, which corresponds to the uniform distribution. Notice that the analytical result is based on the global noise α , for which the numerical simulation given in Fig. 6(b) is based on the gate noise ϵ . In general, α is a (complicated) function of ϵ that depends on both N and D (Table I provides some intuition of this function for $N = 10, D = 24$). Nevertheless, the results of Fig. 6(b) qualitatively agree with the analytical results given by Fig. 7, which also demonstrates that the MPDO simulation delivers reasonable output.

V. TRUNCATION OF BOND AND INNER DIMENSIONS

In this section, we study the effect of truncation in our MPDO for bond and inner dimensions. We consider random circuits with 10 qubits and depth $D = 24$.

We first study the effect of truncating the bond dimension. Notice the bond dimension, in fact, puts an upper bound for the inner dimension. That is, when the bond dimension is

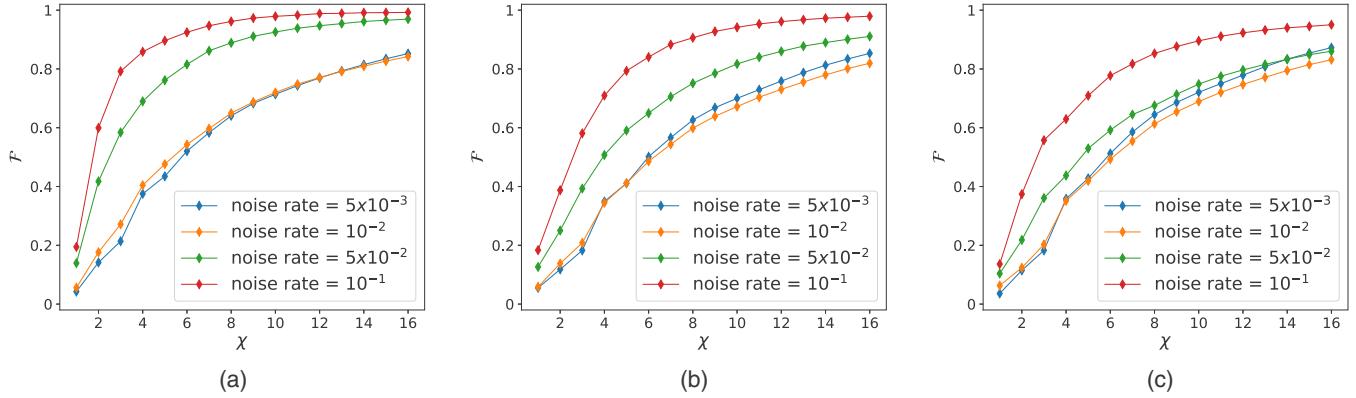


FIG. 8. $\mathcal{F}(\rho_e, \rho_d)$ for different noise rates and different MPDO with maximum bond dimension χ and maximum inner dimension $\kappa = 2\chi$. (a) The dephasing noise; (b) the depolarizing noise; (c) the amplitude damping noise.

truncated to a maximum value of χ , the inner dimension is upper bounded by $2\chi^2$. For each of the error models with different error rates, we choose to truncate the bond dimension to a maximum value χ and also truncate the inner dimension to a maximum value $\kappa = 2\chi$ and compute $\mathcal{F}(\rho_e, \rho_d)$ to see how well the MPDO method approximates the exact noisy result ρ_e . Our results are shown in Fig. 8.

As shown in Fig. 8, when the gate error is larger than some threshold (approximately 0.01 for all the error models), the smaller bond dimension suffices to simulate the noisy circuits, indicating it is the regime that the noise is large enough for the quantum circuit to be “easy” for classical simulation.

We further study the effect of truncating the inner dimensions. For each of the error models with different error rates, we choose to truncate the bond dimension to a maximum value $\chi = 32$ and to truncate the inner dimension to a maximum value of κ . We compute $\mathcal{F}(\rho_e, \rho_d)$ to see how well the MPDO method approximates the exact noisy result ρ_e . Our results are shown in Fig. 9.

