

Variational Neural-Network Ansatz for Continuum Quantum Field Theory

John M. Martyn^{1,2,†}, Khadijeh Najafi,^{3,4} and Di Luo^{1,2,5,*}¹Center for Theoretical Physics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA²The NSF AI Institute for Artificial Intelligence and Fundamental Interactions³IBM Quantum, IBM T. J. Watson Research Center, Yorktown Heights, New York 10598, USA⁴MIT-IBM Watson AI Lab, Cambridge, Massachusetts 02142, USA⁵Department of Physics, Harvard University, Cambridge, Massachusetts 02138, USA (Received 16 December 2022; revised 21 July 2023; accepted 25 July 2023; published 24 August 2023)

Physicists dating back to Feynman have lamented the difficulties of applying the variational principle to quantum field theories. In nonrelativistic quantum field theories, the challenge is to parametrize and optimize over the infinitely many n -particle wave functions comprising the state's Fock-space representation. Here we approach this problem by introducing *neural-network quantum field states*, a deep learning ansatz that enables application of the variational principle to nonrelativistic quantum field theories in the continuum. Our ansatz uses the Deep Sets neural network architecture to simultaneously parametrize all of the n -particle wave functions comprising a quantum field state. We employ our ansatz to approximate ground states of various field theories, including an inhomogeneous system and a system with long-range interactions, thus demonstrating a powerful new tool for probing quantum field theories.

DOI: [10.1103/PhysRevLett.131.081601](https://doi.org/10.1103/PhysRevLett.131.081601)

Introduction.—It is notoriously challenging to solve an interacting quantum field theory (QFT), with analytical solutions limited to mean field systems and few exactly solvable models. While perturbation theory is useful for weakly interacting field theories, much important physics lies in the nonperturbative regime, such as in quantum chromodynamics. Although lattice field theory is a well-established method for studying QFT, it can be computationally expensive to reach the continuum limit [1]. It is therefore an ongoing quest to develop new methods for solving and simulating QFTs in the continuum.

An alternative approach to probing QFT is through applying the variational principle to field theory. However, as pointed out by Feynman [2], a significant barrier to this method is constructing a variational ansatz for a quantum field state whose expectation values can be efficiently computed and optimized. Reference [3] took a key step in this direction by extending the matrix product state (MPS) from lattice models to field theories through the development of the continuous matrix product state (cMPS), which can describe low-energy states of 1D nonrelativistic QFTs [4]. However, unlike the MPS on the lattice, cMPSs suffer substantial drawbacks in their optimization and applicability. While cMPSs can be efficiently optimized in translation-invariant systems [5], in inhomogeneous settings, the direct optimization of cMPSs is prone to numerical instabilities [6,7], and algorithms to better adapt cMPSs to these settings ultimately rely on discretizations and/or interpolations [7–9]. Moreover, approximation techniques are needed to apply cMPSs to systems with generic long-range interactions [10,11], leaving

cMPSs often restricted to contact interactions. Furthermore, extensions of cMPS to higher spatial dimensions are susceptible to UV divergences and thus limited in application [12–15].

Meanwhile, the striking success of machine learning in physically motivated contexts [16] has recently inspired the application of neural networks to problems in quantum physics, leading to a new variational ansatz known as a *neural-network quantum state* (NQS) [17]. This ansatz parametrizes a wave function by neural networks, which are optimized over to approximate a state of interest [18–20] or even simulate real-time dynamics [21–30]. These methods are justified by recent proofs that NQSs can efficiently model highly entangled states and subsume tensor networks in their expressivity [31–36].

However, NQSs have not been generalized to QFT in the continuum, which would necessitate an efficient parametrization of the infinitely many n -particle wave functions comprising a quantum field state's Fock-space representation. Prominent instances of NQSs in continuous space have only studied systems with a fixed number of particles [37–43] and have not yet made the connection to QFT with variable particle number. Applications of machine learning to QFT have only studied lattice field theory—most notably, Monte Carlo sampling in the Lagrangian formulation [44–47] and simulations in the Hamiltonian formulation [28,48,49]—rather than applying the variational principle directly to continuum QFT.

In this paper, we fill this gap by developing *neural-network quantum field states* (NQFSs), extending the range of NQSs to field theories. In this introductory work, we

focus on the nonrelativistic QFT of bosons, while the approach we introduce opens up the opportunity to study generic field theories. Working directly in the continuum, we model a quantum field state by exploiting its Fock-space representation as a superposition of n -particle wave functions. We adopt permutation-invariant methods from deep learning—specifically, the Deep Sets architecture [50]—to simultaneously parametrize the infinitely many n -particle wave functions by a finite number of neural networks. We further apply an algorithm for variational Monte Carlo (VMC) in Fock space, that enables the estimation and optimization of the energy of a quantum field state. The merit of the NQFS lies in its applicability to various field theories, such as inhomogeneous systems and systems with long-range interactions, including both periodic and closed boundary conditions. We demonstrate these properties by employing NQFSs to approximate the ground states of 1D QFTs, benchmarking on the Lieb-Liniger model, the Calogero-Sutherland model, and a regularized Klein-Gordon model.

