Simulating Neutral Atom Quantum Systems with Tensor Network States

James Allen,^{1,*} Matthew Otten,² Stephen Gray,³ and Bryan K. Clark^{1,†}

¹Institute for Condensed Matter Theory and IQUIST and NCSA Center for Artificial Intelligence Innovation and Department of Physics, University of Illinois at Urbana-Champaign, IL 61801, USA ²HRL Laboratories LLC, Malibu, CA, United States

³Argonne National Laboratory (Dated: September 18, 2023)

While abstract models of quantum computation assume a closed system of two-level states, practical quantum devices inevitably couple to the environment in some way, creating sources of noise. Understanding the tolerance to noise of specific quantum algorithms run on specific devices is important for determining the feasibility of quantum computing in the current noisy intermediate scale quantum era. Of particular interest is understanding the noise sensitivity of these devices as more qubits are added to the system. Classical simulations are a useful tool to understand the effects of this noise, but direct classical simulations of open quantum systems are burdened by an exponentially growing cost in the number of qubits and a large local Hilbert space dimension. For one dimensional, shallow circuits, using tensor networks can replace this exponential cost with a linear one and simulate far wider systems than what would normally be available. In this paper, we describe a tensor network simulation of a neutral atom quantum system under the presence of noise, while introducing a new purity-preserving truncation technique that compromises between the simplicity of the matrix product state and the positivity of the matrix product density operator. We apply this simulation to a near-optimized iteration of the quantum approximate optimization algorithm on a transverse field Ising model in order to investigate the influence of large system sizes on the performance of the algorithm. We find that while circuits with a large number of qubits fail more often under noise that depletes the qubit population, their outputs on a successful measurement are just as robust under Rydberg atom dissipation or qubit dephasing as smaller systems. However, such circuits might not perform as well under coherent multi-qubit errors such as Rydberg atom crosstalk. We also find that the optimized parameters are especially robust to noise, suggesting that a noisier quantum system can be used to find the optimal parameters before switching to a cleaner system for measurements of observables.

I. INTRODUCTION

An ideal quantum computer is decoupled from the environment so as to minimize the effects of noise coming from sources such as dephasing and dissipation. Unfortunately, in practice such separation is difficult because quantum circuit operations necessarily couple the system with an outside source. Realistic quantum devices are best thought of as open quantum systems that interact with various sources of environmental noise. This limits the extent to which quantum circuits can be operated, both in terms of circuit size and depth, before breaking down.

Given the current generation of noisy quantum computers, a wide array of quantum algorithms have been developed, such as the variational quantum eigensolver (VQE) and the quantum approximate optimization algorithm (QAOA),^{1–4} which need low circuit depth and hopefully limited coherence. It is still unclear, though, how even in this shallow depth circuit regime, the effect of realistic noise influences the output, such as variational energies and optimized parameters, of these algorithms, especially as we scale circuits to larger system sizes. Therefore, to facilitate our understanding and characterization of quantum devices, we would like to simulate them as best we can using classical algorithms.

Simulations can play an important role in determining the effect of noise on quantum devices. Unfortu-

nately, simulating quantum devices classically is exponentially difficult in general, a feature which is essential to the algorithmic strength of quantum computing. These simulations are even more difficult in the case of open quantum systems where a single wavefunction cannot represent the full state. A traditional approach which directly represents the entire density matrix of the system becomes inefficient very quickly for state-of-the-art sizes such as Sycamore's 53 qubit system,⁵ and stochastic approaches are limited by a poorly scaling signal-to-noise cost.⁶ For a quantum system based on neutral atom arrays, the qubit count can reach even higher, with systems up to 100 qubits being implemented.⁷

Due to their ability to simulate circuits with a very large number of qubits (albeit at low depth), tensor network states (TNS) are particularly well equipped to study the scaling of noise-based errors with system size. Tensor networks are most frequently used to represent a wavefunction. However, with an open quantum system we need to create a TNS that represents the density matrix. Besides the increased computational burden from the extra physical dimensions, this introduces another problem: enforcing the positivity of the density matrix with a TNS is nontrivial.

In this work, our focus is two-fold. First, we develop and validate a new tensor network approach to approximately enforce positivity of the density matrix when simulating open quantum systems. The most naive representation of the density matrix – a vectorized Matrix Product Operator (MPO) – can be modified to enforce its positivity, following the Matrix Product Density Operator (MPDO) scheme. While this type of tensor network is well-behaved for simple channels, it struggles to implement circuit operations that combine the channels of multiple time steps together. This limits the ability of the truncation algorithm to find the most accurate approximate forms – in fact, we found that it did not perform as well as the naive MPO in our simulations. Instead, we have devised an efficient compromise, the Purity-Preserving Truncation (PPT) algorithm, which keeps the purity of the density matrix constant after each truncation, limiting the maximum negativity of the system.

Second, we have used the PPT, combined with an efficient massively parallel code (see Appendix C) to determine the effect of noise on a neutral atom simulation of the QAOA. We find that even in the face of non-trivial dissipation and dephasing noise, both the energy and optimized parameters found in a QAOA simulation of a transverse field Ising model (TFIM) are quite accurate even as the system size grows to 80 qubits. However, this is conditioned on a successful measurement, i.e. one without a qubit in a dark state. While the measured observable values when the system returns a result seem largely insensitive to noise, the probability of an unsuccessful measurement increases with both noise and system size resulting in the need for many more shots to achieve a similar result. We also find that certain coherent effects, such as crosstalk between qubits, can influence the circuit in a way that creates compounding errors over system size, making it difficult to operate large, accurate circuits under these errors.

The rest of the paper is as follows. In Section II, we introduce the dynamics of the neutral atom array, the specific quantum system that we will be simulating.¹⁰ In Section III we outline our MPO simulation approach, explaining the PPT algorithm and showing that it performs better than the bare MPO and MPDO in a heavily truncated Random Circuit Sampling (RCS) algorithm. In Section IV, we apply our new machinery on a near-optimized iteration of a QAOA circuit, where the circuit parameters have been optimized to create a ground state wavefunction of the TFIM. In Section IVa, we demonstrate that under most sources of error, the VQE iteration's final evaluated energy depends only on error strength and not on system size, although coherent errors caused by Rydberg atom crosstalk might create a system size dependence. In Section IVb, we also consider the possibility of the algorithm selecting a spurious value due to errors, and find that this is also mostly independent of system size in the same way as the energy measurement. Moreover, we find that the parameter optimization tends to be more robust to the noise than the energy measurement. Our work opens a new approach for the simulation of large, open quantum systems and validates the efficacy of QAOA algorithms on neutral atom devices at the noisy intermediate scale.

II. LINDBLAD MASTER EQUATION FOR NEUTRAL ATOM ARRAYS

In this work, we focus on modeling a neutral atom array. The neutral atom array is a system for implementing quantum circuits which is well suited for a large number of qubits, with some current systems composed of hundreds of qubits.¹¹ We focus in this work on a one-dimensional geometry, where each atom contains a two-level computational subspace $|0\rangle$, $|1\rangle$ and a high energy Rydberg state $|r\rangle$ as well as an additional set of dark states $|d\rangle$.

In a neutral atom array, entanglement between nearest-neighbor sites can be created via the Rydberg blockade (Fig. 1a),¹ where two neighboring qubits are temporarily promoted to Rydberg states that repulsively interact with each other. In one such scheme, the two active qubits experience simultaneous pulses under a Hamiltonian

$$H_p(t) = H_1(t) \otimes I_2 + I_1 \otimes H_2(t) + B|rr\rangle\langle rr| \tag{1}$$

where B is the Rydberg blockade strength, and H_1, H_2 are the single-qubit components of the Hamiltonian. These components include a Rabi frequency $\Omega_i(t)$ which promotes a $|1\rangle$ qubit to the Rydberg state $|r\rangle$ and a Rydberg detuning $\Delta_i(t)$,

$$H_i = \frac{\Omega_i(t)}{2} [|r\rangle\langle 1| + \text{h.c.}] + \Delta_i(t)|r\rangle\langle r|.$$
 (2)

For all the systems considered in this paper, each pulse is identical, so $\Omega_1(t) = \Omega_2(t)$ and $\Delta_1(t) = \Delta_2(t)$.

