YITP Workshop DFT2022

Recent Advances in Orbital-Free Density Functional Theory and Machine Learning Approaches to Develop Kinetic Energy Functionals

Fumihiro Imoto (Schrödinger Inc.)

Previously at Oshiyama Group (Nagoya Univ.) and Imada Group (Waseda Univ.)

Contents

Introduction: Orbital-free density functional theory (OFDFT)

- Kohn-Sham density functional theory (KSDFT)

$$E_{\mathrm{tot}}[\rho] = T_s[\rho] + E_{\mathrm{H}}[\rho] + E_{\mathrm{xc}}[\rho] + E_{\mathrm{ext}}[\rho]$$

Kinetic energy $T_s\left[\{\phi_i\}_{i=1}^{N}\right] = \frac{1}{2}\sum_{i=1}^{N}\int |\nabla\phi_i(\boldsymbol{r})|^2 d\boldsymbol{r}$

Matrix diagonalization and orbital orthonormalization with $O(N^3)$

- Orbital-free density functional theory (OFDFT), $O(N)$ method

Open problem 1:

✓ Local pseudopotentials (PPs) should be used instead of nonlocal PPs.

✓ ***Ab initio* pseudopotentials are available only for limited elements.**

Open problem 2:

✓ **The optimal method to solve OFDFT Euler equation has not been established.**

Open problem 3:

✓ **Accurate kinetic energy density functional (KEDF) $T_s[\rho]$ has not yet been constructed for non-metallic systems.**

Introduction: OFDFT scheme and KEDFs

- OFDFT scheme

Euler equation $\frac{\delta T_s[\rho]}{\delta\rho(\boldsymbol{r})} + v_{\text{ext}}(\boldsymbol{r}) + \frac{\delta E_{\text{H}}[\rho]}{\delta\rho(\boldsymbol{r})} + \frac{\delta E_{\text{xc}}[\rho]}{\delta\rho(\boldsymbol{r})} = \mu$ Kinetic energy functional derivative (KEFD)

Kohn-Sham (KS) potential $v_{\text{KS}}([\rho];\boldsymbol{r})$

$$\Leftrightarrow \left[-\frac{\nabla^2}{2} + \left(\frac{\delta T_s[\rho]}{\delta\rho(\boldsymbol{r})} + \frac{\nabla^2\sqrt{\rho(\boldsymbol{r})}}{2\sqrt{\rho(\boldsymbol{r})}}\right) + v_{\text{KS}}([\rho];\boldsymbol{r})\right]\sqrt{\rho(\boldsymbol{r})} = \mu\sqrt{\rho(\boldsymbol{r})}$$

Schrödinger-like equation[1] that is solved self-consistently in our work. [1] M. Levy et al., Phys. Rev. A **30**, 2745 (1984).

- Kinetic energy density functionals (KEDFs)

1. Semilocal KEDFs defined as enhancement factors
   - ✓ lattice constants (<2 % error), bulk moduli (<15 % error)

Generalized gradient approximation (GGA) $T_s[\rho] = \int \tau^{\text{TF}}(\boldsymbol{r})F[\rho]d\boldsymbol{r}$ enhancement factor

State-of-the-art semilocal KEDFs:

- **LKT** (Luo-Karasiev-Trickey)[2]
- **PGSLβ** (Pauli-Gaussian Second order and Laplacian with a parameter β)[3]

Thomas-Fermi (TF) KEDF is exact in homogeneous density.

[2] K. Luo et al., Phys. Rev. B **98**, 041111(R) (2018). [3] L. A. Constantin et al., J. Phys. Chem. Lett. **9**, 4385 (2018).

2. Nonlocal KEDFs
   - ✓ lattice constants (<1 % error), bulk moduli (<10 % error)
   - ✓ heavier computational cost with $O(N\log N)$ scaling

Recent advances: Nonlocal pseudopotential (NLPP) energy density functional

- Nonlocal pseudopotentials in OFDFT scheme

$$E_{\mathrm{tot}}[\rho] = T_s[\rho] + E_{\mathrm{H}}[\rho] + E_{\mathrm{xc}}[\rho] + E_{\mathrm{II}}[\{R_a\}] + \int \rho(\mathbf{r}) V_{loc}(\mathbf{r}) dr + E_{nl}[\rho]$$

$$E_{nl}[\rho] = \int\int V_{nl}(\mathbf{r}', \mathbf{r}) \gamma_s(\mathbf{r}, \mathbf{r}') d\mathbf{r} d\mathbf{r}'$$

$$V_{nl}(\mathbf{r}', \mathbf{r}) = \langle \mathbf{r}' | \hat{V}_{nl} | \mathbf{r} \rangle$$

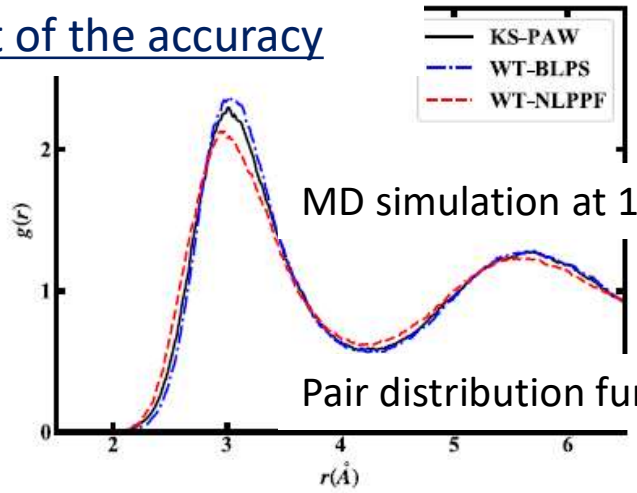
$$\gamma_s(\mathbf{r}, \mathbf{r}') = \sum_i f_i \psi_i(\mathbf{r}) \psi_i^*(\mathbf{r}')$$