Neural-network quantum field states (NQFSs).—We aim to approximate the ground state of a bosonic QFT, whereupon second quantization particles are created and annihilated by the operators $\hat{\psi}^\dagger(x)$ and $\hat{\psi}(x)$, respectively, which obey the commutation relation $[\hat{\psi}(x), \hat{\psi}^\dagger(x')] = \delta(x - x')$. Focusing on nonrelativistic systems in 1D, the corresponding quantum field states live in a Fock space that is a direct sum over all symmetrized n -particle Hilbert spaces, such that an arbitrary state may be expressed as a superposition of permutation-invariant, unnormalized, n -particle wave functions $\varphi_n(\mathbf{x}_n)$:

$$|\Psi\rangle = \bigoplus_{n=0}^{\infty} |\varphi_n\rangle = \bigoplus_{n=0}^{\infty} \int d^n x \varphi_n(\mathbf{x}_n) |\mathbf{x}_n\rangle. \quad (1)$$

Here, we use the shorthand $\mathbf{x}_n = (x_1, x_2, \dots, x_n)$ to denote a vector containing the positions of n particles, and we employ the n -particle basis $|\mathbf{x}_n\rangle := (1/\sqrt{n!}) \hat{\psi}(x_1)^\dagger \times \hat{\psi}(x_2)^\dagger \dots \hat{\psi}(x_n)^\dagger |\Omega\rangle$, where $|\Omega\rangle$ is the vacuum.

We introduce the NQFS as a variational ansatz to parametrize such a quantum field state with neural networks—more specifically, to parametrize each of its n -particle wave functions with a common neural network architecture. This imposes the following constraints on the architecture: First, the architecture must be permutation invariant in order to model bosonic wave functions. And more stringently, to model wave functions of arbitrary particle number, the architecture must also be *variadic*: able to accept an arbitrary number of arguments [i.e., particle positions $\mathbf{x}_n = (x_1, x_2, \dots, x_n)$ for arbitrary n].

Remarkably, we can realize both of these properties with the Deep Sets neural network architecture, which is both permutation invariant and variadic. Reference [50] introduces the Deep Sets architecture to model a permutation-invariant function on a set X and proves that any such

function $f(X)$ can be decomposed as

$$f(X) = \rho\left(\sum_{x \in X} \phi(x)\right) \quad (2)$$

for appropriately chosen constituent functions ρ and ϕ . The function ϕ maps the inputs into a feature space of possibly higher dimension, the sum aggregates the embedded inputs in a permutation-invariant manner, and ρ adds the correlations necessary to output the function of interest. This decomposition is clearly permutation invariant, as each $\phi(x)$ is summed independently; it is also variadic, because an arbitrary number of arguments may be included in the sum.

In the setting of machine learning, this decomposition is used to learn permutation-invariant functions by parametrizing ρ and ϕ as deep neural networks [50]. This approach has recently been used in NQS contexts to model ground states of atomic nuclei [39–41] and bosons at fixed particle number [37]. Building on this prior work, here we are the first to leverage Deep Sets to develop a neural network ansatz for continuum QFT, with an arbitrary number of particles.

For the construction of a NQFS, denoted $|\Psi^{\text{NQFS}}\rangle$, we parametrize each n -particle wave function as a product of two Deep Sets [51]—one for particle positions $\{x_i\}_{i=1}^n$, and the other for particle separations $\{x_i - x_j\}_{i < j}$:

$$|\Psi^{\text{NQFS}}\rangle = \bigoplus_{n=0}^{\infty} \int d^n x \varphi_n^{\text{NQFS}}(\mathbf{x}_n) |\mathbf{x}_n\rangle, \quad \varphi_n^{\text{NQFS}}(\mathbf{x}_n) = \frac{1}{L^{n/2}} \cdot f_1(\{x_i\}_{i=1}^n) \cdot f_2(\{x_i - x_j\}_{i < j}), \quad (3)$$

where the factor of $(1/L^{n/2})$ is included for dimensional consistency. The constituent functions of the Deep Sets f_1 and f_2 (ρ_1, ϕ_1 and ρ_2, ϕ_2 , respectively) are parametrized as feed-forward neural networks, which capture the global properties and pairwise correlations of the state, respectively. Fundamentally, the variadic property of Deep Sets enables us to parametrize *infinitely many* n -particle wave functions with a *finite* number of neural networks, and thus represent an arbitrary state in Fock space. We provide a graphical illustration of a NQFS in Fig. 1.

We employ feature embeddings on inputs to the neural networks. In a closed system with hard walls at $x = 0$ and $x = L$, we use the embedding $x_i \mapsto [(x_i/L), 1 - (x_i/L)]$ for f_1 , and $(x_i - x_j) \mapsto [(x_i - x_j)/L]^2$ for f_2 (note that this embedding must be even in order to achieve permutation invariance). For a periodic system, we reflect this periodicity with the embedding $x_i \mapsto \{\sin[(2\pi/L)x_i], \cos[(2\pi/L)x_i]\}$ for f_1 , and $(x_i - x_j) \mapsto \cos[(2\pi/L)(x_i - x_j)]$ for f_2 .