In any quantum system, coupling between the system and environment introduces noise degrees of freedom that must be accounted for. Provided the coupling is weak enough and the environment is memoryless, we can model the evolution of the reduced density matrix of the system ρ with the Lindblad Master Equation (LME),

$$\frac{\mathrm{d}\rho(t)}{\mathrm{d}t} \equiv -i\mathcal{L}[\rho]$$

$$= -i[H_p(t), \rho(t)] + \sum_i L_i \rho(t) L_i^{\dagger} - \frac{1}{2} \{L_i^{\dagger} L_i, \rho(t)\}.$$
(3)

Here L_i are jump operators representing different noise sources in the neutral atom array. One such source is Rydberg atom dissipation. The Rydberg states have a finite lifetime and can decay to either the qubit states or arbitrary dark states $|d\rangle$ that represent any reachable atomic levels. These dark states cease to interact with the rest of the system and we assume that measuring an atom in a dark state counts as a failure of the entire circuit. The jump operator for this mechanism is

$$L_j = \sqrt{\gamma_{diss}b_j}|j\rangle\langle r| \tag{4}$$

for branching ratios b_0, b_1, b_d , and overall dissipation strength γ_{diss} . If the Zeeman shift between qubit energy levels fluctuates, there is another jump operator to represent qubit dephasing, ^{12,13}

$$L_{deph} = \frac{\gamma_{deph}}{\sqrt{2}} (|0\rangle\langle 0| - |1\rangle\langle 1|). \tag{5}$$

The key operation in a quantum circuit is a universal two qubit gate – we will focus on simulating a CZ gate in this paper. To measure the quality of these gates, we use a metric based on a combined arithmetic-geometric mean fidelity, as introduced in. ¹⁴ We investigate two types of pulses. The first type of pulse we consider is an Adiabatic Rapid Passage (ARP) pulse (Fig. 1c inset). ¹ This pulse attempts to use the Rydberg blockade to prevent simultaneous excitations of both sites into the Rydberg states. Due to a π phase shift acquired by the site after it enters and exits the Rydberg states, this pulse will create a CZ gate in the limit of infinite blockade strength and no noise.

With a realistic¹⁰ blockade strength $B_0 = 2\pi \times 500 \text{MHz}$, time period $T = 0.54 \mu s$ and dissipation $\gamma_{diss} = 0.001 T^{-1}$, we calculated a Bell fidelity of 0.989, with most of the inaccuracy coming from phase errors in the unitary operator caused by finite blockade strength (Fig. 1b).

The second type of pulse uses a Gaussian profile for $\Omega(t)$, and a $\Delta(t)$ which is constant as a function of time. (Fig. 1e inset). 13 Unlike in the ARP pulse, the application of the Gaussian pulse to a single qubit does not complete a full $|1\rangle \rightarrow |r\rangle \rightarrow |1\rangle$ oscillation. Therefore at infinite blockade strength, a π phase shift doesn't happen; instead one must select a specific finite blockade strength which in combination with the single qubit rotation generates a CZ. The Gaussian unitary varies more rapidly over blockade strength compared to the ARP pulse, so fluctuations in the blockade strength create more significant errors. However, we no longer need large blockade strengths to create an accurate unitary operator. For blockade strengths as low as $2\pi \times 60 \text{MHz}$ (Fig. 1d) we can create a unitary with a combined infidelity of 6×10^{-8} , provided there are no additional sources of noise.

III. MATRIX PRODUCT OPERATOR REPRESENTATION

In this section we will describe how to efficiently represent a large number of neutral atom qubits using tensor networks. We will also introduce a new way of maintaining the physicality of an approximate representation of the system under the time evolution of a quantum circuit.

The wavefunction of a one-dimensional chain of N qubits can be represented by a matrix product state (MPS), a string of rank 3 tensors (except for the edges which are rank 2), with each tensor representing an in-

dividual site entangled with its neighbors through virtual/bond dimensions (Fig. 2a),

$$|\psi\rangle_{\sigma_1...\sigma_N} = A_{\sigma_1 i_1}^1 A_{\sigma_2 i_1 i_2}^2 ... A_{\sigma_N i_{N-1}}^N$$
 (6)

where σ_j is a physical dimension and i_j is the bond dimension. Likewise, a density matrix can be represented as a matrix product operator (MPO), where each tensor contains two physical dimensions (Fig. 2b),

$$\rho_{\sigma_1...\sigma_N,\sigma_1'...\sigma_N'} = B_{\sigma_1\sigma_1'i_1}^1 B_{\sigma_2\sigma_2'i_1i_2}^2 ... B_{\sigma_N\sigma_N'i_{N-1}}^N. \quad (7)$$

The MPS representation of the wavefunction requires $O(Nd^2D^2)$ values, where d is the physical dimension and D is the bond dimension. Unlike the statevector representation, which requires $O(d^N)$ values, the MPS representation only grows linearly in N, so it becomes the more tractable representation for any quantum state where the required bond dimension is not expected to be too high. This difference is even more important for the density matrix, where each site requires two physical indices, giving a $O(d^{2N})$ cost in statevectors and $O(Nd^4D^2)$ in MPOs. This harsher scaling makes open quantum system density matrix simulations difficult for systems beyond 12 sites: 15 for example, a 15 site system with the same three levels as the neutral atom array requires approximately 4PB of RAM using a naive implementation.

A. MPO form of the Noisy CZ Gate

We can convert the density matrix into an MPS by vectorizing⁸ the forward and backward physical indices at each site, $\sigma_i \sigma'_i \to \eta_i$. After that, we want to evaluate the time evolution channel of a pulse on the neutral atom array,

$$\rho(t+T) = \mathcal{C}_T(\Omega, \Delta, B)[\rho(t)] \tag{8}$$

with duration T, Rabi frequency $\Omega(t)$, detuning $\Delta(t)$ and Rydberg blockade B in terms of a vectorized MPO. We split the time evolution into small steps of duration $\tau \ll T$ and attempt to find an approximate form for the small time step channel

$$C_{\tau}(t) = e^{-i\int_{t}^{t+\tau} \mathcal{L}(t) dt}.$$
 (9)

With $O\left(\tau^2\right)$ errors we can decompose the Lindbladian (3) into each individual component and apply the time-evolved form of the components separately. If ρ is in a vectorized form, we can interpret channel actions $A\rho B$ as an operator $(A\otimes B)(\rho)$ acting on the forward and backward physical indices of ρ . In this notation,

$$C_{\tau}(t)[\rho(t)] = e^{-i\tau H_{p}(t)\otimes I} e^{i\tau I\otimes H_{p}(t)}$$

$$\left(\prod_{j} e^{L_{i}\otimes L_{i}^{\dagger} - \frac{1}{2}(L_{i}^{\dagger}L_{i}\otimes I + I\otimes L_{i}^{\dagger}L_{i})}\right) (\rho(t)). \quad (10)$$

Note that we wish to evaluate the exponential of each component analytically, instead of taking a Taylor series

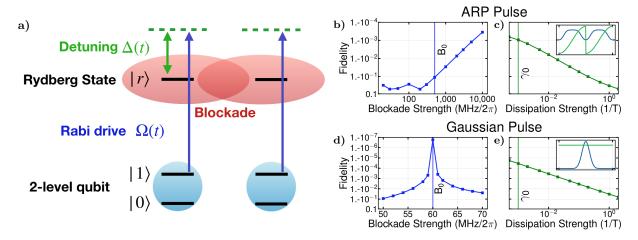


FIG. 1. (a) Rydberg blockade mechanism. A pulse $\Omega(t)$ promotes a $|1\rangle$ qubit state to a Rydberg state $|r\rangle$ with detuning $\Delta(t)$; the two neighboring Rydberg states experience an extra blockade interaction B. (b,c) ARP gate fidelity as a function of (b) blockade strength with dissipation $\gamma_{diss}=0$, and (c) dissipation with blockade $B=2\pi\times 10000$ MHz. $\Omega(t)$ and $\Delta(t)$ follow the blue and green curves of the inset in (c) respectively: $\Omega(t)=\Omega_{max}\big[e^{-(t-t_0)^4/v^4}-a]/(1-a)$ with pulse width parameter v=0.175T, $t_0=T/4$ and $a=e^{-(t_0/v)^4}$, while $\Delta(t)$ follows a split cosine. We use, following Saffman et al, $t=0.54\mu s$, $t=0.54\mu s$

approximation like $e^{\tau \hat{O}} \approx I + \tau \hat{O}$. This is to make sure the approximate channel remains CPTP. The pulse Hamiltonian is further broken down into its respective components

$$e^{-i\tau H_p(t)} = \left(e^{-i\tau \left[\frac{1}{2}\Omega_1(t)(|r\rangle\langle 1| + \text{h.c.}) + \frac{1}{2}\Delta_1(t)|r\rangle\langle r|\right]} \otimes I_2\right)$$
(site 1 \leftrightarrow site 2) \left(e^{-i\tau B|rr\gamma\left(rr|}\right)) (11)

where (site $1 \leftrightarrow \text{site } 2$) refers to the first term on the RHS with sites 1 and 2 exchanged. Each single site operator $\hat{O}_{\sigma_i \sigma_i'}$ can be represented as a rank 2 tensor acting on that site's physical index σ_i . The two-site operator $e^{-i\tau B|rr\rangle\langle rr|}$ becomes a rank 4 tensor acting on both σ_1 and σ_2 . This is the only component of the time evolution channel that entangles the sites.

