

## Neural network solutions of bosonic quantum systems in one dimension

Paulo F. Bedaque<sup>1</sup>,\* Hersh Kumar<sup>1</sup>,† and Andy Sheng<sup>1</sup>,‡

*Department of Physics, University of Maryland, College Park, Maryland 20742, USA*



(Received 28 September 2023; accepted 16 February 2024; published 27 March 2024)

Neural networks have been proposed as efficient numerical wave function *Ansätze*, which can be used to variationally search a wide range of functional forms for ground-state solutions. These neural network methods are also advantageous in that more variational parameters and system degrees of freedom can be easily added. We benchmark the methodology by using neural networks to study several different integrable bosonic quantum systems in one dimension and compare our results to the exact solutions. While testing the scalability of the procedure to systems with many particles, we also introduce using symmetric function inputs to the neural network to enforce exchange symmetries of indistinguishable particles.

DOI: [10.1103/PhysRevC.109.034004](https://doi.org/10.1103/PhysRevC.109.034004)

### I. INTRODUCTION

Solving many-body quantum mechanical systems involves obtaining solutions to a high-dimensional partial differential equation or diagonalizing exponentially large Hamiltonians, both of which pose significant challenges. Thus, analytically solved models are few and far between [1–7].

Various methods have been developed to obtain approximate solutions to complex quantum systems. For one such method called variational Monte Carlo, a variational *Ansatz* for the ground-state wave function is assumed and an estimate for the ground-state energy is obtained by minimizing the energy of the system over the parameters of the chosen approximate wave function form. The use of trial wave functions has been used to study a broad range of many-body systems, from Fermi and Bose gases to quantum chemistry [8–17]. In particular, the Slater-Jastrow [18] wave function *Ansatz* has been employed widely [19–21] because, by construction, it captures correlations between interacting particles and incorporates the correct intuitive functional form of the physical system. Tensor networks and density matrix renormalization group methods [22] were shown to efficiently encode low-lying eigenstates of local, gapped Hamiltonian spin systems using only a polynomial number of parameters. This eliminates the need to search through an exponentially large Hilbert space.

More recently, established machine learning techniques such as neural networks have found their way into quantum physics [23,24]. In the context of studying ground states of quantum systems, neural networks conveniently provide a numerical wave function *Ansatz* for which one can vary to minimize the system energy cost function. The minimization can be done efficiently using established machine learning

methods of back propagation and gradient descent. Another appeal of using neural networks is that they were shown to be universal function approximators [25,26] and thus can be used to describe any general wave function. As a consequence, the number of variational parameters is not intrinsically tied to the function form (i.e., Jastrow) of the *Ansatz* and can be increased easily.

Carleo and Troyer [27] showed that a type of neural network called restricted Boltzmann machines were sufficient to describe ground-state wave functions of spin systems with a reasonable number of parameters. Later, restricted Boltzmann machines was found to map directly to matrix product state approximations of spin states [28,29]. For particles interacting via continuous-space potentials, several systems of indistinguishable particles have also been studied using neural networks including the Calogero-Sutherland model [30], nuclear models [31–35], and chemical systems [36,37].

Quantum systems involving indistinguishable particles must have ground-state wave functions, which satisfy position exchange (anti)symmetries. This poses a challenge in which the constructed neural network trial wave function should satisfy these constraints. In one spatial dimension, Bose symmetry can be satisfied by feeding in ordered coordinates of the bosons into the neural network [30]; however, this method is not easily generalizable to systems of higher dimensions. Pfau *et al.* [37] developed the FermiNet neural network structure to impose fermionic antisymmetry to study electron-electron and electron-ion interactions in atoms. The FermiNet structure constructs the most general fermionic wave function space for which one can search for the ground state through by taking Slater determinants of permutation-equivariant functions. The permutation-equivariance of such functions are enforced through permutation-equivariant neural network layers.

In this paper, we instead propose the use of carefully selected symmetric inputs to enforce the Bose symmetry of the neural network *Ansatz*. By using symmetric functions as inputs to the neural network, we are able to construct the most general Bose symmetric function as we increase the

\*bedaque@umd.edu

†hekumar@terpmail.umd.edu

‡asheng@umd.edu

number of nodes and layers; this gives us full flexibility in changing the architecture of the network. We benchmark the use of symmetric functions by computing ground-state energies and wave functions of several one-dimensional, exactly solvable quantum systems and comparing both quantitative and qualitative results. Our methodology is not restricted to studying exactly solvable systems and we also explore systems in which no analytical solution is known. We further demonstrate the scalability and versatility of our neural network *Ansatz* by calculating ground states of systems with dozens of particles.

In Sec. II, we outline the various models for which we study and compare numerical results to. In Sec. III, we explicitly construct the Bose symmetric neural network functions and outline our procedure for computing ground states. Section IV discusses our results; we find that our neural network *Ansatz* is able to accurately compute ground-state energies and features of the many-body systems considered over their various phases. Then we conclude with Sec. V by exploring prospects of using symmetric function constructions to write general bosonic and fermionic trial wave functions in a higher number of spatial dimensions.

## II. MODELS

To show the validity of our symmetric input neural network solutions, we apply our neural-network-based methodologies to solve for the ground-state wave functions and energies for several different one-dimensional quantum systems. We consider both a system of cold bosons in a harmonic trap and a system of trapped bosons with short- and long-range interactions. Both quantum systems are inspired by the Lieb-Liniger model [1,2], a homogeneous gas of indistinguishable bosons interacting via a contact  $\delta$ -function potential which was shown to be exactly solvable in one dimension through the use of Bethe *Ansatz* [38–41]. These Lieb-Liniger-inspired models have rich phase structure and have been studied through the lenses of both finding exact solutions and performing numerical analyses [7,10,42–45]. Thus, they provide good test beds for comparison of both quantitative and qualitative ground-state results.

### A. Cold bosons in a harmonic trap

The first Hamiltonian that we consider describes spinless bosons in a harmonic trap. The interactions of the condensate are approximated by a contact potential. Using units where  $\hbar = 1$ , the Hamiltonian is

$$\hat{H} = \sum_{i=1}^N \left( -\frac{1}{2m} \frac{\partial^2}{\partial x_i^2} + \frac{1}{2} m \omega^2 x_i^2 \right) + \sum_{i<j}^N g \delta(x_i - x_j). \quad (1)$$

In general, we take  $m = \omega = 1$ . The interaction is repulsive when  $g > 0$  and attractive when  $g < 0$ . When  $g > 0$ , the interaction strength  $g$  is proportional to the two-boson  $s$ -wave scattering length [42]. Although the addition of the harmonic trap causes the system to no longer be analytically solvable [10] for the general case of many particles, the two-boson case has been solved exactly [10,42].

### B. Trapped bosons with short- and long-range interactions

Recently, another quantum system has been shown to be exactly solvable for a general case of  $N$  bosons [7]. The model consists again of particles in harmonic confinement interacting with a  $\delta$ -function potential of strength  $g$ . Additionally, a long-range linear interaction potential of strength  $\sigma$  is added, which, in one dimension, corresponds to either Coulomb repulsion ( $\sigma < 0$ ) or gravitational attraction ( $\sigma > 0$ ).

$$\hat{H} = \sum_{i=1}^N \left( -\frac{1}{2m} \frac{\partial^2}{\partial x_i^2} + \frac{1}{2} m \omega^2 x_i^2 \right) + \sum_{i<j}^N (g \delta(x_i - x_j) + \sigma |x_i - x_j|). \quad (2)$$

Again, we take  $m = \omega = 1$ . The model was found to be integrable when  $\sigma = -m\omega g/2$  [45].<sup>1</sup> In that case,  $g > 0$  corresponds to both interaction terms being repulsive and  $g < 0$  corresponds to both terms being attractive. The coupling strength is again related to the one-dimensional  $s$ -wave scattering length  $a_s = -2/(mg)$ . The exact expression for the ground-state energy of the Hamiltonian (2) was found to be

$$E_0 = \frac{N\omega}{2} - mg^2 \frac{N(N^2 - 1)}{24} \quad (3)$$

with a corresponding (non-normalized) ground-state wave function

$$\psi_0(x_1, \dots, x_N) = \prod_{i<j} e^{-|x_i - x_j|/a_s} \prod_i e^{-x_i^2/(2a_{\text{ho}}^2)}, \quad (4)$$

where  $a_{\text{ho}} = \sqrt{1/(m\omega)}$  denotes the harmonic oscillator length.

