

Neural wave functions for high-pressure atomic hydrogen

David Linteau^{1,2,*}, Saverio Moroni³, Giuseppe Carleo^{1,2} and Markus Holzmann⁴¹*Institute of Physics, École Polytechnique Fédérale de Lausanne (EPFL), CH-1015 Lausanne, Switzerland*²*Center for Quantum Science and Engineering, École Polytechnique Fédérale de Lausanne (EPFL), CH-1015 Lausanne, Switzerland*³*CNR-IOM DEMOCRITOS, Istituto Officina dei Materiali and SISSA Scuola Internazionale Superiore di Studi Avanzati, Via Bonomea 265, I-34136 Trieste, Italy*⁴*Univ. Grenoble Alpes, CNRS, LPMCM, 38000 Grenoble, France*

(Received 16 April 2025; revised 16 November 2025; accepted 4 February 2026; published 7 April 2026)

We leverage the power of neural quantum states to describe the ground-state wave function of solid and liquid atomic hydrogen, including both electronic and protonic degrees of freedom. For static protons, the resulting Born-Oppenheimer energies are consistently comparable to or lower than all previous projector Monte Carlo results for systems containing up to 128 hydrogen atoms. The same level of accuracy is preserved upon inclusion of nuclear quantum effects, thus going beyond the Born-Oppenheimer approximation. In addition, our description overcomes major limitations of current wave functions, notably by avoiding any explicit symmetry assumption on the expected quantum crystal and sidestepping efficiency issues of imaginary time evolution with disparate mass scales. As a first application, we examine crystal formation in an extremely high-density region up to pressure-induced melting.

DOI: [10.1103/PhysRevResearch.8.L022007](https://doi.org/10.1103/PhysRevResearch.8.L022007)

Introduction. Atomic and molecular hydrogen have played a fundamental role in the development of quantum mechanics and quantum chemistry. High-pressure experiments have revealed the existence of several unexpected molecular solid phases [1,2] in the long-lasting search for atomic metallic hydrogen [3]. Despite the apparent simplicity, understanding the properties of high-pressure hydrogen and deuterium remains a formidable challenge for advanced first-principles simulation methods [4,5].

Diffusion Monte Carlo (DMC) calculations [6] have provided some of the most accurate zero-temperature descriptions. However, hydrogen simulations remain challenging for two main reasons. First, the fixed-node approximation—inherent to DMC and required to circumvent the fermion sign problem—relies on trial wave functions of exceptional accuracy to capture the subtle physics of high-pressure phases. Second, inclusion of nuclear degrees of freedom substantially increases the computational cost: The Debye energy scale of atomic motion, much smaller than the Fermi energy scale of electrons, requires longer projection times in imaginary time evolution, and the very large proton-to-electron mass ratio ($m_p/m_e \simeq 1836$) implies much slower diffusion of the heavier particles [7]. These difficulties are shared by related methods like reptation Monte Carlo (RMC) and finite-temperature path integral Monte Carlo (PIMC) [8,9].

In this work, we show that variational Monte Carlo (VMC) with neural quantum states (NQS) can overcome these limita-

tions by using a flexible backflow wave function, enhanced by a graph neural network that couples electrons and protons. This approach naturally incorporates nuclear quantum effects beyond the Born-Oppenheimer approximation (BOA), a regime where conventional methods struggle due to mass-scale disparities.

Furthermore, our wave function avoids *ad hoc* symmetry constraints, a crucial feature for exploring high-pressure hydrogen phases where crystalline structures are unknown. As a first application, we study the onset of crystal formation in an ultra-dense region where pressure-induced melting is expected.

Since the first quantum Monte Carlo (QMC) studies on hydrogen [6,7], only a few works have treated electronic and protonic degrees of freedom on equal footing [10–12]. Advances in QMC stem from decoupling electronic and protonic motions within the BOA. This underlies coupled electron ion Monte Carlo (CEIMC) [13] and *ab initio* molecular dynamics (MD) [14] methods, operating at temperatures substantially below the electronic Fermi temperature, but high enough to describe nuclear motion within the adiabatic approximation (i.e., the BOA). Within the BOA, robust QMC trial wave functions are built from Hartree-Fock or density functional theory (DFT) band structure methods [10,11,15,16], and improved with backflow coordinates [17]. The nuclear motion is then obtained via MD or PIMC [4,18,19], or, more recently, using deep network free energy methods [20,21].

However, so far, CEIMC or MD simulations have been performed based on VMC calculations of the BOA energy surface. Although changes due to more accurate RMC or DMC values can be estimated *a posteriori* via reweighting of selected trajectories [22], or based on machine-learned force fields trained on DMC calculations of selected proton configurations [23–26], the accuracy of the standard trial wave

*Contact author: david.linteau@epfl.ch

functions employed remains a major bottleneck of these methods. Therefore, we first benchmark our NQS on static proton configurations obtaining BOA energies that reach or surpass previous DMC/RMC values. This demonstrates the flexibility of backflow-based NQS for high-pressure hydrogen, as seen in studies of other systems such as atoms and molecules [27–35], as well as electron matter [36–42].

We then address the ground state of the full Schrödinger equation of the combined electron-proton system—the “dynamic” case—thereby going beyond the BOA. Unlike BOA wave functions with orbitals tied to a fixed proton configuration, our NQS explicitly incorporates proton motion via a factor capturing zero-point motion. Whereas projection methods, such as DMC and RMC, are burdened by mass-dependent convergence issues (as the kinetic energy of the propagator ties the diffusion of the particles to their inverse mass), in VMC the convergence does not depend on the nuclear mass, because Monte Carlo updates can be addressed on the appropriate scale for each type of particle. Finally, our ansatz accurately describes liquid and solid phases in a translationally invariant form, without imposing a particular crystal symmetry [43,44], which is crucial for high-pressure hydrogen where the phase diagram is not fully determined. We present preliminary results in the region where atomic hydrogen is expected to melt at ultra-high pressures.

Method. We consider N hydrogen atoms in a simulation cell of volume $V = 4\pi N r_s^3 a_0^3/3$, where $r_s = a/a_0$ parametrizes the electronic number density, with a and a_0 being the mean electronic distance and the Bohr radius, respectively. To simplify the notation, we will focus on the case of a cubic box of extension $L = V^{1/3}$ in the following. We denote the electron and proton coordinates by $\{\mathbf{r}_i\}$ and $\{\mathbf{R}_I\}$, respectively. As we simultaneously sample both sets of coordinates using Monte Carlo, a configuration is denoted by $\mathbf{X} \equiv \{\mathbf{r}, \mathbf{R}\} \equiv \{\mathbf{r}_1, \dots, \mathbf{r}_N, \mathbf{R}_1, \dots, \mathbf{R}_N\}$. We periodize the simulation cell so that the Hamiltonian in atomic units reads

$$H = - \sum_{a=1}^{2N} \frac{1}{2m_a} \nabla_{\mathbf{x}_a}^2 + \frac{1}{2} \sum_{\mathbf{n} \in \mathbb{Z}^3} \sum_{a,b=1}^{2N} \frac{q_a q_b}{|\mathbf{x}_a - \mathbf{x}_b + \mathbf{n}L|}, \quad (1)$$

where $\mathbf{x}_a \in \mathbf{X}$, and m_a and q_a are, respectively, the mass and electric charge of particle a . The restriction on the Coulomb sum, denoted by an apostrophe, specifies that $\mathbf{n} = \mathbf{0}$ is omitted when $a = b$. The Ewald procedure [45] is used to evaluate the (conditionally convergent) second term in Eq. (1).

To construct a ground-state trial wave function, it is convenient to consider the following general form:

$$\Psi(\mathbf{X}) = \Phi(\mathbf{R}) \det[\phi_k(\mathbf{r}_i|\mathbf{X})] e^{-U(\mathbf{X})}, \quad (2)$$

where $U(\mathbf{X})$ denotes a symmetric correlation factor, $\Phi(\mathbf{R})$ is a pure nuclear wave function, and $\{\phi_k(\mathbf{r}_i|\mathbf{X})\}$ is a set of electronic orbitals, with $1 \leq k \leq N$, which may symmetrically depend on all coordinates in $\mathbf{X}$ that are different from $\mathbf{r}_i$. Although this form is general for all fermionic ground-state wave functions [46–48], there is no guarantee that the orbitals $\phi_k(\mathbf{r}_i|\mathbf{X})$ can be efficiently represented numerically.