None of the jump operators we cover in this paper operate on multiple sites at once. However, they do operate on the forward and backward physical indices of the density matrix at the same time. We can write these channels as an operator that acts on the combined vectorized index η_i at a particular site. With all of these channel components combined, the time evolution channel $C_{\tau}(t)$ in the vectorized picture is a rank 4 tensor that acts on the vectorized index of two neighboring sites.

At this point, we can time evolve a vectorized density matrix by applying this channel for each time step. This is accomplished by first multiplying together every tensor involved (Fig. 2c, first step). The resulting tensor now contains the physical indices for two neighboring qubits. We use singular value decomposition (SVD,

Fig. 2c second step) to split the two qubit degrees of freedom into separate tensors. This process also creates a new bond index between the tensors and a diagonal matrix Λ on the bond. Generally, the dimension of this bond index will be larger than the previous bond dimension, but we can truncate it back to the original dimension by removing the least significant diagonal elements in Λ . This process is most efficient when the MPS has been canonized, with its center at one of the active sites of the gate. ¹⁶ Canonization is a gauge transformation that creates a specific center c in the MPS such that the contraction of all sites around that center reduces to the identity,

$$\sum_{i_{k},k < c-1} \sum_{\eta_{l},l < c} B^{1}_{\eta_{1}i_{1}} \dots B^{c-1}_{\eta_{c-1}i_{c-2}i_{c-1}} \sum_{i'_{k},k < c-1} B^{1\dagger}_{\eta_{1}i'_{1}} \dots B^{c-1\dagger}_{\eta_{c-1}i'_{c-2}i'_{c-1}}$$

$$= \delta_{i_{c-1}i'_{c-1}}$$

$$(12)$$

$$\sum_{i_{k},k>c} \sum_{\eta_{l},l>c} B_{\eta_{N}i_{N}}^{N} ... B_{\eta_{c+1}i_{c}i_{c+1}}^{c+1} \sum_{i'_{k},k>c} B_{\eta_{N}i'_{N}}^{N\dagger} ... B_{\eta_{c+1}i'_{c}i'_{c+1}}^{c+1\dagger}$$

$$= \delta_{i_{c}i'_{c}}. \tag{13}$$

The argument in¹⁶ is based on the L2-norm of the MPS being well-behaved, so for a density matrix MPS, as long as the purity is relatively close to 1 (which is the case for any light source of noise), canonization will yield a similar increase in truncation efficiency.

Each application costs a time $O(d^2D^3)$ where d is the physical dimension and D the bond dimension of the vectorized density matrix MPS. As many of the frequencies used in the pulse are very large (particularly the Rydberg blockade), the required time step for accurate simulation is very small, requiring the application of thousands of time step channels for a single pulse. Applying each channel one-by-one to the MPS would be unnecessarily expensive. Instead, we integrate the time evolution channel for the entire pulse directly before applying it to the MPS. In a vectorized picture, this is simply a matter of multiplying all the time step channels together. Once the full time evolution channel of a CZ gate has been assembled, it does not have to be re-evaluated unless one of the dissipation or Hamiltonian parameters change and can be copied onto any CZ gate that appears in the quantum circuit we want to simulate.

The advantage of this method is that we can significantly reduce the amount of times we have to apply an operator directly onto the density matrix, but the disadvantage is that the form of the time evolution channel becomes more complicated to evaluate if it acts on more than two sites. This is the case for a circuit with significant global Rydberg atom crosstalk error.

B. MPO form of the Noisy CZ Gate with crosstalk

In this section we will cover a specific source of coherent error that will be introduced to some (but not all) of the circuit simulations in the rest of this paper. When a pulse excites a qubit to the Rydberg states, it leaves a residual population in that state. If the residual population does not fully decay before the next pulse, there can be unwanted crosstalk between the Rydberg population of the target sites of the pulse and residual populations in neighboring sites. In order to add Rydberg atom crosstalk to our time evolution channel, we need to include the global blockade term

$$H_b = \sum_{i=1}^{N-1} B|rr\rangle\langle rr|_{i,i+1} \bigotimes_{\substack{j=1\\j\neq\{i,i+1\}}}^{N} I_j$$
(14)

to our time evolution channel. In all our simulations of unwanted crosstalk, we assume that there is no noise to combine with the crosstalk, so we time evolve under the Hamiltonian instead of the Lindbladian. All blockade terms commute with each other and the only blockade terms that do not commute with the rest of the Hamiltonian are those that have some overlap with the active sites, i.e. the active site blockade and the nearest neighbor blockade terms. All other terms can be applied as their own local operator onto the MPS, independent of the rest of the MPO.

The blockade between an active site and its nearest non-active neighbor can be interpreted as a shift in the effective detuning of that site, $\Delta \to \Delta + B$, conditioned on whether the non-active neighbor is in a Rydberg state or not. Thus, we can write the gate with nearestneighbor blockade $\mathcal{C}^{NN}_{i,i+1}$ as a combination of unmodified

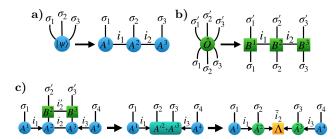


FIG. 2. a) Wavefunction/MPS. b) Operators or Density Matrix/MPO. c) A potential scheme for truncating an MPS while applying a two-site operator. Arrows on the MPS denote a canonization direction. All operator and active site tensors are first combined then resplit into two sites with SVD. This also produces a diagonal tensor Λ between the sites containing the singular values of the squared density matrix. The truncated density matrix is then obtained by removing the lowest singular values of Λ .

two-site gates $C_{i,i+1}(\Delta_i, \Delta_{i+1})$,

$$C_{i,i+1}^{NN} = (I - |r\rangle\langle r|)_{i-1} \otimes (I - |r\rangle\langle r|)_{i+2} \otimes C_{i,i+1}(\Delta_i, \Delta_{i+1})$$

$$+ (|r\rangle\langle r|)_{i-1} \otimes (I - |r\rangle\langle r|)_{i+2} \otimes C_{i,i+1}(\Delta_i + B, \Delta_{i+1})$$

$$+ (I - |r\rangle\langle r|)_{i-1} \otimes (|r\rangle\langle r|)_{i+2} \otimes C_{i,i+1}(\Delta_i, \Delta_{i+1} + B)$$

$$+ (|r\rangle\langle r|)_{i-1} \otimes (|r\rangle\langle r|)_{i+2} \otimes C_{i,i+1}(\Delta_i + B, \Delta_{i+1} + B).$$

$$(15)$$

This tensor gives us the form of the time evolution operator for the active sites and their nearest neighbors, which combined with the two-site blockade operators on the other sites gives the full time evolution of the system as a string of tensors. This assumes that each gate is being applied sequentially, which is not necessarily the case in a real circuit. We could also apply each gate in the same layer simultaneously. However, this would result in a far less tractable MPO, as we would no longer have commutativity of the blockade terms.

C. Matrix Product Density Operators

While the LME by itself preserves the positivity of the density matrix, positivity is not an inherent quality of the MPS. Therefore, if the density matrix is truncated, there is a possibility of introducing negative eigenvalues to the system. In the following sections we will describe two ways to alleviate this negativity problem.

