

# Matrix-Model Simulations Using Quantum Computing, Deep Learning, and Lattice Monte Carlo

Enrico Rinaldi<sup>1,2,3,\*</sup>, Xizhi Han,<sup>4</sup> Mohammad Hassan<sup>5</sup>, Yuan Feng<sup>6</sup>, Franco Nori<sup>1,2,7</sup>,  
Michael McGuigan,<sup>8</sup> and Masanori Hanada<sup>9</sup>

<sup>1</sup>Physics Department, University of Michigan, Ann Arbor, Michigan 48109, USA

<sup>2</sup>Theoretical Quantum Physics Laboratory, Cluster for Pioneering Research, RIKEN, Wako, Saitama 351-0198, Japan

<sup>3</sup>Interdisciplinary Theoretical and Mathematical Sciences Program (iTHEMS), RIKEN, Wako, Saitama 351-0198, Japan

<sup>4</sup>Department of Physics, Stanford University, Stanford, California 94305-4060, USA

<sup>5</sup>City College of New York, New York, New York 10031, USA

<sup>6</sup>Pasadena City College, Pasadena, California 91106, USA

<sup>7</sup>RIKEN Center for Quantum Computing (RQC), Wako, Saitama 351-0198, Japan

<sup>8</sup>Brookhaven National Laboratory, Upton, New York 11973, USA

<sup>9</sup>Department of Mathematics, University of Surrey, Guildford GU2 7XH, United Kingdom



(Received 23 August 2021; accepted 20 December 2021; published 10 February 2022)

Matrix quantum mechanics plays various important roles in theoretical physics, such as a holographic description of quantum black holes, and it underpins the only practical numerical approach to the study of complex high-dimensional supergravity theories. Understanding quantum black holes and the role of entanglement in a holographic setup is of paramount importance for the realization of a quantum theory of gravity. Moreover, a complete numerical understanding of the holographic duality and the emergence of geometric space-time features from microscopic degrees of freedom could pave the way for new discoveries in quantum information science. Euclidean lattice Monte Carlo simulations are the *de facto* numerical tool for understanding the spectrum of large matrix models and have been used to test the holographic duality. However, they are not tailored to extract dynamical properties or even the quantum wave function of the ground state of matrix models. Quantum computing and deep learning provide potentially useful approaches to study the dynamics of matrix quantum mechanics. If successful in the context of matrix models, these rapidly improving numerical techniques could become the new Swiss army knife of quantum gravity practitioners. In this paper, we perform the first systematic survey for quantum computing and deep-learning approaches to matrix quantum mechanics, comparing them to lattice Monte Carlo simulations. These provide baseline benchmarks before addressing more complicated problems. In particular, we test the performance of each method by calculating the low-energy spectrum.

DOI: [10.1103/PRXQuantum.3.010324](https://doi.org/10.1103/PRXQuantum.3.010324)

## I. INTRODUCTION

Gauge-gravity duality [1,2] translates difficult (or intractable) problems in quantum gravity to well-defined problems in nongravitational quantum theories. Although it originated from string–M-theory, connections to various other fields, including quantum information theory,

condensed-matter theory, and cosmology, have been recognized by now, with research programs actively pursued. Quantum mechanics with matrix degrees of freedom (matrix QM, or matrix models in the following) plays an important role in gauge-gravity duality. More specifically, Yang–Mills–type matrix models can provide us with a non-perturbative formulation of superstring–M-theory [2–5]. In certain regions of the parameter space (i.e., strong coupling and large  $N$ ), weakly coupled gravity with small stringy corrections can be described by matrix models. Therefore, matrix models can offer us an ideal setup to study quantum corrections to gravity. By solving the matrix models, it should be possible to study fascinating problems such as the microscopic mechanism of black-hole evaporation.

\*erinaldi.work@gmail.com

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/) license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

Under the gauge-gravity (or holographic) duality setup, we seek a precise understanding of the quantum-mechanical solution to matrix models, including properties of the quantum ground-state wave function, quantum dynamical effects, and quantum entanglement patterns. This is our best shot at advancing our knowledge about puzzling concepts including the information-loss paradox. In the past, Monte Carlo simulations based on the Euclidean path-integral formulation have been used extensively to extract thermodynamic information from several matrix models; for a remarkable agreement between the D0-brane matrix model and type-IIA superstring theory, see, e.g., Ref. [6]. However, there are certain problems that it is hard to access via Monte Carlo methods, such as the Hamiltonian time evolution, which would be needed to understand dynamical phenomena such as graviton scattering. Therefore, it is important to develop alternative methods too. In this paper, we consider quantum simulation [7–11] and deep learning as such alternative approaches. Quantum simulations of lattice gauge theories have recently attracted considerable attention, since the pioneering work of Ref. [12], and deep learning is being used to accelerate simulations of lattice gauge theories similar to QCD [13]. These two rapidly growing numerical techniques hold the promise to transform the landscape of numerical simulations in quantum gravity, under the umbrella of “quantum gravity in the laboratory,” and matrix models are a simple and theoretically well-controlled setup to test them. Another interesting approach, outside the scope of this paper, is related to the numerical bootstrap [14–16].

We briefly introduce the pros and cons of these numerical approaches to matrix models:

*Monte Carlo simulations*—Monte Carlo simulations can be used to study problems that can be accessed by the Euclidean path integral, such as canonical thermodynamics and Euclidean correlation functions. Large-scale parallel simulations on high-performance computing (HPC) systems are doable and are done regularly. It is known how to improve the simulation results systematically, by accumulating more chain trajectories and simulating larger systems.

However, Monte Carlo simulation can only work when there is no sign problem [17].

Due to the sign problem, we cannot study the Hamiltonian time evolution. The sign problem may also appear at the very-low-temperature region of the matrix models. Note also that, even when there is no sign problem, we cannot see the quantum states in the Hilbert space directly. This might be a disadvantage for several interesting applications, such as decoding the emergent bulk geometry in holography along the lines of Ref. [18].

*Quantum computing*—On a quantum computer, quantum states are directly realized. Therefore, quantum computing

does not suffer from the sign problem and the Hamiltonian time evolution can be simulated straightforwardly. Some research directions that could benefit from direct quantum simulations are the formation and evaporation of a black hole, or a wormhole, and quantum teleportation [19].

The biggest challenge is that we do not yet have a reliable large-scale quantum computer. Although we can emulate a quantum computer on a classical computer, these simulations are restricted to systems with only a small number of qubits. Furthermore, the effects of the regularization needed to map a theory on a quantum device are not well understood yet.

The variational quantum eigensolver (VQE) [20–22] is expected to be a useful quantum-classical hybrid algorithm in the noisy intermediate-scale quantum (NISQ) era, because it can provide accurate results using shallower quantum circuits, less prone to noise. Being a variational method, the VQE can have systematic errors: *a priori*, we do not know if we can obtain the right answer, because the simulation might be trapped in a local optimum. This is a general problem for variational methods.

Quantum computing represents the computation in terms of qubits and quantum gates that can represent larger Hamiltonians and Hilbert spaces on quantum hardware than classical computers that are limited by memory constraints. Also, quantum computing represents fermions straightforwardly in a similar way to how bosons are represented and, as mentioned above, it is not susceptible to the sign problem like classical methods using Monte Carlo algorithms.

Some of the disadvantages of NISQ devices [23] include the lack of quantum error correction for quantum gates and the presence of noise, which can lead to inaccurate quantum gates and decoherence of the qubits during the computation. For the hybrid quantum-classical computations that we discuss here, the advantages are that many parts of a hybrid quantum-classical variational calculation, such as optimization, can be done efficiently on a classical computer, while expectation values can be computed on a quantum computer. This leads to a lower depth for the circuits, making them less susceptible to decoherence. However, these methods include complicated state-preparation protocols and variational ansatz for the wave function that need to be specified to have strong overlaps with the true ground state. Such wave functions may need a high-depth quantum circuit, which is difficult to implement on NISQ hardware and is prone to decoherence.

*Deep Learning*—Deep-learning methods have been applied successfully to the study of many-body quantum systems [24]. As one such approach, we consider a variational quantum Monte Carlo method with generative autoregressive flows: this method uses a deep neural network to

represent the probability distribution given by the system wave function [25]. This method may make the calculation on a classical computer much more efficient and we can obtain quantum states directly. This enables us to access a class of physics that cannot be handled via Markov chain Monte Carlo simulations of the Euclidean path integral.

The biggest disadvantage of this approach is that, *a priori*, we do not know if we can get the right answer; that is, systematic errors are difficult to estimate and the simulation might be trapped in a local optimum. This is the same problem as the one we mentioned above for the VQE. However, in practice, this approach has been used extensively across several many-body quantum systems [26–32], with neural quantum states becoming increasingly useful in real-world applications.

*Focus of this paper*—By listing the differences between the available numerical methods, it is clear that quantum computing and deep learning can, in principle, be very useful tools in solving matrix models. These allow a direct representation of quantum states (encoded in qubits or neural networks), which is needed to access the quantum information stored in the wave function. However, we do not know in practice if they can be useful. In this paper, we perform a systematic survey for quantum-simulation and deep-learning approaches that will serve as the first stepping stone for future research directions. Specifically, we focus on two important aspects that need to be addressed, before embarking on more complicated numerical efforts:

- (a) Quantum simulation requires a truncation of the Hilbert space when the underlying degrees of freedom can access an infinite tower of states, such as for bosonic particles. For SU(2) matrix models, we study the truncation effects by determining the low-energy states precisely via an exact diagonalization procedure. We find that the truncation effects are exponentially small with respect to the truncation level. This is not expected *a priori* and it represents the first original outcome of this paper.
- (b) We use the VQE and deep-learning methods to estimate the low-energy spectrum and compare them with other methods (exact diagonalization of a truncated Hamiltonian, lattice Monte Carlo simulation, and some exact results for supersymmetric models). The energy of the ground state is our first benchmark, telling us if the qubit and the neural-network representation of the Hamiltonian is correct. This is a mandatory first step for any future algorithm where an accurate representation of the Hamiltonian is needed, such as time evolution. For the deep-learning method, we observe a reasonable agreement, showing that there are specific neural

networks that are empirically well suited to represent matrix-model states. As for the VQE, the specific architecture that we use does not show a satisfactory performance at strong coupling, perhaps due to the variational forms parametrized by the quantum circuits not adequately probing the full gauge-invariant Hilbert space. This result shows that going beyond the VQE and using more complicated or fully quantum algorithms is not the correct way to approach matrix quantum mechanics for now, because they would require even deeper quantum circuits that are more prone to noise on actual quantum hardware.

This paper is organized as follows. In Sec. II, we define the models: a Yang-Mills–type bosonic matrix model and a supersymmetric matrix model, which we call “minimal BMN.” We can obtain a few exact results for the latter due to supersymmetry. In Sec. III A, we consider the quantum simulation approach. We introduce a simple scheme of the truncation of the Hilbert space and study the truncation effect on the energy spectrum and gauge symmetry. Then, in Sec. IV, we study the VQE as an actual example of the application of a quantum computer. In Sec. V, we estimate the ground-state energy via the deep-learning approach. In Sec. VI, we calculate the ground-state energy of the bosonic matrix model by performing lattice Monte Carlo simulation. We compare the results in Sec. VII.

The codes used to generate the data and make the figures are open source and we provide a web site with additional figures and tables in Ref. [33].

## II. MODELS

In this section, we define the two kinds of matrix models studied in this paper. The first model contains only bosonic degrees of freedom (DOF), while the second contains both bosonic and fermionic DOF in a supersymmetric setup.

### A. Bosonic matrix model

Let us start with the bosonic matrix model, i.e., a matrix model consisting of only bosonic degrees of freedom. We introduce  $d$  traceless Hermitian matrices  $X_1(t), X_2(t), \dots, X_d(t)$ , where  $t$  is time. We consider the following action (in the Minkowski signature):

$$S = N \int dt \text{Tr} \left( \frac{1}{2} \sum_I (D_t X_I)^2 - \frac{m^2}{2} \sum_I X_I^2 + \frac{\lambda}{4} \sum_{I \neq J} [X_I, X_J]^2 \right), \quad (1)$$

where  $\lambda = g^2 N$  and  $D_t$  is the covariant derivative with gauge field  $A_t$ , i.e.,  $D_t X_I = \partial_t X_I - i[A_t, X_I]$ . Via the  $SU(N)$  transformation  $X_I(t) \rightarrow \Omega(t) X_I(t) \Omega^{-1}(t)$ ,  $D_t X_I$  transforms as  $D_t X_I \rightarrow \Omega(t) [D_t X_I(t)] \Omega^{-1}(t)$ ; hence this action has the  $SU(N)$  gauge symmetry. The normalization used in Eq. (1) is convenient when we take the 't Hooft large- $N$  limit. The number of matrices  $d$  can be arbitrary; we consider  $d = 2$  in this paper. We can also use different normalization, by rescaling  $X_I$  with a factor  $\sqrt{N}$ :

$$S = \int dt \text{Tr} \left( \frac{1}{2} (D_t X_I)^2 - \frac{m^2}{2} X_I^2 + \frac{g^2}{4} [X_I, X_J]^2 \right). \quad (2)$$

In the operator formalism, the Hamiltonian is given by

$$\hat{H} = \text{Tr} \left( \frac{1}{2} \hat{P}_I^2 + \frac{m^2}{2} \hat{X}_I^2 - \frac{g^2}{4} [\hat{X}_I, \hat{X}_J]^2 \right), \quad (3)$$

where

$$\hat{P}_I = \sum_{\alpha=1}^{N^2-1} \hat{P}_I^\alpha \tau_\alpha, \quad \hat{X}_I = \sum_{\alpha=1}^{N^2-1} \hat{X}_I^\alpha \tau_\alpha. \quad (4)$$

$\tau_\alpha$  are the generators of  $SU(N)$  normalized as  $\text{Tr}(\tau_\alpha \tau_\beta) = \delta_{\alpha\beta}$ . The canonical commutation relation is

$$[\hat{X}_{I\alpha}, \hat{P}_{J\beta}] = i \delta_{IJ} \delta_{\alpha\beta}. \quad (5)$$

The Hamiltonian and the canonical commutation relation are invariant under the  $SU(N)$  transformation  $\hat{X}_{I,ij} \rightarrow (\Omega \hat{X}_I \Omega^{-1})_{ij}$ ,  $\hat{P}_{I,ij} \rightarrow (\Omega \hat{P}_I \Omega^{-1})_{ij}$ . The physical states are restricted to singlets under this  $SU(N)$  transformation (gauge singlets). We denote the Hilbert space spanned by gauge singlets by  $\mathcal{H}_{\text{inv}}$ . We can also consider a bigger ‘‘extended’’ Hilbert space  $\mathcal{H}_{\text{ext}}$  that contains nonsinglets. Because operators  $\hat{X}_I$  and  $\hat{P}_I$  are not gauge invariant, they are defined naturally on  $\mathcal{H}_{\text{ext}}$ .

It is tedious but straightforward to show the equivalence between the path-integral formalism and the operator formalism; see Appendix A.

## B. Supersymmetric matrix model (minimal BMN)

The BMN matrix model [5] is a one-parameter deformation of the D0-brane quantum mechanics that preserves all supersymmetries. In Ref. [34], low-supersymmetry analogs of the BMN matrix model are listed [35]. The one corresponding to 3D  $\mathcal{N} = 1$  SYM [Eq. (6.3) in Ref. [34]] has the minimal degrees of freedom. So let us denote this model ‘‘minimal BMN.’’

Gamma matrices in three dimensions can be chosen as  $\gamma^0 = i\sigma_3$ ,  $\gamma^1 = \sigma_1$  and  $\gamma^2 = \sigma_2$ . The charge conjugation matrix  $C$ , which satisfies  $C^{-1} \gamma_\mu C = -\gamma_\mu^T$ , can be chosen

as  $C = i\sigma_2$ . The Majorana condition  $C\psi = \gamma_0^T \psi^*$  can be solved by  $\psi = \begin{pmatrix} \xi \\ i\xi^* \end{pmatrix}$ . When  $\psi$  is a matrix,  $\psi = \begin{pmatrix} \xi \\ i\xi^\dagger \end{pmatrix}$ .

Having this convention in mind, the Hamiltonian of minimal BMN is expressed as [34]

$$\begin{aligned} \hat{H} &= \text{Tr} \left( \frac{1}{2} \hat{P}_I^2 - \frac{g^2}{4} [\hat{X}_I, \hat{X}_J]^2 + \frac{g}{2} \hat{\psi} \gamma^I [\hat{X}_I, \hat{\psi}] \right. \\ &\quad \left. - \frac{3i\mu}{4} \hat{\psi} \hat{\psi} + \frac{\mu^2}{2} \hat{X}_I^2 \right) - (N^2 - 1)\mu \\ &= \text{Tr} \left( \frac{1}{2} \hat{P}_I^2 - \frac{g^2}{2} [\hat{X}_1, \hat{X}_2]^2 + \frac{g}{2} \hat{\xi} [-\hat{X}_1 - i\hat{X}_2, \hat{\xi}] \right. \\ &\quad \left. + \frac{g}{2} \hat{\xi}^\dagger [-\hat{X}_1 + i\hat{X}_2, \hat{\xi}^\dagger] + \frac{3\mu}{2} \hat{\xi}^\dagger \hat{\xi} + \frac{\mu^2}{2} \hat{X}_I^2 \right) \\ &\quad - (N^2 - 1)\mu. \end{aligned} \quad (6)$$

The last term  $-(N^2 - 1)\mu$  causes the ground-state energy to be zero [36]. Although the fermion number is not conserved because there are interaction terms including two  $\hat{\xi}$  or two  $\hat{\xi}^\dagger$ , fermion parity (i.e., fermion number even or odd) is conserved.

### 1. Symmetry of the model

The supersymmetry transformation is given by [34]

$$\delta_\epsilon \cdot = [\hat{Q}\epsilon^* + \hat{Q}^\dagger \epsilon, \cdot], \quad (7)$$

where

$$\begin{aligned} \hat{Q} &= -\hat{\xi}_\alpha^\dagger \left[ (\hat{P}_1^\alpha - i\hat{P}_2^\alpha) - i\mu(\hat{X}_1^\alpha - i\hat{X}_2^\alpha) \right] \\ &\quad - \frac{ig}{\sqrt{2}} f_{\alpha\beta\gamma} \hat{\xi}^\alpha \hat{X}_1^\beta \hat{X}_2^\gamma. \end{aligned} \quad (8)$$

Here,  $f_{abc}$  is the structure constant of  $SU(N)$ , which is  $\epsilon_{abc}$  for  $SU(2)$ .

It is convenient to use

$$\hat{Z} = \frac{X_1 - iX_2}{\sqrt{2}}, \quad \hat{P}_Z = \frac{\hat{P}_1 - i\hat{P}_2}{\sqrt{2}}, \quad (9)$$

which satisfy

$$[\hat{Z}, \hat{P}_Z^\dagger] = [\hat{Z}^\dagger, \hat{P}_Z] = i, \quad [\hat{Z}, \hat{P}_Z] = [\hat{Z}^\dagger, \hat{P}_Z^\dagger] = 0. \quad (10)$$

The Hamiltonian is

$$\hat{H} = \text{Tr} \left( \hat{P}_Z \hat{P}_Z^\dagger + \frac{g^2}{2} [\hat{Z}, \hat{Z}^\dagger]^2 - \frac{g}{\sqrt{2}} \hat{\xi} [\hat{Z}^\dagger, \hat{\xi}] - \frac{g}{\sqrt{2}} \hat{\xi}^\dagger [\hat{Z}, \hat{\xi}^\dagger] + \frac{3\mu}{2} \hat{\xi}^\dagger \hat{\xi} + \mu^2 \hat{Z} \hat{Z}^\dagger \right). \quad (11)$$

Hence the SO(2) rotation is

$$\hat{Z} \rightarrow e^{i\theta} \hat{Z}, \quad \hat{\xi} \rightarrow e^{i\theta/2} \hat{\xi}. \quad (12)$$

The generator of SO(2) is

$$\hat{M} = \sum_{\alpha} \left( i(\hat{Z}_{\alpha} \hat{P}_{Z\alpha}^\dagger - \hat{Z}_{\alpha}^\dagger \hat{P}_{Z\alpha}) - \frac{1}{2} \hat{\xi}_{\alpha}^\dagger \hat{\xi}_{\alpha} \right). \quad (13)$$

Note that  $\hat{H}$  and  $\hat{M}$  commute. The supercharge is written as

$$\hat{Q} = -\sqrt{2} \hat{\xi}_{\alpha}^\dagger \left( \hat{P}_{Z\alpha} - i\mu \hat{Z}_{\alpha} \right) - \frac{g}{\sqrt{2}} f_{\alpha\beta\gamma} \hat{\xi}_{\alpha} \hat{Z}_{\beta} \hat{Z}_{\gamma}^\dagger. \quad (14)$$

This satisfies

$$\hat{Q}^2 = -i \hat{Z}_{\alpha} \hat{G}_{\alpha}, \quad \hat{Q}^{\dagger 2} = i \hat{Z}_{\alpha}^\dagger \hat{G}_{\alpha}, \quad (15)$$

and

$$\{\hat{Q}, \hat{Q}^\dagger\} = 2 \left( \hat{H} - \mu \hat{M} \right). \quad (16)$$

This means that  $\hat{H} - \mu \hat{M}$  is positive semidefinite and that the supersymmetric states [also called the Bogomol'nyi-Prasad-Sommerfield (BPS) states [37]] should satisfy

$$\hat{Q} |\text{BPS}\rangle = \hat{Q}^\dagger |\text{BPS}\rangle = \left( \hat{H} - \mu \hat{M} \right) |\text{BPS}\rangle = 0. \quad (17)$$

When restricted to the gauge-invariant Hilbert space,  $\hat{Q}$  and  $\hat{Q}^\dagger$  commute with  $\hat{H} - \mu \hat{M}$ . Hence any non-BPS states form a pair consisting of states in fermion-number even and odd sectors. Therefore, when the coupling is turned on, the BPS states must remain BPS. In particular, the SO(2)-invariant ground state has to stay at zero energy. This property is useful for a sanity check of numerical computations.

### III. HAMILTONIAN TRUNCATION AND QUANTUM SIMULATION

On a quantum computer, the dimension of the Hilbert space is finite, because the number of qubits is finite. On the other hand, bosons require an infinite-dimensional Hilbert space. The matrix models contain bosons and, hence, we have to truncate the Hilbert space. The Hamiltonian is also truncated to a matrix of finite size and, hence, we call this method ‘‘Hamiltonian truncation’’ (HT).

In the Hamiltonian-truncation method, it is important to control the truncation effects. How the full theory is recovered when the truncation level is taken to infinity is a well-known problem in many-body quantum physics [38].

In this section, we test the performance of the Hamiltonian-truncation method. Specifically, we study the truncation effect to estimate the amount of resources needed for precise quantum simulations. As a concrete truncation scheme, we use a simple Fock-space truncation. To confirm the validity of this approach, we implement the regularized Hamiltonian on a classical computer and study its properties numerically. The numerical results in this section are obtained by using QuTiP [39] (QUANTUM TOOLBOX in PYTHON) [40,41].

In Sec. III A, we regularize the Hamiltonian by using the Fock basis and explain how it can be written in terms of qubits. We then discuss the gauge-singlet constraint in Sec. III B. In Sec. III C, we numerically study the regularized Hamiltonian and estimate the truncation effects. The results in these sections tell us how we can control the truncation effects. In Sec. IV, we use the truncation studied in this section together with the VQE, a promising hybrid quantum-classical tool available for NISQ-era hardware.

#### A. Regularization of the Hamiltonian

We start by truncating the infinite-dimensional Hamiltonian to a finite dimension. Here, we closely follow Ref. [42] and use the Fock-space truncation method [43]. As a concrete example, let us consider the bosonic SU( $N$ ) matrix model given in Eq. (3). We can regularize the minimal BMN in a very similar manner, as we see in Sec. III C 2.

We write the matrices as in Eq. (4), where  $\tau_{\alpha}$  in Eq. (4) is the generator of SU( $N$ ) satisfying  $\text{Tr}(\tau_{\alpha} \tau_{\beta}) = \delta_{\alpha\beta}$  and  $[\tau_{\alpha}, \tau_{\beta}] = i f_{\alpha\beta\gamma} \tau_{\gamma}$ . (The values of the structure constant  $f_{\alpha\beta\gamma}$  depend on the detail of the choice of  $\tau_{\alpha}$ .) Then, the Hamiltonian given in Eq. (3) is written as

$$\hat{H} = \sum_{\alpha, I} \left( \frac{1}{2} \hat{P}_{I\alpha}^2 + \frac{m^2}{2} \hat{X}_{I\alpha}^2 \right) + \frac{g^2}{4} \sum_{\gamma, I, J} \left( \sum_{\alpha, \beta} f_{\alpha\beta\gamma} \hat{X}_{I\alpha}^{\alpha} \hat{X}_{J}^{\beta} \right)^2. \quad (18)$$

The canonical commutation relation is given by Eq. (5). We introduce the creation and annihilation operators as

$$\hat{a}_{I\alpha}^{\dagger} = \sqrt{\frac{m}{2}} \hat{X}_{I\alpha} - \frac{i \hat{P}_{I\alpha}}{\sqrt{2m}}, \quad \hat{a}_{I\alpha} = \sqrt{\frac{m}{2}} \hat{X}_{I\alpha} + \frac{i \hat{P}_{I\alpha}}{\sqrt{2m}}. \quad (19)$$

By using the creation and annihilation operators and the number operator  $\hat{n}_{I\alpha} = \hat{a}_{I\alpha}^{\dagger} \hat{a}_{I\alpha}$ , the Hamiltonian can be

expressed as

$$\hat{H} = m \sum_{\alpha, I} \left( \hat{n}_{I\alpha} + \frac{1}{2} \right) + \frac{g^2}{16m^2} \sum_{\gamma, I, J} \left( \sum_{\alpha, \beta} f_{\alpha\beta\gamma} (\hat{a}_{I\alpha} + \hat{a}_{I\alpha}^\dagger) (\hat{a}_{J\beta} + \hat{a}_{J\beta}^\dagger) \right)^2. \quad (20)$$

The generators of the gauge transformation are

$$\hat{G}_\alpha = i \sum_{\beta, \gamma, I} f_{\alpha\beta\gamma} \hat{a}_{I\beta}^\dagger \hat{a}_{I\gamma}. \quad (21)$$

The Hamiltonian is gauge invariant or, equivalently,

$$[\hat{H}, \hat{G}_\alpha] = 0. \quad (22)$$

For each  $(I, \alpha)$ , we can take the Fock vacuum  $|0\rangle_{I\alpha}$ , which satisfies

$$\hat{a}_{I\alpha} |0\rangle_{I\alpha} = 0, \quad (23)$$

and the excited states  $|n\rangle_{I\alpha}$  defined by

$$|n\rangle_{I\alpha} = \frac{(\hat{a}_{I\alpha}^\dagger)^n}{\sqrt{n!}} |0\rangle_{I\alpha}. \quad (24)$$

We can take the Fock states for the matrix model by taking the tensor product, as

$$| \{n_{I\alpha}\} \rangle = \otimes_{I,\alpha} |n\rangle_{I\alpha}. \quad (25)$$

Note that the Fock vacuum is gauge invariant, i.e.,

$$\hat{G}_\alpha (\otimes_{I,\beta} |0\rangle_{I\beta}) = 0 \quad (26)$$

for all gauge generators  $\hat{G}_\alpha$ . As a regularization, we truncate the Hilbert space such that the excitation level of each oscillator is below  $\Lambda$ . We define the truncated raising, lowering, and number operators to be

$$\begin{aligned} \hat{a}_{\text{truncated}}^\dagger &= \sum_{n=0}^{\Lambda-2} \sqrt{n+1} |n+1\rangle \langle n|, & \hat{a}_{\text{truncated}} \\ &= \sum_{n=0}^{\Lambda-2} \sqrt{n+1} |n\rangle \langle n+1|, \end{aligned} \quad (27)$$

and

$$\hat{n}_{\text{truncated}} = \sum_{n=0}^{\Lambda-1} n |n\rangle \langle n|. \quad (28)$$

To define the truncated versions of the Hamiltonian and gauge generators, we replace  $\hat{n}$ ,  $\hat{a}$ , and  $\hat{a}^\dagger$  in Eqs. (20)

and (21) with  $\hat{n}_{\text{truncated}}$ ,  $\hat{a}_{\text{truncated}}$ , and  $\hat{a}_{\text{truncated}}^\dagger$ , respectively. Note that the truncation breaks gauge symmetry manifestly, i.e.,

$$[\hat{H}_{\text{truncated}}, \hat{G}_{\alpha, \text{truncated}}] \neq 0. \quad (29)$$

In the truncated Fock basis, the creation and annihilation operators are expressed as simple Pauli strings (i.e., tensor products of Pauli matrices), which are convenient for quantum computations [42]. In Eqs. (27) and (28), let  $|n\rangle$  ( $n = 0, 1, \dots, \Lambda - 1$ ) be the  $j$ th excited state of an oscillator. We can write  $j$  in terms of binaries as  $n = \sum_{l=0}^{K-1} b_l 2^l$ . By using  $K = \log_2 \Lambda$  qubits, we can rewrite the state  $|j\rangle$  as

$$|n\rangle = |b_0\rangle |b_1\rangle \dots |b_{K-1}\rangle. \quad (30)$$

Writing  $|n\rangle = |b_0\rangle |b_1\rangle \dots |b_{K-1}\rangle$  and  $|n+1\rangle = |b'_0\rangle |b'_1\rangle \dots |b'_{K-1}\rangle$ , we can express  $|n+1\rangle \langle n|$  as an operator in this basis as

$$|n+1\rangle \langle n| = \otimes_{l=0}^{K-1} (|b'_l\rangle \langle b_l|). \quad (31)$$

Note that each  $|b'_l\rangle \langle b_l|$  is a linear combination of the Pauli matrices:

$$\begin{aligned} |0\rangle \langle 0| &= \frac{\mathbf{1}_2 - \sigma_z}{2}, & |1\rangle \langle 1| &= \frac{\mathbf{1}_2 + \sigma_z}{2}, \\ |0\rangle \langle 1| &= \frac{\sigma_x + i\sigma_y}{2}, & |1\rangle \langle 0| &= \frac{\sigma_x - i\sigma_y}{2}. \end{aligned} \quad (32)$$

The annihilation operator  $\hat{a}_{\text{truncated}}$  and the number operator  $\hat{n}_{\text{truncated}}$  have similar forms.

## B. Gauge singlets and nonsinglets

The physical states in gauge theory are gauge singlets (gauge-invariant states). Let us denote the Hilbert space spanned by singlets by  $\mathcal{H}_{\text{inv}}$ . Our construction involves the extended Hilbert space  $\mathcal{H}_{\text{ext}}$ , which contains gauge-nonsinglet states. Depending on the problem under consideration, this may or may not be an issue. First, it is not a problem as long as one considers an exact time evolution of the exact Hamiltonian (without truncation), for the following reasons:

- Time evolution commutes with the gauge transformation, i.e.,  $[\hat{G}_\alpha, e^{-i\hat{H}t}] = 0$ . Therefore, if we take the initial state to be in  $\mathcal{H}_{\text{inv}}$ , it stays in  $\mathcal{H}_{\text{inv}}$ .
- From any nonsinglet state  $|\phi\rangle$ , we can obtain a singlet state by projecting [44] it to  $\mathcal{H}_{\text{inv}}$ . We can regard  $|\phi\rangle$  as a gauge-fixed state. The projection and time evolution commute with each other. Often, the gauge-fixed states are easier to handle.

