


# Ab initio calculation of real solids via neural network ansatz

Received: 24 May 2022

Xiang Li<sup>1</sup>✉, Zhe Li<sup>1</sup> & Ji Chen<sup>2</sup>

Accepted: 13 December 2022

Published online: 22 December 2022

 Check for updates

Neural networks have been applied to tackle many-body electron correlations for small molecules and physical models in recent years. Here we propose an architecture that extends molecular neural networks with the inclusion of periodic boundary conditions to enable ab initio calculation of real solids. The accuracy of our approach is demonstrated in four different types of systems, namely the one-dimensional periodic hydrogen chain, the two-dimensional graphene, the three-dimensional lithium hydride crystal, and the homogeneous electron gas, where the obtained results, e.g. total energies, dissociation curves, and cohesive energies, reach a competitive level with many traditional ab initio methods. Moreover, electron densities of typical systems are also calculated to provide physical intuition of various solids. Our method of extending a molecular neural network to periodic systems can be easily integrated into other neural network structures, highlighting a promising future of ab initio solution of more complex solid systems using neural network ansatz, and more generally endorsing the application of machine learning in materials simulation and condensed matter physics.

Solving the many-body electronic structure of real solids from ab initio is one of the grand challenges in condensed matter physics and materials science<sup>1</sup>. More accurate ab initio solutions can push the limit of our understanding of many fundamental and mysterious emergent phenomena, such as superconductivity, light-matter interaction, and heterogeneous catalysis, to name just a few<sup>2</sup>. The current workhorse method is density functional theory (DFT), whose accuracy depends quite sensitively on the choice of the so-called exchange-correlation functional and unfortunately there lacks a systematic routine towards the exact<sup>3,4</sup>. Other commonly used ab initio quantum chemistry methods, such as the coupled-cluster and configuration interaction theories<sup>5</sup>, can provide more accurate solutions for molecules but face severe difficulty when applied to solid systems due to their high computational complexity. Recently, several breakthroughs have been made in applying these quantum chemistry methods on solids<sup>6,7</sup>, driving the study of solid systems towards a new frontier.

Meanwhile, in the last few years, many attempts to tackle the correlated wavefunction problem in molecules or model Hamiltonians

using neural network-based approaches have been reported by different groups<sup>8–16</sup>. The key idea is to use the neural network as the wavefunction ansatz in quantum Monte Carlo (QMC) simulations. The stochastic nature of QMC enables a considerably economical time scaling and efficient parallelization<sup>6,17–19</sup>. The universal approximation theorem behind neural network-based ansatz significantly improves the accuracy of the traditional QMC method. This strategy has been proved successful in studying small molecules<sup>10–13</sup> in the first and second quantization, and solids in the second quantization<sup>14</sup>. However, how to apply such neural network ansatz for real solids in continuous space, and whether it can describe the long-range electron correlations in extended systems remain as open questions.

Here we propose a powerful periodic neural network ansatz for solids, which combines periodic distance features<sup>20</sup> with existing molecule neural networks<sup>10</sup>. Based on that, we develop a highly efficient QMC method for ab initio calculation of real solid and general periodic systems with high accuracy. We apply our method to periodic hydrogen chains, graphene, lithium hydride (LiH) crystals, and

<sup>1</sup>ByteDance Inc, Zhonghang Plaza, No. 43, North 3rd Ring West Road, Haidian District, Beijing, China. <sup>2</sup>School of Physics, Interdisciplinary Institute of Light-Element Quantum Materials, Frontiers Science Center for Nano-Optoelectronics, Peking University, Beijing 100871, P. R. China.

✉ e-mail: [lixiang.62770689@bytedance.com](mailto:lixiang.62770689@bytedance.com)

homogeneous electron gas. These systems cover a wide range of interests, including materials dimension from one to three, electronic structures from metallic to insulating, and bonding types from covalent to ionic. Standard techniques are employed to reduce finite-size errors. The calculated dissociation curve, cohesive energy and correlation energy, can be compared satisfactorily with available experimental values and other state-of-the-art computational approaches. Electron densities of typical systems are further calculated to test our neural network and explore the underlying physics. All the results demonstrate that our method can achieve accurate electronic structure calculations of solid/periodic systems. In parallel to our work, refs. 21, 22 also developed periodic versions of neural networks to study the homogeneous electron gas system and obtained high-accuracy results. A more detailed comparison is discussed in the following sections.

## Results

### Neural network for a solid system

Periodicity and anti-symmetry are two fundamental properties of the wavefunction of a solid system. The anti-symmetry can be ensured by the Slater determinant, which has been commonly used as the basic block in molecular neural networks. We also approximate the wavefunction by two Slater determinants of one spin-up channel and one spin-down channel,

$$\psi(\mathbf{r}) = \text{Det}_\uparrow [e^{i\mathbf{k}\cdot\mathbf{r}_i} u_{\text{mol}}^\uparrow(d)] \text{Det}_\downarrow [e^{i\mathbf{k}\cdot\mathbf{r}_i} u_{\text{mol}}^\downarrow(d)]. \quad (1)$$

In this regard, our ansatz resembles the structure of FermiNet<sup>10,11</sup>, whereas other neural network wavefunction ansatz may include extra terms in addition to the Slater determinants<sup>12</sup>. Each determinant is then constructed from a set of periodic orbitals, which inherits the physics captured by the generalized collective Bloch function formed by a product of phase factor  $e^{i\mathbf{k}\cdot\mathbf{r}}$  and collective molecular orbital  $u_{\text{mol}}$ . The generalized many-body Bloch function incorporates electron correlations and goes beyond single-electron approximation<sup>18</sup>.

Figure 1 displays more details on the structure of our neural network. Building an efficient and accurate periodic ansatz is the key step in developing ab initio methods for solids. Here we have followed the

recently proposed scheme of Whitehead et al. to construct a set of periodic distance features  $d(\mathbf{r})$ <sup>20</sup> using lattice vectors in real and reciprocal space ( $\mathbf{a}_i, \mathbf{b}_i$ ),

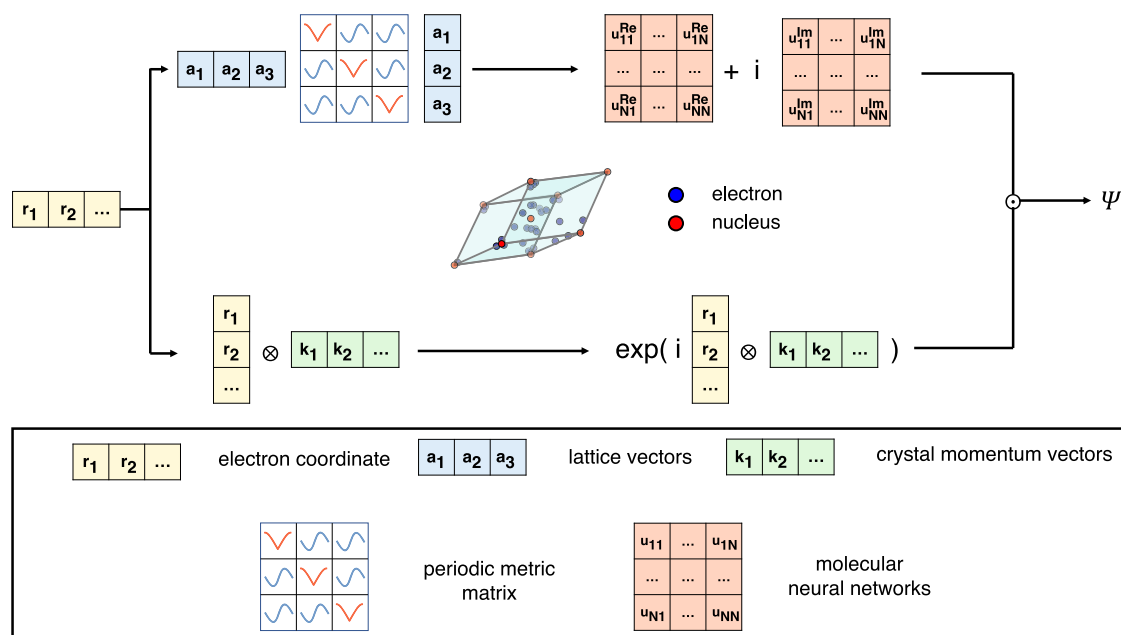
$$d(\mathbf{r}) = \frac{\sqrt{\mathbf{A}\mathbf{M}\mathbf{A}^T}}{2\pi}, \quad \mathbf{A} = (\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3), \quad (2)$$

$$\mathbf{M}_{ij} = f^2(\omega_i)\delta_{ij} + g(\omega_i)g(\omega_j)(1 - \delta_{ij}), \quad \omega_i = \mathbf{r} \cdot \mathbf{b}_i.$$

The periodic metric matrix  $\mathbf{M}$  is constructed by periodic functions  $f, g$ , which retain ordinary distances at the origin and regulate them to periodic ones at far distances, ensuring asymptotic cusp form, continuity, and periodicity requirement at the same time.