The traditional method, described in this section, is a tensor network representation of the density matrix that enforces its positivity at all times. This representation comes from the diagonalized form of the density matrix as a sum over wavefunction projectors,

$$\rho = \sum_{i} p_i |\psi_i\rangle\langle\psi_i|. \tag{16}$$

Each term in this sum can be represented as the outer product of an MPS and its complex conjugate. Instead of accumulating a potentially exponential amount of

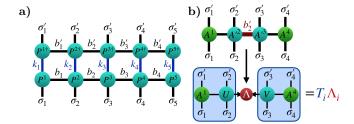


FIG. 3. (a) MPDO Ansatz. The k_i indices are the inner dimensions of the tensor network, playing a similar role to the Kraus indices found in completely positive quantum channels. (b) Setup of the Purity-Preserving Truncation algorithm. Since the Λ tensor is diagonal, the trace can be treated as the inner product of two vectors – the Λ tensor's diagonal entries and its environment.

terms, we can distribute the summation over an extra internal dimension for every site, creating the Matrix Product Density Operator (Fig. 3a). The drawback of this representation is the higher computational cost due to the extra index, and the increased difficulty of applying new gates to the density matrix and truncating it. Unlike in many other works, 9,17 the gate and noise operators are simulated at the pulse level, and occur simultaneously. While the individual time steps can be separated into entangling and noise components, the need to integrate each time step into a single channel, which was caused by the LME parameters requiring large frequencies and short time steps, prevents the gate from being separated this way. Instead, the tensor representing the integrated time evolution channel must be split along its bond and Kraus direction at the same time. In Appendix B we have described such a scheme in the spirit of the Moses move used in isometric tensor networks, ^{18,19} where the split along the Kraus direction is attempted first and then adjusted to optimize its site-wise separation of information. However, such a splitting of information is inefficient when the operators responsible for entanglement and noise do not commute with the other terms of the LME, as is the case for all the noise models we consider in this paper. In addition, applying the time evolution channel to an MPDO affects both its bond and inner dimension at the same time. The fact that four indices must be truncated simultaneously, as opposed to a conventional MPS truncation which only addresses one bond at a time, makes it difficult to concentrate information onto any individual site through a canonization-like scheme.

D. Purity-Preserving Truncation

There is a compromise between the simplicity of the vectorized MPO and the representational faithfulness of the MPDO – we keep the density matrix as a vectorized MPO but ensure that all truncations of the density matrix do not change its purity $tr(\rho^2)$. Specifically, we

define

$$\xi = \frac{\operatorname{tr}(\rho^2)}{\operatorname{tr}(\rho)^2}.\tag{17}$$

In Purity-Preserving Truncation (PPT), after the truncation of the bond's singular values Λ_i to a smaller set $\Lambda'_i \equiv P_T \Lambda_i$ (where P_T is the partial projection that truncates the least significant values), we modify the truncated singular values to $\tilde{\Lambda}_i$, the closest set of values that keeps ξ constant. This is not significantly harder than regular truncation because all terms in the fraction can be represented as polynomial functions of Λ_i and the environment tensors T_i , P_{ij} (Fig. 3b),

$$\operatorname{tr}(\rho) = \sum_{i} T_{i} \Lambda_{i}, \tag{18}$$

$$\operatorname{tr}(\rho^2) = \sum_{i}^{i} P_{ij} \Lambda_i \Lambda_j. \tag{19}$$

Once we have determined T_i and P_{ij} , finding the vector $\tilde{\Lambda}_i$ which is closest to the vector of original singular values Λ_i becomes a constrained optimization problem. If the density matrix is canonized, this problem becomes even simpler, because P_{ij} is the identity matrix. Then $\tilde{\Lambda}_i$ must satisfy

$$\xi = \frac{|\Lambda|^2}{\left(\sum_i T_i \Lambda_i\right)^2} \stackrel{!}{=} \frac{|\tilde{\Lambda}|^2}{\left(\sum_i T_i' \tilde{\Lambda}_i\right)^2}$$
 (20)

where $T' \equiv P_T T$ is the truncated trace environment and $|\tilde{\Lambda}|^2$ is the squared norm of $\tilde{\Lambda}$ treated as a vector. If we define ϕ as the angle between Λ and T, and θ as the angle between $\tilde{\Lambda}$ and T', then we must have

$$\frac{\sec^2 \phi}{|T|^2} = \frac{\sec^2 \theta}{|T'|^2},\tag{21}$$

$$\theta = \cos^{-1}\left(\frac{|T|}{|T'|}\cos\phi\right). \tag{22}$$

To satisfy this bound while maximizing the overlap with both the original singular values Λ_i and its truncation Λ'_i , the new singular values $\tilde{\Lambda}_i$ should become the orthogonal projection of Λ'_i onto the cone of constant angle θ around T'_i . If σ is the angle between the original truncated values Λ'_i and T'_i , we have

$$\tilde{\Lambda}_i = \Lambda_i' + \left(\frac{\tan \theta}{\tan \sigma} - 1\right) \left(\Lambda_i' - \frac{|\Lambda'|}{|T'|} T_i' \cos \sigma\right).$$
 (23)

One caveat is that this bound is not necessarily always achievable. We see from Eq. (22) that if $|T'| < |T| \cos \phi$, then there is no angle θ to satisfy the condition. This is equivalent to the possibility that the original trace environment T, when projected onto the span of the single vector Λ , is longer than its projection T' onto the D environment components corresponding to the largest singular values, where D is the maximum bond dimension. As D increases, this becomes more unlikely. In practice, we only find this occurring under heavy truncation. In these cases, setting $\theta=0$ is the best that we can achieve.

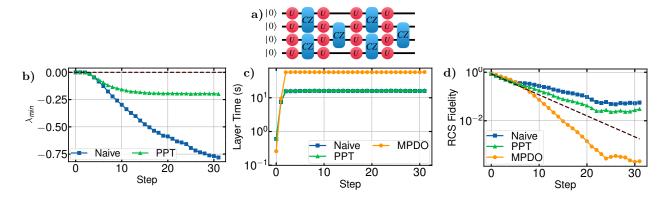


FIG. 4. Simulated random circuit fidelity over layers of a 12 site quantum circuit over 20 samples, with a maximum bond dimension of 64 and a maximum inner dimension of 4, if applicable. (a): Diagram of the random circuit. Each gate U is an independent random 2x2 unitary chosen from the Haar measure, acting on the qubit states. (b): Negativity comparisons of the non-positive MPO with and without PPT. The negativity λ_{min} is the most negative eigenvalue of the density matrix after each layer of the random circuit, determined through DMRG. Without PPT, these eigenvalues increase rapidly, while with PPT they are bounded in absolute value from above. (c): Average wall time required to compute each layer of the random circuit with and without PPT and MPDOs on a single node personal computer. Using PPT introduces a small overhead to the time cost that becomes insignificant for MPS bond dimensions of 64 or above. Given that MPS's typically only become difficult to run at bond dimensions in the thousands, this is a minor cost for most circuits. (d): Random circuit fidelity between systems with and without PPT, including MPDO results. The fidelity is expected to maintain a consistent exponential decay, which is obeyed most closely by the highest bond dimension MPO with PPT – in other circuits we see a sharp deviation from the initial exponential once the bond dimension saturates.

E. Random Circuit Benchmarking of Density Matrix Ansatzes

We use a random circuit architecture to compare the performance of each ansatz. Each layer consists of CZ gates surrounded by Haar random 1-site unitaries applied to every site in an even-odd pattern (Fig. 4a). Given a sufficiently small noise, this circuit should eventually be intractible for any classical simulation, with a complexity growing exponentially in depth.

We first test the positivity of the system with and without PPT under a heavily truncated RCS iteration. We see that PPT does not force the density matrix to be positive, but it does result in less negative eigenvalues for a minor computational cost (Fig. 4b-c). In particular, for the 12-site random circuit of Fig. 4, the minimum eigenvalue of the tensor network with PPT appears bounded at $\lambda_{min} \approx -0.2$. Given a mostly pure density matrix with maximum eigenvalue λ_{max} , we can have a general bound of $\lambda_{\min} \geq -\sqrt{\mathrm{tr}\rho^2 - \lambda_{max}^2}$, however this does not appear to be enough to explain the plateau behavior of the minimum eigenvalue.

