

Unified quantum framework for electrons and ions: The self-consistent harmonic approximation on a neural network curved manifold

Lorenzo Monacelli ^{1,2,*}, Antonio Siciliano ² and Nicola Marzari ¹

¹*Theory and Simulation of Materials (THEOS), and National Centre for Computational Design and Discovery of Novel Materials (MARVEL), École Polytechnique Fédérale de Lausanne (EPFL), 1015 Lausanne, Switzerland*

²*University of Rome “Sapienza,” Piazzale Aldo Moro 5, 00185 Roma, Italy*



(Received 24 October 2024; accepted 29 October 2025; published 9 December 2025)

The numerical solution of the many-body problem, which involves interacting electrons and ions, is a key challenge in condensed matter physics, chemistry, and materials science. Traditional methods to solve the multicomponent quantum Hamiltonian are usually specialized for one kind of particles—electrons or ions—and can suffer from a methodological gap when applied to the other ones. This work extends the self-consistent harmonic approximation, a proven successful technique for simulating quantum ions at finite temperatures in anharmonic crystals, to electrons. The approach minimizes the total free energy by optimizing an ansatz density matrix, solving a fermionic self-consistent harmonic Hamiltonian on a curved manifold parameterized through a neural network. This approach preserves an analytical expression for entropy, enabling the direct computation of free energies and phase diagrams of materials. By benchmarking this technique across several prototypical cases—a double-well potential, the hydrogen atom, and the H₂ dissociation—we demonstrate that it can address both the ground- and excited-state properties of electronic systems, capture quantum tunneling and static electronic correlations, and thereby provide a unified quantum framework of electrons and atomic nuclei.

DOI: [10.1103/7d6s-zz54](https://doi.org/10.1103/7d6s-zz54)

I. INTRODUCTION

Modeling materials from first principles requires the numerical solution of the fundamental equations of quantum mechanics for coupled electrons and ions. This is challenging if more than a few particles are involved. Thus, many approximations are typically employed to perform computational simulations and make predictions [1]. The most common is treating electrons within a mean-field approach, like in density-functional theory (DFT) [2,3], simplifying the electron-electron interaction and significantly reducing complexity. Moreover, the quantum nature of atomic nuclei is usually neglected, as deviations from the classical behavior typically occur only at cryogenic temperatures. Finally, nuclei and electrons are mostly treated adiabatically; i.e., electrons remain in the ground state during the ionic dynamics. However, there are many materials in which these approximations fail, sometimes quite spectacularly. The mean-field approach to the electron-electron interaction is unsuitable in strongly correlated materials [4]. Nuclear quantum effects play a significant role in the thermodynamic phase stability at high pressure or when the atomic mass is small, as is the case for hydrogen [5,6], hydrides [7,8], and water [9–12]. The limits of computer simulations become even more severe

when energy flows between electronic and nuclear excitations, leading to the breakdown of the adiabatic approximation. This is relevant, e.g., when a system is near a conical intersection between electronic states, in metals with a low Fermi velocity [13–15], or in nonradiative electron-hole recombination [16]. The complexity in modeling nonadiabatic phenomena hinders impactful technological advancements in materials design, like unconventional high-temperature superconductors, where low Fermi temperature coexists with a high superconductive critical temperature [17], or preventing the nonradiative electron-hole recombination in metal-halide solar cells [16].

The most common approach to address nonadiabaticity is accounting for multiple electronic excited states as different potential energy landscapes (PES) for nuclei [18]. The Ehrenfest dynamics is a mean-field approach to the electron-ion interactions, where ions move according to average forces on the different PES whose weights are evaluated by projecting the electronic dynamics within a time-dependent framework like TD-DFT. However, it has been shown that the Ehrenfest dynamic does not satisfy the detailed balance and leads to a wrong thermal state [19], thus being ineffective for simulating thermal equilibrium. The solution to this problem is achieved by the so-called “surface hopping” methods [18], in which each ionic trajectory evolves on a single PES with a probability of swapping electronic state during the dynamics [20]. Final observables are then evaluated by averaging many different trajectories. However, these calculations are very expensive, requiring a dynamic treatment of multiple electronic excited states, and do not take into account intrinsically quantum nuclear effects [21], for which a path-integral reformulation is needed [22] that further increases computational complexity. Moreover, these methods are explicitly devised for insulators or molecules. Their application is problematic

*Contact author: lorenzo.monacelli@uniroma1.it

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.

in metallic systems where there is an infinite number of electronically excited states accessible by ionic excitations.

The opposite approach, which does not assume a multiple PES, is to treat electrons and nuclei within the same theory and solve the complete quantum problem. However, this poses substantial challenges, as methods traditionally used to tackle electrons and ions follow opposite strategies. While a mean-field approach is often enough for electrons, ions are intrinsically strongly correlated even in the simplest harmonic crystals; in fact, the lattice excited states, the phonons, are collective excitations where the ionic motion is correlated. Thus, one must choose methods capable of accounting for correlations, like quantum Monte Carlo (QMC) or post-Hartree-Fock. However, these methods were devised to solve pure quantum states at $T = 0$ K, as thermal excitations are not so relevant in electronic systems at room temperature; in fact, a temperature of 300 K corresponds approximately to 0.026 eV, which is usually small compared to the typical electronic excitation energies. In contrast, temperature heavily affects ions. The thermal excitations of nuclei trigger most of the physically relevant phenomena for materials, like phase transitions and thermal expansion. Path-integral molecular dynamics (PIMD) is the state-of-art for nuclei when quantum effects are relevant [23]. The computational complexity required to converge PIMD calculations depends on the temperature, diverging for $T \rightarrow 0$ K. This makes PIMD suitable for studying high-temperature states near the limit where particles behave classically, and it is ideal for nuclei. However, electrons at room temperature are very far from the classical limit. For these reasons, path-integral Monte Carlo simulations have been applied to the electrons-ion plasma far above room temperature [23,24], in the range of 10^4 – 10^9 K. Moreover, all the aforementioned methods suffer from the complexity arising when computing the phase diagram of materials. No direct computable expression for the entropy exists in these methods, and differences between the free energy of the structures need to be evaluated through thermodynamic integration, which is often out of reach even when fast potentials are available within the Born-Oppenheimer (BO) approximation.

In this work, we introduce a framework that can be applied simultaneously to both electrons and ions at room temperature. To this aim, we extend the self-consistent harmonic approximation [25–27] (SCHA), a technique devised initially to solve the adiabatic quantum nuclear problem, to electrons. The SCHA was successfully employed to study many equilibrium properties of matter where quantum atomic effects are dominant, like the phase diagram of high-pressure hydrogen [5,6,28,29], hydrides superconductors [7,8], hydrate clathrates [11], metal-halide perovskite solar cells [30], and the emergency of charge-density waves in 2D and bulk transition metal dichalcogenides [31–34]. The original method approximates the quantum density matrix of nuclei as the Gaussian equilibrium state of a trial quantum harmonic Hamiltonian. Gaussian density matrices provide an appropriate description for nuclei in a crystal as they oscillate around fixed equilibrium positions. However, they are not suited to describe electronic wave functions, as the lighter mass of electrons produces delocalized states where deviations from the Gaussian shape become much more relevant, leading to

the breakdown of the SCHA in electronic systems. Besides, while the indistinguishability of nuclei plays a negligible part in thermodynamics, electronic exchange gives rise to the Pauli exclusion principle, which plays a dominant role in atomic physics.

Here, we extend the SCHA framework to handle indistinguishable particles with spin. Our approach also enhances the density matrix beyond the Gaussian approximation, but it keeps a computational cost comparable to the original SCHA. This is achieved thanks to the recently developed nonlinear SCHA [35,36], where the degrees of freedom of the trial density matrix are augmented by introducing a curved manifold on which the Gaussian density matrix is defined. The curved manifold deforms the density matrix, increasing the variational space available beyond Gaussians for optimizing the density matrix to minimize the free energy. To parameterize and optimize the curved manifold, we introduce a neural network representation, where the systematic addition of hidden layers further improves the variational space of the density matrix. The core of the approach is illustrated in Fig. 1, where the original Gaussian density matrix is deformed through the transformation encoded by the neural network.

We start by briefly reviewing the SCHA in Sec. II, extending the theory in the case of indistinguishable particles, with particular care to the case of fermions (Sec. III). Then, we introduce the curved manifold that allows the wave function to deviate from the original Gaussian shape (Sec. IV). We discuss the neural network parameterization of the curved manifold (Sec. V) and how to encode crystal symmetries within it. Section VI discusses the specific requirements of the neural network parameterization to deal with fermionic degrees of freedom. In Sec. VII, we apply the theory to study several challenging cases where the SCHA fails, such as the profound double-well potential (to model quantum tunneling), the hydrogen atom (to model the electron-ion cusp), and the H_2 dissociation (to model electron-electron correlations), both in the ground bonding state (singlet) and in the excited anti-bonding one (triplet). We then address limitations and future perspectives of this approach in Sec. VIII.

II. SELF-CONSISTENT HARMONIC APPROXIMATION

The self-consistent harmonic approximation is a variational approach to the solution of the time-independent Schrödinger equation:

$$\hat{H}|\psi\rangle = E|\psi\rangle, \quad (1)$$

where $|\psi\rangle$ is the ground-state wave function of a generic Hamiltonian operator \hat{H} and E is its associated energy. We indicate an operator in the Hilbert space with a hat. The SCHA was successfully employed to solve the nuclear quantum problem, where \hat{H} depends only on atomic coordinates, and the electrons' degrees of freedom are integrated out within the BO approximation. Here, we extend the formalism to electrons. Thus, \hat{H} depends on the position operators \hat{R}_i of both nuclei and electrons, overcoming the BO approximation. We neglect the role of spin interactions, such as spin-orbit coupling, and assume that the total spin operator \hat{S}^2 and its projection on one axis \hat{S}_z commute with the Hamiltonian \hat{H} . The SCHA constrains the wave function $|\tilde{\psi}\rangle$ to the ground state of an

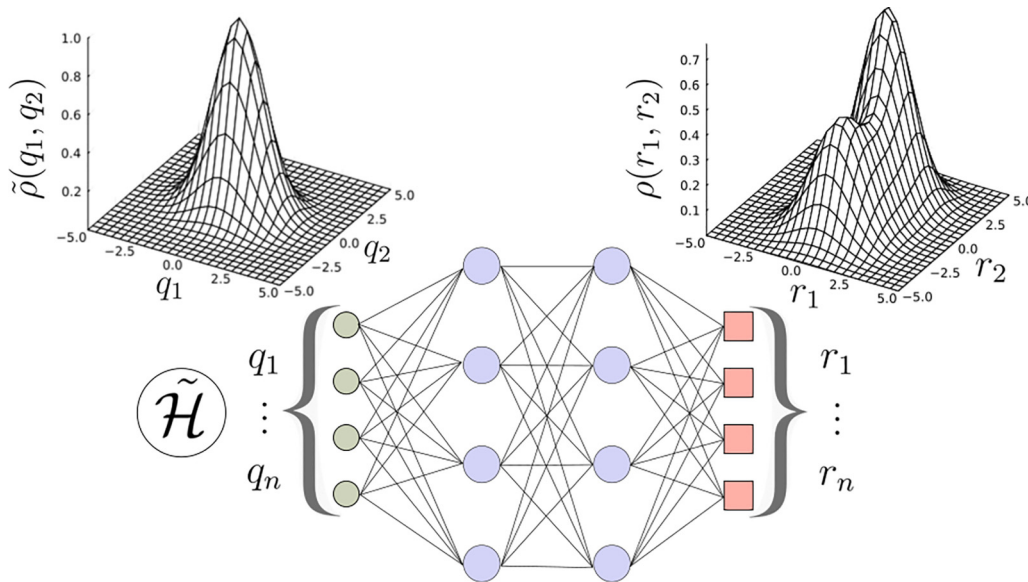


FIG. 1. Schematic representation of the SCHA on a neural network curved manifold [nonlinear SCHA (NLSCHA)]. The Gaussian density matrix $\tilde{\rho}$, solution of the trial harmonic Hamiltonian $\tilde{\mathcal{H}}$ in an auxiliary space \mathbf{q} (the curved manifold) is transformed into real space thanks to the transformation $\mathbf{r}(\mathbf{q})$ defined through a neural network. The nonlinearity of the transformation deforms the Gaussian shape and introduces features on the final density matrix, including correlations beyond the original distribution.

auxiliary harmonic Hamiltonian $\hat{\mathcal{H}}[\mathcal{R}, \Phi]$ depending on the parameters \mathcal{R} and Φ as

$$\hat{\mathcal{H}}[\mathcal{R}, \Phi] = \hat{K} + \frac{1}{2} \sum_{ij} (\hat{R}_i - \mathcal{R}_i) \Phi_{ij} (\hat{R}_j - \mathcal{R}_j), \quad (2)$$

$$\hat{\mathcal{H}}[\mathcal{R}, \Phi] |\tilde{\psi}[\mathcal{R}, \Phi]\rangle = \mathcal{E} |\tilde{\psi}[\mathcal{R}, \Phi]\rangle, \quad (3)$$

where \hat{K} is the kinetic energy operator and \hat{R}_i is the position operator of the i th particle (electron or nucleus). Latin indices indicate both the particle and the Cartesian component, and bold symbols represent tensorial quantities (e.g., vectors and matrices). The vector \mathcal{R}_i indicates the average position of particle i (centroids), while Φ are auxiliary force constant matrices encoding the quantum fluctuations and the correlations between different particles. These parameters must minimize the total energy within the Rayleigh-Ritz variational principle

$$E^{(\text{SCHA})} = \min_{\mathcal{R}, \Phi} \langle \tilde{\psi}[\mathcal{R}, \Phi] | \hat{\mathcal{H}} | \tilde{\psi}[\mathcal{R}, \Phi] \rangle. \quad (4)$$

The minimum of Eq. (4) is satisfied by the self-consistent equations [37]

$$\langle \tilde{\psi}[\mathcal{R}, \Phi] | \frac{d\hat{V}}{dR_i} | \tilde{\psi}[\mathcal{R}, \Phi] \rangle = 0, \quad (5)$$

$$\langle \tilde{\psi}[\mathcal{R}, \Phi] | \frac{d^2\hat{V}}{dR_i dR_j} | \tilde{\psi}[\mathcal{R}, \Phi] \rangle = \Phi_{ij}, \quad (6)$$

where $\hat{V} = V(\hat{\mathbf{R}})$ is the interaction potential as a function of the position operators of the particles. By solving Eqs. (5) and (6), \mathcal{R} and Φ are updated to get an auxiliary Hamiltonian $\hat{\mathcal{H}}[\mathcal{R}^{(\text{next})}, \Phi^{(\text{next})}]$, defining a ground state $|\tilde{\psi}^{(\text{next})}\rangle$ for the next iteration. The process is repeated until convergence of Eqs. (5) and (6).