In the first calculations on dense hydrogen [6,7], $U(\mathbf{X})$ was chosen as a sum of pairwise correlation factors in terms of electron-electron (e-e), electron-proton (e-p), and proton-proton (p-p) pseudopotentials. The orbital part of the

wave function consisted of Slater determinants for spin-up and spin-down electrons occupying either plane wave orbitals with wave vectors $\mathbf{k}$ inside the Fermi surface or Gaussians localized at the crystalline lattice sites, denoted by $\{\mathbf{R}_I^{(0)}\}$. Later calculations [11] have used orbitals from DFT within the local-density approximation (LDA), with the orbitals also pinned at $\{\mathbf{R}_I^{(0)}\}$. Exchange effects for protons (and deuterium) are expected to be small in the crystalline phase [6]. Quantum statistical effects on protons have thus been neglected, and the proton crystal was described by the asymmetric (Nosanow) wave function $\Phi(\mathbf{R}) = \prod_I \varphi(|\mathbf{R}_I - \mathbf{R}_I^{(0)}|)$ [49,50], where φ is a Gaussian function with a width variational parameter to optimize. Although DMC stochastically improves the wave function via imaginary time projection, in practice the results are strongly biased by the explicit dependence on the assumed crystalline structure. Full exploration of the nuclear phase space in imaginary time is difficult due to the large mass imbalance between protons and electrons. We note that the effect is amplified when considering deuterium having twice the proton’s mass, obeying bosonic nuclear statistics. Backflow and three-body correlations in hydrogen [17] have shown to improve energies also for the fully dynamic protons at zero temperature [12], but have not been explored further since.

Our wave function builds upon these previous ones by incorporating two important modifications. First, as electronic orbitals we use

$$\phi_{\mathbf{k}}^\sigma(\mathbf{r}_i) = \sum_{\mathbf{n} \in \mathbb{Z}^3} \sum_{I=1}^N e^{i\mathbf{k} \cdot (\mathbf{R}_I + \mathbf{n}L)} \chi(\mathbf{r}_i^\sigma - \mathbf{R}_I - \mathbf{n}L), \quad (3)$$

which coincide with Bloch functions for static protons on a crystalline lattice. For χ , we have chosen a Gaussian with a free width parameter. While these radially symmetric s orbitals seem to be sufficient at describing the range of densities studied in this work, angular dependence via higher orbitals (p, d, f , etc.) can be included in a straightforward manner.

Second, we increase the expressivity of the trial wave function by adding propagator-like backflow to the electron coordinates,

$$\mathbf{r}_i \rightarrow \mathbf{r}_i + W \mathbf{y}_i^{(b)}, \quad (4)$$

where $\mathbf{y}_i^{(b)}$ is the electron vertex output of a message-passing neural network (MPNN) after b iterations and W is a matrix of complex variational parameters. This electron backflow is only added in the orbital part of the wave function. For $U(\mathbf{X})$, we take advantage of the electron edge output of the MPNN, denoted by $\mathbf{Y}_{ij}^{(b)}$, to create a propagator-like term, similarly as in Ref. [44]. With this additional piece, the correlation factor $U(\mathbf{X})$ reads

$$U(\mathbf{x}) \rightarrow U(\mathbf{x}) + \sum_{i < j} [w \text{MLP}(\mathbf{Y}_{ij}^{(b)})], \quad (5)$$

where MLP stands for multilayer perceptron and $w \in \mathbb{R}$ is a variational parameter. The MPNN operates on two underlying graphs: the electron one and the proton one, which interact at each message-passing iteration. In particular, it involves e-e and e-p contributions, meaning that the resulting backflow involves both the electron and proton coordinates, with the e-p contributions capturing a smaller—though important—part of the correlation (see *Appendix A: Message-passing neural network implementation*).

This corresponds to a modified version of the message-passing neural quantum state architecture introduced in Ref. [51]. In the dynamic case, when the protons are not localized to lattice sites, the zero-point motion of the protons is captured with the edge output $\mathbf{Y}_{ij}^{(b)}$ of a different MPNN than the electronic one, which is processed similarly as in Eq. (5). This latter MPNN operates only on a single graph—the proton one.

Although we can reach large system sizes of around 10^2 particles with our implementation (see the Results section), important finite-size effects remain, which can be addressed from strategies discussed in Ref. [52]. Most important, especially for metallic systems, are shell effects due to the sharp Fermi surface. These shell effects can, however, be strongly reduced by employing twist-averaged boundary conditions (TABCs) [53], which, for a displacement of particle i by one lattice vector $L\mathbf{e}_\alpha$ along direction α , impose the following constraint on the wave function:

$$\Psi(\{\mathbf{r}_j + \delta_{ij}L\mathbf{e}_\alpha\}, \mathbf{R}) = e^{i\vartheta_\alpha} \Psi(\{\mathbf{r}_j\}, \mathbf{R}), \quad (6)$$

where $\mathbf{e}_\alpha$ denotes the unit vector in the direction of $\alpha \in \{x, y, z\}$, with $-\pi \leq \vartheta_\alpha < \pi$, and i, j labels any of the electron coordinates. We remark that we have only imposed TABCs on the electronic degrees of freedom. Periodic boundary conditions (PBCs) correspond to the special case where the twist angle is trivial, that is, $\vartheta \equiv (\vartheta_x, \vartheta_y, \vartheta_z) = \mathbf{0}$. We can impose TABCs on the wave function by selecting the “twisted” wave vector $\mathbf{k} = (2\pi\mathbf{m} + \vartheta)/L$, with $\mathbf{m} \in \mathbb{Z}^3$, for the electronic orbitals. In practice, the electron orbitals are then filled by taking the twisted wave vectors with the smallest norm.

Results. As a first step, we want to establish the accuracy of our NQS compared to previous VMC and DMC results. Prior calculations have almost exclusively focused on BOA ground state of electrons using the external (Coulomb) potential of the static protons. Formally, the electronic Hamiltonian, $H_e(\mathbf{R})$, corresponds to infinite proton mass in Eq. (1),

$$H_e(\mathbf{R}) = -\frac{1}{2m_e} \sum_{i=1}^N \nabla_{\mathbf{r}_i}^2 + \frac{1}{2} \sum_{\mathbf{n} \in \mathbb{Z}^3} \sum_{a,b=1}^{2N} \frac{q_a q_b}{|\mathbf{x}_a - \mathbf{x}_b + \mathbf{n}L|}. \quad (7)$$

To obtain upper bounds to the BOA ground-state energies, we minimize $H_e(\mathbf{R})$ over electronic wave functions $\Psi_e(\mathbf{r}|\mathbf{R})$ with purely parametric dependence on $\mathbf{R}$. We have used the same functional form as in Eq. (2), simply treating the protonic coordinates as (external) static parameters. Additional implicit parametric dependence on $\mathbf{R}$ may be introduced via the optimization parameters.

In Table I, we compare our results for static protons localized on a perfect body-centered cubic (BCC) lattice around metallization, at $r_s = 1.31$, with DMC calculations using Slater-Jastrow (SJ) and backflow wave functions with metallic plane wave orbitals [17], as well as with LDA-DFT orbitals [11] and backflow augmented PBE-DFT orbitals (BF-PBE) [54]. Calculations have been performed for $N = 16, 54$, and 128 electrons under PBCs. Our VMC results systematically lower the energies per atom by ~ 1 mHa compared to the best SJ-DMC calculations reported in the literature, and perform similarly to RMC results with BF-PBE wave functions. The quality of the wave function is further quantified by roughly a fivefold reduction in variance compared to the SJ-VMC wave

TABLE I. Static hydrogen with protons pinned to a BCC lattice at $r_s = 1.31$ using PBCs. The energy per atom, E/N , is given in units of hartree, and σ^2 denotes the energy variance. The NQS energy, obtained from VMC calculations using the ansatz given in Eq. (2), is compared to previous DMC reference energies taken from Ref. [17], based on Slater-Jastrow (SJ-PW) and backflow (BF-PW) plane wave orbitals, as well as Slater-Jastrow calculations using DFT-LDA (SJ-LDA) orbitals, and backflow PBE-DFT (BF-PBE) orbitals. For DMC results, we report the energy variance σ^2 of the underlying VMC wave function. Boldface entries indicate the lowest energy value within each set of compared results.