We then compare the random circuit fidelity between the simulated output of the circuit under noise and the ideal output without noise. This should take the form of an exponential decay in the circuit depth with a coefficient dependent on the noise factors.^{5,20} Deviations in this exponential decay represent the breakdown of the classical ansatz as too much information gets truncated. Fig. 4d demonstrates the differences between the performance of the non-positive MPO, with and without PPT, as well as the MPDO ansatz. The MPO with PPT ap-

pears to be the most stable ansatz under RCS fidelity. The MPDO fidelity decreases much more rapidly than the expected exponential decay once its bond dimension becomes saturated, with the size of the inner dimension having little effect on this deviation (Appendix A). On the other hand, the MPO without PPT drifts above the expected exponential once its bond dimension saturates. This is because it is no longer reporting a fidelity – the negativity increases the value of $\operatorname{tr}(\rho)^2$, which in turn increases the overlap of ρ with the ideal wavefunction. This can even cause the reported fidelity to exceed 1 for deeper circuits.

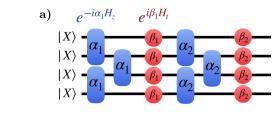
IV. A CANDIDATE CIRCUIT: QAOA ITERATION

We study the ability of the quantum system to use QAOA to generate the ground state of a transverse field Ising model,

$$H_{TFIM} = -J\left(\underbrace{\sum_{\langle ij\rangle} S_i^z S_j^z + b \sum_i S_i^z}_{H_z}\right) - g\underbrace{\sum_i S_i^x}_{H_t}$$
(24)

as one needs to do in a variational quantum eigensolver.

The QAOA ansatz²¹ alternates k times between time evolution of the various terms of Hamiltonian H_z and H_t with weights α_k and β_k respectively. With exact noiseless gates the state $|\psi\rangle$ would evolve under the



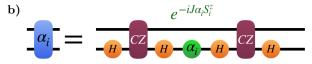


FIG. 5. (a): A circuit diagram of the first two layers of a QAOA iteration for a 4-qubit system. In each layer, blocks of the entangling component H_z of the TFIM Hamiltonian are applied to the circuit for each nearest-neighbor pair, followed by single site operators representing the transverse field component H_t . (b): Each two-site block is composed of two CZ gates simulated on the pulse level, as well as multiple single site operators, following Equation (27).

quantum circuit as

$$|\psi\rangle \to \prod_{k=1}^{K} e^{i\alpha_k H_z} e^{i\beta_k H_t} |\psi_0\rangle.$$
 (25)

The final energy of $|\psi\rangle$ is measured with the Hamiltonian in Equation (24). We use TFIM parameters J=1,b=0.2 and g=1.2 for all our circuits.

A QAOA iteration therefore consists of single site $e^{i\beta_k H_t}$ gates and two site $e^{i\alpha_k H_z}$ gates. We assume the single site gates are comparably easy to perform in a noiseless manner and focus on simulating the two site gates. We can create an arbitrary $e^{i\alpha_k H_z}$ operation from pairs of CZ gates as follows:

$$CNOT_{ij} = I_i \otimes H_j \cdot CZ_{ij} \cdot I_i \otimes H_j \tag{26}$$

$$e^{-iJ\alpha_k S_i^z S_j^z} = \text{CNOT}_{ij} \cdot I_i \otimes e^{-iJ\alpha_k S_j^z} \cdot \text{CNOT}_{ij}$$
 (27)

$$e^{-iJ\alpha_k H_z} = \left(\prod_{\langle ij\rangle} e^{-iJ\alpha_k S_i^z S_j^z}\right) \prod_l e^{-ibJ\alpha_k S_l^z}.$$
 (28)

Each CZ gate is a copy of the one constructed by the LME, Equation (3). Since all terms in the final product (28) commute with each other, our simulation of the QAOA iteration is as follows (Fig. 5). We first assemble each two-site term in Equation (28) using Equation (27). We then apply each term to the circuit, as well as the single site terms. Finally we apply the layer of transverse field gates $e^{i\beta_k H_t}$ for each site l.

We simulate the QAOA iteration using realistic noise sources and Rydberg blockade using a vectorized MPO with Purity-Preserving Truncation and a maximum bond dimension of 768, over different system sizes, with K=8 layers, with the initial state $|\psi_0\rangle=\bigotimes_{i=1}^N|+_X\rangle$ as a product of positive eigenstates of the Pauli X operator. The most expensive calculations were run on

k	α_k	β_k
1	0.11076513	0.75428624
2	0.2771272	0.73016842
3	0.36282021	0.7096901
4	0.40618171	0.68739375
5	0.43256044	0.65733871
6	0.44492256	0.60978220
7	0.42887337	0.51570246
8	0.3225842	0.19145101

TABLE I. Classically optimized time evolution parameters from Eq. (25) for a 10-site, 8-layer TFIM.

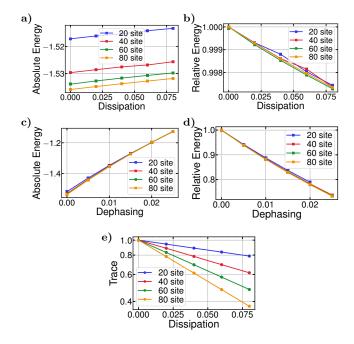


FIG. 6. Final energy of a QAOA iteration with optimized parameters under noise. (a): Absolute energy per site over Rydberg atom dissipation. The energy has very little dependence on dissipation. (b): Relative energy per site $(E-E_0)/E_0$, where E_0 is the energy per site of the noiseless circuit. The relative energy decreases with dissipation but has very little dependence on the system size. (c,d): Absolute (c) and relative (d) energy per site over qubit dephasing; the system size dependence is very small. (e): Trace of the qubit component of the density matrix over system size and dissipation.

the Argonne National Laboratory Computing Resource Center (LCRC) using the distributed-memory Cyclops Tensor Framework (CTF, Appendix C).

For each system size, we use the same parameters α_j , β_j optimized classically on a 10-site system (Table I). On a larger number of sites, these parameters will produce an energy that is close to, but not exactly, the ground state energy.

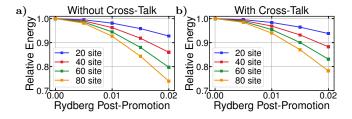


FIG. 7. Relative energy per site over Rydberg atom postpromotion, both without (a) and with (b) crosstalk. Introducing crosstalk in a sequential pattern of gates creates an error that is dependent on both post-promotion level and system size.

A. QAOA Energy in the Presence of Noise

We investigate the performance of the QAOA iteration over two types of incoherent noise: dissipation of the Rydberg population and dephasing of the individual qubit states (Fig. 6). We first fix the Rydberg blockade at $2\pi \times 60$ MHz and adjust the dissipation from 0 to 0.1. We note that the accuracy of the energy gets worse as we increase both the dissipation and dephasing although is more strongly affected by dephasing. Interestingly, the energy per site has almost no system size dependence suggesting that the errors due to both noise sources do not accumulate as the system gets larger. This suggests that a neutral atom experiment could successfully measure the QAOA energy for large systems even in the face of significant dissipation. Unfortunately, while the energy eventually measured by the circuit is independent of system size, the number of circuit iterations required to measure the energy will increase for larger systems, due to the decrease in the qubit population of the density matrix caused by Rydberg atom dissipation (dephasing is irrelevant here as it does not affect the dark states). This creates a larger chance of errant population in the Rydberg atom and dark states, which would make the energy measurement invalid. For large enough system sizes, this makes the energy difficult to evaluate, even if the energy would theoretically be accurate if one were lucky enough to measure it. We can see this effect by looking at the trace of the Rydberg/dark state components of the density matrix and its deviation from 0 as seen in Fig. 6e. The accumulation of non-qubit population is a site-wise independent process - the overall qubit trace of the density matrix is an exponential in dissipation and system size,

$$\operatorname{tr}_{q}(\rho) = e^{-0.1556\gamma_{diss}N}.$$
 (29)

We also simulated the influence of a possible coherent error within the system, that of unwanted Rydberg atom crosstalk. The Rydberg blockade can be problematic if there is a residual Rydberg atom population on sites neighboring those where a gate is being applied, as they can interfere with the dynamics of the gate. These residual populations are normally too small to have an

