npj | computational materials

Article

Published in partnership with the Shanghai Institute of Ceramics of the Chinese Academy of Sciences

https://doi.org/10.1038/s41524-025-01643-0

Orbital-free density functionals based on real and reciprocal space separation

Bishal Thapa[1,2], Tracey G. Oellerich[3], Maria Emelianenko[3], Phanish Suryanarayana[4] & Igor I. Mazin[1,2]

We introduce a general class of orbital-free density functionals (OF-DFT) decomposed into a local part in coordinate space and a local part in reciprocal space. As a demonstration of principle, we choose for the former the Thomas-Fermi-von Weizsäcker (TFW) kinetic energy density functional (KEDF) and for the latter a form derived from the Lindhard function, but with the two system-dependent adjustable parameters. These parameters are machine-learned from Kohn-Sham data using Bayesian linear regression with a kernel method, which employs moments of the Fourier components of the electronic density as the descriptor. Through a number of representative cases, we demonstrate that our machine-learned model provides more than an order-of-magnitude improvement in the accuracy of the frozen-phonon energies compared to the TFW KEDF, with negligible increase in the computational cost. Overall, this work opens an avenue for the construction of accurate KEDFs for OF-DFT.

The foundation of modern first-principles computational materials science is the Hohenberg-Kohn (HK)[1] density functional theory (DFT), wherein the electron density replaces the many-body wavefunction as the fundamental quantity of interest. In particular, the HK theorem states that there exists a universal functional of the electron density $n(\mathbf{r})$, the minimization of which results in the exact ground-state density and energy, i.e., corresponding to the many-body interacting system. While the exact form of the HK functional is unknown, it can be assumed to be partitioned as: $E[n] = T[n] + W[n]$, where T is the electronic kinetic energy, and W is the energy arising from electron-nuclei and electron-electron interactions. Though approximations for the interaction energy functional are well-established[1–3], the same cannot be said for the kinetic energy functional.

This problem can be circumvented by the formally exact Kohn-Sham formalism[4], wherein the real system of interacting electrons is replaced by a fictitious system of non-interacting fermions that generate the same density. In particular, the kinetic energy of the non-interacting fermions T_s is written exactly in terms of the single-particle orbitals, the electron-electron interactions are modeled within the Hartree approximation, and the many-body effects are collected in the so-called exchange-correlation functional, for which a number of approximations at various levels of complexity have been developed[5]. This formalism has found tremendous success over the past few decades, firmly establishing itself as one of the cornerstones of materials science research. However, even with the development of highly efficient and scalable implementations[6], the range of systems that is accessible to a rigorous Kohn-Sham investigation is still limited. This is in significant part due to the cubic scaling with the number of atoms, which arises from the orthogonality constraint on the Kohn-Sham orbitals.

An alternative to the Kohn-Sham formalism, while still in the spirit of HK DFT, is an orbital-free DFT (OF-DFT), where the real system of interacting electrons is replaced with a fictitious system of non-interacting bosons. Since OF-DFT requires the calculation of only a single orbital, which corresponds to the square root of the electron density, it scales linearly with system size, allowing access, compared to Kohn-Sham DFT[7], to systems that are significantly larger. Given the success of Kohn-Sham DFT, it is common to maintain the same form in OF-DFT, while replacing the orbital-dependent electronic kinetic energy with a functional of the density, referred to as kinetic energy density functionals (KEDF). However, the KEDFs that have been developed for OF-DFT, including those discussed below, suffer from limited accuracy and transferability, which has limited the use of OF-DFT in practice. Note that an alternative strategy to developing KEDFs is to instead use the electron density from OF-DFT, which is found to be *close* to that from Kohn-Sham DFT, within the non self-consistent Kohn-Sham formalism for evaluating the energy[8–12]. Though this strategy is found to significantly improve the accuracy, and lower computational cost relative to Kohn-Sham DFT, it still retains the cubic scaling complexity of the Kohn-Sham formalism.

The earliest OF-DFT KEDFs are the Thomas-Fermi model and the von Weizsäcker model[2,5,13], the latter representing a lower bound on T_s. Kirzhnits proposed an interesting strategy to enhance the effectiveness of the KEDF by combining the TF model with the W model, the latter scaled by a parameter λ. Here, $\lambda = 1$ provides the exact KEDF for rapidly varying small density perturbation (large wave-vector $\mathbf{q}$), while $\frac{1}{9}$ provides the exact KEDF for slowly varying densities (i.e., small wave-vector $\mathbf{q}$)[14–16]. In addition to other choices of λ[3], the parameter can also be made to vary spatially[17].

