Supplementary information: *Ab initio* calculation of real solids via neural network ansatz

Xiang Li[1], Zhe Li[1], and Ji Chen[2]

[1]*ByteDance Inc, Zhonghang Plaza, No. 43, North 3rd Ring West Road, Haidian District, Beijing.*
[2]*School of Physics, Interdisciplinary Institute of Light-Element Quantum Materials, Frontiers Science Center for Nano-Optoelectronics, Peking University, Beijing 100871, People's Republic of China*

May 24, 2022

Supplementary Note 1. Hyperparameters for simulations

The recommended hyperparameters are listed in Supplementary Table 1. Some employed hyperparameters of the presented results differ from the recommended ones, which are specially given in Supplementary Table 2.

Hyperparameter	Value	Hyperparameter	Value
Pretrain basis	ccpvdz	Pretrain iterations	1e3
Dimension of one electron layer **V**	256	Dimension of two electron layer **W**	32
Number of layers	4	Number of determinants	8
Optimizer	KFAC	Learning rate	3e-2
Damping	1e-3	Constrained norm of gradient	1e-3
Momentum of optimizer	0.0	Batch size	4096
Number of training steps	2e5	Clipping window of gradient	5
MCMC burn in	1e3	MCMC steps between each iterations	20
MCMC move width	2e-2	Target MCMC acceptance	55%
Precision	Float64	Number of inference steps	5e4

Supplementary Table 1 | Recommended hyperparameters

System	Layer dimension	Layer	Determinants	Batch size	Training steps
Hydrogen chain	(256, 32)	3	8	4096	1e5
Graphene	(256, 32)	4	8	4096	3e5
$2\times2\times2$ Lithium hydride	(256, 32)	4	8	4096	3e5
$3\times3\times3$ Lithium hydride	(256, 32)	4	1	8192	4e5
Homogeneous electron gas	(256, 32)	3	1	4096	3e5