The situation can be different when a quantum simulation is considered. By assuming that the noise is completely

removed, there are two sources of errors. The first is the error associated with the truncation of the Hilbert space. The second is the error associated with the approximation of  $e^{-i\hat{H}t}$ , such as the Trotter-Suzuki error. For each algorithm, in principle, it is straightforward to make the second type of error as small as we want. Hence let us focus on the first type of error. Namely, we assume that  $e^{-i\hat{H}_{\text{truncated}}t}$  is constructed precisely. We need to truncate the generators of the gauge transformation  $\hat{G}_\alpha$  as well. Due to the truncation error, neither  $\hat{G}_\alpha$  nor  $\hat{G}_{\alpha,\text{truncated}}$  commute with  $\hat{H}_{\text{truncated}}$ . Therefore, even if the initial state is taken to be (approximately) gauge invariant, the singlet condition may be broken badly after a long time.

A related problem is that when we calculate low-lying modes of the Hamiltonian, many nonsinglet modes, which are not necessarily of interest, are obtained. We might be able to circumvent these problems by adding a term proportional to  $\hat{G}^2 \equiv \sum_\alpha \hat{G}_\alpha^2$  to the Hamiltonian [45–49]:

$$\hat{H}' = \hat{H} + c \sum_\alpha \hat{G}_\alpha^2. \quad (33)$$

Here,  $c$  is a free parameter.

Let  $|E'\rangle$  be an eigenstate of  $\hat{H}'$  with eigenvalue  $\lambda_{E'}$ , i.e.,  $\hat{H}'|E'\rangle = \lambda_{E'}|E'\rangle$ . Let  $E'$  be the expectation value of the original Hamiltonian  $\hat{H}$  for eigenstate  $|E'\rangle$ , i.e.,  $E' = \langle E'|\hat{H}|E'\rangle$ . Then, by definition,  $\lambda_{E'} - E' = c \sum_\alpha \langle E'|\hat{G}_\alpha^2|E'\rangle$ . We use  $c = \Lambda$  later in this paper. Because gauge-singlet states are zero modes of  $\sum_\alpha \hat{G}_\alpha^2$ , such a deformation makes the nonsinglets heavy and they decouple from the low-energy dynamics. A nontrivial issue here is whether  $\lambda_{E'}$  converges to the energy eigenvalue of the original Hamiltonian. As we will see, this is indeed the case, as  $\sum_\alpha \langle E'|\hat{G}_\alpha^2|E'\rangle$  decays exponentially fast as  $\Lambda$  becomes large.

If the initial state is taken to be sufficiently low energy in terms of  $\hat{H}'_{\text{truncated}}$  and  $e^{-i\hat{H}'_{\text{truncated}}t}$  is calculated precisely, then the singlet condition is preserved with a good precision. Note that even with this deformation, the Hamiltonian can be written in a simple form in terms of the Pauli strings. Therefore, we can study this modified Hamiltonian efficiently on a quantum computer.

## C. Energy spectrum via classical computation

### 1. SU(2) bosonic matrix model

We consider the SU(2) two-matrix model. Following Sec. III A, we can easily write down the Hamiltonian explicitly, by taking  $\tau_\alpha = \sigma_\alpha/\sqrt{2}$  and  $f_{\alpha\beta\gamma} = \sqrt{2}\epsilon_{\alpha\beta\gamma}$ , where  $\epsilon_{123} = \epsilon_{231} = \epsilon_{312} = +1$ ,  $\epsilon_{321} = \epsilon_{213} = \epsilon_{132} = -1$ .

The Hamiltonian is sparse (only a small fraction of the matrix elements are nonzero) and we can make efficient use of optimized sparse linear-algebra solvers from the ARPACK [50] software suite, such as the implicitly

restarted Arnoldi methods [51], to find the eigenvalues and eigenvectors.

First, we use the setup without deformation, i.e.,  $c = 0$  and hence  $\hat{H}' = \hat{H}$  [cf. Eq. (33)]. We obtain the ground state, which we denote by  $|E_0\rangle$ . We estimate the ground-state energy  $E_0 = \langle E_0|\hat{H}|E_0\rangle$  and the violation of the singlet constraint  $\sum_\alpha \langle E_0|\hat{G}_\alpha^2|E_0\rangle$  for the cutoff  $\Lambda = 3, 4, \dots, 14$  and the couplings  $\lambda = g^2N = 0.2, 0.5, 1.0$ , and 2.0. We plot them in Fig. 1 and Fig. 2, respectively. The tables corresponding to the plots are reported in Appendix B.

At each value of the coupling constant, we can see that  $\sum_\alpha \langle E_0|\hat{G}_\alpha^2|E_0\rangle$  scales as  $e^{-a\Lambda}$ , with the same exponent  $a$  but different overall factors for even and odd  $\Lambda$ ; see Fig. 2. This suggests that we can take  $c$  to be an arbitrary power of  $\Lambda$ ; we expect that as long as it is smaller than  $e^{+a\Lambda}$ , it does not affect the ground state and probably the excited states as well. (We will confirm this shortly, by taking  $c = \Lambda$ .)

In Fig. 1, the approach of the ground-state energy to the large- $\Lambda$  limit appears to be exponentially fast. To confirm the exponential decay manifestly, we plot the difference of the ground-state energies for cutoffs differing by one unit,  $E_0^{\text{diff}}(\Lambda) \equiv |E_0(\Lambda) - E_0(\Lambda - 1)|$ , and we can see that it decays exponentially with  $\Lambda$  in Fig. 3.

Next, we consider the setup with deformation  $\hat{H}' = \hat{H} + c\hat{G}^2$  [cf. Eq. (33)]. We take  $c = \Lambda$ , such that the nonsinglet modes completely decouple in the limit  $\Lambda \rightarrow \infty$ . For this deformed Hamiltonian  $\hat{H}'$ , we compute the low-lying eigenstates  $|E'\rangle$ . Let us first set the mass and coupling to  $m^2 = 1$  and  $\lambda = 0.2$ , respectively.

In Fig. 4, we compare the five eigenstates of  $\hat{H}$  with the smallest values of  $\sum_\alpha \langle E|\hat{G}_\alpha^2|E\rangle$  and the five eigenstates of  $\hat{H}'$  with the smallest eigenvalues  $\lambda_{E'} = \langle E'|\hat{H}'|E'\rangle$ , for  $c = \Lambda = 4$ ,  $m^2 = 1$ , and  $\lambda = 0.2$ . We can see the deformation does not affect the low-lying modes; it just removes the nonsinglet modes.

As we can see from Fig. 5, up to the fourth excited level we already achieve a reasonably good convergence at  $\Lambda = 5$  or  $\Lambda = 6$ . Higher excited modes receive larger truncation effects, as expected. We can see that the violation of the singlet constraint  $\sum_\alpha \langle E'_n|\hat{G}_\alpha^2|E'_n\rangle$  decays exponentially with  $\Lambda$ . We can see the same pattern for  $m^2 = 1$ ,  $\lambda = 1.0$  in Fig. 6. Therefore, both  $\lambda_{E'} = \langle E'|\hat{H}'|E'\rangle$  and  $E' \equiv \langle E'|\hat{H}|E'\rangle$  converge to the correct energy eigenvalue.

Because we add  $\hat{G}^2$  to the Hamiltonian, the low-lying modes of  $\hat{H}'$  are gauge-singlet states. In the weak-coupling limit ( $\lambda \rightarrow 0$ ), the five lowest-energy levels in the gauge-singlet sector are  $E'_0 = 3$ ,  $E'_1 = E'_2 = E'_3 = 5$ , and  $E'_4 = 7$ . Values closer to them are observed at  $\lambda = 0.2$  (cf. Fig. 5).

At  $\lambda = 0$ ,  $|E'_{1,2,3}\rangle$  are linear combinations of  $\text{Tr}(\hat{a}_1^\dagger \hat{a}_1^\dagger)|0\rangle$ ,  $\text{Tr}(\hat{a}_1^\dagger \hat{a}_2^\dagger)|0\rangle$ , and  $\text{Tr}(\hat{a}_2^\dagger \hat{a}_2^\dagger)|0\rangle$ . With respect to the SO(2) symmetry mixing  $\hat{a}_1^\dagger$  and  $\hat{a}_2^\dagger$ , they split to the singlet

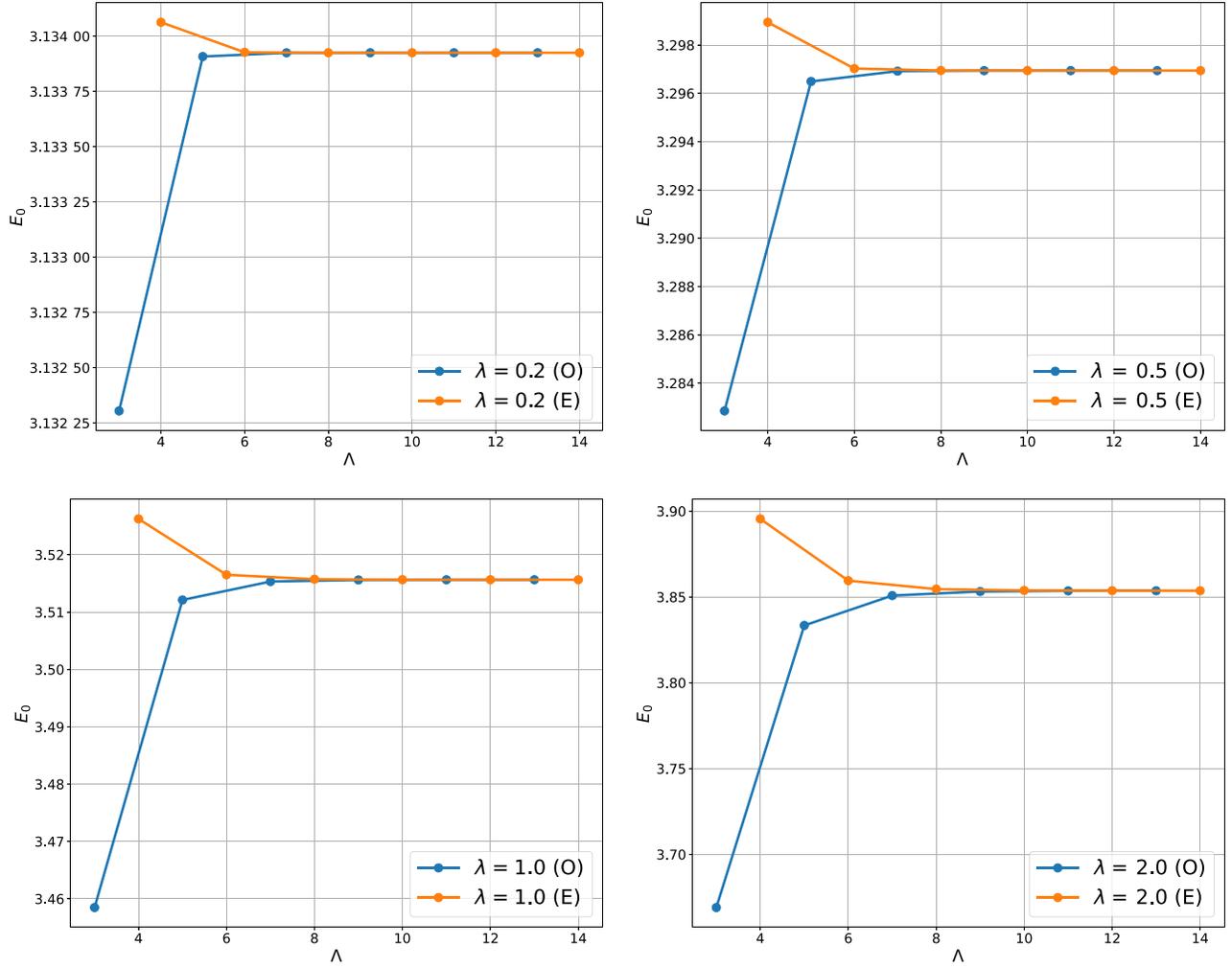


FIG. 1. The ground-state energy  $E_0$  as a function of the cutoff  $\Lambda$  for various couplings  $\lambda = g^2N = 0.2, 0.5, 1.0,$  and  $2.0$  for the  $SU(2)$  bosonic model. Even (E) and odd (O) values of  $\Lambda$  are plotted with different colors. The other parameters are  $m^2 = 1$  and  $c = 0$ .

$\text{Tr}(\hat{a}_1^\dagger \hat{a}_1^\dagger + \hat{a}_2^\dagger \hat{a}_2^\dagger)|0\rangle$  and the doublet. In Fig. 5 ( $\lambda = 0.2$ ) and Fig. 6 ( $\lambda = 1.0$ ), we can see the degeneracy of  $E'_1$  and  $E'_2$  as  $\Lambda \rightarrow \infty$ ; hence it is natural to expect that the first and second excited levels form the doublet, while the third excited level is the singlet. To check this, we define the generator of the  $SO(2)$  rotation  $\hat{M}$ , which acts on the states as  $\hat{M}f(\hat{a}_1^\dagger, \hat{a}_2^\dagger)|0\rangle = f(-\hat{a}_2^\dagger, \hat{a}_1^\dagger)|0\rangle$ , and calculate  $\langle E'_n | \hat{M} | E'_n \rangle$ . We obtain  $+1, -1, -1, +1, +1$  for  $n = 0, 1, 2, 3, 4$ , both for  $\lambda = 0.2$  and  $\lambda = 1.0$ , with good precision for  $\Lambda \geq 4$ .

## 2. $SU(2)$ minimal BMN

The  $SU(2)$  minimal BMN has six bosonic degrees of freedom and three fermionic degrees of freedom. With the truncation scheme for the bosons discussed earlier in this paper, the dimension of the truncated Hilbert space is  $8\Lambda^6$ , because the dimension of the fermionic Hilbert space is

$2^3 = 8$ . The Hamiltonian is

$$\begin{aligned} \hat{H} = & \sum_{\alpha} \left( \frac{\hat{P}_{1\alpha}^2}{2} + \frac{\hat{P}_{2\alpha}^2}{2} + \frac{\mu^2 \hat{X}_{1\alpha}^2}{2} + \frac{\mu^2 \hat{X}_{2\alpha}^2}{2} + \frac{3\mu}{2} \hat{\xi}_{\alpha}^{\dagger} \hat{\xi}_{\alpha} \right) \\ & + g^2 \sum_{\alpha \neq \beta} \hat{X}_{1\alpha}^2 \hat{X}_{2\beta}^2 - 2g^2 \sum_{\alpha < \beta} \hat{X}_{1\alpha} \hat{X}_{1\beta} \hat{X}_{2\alpha} \hat{X}_{2\beta} \\ & + \frac{ig}{\sqrt{2}} \sum_{\alpha, \beta, \gamma} \epsilon_{\alpha\beta\gamma} \left( (-\hat{X}_{1\alpha} - i\hat{X}_{2\alpha}) \hat{\xi}_{\beta}^{\dagger} \hat{\xi}_{\gamma}^{\dagger} \right. \\ & \left. + (-\hat{X}_{1\alpha} + i\hat{X}_{2\alpha}) \hat{\xi}_{\beta} \hat{\xi}_{\gamma} \right) - 3\mu. \end{aligned} \quad (34)$$

The complex fermion operators  $\xi_{\alpha}$  ( $\alpha = 1, 2, 3$ ) satisfy the canonical anticommutation relation  $\{\xi_{\alpha}^{\dagger}, \xi_{\beta}\} = \delta_{\alpha\beta}$ . Regarding the fermions, we can use the standard Fock basis and build states from the Fock vacuum  $\hat{\xi}_{\alpha}|0\rangle = 0$ . For  $SU(2)$ , as we only have three fermions, one can use the

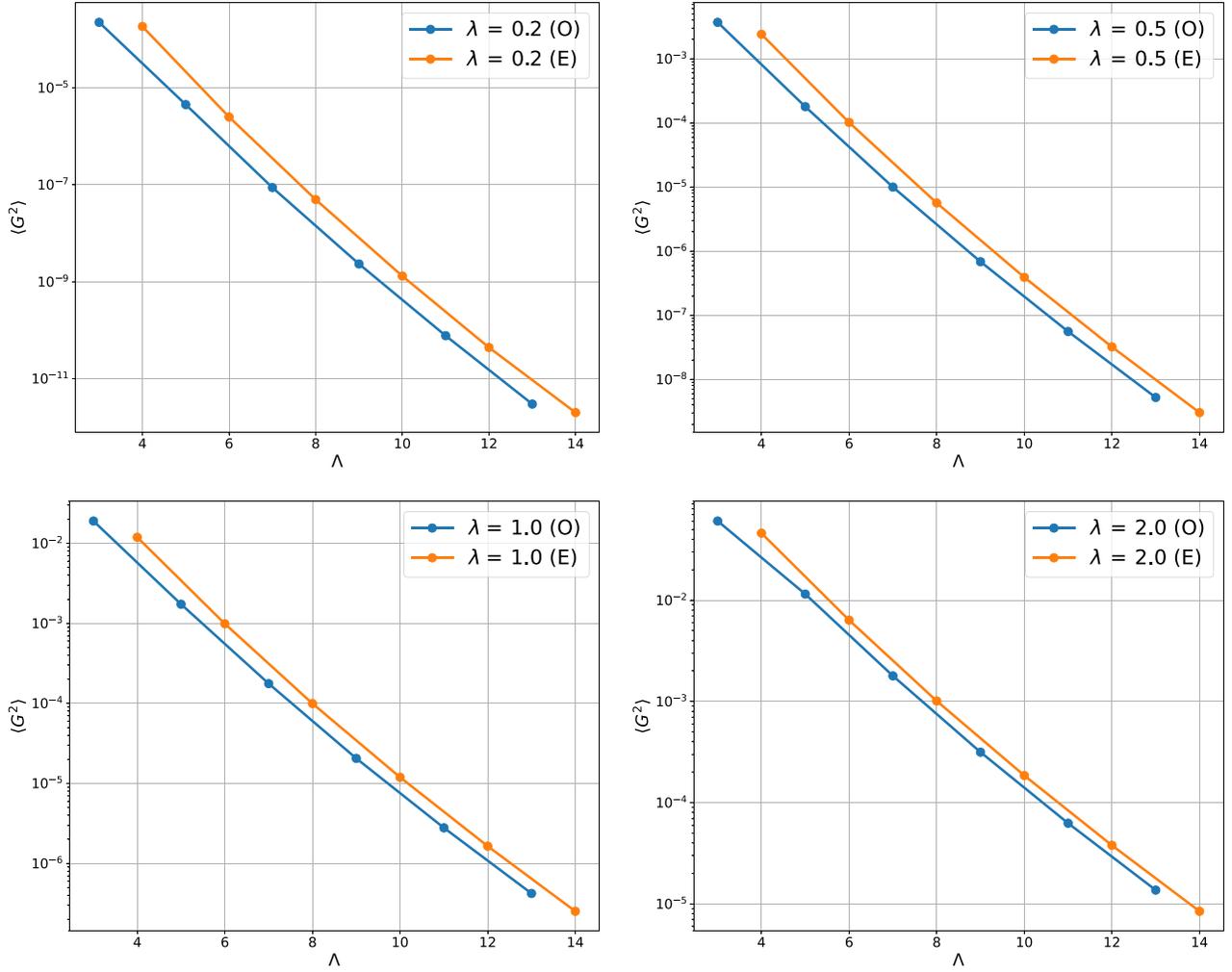


FIG. 2. The violation of the singlet constraint  $\sum_{\alpha} \langle E_0 | \hat{G}_{\alpha}^2 | E_0 \rangle$  as a function of the cutoff  $\Lambda$  for various couplings  $\lambda = g^2 N = 0.2, 0.5, 1.0,$  and  $2.0$  for the  $SU(2)$  bosonic model. Even (E) and odd (O) values of  $\Lambda$  are plotted with different colors in logarithmic scale. The other parameters are  $m^2 = 1$  and  $c = 0$ .

Jordan-Wigner transformation to form fermion operators from Pauli spin matrices. However, for  $SU(N)$  with larger values of  $N$ , the number of fermions increases and it will be more efficient to adapt Bravyi-Kitaev [52] or Bravyi-Kitaev [53] superfast methods of constructing fermion operators. This has been successfully implemented in quantum chemistry applications. In addition, the use of the symmetry present in the Hamiltonian (which has been used in quantum chemistry [54] and other many-body quantum physics systems) can also be used to reduce the number of qubits that will be necessary to represent the fermions for large  $N$ . Such symmetries include discrete symmetries such as parity as well as continuous symmetries such as gauge invariance. In this work, we adopt a simple approach disregarding such symmetries, leading to a less favorable scaling of resources with respect to the size of the system, which can be improved in the future. As already

mentioned, for the simplest case of  $N = 2$ , the annihilation operator for each fermion  $\hat{\xi}_{\alpha}$  is constructed using the Jordan-Wigner prescription in the  $2^3$ -dimensional Hilbert space using Pauli matrices [55]:

$$\hat{\xi}_{\alpha} = \underbrace{\sigma_z \otimes \sigma_z \otimes \cdots \otimes \sigma_z}_{\alpha-1 \text{ times}} \otimes \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \otimes \mathbf{1} \otimes \cdots \otimes \mathbf{1}, \quad (35)$$

where  $\mathbf{1}$  is the  $2 \times 2$  identity matrix. We can define the fermion number by counting how many  $\hat{\xi}_{\alpha}^{\dagger}$  operators act on the Fock vacuum. The Hamiltonian does not mix the fermion-number odd and even sectors.

The generators of the gauge transformation are

$$\hat{G}_{\alpha} = i \sum_{\beta, \gamma} f_{\alpha\beta\gamma} \left( \hat{a}_{1\beta}^{\dagger} \hat{a}_{1\gamma} + \hat{a}_{2\beta}^{\dagger} \hat{a}_{2\gamma} + \hat{\xi}_{\beta}^{\dagger} \hat{\xi}_{\gamma} \right). \quad (36)$$

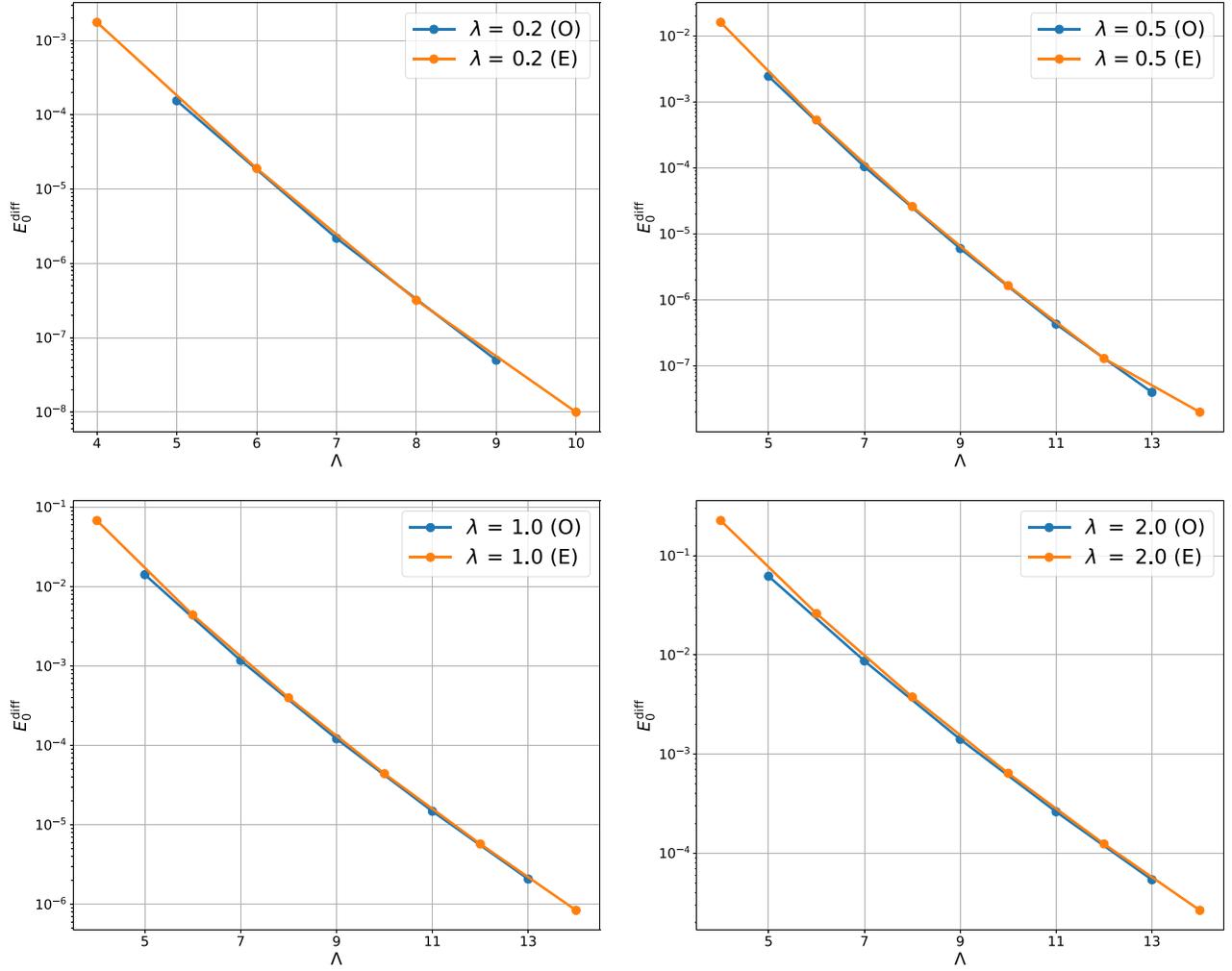


FIG. 3. The absolute value of the ground-state energy difference between successive values of the cutoff (see text for details) as a function of the cutoff  $\Lambda$  for various couplings  $\lambda = g^2 N = 0.2, 0.5, 1.0,$  and  $2.0$  for the SU(2) bosonic model. Even (E) and odd (O) values of  $\Lambda$  are plotted with different colors and the vertical axis is in logarithmic scale. The other parameters are  $m^2 = 1$  and  $c = 0$ .

As before, we consider  $\hat{H}' \equiv \hat{H} + c \sum_{\alpha} \hat{G}_{\alpha}^2$ , in order to eliminate the nonsinglet modes from the dynamics. Note that Eq. (36) has an extra fermionic contribution compared to the generator in the bosonic theory defined by Eq. (21) and this is reflected in a more complicated form of the deformed Hamiltonian  $\hat{H}'$  with fermionic variables in both terms.

We consider gauge-invariant and fixed-angular-momentum sectors. Hence, just for numerical purposes, we add  $c'(\hat{M} - J)^2$  to the Hamiltonian. In addition, we keep  $c \sum_{\alpha} \hat{G}_{\alpha}^2$  to penalize gauge nonsinglets. To summarize, we use a modified Hamiltonian  $\hat{H}'$  defined by

$$\hat{H}' = \hat{H} + c \sum_{\alpha} \hat{G}_{\alpha}^2 + c'(\hat{M} - J)^2. \quad (37)$$

Here,  $J$  is the angular momentum of the states that we want to consider. We take  $c = \Lambda$  as before and several choices

of  $c'$ , including  $c' = 10\Lambda$  for two values of  $J = 0, \frac{1}{2}$ . If we consider only the low-lying modes of  $\hat{H}'$ , states with  $G^2 > 0$  or  $M \neq J$  are removed because their energies are increased.

Let us consider the case of zero coupling as a sanity check. At zero coupling in the  $J = 0$  sector, the ground state is the Fock vacuum  $|0\rangle$  and the first excited mode is  $\text{Tr}(\hat{a}_1^{\dagger 2} + \hat{a}_2^{\dagger 2})|0\rangle$ . The energies are 0 and  $2\mu$  (recall that  $\mu$  and  $m$  are equivalent). The ground state is BPS, so the ground-state energy has to be zero at any coupling. On the other hand, at zero coupling in the  $J = 1/2$  sector, the lowest-energy mode is  $\text{Tr}[\hat{\xi}^{\dagger}(\hat{a}_1^{\dagger} + i\hat{a}_2^{\dagger})]|0\rangle$ , the energy of which is  $5\mu/2$ . The second lightest modes are made of  $\hat{\xi}^{\dagger}$ , two  $(\hat{a}_1^{\dagger} + i\hat{a}_2^{\dagger})$ , and one  $(\hat{a}_1^{\dagger} - i\hat{a}_2^{\dagger})$ , such as  $\text{Tr}[\hat{\xi}^{\dagger}(\hat{a}_1^{\dagger} + i\hat{a}_2^{\dagger})(\hat{a}_1^{\dagger 2} + \hat{a}_2^{\dagger 2})]|0\rangle$ , and the energy is  $9\mu/2$ . None of them is BPS; hence their energy can change when the interaction is turned on. We show the

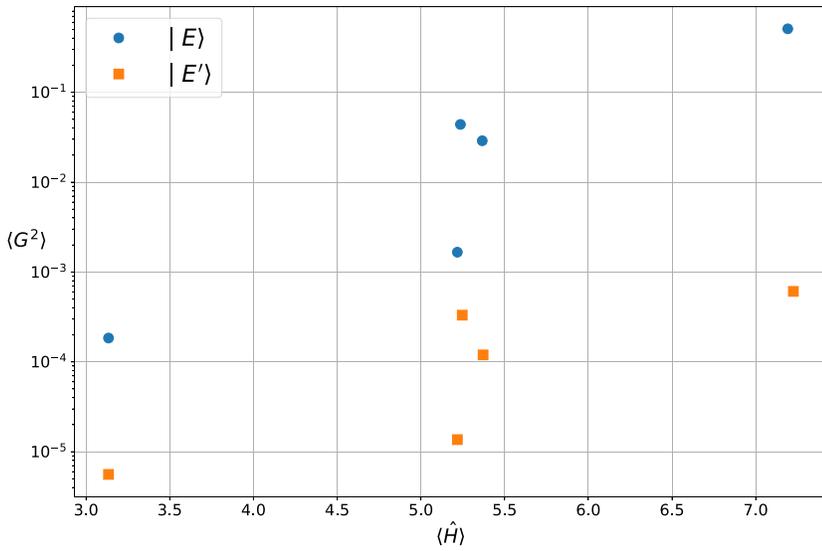


FIG. 4. We take the five eigenstates of  $\hat{H}$  ( $|E_j\rangle$ ) with the smallest values of  $\sum_{\alpha} \langle E_j | \hat{G}_{\alpha}^2 | E_j \rangle$  and the five lowest eigenstates for  $\hat{H}'$  ( $|E'_j\rangle$ ) with  $j = 0, \dots, 4$  and we plot them as a function of their energy ( $\langle \hat{H} \rangle$ ) for  $c = \Lambda = 4$ ,  $m^2 = 1$ ,  $\lambda = g^2 N = 0.2$  in the SU(2) bosonic model. We can see the deformation does not affect the low-lying modes; it just removes the nonsinglet modes.

zero-coupling case explicitly in Fig. 7, where the left panel corresponds to  $J = 0$  and the right panel to  $J = 1/2$  for  $\mu = 1$ .

Let the eigenstates of  $\hat{H}'$  in the sector of angular momentum  $J$  be  $|E'_{0,J}\rangle, |E'_{1,J}\rangle, \dots$ . The amount of breaking of gauge symmetry and SO(2) rotational symmetry

due to the truncation can be seen from  $\langle E'_{i,J} | \hat{G}^2 | E'_{i,J} \rangle$  and  $\langle E'_{i,J} | \hat{M} | E'_{i,J} \rangle - J$ . These quantities are plotted in Figs. 8 and 9. We can see the quick restoration of these symmetries as the cutoff increases, for all couplings.