VMC in Fock space.—Our approach to ground-state approximation employs the variational principle—that the

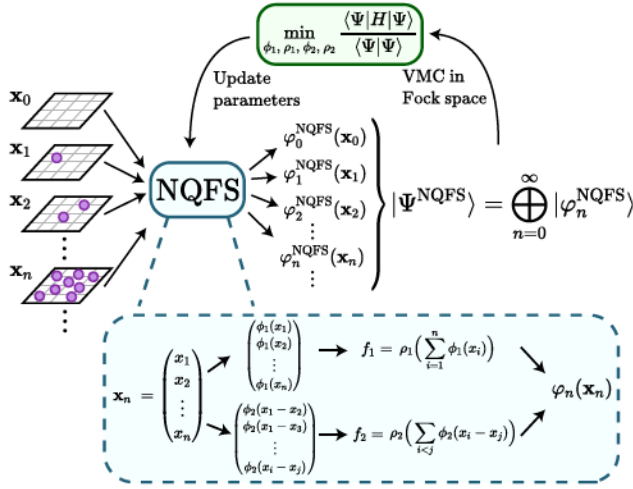


FIG. 1. Diagram of a NQFS: Particle positions $\mathbf{x}_n = (x_1, x_2, \dots, x_n)$ are input to evaluate the corresponding n -particle wave functions $\phi_n^{\text{NQFS}}(\mathbf{x}_n)$, of which the NQFS $|\Psi^{\text{NQFS}}\rangle$ is modeled as a direct sum. Computation of $\phi_n^{\text{NQFS}}(\mathbf{x}_n)$ is shown in the inset as a product of two Deep Sets, f_1 and f_2 . Optimization of the NQFS is illustrated in the feedback loop; the variational parameters are updated via VMC in Fock space to minimize the energy.

energy of an arbitrary quantum state is greater than or equal to the ground-state energy. To estimate and optimize the energy of $|\Psi^{\text{NQFS}}\rangle$, we generalize traditional VMC to *VMC in Fock space*. As we illustrate in the Supplemental Material [52], the energy of a quantum field state $|\Psi\rangle$ can be expressed as an expectation value of an n -particle local energy $E_n^{\text{loc}}(x)$, taken jointly over a probability distribution of the particle number, denoted P_n , and the probability distribution of the n -particle wave function, denoted $|\bar{\varphi}_n(x)|^2$:

$$E(|\Psi\rangle) = \frac{\langle\P|H|\Psi\rangle}{\langle\P|\Psi\rangle} = \mathbb{E}_{n \sim P_n} \mathbb{E}_{\mathbf{x}_n \sim |\bar{\varphi}_n|^2} [E_n^{\text{loc}}(\mathbf{x}_n)]. \quad (4)$$

Explicitly, the n -particle local energy is $E_n^{\text{loc}}(\mathbf{x}_n) = \langle\mathbf{x}_n|H|\Psi\rangle/\langle\mathbf{x}_n|\varphi_n\rangle$, the probability distribution of the particle number is $P_n := \langle\varphi_n|\varphi_n\rangle/\langle\P|\Psi\rangle$, and the probability distribution of the n -particle wave function is $|\bar{\varphi}_n(\mathbf{x}_n)|^2 := |\varphi_n(\mathbf{x}_n)|^2/\langle\varphi_n|\varphi_n\rangle$.

We estimate the above expectation value and its standard deviation by drawing samples from P_n and $|\bar{\varphi}_n(x)|^2$ jointly via the *Markov chain Monte Carlo (MCMC) in Fock space* algorithm described in the Supplemental Material [52]. This algorithm employs Metropolis-Hastings sampling with a proposal function that allows the particle number to increase or decrease, such that P_n and $|\bar{\varphi}_n|^2$ are sampled simultaneously. The standard deviation of the expectation value is controllable and decays as the inverse square root of the number of samples.

We use gradient-based optimization to optimize a NQFS. As shown in the Supplemental Material [52], the derivative of the energy with respect to a variational parameter—say, a parameter of the neural network—may also be expressed as a combination of expectation values over P_n and $|\bar{\varphi}_n(x)|^2$. We again estimate these as empirical means over MCMC samples, and we feed the resulting gradient estimate into the ADAM algorithm [70] to minimize the energy. The systematic uncertainty of the optimized NQFS can be studied with variance extrapolation [71,72], as explained in the Supplemental Material [52].

It is useful here to employ regularization and system-dependent modifications that assist with the learning procedure (see the Supplemental Material [52] for specific details). To facilitate smooth optimization over particle number, we multiply $\phi_n^{\text{NQFS}}(\mathbf{x}_n)$ by a parametrizable regularization factor that keeps the probability distribution over particle number well behaved. Moreover, for systems with hard wall boundary conditions, we multiply $\phi_n^{\text{NQFS}}(\mathbf{x}_n)$ by a cutoff factor [73–76] that forces it to vanish at the boundaries. Lastly, for an interaction potential that diverges as two particles approach each other $x_i \rightarrow x_j$, the eigenstates exhibit a cusp at $x_i = x_j$ that is quantified by Kato’s cusp condition [77,78]; we account for this cusp by multiplying $\phi_n^{\text{NQFS}}(\mathbf{x}_n)$ by a Jastrow factor that satisfies Kato’s cusp condition.

To highlight applicability, we apply NQFSs to the Lieb-Liniger model in an inhomogenous system with hard walls, the Calogero-Sutherland model which has long-range interactions, and a regularized Klein-Gordon model that does not conserve particle number. We select these models because they have exactly solvable ground states to benchmark against (details provided in the Supplemental Material [52]). These models are nonperturbative, in that their interactions strongly influence the ground states and cannot be treated perturbatively. Beyond these models, NQFSs can be straightforwardly applied to nonexactly solvable field theories, which we exemplify in the Supplemental Material [52].