[1]Department of Physics and Astronomy, George Mason University, Fairfax, VA, 22030, USA. [2]Quantum Science and Engineering Center, George Mason University, Fairfax, VA, 22030, USA. [3]Department of Mathematical Sciences, George Mason University, Fairfax, VA, 22030, USA. [4]College of Engineering, Georgia Institute of Technology, Atlanta, GA, 30332, USA. ✉e-mail: imazin2@gmu.edu

Given the limited accuracy of these KEDFs, a number of semilocal functionals have been developed that still satisfy the constraint arising from the aforementioned lower bound[18–23]. Although these functionals demonstrate improved accuracy, they are semi-empirical and are not able to reproduce the Lindhard response function in the homogeneous electron gas limit[24]. This limitation has driven the development of KEDFs that are nonlocal in coordinate space, i.e., depend on the density correlation at finite distances, which also reproduce the uniform electron gas limit[15,16,25,25–39]. Though such functionals have found success in particular applications[25,32–38], they still have found limited use in practice due to significantly larger computational expense, while still needing specialized kernels to be developed for different materials systems[25,34,36,37,39]. This has motivated the recent development of a number of machine-learned KEDFs[40–48]. However, the models so developed lack accuracy and transferability, have been mainly tested on model systems. Moreover, they are extremely complex and therefore do not provide the physical insight necessary for future developments.

In this work, we discuss a particular class of KEDFs for OF-DFT: those that can be decomposed into a local part in coordinate space and a local part in reciprocal space. Some of the proposals discussed above can be shown to belong to this class. As a demonstration of principle, we choose for the former the TFW KEDF and try for the latter a functional form derived from the Lindhard response function, but with the two adjustable parameters made system-dependent. These parameters are machine-learned from Kohn-Sham data using Bayesian linear regression in conjunction with the kernel method, while employing moments of the Fourier components of the electronic density as the descriptor. Through representative cases of metals and binary alloys, we demonstrate that when the part of the functional that is local in reciprocal space is applied non-self-consistently for frozen phonon calculations, there is more than an order-of-magnitude improvement in the accuracy of the energy relative to the TFW KEDF, with negligible increase in the cost.

Results

Formalism

The KEDFs can typically be decomposed into local and non-local components in **r**-space, such that:

$$\begin{aligned} T_s[n] &= T_{\mathrm{TF}\lambda\mathrm{W}}[n, \nabla n] + T_{\mathrm{NL}}[n] \\ &= T_{\mathrm{TF}}[n] + \lambda \cdot T_{\mathrm{W}}[n, \nabla n] + T_{\mathrm{NL}}[n] \end{aligned} \tag{1}$$

where

$$\begin{aligned} T_{\mathrm{TF}\lambda\mathrm{W}} = &\int \left(\frac{3}{10}(3\pi^2)^{\frac{2}{3}} n(\mathbf{r})^{\frac{5}{3}}\right) d\mathbf{r} \\ &+ \int \left(\frac{\lambda}{8}\frac{|\nabla n(\mathbf{r})|^2}{n(\mathbf{r})}\right) d\mathbf{r}. \end{aligned}$$

The choice of λ has been a point of contention for years. While the original choice of $\lambda = 1$[49,50] represents a lower bound on T_s and is correct in the limit of small and rapid (large wavevector) density variations, Kirzhnits' version, $\lambda = 1/9$, provides the exact response in the limit of slow but not necessarily small variations[14]. Regardless of the value of λ, the functional TFλW generally provides a poor representation of the linear response function and, consequently, of the KEDF[3,17,51]. This shortcoming necessitates development of more accurate functionals with an improved response function, motivating the incorporation of non-local KEDFs[8–12]. From a broader perspective, these nonlocal KEDFs are density-driven and employ specific forms, often constrained by system-dependent physical considerations, to develop density-weighted kernels, $\mathcal{K}[n](\mathbf{r}, \mathbf{r}')$[7]. The general formulation of nonlocal KEDFs in terms of a kernel $\mathcal{K}[n](\mathbf{r}, \mathbf{r}')$ can be expressed as:

$$T_{\mathrm{NL}} = \int d\mathbf{r}d\mathbf{r}' f(n(\mathbf{r}))\mathcal{K}[n](\mathbf{r}, \mathbf{r}')f(n(\mathbf{r}')), \tag{2}$$

where the function $f(n)$ can take the various forms[7,15,25,32] etc. While showing good results for particular problems, such nonlocal KEDF have found rather limited use in practice due to greater computational expense and the need for specialized kernels to be developed for different materials systems[7]. Our previous study on this issue suggests that the primary challenge does not stem from the self-consistent density distribution resulting from orbital-free functional, but rather from the error of calculating the total energy from a given density distribution using local orbital-free functionals, that is, from poor description of the charge susceptibility[8]. Since the most common application of DFT is on calculated energy differences upon the external perturbation, we first propose to seek for a density functional in terms of the density difference, i.e.,

$$\Delta T_{\mathrm{NL}} = \int d\mathbf{r}d\mathbf{r}'\, \Delta n(\mathbf{r})\, \mathcal{K}[n](\mathbf{r}, \mathbf{r}')\, \Delta n(\mathbf{r}'), \tag{3}$$

