

Machine learning based nonlocal kinetic energy density functional for simple metals and alloysLiang Sun¹ and Mohan Chen^{1,2,*}¹*HEDPS, CAPT, School of Physics and College of Engineering, Peking University, Beijing 100871, People's Republic of China*²*AI for Science Institute, Beijing 100080, People's Republic of China*

(Received 24 October 2023; accepted 27 February 2024; published 18 March 2024)

Developing an accurate kinetic energy density functional (KEDF) remains a major hurdle in orbital-free density functional theory. We propose a machine-learning-based physical-constrained nonlocal (MPN) KEDF and implement it with the usage of the bulk-derived local pseudopotentials and plane wave basis sets in the ABACUS package. The MPN KEDF is designed to satisfy three exact physical constraints: the scaling law of electron kinetic energy, the free electron gas limit, and the non-negativity of Pauli energy density. The MPN KEDF is systematically tested for simple metals, including Li, Mg, Al, and 59 alloys. We conclude that incorporating nonlocal information for designing new KEDFs and obeying exact physical constraints are essential to improve the accuracy, transferability, and stability of ML-based KEDF. These results shed new light on the construction of ML-based functionals.

DOI: [10.1103/PhysRevB.109.115135](https://doi.org/10.1103/PhysRevB.109.115135)**I. INTRODUCTION**

Kohn-Sham density functional theory (KSDFT) is a widely used *ab initio* method in materials science [1,2]. However, its computational complexity of $O(N^3)$, where N is the number of atoms, poses significant challenges for large systems. Alternatively, orbital-free density functional theory (OFDFT) [3,4] calculates the noninteracting electron kinetic energy T_s directly from the charge density instead of relying on the one-electron Kohn-Sham orbitals. As a result, OFDFT achieves a more affordable computational complexity of typically $O(N \ln N)$ or $O(N)$ [5–8]. Given that T_s is comparable in magnitude to the total energy, the accuracy of OFDFT largely depends on the approximated form of the kinetic energy density functional (KEDF). However, developing an accurate KEDF has been a major hurdle in the field of OFDFT for decades.

Over the past few decades, continuous efforts have been devoted to developing analytical KEDFs [4,9]. In general, KEDFs can be classified into two categories. The first category comprises local and semilocal components in KEDFs, where the kinetic energy density is a function of the charge density, the charge density gradient, the Laplacian of charge density, or even higher-order derivatives of the charge density [10–15]. The second category consists of nonlocal forms of KEDFs, where the kinetic energy density is a functional of charge density, such that the kinetic energy density at each point in real space depends on the nonlocal charge density [16–20]. Typically, semilocal KEDFs are more computationally efficient, while nonlocal ones offer a higher accuracy. However, since most of the existing nonlocal KEDFs are constructed based on the Lindhard response function, which is accurate for nearly free electron gas, they are

mainly adequate for simple metals [16,17]. Some KEDFs were proposed to describe semiconductor systems, but they cannot work well for simple metals [18–20]. As a result, a KEDF that works for both simple metal and semiconductor systems is still lacking, and it is still unclear how to construct it systematically.

In recent years, machine learning (ML) techniques have been involved in the developments of computational physics [21]. In particular, the remarkable fitting ability of ML models has been demonstrated in various applications, including the fitting of potential energy surfaces in molecular dynamics [22,23], as well as fitting exchange-correlation functionals [24–27] and Hamiltonian matrices [28] within the framework of density functional theory (DFT) [1,2]. Additionally, there have been endeavors to construct ML-based KEDFs within the framework of OFDFT [29–35]. For example, Imoto *et al.* implemented a semilocal ML-based KEDF, which takes dimensionless gradient and dimensionless Laplacian of charge density as descriptors and puts the enhancement factor of kinetic energy density as the output of neural network (NN) [33]. This model exhibits convergence and satisfies the scaling law, but it overlooks nonlocal information crucial for improving the accuracy of KEDFs. Ryczko *et al.* implemented a nonlocal ML-based KEDF, utilizing a voxel deep NN, but this model could not achieve convergence in OFDFT computations [34]. Thus, it is still a formidable task to construct an accurate, transferable, and computationally stable ML-based KEDF.

In this work, as the first step to construct an ML-based KEDF that works for both simple metal and semiconductor systems, we construct an ML-based physical-constrained nonlocal KEDF (MPN KEDF) for simple metals and their alloys, which (i) contains nonlocal information, (ii) obeys a series of exact physical constraints, and (iii) achieves convergence via careful design of descriptors, NN output, postprocessing, and loss function, etc. The performance of the MPN KEDF is

*mohanchen@pku.edu.cn

systematically evaluated by testing on a series of simple metals, including lithium (Li), magnesium (Mg), aluminum (Al), and their alloys. In particular, incorporating nonlocal information and exact physical constraints is crucial to improving the accuracy, transferability, and stability of ML-based KEDFs [31].

The rest of this paper is organized as follows. In Sec. II, we propose an ML-based KEDF that satisfies physical constraints and introduces numerical details of KSDFT and OFDFT calculations. In Sec. III, we analyze the performances of the MPN KEDF and discuss the results. Finally, the conclusions are drawn in Sec. IV.

II. METHODS

A. Pauli energy and Pauli potential

In general, the noninteracting kinetic energy T_s can be divided into two parts [36],

$$T_s = T_{vW} + T_\theta, \quad (1)$$

where

$$T_{vW} = \frac{1}{8} \int \frac{|\nabla \rho(r)|^2}{\rho(r)} d^3r \quad (2)$$

is the von Weizsäcker (vW) KEDF, [12] a rigorous lower bound to the T_s , with $\rho(r)$ being the charge density. The second term T_θ represents the Pauli energy, which takes the form of

$$T_\theta = \int \tau_{TF} F_\theta d^3r, \quad (3)$$

where the Thomas-Fermi (TF) kinetic energy density [10,11] term is

$$\tau_{TF} = \frac{3}{10} (3\pi^2)^{2/3} \rho^{5/3}. \quad (4)$$

Additionally, F_θ denotes the enhancement factor. The corresponding Pauli potential is given by

$$V_\theta(r) = \delta E_\theta / \delta \rho(r). \quad (5)$$

The Pauli energy and Pauli potential satisfy several exact physical constraints. For example, first, the scaling law is

$$T_\theta[\rho_\lambda] = \lambda^2 T_\theta[\rho], \quad (6)$$

where $\rho_\lambda = \lambda^3 \rho(\lambda r)$ and λ is a positive number [36].