Supplementary Table 2 | Some system dependent hyperparameters

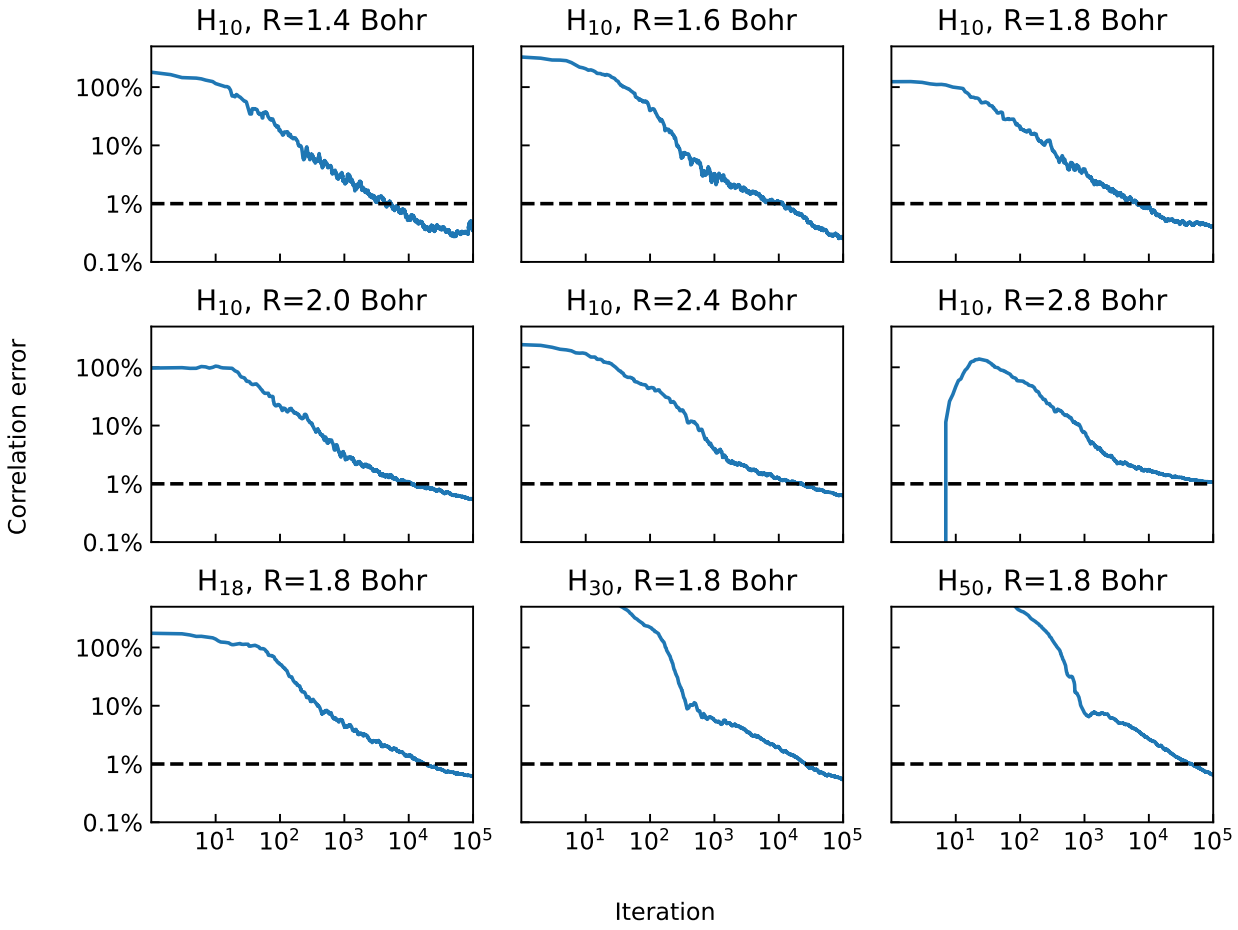
Supplementary Note 2. Hydrogen chain

Supplementary Note 2.1 Training curve

The training curve of H_{10} in PBC is plotted in Supplementary Fig. 1. The correlation error is defined as

$$\text{Correlation error} = \left(1 - \frac{E_{\text{Net}} - E_{\text{HF}}}{E_{\text{DMC}} - E_{\text{HF}}}\right) \times 100\% \,, \tag{1}$$

where E_{HF} is calculated using the ccpvdz basis set and E_{DMC} is taken from Ref. [1].



Supplementary Figure 1 | Hydrogen chain training curve. For clarity, at each iteration number, we plot the median correlation error of the last 10% of the corresponding iteration.

Bond length (Å)	Net	LR-DMC(LDA)	VMC(LDA)
1.4	-0.551677(1)	-0.55178(1)	-0.55049(1)
1.6	-0.568740(1)	-0.56881(1)	-0.56752(1)
1.8	-0.572922(1)	-0.57304(1)	-0.57172(1)
2.0	-0.570401(1)	-0.57055(1)	-0.56911(1)
2.4	-0.556861(1)	-0.55703(1)	-0.55522(1)
2.8	-0.540783(1)	-0.54102(1)	-0.53831(1)

Supplementary Table 3 | Energy of H_{10} chain.

Supplementary Note 2.2 H_{10} dissociation curve

Energy of H_{10} chain per atom is given in Supplementary Table 3, LR-DMC and VMC results from Ref. [1] are also listed for comparison.

Supplementary Note 2.3 Finite-size error extrapolation

Energies of different hydrogen chains are given in Supplementary Table 4.

Size	Net	LR-DMC(LDA)	VMC(LDA)
10	-0.572922(1)	-0.57304(1)	-0.57172(1)
18	-0.567776(1)	-0.56796(1)	-0.56644(1)
30	-0.566114(1)	-0.56627(1)	-0.56478(1)
50	-0.565419(1)	-0.56560(1)	-0.56409(1)

Supplementary Table 4 | Energies of different hydrogen chains, energies are given in Hartree and the bond length of hydrogen chain is fixed at 1.8 Bohr.

Supplementary Note 3. Graphene

Atom	Position (Å)	Lattice vector	Position (Å)
C1	(1.421, 0.0, 0.0)	$\mathbf{a}_1$	(2.1315, -1.2306, 0.0)
C2	(2.842, 0.0, 0.0)	$\mathbf{a}_2$	(2.1315, 1.2306, 0.0)
		$\mathbf{a}_3$	(0, 0, 52.9177)

Supplementary Table 5 | Geometry of Graphene

Supplementary Note 3.1 Geometry

The primitive cell lattice vectors as well as carbon atom coordinates are given in Supplementary Table 5. The size of supercell is 2×2.