As shown in Fig. 9, in case of weak system noise, the smaller inner dimension suffices to simulate the noisy circuits. In this case, the memory cost of MPDO is significantly less than the MPO method by directly using the M tensor.

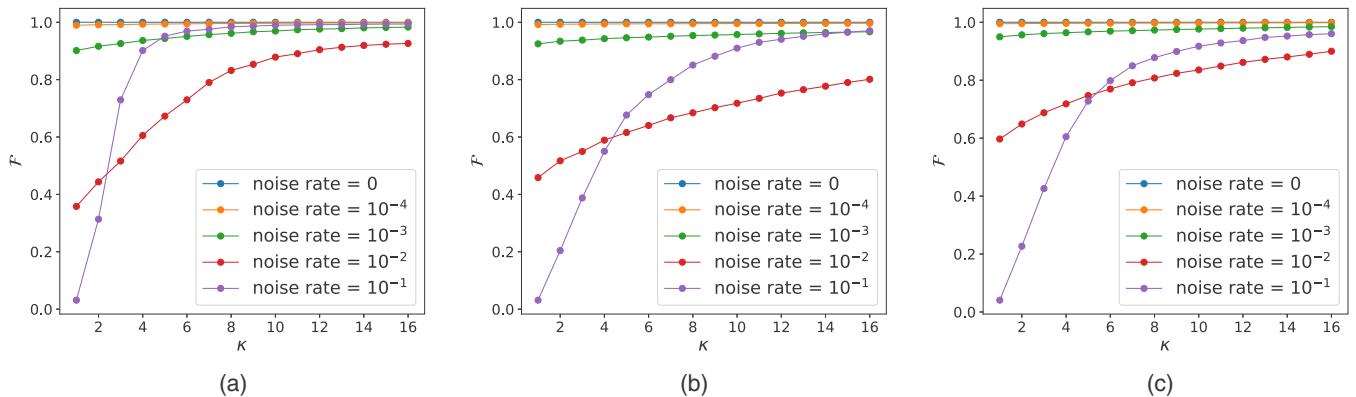


FIG. 9. $\mathcal{F}(\rho_e, \rho_d)$ for different noise rates and different MPDO with inner dimension κ . (a) The dephasing noise; (b) the depolarizing noise; (c) the amplitude damping noise.

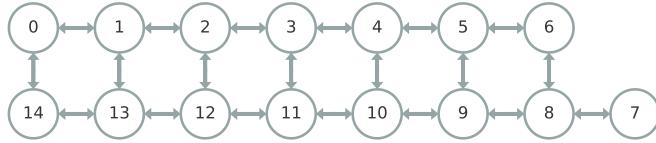
VI. EXPERIMENTS ON IBM QUANTUM DEVICES

To test our MPDO method with real quantum computers, we run several 1D random circuits on the IBM device. We use the 15-qubit device `ibmq_16_melbourne` [36]. The structure of `ibmq_16_melbourne` is shown in Fig. 10.

We run 10-qubit random circuits on a chain, which consists of the qubits 0, 1, 2, 3, 4, 5, 6, 8, 9, 10, and simulate these circuits with our MPDO method. The parameters of `ibmq_16_melbourne` are given in Table II. The rightmost column shows the CNOT error rates. $c_{xi,j}$ represents the error rate for CNOT operation of control qubit- i and target qubit- j . $c_{xi,j} = c_{xj,i}$ always holds, so we just list one of them.

Consider three 10-qubit random circuits with $D = 4, 8, 12, 16$ layers, respectively. Each circuit is run and measured in the computational basis on `ibmq_16_melbourne` for 8192 times. We assume the noise model is depolarizing noise, and simulate these circuits via MPDO according to the noise rates given in Table II.

Denote $P(x_i)$ as the probability of bit string x_i in experiment, and $P_s(x_i)$ as the probability of x_i in our MPDO simulation. To measure the similarity between P and P_s , we use the Kullback-Leibler divergence between the distributions

FIG. 10. Structure of `ibmq_16_melbourne`.