Second, in the free electron gas (FEG) limit, the TF KEDF is exact, and the vW part vanishes so that the enhancement factor in the FEG limit takes the form of

$$F_\theta(r)|_{\text{FEG}} = 1. \quad (7)$$

In addition, the Pauli potential returns to the potential of TF KEDF $V_{\text{TF}}(r)$

$$V_\theta(r)|_{\text{FEG}} = V_{\text{TF}}(r) = \frac{1}{2} (3\pi^2)^{2/3} \rho^{2/3}. \quad (8)$$

Third, the non-negativity ensures

$$F_\theta(r) \geq 0 \quad (9)$$

and

$$V_\theta(r) \geq 0. \quad (10)$$

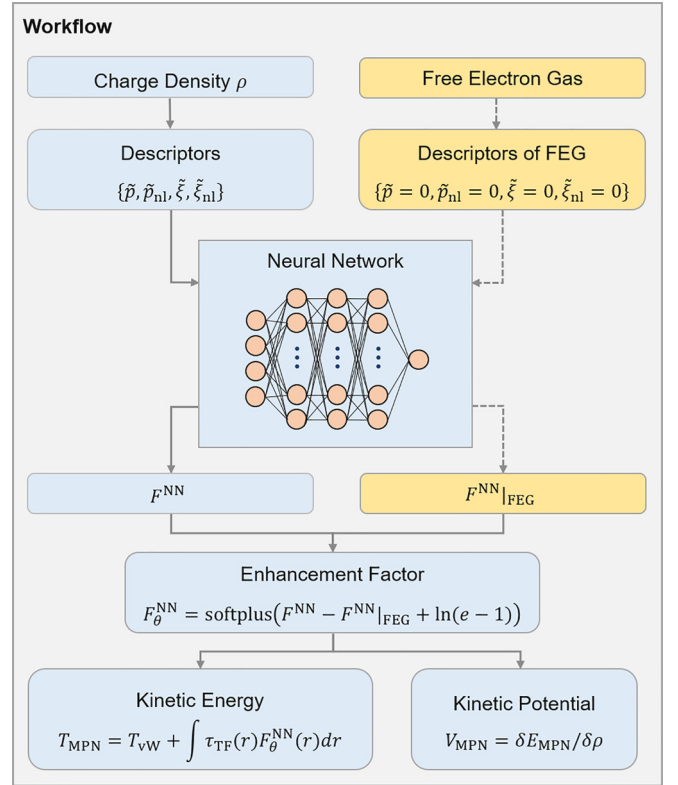


FIG. 1. Workflow of the MPN KEDF. $F^{\text{NN}}(r)$ is the enhancement factor obtained by the deep neural network (NN), and $F^{\text{NN}}|_{\text{FEG}}$ denotes the enhancement factor under the free electron gas (FEG) limit. In order to ensure both the FEG limit and the non-negativity of Pauli energy density are satisfied, the enhancement factor of Pauli energy is defined as $F_\theta^{\text{NN}} = \text{softplus}[F^{\text{NN}} - F^{\text{NN}}|_{\text{FEG}} + \ln(e-1)]$, where $\text{softplus}(x) = \ln(1 + e^x)$ is an activation function commonly used in machine learning with $\text{softplus}(x)|_{x=\ln(e-1)} = 1$. The defined formulas are used to evaluate the kinetic energy and kinetic potential.

In order to train the MPN KEDF, we collect the Pauli energy and Pauli potential data from KSDFT calculations performed on a set of selected systems. In detail, with the help of the Kohn-Sham orbitals and eigenvalues, in a spin degenerate system, the Pauli energy density can be analytically expressed by [36]

$$\tau_\theta^{\text{KS}} = \sum_{i=1}^M f_i |\nabla \psi_i(r)|^2 - \frac{|\nabla \rho|^2}{8\rho}, \quad (11)$$

while the Pauli potential has the form of

$$V_\theta^{\text{KS}} = \rho^{-1} \left(\tau_\theta^{\text{KS}} + 2 \sum_{i=1}^M f_i (\varepsilon_M - \varepsilon_i) \psi_i^* \psi_i \right), \quad (12)$$

where $\psi_i(r)$ denotes an occupied Kohn-Sham orbital with index i , while ε_i and f_i are the corresponding eigenvalue and occupied number, respectively. In addition, M represents the highest occupied state, and ε_M is the eigenvalue of $\psi_M(r)$, i.e., the chemical potential.

B. Design neural network based on exact physical constraints

The workflow of the MPN KEDF is summarized in Fig. 1. The major structure of the MPN KEDF is an NN composed of one input layer consisting of four nodes, three hidden layers with ten nodes in each layer, and an output layer with one node. The activation functions used in the hidden layers are chosen to be hyperbolic tangent functions, i.e., $\tanh(x)$. In order to ensure that the calculated Pauli energy and potential obey the physical constraints mentioned above, the output of the NN is chosen as the enhancement factor F_θ for each real-space grid point r , which is denoted as $F^{\text{NN}}(r)$. Next, we elucidate how nonlocal information and exact physical constraints can be incorporated into the NN to improve its accuracy and reliability.

As shown in Fig. 1, we define four descriptors $\{\tilde{p}, \tilde{p}_{\text{nl}}, \tilde{\xi}, \tilde{\xi}_{\text{nl}}\}$ (*vide infra*) as the input of the NN for the MPN KEDF. The first descriptor $\tilde{p}(r)$ is semilocal, while the other three are nonlocal. First, the semilocal descriptor is the normalized dimensionless gradient of the charge density given by

$$\tilde{p}(r) = \tanh(\chi_p p(r)), \quad (13)$$

where the parameter $p(r)$ is evaluated via

$$p(r) = |\nabla \rho(r)|^2 / [2(3\pi^2)^{1/3} \rho^{4/3}(r)]^2. \quad (14)$$

Here, χ_p is a hyperparameter to control the distribution of \tilde{p} .

Second, we propose a nonlocal descriptor of \tilde{p} , which is defined as

$$\tilde{p}_{\text{nl}}(r) = \int w(r-r') \tilde{p}(r') d^3 r', \quad (15)$$

where $w(r-r')$ is the kernel function similar to the Wang-Teter [16] kernel function, satisfying

$$\int w(r-r') d^3 r' = 0. \quad (16)$$

The kernel function is defined in reciprocal space as

$$w(\eta) = \left(\frac{1}{2} + \frac{1-\eta^2}{4\eta} \ln \left| \frac{1+\eta}{1-\eta} \right| \right)^{-1} - 3\eta^2 - 1. \quad (17)$$

Here $\eta = \frac{k}{2k_F}$ is a dimensionless reciprocal space vector, while $k_F = (3\pi^2 \rho_0)^{1/3}$ is the Fermi wave vector with ρ_0 being the average charge density.

The third and fourth nonlocal descriptors represent the distribution of charge density and take the form of

$$\tilde{\xi}(r) = \tanh \left(\frac{\int w(r-r') \rho^{1/3}(r') d^3 r'}{\rho^{1/3}(r)} \right), \quad (18)$$

and

$$\tilde{\xi}_{\text{nl}}(r) = \int w(r-r') \tilde{\xi}(r') d^3 r', \quad (19)$$

respectively.

In summary, the MPN KEDF is characterized by the above four descriptors: $\{\tilde{p}, \tilde{p}_{\text{nl}}, \tilde{\xi}, \tilde{\xi}_{\text{nl}}\}$, with $\chi_p = 0.2$ being an empirical parameter adopted in all calculations. Next, we propose three physical constraints that are met by our ML-based MPN KEDF.