Supplementary Note 3.2 Twist average boundary condition (TABC)

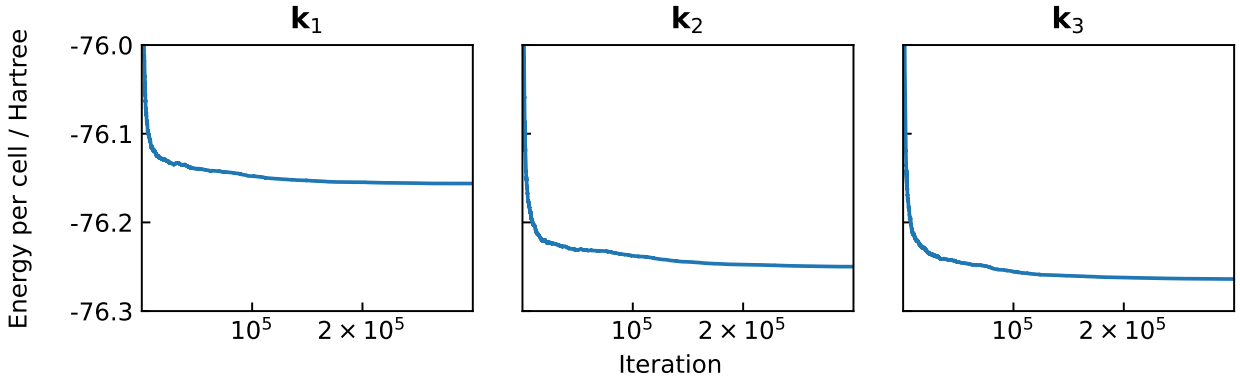
A 3×3 Monkhorst-Pack mesh in the first Brillouin zone of the supercell reciprocal space with Γ point centered is used to approximate the twist average integral, which reads

$$E_{\mathrm{TABC}} = \frac{\Omega_S}{(2\pi)^3} \int_{1.\mathrm{B.Z.}} d^3\mathbf{k}_S \frac{\Psi^*_{\mathbf{k}_S} \hat{H}_S \Psi_{\mathbf{k}_S}}{\Psi^*_{\mathbf{k}_S} \Psi_{\mathbf{k}_S}} \approx \frac{1}{9} E_{\mathbf{k}_1} + \frac{2}{3} E_{\mathbf{k}_2} + \frac{2}{9} E_{\mathbf{k}_3}, \tag{2}$$
$$\mathbf{k}_1 = 0,\ \mathbf{k}_2 = \frac{1}{3}\mathbf{b}^S_1 + \frac{1}{3}\mathbf{b}^S_2,\ \mathbf{k}_3 = \frac{2}{3}\mathbf{b}^S_1 + \frac{1}{3}\mathbf{b}^S_2,$$

and the weight factors origin from the different number of symmetry equivalent $\mathbf{k}$ points.

Supplementary Note 3.3 Training curves

Training curves at different $\mathbf{k}_S$ are plotted in Supplementary Fig. 2.



Supplementary Figure 2 | 2×2 **Graphene training curve.** For clarity, at each iteration number, we show the median energy per primitive cell over the last 10% of iteration.

The final results are listed in Supplementary Table 6. The energy of an isolated carbon atom is taken from Ref. [2], $E = -37.84471$ Hartree.

	$\mathbf{k}_1$	$\mathbf{k}_2$	$\mathbf{k}_3$
Energy (Hartree)	-76.15588(6)	-76.24949(5)	-76.26314(5)