N	Wave function	E/N	σ^2/N
16	SJ-PW (DMC)	−0.4857(1)	0.0773(25)
	SJ-LDA (DMC)	−0.4890(5)	
	BF-PW (DMC)	−0.4905(1)	0.0232(1)
	NQS (VMC)	−0.491 54(1)	0.0062(1)
54	SJ-PW (DMC)	−0.5329(1)	0.0642(9)
	BF-PW (DMC)	−0.5382(1)	0.0222(2)
	SJ-LDA (DMC)	−0.5390(5)	
	BF-PBE (RMC)	−0.540 09(2)	0.005 65(1)
	NQS (VMC)	−0.540 09(2)	0.004 61(3)
128	SJ-PW (DMC)	−0.4900(2)	0.0656(23)
	BF-PW (DMC)	−0.4978(4)	0.030(1)
	SJ-LDA (DMC)	−0.4978(2)	
	BF-PBE (RMC)	−0.499 22(2)	0.006 66(2)
	NQS (VMC)	−0.499 62(2)	0.007 21(2)

function underlying the DMC calculation. Notably, about the same improvement is obtained for all three system sizes using a total of 2889 parameters, illustrating the size consistency of our wave function and of the results.

In Table III (see Appendix B: *Energy comparison for various crystal structures*), we further provide explicit comparisons of twist-averaged calculations for static protons in different crystal structures obtained with RMC in Ref. [16], which also provide DFT orbitals augmented with backflow wave functions, as used in CEIMC calculations. Again, our NQS is able to compete with the best projector Monte Carlo results, in this case RMC based on LDA or optimized independent particle (IPP) orbitals augmented with backflow.

So far, we have shown that our NQS wave function reaches and improves on previous DMC energies from the literature by focusing on high-symmetry configurations of atomic hydrogen. Our NQS state can therefore serve to enhance the accuracy of machine-learned effective potentials for hydrogen atoms [23–26,55–59], to benchmark DFT functionals on representative proton configurations in the atomic regime of the high-pressure phase diagram [60,61], and to further achieve DMC/RMC-level precision of the BOA energy surface in future CEIMC or MD calculations.

We have further explored the expressivity of our ansatz in the molecular phases at lower pressures. Comparing to generic configurations with $N = 96$ protons provided in the dense hydrogen DMC database from Refs. [24,62], spanning molecular solid and liquid configurations at pressures from 50 to 200 GPa, our present NQS architecture remains roughly 2–5 mHa above the DMC reference energies, indicating limitations of our orbitals in the molecular phases. Further work is needed to reduce the systematic bias of our ansatz in the

TABLE II. Ground-state energies for dynamic hydrogen at $r_s = 1.31$, for different wave functions obeying PBCs. The reference VMC and DMC energies for Slater-Jastrow with LDA-DFT orbitals (SJ-LDA) and backflow plane wave orbitals (BF-PW) are taken from Ref. [12]. For all calculations, the nuclear part of the trial wave function is given by Gaussians localized around BCC lattice sites. Boldface entries indicate the lowest energy value within each set of compared results.

N	Wave function	E/N	σ^2/N
16	SJ-LDA (VMC)	-0.467 85(2)	0.030(2)
	BF-PW (VMC)	-0.4724(1)	
	BF-PW (DMC)	-0.4792(1)	
	NQS (VMC)	-0.480 91(5)	
54	SJ-LDA (VMC)	-0.5195(2)	0.025(1)
	SJ-LDA (DMC)	-0.524 15(5)	
	BF-PW (VMC)	-0.521 94(5)	
	BF-PW (DMC)	-0.526 10(7)	
	NQS (VMC)	-0.528 54(9)	
			0.0088(1)

molecular region to simultaneously capture both molecular and atomic phases using a universal ansatz, similar to recent foundation models [31,32,63]. Details on the global optimization scheme that can be used for such universal ansätze, as well as preliminary results, are provided in *Appendix C: Global optimization of a universal wave function*.

We finally present the calculations of ground-state energies with dynamical protons without relying on the BOA. There, electronic and protonic degrees of freedom are sampled according to the full wave function, given in Eq. (2), and energies are obtained from minimizing the expectation value of H , given in Eq. (1). Effects beyond BOA are naturally included in the evaluation of the protons' kinetic energy, acting on the purely nuclear part $\Phi(\mathbf{R})$, as well as on the electronic wave function $\Psi_e(\mathbf{r}|\mathbf{R})$. As a first benchmark, we consider hydrogen at $r_s = 1.31$, with localized Gaussians for the nuclear wave function $\Phi(\mathbf{R})$, so that the proton positions fluctuate around BCC lattice sites. This allows us to faithfully compare in Table II with VMC and DMC results from Ref. [12], based on BF-PW and SJ-LDA trial wave functions. Again, improvement is consistently obtained. To provide more insights into the structure of the optimized wave function, we further show in Fig. 1 pair correlation functions between different particle types (electrons, protons, and spin-up electrons), as well as Monte Carlo configurations in position space showcasing the BCC lattice structure.

We finally fully relax the nuclear wave function—keeping the wave function fully translationally invariant—and explore the ground-state structures of $N = 8$ atoms under increasing compression. The associated nuclear wave function $\Phi(\mathbf{R})$ is constructed with $\mathbf{Y}_{JJ}^{(b)}$, as described above, as well as a p-p pseudopotential. Importantly, it does not involve a product of Gaussian functions localizing the protons at lattice sites. Figure 2 shows the (proton-proton) pair correlation function in three-dimensional space, at various pressures near the expected melting transition. The method used is clearly able to find structural changes when increasing the pressure (or lowering r_s) and the system turns liquid around $r_s \approx 0.2$. However, studies at larger system sizes and different cell geometries are required to precisely determine the crys-

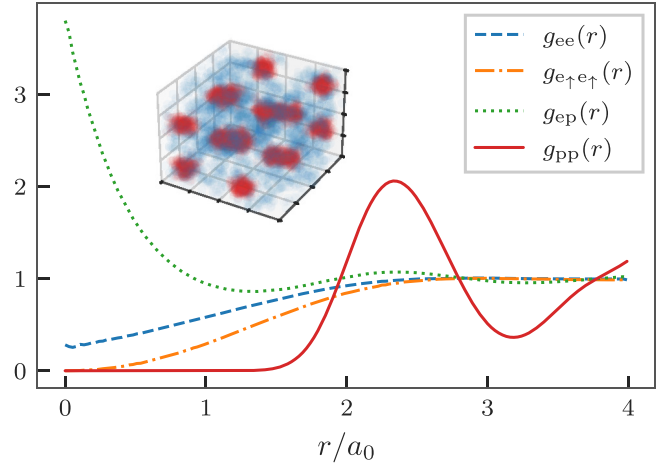


FIG. 1. Four pair correlation functions are shown for $N = 54$ between three different types of particles: “e”, “p”, and “e $_{\uparrow}$ ”, corresponding to electrons, protons, and spin-up electrons, respectively. An inset with 300 Monte Carlo configurations in the BCC crystal is depicted (for $N = 16$, for clarity), where protons form the localized red balls while electrons form the delocalized blue cloud.

talline structures as well as the exact melting density. So far, pressure-induced melting of hydrogen has only been studied from approximate matching of hydrogen to calculations for the one-component plasma [64], or based on an approximate Thomas-Fermi screening model that uses a Yukawa pair potential for the proton-proton interaction [65–67].

The NQS-based approach presented here enables the direct identification and characterization of atomic crystalline hydrogen structures and their melting transition, and is only limited in this study by finite-size effects.