When  $g > 0$ , the ground-state wave function is maximized at  $|x_i - x_j| = Na_{\text{ho}}^2/a_s$  and the bosons tend to spread out with equal spacing about the center of the harmonic trap. When a weak repulsion is turned on, the system forms an incompressible liquid of flat density [46,47]. In the limit of strong repulsion  $g \gg 1$  the probability of overlap between any two particles drops to zero and the system forms a Wigner crystal [48]. In the other regime, when  $g < 0$ , the bosons collapse into the center of the trap. In the limit of  $g \rightarrow -\infty$ , the ground-state wave function of (2) becomes

$$\psi_0(x_1, \dots, x_N) = \prod_{i<j} e^{-|x_i - x_j|/a_s} \quad (5)$$

resembling a McGuire bound-state solution [41] describing a bright soliton [49]. It is surprising that despite the fact that different signs of the interaction strength  $g$  yield such different behaviors, the expression (3) for the ground-state energy of the system implies that the ground-state energy does not depend on the sign of  $g$ .

<sup>1</sup>References [7,45] claim the solution found is for the regime where  $\sigma = -m\omega g$ . However, we analytically find that the provided solution is correct not for  $\sigma = -m\omega g$ , but for  $\sigma = -m\omega g/2$ .

### III. METHODS

Given a Hamiltonian describing a quantum system of  $N$  indistinguishable bosons, we would like to solve the time-independent Schrödinger equation for the ground-state energy and wave function. To do this, we employ the variational principle; any arbitrary wave function  $\psi$  must have an expectation value of the energy at least as large as the ground-state energy.

$$\langle E \rangle_0 \leq \langle E(\theta) \rangle = \frac{\int dX \psi^\dagger(X, \theta) \hat{H} \psi(X, \theta)}{\int dX \psi^\dagger(X, \theta) \psi(X, \theta)}. \quad (6)$$

Here  $X$  stands for the positions of the  $N$  bosons. Typically, one begins with an *Ansatz* for the wave function  $\psi(X, \theta)$ , where the  $\theta$  are its variational parameters and then minimize  $\langle E(\theta) \rangle$  with respect to  $\theta$  to obtain an estimate for the ground state of the quantum system.

#### A. Neural network wave function *Ansätze*

The quality of the estimate for the ground-state energy and wave function depends both on the number of variational parameters one uses as well as the particular functional form of  $\psi$ . Although one could successfully make well-posed guesses as to the correct functional form of the ground-state wave function based on expectations on the underlying physics of the system, we would like to more systematically search the space of potential ground-state wave functions. In this context, the application of neural networks is powerful.

##### 1. Neural networks

We define a neural network with  $L$  layers as a function, which takes  $n_1 = N$  inputs and produces  $n_L = M$  outputs, defined in the following way:

$$I_{i+1} = O_i = f(\mathbf{W}_i \vec{I}_i + \vec{b}_i), \quad 1 \leq i \leq L-1$$

$$O_L = \mathbf{W}_L \vec{I}_L + \vec{b}_L. \quad (7)$$

The inputs  $I_i$  at layer  $i$  are multiplied by an  $n_{i+1} \times n_i$  matrix of weights  $\mathbf{W}_i$ .<sup>2</sup> A bias vector  $\vec{b}_i$  of length  $n_{i+1}$  is added before applying an activation function  $f$  to each resulting vector element. The output  $O_i$  of the  $i$ th layer then becomes the input for the  $(i+1)$ th layer, where the same procedure is repeated with a different set of weights and biases. While in general, a different activation function can be used at each layer, we choose to keep the same function throughout. To ensure that the neural network produces an output whose range is over all reals, an activation function is not used in the final output layer.

Neural networks are useful in the context of writing down a variational *Ansatz* for the wave function because multilayer neural networks have been shown to be universal function approximators [25, 26]. The advantage is that, without needing to change the architecture of the neural network function, we can explore a general space of functions to find the ground-state wave function by varying the parameters  $\theta = \{\mathbf{W}_i, \vec{b}_i\}$ .

<sup>2</sup>For clarity,  $n_1 = N$  and  $n_L = M$ ;  $n_i$  is the number of nodes at layer  $i$ .

The efficiency of the functional expressivity for many-body wave functions, however, is not clear and is a property we seek to explore. We denote our *Ansatz* neural network function as  $\mathcal{A}(I_1, \dots, I_N)$ ; it takes in  $N$  inputs and returns a single output ( $M = 1$ ).

#### 2. Enforcing ground-state symmetries

A neural network *Ansatz* for the ground-state wave function of a system of  $N$  indistinguishable bosons should satisfy the following two constraints. First, the wave function must obey Bose exchange symmetry.

$$\psi(\dots, x_i, \dots, x_j, \dots) = \psi(\dots, x_j, \dots, x_i, \dots), \quad (8)$$

where the  $x_n$  are the positions of the bosons. Thus we must construct our neural network such that it is fully symmetric under exchange of particle positions. To do this, we look for a bijection between the set of particle positions  $\{x_i\}$  and a set of functions of  $\{x_i\}$ , which are symmetric under  $x_i \leftrightarrow x_j$ . The simplest such bijection generates the set of elementary symmetric polynomials

$$e_i = \sum_{1 \leq j_1 < j_2 < \dots < j_i \leq N} x_{j_1} x_{j_2} \dots x_{j_i}. \quad (9)$$

For instance, for  $N = 3$ :

$$e_1 = x_1 + x_2 + x_3,$$

$$e_2 = x_1 x_2 + x_1 x_3 + x_2 x_3, \quad (10)$$

$$e_3 = x_1 x_2 x_3.$$

We can see this by noting that the  $e_i$  are the coefficients of a degree- $N$  polynomial over an auxiliary variable  $t$  whose  $N$  roots are the positions  $x_i$ , namely

$$f(t) = (t - x_1)(t - x_2) \dots (t - x_N). \quad (11)$$

For ease of implementation, however, we note that the Newton-Girard identities define a further bijection between the set of  $\{e_i\}$  to the set of power sums

$$\xi_i = \sum_{k=1}^N x_k^i. \quad (12)$$

Since the set of  $\{\xi_i\}$  are also fully symmetric under exchange of  $x_i$ , we can construct a Bose symmetric neural network function by feeding in the  $\{\xi_i\}$  as inputs ( $I_i = \xi_i$ ). In practice, for large  $N$ ,  $|\xi_i|$  can easily become extremely large; thus, we normalize the initial particle positions by a factor  $w$

$$\tilde{\xi}_i = \sum_{k=1}^N \left( \frac{x_k}{w} \right)^i \quad (13)$$

such that the  $\tilde{\xi}_i$  are at most  $\mathcal{O}(1)$  throughout the search for the ground-state wave function. The coordinate rescaling factor  $w$  is chosen to be an estimate of the magnitude of the largest typical  $x_k$  sampled from  $\pi(X) \propto \psi_0(X)^2$ . Because the ground-state wave function is not known *a priori* for systems without analytic solutions, the proper  $w$  must often be obtained through trial and error. Insufficiently large  $w$  manifest as numerical instabilities when computing the neural network

outputs

$$\psi(x_1, \dots, x_N) \sim \mathcal{A}(\tilde{\xi}_1, \dots, \tilde{\xi}_N). \quad (14)$$

The  $\{\tilde{\xi}_i\}$  retains information about the original particle coordinates but loses information about their ordering, as desired.