where $\Delta n(\mathbf{r}) = n(\mathbf{r}) - n_{\mathrm{eq}}(\mathbf{r})$ is the deviation of the density distribution from the equilibrium unperturbed density $n_{\mathrm{eq}}(\mathbf{r})$ due to the frozen phonon-type perturbation, and ΔT_{NL} is a nonlocal KEDF correction to energy difference Δ, when going from $n_{\mathrm{eq}}(\mathbf{r})$ to $n(\mathbf{r})$, rather than a total-energy functional. Note that in general Eq. (3) is nonlocal not only in the real space, but also in the reciprocal space after the Fourier transform. Assuming that $\mathcal{K}[n](\mathbf{r}, \mathbf{r}')$ depends only on the coordinate difference, $\mathcal{K}[n](\mathbf{r} - \mathbf{r}')$, it can be made local, i.e., involve a single reciprocal lattice summation. Importantly, this is still a much better approximation than a local-in-real-space functional. For a uniform electron gas the function ($\mathcal{K}$) is nothing but the Lindhard function, which is strongly nonlocal in the coordinate space, but strictly local in the momentum space. Furthermore, we believe that for the density differences it is a much better approximation than for the total density functionals[7,15,25,32]. Thus, we assume that

$$\Delta T_{\mathrm{NL}} = \int d\mathbf{r}d\mathbf{r}'\, \Delta n(\mathbf{r})\, \mathcal{K}[n](\mathbf{r} - \mathbf{r}')\, \Delta n(\mathbf{r}'). \tag{4}$$

For a homogeneous electron gas the functional can be further constrained to depend only on $|\mathbf{r} - \mathbf{r}'|$, leading to the following nonlocal KEDF correction:

$$\Delta T_{\mathrm{NL}} = \int d\mathbf{r}d\mathbf{r}'\, \Delta n(\mathbf{r})\, \tilde{\chi}^{-1}_{\mathrm{NL}}(|\mathbf{r} - \mathbf{r}'|)\, \Delta n(\mathbf{r}'), \tag{5}$$

where $\tilde{\chi}^{-1}_{\mathrm{NL}}$ is density-dependent susceptibility kernel derived from Lindhard response function. Similar ideas have been explored in previous works, notably in the Perrot functional[52] and in the ansatz proposed by Chai and Weeks[53]. While the expression for $\mathcal{K}$ in Eq. (4) does not have to be isotropic as in Eq. (5), we will accept in the following the latter form, as a reasonable starting point.

In a periodic system, $\Delta n(\mathbf{r}) = \sum_{\mathbf{G}} \delta n_{\mathbf{G}} e^{i\mathbf{G}\cdot\mathbf{r}}$ represents the density expansion in the reciprocal basis set **G**, and $\delta n_{\mathbf{G}}$ denotes the density variation. In the following we will consider potential perturbations that do not break periodicity ("frozen phonons"). By construction, the non-local correction T_{NL} is a functional of the total density n obtained from OF-DFT and is assumed to be local in **q**-space. The discretized version of the nonlocal KEDF in **q**-space is obtained by taking the Fourier transform of Eq. (5), as given by:

$$\Delta T_{\mathrm{NL}}[n(\mathbf{G})] = \Omega \sum_{\mathbf{G}} \delta n^2_{\mathbf{G}} \left(\tilde{\chi}^{-1}_{\mathrm{NL}}[n]\right)_{\mathbf{G}}. \tag{6}$$

Here, Ω is the volume of the entire cell in **r**-space. The summation spans over discrete reciprocal vectors **G** and $\delta n_{\mathbf{G}}$ is the density variation.

At this point, $\tilde{\chi}^{-1}_{\mathrm{NL}}[n]$ is an arbitrary functional, which needs to be parametrized in terms of $n_{\mathbf{G}}$; a suitable functional form needs to be selected, and the parameters entering this form will be functions of $n_{\mathbf{G}}$, as determined through machine learning. To guess an optimal functional form is a challenging task that can only be solved when a lot of experience will have been accumulated. Keeping in mind the case of uniform electron gas, we select somewhat arbitrarily, the functional form corresponding to the difference between the Lindhard function and the quadratic response in the generalized TFλW functional. This form includes just two parameters, the free-electron-gas Fermi vector, and the coefficient in front of the von Weizsacker term, λ.

In this case, the nonlocal KEDF kernel, $\tilde{\chi}_{\mathrm{NL}}^{-1}$, will be written as

$$\tilde{\chi}_{\mathrm{NL}}^{-1}[n](\mathbf{G}) \quad = \tilde{\chi}_{\mathrm{Lind}}^{-1} - \tilde{\chi}_{\mathrm{TF}\lambda\mathrm{W}}^{-1} \tag{7}$$

where:

$$\begin{aligned}\tilde{\chi}_{\mathrm{Lind}}^{-1}(\eta'(n(\mathbf{G}))) \quad = \quad & -\tfrac{\pi^2}{k_{\mathrm{F}}}\left(\tfrac{1}{2}\left(1+\tfrac{1-\eta'^2}{2\eta'}\right.\right.\\ & \left.\left.\times \ln\left|\tfrac{1+\eta'}{1-\eta'}\right|\right)\right)^{-1},\end{aligned} \tag{8}$$

$$\tilde{\chi}_{\mathrm{TF}\kappa\mathrm{W}}^{-1}(\eta'(n(\mathbf{G}))) = -\frac{\pi^2}{k_{\mathrm{F}}}\left(\frac{1}{1+3\kappa\eta'^2}\right)^{-1}. \tag{9}$$