P and P_s as given by

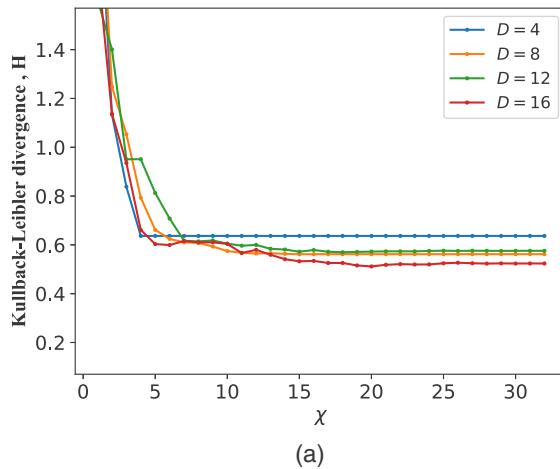
$$H(P, P_s) = \sum P(x) \log P(x) - \sum P(x) \log P_s(x). \quad (6.1)$$

We first show the results between MPS models and experiments in Fig. 11(a). We truncate the bond dimension of MPS to a maximum value χ . However, the K-L divergence of the MPS model no longer drops with increasing χ when it reaches ~ 0.5 , which indicated that the MPS model fails to express the real physical noise in quantum circuits. For our MPDO simulation, we truncate the bond dimension to a maximum value χ and also truncate the inner dimension to a maximum value $\kappa = 2\chi$. The results of $H(P, P_s)$ are shown in Fig. 11(b). Then we truncate the bond dimension χ to a maximum value 32 and the inner dimension to a maximum value κ ; the results of K-L divergence $H(P, P_s)$ versus κ are shown in the inset of Fig. 11(b). These results not only show that the MPDO model is better than the MPS model in simulating a real quantum computer, but also shows that relatively small χ and κ can already simulate the noisy random circuit efficiently.

VII. SIMULATING QUANTUM ERROR CORRECTING CODE

In this section, we give an example of applying the MPDO method to simulate quantum error correction codes. We consider the perfect five-qubit code [37], whose encoding circuit is shown in Fig. 12.

For a single-qubit state in the ensemble $\{|0\rangle, |1\rangle, (|0\rangle + |1\rangle)/\sqrt{2}, (|0\rangle - |1\rangle)/\sqrt{2}, (|0\rangle + i|1\rangle)/\sqrt{2}, (|0\rangle - i|1\rangle)/\sqrt{2}\}$,



(a)

TABLE II. Noise rates of `ibmq_16_melbourne`.

Qubit	Readout error	Single-qubit U2 error rate	CNOT error rate
0	0.0185	5.48×10^{-4}	<code>cx0_1</code> : 0.0236
1	0.0915	2.78×10^{-3}	<code>cx1_2</code> : 0.0165
2	0.0395	8.90×10^{-4}	<code>cx2_3</code> : 0.0171
3	0.0475	3.78×10^{-4}	<code>cx3_4</code> : 0.0169
4	0.0595	1.06×10^{-3}	<code>cx4_5</code> : 0.0295
5	0.0615	2.34×10^{-3}	<code>cx5_6</code> : 0.0467
6	0.027	1.27×10^{-3}	<code>cx6_8</code> : 0.0322
8	0.283	8.13×10^{-4}	<code>cx8_9</code> : 0.0346
9	0.05	9.66×10^{-3}	<code>cx9_10</code> : 0.0510
10	0.03	1.60×10^{-3}	

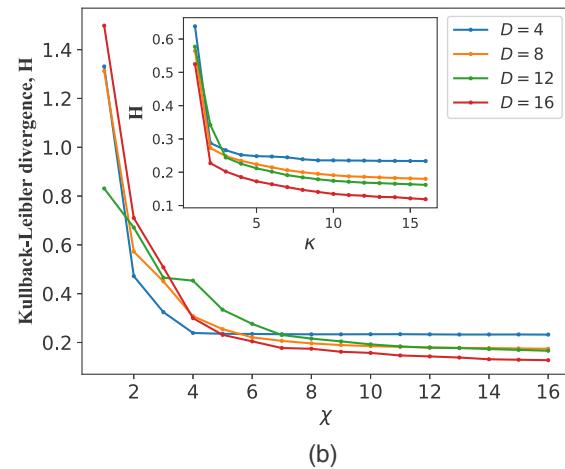
we encode it to five qubits via the encoding circuit, apply a local depolarizing noise channel to all five qubits with noise rate 0.05, then decode and recover the initial state. Denote the average fidelity between the initial state and the recovered state as $\mathcal{F}(\rho, \mathcal{R} \circ \mathcal{E})$.