Modified Gaussian form derived from the second-order Taylor expansions of the density matrix:

$$\gamma_s^{MG}[\rho](\mathbf{r}, \mathbf{r}') = \rho(\bar{\mathbf{r}}) \exp\left(-\frac{s^2}{2\beta(\mathbf{r})}\right)\left[1 + A\left(\frac{s^2}{2\beta(\bar{\mathbf{r}})}\right)\right]$$

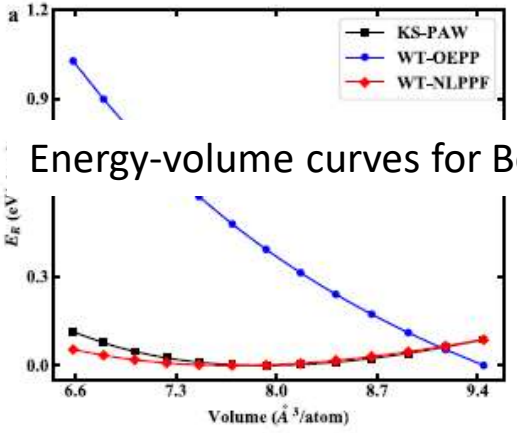
$$s = |\mathbf{r} - \mathbf{r}'|, \bar{\mathbf{r}} = (\mathbf{r} + \mathbf{r}')/2$$

Assessment of the accuracy



MD simulation at 1000 K.

Pair distribution function for $Li_{54}Mg_{54}$ alloy.



Energy-volume curves for Be-HCP.

[Ref.] Q. Xu et al., Nature Comm. **13**, 1385 (2022).

Recent advances: Nonlocal pseudopotential (NLPP) energy density functional

- Nonlocal pseudopotentials in OFDFT scheme

Assessment of the accuracy

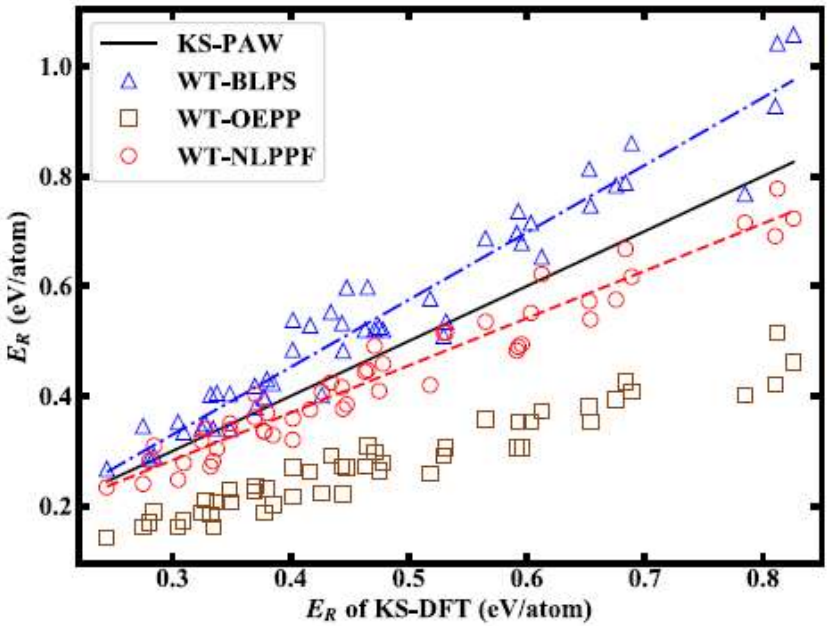


Fig. 2 Relative energies for random structures of elemental Li. The results are calculated by OF-DFT using BLPS, OEPP and NLPPF in comparison with that by KS-DFT using the PAW method. Blue dash-dotted and red dashed lines are the least-square fittings of WT-BLPS and WT-NLPPF results, respectively.

Table 1 B_0 (GPa), E_R (eV/atom), and V_0 (Å^3/atom) for bulk Li, Mg, and Be by KS-DFT and OF-DFT.

		Method	HCP	FCC	BCC	SC
Li	B_0	KS-PAW	13.9	13.6	13.9	12.1
		WT-NLPPF	13.5	13.5	13.7	11.0
	V_0	KS-PAW	20.280	20.372	20.396	20.580
		WT-NLPPF	19.483	19.462	19.352	20.844
	E_R	KS-PAW	0.000	0.000	0.001	0.120
		WT-NLPPF	0.000	0.000	0.001	0.152
Mg	B_0	KS-PAW	35.8	35.5	34.8	22.7
		WT-NLPPF	33.0	31.3	31.3	21.2
	V_0	KS-PAW	22.838	23.071	22.826	27.478
		WT-NLPPF	23.194	23.924	23.730	28.274
	E_R	KS-PAW	0.000	0.012	0.029	0.382
		WT-NLPPF	0.000	0.011	0.031	0.372
Be	B_0	KS-PAW	123.3	119.7	124.1	74.5
		WT-NLPPF	91.5	90.5	87.2	63.3
	V_0	KS-PAW	7.910	7.875	7.822	10.274
		WT-NLPPF	7.690	7.942	7.798	10.160
	E_R	KS-PAW	0.000	0.080	0.099	1.004
		WT-NLPPF	0.000	0.058	0.082	0.561

[Ref.] Q. Xu et al., Nature Comm. **13**, 1385 (2022).