Second, the ground-state wave function of a system of indistinguishable bosons can be taken to be real and positive. Thus, we constrain our neural network weights and biases to be real valued. We can also, without loss of generality, write the wave function *Ansatz* as

$$\psi(x_1, \dots, x_N) = e^{-\mathcal{A}(\tilde{\xi}_1, \dots, \tilde{\xi}_N)} \times e^{-\Omega \sum_{i=1}^N x_i^2}. \quad (15)$$

Since we are looking for bound ground-state solutions of the Hamiltonians models we consider, we want to search the space of wave functions, which vanish at infinity. We multiply our neural network function by the Gaussian factor to ensure that this property is satisfied even when the neural network function itself may be relatively flat over its inputs for a given set of weights and biases. Because the bosons in the quantum systems we study are confined in a harmonic trap with  $m = \omega = 1$ , we choose  $\Omega = 0.5$ . The neural network function  $\mathcal{A}$ , then, is purely responsible for capturing the complexities of particle interactions. At zero interparticle interaction strength,  $\mathcal{A}$  should minimize trivially to zero everywhere in the ground state.

### B. Neural network architecture

To search for the ground-state energy and wave function of a quantum system, we first construct a neural network wave function *Ansatz* of the form in (15). For the neural network, we choose to use the continuously differentiable CELU as our activation function  $f$ .

$$\text{CELU}(x) = \begin{cases} x & x > 0 \\ e^x - 1 & x \leq 0 \end{cases}. \quad (16)$$

Furthermore, we initialize the real-valued weights and biases of the neural network to be  $\mathcal{O}(1)$  and nonzero, drawn randomly from a normal distribution centered around zero with standard deviation  $\sigma = 1$ . Using  $\langle E \rangle$  as a cost function, we perform gradient descent using the standard Adam optimizer [50] to approach the ground state. Adam utilizes an adaptive learning rate that allows for a quicker convergence toward minima at a reduced training cost relative to other gradient descent methods. We use Adam step sizes of between  $10^{-3}$  and  $10^{-6}$ , decreasing the step size as the minimization approaches closer to convergence. Using relatively large step sizes in certain regions of parameter space close to convergence can lead to being stuck in orbit cycles and exhibit oscillatory behavior in the cost function [51]. The other hyperparameters of the Adam optimizer describing the decay rates of the moments and ensuring numerical stability are kept at their default values:  $\beta_1 = 0.9$ ,  $\beta_2 = 0.999$ ,  $\epsilon = 10^{-8}$ .

Because the integrals involved in obtaining  $\langle E \rangle$  and  $\langle \partial E / \partial \theta \rangle$  are difficult to compute exactly for a complicated neural network function and for many particles  $N$ , we estimate the energy and the gradient using Metropolis importance sampling, drawing from the distribution  $\pi(X) \sim \psi(X, \theta)^2 / \int dX \psi(X, \theta)^2$ . However, the Hamiltonians that we

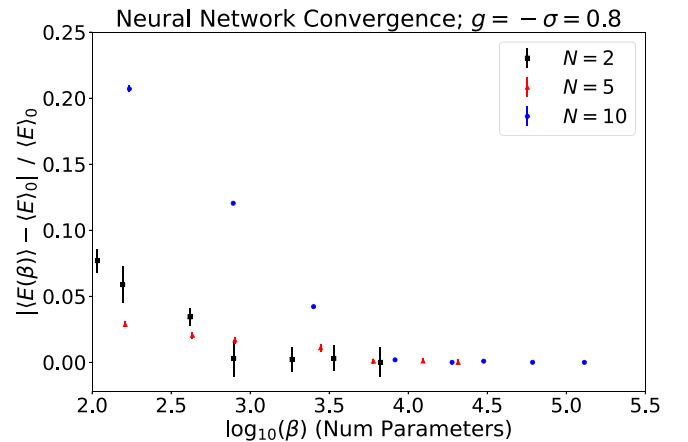


FIG. 1. Shows the relative error of the ground-state energies for the quantum system of  $N$  bosons described by the Hamiltonian in (2) as the number of neural network parameters  $\beta$  is increased. Eight layers were used in the neural network throughout. The statistical error bars from the Monte Carlo calculations on some points are too small to be seen.  $\langle E \rangle_0$  here is taken to be the value of the ground-state energy computed with the largest  $\beta$ .

study contain potential terms, which involve  $\delta$  functions in position. This creates a zero-overlap problem in the Monte Carlo evaluation of this integral in which none of the sampled position configurations induce a contribution to the energy or gradient from those  $\delta$ -function terms. We introduce a remedy in Appendix. Although computing the gradient using Monte Carlo methods is not as precise as an analytic calculation, the stochasticity tends to aid in avoiding being stuck in local minima during the minimization [52–54].

In general, there has not been a systematic solution to choosing an optimal number of layers and nodes of a neural network for any given problem [55]. Here, we arbitrarily choose to construct our neural network function with between 4–10 layers. Although a more careful analysis in choosing the number of layers has yet to be done, we decide to use more layers when solving quantum systems of more particles since we expect that more layers are needed to capture the ground state of a more complicated system. Keeping the same number of layers in solving a particular system, we systematically increase the number of nodes at each layer until we find that our prediction for the ground-state energy ceases to decrease within error bars given by the finite Monte Carlo sampling. Figure 1 shows convergence in  $\langle E(\beta) \rangle$  as we increase the total number of variational parameters (weights and biases) in the neural network. We find that an increasing number of parameters for systems of more particles is needed; however, this dependence appears to be slower than exponential in number of particles.

Finally, to perform the neural network computations and gradient descent minimizations, we utilize Google’s JAX [56] package in PYTHON. JAX has built-in automatic differentiation, just-in-time compilation speed up, and vectorization features, which make it an appealing and convenient choice to run neural network computations.



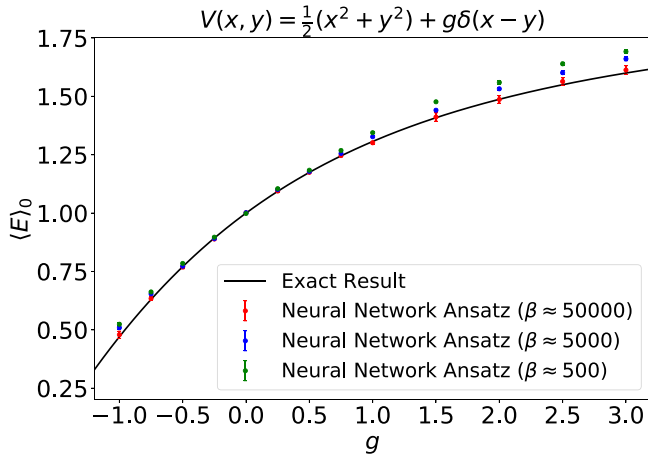


FIG. 2. Shows the ground-state energies of the quantum system described in Sec. II A. We plot the results computed from our neural network *Ansatz* (red) against the exact results (black) given by Busch *et al.* [42]. The statistical error bars from the Monte Carlo integration of the energy in most cases are too small to be seen. Convergence towards the exact result is seen as the number of parameters  $\beta$  in the neural network *Ansatz* is increased.