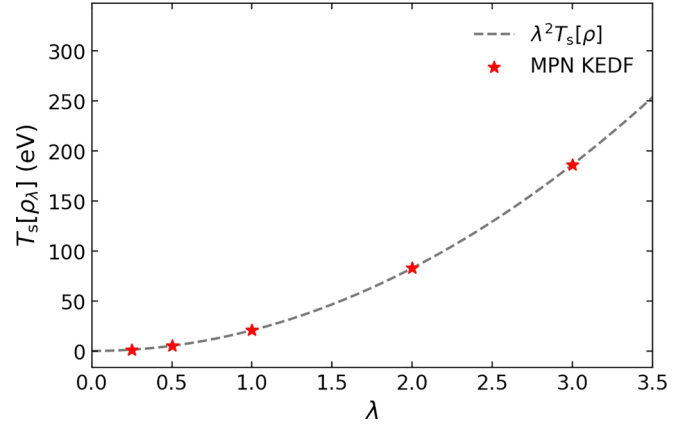


FIG. 2. Illustration of the scaling law introduced in Eq. (6). The gray line represents the function of $\lambda^2 T_s[\rho]$, where ρ denotes the ground charge density of face-centered cubic (fcc) Al as obtained by the MPN KEDF. The red stars denote the kinetic energies of $\rho_\lambda = \lambda^3 \rho(\lambda r)$ computed using the MPN KEDF for different values of λ , namely 0.25, 0.5, 1.0, 2.0, and 3.0. All the red stars fall on the gray line, indicating that the scaling law $T_s[\rho_\lambda] = \lambda^2 T_s[\rho]$ is exactly obeyed by the MPN KEDF.

First, the scaling law of noninteracting electron kinetic energy is ensured when we design the above descriptors. In detail, under the scaling translation $\rho(r) \rightarrow \rho_\lambda = \lambda^3 \rho(\lambda r)$, the descriptors $\{\tilde{p}(r), \tilde{p}_{\text{nl}}(r), \tilde{\xi}(r), \tilde{\xi}_{\text{nl}}(r)\}$ become $\{\tilde{p}(\lambda r), \tilde{p}_{\text{nl}}(\lambda r), \tilde{\xi}(\lambda r), \tilde{\xi}_{\text{nl}}(\lambda r)\}$, i.e., the descriptors are invariant under the scaling transformation, and the detailed derivation can be found in the Supplemental Material (SM) [37]. Since the T_{vW} term satisfies the scaling law, we have

$$\begin{aligned} T_{\text{MPN}}[\rho_\lambda] &= T_{\text{vW}}[\rho_\lambda] + \lambda^5 \int \tau_{\text{TF}}(\lambda r) \\ &\quad \times F_\theta^{\text{NN}}(\tilde{p}(\lambda r), \tilde{p}_{\text{nl}}(\lambda r), \tilde{\xi}(\lambda r), \tilde{\xi}_{\text{nl}}(\lambda r)) d^3 r \\ &= \lambda^2 \left[T_{\text{vW}}[\rho] + \int \tau_{\text{TF}}(\lambda r) \right. \\ &\quad \left. \times F_\theta^{\text{NN}}(\tilde{p}(\lambda r), \tilde{p}_{\text{nl}}(\lambda r), \tilde{\xi}(\lambda r), \tilde{\xi}_{\text{nl}}(\lambda r)) d^3(\lambda r) \right] \\ &= \lambda^2 T_{\text{MPN}}[\rho]. \end{aligned} \quad (20)$$

In order to verify the scaling law, we obtain the ground-state charge density $\rho(r)$ of fcc Al with the MPN KEDF, then the kinetic energy of $\rho_\lambda = \lambda^3 \rho(\lambda r)$ with various λ (0.25, 0.5, 1.0, 2.0, and 3.0) are calculated by the MPN KEDF. As displayed in Fig. 2, all of the $T_{\text{MPN}}[\rho_\lambda]$ s computed by the MPN KEDF fall on the line of $f(\lambda) = \lambda^2 T_s[\rho]$, demonstrating that the MPN KEDF obeys the scaling law.

The second and third constraints, i.e., the FEG limit and the non-negativity of Pauli energy density, are introduced through postprocessing of the deep neural network. In the FEG limit, all four descriptors become zero, and hence we define the output of the NN in this limit as $F^{\text{NN}}|_{\text{FEG}}$. In addition, the enhancement factor of Pauli energy is defined as

$$F_\theta^{\text{NN}} = \text{softplus}(F^{\text{NN}} - F^{\text{NN}}|_{\text{FEG}} + \ln(e-1)), \quad (21)$$

where F^{NN} is the output of NN, and

$$\text{softplus}(x) = \ln(1 + e^x) \quad (22)$$

is an activation function commonly used in machine learning, satisfying

$$\text{softplus}(x) \geq 0 \quad (23)$$

and

$$\text{softplus}(x)|_{x=\ln(e-1)} = 1. \quad (24)$$

By construction, the non-negativity constraint is satisfied as

$$F_{\theta}^{\text{NN}} \geq 0, \quad (25)$$

and in the FEG limit where the charge density is a constant, we have

$$\begin{aligned} F_{\theta}^{\text{NN}}|_{\text{FEG}} &= \text{softplus}(F^{\text{NN}}|_{\text{FEG}} - F^{\text{NN}}|_{\text{FEG}} + \ln(e-1)) \\ &= 1, \end{aligned} \quad (26)$$

thereby ensuring that the FEG limit is also exactly satisfied. We note that the selection of kernel function and descriptors guarantees that once the FEG limit of Pauli energy is met, the FEG limit of Pauli potential is automatically satisfied, as discussed in Sec. III of SM [37].

Figure 1 summarizes the workflow of the MPN KEDF, which involves the above-mentioned physical constraints. First, for each real-space grid point, the descriptors of charge density $\rho(r)$ ($\{\tilde{p}, \tilde{p}_{\text{nl}}, \tilde{\xi}, \tilde{\xi}_{\text{nl}}\}$) are entered into NN to get the corresponding enhancement factor $F^{\text{NN}}(r)$. Second, the descriptors of FEG ($\{\tilde{p} = 0, \tilde{p}_{\text{nl}} = 0, \tilde{\xi} = 0, \tilde{\xi}_{\text{nl}} = 0\}$) are fed into the NN, and the enhancement factor of FEG $F^{\text{NN}}|_{\text{FEG}}$ is obtained. Third, to ensure both the FEG limit and the non-negativity of Pauli energy density are satisfied, the enhancement factor of Pauli energy is defined as $F_{\theta}^{\text{NN}} = \text{softplus}(F^{\text{NN}} - F^{\text{NN}}|_{\text{FEG}} + \ln(e-1))$. Finally, the kinetic energy and kinetic potential are calculated by the MPN KEDF using the defined formulas.