Suppose the CZ gate in the encoder and decoder is noisy with noise rate ϵ ; we considered four different noise models in total. In addition to the previous dephasing, depolarizing, and amplitude damping noise, we also added an example of neighboring two-qubit noise, the collective dephasing noise [38], which is defined as

$$\rho \rightarrow \mathcal{E}_{CD}(\rho) = (1 - \epsilon)\rho + \epsilon Z_c \rho Z_c^\dagger, \quad (7.1)$$

where $Z_c = \text{diag}(1, -1, -1, 1)$.

Applying a two-qubit noise on the MPDO is similar to the single-qubit case. We only need to contract those two neighboring tensors $T^{[i]}$ and $T^{[i+1]}$ to a merged tensor W with a physical dimension of 4, and then directly apply the two-qubit noise gate to W as the way of the single-qubit noise gate. Finally, the SVD decomposition is used to separate the W into two new tensors. The product of the inner dimension



(b)

FIG. 11. K-L divergence between simulations and experiments on the `ibmq_16_melbourne` device. (a) K-L divergence H between experimental results and simulation results of MPS with maximum bond dimension χ for circuits with D layers. (b) K-L divergence H between experimental results and simulation results of MPDO with maximum bond dimension χ and maximum inner dimension $\kappa = 2\chi$ for circuits with D layers. (Inset) K-L divergence H between experimental results and simulation results of MPDO with maximum inner dimension κ and maximum bond dimension $\chi = 32$ for circuits with D layers.

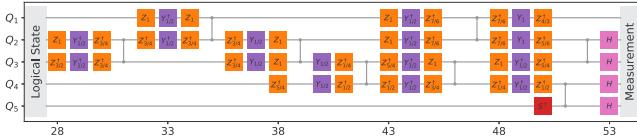


FIG. 12. Encoder of [5,1,3] code.

of those two tensors is equal to m times of the inner dimension of W , where m is the number of terms of the noise model.

We simulate the aforementioned quantum error correcting process with exact density matrix simulation and MPDO simulation. The maximum bond dim and inner dim in MPDO are $\chi = 16$ and $\kappa = 32$. The results of recovered fidelity versus gate noise ϵ for different noise models are shown in Fig. 13. The simulation results of the MPDO are close to the exact ones even with a relative small χ and κ .