Lieb-Liniger model.—Let us illustrate the application of NQFSs to the quintessential nonrelativistic QFT, the Lieb-Liniger model [79,80]:

$$H_{\text{LL}} = \frac{1}{2m} \int dx \frac{d\hat{\psi}^\dagger(x)}{dx} \frac{d\hat{\psi}(x)}{dx} - \mu \int dx \hat{\psi}^\dagger(x) \hat{\psi}(x) + g \int dx \hat{\psi}^\dagger(x) \hat{\psi}^\dagger(x) \hat{\psi}(x) \hat{\psi}(x). \quad (5)$$

This model describes bosons of chemical potential μ interacting via a contact interaction of strength g : $V(x-y) = 2g\delta(x-y)$. Experimentally, the Lieb-Liniger model has been realized with ultracold atoms [81] and in optical lattices [82,83]. The model is integrable and thus provides a useful benchmark for numerics; as the

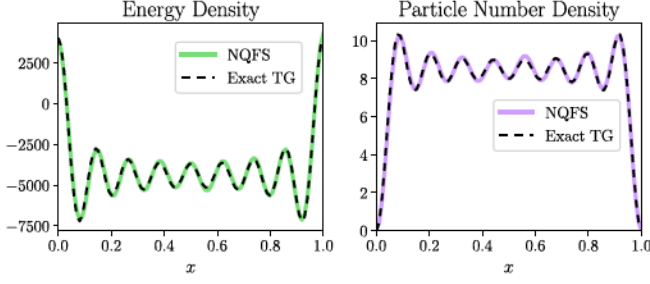


FIG. 2. The energy density (left) and particle number density (right) of the NQFS (at $g = 10^6$) and the exact Tonks-Girardeau ground state of the Lieb-Liniger model.

Hamiltonian conserves particle number, its ground state lies in a definite particle number sector, which we denote by n_0 .

With our NQFS ansatz, we probe the ground state of an inhomogeneous system with hard walls at $x = 0$ and $x = L$, accommodating these boundary conditions with a cutoff factor as mentioned above, and accounting for the delta function potential with an appropriate Jastrow factor. We first consider the Tonks-Girardeau limit, in which $g \rightarrow \infty$ and the model can be mapped to noninteracting spinless fermions [7,84]. We simulate this limit with parameters $L = 1$, $m = 1/2$, $\mu = (8.75\pi)^2$, and $g = 10^6$, in which case the ground state has energy $E_0 = -4031.79$ and $n_0 = 8$ particles. After optimization, we obtain an energy $E = -4030.9 \pm 0.1$ and particle number $n = 8.00 \pm 0.01$, showcasing that NQFSs are capable of finding both the ground state and the particle number sector in which it lies. We also compare the energy density and the particle number density of the NQFS with the Tonks-Girardeau solution in Fig. 2, the clear agreement of which further justifies that the NQFS has learned the ground state correctly.

To study performance at smaller values of g , we set $\mu = 185$ and $g = 10$, in which case $E_0 = -954.60$ and $n_0 = 10$. Our NQFS finds $E = -952.20 \pm 0.09$ and $n = 10.00 \pm 0.03$, indicating that good performance is maintained away from the Tonks-Girardeau limit.

Calogero-Sutherland model.—We further aim to study a system with long-range interactions, for which purpose we consider the Calogero-Sutherland model on a ring of length L [85,86]:

$$H_{\text{CS}} = \frac{1}{2m} \int dx \frac{d\hat{\psi}^\dagger(x)}{dx} \frac{d\hat{\psi}(x)}{dx} - \mu \int dx \hat{\psi}^\dagger(x) \hat{\psi}(x) \\ + \frac{g\pi^2}{2L^2} \int dx dy \hat{\psi}^\dagger(x) \hat{\psi}^\dagger(y) \hat{\psi}(y) \hat{\psi}(x) \\ \times \left[\sin\left(\frac{\pi}{L}(x-y)\right) \right]^{-2}. \quad (6)$$

This model describes particles interacting via an inverse square sinusoidal potential with interaction strength g . The ground-state energy is $E_0 = (\pi^2 \lambda^2 / 6mL^2) n_0(n_0^2 - 1) - \mu n_0$, where $\lambda = \frac{1}{2}(1 + \sqrt{1 + 4mg})$, and n_0 minimizes E_0 and is the particle number of the ground state.

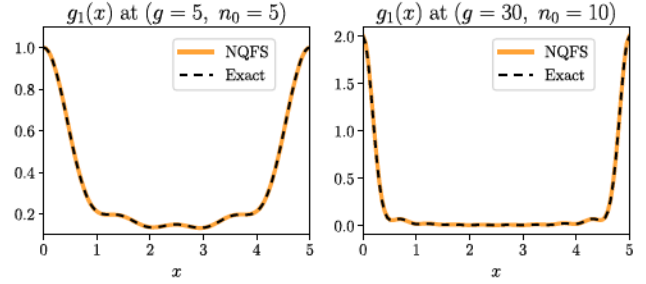


FIG. 3. The NQFS and exact one-body density matrices $g_1(x)$ of the Calogero-Sutherland model at $(g = 5, n_0 = 5)$ (left) and $(g = 30, n_0 = 10)$ (right).