Previous works and aims of the work[Ref.]

Previous neural network (NN) KEDF

- ✓ Semilocal NN KEDFs $T^{\mathrm{NN}}[\rho]$ were developed as the enhancement factors $F^{\mathrm{NN}}[\rho]$.
- ✓ OFDFT Euler eq. was not solved with the functional derivative of the NN KEDF.

$$T^{\mathrm{NN}}[\rho] = \int \tau^{\mathrm{TF}}(\boldsymbol{r}) F^{\mathrm{NN}}[\rho] d\boldsymbol{r}$$

Training data was taken from atoms/molecules[1-4] or periodic solids[5].
Refs. [4,5] use inverse distances from nuclei positions $\{\boldsymbol{R}_m\}$, $1/|\boldsymbol{r} - \boldsymbol{R}_m|$, as inputs for NN.

[1] K. Yao and J. Parkhill, J. Chem. Theory Comput. **12**, 1139 (2016). [2] J. Seino *et al.*, J. Chem. Phys. **148**, 241705 (2018). [3] J. Seino *et al.*, Chem. Phys. Lett. **734**, 136732 (2019). [4] M. Fujinami *et al.*, Chem. Phys. Lett. **748**, 137358 (2020). [5] P. Golub and S. Manzhos, Phys. Chem. Chem. Phys. 21, 378 (2019).

- Develop a scheme to solve OFDFT Euler eq. using the functional derivative of the NN KEDF $T^{\mathrm{NN}}[\rho]$.

Schrödinger-like eq. that is equivalent to the Euler eq.

$$\left[-\frac{\nabla^2}{2} + \left(\frac{\delta T^{\mathrm{NN}}[\rho]}{\delta \rho(\boldsymbol{r})} + \frac{\nabla^2\sqrt{\rho(\boldsymbol{r})}}{2\sqrt{\rho(\boldsymbol{r})}}\right) + v_{\mathrm{KS}}([\rho];\boldsymbol{r})\right]\sqrt{\rho(\boldsymbol{r})} = \mu\sqrt{\rho(\boldsymbol{r})}$$

- Predict structural parameters of semiconductors via the solution of OFDFT Euler eq.

[Ref.] F. Imoto, M. Imada, and A. Oshiyama, Phys. Rev. Research **3**, 033198 (2021).

Methods: Neural network (NN) used for our KEDF

- General definition of NN

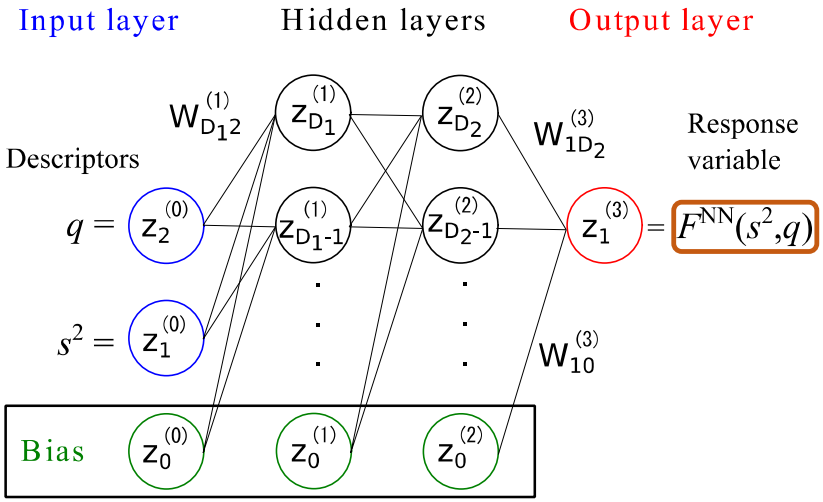
$$z_j^{(l)} = \sigma^{(l)}\left(a_j^{(l)}\right) = \sigma^{(l)}\left(\sum_{k=0}^{D_{l-1}} W_{jk}^{(l)} z_k^{(l-1)}\right)$$

$\sigma^{(l)}\left(a_j^{(l)}\right)$ is ELU for each layer with $1 \le l \le N-1$ and an identity function for an output layer ($l = N$).

- Definition of our NN KEDF

$$T^{\mathrm{NN}}[\rho] = \int \tau^{\mathrm{TF}}(\boldsymbol{r}) F^{\mathrm{NN}}(s^2, q; \boldsymbol{W}) d\boldsymbol{r}$$

NN KEDF as the enhancement factor



$z_j^{(l)}$: output from the j-th neuron in the l-th layer
l: layer index
D_l: number of neurons in the l-th layer
$\{W_{jk}^{(l)}\}$: fitting parameters trained in NN

ELU (exponential linear unit):

$$\sigma\left(a_j^{(l)}\right) = \begin{cases} a_j^{(l)}, & \text{if } a_j^{(l)} > 0 \\ \exp\left(a_j^{(l)}\right) - 1, & \text{if } a_j^{(l)} \le 0 \end{cases}$$

Dimensionless gradients and Laplacian:

$$s = |\nabla\rho| / \left[2(3\pi^2)^{1/3}\rho^{4/3}\right]$$
$$q = \nabla^2\rho / \left[4(3\pi^2)^{2/3}\rho^{5/3}\right]$$