Although melting occurs at ultra-high densities, the interprotonic distance a^p remains larger than the protonic Bohr radius ($a_0^p \equiv a_0 m_e/m_p$), that is, $a^p/a_0^p \simeq 1836 r_s \gg 1$. This justifies neglecting fermionic statistics in the solid phase, which means in practice that no protonic determinants are included in the wave function. Finite-size effects, by imposing TABCs for instance, as well as Fermi statistics for protons and isotope effects, will be addressed in future work.

Conclusion. We have introduced a neural quantum state that accurately describes electron-ion systems. Validated against RMC and DMC results, our method achieves high precision for high-pressure atomic hydrogen both in the Born-Oppenheimer regime and for the full zero-temperature ground state beyond BOA. Our preliminary study demonstrates that our NQS can capture the zero-temperature melting of atomic hydrogen while treating protons and electrons on equal footing. These attributes pave the way for an extensive exploration of the high-pressure phase diagram for both hydrogen and deuterium. Moreover, the incorporation of nuclear quantum statistics opens promising avenues for studying isotope effects and nuclear spin phenomena such as the *ortho-para* transitions in molecular hydrogen.

The simulations were performed using NETKET [68,69], built on JAX [70] and MPI4JAX [71]. The data underlying the figures, together with the code used to generate them, are publicly available in the GitHub repository [72]. The Laplacian was evaluated using the forward Laplacian

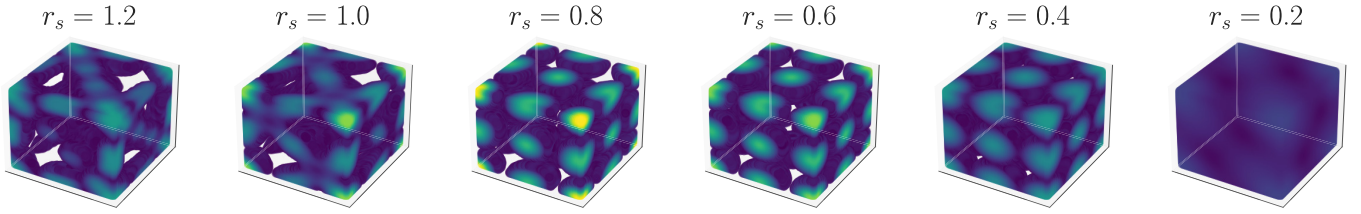


FIG. 2. Pair correlation function for $N = 8$ hydrogen atoms at different densities under PBCs. The same color scale is used in all subplots, globally adjusted with a power law normalization to enhance visibility of the structural features. A fixed threshold is manually set to prevent plotting areas where the pair correlation function has low values. At $r_s = 0.2$, the structure is drastically reduced suggesting that the system turned liquid.

method [73], implemented via either FOLX [74] or FWD-LAP [75].

Acknowledgments. We thank Y. Yang for providing help with the dense hydrogen DMC database [24,62], C. Pierleoni for providing BF-PBE results, and acknowledge useful discussions with G. Mazzola. The authors acknowledge support from SEFRI under Grant No. MB22.00051 (NEQS—Neural Quantum Simulation). This work was supported as part of the “Swiss AI initiative” by a grant from the Swiss National Supercomputing Centre (CSCS) under project ID a117 on Alps.

Data availability. The data that support the findings of this article are openly available [72].

Appendix A: Message-passing neural network implementation. The implementation used in this work is detailed in Algorithm 1. Other recent MPNN implementations for NQS were presented, for instance, in Refs. [44,51,76,77].

Appendix B: Energy comparison for various crystal structures. In Table III, we compare twist-averaged (TABC) calculations for static protons with different crystal structures, as well as different N and r_s , with the best RMC results from Refs. [16,78], obtained either by using backflow augmented metallic (BF-PW), DFT (BF-LDA), or optimized independent particle (BF-IPP) orbitals. All TABC results are based on a $6 \times 6 \times 6$ twist grid. The NQS results were obtained with a $6 \times 6 \times 6$ Monkhorst-Pack grid [79], providing twist grid convergence within ~ 0.5 mHa. We further provide PBC energies to facilitate the comparison for future work. We find that the NQS generally performs as good or better than the RMC wave functions, except for the diamond lattice where NQS energies are slightly higher.

Appendix C: Global optimization of a universal wave function. In developing a foundation model for electronic structure, the objective is to optimize a single set of variational parameters, θ , so that the electronic ground-state wave function $\Psi_\theta(\mathbf{R})$ accurately describes a set of electronic Hamiltonians $\{H_e(\mathbf{R})\}_{\mathbf{R}}$ [with $H_e(\mathbf{R})$ defined in Eq. (7)] associated with different static nuclear configurations $\{\mathbf{R}\}$, with $\mathbf{R} = \{\mathbf{R}_1, \dots, \mathbf{R}_N\}$. In the most general form, the associated global variational energy $E_{\text{global}}(\theta)$ is given by

$$E_{\text{global}}(\theta) = \int d\mathbf{R} \mathcal{P}(\mathbf{R}) \frac{\langle \Psi_\theta(\mathbf{R}) | H_e(\mathbf{R}^{(i)}) | \Psi_\theta(\mathbf{R}) \rangle}{\langle \Psi_\theta(\mathbf{R}) | \Psi_\theta(\mathbf{R}) \rangle}, \quad (\text{C1})$$

where the expectation value $\langle \dots \rangle$ is taken over a sufficiently large set of electronic configurations $\{\mathbf{r}\}$, with $\mathbf{r} = \{\mathbf{r}_1, \dots, \mathbf{r}_N\}$ (assuming the same number of electrons as nuclei without loss of generality), from the conditional probability density $P_\theta(\mathbf{r}|\mathbf{R}) = |\langle \Psi_\theta(\mathbf{R}) | \mathbf{r} \rangle|^2 / \langle \Psi_\theta(\mathbf{R}) | \Psi_\theta(\mathbf{R}) \rangle$. In

practice, the integration over the reference probability distribution $\mathcal{P}(\mathbf{R})$ is performed through Monte Carlo averages obtained using a finite set of M independent and identically distributed (i.i.d.) nuclear configurations drawn from $\mathcal{P}(\mathbf{R})$, that is, $\mathbf{R} \sim \mathcal{P}(\mathbf{R})$. Assuming the nuclear configurations are resampled uniformly from an existing dataset of configurations, $\mathcal{D}$, the integral reduces to the following discrete form:

$$\begin{aligned} E_{\text{global}}(\theta) &= \frac{1}{M} \sum_{i=1}^M \frac{\langle \Psi_\theta(\mathbf{R}^{(i)}) | H_e(\mathbf{R}^{(i)}) | \Psi_\theta(\mathbf{R}^{(i)}) \rangle}{\langle \Psi_\theta(\mathbf{R}^{(i)}) | \Psi_\theta(\mathbf{R}^{(i)}) \rangle} \\ &\equiv \frac{1}{M} \sum_{i=1}^M E(\mathbf{R}^{(i)}, \theta), \end{aligned} \quad (\text{C2})$$

where a uniform distribution $\mathcal{P}(\mathbf{R}) = 1/M$ is implicitly used and each nuclear configuration $\mathbf{R}^{(i)}$ is part of the dataset $\mathcal{D}$, i.e., $\mathbf{R}^{(i)} \in \mathcal{D}$. We further implicitly defined the expectation value of $H_e(\mathbf{R}^{(i)})$ with respect to the globally optimized wave function Ψ_θ , evaluated at the set of (global) variational parameters θ and at one nuclear configuration $\mathbf{R}^{(i)}$, as $E(\mathbf{R}^{(i)}, \theta)$. The global optimization strategy we propose, analogous to the framework described in the foundation neural-network quantum states paper [63], leverages an extended stochastic reconfiguration (SR) scheme [80–82] for multiple auxiliary systems (here each labeled by the nuclear configuration $\mathbf{R}$). For completeness, we rewrite this procedure, simply adapting the notation, which states that the variational parameter update $\delta\theta$ is given by the following matrix equation:

$$S \delta\theta = -\eta \mathbf{F}, \quad (\text{C3})$$

where the quantum geometric tensor S has size $P \times P$, with P the total number of variational parameters, η is the learning rate, and $\mathbf{F} = \nabla_\theta E_{\text{global}}(\theta)$ is the gradient of the loss function. The components of the latter $\mathcal{F}_k \in \mathbf{F}$, with $k \in 1, \dots, P$, are formally defined as $\mathcal{F}_k = \int d\mathbf{R} \mathcal{P}(\mathbf{R}) F_k(\mathbf{R})$, where $F_k(\mathbf{R})$ can be written as a covariance of the form