C. Training details

Before training the MPN KEDF, the loss function is defined as

$$\begin{aligned} L &= \frac{1}{N} \sum_r \left[\left(\frac{F_{\theta}^{\text{NN}} - F_{\theta}^{\text{KS}}}{\bar{F}_{\theta}^{\text{KS}}} \right)^2 + \left(\frac{V_{\theta}^{\text{MPN}} - V_{\theta}^{\text{KS}}}{\bar{V}_{\theta}^{\text{KS}}} \right)^2 \right] \\ &+ [F^{\text{NN}}|_{\text{FEG}} - \ln(e-1)]^2. \end{aligned} \quad (27)$$

where N is the number of grid points, and $\bar{F}_{\theta}^{\text{KS}}$ ($\bar{V}_{\theta}^{\text{KS}}$) represents the mean of F_{θ}^{KS} (V_{θ}^{KS}). The first term helps NN to learn information from the Pauli energy, while the second term emphasizes the significance of reproducing the correct Pauli potential. We emphasize that fitting the Pauli potential is crucial in determining the optimization direction and step length during the OFDFT calculations, and V_{θ}^{MPN} can be obtained through the back propagation of NN, as derived in SM [37]. The last term is a penalty term to reduce the magnitude of the FEG correction, which improves the stability of the MPN KEDF.

The training set consists of eight metallic structures, namely bcc Li, fcc Mg, fcc Al, as well as five alloys: Li₃Mg

(mp-976254), LiMg (mp-1094889), Mg₃Al (mp-978271), β'' MgAl₃ [38], LiAl₃ (mp-10890), where the numbers in brackets are the Materials Project IDs [39]. We performed KSDFT calculations to obtain the ground charge density and calculate the corresponding descriptors. Additionally, the Pauli energy and potential are calculated using Eqs. (11) and (12), respectively. These calculations are performed on a $27 \times 27 \times 27$ grid, resulting in a total of 157 464 grid points in the training set.

D. Numerical details

We have employed the ABACUS 3.0.4 packages [40] to carry out OFDFT and KSDFT calculations, while for OFDFT with the Wang-Govind-Carter (WGC) KEDF [17], we have utilized the PROFESS 3.0 package [7]. The MPN KEDF is implemented in ABACUS using the LIBTORCH package [41], and the LIBNPY package is adopted to dump the data. Table S1 [37] lists the plane-wave energy cutoffs employed in both OFDFT and KSDFT calculations, as well as the Monkhorst-Pack k -point samplings [42] utilized in KSDFT. For both OFDFT and KSDFT calculations, we used the Perdew-Burke-Ernzerhof (PBE) [43] and bulk-derived local pseudopotentials (BLPS) [44]. Additionally, we used the Gaussian smearing method with a smearing width of 0.1 eV in our KSDFT calculations.

In order to calculate the ground-state bulk properties, we first optimize the crystal structures until the stress tensor elements are below 5×10^{-7} Hartree/Bohr³, then compress and expand the lattice constant of the unit cell from $0.99a_0$ to $1.01a_0$, where a_0 is the equilibrium lattice constant. Once the energy-volume curve is obtained, the bulk modulus B of a given system is fitted by Murnaghan's equation of state [45].

We compare the results obtained by the MPN KEDF to those obtained from OFDFT calculations with traditional KEDFs. Specifically, we have employed semilocal KEDFs such as the TF λ vW [46] and the Luo-Karasiev-Trickey (LKT) KEDFs [13], as well as the nonlocal ones including the Wang-Teter (WT) [16] and WGC KEDFs. The parameter λ of TF λ vW has been set to be 0.2, and the parameter a of the LKT KEDF is set to be 1.3, as in the original work [13]. In addition, we set $\alpha = \frac{5+\sqrt{5}}{6}$, $\beta = \frac{5-\sqrt{5}}{6}$ and $\gamma = 2.7$ in the WGC KEDF [17], as well as $\alpha = \frac{5}{6}$, $\beta = \frac{5}{6}$ in the WT KEDF [16].

The formation energy E_f of Li-Mg-Al alloy is defined as

$$E_f = \frac{1}{N} (E_{\text{total}} - n_{\text{Li}} E_{\text{Li}} - n_{\text{Mg}} E_{\text{Mg}} - n_{\text{Al}} E_{\text{Al}}), \quad (28)$$

where E_{total} is the total energy of the alloy, and E_{Li} , E_{Mg} , and E_{Al} denote the equilibrium energy of the bcc Li, hcp Mg, and fcc Al structures, respectively. Furthermore, n_{Li} , n_{Mg} , and n_{Al} depict the number of Li, Mg, and Al atoms, respectively. $N = n_{\text{Li}} + n_{\text{Mg}} + n_{\text{Al}}$ denotes the total number of atoms of the alloy.

The mean absolute relative error (MARE) and mean absolute error (MAE) of property x are respectively defined as

$$\text{MARE} = \frac{1}{N} \sum_i \left| \frac{x_i^{\text{OF}} - x_i^{\text{KS}}}{x_i^{\text{KS}}} \right|, \quad (29)$$

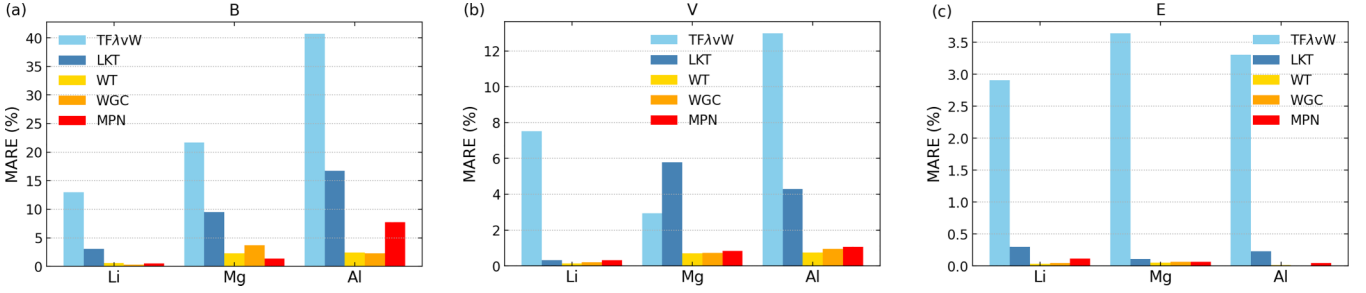


FIG. 3. MAREs of bulk properties of Li, Mg, and Al systems, i.e., (a) the bulk moduli (B in GPa), (b) the equilibrium volumes (V_0 in $\text{\AA}^3/\text{atom}$), and (c) the total energies of given systems (E_0 in eV/atom). The MARE defined in Eq. (29) is obtained by comparing OFDFT to KS-BLPS results. We use body-centered cubic (bcc), fcc, simple cubic (sc), and cubic diamond (CD) structures of Li. We also adopt hexagonal close-packed (hcp), fcc, bcc, and sc structures of Mg. For Al systems, we take fcc, hcp, bcc, and sc structures.

$$\text{MAE} = \frac{1}{N} \sum_i^N |x_i^{\text{OF}} - x_i^{\text{KS}}|. \quad (30)$$

Here N is the number of data points, x_i^{OF} and x_i^{KS} are properties obtained from OFDFT and KSDFT calculations, respectively.