Here, $k_{\mathrm{F}} = (3\pi^2 n_{\mathbf{G}=0})^{1/3}$ is the Fermi wave vector of the uniform electron gas with the average density $n_{\mathbf{G}=0}$ and we will use it as an adjustable parameter in ML, i.e., $\eta' = \frac{\eta}{\alpha}$, with $\eta = \frac{G}{2k_{\mathrm{F}}}$ is now the reduced momentum. The κ parameter corresponds to λ in the TFλW functional, and we shall use it as a second ML parameter, as follows:

- For $\alpha = 1$ and $\kappa = 1$, $\tilde{\chi}_{\mathrm{Lind}} = \chi_{\mathrm{Lind}}$, where χ_{Lind} denotes the Lindhard response function for the uniform electron gas[2,31]. Similarly, $\tilde{\chi}_{\mathrm{TF}\kappa\mathrm{W}} \rightarrow \chi_{\mathrm{TF}W}$, where $\chi_{\mathrm{TF}W}$ is the TFW response function, corresponding to the second-order gradient expansion approximation with $\lambda = 1$. This represents the limiting case of the Lindhard response function as $\eta \longrightarrow \infty$[49,50].
- For $\alpha = 1$ and $\kappa = \frac{1}{9}$, $\tilde{\chi}_{\mathrm{TF}\kappa\mathrm{W}} \rightarrow \chi_{\mathrm{TF}\frac{1}{9}\mathrm{W}}$, where $\chi_{\mathrm{TF}\frac{1}{9}\mathrm{W}}$ is the TF$\frac{1}{9}$W response function, corresponding to the second-order gradient expansion approximation with $\lambda = \frac{1}{9}$. This provides the limiting case of the Lindhard response function as $\eta \longrightarrow 0$[49,54].

The behavior of these response functions, corresponding to the Lindhard and TFλW response functions, is illustrated in Fig. 1.

We want to emphasize again that the chosen functional form, inspired by uniform electron gas, is somewhat arbitrary and likely not the best choice, but appears to be a good starting point. Indeed, selection of the two adjustable parameters is not physically motivated and should be looked at as a purely mathematical construct. However, our choice was motivated by the following considerations: (i) we want to retain some aspects of the Lindhard function, as linear response is crucial; (ii) we know that OF-DFT can deviate from KS-DFT in either direction for different materials, so we needed a method that allows for a sign change and includes parameters that control this sign change. Scaling the independent variable in the Lindhard function provides these desired properties.

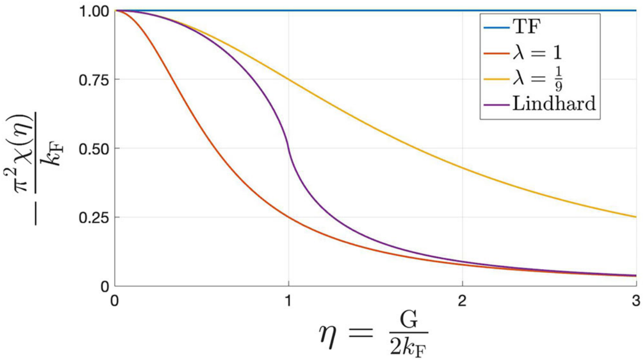


Fig. 1 | Lindhard correspond to the response function of uniform system of independent fermions, and TF model shows the constant response. The factor $\lambda = 1$ and $\frac{1}{9}$ corresponding to response of second order gradient expansion approximations.

The key remaining task is to investigate the density dependence of the scaling parameters α and κ, which are determined using data-driven machine learning (ML) techniques. In particular, a kernel-based Bayesian regression technique[55] is employed to predict α as follows:

$$\alpha = \sum_{l=1}^{N_\tau} \mathbf{w}_l \mathsf{k}(\mathbf{X}_i, \tilde{\mathbf{X}}_\tau), \tag{10}$$

where the kernel k measures the distance between the descriptor vectors $\mathbf{X}_i$ and $\tilde{\mathbf{X}}_\tau$, corresponding to different sets in the training data. The weights $\mathbf{w}_l$ are trained to DFT dataset and summed over the index l, with the size of the training dataset denoted by N_τ. The descriptor vectors are defined in terms of the Fourier components of the moments of the electronic density,

$$\overrightarrow{n}^m_G = \sum_{\mathbf{G}\backslash\{\mathbf{G}=0\}} |n_{\mathbf{G}}| \parallel \mathbf{G}\parallel_2^m \quad \& \quad m \in \mathbb{Z},$$

such that description vector are generalized as

$$\mathbf{X} = \begin{bmatrix} \overrightarrow{n}^m_G \\ \kappa \end{bmatrix}.$$

Another parameter, κ, is obtained through linear regression against the average density $n_{\mathbf{G}=0}$. However, the constraint imposed on κ through this regression can be too strict for training α. To develop a more effective model, one could consider a multi-variable kernel-based Bayesian regression technique or another algorithm from the family of manifold learning methods. For now, this is our initial prototype, and we are testing the effectiveness of Bayesian regression model. While a fully developed functional could potentially involve more parameters beyond just α and κ, our goal is to demonstrate that machine learning (ML) methods can effectively be use to learn these two parameters, and significantly improving accuracy.