$\langle E'_{0,J=0} | \hat{H} | E'_{0,J=0} \rangle$  is plotted in Fig. 10. Given that the breaking of gauge symmetry and rotational symmetry is

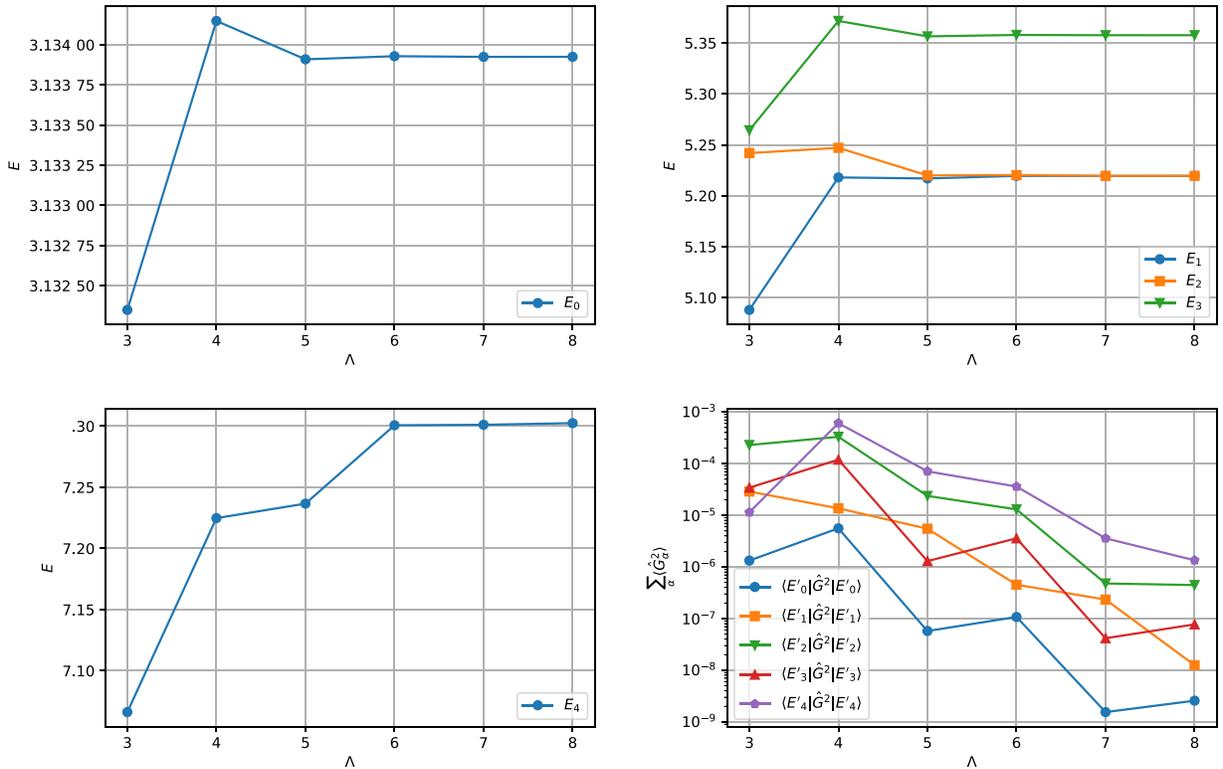


FIG. 5. The energy  $\langle E'_n | \hat{H} | E'_n \rangle$  and the violation of the singlet constraint  $\sum_{\alpha} \langle E'_n | \hat{G}_{\alpha}^2 | E'_n \rangle$  in the SU(2) bosonic model. The parameters are  $c = \Lambda$ ,  $m^2 = 1$ , and  $\lambda = g^2 N = 0.2$ .

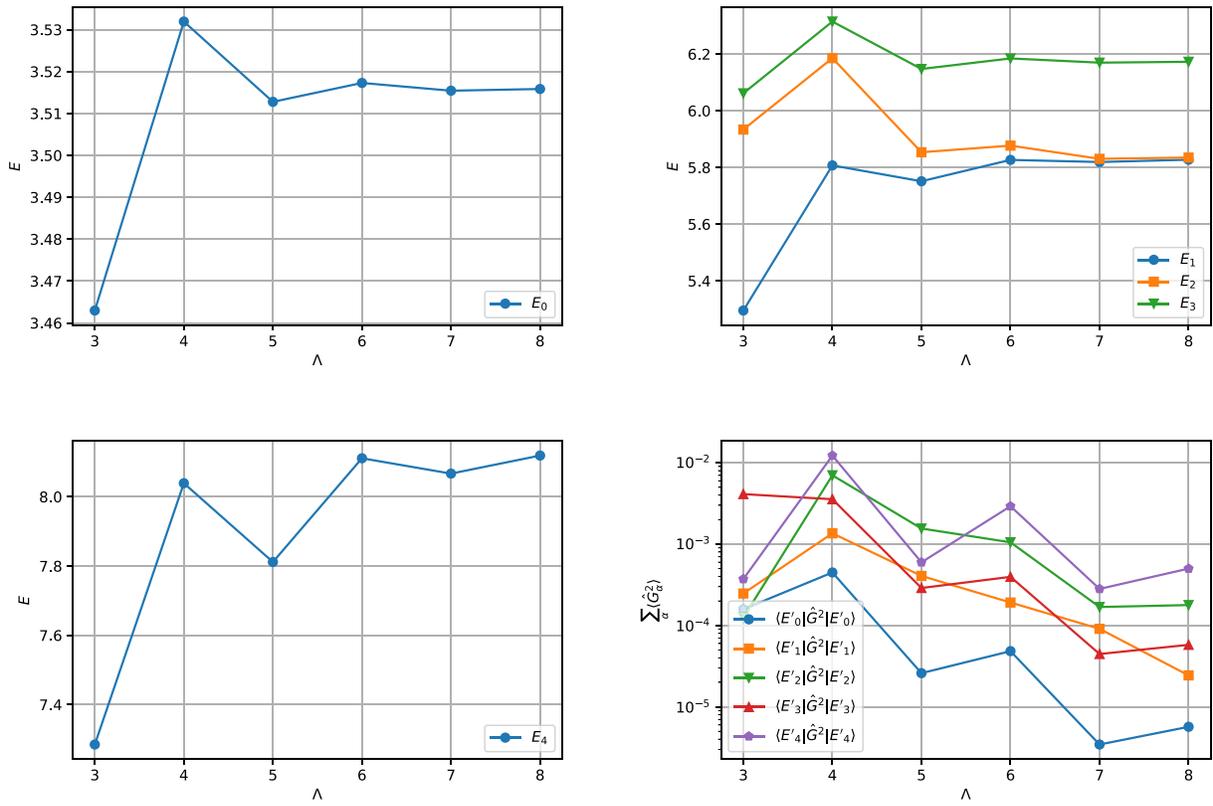


FIG. 6. The energy  $\langle E'_n | \hat{H} | E'_n \rangle$  and the violation of the singlet constraint  $\sum_{\alpha} \langle E'_n | \hat{G}_{\alpha}^2 | E'_n \rangle$ . The parameters are  $c = \Lambda$ ,  $m^2 = 1$ , and  $\lambda = g^2 N = 1.0$  in the SU(2) bosonic model.

small, this is very close to the ground-state energy  $E_{0,J=0}$ , which has to be zero at  $\Lambda = \infty$  due to the BPS condition. We can see an exponentially fast approach to zero as  $\Lambda$  increases. A few of the low-lying modes (up to the third

excited state) for the SU(2) model with  $\lambda = 0.2$  are plotted in Fig. 11 for the  $J = 0$  and the  $J = 1/2$  sectors. As the cutoff  $\Lambda$  increases, the ground states reach their asymptotic energy value, which is zero for the  $J = 0$  sector where the

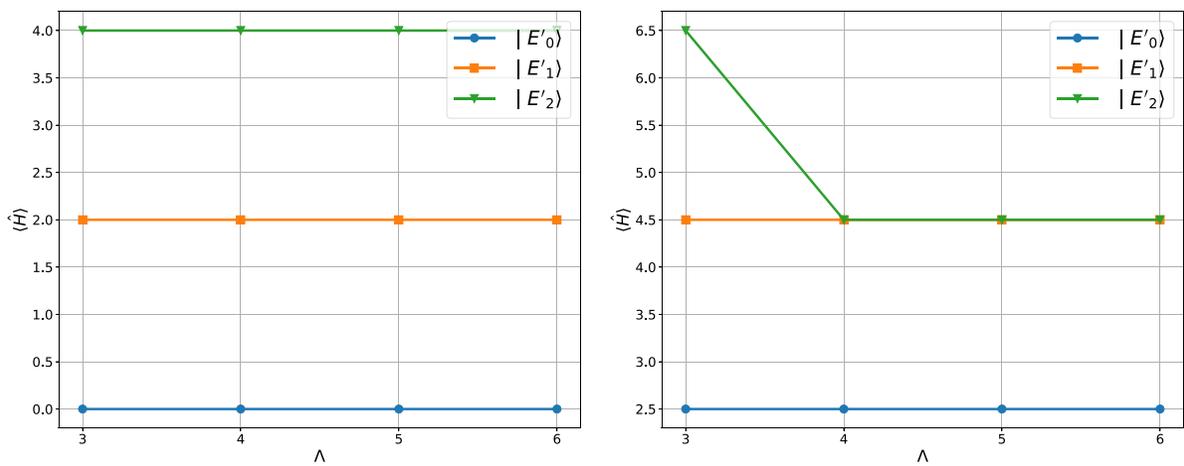


FIG. 7.  $\langle E'_{i,J} | \hat{H} | E'_{i,J} \rangle$  in the SU(2) minimal-BMN model at zero coupling, for a few low-lying excited modes. The modes are ordered by their  $\lambda_{E'}$  eigenvalues of  $\hat{H}'$  and their order can change depending on the cutoff  $\Lambda$ . The left panel is for the sector with  $J = 0$  and the right panel is for  $J = 1/2$ . The other parameters are  $\lambda = g^2 N = 0$ ,  $\mu = 1$ ,  $c = \Lambda$ , and  $c' = 100$ . (This large value of  $c'$  ensures that the low-cutoff states are in the right sector.)

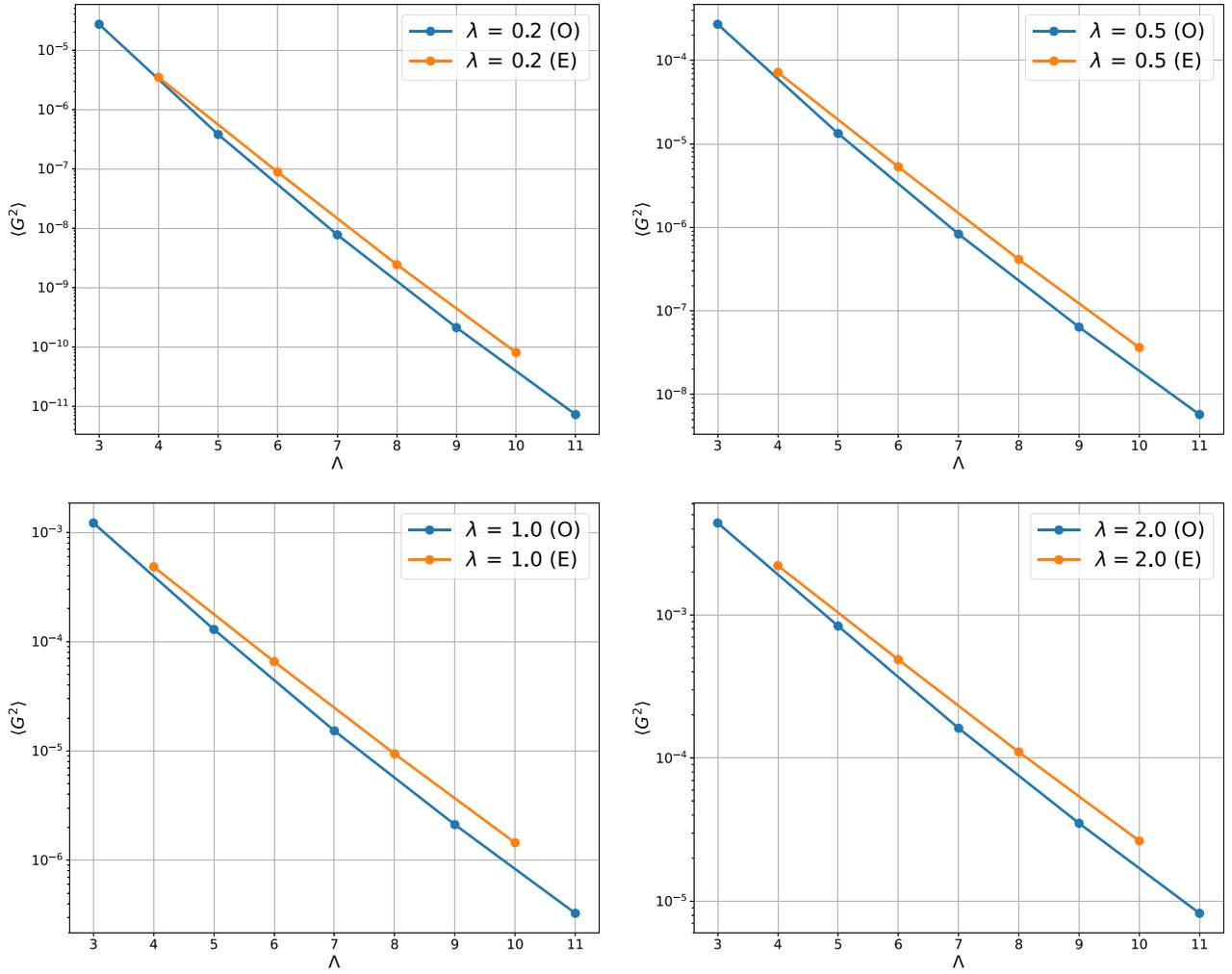


FIG. 8. The violation of the singlet constraint for the ground state of the Hamiltonian  $\hat{H}'$  as a function of the cutoff  $\Lambda$  for various couplings  $\lambda = g^2N = 0.2, 0.5, 1.0,$  and  $2.0$  in the SU(2) minimal-BMN model. Even (E) and odd (O) values of  $\Lambda$  are plotted with different colors in a logarithmic scale to show that the trend to zero is exponentially fast. The other parameters are  $\mu = 1, c = \Lambda, c' = 1,$  and  $J = 0$  (a low value of  $c'$  is enough for the ground state only).

ground state is BPS but is different from zero for the  $J = 1/2$  sector.

#### IV. QUANTUM-CLASSICAL HYBRID ALGORITHM (VQE)

The VQE algorithm is expected to be a practically useful tool in the NISQ era and has been shown to be extremely successful in finding the ground state of large physical systems such as molecules. The algorithm is a hybrid quantum-classical algorithm and this makes it more robust to noise on NISQ hardware. The expectation value of the Hamiltonian is efficiently computed on quantum hardware for quantum states represented by parametrized quantum circuits that mimic wave functions with variational parameters. The parameters of the quantum circuits or, in other words, of the wave function are optimized using classical

algorithms (e.g., steepest descent) on classical hardware such that each step of the optimization is moving toward convergence to the lower bound of the energy expectation value. The algorithm returns an upper bound for the ground-state energy  $E_{\text{var}}$ :

$$E_0 \leq E_{\text{var}} = \frac{\langle \psi(\theta_i) | H | \psi(\theta_i) \rangle}{\langle \psi(\theta_i) | \psi(\theta_i) \rangle}, \quad (38)$$

where  $E_0$  is the true ground-state energy and  $\theta_i$  are variational parameters of the trial wave function, which is represented in terms of parametrized quantum gates. More details about the VQE hybrid quantum-classical algorithm can be found in Ref. [21]. There is also a fully quantum algorithm for computing eigenvalues described in Ref. [56].

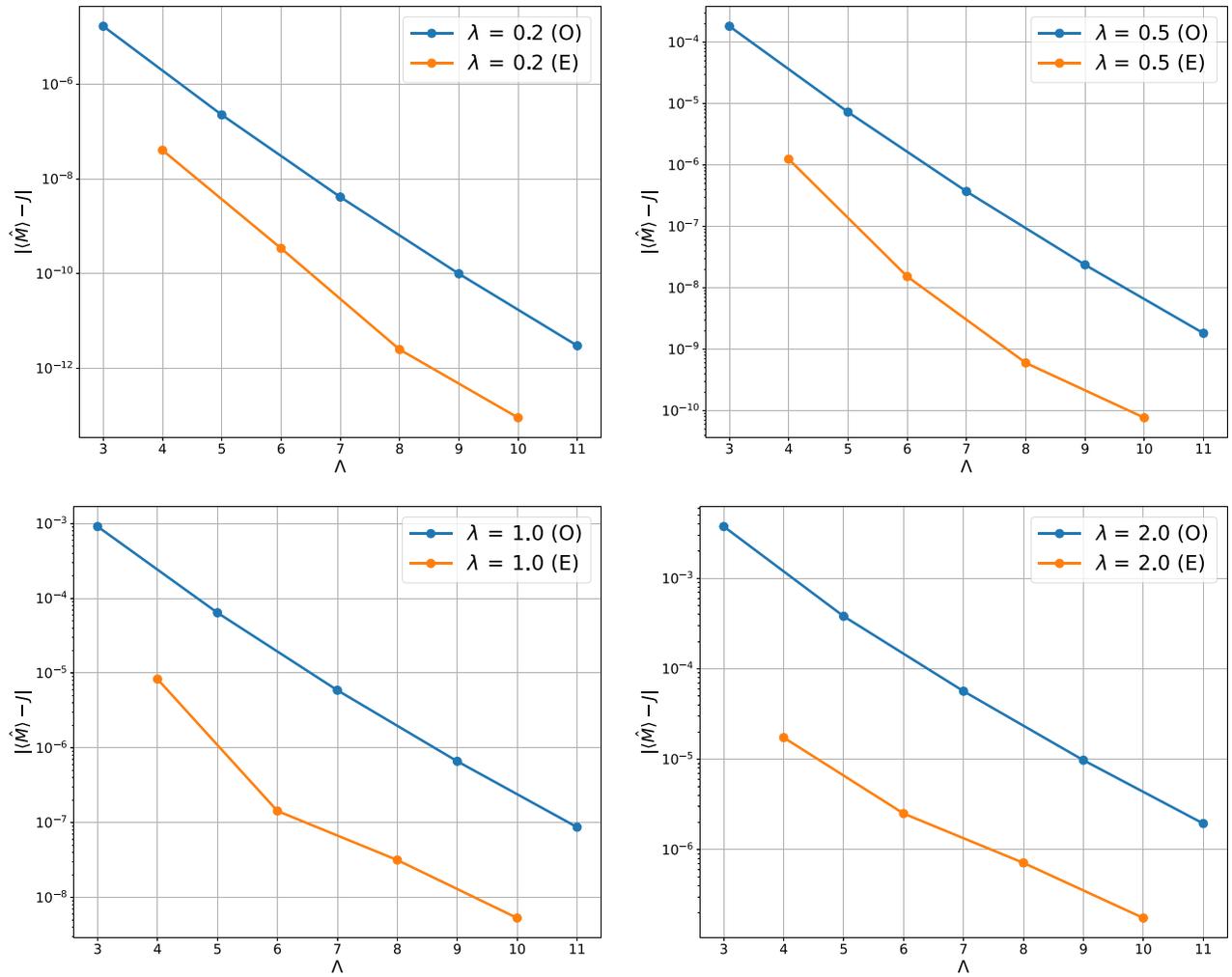


FIG. 9. The  $|\langle E'_{i,j} | \hat{M} | E'_{i,j} \rangle - J|$  for the ground state of the Hamiltonian  $\hat{H}'$  as a function of the cutoff  $\Lambda$  for various couplings  $\lambda = g^2 N = 0.2, 0.5, 1.0,$  and  $2.0$  in the  $SU(2)$  minimal-BMN model. Even (E) and odd (O) values of  $\Lambda$  are plotted with different colors in logarithmic scale. The other parameters are  $m^2 = 1, c = \Lambda, c' = 1,$  and  $J = 0$  (a low value of  $c'$  is enough for the ground state only).

On NISQ devices, it is of paramount importance to test how well the VQE performs for different systems [57]. So far, research advances in VQE have only been tested on simple systems, such as the hydrogen molecule, or with shallow parametrized circuits. However, VQE has the potential to become a successful tool for many quantum physical systems, including lattice models or lattice field theories. Matrix models are a natural test bed for this hybrid quantum-classical algorithm, because they are formally well defined and simpler than many useful quantum field theories in higher dimensions. In this section, we test the potential performance of the VQE approach in estimating the ground states of matrix models for the first time. First, we construct a truncated Hamiltonian that maps the Hamiltonian of the matrix model of interest into strings of Pauli matrices that can be implemented in terms of gates on quantum computers. We use the IBM QISKIT [58]

software framework and we use the VQE algorithm to place an upper bound on the ground-state energy of the matrix model. The QISKIT software allows us to write a single code for performing both a classical simulation of the VQE and a real quantum-classical run on IBM hardware. Moreover, all the necessary ingredients for building the algorithm are already implemented, thus lowering the barrier for future experimentations on matrix-model systems.

The performance of the VQE can depend on the choice of the ansatz for the wave function and the optimizer used for the minimization of the energy. We try several different combinations of classical optimizers and ansätze for the wave function. Even if the same ansatz is used, different upper bounds might be obtained depending on the optimizers because the convergence to the true minimum can be sabotaged by restricting the maximum number of iterations

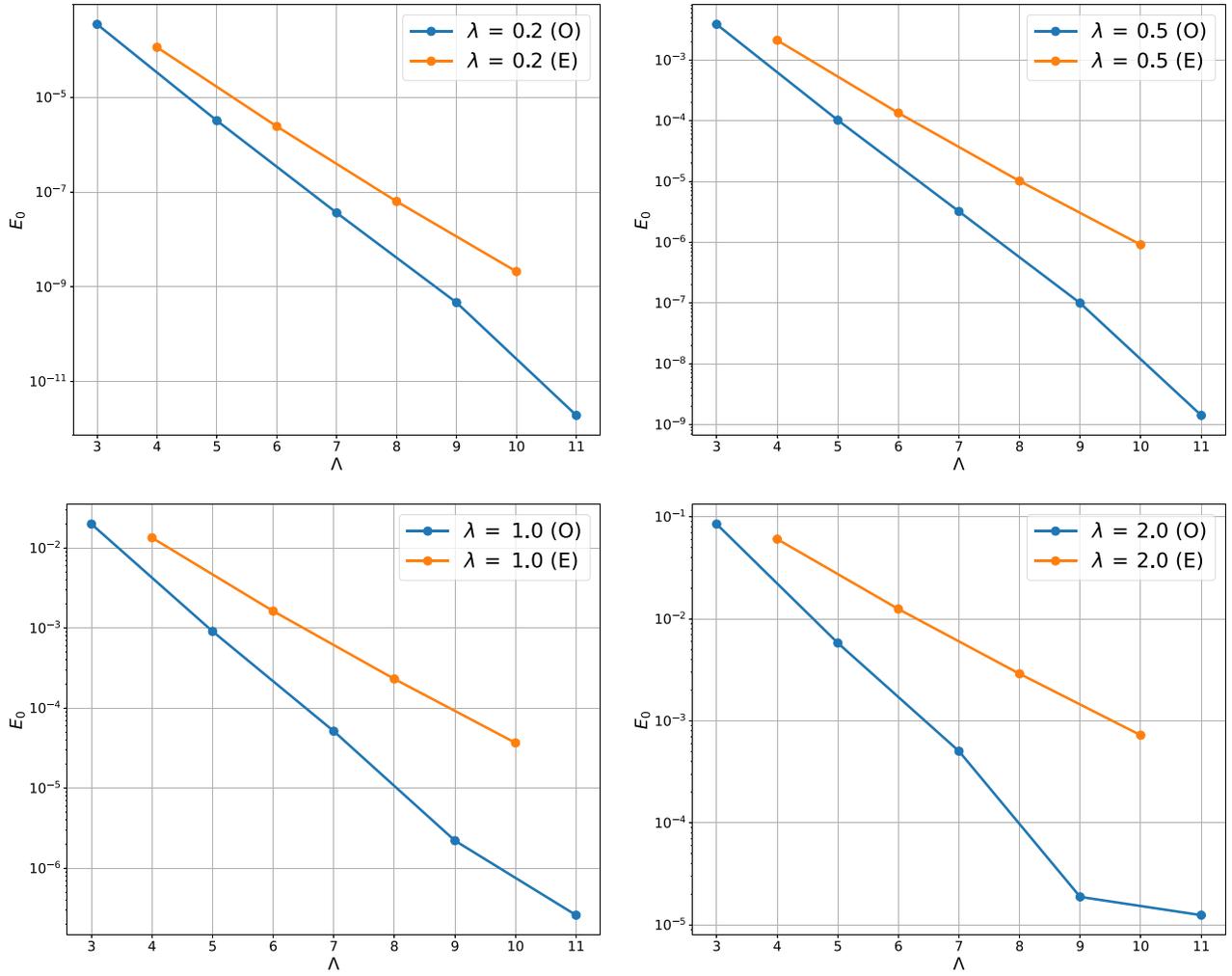


FIG. 10. The energy expectation value  $|\langle E'_{0,J=0} | \hat{H} | E'_{0,J=0} \rangle|$  of the ground state of the Hamiltonian  $\hat{H}$  as a function of the cutoff  $\Lambda$  for various couplings  $\lambda = g^2 N = 0.2, 0.5, 1.0,$  and  $2.0$  in the SU(2) minimal-BMN model. Even (E) and odd (O) values of  $\Lambda$  are plotted with different colors in a logarithmic scale. The approach to the  $\Lambda \rightarrow \infty$  limit value of zero is exponentially fast. The other parameters are  $m^2 = 1, c = \Lambda,$  and  $c' = 1$ .

or starting from an unfavorable point in parameter space, i.e., getting trapped in local minima. The ansatz we use in this section are not designed specifically for matrix models and, hence, there is no reason to expect good results. A first step toward identifying the best ansatz within a certain class of parametrized quantum circuits is shown at the end of this section but we leave a more comprehensive search to future work. In fact, we find larger deviation from the exact result at stronger couplings, which suggests a need to find a better ansatz for the wave function as we change the coupling of the matrix quantum mechanics model. This is expected, since the ground state at weak coupling is expected to exhibit different properties than at strong coupling.

A VQE solver obtains an upper bound for the ground-state energy by starting from a parametrized wave function in the form of quantum circuits made of parametrized

gates. Such circuits can be chosen arbitrarily and, typically, the final energy will depend on the choice of the initial form of the wave function. In Fig. 12, we show a single block of the variational circuit of six qubits that we use in the next section, where the parameters are encoded in gates that rotate each qubit about a certain axis. In the figure, we show specifically  $R_y$  gates, which are single-qubit rotations about the  $Y$  axis, so that

$$R_y(\theta) = \exp\left(-i\frac{\theta}{2}\right) = \begin{pmatrix} \cos\frac{\theta}{2} & -\sin\frac{\theta}{2} \\ \sin\frac{\theta}{2} & \cos\frac{\theta}{2} \end{pmatrix}. \quad (39)$$

We also consider a similar circuit obtained by replacing an  $R_y$  with a sequence of  $R_y$  and  $R_z$  gates.

These blocks can be composed into multiple “layers” to make deeper circuits, which allow the representation of

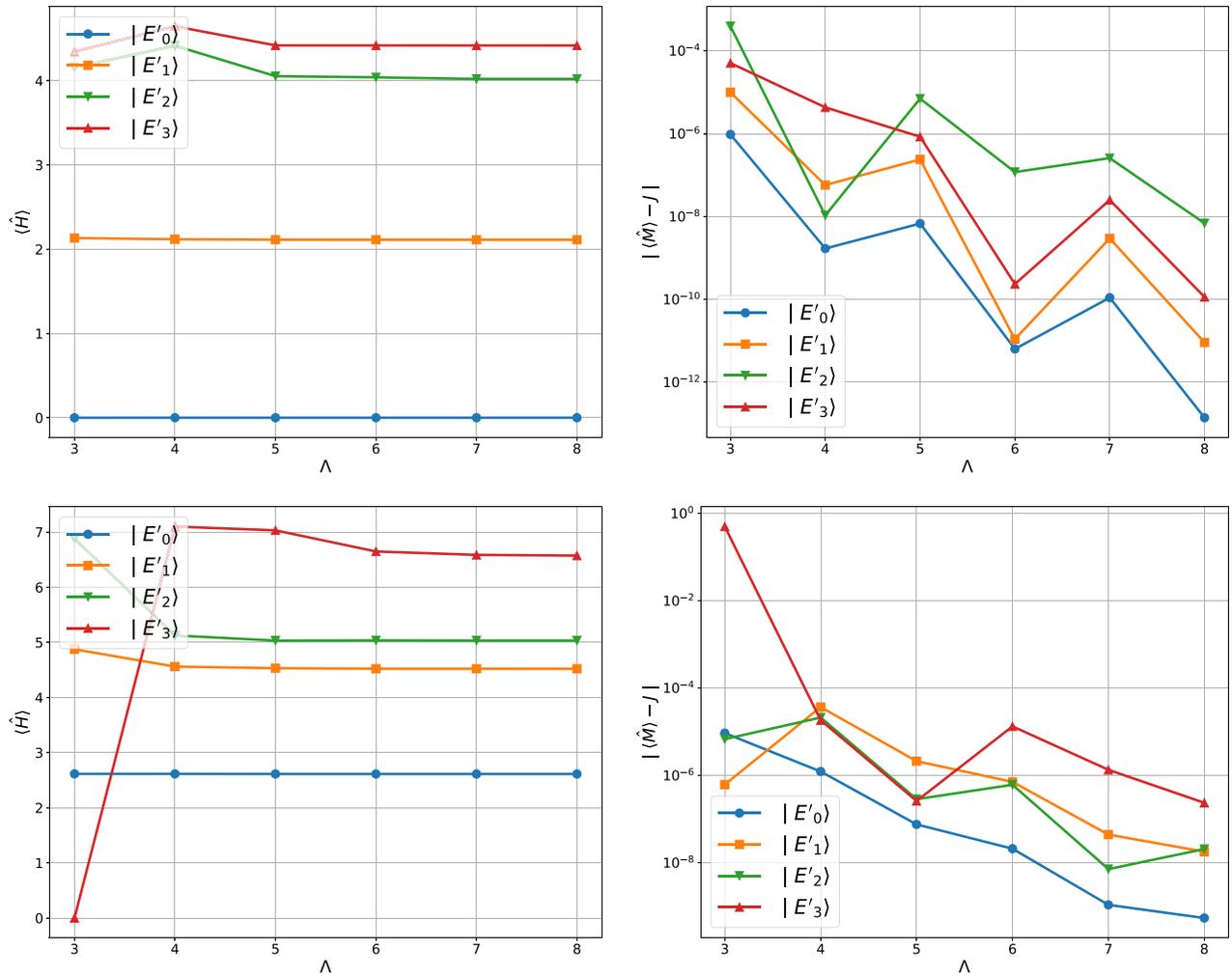


FIG. 11.  $\langle E'_{i,J} | \hat{H} | E'_{i,J} \rangle$  and  $|\langle E'_{i,J} | \hat{M} - J | E'_{i,J} \rangle|$  in SU(2) minimal BMN, for a few low-lying excited modes. The modes are ordered by their  $\lambda_{E'}$  eigenvalues of  $\hat{H}'$  and their order can change depending on the cutoff  $\Lambda$ , because the deformed parameters of the Hamiltonian in  $\hat{H}'$  are changing. The upper row is for the sector with  $J = 0$  and the lower row is for  $J = 1/2$ . The other parameters are  $\lambda = g^2 N = 0.2$ ,  $m^2 = 1$ ,  $c = \Lambda$ , and  $c' = 10\Lambda$ .

more expressive wave functions. By repeating the same blocks multiple times in a sequence, we also increase the number of parameters to optimize. For most optimizers in the following sections, we use three repetitions of the  $R_y$  fully entangled block, as shown in Fig. 13, and we refer to this variational ansatz as having a depth of 3 (just counting the number of layers). This amounts to  $8 \times 3 = 24$  parameters in total, which are the angles  $\theta[0], \theta[1], \dots, \theta[23]$  for the rotation gates. If the variational ansatz requires more qubits because, for example, we want to use a larger cutoff of the Hilbert space, then we would need a larger number of parameters. In general, having access to more parameters to optimize will allow for a more expressive wave function and we can expect a better overlap with the true ground-state wave function. However, the optimization landscape in a high-dimensional space (for a

large number of parameters) becomes more complex and some optimizers might fail in finding a true minimum or might take too long to converge. As a final remark, note that these parametrized circuits are heuristically designed without considering any prior knowledge about the physical system that we target. We choose them because they have been shown to be efficient at reproducing good quantum states for a large range of physical problems. They are our first step toward understanding the applicability of VQE to matrix models. In the spirit of keeping our variational approach agnostic with respect to the symmetries of the system, we do not attempt to include information about the discrete or continuous symmetries of the states into our quantum circuit. However, we note that this can be done [59] and that it may improve the variational results.