The SCHA works at finite temperature by replacing the wave function with the density matrix $\hat{\rho}$ and the averages

as traces. In this case, the auxiliary Hamiltonian defines the density matrix through the usual equilibrium relationship

$$\hat{\rho}[\mathcal{R}, \Phi] = \frac{\exp(-\beta \hat{\mathcal{H}}[\mathcal{R}, \Phi])}{\mathcal{Z}}, \quad (7)$$

$$\mathcal{Z} = \text{Tr}[\exp(-\beta \hat{\mathcal{H}}[\mathcal{R}, \Phi])]. \quad (8)$$

We omit the explicit dependence of $\hat{\rho}$ from the \mathcal{R} and Φ in the following equations. Indeed, the equilibrium solution is obtained by minimizing the free energy functional $\mathcal{F}[\hat{\rho}]$ [37],

$$\mathcal{F}[\hat{\rho}] = \langle \hat{H} \rangle_{\hat{\rho}} - TS[\hat{\rho}], \quad (9)$$

where T is the temperature and $S[\hat{\rho}]$ is the entropy functional, defined as

$$S[\hat{\rho}] = -k_b (\ln \hat{\rho})_{\hat{\rho}}, \quad (10)$$

where the quantum averages are

$$\langle \hat{H} \rangle_{\hat{\rho}} = \text{Tr}[\hat{\rho} \hat{H}] \quad (11)$$

and k_b is the Boltzmann constant. The equilibrium free energy F is always lower than the free energy functional $\mathcal{F}[\hat{\rho}]$ for any $\hat{\rho}$ that is not the exact equilibrium one. Thus, the self-consistent equations to update \mathcal{R} and Φ are obtained by minimizing Eq. (9),

$$\left\langle \frac{dV}{dR_i} \right\rangle_{\hat{\rho}} = \text{Tr} \left[\hat{\rho} \frac{dV}{dR_i} \right] = 0, \quad (12)$$

$$\left\langle \frac{d^2V}{dR_i dR_j} \right\rangle_{\hat{\rho}} = \text{Tr} \left[\hat{\rho} \frac{d^2V}{dR_i dR_j} \right] = \Phi_{ij}[\hat{\rho}]. \quad (13)$$

III. SCHA FOR ELECTRONS

In this section, we apply the SCHA to electrons, showing how it is possible to deal with a system of indistinguishable fermions within the SCHA framework. To describe a

fermionic (or bosonic) system, the auxiliary harmonic Hamiltonian must commute with the operator S_{ij} that swaps the position of particle i with j , for any i and j . This imposes the constraints on \mathcal{R} and Φ ,

$$\forall i, j \quad S_{ij}\mathcal{R} = \mathcal{R}, \quad S_{ij}\Phi S_{ij}^\dagger = \Phi. \quad (14)$$

The result is that the centroid position must be the same for all electrons, reducing its independent components from $3N$ to 3. The only surviving elements of the Φ are the 6 independent degrees of freedom that encode the coupling between an electron with itself ($\Phi^{(\text{same})}$) and the 9 degrees of freedom that couple two different electrons ($\Phi^{(\text{diff})}$), for a total of 15 degrees of freedom [versus $3N(3N+1)/2$ for distinguishable particles],

$$\begin{aligned} \hat{\mathcal{H}} = & \hat{K} + \frac{1}{2} \sum_{i\alpha\beta} (\hat{R}_{i\alpha} - \mathcal{R}_\alpha) \Phi_{\alpha\beta}^{(\text{same})} (\hat{R}_{i\beta} - \mathcal{R}_\beta) \\ & + \frac{1}{2} \sum_{\substack{i \neq j \\ \alpha\beta}} (\hat{R}_{i\alpha} - \mathcal{R}_\alpha) \Phi_{\alpha\beta}^{(\text{diff})} (\hat{R}_{j\beta} - \mathcal{R}_\beta). \end{aligned} \quad (15)$$

In this section, we explicitly encode with $i\alpha$ and $j\beta$ the index of the particle and the Cartesian coordinate, respectively. While $\Phi^{(\text{same})}$ is Hermitian, the same is not necessarily valid for $\Phi^{(\text{diff})}$. Notably, $\Phi^{(\text{diff})}$ is a coupling term between electrons. Therefore, the SCHA differs from other mean-field approaches like Hartree-Fock (HF) and DFT because the auxiliary Hamiltonian already accounts for a certain degree of electronic correlation. Equation (15) can be diagonalized exactly, and the resulting eigenstates cannot be separated into products of independent particle wave functions. However, it is challenging to disentangle the bosonic-fermionic statistics of the states. Also, we prove in Appendix H that $\Phi^{(\text{diff})} \rightarrow 0$ in the $N \rightarrow \infty$ limit. For these reasons, we further restrict the problem to noninteracting fermions in a harmonic potential, neglecting the correlations in the harmonic Hamiltonian. We show in Sec. V that this correlation can be restored on the density matrix *a posteriori* with a linear deformation of the curved manifold.

Without $\Phi^{(\text{diff})}$, the Hamiltonian Eq. (15) is noninteracting, and it can be diagonalized by identifying the three eigenvectors of $\Phi^{(\text{same})}$. Introducing the standard creation a_μ^\dagger and annihilation a_μ operators of the harmonic oscillator along each of these three modes μ , we have

$$\hat{\mathcal{H}} = \frac{1}{2} \sum_{\mu=1}^3 \omega_\mu (2a_\mu^\dagger a_\mu + 1). \quad (16)$$

A three-state vector $|n_1 n_2 n_3\rangle$ uniquely identifies the state, encoding the excited level on each polarization mode. The overall fermionic wave function is obtained by constructing a Slater determinant with these orbitals, similar to standard mean-field approaches, constraining the rules for populating states given by the Pauli exclusion principle, as shown in Fig. 2. Notably, the electronic oscillator's ground state is a Gaussian only up to two electrons. Excited orbitals are occupied for more than two electrons, and the many-body ground

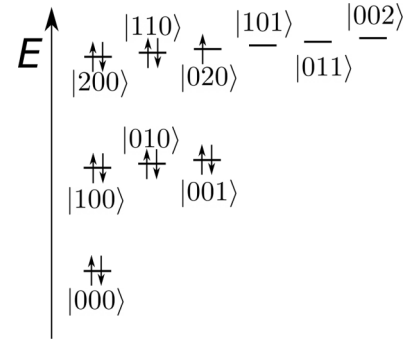


FIG. 2. States of the noninteracting fermionic oscillator in 3D. Electrons populate the excited states of the harmonic Hamiltonian $|n_1 n_2 n_3\rangle$ according to the Fermi-Dirac statistics and the spin multiplicity.

state acquires the nodes typical of fermionic wave functions. In the variational frameworks of the SCHA, the nodes are not fixed and depend on the parameters \mathcal{R} and $\Phi^{(\text{same})}$. Thus, the electronic SCHA does not rely on any fixed-node approximation necessary for applying path-integral Monte Carlo or diffusion Monte Carlo to fermions [38].

IV. SCHA IN A CURVED MANIFOLD

While the SCHA has been shown to work exceptionally well in investigating the anharmonic quantum motion of nuclei in crystals [5–8,39–43], its limitations arise from constraining the density matrix to solve a harmonic Hamiltonian. In the case of distinguishable particles like atomic nuclei, the real-space expression $\rho(\mathbf{R}, \mathbf{R}') = \langle \mathbf{R}' | \hat{\rho} | \mathbf{R} \rangle$ of the SCHA density matrix is a Gaussian [26]. A Gaussian distribution fits the fluctuations of nuclei vibrating in a solid lattice. Still, it fails to capture ionic diffusion, molecular rotations, quantum tunneling, or cases where the wave function extends beyond the interatomic distance. These cases are significant for electrons, where their lighter mass by at least three orders of magnitude causes a quantum delocalization across a broad region of space, often encompassing many atoms by forming chemical bonds.

Overcoming the SCHA has proven more difficult than expected: Any attempt to augment the auxiliary harmonic Hamiltonian with extra anharmonic terms prevents the analytical solvability of the auxiliary system. This results in the loss of many of the technique's advantages, among which is the analytical expression for the entropy [37], fundamental to computing the phase diagram of materials. An alternative strategy consists of maintaining the auxiliary Hamiltonian harmonic, introducing a nonlinear change of variables between the Cartesian positions \mathbf{R} and the degrees of freedom of the auxiliary Hamiltonian \mathbf{q} [35,36]. This defines the auxiliary harmonic Hamiltonian on a manifold whose curvature is parameterized by the invertible transformation $\mathbf{R}(\mathbf{q})$. This approach is similar to what was proposed by Gygi [44] and Hamann [45] in the context of density functional theory to improve the basis set variationally. In this case, however, we utilize the nonlinear change of variable to capture both deviations from Gaussianity and many-body correlation effects.

The expression of physical observables must be the same on the curved manifold \mathbf{q} and real space \mathbf{R} ,

$$\langle \hat{O} \rangle_{\tilde{\rho}} = \text{Tr}[\tilde{\rho} \hat{O}], \quad (17)$$

$$\langle \hat{O} \rangle_{\tilde{\rho}} = \int d\mathbf{R} d\mathbf{R}' O(\mathbf{R}, \mathbf{R}') \rho(\mathbf{R}', \mathbf{R}), \quad (18)$$

$$\langle \hat{O} \rangle_{\tilde{\rho}} = \int d\mathbf{q} d\mathbf{q}' \tilde{O}(\mathbf{q}, \mathbf{q}') \tilde{\rho}(\mathbf{q}', \mathbf{q}), \quad (19)$$

where we indicated with a tilde $\tilde{\cdot}$ quantities evaluated on the curved manifold. By imposing that the density matrix conserves the probability density in the position space, we can prove that diagonal observables like $O(\mathbf{R}, \mathbf{R}') \propto \delta(\mathbf{R} - \mathbf{R}')$ transform like

$$\tilde{O}(\hat{\mathbf{q}}) = O(\mathbf{R}(\hat{\mathbf{q}})) \quad (20)$$

and the density matrix as

$$\tilde{\rho}(\mathbf{q}, \mathbf{q}') = \rho(\mathbf{R}(\mathbf{q}), \mathbf{R}(\mathbf{q}')) \sqrt{\det \mathcal{J}(\mathbf{R}(\mathbf{q})) \det \mathcal{J}(\mathbf{R}(\mathbf{q}'))}, \quad (21)$$

where \mathcal{J} is the Jacobian of the transformation

$$\mathcal{J}_{ab}(\mathbf{q}) = \frac{\partial R_a}{\partial q_b}. \quad (22)$$

The transformation of Eq. (21) preserves the probability density in the position space. A more complex transformation for the observable needs to be derived if $O(\mathbf{R}, \mathbf{R}')$ is not diagonal in the position basis, e.g., for the kinetic energy.

We have, thus, two different density matrices: $\tilde{\rho}(\mathbf{q}, \mathbf{q}')$, defined on the curved manifold and solution of the auxiliary harmonic Hamiltonian Eq. (8), and $\rho(\mathbf{R}, \mathbf{R}')$, the transformed density matrix in Cartesian space. Notably, $\rho(\mathbf{R}, \mathbf{R}')$ is the real density matrix of our solution, which depends on the parameters that define the auxiliary Hamiltonian (\mathcal{R} and Φ) and on the curvature of the manifold, defined by the metric tensor g_{ab} as

$$g_{ab}(\mathbf{q}) = \sum_c \frac{\partial R_c}{\partial q_a} \frac{\partial R_c}{\partial q_b} = \sum_c \mathcal{J}_{ca} \mathcal{J}_{cb}, \quad (23)$$

$$\sqrt{\det \mathbf{g}} = \sqrt{\det \mathcal{J}(\mathbf{R}(\mathbf{q})) \det \mathcal{J}(\mathbf{R}(\mathbf{q}'))}. \quad (24)$$

This allows us to increase the degrees of freedom of the ansatz density matrix thanks to the metric tensor \mathbf{g} while keeping the auxiliary system harmonic. This approach maintains all the advantages of the original SCHA. In fact, all observables can be evaluated in the curved manifold, where the density matrix is harmonic. For convenience, we name the method nonlinear SCHA (or NLSCHA) throughout the text.

Free energy in the curved manifold

The target of the SCHA as a variational theory is to optimize the trial density matrix $\rho(\mathbf{R}, \mathbf{R}')$ to minimize the free energy. The free energy $F[\tilde{\rho}]$ is a functional of the density matrix comprising the internal energy $U[\tilde{\rho}] = \langle \hat{H} \rangle_{\tilde{\rho}}$ and the entropy $S[\tilde{\rho}]$. The most complex part of the free energy is the entropy, for which the expression is challenging to be converged numerically,

$$S[\tilde{\rho}] = -k_b \langle \ln \tilde{\rho} \rangle. \quad (25)$$

In fact, due to the logarithm, even very high-energy states are important in the average of Eq. (25), and it cannot be computed as an average on the sampled configurations. However, if the nonlinear transformation $\mathbf{R}(\mathbf{q})$ is bijective on $\mathbb{R}^{3N} \rightarrow \mathbb{R}^{3N}$, the entropy of $\rho(\mathbf{R}, \mathbf{R}')$ coincides with the one of $\tilde{\rho}(\mathbf{q}, \mathbf{q}')$ [35]. Thus, $S[\tilde{\rho}]$ can be evaluated conveniently on the curved manifold, where the density matrix solves the auxiliary harmonic Hamiltonian $\hat{\mathcal{H}}[\mathcal{R}, \Phi]$ for which the energy levels are known analytically. In the case of distinguishable particles, like atomic nuclei, the entropy has a simple expression depending on the frequencies ω_μ of the auxiliary Hamiltonian [46],

$$\sum_b \frac{\Phi_{ab}}{\sqrt{m_a m_b}} e_\mu^a = \omega_\mu^2 e_\mu^b, \quad (26)$$

$$S[\rho(\mathbf{R}, \mathbf{R}')] = \frac{k_b}{2} \sum_\mu \left[\frac{\beta \hbar \omega_\mu}{\tanh \frac{\beta \hbar \omega_\mu}{2}} - 2 \log \left(\sinh \frac{\beta \hbar \omega_\mu}{2} \right) \right], \quad (27)$$

where m_a is the mass of the a particle and $\beta = (k_b T)^{-1}$ is the Boltzmann factor. In the more general case of mixtures between distinguishable and indistinguishable particles, the energy levels of the auxiliary harmonic Hamiltonian must be appropriately populated, accounting for the specific Fermi-Dirac or Bose-Einstein statistics and the particles' spin. Thankfully, the auxiliary harmonic Hamiltonian can be decoupled between each different kind of particle, separating the indistinguishable fermions (electrons) and the distinguishable ions. The electronic entropy is thus evaluated in a noninteracting harmonic system, as discussed in Sec. III. Let f_i be the occupation number of the i state of the auxiliary harmonic Hamiltonian, the resulting electronic entropy is

$$S[\tilde{\rho}] = -k_b \sum_i [f_i \ln(f_i) + (1 - f_i) \ln(1 - f_i)]. \quad (28)$$

The internal energy $U[\tilde{\rho}]$ accounts for kinetic and potential energy. The potential energy is a diagonal operator in the position of the particles $\hat{\mathbf{R}}$, and can be evaluated on the curved manifold following Eq. (20):

$$\langle \hat{V} \rangle_{\tilde{\rho}} = \int d\mathbf{q} V(\mathbf{R}(\mathbf{q})) \tilde{\rho}(\mathbf{q}). \quad (29)$$

Equation (29) is the average interacting potential in the curved manifold. The practical computation of $\langle \hat{V} \rangle_{\tilde{\rho}}$ through Monte Carlo consists of generating a random ensemble of particles on the positions \mathbf{q} in the manifold exploiting the Gaussian shape of the auxiliary probability distribution given by $\tilde{\rho}(\mathbf{q})$, transforming the coordinates from the manifold \mathbf{q} to Cartesian space \mathbf{R} with the transformation $\mathbf{R}(\mathbf{q})$, and averaging the values of the potential $V(\mathbf{R})$ in Cartesian space. Compared with other approaches like PIMD or QMC, the generation of the ensemble is faster and does not require any Metropolis algorithm or thermalization process, as it occurs on a Gaussian distribution in the curved manifold.