$$\begin{aligned} F_k(\mathbf{R}) &= -2\text{Re}\{\langle \Delta O_k^\dagger(\mathbf{R}) \Delta H_e(\mathbf{R}) \rangle\} \\ &= -2\text{Re}\{\langle O_k^\dagger(\mathbf{R}) H_e(\mathbf{R}) \rangle - \langle H_e(\mathbf{R}) \rangle \langle O_k^\dagger(\mathbf{R}) \rangle\}, \end{aligned} \quad (\text{C4})$$

where $\Delta A = A - \langle A \rangle$ corresponds to the shifted-mean operator of an arbitrary operator A , and $O_k(\mathbf{R}) = \partial \ln \Psi_\theta(\mathbf{r}, \mathbf{R}) / \partial \theta_k$, with $\theta_k \in \theta$. The extension of the quantum geometric tensor to the extended space is performed similarly, so that $S_{kl} = \int d\mathbf{R} \mathcal{P}(\mathbf{R}) S_{kl}(\mathbf{R})$, with

$$\begin{aligned} S_{kl}(\mathbf{R}) &= \text{Re}\{\langle \Delta O_k^\dagger(\mathbf{R}) \Delta O_l(\mathbf{R}) \rangle\} \\ &= \text{Re}\{\langle O_k^\dagger(\mathbf{R}) O_l(\mathbf{R}) \rangle - \langle O_k^\dagger(\mathbf{R}) \rangle \langle O_l(\mathbf{R}) \rangle\}. \end{aligned} \quad (\text{C5})$$

ALGORITHM 1. MPNN.

Data– Given a Monte Carlo configuration $((\mathbf{r}_1, s_1), \dots, (\mathbf{r}_N, s_N), \mathbf{R}_1, \dots, \mathbf{R}_N)$, where s_i is an electron spin label, the following electron-electron and electron-proton input tensors can be constructed:

$$\mathbf{I}_{ij} = [\sin(2\pi \mathbf{r}_{ij}/L), \cos(2\pi \mathbf{r}_{ij}/L), |\sin(\pi \mathbf{r}_{ij}/L)|, s_i s_j] \in \mathbb{R}^{2d+2}, \quad (\text{A1})$$

$$\mathbf{I}_{il} = [\sin(2\pi \mathbf{r}_{il}/L), \cos(2\pi \mathbf{r}_{il}/L), |\sin(\pi \mathbf{r}_{il}/L)|] \in \mathbb{R}^{2d+1}, \quad (\text{A2})$$

where $\mathbf{r}_{ij} = \mathbf{r}_i - \mathbf{r}_j$, $\mathbf{r}_{il} = \mathbf{r}_i - \mathbf{R}_l$ and the square brackets correspond to the concatenation operation. The following hyperparameters have to be fixed by the user: the hidden dimension h , the attention dimension a , the MLP output dimension e , and the number of message-passing iteration(s) b .

(The vector or matrix dimension is shown below as a comment on the right.)

Initialization– instantiate “hidden” embedding parameters $\mathbf{h}_i^{(0)}, \mathbf{H}_{ij}^{(0)}, \mathbf{H}_{il}^{(0)}$ $\triangleright h$

Define vertex states $\mathbf{y}_i^{(0)} = \mathbf{h}_i^{(0)}$ $\triangleright h$

Define edge states $\mathbf{Y}_{ij}^{(0)} = [\mathbf{I}_{ij}, \mathbf{H}_{ij}^{(0)}], \mathbf{Y}_{il}^{(0)} = [\mathbf{I}_{il}, \mathbf{H}_{il}^{(0)}]$ $\triangleright 2d + 2 + h, 2d + 1 + h$

for $0 \leq \ell \leq b$ **do**

Electron-proton edge update:

Initialize query and key matrices $\mathbf{Q}_{ep}^{(\ell)}, \mathbf{K}_{ep}^{(\ell)}$ $\triangleright \dim(\mathbf{Y}_{il}^{(\ell)}) \times a$

Compute queries and keys $\mathbf{Q}_{il}^{(\ell)} = \mathbf{Q}_{ep}^{(\ell)} \mathbf{Y}_{il}^{(\ell)}, \mathbf{K}_{il}^{(\ell)} = \mathbf{K}_{ep}^{(\ell)} \mathbf{Y}_{il}^{(\ell)}$ $\triangleright a$

Compute attention weights $\mathbf{W}_{il}^{(\ell)} = \text{MLP}_{ep,1}^{(\ell)}(\mathbf{Q}_{il}^{(\ell)} \mathbf{K}_{il}^{(\ell)} / \sqrt{a})_{il}$ $\triangleright e$

Compute messages $\mathbf{M}_{il}^{(\ell)} = \mathbf{W}_{il}^{(\ell)} \odot \text{MLP}_{ep,2}^{(\ell)}(\mathbf{Y}_{il}^{(\ell)})$ $\triangleright e$

Update edge hidden states $\mathbf{H}_{il}^{(\ell+1)} = \text{MLP}_{ep,3}^{(\ell)}([\mathbf{Y}_{il}^{(\ell)}, \mathbf{M}_{il}^{(\ell)}])$ $\triangleright e$

Update edge states $\mathbf{Y}_{il}^{(\ell+1)} = [\mathbf{I}_{il}, \mathbf{H}_{il}^{(\ell+1)}]$ $\triangleright 2d + 1 + e$

Electron-electron edge and vertex update:

Initialize query and key matrices $\mathbf{Q}_{ee}^{(\ell)}, \mathbf{K}_{ee}^{(\ell)}$ $\triangleright \dim(\mathbf{Y}_{ij}^{(\ell)}) \times a$

Compute queries and keys $\mathbf{Q}_{ij}^{(\ell)} = \mathbf{Q}_{ee}^{(\ell)} \mathbf{Y}_{ij}^{(\ell)}, \mathbf{K}_{ij}^{(\ell)} = \mathbf{K}_{ee}^{(\ell)} \mathbf{Y}_{ij}^{(\ell)}$ $\triangleright a$

Compute attention weights $\mathbf{W}_{ij}^{(\ell)} = \text{MLP}_{ee,1}^{(\ell)}(\mathbf{Q}_{ij}^{(\ell)} \mathbf{K}_{ij}^{(\ell)} / \sqrt{a})_{ij}$ $\triangleright e$

Compute messages $\mathbf{M}_{ij}^{(\ell)} = \mathbf{W}_{ij}^{(\ell)} \odot \text{MLP}_{ee,2}^{(\ell)}(\mathbf{Y}_{ij}^{(\ell)})$ $\triangleright e$

Compute vertex cross-species contributions: $\mathbf{C}_i^{(\ell)} = \sum_j \text{MLP}_{ee,3}^{(\ell)}(\mathbf{Y}_{ij}^{(\ell+1)})$ $\triangleright e$

Compute edge cross-species contributions: $\mathbf{C}_{ij}^{(\ell)} = \sum_l \text{MLP}_{ee,4}^{(\ell)}(\mathbf{Y}_{il}^{(\ell+1)}) \odot \text{MLP}_{ee,5}^{(\ell)}(\mathbf{Y}_{jl}^{(\ell+1)})$ $\triangleright e$

Update vertex hidden states $\mathbf{h}_i^{(\ell+1)} = \text{MLP}_{ee,6}^{(\ell)}([\mathbf{h}_i^{(\ell)}, \sum_j \mathbf{M}_{ij}^{(\ell)}, \mathbf{C}_i^{(\ell)}])$ $\triangleright e$

Update edge hidden states $\mathbf{H}_{ij}^{(\ell+1)} = \text{MLP}_{ee,7}^{(\ell)}([\mathbf{Y}_{ij}^{(\ell)}, \mathbf{M}_{ij}^{(\ell)}, \mathbf{C}_{ij}^{(\ell)}])$ $\triangleright e$

Update vertex states $\mathbf{y}_i^{(\ell+1)} = \mathbf{h}_i^{(\ell+1)}$ $\triangleright e$

Update edge states $\mathbf{Y}_{ij}^{(\ell+1)} = [\mathbf{I}_{ij}, \mathbf{H}_{ij}^{(\ell+1)}]$ $\triangleright 2d + 2 + e$

end for

return final vertex and edge states $\mathbf{y}_i^{(b)}, \mathbf{Y}_{ij}^{(b)}, \mathbf{Y}_{il}^{(b)}$

This global optimization scheme contrasts sharply with the conventional approach of individually optimizing the electronic wave function for each nuclear configuration. Although the latter will generally provide more accurate absolute energies, it is exceedingly inefficient for large datasets and requires extensive computational resources and human intervention.