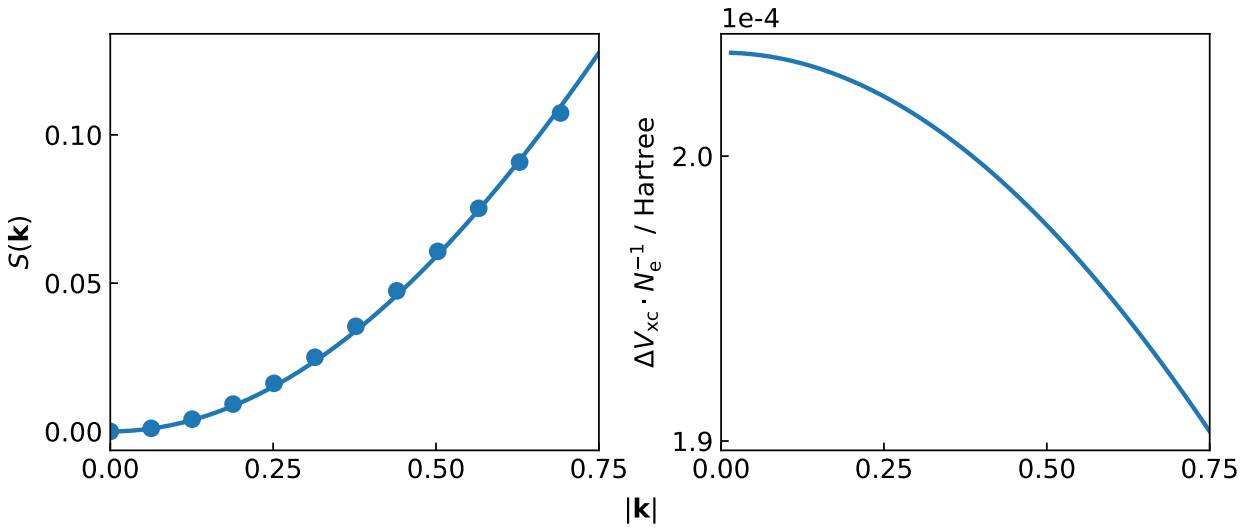
Supplementary Table 6 | Energy of graphene at different twists

Supplementary Note 3.4 Structure factor correction

TABC technique is usually combined with structure factor corrections [3], and the combination is now seen as the standard scheme of applying QMC to solids. Structure factor $S(\mathbf{k})$ is calculated to correct the exchange-correlation part, namely V_{xc}, of the total potential energy, which reads

$$\begin{aligned} \frac{\Delta V_{\mathrm{xc}}}{N_{\mathrm{e}}} &= \frac{2\pi}{\Omega_S} \lim_{\mathbf{k}\to 0} \frac{S(\mathbf{k})}{\mathbf{k}^2} , \\ S(\mathbf{k}) = \frac{1}{N_{\mathrm{e}}} [\langle \rho(\mathbf{k})\rho^*(\mathbf{k})\rangle - \langle \rho(\mathbf{k})\rangle\langle \rho^*(\mathbf{k})\rangle] &, \ \rho(\mathbf{k}) = \sum_i \exp(i\mathbf{k}\cdot\mathbf{r}_i) , \end{aligned} \tag{3}$$

where $\mathbf{r}_i$ refers to the coordinate of each electron, and N_{e} denotes the number of electrons in the simulation cell. The calculated $S(\mathbf{k})$ and corresponding ΔV_{xc} of Γ point is plotted in Supplementary Fig. 3, and corrections of all twists are quite close to each other.



Supplementary Figure 3 | Structure factor correction of Graphene. The lines are fitted with the formula: $S(\mathbf{k}) = 1 - \exp(-a \cdot \mathbf{k}^2)$.

The final correction from structure factor is 0.00122 Hartree / atom.

Supplementary Note 4. Lithium hydride

Supplementary Note 4.1 Geometry

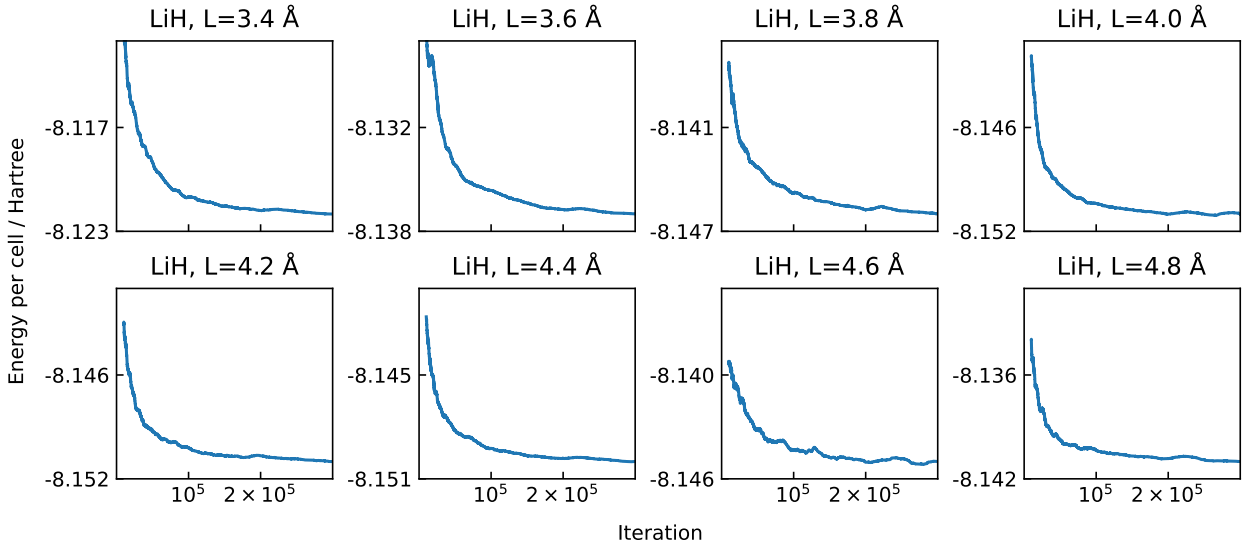
Lithium hydride crystal has a rock-salt structure, whose lattice vectors and atom positions are given in Supplementary Table 7.