#### IV. RESULTS

To check that our variational neural network wave function *Ansatz* (15) is universal enough to capture the ground-state behaviors of many different quantum systems, we use our *Ansatz* to compute the ground states in various phases of the Hamiltonian models described in Sec. II. We compare our results to the analytic solutions of the exactly solvable models. We then also explore an unsolved regime of one of these models.

##### A. Cold bosons in a harmonic trap

Preliminarily, we first study a one-dimensional system of two indistinguishable bosons ( $N = 2$ ) in a harmonic trap, interacting via a contact potential (Sec. II A). An exact expression for the ground-state energy of the system was given by Busch *et al.* [42]. Using our neural network *Ansatz*, we compute the ground-state energies of the system at various interaction strengths  $g$ . Figure 2 shows a comparison of the two results. As we increase the number of parameters  $\beta$  in the neural network *Ansatz*, we find that, over the wide range of  $g$ , the neural network ground-state energies converge towards agreement with those of the analytic solution to within error bars. The mean values are also consistent with the exact expression to approximately 1–2 % precision. Throughout the final calculations for this model, we fixed our neural network to have four layers and roughly  $5 \times 10^4$  variational parameters. Although this likely uses more parameters than needed to describe the simpler two-boson system, the results give confidence that our Bose symmetric neural network *Ansatz*, as desired, is able to capture the ground-state physics of the different systems without needing to specify a particular functional form for the *Ansatz* or adjusting the architecture of the neural network.

##### B. Trapped bosons with short- and long-range interactions

Next, we examine the model in which the bosons interact with a long-range linear potential along with the  $\delta$ -function potential whose properties are summarized in Sec. II B. We are able to use the neural network *Ansatz* to study both the exactly solvable regime of  $\sigma = -g/2$  solved by Beau *et al.* [7] as well as the regime  $\sigma = -g$  in which no analytic solution is known.

###### 1. Exactly solvable regime ( $\sigma = -g/2$ )

We again begin with a simpler system of just two particles. Using our neural network *Ansatz*, we compute the ground-state energies and wave functions for various values of the coupling in the exactly solvable regime  $g = -\sigma/2$  and compare the results to the analytic solution. Figure 3 displays the ground-state wave function of the system as computed by the neural network method while Fig. 4 shows a comparison of the ground-state energies. We find that the ground states given by the neural network *Ansatz* give the same behavior as described by the exact ground-state wave function (4). When  $g > 0$  and the interactions are both repulsive, the two bosons separate an equal distance apart from the center of the harmonic trap. Alternatively, when  $g < 0$  and the interactions are both attractive, the bosons collapse towards the center of the trap and form a bright soliton. Quantitatively, the ground-state energies computed using the variational method also agree with the analytic expression (3) to within 2% statistical errors. The unintuitive quadratic dependence of the ground-state energy on  $g$  is also confirmed.

Having shown the ability of our neural network *Ansatz* to accurately compute ground-state wave functions and energies of smaller, simpler systems, we further demonstrate the versatility and generalizability of the neural network methodology by computing ground-state properties of the quantum system in (2) with 50 indistinguishable bosons.<sup>3</sup> We study both a repulsive system ( $g = 0.05$ ) and an attractive system ( $g = -0.03$ ). Using a neural network *Ansatz* with  $\beta \approx 2 \times 10^5$ , we find that the computed ground-state energies fall within 1% of those predicted by the exact  $E_0$  (3); the calculations yielded for  $g = 0.05$ ,  $E_0^{\text{NN}} = 12.083(72)$  as compared with  $E_0 = 11.984\,375$  and for  $g = -0.03$ ,  $E_0^{\text{NN}} = 20.367(76)$  as compared with  $E_0 = 20.314\,375$ .

To comprehensively check of the ground-state features of the system, we also calculate the local density profile

$$n(x) = \int dx_2 \dots dx_N \psi_0(x = x_1, x_2, \dots, x_N)^2 \quad (17)$$

and the density-density correlation function

$$g_2(y) = \int dx_1 \dots dx_N \delta(y - |x_1 - x_2|) \psi_0(x_1, \dots, x_N)^2 \quad (18)$$

<sup>3</sup>The calculations with  $N = 50$ ,  $\beta \approx 2 \times 10^5$  took roughly 2 min/training step CPU run time using  $8 \times 10^5$  Monte Carlo steps/training step on a 32-core Lambda workstation.

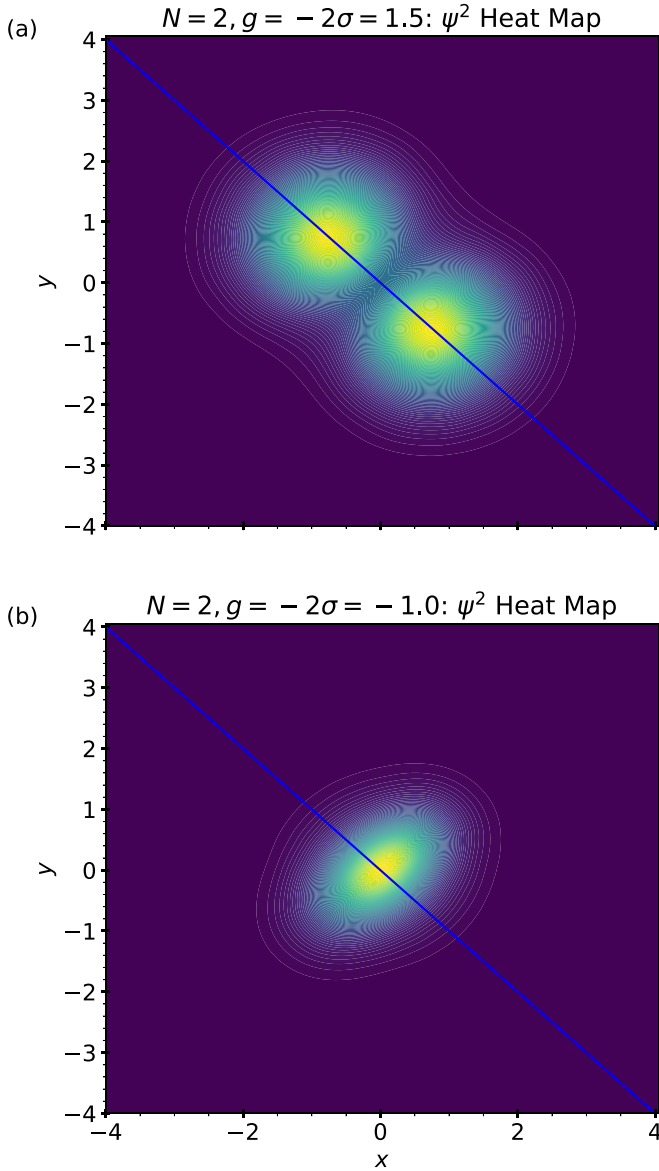


FIG. 3. Shows ground-state behaviors of two bosons interacting with short- and long-range interactions described in Sec. II B in the exactly solvable regime  $\sigma = -g/2$ . We plot the position-space probability density; brighter colors indicate a larger value of  $\psi(x, y)^2$  and the bosons have a higher probability of being there. The blue lines, which run through  $y = -x$ , are shown to guide the eye. The top shows the system with a repulsive potential ( $g = 1.5$ ) while the bottom shows the system with an attractive potential ( $g = -1.0$ ).

for both the analytic  $\psi_0$  (4) and the ground-state wave function given by the neural network *Ansatz*  $\psi_0^{\text{NN}}$ .<sup>4</sup> The density functions  $n(x)$  and  $g_2(y)$  are computed by generating histograms of positions  $x_1$  and distances  $x_1 - x_2$ , respectively, both taken from Monte Carlo samples drawn from the distribution  $\pi(X) \sim \psi_0(X)^2 / \int dX \psi_0(X)^2$ . Figure 5 shows the

<sup>4</sup>The choices  $x = x_1$  and  $\delta(y - |x_1 - x_2|)$  in  $n(x)$  and  $g_2(y)$ , respectively, are arbitrary because  $\psi_0$  is Bose symmetric.