In the context of high-pressure hydrogen, we considered the Dense hydrogen DMC database [24,62], that we denote as $\mathcal{D}$, containing $|\mathcal{D}| = 17\,528$ configurations spanning pressures from 50 to 200 GPa and temperatures from 600 to 2200 K. For preliminary testing, we restrict our attention to the 50 GPa subset $\mathcal{D}_{50} = \{x \in \mathcal{D} : P(x) = 50 \text{ GPa}\}$ with $|\mathcal{D}_{50}| = 2\,469$. A training set $\mathcal{C} \subset \mathcal{D}_{50}$ of size $M = 512$ is selected without imposing any temperature bias, leaving $|\mathcal{D}_{50} \setminus \mathcal{C}| = 1\,957$ configurations outside the learning set. The pressure of $P = 50$ GPa was chosen as it lies farthest from the metallic regime, allowing us to test the wave function beyond the metallic phase where it performs well. We then obtained an optimized set of parameters θ^* by minimizing Eq. (C2) over this subset of configurations $\mathcal{C}$. In particular, the SR scheme

described above was used to optimize all the 2889 parameters included in the trial wave function.

Using a $4 \times 4 \times 4$ twist grid (matching the protocol of Refs. [24,62]), the average difference between the global NQS and DMC reference energies (per electron), given by

$$\Delta E = \frac{1}{M} \sum_{i=1}^M \left(\frac{E(\mathbf{R}^{(i)}, \theta^*)}{N} - \frac{E_{\text{ref}}(\mathbf{R}^{(i)})}{N} \right), \quad (\text{C6})$$

is $\Delta E \sim 4$ mHa for configurations included in the training set and $\Delta E \sim 5$ mHa for configurations excluded from the training set but still tagged with a DFT pressure of $P = 50$ GPa (defining a so-called test set). (Note that we have not corrected for two-body finite-size effects [52] in both NQS and database DMC energies.) As hinted in the main text, the global optimization results’ accuracy is therefore not yet satisfactory, being on average more than 4 mHa away from the reference energies, leaving a lot of place for improvement.

However, in terms of the computation time, the global optimization routine ran once on a single NVIDIA H100 GPU for approximately 72 hours, with an additional 15 hours required

TABLE III. Hydrogen with static protons for various crystal structures. The lattices considered include BCC, face-centered cubic (FCC), diamond (DIAM), and simple cubic (SC). The following acronyms are used: “BF-PW” corresponds to a metallic wave function using backflow plane wave orbitals, “BF-LDA” denotes LDA-DFT orbitals with backflow, “BF-IPP” stands for Independent Particle Potential (IPP) with backflow where the plane wave coefficients of the orbitals are obtained by solving an eigenvalue problem with an effective electron-ion potential. Those reference energies, taken from Refs. [16,78], were obtained with RMC and a $6 \times 6 \times 6$ twist grid. Our VMC results of the NQS wave function are obtained using a $6 \times 6 \times 6$ Monkhorst-Pack [79] twist grid when applying TABCs; values at the Γ -point (PBCs) are provided as reference. Boldface entries indicate the lowest energy value within each set of compared results.

System				$r_s = 1$		$r_s = 1.4$	
N	Lattice	Boundary condition	Wave function	E/N	σ^2/N	E/N	σ^2/N
54	BCC	PBC	NQS (VMC)	-0.422 32(2)	0.0101(1)	-0.550 03(3)	0.003 92(5)
		TABC	BF-PW (RMC)	-0.3721(1)	0.0182(7)		
			BF-IPP (RMC)			-0.5228(1)	0.01413(7)
32	FCC	PBC	NQS (VMC)	-0.375 19(6)	0.012 98(3)	-0.526 13(3)	0.006 56(2)
		TABC	NQS (VMC)	-0.359 27(1)	0.009 83(3)	-0.519 56(7)	0.003 65(2)
			BF-PW (RMC)	-0.3792(1)	0.015 43(4)		
64	DIAM	PBC	BF-IPP (RMC)			-0.5280(1)	0.013 52(5)
			NQS (VMC)	-0.3820(1)	0.012 76(5)	-0.530 83(5)	0.006 04(1)
		TABC	NQS (VMC)	-0.354 73(7)	0.0176(1)	-0.518 87(7)	0.010 27(4)
8	DIAM	PBC	BF-LDA (RMC)	-0.3635(1)	0.0406(1)		
		TABC	BF-IPP (RMC)			-0.5346(1)	0.01740(7)
			NQS (VMC)	-0.359 93(4)	0.012 25(5)	-0.528 26(4)	0.006 84(1)
8	SC	PBC	NQS (VMC)	-0.100 57(7)	0.008 21(3)		
		TABC	BF-PW (RMC)	-0.413 68(6)	0.010 32(2)		
			NQS (VMC)	-0.4223(3)	0.010 26(4)		
8	SC	PBC	NQS (VMC)	-0.251 79(3)	0.010 51(8)		
		TABC	NQS (VMC)	-0.421 2(3)	0.014 74(4)		

to compute 500 test-set energies, for instance, amounting to roughly 100 GPU hours in total. In contrast, a conventional DMC procedure, in which the electronic wave function is

independently optimized for each configuration, can easily require several orders of magnitude more computational time as well as more human oversight.

- [1] H.-K. Mao and R. J. Hemley, Ultrahigh-pressure transitions in solid hydrogen, *Rev. Mod. Phys.* **66**, 671 (1994).
- [2] A. F. Goncharov, R. J. Hemley, and H.-K. Mao, Vibron frequencies of solid H₂ and D₂ to 200 GPa and implications for the P - T phase diagram, *J. Chem. Phys.* **134**, 174501 (2011).
- [3] E. Gregoryanz, C. Ji, P. Dalladay-Simpson, B. Li, R. T. Howie, and H.-K. Mao, Everything you always wanted to know about metallic hydrogen but were afraid to ask, *Matter Radiat. Extremes* **5**, 038101 (2020).
- [4] J. M. McMahon, M. A. Morales, C. Pierleoni, and D. M. Ceperley, The properties of hydrogen and helium under extreme conditions, *Rev. Mod. Phys.* **84**, 1607 (2012).
- [5] M. Bonitz, J. Vorberger, M. Bethkenhagen, M. P. Böhme, D. M. Ceperley, A. Filinov, T. Gawne, F. Graziani, G. Gregori, P. Hamann *et al.*, Toward first principles-based simulations of dense hydrogen, *Phys. Plasmas* **31**, 110501 (2024).
- [6] D. Ceperley and B. Alder, The calculation of the properties of metallic hydrogen using Monte Carlo, *Physica B+C* **108**, 875 (1981).
- [7] D. M. Ceperley and B. J. Alder, Ground state of solid hydrogen at high pressures, *Phys. Rev. B* **36**, 2092 (1987).
- [8] S. Baroni and S. Moroni, Reptation quantum Monte Carlo: A method for unbiased ground-state averages and imaginary-time correlations, *Phys. Rev. Lett.* **82**, 4745 (1999).
- [9] C. Pierleoni, D. Ceperley, B. Bernu, and W. Magro, Equation of state of the hydrogen plasma by path integral Monte Carlo simulation, *Phys. Rev. Lett.* **73**, 2145 (1994).
- [10] V. Natoli, R. M. Martin, and D. M. Ceperley, Crystal structure of atomic hydrogen, *Phys. Rev. Lett.* **70**, 1952 (1993).
- [11] V. Natoli, R. M. Martin, and D. Ceperley, Crystal structure of molecular hydrogen at high pressure, *Phys. Rev. Lett.* **74**, 1601 (1995).
- [12] M. Holzmann, C. Pierleoni, and D. M. Ceperley, Coupled electron-ion Monte Carlo calculations of atomic hydrogen, *Comput. Phys. Commun.* **169**, 421 (2005).
- [13] C. Pierleoni, D. Ceperley, and M. Holzmann, Coupled electron-ion Monte Carlo calculations of dense metallic hydrogen, *Phys. Rev. Lett.* **93**, 146402 (2004).
- [14] C. Attaccalite and S. Sorella, Stable liquid hydrogen at high pressure by a novel *ab initio* molecular-dynamics calculation, *Phys. Rev. Lett.* **100**, 114501 (2008).
- [15] X. W. Wang, J. Zhu, S. G. Louie, and S. Fahy, Magnetic structure and equation of state of bcc solid hydrogen: A variational quantum Monte Carlo study, *Phys. Rev. Lett.* **65**, 2414 (1990).
- [16] C. Pierleoni, K. T. Delaney, M. A. Morales, D. M. Ceperley, and M. Holzmann, Trial wave functions for high-pressure metallic hydrogen, *Comput. Phys. Commun.* **179**, 89 (2008).