Analysis of Δ-correction

By construction, the nonlocal KEDF can be either negative or positive, depending on the TF$\frac{1}{9}$W functional as compared to the "exact" Kohn-Sham kinetic energy density functional. According to Eq. (1), the non-local KEDF relationship is expressed as:

$$T_{\mathrm{NL}} = T_s - T_{\mathrm{TF}}\frac{1}{9}W. \tag{11}$$

Now, the variation of the non-local KEDF, i.e., the non-local correction, $\Delta T_{\mathrm{NL}}[n(\mathbf{G})]$, is derived as a functional of the density obtained self-consistently through OF-calculations. Consequently, the non-local KEDF correction, $\Delta T_{\mathrm{NL}}[n(\mathbf{G})]$, must account for both positive and negative deviations, as the TF$\frac{1}{9}$W model can differ across material systems, either exceeding or falling short of the Kohn-Sham benchmark. Note that in Fig. 1, $\lambda = 1$ always needs a positive correction, and $\lambda = 1/9$a negative one. From that fact it is clear that in the proposed scheme λ must be system-dependent. To account for this, the KEDF kernel, $\tilde{\chi}_{\mathrm{NL}}^{-1}[n]$, is scaled and incorporates two parameters, α and κ as in Eq. (8) and (9) where κ now plays the role of an adjustable λ.

We use the relative error between OF and KS calculations, both with and without non-local correction, integrated across the perturbation parameters g, as our figure of merit, or "cost". Analyzing the behavior of this cost,

$$\frac{\Delta_{\mathrm{NL}-g}[n]}{\Delta_g[n]} = \frac{\int\left[(\Delta_{\mathrm{OF}} + \Delta T_{\mathrm{NL}}) - \Delta_{\mathrm{KS}}\right]^2 dg}{\int\left[\Delta_{\mathrm{OF}} - \Delta_{\mathrm{KS}}\right]^2 dg}, \tag{12}$$