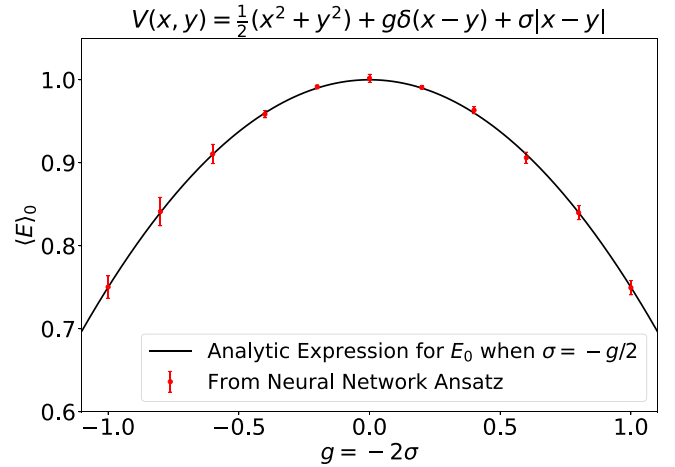


FIG. 4. Shows a comparison of the ground-state energies of the system (2) with two bosons at potential strengths  $\sigma = -g/2$ . The black line shows the analytic expression for the ground-state energy  $E_0$  from Beau *et al.* (3) and the red points are ground-state energies computed from the neural network *Ansatz*.

density functions from both  $\psi_0$  and  $\psi_0^{\text{NN}}$  plotted on top of one another. We find close agreement between the two in both phases of the model ( $g > 0$  and  $g < 0$ ). This gives evidence that the neural network *Ansatz* is not only able to provide an accurate estimate for the ground-state energy but can also faithfully encode information about the ground-state wave function for bosonic systems in one dimension.

## 2. Nonintegrable regime ( $\sigma = -g$ )

The variational approach using neural network *Ansätze* is not restricted to computing integrable systems. We also apply the methodology to explore the system of bosons interacting with short- and long-range interactions (2) with symmetric coupling strengths  $\sigma = -g$  for which no exact ground state has been found. We then compare our computed ground-state properties to those of the exactly solvable system  $\sigma = -g/2$  (see Ref. [7]). The results show that although the ground-state energies behave quantitatively very differently, the general phase structures of the two regimes are similar in nature.

We begin by looking at a system of four bosons with a strong repulsive coupling  $g = -\sigma = 2.5$  and another with an attractive coupling  $g = -\sigma = -1.0$ . Figure 6 shows the local density and the pair-correlation function for both four-boson systems. Just as with the exactly solvable model, a crystal structure is clearly seen in the case of strong repulsion. The local density shows four distinct equally spaced peaks. The pair-correlation function shows that the probability of two particles being in the same position is close to zero. Instead, the correlation function is maximal at distinct separations corresponding to  $y = \pm|x_1 - x_2| = nd$ ,  $n \in \mathbb{Z}$  where  $d$  is the interparticle spacing between two nearest neighbors in the crystal. In the case of attraction on the other hand, both the density profile and the pair density peak strongly at  $y = 0$ ; again, all of the particles clump towards the center of the harmonic trap and form a bright soliton [49].

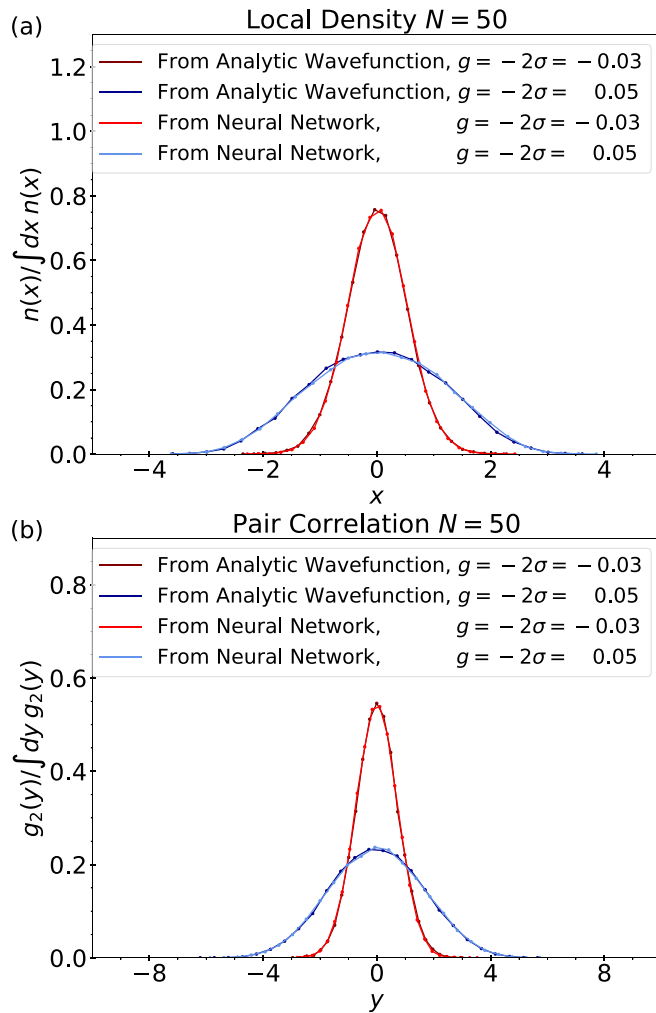


FIG. 5. Shows the ground-state density functions for the system of 50 bosons with  $\sigma = -g/2$ . The top panel displays the normalized local density while the bottom displays the normalized pair correlation. The points are heights of the position/distance binning and are linearly interpolated to guide the eye. The dark/light red show the density functions from the analytic/neural network wave function, respectively, in the attractive case while the dark/light blue show the density function from the analytic/neural network wave function in the repulsive case.

We further use our neural network *Ansatz* to search for the ground states of systems of 10 and 50 particles at various coupling strengths  $g = -\sigma$ . Figure 7 shows the computed ground-state energies as a function of  $g$  while Fig. 8 shows the single-particle and pair density profiles. We find that here, the energies exhibit the same asymptotic behavior of  $E_0 \rightarrow -\infty$  as  $|g| \rightarrow \infty$  as in the exactly solvable model. However, the ground-state energy is no longer independent of the sign of  $g$  and changes asymmetrically as the interactions in the system flip from repulsion to attraction. When  $g > 0$ , the energy drops more rapidly than in the exactly solvable case while when  $g < 0$ , the energy of the resulting soliton initially rises before dropping again at higher attractive coupling.