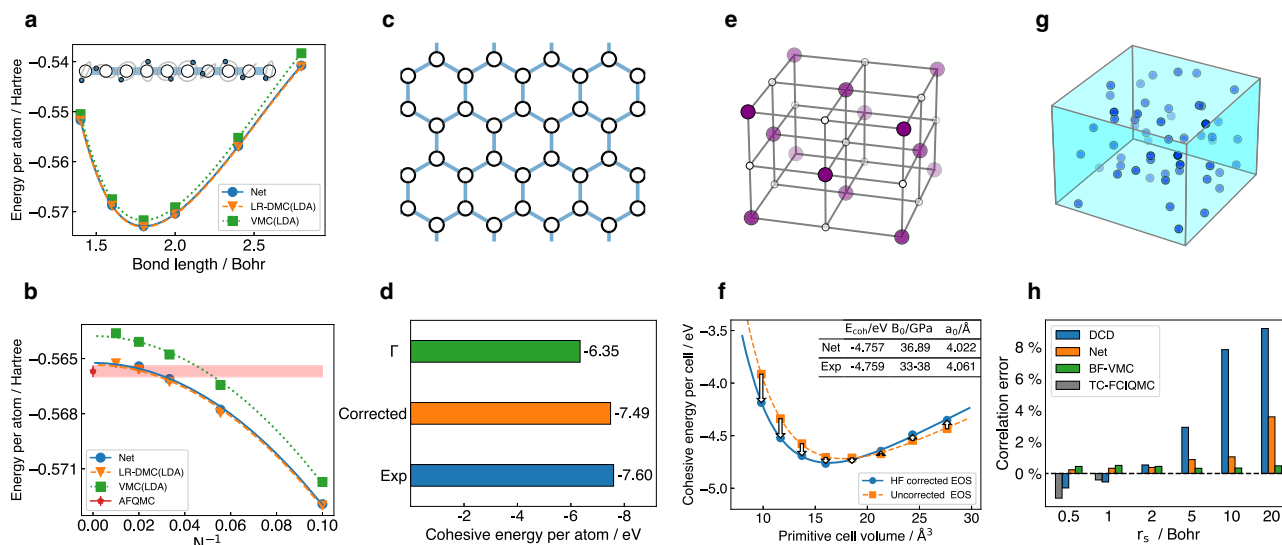
The constructed periodic distance features  $d(\mathbf{r})$  can then be fed into molecular neural networks to form collective orbitals  $u_{\text{mol}}$ . Specifically, in this work, we represent the molecular networks with FermiNet<sup>10</sup>, which incorporates electron–electron interactions. The inclusion of all-electron features promotes the traditional single-particle orbitals to the collective ones, and hence the description of wavefunction and correlation effects can be improved while fewer Slater determinants are required. In addition, the wavefunction of solid systems is necessarily complex-valued, and we introduce two sets of molecular orbitals to represent the real and imaginary parts of the solid wavefunction, respectively. The plane-wave phase factors  $e^{i\mathbf{k}\cdot\mathbf{r}}$  in Fig. 1 are used to construct the Bloch function-like orbitals, and the corresponding  $\mathbf{k}$  points are selected to minimize the Hartree–Fock (HF) energy.

Based on the variational principle, our neural network is trained using the variational Monte Carlo (VMC) approach. To efficiently optimize the network, a Kronecker-factored curvature estimator (KFAC) optimizer<sup>23</sup> implemented by DeepMind team<sup>24</sup> is modified and adopted, which significantly outperforms traditional energy minimization algorithms. Calculations are also ensured by efficient and massive parallelization on multiple nodes of high-performance GPUs. More details on the theories, methods, and computations are included in the Methods section and the supplementary information.



**Fig. 1 | Sketch of neural network architecture.** The electron coordinates  $\mathbf{r}_i$  are passed to two channels. In the first one, they build the periodic distance features  $d(\mathbf{r})$  using the periodic metric matrix  $\mathbf{M}$  and the lattice vectors  $\mathbf{a}$ , and then  $d(\mathbf{r})$  features are fed into two molecular neural networks, that represent separately the

real and the imaginary part of the wavefunction. In the second channel,  $\mathbf{r}_i$  constructs the plane-wave phase factors on a selected set of crystal momentum vectors. The total wavefunctions of solids are constructed by the two channels following the expression of Eq. (1).



**Fig. 2 | Calculated results of neural network.** Our results are all labeled as Net. Statistical errors are negligible for the presented data. **a**  $H_{10}$  dissociation curve is plotted. **b** energy of different hydrogen chain sizes  $N$ , the bond length of the hydrogen chain is fixed at 1.8 Bohr. LR-DMC and VMC both use the cc-pVTZ basis set, and the one-body Jastrow function uses orbitals from LDA calculations. AFQMC is pushed to complete the basis limit. All the comparison results are taken from ref. 25. **c** Structure of graphene. **d** the cohesive energy per atom of  $\Gamma$  point and finite-size error corrected result is plotted. Experiment cohesive energy is from ref. 29. Graphene is calculated at its equilibrium length 1.421 Å. **e** Structure of rock-salt

lithium hydride crystal. **f** Equation of state of LiH crystal is plotted, fitted Birch–Murnaghan parameters and experimental data are also given. HF corrections are calculated using cc-pVDZ basis, and  $E_{\infty}^{\text{HF}}$  is approximated by  $E_{N=8}^{\text{HF}}$ . The arrows denote the corresponding HF corrections. **g** Plot of homogeneous electron gas system. **h** Correlation error of 54-electrons HEG systems at different  $r_s$ . Correlation error is defined as  $[1 - (E - E_{\text{HF}})/(E_{\text{ref}} - E_{\text{HF}})] \times 100\%$ , and  $E_{\text{HF}}$  is taken from ref. 33. DCD, BF-VMC, and TC-FCIQMC results are displayed for comparison, and BF-DMC data were used as reference<sup>33,34</sup>.

## Hydrogen chain

Hydrogen chain is one of the simplest models in condensed matter research. Despite its simplicity, the hydrogen chain is a challenging and interesting system, serving as a benchmark system for electronic structure methods and featuring intriguing correlated phenomena<sup>25</sup>. The calculated energy of the periodic  $H_{10}$  chain as a function of the bond length is shown in Fig. 2a. The results from lattice-regularized diffusion Monte Carlo (LR-DMC) and traditional VMC are also plotted for comparison<sup>25</sup>. We can see that our results nearly coincide with the LR-DMC results and significantly outperform traditional VMC (see Supplementary Table 3). In Fig. 2b, the energy of hydrogen chains of different atom numbers are calculated for extrapolation to the thermodynamic limit (TDL). The shaded bar in Fig. 2b illustrates the extrapolated energy of the periodic hydrogen chain at TDL from auxiliary field quantum Monte Carlo (AFQMC), which is considered as the current state-of-the-art along with LR-DMC. Our TDL result is comparable with both AFQMC and LR-DMC (see Supplementary Table 4).