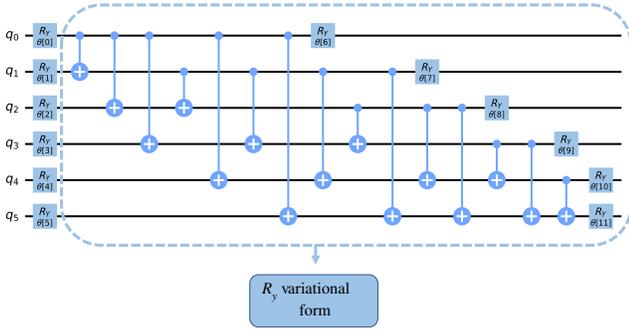


FIG. 12. A single block of the variational wave function in terms of parametrized quantum circuits. A parametrized  $R_y$  gate is applied to each qubit and each qubit is entangled with every other qubit using CNOT gates. We also consider a similar block ( $R_y, R_z$  variational form) in which the  $R_y$  gate is replaced with a sequence of  $R_y$  and  $R_z$  gates.

After choosing the variational form of the quantum circuit, we have to choose a classical optimizer that will be used to find the correct set of parameters to minimize the energy. We use four different optimizers: a sequential least-squares programming (SLSQP) optimizer, a constrained optimization by linear approximation (COBYLA) optimizer, a limited-memory Broyden-Fletcher-Goldfarb-Shanno- (BFGS) bound (L-BFGS-B) optimizer, and a Nelder-Mead optimizer. More information on these optimizers can be found in Ref. [60]. For a series of  $N_r$  runs starting from different initial parameter values  $\theta[i]$ , the least upper bound (minimum) gives the closest value to the ground state. For our optimizers, we limit the maximum number of iterations to  $10^4$ , which is sufficient to reach convergence in most of the cases that we study in the following sections, and we discuss the case of more iterations in a dedicated section featuring deeper parametrized circuits. To give a summary of the performance of each optimizer for a fixed variational form of the quantum circuit, we report the minimum value of the energy, the maximum value, the mean value, and the standard deviation across  $N_r$  runs. These optimizers can work

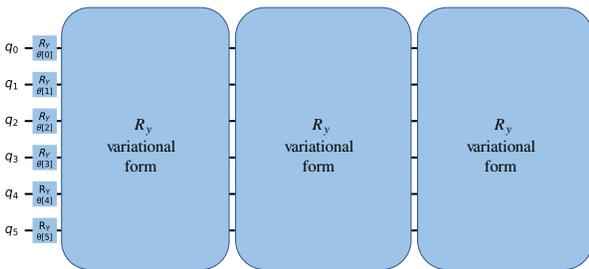


FIG. 13. The variational ansatz used in the VQE simulation. The circuit has a depth of 3 (i.e., the  $R_y$  variational form is repeated three times). We also consider the  $R_y, R_z$  variational form.

well in the absence of noise, such as in our state-vector simulation, but we would have to explore different algorithms once we take quantum simulations of matrix models to real quantum hardware [61].

### A. SU(2) bosonic matrix model

In this section, we consider the bosonic matrix model defined in Sec. II A. We study the case of two SU(2) matrices ( $N = 2, d = 2$ ) and different truncation levels:  $\Lambda = 2$  and  $\Lambda = 4$ . They require only six and 12 qubits, respectively and, hence, the emulation on a classical computer is straightforward. IBM QISKIT provides several quantum simulators as well as access to fully quantum hardware resources. The QASM simulator can be used to simulate shot noise on an actual quantum device and the state-vector simulator holds the value of the quantum state in computer memory through the computation. Moreover, it is straightforward to add different models of noise for each specific IBM quantum hardware resource on top of the classical simulation, thus opening up the possibility of understanding the limitations of VQE in the presence of noise for matrix models in the future. Using a NISQ device with the current knowledge we have on the best variational wave function for matrix models would not be an efficient use of resources. For this reason, in this paper we use the state-vector simulator on a classical computer, in the absence of any source of noise, therefore focusing on an ideal scenario as the first ever quantum simulation of matrix models. We then focus on identifying good variational ansatz. We would like to use the QASM simulator and hardware resources in upcoming works when we can get the most out of them. Moreover, we would invite researchers benchmarking different quantum algorithms for physical systems to take matrix models into account as a challenging test bed. As we will clearly see below, changing the number and type of the degrees of freedom (the gauge group size  $N$  and the bosonic and fermionic particles), changing the gauge coupling strength  $\lambda$ , and changing the truncation level  $\Lambda$  are all tunable knobs to increase or decrease the difficulty of the problem.

#### 1. One qubit for each boson

When the cutoff for the truncated Hilbert space is  $\Lambda = 2$ , we can represent each degree of freedom in the matrix model with one qubit. In the SU(2) two-matrix model, we have six degrees of freedom,  $X_{I=1,2}^{\alpha=1,2,3}$ . We denote  $X_1^1, X_1^2, X_1^3, X_2^1, X_2^2,$  and  $X_2^3$  in Eq. (18) as  $x_1, x_2, \dots, x_6$ . There are six annihilation operators  $\hat{a}_{i=1,\dots,6}$ , which can be represented in terms of tensor products as

$$\hat{a}_i = \hat{I}_1 \otimes \dots \otimes \hat{I}_{i-1} \otimes \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \otimes \hat{I}_{i+1} \otimes \dots \otimes \hat{I}_6, \quad (40)$$

TABLE I. VQE results for the bosonic BMN matrix model with SU(2) and  $\lambda = 0.2$ , where each boson is represented by one qubit. We use four optimizers and two variational forms (both with depth 3). Each optimizer begins from the same initial point. The minimum, maximum, average, and standard deviation (SD) of the best results across  $N_r = 100$  runs are reported. The exact result from the Hamiltonian truncation is  $E_0 = 3.14808$  and the best results are in bold type for each variational form. We limit the maximum number of iterations to  $10^4$ .

Optimizer	Variational form $R_y$				Variational form $R_y R_z$			
	Minimum	Maximum	Mean	SD	Minimum	Maximum	Mean	SD
COBYLA	3.149 370	4.147 156	3.159 740	0.099 739	<b>3.149 157</b>	3.150 034	3.149 862	0.000 202
L-BFGS-B	3.149 268	4.150 000	3.159 886	0.100 012	3.149 375	4.148 751	3.159 925	0.099 882
SLSQP	3.149 397	4.150 000	3.164 968	0.111 340	3.149 377	4.149 946	3.164 980	0.111 349
Nelder-Mead	<b>3.148 972</b>	3.195 922	3.150 774	0.005 065	3.149 516	4.149 891	3.171 468	0.140 469

where  $\hat{I}_i$  is the  $2 \times 2$  identity matrix for  $\Lambda = 2$ . The truncated versions of the Hamiltonian  $\hat{H}$  and the gauge generators  $\hat{G}_\alpha$  are defined in the same way as in Sec. III A, by using these  $\hat{a}_i$  qubit operators as the truncated annihilation operators. The truncated creation operators are defined as the Hermitian conjugate of the truncated annihilation operators and the truncated number operators are defined by  $\hat{n}_i = \hat{a}_i^\dagger \hat{a}_i$ . This procedure changes appropriately as we include more qubits in our basis for larger cutoff values  $\Lambda$ .

The regularized Hamiltonian in Eq. (18) for  $\Lambda = 2$  is a  $2^6 \times 2^6 = 64 \times 64$  matrix the smallest eigenvalue of which at coupling  $\lambda = 0.2$  is  $E_0 = 3.14808$ . This is the exact result from the HT approach. We compare the VQE result for the truncated Hamiltonian, which is only an upper bound, to this exact HT result. Note that we do not add a penalty term  $c \sum \hat{G}_\alpha^2$  to the Hamiltonian during the VQE process to lift the energy of the nonsinglet states. This is done in order to keep this first experiment simple and compare to the exact diagonalization results for  $c = 0$ . Moreover, this simple comparison allows us to focus on the performance of the classical optimizers and the

variational parametrized quantum circuit for this challenging new matrix-model system. In Table I, the statistics of the results for the four different VQE solvers are shown for  $N_r = 100$  runs and variational forms  $R_y$ ,  $R_y R_z$  of depth 3. The best results are  $E_{\text{var}} = 3.148 972$  and  $E_{\text{var}} = 3.149 157$  for  $R_y$  and  $R_y R_z$ , respectively. The  $R_y R_z$  variational form is an extension of the  $R_y$  form, i.e., one can set the  $R_z$  gate angles to zero, but the high dimensionality of the parameter space does not allow the optimizers to find an optimal point that is better unless the starting point is carefully chosen.

The upper bounds we find are very close to the exact value of  $E_0$ , even though these variational quantum circuits have not been specifically defined for the bosonic matrix models. The number of steps needed for the convergence of the solvers depends on the optimizer, as shown in Fig. 14. Among the solvers we use, COBYLA, L-BFGS-B, and SLSQP exhibit faster convergence than Nelder-Mead.

## 2. Two qubits for each boson

When we use a cutoff  $\Lambda = 4$ , we can represent each matrix degree of freedom with two qubits. The SU(2)

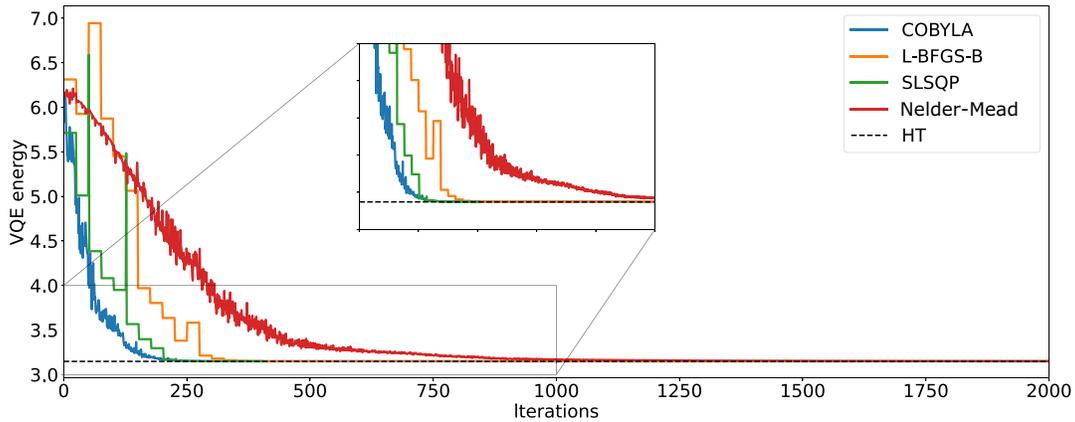


FIG. 14. The convergence of the VQE results for the bosonic matrix model with SU(2) and  $\lambda = 0.2$ , where each boson is represented by one qubit. Each curve represents a different classical optimizer (SLSQP, COBYLA, L-BFGS-B, and Nelder-Mead) and we only show the first 2000 iterations out of 10 000. The VQE result represents the least upper bound from ten runs of each optimizer with an  $R_y$  variational form of depth 3. The exact energy from the truncated Hamiltonian (HT) is represented by the dotted line at  $E_0 = 3.14808$ .

TABLE II. VQE results for the bosonic matrix model with SU(2) and  $\lambda = 0.2$ , where each boson is represented by two qubits. We use four optimizers and two variational forms (both with depth 3). Each optimizer began from the same initial point. The minimum, maximum, average, and standard deviation (SD) of the best results across  $N_r = 100$  runs are reported. The exact result from the Hamiltonian truncation is  $E_0 = 3.13406$  and the best results are in bold type for each variational form. We limited the maximum number of iterations to  $10^4$ .

Optimizer	Variational form $R_y$				Variational form $R_y R_z$			
	Minimum	Maximum	Mean	SD	Minimum	Maximum	Mean	SD
COBYLA	<b>3.137 059</b>	4.769 101	3.251 414	0.347 646	3.137 237	4.782 013	3.378 628	0.472 015
L-BFGS-B	<b>3.137 059</b>	5.769 553	3.283 462	0.434 162	<b>3.137 050</b>	4.286 367	3.243 110	0.307 549
SLSQP	3.137 060	5.769 554	3.327 706	0.471 957	3.137 059	4.232 419	3.236 925	0.290 855
Nelder-Mead	3.137 471	5.713 976	3.492 673	0.478 810	3.273 614	6.443 055	4.428 032	0.758 732

bosonic matrix model will then have each boson represented by a  $4 \times 4$  matrix (corresponding to two qubits). This is similar to Sec. 1 but with the annihilation operators represented by the tensor product

$$\hat{a}_i = \hat{I}_1 \otimes \dots \otimes \hat{I}_{i-1} \otimes \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & \sqrt{2} & 0 \\ 0 & 0 & 0 & \sqrt{3} \\ 0 & 0 & 0 & 0 \end{pmatrix} \otimes \hat{I}_{i+1} \otimes \dots \otimes \hat{I}_6, \quad (41)$$

where the identity matrix  $\hat{I}_i$  that we use is now  $4 \times 4$ . Again, the truncated versions of the Hamiltonian  $\hat{H}$  and gauge generators  $\hat{G}_\alpha$  are defined in the same way as in Sec. III A, by using these  $\hat{a}_i$  matrices as the truncated annihilation operators. The truncated Hamiltonian will then be a  $2^{12} \times 2^{12} = 4096 \times 4096$  matrix (on 12 qubits) and the exact ground-state energy is  $E_0 = 3.13406$ . In this case, we also focus on the results of the VQE without a gauge penalty term, i.e., for  $c = 0$ .

We test the performance of the same four types of VQE solvers as before and we use the 12-qubit equivalent of

the parametrized variational form of Fig. 13, which has 48 parameters,  $\theta[0], \theta[1], \dots, \theta[47]$ . In Fig. 15, we plot the convergence of the best run out of  $N_r = 100$  for each optimizer. After a certain number of optimization steps (which depends on the optimizer and the Hamiltonian), the estimate for the upper bound of the ground-state energy flattens out and this yields the best value for that run. In Table II, we report the statistics for the VQE results with  $N_r = 100$ . The best results are  $E_{\text{var}} = 3.137059$  and  $E_{\text{var}} = 3.137050$  for  $R_y$  and  $R_y R_z$ , respectively, and are again very close to the exact value of  $E_0 = 3.13406$ .

### 3. VQE for different coupling constants

To test the accuracy of the VQE results with respect to the coupling, we run the VQE while varying the coupling for the SU(2) Hamiltonian for both cases where the bosons are represented by one qubit or by two qubits. A graph of the VQE results with respect to coupling is shown in Fig. 16, with the values tabulated in Table III, for both cases where bosons are represented by one qubit or two qubits. In both cases, we use the same variational forms as before and we compare them using all of the solvers.

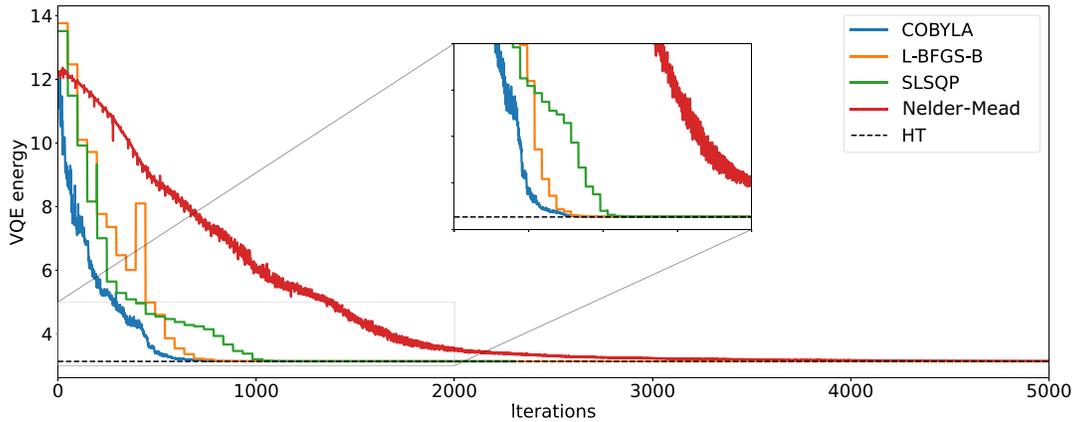


FIG. 15. The convergence of the VQE results for the SU(2) bosonic matrix model with  $\lambda = 0.2$ , where each boson is represented by two qubits. Each curve represents a different classical optimizer (SLSQP, COBYLA, L-BFGS-B, and Nelder-Mead), and we only show the first 5000 iterations out of 10 000. The VQE result represents the least upper bound from ten runs of each optimizer with an  $R_y$  variational form of depth 3. The exact energy from the truncated Hamiltonian (HT) is represented by the dotted line at  $E_0 = 3.13406$ .

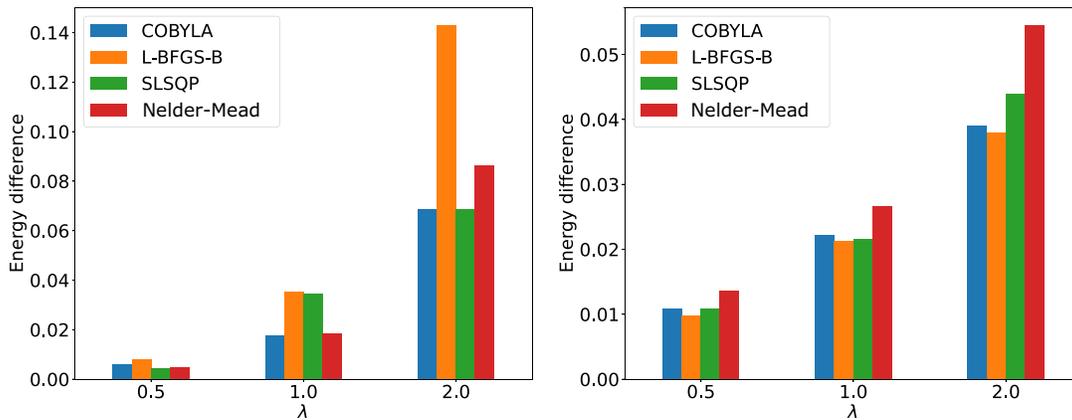


FIG. 16. The energy difference between the VQE upper bound for the SU(2) bosonic matrix model using the three best optimizers for coupling  $\lambda = 0.5, 1.0,$  and  $2.0$ . The left plot is for bosons represented by one qubit and the right plot is for bosons represented by two qubits. The  $R_y$  ansatz with depth 3 is used and the best of  $N_r = 100$  runs is displayed. We limited the maximum number of iterations to  $10^4$ .

We conclude our VQE studies of the bosonic matrix model with a discussion of the dependence of the upper bound determined by the VQE for different couplings. Clearly, the gap between the upper bound determined by the VQE and the exact value grows with the coupling for both the one-qubit bosons and the two-qubit bosons. This is consistent with what has been found in studying the anharmonic oscillator and the supersymmetric anharmonic oscillator in Ref. [62]. The most likely reason for the increase in the gap is that the ansatz of the variational wave function does not have as large an overlap with the ground-state wave function at strong coupling as it does at weak coupling. Some possibilities to improve this include using specially designed ansatz tailored to the Hamiltonian, as is done with coupled-cluster ansatz for chemistry or with gauge-invariant forms for lattice gauge theories [59], adaptive methods to obtain better ansatz described in Ref. [63], and machine-learning methods to improve overlap with the ground state in Ref. [30].

## B. SU(2) minimal BMN

### 1. One qubit for each boson

For the minimal-BMN matrix model, we use the same representation as above, with each boson represented by one qubit (a  $2 \times 2$  matrix) and each fermion also represented by one qubit. In total, we have  $9^\circ$  of freedom to be described by nine qubits. The fermionic space operators have to be tensored with the bosonic space, which is  $2^6 = 64$  dimensional. Therefore, we define three fermionic annihilation operators as in Eq. (35) and they are tensored with the identity matrix in the bosonic space:

$$c_1 = \hat{I}_{64} \otimes \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad (42)$$

$$c_2 = \hat{I}_{64} \otimes \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \otimes \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad (43)$$

TABLE III. VQE results for the bosonic BMN matrix model with the SU(2) group at different couplings, with each boson represented by one qubit and two qubits. The  $R_y$  ansatz with depth 3 is used and the best of  $N_r = 100$  runs is reported. The last row is the exact result obtained from the HT. The best results are reported in bold type. We limit the maximum number of iterations to  $10^4$ .

Optimizer	One qubit			Two qubits		
	$\lambda = 0.5$	$\lambda = 1.0$	$\lambda = 2.0$	$\lambda = 0.5$	$\lambda = 1.0$	$\lambda = 2.0$
COBYLA	3.368 59	<b>3.714 63</b>	4.336 38	3.309 75	3.548 39	3.934 52
L-BFGS-B	3.370 43	3.732 59	4.411 03	<b>3.308 69</b>	<b>3.547 48</b>	<b>3.933 48</b>
SLSQP	<b>3.366 75</b>	3.731 79	<b>4.336 36</b>	3.309 74	3.547 76	3.939 46
Nelder-Mead	3.367 18	3.715 47	4.354 11	3.312 55	3.552 92	3.950 03
HT (exact)	3.362 54	3.697 22	4.267 95	3.298 94	3.526 25	3.895 48

TABLE IV. VQE results for the minimal-BMN matrix model with SU(2) and  $\lambda = 0.2$ , where each boson is represented by one qubit. The VQE optimizers are run  $N_r = 100$  times with two variational forms,  $R_y$  and  $R_y R_z$  (both with depth 3). We limit the maximum number of iterations to  $10^4$ . Each optimizer begins from the same initial point. The minimum, maximum, average, and standard deviation (SD) of the best results across all runs are reported. The exact value is  $E_0 = 0.003\,287$ . The best results are in bold type.

Optimizer	Variational form $R_y$				Variational form $R_y R_z$			
	Minimum	Maximum	Mean	SD	Minimum	Maximum	Mean	SD
COBYLA	<b>0.100 79</b>	0.150 02	0.130 20	0.011 16	<b>0.077 83</b>	1.34775	0.202 78	0.278 19
L-BFGS-B	0.124 52	1.357 61	0.187 85	0.194 27	0.100 68	1.383 51	0.182 51	0.187 85
SLSQP	0.124 56	1.650 00	0.235 70	0.287 36	0.100 68	1.367 07	0.191 68	0.210 56
Nelder-Mead	0.124 96	1.358 08	0.194 79	0.200 55	0.132 18	1.888 56	0.320 12	0.319 14

$$c_3 = \hat{I}_{64} \otimes \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \otimes \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad (44)$$

where  $\hat{I}_{64}$  is the identity matrix for the bosonic space. These qubit operators representing the  $\hat{\xi}_\alpha$  creation operators for the fermionic DOF are used in the construction of the truncated Hamiltonian and the gauge transformation operators  $\hat{G}_\alpha$ . The bosonic annihilation operators are the same as those defined in Eq. (40), except that they are tensored with the identity matrix  $\hat{I}_8$  for the  $2^3 = 8$  dimensional fermionic space.

The exact ground-state energy for the minimal-BMN matrix model is zero. The truncated Hamiltonian for this model is represented by a  $2^9 \times 2^9 = 512 \times 512$  matrix the exact ground-state energy of which for  $\lambda = 0.2$  is  $E_0 = 0.003\,287$ . We repeat the same strategy as used in the bosonic BMN model and report the performance statistics for various optimizers and variational forms in Table IV.

For the minimal-BMN model, we find that the least upper bound is further from the exact value than in the bosonic BMN if we use the same type of variational wave

functions. In this case, the parametrization of the variational wave function does not have as strong an overlap with the true ground-state wave function as is found in the bosonic matrix-model case. It would be interesting to investigate the form of the variational ground states to further examine the differences between the two models and in the following we show results for deeper parametrized quantum circuits. We see this as a low-hanging fruit for future improvements: similarly to advances made by the VQE in quantum chemistry using parametrized circuits designed with domain knowledge, we expect a great deal of improvement in matrix models once these variational forms are carefully designed, possibly with help from machine-learning methods. A plot of the convergence for the different optimizers is shown in Fig. 17, where only the best out of  $N_r = 100$  is plotted for each classical optimizer. In some cases, the optimizers seem to have converged to a minimum energy but during some iterations they briefly move away (larger energy) from the minimum before moving closer to it again. This behavior affects the total number of iterations needed for convergence and it depends on the path taken by the optimizer in parameter

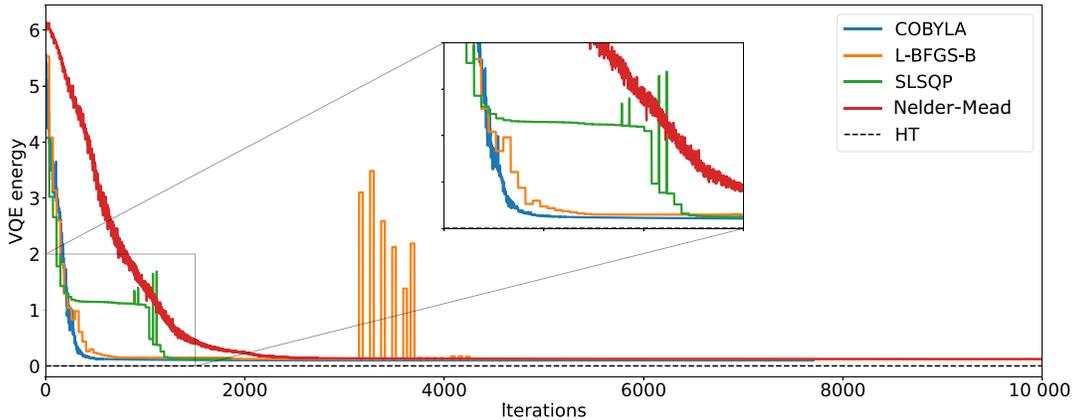


FIG. 17. The convergence of the VQE results for the minimal-BMN matrix model with SU(2) and  $\lambda = 0.2$ , where each boson is represented by one qubit. Each curve represents a different classical optimizer that is used with the VQE algorithm (SLSQP, COBYLA, L-BFGS-B, and Nelder-Mead). The VQE result represents the least upper bound from 100 runs of each optimizer. The exact energy of the truncated Hamiltonian (HT) is represented by the dotted line at  $E_0 = 0.003\,287$ . We limit the maximum number of iterations to  $10^4$ .

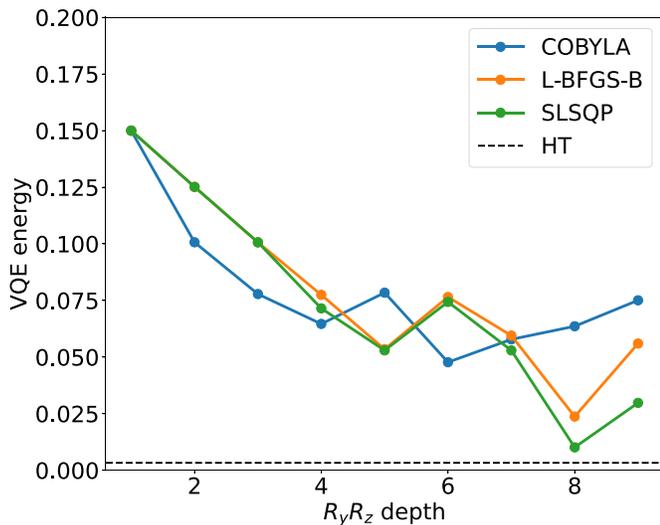


FIG. 18. VQE results with respect to depth for the minimal-BMN matrix model with  $SU(2)$  and  $\lambda = 0.2$ . The VQE is run using various optimizers, for an  $R_y R_z$  variational form of depth ranging from 1 to 9. The least upper bound across 100 different runs (with  $10^4$  iterations) is reported. The exact ground-state energy of the truncated Hamiltonian is represented by the dotted line.

space to navigate the energy-function landscape and find the optimal point.

## 2. VQE computations at different depths

Increasing the depth of a quantum circuit allows for access to a larger number of attainable states over which the expectation value of the Hamiltonian can be calculated. To understand how the results change with respect to depth, we run the VQE algorithm for the minimal-BMN model over various depths. We vary the depth from 1 to 9, with an increment of 1.

TABLE V. VQE results for the minimal-BMN model with  $SU(2)$  and  $\lambda = 0.2$ . The VQE is run using various optimizers, for an  $R_y R_z$  variational form of depth ranging from 1 to 9. The least upper bound across 100 different runs (with 10 000 iterations) is reported. The exact value is  $E_0 = 0.003287$ .

Depth	COBYLA	L-BFGS-B	SLSQP
1	0.150 000	0.150 000	0.150 000
2	0.100 689	0.125 154	0.125 158
3	0.077 833	0.100 683	0.100 683
4	0.064 637	0.077 476	0.071 500
5	0.078 448	0.053 471	0.052 959
6	0.047 723	0.076 528	0.074 366
7	0.057 819	0.059 495	0.052 949
8	0.063 603	0.023 726	0.010 126
9	0.075 062	0.055 973	0.029 769

The results are shown in Fig. 18 and are tabulated in Table V. For each depth, the initial point for each optimizer is the same, which ensures a fair comparison of the ability of each optimizer to converge. Additionally, we use an  $R_y R_z$  variational form to further increase the number of attainable states by starting already with a larger number of parameters. We use the least upper bound out of  $N_r = 100$  different runs. As expected, a more expressive variational form allows us to obtain results that are increasingly closer to the exact diagonalization result for this truncated Hamiltonian. For a depth of 8, the SLSQP optimizer gives the best result, a least upper bound of  $E_{\text{var}} = 0.010126$  to be compared with the exact result of  $E_0 = 0.003287$ . We also note that the algorithms take more iterations to reach convergence when more parameters are included in the variational quantum circuits representing the wave-function ansatz. This is shown in Fig. 19 for an  $R_y R_z$  variational form of depth 9. These results imply that increasing the number of maximum iterations above  $10^4$  for some of the solvers might allow them to find better optimal points. For example, if we let the SLSQP optimizer run until convergence, with no limit on the number of iterations, at depth 9, it will get to a least upper bound of  $E_{\text{var}} = 0.004755$  but it would need five times more iterations (approximately  $5 \times 10^4$ ). In order to reduce the number of iterations, one might consider metalearning approaches to find good initialization heuristics for the parameters [64,65].

## 3. VQE for different coupling constants

As is done with the  $SU(2)$  bosonic model, we test the accuracy of the VQE results with respect to coupling for the minimal-BMN matrix model by running the VQE for different coupling constants. We use an  $R_y R_z$  ansatz with depth 5 to expand the number of attainable states (by including more variational parameters), since we see this to be advantageous in the previous section at  $\lambda = 0.2$ . The VQE results at different coupling constants are tabulated in Table VI. The accompanying optimizer convergence plots are shown in Fig. 20. It is evident from the convergence plots that the Nelder-Mead optimizers needs more iterations to converge.

## V. DEEP LEARNING

### A. Overview of variational quantum Monte Carlo

Next, we discuss the deep-learning approach that has been applied to the minimal-BMN matrix model in Ref. [25]. Specifically, neural networks are used as the wave-function ansatz in the variational quantum Monte Carlo method for the ground state. The algorithm can be roughly decomposed into four steps:

- (i) Parametrize a quantum state  $|\psi_\theta\rangle$  by a set of parameters  $\theta$ . Specifically, we construct a function that is

TABLE VI. VQE results for the minimal-BMN matrix model, where each boson and fermion are represented by one qubit, at different couplings. The  $R_y R_z$  ansatz with depth 5 is used and the least upper bound across  $N_r = 100$  runs is reported (with  $10^4$  iterations). In the “Best” column, we report the least upper bound that we obtain from a depth 9  $R_y R_z$  ansatz using the best optimizer for ten times more iterations. The “HT” column is the exact energy of the truncated Hamiltonian.