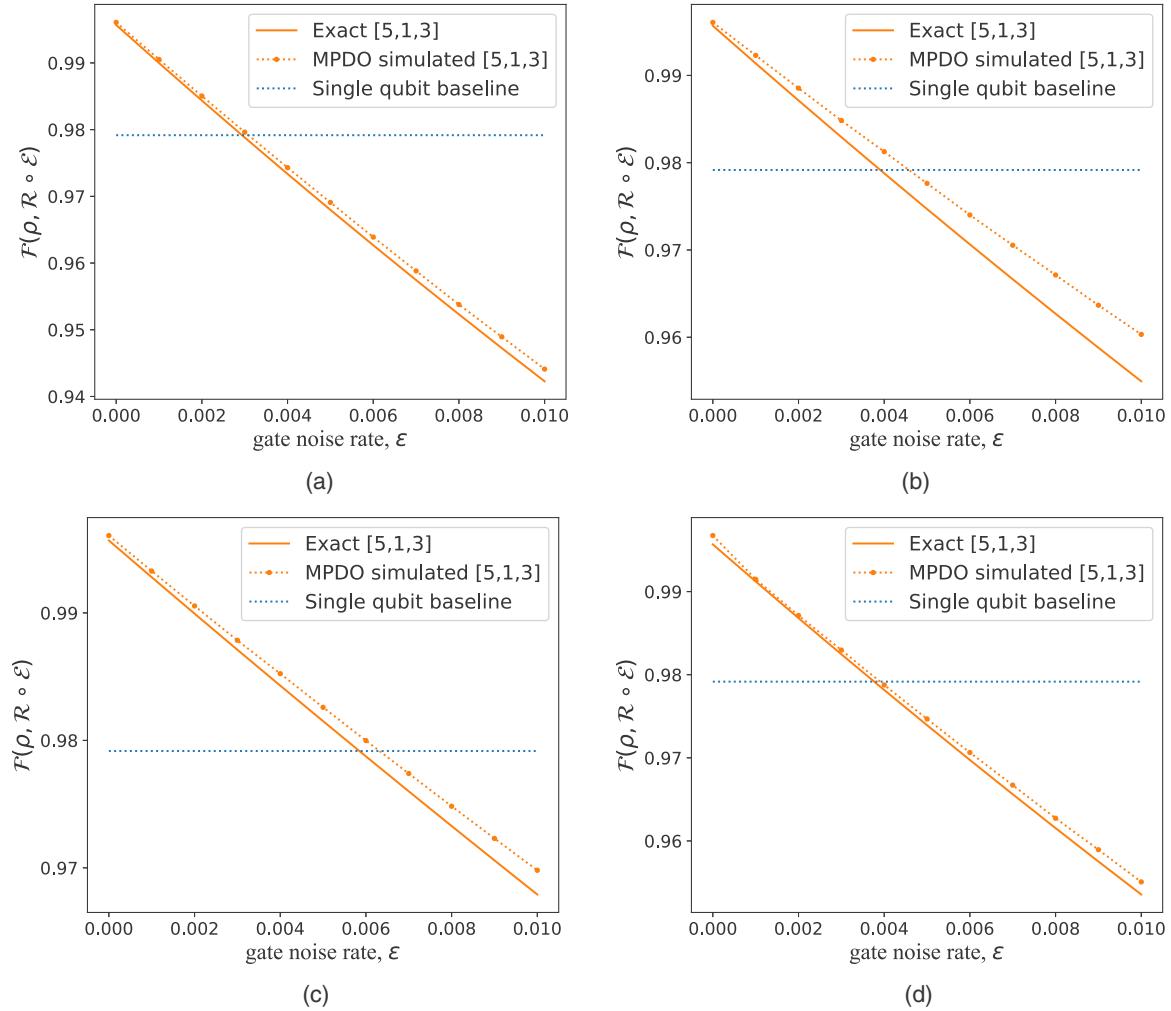
VIII. DISCUSSION

In this work, we have developed a method to simulate noisy quantum circuits based on MPDO. We show that our method approximates the noisy output states well. The method based

on MPS bond dimension truncation, failed to approximate the noisy output states for any of the noise models considered, indicating that the MPS method might not represent any local noise model in real physical systems.

Our MPDO method exhibits the following advantages.

- (i) It reflects a clear physical picture, with inner indices taking care of noise simulation, and bond indices taking care of two-qubit gate simulation.
- (ii) Both bond and inner dimensions can be truncated using the SVD method, adaptive to the need of different situations of the noise simulation.
- (iii) In case of strong system noise, small bond dimensions are sufficient to simulate the noisy circuits.
- (iv) In case of weak system noise, the memory cost of MPDO is significantly less than the MPO method.
- (v) With an effective tensor update scheme that truncates the inner dimension up to a maximum value κ and bond dimension up to a maximum value χ , performed after each layer of the circuit, the cost of our simulation scales as $\sim N D \kappa^3 \chi^3$, for simulating an N -qubit circuit with depth D .
- (vi) Experimental results on IBM devices demonstrate that relatively small χ and κ can simulate the noisy random circuit efficiently.

FIG. 13. Recovered fidelity versus gate noise ϵ for [5,1,3] perfect code. The blue line is the fidelity without quantum error correction. (a) The dephasing gate noise; (b) the depolarizing gate noise; (c) the amplitude damping gate noise; (d) the collective dephasing gate noise.