## Graphene

Graphene is arguably the most famous two-dimensional system (Fig. 2c) receiving broad attention in the past two decades for its mechanical, electronic, and chemical applications<sup>26</sup>. Here we carry out simulations to estimate its cohesive energy, which measures the strength of C-C chemical bonding and long-range dispersion interactions. The calculations are performed on a  $2 \times 2$  supercell of graphene using twist average boundary condition (TABC)<sup>27</sup> in conjunction with structure factor  $S(\mathbf{k})$  correction<sup>28</sup> (see Supplementary Fig. 3) to reduce the finite-size error. The calculated results are plotted in Fig. 2d along with the experimental value<sup>29</sup>, and it shows that our neural network can deal with graphene very well, producing a cohesive energy of graphene within 0.1 eV/atom to the experimental reference (see Supplementary Table 6). We also plotted the results with periodic boundary conditions (PBC), namely the  $\Gamma$  point-only result, which deviates from the experiment data by 1.25 eV/atom.

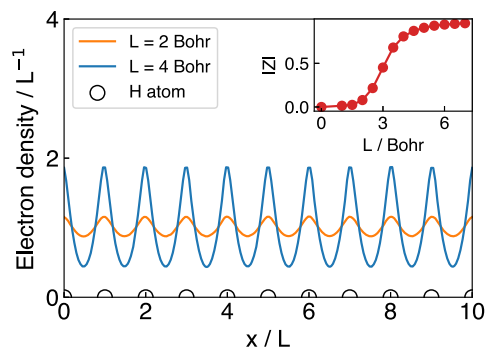
## Lithium hydride crystal

For a three-dimensional system, we consider the LiH crystal with a rock-salt structure (Fig. 2e), another benchmark system for accurate ab initio methods<sup>6,30,31</sup>. Despite consisting of only simple elements, LiH represents typical ionic and covalent bonds upon changing the lattice constants. Using our neural network, we first simulate the equation of the state of LiH on a  $2 \times 2 \times 2$  supercell, as shown in Fig. 2f. In addition, we employ a standard finite-size correction based on Hartree–Fock calculations of a large supercell (see Supplementary Fig. 5). In Fig. 2f we also show the Birch–Murnaghan fitting to the equation of state, based on which we can obtain thermodynamic quantities such as the cohesive energy, the bulk modulus, and the equilibrium lattice constant of LiH. As shown in the inset, our results on the thermodynamic quantities agree very well with experimental data<sup>30</sup> (see Supplementary Table 8, 9).

For further validation, we have also computed directly the  $3 \times 3 \times 3$  supercell of LiH at its equilibrium length of 4.061 Å, which contains 108 electrons. To the best of our knowledge, this is the largest electronic system computed using a high-quality neural network ansatz. The  $3 \times 3 \times 3$  supercell calculation predicts the total energy per unit cell of LiH is  $-8.160$  Hartree and the cohesive energy per unit cell is  $-4.770$  eV after the finite-size correction (see Supplementary Table 10), which is also very close to the experimental value  $-4.759$  eV<sup>30</sup>.

## Homogeneous electron gas

In addition to the solids containing nuclei, our computational framework can also apply straightforwardly to model systems such as homogeneous electron gas (HEG). HEG has been studied for a long time to understand the fundamental behavior of metals and electronic phase transitions<sup>32</sup>. Several seminal ab initio works have reported the energy of HEG at different densities<sup>21,22,32–35</sup>. Recently two other works have extended neural network ansatz to study HEG<sup>21,22</sup>. Although our computational framework is independently designed for solids, the network structure between this work and refs. 21, 22 employ similar ideas. Different physics-inspired envelope functions and periodic



**Fig. 3 | Electron density of  $H_{10}$  chains.** The horizontal axis is scaled by the corresponding bond length. Complex polarization modulus  $|Z|$  as a function of bond length is plotted in the inset.

features are used in these works, which suit the features of solids and homogeneous electron gas respectively. We make comparisons between these networks and ours on HEG, and observe consistent performance, which further proves the effectiveness of neural network-based QMC works. In this section, we present the results calculated on a simple cubic cell containing 54 electrons in a closed-shell configuration, the largest HEG system studied in this work (Fig. 2g). More results and comparisons with other works on smaller systems are discussed in the section “Network comparison” and Supplementary Table 13.

Figure 2h shows our calculated correlation error on the 54-electrons HEG at six different densities from  $r_s = 0.5$  Bohr to 20.0 Bohr. The state-of-the-art results, namely VMC with backflow correlation (BF)<sup>33</sup>, distinguishable cluster with double excitations (DCD)<sup>34</sup>, and transcorrelated full configuration interaction quantum Monte Carlo (TC-FCIQMC)<sup>35</sup> are also plotted for comparison, and BF-DMC result is often used as the reference energy of correlation error. Overall, our neural network performs very well, with an error of less than 1% in a wide range of density (see Supplementary Table 14). Generally, the correlation error increases as the density of HEG decreases when the correlation effects become larger, which also appears in DCD calculations.

### Electron density

Besides the total energy of solid systems, the electron density is also a key quantity to be calculated. For example, the electron density is crucial for characterizing different solids, such as the difference between insulators and conductors, and the distinct nature of ionic and covalent crystals. In DFT the one-to-one correspondence between many-body wavefunction and electron density proved by Hohenberg and Kohn in 1964 suggests that electron density is a fundamental quantity of materials. However, an interesting survey found that while new functionals in DFT improve the energy calculation the obtained density somehow can deviate from the exact<sup>36</sup>. Here, with our accurate neural network wavefunction, we can also obtain accurate electron density of solids and provide a valuable benchmark and guidance for method development.

A conductor features free-moving electrons, which would have macroscopic movements under external electric fields. In contrast, electrons are localized and constrained in insulators and cause considerable electron resistance. In Fig. 3, as an example, we show the calculated electron density of the hydrogen chain at two different bond lengths. As we can see, for the compressed hydrogen chain ( $L = 2$  Bohr), the electron density is rather uniform and has small fluctuations. As the chain is stretched, the electrons are more localized and the density profile has larger variations. The observation is consistent with the well-known insulator-conductor transition on the hydrogen chain by varying the bond length. To measure the transition more

quantitatively, we further calculate the complex polarization  $Z$  as the order parameter for insulator-conductor transition<sup>37</sup>. A conducting state is characterized by a vanishing complex polarization modulus  $|Z| \rightarrow 0$ , while an insulating state has a finite  $|Z| \rightarrow 1$ . We can see that the insulator-conductor transition bond length of the hydrogen chain is around 3 Bohr according to the calculated results, which is also consistent with the previous studies<sup>37</sup>.

Ionic and covalent bonds are the most fundamental chemical bonds in solids. While the physical pictures of these two types of bonding are very different, they both lie in the behavior of electrons as the “quantum glue” and electron density distribution is a simple way to visualize different bonding types. Here we choose to calculate the electron density of the diamond-structured Si, rock-salt NaCl and LiH crystals at their equilibrium position. They are representative of covalent and ionic crystals, and have also been investigated by other high-level wavefunction methods, e.g., AFQMC<sup>38</sup>. Note that in the calculations of NaCl and Si, correlation-consistent effective core potential (ccECP) is employed to reduce the cost, which removes the inertia of core electrons and keeps the behavior of active valence electrons<sup>15,39</sup>.