- [17] M. Holzmann, D. M. Ceperley, C. Pierleoni, and K. Esler, Backflow correlations for the electron gas and metallic hydrogen, *Phys. Rev. E* **68**, 046707 (2003).
- [18] C. Pierleoni and D. Ceperley, The coupled electron-ion Monte Carlo method, in *Computer Simulations in Condensed Matter Systems: From Materials to Chemical Biology*, edited by M. Ferrario, G. Ciccotti, and K. Binder (Springer, Berlin, 2006), Vol. 1, pp. 641–683.
- [19] M. Holzmann, High-pressure phases of hydrogen, in *Correlations and Phase Transitions*, edited by E. Pavarini and E. Koch (Verlag des Forschungszentrums Jülich, Jülich, 2024).
- [20] H. Xie, Z.-H. Li, H. Wang, L. Zhang, and L. Wang, Deep variational free energy approach to dense hydrogen, *Phys. Rev. Lett.* **131**, 126501 (2023).
- [21] X. Dong, H. Xie, Y. Chen, W. Liang, L. Zhang, L. Wang, and H. Wang, Deep variational free energy prediction of dense hydrogen solid at 1200 K, *Phys. Rev. B* **111**, 214118 (2025).
- [22] M. Ruggeri, M. Holzmann, D. M. Ceperley, and C. Pierleoni, Quantum Monte Carlo determination of the principal Hugoniot of deuterium, *Phys. Rev. B* **102**, 144108 (2020).
- [23] A. Tirelli, G. Tenti, K. Nakano, and S. Sorella, High-pressure hydrogen by machine learning and quantum Monte Carlo, *Phys. Rev. B* **106**, L041105 (2022).
- [24] H. Niu, Y. Yang, S. Jensen, M. Holzmann, C. Pierleoni, and D. M. Ceperley, Stable solid molecular hydrogen above 900 K from a machine-learned potential trained with diffusion quantum Monte Carlo, *Phys. Rev. Lett.* **130**, 076102 (2023).
- [25] S. Goswami, S. Jensen, Y. Yang, M. Holzmann, C. Pierleoni, and D. M. Ceperley, High temperature melting of dense molecular hydrogen from machine-learning interatomic potentials trained on quantum Monte Carlo, *J. Chem. Phys.* **162**, 054118 (2025).
- [26] K. K. Ly and D. M. Ceperley, Melting curves of atomic hydrogen and deuterium calculated using path-integral Monte Carlo, *Phys. Rev. B* **111**, 104102 (2025).
- [27] M. Holzmann and S. Moroni, Orbital-dependent backflow wave functions for real-space quantum Monte Carlo, *Phys. Rev. B* **99**, 085121 (2019).
- [28] D. Pfau, J. S. Spencer, A. G. Matthews, and W. M. C. Foulkes, *Ab initio* solution of the many-electron Schrödinger equation with deep neural networks, *Phys. Rev. Res.* **2**, 033429 (2020).
- [29] J. Hermann, Z. Schätzle, and F. Noé, Deep-neural-network solution of the electronic Schrödinger equation, *Nat. Chem.* **12**, 891 (2020).
- [30] N. Gao and S. Günnemann, Generalizing neural wave functions, in *Proceedings of the 40th International Conference on Machine Learning*, edited by A. Krause, and E. Brunskill, and K. Cho, and B. Engelhardt, and S. Sabato, and J. Scarlett, Proceedings of Machine Learning Research (PMLR, 2023), Vol. 202, pp. 10708–10726.
- [31] N. Gao and S. Günnemann, Neural Pfaffians: Solving many many-electron Schrödinger equations, [arXiv:2405.14762](https://arxiv.org/abs/2405.14762).
- [32] M. Scherbela, L. Gerard, and P. Grohs, Towards a transferable fermionic neural wavefunction for molecules, *Nat. Commun.* **15**, 120 (2024).
- [33] Y. Zhang, B. Jiang, and H. Guo, Schrödingernet: A universal neural network solver for the Schrödinger equation, *J. Chem. Theory Comput.* **21**, 670 (2025).
- [34] J. Hermann, J. Spencer, K. Choo, A. Mezzacapo, W. M. C. Foulkes, D. Pfau, G. Carleo, and F. Noé, *Ab initio* quantum chemistry with neural-network wavefunctions, *Nat. Rev. Chem.* **7**, 692 (2023).
- [35] M. Scherbela, N. Gao, P. Grohs, and S. Günnemann, Accurate *ab-initio* neural-network solutions to large-scale electronic structure problems, [arXiv:2504.06087](https://arxiv.org/abs/2504.06087).
- [36] M. Holzmann and S. Moroni, Itinerant-electron magnetism: The importance of many-body correlations, *Phys. Rev. Lett.* **124**, 206404 (2020).
- [37] M. Wilson, S. Moroni, M. Holzmann, N. Gao, F. Wudarski, T. Vegge, and A. Bhowmik, Neural network ansatz for periodic wave functions and the homogeneous electron gas, *Phys. Rev. B* **107**, 235139 (2023).
- [38] X. Li, Z. Li, and J. Chen, *Ab initio* calculation of real solids via neural network ansatz, *Nat. Commun.* **13**, 7895 (2022).
- [39] G. Cassella, H. Sutterud, S. Azadi, N. D. Drummond, D. Pfau, J. S. Spencer, and W. M. C. Foulkes, Discovering quantum phase transitions with fermionic neural networks, *Phys. Rev. Lett.* **130**, 036401 (2023).
- [40] C. Smith, Y. Chen, R. Levy, Y. Yang, M. A. Morales, and S. Zhang, Unified variational approach description of ground-state phases of the two-dimensional electron gas, *Phys. Rev. Lett.* **133**, 266504 (2024).
- [41] D. Luo, D. D. Dai, and L. Fu, Simulating moiré quantum matter with neural network, [arXiv:2406.17645](https://arxiv.org/abs/2406.17645).
- [42] D. Luo, T. Zaklaman, and L. Fu, Solving fractional electron states in twisted MoTe₂ with deep neural networks, [arXiv:2503.13585](https://arxiv.org/abs/2503.13585).
- [43] M. Ruggeri, S. Moroni, and M. Holzmann, Nonlinear network description for many-body quantum systems in continuous space, *Phys. Rev. Lett.* **120**, 205302 (2018).
- [44] D. Linteau, G. Pescia, J. Nys, G. Carleo, and M. Holzmann, Phase diagram and crystal melting of helium-4 in two dimensions, *Phys. Rev. Lett.* **134**, 246001 (2025).
- [45] P. P. Ewald, Die Berechnung optischer und elektrostatischer Gitterpotentiale, *Ann. Phys. (NY)* **369**, 253 (1921).
- [46] M. Taddei, M. Ruggeri, S. Moroni, and M. Holzmann, Iterative backflow renormalization procedure for many-body ground-state wave functions of strongly interacting normal Fermi liquids, *Phys. Rev. B* **91**, 115106 (2015).
- [47] D. M. Ceperley, Fermion nodes, *J. Stat. Phys.* **63**, 1237 (1991).
- [48] D. Luo and B. K. Clark, Backflow transformations via neural networks for quantum many-body wave functions, *Phys. Rev. Lett.* **122**, 226401 (2019).
- [49] L. H. Nosanow, Theory of crystalline He³ at 0 °K, *Phys. Rev. Lett.* **13**, 270 (1964).
- [50] J.-P. Hansen and D. Levesque, Ground state of solid helium-4 and -3, *Phys. Rev.* **165**, 293 (1968).
- [51] G. Pescia, J. Nys, J. Kim, A. Lovato, and G. Carleo, Message-passing neural quantum states for the homogeneous electron gas, *Phys. Rev. B* **110**, 035108 (2024).
- [52] M. Holzmann, R. C. Clay, M. A. Morales, N. M. Tubman, D. M. Ceperley, and C. Pierleoni, Theory of finite size effects for electronic quantum Monte Carlo calculations of liquids and solids, *Phys. Rev. B* **94**, 035126 (2016).
- [53] C. Lin, F. H. Zong, and D. M. Ceperley, Twist-averaged boundary conditions in continuum quantum Monte Carlo algorithms, *Phys. Rev. E* **64**, 016702 (2001).