The kinetic energy is more complex as it is nondiagonal in position space. Its expression $\tilde{K}(\mathbf{q}, \mathbf{q}')$ on the curved manifold is derived by transforming the integration variables from

Cartesian space

$$\langle K \rangle_{\tilde{\rho}} = - \sum_i \frac{\hbar^2}{2m_i} \int d\mathbf{R} d\mathbf{R}' \delta(\mathbf{R} - \mathbf{R}') \frac{\partial^2}{\partial R'_i} \rho(\mathbf{R}, \mathbf{R}'). \quad (30)$$

In Appendix A, we derive the explicit expression for the kinetic energy in the curved manifold. In practice, it can be computed as

$$\begin{aligned} \langle K \rangle_{\tilde{\rho}} = & \sum_i \frac{1}{2m_i} \left[\sum_{ab} \int d\mathbf{q} d\mathbf{q}' \gamma_i^{ab}(\mathbf{q}) \frac{\partial^2 \tilde{\rho}(\mathbf{q}, \mathbf{q}')}{\partial q_a \partial q'_b} \delta(\mathbf{q} - \mathbf{q}') \right. \\ & \left. + \frac{1}{2} \sum_a \left\langle \beta_i^a(\mathbf{q}) \frac{d \ln \tilde{\rho}(\mathbf{q})}{dq_a} \right\rangle_{\tilde{\rho}} + \langle \alpha_i(\mathbf{q}) \rangle_{\tilde{\rho}} \right]. \quad (31) \end{aligned}$$

Here, $\tilde{\rho}(\mathbf{q})$ is the probability density in the auxiliary flat space: the diagonal elements in real space of the density matrix, i.e., $\tilde{\rho}(\mathbf{q}) = \tilde{\rho}(\mathbf{q}, \mathbf{q})$. $\langle f(\mathbf{q}) \rangle_{\tilde{\rho}}$ indicates an expected value of the diagonal operator \hat{f} in the positions $\hat{\mathbf{q}}$ on the auxiliary flat space, and $\alpha(\mathbf{q})$, $\beta(\mathbf{q})$, and $\gamma(\mathbf{q})$ are operators that depend only the local curvature of the manifold (the nonlinear transformation). Their explicit expression is reported in Appendix A. Indeed, in the case of a flat manifold, $\alpha = \beta = 0$, $\gamma_i^{ab}(\mathbf{q}) = \delta_{ai} \delta_{bi}$, and we recover the kinetic energy of a flat space [Eq. (30)]. The complete kinetic energy on the curved manifold comprises a term (the first one) that represents the kinetic energy of the auxiliary flat space, where the masses are contracted with a vector that depends on the metric tensor: $\gamma_i^{ab}(\mathbf{q})$. This can be interpreted as a change in the mass of the particles due to the nonlinear transformations (see Appendix I for more details). Besides this term, the kinetic energy also includes a correction that is diagonal in position space, thus behaving like an effective potential energy. This extra term encodes the correlations introduced by the curved manifold, similarly to the exchange-correlation potential defined within DFT. Notably, Eq. (31) differs from the equation presented in Ref. [35], as here we do not perform any hypothesis on the Gaussian shape of the auxiliary density matrix $\tilde{\rho}(\mathbf{q}, \mathbf{q}')$. This way, Eq. (31) can also be applied to electrons, excited states, or any nonharmonic auxiliary Hamiltonian.

V. NEURAL NETWORK PARAMETERIZATION OF THE CURVED MANIFOLD

In Sec. IV A, we derived the free energy expression on the curved manifold, where the density matrix is the ground state of an auxiliary harmonic Hamiltonian, and the free energy can be computed efficiently. To solve the quantum problem beyond the SCHA, we need to minimize the free energy by optimizing the parameters defining the auxiliary Hamiltonian and the manifold. The auxiliary Hamiltonian depends on the centroids \mathcal{R} and the effective force constants Φ . Reference [36] devised a specific manifold from a cartographic representation of spherical coordinates to describe molecular rotations. While it provides an excellent solution to enhance the SCHA by enabling molecules to rotate thanks to the manifold's curvature, it is too restrictive to capture the complex features of the electronic wave function. Here, we introduce a general transformation that can be systematically improved to

parameterize any possible curved manifold, thus significantly enhancing the variational space of density matrices.

The manifold is defined through the transformation $\mathbf{R}(\mathbf{q})$. This function maps $\mathbb{R}^N \rightarrow \mathbb{R}^N$, is bijective, continuous, and differentiable. We define the Gaussian stretchers as the class of manifolds parameterized by the transformation

$$\mathbf{R}(\mathbf{q}) = \mathcal{Q} + \tilde{\mathbf{q}} \left[1 + (e^\zeta - 1) \exp \left(-\frac{1}{2} \sum_{cd} \tilde{q}_c \Gamma_{cd} \tilde{q}_d \right) \right], \quad (32)$$

where

$$\tilde{\mathbf{q}} = \mathbf{q} - \mathcal{Q} \quad (33)$$

and Γ is positive definite. As the name suggests, this transformation stretches the space, deforming the density matrix in the position identified by \mathcal{Q} along the principal axis of the Γ^{-1} covariance matrix, with intensity ζ . Outside the central position \mathcal{Q} and the range of the covariance matrix Γ^{-1} , the exponential of Eq. (32) becomes small and the manifold flat ($\mathbf{R} = \mathbf{q}$). Therefore, the Gaussian stretcher only warps the space inside the Gaussian by stretching the density matrix with a positive curvature if $\zeta > 0$ (reducing the probability distribution at the center and accumulating it on the edges) or compressing it with a negative curvature if $\zeta < 0$ (reducing the probability on the edges to accumulate it on the center). The parameterization of the Gaussian stretcher is only valid for $\zeta < \zeta_c \approx 1.1$, above which the transformation defined in Eq. (32) is no longer bijective.

Each Gaussian stretcher has the same number of degrees of freedom as the original SCHA algorithm, plus ζ (the Γ symmetric two-rank tensor is equivalent to the auxiliary force constant matrix Φ , and the \mathcal{Q} vector has the same length as the centroids \mathcal{R}). Moreover, multiple Gaussian stretchers can be stacked together in layers as in neural networks:

$$\mathbf{q}_{n+1} = \mathcal{Q}_n + \tilde{\mathbf{q}}_n \left[1 + (e^{\zeta_n} - 1) \exp \left(-\frac{1}{2} \sum_{cd} \tilde{q}_{nc} \Gamma_{cd}^{(n)} \tilde{q}_{nd} \right) \right], \quad (34)$$

$$\mathbf{q}_0 = \mathbf{R}, \quad \mathcal{Q}_0 = \mathcal{R}. \quad (35)$$

Each layer introduces three parameters: ζ_n , \mathcal{Q}_n , and $\Gamma^{(n)}$. In the case of concatenated Gaussian stretchers, the resulting manifold exhibits a complex curvature that is fine-tuned in any position by a layer with a nearby centroid \mathcal{Q}_n . The transformation resembles a neural network where the activation function for each neuron is given by Eq. (32). In Appendix C, we evaluate the metric tensor and the expression of the kinetic energy. The manifold parameters are optimized through a standard ADAM algorithm [47], where the gradient of the quantum free energy is evaluated with backpropagation, a technique commonly employed in neural networks.

Equivariance under specific symmetries can be enforced on the Gaussian stretcher manifold. This is pivotal for ensuring compliance with fundamental laws of physics, like the acoustic sum rule (equivariance with respect to global translations) and the permutation of indistinguishable particles. Moreover, this characteristic extends to any symmetry operation commuting with the original Hamiltonian as the crystallographic symmetry group. In particular, let \mathcal{S} be the symmetry operator;

the manifold is equivariant under S if

$$\mathbf{R}(\mathbf{S}\mathbf{q}) = \mathbf{S}\mathbf{R}(\mathbf{q}). \quad (36)$$

This condition can be applied layer by layer on the Gaussian stretcher as

$$\mathbf{S}\mathbf{q}_{n+1} = \mathbf{Q}_n + (\mathbf{S}\mathbf{q}_n - \mathbf{Q}_n)f_n(\mathbf{S}\mathbf{q}_n), \quad (37)$$

where

$$f_n(\mathbf{q}) = 1 + (e^{\xi_n} - 1) \exp\left[-\frac{1}{2}(\mathbf{q} - \mathbf{Q}_n)\mathbf{\Gamma}^{(n)}(\mathbf{q} - \mathbf{Q}_n)\right], \quad (38)$$

resulting in the following conditions that restrict the degrees of freedom of \mathbf{Q}_n and $\mathbf{\Gamma}^{(n)}$:

$$\mathbf{S}\mathbf{Q}_n = \mathbf{Q}_n, \quad f_n(\mathbf{q}) = f_n(\mathbf{S}\mathbf{q}). \quad (39)$$

The constraint on \mathbf{Q}_n enforces the centroid of the network to be a high-symmetry point of the crystal (Wyckoff position), while the constraint on f_n enforces the symmetry on the $\mathbf{\Gamma}$ matrix as

$$\mathbf{S}^\dagger \mathbf{\Gamma}_n \mathbf{S} = \mathbf{\Gamma}_n. \quad (40)$$

These conditions coincide with the ones complied by \mathcal{R} and Φ in the standard harmonic Hamiltonian to preserve the symmetry of the density matrix. Therefore, imposing symmetries on the Gaussian stretchers is equivalent to constraining symmetries of positions and dynamical matrices.

Imposing crystal symmetries by constraining each layer of the nonlinear transformation suppresses the size of the variational space spanned by the Gaussian stretchers. Therefore, we resort to this scheme only for symmetries that preserve the physical meaning of the resulting density matrix, like the exchange between indistinguishable particles and the total translational symmetry. Any further symmetry constraint is imposed *a posteriori* throughout a Lagrange multiplier like

$$\mathcal{L} = \sum_{i=1}^N \|\mathbf{S}\mathbf{q}^{(n)}(\mathbf{q}_i^{(1)}) - \mathbf{q}^{(n)}(\mathbf{S}\mathbf{q}_i^{(1)})\|^2. \quad (41)$$

By including \mathcal{L} in the cost function to minimize the free energy, symmetries can be imposed *a posteriori* only between the first and the last layer of the Gaussian stretcher, thus increasing the variational degrees of freedom of the solution.

Since multiple transformations can be stacked layer by layer, it is also convenient to employ a linear transformation. In normal circumstances, this is redundant with the SCHA itself. the SCHA centroids \mathcal{R} and auxiliary force constant matrix Φ act as a linear transformation on an uncorrelated ensemble of normalized Gaussians. The transformation reads as

$$\mathbf{q}^{(1)} = \mathcal{R} + \mathbf{L}\mathbf{q}^{(0)}, \quad (42)$$

$$\mathbf{L}\mathbf{L}^\dagger = \mathbf{\Upsilon}^{-1}, \quad (43)$$

where $\mathbf{\Upsilon}^{-1}$ is the covariance matrix of the SCHA distribution,

$$\Upsilon_{ab} = \sqrt{m_a m_b} \sum_{\mu} \frac{2\omega_{\mu}}{2n_{\mu} + 1} e_{\mu}^a e_{\mu}^b. \quad (44)$$

Using a first (or last) linear layer is a feature familiar to standard machine learning neural networks. Besides the

resemblance with machine learning, using the linear transformation matrix \mathbf{L} as a variational degree of freedom instead of the auxiliary force constant matrix Φ ensures the continuity of the transformation of the ensemble even in the presence of mode crossing and degeneracies if the original ensemble $q^{(0)}$ is kept fixed. This is essential to have a smooth and differentiable energy landscape (cost function). Moreover, in the case of indistinguishable particles, this linear transformation can restore the harmonic correlation lost by considering $\Phi^{(\text{diff})} \rightarrow 0$ in Eq. (15) to simplify the population of fermionic degrees of freedom. Without loss of generality, \mathbf{L} can be written as

$$L_{ab} = \frac{1}{\sqrt{m_a}} \sum_{\mu} \sqrt{\frac{2\omega_{\mu}}{2n_{\mu} + 1}} e_{\mu}^a e_{\mu}^b, \quad (45)$$

thus satisfying Eq. (43).