III. RESULTS AND DISCUSSION

In order to examine the precision and transferability of the MPN KEDF, we prepared two testing sets. The first set comprises four structures of Li (bcc, fcc, sc, and CD), four structures of Mg (hcp, fcc, bcc, and sc), and four structures of Al (fcc, hcp, bcc, and sc). We evaluated the properties of these bulk systems, including the bulk moduli, the equilibrium volumes, and the equilibrium energies using various KEDFs. For the second testing set, we selected 59 alloys obtained from the Materials Project database [39], including 20 Li-Mg alloys, 20 Mg-Li alloys, 10 Li-Al alloys, and 9 Li-Mg-Al alloys, and the detailed information of these alloys are listed in Table S2 [37].

Notably, most of the structures in the two testing sets do not appear in the training set, allowing for an unbiased comparison. We systematically compared the total energies, the formation energies, and the charge densities of alloys within

the second testing set. We also trained another semilocal ML-based KEDF with descriptors as $\{\tilde{\rho}, \tilde{q}\}$ with $\tilde{q} = \tanh(0.1q)$, where $q = \nabla^2 \rho / [4(3\pi^2)^{2/3} \rho^{5/3}]$. However, we found this semilocal ML-based KEDF cannot achieve convergence in all tested systems.

A. Simple metals

Figure 3 displays the MAREs of bulk properties of Li, Mg, and Al systems. Compared to the nonlocal WT and WGC KEDFs, the semilocal KEDFs (the $\text{TF}\lambda\text{vW}$ and LKT KEDFs) yield larger MAREs across all the properties in all three systems, indicating that the nonlocal information is crucial to enhance the accuracy of KEDF. Comparatively, the MPN KEDF yields MAREs slightly larger than those of the WT and WGC KEDFs but does not exceed those of semilocal ones. Notably, the MPN KEDF achieves a lower MARE of 1.37% for the bulk modulus of Mg, outperforming WT and WGC KEDFs, which exhibit MAREs of 2.32% and 3.73%, respectively. On the other hand, the poorest results obtained by the MPN KEDF are the bulk modulus of Al, where it exhibits a MARE of 7.75%, nearly three times than those from the WT (2.41%) and WGC (2.26%) KEDFs. This may be caused by the fact that we did not include more Al structures with different densities in the training set. However, even in this case, the MAREs obtained by the $\text{TF}\lambda\text{vW}$ (40.72%) and LKT (16.69%) KEDFs are almost five and two times higher

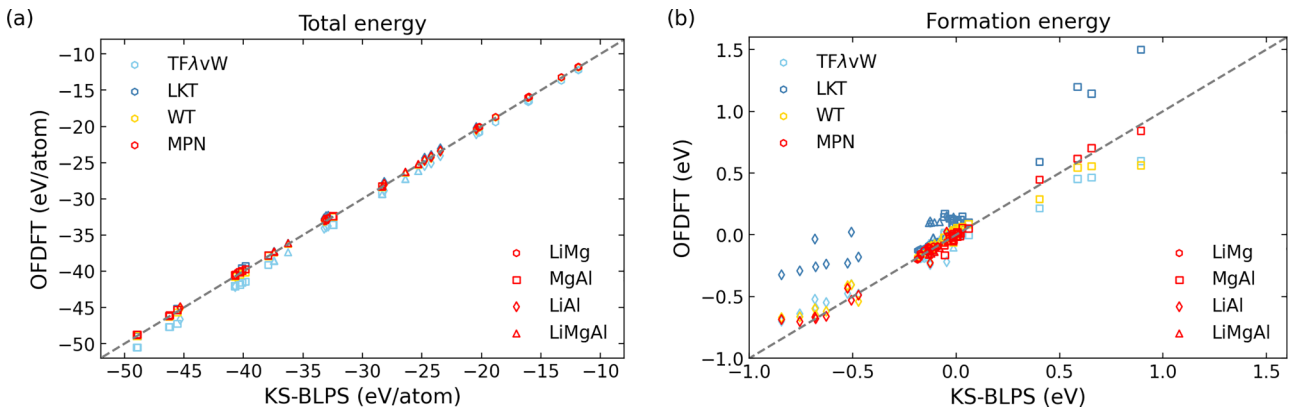


FIG. 4. (a) Total energies (in eV/atom) and (b) formation energies (in eV) of 59 alloys, including 20 Li-Mg alloys, 20 Mg-Li alloys, 10 Li-Al alloys, and 9 Li-Mg-Al alloys. Different colors indicate the formation energies from different KEDFs ($\text{TF}\lambda\text{vW}$, LKT, WT, and MPN), while different shapes of markers indicate different alloys.

TABLE I. MAEs [Eq. (30)] of the total energies and formation energies of 59 alloys obtained by comparing various KEDFs (TF λ vW, LKT, WT, and MPN) in OFDFT to KS-BLPS results. The systems include 20 Li-Mg alloys, 20 Mg-Li alloys, 10 Li-Al alloys, and 9 Li-Mg-Al alloys.

MAE of total energy (eV/atom)					
	LiMg	MgAl	LiAl	LiMgAl	Total
TF λ vW	0.540	1.330	0.873	0.999	0.934
LKT	0.040	0.156	0.351	0.124	0.145
WT	0.013	0.059	0.082	0.031	0.043
MPN	0.078	0.163	0.146	0.106	0.123
MAE of formation energy (eV)					
	LiMg	MgAl	LiAl	LiMgAl	Total
TF λ vW	0.022	0.061	0.103	0.036	0.051
LKT	0.041	0.189	0.397	0.138	0.166
WT	0.005	0.050	0.077	0.023	0.035
MPN	0.015	0.027	0.056	0.024	0.028

than that of the MPN KEDF. As a result, we conclude that the MPN KEDF yields reasonable accuracy when compared to other nonlocal KEDFs.

It is noteworthy that the energy difference between the fcc and hcp structures of bulk Al is small, which is 0.025 eV/atom as predicted by KSDFT, and it is sensitive to the accuracy of KEDF [47]. Both TF λ vW and LKT KEDFs, as semilocal KEDFs, fail to distinguish this subtle energy difference and predict it as 0.000 eV/atom. In contrast, the nonlocal WT and WGC KEDFs yield nonzero energy differences of 0.018 and 0.016 eV/atom, respectively. Moreover, the MPN KEDF predicts the energy difference to be 0.021 eV/atom, which is close to the result of 0.025 eV/atom obtained by KSDFT and is more accurate than the WT and WGC KEDFs. This result emphasizes the importance of involving the nonlocal information again, which enables the MPN KEDF to distinguish the subtle difference between similar crystal structures [47].

B. Alloys

Figure 4 illustrates the total energies and the formation energies of 59 alloys obtained by different KEDFs in OFDFT calculations, and their corresponding MAEs are listed in Table I. Notably, the WGC KEDF failed to achieve convergence for nine alloys; therefore, we have excluded the WGC results from Table I. Regarding the prediction of total energy shown in Fig. 4(a), the TF λ vW KEDF consistently underestimates the values compared to those obtained by KSDFT,

resulting in a large MAE of 0.934 eV/atom. In contrast, the LKT KEDF performs better with a reduced MAE of 0.145 eV/atom, while the nonlocal WT KEDF yields a lower MAE of 0.043 eV/atom. The MPN KEDF yields a higher MAE (0.123 eV/atom) than the WT KEDF but still outperforms the TF λ vW and LKT KEDFs. As for the formation energies shown in Fig. 4(b), we observe that the LKT KEDF consistently yields larger values compared to KSDFT, resulting in a high MAE of 0.166 eV, which is much larger than the MAEs obtained by the TF λ vW KEDF (0.051 eV) and WT KEDFs (0.035 eV). Remarkably, the MPN KEDF exhibits an even lower MAE (0.028 eV) than the WT KEDF. Overall, these results demonstrate the promising potential of the MPN KEDF in predicting the energies of complex alloy systems with a high accuracy.