We first study this model with system size $L = 5$ at $g = 5$ and $\mu = 3 \cdot 5^2 \cdot (\pi^2 \lambda^2 / 6mL^2)$, which has the exact ground-state energy $E_0 = -156.317$ and particle number $n_0 = 5$. Employing the NQFS with an appropriate Jastrow factor, we obtain $E = -156.291 \pm 0.003$ and $n = 5.000 \pm 0.007$, indicative of great performance. We also look at the one-body density matrix $g_1(x) = n \int d^{n-1} x \varphi_n^*(x, x_2, \dots, x_n) \times \varphi_n(0, x_2, \dots, x_n)$ and compare it to the exact solution in Fig. 3.

We next consider a more strongly interacting limit with $g = 30$ and $\mu = 3 \times 10^2 \cdot (\pi^2 \lambda^2 / 6mL^2)$, in which case $E_0 = -5132.76$ and $n_0 = 10$. The NQFS yields $E = -5131.77 \pm 0.09$ and $n = 10.000 \pm 0.004$, and we again illustrate the one-body density matrix in Fig. 3. The agreement with the exact solution in both of these cases evidences the NQFS's applicability to systems with long-range interactions.

Regularized Klein-Gordon model.—The previous Hamiltonians conserve particle number and have ground states that lie in a single particle number sector. We are also interested in Hamiltonians that violate particle number conservation and whose ground states are necessarily superpositions of n -particle wave functions.

For this purpose, we consider the Klein-Gordon model: $H_{\text{KG}} = \frac{1}{2} \int dx |\hat{\pi}(x)|^2 + |\nabla \hat{\phi}(x)|^2 + m^2 |\hat{\phi}(x)|^2$, where $\hat{\phi}(x)$ and $\hat{\pi}(x)$ are canonically conjugate fields. To prevent this Hamiltonian from diverging, it is necessary to introduce a momentum cutoff Λ and regularize H_{KG} ; a convenient way to do so is by adding a counterterm $(1/\Lambda^2) |\nabla \hat{\pi}(x)|^2$ [87]. With this modification, we can introduce creation and annihilation operators by the change of variables $\hat{\phi}(x) = (1/\sqrt{2\Lambda})(\hat{\psi}(x) + \hat{\psi}^\dagger(x))$ and $\hat{\pi}(x) = -i\sqrt{(\Lambda/2)}(\hat{\psi}(x) - \hat{\psi}^\dagger(x))$, upon which the regularized Hamiltonian is mapped to the following quadratic Hamiltonian [14,87]:

$$H_{\text{Quad}} = \int dx \frac{d\hat{\psi}^\dagger(x)}{dx} \frac{d\hat{\psi}(x)}{dx} + v \int dx \hat{\psi}^\dagger(x) \hat{\psi}(x) \\ + \lambda \int dx (\hat{\psi}^\dagger(x) \hat{\psi}^\dagger(x) + \hat{\psi}(x) \hat{\psi}(x)), \quad (7)$$

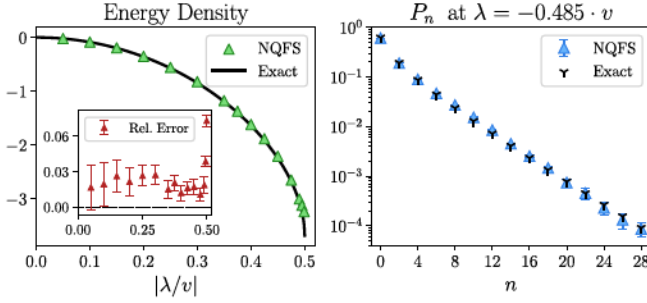


FIG. 4. *Left:* The NQFS and exact energy densities of the regularized Klein-Gordon model across a range of λ , as well as the corresponding relative errors and standard deviations in the inset. *Right:* The NQFS and exact particle number distributions at $\lambda = -0.485 \cdot v$.

where the coefficients are $v = \frac{1}{2}(m^2 + \Lambda^2)$ and $\lambda = \frac{1}{4}(m^2 - \Lambda^2)$, and the last term violates particle number conservation. In this formulation, $|\lambda/v| \leq 1/2$ in order for the Hamiltonian to be well-defined; as $|\lambda/v| \rightarrow 1/2$, the spectrum becomes gapless. The ground state of H_{Quad} can be exactly determined by moving to momentum space and performing a Bogoliubov transformation.

We study a finite system of size $L = 1$ with coefficient $v = 6$ over a range of λ from $\lambda = 0$ to the critical point $\lambda = -3$. As suggested by the particle-number-nonconserving term in the Hamiltonian, we restrict our NQFS to n -particle wave functions with an even number of particles. We compare the resulting energy density of the NQFS with the exact solution in Fig. 4. This plot indicates good agreement with the exact solution, achieving errors around a few percent that increase as the critical point is approached. To emphasize that the NQFS has truly learned a superposition of n -particle wave functions, in Fig. 4 we plot the particle number distributions P_n at $\lambda = -0.485 \cdot v$. We see clear agreement between the NQFS and exact distributions, evidencing that NQFSs can accurately represent such a superposition.