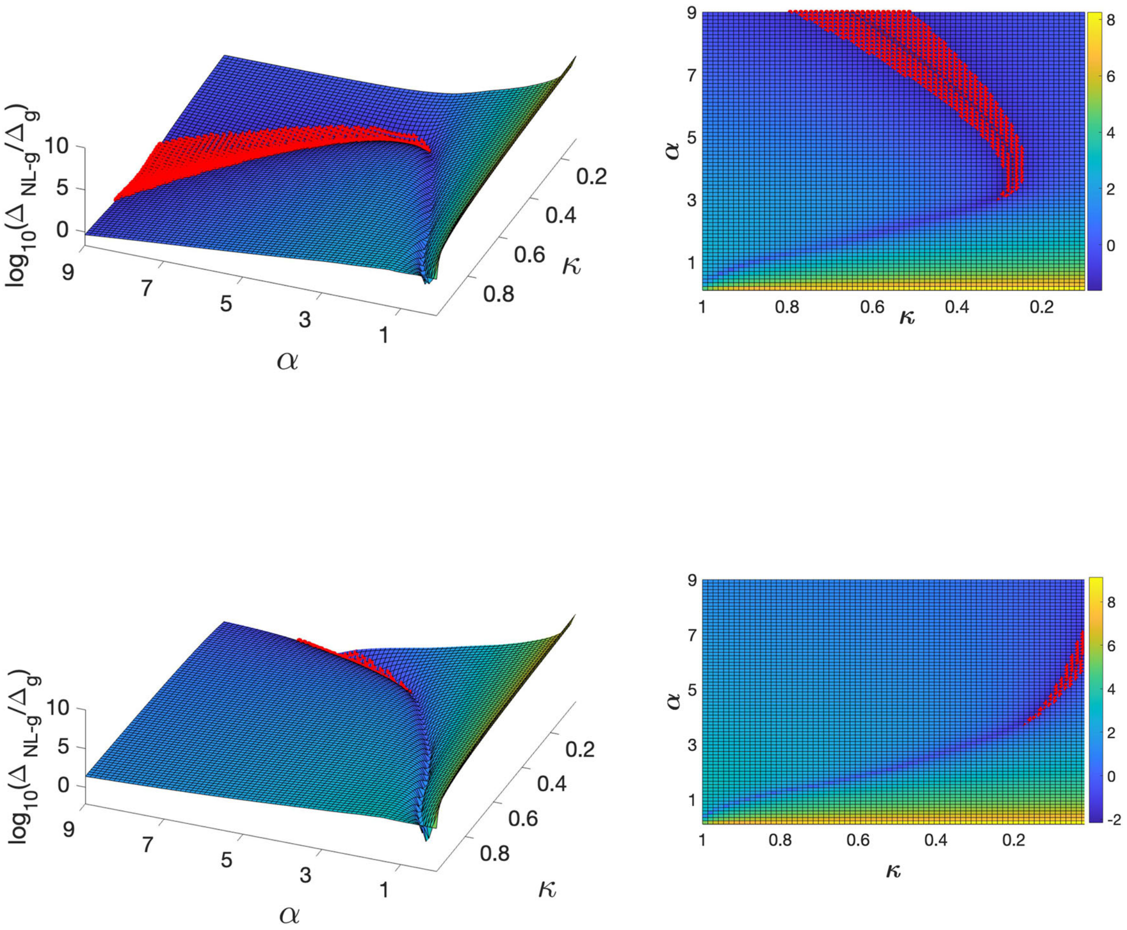


Fig. 2 | Figure shows the cost profiles in the α-κ space for FCC Al (first row) and HCP Mg (second row). The left panel displays the profiles at a specific orientation, while the right panel provides a top view of the same profiles, with elevation indicated by the color bar. Red regions highlight areas with costs less than or equal to—1 (in log units). These are the representative funnel shapes, and the differing funnel shapes correlate directly with positive and negative non-local corrections. For details, see Supplementary Figs. 2–5.

within the α-κ space, each material and perturbation pattern, reveals a distinctive funnel-shaped region highlighted in red. This is illustrated in Fig. 2 for representative cases. This red funnel region highlights a significant dip in the $(\frac{\Delta_{\mathrm{NL}-g}}{\Delta_g})$ surface for specific values of the KEDF kernel scaled parameters α and κ. The characteristic red funnel exhibits subtle variations in shape across the different systems under investigation. These variations can be attributed to the fact that the non-local correction $T_{\mathrm{NL}}[n(\mathbf{G})]$ can be either positive or negative, depending on the different phases of un-distorted and slightly distorted structures in our study.

To effectively improve the cost $(\frac{\Delta_{\mathrm{NL}-g}}{\Delta_g})$ and ensure that results remain within the red funnel region, selecting an appropriate value of α for a given κ is crucial. This prediction is achieved by employing linear regression against the mean density per unit cell, $n_{\mathbf{G}=0}$, resulting in κ_{reg} as demonstrated in Fig. 3. In the figure, different markers correspond to different material systems. Each point represents the average of all possible κ values within the red funnel, which exists for the case when the cost $(\frac{\Delta_{\mathrm{NL}-g}}{\Delta_g})$ is at-least lower by the factor of 10, and encompasses various phases of both distorted and undistorted conventional cells. The crude linear regression method is used to estimate the relationship between the scaled parameter κ in terms of the regression variable κ_{reg}, with the average density per unit cell, denoted as $n_{\mathbf{G}=0}$, serving as the independent regression variable. This approach allows to obtain a tentative estimation of the scaled parameter κ in terms of κ_{reg}. The purpose of this simple linear regression is not to achieve a perfect fit, but to ensure that the value of κ_{reg} lies within the funnel-shaped structure in the α-κ space, which our analysis confirms.

Subsequently, for a designated parameter κ_{reg}, the optimized value, α_{opt}, is computed using an optimization algorithm `fminbnd` implemented in `Matlab`, which uses a combination of golden section search and parabolic interpolation over a specified interval. Hence, we aim to minimize the cost function, $(\frac{\Delta_{\mathrm{NL}-g}}{\Delta_g})$, with the cost function tolerance threshold of 1×10^{-13} and setting the constraint on the α_{opt} from 0.1 to 9. Both α and κ are scaling parameters, so we constraint them to vary only within one order of magnitude. It is worth noting that these optimized values of α_{opt} typically fall within the range delineated by the red funnel in Fig. 2. As an example, Fig. 3 includes an inset figure showing the cost surface along with the corresponding values of α and κ for hexagonal closed-packed structure (HCP) of magnesium (Mg). These values—as highlighted in the offset—are close to κ_{reg} and α_{opt}, which were obtained through a non-linear optimization scheme in `Matlab`.

The improvement in overall calculation performance is quantitatively evaluated through the deviation, defined as:

$$\Delta_{\mathrm{model}} = E_{\mathrm{model}}(g) - E_{\mathrm{model}}(g=0), \tag{13}$$

where E_{model} denotes the energy computed using the corresponding ground-state density. Specifically, when model = KS, the ground-state KS density is used, whereas for model ∈{OF, NL}, the ground-state OF density is employed. The deviation, Δ_{model}, takes values from the set $\{\Delta_{\mathrm{KS}}, \Delta_{\mathrm{OF}}, \Delta_{\mathrm{NL}}\}$. Here, $g=0$ corresponds to the equilibrium case (no frozen phonon perturbation), and model refers to KS, OF, or NL. The direction of the perturbation is indicated in parentheses through the value of g. The behavior of Δ_{model} is illustrated in Fig. 4 for κ_{reg} and α_{opt}. These model sets provide a reference for the comparison, and validates the efficacy of our non-local model. The figure illustrates the deviation Δ_{model} for each perturbation parameter, 'g', corresponding to the cases depicted in Fig. 2). Notably, the incorporation of the correction $T_{\text{NL}}[n(\mathbf{G})]$ via the non-local KEDF kernel yielding substantially better agreement with KS calculations. Additionally, the non-local model is robust enough to distinguish between positive and negative error correction. Notably, the average improvement in accuracy for the optimal choice of parameters is $\mathcal{O}(200)$ (see Supplementary Fig. 1), whereas the improvement for the optimized parameters obtained above is $\mathcal{O}(60)$ (see Supplementary Tables 1–6).