$\lambda$	COBYLA	L-BFGS-B	SLSQP	Nelder-Mead	Best	HT (exact)
0.5	0.088 492	0.139 702	0.134 517	0.406 003	0.027 44	0.016 90
1.0	0.135 800	0.219 268	0.308 781	0.752 459	0.079 00	0.048 29
2.0	0.387 977	0.622 704	0.522 396	1.271 939	0.176 88	0.083 85

interpreted as the wave function  $\psi_\theta(X) = \langle X | \psi_\theta \rangle$  by using a neural network. The explicit form of  $\psi_\theta(X)$  is given later in Sec. V B.

- (ii) Estimate the energy from Monte Carlo samples of the wave function. The energy of the state  $|\psi_\theta\rangle$  is defined by

$$E_\theta \equiv \langle \psi_\theta | \hat{H} | \psi_\theta \rangle = \int dX |\psi_\theta(X)|^2 \cdot \frac{\langle X | \hat{H} | \psi_\theta \rangle}{\psi_\theta(X)} = \mathbf{E}_{X \sim |\psi_\theta|^2} [\epsilon_\theta(X)]. \tag{45}$$

In the last expression, the samples of matrices are drawn from the probability distribution  $|\psi_\theta(X)|^2$ ,  $\epsilon_\theta(X)$  is defined as  $\langle X | \hat{H} | \psi_\theta \rangle / \psi_\theta(X)$ , and  $E_\theta$  is estimated as the mean of  $\epsilon_\theta(X)$  from these samples. The generative nature of our wave-function ansatz, to be discussed in detail later, allows for both efficient sampling from  $|\psi_\theta|^2$  and evaluation of  $\epsilon_\theta$ .

- (iii) Compute the gradient of the energy with respect to model parameters  $\theta$ . Note that the sampling distribution depends on  $\theta$  as well and hence

$$\nabla_\theta E_\theta = \mathbf{E}_{X \sim |\psi_\theta|^2} [\nabla_\theta \epsilon_\theta(X)] + \mathbf{E}_{X \sim |\psi_\theta|^2} \times [\epsilon_\theta(X) \nabla_\theta \ln |\psi_\theta|^2]. \tag{46}$$

- (iv) Update the model parameters via the gradient descent as

$$\theta' = \theta - \beta \nabla_\theta E_\theta, \tag{47}$$

where the step size  $\beta$  is the “learning rate.” The energy in Eq. (45) is then minimized within the family of wave functions parametrized by  $\theta$ . The (local) minimum found is then a variational upper bound for the ground-state energy.

**B. Form of the wave-function ansatz**

First, we focus on the wave-function form in the  $SU(N)$  bosonic matrix quantum mechanics. The quantum wave function is a complex function  $\psi(X) = |\psi(X)| e^{i\theta(X)}$  of Hermitian matrices  $X$ . The wave-function norm  $|\psi(X)|$  and the phase  $\theta(X)$  are modeled separately.

The wave-function norm is parametrized as an autoregressive flow model. Flow models are extremely attractive in the context of density estimation, i.e., sampling from a complicated probability distribution over a large number of variables, because of their flexibility and tractability as generative models. We choose an autoregressive type of generative model because it allows for extremely fast sampling of the variables according to their probability

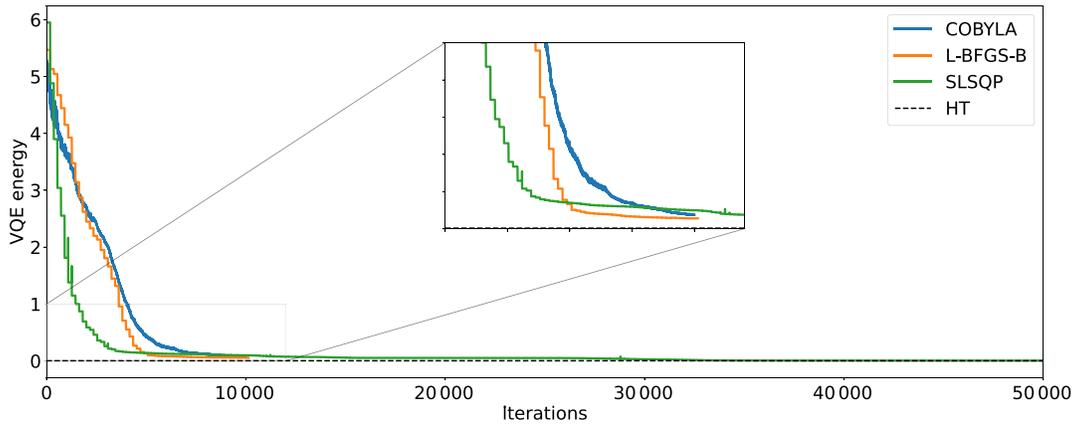


FIG. 19. The convergence of the VQE results for the minimal-BMN matrix model with  $SU(2)$  and  $\lambda = 0.2$ , where each boson is represented by one qubit. The VQE results are the least upper bound from 100 runs of each optimizer and an  $R_y R_z$  variational form of depth 9 is used. The SLSQP optimizer is allowed to run for five times more iterations than the other two, which are stopped at  $10^4$  iterations. The exact energy of the truncated Hamiltonian (HT) is represented by the dotted line at  $E_0 = 0.003 287$ .

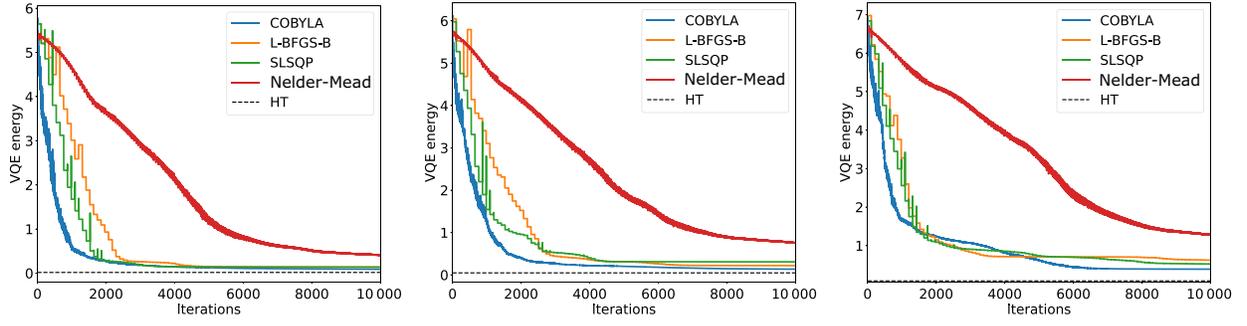


FIG. 20. Convergence plots of the results of the VQE computation for the minimal-BMN matrix model with SU(2) for coupling  $\lambda = 0.5$  (left),  $\lambda = 1.0$  (middle), and  $\lambda = 2.0$  (right). The  $R_y R_z$  ansatz with depth 5 is used and the best of  $N_r = 100$  runs is displayed up to  $10^4$  iterations.

distribution. In the following, we review the major components of an autoregressive flow. Let  $|\psi(X)\rangle = \sqrt{p_\theta(X)}$ , where  $p_\theta(X)$ , the wave-function probability distribution, takes the following form:

$$p_\theta(X) = p(x_1; F_\theta^0) p(x_2; F_\theta^1(x_1)) p(x_3; F_\theta^2(x_1, x_2)) \dots, \quad (48)$$

in which  $p(x; F)$  is a function of  $x$  specified by parameters  $F$ . Here,  $x_i$  are the components of the Hermitian matrices expanded onto some basis. For the SU( $N$ ) bosonic matrix model,  $i$  runs from 1 to  $N^2 - 1$  times the number of matrices. The  $x_i$  are ordered in some arbitrary way in this ansatz, due to the autoregressive nature of the flow, and neither the choice of the ordering nor that of the basis of the  $\mathfrak{su}(N)$  algebra used for expansion significantly affects the final results. As an example, in Eq. (48), the  $p$  on the right side could be normal distributions, which are parametrized by their locations and scales. The location and the scale parameters of the coordinate  $x_i$  depend on  $x_j$  with  $j < i$ . The dependence is parametrized by function  $F_\theta^i$ , which is discussed later. This autoregressive ordering of coordinates allows for easy sampling from  $p_\theta(X)$ , as we can first sample  $x_1$  independent of any other variables, then sample  $x_2$  dependent on  $x_1$ ,  $x_3$  dependent on  $x_1$  and  $x_2$ , and so on. Recall that sampling is a crucial step in the variational quantum Monte Carlo method, both for estimating the cost function and for its gradients, therefore dictating how well the learning process can be tackled. Moreover, it takes up a considerable amount of time within the full algorithm. This form of the neural ansatz is therefore the more appropriate one for our task.

The final component in Eq. (48) is the form of  $F_\theta^i$ . If the distributions  $p$  in Eq. (48) are normal with two parameters,  $F_\theta^i$  is then a vector function from  $i$  to two real numbers. For each  $i$ ,  $F_\theta^i$  is given by a fully connected neural network ( $\circ$  denotes a concatenation of functions):

$$F_\theta^i = A_\theta^{i,m} \circ \tanh \circ A_\theta^{i,m-1} \circ \tanh \circ \dots \circ A_\theta^{i,2} \circ \tanh \circ A_\theta^{i,1}, \quad (49)$$

where for  $a = 1, 2, \dots, m$ ,

$$A_\theta^{i,a}(\vec{x}) = M_\theta^{i,a} \vec{x} + \vec{b}_\theta^{i,a} \quad (50)$$

is an affine transformation. The weight matrices  $M_\theta^{i,a}$  and the bias vectors  $\vec{b}_\theta^{i,a}$  are trainable parameters, not shared among  $i$ . The input and output dimensions are determined ( $i$  and 2 in the Gaussian example) and the dimensions of the intermediate layers are hyperparameters.

Of course, if the distribution  $p$  in Eq. (48) is only Gaussian, the ansatz cannot describe non-Gaussian conditional distributions, although it is still much more flexible than a free ansatz. For more flexibility, in this work we use the neural autoregressive flow (NAF) [66] (and its B-NAF variant [67]), which is an extension of the simple Gaussian example discussed. We use  $m = 3$  for the fully connected network in NAF and denote the ratio of the intermediate (hidden) dimension to the input dimension as  $\alpha$ .

The phase of the wave function  $\theta(X)$  is parametrized by another fully connected neural network as in Eq. (49). The input dimension of the network is  $N^2 - 1$  times the number of matrices. The output dimension is one, because the phase is a scalar number. For ground states of the bosonic models, the phase is not strictly necessary.

For the minimal-BMN model, the quantum wave function  $\psi(X, \xi)$  depends on fermionic matrices as well. We expand fermionic matrices onto some SU( $N$ ) basis and choose the fermion-number basis as the basis for quantum states. Then the coordinates  $(X, \xi)$  for the wave function contain  $2(N^2 - 1)$  real numbers and  $(N^2 - 1)$  binaries. The wave function is parametrized as  $\sqrt{p(X)} [f_\xi(X) + ig_\xi(X)]$ , where  $p(X)$  is given by the NAF as previously discussed and  $f_\xi$  and  $g_\xi$  are real functions given by fully connected neural networks as in Eq. (49). The weights and biases in  $f_\xi$  and  $g_\xi$  depend on the binary vector  $\xi$ , via additional fully connected networks as well. We test other types of neural networks, such as normalizing flows, finding them less accurate for our purposes. In all of them, we do not use any domain knowledge, such as the presence of symmetries between degrees of freedom. In other

words, we seek to start testing this algorithm in the context of matrix models without any prior bias about the type of solution we expect for the ground state. This is similar to what we do on quantum computers with the VQE, where our variational ansatz is not tailored to matrix models. However, gauge symmetries are important in these models, even when not included in the variational ansatz. For this reason, we discuss them in detail in the next section.

### C. Gauge constraints

Physically, we may hope to keep only  $SU(N)$  gauge-invariant states. In Ref. [25], this is achieved by gauge fixing, which scales to large  $SU(N)$  symmetry groups. Here, we propose two more methods of imposing gauge invariance (cf. Sec. III B) and we discuss future relevant directions that are not considered in our present study.

The first alternative is to add a gauge Casimir penalty to the Hamiltonian:

$$\hat{H}' = \hat{H} + c \sum_{\alpha} \hat{G}_{\alpha}^2. \quad (51)$$

Here,  $c > 0$  is a hyperparameter. We introduce this modified Hamiltonian in Sec. III B, where we use exact diagonalization to find the eigenstates. The penalty term favors gauge-singlet states and imposes strict gauge invariance as  $c \rightarrow \infty$ : the lowest-energy states will be the gauge-invariant ones. However, in practice, as our ansatz is not strictly gauge invariant, very large values of  $c$  will interfere with minimization of the energy. For this approach, we define the gauge Casimir operator, using its differential form as in Ref. [25], because this is easily computed with automatic differentiation for the neural wave function with both bosons and fermions that we adopt in Sec. V B.

The second approach is to project the wave function onto the singlet sector. Let  $\psi(X)$  be a variational wave function, not necessarily gauge invariant. The projection onto the singlet sector can be written as an average over  $SU(N)$  gauge transformations:

$$\hat{P}|\psi\rangle = \int dU \hat{U}|\psi\rangle, \quad (52)$$

where the integration is done with the Haar measure. For any gauge-invariant observable  $\hat{O}$ , its expectation value in the projected (unnormalized) state is

$$\langle \hat{O} \rangle_{\text{singlet}} = \frac{\langle \psi | \hat{P} \hat{O} | \psi \rangle}{\langle \psi | \hat{P} | \psi \rangle}. \quad (53)$$

TABLE VII. The neural variational quantum Monte Carlo ground-state energy and gauge Casimir for the bosonic  $SU(2)$  matrix quantum mechanics, at coupling  $\lambda = g^2 N = 0.2, 0.5, 1.0,$  and  $2.0$  and gauge penalty  $c = 0$  and  $10$  [cf. Eq. (33)]. Here,  $E_{0,\text{var}} = \langle \hat{H} \rangle$  in the variational state for  $N = 2$ ,  $\hat{H}$  in Eq. (3), and  $E_{0,\text{singlet}} = \langle \hat{H} \rangle_{\text{singlet}}$  as in Eq. (53). The ratio of the hidden dimension and the input dimension  $\alpha = 20$ .

$c = 0$				
$\lambda$	0.2	0.5	1.0	2.0
$E_{0,\text{var}}$	3.137(2)	3.299(2)	3.518(2)	3.856(3)
$G_{\text{var}}^2$	0.0028(4)	0.0059(6)	0.0062(7)	0.0122(9)
$E_{0,\text{singlet}}$	3.135(2)	3.297(2)	3.520(2)	3.859(3)
$c = 10$				
$\lambda$	0.2	0.5	1.0	2.0
$E_{0,\text{var}}$	3.137(2)	3.309(2)	3.545(3)	3.912(3)
$G_{\text{var}}^2$	0.000 11(8)	0.000 19(7)	0.000 22(11)	0.000 21(8)
$E_{0,\text{singlet}}$	3.139(2)	3.307(2)	3.544(3)	3.908(3)

Note that  $\hat{U}|X\rangle = |UXU^\dagger\rangle$ , so the numerator is

$$\begin{aligned} \langle \psi | \hat{P} \hat{O} | \psi \rangle &= \int dX \langle \psi | \hat{P} | X \rangle \langle X | \hat{O} | \psi \rangle \\ &= \int dU dX \psi^*(UXU^\dagger) \langle X | \hat{O} | \psi \rangle \\ &= \mathbf{E}_{U,X \sim |\psi|^2} \left[ \frac{\langle X | \hat{O} | \psi \rangle}{\psi(X)} \frac{\psi^*(UXU^\dagger)}{\psi^*(X)} \right]. \quad (54) \end{aligned}$$

The denominator is simply  $\hat{O} = 1$ . Hence the singlet expectation values can be estimated with Monte Carlo samples of  $U$  and  $X$  as well.

However, for ground states of bosonic models, it is not necessary to impose gauge invariance, as the ground state is always a singlet. This is numerically verified in Table VII, where the gauge Casimir in the variational ground states is less than  $10^{-3}$  even at  $c = 0$ . With  $c = 10$ , the expectation value of  $G^2$  is less than  $10^{-4}$  and the variational ground-state energy is not significantly increased.

As we have explained before, we make the choice of avoiding the inclusion of domain knowledge into the structure of the variational ansatz in this initial benchmark. For this reason, the two methods above do not require a change in the ansatz, which is allowed to represent noninvariant states. Recent studies of quantum lattice field theory models and condensed-matter models with gauge symmetries have shed light on the importance of realizing these symmetries exactly in the architecture of the variational neural quantum states (NQSs). This is an interesting future direction to explore. Variational NQSs with gauge-invariant or equivariant architectures have been shown to increase the performance to convergence to the physical ground-state wave function [68,69] and to help in extracting excited states as well. Moreover, generative models with gauge

TABLE VIII. The neural variational quantum Monte Carlo ground-state energy for the bosonic SU(2) matrix quantum mechanics, at coupling  $\lambda = g^2N = 0.2, 0.5, 1.0,$  and  $2.0$ . The neural autoregressive flow ansatz contains a fully connected network with an input and an output layer of  $2(N^2 - 1) = 6$  units and a hidden layer of  $\alpha \times 2(N^2 - 1)$  units. Hence the  $\alpha$  here is the ratio of the hidden dimension and the input dimension. A larger  $\alpha$  means more parameters and a more flexible ansatz. The exact result of the HT at high truncation level  $\Lambda = 11$  is reported in the last column.

$\alpha$	1	2	5	10	20	50	HT (exact)
$\lambda = 0.2$	3.137(2)	3.137(2)	3.140(2)	3.138(2)	3.137(2)	3.135(2)	3.134
$\lambda = 0.5$	3.313(2)	3.312(2)	3.308(2)	3.307(2)	3.302(2)	3.305(2)	3.297
$\lambda = 1.0$	3.544(3)	3.544(2)	3.541(3)	3.528(2)	3.519(2)	3.520(2)	3.516
$\lambda = 2.0$	3.914(3)	3.910(3)	3.892(3)	3.872(3)	3.857(3)	3.859(3)	3.854

equivariant neural networks have been used for sampling configurations of non-Abelian gauge theories on the lattice [13,70], with and without fermions. This is an extremely challenging endeavor and it is an area of active research [71]. In the future, matrix models could equally benefit from these techniques, in particular when the field is mature regarding the exact realization of supersymmetries in neural quantum states.

## D. Ground-state energy and observables

### 1. Bosonic matrix model

We start with the ground-state energy in the SU(2) bosonic two-matrix quantum mechanics. The variational energies will be compared to Monte Carlo results and they agree within numerical uncertainties. The accuracy of the variational energies can also be assessed by increasing the number of hidden units in the neural-network ansatz. With more parameters, we expect that the ansatz will be more flexible and accurate, and indeed convergence to exact (Hamiltonian-truncation) results is observed in Table VIII.

No significant improvement is observed from  $\alpha = 20$  to  $\alpha = 50$  in Table VIII. Similarly, for the SU(3) bosonic matrix model we summarize the results in Table IX

### 2. SU(2) minimal BMN

In the supersymmetric model, the ground-state energy should be zero, and indeed we see vanishing variational energies. In Table X for  $N = 2$  and  $\lambda = \mu = 1$ , the variational energy converges to zero as we increase  $\alpha$ . Other observables in the variational state are accurate as well.

TABLE IX. The neural variational quantum Monte Carlo ground-state energy for the bosonic SU(3) matrix quantum mechanics, at coupling  $\lambda = g^2N = 0.5, 1.0,$  and  $2.0$ . The neural autoregressive flow ansatz contains a fully connected network with an input and an output layer of  $2(N^2 - 1) = 16$  units and a hidden layer of  $\alpha \times 2(N^2 - 1)$  units. Hence the  $\alpha$  here is the ratio of the hidden dimension and the input dimension. A larger  $\alpha$  means more parameters and a more flexible ansatz. The results of the lattice Monte Carlo (MC) simulations are reported in the last column.

$\alpha$	1	2	5	10	20	50	MC
$\lambda = 0.5$	8.833(10)	8.828(7)	8.829(7)	8.835(7)	8.824(7)	8.826(7)	8.84(4)
$\lambda = 1.0$	9.459(7)	9.447(7)	9.468(8)	9.466(10)	9.432(7)	9.440(8)	9.38(4)
$\lambda = 2.0$	10.447(12)	10.448(9)	10.441(9)	10.457(12)	10.426(8)	10.397(13)	10.24(4)

The results at  $\alpha = 20$  for other couplings are summarized in Table XI. Note that a natural energy scale of the theory is of order one. For example, the ground-state energies of bosonic and fermionic oscillators are  $\pm(N^2 - 1) = \pm 3$ , and they cancel out. The results at  $\alpha = 20$  mean that the cancellation is reproduced up to a percent-order error.

## VI. MONTE CARLO SIMULATION ON A EUCLIDEAN LATTICE

In this section, we employ the lattice Monte Carlo method to study the bosonic matrix model and compare its results to the methods described in the previous sections. For the minimal-BMN model, since several exact relations are known about the ground state due to supersymmetry, we do not repeat the study using lattice Monte Carlo simulations. For an introductory review, see, e.g., Ref. [72].

We use the following Euclidean action on the circle:

$$S = N \int_0^\beta dt \text{Tr} \left( \frac{1}{2} (D_t X_I)^2 + \frac{m^2}{2} X_I^2 - \frac{\lambda}{4} [X_I, X_J]^2 \right), \quad (55)$$

where  $\lambda = g^2N$ . The circumference of the temporal circle,  $\beta$ , is the inverse temperature  $\beta = 1/T$  of the system. The matrices  $X_I$  are traceless Hermitian.

The lattice regularization that we use is the tree-level improved action, which is essentially the same as the one used in Ref. [73]. The discretized lattice action can be

TABLE X. The neural variational quantum Monte Carlo ground-state observables for the SU(2) minimal-BMN model, at coupling  $\lambda = g^2 N = 1.0$  and  $\mu = 1$ . The observables are the Hamiltonian  $H$ , the gauge Casimir  $G^2$ , the SO(2) angular momentum  $M$ , and the fermion number  $F$ . A larger  $\alpha$  means more parameters and a more flexible ansatz. The last column contains the exact results from the HT at  $\Lambda = 8$ .

$\alpha$	1	2	5	10	20	50	HT (exact)
$H$	0.058(6)	0.053(6)	0.041(6)	0.031(6)	0.014(6)	0.005(6)	0.000
$G^2$	0.007(8)	-0.008(8)	0.014(8)	0.007(9)	0.022(9)	0.012(9)	0.000
$M$	-0.0003(3)	-0.0004(3)	-0.0001(4)	0.0001(4)	-0.0003(5)	-0.0001(4)	0.0000
$F$	0.1844(6)	0.1833(6)	0.1895(6)	0.1922(6)	0.1946(7)	0.1935(7)	0.2034

written as

$$S_{\text{lattice}} = N a \sum_{t=1}^{n_t} \text{Tr} \left( \frac{1}{2} (D_t X)_{I,t}^2 + \frac{m^2}{2} X_{I,t}^2 - \frac{\lambda}{4} [X_{I,t}, X_{J,t}]^2 \right), \quad (56)$$

where  $n_t$  is the number of lattice points used for discretizing the circle and  $a = \beta/n_t$  is the lattice spacing. The lattice spacing has to be removed because it introduces systematic effects and we are interested in the so-called ‘‘continuum-limit’’ system, when  $a \rightarrow 0$ .

The covariant derivative  $D_t$  is defined by

$$(D_t X)_{I,t} = \frac{1}{a} \left( -\frac{1}{2} U_t U_{t+a} X_{I,t+2a} U_{t+a}^\dagger U_t^\dagger + 2 U_t X_{I,t+a} U_t^\dagger - \frac{3}{2} X_{I,t} \right). \quad (57)$$

Here, we adopt periodic boundary conditions for the matrices,  $X_{I,n_t+1} = X_{I,1}$ . If we set the unitary link variable  $U_t$  to be the identity, the ungauged version of the model can be studied as well [73]. Note that  $\beta = a n_t$ .

In lattice Monte Carlo, a sequence of lattice configurations  $\{X^{(1)}, U^{(1)}\} \rightarrow \{X^{(2)}, U^{(2)}\} \rightarrow \dots \rightarrow \{X^{(k)}, U^{(k)}\} \rightarrow \dots$  is generated, such that their probability distribution converges to  $e^{-S_{\text{lattice}}(X,U)}$ . Then the average over lattice configurations converges to the expectation value in the

TABLE XI. The neural variational quantum Monte Carlo ground-state observables for the SU(2) minimal BMN, at coupling  $\lambda = g^2 N = 0.5, 1.0$ , and  $2.0$ , using  $\mu = 1$ .

	$\lambda = 0.5$	$\lambda = 1.0$	$\lambda = 2.0$
$H$	0.009(5)	0.014(6)	0.034(7)
$G^2$	0.010(6)	0.022(9)	0.038(14)
$M$	-0.0002(3)	-0.0003(5)	0.0006(7)
$F$	0.1224(4)	0.1946(7)	0.2729(9)

Euclidean path integral:

$$\begin{aligned} \langle f \rangle &\equiv \frac{\int dX dU f(X, U) e^{-S_{\text{lattice}}(X,U)}}{\int dX dU e^{-S_{\text{lattice}}(X,U)}} \\ &= \lim_{K \rightarrow \infty} \frac{1}{K} \sum_{k=1}^K f(X^{(k)}, U^{(k)}). \end{aligned} \quad (58)$$

To generate configurations with the correct probability distribution, we use the hybrid Monte Carlo (HMC) algorithm [74], which allows us to sample configurations along molecular dynamics trajectories in a fictitious Monte Carlo time. The configurations sampled along these HMC trajectories are autocorrelated because of the dynamics used to select the next step in a trajectory and successive configurations must be discarded to reduce such autocorrelations.

To determine the energy of the system in the ground state using this path-integral formulation, it is convenient to use the virial theorem, which relates the kinetic ( $K$ ) and potential terms ( $V$ ) in the Hamiltonian:

$$\langle K \rangle = \left\langle \frac{1}{2} \sum_i x_i \frac{\partial V}{\partial x_i} \right\rangle. \quad (59)$$

Hence the total energy  $E$  at temperature  $T = \beta^{-1}$  is evaluated as

$$\begin{aligned} E &= \left\langle \frac{1}{\beta} \int_0^\beta dt \left( V + \frac{1}{2} \sum_i x_i \frac{\partial V}{\partial x_i} \right) \right\rangle \\ &= \left\langle \frac{N}{\beta} \int_0^\beta dt \left( m^2 X_I^2 - \frac{3\lambda}{4} [X_I, X_J]^2 \right) \right\rangle. \end{aligned} \quad (60)$$

On the discretized lattice circle, this becomes

$$E_{\text{lattice}} = \left\langle \frac{N}{n_t} \sum_{t=1}^{n_t} \left( m^2 X_{I,t}^2 - \frac{3\lambda}{4} [X_{I,t}, X_{J,t}]^2 \right) \right\rangle. \quad (61)$$

We compute the energy for the bosonic two-matrix model with gauge groups SU(2) and SU(3) at three values of the coupling  $\lambda = 0.5, 1.0$ , and  $2.0$ , with fixed  $m^2 = 1$ . The simulations are performed at temperatures  $0.025 \leq$

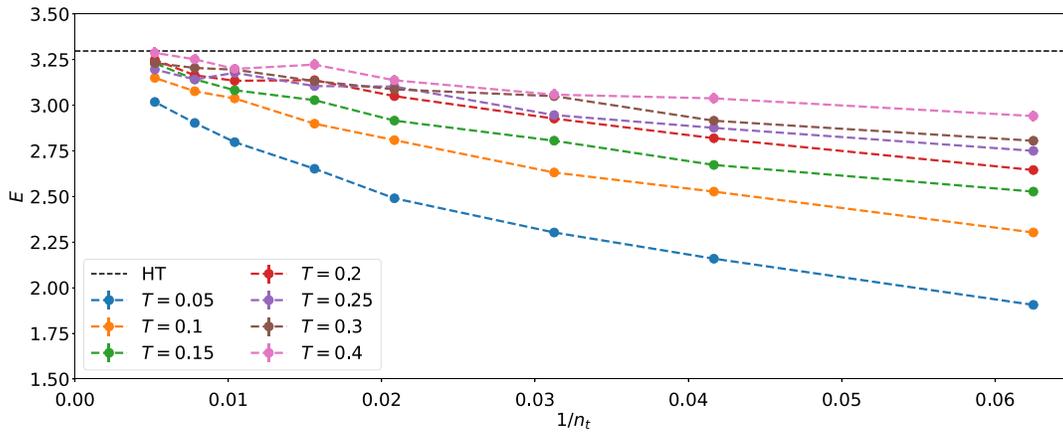


FIG. 21. The energy of the SU(2) bosonic model at  $m^2 = 1$  and  $\lambda = g^2 N = 0.5$  for various temperatures  $T$  and lattice sizes  $n_t$ . The black dashed line is the exact result from the HT approach at cutoff  $\Lambda = 14$ ,  $E_0 = 3.297$ . Higher temperatures approach it from below.

$T \leq 0.4$  and lattice size  $n_t$  ranging from 16 to 192. For each parameter set, we simulate about one million HMC trajectories of length between 0.5 and 1 (a length of 1 corresponds to one molecular dynamics time unit, MDTU). We discard 1000 MDTU at the beginning of each trajectory to remove thermalization (burn-in) effects and have configurations that are from the correct distribution obtained when the HMC algorithm converges. We also save the configuration every 10–50 MDTU to reduce autocorrelation times. *A posteriori*, we check that the integrated autocorrelation time of the energy is always around unity.

In Fig. 21, we show the energy of the SU(2) theory with  $m^2 = 1$  and  $\lambda = g^2 N = 0.5$  at several temperature values between  $T = 0.4$  and  $T = 0.05$ . The horizontal axis is the inverse lattice size,  $1/n_t$ . In Fig. 22, we change the horizontal axis to be proportional to the lattice spacing  $a = 1/(Tn_t)$ . In this representation, the data do not

show a significant  $T$  dependence and, at fixed lattice spacing  $a$ , different temperatures mostly agree within statistical uncertainties. The black horizontal line is the value coming from the HT of Sec. III A computed at a cutoff of  $\Lambda = 14$ . We can see that the lattice data toward the continuum limit  $a \rightarrow 0$  are consistent with this value but the comparison requires a careful extrapolation of the lattice data. The continuum limit is taken using energy data in the range  $a = 1/(Tn_t) \in [0, 0.25]$  and extrapolated with simple polynomial functions of the lattice spacing.

We fit the data using Bayesian least-squares fits of Gaussian-distributed variables implemented in LSQFIT [75] and using uninformative priors for the parameters. A way to assess the systematic effects of the continuum-limit extrapolation is to try cutting the data used at different maximum lattice spacings—that is, using  $a \leq a_{\max}$ —and also to increase the order  $n_p$  of the fitting functions (polynomials) to include higher powers of  $a$ . We use the fitting

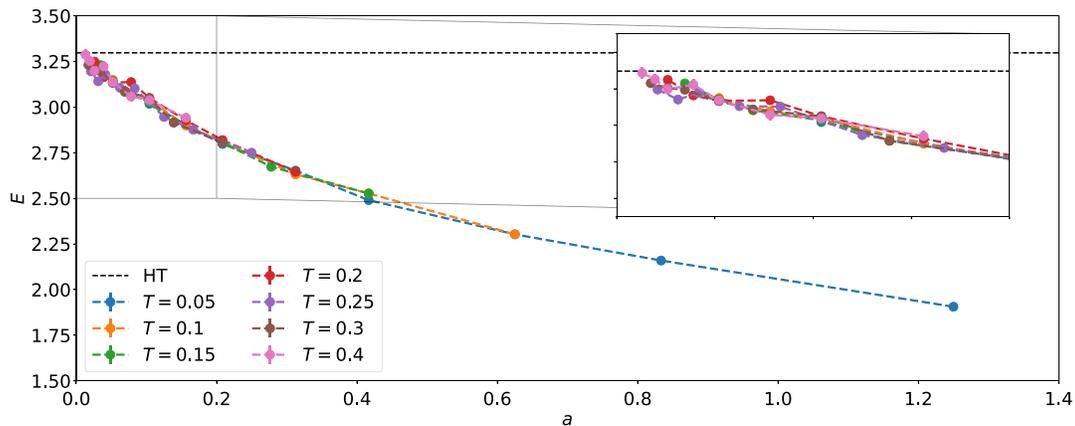


FIG. 22. The energy of the SU(2) bosonic model at  $m^2 = 1$  and  $\lambda = g^2 N = 0.5$  for various temperatures  $T$  and lattice sizes  $n_t$ . The horizontal axis is the lattice spacing  $a = 1/(Tn_t)$ . The black dashed line is the exact result from the HT approach at cutoff  $\Lambda = 14$ ,  $E_0 = 3.297$ .