Thomas-Fermi (TF) KEDF:

$$\tau^{\mathrm{TF}} = (3/10)(3\pi^2)^{2/3}\rho^{5/3}$$

exact form in homogeneous density

Methods: Training of NN toward KEFD

- We minimize the cost function for kinetic energy functional derivative (KEFD).

$$L = \frac{1}{N_t}\sum_{n=1}^{N_t}\frac{1}{2}\left[\frac{\delta T^{\mathrm{NN}}(\boldsymbol{r}_n)}{\delta\rho} - \frac{\delta T^{\mathrm{KS}}(\boldsymbol{r}_n)}{\delta\rho}\right]^2$$

N_t: total number of of the training data
13,824 real-space grid points $\boldsymbol{r}_n$ in diamond

- The training data $\delta T^{\mathrm{KS}}(\boldsymbol{r})/\delta\rho$ is computed from Kohn-Sham (KS) eigenvalues ε_i and KS orbitals $\phi_i(\boldsymbol{r})$.

$$\frac{\delta T^{\mathrm{KS}}(\boldsymbol{r})}{\delta\rho} = \frac{1}{\rho^{\mathrm{KS}}(\boldsymbol{r})}\sum_i f_i\left[-\frac{1}{2}\phi_i^*(\boldsymbol{r})\nabla^2\phi_i(\boldsymbol{r}) + \left(\varepsilon^{\mathrm{HOKS}} - \varepsilon_i\right)|\phi_i(\boldsymbol{r})|^2\right]$$

$\varepsilon^{\mathrm{HOKS}}$: the highest-occupied KS eigenvalue
ε_i: KS eigenvalue of the i-th KS orbital
f_i: occupation number of the i-th KS orbital
$\rho^{\mathrm{KS}}(\boldsymbol{r}) = \sum_i f_i\,|\phi_i(\boldsymbol{r})|^2$: KS density

✓ $\boldsymbol{F^{\mathrm{NN}}}$ **is trained** by also using the following analytical expression of KEFD

$$\frac{\delta T^{\mathrm{NN}}}{\delta\rho} = c_0\rho^{2/3}\left[\frac{5}{3}F^{\mathrm{NN}} - \frac{8}{3}s^2\frac{\partial F^{\mathrm{NN}}}{\partial(s^2)} - \frac{5}{3}q\frac{\partial F^{\mathrm{NN}}}{\partial q}\right] - \frac{3}{20}\nabla\cdot\left[\frac{\partial F^{\mathrm{NN}}}{\partial(s^2)}\frac{\nabla\rho}{\rho}\right] + \frac{3}{40}\nabla^2\left(\frac{\partial F^{\mathrm{NN}}}{\partial q}\right)$$

$c_0 = 3(3\pi^2)^{2/3}/10$

Next, we apply physical constraints on NN KEDF.

Methods: Training of NN toward KEFD

- Functional derivative of subsystem NN KEDFs

$$\frac{\delta \tilde{T}_s}{\delta \rho} = c_0 \rho^{2/3}\left[\frac{5}{3}\tilde{F} - \frac{8}{3}s^2\frac{\partial \tilde{F}}{\partial (s^2)} - \frac{10}{3}q^2\frac{\partial \tilde{F}}{\partial (q^2)}\right] - \frac{3}{20}\nabla\cdot\left[\frac{\partial \tilde{F}}{\partial (s^2)}\frac{\nabla\rho}{\rho}\right] + \frac{3}{20}\nabla^2\left(q\frac{\partial \tilde{F}}{\partial (q^2)}\right)$$

- Training meta-GGA functional derivative

$$L\left(\gamma_{11}^{(L)}, \gamma_{12}^{(L)}\right) = \frac{1}{N_\mathrm{t}}\sum_{i=1}^{N_\mathrm{t}}\frac{1}{2}\left[\frac{\delta T_s^{\mathrm{NN}}(\mathbf{r}_i)}{\delta\rho} - \frac{\delta T_s^{\mathrm{KS}}(\mathbf{r}_i)}{\delta\rho}\right]^2$$

$$\frac{\partial L}{\partial w_{ji}^{(l)}} = \frac{1}{N_\mathrm{t}}\sum_{i=1}^{N_\mathrm{t}}\left[\frac{\delta T_s^{\mathrm{NN}}(\mathbf{r}_i)}{\delta\rho} - \frac{\delta T_s^{\mathrm{KS}}(\mathbf{r}_i)}{\delta\rho}\right]\frac{\partial}{\partial w_{ji}^{(l)}}\left(\frac{\delta T_s^{\mathrm{NN}}(\mathbf{r}_i)}{\delta\rho}\right)$$

$$\frac{\partial}{\partial w_{ji}^{(l)}}\left(\frac{\delta T_s^{\mathrm{NN}}(\mathbf{r}_i)}{\delta\rho}\right) = c_0\rho^{2/3}\left[\frac{5}{3}\frac{\partial z_1^{(L)}}{\partial w_{ji}^{(l)}} - \frac{8}{3}s^2\frac{\partial \gamma_{11}^{(L)}}{\partial w_{ji}^{(l)}} - \frac{10}{3}q^2\left(\frac{\partial \gamma_{12}^{(L)}}{\partial w_{ji}^{(l)}}\right)\right] - \frac{3}{20}\nabla\cdot\left[\frac{\partial \gamma_{11}^{(L)}}{\partial w_{ji}^{(l)}}\frac{\nabla\rho}{\rho}\right] + \frac{3}{20}\nabla^2\left(\frac{\partial \gamma_{12}^{(L)}}{\partial w_{ji}^{(l)}}\right)$$

$\frac{\partial F^{\mathrm{NN}}}{\partial (s^2)}$ $\frac{\partial F^{\mathrm{NN}}}{\partial (q^2)}$

For each training data, we also need to "feedforward" & "backpropagate" using "neighbor data".