It remains an interesting open question to understand further the relationship between bond/inner dimensions and entanglement/classical entropy. It is also highly desired to generalize our method to simulate noisy circuits in two spatial dimensions so that we can more directly compare it to 2D experimental data from, e.g., Google's experiments [7].

ACKNOWLEDGMENTS

We acknowledge the use of IBM Quantum services for this work. S.C. and Y.L. are supported by the National Science Foundation of China (Grants No. 12004205 and No. 11701536, respectively).

APPENDIX: DEPOLARIZING NOISE

Consider a simple mode of the depolarizing noise, which is widely used by experiments and known to be a good model for simulating many real quantum systems [7]. Consider a system of N qubits. If the final error rate is α , then the final noisy state can be written as

$$\rho = (1 - \alpha)|\psi\rangle\langle\psi| + \alpha\frac{I}{M}, \quad (\text{A1})$$

where $M = 2^N$ and $|\psi\rangle$ is the ideal final state with the distribution,

$$Pr(p) = (M - 1)(1 - p)^{M-2}. \quad (\text{A2})$$

We notice (A1), where the first term $|\psi\rangle\langle\psi|$ corresponds to the exact output, while the second term $\frac{I}{M}$ corresponds to the noise with uniform distribution, which is the white noise.

For the first term, we already know that the random variable $p_1 = |\langle\psi|\phi\rangle|^2$, with the state $|\psi\rangle$ chosen uniformly at random in the full space, satisfies Porter-Thomas distribution, i.e., (A2). For the second term, by similar manipulation, we have that p_2 satisfies the single point distribution, with the

density function to be

$$Psp(p) = \delta\left(p - \frac{1}{M}\right). \quad (\text{A3})$$

Then we calculate the probability with noise. Note that $p \in [0, 1]$ in (A2) and (A3), however, for convenience; we should make some extension of the function $Pr(p)$ and $Psp(p)$. We define

$$\overline{Pr}(p) := \begin{cases} 0, & \text{if } p < 0, \\ (M - 1)(1 - p)^{M-2}, & \text{if } 0 \leq p \leq 1, \\ 0, & \text{if } p > 1. \end{cases}$$

For $Psp(p)$, we directly make the 0 extension.

According to (A1), since the assumption that white noise is independent of the exact result, the probability with noise should be

$$\begin{aligned} & P((1 - \alpha)p_1 + \alpha p_2 < p) \\ &= \iint_{(1-\alpha)x+\alpha y < p} \overline{Pr}(x)\overline{Psp}(y)dx dy \\ &= \int_{-\infty}^{+\infty} \int_{-\infty}^{\frac{p-\alpha y}{1-\alpha}} \overline{Pr}(x)\delta\left(y - \frac{1}{M}\right)dx dy \\ &= \int_{-\infty}^{+\infty} \int_{-\infty}^p \frac{1}{1-\alpha} \overline{Pr}\left(\frac{z - \alpha y}{1-\alpha}\right)\delta\left(y - \frac{1}{M}\right)dz dy \\ &= \int_{-\infty}^p \int_{-\infty}^{+\infty} \frac{1}{1-\alpha} \overline{Pr}\left(\frac{z - \alpha y}{1-\alpha}\right)\delta\left(y - \frac{1}{M}\right)dy dz \\ &= \int_{-\infty}^p \frac{1}{1-\alpha} \overline{Pr}\left(\frac{z - \alpha/M}{1-\alpha}\right)dz. \end{aligned} \quad (\text{A4})$$

By further calculation, we obtain the result

$$\begin{aligned} & P((1 - \alpha)p_1 + \alpha p_2 < p) \\ &= \begin{cases} 0, & \text{if } 0 \leq p \leq \frac{\alpha}{M}, \\ 1 - \left(1 - \frac{p - \alpha/M}{1-\alpha}\right)^{M-1}, & \text{if } \frac{\alpha}{M} < p < (1 - \alpha) + \frac{\alpha}{M}, \\ 1, & \text{if } (1 - \alpha) + \frac{\alpha}{M} \leq p \leq 1. \end{cases} \end{aligned} \quad (\text{A5})$$

[1] M. A. Nielsen and I. L. Chuang, *Quantum Information and Quantum Computation*, Vol. 2 (Cambridge University Press, Cambridge, 2000).