While these results are promising, we note that NQFSs struggle on systems of larger size and greater couplings v and λ . As we note in the Supplemental Material [52], one reason for this is that the local energy of the term $\lambda \int dx \psi^\dagger(x) \psi^\dagger(x) + \text{H.c.}$ imposes a sort of “cusp condition” that is nontrivial to build into a variational ansatz. As we do not explicitly account for this condition here, the energy suffers a large variance that hinders its optimization and can result in a subpar ground-state approximation; this effect becomes most pronounced near the critical point. Improving performance near the critical point provides an interesting direction for future work.

Conclusion.—We have developed neural-network quantum field states as a variational ansatz for quantum field theories in the continuum, and demonstrated their application to an inhomogeneous system, a system with long-range interactions, and a system that breaks particle number

conservation. Our work opens the door to new variational techniques in QFT.

Beyond the applications demonstrated here, there are ample refinements and extensions of NQFSs. From an algorithmic point of view, optimization over the particle number can be hindered by n -particle wave functions whose magnitudes differ significantly across particle number sectors. While we mitigated this issue by incorporating a regularization factor, it would be beneficial to develop more precise algorithms for stable optimization over particle number. Furthermore, another direction for improving performance is to properly account for the cusp conditions imposed by QFT Hamiltonians—in particular, those that violate particle number conservation.

From a physics perspective, a particularly enticing idea is to extend NQFSs to relativistic QFT, which would require a move to the thermodynamic limit and also a resolution to Feynman’s warning about the sensitivity to high frequencies suffered by variational methods in relativistic QFT [2]. A generalization to fermionic fields would also be fascinating; to enforce antisymmetry here, one could use a Deep Sets with an antisymmetric aggregation function, such as a Slater determinant. Another interesting avenue is application to real-time dynamics of QFT, which could provide physics simulations beyond the Euclidean time formulation. Moreover, the NQFS framework can be straightforwardly adapted to $d \geq 2$ spatial dimensions by treating positions x_i as d -dimensional vectors. This could enable simulation of higher-dimensional QFTs and surpass the capabilities of traditional methods like continuous tensor networks [12,13].

Codes to run the simulations in this Letter are available on GitHub at Ref. [88].

The authors acknowledge useful feedback and discussions from Isaac Chuang, Martin Ganahl, Bryan Clark, William Detmold, Alexander Zlokapa, Trevor McCourt, Rikab Gambhir, Aldi Maraya, Phiala Shanahan, and James Stokes. J. M. M. acknowledges the MIT SuperCloud and Lincoln Laboratory Supercomputing Center for providing the computing resources used in this paper. J. M. M. acknowledges support from the National Science Foundation Graduate Research Fellowship under Grant No. 2141064. D. L. acknowledges support from the NSF AI Institute for Artificial Intelligence and Fundamental Interactions (IAIFI). This material is based upon work supported 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. DE-SC0012704.

*Corresponding author.