- [54] M. A. Morales, R. Clay, C. Pierleoni, and D. M. Ceperley, First principles methods: A perspective from quantum Monte Carlo, *Entropy* **16**, 287 (2014).
- [55] H. Zong, H. Wiebe, and G. J. Ackland, Understanding high pressure molecular hydrogen with a hierarchical machine-learned potential, *Nat. Commun.* **11**, 5014 (2020).
- [56] B. Cheng, G. Mazzola, C. J. Pickard, and M. Ceriotti, Evidence for supercritical behaviour of high-pressure liquid hydrogen, *Nature (London)* **585**, 217 (2020).
- [57] V. V. Karasiev, J. Hinz, S. Hu, and S. Trickey, On the liquid–liquid phase transition of dense hydrogen, *Nature (London)* **600**, E12 (2021).
- [58] M. Istas, S. Jensen, Y. Yang, M. Holzmann, C. Pierleoni, and D. M. Ceperley, Liquid-liquid phase transition of hydrogen and its critical point: Analysis from *ab initio* simulation and a machine-learned potential, *Phys. Rev. E* **111**, 045307 (2025).
- [59] G. Tenti, B. Jäckl, K. Nakano, M. Rupp, and M. Casula, Hydrogen liquid-liquid transition from first principles and machine learning, *Phys. Rev. B* **112**, 104208 (2025).
- [60] R. C. Clay, M. Holzmann, D. M. Ceperley, and M. A. Morales, Benchmarking density functionals for hydrogen-helium mixtures with quantum Monte Carlo: Energetics, pressures, and forces, *Phys. Rev. B* **93**, 035121 (2016).
- [61] C. Cozza, K. Nakano, S. Howard, H. Xie, R. Helled, and G. Mazzola, Denser hydrogen inferred from first-principles simulations challenges Jupiter’s interior models, *Phys. Rev. Res.* **8**, 013089 (2026).
- [62] Y. Yang, S. Jensen, D. Ceperley, K. Kowalik, and M. Turk, Dense hydrogen diffusion Monte Carlo database 2022, <https://qmc-hamm.hub.yt/data.html>.
- [63] R. Rende, L. L. Viteritti, F. Becca, A. Scardicchio, A. Laio, and G. Carleo, Foundation neural-networks quantum states as a unified ansatz for multiple Hamiltonians, *Nat. Commun.* **16**, 7213 (2025).
- [64] M. D. Jones and D. M. Ceperley, Crystallization of the one-component plasma at finite temperature, *Phys. Rev. Lett.* **76**, 4572 (1996).
- [65] K. K. Mon, G. V. Chester, and N. W. Ashcroft, Simulation studies of a model of high-density metallic hydrogen, *Phys. Rev. B* **21**, 2641 (1980).
- [66] D. M. Ceperley, Quantum Monte Carlo simulation of systems at high pressure, in *Simple Molecular Systems at Very High Density*, edited by P. L. A. Polian and N. Boccara (Plenum Press, New York, 1989).
- [67] B. Militzer and R. L. Graham, Simulations of dense atomic hydrogen in the Wigner crystal phase, *J. Phys. Chem. Solids* **67**, 2136 (2006).
- [68] G. Carleo, K. Choo, D. Hofmann, J. E. Smith, T. Westerhout, F. Alet, E. J. Davis, S. Efthymiou, I. Glasser, S.-H. Lin, M. Mauri, G. Mazzola, C. B. Mendl, E. van Nieuwenburg, O. O’Reilly, H. Théveniaut, G. Torlai, F. Vicentini, and A. Wietek, NetKet: A machine learning toolkit for many-body quantum systems, *SoftwareX* **10**, 100311 (2019).
- [69] F. Vicentini, D. Hofmann, A. Szabó, D. Wu, C. Roth, C. Giuliani, G. Pescia, J. Nys, V. Vargas-Calderón, N. Astrakhantsev, and G. Carleo, NetKet 3: Machine learning toolbox for many-body quantum systems, *SciPost Phys. Codebases* **7** (2022).
- [70] J. Bradbury, R. Frostig, P. Hawkins, M. J. Johnson, C. Leary, D. Maclaurin, G. Necula, A. Paszke, J. VanderPlas, S. Wanderman-Milne, and Q. Zhang, JAX: composable transformations of Python+NumPy programs (2025), <http://github.com/google/jax/>.
- [71] D. Häfner and F. Vicentini, mpi4jax: Zero-copy MPI communication of JAX arrays, *J. Open Source Software* **6**, 3419 (2021).
- [72] D. Linteau, Metallic hydrogen (2026), <https://github.com/cqsl/metallic-hydrogen>.
- [73] R. Li, H. Ye, D. Jiang, X. Wen, C. Wang, Z. Li, X. Li, D. He, J. Chen, W. Ren, and L. Wang, A computational framework for neural network-based variational Monte Carlo with forward Laplacian, *Nat. Mach. Intell.* **6**, 209 (2024).
- [74] N. Gao, J. Köhler, and A. Foster, FOLX—Forward Laplacian for JAX (2023), <http://github.com/microsoft/folx>.
- [75] Y. Chen, Poor man’s forward Laplacian (using JAX tracer!) (2024), <https://github.com/y1xiaoc/fwdlap>.
- [76] J. Kim, G. Pescia, B. Fore, J. Nys, G. Carleo, S. Gandolfi, M. Hjorth-Jensen, and A. Lovato, Neural-network quantum states for ultra-cold Fermi gases, *Commun. Phys.* **7**, 148 (2024).
- [77] J. Nys, G. Pescia, A. Sinibaldi, and G. Carleo, *Ab-initio* variational wave functions for the time-dependent many-electron Schrödinger equation, *Nat. Commun.* **15**, 9404 (2024).
- [78] C. Pierleoni, K. T. Delaney, M. A. Morales, D. M. Ceperley, and M. Holzmann, Progress in coupled electron-ion Monte Carlo simulations of high-pressure hydrogen, in *Quantum Simulations of Complex Many-Body Systems: From Theory to Algorithms*, Series on Advances in Quantum Many-Body Theory, Vol. 10, edited by J. Grotendorst, and D. Marx, and A. Muramatsu, (World Scientific Publishing, Singapore, 2008), pp. 217–232.
- [79] H. J. Monkhorst and J. D. Pack, Special points for Brillouin-zone integrations, *Phys. Rev. B* **13**, 5188 (1976).
- [80] S. Sorella, Green function Monte Carlo with stochastic reconfiguration, *Phys. Rev. Lett.* **80**, 4558 (1998).
- [81] S. Sorella, Wave function optimization in the variational Monte Carlo method, *Phys. Rev. B* **71**, 241103(R) (2005).
- [82] F. Becca and S. Sorella, *Quantum Monte Carlo Approaches for Correlated Systems* (Cambridge University Press, Cambridge, UK, 2017).