VI. FERMIONIC STATES

In this section, we present the constraints that the Gaussian stretcher has to satisfy to preserve the fermionic (or bosonic) character of the auxiliary density matrix $\tilde{\rho}(\mathbf{q}, \mathbf{q}')$ when it is transformed into the real-space density matrix $\hat{\rho}(\mathbf{R}, \mathbf{R}')$. This is the fundamental requirement to treat indistinguishable particles within the formalism presented in this work. In Sec. III, we introduced the noninteracting fermionic wave function solving the auxiliary harmonic Hamiltonian. The manifold curvature mixes the degrees of freedom of different particles, leading to a correlated state. To preserve the antisymmetric characteristic of the underlying density matrix across the nonlinear transformation $\mathbf{R}(\mathbf{q})$, the manifold should preserve the exchange operation between different particles. It is trivial to show that $\rho(\mathbf{R}, \mathbf{R}')$ is an antisymmetric if $\tilde{\rho}(\mathbf{q}, \mathbf{q}')$ is antisymmetric and $\mathbf{R}(\mathbf{q})$ satisfies the condition

$$\mathbf{R}_1(\mathbf{q}_1, \mathbf{q}_2, \dots) = \mathbf{R}_2(\mathbf{q}_2, \mathbf{q}_1, \dots), \quad (46)$$

where with \mathbf{q}_1 (and \mathbf{R}_1) we indicate the components of \mathbf{q} (and \mathbf{R}) on particle 1, 1 and 2 are any two fermions, and the \dots indicates that the \mathbf{q} vectors of the other particles have not been exchanged. Equation (46) is fulfilled by a transformation of the kind

$$\mathbf{R}(\mathbf{q}) = \begin{cases} \mathbf{R}_1(\mathbf{q}) = f(\mathbf{q})\mathbf{q}_1 \\ \mathbf{R}_2(\mathbf{q}) = f(\mathbf{q})\mathbf{q}_2 \\ \vdots \end{cases} \quad (47)$$

where $f(\mathbf{q})$ is a generic symmetric function with respect to particle changes. The Gaussian stretchers meet this condition if \mathcal{Q} is the same for all electrons and most of $\mathbf{\Gamma}$ parameters are constrained so that

$$\mathcal{Q}_{(i,\alpha)} = \mathcal{Q}_{(j,\alpha)}, \quad (48)$$

$$\Gamma_{(i,\alpha)(i,\beta)} = \Gamma_{(j,\alpha)(j,\beta)}, \quad (49)$$

$$\Gamma_{(i,\alpha)(j,\beta)} = \Gamma_{(j,\alpha)(k,\beta)}, \quad (50)$$

where the index (i, α) indicates atom i and Cartesian coordinate α , for any choice of the atoms i, j, k . These are the same conditions as the \mathcal{R} and Φ matrix of the auxiliary harmonic

Hamiltonian for fermions we derived in Sec. III. Thanks to the cross-diagonal terms of Γ Eq. (50), the transformation can still couple different particles, introducing correlations at each layer of the nonlinear transformation defining the curved manifold. Another consequence of the exchange symmetry of the manifold is the combinatorial number of constraints imposed on the degrees of freedom that counterbalance their growth when the number of particles increases: above two electrons, the degrees of freedom do not increase with the number of particles on each layer of the Gaussian stretcher.

VII. APPLICATIONS AND EXAMPLES

In this section, we illustrate some applications to show how the NLSCHA with the Gaussian stretcher systematically improves the SCHA result. We start with a one-particle problem, where deviations from the Gaussian ground state are significant, like in the case of a profound double-well potential for both the ground (Sec. VII A) and excited state (Sec. VII B), a Coulomb potential (hydrogen atom) in Sec. VII C, and finally for interacting electrons in the dissociation of the H_2 molecule (Sec. VII D). We discuss and display also the capabilities of the method to capture excited states by exciting electrons in the auxiliary harmonic Hamiltonian for the 1D double potential and the H_2 dissociation in the antibonding triplet electronic configuration.

A. Double-well potential

The double-well potential is a classic example of a strongly anharmonic system. It is also a good prototype for testing quantum tunneling, a regime in which the SCHA is known to fail [35]. We represent the double-well potential of a 1D particle as

$$V(R) = aR^4 - bR^2. \quad (51)$$

For values of $b > 0$, the harmonic approximation presents an imaginary frequency in $R = 0$, and the system becomes highly anharmonic. Here, we compare the performance of the SCHA and the NLSCHA, with the multilayer Gaussian stretcher manifold introduced in this work, against the exact (numerical) solution.

The double-well potential has three regimes as b varies. When $b^3 \gg \frac{2a^2}{m^2}$ (in Hartree atomic units), the two wells are far apart, the energy barrier is too high, and the solution localizes into one of the minima (broken symmetry solution). The SCHA describes this regime well if we allow the Gaussian solution to break the inversion symmetry and localize in one of the two minima, similarly to how unrestricted Hartree-Fock correctly captures molecular dissociation by breaking the spin symmetry. The opposite occurs when the barrier is small ($b^3 \ll \frac{2a^2}{m^2}$); in this case, quantum/thermal fluctuations entirely overcome the energy barrier, and the probability density is located in the saddle point $R = 0$. This phase is also well captured by the SCHA. However, when the energy barrier is comparable with the quantum/thermal fluctuations ($b^3 \approx \frac{2a^2}{m^2}$), tunneling and thermal hopping between the two minima occur on the same timescale as the lattice vibrations. This is the regime where we expect the SCHA to fail. Figure 3(a) reports the comparison between the ground state electronic densities

of the SCHA and NLSCHA applying a different number of layers with the exact (numerical) solution. This test is performed for $b = 6E_h/a_0^2$, constraining the solution to satisfy the inversion symmetry (restricted SCHA). This potential is deep in the regime of quantum tunneling between the two minima.

Applying multiple layers of the Gaussian stretcher improves the quality of the results. One layer of Gaussian stretcher is not enough to split the wave function into the two peaks around the minima of the potential, as it is energetically more convenient to reproduce the distribution's tails due to the fast rise of the quartic potential. However, adding extra layers enables the separation of the peaks. In this case, as with any 1D system, it is easy to show that the NLSCHA converges to the exact solution. In fact, we can prove that a nonlinear transformation $q(r)$ that maps a Gaussian into any possible ground-state wave function always exists. In particular, $q(r)$ solves the differential equation

$$\frac{dq}{dr} = \left(\frac{\psi(r)}{\phi[q(r)]} \right)^2, \quad (52)$$

which is possible if $\psi(r)$ and $\phi[q(r)]$ share the same phase for all values of r (apart from a global phase factor). In 1D with the time-reversal symmetry, the ground state can always be chosen as a real positive function, and the square in Eq. (52) becomes the ratio between probability densities. Then, the Bernier's theorem for optimal transport [48] ensures that always exists a solution for Eq. (52), and the $q(r)$ can be expressed as the gradient of a convex scalar function. In Fig. 3(c), we solved Eq. (52) for the map between the exact ground state of the double-well potential and the normal distribution, and compared to the neural network approximations of $q(r)$ obtained by variationally minimizing the ground-state energy.

In Fig. 3(d), we show the error on the ground-state energy as a function of the number of layers and the b parameter. We impose inversion symmetry, preventing the SCHA from localizing the Gaussian around one of the two minima, which is energetically more favorable for these values of b .

The NLSCHA's ability to reproduce the splitting between wave function peaks and to minimize the error in the nonbroken-symmetry solution has significant repercussions for the simulation of quantum tunneling, which was previously impossible within the standard SCHA or other harmonic-based approaches.

B. Excited states

The formalism introduced in this work also allows populating excited states in the auxiliary harmonic Hamiltonian, enabling the simulation of wave functions with nodes to optimize the total energy. This approach has already been successfully employed in variational Monte Carlo to study optical gaps in solid-state systems [49,50], and transfers naturally to SCHA. As a simple benchmark, we apply the NLSCHA to unveil the wave function of the first excited state of the double-well potential in the strongly anharmonic regime. The comparison with the exact diagonalization is reported in Fig. 4. The wave function agrees exceptionally well with the exact result already with just two layers of Gaussian stretcher

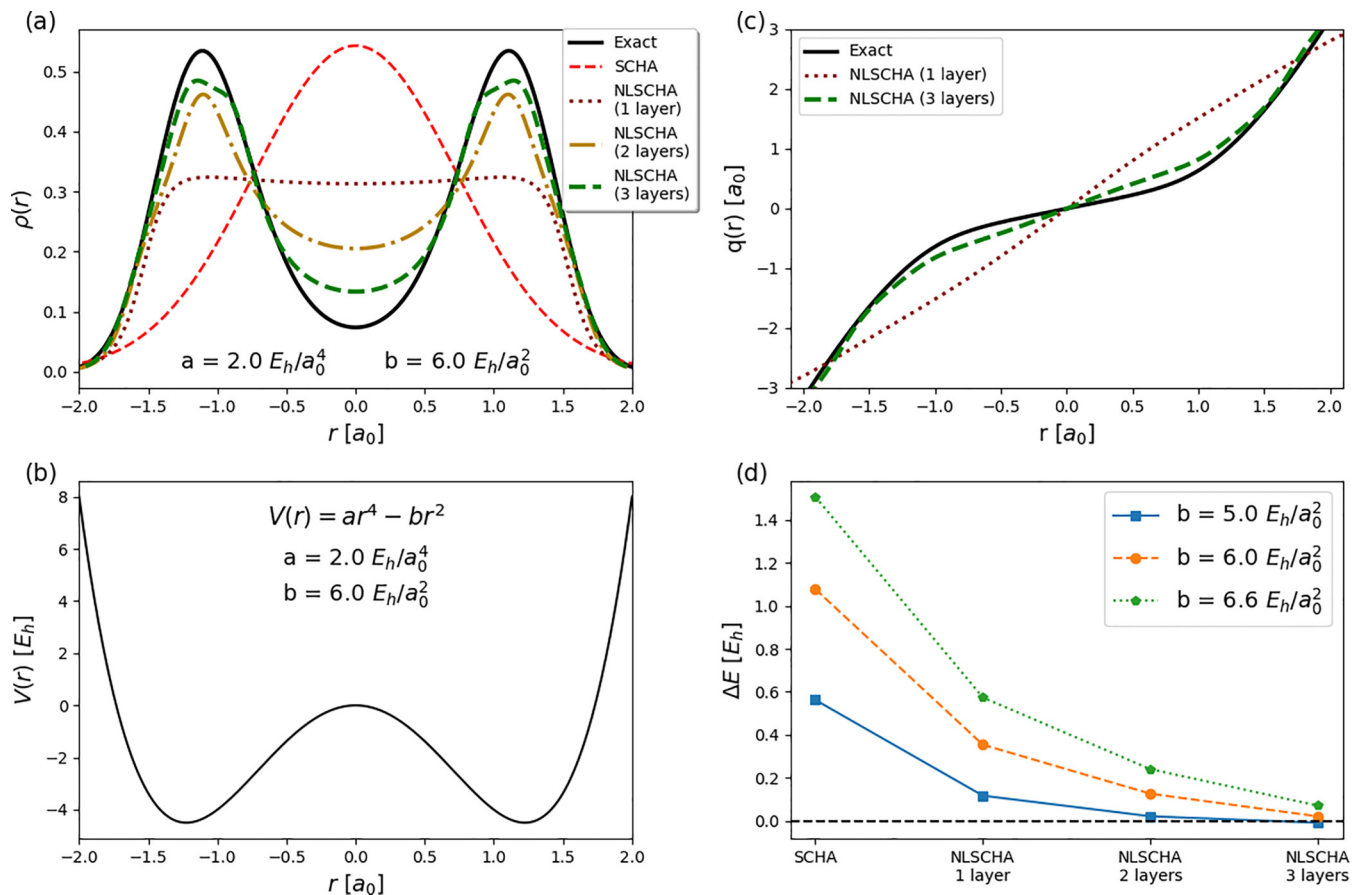


FIG. 3. Comparison between the SCHA, NLSCHA, and exact (numerical) solution for a double-well potential. (a) Probability density of the ground state wave function at $b = 6.0 E_h/a_0^2$ by varying the number of layers of the nonlinear transformation. (b) Potential energy landscape of the double well. (c) Comparison between the nonlinear transformations $q(r)$ encoded by the neural network and the one that exactly maps the SCHA Gaussian wave function to the true ground state of the double-well potential. (d) The energy error at several values of b and varying the number of layers. All the values are in the deep quantum regime $b > 2E_h/a_0^2$. The wave function is constrained to keep the inversion symmetry (inhibiting the transition into the localized state). The exact ground state of the double-well potential has been evaluated numerically with the implicitly restarted Lanczos algorithm [51] as implemented in SciPy [52].

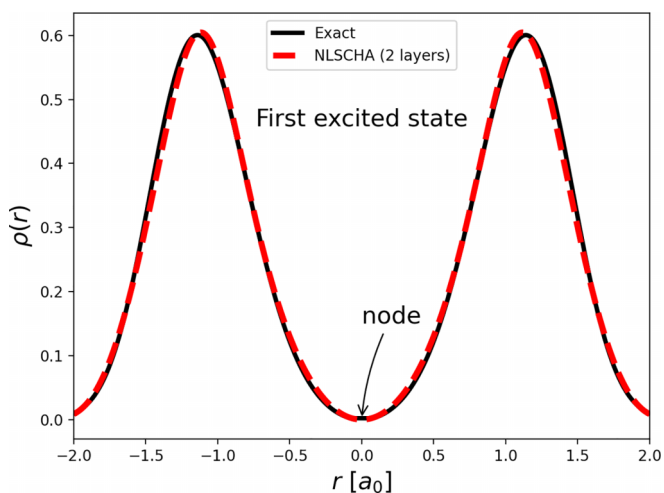


FIG. 4. Comparison between the exact and NLSCHA (two layers) excited state density. The potential is a double well of the form $V(r) = ar^4 - br^2$, where $a = 2E_h/a_0^4$ and $b = -6E_h/a_0^2$, as in Fig. 3(c). We highlight the position of the node of the wave function in the origin.

manifold, outmatching even the precision in the ground state thanks to the presence of a node in the original wave function.

C. Hydrogen atom: the electron-ion cusp

While the NLSCHA performs exceptionally well on ionic systems and solves many issues of the SCHA, like molecular rotations, delocalization, and tunneling, its applicability in the electronic problem has to be proven. Electronic correlated wave functions are strongly non-Gaussian, with cusps in the overlap between different electrons and electron-ion due to the divergence of the Coulomb potential [53]. While some of the problems could be accounted for with the use of pseudo-potentials, like done in mean-field approaches with a plane-wave basis set, in this section, we benchmark the expressibility of the transformed wave function on the hydrogen atom, a prototypical system with an electron-ion cusp that is highly challenging to all methods that do not explicitly account for the cusp in the wave function (as it occurs in plane-wave codes).