Atom	Position	lattice vector	Position
Li	(0.0, 0.0, 0.0)	$\mathbf{a}_1$	(0.0, L/2, L/2)
H	(L/2, L/2, L/2)	$\mathbf{a}_2$	(L/2, 0.0, L/2)
		$\mathbf{a}_3$	(L/2, L/2, 0.0)

Supplementary Table 7 | Geometry of LiH crystal

Supplementary Note 4.2 Training curves

Training curves of the $2 \times 2 \times 2$ LiH crystal is plotted in Supplementary Fig. 4.



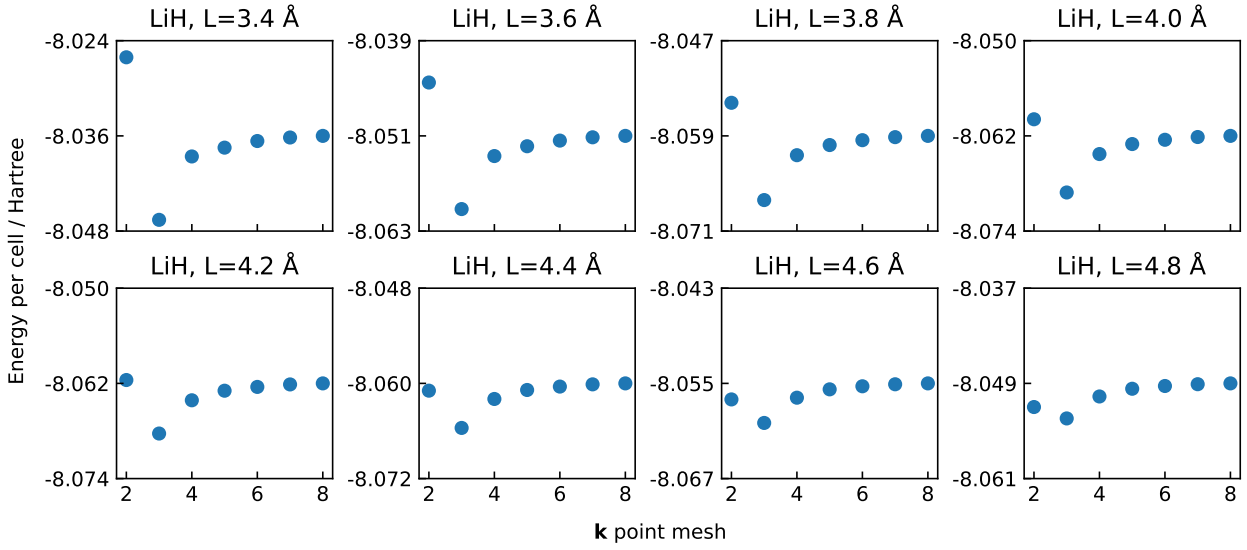
Supplementary Figure 4 | $2 \times 2 \times 2$ LiH training curve. For clarity, at each iteration number, we show the median energy of primitive cell over the last 10% of iteration.

Supplementary Note 4.3 Dissociation curve

The energy of $2 \times 2 \times 2$ LiH is listed in Supplementary Table 8. The energy of an isolated lithium atom is taken from Ref. [2], $E = -7.47798$ Hartree. Corresponding Hatree-Fock corrections are calculated with the ccpvdz basis set and the convergence behavior of HF calculation is plotted in Supplementary Fig. 5.