N	$\frac{\mathrm{d}p}{\mathrm{d}\gamma_{diss}}$	$\frac{\mathrm{d}p}{\mathrm{d}\gamma_{deph}}$	$\frac{\mathrm{d}p}{\mathrm{d}\delta_{PP}}$ (no x-talk)	$\frac{\mathrm{d}p}{\mathrm{d}\delta_{PP}}$ (with x-talk)
	4.963×10^{-3}		0.4152	0.4569
40	4.910×10^{-3}	-0.7955	0.6536	0.7768
60	6.594×10^{-3}	-0.8008	0.8487	1.022
80	6.926×10^{-3}	-0.7959	0.9976	1.241

TABLE II. Error in the α_3 parameter that minimizes the energy over noise for different system sizes.

observable effect during the normal execution of a circuit, so we introduce a post-promotion term

$$\hat{A}_{PP} = e^{-i\delta_{PP}(|1\rangle\langle r| + |r\rangle\langle 1|)} \tag{30}$$

to increase the Rydberg atom population after a gate is applied. Fig. 7 shows the effects on the relative energy error of this post-promotion, with and without any crosstalk between sites. Post-promotion introduces an error in the relative energy per site that increases with both post-promotion strength δ_{PP} and system size. However, the effects of introducing crosstalk are not only minor, the crosstalk even appears to cause a slight improvement in the energy (Fig. 7b).

B. QAOA Parameter Optimization over errors

In addition to measuring the robustness of the noisy neutral atom device with respect to the energy, we would also like to understand if the parameters α_k , β_k , one would find during optimization of the QAOA process are robust to noise. To study this, we start with a circuit that is near-optimized, by fixing all α_i and β_i parameters at the pre-optimized values except for one α_3 , which we vary by a multiplicative phase factor p, and then measure the circuit energy over that phase factor (Fig. 8a-b). As the noise factor γ is increased, the optimal phase factor should vary according to some function $p(\gamma)$. We can measure the overall degree to which this phase factor changes, $\frac{dp}{d\gamma}$.

The dissipation of the system changes the optimal value of α_3 to a much lesser degree than it changes the final energy. For example, at a dephasing of 0.01 with 60 qubits, the optimal α_3 parameter decreases by a factor of 8.6×10^{-3} ; this error in α_3 would affect the final energy per site by 9.8×10^{-5} . The effect on the energy from the changed parameter is much smaller then the energy error measured due to dephasing which, at this level of dephasing, is at 0.18. Therefore, the energy errors induced by selecting an improper optimization parameter are not the main source of inaccuracy.

This suggests that if we only want to find the optimal parameters of this model, we do not need a very clean system. Therefore, we can consider a protocol where we use a cheap, noisy system to determine the optimal parameters of the circuit, then use a cleaner, more expensive circuit to measure the expectation value of operators over the optimized wavefunction. Provided

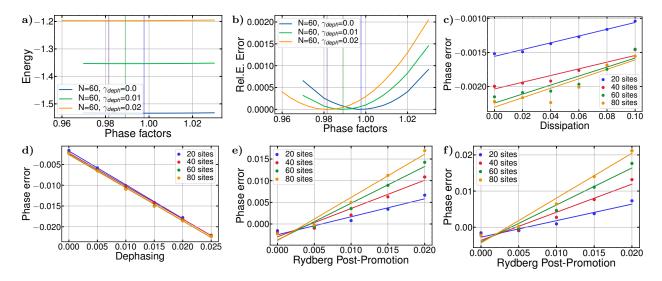


FIG. 8. Final absolute (a) or relative error (b) energy of a QAOA iteration at 60 sites and various levels of dephasing, where the circuit gates differ from the pre-optimized minimum parameters by a phase factor applied to α_3 . The vertical lines in the graph are located at the phase factor that gives the minimal energy for each circuit (the difference in the optimal phase factor over each circuit is not visible at this scale of the graph). (c-f): Errors p-1 in the energy-minimizing phase factor p of p0 of p0 of p0 setimated by noisy gates, over the relevant noise parameter and system size. The noise parameters are dissipation (c), dephasing (d), and Rydberg Post-Promotion without crosstalk (e) and with crosstalk (f).

that the higher noise does not reduce the qubit population of the density matrix to the extent that most circuits are immediately rejected, this can save time on the QAOA optimization process.

There is no discernible dependence of the optimization error $\frac{\mathrm{d}p}{\mathrm{d}\gamma}$ with system size for incoherent noise types or environmental effects encoded by the LME (Fig. 8c-f, Table II). For Rydberg post-promotion, however, the slope increases with system size without crosstalk and, unlike the relative energy measurements, becomes even worse with crosstalk. Therefore, larger systems will be proportionally more difficult to optimize parameters over in the presence of this error.

V. OUTLOOK

We have constructed a tensor-network based pulselevel simulation of large-scale, one-dimensional neutral atom arrays, using vectorized MPO's to represent the density matrix and two-site gate created by integrating a Linblad Master Equation with the Rydberg blockade. We have developed a new algorithmic approach, the PPT, to help maintain the physicality of the density matrix as a quantum circuit is acted on it. We have benchmarked the PPT and found that it is more efficient then MPDO while having only a minimal number of negative eigenvalues in the resulting density matrix. In practice, because we can go to larger bond dimensions than an MPDO at similar computational cost, we are closest to the ground truth using PPT. We proceeded to then use this machinery to simulate the pulse-level dynamics of open quantum circuits on a neutral atom array for QAOA on a transverse field Ising model. We find that at fixed depth there is little to no dependence on the system size of the accuracy of the circuit under noisy errors, although there is a non-trivial increase in the failure rate under such errors. Under coherent errors, there is a possible decrease in accuracy (specifically final energy and the correct optimization parameters) as the number of qubits is increased.

If we extrapolate our findings to systems of arbitrary size, the failure rate of a QAOA iteration under a dissipation of $0.001T^{-1}$ such as in¹⁰ will yield a qubit trace of $\mathrm{tr}_q(\rho) = e^{-1.556 \times 10^{-4} N}$. At N=200 this would result in a 3% trace error. The expected error in relative energy is extremely minor at only -3.3×10^{-5} , and a near-optimized system would find a parameter error of order 10^{-5} . Dephasing is a much more significant error under a dephasing of $0.001T^{-1}$, the relative energy error becomes 0.011, and the parameter error becomes order 10^{-3} . This is possibly due to the limited amount of time in which Rydberg atom populations are large enough that dissipation is allowed to act. This suggests that it would be possible to run QAOA iterations under this noise model with relatively low error and failure rates.

Given the harsher scaling of coherent errors, an interesting open question is whether coherent errors are generically the dominant error source in open quantum devices. In fact, even for noise sources such as dissipation there is a coherent and incoherent piece and it is plausible that even in dissipation, the coherent piece is driving errors (see Appendix D for a discussion of this).

While we mainly focused on parameters that were close to the optimal value, it remains to determine how

the QAOA behaves under noise at any stage of optimization, including completely random starting parameters and semi-optimized parameters. This would mainly be useful for problems where even the roughest optimization is classically unfeasible, which does not apply to the current TFIM.

The ability to determine the effect of errors on realistic quantum devices at scale with respect to interesting algorithms is important to making progress in the field of quantum computers. We have demonstrated a particular example of this in neutral atoms systems and believe our new techniques are an important step for-

ward for future applications.

VI. ACKNOWLEDGEMENTS

We acknowledge useful discussions with Mark Saffman, Martin Suchara and Xiaoyu Jiang. This material is based upon work supported by the U.S. Department of Energy, Office of Science, National Quantum Information Science Research Centers. We gratefully acknowledge the computing resources provided on Bebop, a high-performance computing cluster operated by the Laboratory Computing Resource Center at Argonne National Laboratory.