In order to verify the results, we compared the deviation $\underline{\Delta}(\alpha) = \alpha_{\text{opt}} - \alpha_{\text{pert}}$, where α_{pert} ranges from 15% to 15% of α_{opt}. Additionally, we analyzed the change in the normalized deviation, $\underline{\Delta}_{\text{N}} = (\frac{\Delta_{\text{NL}-g}(\alpha_{\text{pert}}) - \Delta_{\text{NL}-g}(\alpha_{\text{opt}})}{\Delta_g})$, as illustrated in Fig. 5. The normalized deviation $\underline{\Delta}_{\text{N}}$ become pronounce for some cases when α_{pert} is considerably different than α_{opt}. This analysis serves as our test of α-sensitivity to the overall cost. The results indicate that $\Delta_{\text{NL}-g}(\alpha_{\text{pert}}) \geq \Delta_{\text{NL}-g}(\alpha_{\text{opt}})$, underscoring our motivation to minimize the deviation Δ_{NL} as much as possible. There is an outlier where $\Delta_{\text{NL}-g}(\alpha_{\text{pert}}) < \Delta_{\text{NL}-g}(\alpha_{\text{opt}})$, for which the method yields worse results compared to the original deviation (Δ_g). This corresponds to frozen phonons in heavily strained aluminum (5% isochoric c—compression of the ideal hexagonal close-packed, HCP, lattice). The outlier is excluded from our Bayesian training algorithm to prevent over-fitting, as its inclusion, though marginal, reduces the model's predictive accuracy. Additionally, in the case of magnesium (Mg), our model does not improve upon the existing TF$\frac{1}{9}$W model, as the designated funnel corresponds to a very narrow range of α_{opt} for the fixed κ obtained through regression. However, it does not worsen the result either, as the original deviation (Δ_g) remains ~10^{-4} ha. Overall, this demonstrates that a simple partitioning of the nonlocal functional in Fourier space can improve accuracy.

The result presented demonstrates the feasibility and accuracy of our proposed hypothesis for decomposing nonlocal KEDFs, and these findings should be robust enough to transfer to larger systems for studying electronic structure calculations which is our primary interest.

Machine learning

In the earlier section, we demonstrated that our selected functional form, which depends on two parameters, α and κ, is indeed a powerful tool. The only challenge lies in accurately determining these parameters to bypass the need for Kohn-Sham (KS) calculations. This section presents results from the calculation of kinetic energy density functional (KEDF) with a nonlocal contribution using machine learned model, in addition to the local TF$\frac{1}{9}$W model, and compares the result with the "benchmark" Kohn-Sham (KS) calculation. A significant technical challenge arises when handling density changes, $\delta n_{\mathbf{G}}$, that alter the crystal unit cell. Existing Kohn-Sham (KS) codes[56] generate $n_{\text{KS}}(\mathbf{r})$ and $n_{\text{KS}}(\mathbf{G})$ on regular grids aligned with the crystal symmetry, necessitating a sophisticated mathematical model for interpolation scheme. To address this issue, we propose a straightforward test utilizing a frozen phonon type of perturbation. In this case, the positions of the atoms are held fixed with respect to the perturbation parameter, (g), thereby avoiding potential pitfalls associated with mesh changes. The machine learning (ML) algorithm used for training and testing is based on kernel-based Bayesian regression. In this approach, the optimal values, denoted by α_{opt}, are used to update the weight vector $\mathbf{w}$. Further details are provided in the "Machine Learned Model" subsection of the Methods section. The test/train set partitioning is achieved through five-fold stratified Monte-Carlo Cross Validation techniques. In Bayesian regression, the kernel is defined using descriptor vectors that encapsulate information regarding the density moments, and κ_{reg}. Utilizing this ML approach, predictions of α values for the test cases result in a significant enhancement, defined as the inverse of the "cost", in overall accuracy, without introducing additional computational complexity. The details of these improvements are depicted in Fig. 6, with the ML technique applied across 200 test trials (only 100 sample trails are shown). The results are presented on a log_{10} scale, and for the given trials, our ML method improves accuracy on average by up to 1.44 on the log_{10} scale. This validates the robustness of our algorithm in predicting α, referred to as α_{pred}, thereby improving the overall TFλW model by at least a factor of 10.