In order to further evaluate the accuracy of the MPN KEDF, we computed the charge densities of 59 alloys and calculated the mean MAREs, listed in Table II. As expected, the semilocal TF λ vW and LKT KEDFs failed to reproduce the charge density obtained by KSDFT, exhibiting mean MAREs of 14.30% and 7.34%, respectively. These MAREs are considerably higher than the mean MARE obtained by the nonlocal WT KEDF (2.38%). Impressively, the MPN KEDF yields a mean MARE of 3.30%, which is slightly higher than that of the WT KEDF but significantly lower than those of the TF λ vW and LKT KEDFs. We note that the above 59 alloys are not present in the training set, and there are even no Li-Mg-Al alloys in the training set, so the above results not only indicate a high accuracy but also excellent transferability of the MPN KEDF.

Figure 5 shows the charge densities of four typical structures, one taken from the training set and the other three from the testing set. The first structure is Li₃Mg (mp-976254) from the training set, containing four atoms. The MPN KEDF yields a MARE of 2.73%, which is slightly larger than that obtained by the WT KEDF (1.03%) but significantly lower than those obtained by the TF λ vW (13.85%) and LKT KEDFs (5.28%), demonstrating the efficiency of the training process.

The second structure is Li(Mg₄Al₃)₄ (mp-1185175) with 87 atoms, which is the largest system among the testing set. Notably, the MPN KEDF achieves convergence to yield a smooth ground-state density, which is close to the result obtained by KSDFT, indicating an excellent stability in optimizing the electron charge density. In contrast, the WGC KEDF fails to reach convergence for this structure.

The last two crystal structures are Mg₃Al (mp-1094666, 16 atoms) and LiAl (mp-1191737, 48 atoms) from the testing set, for which the MPN KEDF yields the lowest MARE and largest MARE among the testing set, respectively. For the

TABLE II. Mean MAREs [Eq. (29)] of charge densities of 59 alloys, including 20 Li-Mg alloys, 20 Mg-Li alloys, 10 Li-Al alloys, and 9 Li-Mg-Al alloys. MAREs are obtained by comparing various KEDFs (TF λ vW, LKT, WT, and MPN) in OFDFT to KS-BLPS results.

mean MARE of charge density (%)	LiMg	MgAl	LiAl	LiMgAl	Total
TF λ vW	12.40	16.11	13.26	15.66	14.30
LKT	5.26	7.44	11.61	6.98	7.34
WT	1.06	2.57	4.98	2.04	2.38
MPN	2.41	3.12	5.81	2.89	3.30