[Note] $\partial \gamma_{1k}^{(L)}/\partial w_{ji}^{(l)}$ can be calculated by the "generalized backpropagation".

[Ref.] IEEE Trans. Neural. Netw. **22** , 936-947 (2011).

Methods: Training of NN toward KEFD

Natural gradient method

$$G_{ij}(\{w\}) = \frac{1}{N_b}\sum_{p=1}^{N_b}\frac{\partial}{\partial w^i}\left(\frac{\delta T_s^{\mathrm{NN}}(\mathbf{r}_p)}{\delta\rho}\right)\frac{\partial}{\partial w^j}\left(\frac{\delta T_s^{\mathrm{NN}}(\mathbf{r}_p)}{\delta\rho}\right)$$

w^i: one-dimentionally arranged NN parameter {**w**}, # of parameters=N_w

Cost function: $L = \frac{1}{N_b}\sum_{p=1}^{N_b}\frac{1}{2}\left[\frac{\delta T_s^{\mathrm{NN}}(\mathbf{r}_p)}{\delta\rho} - \frac{\delta T_s^{\mathrm{KS}}(\mathbf{r}_p)}{\delta\rho}\right]^2$

- Training using the metric by subsystem NN KEDF

Directly optimize the subsystem KEDF $\tilde{T}_s$

$$\tilde{F}(s,q) \equiv X(q^4)F^{(0)}(s^2,q^2) + [1 - X(q^4)]F^{\mathrm{NN}}(s^2,q^2)$$

$$G_{ij}(\{w\}) = \frac{1}{N_b}\sum_{p=1}^{N_b}\frac{\partial}{\partial w^i}\left(\frac{\delta\tilde{T}_s(\mathbf{r}_p)}{\delta\rho}\right)\frac{\partial}{\partial w^j}\left(\frac{\delta\tilde{T}_s(\mathbf{r}_p)}{\delta\rho}\right)$$

Cost function: $L = \frac{1}{N_b}\sum_{p=1}^{N_b}\frac{1}{2}\left[\frac{\delta\tilde{T}_s(\mathbf{r}_p)}{\delta\rho} - \frac{\delta T_s^{\mathrm{KS}}(\mathbf{r}_p)}{\delta\rho}\right]^2$

Methods: Augmentation of the NN enhancement factor

- We augment NN enhancement factor as $\tilde{F}^{\mathrm{NN}}(s^2,q)$ by requiring to satisfy known physical constraints.

$$\tilde{F}^{\mathrm{NN}}(s^2,q) \equiv X(q)F^{(0)}(s^2,q\,) + [1-X(q)]F^{\mathrm{NN}}(s^2,q\,),$$

$$\text{where } X(q) = \exp(-Aq^4) \text{ and } F^{(0)} = \frac{5}{3}s^2 + e^{-(40/27)s^2} + \beta q^2.$$

$\tilde{F}^{\mathrm{NN}}(s^2,q)$ satisfies the following limits:

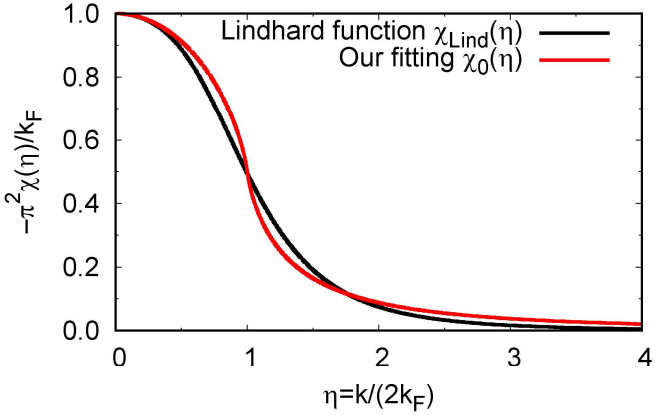
(a) $s \to 0, q \to 0$: $\tilde{F}^{\mathrm{NN}} \to 1 + 5s^2/27$

(b) $s \to \infty, q \to 0$: $\tilde{F}^{\mathrm{NN}} \to 5s^2/3$

We determine β as 0.382 so that the inverse of the response function computed by $\tilde{F}^{\mathrm{NN}}(s^2,q)$ reproduces the Lindhard function in $s \to 0, q \to 0$ limit.

Inverse response function by $\tilde{F}^{\mathrm{NN}}(s^2,q)$ in $s \to 0, q \to 0$:

$$-\frac{1}{\chi_0(\eta)} = \frac{\pi^2}{k_F}\left(1 + \frac{\eta^2}{3} + \frac{9}{5}\beta\eta^4\right)$$



$\eta = k/(2k_F)$, where k_F is the Fermi wavenumber.

[Note] A is determined as 31.6 so that self-consistent field (SCF) density becomes minimum.