We constrain the wave function for the hydrogen atom potential with a fully rotational symmetry group centered

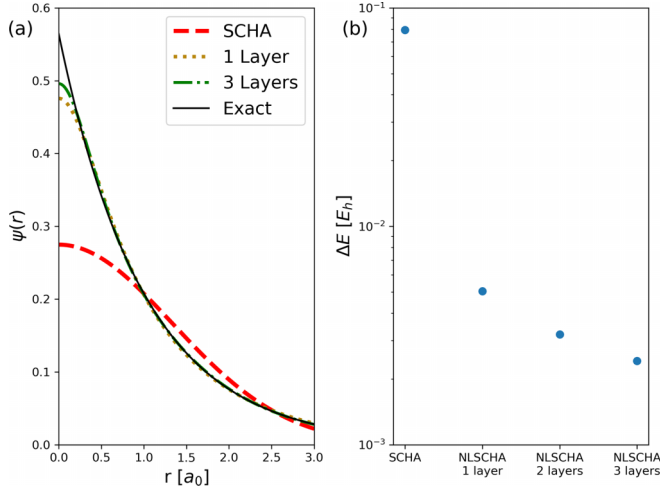


FIG. 5. Hydrogen atom, comparison between the SCHA, nonlinear SCHA (NLSCHA), and exact (analytical) solution. (a) Comparison between the radial wave function of the $1s$ orbital around the electron-ion cusp in the origin. (b) Error on the ground-state energy compared to the analytical result (log scale).

around the origin and the potential $V(r) = -1/r$, as discussed in Sec. V. We solve the system by progressively increasing the number of layers of the Gaussian stretchers. Thanks to the spherical symmetry constraints, each layer has only 2 degrees of freedom (ζ is free, Γ is proportional to the identity matrix, and $\mathcal{Q} = 0$). The final wave functions and energies are reported in Fig. 5. Already with one layer, the long-range behavior of the wave function is very good ($r > 0.5a_0$), while increasing the number of layers further improves the $r \rightarrow 0$ limit (the cusp). In contrast to the 1D double-well potential—where adding additional layers leads to an exponential convergence toward the exact solution—we observe only more gradual improvement beyond the first transformation layer in Fig. 5(b). This limited improvement rate is also evident in the corresponding wave-function profiles, Fig. 5(a), where the overall shape changes only slightly between one and three layers. The underlying reason lies in the functional form of the Gaussian stretcher: its transformed wave function is inherently C_∞ , and thus a considerable number of layers would be required to reproduce the derivative discontinuity introduced by the divergent Coulomb potential. Nevertheless, the hydrogen atom’s ground state can be expressed as a real wave function; thus, there exists a mathematical transformation that solves Eq. (52), and the NLSCHA is exact in this case. However, achieving chemical accuracy requires overcoming this limitation of Gaussian stretchers, likely through *ad hoc* nonlinear transformations that introduce controlled derivative discontinuities.

D. H_2 dissociation

We have tested the NLSCHA on problems with only one electron, where mean-field approaches are exact. Here, we tackle the H_2 molecule under dissociation, a fully interacting many-body system, and one of the most challenging problems for mean-field approaches like Hartree-Fock, DFT, or many-body diagrammatic expansions [54,55]. H_2 is one of the few

systems with strong electron-electron correlation that can be efficiently solved numerically, thus being the best benchmark for different methods. When the two H nuclei are far apart, the trivial solution (two hydrogen atoms with electrons localized on different nuclei) is correlated, as one electron’s location on an atom determines the atom on which the other electron must be. Thus, the ground state is composed of a linear combination of two Slater determinants. Here, we benchmark the ground and the first excited state of the H_2 during dissociation, preserving the correct spin state (singlet and triplet) along the process.

To model the H_2 molecule, we employ a soft-core Coulomb potential that alleviates the electron-ion cusp and admits a nontrivial solution in one dimension while keeping all the essential correlation properties of the original Hamiltonian. The H_2 soft-core Coulomb potential is composed of the one-particle electron-ion interaction $V(r)$, the electron-electron interaction, and the ion-ion interaction (which is constant):

$$V(r_1, r_2) = V(r_1) + V(r_2) + \frac{1}{\sqrt{(r_1 - r_2)^2 + \eta^2}} + \frac{1}{\sqrt{d^2 + \eta^2}}, \quad (53)$$

$$V(r) = -\frac{1}{\sqrt{(r - d/2)^2 + \eta^2}} - \frac{1}{\sqrt{(r + d/2)^2 + \eta^2}}, \quad (54)$$

where d is the interatomic distance, η the softening factor of the Coulomb potential, and r_1, r_2 are the coordinates of the two electrons. It is known that the soft-Coulomb potential fundamentally changes the chemistry of bonding, as it affects the Coulomb potential even for $r \gg \eta$ [56,57]; however, in production simulations, this problem can be efficiently dealt with proper treatment of pseudopotentials, and allows us to separate the numerical instabilities originating from the electron-ion cusp (treated explicitly in Sec. VIIC) with the error of the method in describing electronic correlations.

The results are reported in Fig. 6, where we compare in red the NLSCHA with a six-layer transformation to the numerical exact diagonalization and the mean-field Hartree-Fock (HF) solution. As expected, HF reflects the problem of all mean-field approaches (as density functional theory) suffering from the static correlation error arising when the electrons localize on the atoms. The NLSCHA delivers an excellent agreement with the exact (numerical) result within the stochastic accuracy of the averages ($0.001E_h$ in the ground state).

We benchmarked the antibonding excited state solution by promoting an electron to the first excited state of the auxiliary harmonic Hamiltonian in a triplet spin configuration. Notably, the H_2 antibonding state is strongly affected by electron-hole interactions, as after the promotion of the electron in the excited state of the auxiliary Hamiltonian, the system significantly relaxes with dynamics similar to those giving rise to excitonic states in condensed matter systems. Therefore, the method displays promising potential for applications in materials with excitonic effects. Indeed, for a complete characterization of excitons, a dynamical extension of the method, in line with what is done for time-dependent SCHA [26], is required.

As the method is based on a first-principles quantization with a wave-function approach, like VMC, we have an ana-

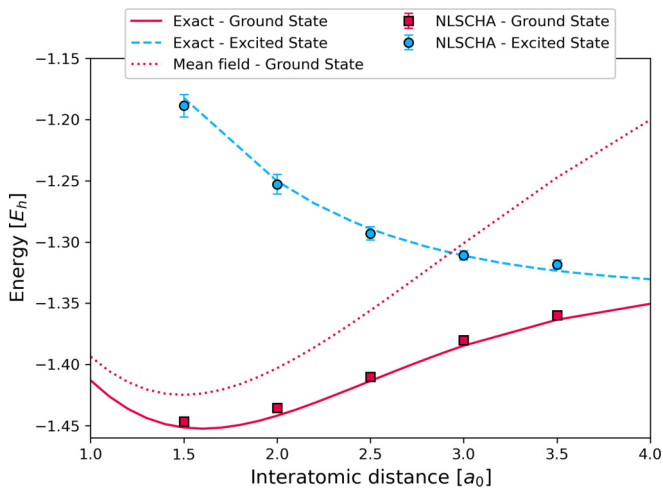


FIG. 6. Dissociation energies of H_2 . We compare the exact results obtained with numerical diagonalization of the Hamiltonian with the NLSCHA. The ground state is obtained in the singlet wave function, where both electrons have opposite spins in the ground state of the auxiliary harmonic oscillator. The excited state is the antibonding triplet solution, where one electron is in the ground state of the auxiliary harmonic Hamiltonian while the other is in the first excited state. The ground state's mean-field energy (Hartree-Fock) is also reported as a reference. The deviation in the dissociation limit to the expected energy for H_2 ($-1.0E_h$) is due to the soft-core Coulomb potential [Eq. (53)].

lytical expression for the resulting many-body wave function that we compare in Fig. 7 with the exact (full diagonalization) solution. The absence of spots in the electronic density located when $r_1 = r_2$ over the hydrogen nuclei confirms that the NLSCHA correctly accounts for static correlation, localizing the two electrons always in the two different ions. This is achieved thanks to the off-diagonal elements in the Γ matrix of the Gaussian stretcher that couple the wave function of different electrons (Sec. VI), like a backflow transformation employed in VMC.

VIII. LIMITS OF THE CURRENT IMPLEMENTATION

In this section, we present the limits of the current implementation and discuss how they could be overcome.

As shown in Sec. VII C, the NLSCHA struggles to reproduce the electron-ion cusp, especially when large unscreened nuclear charges are in play (like for the case of core electrons), requiring a diverging number of hidden layers in the curved manifold. However, this is a common problem in electronic-structure simulations, e.g., in plane-wave basis sets, and has been addressed *a posteriori* by employing pseudopotentials that smear out the electron-ion cusp near the nucleus. Also, the electron-electron cusp is likely to lead to numerical instabilities, and it is less clear how to fix it. However, due to its repulsive nature, the contribution of the exact wave-function shape impacts the total energy much less as the many-body electron density goes to zero on the cusp. An alternative approach would be to design an *ad hoc* transformation whose Jacobian introduces a Jastrow-like factor that encodes the correct electron-ion and electron-electron cusp.

Different network architectures can be used to parameterize the curved manifold, like the neural ordinary differential equations [58] and eventually stacked on top of the Gaussian stretchers to improve accuracy.

Our current implementation does not yet support periodic boundary conditions. Periodic system of distinguishable particles (like atomic nuclei) can be trivially treated by enforcing translational equivariance in the Gaussian stretcher and periodicity of the auxiliary harmonic Hamiltonian by imposing Eqs. (39) and (40). These are precisely the same conditions on Γ that a standard phononic force-constant matrix must satisfy in the presence of periodic boundary conditions. However, these requirements become too strict when combined with the antisymmetry constraints discussed in Sec. VI. To overcome this limitation for fermionic systems, the nonlinear transformation must be generalized, e.g., operating in relative coordinates. The proper definition and characterization of such a case go beyond the scope of this work.

Our framework allows, in principle, to describe systems at finite temperatures, as we optimize the density matrix to minimize the full Helmholtz free energy in the canonical ensemble. This is achieved thanks to the analytical expressions for the entropy in Eqs. (27) and (28). However, while our mapping is exact for pure states (at least for real and positive wave functions), it is not at finite temperatures. In fact, the curved manifold maps uniformly all the excited states of the auxiliary Hamiltonian into the real Cartesian space without affecting their energy. Therefore, the energy spacing between excited states of the transformed auxiliary system is preserved, not matching in general the real Hamiltonian one. Since the entropy depends solely on the occupations of states, which rely on their relative energy, the excited states of the auxiliary and real Hamiltonian are populated differently. To solve this problem, Ref. [35] proposed to introduce a variational parameter: the mass-tensor of the auxiliary Hamiltonian. This mass tensor effectively operates a linear transformation on the momentum operator, providing a way to rescale the energy level of excited states in the auxiliary Hamiltonian to resemble that of the real Hamiltonian more closely. A different approach is to use all the occupation numbers f_i in Eq. (28) as variational parameters, similarly to the scheme proposed for ions by Ref. [59]. This introduces a much higher number of degrees of freedom (infinite, in principle, for every mode), whose optimization can be discontinuous in correspondence with avoided crossing and hybridization of excitations. However, despite the challenges of reaching an exact solution at finite temperatures, the nonlinear SCHA with a variational mass is already expected to yield excellent results, particularly for finite temperature electron-ion systems, as most thermal effects impact the ionic density matrix, for which even the standard SCHA provides a very good finite temperature description [37].

IX. CONCLUSIONS

The introduction of the SCHA in a curved manifold enabled systematic improvement of the variational solution of the SCHA while preserving the analytical solution of the auxiliary Harmonic system. In this work, we established how

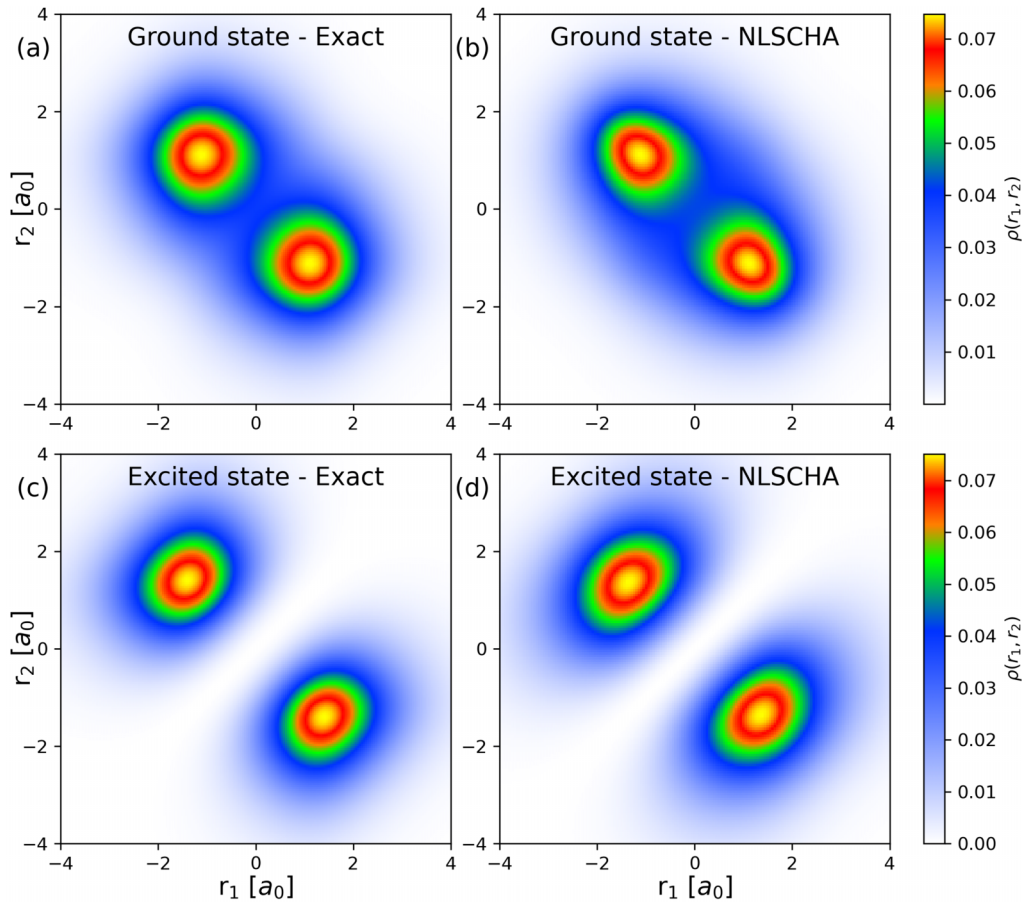


FIG. 7. Two-body electron density $\rho(r_1, r_2)$ for the H_2 in dissociation ($d = 2.5a_0$) in the 1D soft-core Coulomb potential. (a) The ground-state density evaluated with exact (numerical) diagonalization. (b) Ground-state density evaluated within the NLSCHA, employing six layers in the nonlinear transformation from the Gaussian. (c) First excited state with exact (numerical) diagonalization. This is an antibonding state with the spins in the triplet configuration. (d) Excited state density computed with the NLSCHA, employing two layers in the nonlinear transformation from the Gaussian. The error measured as the difference between the exact and NLSCHA probability density is reported in Appendix G (Fig. 8).

to build a curved manifold that can be iteratively improved by adding multiple hidden layers, like in a deep neural network.