- * jamesza2@illinois.edu
- † bkclark@illinois.edu
- ¹ T. M. Graham, Y. Song, J. Scott, C. Poole, L. Phuttitarn, K. Jooya, P. Eichler, X. Jiang, A. Marra, B. Grinkemeyer, M. Kwon, M. Ebert, J. Cherek, M. T. Lichtman, M. Gillette, J. Gilbert, D. Bowman, T. Ballance, C. Campbell, E. D. Dahl, O. Crawford, N. S. Blunt, B. Rogers, T. Noel, and M. Saffman. Multi-qubit entanglement and algorithms on a neutral-atom quantum computer. *Nature*, 604(7906):457–462, 2022.
- ² Edward Farhi, Jeffrey Goldstone, and Sam Gutmann. A Quantum Approximate Optimization Algorithm. 2014.
- ³ Suguru Endo, Zhenyu Cai, Simon C. Benjamin, and Xiao Yuan. Hybrid quantum-classical algorithms and quantum error mitigation. *Journal of the Physical Society of Japan*, 90(3), 2 2021.
- ⁴ Kishor Bharti, Alba Cervera-Lierta, Thi Ha Kyaw, Tobias Haug, Sumner Alperin-Lea, Abhinav Anand, Matthias Degroote, Hermanni Heimonen, Jakob S. Kottmann, Tim Menke, Wai Keong Mok, Sukin Sim, Leong Chuan Kwek, and Alán Aspuru-Guzik. Noisy intermediate-scale quantum algorithms. Reviews of Modern Physics, 94(1), 2022.
- Frank Arute, Kunal Arva, Ryan Babbush, Dave Bacon, Joseph C. Bardin, Rami Barends, Rupak Biswas, Sergio Boixo, Fernando G.S.L. Brandao, David A. Buell, Brian Burkett, Yu Chen, Zijun Chen, Ben Chiaro, Roberto Collins, William Courtney, Andrew Dunsworth, Edward Farhi, Brooks Foxen, Austin Fowler, Craig Gidney, Marissa Giustina, Rob Graff, Keith Guerin, Steve Habegger, Matthew P. Harrigan, Michael J. Hartmann, Alan Ho, Markus Hoffmann, Trent Huang, Travis S. Humble, Sergei V. Isakov, Evan Jeffrey, Zhang Jiang, Dvir Kafri, Kostyantyn Kechedzhi, Julian Kelly, Paul V. Klimov, Sergey Knysh, Alexander Korotkov, Fedor Kostritsa, David Landhuis, Mike Lindmark, Erik Lucero, Dmitry Lyakh, Salvatore Mandrà, Jarrod R. McClean, Matthew McEwen, Anthony Megrant, Xiao Mi, Kristel Michielsen, Masoud Mohseni, Josh Mutus, Ofer Naaman, Matthew Neeley, Charles Neill, Murphy Yuezhen Niu, Eric Ostby, Andre Petukhov, John C. Platt, Chris Quintana, Eleanor G. Rieffel, Pedram Roushan, Nicholas C. Rubin, Daniel Sank, Kevin J. Satzinger, Vadim Smelyanskiy, Kevin J. Sung, Matthew D. Trevithick, Amit Vainsencher, Benjamin Villalonga, Theodore White,

- Z. Jamie Yao, Ping Yeh, Adam Zalcman, Hartmut Neven, and John M. Martinis. Quantum supremacy using a programmable superconducting processor. *Nature*, 574(7779):505–510, 2019.
- William Berquist, Danylo Lykov, Minzhao Liu, and Yuri Alexeev. Stochastic Approach for Simulating Quantum Noise Using Tensor Networks. Proceedings of QCS 2022: 3rd International Workshop on Quantum Computing Software, Held in conjunction with SC 2022: The International Conference for High Performance Computing, Networking, Storage and Analysis, pages 107–113, 2022.
- ⁷ Ettore Canonici, Stefano Martina, Riccardo Mengoni, Daniele Ottaviani, and Filippo Caruso. Machine-learning based noise characterization and correction on neutral atoms NISQ devices. 2023.
- ⁸ Kyungjoo Noh, Liang Jiang, and Bill Fefferman. Efficient classical simulation of noisy random quantum circuits in one dimension. *Quantum*, 4, 2020.
- ⁹ 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. *Physical Review Letters*, 116(23), 2016.
- M. Saffman, I. I. Beterov, A. Dalal, E. J. Páez, and B. C. Sanders. Symmetric rydberg controlled- z gates with adiabatic pulses. *Physical Review A*, 101(6), 2020.
- Loïc Henriet, Lucas Beguin, Adrien Signoles, Thierry Lahaye, Antoine Browaeys, Georges Olivier Reymond, and Christophe Jurczak. Quantum computing with neutral atoms. Quantum, 4:327, 6 2020.
- ¹² X. L. Zhang, A. T. Gill, L. Isenhower, T. G. Walker, and M. Saffman. Fidelity of a Rydberg-blockade quantum gate from simulated quantum process tomography. *Physical Review A - Atomic, Molecular, and Optical Physics*, 85(4), 2012.
- ¹³ F. Robicheaux, T. M. Graham, and M. Saffman. Photonrecoil and laser-focusing limits to Rydberg gate fidelity. *Physical Review A*, 103(2), 2021.
- Daniel M. Reich, Giulia Gualdi, and Christiane P. Koch. Minimum number of input states required for quantum gate characterization. *Physical Review A - Atomic, Molecular, and Optical Physics*, 88(4), 5 2013.
- Matthew Otten and Stephen K. Gray. Accounting for errors in quantum algorithms via individual error reduction. npj Quantum Information, 5(1), 2019.

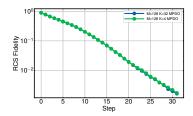


FIG. 9. 12-site RCS fidelity of two MPDO's with the same bond dimension M=128 but different inner dimensions K=4,32. The two fidelities only deviate slightly for very large time steps.

- Yifan Zhang and Edgar Solomonik. On Stability of Tensor Networks and Canonical Forms. Technical report, 2020.
- ¹⁷ Song Cheng, Chenfeng Cao, Chao Zhang, Yongxiang Liu, Shi Yao Hou, Pengxiang Xu, and Bei Zeng. Simulating noisy quantum circuits with matrix product density operators. *Physical Review Research*, 3(2):23005, 2021.
- ¹⁸ Michael P. Zaletel and Frank Pollmann. Isometric Tensor Network States in Two Dimensions. Technical Report 3, 2020.
- ¹⁹ Johannes Hauschild, Eyal Leviatan, Jens H. Bardarson, Ehud Altman, Michael P. Zaletel, and Frank Pollmann. Finding purifications with minimal entanglement. Technical Report 23, 2018.
- Yunchao Liu, Matthew Otten, Roozbeh Bassirianjahromi, Liang Jiang, and Bill Fefferman. Benchmarking near-term quantum computers via random circuit sampling. 2021.
- Wen Wei Ho and Timothy H. Hsieh. Efficient variational simulation of non-trivial quantum states. SciPost Physics, 6(3), 2019.
- Michael P Zaletel and Frank Pollmann. Isometric Tensor Network States in Two Dimensions. Technical report, 2019
- ²³ Adam Bouland, Bill Fefferman, Chinmay Nirkhe, and Umesh Vazirani. On the complexity and verification of quantum random circuit sampling. *Nature Physics*, 15, 2019.
- ²⁴ Ryan Levy, Edgar Solomonik, and Bryan K. Clark. Distributed-memory DMRG via sparse and dense parallel tensor contractions. *International Conference for High Performance Computing, Networking, Storage and Analysis, SC*, 2020-Novem:319–332, 2020.
- ²⁵ Edgar Solomonik, Devin Matthews, Jeff R. Hammond, John F. Stanton, and James Demmel. A massively parallel tensor contraction framework for coupled-cluster computations. *Journal of Parallel and Distributed Comput*ing, 74(12):3176–3190, 2014.

A. ENFORCING THE INNER DIMENSION OF AN MPDO

We have both strong numerical evidence and an intuitive suggestion that a high inner dimension is not required to simulate the MPDO. The latter is a constructive proof that any density matrix can be represented as an MPDO with low inner dimension. Starting with the standard decomposition of a density matrix into orthogonal wavefunctions

$$\rho = \sum_{i=0}^{d^N - 1} p_i |\psi_i\rangle\langle\psi_i|, \tag{31}$$

each wavefunction $|\psi_i\rangle$ can be represented as an MPS with tensors A^i_j on site j. We can construct an MPS for the sum of wavefunctions

$$|\Psi\rangle = \sum_{i=0}^{d^N - 1} \sqrt{p_i} |\psi_i\rangle \tag{32}$$

by turning each site tensor B_j into a block-diagonal composition of the individual site tensors A_j^i . If we consider the projection

$$|\Psi\rangle\langle\Psi| = \sum_{i,j=0}^{d^N - 1} \sqrt{p_i p_j} |\psi_i\rangle\langle\psi_j| \tag{33}$$

we obtain the original density matrix ρ as long as we discard all cross-terms $|\psi_i\rangle\langle\psi_j|$, $i\neq j$. We can achieve this by using the inner indices at each site of the MPDO. For each ψ_i , we consider the base-d decomposition of i,

$$i \to k_{i1}k_{i2}...k_{iN}. \tag{34}$$

We can then construct the positive TNS of ρ with inner dimension d at every site, by starting with B_j and setting each block A^i_j to 0 unless the jth digit k_{ij} matches the inner index. The total contraction of this TNS therefore only keeps the terms $|\psi_i\rangle\langle\psi_j|$ where i and j have the exact same digits at every site, i.e. when i=j exactly.