diluo@mit.edu

†jmmartyn@mit.edu

- [1] H. J. Rothe, *Lattice Gauge Theories: An Introduction* (World Scientific Publishing Company, Singapore, 2012).
- [2] R. P. Feynman, in *Variational Calculations in Quantum Field Theory* (World Scientific Publishing, Singapore, 1988), pp. 28–40.
- [3] F. Verstraete and J. I. Cirac, *Phys. Rev. Lett.* **104**, 190405 (2010).
- [4] J. Haegeman, J. I. Cirac, T. J. Osborne, and F. Verstraete, *Phys. Rev. B* **88**, 085118 (2013).
- [5] M. Ganahl, J. Rincón, and G. Vidal, *Phys. Rev. Lett.* **118**, 220402 (2017).
- [6] J. Haegeman, D. Draxler, V. Stojevic, I. Cirac, T. Osborne, and F. Verstraete, *SciPost Phys.* **3**, 006 (2017).
- [7] B. Tuybens, J. De Nardis, J. Haegeman, and F. Verstraete, *Phys. Rev. Lett.* **128**, 020501 (2022).
- [8] M. Ganahl, [arXiv:1712.01260](https://arxiv.org/abs/1712.01260).
- [9] M. Ganahl and G. Vidal, *Phys. Rev. B* **98**, 195105 (2018).
- [10] I. V. Lukin and A. G. Sotnikov, *Phys. Rev. B* **106**, 144206 (2022).
- [11] J. Rincón, M. Ganahl, and G. Vidal, *Phys. Rev. B* **92**, 115107 (2015).
- [12] A. Tilloy and J. I. Cirac, *Phys. Rev. X* **9**, 021040 (2019).
- [13] D. Jennings, C. Brockett, J. Haegeman, T. J. Osborne, and F. Verstraete, *New J. Phys.* **17**, 063039 (2015).
- [14] T. D. Karanikolaou, P. Emonts, and A. Tilloy, *Phys. Rev. Res.* **3** (2021).
- [15] T. Shachar and E. Zohar, *Phys. Rev. D* **105**, 045016 (2022).
- [16] H. W. Lin, M. Tegmark, and D. Rolnick, *J. Stat. Phys.* **168**, 1223 (2017).
- [17] G. Carleo and M. Troyer, *Science* **355**, 602 (2017).
- [18] R. G. Melko, G. Carleo, J. Carrasquilla, and J. I. Cirac, *Nat. Phys.* **15**, 887 (2019).
- [19] K. Choo, G. Carleo, N. Regnault, and T. Neupert, *Phys. Rev. Lett.* **121**, 167204 (2018).
- [20] G. Torlai, G. Mazzola, J. Carrasquilla, M. Troyer, R. Melko, and G. Carleo, *Nat. Phys.* **14**, 447 (2018).
- [21] I. L. Gutiérrez and C. B. Mendl, *Quantum* **6**, 627 (2022).
- [22] M. Schmitt and M. Heyl, *Phys. Rev. Lett.* **125**, 100503 (2020).
- [23] F. Vicentini, A. Biella, N. Regnault, and C. Ciuti, *Phys. Rev. Lett.* **122**, 250503 (2019).
- [24] N. Yoshioka and R. Hamazaki, *Phys. Rev. B* **99**, 214306 (2019).
- [25] J. Carrasquilla, D. Luo, F. Pérez, A. Milsted, B. K. Clark, M. Volkovs, and L. Aolita, *Phys. Rev. A* **104**, 032610 (2021).
- [26] M. J. Hartmann and G. Carleo, *Phys. Rev. Lett.* **122**, 250502 (2019).
- [27] A. Nagy and V. Savona, *Phys. Rev. Lett.* **122**, 250501 (2019).
- [28] D. Luo, Z. Chen, K. Hu, Z. Zhao, V. M. Hur, and B. K. Clark, *Phys. Rev. Res.* **5**, 013216 (2023).
- [29] D. Luo, Z. Chen, J. Carrasquilla, and B. K. Clark, *Phys. Rev. Lett.* **128**, 090501 (2022).
- [30] M. Reh, M. Schmitt, and M. Gärtner, *Phys. Rev. Lett.* **127**, 230501 (2021).
- [31] D.-L. Deng, X. Li, and S. Das Sarma, *Phys. Rev. X* **7**, 021021 (2017).
- [32] J. Chen, S. Cheng, H. Xie, L. Wang, and T. Xiang, *Phys. Rev. B* **97**, 085104 (2018).
- [33] X. Gao and L.-M. Duan, *Nat. Commun.* **8**, 1 (2017).
- [34] I. Glasser, N. Pancotti, M. August, I. D. Rodriguez, and J. I. Cirac, *Phys. Rev. X* **8**, 011006 (2018).
- [35] Y. Levine, O. Sharir, N. Cohen, and A. Shashua, *Phys. Rev. Lett.* **122**, 065301 (2019).
- [36] O. Sharir, A. Shashua, and G. Carleo, *Phys. Rev. B* **106**, 205136 (2022).
- [37] G. Pescia, J. Han, A. Lovato, J. Lu, and G. Carleo, *Phys. Rev. Res.* **4**, 023138 (2022).
- [38] M. Wilson, S. Moroni, M. Holzmann, N. Gao, F. Wudarski, T. Vegge, and A. Bhowmik, *Phys. Rev. B* **107**, 235139 (2023).
- [39] A. Lovato, C. Adams, G. Carleo, and N. Rocco, *Phys. Rev. Res.* **4**, 043178 (2022).
- [40] A. Gnech, C. Adams, N. Brawand, G. Carleo, A. Lovato, and N. Rocco, *Few-Body Syst.* **63**, 7 (2021).
- [41] C. Adams, G. Carleo, A. Lovato, and N. Rocco, *Phys. Rev. Lett.* **127**, 022502 (2021).
- [42] D. Pfau, J. S. Spencer, A. G. D. G. Matthews, and W. M. C. Foulkes, *Phys. Rev. Res.* **2**, 033429 (2020).
- [43] J. Hermann, Z. Schätzle, and F. Noé, *Nat. Chem.* **12**, 891 (2020).
- [44] M. S. Albergo, G. Kanwar, and P. E. Shanahan, *Phys. Rev. D* **100**, 034515 (2019).
- [45] K. A. Nicoli, C. J. Anders, L. Funcke, T. Hartung, K. Jansen, P. Kessel, S. Nakajima, and P. Stornati, *Phys. Rev. Lett.* **126**, 032001 (2021).
- [46] M. S. Albergo, G. Kanwar, S. Racanière, D. J. Rezende, J. M. Urban, D. Boyda, K. Cranmer, D. C. Hackett, and P. E. Shanahan, *Phys. Rev. D* **104**, 114507 (2021).
- [47] G. Kanwar, M. S. Albergo, D. Boyda, K. Cranmer, D. C. Hackett, S. Racanière, D. J. Rezende, and P. E. Shanahan, *Phys. Rev. Lett.* **125**, 121601 (2020).
- [48] D. Luo, S. Yuan, J. Stokes, and B. K. Clark, [arXiv:2211.03198](https://arxiv.org/abs/2211.03198).
- [49] D. Luo, G. Carleo, B. K. Clark, and J. Stokes, *Phys. Rev. Lett.* **127**, 276402 (2021).
- [50] M. Zaheer, S. Kottur, S. Ravanbakhsh, B. Póczos, R. Salakhutdinov, and A. Smola, [arXiv:1703.06114](https://arxiv.org/abs/1703.06114).
- [51] While in principle, f_1 suffices to model an arbitrary state as per Ref. [50], we found it helpful to also include f_2 to better capture pairwise interactions.
- [52] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/PhysRevLett.131.081601> for a review of relevant topics, derivations of the Monte Carlo algorithms employed, and details on the models studied in the main text. The Supplemental Material also includes Refs. [53–69].
- [53] A. Altland and B. D. Simons, *Condensed Matter Field Theory* (Cambridge University Press, Cambridge, England, 2010).
- [54] A. Tilloy, Notes on continuous tensor network states (2021), <https://atilloy.com/2021/09/30/notes-on-continuous-tensor-networks/>.
- [55] F. Vicentini, D. Hofmann, A. Szabó, D. Wu, C. Roth, C. Giuliani, G. Pescia, J. Nys, V. Vargas-Calderón, N. Astrakhantsev, and G. Carleo, *SciPost Phys. Codebases* **007** (2022).