We not only introduced a method to simulate electrons and ions within the same theoretical framework, but we also unveiled its promising perspective to study correlated systems. The NLSCHA outperforms mean-field approaches by capturing static correlation without breaking the spin symmetry, both in the ground and the excited state of H_2 . The similarity of the nonlinear transformation to a backflow and its Jacobian to a Jastrow factor suggests that it can also efficiently describe dynamical correlations, as in established variational Monte Carlo approaches. Ultimately, this work transforms a robust theory for atomic vibrations into a versatile framework for electronic structure, opening a chapter in the unified, first-principles simulation of quantum matter.

ACKNOWLEDGMENTS

L.M. acknowledges the European Union and the H2020 program for funding this project under the MSCA-IF, project ID 101018714. This research was supported by the NCCR MARVEL, a National Centre of Competence in Research,

funded by the Swiss National Science Foundation (Grant No. 205602).

DATA AVAILABILITY

The raw data used to generate the figures presented in this work are available from the corresponding author upon reasonable request.

APPENDIX A: KINETIC ENERGY

Here, we derive the kinetic energy

$$K = - \sum_{i=1}^N \int \prod_j dR_j \psi^\dagger(\mathbf{R}) \frac{\partial^2 \psi}{\partial R_i^2}. \quad (\text{A1})$$

Integrating by parts, we get

$$K = \sum_{i=1}^N \int \prod_j dR_j \frac{\partial \psi^\dagger}{\partial R_i} \frac{\partial \psi}{\partial R_i}. \quad (\text{A2})$$

We can bring the density outside as

$$K = \sum_{i=1}^N \frac{1}{2m_i} \int \prod_j dR_j \rho(\mathbf{R}) \frac{\partial \ln \psi^\dagger}{\partial R_i} \frac{\partial \ln \psi}{\partial R_i}. \quad (\text{A3})$$

From which, we have

$$\begin{aligned} \ln \psi &= \ln \tilde{\psi} + \frac{1}{2} \ln \det \mathbf{J}, \\ \frac{d \ln \psi}{dR_a} &= \sum_b \frac{d \ln \tilde{\psi}}{dq_b} \frac{dq_b}{dR_a} + \frac{1}{2} \frac{d \ln \det \mathbf{J}}{dR_a} \\ &= \sum_b \frac{d \ln \tilde{\psi}}{dq_b} J_{ba} + \frac{1}{2} \sum_{cd} \mathcal{J}_{cd} J_{ca}^d. \end{aligned} \quad (\text{A4})$$

The overall kinetic energy is

$$\begin{aligned} K &= \sum_{i=1}^N \frac{1}{2m_i} \left[\sum_{bc} \left\langle J_{bi} J_{ci} \frac{d \ln \tilde{\psi}^\dagger}{dq_b} \frac{d \ln \tilde{\psi}}{dq_c} \right\rangle \right. \\ &\quad + \frac{1}{2} \left\langle \left(\sum_{cd} \mathcal{J}_{cd} J_{ci}^d \right) \sum_e \frac{d \ln \tilde{\rho}}{dq_e} J_{ei} \right\rangle \\ &\quad \left. + \frac{1}{4} \left\langle \left(\sum_{cd} \mathcal{J}_{cd} J_{ci}^d \right)^2 \right\rangle \right], \end{aligned} \quad (\text{A6})$$

where $\tilde{\rho}(\mathbf{q}) = \tilde{\psi}^\dagger(\mathbf{q}) \tilde{\psi}(\mathbf{q})$. From Eq. (A20), we get the three observables:

$$\alpha_i(\mathbf{q}) = \frac{1}{4} \left[\sum_{cd} \mathcal{J}_{cd}(\mathbf{q}) J_{ci}^d(\mathbf{q}) \right]^2, \quad (\text{A7})$$

$$\beta_i^e(\mathbf{q}) = J_{ei}(\mathbf{q}) \left[\sum_{cd} \mathcal{J}_{cd}(\mathbf{q}) J_{ci}^d(\mathbf{q}) \right], \quad (\text{A8})$$

$$\gamma_i^{bc}(\mathbf{q}) = J_{bi}(\mathbf{q}) J_{ci}(\mathbf{q}), \quad (\text{A9})$$

by exploiting the relation Eq. (D1), we get

$$\sum_{cd} \mathcal{J}_{cd} J_{ci}^d = - \sum_{ahk} J_{ha} J_{ki} J_{kh}^a. \quad (\text{A10})$$

This, indeed, works only for $T = 0$ K; however, a similar approach can be derived also at finite temperature recognizing that

$$K = - \sum_i \frac{1}{2m_i} \int \prod_j dR_j dR'_j \delta(\mathbf{R} - \mathbf{R}') \frac{\partial^2 \rho(\mathbf{R}, \mathbf{R}')}{\partial R_i'^2}. \quad (\text{A11})$$

We can then exploit the integration by parts, so that

$$K = \sum_i \frac{1}{2m_i} \int \prod_j dR_j dR'_j \frac{\partial \delta(\mathbf{R} - \mathbf{R}')}{\partial R_i'} \frac{\partial \rho(\mathbf{R}, \mathbf{R}')}{\partial R_i'}. \quad (\text{A12})$$

Now, we can exchange the derivative in the Dirac δ with R ,

$$K = - \sum_i \frac{1}{2m_i} \int \prod_j dR_j dR'_j \frac{\partial \delta(\mathbf{R} - \mathbf{R}')}{\partial R_i} \frac{\partial \rho(\mathbf{R}, \mathbf{R}')}{\partial R_i'}. \quad (\text{A13})$$

And integrating back by parts,

$$K = \sum_i \frac{1}{2m_i} \int \prod_j dR_j dR'_j \delta(\mathbf{R} - \mathbf{R}') \frac{\partial^2 \rho(\mathbf{R}, \mathbf{R}')}{\partial R_i' \partial R_i}. \quad (\text{A14})$$

This expression allows us to recover a symmetry and a single derivative on each variable that we can exploit:

$$\begin{aligned} \frac{\partial^2 \rho(\mathbf{R}, \mathbf{R}')}{\partial R_i' \partial R_i} &= \sum_h \lambda_h \frac{\partial \psi_h}{\partial R_i}(\mathbf{R}) \frac{\partial \psi_h^\dagger}{\partial R_i}(\mathbf{R}') \\ &= \sum_h \lambda_h |\psi_h|^2 \frac{\partial \ln \psi_h}{\partial R_i}(\mathbf{R}) \frac{\partial \ln \psi_h^\dagger}{\partial R_i}(\mathbf{R}'). \end{aligned} \quad (\text{A15})$$

Now, the same procedure as in Eq. (A5) can be exploited, where we end up with the same integrals as in Eq. (A20), but where the averages are taken over each state in the mixture defined by the density matrix:

$$\begin{aligned} K &= \sum_{i=1}^N \frac{1}{2m_i} \left[\sum_{bc} \int d\mathbf{q} J_{bi} J_{ci} \sum_h \lambda_h |\tilde{\psi}_h|^2 \frac{d \ln \tilde{\psi}_h^\dagger}{dq_b} \frac{d \ln \tilde{\psi}_h}{dq_c} \right. \\ &\quad + \int d\mathbf{q} \frac{1}{2} \left(\sum_{cd} \mathcal{J}_{cd} J_{ci}^d \right) \sum_e J_{ei} \sum_h \lambda_h |\tilde{\psi}_h|^2 \frac{d \ln \tilde{\rho}_h}{dq_e} \\ &\quad \left. + \frac{1}{4} \int d\mathbf{q} \sum_h \lambda_h |\tilde{\psi}_h|^2 \left(\sum_{cd} \mathcal{J}_{cd} J_{ci}^d \right)^2 \right], \end{aligned} \quad (\text{A16})$$

where

$$\sum_h \lambda_h |\tilde{\psi}_h|^2 \frac{d \ln \tilde{\psi}_h^\dagger}{dq_b} \frac{d \ln \tilde{\psi}_h}{dq_c} = \int d\mathbf{q}' \delta(\mathbf{q} - \mathbf{q}') \frac{\partial^2 \tilde{\rho}(\mathbf{q}, \mathbf{q}')}{\partial q_b \partial q_c'}, \quad (\text{A17})$$

$$\sum_h \lambda_h |\tilde{\psi}_h|^2 \frac{d \ln \tilde{\rho}_h}{dq_e} = \tilde{\rho}(\mathbf{q}) \frac{d \ln \tilde{\rho}}{dq_e}, \quad (\text{A18})$$

$$\sum_h \lambda_h |\tilde{\psi}_h|^2 = \tilde{\rho}(\mathbf{q}), \quad (\text{A19})$$

$$\begin{aligned} K &= \sum_{i=1}^N \frac{1}{2m_i} \left[\sum_{bc} \int d\mathbf{q} J_{bi} J_{ci} \int d\mathbf{q}' \delta(\mathbf{q} - \mathbf{q}') \frac{\partial^2 \tilde{\rho}(\mathbf{q}, \mathbf{q}')}{\partial q_b \partial q_c'} \right. \\ &\quad + \frac{1}{2} \int d\mathbf{q} \left(\sum_{cd} \mathcal{J}_{cd} J_{ci}^d \right) \sum_e J_{ei} \tilde{\rho}(\mathbf{q}) \frac{d \ln \tilde{\rho}}{dq_e} \\ &\quad \left. + \frac{1}{4} \left\langle \left(\sum_{cd} \mathcal{J}_{cd} J_{ci}^d \right)^2 \right\rangle \right]. \end{aligned} \quad (\text{A20})$$

Equation (A20) is the complete expression for the free energy, regardless of the auxiliary density matrix of choice $\tilde{\rho}(\mathbf{q}, \mathbf{q}')$. In the particular case of the Gaussian density matrix,

$$\begin{aligned} &\int d\mathbf{q}' \delta(\mathbf{q} - \mathbf{q}') \frac{\partial^2 \tilde{\rho}(\mathbf{q}, \mathbf{q}')}{\partial q_a \partial q_b} \\ &= \tilde{\rho}(\mathbf{q}) \left(A_{ab} + \sum_{cd} \frac{\Upsilon_{ac} \Upsilon_{bd}}{4} (q_c - Q_c)(q_d - Q_d) \right), \end{aligned} \quad (\text{A21})$$

$$\rho(\mathbf{q}) \frac{d \ln \tilde{\rho}}{dq_e} = -\tilde{\rho}(\mathbf{q}) \sum_b \Upsilon_{eb} (q_b - Q_b), \quad (\text{A22})$$

that can be fit into the expression of Eq. (A20) to evaluate numerically the kinetic energy.

APPENDIX B: HIGH-ORDER JACOBIANS OF THE SINGLE-LAYER GAUSSIAN STRETCHER

To compute explicitly all the expressions for the kinetic energy and the free energy gradients in the specific case of the Gaussian Stretcher transformation, we need to compute the Jacobian matrix up to the fourth order. We start with the single Gaussian stretcher

$$\mathcal{J}_{ab} = \frac{dR_a}{dq_b}, \quad (\text{B1})$$

$$\begin{aligned} \mathcal{J}_{ab} = & \delta_{ab} [1 + (e^\zeta - 1)e^{-\frac{1}{2}\tilde{\mathbf{q}}\tilde{\mathbf{\Gamma}}\tilde{\mathbf{q}}} \\ & - (e^\zeta - 1)e^{-\frac{1}{2}\tilde{\mathbf{q}}\tilde{\mathbf{\Gamma}}\tilde{\mathbf{q}}} \sum_c \Gamma_{bc}(q_c - Q_c)(q_a - Q_a). \end{aligned} \quad (\text{B2})$$

We can go on with further derivatives:

$$\mathcal{J}_{bc}^a = \frac{dR_a}{dq_b dq_c}, \quad (\text{B3})$$

$$\begin{aligned} \frac{\mathcal{J}_{bc}^a e^{\frac{1}{2}\tilde{\mathbf{q}}\tilde{\mathbf{\Gamma}}\tilde{\mathbf{q}}}}{(e^\zeta - 1)} = & -\delta_{ab} \sum_h \Gamma_{ch}(q_h - Q_h) - \delta_{ac} \sum_h \Gamma_{bh}(q_h - Q_h) \\ & + \sum_h \Gamma_{bh}(q_h - Q_h) \sum_k \Gamma_{ck}(q_k - Q_k)(q_a - Q_a) \\ & - \Gamma_{bc}(q_a - Q_a), \end{aligned} \quad (\text{B4})$$

$$\begin{aligned} \frac{\mathcal{J}_{bcd}^a e^{\frac{1}{2}\tilde{\mathbf{R}}\tilde{\mathbf{\Gamma}}\tilde{\mathbf{R}}}}{(e^\zeta - 1)} = & -\delta_{ab}\Gamma_{cd} - \delta_{ac}\Gamma_{bd} \\ & - \delta_{ad}\Gamma_{bc} + \delta_{ad}(\tilde{\mathbf{\Gamma}}\tilde{\mathbf{q}})_b(\tilde{\mathbf{\Gamma}}\tilde{\mathbf{q}})_c \\ & + \tilde{q}_a\Gamma_{bd}(\tilde{\mathbf{\Gamma}}\tilde{\mathbf{q}})_c + \tilde{q}_a\Gamma_{cd}(\tilde{\mathbf{\Gamma}}\tilde{\mathbf{q}})_b \\ & + \delta_{ab}(\tilde{\mathbf{\Gamma}}\tilde{\mathbf{q}})_d(\tilde{\mathbf{\Gamma}}\tilde{\mathbf{q}})_c + \delta_{ac}(\tilde{\mathbf{\Gamma}}\tilde{\mathbf{q}})_b(\tilde{\mathbf{\Gamma}}\tilde{\mathbf{q}})_d \\ & + \tilde{q}_a\Gamma_{bc}(\tilde{\mathbf{\Gamma}}\tilde{\mathbf{q}})_d - \tilde{q}_a(\tilde{\mathbf{\Gamma}}\tilde{\mathbf{q}})_b(\tilde{\mathbf{\Gamma}}\tilde{\mathbf{q}})_c(\tilde{\mathbf{\Gamma}}\tilde{\mathbf{q}})_d. \end{aligned} \quad (\text{B5})$$

APPENDIX C: HIGH-ORDER JACOBIANS OF THE MULTILAYER GAUSSIAN STRETCHER

As suggested in the main text, nonlinear transformations can be concatenated to form more complex transformations. We can derive the Jacobian by using the composite transformation:

$$\frac{d}{dq_j} q_i^{(m)}(q_1^{(m-1)}, \dots, q_n^{(m-1)}) = \sum_{a=1}^n \frac{dq_i^{(m)}}{dq_a^{(m-1)}} \frac{dq_a^{(m-1)}}{dq_j}. \quad (\text{C1})$$

From which, we get a recursive equation

$$\mathcal{J}_{ij}^{(m)} = \sum_a \mathcal{J}_{ia}(q^{(m-1)}) \mathcal{J}_{aj}^{(m-1)}(q^{(1)}), \quad (\text{C2})$$

which means that the Jacobian is the matrix product of the Jacobians after each iteration. The second-order Jacobian becomes

$$\begin{aligned} (\mathcal{J}^{(m)})_{jk}^i = & \sum_{ab} \mathcal{J}_{ab}^i(q^{(m-1)}) \mathcal{J}_{aj}^{(m-1)} \mathcal{J}_{bk}^{(m-1)} \\ & + \sum_a \mathcal{J}_{ia}(q^{(m-1)}) (\mathcal{J}^{(m-1)})_{jk}^a. \end{aligned} \quad (\text{C3})$$

And the higher-order derivative becomes

$$\begin{aligned} (\mathcal{J}^{(m)})_{jkh}^i = & \sum_{abc} \mathcal{J}_{abc}^i(q^{(m-1)}) \mathcal{J}_{aj}^{(m-1)} \mathcal{J}_{bk}^{(m-1)} \mathcal{J}_{ch}^{(m-1)} \\ & + \sum_{ab} \mathcal{J}_{ab}^i(q^{(m-1)}) (\mathcal{J}^{(m-1)})_{jh}^a \mathcal{J}_{bk}^{(m-1)} \\ & + \sum_{ab} \mathcal{J}_{ab}^i(q^{(m-1)}) \mathcal{J}_{aj}^{(m-1)} (\mathcal{J}^{(m-1)})_{kh}^b \\ & + \sum_{ab} \mathcal{J}_{ab}^i(q^{(m-1)}) \mathcal{J}_{bh}^{(m-1)} (\mathcal{J}^{(m-1)})_{jk}^a \\ & + \sum_a \mathcal{J}_{ia}(q^{(m-1)}) (\mathcal{J}^{(m-1)})_{jkh}^a. \end{aligned} \quad (\text{C4})$$

This chain rule allows the computation of all the Jacobians as a forward-propagation algorithm. This scheme can be seen as a neural network with a Gaussian activation function.