Our NN KEDFs
- uninterpolated KEDF NN[bare]: $F^{\mathrm{NN}}(s^2,q)$
- augmented KEDF by interpolation NN: $\tilde{F}^{\mathrm{NN}}(s^2,q)$

Results: Accuracy of NN KEDF

- Self-consistent-field (SCF) electron density compared with the Kohn-Sham (KS) density

Root mean square error (RMSE) of SCF density in $10^{-2} \times \mathrm{bohr}^{-3}$ with respect to the KS density

	diamond(trained)	graphene	ds-Si	fcc-Si	β-tin Si	3C-SiC	bcc-Li	fcc-Al
NN	1.15	0.65	0.36	0.12	0.24	0.62	4.19	0.10
NN[bare]	0.94	0.48	0.31	0.09	0.21	0.61	2.15	0.08
PGSL0.25	1.29	0.71	0.44	0.16	0.26	0.71	5.30	0.12
LKT	1.68	0.67	0.51	0.11	0.31	0.96	5.57	0.11
TF(1/5)vW	2.06	0.32	0.55	0.58	0.49	1.20	5.76	0.35

ds: diamond structured

	fcc-Cu	bcc-Na	NaCl	ratio
NN	8.42	2.88	2.20	1
NN[bare]	8.78	1.65	1.28	0.78
PGSL0.25	11.2	4.12	3.47	1.26
LKT	12.9	4.30	3.78	1.36
TF(1/5)vW	25.8	5.08	3.18	2.15

"ratio" indicates the ratio of RMSE to NN's RMSE averaged over all 11 solids.

✓ State-of-the-art KEDFs:

1. PGSL0.25[1]
2. LKT[2]

✓ Conventional KEDF:

TF(1/5)vW (Thomas-Fermi + (1/5) von Weizsäcker)[3]

[1] L. A. Constantin et al., J. Phys. Chem. Lett. **9**, 4385 (2018). [2] K. Luo et al., Phys. Rev. B **98**, 041111(R) (2018). [3] A. Berk, Phys. Rev. A **28**, 1908 (1983).

Results: Accuracy of NN KEDF

- SCF electron density compared with the Kohn-Sham density

  ✓ Even for 7 atoms (Li, C, Na, Al, Si, Cl, Cu) and 6 molecules (Li_2, C_2, Na_2, Al_2, Si_2, Cl_2), "ratio" for NN is smaller than the previous KEDFs.

	ratio		
	11 solids	7 atoms	6 molecules
NN	1	1	1
$NN^{[bare]}$	0.78	1.95	0.83
PGSL0.25	1.26	3.26	1.28
LKT	1.36	1.74	1.45
TF(1/5)vW	2.15	1.03	1.46

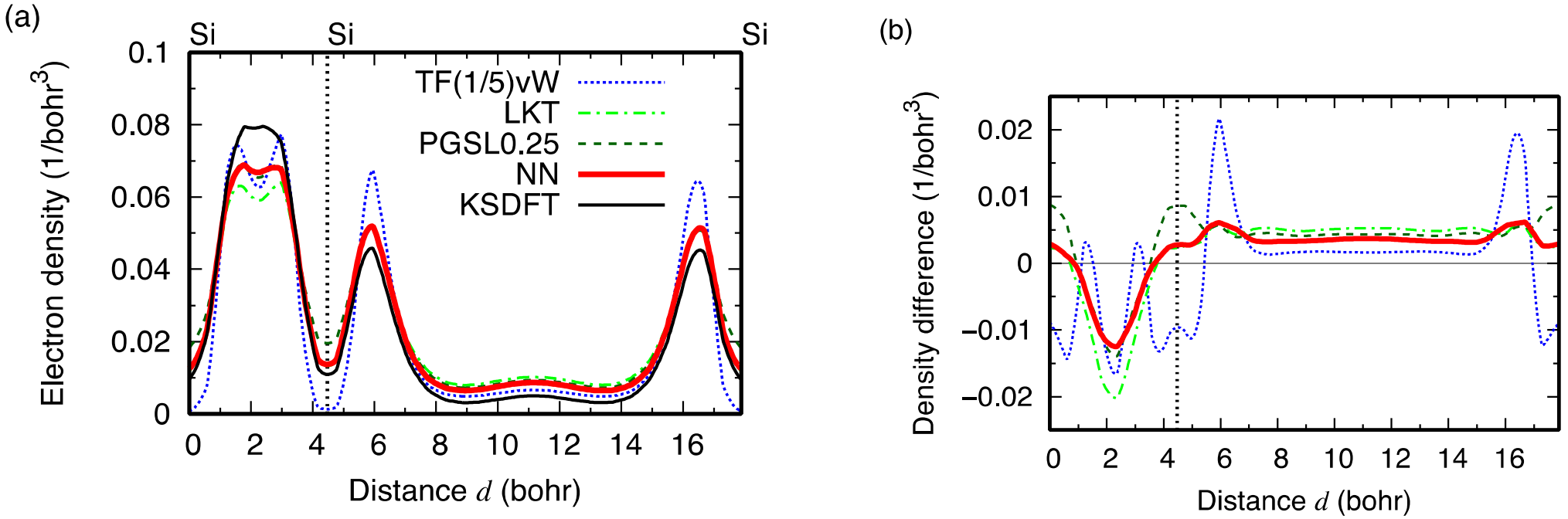
"ratio" indicates the ratio of RMSE to NN's RMSE averaged over all systems.