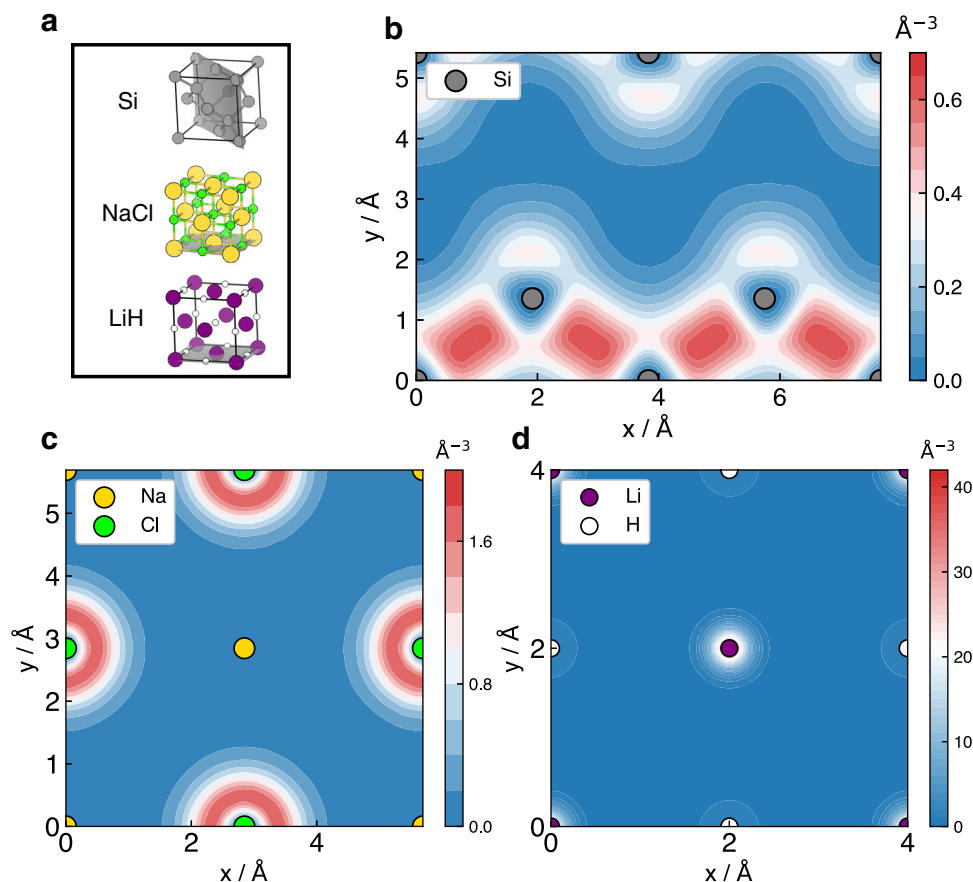
The electron density of diamond-structured Si in its (011) plane is plotted in Fig. 4b. We can see that valence electrons are shared by the nearest Si atoms, forming apparent Si-Si covalent bonds. In contrast, valence electrons are located around atoms in NaCl crystal as Fig. 4c shows. All the valence electrons are attracted around Cl atoms, forming effective  $Na^+$  and  $Cl^-$  ions in the crystal. Moreover, the electron density of LiH crystal is also calculated and plotted in Fig. 4d. LiH crystal is a moderate system between a typical ionic and covalent crystal. According to the result, the electrons are nearly equally distributed near Li and H atoms for our network. Detailed Bader charge analysis<sup>40</sup> manifests the atoms in the crystal become  $Li^{0.67+}$  and  $H^{0.67-}$  ions, respectively (resolution  $\sim 0.015$  Å), which slightly deviates from the stable closed-shell configuration (see Supplementary Note 7 for more details).

### Network comparison

In refs. 21, 22, neural networks are also used to simulate homogeneous electron gas system, employing a different choice of periodic feature function. In Fig. 5 we plot the correlation error computed on the 14-electrons HEG system, which can be compared with the results of other works. We can see that all three networks can go beyond the BF-DMC level for high-density systems. For all systems tested, our correlation errors are about 2% with the TC-FCIQMC result as the reference<sup>35</sup>, whereas the results of refs. 21, 22 are within 1%. It is understandable that the networks of refs. 21, 22 are specially designed for HEG systems, so slightly better accuracy can be achieved than our network. In their works, multiple phase factors  $e^{ikr}$  are used in the constructed orbitals, which improve the expressiveness of the network. In comparison, our network contains an additional exponential decay term, which simulates the attraction between atoms and electrons in solids containing nuclei (see Methods section for more details). Furthermore, the choice of periodic distance, as well as the domains of the constructed wavefunction (complex or real-valued), are also different in these three works, which may add differences to their performance. In the future, it would be interesting to combine the insights learned from these three works and design a better network ansatz for periodic systems.

### Metallic lithium

We have also carried out preliminary calculations on metallic lithium. The real metal system remains a notoriously difficult task for accurate wavefunction approaches<sup>7,41–44</sup>. The zero gap of metal leads to a discontinuity in the Brillouin zone integral. As a consequence, a significantly larger simulation cell is required for metals than insulators to



**Fig. 4 | Electron density of solids.** **a** Structures of solids, where the lattice planes for plotting electron densities are indicated. **b** Electron density of diamond-structured Si in its (011) plane, ccECP[Ne] is employed, and the bond length of Si

equals 5.42 Å. **c** Electron density of NaCl crystal in its  $xy$ -plane, ccECP[Ne] is employed, and the bond length of NaCl equals 5.7 Å. **d**, the electron density of LiH crystal in its  $xy$ -plane, and the bond length of LiH equals 4.0 Å.

reach the thermodynamic limit. Shortcut approaches to simulate metals are proposed via employing a special twist angle<sup>7,43</sup>, which helps to reduce the simulation size and finite-size error. Here we employ our network to simulate lithium with a body-centered cubic (bcc) structure, which is a typical metal with zero gap. A  $2 \times 2 \times 2$  conventional cell of bcc-Li at  $\Gamma$  point is employed (see Supplementary Table 11). In Supplementary Table 12, we list the total energy and the cohesive energy computed. As expected, the error in cohesive energy of lithium with such a limited supercell is larger than in non-metallic solids such as LiH, and further developments are desired to treat the large finite-size errors in metal.

## Discussion

The construction of a wavefunction for solid systems is a crucial but unsolved problem in the neural network community. The core mechanism of our neural network is the use of the periodic distance feature, which promotes molecule neural networks elegantly to the corresponding periodic ones and avoids time-consuming lattice summation. Considering the high-accuracy results obtained in this work, our neural network can be further applied to study more delicate physics and materials problems, such as the phase transitions of solids, surfaces, interfaces, and disordered systems, to name just a few. Our ansatz also offers a flexible extension to other neural networks and an easy integration into traditional computational techniques. The naturally evolved many-body wavefunction from the neural network may provide more physical and chemical insights into emergent phenomena of complex materials.

For further development of neural network-based QMC, the most crucial task is to enlarge its simulation size while retaining a

reasonable accuracy, which allows a more accurate simulation of metals and high-temperature superconductors. Employing pseudopotential is helpful to enlarge the simulation size<sup>15</sup>, while a better solution is a more efficient neural network, and related works are under progress.

## Methods

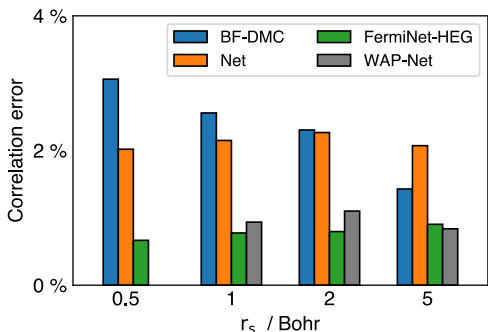
### Supercell approximation

Simulating a solid system requires solving the Schrödinger equation of many electrons within a large bulk. Supercell approximation is usually adopted to simplify the problem, considering a finite number of electrons and nuclei with periodic boundary conditions, whose Hamiltonian reads

$$\hat{H}_S = \sum_i -\frac{1}{2}\Delta_i + \frac{1}{2} \sum'_{\mathbf{L}_S, i, j} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j + \mathbf{L}_S|} - \sum_{\mathbf{L}_S, i, l} \frac{Z_l}{|\mathbf{r}_i - \mathbf{R}_l + \mathbf{L}_S|} + \frac{1}{2} \sum'_{\mathbf{L}_S, l, j} \frac{Z_l Z_j}{|\mathbf{R}_l - \mathbf{R}_j + \mathbf{L}_S|}, \quad (3)$$

where  $\mathbf{r}_i$  denotes the spatial position of  $i$ th electron in the supercell.  $\mathbf{R}_l$ ,  $Z_l$  are the spatial position and charge of the  $l$ th nucleus and  $\{\mathbf{L}_S\}$  is the set of supercell lattice vectors, which is usually a subset of primitive cell lattice vectors  $\{\mathbf{L}_p\}$ . In order to simulate the real environments of electrons in solids, the interactions between the particles and their images are also included in  $\hat{H}_S$ , and the prime symbol in summation means  $i=j$  terms are omitted for  $\mathbf{L}_S = 0$ .