[2] P. W. Shor, Scheme for reducing decoherence in quantum computer memory, *Phys. Rev. A* **52**, R2493 (1995).

[3] A. Steane, Multiple-particle interference and quantum error correction, *Proc. R. Soc. London A* **452**, 2551 (1996).

[4] L. S. Bishop, S. Bravyi, A. Cross, J. M. Gambetta, and J. Smolin, Quantum Volume, Technical Report (2017), <https://www.semanticscholar.org/paper/Quantum-Volume-Bishop-Bravyi/650c3fa2a231cd77cf3d882e1659ee14175c01d5#citing-papers>.

[5] Quantum takes flight: Moving from laboratory demonstrations to building systems, <https://www.ibm.com/blogs/research/2020/01/quantum-volume-32/>.

[6] J. Gaebler, B. Bjork, D. Stack, M. Swallows, M. Fabrikant, A. Reed, B. Spaun, J. Pino, J. Dreiling, and C. Figgatt, Progress toward scalable quantum computing at honeywell quantum solutions, *Bull. Am. Phys. Soc.* **64**, S01.103 (2019).

[7] F. Arute, K. Arya, R. Babbush, D. Bacon, J. C. Bardin, R. Barends, R. Biswas, S. Boixo, F. G. S. L. Bradao, D. A. Buell *et al.*, Quantum supremacy using a programmable superconducting processor, *Nature (London)* **574**, 505 (2019).

[8] S. Aaronson, Quantum computing, postselection, and probabilistic polynomial-time, *Proc. R. Soc. A* **461**, 3473 (2005).

[9] M. J. Bremner, R. Jozsa, and D. J. Shepherd, Classical simulation of commuting quantum computations implies collapse of the polynomial hierarchy, *Proc. R. Soc. A* **467**, 459 (2011).

[10] S. Aaronson and A. Arkhipov, The computational complexity of linear optics, in *Proceedings of the Forty-third Annual ACM*

Symposium on Theory of Computing (ACM, New York, 2011), pp. 333–342.

[11] K. Fujii and T. Morimae, Commuting quantum circuits and complexity of Ising partition functions, *New J. Phys.* **19**, 033003 (2017).

[12] M. J. Bremner, A. Montanaro, and D. J. Shepherd, Average-Case Complexity Versus Approximate Simulation of Commuting Quantum Computations, *Phys. Rev. Lett.* **117**, 080501 (2016).

[13] S. Aaronson and L. Chen, Complexity-theoretic foundations of quantum supremacy experiments, [arXiv:1612.05903](https://arxiv.org/abs/1612.05903).

[14] R. LaRose, Overview and comparison of gate level quantum software platforms, *Quantum* **3**, 130 (2019).

[15] H. De Raedt, F. Jin, D. Willsch, M. Willsch, N. Yoshioka, N. Ito, S. Yuan, and K. Michielsen, Massively parallel quantum computer simulator, eleven years later, *Comput. Phys. Commun.* **237**, 47 (2019).

[16] M. Smelyanskiy, N. P. D. Sawaya, and A. Aspuru-Guzik, QHIPSTER: the quantum high performance software testing environment, [arXiv:1601.07195](https://arxiv.org/abs/1601.07195).

[17] T. Jones, A. Brown, i. Bush, and S. C. Benjamin, QuEST and high performance simulation of quantum computers, *Sci. Rep.* **9**, 10736 (2019).

[18] E. Pednault, J. A. Gunnels, G. Nannicini, L. Horesh, and R. Wisnieff, Leveraging secondary storage to simulate deep 54-qubit sycamore circuits, [arXiv:1910.09534](https://arxiv.org/abs/1910.09534).

[19] Y. Zhou, E. M. Stoudenmire, and X. Waintal, What Limits the Simulation of Quantum Computers? *Phys. Rev. X* **10**, 041038 (2020).

[20] F. Verstraete, V. Murg, and J. I. Cirac, Matrix product states, projected entangled pair states, and variational renormalization group methods for quantum spin systems, *Adv. Phys.* **57**, 143 (2008).