L (Å)	Net	HF correction	L (Å)	Net	HF correction
3.4	-8.12185(1)	-0.0099	4.2	-8.15112(1)	-0.0004
3.6	-8.13738(1)	-0.0067	4.4	-8.14967(1)	0.0009
3.8	-8.146147(1)	-0.0042	4.6	-8.14502(1)	0.0020
4.0	-8.15096(1)	-0.0021	4.8	-8.14094(1)	0.0030

Supplementary Table 8 | Energy of $2 \times 2 \times 2$ LiH crystal. Energies are all given in Hartree.



Supplementary Figure 5 | Hartree-Fock corrections. The convergence behavior of HF calculations with respect to the number of **k** points.

Supplementary Note 4.4 Birch-Murnaghan fit

The third order Birch-Murnaghan equation of state is employed to fit the dissociation curve, which reads

$$E(V) = E_0 + \frac{9V_0B_0}{16}\left\{\left[\left(\frac{V_0}{V}\right)^{2/3} - 1\right]^3 B_0' + \left[\left(\frac{V_0}{V}\right)^{2/3} - 1\right]^2\left[6 - 4\left(\frac{V_0}{V}\right)^{2/3}\right]\right\}, \tag{4}$$

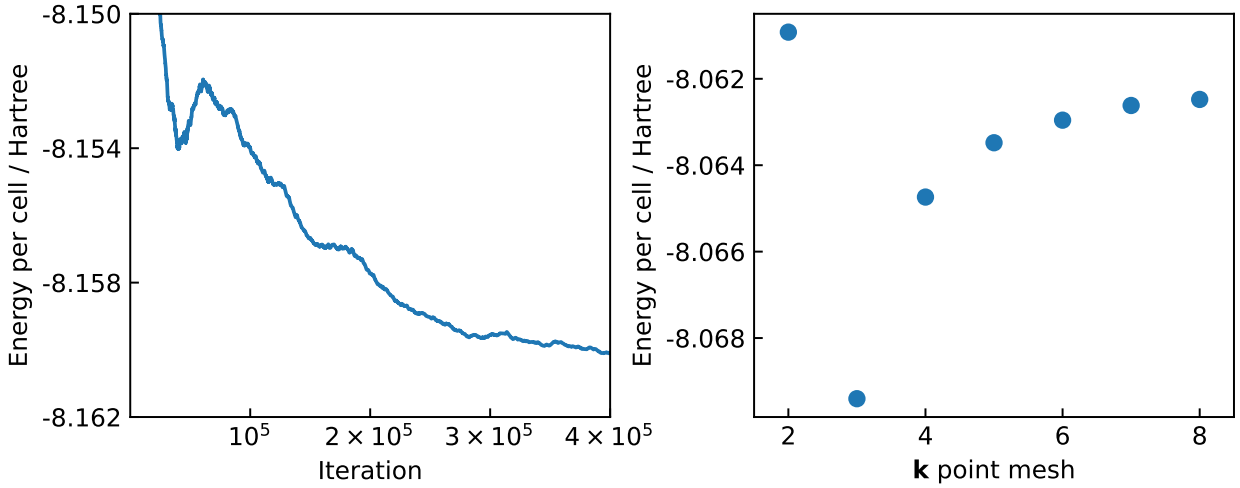
where E_0, V_0, B_0, B_0' are fitted quantities, their results and corresponding experiment data [4] are listed in Supplementary Table 9.

	a_0 (Å)	B_0 (GPa)	E_{coh} (eV)
Net	4.022	36.89	-4.757
Exp	4.061(1)	33-38	-4.778,-4.759

Supplementary Table 9 | Parameters of Birch-Murnaghan equation of state

Supplementary Note 4.5 $3\times3\times3$ LiH

The training curve of the $3\times3\times3$ LiH crystal at its equilibrium lattice constant L = 4.061Å is plotted in Supplementary Fig. 6, corresponding Hartree-Fock corrections are also given. The final inference results from neural network are listed in Supplementary Table 10.



Supplementary Figure 6 | $3\times3\times3$ LiH. Left panel plots the training curve of the $3\times3\times3$ LiH. For clarity, at each iteration number, we show the median energy per unit cell over the last 10% of iteration. Right panel plots the corresponding Hartree-Fock corrections with the ccpvdz basis set.