The qualitative features, however, again match the behaviors of the system at  $\sigma = -g/2$ . From the local density profile,

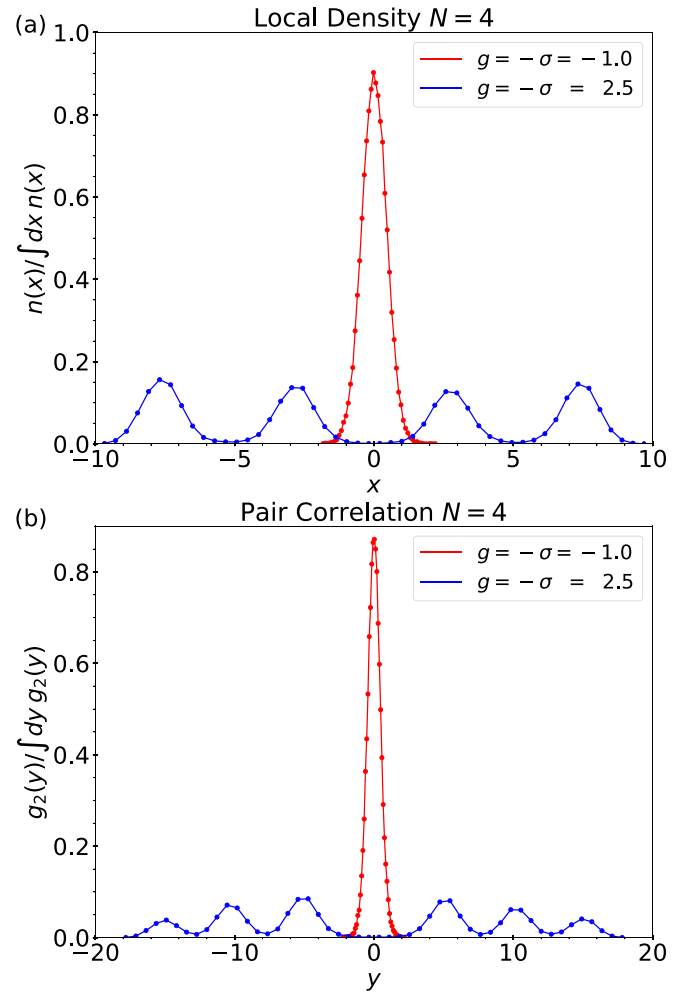


FIG. 6. Shows the ground-state density functions for the system of four bosons. The top panel displays the normalized local density while the bottom displays the normalized pair correlation. Each panel contains results for the attractive (red) and the repulsive (blue) systems.

we see that as we turn on a repulsive potential ( $g > 0$ ), the bosons begin to spread symmetrically about the center of the harmonic trap. As the repulsion becomes stronger, the particle density shows a mesalike feature of finite width  $w$ . This resembles, for example, the behavior of nuclei at saturation density. The addition of particles to the system does not change the density but instead increases the droplet size. The two-body correlation for many particles at weak repulsion increases monotonically from the maximal separation  $y = \pm w$  inwards towards  $y = 0$  since more pairs of particles have smaller rather than larger separations. In the vicinity of zero separation, the probability of finding two particles in the same location is decreased and we expect a slight dip in the pair-correlation function. This is clearly seen in the system with ten particles where the repulsion was stronger; the dip is not seen in the 50-particle system because the weaker repulsion allows the bosons to retain a tendency of remaining close to one another. On the other hand, we again find that if we instead turn on an attractive interaction between the

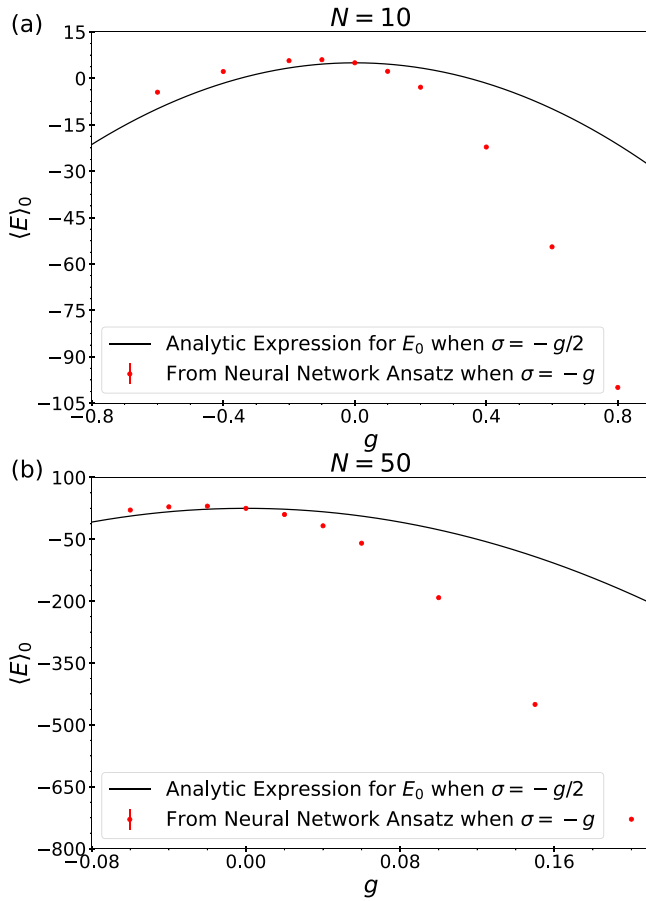


FIG. 7. Shows a comparison of the ground-state energies of the Hamiltonian (2) with  $N = 10$  and 50 bosons. The red points are energies computed from the neural network *Ansatz* at the nonintegrable regime  $\sigma = -g$ . The statistical error are too small to be seen. The black line shows the analytic expression for  $E_0$  when  $\sigma = -g/2$  for comparison.

particles, the local density as well as the pair-correlation peak sharply at the center of the harmonic trap and zero separation, respectively. This structure is again characteristic to that of a bright soliton.

## V. CONCLUSIONS

Neural networks and machine learning techniques provide powerful tools for the variational Monte Carlo approach to solving many-body quantum systems. We used a neural network to parametrize a variational wave function *Ansatz* for the ground state of systems of indistinguishable bosons in one dimension. Then using the *Ansatz*, we were able to successfully numerically compute the ground-state energies and wave functions of several quantum systems of various qualitative phases, ranging from 2–50 particles. The advantages of using a neural network *Ansatz* include being able to search the most general space of possible ground-state wave functions without needing to limit to specific functional forms. Neural networks also have convenient scalability; it is straightforward both to increase the number of variational parameters in the *Ansatz*

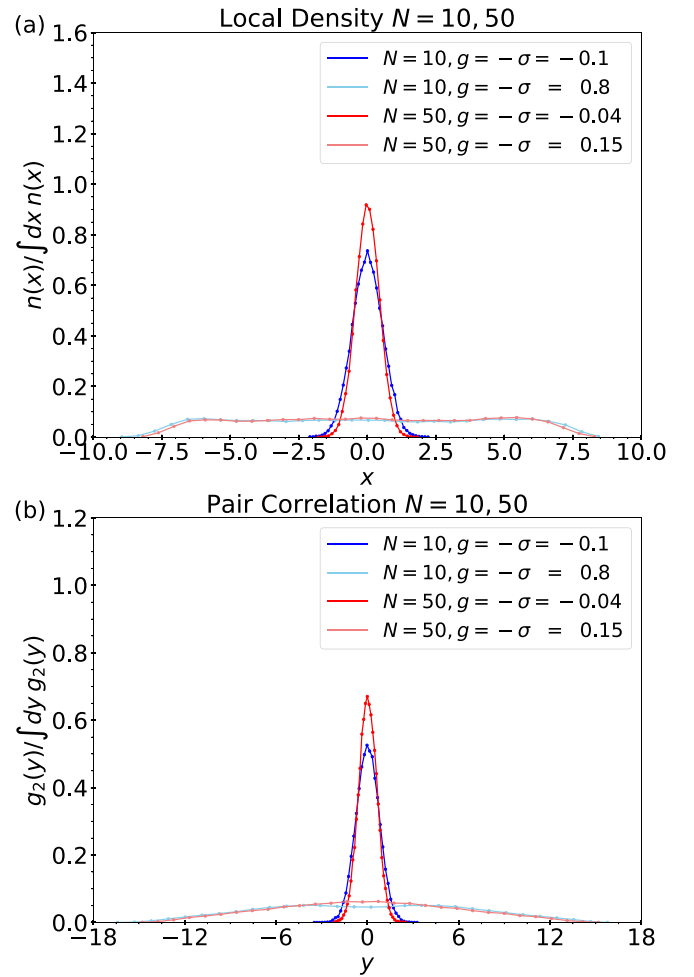


FIG. 8. Shows the ground-state density functions for systems of 10 and 50 bosons. In each panel, the (light) blue shows the (repulsive) attractive system for  $N = 10$  and the (light) red shows the (repulsive) attractive system for  $N = 50$ .

as well as increase the number of particles in the system. The application of such *Ansätze* also does not need to be restricted to solvable systems and can be used to explore properties of nonintegrable systems.