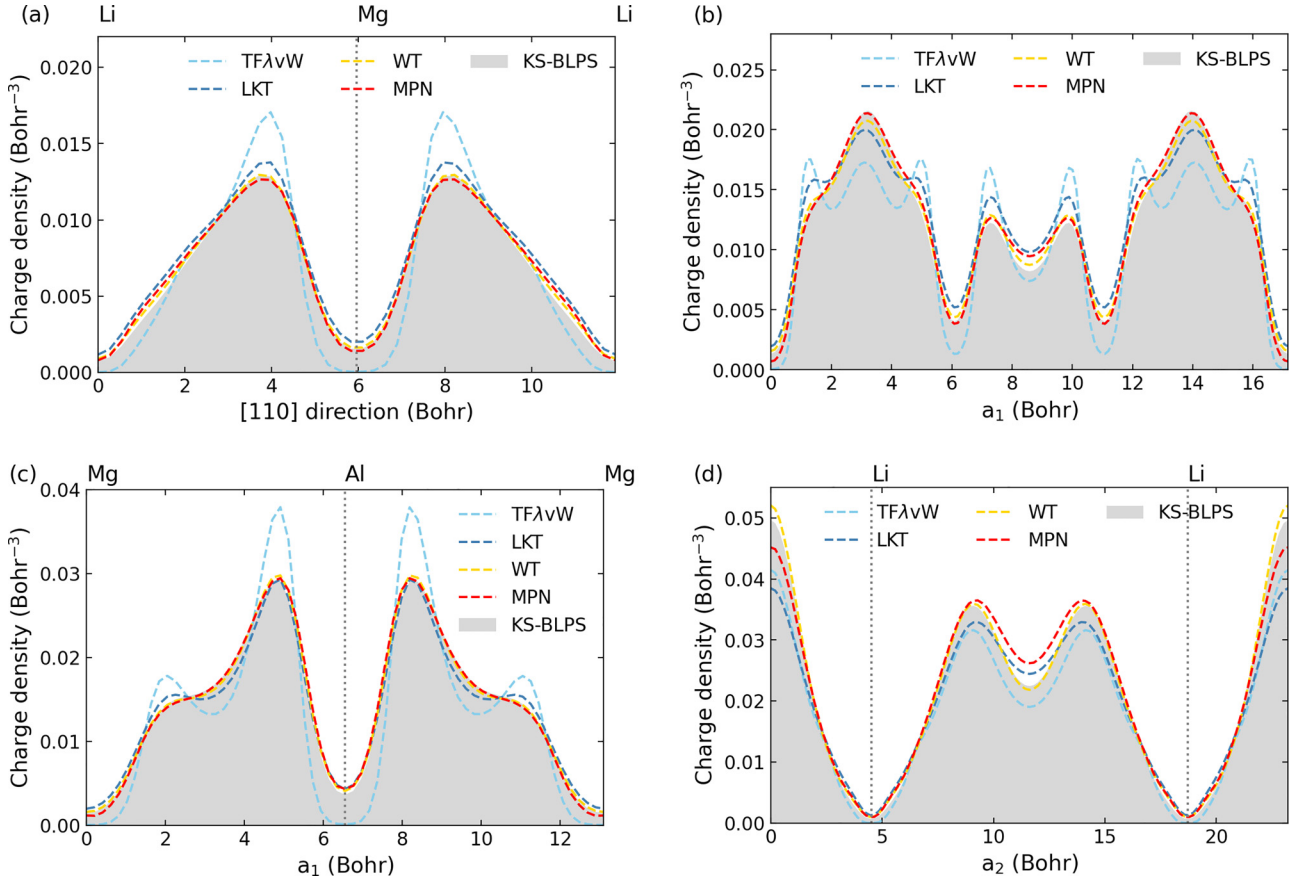


FIG. 5. Charge densities of four typical alloys. (a) Li_3Mg (mp-976254, 4 atoms) from the training set. The MARE of charge density obtained by the $\text{TF}\lambda\text{vW}$, LKT, WT, and MPN KEDFs are 13.85%, 5.28%, 1.03%, and 2.73%, respectively. (b) $\text{Li}(\text{Mg}_4\text{Al}_3)_4$ (mp-1185175), the largest system in the testing set, containing 87 atoms. The MARE of charge density obtained by the $\text{TF}\lambda\text{vW}$, LKT, WT, and MPN KEDFs are 16.15%, 7.76%, 2.54%, and 3.08%, respectively. (c) Mg_3Al (mp-1094666, 16 atoms) from the testing set, in which the MPN KEDF yields the lowest MARE among the testing set. The MARE of charge density obtained by the $\text{TF}\lambda\text{vW}$, LKT, WT, and MPN KEDFs are 16.73%, 6.42%, 1.65%, and 1.57%, respectively. (d) LiAl (mp-1191737, 48 atoms) from the testing set, in which the MPN KEDF obtains the largest MARE among the testing set. The MARE of charge density obtained by the $\text{TF}\lambda\text{vW}$, LKT, WT, and MPN KEDFs are 13.51%, 15.41%, 6.98%, and 8.16%, respectively. The labels a_1 and a_2 denote the first and second lattice vectors of the corresponding structures, respectively.

Mg_3Al structure, the MPN KEDF exhibits a better accuracy than the WT KEDF, yielding a MARE of 1.57%, lower than the 1.65% obtained by the WT KEDF. For the LiAl structure, although the MPN KEDF yields the largest MARE of 8.16%, it is still much lower than those obtained by the $\text{TF}\lambda\text{vW}$ (13.51%) and LKT KEDFs (15.41%). Overall, the MPN KEDF outperforms the semilocal KEDFs in terms of accuracy and achieves comparable accuracy to the other nonlocal KEDFs. Additionally, the stability of the MPN KEDF is evidenced by reaching convergence and obtaining smooth charge densities for all alloys in the testing set.

What is more, in order to further test the transferability and stability of the MPN KEDF, we have generated 45 hypothetical Mg-Al alloys and calculated the total energies and formation energies of these alloys. Similar to the phenomenon described above, the MPN KEDF gives worse total energies than the WT KEDF, but it yields substantially more accurate formation energies than the WT KEDF, and detailed results can be found in SM [37]. In conclusion, the above results imply the good transferability of the MPN KEDF for Mg-Al alloys.

IV. CONCLUSIONS

In this work, based on the framework of deep neural networks, we proposed an ML-based physical-constrained nonlocal (MPN) KEDF. Our proposed method relied on four descriptors, i.e., $\{\tilde{p}, \tilde{p}_{\text{nl}}, \tilde{\xi}, \tilde{\xi}_{\text{nl}}\}$, in which \tilde{p} was a semilocal descriptor, and the other three captured the nonlocal information of charge density. Importantly, the MPN KEDF was subject to three crucial physical constraints, including the scaling law of Eq. (6), the FEG limit shown in Eq. (7) and the non-negativity of Pauli energy density. The MPN KEDF was implemented in the ABACUS package [40].

We systematically evaluated the performance of various KEDFs on simple metals, including bulk Li, Mg, and Al, by calculating their bulk properties, i.e., the bulk moduli, the equilibrium volumes, and the equilibrium energies. Additionally, we tested 59 alloys consisting of 20 Li-Mg alloys, 20 Mg-Li alloys, 10 Li-Al alloys, and 9 Li-Mg-Al alloys. Overall, our results demonstrated that the MPN KEDF exceeded the accuracy of semilocal KEDFs and approached the accuracy of nonlocal KEDFs for all of the tested systems. Additionally, the

proposed MPN KEDF exhibited satisfactory transferability and stability during density optimization. In the future, our proposed approach sheds new light on generating KEDFs for semiconductors or molecular systems, and may also serve as a reference for developing ML-based exchange-correlation functionals.

ACKNOWLEDGMENTS

The work of L.S. and M.C. was supported by the National Science Foundation of China under Grants No. 12074007 and No. 12122401. The numerical simulations were performed on the High-Performance Computing Platform of CAPT.