TABLE XII. The results of systematic fitting for the SU(2) bosonic model with  $m^2 = 1$  and  $\lambda = g^2N = 0.5$ .

$a_{\max}$	$n_p$	$E$	$\chi^2$ per DOF	$a_{\max}$	$n_p$	$E$	$\chi^2$ per DOF
0.05	1	3.281(24)	1.70	0.30	1	3.2590(53)	1.43
0.05	2	3.317(57)	1.67	0.30	2	3.2841(88)	1.15
0.05	3	3.317(57)	1.67	0.30	3	3.285(14)	1.15
0.10	1	3.278(11)	1.46	0.40	1	3.2416(46)	2.31
0.10	2	3.312(26)	1.37	0.40	2	3.2878(76)	1.12
0.10	3	3.311(27)	1.37	0.40	3	3.282(12)	1.12
0.20	1	3.2725(66)	1.29	0.50	1	3.2202(41)	4.13
0.20	2	3.280(12)	1.28	0.50	2	3.2823(65)	1.23
0.20	3	3.287(22)	1.27	0.50	3	3.2900(98)	1.21

function

$$F(T, n_t) = E + \sum_{i=1}^{n_p} a_i \left( \frac{1}{Tn_t} \right)^i, \quad (62)$$

where  $E$  is the energy value in the continuum limit that we are seeking. We repeat the fit to Eq. (62) for  $n_p = 1, 2$ , and 3, while sliding  $a_{\max}$  from values closer to the continuum limit all the way to the largest that we have for each matrix model. The results for the SU(2) bosonic matrix model at  $\lambda = 0.5$  are summarized in Table XII.

We choose the final result based on where the results stabilize and become consistent across models with polynomials of different orders. This can be seen in Fig. 23, where a linear-function fit only works very close to the continuum limit, while higher-order polynomials can model the data at larger lattice spacing. We repeat the above procedure for all the couplings.

The results from this systematic procedure, choosing the  $n_p = 2$  polynomial model for the continuum limit and

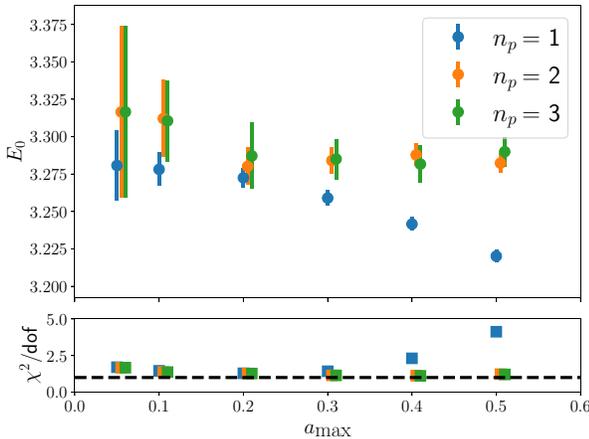


FIG. 23. The results of systematic fitting for the SU(2) bosonic model with  $m^2 = 1$  and  $\lambda = g^2N = 0.5$  using different data portions with polynomials of different order  $n_p$ . The lower panel shows the reduced  $\chi^2$  per DOF, which becomes very large for the low-order polynomials when larger lattice spacings are included, going well above the black dashed line that represents 1.

cutting the data at  $a_{\max} = 0.1$ , are summarized in Tables XIII and XIV for the SU(2) and the SU(3) bosonic matrix model, respectively.

As an alternative analysis to study fitting systematic effects, we can also fit the energy data at each temperature  $T$  with the function

$$F(T, n_t) = E(T) + \sum_i^{n_p} a_i(T) \left( \frac{1}{n_t} \right)^i, \quad (63)$$

where the  $E(T)$  energy in the large- $n_t$  limit has  $T$  dependence. Figure 24 shows that the fit for the lowest temperature is of poor quality and that makes the result for  $E(T)$  deviate from the expected result. This is due to lower temperatures requiring much larger values of  $n_t$  to achieve small discretization effects. We summarize the results of the fits (including only lattice sizes  $n_t > n_t^{\text{cut}} = 16$ ) in Table XV.

## VII. COMPARISON OF DIFFERENT APPROACHES

In this section, we compare the approaches investigated in this paper.

### A. Computational cost

In principle, the Hamiltonian-truncation method on a classical computer is the most straightforward and most reliable approach, because we are guaranteed to obtain the precise value at large cutoffs  $\Lambda$ . As we have seen for the SU(2) matrix models, the cutoff dependence of the ground-state energy is exponentially suppressed. However, this method is not tractable for SU( $N$ ) models with  $N \geq 3$ ,

TABLE XIII. The ground-state energy for the SU(2) bosonic model with  $m^2 = 1$  at coupling  $\lambda = g^2N = 0.5, 1.0$ , and 2.0. These are selected with an order  $n_p = 2$  polynomial and by removing lattice spacing larger than 0.1.

	$\lambda = 0.5$	$\lambda = 1.0$	$\lambda = 2.0$
$E_{0,MC}$	3.312(26)	3.497(33)	3.847(30)

TABLE XIV. The ground-state energy for the SU(3) bosonic model with  $m^2 = 1$  at coupling  $\lambda = g^2N = 0.5, 1.0, \text{ and } 2.0$ . These are selected with an order  $n_p = 2$  polynomial and by removing lattice spacing larger than 0.1.

	$\lambda = 0.5$	$\lambda = 1.0$	$\lambda = 2.0$
$E_{0,MC}$	8.836(38)	9.381(38)	10.236(41)

because the dimension of the Hilbert space is  $\Lambda^{2(N^2-1)}$  for the bosonic two-matrix model and  $2^{N^2-1}\Lambda^{2(N^2-1)}$  for the minimal-BMN model. In other words, the classical resources needed to solve the problem scale exponentially with the size of the system (number of degrees of freedom). In contrast to this, the quantum resources (number of qubits) scale polynomially with the size and logarithmically with the cutoff. Even with this favorable scaling, the necessary classical resources are too large for current technology when  $N \geq 3$ , but the situation could be different if we could use an error-corrected quantum computer on which adiabatic state preparation and quantum phase estimation could be used reliably. The VQE can circumvent the quantum-noise issue if a quantum computer with sufficiently many qubits becomes available.

The deep-learning method and the lattice Monte Carlo method do not suffer from the quick growth of the dimension of the Hilbert space. The deep-learning approach that we have adopted requires computational resources to study hyperparameters for the neural-network parametrizations of the wave function that will give the best (in a variational sense) energy. One of the advantages is that the information on the entire wave function can be obtained instead of just a few observables. The computational cost scales with the dimension of the neural network (the number of training parameters): the input dimension of the neural networks scales as  $O(N^2)$  for SU( $N$ ) matrices and the number of parameters in hidden layers scales as  $O(N^4)$ . For example, in bosonic models  $N = 20$  (about 400 bosons) and in the supersymmetric case  $N = 12$  (about 150 bosons with 150 fermions) are readily accessible on a laptop [76].

TABLE XV. The results of systematic fitting for the SU(2) matrix model with  $m^2 = 1$  and  $\lambda = g^2N = 0.5$  using a  $T$ -dependent energy function.

$T$	$n_t^{\text{cut}}$	$n_p$	$E$	$\chi^2$ per DOF
0.05	16	2	3.219(13)	1.54
0.10	16	2	3.294(18)	0.33
0.15	16	2	3.315(23)	0.99
0.20	16	2	3.274(27)	1.21
0.25	16	2	3.221(29)	1.05
0.30	16	2	3.263(31)	0.41
0.40	16	2	3.338(38)	0.52

TABLE XVI. The ground-state energy in the SU(2) bosonic two-matrix model obtained from the Hamiltonian truncation ( $E_{0,HT}$ ), the VQE ( $E_{0,VQE}$ ), deep learning ( $E_{0,DL}$ ), and lattice Monte Carlo ( $E_{0,MC}$ ) at coupling  $\lambda = g^2N = 0.5, 1.0, \text{ and } 2.0$ . The apparent poor performance of the VQE may come from using a small cutoff  $\Lambda = 4$  compared to the Hamiltonian truncation. The Hamiltonian-truncation results at  $\Lambda = 14$  have a negligible truncation error at this number of digits.

	$\lambda = 0.5$	$\lambda = 1.0$	$\lambda = 2.0$
$E_{0,HT}$	3.297	3.516	3.855
$E_{0,DL}$	3.302(2)	3.519(2)	3.857(3)
$E_{0,MC}$	3.312(26)	3.497(33)	3.847(30)
$E_{0,VQE}$	3.309	3.547	3.933

The lattice Monte Carlo simulation can be used even when  $N$  is very large, with the computational time [77] scaling roughly with  $N^3$ . However, we can only obtain expectation values of observables related to the ground state or a few low-lying states. For a reliable estimation of these observables, we have to study sufficiently low temperatures at various (large) lattice sizes, such that discretization effects become small or can be modeled accurately. Technically, the most complicated part is the uncertainty of the continuum extrapolation. Note that it is difficult to determine the spectrum of the excited modes precisely when using lattice Monte Carlo methods. Moreover, we should remark that typical lattice field configurations are not related to the wave function.

## B. Results

### 1. SU(2) bosonic model

We compare the ground-state energy obtained with different methods for the SU(2) bosonic matrix model. Very good agreement can be seen between the Hamiltonian-truncation method (at large cutoff  $\Lambda = 14$ ), the VQE method, the deep-learning (DL) method, and the lattice Monte Carlo (MC) method. The level of agreement deteriorates for the VQE as the model is explored at larger gauge couplings  $\lambda$ , because the VQE has a large truncation error due to the small cutoff  $\Lambda = 4$  used. The agreement might improve if a larger cutoff (larger number of qubits) is used or if more expressive variational wave functions are employed for the ground state. Both these directions require an increased computational time.

The values for the ground-state energy coming from all the approaches are compared in Table XVI. Note the error bars in the DL and MC approaches because of their stochastic Monte Carlo estimation nature. The DL-method error bars are statistical and they can be reduced with more Monte Carlo samples. On the other hand, the MC-method error bars are mostly from a parameter-inference procedure to extract the continuum-limit-extrapolated energy. These error bars can be reduced but there will be a lower

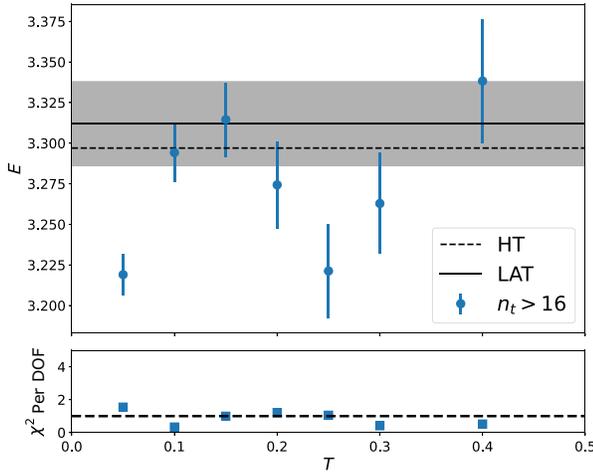


FIG. 24. The results of fitting different temperatures with a second-order polynomial and including all values of  $n_t > 16$  for the SU(2) matrix model with  $m^2 = 1$  and  $\lambda = g^2N = 0.5$ . The result from the HT is the black dashed line. The result from the original fitting strategy in Table XIII with the corresponding error bar is the black line (LAT) and the gray-shaded area. The lower panel shows the reduced  $\chi^2$  per DOF, with the dashed line representing 1.

limit given by the systematic effects of finding the right extrapolation model.

2. SU(3) bosonic model

We cannot treat the truncated Hamiltonian exactly (or with the VQE approach) due to the growth of the Hilbert space and the corresponding increase in the computational resources for the SU(3) gauge group. However, the DL and lattice MC techniques do not suffer from this unfavorable scaling. We can see very good agreement between these two approaches for the bosonic SU(3) matrix model. At the large coupling  $\lambda = 2.0$ , the ML bound is higher than the MC estimation in the continuum limit. At smaller coupling, they are compatible within errors. The results are summarized in Table XVII.

3. SU(2) minimal BMN

For the minimal-BMN matrix model with the SU(2) group, we know the exact value for the ground-state energy at each coupling  $\lambda$ . This energy is zero due

TABLE XVII. The ground-state energy in the SU(3) bosonic two-matrix model at coupling  $\lambda = g^2N = 0.5, 1.0, \text{ and } 2.0$ . The neural variational quantum Monte Carlo (DL) and lattice Monte Carlo (MC) techniques are compared.

	$\lambda = 0.5$	$\lambda = 1.0$	$\lambda = 2.0$
$E_{0,DL}$	8.824(7)	9.432(7)	10.426(8)
$E_{0,MC}$	8.836(38)	9.381(38)	10.236(41)

TABLE XVIII. The ground-state energy in the SU(2) minimal-BMN model obtained from the Hamiltonian truncation ( $E_{0,HT}$ ), the VQE ( $E_{0,VQE}$ ), and deep learning ( $E_{0,DL}$ ) at coupling  $\lambda = g^2N = 0.5, 1.0, \text{ and } 2.0$ . The VQE results are obtained at cutoff  $\Lambda = 2$  with the best optimizer and variational form but they may improve at larger cutoff. The Hamiltonian-truncation results at  $\Lambda = 11$  cannot be distinguished from the exact supersymmetric results of zero, at this number of digits.

	$\lambda = 0.5$	$\lambda = 1.0$	$\lambda = 2.0$
$E_{0,HT}$	0.000	0.000	0.000
$E_{0,DL}$	0.009(5)	0.014(6)	0.034(7)
$E_{0,VQE}$	0.027	0.079	0.177

to supersymmetry; hence this case is perfectly suitable to benchmark our computational approaches. The VQE approach is limited by the size of the Hilbert space that can be explored with a small number of qubits and the ground-state energy variational upper bound is close to zero for small coupling but much higher for larger couplings. Similarly, the DL method is compatible with the exact energy at small coupling and it becomes less compatible at larger couplings, due to the difficulty in determining a good variational ansatz for a strongly coupled system. We summarize the results in Table XVIII.

VIII. DISCUSSION AND CONCLUSIONS

Matrix models are simple (0+1)-dimensional theories that contain similar features as complicated high-dimensional quantum field theories related to quantum gravity via the holographic duality. In this paper, we conduct a systematic survey of quantum-simulation and deep-learning approaches to matrix models for the first time, as a first step to gauge the applicability and scalability of these modern and rapidly growing numerical techniques.

TABLE XIX. The ground-state energy in the SU(2) bosonic model at different coupling constants  $\lambda$  for various cutoffs  $\Lambda$ . The precision of the iterative sparse eigensolver is set to  $10^{-8}$ .

$\Lambda$	$\langle E_0   \hat{H}   E_0 \rangle$			
	$\lambda = 0.2$	$\lambda = 0.5$	$\lambda = 1.0$	$\lambda = 2.0$
3	3.132 304 65	3.282 851 59	3.458 479 92	3.669 040 08
4	3.134 063 07	3.298 943 63	3.526 254 44	3.895 478 37
5	3.133 908 03	3.296 492 79	3.512 117 72	3.833 392 47
6	3.133 927 06	3.297 025 15	3.516 507 58	3.859 506 13
7	3.133 924 87	3.296 920 58	3.515 335 49	3.850 814 24
8	3.133 925 19	3.296 946 63	3.515 733 42	3.854 575 65
9	3.133 925 14	3.296 940 61	3.515 612 21	3.853 167 16
10	3.133 925 15	3.296 942 26	3.515 656 35	3.853 808 64
11	3.133 925 15	3.296 941 83	3.515 641 51	3.853 546 47
12	3.133 925 15	3.296 941 96	3.515 647 26	3.853 670 79
13	3.133 925 15	3.296 941 92	3.515 645 18	3.853 616 68
14	3.133 925 15	3.296 941 94	3.515 646 02	3.853 643 30

TABLE XX. The ground-state violation of the singlet constraint in the SU(2) bosonic model at different coupling constants  $\lambda$  for various cutoffs  $\Lambda$ . The precision of the iterative sparse eigensolver is set to  $10^{-8}$ .

$\Lambda$	$\sum_{\alpha} \langle E_0   \hat{G}_{\alpha}^2   E_0 \rangle$			
	$\lambda = 0.2$	$\lambda = 0.5$	$\lambda = 1.0$	$\lambda = 2.0$
3	0.000 224 293 831	0.003 688 212 672	0.018 964 932 725	0.060 510 041 162
4	0.000 184 221 658	0.002 400 096 378	0.011 889 840 764	0.045 924 094 777
5	0.000 004 511 062	0.000 178 633 922	0.001 732 981 158	0.011 493 185 261
6	0.000 002 506 324	0.000 101 373 665	0.000 986 724 060	0.006 352 557 866
7	0.000 000 087 650	0.000 009 947 253	0.000 176 500 569	0.001 788 911 061
8	0.000 000 049 542	0.000 005 629 387	0.000 099 441 023	0.001 011 302 266
9	0.000 000 002 327	0.000 000 683 314	0.000 020 545 750	0.000 315 815 031
10	0.000 000 001 312	0.000 000 389 452	0.000 011 932 039	0.000 185 582 780
11	0.000 000 000 077	0.000 000 055 646	0.000 002 785 207	0.000 062 685 672
12	0.000 000 000 044	0.000 000 032 055	0.000 001 643 162	0.000 037 949 122
13	0.000 000 000 003	0.000 000 005 217	0.000 000 421 954	0.000 013 712 553
14	0.000 000 000 002	0.000 000 003 031	0.000 000 252 683	0.000 008 471 522

Specifically, we consider a bosonic two-matrix model and its supersymmetric extension (minimal-BMN model) and compute the spectrum of the low-lying modes. In fact, efficiently and precisely computing the energy spectrum—e.g., for the ground state—is a necessary requirement for any future algorithm aiming at solving the time evolution of the system based on the quantum Hamiltonian. We use the Hamiltonian-truncation method (in the Fock basis), the VQE method, and variational quantum Monte Carlo with a neural-network ansatz for the wave function. The choice of these methods is dictated by the success they have had in many different physical sciences, from quantum chemistry to complicated many-body models such as spin glasses. We also perform lattice Monte Carlo simulations based on the Euclidean path integral for the bosonic model as a comparison with more traditional methods.

With the Hamiltonian-truncation method, we consider the matrix model with a fixed number of modes for each boson up to a cutoff value that is slowly taken to infinity in order to remove truncation effects. The spectrum of the

truncated Hamiltonian is obtained using QuTiP on classical computers. We explore how quickly the correct answer at infinite cutoff can be recovered. With this setup, we only study the bosonic matrix model with the SU(2) gauge group, because the SU(3) model contains more bosonic degrees of freedom (16 compared to six), resulting in a much larger Hilbert space: this would require computational resources not available to us at the moment. We find that the truncation effects to the low-energy spectrum are suppressed exponentially with respect to the cutoff  $\Lambda$ . This is a favorable scaling and we are able to prove it empirically for the first time in this work. Note that the number of qubits in quantum simulation scales as  $\log_2 \Lambda$ ; hence the truncation error can be suppressed double exponentially with respect to the number of qubits. It would be natural to assume the same scaling of the truncation errors for larger systems (larger gauge group  $N$  or more matrices  $D$ ). Given a fixed threshold for the truncation effects that one is willing to accept, the number of qubits needed to study a matrix model with a SU( $N$ ) gauge group and  $D$  bosonic matrices is polynomial in  $D \cdot (N^2 - 1)$ . Moreover,

TABLE XXI. The energy expectation value  $\langle E'_0 | \hat{H} | E'_0 \rangle$  for the ground state of the  $\hat{H}'$  Hamiltonians of the SU(2) minimal-BMN model at different coupling constants  $\lambda$  for various cutoffs  $\Lambda$ . Only the first 12 digits are shown.

$\Lambda$	$\langle E'_0   \hat{H}   E'_0 \rangle$			
	$\lambda = 0.2$	$\lambda = 0.5$	$\lambda = 1.0$	$\lambda = 2.0$
3	-0.000 348 435 200	-0.003 873 948 083	-0.019 907 205 965	-0.084 936 973 789
4	0.000 114 126 215	0.002 116 374 610	0.013 418 689 187	0.060 446 205 687
5	-0.000 003 216 759	-0.000 102 255 010	-0.000 909 125 958	-0.005 803 433 350
6	0.000 002 429 833	0.000 134 468 904	0.001 628 985 778	0.012 489 096 584
7	-0.000 000 036 383	-0.000 003 214 703	-0.000 051 505 094	-0.000 504 975 266
8	0.000 000 063 447	0.000 010 165 342	0.000 231 428 987	0.002 895 837 534
9	-0.000 000 000 464	-0.000 000 099 555	-0.000 002 210 416	-0.000 018 843 848
10	0.000 000 002 099	0.000 000 910 719	0.000 036 818 848	0.000 724 187 898
11	-0.000 000 000 002	0.000 000 001 411	0.000 000 259 748	0.000 012 480 441

TABLE XXII. The gauge-constraint violation  $\sum_{\alpha} \langle E'_0 | \hat{G}_{\alpha}^2 | E'_0 \rangle$  for the ground state of the  $\hat{H}'$  Hamiltonians of the SU(2) minimal-BMN model at different coupling constants  $\lambda$  for various cutoffs  $\Lambda$ . Only the first 12 digits are shown.

$\Lambda$	$\sum_{\alpha} \langle E'_0   \hat{G}_{\alpha}^2   E'_0 \rangle$			
	$\lambda = 0.2$	$\lambda = 0.5$	$\lambda = 1.0$	$\lambda = 2.0$
3	0.000 027 144 384	0.000 270 665 767	0.001 217 678 752	0.004 391 570 468
4	0.000 003 466 155	0.000 071 748 567	0.000 483 864 489	0.002 211 074 147
5	0.000 000 377 832	0.000 013 303 385	0.000 129 198 647	0.000 837 166 225
6	0.000 000 087 652	0.000 005 300 120	0.000 065 635 237	0.000 487 463 732
7	0.000 000 007 701	0.000 000 831 344	0.000 015 273 650	0.000 161 843 291
8	0.000 000 002 422	0.000 000 413 367	0.000 009 393 689	0.000 110 264 245
9	0.000 000 000 211	0.000 000 063 891	0.000 002 115 537	0.000 035 041 177
10	0.000 000 000 080	0.000 000 036 382	0.000 001 444 227	0.000 026 392 551
11	0.000 000 000 007	0.000 000 005 736	0.000 000 326 893	0.000 008 241 117

increasing the cutoff level can be done efficiently due to the logarithmic scaling with  $\Lambda$  together with the exponential suppression of truncation errors. Such a favorable scaling would be good news for quantum simulation of quantum gravity via holography. Quantum simulations of matrix models can naturally lead to “quantum gravity in the lab” [19,78,79]. So far, toy models such as the Sachdev-Ye-Kitaev (SYK) models have been considered (regarding quantum simulation protocols of the SYK model on universal quantum computers, see Refs. [80,81]) but matrix models can be a much better, theoretically well-controlled, setup for quantum gravity. Moreover, understanding these matrix models through quantum simulations could help in the context of quantum error-correction studies due to their connection with the holographic conjecture [82–84].

We also want to understand how quantum hardware can be utilized efficiently to study matrix models in the NISQ era. In particular, we want to see if currently available algorithms and software frameworks allow us to correctly reproduce the results of the Hamiltonian truncation. This is a necessary step for any future endeavor in this direction. For this reason, we resort to a hybrid quantum-classical algorithm, the VQE, to extract the ground state of the models. We use the IBM QISKIT software to estimate the

ground-state energy at small system sizes and compare the numbers with those from exact diagonalization of the truncated Hamiltonian using classical hardware. In the future, we plan to use quantum hardware for the simulations, but first we need to investigate the performance of the variational quantum circuits and the classical optimizers. Our results for the bosonic model show very good agreement at weak coupling but there is a growing discrepancy between the least upper bound for the energy from the VQE and the exact diagonalization value of the truncated Hamiltonian as the coupling becomes larger. This is not surprising, because the quantum circuits we use to approximate the ground-state wave function are not crafted for matrix models. We need to find better circuits and for that purpose, the deep-learning approach may give us some hints. Moreover, we note that the upper bound given by the VQE for the bosonic model is usually better than the one in the minimal-BMN model. In the case of the minimal-BMN model, the energy of the ground state is zero up to exponentially small truncation effects. However, we note that the VQE is giving an upper bound that is nonzero and the difference with the exact value grows with the coupling. These results point out an interesting possibility: the use of matrix models as a challenging test bed for quantum

TABLE XXIII. The angular-momentum-constraint violation  $|\sum_{\alpha} \langle E'_0 | \hat{M}_{\alpha} | E'_0 \rangle - J|$  for the ground state of the  $\hat{H}'$  Hamiltonians of the SU(2) minimal-BMN model at different coupling constants  $\lambda$  for various cutoffs  $\Lambda$ . Only the first 12 digits are shown and  $J = 0$ .

$\Lambda$	$ \sum_{\alpha} \langle E'_0   \hat{M}_{\alpha}   E'_0 \rangle - J $			
	$\lambda = 0.2$	$\lambda = 0.5$	$\lambda = 1.0$	$\lambda = 2.0$
3	0.000 016 801 483	0.000 182 372 511	0.000 913 116 520	0.003 744 814 441
4	0.000 000 040 458	0.000 001 248 745	0.000 008 311 970	0.000 017 373 811
5	0.000 000 225 739	0.000 007 305 553	0.000 064 178 889	0.000 381 297 761
6	0.000 000 000 342	0.000 000 015 301	0.000 000 143 203	0.000 002 510 586
7	0.000 000 004 147	0.000 000 371 546	0.000 005 900 379	0.000 056 563 759
8	0.000 000 000 002	0.000 000 000 600	0.000 000 031 637	0.000 000 714 895
9	0.000 000 000 099	0.000 000 023 729	0.000 000 662 279	0.000 009 762 673
10	0.000 000 000 000	0.000 000 000 077	0.000 000 005 308	0.000 000 175 686
11	0.000 000 000 003	0.000 000 001 820	0.000 000 087 067	0.000 001 946 922

TABLE XXIV. The results for the internal energy of the two-matrix bosonic model with gauge group SU(3) and  $\lambda = 0.5$ .  $N_{\text{cfgs}}$  is the number of configurations used in the analysis and  $N_{\text{drop}}$  is the number of molecular dynamics time units (MDTUs) between successive configurations.  $\tau$  is the integrated autocorrelation time (in units of  $N_{\text{drop}}$ ) and it is always close to unity, indicating that there are no strong autocorrelations.

$T$	$n_t$	$E$	$N_{\text{cfgs}}$	$N_{\text{drop}}$	$\tau$
0.400	16	7.845(70)	980	50	0.810
0.350	16	7.639(41)	1980	50	0.950
0.300	16	7.481(54)	980	50	1.040
0.250	16	7.354(52)	990	100	0.900
0.200	16	7.072(44)	990	100	0.850
0.150	16	6.710(38)	988	80	1.130
0.100	16	6.131(29)	988	80	0.750
0.050	16	5.071(20)	990	100	0.980
0.025	16	3.873(13)	990	100	1.020
0.400	24	7.912(60)	990	100	0.990
0.350	24	8.036(43)	1980	50	1.000
0.300	24	7.759(57)	990	100	0.990
0.250	24	7.773(52)	990	100	1.120
0.200	24	7.558(44)	990	100	1.450
0.150	24	7.134(38)	990	100	0.800
0.100	24	6.669(31)	990	100	0.930
0.050	24	5.670(20)	990	100	1.170
0.025	24	4.613(13)	990	100	0.910
0.400	32	8.284(69)	990	100	0.860
0.350	32	8.162(44)	1980	50	0.940
0.300	32	8.113(57)	990	100	1.000
0.250	32	8.001(49)	990	100	1.030
0.200	32	7.720(45)	990	100	0.990
0.150	32	7.513(39)	990	100	0.770
0.100	32	7.053(31)	990	100	0.820
0.050	32	6.137(22)	990	100	1.130
0.025	32	5.064(14)	990	100	0.820
0.400	48	8.430(68)	980	50	0.980
0.350	48	8.334(44)	1980	50	0.860
0.300	48	8.191(58)	980	50	0.790
0.250	48	8.219(51)	988	80	0.990
0.200	48	7.986(43)	988	80	0.680
0.150	48	7.809(41)	988	80	1.020
0.100	48	7.489(30)	990	100	0.960
0.050	48	6.697(21)	990	100	1.130
0.025	48	5.723(15)	991	102	1.030
0.400	64	8.455(62)	980	50	0.930
0.350	64	8.456(43)	1980	50	1.150
0.300	64	8.418(57)	980	50	0.790
0.250	64	8.258(50)	980	50	1.020
0.200	64	8.192(45)	980	50	1.020
0.150	64	8.024(38)	980	50	0.790
0.100	64	7.810(31)	980	50	0.830
0.050	64	7.036(22)	980	50	1.030
0.025	64	6.125(15)	980	50	1.250
0.400	96	8.654(68)	980	50	1.140
0.350	96	8.518(43)	1980	50	0.940
0.300	96	8.630(56)	980	50	1.250
0.250	96	8.405(30)	2760	25	1.380
0.200	96	8.458(48)	980	50	0.870

*Continued*

TABLE XXIV. Continued

$T$	$n_t$	$E$	$N_{\text{cfgs}}$	$N_{\text{drop}}$	$\tau$
0.150	96	8.215(39)	980	50	0.930
0.100	96	8.009(31)	980	50	0.900
0.050	96	7.475(22)	980	50	0.930
0.025	96	6.672(15)	980	50	0.860
0.400	128	8.740(54)	1267	12	5.510
0.350	128	8.568(46)	1980	50	0.990
0.300	128	8.636(59)	925	8	4.250
0.250	128	8.520(52)	1025	8	3.150
0.200	128	8.481(32)	2000	20	1.350
0.150	128	8.413(29)	1950	20	1.700
0.100	128	8.222(14)	4900	10	1.300
0.050	128	7.761(16)	2000	20	1.140
0.025	128	7.058(11)	2000	20	1.340
0.400	192	8.754(47)	1980	50	1.470
0.350	192	8.622(43)	1980	50	0.980
0.300	192	8.654(41)	1980	50	1.250
0.250	192	8.651(37)	1980	50	1.180
0.200	192	8.560(33)	1980	50	0.950
0.150	192	8.512(28)	1980	50	1.020
0.100	192	8.369(24)	1980	50	0.920
0.050	192	8.028(16)	1980	50	0.980
0.025	192	7.478(11)	1980	50	1.120

algorithms and quantum devices. Using the VQE with a state-vector simulator as a prototypical best-case scenario, we can tackle progressively harder problems within the context of matrix models by simply tuning the gauge coupling strength (keeping the number of qubits fixed) or by adding fermions and bosons (increasing the number of qubits). The same possibility holds for scenarios where noise is introduced in the simulation or using real quantum hardware, thus allowing one to test error-mitigation strategies using matrix models.