- [56] M. Hibat-Allah, M. Ganahl, L. E. Hayward, R. G. Melko, and J. Carrasquilla, *Phys. Rev. Res.* **2**, 023358 (2020).
- [57] H. Saito and M. Kato, *J. Phys. Soc. Jpn.* **87**, 014001 (2018).
- [58] O. Sharir, Y. Levine, N. Wies, G. Carleo, and A. Shashua, *Phys. Rev. Lett.* **124**, 020503 (2020).
- [59] G. E. Astrakharchik, D. M. Gangardt, Y. E. Lozovik, and I. A. Sorokin, *Phys. Rev. E* **74**, 021105 (2006).
- [60] D. Frenkel and B. Smit, *Understanding Molecular Simulation: From Algorithms to Applications* (Elsevier, New York, 2001), Vol. 1.
- [61] J. P. Valleau and L. K. Cohen, *J. Chem. Phys.* **72**, 5935 (1980).
- [62] H. Evertz, Computer simulations (2020), <https://itp.tugraz.at/~evertz/Computersimulations/cs2020.pdf>.
- [63] Z. Chen, D. Luo, K. Hu, and B. K. Clark, *arXiv:2212.06835*.
- [64] M. Gaudin, *Phys. Rev. A* **4**, 386 (1971).
- [65] M. T. Batchelor, X. W. Guan, N. Oelkers, and C. Lee, *J. Phys. A* **38**, 7787 (2005).
- [66] A. Selberg, *Norsk Mat. Tidsskr.* **26**, 71 (1944).
- [67] R. M. Fernandes, Lecture notes: BCS theory of superconductivity (2015), http://materias.df.uba.ar/sa2020c2/files/2012/07/Lecture-BCS_Fernandes.pdf.
- [68] E. Demler, Strongly correlated systems in atomic and condensed matter physics (2018), <http://cmt.harvard.edu/demler/TEACHING/Physics284/physics284.html>.
- [69] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen *et al.*, *Adv. Neural Inform. Proces. Syst.* **32**, 8026 (2019).
- [70] D. P. Kingma and J. Ba, *arXiv:1412.6980*.
- [71] S. Sorella, *Phys. Rev. B* **64** (2001).
- [72] T. Kashima and M. Imada, *J. Phys. Soc. Jpn.* **70**, 2287 (2001).
- [73] A. Sarsa and C. Le Sech, *J. Chem. Theory Comput.* **7**, 2786 (2011).
- [74] A. Flores-Riveros and A. Rodriguez-Contreras, *Phys. Lett. A* **372**, 6175 (2008).
- [75] A. Sarsa, E. Buendía, and F. Gálvez, *J. Phys. B* **47**, 185002 (2014).
- [76] C. Laughlin and S.-I. Chu, *J. Phys. A* **42**, 265004 (2009).
- [77] T. Kato, *Commun. Pure Appl. Math.* **10**, 151 (1957).
- [78] W. M. C. Foulkes, L. Mitás, R. J. Needs, and G. Rajagopal, *Rev. Mod. Phys.* **73**, 33 (2001).
- [79] E. H. Lieb and W. Liniger, *Phys. Rev.* **130**, 1605 (1963).
- [80] E. H. Lieb, *Phys. Rev.* **130**, 1616 (1963).
- [81] E. H. Lieb, R. Seiringer, and J. Yngvason, *Phys. Rev. Lett.* **91** (2003).
- [82] B. Paredes, A. Widera, V. Murg, O. Mandel, S. Fölling, I. Cirac, G. V. Shlyapnikov, T. W. Hänsch, and I. Bloch, *Nature (London)* **429**, 277 (2004).
- [83] T. Kinoshita, T. Wenger, and D. S. Weiss, *Science* **305**, 1125 (2004).
- [84] M. Girardeau, *J. Math. Phys. (N.Y.)* **1**, 516 (1960).
- [85] B. Sutherland, *Phys. Rev. A* **4**, 2019 (1971).
- [86] B. Sutherland, *Phys. Rev. A* **5**, 1372 (1972).
- [87] V. Stojevic, J. Haegeman, I. P. McCulloch, L. Tagliacozzo, and F. Verstraete, *Phys. Rev. B* **91** (2015).
- [88] J. M. Martyn and D. Luo, Neural-Network Quantum Field States (2023), <https://github.com/jmmartyn/Neural-Network-Quantum-Field-States>.