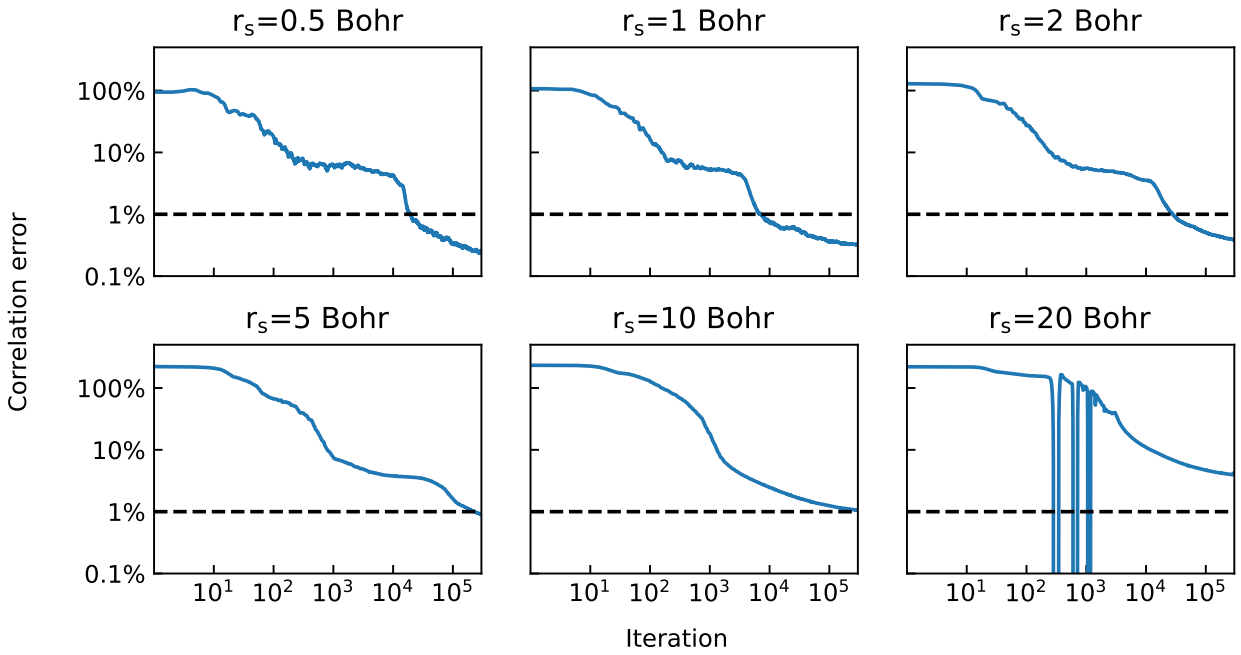
L (Å)	Net	HF correction
4.061	-8.16020(2)	0.0069

Supplementary Table 10 | Energy of the $3\times3\times3$ LiH crystal. Energies are all given in Hartree.

Supplementary Note 5. Homogeneous electron gas

Supplementary Note 5.1 Training curve

The training curve of HEG system containing 54 electrons is plotted in Supplementary Fig. 7. E_{HF} and E_{DMC} are taken from Ref. [5]. Final results of neural network, BF-DMC, BF-VMC and DCD [5, 6] are listed in Supplementary Table 11.



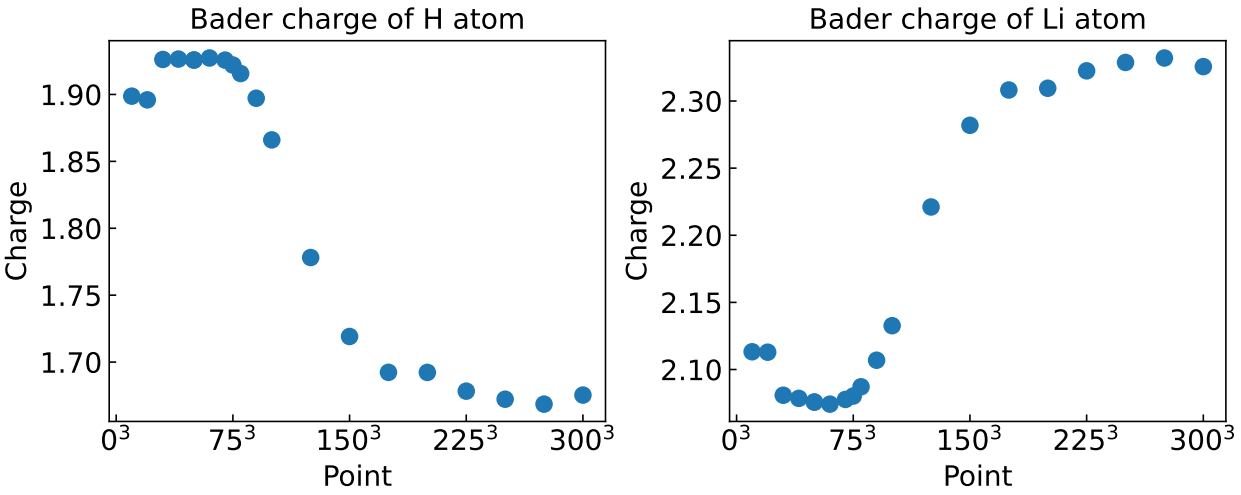
Supplementary Figure 7 | Homogeneous electron gas. For clarity, at each iteration number, we show the median correlation error over the last 10% of iteration.