To explore what deep learning can do for the matrix models, we use a variational quantum Monte Carlo approach with an autoregressive neural flow as the parametrization for the wave function of the ground state. This approach can be readily applicable to larger system sizes and larger gauge groups. We study the SU(2) bosonic model and the SU(2) minimal-BMN model and we also look at the SU(3) bosonic model. Because we do not study the SU(3) bosonic model via the Hamiltonian truncation due to the large Hilbert space, a comparison with lattice Monte Carlo is crucial to establish a baseline for the results. This neural-network-based variational quantum Monte Carlo approach is able to obtain a good approximation of the ground-state energy for all the parameters that we study in the bosonic and minimal-BMN matrix models. This is a very promising outcome, which is demonstrated here for the first time. We believe that this deep-learning approach is readily applicable to various problems such as real-time evolution or complex models such as the full BMN or Banks-Fischler-Shenker-Susskind (BFSS) matrix

TABLE XXV. The results for the internal energy of the two-matrix bosonic model with gauge group SU(3) and  $\lambda = 1.0$ .  $N_{\text{cfs}}$  is the number of configurations used in the analysis and  $N_{\text{drop}}$  is the number of molecular dynamics time units (MDTUs) between successive configurations.  $\tau$  is the integrated autocorrelation time (in units of  $N_{\text{drop}}$ ) and it is always close to unity, indicating that there are no strong autocorrelations.

$T$	$n_t$	$E$	$N_{\text{cfs}}$	$N_{\text{drop}}$	$\tau$
0.400	16	8.084(49)	1980	50	0.930
0.350	16	7.994(45)	1980	50	0.930
0.300	16	7.880(41)	1980	50	0.850
0.250	16	7.608(37)	1980	50	0.920
0.200	16	7.291(32)	1980	50	0.980
0.150	16	6.936(28)	1980	50	0.950
0.100	16	6.352(22)	1980	50	1.280
0.050	16	5.170(15)	1980	50	1.070
0.025	16	3.9119(91)	1980	50	1.040
0.400	24	8.433(49)	1980	50	1.260
0.350	24	8.310(45)	1980	50	1.030
0.300	24	8.214(42)	1980	50	0.950
0.250	24	8.052(38)	1980	50	0.820
0.200	24	7.807(33)	1980	50	0.840
0.150	24	7.470(29)	1980	50	0.980
0.100	24	6.943(22)	1980	50	1.130
0.050	24	5.880(15)	1980	50	1.130
0.025	24	4.6610(99)	1980	50	0.910
0.400	32	8.759(50)	1980	50	0.840
0.350	32	8.572(46)	1980	50	1.130
0.300	32	8.521(42)	1980	50	0.980
0.250	32	8.247(38)	1980	50	1.060
0.200	32	8.151(34)	1980	50	0.890
0.150	32	7.857(29)	1980	50	0.870
0.100	32	7.328(23)	1980	50	1.160
0.050	32	6.311(15)	1980	50	1.000
0.025	32	5.182(10)	1980	50	0.850
0.400	48	8.834(50)	1980	50	0.770
0.350	48	8.808(48)	1980	50	0.960
0.300	48	8.654(42)	1980	50	1.090
0.250	48	8.654(39)	1980	50	0.930
0.200	48	8.568(35)	1980	50	0.950
0.150	48	8.251(29)	1980	50	0.820
0.100	48	7.825(24)	1980	50	1.090
0.050	48	6.972(16)	1980	50	0.900
0.025	48	5.877(11)	1980	50	1.170
0.400	64	9.033(50)	1980	50	0.870
0.350	64	8.958(48)	1980	50	1.020
0.300	64	8.934(43)	1980	50	0.910
0.250	64	8.768(39)	1980	50	0.840
0.200	64	8.605(35)	1980	50	0.970
0.150	64	8.482(30)	1980	50	0.960
0.100	64	8.139(25)	1980	50	0.970
0.050	64	7.362(17)	1980	50	0.930
0.025	64	6.324(11)	1980	50	0.970
0.400	96	9.173(49)	1980	50	1.080
0.350	96	9.092(46)	1980	50	0.810
0.300	96	8.997(43)	1980	50	1.340
0.250	96	8.981(38)	1980	50	0.900
0.200	96	8.851(35)	1980	50	1.350

*Continued*

TABLE XXV. Continued

$T$	$n_t$	$E$	$N_{\text{cfs}}$	$N_{\text{drop}}$	$\tau$
0.150	96	8.748(30)	1980	50	0.910
0.100	96	8.476(25)	1980	50	1.010
0.050	96	7.844(17)	1980	50	0.930
0.025	96	6.928(12)	1980	50	0.970
0.400	128	9.200(50)	1980	50	0.940
0.350	128	9.104(47)	1980	50	0.840
0.300	128	9.143(43)	1980	50	0.990
0.250	128	9.112(40)	1980	50	0.960
0.200	128	8.962(35)	1980	50	0.850
0.150	128	8.866(31)	1980	50	0.950
0.100	128	8.715(24)	1980	50	0.990
0.050	128	8.138(17)	1980	50	1.130
0.025	128	7.326(12)	1980	50	1.100
0.400	192	9.238(51)	1980	50	1.380
0.350	192	9.252(48)	1980	50	1.110
0.300	192	9.182(43)	1980	50	0.980
0.250	192	9.148(40)	1980	50	0.790
0.200	192	9.068(35)	1980	50	1.130
0.150	192	9.012(30)	1980	50	0.940
0.100	192	8.905(26)	1820	50	0.790
0.050	192	8.469(17)	1980	50	1.100
0.025	192	7.834(12)	1980	50	1.050

models. Moreover, the ability to incorporate discrete and continuous symmetries of the Hamiltonian into the neural parametrization of the wave function, demonstrated in other systems such as lattice models or quantum chemistry, will lead to future improvements.

## ACKNOWLEDGMENTS

We wish to thank David Berenstein for useful suggestions and discussions on matrix models and Fabrizio Minganti for discussions about QuTiP. We also thank Antonio Mezzacapo for useful suggestions on quantum computing applied to gauge theories. This material is based upon work supported in part by the U.S. Department of Energy, Office of Science, National Quantum Information Science Research Centers, Co-design Center for Quantum Advantage (C2QA) under Contract No. DESC0012704. This project was supported in part by the U.S. Department of Energy, Office of Science, Office of Workforce Development for Teachers and Scientists (WDTS) under the Science Undergraduate Laboratory Internships Program (SULI). F.N. is supported in part by: Nippon Telegraph and Telephone Corporation (NTT) Research; the Japan Science and Technology Agency (JST) [via the Quantum Leap Flagship Program (Q-LEAP), the Moonshot R&D Grant No. JPMJMS2061, and the Centers of Research Excellence in Science and Technology (CREST) Grant No. JPMJCR1676]; the Japan Society for the Promotion of Science (JSPS) [via the Grants-in-Aid for Scientific Research (KAKENHI) Grant No. JP20H00134 and

TABLE XXVI. The results for the internal energy of the two-matrix bosonic model with gauge group SU(3) and  $\lambda = 2.0$ .  $N_{\text{cfgs}}$  is the number of configurations used in the analysis and  $N_{\text{drop}}$  is the number of molecular dynamics time units (MDTUs) between successive configurations.  $\tau$  is the integrated autocorrelation time (in units of  $N_{\text{drop}}$ ) and it is always close to unity, indicating that there are no strong autocorrelations.

$T$	$n_t$	$E$	$N_{\text{cfgs}}$	$N_{\text{drop}}$	$\tau$
0.400	16	8.831(54)	1980	50	1.120
0.350	16	8.510(48)	1980	50	0.900
0.300	16	8.275(44)	1980	50	0.990
0.250	16	8.104(40)	1980	50	0.960
0.200	16	7.780(35)	1980	50	0.830
0.150	16	7.322(29)	1980	50	1.060
0.100	16	6.614(24)	1980	50	0.950
0.050	16	5.316(15)	1980	50	0.840
0.025	16	3.9537(95)	1980	50	0.960
0.400	24	9.186(54)	1980	50	1.100
0.350	24	9.069(51)	1980	50	0.910
0.300	24	8.864(46)	1980	50	0.910
0.250	24	8.606(41)	1980	50	0.840
0.200	24	8.312(38)	1980	50	0.730
0.150	24	7.901(32)	1980	50	1.050
0.100	24	7.349(24)	1980	50	0.960
0.050	24	6.064(16)	1980	50	0.950
0.025	24	4.762(10)	1980	50	1.140
0.400	32	9.394(54)	1980	50	0.980
0.350	32	9.097(48)	1980	50	0.870
0.300	32	9.238(47)	1980	50	1.150
0.250	32	8.954(40)	1980	50	1.070
0.200	32	8.714(37)	1980	50	1.120
0.150	32	8.337(31)	1980	50	0.910
0.100	32	7.761(25)	1980	50	1.050
0.050	32	6.604(17)	1980	50	1.150
0.025	32	5.321(11)	1980	50	1.000
0.400	48	9.627(54)	1980	50	1.000
0.350	48	9.552(52)	1980	50	0.870
0.300	48	9.385(47)	1980	50	0.980
0.250	48	9.348(42)	1980	50	0.960
0.200	48	9.184(38)	1980	50	1.000
0.150	48	8.813(32)	1980	50	0.890
0.100	48	8.331(25)	1980	50	1.130
0.050	48	7.315(18)	1980	50	0.950
0.025	48	6.083(11)	1980	50	1.050
0.400	64	9.829(55)	1980	50	0.990
0.350	64	9.711(51)	1980	50	0.860
0.300	64	9.570(47)	1980	50	0.950
0.250	64	9.470(42)	1980	50	0.940
0.200	64	9.325(37)	1980	50	0.890
0.150	64	9.112(33)	1980	50	0.910
0.100	64	8.669(26)	1980	50	1.080
0.050	64	7.756(18)	1980	50	0.920
0.025	64	6.605(12)	1980	50	0.980
0.400	96	9.882(54)	1980	50	1.130
0.350	96	9.927(52)	1980	50	0.980
0.300	96	9.822(46)	1980	50	1.040
0.250	96	9.778(43)	1980	50	1.120
0.200	96	9.647(39)	1980	50	0.990

Continued

TABLE XXVI. Continued

$T$	$n_t$	$E$	$N_{\text{cfgs}}$	$N_{\text{drop}}$	$\tau$
0.150	96	9.414(33)	1980	50	0.940
0.100	96	9.114(26)	1980	49	1.110
0.050	96	8.312(18)	1980	49	1.010
0.025	96	7.290(12)	1980	50	1.130
0.400	128	9.920(53)	1980	50	1.160
0.350	128	10.017(51)	1980	50	1.150
0.300	128	10.047(49)	1980	50	0.920
0.250	128	9.864(42)	1980	50	0.880
0.200	128	9.797(39)	1980	50	0.990
0.150	128	9.656(34)	1980	50	0.950
0.100	128	9.338(27)	1980	50	1.200
0.050	128	8.677(18)	1980	50	0.960
0.025	128	7.748(12)	1980	50	0.940
0.400	192	10.129(55)	1980	50	0.980
0.350	192	10.124(52)	1980	50	0.970
0.300	192	10.074(46)	1980	50	0.980
0.250	192	9.915(44)	1980	50	0.940
0.200	192	9.891(40)	1980	50	1.120
0.150	192	9.848(34)	1980	50	1.090
0.100	192	9.659(27)	1980	50	1.120
0.050	192	9.128(19)	1980	50	0.970
0.025	192	8.335(13)	1980	50	1.080

the JSPS-RFBR Grant No. JPJSBP120194828]; the Army Research Office (ARO) (Grant No. W911NF-18-1-0358); the Asian Office of Aerospace Research and Development (AOARD) (via Grant No. FA2386-20-1-4069); and the Foundational Questions Institute Fund (FQXi) via Grant No. FQXi-IAF19-06. E.R. is supported by Nippon Telegraph and Telephone Corporation (NTT) Research. We gratefully acknowledge the use of the Supercomputer HOKUSAI BigWaterfall of the Information Systems Division of RIKEN.

## APPENDIX A: OPERATOR FORMALISM AND PATH-INTEGRAL FORMALISM

### 1. Hilbert space with and without singlet constraint

When quantum gauge theory is treated in the Hamiltonian formulation in the temporal gauge ( $A_t = 0$  gauge), it is often said that “physical states are gauge singlets.” The canonical partition function can be written as a trace over the space of gauge-invariant states  $\mathcal{H}_{\text{inv}}$ ,

$$Z(T) = \text{Tr}_{\mathcal{H}_{\text{inv}}} \left( e^{-\hat{H}/T} \right). \quad (\text{A1})$$

There is another, but equivalent, formulation with nonsinglet Hilbert space (“extended” Hilbert space)  $\mathcal{H}_{\text{ext}}$ , in which the quantum states that map to each other via gauge transformation are identified. The partition function is written as

$$Z(T) = \frac{1}{\text{vol}(G)} \int_G dg \text{Tr}_{\mathcal{H}_{\text{ext}}} \left( \hat{g} e^{-\hat{H}/T} \right). \quad (\text{A2})$$

TABLE XXVII. The results for the internal energy of the two-matrix bosonic model with gauge group  $SU(2)$  and  $\lambda = 2.0$ .  $N_{\text{cfigs}}$  is the number of configurations used in the analysis and  $N_{\text{drop}}$  is the number of molecular dynamics time units (MDTUs) between successive configurations.  $\tau$  is the integrated autocorrelation time (in units of  $N_{\text{drop}}$ ) and it is always close to unity, indicating that there are no strong autocorrelations.

$T$	$n_t$	$E$	$N_{\text{cfigs}}$	$N_{\text{drop}}$	$\tau$
0.400	16	3.296(32)	1980	50	0.960
0.300	16	3.121(26)	1980	50	0.930
0.250	16	3.017(24)	1980	50	0.820
0.200	16	2.924(22)	1980	50	0.920
0.150	16	2.750(18)	1980	50	0.780
0.100	16	2.477(15)	1980	50	1.120
0.050	16	1.9920(94)	1980	50	1.080
0.400	24	3.428(32)	1980	50	1.120
0.300	24	3.289(28)	1980	50	0.920
0.250	24	3.237(25)	1980	50	1.240
0.200	24	3.101(22)	1980	50	1.060
0.150	24	2.935(19)	1980	50	1.120
0.100	24	2.722(14)	1980	50	0.990
0.050	24	2.2931(99)	1980	50	1.050
0.400	32	3.531(32)	1980	50	0.940
0.300	32	3.373(28)	1980	50	1.000
0.250	32	3.394(27)	1980	50	1.150
0.200	32	3.258(22)	1980	50	0.990
0.150	32	3.157(20)	1980	50	1.000
0.100	32	2.922(16)	1980	50	0.880
0.050	32	2.464(10)	1980	50	0.990
0.400	48	3.649(34)	1980	50	1.160
0.300	48	3.524(28)	1980	50	0.980
0.250	48	3.518(26)	1980	50	1.040
0.200	48	3.446(24)	1980	50	1.120
0.150	48	3.328(20)	1980	50	1.140
0.100	48	3.129(16)	1980	50	0.930
0.050	48	2.721(10)	1980	50	0.970
0.400	64	3.666(34)	1980	50	1.040
0.300	64	3.697(30)	1980	50	0.930
0.250	64	3.560(25)	1980	50	0.960
0.200	64	3.512(23)	1980	50	1.180
0.150	64	3.398(19)	1980	50	0.820
0.100	64	3.246(16)	1980	50	1.010
0.050	64	2.916(11)	1980	50	1.100
0.400	96	3.736(34)	1980	50	1.170
0.300	96	3.733(29)	1980	50	1.100
0.250	96	3.658(27)	1980	50	0.800
0.200	96	3.630(24)	1980	50	1.040
0.150	96	3.530(20)	1980	50	0.870
0.100	96	3.379(16)	1980	50	0.880
0.050	96	3.122(11)	1980	50	0.950
0.400	128	3.796(34)	1980	50	0.870
0.300	128	3.725(28)	1980	50	1.170
0.250	128	3.700(26)	1980	50	1.030
0.200	128	3.640(23)	1980	50	1.050
0.150	128	3.617(20)	1980	50	0.950
0.100	128	3.507(16)	1980	50	0.990
0.050	128	3.262(12)	1980	50	1.170
0.400	192	3.741(34)	1980	50	0.940

Continued

TABLE XXVII. Continued

$T$	$n_t$	$E$	$N_{\text{cfigs}}$	$N_{\text{drop}}$	$\tau$
0.300	192	3.740(28)	1980	50	1.060
0.250	192	3.762(26)	1980	50	0.990
0.200	192	3.737(23)	1980	50	0.900
0.150	192	3.705(20)	1980	50	0.960
0.100	192	3.581(16)	1980	50	0.950
0.050	192	3.425(12)	1980	50	1.170

Here,  $G = SU(N)$  and the integral is taken with the Haar measure. Note that

$$\hat{\mathcal{P}} \equiv \frac{1}{\text{vol}(G)} \int_G dg \hat{g} \quad (\text{A3})$$

is a projection operator from  $\mathcal{H}_{\text{ext}}$  to  $\mathcal{H}_{\text{inv}}$  that satisfies  $\hat{\mathcal{P}}^2 = \hat{\mathcal{P}}$ .

We show the equivalence of Eqs. (A1) and (A2). This is a well-known fact but we present the derivation to make the paper self contained. Essentially, we repeat the materials in Refs. [18,85].

Suppose that an energy eigenstate  $|\Phi\rangle \in \mathcal{H}_{\text{ext}}$  is invariant only under  $G_\Phi \subset G$ . From  $|\Phi\rangle$ , we can obtain  $|\Phi\rangle_{\text{inv}} \in \mathcal{H}_{\text{inv}}$  as

$$|\Phi\rangle_{\text{inv}} = \frac{1}{\sqrt{C_\Phi}} \times \frac{1}{\text{vol}(G)} \int_G dg (\hat{g}|\Phi\rangle). \quad (\text{A4})$$

The normalization factor  $C_\Phi$  is determined such that the norm of  $|\Phi\rangle_{\text{inv}}$  is 1. Namely,

$$\begin{aligned} C_\Phi &= \frac{1}{[\text{vol}(G)]^2} \int_G dg \int_G dg' \langle \Phi | \hat{g}^{-1} \hat{g}' | \Phi \rangle \\ &= \frac{1}{\text{vol}(G)} \int_G dg \langle \Phi | \hat{g} | \Phi \rangle \\ &= \frac{\text{vol}(G_\Phi)}{\text{vol}(G)}. \end{aligned} \quad (\text{A5})$$

Again, this is the inverse of the overcounting factor. Therefore,

$$\begin{aligned} \text{Tr}_{\mathcal{H}_{\text{inv}}} \left( e^{-\hat{H}/T} \right) &= \sum_\Phi \frac{\text{vol}(G_\Phi)}{\text{vol}(G)}_{\text{inv}} \langle \Phi | e^{-\hat{H}/T} | \Phi \rangle_{\text{inv}} \\ &= \sum_\Phi \frac{\text{vol}(G_\Phi)}{\text{vol}(G)} \times \frac{1}{C_\Phi} \frac{1}{[\text{vol}(G)]^2} \int_G dg \\ &\quad \int_G dg' \langle \Phi | \hat{g}^{-1} e^{-\hat{H}/T} \hat{g}' | \Phi \rangle \\ &= \sum_\Phi \frac{1}{\text{vol}(G)} \int_G dg \langle \Phi | \hat{g} e^{-\hat{H}/T} | \Phi \rangle \\ &= \frac{1}{\text{vol}(G)} \int_G dg \text{Tr}_{\mathcal{H}_{\text{ext}}} \left( \hat{g} e^{-\hat{H}/T} \right). \end{aligned} \quad (\text{A6})$$

TABLE XXVIII. The results for the internal energy of the two-matrix bosonic model with gauge group SU(2) and  $\lambda = 1.0$ .  $N_{\text{cfigs}}$  is the number of configurations used in the analysis and  $N_{\text{drop}}$  is the number of molecular dynamics time units (MDTUs) between successive configurations.  $\tau$  is the integrated autocorrelation time (in units of  $N_{\text{drop}}$ ) and it is always close to unity, indicating that there are no strong autocorrelations.

$T$	$n_t$	$E$	$N_{\text{cfigs}}$	$N_{\text{drop}}$	$\tau$
0.400	16	3.070(31)	1980	50	1.050
0.300	16	2.973(26)	1980	50	1.010
0.250	16	2.880(23)	1980	50	0.980
0.200	16	2.799(20)	1980	50	0.920
0.150	16	2.615(17)	1980	50	1.090
0.100	16	2.359(13)	1980	50	0.760
0.050	16	1.9534(87)	1980	50	0.950
0.400	24	3.169(30)	1980	50	1.180
0.300	24	3.103(26)	1980	50	0.880
0.250	24	3.017(22)	1980	51	0.970
0.200	24	2.927(20)	1981	51	0.890
0.150	24	2.819(18)	1980	51	1.050
0.100	24	2.592(14)	1980	50	1.000
0.050	24	2.2188(93)	1980	50	1.380
0.400	32	3.273(30)	1980	50	1.090
0.300	32	3.160(26)	1980	50	1.150
0.250	32	3.111(24)	1980	50	1.040
0.200	32	3.059(21)	1980	50	1.050
0.150	32	2.918(17)	1981	51	0.980
0.100	32	2.738(14)	1981	51	0.970
0.050	32	2.3711(94)	1980	50	1.100
0.400	48	3.374(32)	1980	50	1.000
0.300	48	3.281(27)	1980	50	0.970
0.250	48	3.209(23)	1980	50	1.020
0.200	48	3.194(21)	1980	50	1.000
0.150	48	3.081(18)	1980	50	0.980
0.100	48	2.910(14)	1980	51	0.970
0.050	48	2.5936(95)	1980	50	0.870
0.400	64	3.393(32)	1980	50	0.980
0.300	64	3.354(26)	1980	50	0.920
0.250	64	3.279(24)	1980	50	1.160
0.200	64	3.262(21)	1980	50	0.950
0.150	64	3.163(18)	1980	50	0.880
0.100	64	3.040(15)	1980	50	1.000
0.050	64	2.7500(99)	1980	50	0.950
0.400	96	3.430(22)	3980	50	1.030
0.300	96	3.341(27)	1980	50	1.080
0.250	96	3.373(24)	1980	50	1.040
0.200	96	3.339(22)	1980	50	1.380
0.150	96	3.288(19)	1980	50	1.250
0.100	96	3.179(15)	1980	50	0.870
0.050	96	2.941(10)	1981	51	0.980
0.300	128	3.420(27)	1980	50	1.070
0.250	128	3.418(24)	1980	50	0.990
0.200	128	3.367(22)	1980	50	1.270
0.150	128	3.334(18)	1980	50	0.990
0.100	128	3.237(15)	1980	50	0.950
0.050	128	3.029(10)	1980	50	0.980
0.300	192	3.461(27)	1980	50	1.390
0.250	192	3.417(24)	1980	50	1.090

Continued

TABLE XXVIII. Continued

$T$	$n_t$	$E$	$N_{\text{cfigs}}$	$N_{\text{drop}}$	$\tau$
0.200	192	3.437(22)	1980	50	1.130
0.150	192	3.378(18)	1980	50	1.110
0.100	192	3.332(15)	1980	50	0.940
0.050	192	3.176(11)	1980	50	0.940

Here, the sum over  $\Phi$  is taken over the extended Hilbert space.

## 2. Relation between canonical quantization and path-integral formalism

In the canonical partition function defined on the extended Hilbert space, the group element  $g$  is interpreted as the Polyakov loop. It can be seen as follows. The expression given in Eq. (A2) can be rewritten as

$$\begin{aligned}
Z(T) &= \frac{1}{[\text{vol}(G)]^K} \int \left( \prod_{k=1}^K dU_{(k)} \right) \text{Tr}_{\mathcal{H}_{\text{ext}}} \\
&\quad \times \left( \hat{U}_{(K)} e^{-\frac{H(\hat{P}, \hat{X})}{TK}} \hat{U}_{(K-1)}^{-1} \hat{U}_{(K-1)} \right) \\
&\quad \times e^{-\frac{H(\hat{P}, \hat{X})}{TK}} \hat{U}_{(K-2)}^{-1} \hat{U}_{(K-2)} \cdots \hat{U}_{(1)}^{-1} \hat{U}_{(1)} e^{-\frac{H(\hat{P}, \hat{X})}{TK}} \Big) \\
&= \frac{1}{[\text{vol}(G)]^K} \int \left( \prod_{k=1}^K dU_{(k)} \right) \int \left( \prod_{k=1}^K dX_{(k)} \right) \\
&\quad \langle X_{(K)} | \hat{U}_{(K)} e^{-H(\hat{P}, \hat{X})/TK} \hat{U}_{(K-1)}^{-1} | X_{(K-1)} \rangle \\
&\quad \times \langle X_{(K-1)} | \hat{U}_{(K-1)} e^{-H(\hat{P}, \hat{X})/TK} \hat{U}_{(K-2)}^{-1} | X_{(K-2)} \rangle \\
&\quad \times \cdots \times \langle X_{(1)} | \hat{U}_{(1)} e^{-H(\hat{P}, \hat{X})/TK} | X_{(K)} \rangle. \quad (\text{A7})
\end{aligned}$$

At the first line, we simply insert  $\hat{U}^{-1} \hat{U} = \hat{1}$  in many places.  $\hat{U}_{(K)}$  corresponds to  $\hat{g}$ . At the next line, we insert  $\int dX_{(k)} |X_{(k)}\rangle \langle X_{(k)}| = \hat{1}$  in many places:

$$\begin{aligned}
&\langle X_{(k)} | \hat{U}_{(k)} e^{-H(\hat{P}, \hat{X})/TK} \hat{U}_{(k-1)}^{-1} | X_{(k-1)} \rangle \\
&= \langle U_{(k)} X_{(k)} U_{(k)}^{-1} | e^{-H(\hat{P}, \hat{X})/TK} | U_{(k-1)} X_{(k-1)} U_{(k-1)}^{-1} \rangle \\
&= \int dP \langle U_{(k)} X_{(k)} U_{(k)}^{-1} | e^{-H(\hat{P}, \hat{X})/TK} | P \rangle \\
&\quad \times \langle P | U_{(k-1)} X_{(k-1)} U_{(k-1)}^{-1} \rangle \\
&= \int dP e^{i\text{Tr}[P(U_{(k)} X_{(k)} U_{(k)}^{-1} - U_{(k-1)} X_{(k-1)} U_{(k-1)}^{-1})]} \\
&\quad \times e^{-H(P, U_{(k)} X_{(k)} U_{(k)}^{-1})/TK} \\
&= e^{-KT\text{Tr}[(U_{(k)} X_{(k)} U_{(k)}^{-1} - U_{(k-1)} X_{(k-1)} U_{(k-1)}^{-1})^2]} \\
&\quad \times e^{-V(U_{(k)} X_{(k)} U_{(k)}^{-1})/TK}
\end{aligned}$$

TABLE XXIX. The results for the internal energy of the two-matrix bosonic model with gauge group SU(2) and  $\lambda = 0.5$ .  $N_{\text{cfgs}}$  is the number of configurations used in the analysis and  $N_{\text{drop}}$  is the number of molecular dynamics time units (MDTUs) between successive configurations.  $\tau$  is the integrated autocorrelation time (in units of  $N_{\text{drop}}$ ) and it is always close to unity, indicating that there are no strong autocorrelations.

$T$	$n_t$	$E$	$N_{\text{cfgs}}$	$N_{\text{drop}}$	$\tau$
0.400	16	2.941(28)	1980	50	1.000
0.300	16	2.805(24)	1980	50	0.870
0.250	16	2.750(21)	1980	50	1.000
0.200	16	2.645(19)	1980	50	0.850
0.150	16	2.527(16)	1980	50	1.070
0.100	16	2.303(13)	1980	50	1.080
0.050	16	1.9063(85)	1980	50	1.010
0.400	24	3.038(28)	1980	50	1.080
0.300	24	2.915(24)	1980	50	0.770
0.250	24	2.876(22)	1980	50	0.990
0.200	24	2.819(20)	1980	50	1.100
0.150	24	2.673(16)	1980	50	1.130
0.100	24	2.527(14)	1980	50	1.160
0.050	24	2.1593(88)	1980	50	1.080
0.400	32	3.058(27)	1980	50	1.150
0.300	32	3.050(24)	1980	50	0.940
0.250	32	2.946(22)	1980	50	0.980
0.200	32	2.927(19)	1980	50	0.930
0.150	32	2.806(17)	1980	50	1.040
0.100	32	2.631(13)	1980	50	1.190
0.050	32	2.3029(89)	1980	50	0.930
0.400	48	3.137(28)	1980	50	1.180
0.300	48	3.087(25)	1980	50	1.150
0.250	48	3.102(23)	1980	50	0.940
0.200	48	3.050(20)	1980	50	0.970
0.150	48	2.916(17)	1980	50	1.100
0.100	48	2.809(14)	1980	50	1.050
0.050	48	2.4901(91)	1980	50	0.990
0.400	64	3.223(29)	1980	50	0.820
0.300	64	3.133(24)	1980	50	1.000
0.250	64	3.105(22)	1980	50	1.110
0.200	64	3.137(20)	1980	50	0.940
0.150	64	3.028(17)	1980	50	1.050
0.100	64	2.899(14)	1980	50	1.060
0.050	64	2.6519(93)	1980	50	1.020
0.400	96	3.199(28)	1980	50	1.130
0.300	96	3.195(24)	1980	49	0.730
0.250	96	3.178(22)	1980	50	1.060
0.200	96	3.134(20)	1980	50	1.090
0.150	96	3.082(17)	1980	50	0.860
0.100	96	3.037(14)	1980	50	1.150
0.050	96	2.7978(97)	1980	50	1.190
0.400	128	3.252(29)	1980	50	1.040
0.300	128	3.204(24)	1980	50	0.970
0.250	128	3.141(22)	1980	49	0.710
0.200	128	3.164(20)	1980	50	1.100
0.150	128	3.142(18)	1980	50	0.940
0.100	128	3.077(14)	1980	50	0.900
0.050	128	2.903(10)	1980	50	0.990
0.400	192	3.287(30)	1980	50	1.310
0.300	192	3.231(23)	1980	50	1.200

Continued

TABLE XXIX. Continued

$T$	$n_t$	$E$	$N_{\text{cfgs}}$	$N_{\text{drop}}$	$\tau$
0.250	192	3.196(22)	1980	50	1.070
0.200	192	3.249(21)	1980	50	1.070
0.150	192	3.230(18)	1980	50	0.940
0.100	192	3.150(14)	1980	50	0.790
0.050	192	3.018(10)	1980	50	0.910

TABLE XXX. Systematic fitting for the bosonic SU(3) model with  $\lambda = 0.5$ .

$a_{\text{max}}$	$n_p$	$E$	$\chi^2$ per DOF	$a_{\text{max}}$	$n_p$	$E$	$\chi^2$ per DOF
0.05	1	8.823(33)	1.11	0.30	1	8.6919(87)	2.13
0.05	2	8.822(68)	1.11	0.30	2	8.773(15)	1.30
0.05	3	8.822(68)	1.11	0.30	3	8.809(22)	1.20
0.10	1	8.768(17)	1.12	0.40	1	8.6369(74)	4.49
0.10	2	8.836(38)	0.99	0.40	2	8.767(12)	1.24
0.10	3	8.836(38)	0.99	0.40	3	8.794(20)	1.19
0.20	1	8.722(11)	1.82	0.50	1	8.5880(67)	7.94
0.20	2	8.807(19)	1.24	0.50	2	8.747(10)	1.36
0.20	3	8.809(30)	1.24	0.50	3	8.789(15)	1.14

TABLE XXXI. Systematic fitting for the bosonic SU(3) model with  $\lambda = 1.0$ .