APPENDIX D: SOME USEFUL RELATIONS

We simplify some explicit multiplication between Jacobians that often appear in the derivations. For example, in the expression for $\beta_a^i(\mathbf{q})$, we find the multiplication:

$$\sum_{cd} \mathcal{J}_{cd} \mathcal{J}_{ca}^d = - \sum_{hkl} J_{kh} J_{la} \mathcal{J}_{kl}^h. \quad (\text{D1})$$

In the same way, for α is important, the following multiplication:

$$\begin{aligned} \sum_{bc} \mathcal{J}_{bc} \mathcal{J}_{ba}^c = & - \sum_{hkl} \left[J_{kh} J_{la} \sum_m \mathcal{J}_{lkm}^h J_{ma} \right. \\ & \left. + 2J_{ha}^k J_{la} \mathcal{J}_{lk}^h + J_{kh} J_{aa}^l \mathcal{J}_{lk}^h \right]. \end{aligned} \quad (\text{D2})$$

APPENDIX E: SCALING

The computational scaling of the free-energy optimization algorithm with respect to the number of degrees of freedom n is a critical aspect of its performance. The dominant cost arises from the kinetic energy term, which depends on the Jacobian derivatives of the multilayer coordinate transformation. While these terms are in principle highly costly to compute and store in memory, they can be systematically decomposed into a series of multiplications between $n \times n$ matrices, where n is the number of degrees of freedom, as is evident from the expressions in Eqs. (B2) and (B4). The scaling is therefore dictated by the complexity of these underlying matrix operations. For a general Gaussian stretcher where the covariance matrix Γ is assumed to be dense, the resulting Jacobian and its derivative-related matrices are also dense. Consequently, fundamental operations such as matrix inversion or multiplication have a computational cost of $O(n^3)$. As shown by substituting Eq. (B4) into Eq. (D1), even the inverse of higher-order tensors can be constructed from these $n \times n$ matrix products. A multilayer architecture composed of m chained transformations requires these matrix operations to be performed for each layer (see Appendix C), introducing a linear scaling with m . The optimization step, which relies on backpropagation to

compute gradients, shares the same asymptotic complexity as the forward pass. Therefore, the overall theoretical cost of the algorithm scales as

$$O(m \cdot n^3).$$

Notably, this scaling is comparable to the best *ab initio* techniques, such as DFT and VMC. This scaling analysis hinges on the critical assumption of a dense Γ matrix. It is plausible that employing a sparse or structured representation for Γ could significantly reduce the computational cost, potentially leading to a more favorable scaling law.

APPENDIX F: ANHARMONICITY IN THE DOUBLE-WELL POTENTIAL

The anharmonicity in the double-well potential [Eq. (51)] is introduced by the presence of the quartic term R^4 , which has a as a coefficient. However, the lowest-order correction to the average curvature of the potential is

$$\left\langle \frac{d^2V}{dR^2} \right\rangle = 12a\sigma^2 - 2b, \quad (\text{F1})$$

where $\sigma^2 = \langle RR \rangle$. Within the Harmonic approximation, we have at $T = 0$ K:

$$\sigma^2 = \frac{\hbar}{2m\omega}, \quad \omega^2 = -\frac{b}{m}.$$

This result indicates that the lowest-order correction to the harmonic frequency due to anharmonicity diverges as $b \rightarrow 0$, signaling the onset of a nonperturbative regime for positive values of b . For this reason, we analyze the double-well potential as a function of b rather than a : the parameter b more directly quantifies the degree of anharmonicity and represents a more demanding case for theoretical analysis.

APPENDIX G: ERROR ON THE H₂ DENSITY

Here, we report the error for the calculation of the H₂ density. The error is evaluated as

$$\text{Err}(r_1, r_2) = ||\Psi_{\text{exact}}(r_1, r_2)|^2 - |\Psi_{\text{NLSCHA}}(r_1, r_2)|^2|. \quad (\text{G1})$$

The results for the ground and excited states corresponding to the density reported in Fig. 7 are presented in Fig. 8.

APPENDIX H: HARMONIC OSCILLATOR OF FERMIONS

Here, we report the solution of the harmonic Hamiltonian for fermions. The system is composed of N fermions interacting within a harmonic potential. For simplicity, we consider a one-dimensional space. Due to the exchange symmetry in 1D there are only two degrees of freedom: the self-interaction a and electronic repulsion b ,

$$V(q_1, \dots, q_N) = \frac{1}{2}a \sum_{i=1}^N q_i^2 - \frac{1}{2}b \sum_{ij} (q_i - q_j)^2. \quad (\text{H1})$$

Reorganizing the terms of the equation, we get

$$V(q_1, \dots, q_N) = \frac{1}{2}(a-b) \sum_{i=1}^N q_i^2 - b \sum_{i \neq j} q_i q_j, \quad (\text{H2})$$

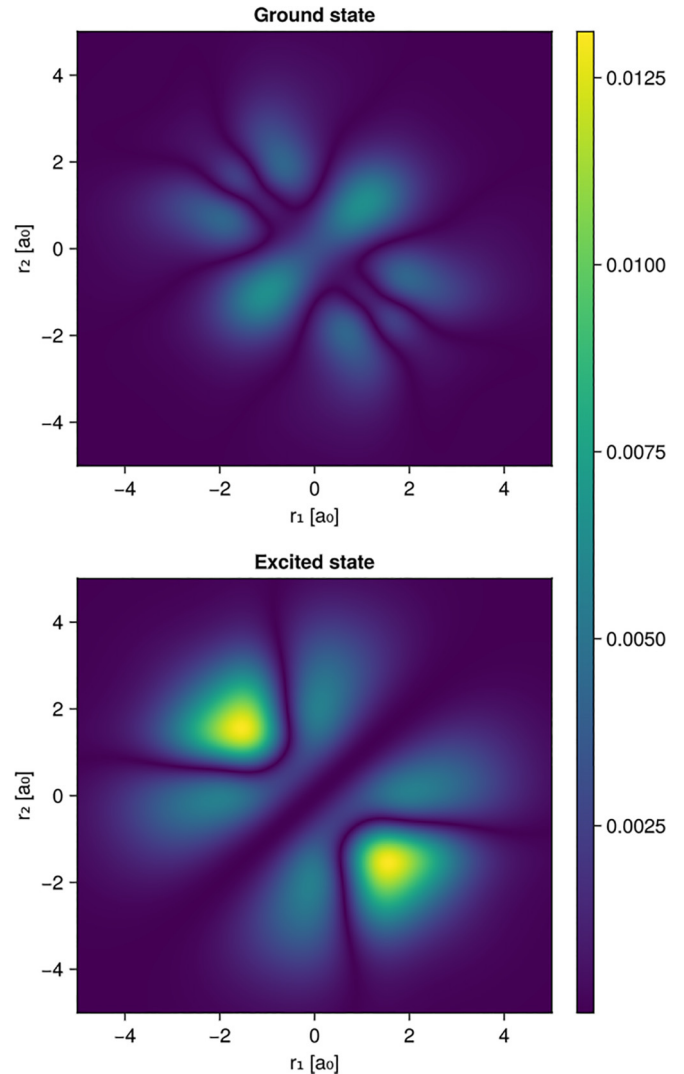


FIG. 8. Error on the probability density of the NLSCHA wave function compared to the exact result for H₂ with an interatomic distance of $2.5a_0$ (Fig. 7). The error is evaluated using Eq. (G1).

$$V(\mathbf{q}) = \frac{1}{2}(q_1 \cdots q_N) \begin{pmatrix} (a-b) & -b & \cdots & -b \\ -b & (a-b) & \cdots & -b \\ \vdots & \vdots & \ddots & \vdots \\ -b & \cdots & -b & (a-b) \end{pmatrix} \begin{pmatrix} q_1 \\ \vdots \\ q_N \end{pmatrix}. \quad (\text{H3})$$

This is an interacting problem that can be trivially solved by diagonalizing the matrix. The solution have two eigenvalues:

$$\lambda_1 = a - Nb, \quad \mathbf{e}_1 = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}, \quad (\text{H4})$$

$$\lambda_2, \dots, \lambda_n = a, \quad \mathbf{e}_2 = \begin{pmatrix} 1 \\ -1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \dots, \quad \mathbf{e}_N = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ -1 \end{pmatrix}. \quad (\text{H5})$$

The eigenvector e_1 is already symmetric under the exchange of any couple of particles, each e_i with $i \geq 2$ is antisymmetric for exchanging two specific particles and symmetric for exchanging any other two. Unfortunately, for more than two particles, it is impossible to build a basis from e_2, \dots, e_n that commutes with all the exchange operators S_{ij} . While we know the spectrum of the problem, this makes it very difficult to find exact selection rules to populate states in the case of fermions and bosons for more than two particles, in particular, if we also account for the spin degrees of freedom. Another problem arises in the thermodynamic limit $N \rightarrow \infty$. In this case, the spectrum, and in particular λ_1 [Eq. (H4)], becomes not positive definite for any small positive value of b . Thus, even if the harmonic Hamiltonian may describe the correlation of molecules or systems with a finite number of electrons, $b \rightarrow 0$ in the thermodynamic limit, we recover a noninteracting Hamiltonian. For this reason, in the main text, we restrict to the noninteracting case where $b = 0$.

APPENDIX I: POSITION-DEPENDENT EFFECTIVE MASS TENSOR

As in classical Lagrange equations, the constrained motion on a curved manifold affects the effective masses of the systems. In classical mechanics, the kinetic energy after the change of variable is obtained as

$$K_{\text{cl}} = \sum_i \frac{1}{2m_i} \langle J_{ai} J_{bi} p_a p_b \rangle, \quad (11)$$

where p_a is the a th component of the moment of the auxiliary variable q_a . The transformation of the variable can thus be seen as a mass that depends on the metric tensor

$$M_{ab}^{-1} = \sum_i \frac{J_{ai} J_{bi}}{m_i}, \quad (12)$$

$$K_{\text{cl}} = \frac{1}{2} \sum_{ab} \langle p_a M_{ab}^{-1} p_b \rangle. \quad (13)$$

Unfortunately, in quantum mechanics, it is not possible to quantize directly the auxiliary variables and thus express everything as a simple change of masses. However, we can always write the kinetic energy as

$$K = K_{\text{cl}} + K_{\text{cor}}, \quad (14)$$

where K_{cor} is the quantum correction to the kinetic energy. In particular, from Eq. (31), the term that contains Eq. (13) is the one that multiplies the γ tensor.

$$K_{\text{cor}} = - \sum_{a=1}^{3N} \frac{\hbar^2}{2m_a} \left\langle \alpha^a(\mathbf{q}) + \sum_i \beta_i^a(\mathbf{q}) \hat{P}_i \right\rangle. \quad (15)$$

Since in practice, the classical kinetic energy is both the biggest part of the kinetic energy and the one with the highest noise, it is convenient to rewrite it as the masses are constant (linear transformation) and average only the curvature:

$$\bar{M}_{ab}^{-1} = \langle M_{ab}^{-1}(\mathbf{q}) \rangle, \quad \Delta M_{ab}^{-1}(\mathbf{q}) = M_{ab}^{-1}(\mathbf{q}) - \bar{M}_{ab}^{-1} \quad (16)$$

$$K_{\text{cl}} = \frac{1}{2} \sum_{ab} \bar{M}_{ab}^{-1} \langle p_a p_b \rangle + \frac{1}{2} \sum_{ab} \langle p_a \Delta M_{ab}^{-1}(\mathbf{q}) p_b \rangle, \quad (17)$$

where the first term is analytical when Gaussian wave functions are used in the auxiliary system. The evaluation of the classical kinetic energy through Eq. (17) strongly suppresses the stochastic noise compared with Eq. (11), but it has the same expected value.

APPENDIX J: FERMIONIC AND BOSONIC WAVE FUNCTION

Up to two electrons, the wave function remains symmetric as it can be described by a singlet state, where the antisymmetric part of the wave function is trivially encoded by their opposed spin. However, the factorization of space and spin for the exchange symmetry is only possible for two electrons. Therefore, to solve a realistic system beyond the helium atom or the H_2 molecule it is necessary to devise a strategy to encode the antisymmetry of the wave function.

In the opening of Sec. VI, we derived the constraint on the nonlinear transformation to keep the exchange symmetry of the wave function in the auxiliary variables throughout the nonlinear transformation operated by the neural network. Here, we devise a way to account for a fermionic (or bosonic) wave function directly in the auxiliary space.

The only requirements are to be able to

- (1) extract random configurations according to the correct statistics,
- (2) evaluate the modified kinetic energy average in the auxiliary space [Eqs. (15) and (13)].

To extract random configurations distributed according to the Fermi-Dirac, we employ a standard quantum Monte Carlo algorithm sampling the wave function $\Psi(q)$ as the Slater determinant of the harmonic eigenfunctions.