11 solids:
diamond, graphene, diamond-structured Si, fcc-Si, β-tin Si, 3C-SiC, bcc-Li, fcc-Al, fcc-Cu, bcc-Na, NaCl

Results: Accuracy of NN KEDF

- SCF electron density compared with the Kohn-Sham density

  Densities in diamond-structured (ds) Si



d: distance from a Si atom along [111]

In ds-Si, NN is more accurate than { PGSL0.25 and TF(1/5)vW at the nuclear site ($d = 4.47$ bohr).
LKT in the bonding region ($0 < d < 4.47$ bohr).

Results: Accuracy of NN KEDF

- Equilibrium lattice constants

The relative errors in % with respect to the KS values.

The Blank "—" indicates the total energy monotonically decreases with respect to the volume expansion.

	diamond (trained)	graphene	ds-Si	fcc-Si	β-tin Si a_0 (c_0/a_0)	3C-SiC	bcc-Li
NN	-2.53	0.29	-0.46	-0.22	0.17 (-1.67)	2.28	-0.83
NN[bare]	-11.0	-2.62	-3.80	-3.32	—	-3.94	-5.10
PGSL0.25	-2.47	-0.32	-0.15	1.42	1.69 (-1.67)	1.65	-0.33
LKT	-4.95	-1.59	-1.04	-0.16	-0.47 (-1.12)	1.42	-0.46
TF(1/5)vW	—	6.24	5.86	5.40	—	10.9	-3.54

	fcc-Al	fcc-Cu	bcc-Na	NaCl	MARE
NN	1.68	2.33	-0.91	3.61	**1.39**
NN[bare]	-5.16	6.64	-5.18	-5.35	4.74
PGSL0.25	3.62	4.12	-0.38	2.10	1.66
LKT	2.32	3.21	-0.50	2.12	1.66
TF(1/5)vW	4.76	4.21	-3.51	10.5	6.10

MARE:
Mean absolute relative errors in % with respect to the KS values averaged over the systems.

✓ NN performs best on average (minimum MARE)

Results: Accuracy of NN KEDF

- Bulk moduli

The relative errors in % with respect to the KS values.

The Blank "—" indicates the total energy monotonically decreases with respect to the volume expansion.

	diamond (trained)	ds-Si	fcc-Si	β-tin Si	3C-SiC	bcc-Li	fcc-Al
NN	25.7	-14.6	5.93	5.10	-36.2	4.03	2.06
NN[bare]	165	-16.5	10.4	—	2.85	26.8	66.5
PGSL0.25	32.4	-6.60	-13.0	-12.7	-11.8	3.78	-13.0
LKT	76.8	-4.00	19.5	18.5	-7.72	2.52	10.9
TF(1/5)vW	—	-60.4	-60.0	—	-74.3	10.1	-43.2

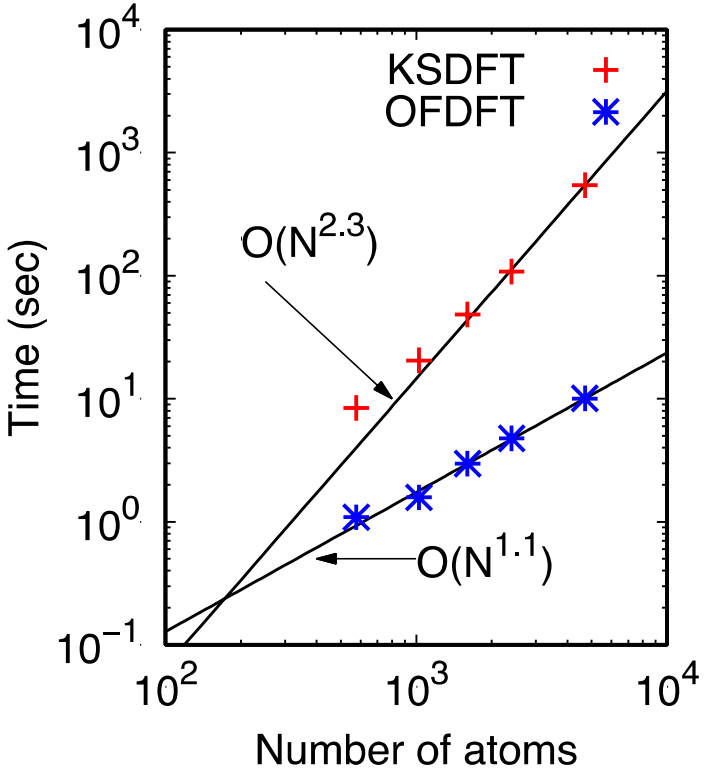
	fcc-Cu	bcc-Na	NaCl	MARE
NN	6.96	3.50	6.94	**11.1**
NN[bare]	44.3	26.2	25.1	38.4
PGSL0.25	-12.4	3.75	-11.5	12.1
LKT	10.7	2.61	-8.20	14.5
TF(1/5)vW	-44.1	10.1	-73.9	47.0

MARE:
Mean absolute relative errors in % with respect to the KS values averaged over the systems.

- ✓ NN performs best on average (minimum MARE)
- ✓ Poor performance of NN[bare] indicates the importance of the physical constraints.

Results: *O*(*N*) scaling of OFDFT computation

- Computational time for a single SCF iteration
  - ✓ 4H-SiC supercell with 576–4704 atoms
  - ✓ We implemented the OFDFT Euler eq. solver in real space using RSDFT code[1-3].



[1] J.-I. Iwata et al., J. Comput. Phys. **229**, 2339 (2010). [2] Y. Hasegawa et al., Int. J. High Perform. Comput. Appl. **28**, 335 (2014). [3] J.-I. Iwata, https://github.com/j-iwata/RSDFT.