r_s	Net	BF-DMC	BF-VMC	DCD
0.5	3.221226(2)	3.22112(4)	3.22132(7)	3.22052
1	0.530019(1)	0.52989(4)	0.53009(3)	0.53001
2	-0.013840(1)	-0.013966(9)	-0.01382(2)	-0.01286
5	-0.0788354(2)	-0.079036(3)	-0.078961(5)	-0.07655
10	-0.0542785(1)	-0.054443(2)	-0.054389(2)	-0.05157
20	-0.0316886(1)	-0.032047(2)	-0.0319984(8)	-0.02925

Supplementary Table 11 | Energy per electron of HEG at different mean radius of electrons r_s. HEG system contains 54 electrons and energies are all given in Hartree, r_s is given in Bohr.

Supplementary Note 6. Bader charge analysis

Detailed Bader charge analysis [7] is applied to conventional LiH crystal, and the result is plotted in Fig. 8.



Supplementary Figure 8 | Bader charge of LiH crystal. Calculated Bader charge of atoms in LiH crystal, point denotes the number uniformly dividing the crystal.

According to the result, Li and H atoms in LiH become $Li^{0.67+}$ and $H^{0.67-}$ ions respectively.

References

[1] Mario Motta, David M. Ceperley, Garnet Kin-Lic Chan, John A. Gomez, Emanuel Gull, Sheng Guo, Carlos A. Jiménez-Hoyos, Tran Nguyen Lan, Jia Li, Fengjie Ma, Andrew J. Millis, Nikolay V. Prokof'ev, Ushnish Ray, Gustavo E. Scuseria, Sandro Sorella, Edwin M. Stoudenmire, Qiming Sun, Igor S. Tupitsyn, Steven R. White, Dominika Zgid, and Shiwei Zhang. 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, Sep 2017.

[2] David Pfau, James S. Spencer, Alexander G. D. G. Matthews, and W. M. C. Foulkes. Ab initio solution of the many-electron schrödinger equation with deep neural networks. *Phys. Rev. Research*, 2:033429, Sep 2020.

[3] Simone Chiesa, David M. Ceperley, Richard M. Martin, and Markus Holzmann. Finite-size error in many-body simulations with long-range interactions. *Phys. Rev. Lett.*, 97:076404, Aug 2006.

[4] S. J. Binnie, S. J. Nolan, N. D. Drummond, D. Alfè, N. L. Allan, F. R. Manby, and M. J. Gillan. Bulk and surface energetics of crystalline lithium hydride: Benchmarks from quantum monte carlo and quantum chemistry. *Phys. Rev. B*, 82:165431, Oct 2010.

[5] P. López Ríos, A. Ma, N. D. Drummond, M. D. Towler, and R. J. Needs. Inhomogeneous backflow transformations in quantum monte carlo calculations. *Phys. Rev. E*, 74:066701, Dec 2006.

[6] Ke Liao, Thomas Schraivogel, Hongjun Luo, Daniel Kats, and Ali Alavi. Towards efficient and accurate ab initio solutions to periodic systems via transcorrelation and coupled cluster theory. *Phys. Rev. Research*, 3:033072, Jul 2021.

[7] W Tang, E Sanville, and G Henkelman. A grid-based bader analysis algorithm without lattice bias. *Journal of Physics: Condensed Matter*, 21(8):084204, jan 2009.