- [1] P. Hohenberg and W. Kohn, Inhomogeneous electron gas, *Phys. Rev.* **136**, B864 (1964).
- [2] W. Kohn and L. J. Sham, Thermal properties of the inhomogeneous electron gas, *Phys. Rev.* **140**, A1133 (1965).
- [3] Y. A. Wang and E. A. Carter, Orbital-free kinetic-energy density functional theory, *Theoretical Methods in Condensed Phase Chemistry* (Springer Netherlands, Dordrecht, 2002), pp. 117–184.
- [4] W. C. Witt, G. Beatriz, J. M. Dieterich, and E. A. Carter, Orbital-free density functional theory for materials research, *J. Mater. Res.* **33**, 777 (2018).
- [5] G. S. Ho, V. L. Lignères, and E. A. Carter, Introducing profess: A new program for orbital-free density functional theory calculations, *Comput. Phys. Commun.* **179**, 839 (2008).
- [6] L. Hung, C. Huang, I. Shin, G. S. Ho, V. L. Lignères, and E. A. Carter, Introducing PROFESS 2.0: A parallelized, fully linear scaling program for orbital-free density functional theory calculations, *Comput. Phys. Commun.* **181**, 2208 (2010).
- [7] M. Chen, J. Xia, C. Huang, J. M. Dieterich, L. Hung, I. Shin, and E. A. Carter, Introducing PROFESS 3.0: An advanced program for orbital-free density functional theory molecular dynamics simulations, *Comput. Phys. Commun.* **190**, 228 (2015).
- [8] W. Mi, X. Shao, C. Su, Y. Zhou, S. Zhang, Q. Li, H. Wang, L. Zhang, M. Miao, Y. Wang *et al.*, ATLAS: A real-space finite-difference implementation of orbital-free density functional theory, *Comput. Phys. Commun.* **200**, 87 (2016).
- [9] V. V. Karasiev and S. B. Trickey, Issues and challenges in orbital-free density functional calculations, *Comput. Phys. Commun.* **183**, 2519 (2012).
- [10] L. H. Thomas, The calculation of atomic fields, in *Mathematical proceedings of the Cambridge philosophical society*, Vol. 23 (Cambridge University Press, Cambridge, 1927), pp. 542–548.
- [11] E. Fermi, Statistical method to determine some properties of atoms, *Rend. Accad. Naz. Lincei* **6**, 602 (1927).
- [12] C. F. v. Weizsäcker, Zur theorie der kernmassen, *Z. Phys.* **96**, 431 (1935).
- [13] K. Luo, V. V. Karasiev, and S. B. Trickey, A simple generalized gradient approximation for the noninteracting kinetic energy density functional, *Phys. Rev. B* **98**, 041111(R) (2018).
- [14] L. A. Constantin, E. Fabiano, and F. Della Sala, Semilocal pauli–gaussian kinetic functionals for orbital-free density functional theory calculations of solids, *J. Phys. Chem. Lett.* **9**, 4385 (2018).
- [15] D. Kang, K. Luo, K. Runge, and S. Trickey, Two-temperature warm dense hydrogen as a test of quantum protons driven by orbital-free density functional theory electronic forces, *Matter Radiat. Extremes* **5**, 064403 (2020).
- [16] L.-W. Wang and M. P. Teter, Kinetic-energy functional of the electron density, *Phys. Rev. B* **45**, 13196 (1992).
- [17] Y. A. Wang, N. Govind, and E. A. Carter, Orbital-free kinetic-energy density functionals with a density-dependent kernel, *Phys. Rev. B* **60**, 16350 (1999).
- [18] C. Huang and E. A. Carter, Nonlocal orbital-free kinetic energy density functional for semiconductors, *Phys. Rev. B* **81**, 045206 (2010).
- [19] W. Mi, A. Genova, and M. Pavanello, Nonlocal kinetic energy functionals by functional integration, *J. Chem. Phys.* **148**, 184107 (2018).
- [20] X. Shao, W. Mi, and M. Pavanello, Revised huang-carter non-local kinetic energy functional for semiconductors and their surfaces, *Phys. Rev. B* **104**, 045118 (2021).
- [21] B. Huang, G. F. von Rudorff, and O. A. von Lilienfeld, The central role of density functional theory in the AI age, *Science* **381**, 170 (2023).
- [22] L. Zhang, J. Han, H. Wang, R. Car, and E. Weinan, Deep potential molecular dynamics: A scalable model with the accuracy of quantum mechanics, *Phys. Rev. Lett.* **120**, 143001 (2018).
- [23] Y. Zhang, H. Wang, W. Chen, J. Zeng, L. Zhang, H. Wang, and E. Weinan, DP-GEN: A concurrent learning platform for the generation of reliable deep learning based potential energy models, *Comput. Phys. Commun.* **253**, 107206 (2020).
- [24] Y. Chen, L. Zhang, H. Wang, and W. E. DeePKS: A comprehensive data-driven approach toward chemically accurate density functional theory, *J. Chem. Theory Comput.* **17**, 170 (2021).
- [25] M. F. Kasim and S. M. Vinko, Learning the exchange-correlation functional from nature with fully differentiable density functional theory, *Phys. Rev. Lett.* **127**, 126403 (2021).
- [26] J. Kirkpatrick, B. McMorro, D. H. Turban, A. L. Gaunt, J. S. Spencer, A. G. D. G. Matthews, A. Obika, L. Thiry, M. Fortunato, D. Pfau *et al.*, Pushing the frontiers of density functionals by solving the fractional electron problem, *Science* **374**, 1385 (2021).
- [27] R. Nagai, R. Akashi, and O. Sugino, Machine-learning-based exchange correlation functional with physical asymptotic constraints, *Phys. Rev. Res.* **4**, 013106 (2022).
- [28] H. Li, Z. Wang, N. Zou, M. Ye, R. Xu, X. Gong, W. Duan, and Y. Xu, Deep-learning density functional theory hamiltonian for efficient ab initio electronic-structure calculation, *Nature Comput. Sci.* **2**, 367 (2022).
- [29] J. C. Snyder, M. Rupp, K. Hansen, K.-R. Müller, and K. Burke, Finding density functionals with machine learning, *Phys. Rev. Lett.* **108**, 253002 (2012).
- [30] J. Seino, R. Kageyama, M. Fujinami, Y. Ikabata, and H. Nakai, Semi-local machine-learned kinetic energy density functional with third-order gradients of electron density, *J. Chem. Phys.* **148**, 241705 (2018).
- [31] J. Hollingsworth, L. Li, T. E. Baker, and K. Burke, Can exact conditions improve machine-learned density functionals? *J. Chem. Phys.* **148**, 241743 (2018).

- [32] R. Meyer, M. Weichselbaum, and A. W. Hauser, Machine learning approaches toward orbital-free density functional theory: Simultaneous training on the kinetic energy density functional and its functional derivative, *J. Chem. Theory Comput.* **16**, 5685 (2020).
- [33] F. Imoto, M. Imada, and A. Oshiyama, Order- N orbital-free density-functional calculations with machine learning of functional derivatives for semiconductors and metals, *Phys. Rev. Res.* **3**, 033198 (2021).
- [34] K. Ryczko, S. J. Wetzel, R. G. Melko, and I. Tamblin, Toward orbital-free density functional theory with small data sets and deep learning, *J. Chem. Theory Comput.* **18**, 1122 (2022).
- [35] P. d. Mazo-Sevillano and J. Hermann, Variational principle to regularize machine-learned density functionals: The non-interacting kinetic-energy functional, *J. Chem. Theory Comput.* **159**, 194107 (2023).
- [36] M. Levy and H. Ou-Yang, Exact properties of the Pauli potential for the square root of the electron density and the kinetic energy functional, *Phys. Rev. A* **38**, 625 (1988).
- [37] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/PhysRevB.109.115135> for the detailed parameters used in OFDFT and KSDFT calculations, the formula of Pauli potential of the MPN KEDF, the derivation of the FEG limit of Pauli potential and the scaling invariance of descriptors, the details of alloy testing set, and the test result on random Mg-Al alloys.
- [38] K. M. Carling and E. A. Carter, Orbital-free density functional theory calculations of the properties of Al, Mg and Al-Mg crystalline phases, *Model. Simul. Mater. Sci. Eng.* **11**, 339 (2003).
- [39] A. Jain, S. P. Ong, G. Hautier, W. Chen, W. D. Richards, S. Dacek, S. Cholia, D. Gunter, D. Skinner, G. Ceder *et al.*, Commentary: The materials project: A materials genome approach to accelerating materials innovation, *APL Mater.* **1**, 011002 (2013).
- [40] P. Li, X. Liu, M. Chen, P. Lin, X. Ren, L. Lin, C. Yang, and L. He, Large-scale *ab initio* simulations based on systematically improvable atomic basis, *Comput. Mater. Sci.* **112**, 503 (2016).
- [41] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimeshin, L. Antiga *et al.*, PyTorch: An Imperative Style, High-Performance Deep Learning Library, *Advances in Neural Information Processing Systems*, edited by H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, and R. Garnett, Vol. 32 (Curran Associates, Inc., 2019).
- [42] H. J. Monkhorst and J. D. Pack, Special points for brillouin-zone integrations, *Phys. Rev. B* **13**, 5188 (1976).
- [43] J. P. Perdew, K. Burke, and M. Ernzerhof, Generalized gradient approximation made simple, *Phys. Rev. Lett.* **77**, 3865 (1996).
- [44] C. Huang and E. A. Carter, Transferable local pseudopotentials for magnesium, aluminum and silicon, *Phys. Chem. Chem. Phys.* **10**, 7109 (2008).
- [45] F. Murnaghan, The compressibility of media under extreme pressures, *Proc. Natl. Acad. Sci. USA* **30**, 244 (1944).
- [46] A. Berk, Lower-bound energy functionals and their application to diatomic systems, *Phys. Rev. A* **28**, 1908 (1983).
- [47] L. Sun, Y. Li, and M. Chen, Truncated nonlocal kinetic energy density functionals for simple metals and silicon, *Phys. Rev. B* **108**, 075158 (2023).