$a_{\text{max}}$	$n_p$	$E$	$\chi^2$ per DOF	$a_{\text{max}}$	$n_p$	$E$	$\chi^2$ per DOF
0.05	1	9.359(35)	0.57	0.30	1	9.2273(86)	2.72
0.05	2	9.375(70)	0.57	0.30	2	9.337(14)	0.99
0.05	3	9.375(70)	0.57	0.30	3	9.367(21)	0.92
0.10	1	9.328(17)	1.00	0.40	1	9.1622(73)	6.30
0.10	2	9.381(38)	0.93	0.40	2	9.326(12)	1.04
0.10	3	9.381(39)	0.93	0.40	3	9.361(19)	0.95
0.20	1	9.281(11)	1.45	0.50	1	9.0787(64)	14.87
0.20	2	9.355(19)	1.01	0.50	2	9.307(10)	1.21
0.20	3	9.359(31)	1.01	0.50	3	9.348(15)	0.99

TABLE XXXII. Systematic fitting for the bosonic SU(3) model with  $\lambda = 2.0$ .

$a_{\text{max}}$	$n_p$	$E$	$\chi^2$ per DOF	$a_{\text{max}}$	$n_p$	$E$	$\chi^2$ per DOF
0.05	1	10.235(38)	1.09	0.30	1	10.0660(93)	4.43
0.05	2	10.265(73)	1.08	0.30	2	10.227(15)	1.34
0.05	3	10.265(73)	1.08	0.30	3	10.264(23)	1.25
0.10	1	10.247(19)	1.10	0.40	1	9.9657(80)	11.43
0.10	2	10.236(41)	1.09	0.40	2	10.213(13)	1.31
0.10	3	10.237(42)	1.09	0.40	3	10.256(21)	1.19
0.20	1	10.142(12)	2.21	0.50	1	9.8484(70)	25.52
0.20	2	10.251(21)	1.41	0.50	2	10.176(11)	1.78
0.20	3	10.288(33)	1.36	0.50	3	10.249(16)	1.21

TABLE XXXIII. Systematic fitting for the bosonic SU(2) model with  $\lambda = 1.0$ .

$a_{\max}$	$n_p$	$E$	$\chi^2$ per DOF	$a_{\max}$	$n_p$	$E$	$\chi^2$ per DOF
0.05	1	3.519(31)	0.57	0.30	1	3.4579(58)	1.80
0.05	2	3.526(77)	0.57	0.30	2	3.5159(98)	0.58
0.05	3	3.526(77)	0.57	0.30	3	3.516(16)	0.58
0.10	1	3.507(13)	0.55	0.40	1	3.4351(50)	3.17
0.10	2	3.497(33)	0.55	0.40	2	3.5074(84)	0.74
0.10	3	3.498(34)	0.55	0.40	3	3.523(14)	0.70
0.20	1	3.4854(72)	0.64	0.50	1	3.4080(44)	5.46
0.20	2	3.514(14)	0.50	0.50	2	3.4923(71)	0.95
0.20	3	3.513(26)	0.50	0.50	3	3.522(11)	0.68

$$\begin{aligned} &\simeq e^{-L[D_t(U_{(k)}X_{(k)}U_{(k)}^{-1}), (U_{(k)}X_{(k)}U_{(k)}^{-1})]/(TK)} \\ &= e^{-L[D_tX_{(k)}-X_{(k)}]/(TK)}. \end{aligned} \quad (\text{A8})$$

Here, we use

$$U_{(k-1)}^{-1}U_{(k)} \equiv e^{iA_{(k)}/(KT)} \quad (\text{A9})$$

and

$$X_{(k)} - (U_{(k-1)}U_{(k)}^{-1})^{-1}X_{(k-1)}(U_{(k-1)}U_{(k)}^{-1}) \simeq \frac{D_tX_{(k)}}{KT}. \quad (\text{A10})$$

$L[D_tX, X]$  is the Lagrangian with the Euclidean signature. By taking  $K \rightarrow \infty$  limit, we obtain

$$Z(T) = \int [dA][dX] e^{-\int dt L[D_tX, X]}. \quad (\text{A11})$$

The temporal direction is compactified with the circumference  $1/T$  and the periodic boundary condition is imposed on  $X_I$  and  $A_t$ . Therefore,

$$\begin{aligned} g &= U_K \\ &= U_K(U_{K-1}^{-1}U_{K-1})(U_{K-2}^{-1}U_{K-2}) \cdots (U_1^{-1}U_1) \end{aligned}$$

TABLE XXXIV. Systematic fitting for the bosonic SU(2) model with  $\lambda = 2.0$ .

$a_{\max}$	$n_p$	$E$	$\chi^2$ per DOF	$a_{\max}$	$n_p$	$E$	$\chi^2$ per DOF
0.05	1	3.836(28)	0.68	0.30	1	3.7745(62)	2.25
0.05	2	3.775(62)	0.59	0.30	2	3.831(10)	1.19
0.05	3	3.775(62)	0.59	0.30	3	3.864(16)	1.03
0.10	1	3.835(13)	1.02	0.40	1	3.7414(54)	4.63
0.10	2	3.847(30)	1.01	0.40	2	3.8341(89)	1.12
0.10	3	3.845(31)	1.01	0.40	3	3.848(14)	1.09
0.20	1	3.8028(77)	1.55	0.50	1	3.7056(47)	7.96
0.20	2	3.854(14)	1.10	0.50	2	3.8147(76)	1.43
0.20	3	3.849(25)	1.10	0.50	3	3.851(11)	1.07

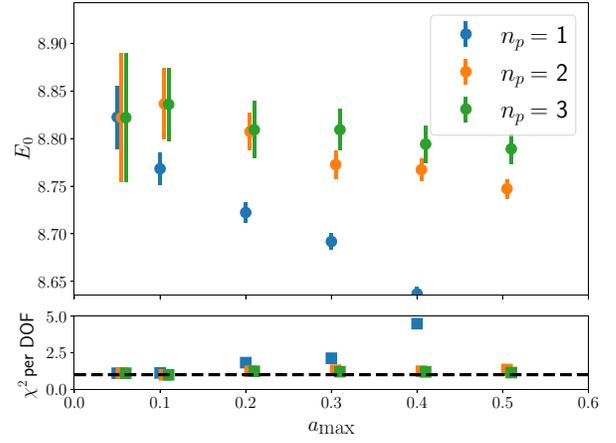


FIG. 25. The bosonic SU(3) model with  $\lambda = 0.5$ . The results of fitting different data portions with polynomials of different order  $n_p$ . The lower panel shows the reduced  $\chi^2$ , which becomes very large for the low-order polynomials when larger lattice spacings are included.

$$\begin{aligned} &= (U_K U_{K-1}^{-1})(U_{K-1} U_{K-2}^{-1}) \cdots (U_2 U_1^{-1}) U_1 \\ &= P e^{i \int_0^{1/T} dt A_t} \end{aligned} \quad (\text{A12})$$

is the Polyakov loop. (Here,  $P$  denotes the path-ordered product along the time direction.)

Note that unless the time direction is noncompact or a special boundary condition is taken, the condition  $A_t = 0$  cannot be imposed in the literal sense. In the case of the canonical partition function, we cannot eliminate the degrees of freedom of the Polyakov loop. The integration over the Polyakov loop leads to the projector  $1/\text{vol}(G) \int dg \hat{g}$  in the operator formalism.

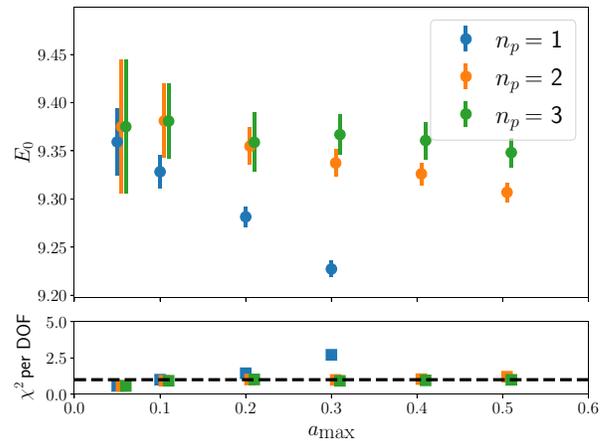


FIG. 26. The bosonic SU(3) model with  $\lambda = 1.0$ . The results of fitting different data portions with polynomials of different order  $n_p$ . The lower panel shows the reduced  $\chi^2$ , which becomes very large for the low-order polynomials when larger lattice spacings are included.

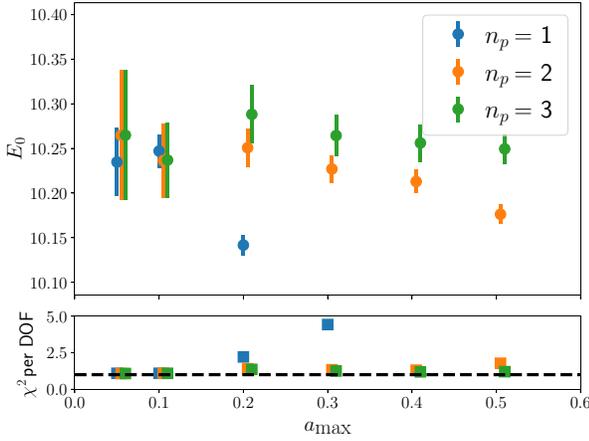


FIG. 27. The bosonic SU(3) model with  $\lambda = 2.0$ . The results of fitting different data portions with polynomials of different order  $n_p$ . The lower panel shows the reduced  $\chi^2$ , which becomes very large for the low-order polynomials when larger lattice spacings are included.

## APPENDIX B: TABLES OF RESULTS OF HAMILTONIAN TRUNCATION

### 1. SU(2) bosonic matrix model

We tabulate the values of the ground-state energy for the SU(2) bosonic model at four coupling constants  $\lambda$  in Table XIX. The values of the singlet constraint violation corresponding to the same couplings are tabulated in Table XX. For the SU(2) bosonic model we study up to cutoff  $\Lambda = 14$ , corresponding to a Hamiltonian matrix with 7529536 rows.

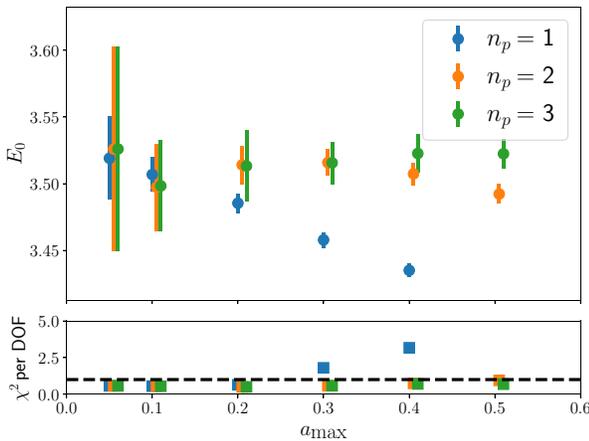


FIG. 28. The bosonic SU(2) model with  $\lambda = 1.0$ . The results of fitting different data portions with polynomials of different order  $n_p$ . The lower panel shows the reduced  $\chi^2$ , which becomes very large for the low-order polynomials when larger lattice spacings are included.

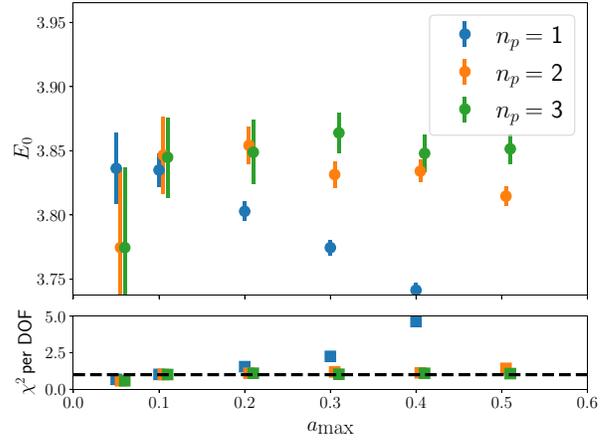


FIG. 29. The bosonic SU(2) model with  $\lambda = 2.0$ . The results of fitting different data portions with polynomials of different order  $n_p$ . The lower panel shows the reduced  $\chi^2$ , which becomes very large for the low-order polynomials when larger lattice spacings are included.

### 2. SU(2) minimal BMN

We tabulate the values of the energy measured on the ground-state of the modified Hamiltonian for the SU(2) minimal-BMN model at four coupling constants  $\lambda$  in Table XXI. The values of the singlet constraint violation corresponding to the same couplings are tabulated in Table XXII. In Table XXIII we also report the angular-momentum-constraint violation for the same states. For the SU(2) minimal-BMN model we study up to cutoff  $\Lambda = 11$ , corresponding to a Hamiltonian matrix with 14172488 rows.

## APPENDIX C: TABLES OF RESULTS OF LATTICE MONTE CARLO SIMULATIONS

### 1. Tables for the results of the ground-state energy

An example of raw results from the analysis of the Markov chain Monte Carlo histories is shown in Table XXIV. For each parameter set of  $T$  and  $n_t$ , we thermalize the MCMC chain and remove a burn-in subset of trajectories (in units of MDTU) of 1000 MDTU at the beginning of each chain. We save the energy observable every  $N_{\text{drop}}$  MDTU and we collect the observables on a total of  $N_{\text{cfgs}}$  configurations. We also measure the integrated autocorrelation time  $\tau$  of the energy observable using the Madras-Sokal windowing algorithm. Results are tabulated in Tables XXIV–XXVI for SU(3) and in Tables XXVII–XXIX for SU(2).

### 2. Tables and figures for the results of the systematic fitting of the continuum limit

The results of the systematic fitting are tabulated in Tables XXX–XXXII for SU(3) and in Tables XXXIII and

XXXIV for SU(2) (in addition to Table XII in the main text). The results are plotted in Figs. 25–27 for SU(3) and in Figs. 28 and 29 for SU(2) (in addition to Fig. 23 in the main text).

- 
- [1] J. M. Maldacena, The large  $N$  limit of superconformal field theories and supergravity, *Adv. Theor. Math. Phys.* **2**, 231 (1998).
- [2] N. Itzhaki, J. M. Maldacena, J. Sonnenschein, and S. Yankielowicz, Supergravity and the large  $N$  limit of theories with sixteen supercharges, *Phys. Rev. D* **58**, 046004 (1998).
- [3] T. Banks, W. Fischler, S. Shenker, and L. Susskind, M theory as a matrix model: A conjecture, *Phys. Rev. D* **55**, 5112 (1997).
- [4] B. de Wit, J. Hoppe, and H. Nicolai, On the quantum mechanics of supermembranes, *Nucl. Phys. B* **305**, 545 (1988).
- [5] D. E. Berenstein, J. M. Maldacena, and H. S. Nastase, Strings in flat space and pp waves from  $N = 4$  super Yang-Mills, *JHEP* **04**, 013.
- [6] E. Berkowitz, E. Rinaldi, M. Hanada, G. Ishiki, S. Shimasaki, and P. Vranas, Precision lattice test of the gauge/gravity duality at large- $N$ , *Phys. Rev. D* **94**, 094501 (2016).
- [7] I. Buluta and F. Nori, Quantum simulators, *Science* **326**, 108 (2009).
- [8] I. M. Georgescu, S. Ashhab, and F. Nori, Quantum simulation, *Rev. Mod. Phys.* **86**, 153 (2014).
- [9] M. C. Bañuls, R. Blatt, J. Catani, A. Celi, J. I. Cirac, M. Dalmonte, L. Fallani, K. Jansen, M. Lewenstein, S. Montangero, and C. A. Muschik, Simulating lattice gauge theories within quantum technologies, *Eur. Phys. J. D* **74**, 165 (2020).
- [10] I. Buluta, S. Ashhab, and F. Nori, Natural and artificial atoms for quantum computation, *Rep. Prog. Phys.* **74**, 104401 (2011).
- [11] Z. Davoudi, M. Hafezi, C. Monroe, G. Pagano, A. Seif, and A. Shaw, Towards analog quantum simulations of lattice gauge theories with trapped ions, *Phys. Rev. Res.* **2**, 023015 (2020).
- [12] T. Byrnes and Y. Yamamoto, Simulating lattice gauge theories on a quantum computer, *Phys. Rev. A* **73**, 022328 (2006).
- [13] M. S. Albergo, G. Kanwar, S. Racanière, D. J. Rezende, J. M. Urban, D. Boyda, K. Cranmer, D. C. Hackett, and P. E. Shanahan, Flow-based sampling for fermionic lattice field theories (2021), [ArXiv:2106.05934](https://arxiv.org/abs/2106.05934).
- [14] X. Han, S. A. Hartnoll, and J. Kruthoff, Bootstrapping Matrix Quantum Mechanics, *Phys. Rev. Lett.* **125**, 041601 (2020).
- [15] D. Berenstein and G. Hulsey, Bootstrapping Simple QM Systems (2021), [ArXiv:2108.08757](https://arxiv.org/abs/2108.08757).
- [16] R. d. M. Koch, A. Jevicki, X. Liu, K. Mathaba, and J. a. P. Rodrigues, Large  $N$  optimization for multi-matrix systems (2021), [ArXiv:2108.08803](https://arxiv.org/abs/2108.08803).
- [17] Monte Carlo methods provide us a way to generate a probability distribution. If the path-integral weight  $e^{-S}$  is real and non-negative, the lattice configurations can be generated in such a way that the probability distribution is proportional to  $e^{-S}$ . If the path-integral weight is not real and non-negative, it cannot be a “probability distribution,” and hence Monte Carlo methods cannot be applied. This is a manifestation of the so-called sign problem. With the Minkowski signature, the path-integral weight is  $e^{iS}$ , which is purely imaginary.
- [18] M. Hanada, Bulk geometry in gauge/gravity duality and color degrees of freedom, *Phys. Rev. D* **103**, 106007 (2021).
- [19] A. R. Brown, H. Gharibyan, S. Leichenauer, H. W. Lin, S. Nezami, G. Salton, L. Susskind, B. Swingle, and M. Walter, Quantum Gravity in the Lab: Teleportation by Size and Traversable Wormholes (2019), [ArXiv:1911.06314](https://arxiv.org/abs/1911.06314).
- [20] A. Peruzzo, J. McClean, P. Shadbolt, M.-H. Yung, X.-Q. Zhou, P. J. Love, A. Aspuru-Guzik, and J. L. O’Brien, A variational eigenvalue solver on a photonic quantum processor, *Nat. Commun.* **5**, eid 4213 (2014).
- [21] N. Moll, P. Barkoutsos, L. S. Bishop, J. M. Chow, A. Cross, D. J. Egger, S. Filipp, A. Fuhrer, J. M. Gambetta, M. Ganzhorn, A. Kandala, A. Mezzacapo, P. Müller, W. Riess, G. Salis, J. Smolin, I. Tavernelli, and K. Temme, Quantum optimization using variational algorithms on near-term quantum devices, *Quantum Sci. Technol.* **3**, 030503 (2018).
- [22] S. Endo, Z. Cai, S. C. Benjamin, and X. Yuan, Hybrid quantum-classical algorithms and quantum error mitigation, *J. Phys. Soc. Jpn* **90**, 032001 (2021).
- [23] J. Preskill, Quantum computing in the NISQ era and beyond, *Quantum* **2**, 79 (2018).
- [24] G. Carleo, I. Cirac, K. Cranmer, L. Daudet, M. Schuld, N. Tishby, L. Vogt-Maranto, and L. Zdeborová, Machine learning and the physical sciences, *Rev. Mod. Phys.* **91**, 045002 (2019).
- [25] X. Han and S. A. Hartnoll, Deep Quantum Geometry of Matrices, *Phys. Rev. X* **10**, 011069 (2020).
- [26] X. Gao and L.-M. Duan, Efficient representation of quantum many-body states with deep neural networks, *Nat. Commun.* **8**, 662 (2017).
- [27] G. Carleo, Y. Nomura, and M. Imada, Constructing exact representations of quantum many-body systems with deep neural networks, *Nat. Commun.* **9**, 00 (2018).
- [28] Z. Jia, B. Yi, R. Zhai, Y. Wu, G. Guo, and G. Guo, Quantum neural network states: A brief review of methods and applications, *Adv. Quantum Technol.* **2**, 1800077 (2019).
- [29] J. Hermann, Z. Schätzle, and F. Noé, Deep-neural-network solution of the electronic Schrödinger equation, *Nat. Chem.* **12**, 891 (2020).
- [30] K. Choo, A. Mezzacapo, and G. Carleo, Fermionic neural-network states for ab-initio electronic structure, *Nat. Commun.* **11**, 2368 (2020).
- [31] N. Yoshioka, W. Mizukami, and F. Nori, Solving quasi-particle band spectra of real solids using neural-network quantum states, *Commun. Phys.* **4**, 106 (2021).
- [32] Y. Nomura, N. Yoshioka, and F. Nori, Purifying Deep Boltzmann Machines for Thermal Quantum States, *Phys. Rev. Lett.* **127**, 060601 (2021).
- [33] The codes used to generate the data and make the figures are available at <https://erinaldi.github.io/mm-qc-dl-supplemental>.

- [34] N. Kim and J.-H. Park, Massive super Yang-Mills quantum mechanics: Classification and the relation to supermembrane, *Nucl. Phys. B* **759**, 249 (2006).
- [35] A. Yuvaraj, The D0-brane quantum mechanics is the dimensional reduction of 10D  $\mathcal{N} = 1$  SYM, which has 16 real supercharges. By dimensionally reducing 6D, 4D, and 3D  $\mathcal{N} = 1$  SYM, similar matrix quantum mechanics with eight, four, and two real supercharges can be constructed. They admit the BMN-like one-parameter deformations.
- [36] For simplicity, we assume that  $\mu \geq 0$ . This assumption is not essential.
- [37] E. Witten and D. I. Olive, Supersymmetry algebras that include topological charges, *Phys. Lett. B* **78**, 97 (1978).
- [38] T. Cubitt, D. Perez-Garcia, and M. M. Wolf, Undecidability of the spectral gap (short version), *Nature* **528**, 207 (2015).
- [39] See <https://qutip.org/> for the official web page of QuTiP. For low-dimensional Hilbert spaces, we cross-check the results using independent codes written in PYTHON and FORTRAN.
- [40] J. R. Johansson, P. D. Nation, and F. Nori, QuTIP: An open-source PYTHON framework for the dynamics of open quantum systems, *Comput. Phys. Commun.* **183**, 1760 (2012).
- [41] J. R. Johansson, P. D. Nation, and F. Nori, QuTIP 2: A PYTHON framework for the dynamics of open quantum systems, *Comput. Phys. Commun.* **184**, 1234 (2013).
- [42] H. Gharibyan, M. Hanada, M. Honda, and J. Liu, Toward simulating superstring/M-theory on a quantum computer, *JHEP* **07**, 140, [ArXiv:2011.06573](https://arxiv.org/abs/2011.06573).
- [43] Another common method uses the coordinate basis [42,86–88] or the conformal truncation [89,90].
- [44] Specifically, we take the average over all gauge transformation  $1/\sqrt{\mathcal{N}_\phi} \int dU(\hat{U}|\phi)$ , where the operator  $\hat{U}$  acts as the  $SU(N)$  transformation and  $\mathcal{N}_\phi$  is the normalization factor.
- [45] E. Zohar and B. Reznik, Confinement and Lattice QED Electric Flux-Tubes Simulated with Ultracold Atoms, *Phys. Rev. Lett.* **107**, 275301 (2011).
- [46] E. Zohar, J. I. Cirac, and B. Reznik, Simulating Compact Quantum Electrodynamics with Ultracold Atoms: Probing Confinement and Nonperturbative Effects, *Phys. Rev. Lett.* **109**, 125302 (2012).
- [47] J. C. Halimeh, H. Lang, and P. Hauke, Gauge protection in non-Abelian lattice gauge theories (2021), [ArXiv:2106.09032](https://arxiv.org/abs/2106.09032).
- [48] M. Dalmonte and S. Montangero, Lattice gauge theory simulations in the quantum information era, *Contemp. Phys.* **57**, 388 (2016).
- [49] S. V. Mathis, G. Mazzola, and I. Tavernelli, Toward scalable simulations of lattice gauge theories on quantum computers, *Phys. Rev. D* **102**, 094501 (2020).
- [50] ARPACK Software, ARPACK software web site.
- [51] R. B. Lehoucq, D. C. Sorensen, and C. Yang, ARPACK user’s guide: Solution of large scale eigenvalue problems by implicitly restarted Arnoldi methods (1997).
- [52] S. B. Bravyi and A. Y. Kitaev, Fermionic quantum computation, *Ann. Phys. (N. Y.)* **298**, 210 (2002).
- [53] K. Setia, J. D. Whitfield, Bravyi-Kitaev superfast simulation of electronic structure on a quantum computer, *J. Chem. Phys.* **148**, 164104 (2018).
- [54] K. Setia, R. Chen, J. E. Rice, A. Mezzacapo, M. Pistoia, and J. D. Whitfield, Reducing qubit requirements for quantum simulations using molecular point group symmetries, *J. Chem. Theory Comput.* **16**, 6091 (2020).
- [55] P. Woit, *Quantum Theory, Groups and Representations* (The American Chemical Society, Washington, DC, 2017) 1.
- [56] S. Wei, H. Li, and G. Long, A full quantum eigensolver for quantum chemistry simulations (2020), [ArXiv:1908.07927](https://arxiv.org/abs/1908.07927).
- [57] A. Kardashin, A. Pervishko, J. Biamonte, and D. Yudin, Benchmarking variational quantum simulation against an exact solution (2021), [ArXiv:2105.06208](https://arxiv.org/abs/2105.06208).
- [58] M. Treinish, *et al.*, Qiskit: QISKIT 0.29.0 (2021).
- [59] G. Mazzola, S. V. Mathis, G. Mazzola, and I. Tavernelli, Gauge invariant quantum circuits for  $u(1)$  and Yang-Mills lattice gauge theories (2021), [ArXiv:2105.05870](https://arxiv.org/abs/2105.05870).
- [60] J. R. Johansson, *Numerical PYTHON: A Practical Techniques Approach for Industry* (Apress, 2015) 1.
- [61] J. Gacon, C. Zoufal, G. Carleo, and S. Woerner, Simultaneous perturbation stochastic approximation of the quantum Fisher information, *Quantum* **5**, 567 (2021).
- [62] C. Kane and M. McGuigan, in *2018 New York Scientific Data Summit (NYS DS)* (2018), p. 1.
- [63] H. R. Grimsley, S. E. Economou, E. Barnes, and N. J. Mayhall, An adaptive variational algorithm for exact molecular simulations on a quantum computer, *Nat. Commun.* **10**, 00 (2019).
- [64] G. Verdon, M. Broughton, J. R. McClean, K. J. Sung, R. Babbush, Z. Jiang, H. Neven, and M. Mohseni, Learning to learn with quantum neural networks via classical neural networks (2019), [ArXiv:1907.05415](https://arxiv.org/abs/1907.05415).
- [65] F. Sauvage, S. Sim, A. A. Kunitsa, W. A. Simon, M. Mauri, and A. Perdomo-Ortiz, FLIP: A flexible initializer for arbitrarily-sized parametrized quantum circuits (2021), [ArXiv:2103.08572](https://arxiv.org/abs/2103.08572).
- [66] C.-W. Huang, D. Krueger, A. Lacoste, and A. Courville, in *International Conference on Machine Learning* (PMLR, 2018), p. 2078.
- [67] N. D. Cao, I. Titov, and W. Aziz, Block neural autoregressive flow (2019), [ArXiv:1904.04676](https://arxiv.org/abs/1904.04676).
- [68] D. Luo, G. Carleo, B. K. Clark, and J. Stokes, Gauge equivariant neural networks for quantum lattice gauge theories (2020), [ArXiv:2012.05232](https://arxiv.org/abs/2012.05232).
- [69] D. Luo, Z. Chen, K. Hu, Z. Zhao, V. M. Hur, and B. K. Clark, Gauge Invariant Autoregressive Neural Networks for Quantum Lattice Models (2021), [ArXiv:2101.07243](https://arxiv.org/abs/2101.07243).
- [70] D. Boyda, G. Kanwar, S. Racanière, D. J. Rezende, M. S. Albergo, K. Cranmer, D. C. Hackett, and P. E. Shanahan, Sampling using  $SU(n)$  gauge equivariant flows, *Phys. Rev. D* **103**, 074504 (2021).
- [71] M. Finzi, M. Welling, and A. G. Wilson, A practical method for constructing equivariant multilayer perceptrons for arbitrary matrix groups (2021), [ArXiv:2104.09459](https://arxiv.org/abs/2104.09459).
- [72] M. Hanada, Markov Chain Monte Carlo for Dummies (2018), [ArXiv:1808.08490](https://arxiv.org/abs/1808.08490).
- [73] E. Berkowitz, M. Hanada, E. Rinaldi, and P. Vranas, Gauged and Ungauged: A Nonperturbative Test, *JHEP* **06**, 124, [ArXiv:1802.02985](https://arxiv.org/abs/1802.02985).

- [74] S. Duane, A. D. Kennedy, B. J. Pendleton, and D. Roweth, Hybrid Monte Carlo, *Phys. Lett. B* **195**, 216 (1987).
- [75] P. Lepage and C. Gohlke, GPLEPAGE: LSQFIT version 11.8 (2021).
- [76] The training process of the deep neural network might be affected by slow down if some hyperparameters are not chosen correctly, such as the learning rate or the optimizer.
- [77] The molecular dynamics time units might require more steps as a function of the temperature or lattice size, therefore increasing the computational time to obtain uncorrelated configurations along a Markov chain.
- [78] I. Danshita, M. Hanada, and M. Tezuka, Creating and probing the Sachdev-Ye-Kitaev model with ultracold gases: Towards experimental studies of quantum gravity, *PTEP* **2017**, 083I01 (2017). [ArXiv:1606.02454](https://arxiv.org/abs/1606.02454)
- [79] S. Nezami, H. W. Lin, A. R. Brown, H. Gharibyan, S. Leichenauer, G. Salton, L. Susskind, B. Swingle, and M. Walter, Quantum Gravity in the Lab: Teleportation by Size and Traversable Wormholes, Part II (2021), [ArXiv:2102.01064](https://arxiv.org/abs/2102.01064).
- [80] L. García-Álvarez, I. L. Egusquiza, L. Lamata, A. del Campo, J. Sonner, and E. Solano, Digital Quantum Simulation of Minimal AdS/CFT, *Phys. Rev. Lett.* **119**, 040501 (2017).
- [81] R. Babbush, D. W. Berry, and H. Neven, Quantum simulation of the Sachdev-Ye-Kitaev model by asymmetric qubitization, *Phys. Rev. A* **99**, 040301 (2019).
- [82] A. Almheiri, X. Dong, and D. Harlow, Bulk Locality and Quantum Error Correction in AdS/CFT, *JHEP* **04**, 163.
- [83] F. Pastawski, B. Yoshida, D. Harlow, and J. Preskill, Holographic quantum error-correcting codes: Toy models for the bulk/boundary correspondence, *JHEP* **06**, 149.
- [84] P. Faist, S. Nezami, V. V. Albert, G. Salton, F. Pastawski, P. Hayden, and J. Preskill, Continuous Symmetries and Approximate Quantum Error Correction, *Phys. Rev. X* **10**, 041018 (2020).
- [85] M. Hanada, H. Shimada, and N. Wintergerst, Color confinement and Bose-Einstein condensation, *JHEP* **08**, 039.
- [86] S. P. Jordan, K. S. Lee, and J. Preskill, Quantum algorithms for quantum field theories, *Science* **336**, 1130 (2012).
- [87] S. P. Jordan, K. S. Lee, and J. Preskill, Quantum computation of scattering in scalar quantum field theories, *Quant. Inf. Comput.* **14**, 1014 (2014).
- [88] N. Klco and M. J. Savage, Digitization of scalar fields for quantum computing, *Phys. Rev. A* **99**, 052335 (2019).
- [89] N. Anand, A. L. Fitzpatrick, E. Katz, Z. U. Khandker, M. T. Walters, and Y. Xin, Introduction to Lightcone Conformal Truncation: QFT Dynamics from CFT Data (2020).
- [90] J. Liu and Y. Xin, Quantum simulation of quantum field theories as quantum chemistry, *JHEP* **12**, 011.