Many important many-body systems, including condensates, atoms, and nuclei involve indistinguishable particles. Wave functions of such systems must satisfy Fermi or Bose exchange symmetry. We advocated and demonstrated that using symmetric functions of the original particle coordinates as inputs to the neural network aptly enforces Bose exchange symmetry in one dimension. The symmetric inputs are chosen to be bijective to the original set of coordinates and so we do not limit the generality of the wave function space spanned by the neural network. In addition, we propose a method to deal with  $\delta$ -function interactions without the need to regularize them. Our study of applying neural network methods, using symmetrized inputs, to solve one-dimensional many-body systems of bosons can be regarded as a stepping stone towards solving three-dimensional systems of indistinguishable bosons and fermions. The successful development of such machine learning techniques would enable the computation of



important properties of nuclear many-body systems, including binding energies, spectra, and decay rates.

The use of symmetric inputs to enforce exchange symmetries of the system can be generalized to use to study systems in a higher number of spatial dimensions in a straightforward way. To enforce Bose symmetry in three dimensions, one needs to find a bijection between the set of vector coordinates  $\{\vec{r}_i\}$ ,  $i = 1, \dots, N$  and symmetric functions  $\{s_j(r_1^x, r_1^y, r_1^z, \dots, r_N^x, r_N^y, r_N^z)\}$ ,  $j = 1, \dots, M$  where  $M$  may be greater than  $3N$ . The  $\{s_j\}$  must be symmetric in exchanges  $\vec{r}_n \leftrightarrow \vec{r}_m$  but not in exchanges of individual coordinate directions (i.e.,  $r_n^x \leftrightarrow r_n^y$ ). Given the set of  $\{s_j\}$ , one should be able to reconstruct the set of  $\{r_i^x, r_i^y, r_i^z\}$  including the fact that particular sets of three coordinates specify the position of individual particles. To write a Fermi antisymmetric function in three dimensions, one theoretically simply needs to take a single Slater determinant of functions  $\phi_i(\vec{r}_j; \{\vec{r}_{j'}\})$ , where  $\phi_i$  is Bose symmetric under exchange of any  $\vec{r}_n \rightarrow \vec{r}_m$ ,  $m, n \neq j$  [37]. Thus if one finds a bijection to a set of symmetric functions necessary for enforcing Bose symmetry in three dimensions, one can use the same bijection to enforce Fermi symmetry. Work to find such a mapping is being pursued.

#### ACKNOWLEDGMENTS

We thank Andrei Alexandru and Scott Lawrence for helpful discussions and comments. This work was supported in part by the U.S. Department of Energy, Office of Nuclear Physics under Award No. DE-FG02-93ER40762.

#### APPENDIX: IMPORTANCE SAMPLING FOR $\delta$ -FUNCTION OBSERVABLES

Consider the Lieb-Liniger Hamiltonian [1,2] for a system of  $N$  indistinguishable bosons

$$\hat{H} = \sum_{i=1}^N -\frac{1}{2} \frac{\partial^2}{\partial x_i^2} + \sum_{i<j}^N g\delta(x_i - x_j) \quad (\text{A1})$$

and a positive and real wave function

$$\psi(X) = e^{-A(x_1, \dots, x_N)}, \quad (\text{A2})$$

We would like to compute the energy associated with the wave function.

$$\langle E \rangle = \frac{\int dx_1 \dots dx_N \psi \hat{H} \psi}{\int dx_1 \dots dx_N \psi^2} = K + V, \quad (\text{A3})$$

where

$$K = \frac{\int dx_1 \dots dx_N e^{-2A(X)} \sum_i \frac{1}{2} \left( \frac{\partial^2 A}{\partial x_i^2} - \left( \frac{\partial A}{\partial x_i} \right)^2 \right)}{\int dx_1 \dots dx_N e^{-2A(X)}} \quad (\text{A4})$$

$$V = \frac{\int dx_1 \dots dx_N e^{-2A(X)} \sum_{i<j} g\delta(x_i - x_j)}{\int dx_1 \dots dx_N e^{-2A(X)}}.$$

The kinetic term  $K$  can be computed easily using importance sampling

$$K = \left\langle \sum_{i=1}^N \frac{1}{2} \left( \frac{\partial^2 A}{\partial x_i^2} - \left( \frac{\partial A}{\partial x_i} \right)^2 \right) \right\rangle_{\psi^2}, \quad (\text{A5})$$

where  $\langle \dots \rangle_{\psi^2}$  denotes an average over samples drawn from the distribution

$$\Pi_{\psi^2}(x_1, \dots, x_N) = \frac{e^{-2A(X)}}{\int dx_1 \dots dx_N e^{-2A(X)}}. \quad (\text{A6})$$

Using the same importance sampling technique to compute the potential term  $V$  leads to a zero-overlap problem as described in Sec. III B. Instead we restructure the integral involved in  $V$  to become amenable to importance sampling using the following manipulations.

We first explicitly integrate out the  $x_2$  coordinate using the  $\delta$  functions. The choice of  $x_2$  is arbitrary due to the Bose exchange symmetry of the wave function

$$V = g \frac{N(N-1)}{2} I$$

$$I = \frac{\int dx_1 dx_3 \dots dx_N e^{-2A(\tilde{X})}}{\int dx_1 dx_2 \dots dx_N e^{-2A(X)}}, \quad (\text{A7})$$

where  $\tilde{X}$  denotes the set of positions where  $x_2$  is replaced by  $x_1$  (i.e.,  $\{x_1, x_1, x_3, \dots, x_N\}$ ). We use the Bose symmetry of the wave function to eliminate the sum over pairs of particles. Next, we rewrite the integral  $I$  as

$$I = \frac{\int dx_1 \dots dx_N \frac{\exp[-2A(\tilde{X})]}{\exp[-2A(X)]} \mathcal{D}(x_2) \exp[-2A(X)]}{\int dx_1 \dots dx_N \exp[-2A(X)]}, \quad (\text{A8})$$

where  $\mathcal{D}(x_2)$  is any function such that

$$\int_{-\infty}^{\infty} dx_2 \mathcal{D}(x_2) = 1. \quad (\text{A9})$$

In practice, we choose a Gaussian form

$$\mathcal{D}(x_2) = \frac{1}{\alpha \sqrt{\pi}} e^{-x_2^2/\alpha^2}. \quad (\text{A10})$$

The parameter  $\alpha$ , which is different for each different  $\psi$ , is chosen arbitrarily such that  $\mathcal{D}(y_*) \approx 10^{-5}$  where  $y_*$  is the maximum value of  $|x_2|$  over many position samples taken from  $\Pi_{\psi^2}(X)$ . In this way,  $\mathcal{D}(x_2)$  is neither too wide such that we emphasize unimportant regions of  $\psi^2$  not sampled by the Monte Carlo method nor too narrow such that we do not capture all of the important regions of  $\psi^2$ .

Now, we can also compute  $V$  by drawing samples from  $\Pi_{\psi^2}(X)$ , just as we did for  $K$ .

$$V = g \frac{N(N-1)}{2} \left\langle \frac{e^{-2A(x_1, x_1, x_3, \dots, x_N)}}{e^{-2A(x_1, x_2, \dots, x_N)}} \mathcal{D}(x_2) \right\rangle_{\psi^2}. \quad (\text{A11})$$

We note that this method of avoiding the zero-overlap problem occasionally produces large variances and may fail to converge if  $e^{-2A(X)}$  is small for many different sampled  $X$ . However, the method is still overall reliable because this is precisely the region of  $e^{-2A(X)}$  that is sampled the least by the Metropolis algorithm.