Supercell Hamiltonian  $\hat{H}_S$  is invariant under the translation of any electron by a vector in  $\{\mathbf{L}_S\}$  as well as a simultaneous translation of



**Fig. 5 | Correlation error of 14-electrons HEG system at different  $r_s$ .** Correlation error is defined as  $[1 - (E - E_{\text{HF}})/(E_{\text{ref}} - E_{\text{HF}})] \times 100\%$ . WAP-Net refers to ref. 21 and FermiNet-HEG refers to ref. 22. BF-DMC results<sup>33,34</sup> are displayed for comparison, and TC-FCIQMC data were used as reference<sup>35</sup>.

all-electrons by a vector in  $\{\mathbf{L}_p\}$ . As a consequence, two periodic conditions are required for the ground-state wavefunction  $\psi^{45}$ ,

$$\begin{aligned}\psi(\mathbf{r}_1 + \mathbf{L}_p, \dots, \mathbf{r}_N + \mathbf{L}_p) &= \exp(i\mathbf{k}_p \cdot \mathbf{L}_p)\psi(\mathbf{r}_1, \dots, \mathbf{r}_N), \\ \psi(\mathbf{r}_1 + \mathbf{L}_s, \dots, \mathbf{r}_N) &= \exp(i\mathbf{k}_s \cdot \mathbf{L}_s)\psi(\mathbf{r}_1, \dots, \mathbf{r}_N),\end{aligned}\quad (4)$$

where  $\mathbf{k}_s, \mathbf{k}_p$  denote the momentum vectors reduced in the first Brillouin zone of the supercell and the primitive cell, respectively. Eq. (4) and the anti-symmetry condition together form the fundamental requirements for  $\psi$ . As the size of the supercell increases, simulation results gradually converge to the thermodynamic limit of a real solid system.

### Wavefunction ansatz

In conventional QMC simulation of solids, Hartree–Fock type wavefunction ansatz composed of Bloch functions is often used, which reads

$$\psi_{\mathbf{k}_s, \mathbf{k}_p}^{\text{HF}}(\mathbf{r}) = \text{Det} \begin{pmatrix} e^{i\mathbf{k}_1 \cdot \mathbf{r}_1} u_{\mathbf{k}_1}(\mathbf{r}_1) & \dots & e^{i\mathbf{k}_N \cdot \mathbf{r}_1} u_{\mathbf{k}_N}(\mathbf{r}_1) \\ \vdots & & \vdots \\ e^{i\mathbf{k}_1 \cdot \mathbf{r}_N} u_{\mathbf{k}_1}(\mathbf{r}_N) & \dots & e^{i\mathbf{k}_N \cdot \mathbf{r}_N} u_{\mathbf{k}_N}(\mathbf{r}_N) \end{pmatrix}. \quad (5)$$

In order to satisfy Eq. (4),  $\mathbf{k}_i$  in the determinant should lie on the grid of supercell reciprocal lattice vectors  $\{\mathbf{G}_S\}$  offset by  $\mathbf{k}_s$  within the first Brillouin zone of the primitive cell. Moreover,  $u_{\mathbf{k}}$  functions in Eq. (5) should satisfy the translation invariant condition by the primitive cell lattice vectors,

$$u_{\mathbf{k}}(\mathbf{r} + \mathbf{L}_p) = u_{\mathbf{k}}(\mathbf{r}). \quad (6)$$

Following the strategy of FermiNet<sup>10</sup>, Bloch functions in Eq. (5) can be promoted with collective distances,

$$e^{i\mathbf{k} \cdot \mathbf{r}_i} u_{\mathbf{k}}(\mathbf{r}_i) \rightarrow e^{i\mathbf{k} \cdot \mathbf{r}_i} u_{\mathbf{k}}(\mathbf{r}_i; \mathbf{r}_{\neq i}), \quad (7)$$

where  $\mathbf{r}_{\neq i}$  denotes all the electron coordinates except  $\mathbf{r}_i$ . These collective orbitals are constructed to achieve the equivalence of electron permutations  $P$ ,

$$P_{ij} u_{\mathbf{k}_i}(\mathbf{r}_j; \mathbf{r}_{\neq j}) = u_{\mathbf{k}_j}(\mathbf{r}_i; \mathbf{r}_{\neq i}), \quad (8)$$

which combined with the Slater determinant ensures the anti-symmetry nature of electrons. Moreover, we use the periodic distance features  $d(\mathbf{r})$  in Eq. (2) to substitute ordinary  $|\mathbf{r}|$  in the molecular neural

network. The periodic functions  $f, g$  used in Eq. (2) read

$$\begin{aligned}f(\omega) &= |\omega| \left( 1 - \frac{|\omega/\pi|^3}{4} \right), \\ g(\omega) &= \omega \left( 1 - \frac{3}{2} |\omega/\pi| + \frac{1}{2} |\omega/\pi|^2 \right),\end{aligned}\quad (9)$$

and their arguments  $\omega$  are reduced into  $[-\pi, \pi]$ . Eq. (6) can then be satisfied without causing discontinuity<sup>20</sup>. The constructed periodic features  $\{\sum_i g(\omega_i) \mathbf{a}_i, d(\mathbf{r})\}$  are substituted into FermiNet<sup>10</sup> to build a periodic wavefunction. Specifically, electron-atom features  $\mathbf{h}_e$  and electron–electron features  $\mathbf{h}_{e,e'}$  are constructed as follows,

$$\begin{aligned}\mathbf{h}_e &= \left\{ \sum_{i=1}^3 g(\omega_{e,i}^j) \mathbf{a}_i^p, d(\omega_{e,i}^j) \right\}, \\ \mathbf{h}_{e,e'} &= \left\{ \sum_{i=1}^3 g(\omega_{e,e'}^i) \mathbf{a}_i^s, d(\omega_{e,e'}^i) \right\},\end{aligned}\quad (10)$$

where  $\omega_{e,i}, \omega_{e,e'}$  are defined as

$$\begin{aligned}\omega_{e,i} &= (\mathbf{r}_e - \mathbf{R}_i) \cdot \{ \mathbf{b}_1^p, \mathbf{b}_2^p, \mathbf{b}_3^p \}, \\ \omega_{e,e'} &= (\mathbf{r}_e - \mathbf{r}_{e'}) \cdot \{ \mathbf{b}_1^s, \mathbf{b}_2^s, \mathbf{b}_3^s \},\end{aligned}\quad (11)$$

and superscripts  $p, s$  denote the primitive cell and supercell respectively. A permutation equivalent feature  $\mathbf{f}_e^\alpha$  are further constructed from  $\mathbf{h}_e, \mathbf{h}_{e,e'}$ ,

$$\mathbf{f}_e^\alpha = \text{concat}(\mathbf{h}_e, \mathbf{g}^\uparrow, \mathbf{g}^\downarrow, \mathbf{g}_e^{\alpha,\uparrow}, \mathbf{g}_e^{\alpha,\downarrow}), \quad (12)$$

where  $\alpha$  denotes the spin index ( $\uparrow, \downarrow$ ).  $\mathbf{g}^\uparrow, \mathbf{g}^\downarrow$  and  $\mathbf{g}_e^{\alpha,\uparrow}, \mathbf{g}_e^{\alpha,\downarrow}$  read

$$\begin{aligned}(\mathbf{g}^\uparrow, \mathbf{g}^\downarrow) &= \left( \frac{1}{n^\uparrow} \sum_e \mathbf{h}_e^\uparrow, \frac{1}{n^\downarrow} \sum_e \mathbf{h}_e^\downarrow \right), \\ (\mathbf{g}_e^{\alpha,\uparrow}, \mathbf{g}_e^{\alpha,\downarrow}) &= \left( \frac{1}{n^\uparrow} \sum_{e'} \mathbf{h}_{e,e'}^{\alpha,\uparrow}, \frac{1}{n^\downarrow} \sum_{e'} \mathbf{h}_{e,e'}^{\alpha,\downarrow} \right).\end{aligned}\quad (13)$$

$\mathbf{f}_e^\alpha$  and  $\mathbf{h}_{e,e'}$  are subsequently substituted into a series of fully connected layers recursively

$$\begin{aligned}\mathbf{h}_e^{l+1,\alpha} &= \tanh(\mathbf{V}^l \cdot \mathbf{f}_e^{l,\alpha} + \mathbf{b}^l) + \mathbf{h}_e^{l,\alpha}, \\ \mathbf{h}_{e,e'}^{l+1,\alpha,\beta} &= \tanh(\mathbf{W}^l \cdot \mathbf{h}_{e,e'}^{l,\alpha,\beta} + \mathbf{c}^l) + \mathbf{h}_{e,e'}^{l,\alpha,\beta},\end{aligned}\quad (14)$$

where  $l$  denotes the number of layers, and  $\{\mathbf{V}_l, \mathbf{b}_l\}, \{\mathbf{W}_l, \mathbf{c}_l\}$  denote corresponding weight and bias of  $l$ -layer.