Storing for each configuration in q space, the gradient of the wave function, we can easily compute the averages of kinetic operators that depend on both P and q :

$$i\hbar \langle \beta_a(\mathbf{q}) P_a \rangle = -\hbar^2 \left\langle \beta_a(\mathbf{q}) \frac{\partial \ln \Psi}{\partial q_a} \right\rangle, \quad (J1)$$

$$\langle P_a \gamma_{ab}(\mathbf{q}) P_b \rangle = \hbar^2 \left\langle \gamma_{ab}(\mathbf{q}) \frac{\partial \ln \Psi^*}{\partial q_a} \frac{\partial \ln \Psi}{\partial q_b} \right\rangle, \quad (J2)$$

which is equivalent to substituting the $\Upsilon \mathbf{q}$ with $\nabla_q \ln \Psi$ from the equations of the standard SSCHA. Interestingly, the derivative of the logarithm of the wave function is independent on its global phase, thus, Eq. (J1) satisfies the Gauge invariance. Indeed, if the global phase depends on the position, care must be taken as Eq. (J1) could be related to topological phenomena. Therefore, we must store, for each point in the auxiliary ensemble, the value of the gradient of the logarithm of the Slater determinant, which can be evaluated efficiently using algorithmic differentiation. The analytical value of the kinetic energy in a pure state can be evaluated as

$$\langle P_a P_b \rangle = \frac{1}{N_e} \delta_{ab} \frac{\hbar}{2} \sum_{\mu=0}^{\infty} [2n_{\mu} + 1] f_{\mu}, \quad (J3)$$

where f_{μ} now indicates the number of electrons occupying the n_{μ} state of the auxiliary harmonic Hamiltonian along the a or b direction.

- [1] N. Marzari, A. Ferretti, and C. Wolverton, Electronic-structure methods for materials design, *Nat. Mater.* **20**, 736 (2021).
- [2] P. Hohenberg and W. Kohn, Inhomogeneous electron gas, *Phys. Rev.* **136**, B864 (1964).
- [3] K. Burke and L. O. Wagner, DFT in a nutshell, *Int. J. Quantum Chem.* **113**, 96 (2012).
- [4] A. Georges, G. Kotliar, W. Krauth, and M. J. Rozenberg, Dynamical mean-field theory of strongly correlated fermion systems and the limit of infinite dimensions, *Rev. Mod. Phys.* **68**, 13 (1996).
- [5] L. Monacelli, I. Errea, M. Calandra, and F. Mauri, Black metal hydrogen above 360 GPa driven by proton quantum fluctuations, *Nat. Phys.* **17**, 63 (2021).
- [6] L. Monacelli, M. Casula, K. Nakano, S. Sorella, and F. Mauri, Quantum phase diagram of high-pressure hydrogen, *Nat. Phys.* **19**, 845 (2023).
- [7] I. Errea, M. Calandra, C. J. Pickard, J. R. Nelson, R. J. Needs, Y. Li, H. Liu, Y. Zhang, Y. Ma, and F. Mauri, Quantum hydrogen-bond symmetrization in the superconducting hydrogen sulfide system, *Nature (London)* **532**, 81 (2016).
- [8] I. Errea, F. Belli, L. Monacelli, A. Sanna, T. Koretsune, T. Tadano, R. Bianco, M. Calandra, R. Arita, F. Mauri, and J. A. Flores-Livas, Quantum crystal structure in the 250-Kelvin superconducting lanthanum hydride, *Nature (London)* **578**, 66 (2020).
- [9] J. A. Morrone and R. Car, Nuclear quantum effects in water, *Phys. Rev. Lett.* **101**, 017801 (2008).
- [10] M. Cherubini, L. Monacelli, and F. Mauri, The microscopic origin of the anomalous isotopic properties of ice relies on the strong quantum anharmonic regime of atomic vibration, *J. Chem. Phys.* **155**, 184502 (2021).
- [11] U. Ranieri, S. Di Cataldo, M. Rescigno, L. Monacelli, R. Gaal, M. Santoro, L. Andriambariarijaona, P. Parisiades, C. De Michele, and L. E. Bove, Observation of the most H₂-dense filled ice under high pressure, *Proc. Natl. Acad. Sci. USA* **120**, e2312665120 (2023).
- [12] M. Cherubini, L. Monacelli, B. Yang, R. Car, M. Casula, and F. Mauri, Quantum effects in the H-bond symmetrization and in the thermodynamic properties of high pressure ice, *Phys. Rev. B* **110**, 014112 (2024).
- [13] L. Binci, P. Barone, and F. Mauri, First-principles theory of infrared vibrational spectroscopy of metals and semimetals: Application to graphite, *Phys. Rev. B* **103**, 134304 (2021).
- [14] G. Marchese, F. Macheda, L. Binci, M. Calandra, P. Barone, and F. Mauri, Born effective charges and vibrational spectra in superconducting and bad conducting metals, *Nat. Phys.* **20**, 88 (2023).
- [15] N. Giroto and D. Novko, Dynamical renormalization of electron-phonon coupling in conventional superconductors, *Phys. Rev. B* **107**, 064310 (2023).
- [16] C.-J. Tong, X. Cai, A.-Y. Zhu, L.-M. Liu, and O. V. Prezhdo, How hole injection accelerates both ion migration and nonradiative recombination in metal halide perovskites, *J. Am. Chem. Soc.* **144**, 6604 (2022).
- [17] Y. J. Uemura *et al.*, Universal correlations between T_c and $\frac{n_s}{m^*}$ (carrier density over effective mass) in high- T_c cuprate superconductors, *Phys. Rev. Lett.* **62**, 2317 (1989).
- [18] J. C. Tully, Molecular dynamics with electronic transitions, *J. Chem. Phys.* **93**, 1061 (1990).
- [19] P. Nijjar, J. Jankowska, and O. V. Prezhdo, Ehrenfest and classical path dynamics with decoherence and detailed balance, *J. Chem. Phys.* **150**, 204124 (2019).
- [20] C. F. Craig, W. R. Duncan, and O. V. Prezhdo, Trajectory surface hopping in the time-dependent Kohn-Sham approach for electron-nuclear dynamics, *Phys. Rev. Lett.* **95**, 163001 (2005).
- [21] L. Wang, A. Akimov, and O. V. Prezhdo, Recent progress in surface hopping: 2011-2015, *J. Phys. Chem. Lett.* **7**, 2100 (2016).
- [22] P. Shushkov, R. Li, and J. C. Tully, Ring polymer molecular dynamics with surface hopping, *J. Chem. Phys.* **137**, 22A549 (2012).
- [23] D. M. Ceperley, Path integrals in the theory of condensed helium, *Rev. Mod. Phys.* **67**, 279 (1995).
- [24] K. P. Driver and B. Militzer, All-electron path integral Monte Carlo simulations of warm dense matter: Application to water and carbon plasmas, *Phys. Rev. Lett.* **108**, 115502 (2012).
- [25] I. Errea, M. Calandra, and F. Mauri, Anharmonic free energies and phonon dispersions from the stochastic self-consistent harmonic approximation: Application to platinum and palladium hydrides, *Phys. Rev. B* **89**, 064302 (2014).
- [26] L. Monacelli and F. Mauri, Time-dependent self-consistent harmonic approximation: Anharmonic nuclear quantum dynamics and time correlation functions, *Phys. Rev. B* **103**, 104305 (2021).
- [27] M. Miotto and L. Monacelli, Fast prediction of anharmonic vibrational spectra for complex organic molecules, *npj Comput. Mater.* **10**, 240 (2024).
- [28] M. Borinaga, P. Riego, A. Leonardo, M. Calandra, F. Mauri, A. Bergara, and I. Errea, Anharmonic enhancement of superconductivity in metallic molecular Cmca-4 hydrogen at high pressure: A first-principles study, *J. Phys.: Condens. Matter* **28**, 494001 (2016).
- [29] M. Borinaga, I. Errea, M. Calandra, F. Mauri, and A. Bergara, Anharmonic effects in atomic hydrogen: Superconductivity and lattice dynamical stability, *Phys. Rev. B* **93**, 174308 (2016).
- [30] L. Monacelli and N. Marzari, First-principles thermodynamics of CsSnI₃, *Chem. Mater.* **35**, 1702 (2023).
- [31] J. S. Zhou, R. Bianco, L. Monacelli, I. Errea, F. Mauri, and M. Calandra, Theory of the thickness dependence of the charge density wave transition in 1 t-TiTe₂, *2D Mater.* **7**, 045032 (2020).
- [32] R. Bianco, I. Errea, L. Monacelli, M. Calandra, and F. Mauri, Quantum enhancement of charge density wave in NbS₂ in the two-dimensional limit, *Nano Lett.* **19**, 3098 (2019).
- [33] R. Bianco, L. Monacelli, M. Calandra, F. Mauri, and I. Errea, Weak dimensionality dependence and dominant role of ionic fluctuations in the charge-density-wave transition of NbSe₂, *Phys. Rev. Lett.* **125**, 106101 (2020).
- [34] J. Diego, A. H. Said, S. K. Mahatha, R. Bianco, L. Monacelli, M. Calandra, F. Mauri, K. Rossnagel, I. Errea, and S. Blanco-Canosa, Van der Waals driven anharmonic melting of the 3D charge density wave in VS₂, *Nat. Commun.* **12**, 598 (2021).
- [35] A. Siciliano, L. Monacelli, and F. Mauri, Beyond Gaussian fluctuations of quantum anharmonic nuclei, *Phys. Rev. B* **110**, 134111 (2024).

- [36] A. Siciliano, L. Monacelli, and F. Mauri, Beyond Gaussian fluctuations of quantum anharmonic nuclei: The case of rotational degrees of freedom, *Phys. Rev. B* **110**, 144101 (2024).
- [37] L. Monacelli, R. Bianco, M. Cherubini, M. Calandra, I. Errea, and F. Mauri, The stochastic self-consistent harmonic approximation: Calculating vibrational properties of materials with full quantum and anharmonic effects, *J. Phys.: Condens. Matter* **33**, 363001 (2021).
- [38] W. M. C. Foulkes, L. Mitas, R. J. Needs, and G. Rajagopal, Quantum Monte Carlo simulations of solids, *Rev. Mod. Phys.* **73**, 33 (2001).
- [39] U. Aseginolaza, R. Bianco, L. Monacelli, L. Paulatto, M. Calandra, F. Mauri, A. Bergara, and I. Errea, Phonon collapse and second-order phase transition in thermoelectric SnSe, *Phys. Rev. Lett.* **122**, 075901 (2019).
- [40] C. Verdi, L. Ranalli, C. Franchini, and G. Kresse, Quantum paraelectricity and structural phase transitions in strontium titanate beyond density functional theory, *Phys. Rev. Mater.* **7**, L030801 (2023).
- [41] D. Romanin, L. Monacelli, R. Bianco, I. Errea, F. Mauri, and M. Calandra, Dominant role of quantum anharmonicity in the stability and optical properties of infinite linear acetylenic carbon chains, *J. Phys. Chem. Lett.* **12**, 10339 (2021).
- [42] L. Ranalli, C. Verdi, L. Monacelli, G. Kresse, M. Calandra, and C. Franchini, Temperature-dependent anharmonic phonons in quantum paraelectric KTaO_3 by first principles and machine-learned force fields, *Adv. Quantum Technol.* **6**, 2200131 (2023).
- [43] A. Pedrielli, P. E. Trevisanutto, L. Monacelli, G. Garberoglio, N. M. Pugno, and S. Taioli, Understanding anharmonic effects on hydrogen desorption characteristics of Mg_nH_2 nanoclusters by *ab initio* trained deep neural network, *Nanoscale* **14**, 5589 (2022).
- [44] F. Gygi, Adaptive Riemannian metric for plane-wave electronic-structure calculations, *Europhys. Lett.* **19**, 617 (1992).
- [45] D. R. Hamann, Application of adaptive curvilinear coordinates to the electronic structure of solids, *Phys. Rev. B* **51**, 7337 (1995).
- [46] N. W. Ashcroft and N. D. Mermin, *Solid State Physics*, 1 ed. (Brooks Cole, California, US, 1976).
- [47] D. P. Kingma and J. Ba, Adam: A Method for Stochastic Optimization, [arXiv:1412.6980](https://arxiv.org/abs/1412.6980).
- [48] Y. Brenier, Polar factorization and monotone rearrangement of vector-valued functions, *Commun. Pure Appl. Math.* **44**, 375 (1991).
- [49] A. Cuzzocrea, A. Scemama, W. J. Briels, S. Moroni, and C. Filippi, Variational principles in quantum Monte Carlo: The troubled story of variance minimization, *J. Chem. Theory Comput.* **16**, 4203 (2020).
- [50] M. Dash, J. Feldt, S. Moroni, A. Scemama, and C. Filippi, Excited states with selected configuration interaction-quantum Monte Carlo: Chemically accurate excitation energies and geometries, *J. Chem. Theory Comput.* **15**, 4896 (2019).
- [51] R. B. Lehouc, D. C. Sorensen, and C. Yang, *ARPACK Users Guide: Solution of Large Scale Eigenvalue Problems by Implicitly Restarted Arnoldi Methods* (SIAM, Philadelphia, PA, 1998).
- [52] P. Virtanen *et al.*, and SciPy 1.0 Contributors, SciPy 1.0: Fundamental algorithms for scientific computing in Python, *Nat. Methods* **17**, 261 (2020).
- [53] T. Kato, On the eigenfunctions of many-particle systems in quantum mechanics, *Commun. Pure Appl. Math.* **10**, 151 (1957).
- [54] T. Olsen and K. S. Thygesen, Static correlation beyond the random phase approximation: Dissociating H_2 with the Bethe-Salpeter equation and time-dependent GW, *J. Chem. Phys.* **140**, 164116 (2014).
- [55] K. J. H. Giesbertz, A.-M. Uimonen, and R. van Leeuwen, Approximate energy functionals for one-body reduced density matrix functional theory from many-body perturbation theory, *Eur. Phys. J. B* **91**, 282 (2018).
- [56] R. L. Hall, N. Saad, K. D. Sen, and H. Ciftci, Energies and wave functions for a soft-core Coulomb potential, *Phys. Rev. A* **80**, 032507 (2009).
- [57] M. Iñarra, V. Lanchares, J. F. Palacián, A. I. Pascual, J. P. Salas, and P. Yanguas, Effects of a soft-core Coulomb potential on the dynamics of a hydrogen atom near a metal surface, *Commun. Nonlinear Sci. Numer. Simul.* **68**, 94 (2019).
- [58] R. T. Q. Chen, Y. Rubanova, J. Bettencourt, and D. K. Duvenaud, in 32nd Conference on Neural Information Processing Systems (NeurIPS), Montréal, Canada (2018).
- [59] Q. Zhang, X. Wang, R. Shi, X. Ren, H. Wang, and L. Wang, Neural canonical transformations for quantum anharmonic solids of lithium, *Phys. Rev. Lett.* **134**, 246101 (2025).