- [1] E. H. Lieb and W. Liniger, *Phys. Rev.* **130**, 1605 (1963).
- [2] E. H. Lieb, *Phys. Rev.* **130**, 1616 (1963).
- [3] B. Sutherland, *Phys. Rev. B* **38**, 6689 (1988).
- [4] F. Calogero, *J. Math. Phys.* **10**, 2197 (1969).
- [5] B. Sutherland, *Phys. Rev. A* **4**, 2019 (1971).
- [6] J. Moser, *Adv. Math.* **16**, 197 (1975).
- [7] M. Beau, S. M. Pittman, G. E. Astrakharchik, and A. del Campo, *Phys. Rev. Lett.* **125**, 220602 (2020).
- [8] A. L. Fetter, *J. Low Temp. Phys.* **106**, 643 (1997).
- [9] D. Ceperley, *Phys. Rev. B* **18**, 3126 (1978).
- [10] B. Wilson, A. Foerster, C. Kuhn, I. Roditi, and D. Rubeni, *Phys. Lett. A* **378**, 1065 (2014).
- [11] C. J. Umrigar, K. G. Wilson, and J. W. Wilkins, *Phys. Rev. Lett.* **60**, 1719 (1988).
- [12] W. M. C. Foulkes, L. Mitás, R. J. Needs, and G. Rajagopal, *Rev. Mod. Phys.* **73**, 33 (2001).
- [13] K. Ido, T. Ohgoe, and M. Imada, *Phys. Rev. B* **92**, 245106 (2015).
- [14] P. Kościuk, M. Płodzień, and T. Sowiński, *Europhys. Lett.* **123**, 36001 (2018).
- [15] M. H. Kalos, *Phys. Rev.* **128**, 1791 (1962).
- [16] J. Carlson, V. Pandharipande, and R. Wiringa, *Nucl. Phys. A* **401**, 59 (1983).
- [17] J. Carlson, K. E. Schmidt, and M. H. Kalos, *Phys. Rev. C* **36**, 27 (1987).
- [18] R. Jastrow, *Phys. Rev.* **98**, 1479 (1955).
- [19] Y. Lutsyshyn, *J. Chem. Phys.* **146**, 124102 (2017).
- [20] P. Kościuk, *Few-Body Syst.* **58**, 59 (2017).
- [21] P. Kościuk and T. Sowiński, *New J. Phys.* **22**, 093053 (2020).
- [22] S. R. White, *Phys. Rev. Lett.* **69**, 2863 (1992).
- [23] S. Das Sarma, D.-L. Deng, and L.-M. Duan, *Phys. Today* **72**(3), 48 (2019).
- [24] G. Carleo, I. Cirac, K. Cranmer, L. Daudet, M. Schuld, N. Tishby, L. Vogt-Maranto, and L. Zdeborová, *Rev. Mod. Phys.* **91**, 045002 (2019).
- [25] K. Hornik, M. Stinchcombe, and H. White, *Neural Networks* **2**, 359 (1989).
- [26] K. Hornik, *Neural Networks* **4**, 251 (1991).
- [27] G. Carleo and M. Troyer, *Science* **355**, 602 (2017).
- [28] J. Chen, S. Cheng, H. Xie, L. Wang, and T. Xiang, *Phys. Rev. B* **97**, 085104 (2018).
- [29] Y. Zheng, H. He, N. Regnault, and B. A. Bernevig, *Phys. Rev. B* **99**, 155129 (2019).
- [30] H. Saito, *J. Phys. Soc. Jpn.* **87**, 074002 (2018).
- [31] A. Gnech, C. Adams, N. Brawand, G. Carleo, A. Lovato, and N. Rocco, *Few-Body Syst.* **63** 7 (2021).
- [32] C. Adams, G. Carleo, A. Lovato, and N. Rocco, *Phys. Rev. Lett.* **127**, 022502 (2021).
- [33] A. Lovato, C. Adams, G. Carleo, and N. Rocco, *Phys. Rev. Res.* **4**, 043178 (2022).
- [34] Y. Yang and P. Zhao, *Phys. Lett. B* **835**, 137587 (2022).
- [35] Y. L. Yang and P. W. Zhao, *Phys. Rev. C* **107**, 034320 (2023).
- [36] E. M. Nordhagen, J. M. Kim, B. Fore, A. Lovato, and M. Hjorth-Jensen, *Frontiers Phys.* **11**, 1061580 (2023).
- [37] D. Pfau, J. S. Spencer, A. G. D. G. Matthews, and W. M. C. Foulkes, *Phys. Rev. Res.* **2**, 033429 (2020).
- [38] C. N. Yang, *Phys. Rev. Lett.* **19**, 1312 (1967).
- [39] C. P. Yang, *Phys. Rev. A* **2**, 154 (1970).
- [40] C. N. Yang and C. P. Yang, *J. Math. Phys.* **10**, 1115 (1969).
- [41] J. B. McGuire, *J. Math. Phys.* **5**, 622 (1964).
- [42] T. Busch, B.-G. Englert, K. Rzazewski, and M. Wilkens, *Found. Phys.* **28**, 549 (1998).
- [43] A. Minguzzi and P. Vignolo, *AVS Quantum Sci.* **4**, 027102 (2022).
- [44] M. Beau and A. del Campo, *SciPost Phys. Core* **4**, 030 (2021).
- [45] J. Yang and A. del Campo, *Phys. Rev. Lett.* **129**, 150601 (2022).
- [46] R. B. Laughlin, *Phys. Rev. Lett.* **50**, 1395 (1983).
- [47] E. H. Lieb, N. Rougerie, and J. Yngvason, *J. Stat. Phys.* **172**, 544 (2018).
- [48] E. Wigner, *Phys. Rev.* **46**, 1002 (1934).
- [49] B. Kalinikos and A. Ustinov, *Sol. State Phys.* **64**, 193 (2013).
- [50] D. P. Kingma and J. Ba, Adam: A method for stochastic optimization, in *3rd International Conference on Learning Representations*, edited by Y. Bengio and Y. LeCun (ICLR, San Diego, CA, USA, 2015), Conference Track Proceedings (2015).
- [51] K. Nar and S. S. Sastry, Step size matters in deep learning, in *Proceedings of the 32nd International Conference on Neural Information Processing Systems*, NIPS'18 (Curran Associates Inc., Red Hook, NY, USA, 2018), pp. 3440–3448.
- [52] L. Bottou *et al.*, Stochastic gradient descent tricks, *Neural Networks: Tricks of the Trade: Second Edition*, edited by G. Montavon, G. B. Orr, and K.-R. Müller (Springer, Berlin, Heidelberg, 2012), pp. 421–436.
- [53] S.-I. Amari, *Neurocomputing* **5**, 185 (1993).
- [54] B. Kleinberg, Y. Li, and Y. Yuan, An alternative view: When does {SGD} escape local minima? in *Proceedings of 35th International Conference*, proceedings of machine learning, edited by J. Dy and A. Krause (PMLR, 2018), Vol. 80, pp. 2698–2707.
- [55] D. Stathakis, *Int. J. Remote Sensing* **30**, 2133 (2009).
- [56] J. Bradbury, R. Frostig, P. Hawkins, M. J. Johnson, C. Leary, D. Maclaurin, G. Necula, A. Paszke, J. VanderPlas, S. Wanderman-Milne, and Q. Zhang, JAX: composable transformations of Python+NumPy programs, 2018.