Functions  $u$  in Eq. (7) are built using the  $\mathbf{h}_e^l$  from the last  $L$ -layer,

$$u = \text{Orb}^{\text{Re}} \cdot \mathbf{h}_e^L + \mathbf{i} \times \text{Orb}^{\text{Im}} \cdot \mathbf{h}_e^L, \quad (15)$$

where  $\text{Orb}^{\text{Re,Im}}$  denote the weight parameters of the real part and the imaginary part respectively.

Moreover,  $u$  function is multiplied by an additional phase factor  $\exp(i\mathbf{k} \cdot \mathbf{r})$ , which mimics Bloch functions and encodes the occupied  $\mathbf{k}$ -point information from HF calculation. Inspired by the tight-binding model in solid physics, a periodic-generalized envelope term  $\exp[-d(\mathbf{r})]$  is also added to the molecule orbitals, which considers an attractive interaction effect between atoms and electrons. The final molecule orbitals  $\phi$  reads

$$\phi(\mathbf{r}_i; \mathbf{r}_{\neq i}) = \exp(i\mathbf{k} \cdot \mathbf{r}_i) \exp[-d(\mathbf{r}_i)] u(\mathbf{r}_i; \mathbf{r}_{\neq i}). \quad (16)$$

For an overall sketch of the neural network, see section “Pseudocode of network”. Note that the distance between electrons and nuclei is omitted for the HEG system since it does not contain any nucleus. Specific hyperparameters of each system are listed in Supplementary Note 1.

**Pseudocode of network**

For clarity, the pseudocode of network reads below:

```

Require: electron positions  $\{\mathbf{r}_1^\uparrow, \dots, \mathbf{r}_n^\uparrow, \mathbf{r}_1^\downarrow, \dots, \mathbf{r}_n^\downarrow\}$ 
Require: nuclear positions  $\{\mathbf{R}_j\}$  in the primitive cell
Require: lattice vector  $\{\mathbf{a}_1^{p,S}, \mathbf{a}_2^{p,S}, \mathbf{a}_3^{p,S}\}$  of primitive cell and supercell
Require: reciprocal lattice vector  $\{\mathbf{b}_1^{p,S}, \mathbf{b}_2^{p,S}, \mathbf{b}_3^{p,S}\}$  of primitive cell and supercell
Require: occupied  $\{\mathbf{k}_j\}$  points offered by Hartree–Fock method
For each electron  $e$ , atom  $l$ :
 $\omega_{e,l} = (\mathbf{r}_e - \mathbf{R}_l) \cdot \{\mathbf{b}_1^p, \mathbf{b}_2^p, \mathbf{b}_3^p\}$ 
 $\omega_{e,e'} = (\mathbf{r}_e - \mathbf{r}_{e'}) \cdot \{\mathbf{b}_1^p, \mathbf{b}_2^p, \mathbf{b}_3^p\}$ 
End For
For each electron  $e$ :
 $\mathbf{h}_e = \{\sum_{i=1}^3 g(\omega_{e,l}^i) \mathbf{a}_i^p, d(\omega_{e,l})\}$ 
 $\mathbf{h}_{e,e'} = \{\sum_{i=1}^3 g(\omega_{e,e'}^i) \mathbf{a}_i^p, d(\omega_{e,e'})\}$ 
End For
For each layer  $l$ :
 $\mathbf{g}^{l,\uparrow} = \frac{1}{n^l} \sum_e \mathbf{h}_e^{l,\uparrow}$ 
 $\mathbf{g}^{l,\downarrow} = \frac{1}{n^l} \sum_e \mathbf{h}_e^{l,\downarrow}$ 
For each electron  $e$ , spin  $\alpha$ :
 $\mathbf{g}_e^{l,\alpha,\uparrow} = \frac{1}{n^l} \sum_{e'} \mathbf{h}_{e,e'}^{l,\alpha,\uparrow}$ 
 $\mathbf{g}_e^{l,\alpha,\downarrow} = \frac{1}{n^l} \sum_{e'} \mathbf{h}_{e,e'}^{l,\alpha,\downarrow}$ 
 $\mathbf{f}_e^{l,\alpha} = \text{concat}(\mathbf{h}_e^{l,\alpha}, \mathbf{g}^{l,\uparrow}, \mathbf{g}^{l,\downarrow}, \mathbf{g}_e^{l,\alpha,\uparrow}, \mathbf{g}_e^{l,\alpha,\downarrow})$ 
 $\mathbf{h}_{e,e'}^{l+1,\alpha} = \tanh(\mathbf{V}^l \cdot \mathbf{f}_e^{l,\alpha} + \mathbf{b}^l) + \mathbf{h}_{e,e'}^{l,\alpha}$ 
 $\mathbf{h}_{e,e'}^{l+1,\alpha,\beta} = \tanh(\mathbf{W}^l \cdot \mathbf{h}_{e,e'}^{l,\alpha,\beta} + \mathbf{c}^l) + \mathbf{h}_{e,e'}^{l,\alpha,\beta}$ 
End For
End For
For each orbital  $i$ :
For each electron  $e$ , spin  $\alpha$ :
 $u_{i,e}^\alpha = \text{Orb}_{i,\alpha}^{\text{Re}} \cdot \mathbf{h}_e^L + \mathbf{i} \times \text{Orb}_{i,\alpha}^{\text{Im}} \cdot \mathbf{h}_e^L$ 
 $p_{i,e}^\alpha = \exp(\mathbf{i} \mathbf{k}_i \cdot \mathbf{r}_e^\alpha)$ 
 $\text{enve}_{i,e}^\alpha = \sum_l \pi_l^{l,\alpha} \exp[-\sigma_l^{l,\alpha} d(\omega_{e,l})]$ 
 $\phi_{i,e}^\alpha = p_{i,e}^\alpha u_{i,e}^\alpha \text{enve}_{i,e}^\alpha$ 
End For
End For
 $\psi = \text{Det}[\phi^\uparrow] \text{Det}[\phi^\downarrow]$ 

```

**Neural network optimization**

Parameters  $\theta$  within the neural network can be optimized to minimize the energy expectation value  $\langle E_l \rangle$ , and the gradient  $\nabla_\theta \langle E_l \rangle$  reads

$$\nabla_\theta \langle E_l \rangle = \text{Re}[\langle E_l \nabla_\theta \ln \psi^* \rangle - \langle E_l \rangle \langle \nabla_\theta \ln \psi^* \rangle], \tag{17}$$

$$E_l = \psi^{-1} \hat{H}_S \psi,$$

where  $E_l$  denotes the local energy of neural network ansatz  $\psi$ . Besides energy minimization, stochastic reconfiguration optimization<sup>46</sup> has also been widely adopted and proved to be much more efficient, whose gradient reads

$$\text{Grad} = F^{-1} \nabla_\theta \langle E_l \rangle, \tag{18}$$

$$F_{ij} = \text{Re} \left[ \left\langle \frac{\partial \ln \psi^*}{\partial \theta_i} \frac{\partial \ln \psi}{\partial \theta_j} \right\rangle - \left\langle \frac{\partial \ln \psi^*}{\partial \theta_i} \right\rangle \left\langle \frac{\partial \ln \psi}{\partial \theta_j} \right\rangle \right].$$