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

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

[23] B. Pirvu, V. Murg, J. I. Cirac, and F. Verstraete, Matrix product operator representations, *New J. Phys.* **12**, 025012 (2010).

[24] F. Verstraete, J. J. Garcia-Ripoll, and J. I. Cirac, Matrix Product Density Operators: Simulation of Finite-Temperature and Dissipative Systems, *Phys. Rev. Lett.* **93**, 207204 (2004).

[25] K. Woolfe, *Matrix Product Operator Simulations of Quantum Algorithms*, Tech. Rep. (University of Melbourne School of Physics, Melbourne, Australia, 2015).

[26] K. Noh, J. Liang, and B. Fefferman, Efficient classical simulation of noisy random quantum circuits in one dimension, *Quantum* **4**, 318 (2020).

[27] J. G. Jarkovsky, A. Molnar, N. Schuch, and J. I. Cirac, Efficient description of many-body systems with matrix product density operators, *PRX Quantum* **1**, 010304 (2020).

[28] A. H. Werner, D. Jaschke, P. Silvi, M. Kliesch, T. Calarco, J. Eisert, and S. Montangero, Positive Tensor Network Approach for Simulating Open Quantum Many-Body Systems, *Phys. Rev. Lett.* **116**, 237201 (2016).

[29] J. Emerson, Y. S. Weinstein, M. Saraceno, S. Lloyd, and D. G. Cory, Pseudo-random unitary operators for quantum information processing, *Science* **302**, 2098 (2003).

[30] R. Oliveira, O. C. O. Dahlsten, and M. B. Plenio, Generic Entanglement can be Generated Efficiently, *Phys. Rev. Lett.* **98**, 130502 (2007).

[31] A. W. Harrow and R. A. Low, Random quantum circuits are approximate 2-designs, *Commun. Math. Phys.* **291**, 257 (2009).

[32] S. Boixo, S. V. Isakov, V. N. Smelyanskiy, R. Babbush, N. Ding, Z. Jiang, M. J. Bremner, J. M. Martinis, and H. Neven, Characterizing quantum supremacy in near-term devices, *Nat. Phys.* **14**, 595 (2018).

[33] C. Neill, P. Roushan, K. Kechedzhi, S. Boixo, S. V. Isakov, V. Smelyanskiy, A. Megrant, B. Chiaro, A. Dunsworth, K. Arya, R. Barends, B. Burkett, Y. Chen, Z. Chen, A. Fowler, B. Foxen, M. Giustina, R. Graff, E. Jeffrey, T. Huang, J. Kelly, P. Klimov, E. Lucero, J. Mutus, M. Neeley, C. Quintana, D. Sank, A. Vainsencher, J. Wenner, T. C. White, H. Neven, and J. M. Martinis, A blueprint for demonstrating quantum supremacy with superconducting qubits, *Science* **360**, 195 (2018).

[34] State-Vector Simulation with Decision Diagrams, in *Quantum Circuit Simulation* (Springer Netherlands, Dordrecht, 2009), pp. 71–101.

[35] Density-matrix simulation with QuIDDs, in *Quantum Circuit Simulation* (Springer Netherlands, Dordrecht, 2009), pp. 103–114.

[36] IBMQ_16_MELBOURNE v2.1.0, IBM Quantum Team (IBM, Poughkeepsie, 2020), <https://quantum-computing.ibm.com>.

[37] M. Gong, X. Yuan, S. Wang, Y. Wu, Y. Zhao, C. Zha, S. Li, Z. Zhang, Q. Zhao, Y. Liu, F. Liang, J. Lin, Y. Xu, H. Deng, H. Rong, H. Lu, S. C. Benjamin, C.-Z. Peng, X. Ma, Y.-A. Chen, X. Zhu, and J.-W. Pan, Experimental verification of five-qubit quantum error correction with superconducting qubits, *National Sci. Rev.* nwab011 (2021).

[38] X.-H. Li, F.-G. Deng, and H.-Y. Zhou, Efficient quantum key distribution over a collective noise channel, *Phys. Rev. A* **78**, 022321 (2008).