Schrödinger-like eq. that is equivalent to the Euler eq.:

$$\left[-\frac{\nabla^2}{2}+\left(\frac{\delta T_s[\rho]}{\delta\rho(\boldsymbol{r})}+\frac{\nabla^2\sqrt{\rho(\boldsymbol{r})}}{2\sqrt{\rho(\boldsymbol{r})}}\right)+v_{\mathrm{KS}}([\rho];\boldsymbol{r})\right]\sqrt{\rho(\boldsymbol{r})}=\mu\sqrt{\rho(\boldsymbol{r})}$$

- We use 36 eigenvectors for the diagonalization.
- real-space grid-spacing = 0.39 Å
- $\sqrt{\rho(\boldsymbol{r})}$ is the eigenvector with the lowest eigenenergy.

Computational time in OFDFT scales as *O*(*N*).

Results: *O*(*N*) scaling of OFDFT computation

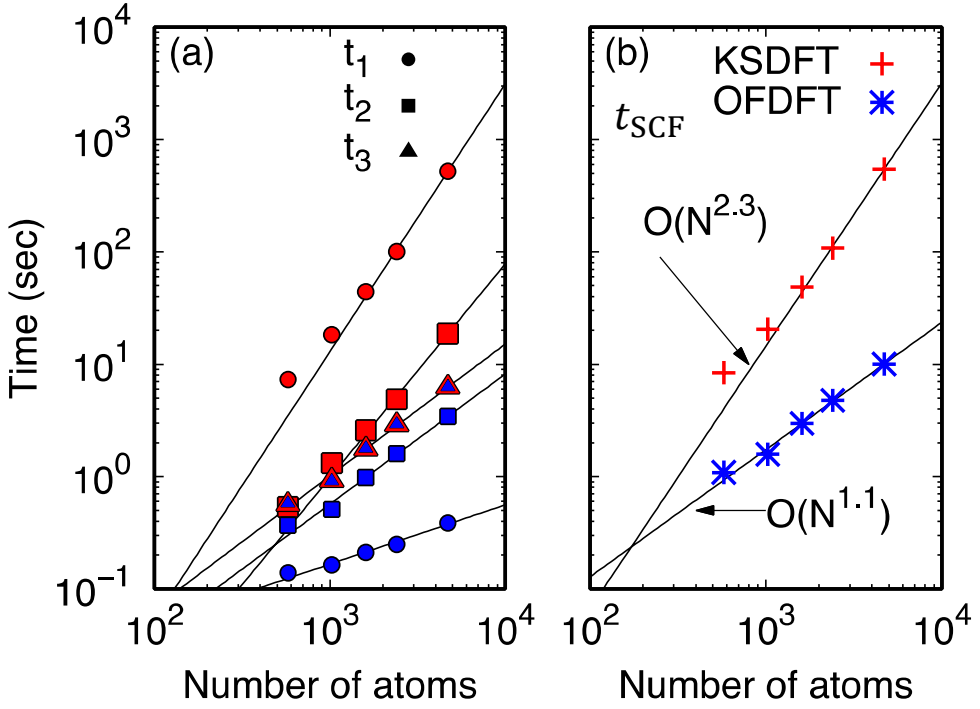
- Division of "SCF time/iteration" t_{SCF} into some segments.

$$t_{\mathrm{SCF}} = t_1 + t_2 + t_3 + t_{\mathrm{others}}$$

t_1: Subspace Diagonalization, Conjugate Gradient, Gram-Schmidt

t_2: Update density, Hartree, exchange-correlation potentials, total energy

t_3: Mixing procedure $\quad$ t_{others}: MPI communications, etc.



4H-SiC supercell with 576–4704 atoms.

In both (a) and (b), data points obtained by KSDFT and OFDFT are plotted in red and blue, respectively.

t_{SCF} in OFDFT scales as *O*(*N*).

Summary

- For 24 different systems including semiconductors, metals, an ionic meterial, atoms, and molecules, neural network (NN) kinetic energy density functional (KEDF) reproduces density better than previous KEDFs.
- For 11 solids, NN KEDF provides better structural properties than previous KEDFs.
- Our implementation of orbital-free density functional theory (OFDFT) shows *O*(*N*) scaling.

The choice of the kinetic energy functional derivative (KEFD) as training data is important.

✓ KEFD appears directly in the Euler equation.

$$\left[-\frac{\nabla^2}{2}+\left(\frac{\delta T_s[\rho]}{\delta\rho(\boldsymbol{r})}+\frac{\nabla^2\sqrt{\rho(\boldsymbol{r})}}{2\sqrt{\rho(\boldsymbol{r})}}\right)+v_{\mathrm{KS}}([\rho];\boldsymbol{r})\right]\sqrt{\rho(\boldsymbol{r})}=\mu\sqrt{\rho(\boldsymbol{r})}$$

✓ KEFD learning assists reproducing the physical quantity related to the first derivative of the total energy.

$$\frac{\delta E_{\mathrm{tot}}[\rho]}{\delta\rho(\boldsymbol{r})}=\frac{\delta T_s[\rho]}{\delta\rho(\boldsymbol{r})}+v_{\mathrm{ext}}(r)+\frac{\delta E_{\mathrm{H}}[\rho]}{\delta\rho(\boldsymbol{r})}+\frac{\delta E_{\mathrm{xc}}[\rho]}{\delta\rho(\boldsymbol{r})}$$

END