In this work, we adopt a modified KFAC optimizer, which approximates  $F$  as

$$F = \text{Re} \left[ \left\langle \frac{\partial \ln \psi^*}{\partial \text{vec}(W_l)} \frac{\partial \ln \psi^T}{\partial \text{vec}(W_l)} \right\rangle - \left\langle \frac{\partial \ln \psi^*}{\partial \text{vec}(W_l)} \right\rangle \left\langle \frac{\partial \ln \psi^T}{\partial \text{vec}(W_l)} \right\rangle \right] \tag{19}$$

$$= \text{Re} \left[ \langle (a_l \otimes e_i^*)(a_l \otimes e_j)^T \rangle - \langle (a_l \otimes e_i^*) \rangle \langle (a_l \otimes e_j)^T \rangle \right]$$

$$\approx \text{Re} [(a_l a_l^T) \otimes (e_i^* e_j^T)],$$

where  $W_l$  denotes the weight parameters of layer  $l$ , and  $\text{vec}$  means vectorized form.  $a_l, e_l$  denote the activation and sensitivity of layer  $l$  respectively. Note that activation  $a_l$  is always real-valued, which explains the absence of conjugation of  $a_l$  in the second line. The first term in the bracket of Eq. (19) is approximated as the Kronecker product of the expectation values, and the second term is omitted for simplification.

**Twist average boundary condition**

TABC is a conventional technique to reduce the finite-size error due to the constrained size of the supercell<sup>27</sup>. It averages the contributions from different periodic images of the supercell and improves the convergence of the total energy. The formula reads

$$E_{\text{TABC}} = \frac{\Omega_S}{(2\pi)^3} \int_{\text{1.B.Z.}} d^3 \mathbf{k}_S \frac{\psi_{\mathbf{k}_S}^* \hat{H}_S \psi_{\mathbf{k}_S}}{\psi_{\mathbf{k}_S}^* \psi_{\mathbf{k}_S}}, \tag{20}$$

where 1.B.Z. denotes the first Brillouin zone of supercell and the integral is practically approximated by a discrete sum of a Monkhorst-Pack mesh (see Supplementary Note 3.2 for more details).

**Structure factor correction**

Finite-size error can be further reduced via the structure factor  $S(\mathbf{k})$  correction<sup>28</sup>, which is usually calculated to correct the exchange-correlation potential  $V_{xc}$  and the formula reads

$$\frac{\Delta V_{xc}}{N_e} = \frac{2\pi}{\Omega_S} \lim_{\mathbf{k} \rightarrow 0} \frac{S(\mathbf{k})}{\mathbf{k}^2}, \tag{21}$$

$$S(\mathbf{k}) = \frac{1}{N_e} \left[ \langle \rho(\mathbf{k}) \rho^*(\mathbf{k}) \rangle - \langle \rho(\mathbf{k}) \rangle \langle \rho^*(\mathbf{k}) \rangle \right],$$

where  $\lim_{\mathbf{k} \rightarrow 0}$  is practically estimated via interpolation (see Supplementary Note 3.4 for more details).

**Empirical correction formula**

Empirical formulas are also commonly employed to reduce the finite-size error<sup>18</sup>, one of which reads

$$E_\infty = E_N^{\text{Net}} + (E_\infty^{\text{HF}} - E_N^{\text{HF}}). \tag{22}$$

The simulation size of high-accuracy methods is usually limited due to high computational costs. Hence methods with a much more practical time scale, such as HF, is usually used to give a posterior estimation of the finite-size error. All the results of LiH are corrected using this empirical formula with HF results in a cc-pVDZ basis (see Supplementary Note 4.3 for more details).

**Electron density analysis**

Electron density  $\rho(\mathbf{r})$  is defined as

$$\rho(\mathbf{r}) = N \int d^3 \mathbf{r}_2 \dots d^3 \mathbf{r}_N |\psi(\mathbf{r}, \mathbf{r}_2, \dots, \mathbf{r}_N)|^2, \tag{23}$$

and it's practically evaluated by accumulating Monte Carlo samples of electrons on a uniform grid over the simulation cell. As for the complex polarization  $Z$ , it is defined as<sup>37</sup>

$$Z = \left\langle \exp \left( i \sum_i \frac{2\pi}{L} \mathbf{r}_i^\parallel \right) \right\rangle, \tag{24}$$

where  $\mathbf{r}^\parallel$  denotes the projection of electron coordinate along the chain direction. Moreover, Bader charge is employed to estimate the charge partition on each atom<sup>40</sup>. The convergence test of Bader charge is shown in the Supplementary Fig. 8.

### Workflow and computational details

This work is developed upon open-source [FermiNet](#)<sup>47</sup> and [PyQMC](#)<sup>48</sup> on Github, deep learning framework [JAX](#)<sup>49</sup> is used which supports flexible and powerful complex number calculation. Ground-state energy calculations are performed with all-electrons. Diamond-structured Si and NaCl crystal are simulated with ccECP[Ne]<sup>39</sup>. The neural network is pretrained by Hartree–Fock ansatz, obtained with PySCF software<sup>50</sup>. All the used  $\mathbf{k}$  points are the occupied  $\mathbf{k}$  points from Hartree–Fock calculation using Monkhorst-Pack mesh offset by  $\mathbf{k}_s$  in cc-pVDZ basis, and the mesh size is the same as the supercell. All the expectation values for distribution  $|\psi|^2$  are evaluated via the Monte Carlo approach, and then the energy and wavefunction is optimized using the modified KFAC optimizer<sup>24</sup> (see Supplementary Figs. 1, 2, 4, 6, 7). The Ewald summation technique is implemented for the lattice summation in energy calculation. After training is converged, energy is calculated in a separate inference phase.

### Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

### Data availability

The data generated in this study are provided in the Supplementary Information.

### Code availability

The concrete code of this work is developed on Github (<https://github.com/bytedance/DeepSolid>).