The bond dimension of such a MPDO would generally be quite large at $O(d^{2N})$. This is expected, as we had not made any assumptions about the form of the density matrix. While we do not have a general scheme for reducing the inner dimension while keeping the bond dimension relatively low, we do observe that the maximum allowed inner dimension of a MPDO can be as low as 4 without significantly changing the RCS fidelity (Fig. 9).

B. THE FOUR-TENSOR SPLIT

An operator with both entangling and mixing terms needs to be split along both directions before being applied to the MPDO. For a two-site operator $\hat{O}_{i_f,j_f,i_b,j_b}^{i_f,j_f',i_b',j_b'}$, where i_f,j_f,i_b,j_b are the forward and backward physical indices at sites i and j, we must identify two bond dimensions l,m and two Kraus dimensions k_i,k_j , with a split of the form

$$\hat{O}_{i_f,j_f,i_h,j_h}^{i'_f,j'_f,i'_b,j'_b} \to \hat{A}_{i_f,l,k_i}^{i'_f} \hat{B}_{j_f,l,k_i}^{j'_f} \hat{A}_{i_h,m,k_i}^{\dagger i'_b} \hat{B}_{j_h,m,k_i}^{\dagger j'_b}. \tag{35}$$

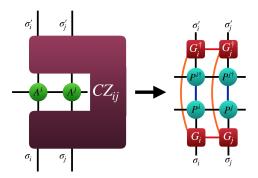


FIG. 10. A two-site operator acting on a density matrix in MPDO form must be split along both the bond and inner dimension before being applied to the MPDO, creating a four-tensor split.

 \hat{A} and \hat{B} are the operators that will be applied to sites i and j respectively of the MPDO, with \hat{A}^{\dagger} and \hat{B}^{\dagger} implicitly applied to the complex conjugated half of the MPDO. We first separate O along the Kraus direction, creating a combined Kraus index k for both sites with maximum dimension K^2 , before splitting this index into two indices k_i , k_j with maximum dimension K each. Initially, these indices have little relation to the physical qualities of the site they are supposed to represent. As a result, there is potentially a large amount of entanglement between the sites. We use an entanglement reduction algorithm between the indices of sites i and j, by applying unitary operators that act on k_i and k_i . This is the same algorithm that enabled the Moses move, ²² and can be found in Ref. 19. We attempt to reduce the entanglement by minimizing the second Renyi entropy across the sites, $S_2 = -\log(\operatorname{tr}(\rho_G^2))$, where ρ_G is an effective density matrix across the gate. The term in the log is a quartic function of the unitary, so its derivative over that unitary can be easily obtained, and from that derivative we can iteratively optimize the entropy.

C. DISTRIBUTED-MEMORY SPEEDUP

For an MPO of bond dimension D, the cost of operations like contraction or circuit evolution scale as $O(D^3)$, with an $O(D^2)$ space requirement. MPDOs add an extra factor of the inner dimension K to both operations. For computationally difficult circuits like RCS, 23 the ideal dimensions required to maintain accuracy grow exponentially with depth, resulting in increasingly costly operations. We use the tensor-tools library 24 built on top of the Cyclops Tensor Framework (CTF) 25 to reduce the time cost of these simulations by distributing the data of individual tensors over multiple nodes.

We tested this framework by performing a 12-site, 20-layer RCS simulation on one sample with maximum bond dimension 1024 over multiple nodes (Fig. 11a). As we allot more nodes to the program, the overall wall

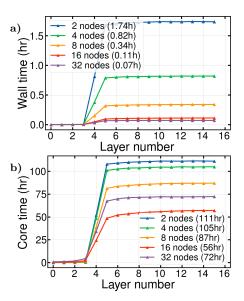


FIG. 11. Stress test of a 32-site, 16-layer random circuit with 1024 bond dimension. (a): Wall time of each process over each layer. (b): Total cost in core-hours of each process. Due to the easier handling of large memories, the program is the most efficient at 16 nodes.

time of the operations decreases. This speedup was better than proportional at some node sizes, resulting in an overall reduction of total core hours as well as wall time (Fig. 11b). This may be due to the distributed-memory nature of the program reducing the size of the tensors each node has to manipulate, removing costs associated with tensors that are too large for the node memory to properly handle.

D. COHERENT COMPONENTS OF INCOHERENT ERRORS

An interesting open question is to better understand whether the coherent components of incoherent errors like Rydberg atom dissipation are the main way they reduce the accuracy of the TFIM. Specifically, a pulse under Rydberg atom dissipation γ_{diss} experiences single and double site population depletion

$$\begin{pmatrix} \rho_{|00\rangle\langle00|} & \rho_{|00\rangle\langle01|} & \rho_{|01\rangle\langle00|} & \rho_{|01\rangle\langle01|} \\ \rho_{|00\rangle\langle10|} & \rho_{|00\rangle\langle11|} & \rho_{|01\rangle\langle10|} & \rho_{|01\rangle\langle11|} \\ \rho_{|10\rangle\langle00|} & \rho_{|10\rangle\langle01|} & \rho_{|11\rangle\langle00|} & \rho_{|11\rangle\langle01|} \\ \rho_{|10\rangle\langle10|} & \rho_{|10\rangle\langle11|} & \rho_{|11\rangle\langle10|} & \rho_{|11\rangle\langle11|} \end{pmatrix} \rightarrow \\ \begin{pmatrix} \rho_{|00\rangle\langle00|} & -\alpha\rho_{|00\rangle\langle01|} & -\alpha\rho_{|01\rangle\langle00|} & -\beta^*\rho_{|01\rangle\langle01|} \\ -\alpha\rho_{|00\rangle\langle10|} & \alpha^2\rho_{|00\rangle\langle11|} & \alpha^2\rho_{|01\rangle\langle10|} & \alpha\beta^*\rho_{|01\rangle\langle11|} \\ -\alpha\rho_{|10\rangle\langle00|} & -\alpha^2\rho_{|10\rangle\langle01|} & \alpha^2\rho_{|11\rangle\langle00|} & \alpha\beta^*\rho_{|11\rangle\langle01|} \\ -\beta\rho_{|10\rangle\langle10|} & \alpha\beta\rho_{|10\rangle\langle11|} & \alpha\beta\rho_{|11\rangle\langle10|} & |\beta|^2\rho_{|11\rangle\langle11|} \end{pmatrix}$$

$$(36)$$

where

$$\alpha = e^{-0.012291\gamma_{diss}} \tag{37}$$

$$\beta = e^{-0.0185454\gamma_{diss}}. (38)$$

This in turn causes the Ising phase gates $e^{iS_z^iS_z^j}$ to experience extra phase shifts of the form

$$e^{i\theta S_z^i S_z^j} \to e^{i\theta (S_z^i S_z^j + \Delta\theta_1 (S_z^i + S_z^j) + \Delta\theta_2 S_z^i S_z^j)}$$
 (39)

$$\Delta\theta_1 = \left(\frac{1-\alpha}{1+\alpha}\right)^2 \qquad \Delta\theta_2 = \left(\frac{\alpha-\beta}{\alpha+\beta}\right)^2.$$
 (40)

Which suggests that we can mimic this part of the error by modifying these phases directly, without implementing any noisy pulse-level simulations in the first place. If the energy and parameter profiles of this system match the respective system with Rydberg atom dissipation, we can deduce that the main source of error in the dissipative system comes from this coherent component. We can also separate this error from the rest of the dissipation by reversing all the phase shifts afterwards, creating a channel which only contains the incoherent part of the dissipation error.