### References

- Kohn, W. Nobel lecture: electronic structure of matter—wave functions and density functionals. *Rev. Mod. Phys.* **71**, 1253–1266 (1999).
- Martin, R. M. *Electronic Structure: Basic Theory and Practical Methods* (Cambridge Univ. Press, 2004).
- Jones, R. O. Density functional theory: Its origins, rise to prominence, and future. *Rev. Mod. Phys.* **87**, 897–923 (2015).
- Kirkpatrick, J. et al. Pushing the frontiers of density functionals by solving the fractional electron problem. *Science* **374**, 1385–1389 (2021).
- Williams, K. T. et al. Direct comparison of many-body methods for realistic electronic Hamiltonians. *Phys. Rev. X* **10**, 011041 (2020).
- Booth, G. H., Grüneis, A., Kresse, G. & Alavi, A. Towards an exact description of electronic wavefunctions in real solids. *Nature* **493**, 365–370 (2013).
- Mihm, T. N. et al. A shortcut to the thermodynamic limit for quantum many-body calculations of metals. *Nat. Comput. Sci.* **1**, 801–808 (2021).
- Han, J., Zhang, L. & Weinan, E. Solving many-electron Schrödinger equation using deep neural networks. *J. Comput. Phys.* **399**, 108929 (2019).
- Carleo, G. & Troyer, M. Solving the quantum many-body problem with artificial neural networks. *Science* **355**, 602–606 (2017).
- Pfau, D., Spencer, J. S., Alexander, G., Matthews, D. G. & Foulkes, W. M. C. Ab initio solution of the many-electron Schrödinger equation with deep neural networks. *Phys. Rev. Res.* **2**, 033429 (2020).
- Spencer, J. S., Pfau, D., Botev, A. & Foulkes, W. M. C. Better, faster fermionic neural networks. Preprint at [arXiv:2011.07125](https://arxiv.org/abs/2011.07125) (2020).
- Hermann, J., Schätzle, Z. & Noé, F. Deep-neural-network solution of the electronic Schrödinger equation. *Nat. Chem.* **12**, 891–897 (2020).
- Choo, K., Mezzacapo, A. & Carleo, G. Fermionic neural-network states for ab-initio electronic structure. *Nat. Commun.* **11**, 2368 (2020).
- Yoshioka, N., Mizukami, W. & Nori, F. Solving quasiparticle band spectra of real solids using neural-network quantum states. *Commun. Phys.* **4**, 1–8 (2021).
- Li, X., Fan, C., Ren, W. & Chen, J. Fermionic neural network with effective core potential. *Phys. Rev. Res.* **4**, 013021 (2022).
- Ren, W., Fu, W. & Chen, J. Towards the ground state of molecules via diffusion monte carlo on neural networks. Preprint at [arXiv:2204.13903](https://arxiv.org/abs/2204.13903) (2022).
- Guther, K. et al. NECI: N-electron configuration interaction with an emphasis on state-of-the-art stochastic methods. *J. Chem. Phys.* **153**, 034107 (2020).
- Foulkes, W. M. C., Mitas, L., Needs, R. J. & Rajagopal, G. Quantum monte carlo simulations of solids. *Rev. Mod. Phys.* **73**, 33 (2001).
- Shi, H. & Zhang, S. Some recent developments in auxiliary-field quantum monte carlo for real materials. *J. Chem. Phys.* **154**, 024107 (2021).
- Whitehead, T. M., Michael, M. H. & Conduit, G. J. Jastrow correlation factor for periodic systems. *Phys. Rev. B* **94**, 035157 (2016).
- Wilson, M. et al. Wave function ansatz (but periodic) networks and the homogeneous electron gas. Preprint at [arXiv:2202.04622](https://arxiv.org/abs/2202.04622) (2022).
- Cassella, G. et al. Discovering quantum phase transitions with fermionic neural networks. Preprint at [arXiv:2202.05183](https://arxiv.org/abs/2202.05183) (2022).
- Martens, J. & Grosse, R. Optimizing neural networks with kronecker-factored approximate curvature. In *Proc. 32nd International Conference on International Conference on Machine Learning* 2408–2417. (JMLR.org, 2015).
- Botev, A. & Martens, J. KFAC-JAX. (2022).
- Motta, M. et al. Towards the solution of the many-electron problem in real materials: equation of state of the hydrogen chain with state-of-the-art many-body methods. *Phys. Rev. X* **7**, 031059 (2017).
- Geim, A. K. Nobel lecture: random walk to graphene. *Rev. Mod. Phys.* **83**, 851–862 (2011).
- Lin, C., Zong, F. H. & Ceperley, D. M. Twist-averaged boundary conditions in continuum quantum monte carlo algorithms. *Phys. Rev. E* **64**, 016702 (2001).
- Chiesa, S., Ceperley, D. M., Martin, R. M. & Holzmann, M. Finite-size error in many-body simulations with long-range interactions. *Phys. Rev. Lett.* **97**, 076404 (2006).
- Dappe, Y. et al. Local-orbital occupancy formulation of density functional theory: application to si, c, and graphene. *Phys. Rev. B* **73**, 235124 (2006).
- Nolan, S. J., Gillan, M. J., Alfè, D., Allan, N. L. & Manby, F. R. Calculation of properties of crystalline lithium hydride using correlated wave function theory. *Phys. Rev. B* **80**, 165109 (2009).
- Binnie, S. J. et al. Bulk and surface energetics of crystalline lithium hydride: benchmarks from quantum monte carlo and quantum chemistry. *Phys. Rev. B* **82**, 165431 (2010).
- Ceperley, D. M. & Alder, B. J. Ground state of the electron gas by a stochastic method. *Phys. Rev. Lett.* **45**, 566–569 (1980).
- López Ríos, P., Ma, A., Drummond, N. D., Towler, M. D. & Needs, R. J. Inhomogeneous backflow transformations in quantum monte carlo calculations. *Phys. Rev. E* **74**, 066701 (2006).
- Liao, K., Schraivogel, T., Luo, H., Kats, D. & Alavi, A. Towards efficient and accurate ab initio solutions to periodic systems via transcorrelation and coupled cluster theory. *Phys. Rev. Res.* **3**, 033072 (2021).
- Luo, H. & Alavi, A. Combining the transcorrelated method with full configuration interaction quantum monte carlo: Application to the homogeneous electron gas. *J. Chem. Theory Comput.* **14**, 1403–1411 (2018).
- Medvedev, M. G., Bushmarinov, I. S., Sun, J., Perdew, J. P. & Lysenko, K. A. Density functional theory is straying from the path toward the exact functional. *Science* **355**, 49–52 (2017).
- Stella, L., Attaccalite, C., Sorella, S. & Rubio, A. Strong electronic correlation in the hydrogen chain: a variational monte carlo study. *Phys. Rev. B* **84**, 245117 (2011).
- Chen, S., Motta, M., Ma, F. & Zhang, S. Ab initio electronic density in solids by many-body plane-wave auxiliary-field quantum monte carlo calculations. *Phys. Rev. B* **103**, 075138 (2021).



39. Annaberdiyev, A., Melton, C. A., Bennett, M. C., Wang, G. & Mitas, L. Accurate atomic correlation and total energies for correlation consistent effective core potentials. *J. Chem. Theory Comput.* **16**, 1482–1502 (2020).
40. Tang, W., Sanville, E. & Henkelman, G. A grid-based bader analysis algorithm without lattice bias. *J. Phys. Condens. Matter* **21**, 084204 (2009).
41. Yao, G., Xu, J. G. & Wang, X. W. Pseudopotential variational quantum monte carlo approach to bcc lithium. *Phys. Rev. B* **54**, 8393–8397 (1996).
42. Sugiyama, G., Zerah, G. & Alder, B. J. Ground-state properties of metallic lithium. *Phys. A Stat. Mech. Appl.* **156**, 144–168 (1989).
43. Dagrada, M., Karakuzu, S., Vildosola, Verónica Laura, Casula, M. & Sorella, S. Exact special twist method for quantum monte carlo simulations. *Phys. Rev. B* **94**, 245108 (2016).
44. Azadi, S. & Foulkes, W. M. C. Efficient method for grand-canonical twist averaging in quantum monte carlo calculations. *Phys. Rev. B* **100**, 245142 (2019).
45. Rajagopal, G., Needs, R. J., James, A., Kenny, S. D. & Foulkes, W. M. C. Variational and diffusion quantum monte carlo calculations at nonzero wave vectors: theory and application to diamond-structure germanium. *Phys. Rev. B* **51**, 10591–10600 (1995).
46. Sorella, S. Green function monte carlo with stochastic reconfiguration. *Phys. Rev. Lett.* **80**, 4558–4561 (1998).
47. Pfau, D., Spencer, J. S. & Contributors, FermiNet. FermiNet. (2020).
48. Wheeler, W. A. et al. Pyqmc: an all-python real-space quantum monte carlo module in pyscf. Preprint at arXiv:2212.01482 (2022).
49. Bradbury, J. et al. JAX: composable transformations of Python +NumPy programs. (2018).
50. Sun, Q. et al. Pyscf: the python-based simulations of chemistry framework. *WIREs Comput. Mol. Sci.* **8**, e1340 (2018).

## Acknowledgements

The authors thank Matthew Foulkes, David Ceperley, Lucas Wagner, Gareth Conduit, Mario Motta, and Ke Liao for helpful discussions. We thank Gino Cassella for providing Hartree–Fock energies of HEG. We thank the ByteDance AML team specially for their technical and computing support. We also thank ByteDance AI-Lab LIT Group and the rest of the ByteDance AI-Lab research team for inspiration and encouragement. This work is directed and supported by Hang Li and ByteDance AI-Lab. J.C. is supported by the National Natural Science Foundation of China under Grant No. 92165101.

## Author contributions

X.L. and J.C. conceived the study; X.L. developed the method, performed implementations, simulations, and data analyses; Z.L. contributed to the code development and simulation of HEG; J.C. supervised the project. X.L., Z.L., and J.C. wrote the paper.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41467-022-35627-1>.

**Correspondence** and requests for materials should be addressed to Xiang Li.

**Peer review information** *Nature Communications* thanks the anonymous reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022