While Fig. 6 contains all information about the improvement that our formalism provides, it is also instructive to show the improvement in

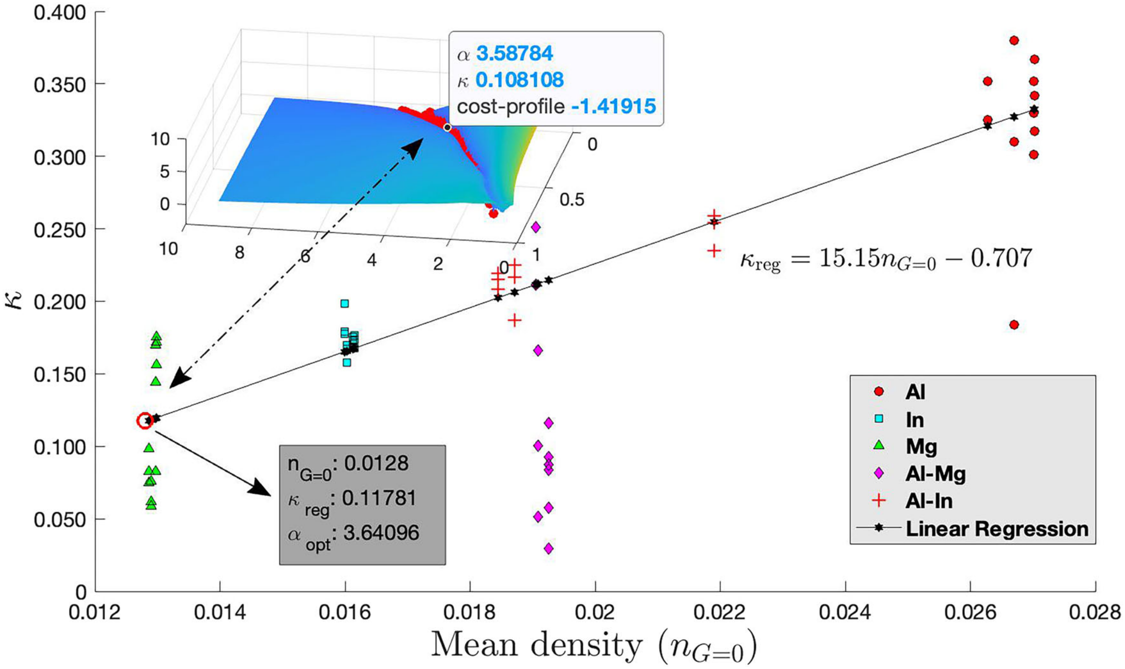


Fig. 3 | This figure illustrates the linear regression between the parameter κ and the mean density per unit cell ($n_{G=0}$). Each marker represents the average of all κ values within the red funnel shown in Fig. 2. The calculation is performed for four different phases, including three slightly distorted conventional phases for each material. The purpose of the linear fit is not to precisely match all data points but to provide a predictive model for κ that aligns with the target funnel. For example, the highlighted region in the offset surface plot is for hexagonal close-packed structure of magnesium (Mg).

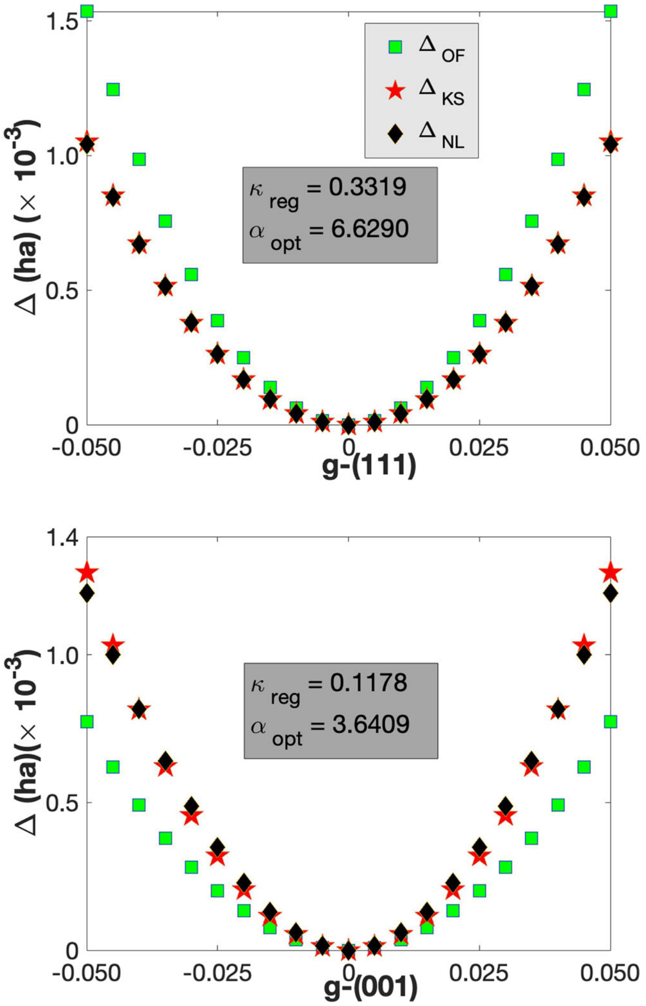


Fig. 4 | The plot illustrates the energy deviations (Δ) due to frozen phonon perturbation g along different crystallographic directions: (111) for FCC Al, (001) for HCP Mg. For Al Δ_{KS} is $<\Delta_{OF}$, while for Mg, Δ_{KS} is $>\Delta_{OF}$. The nonlocal correction algorithm improves the deviations, and the corresponding optimized parameters κ and α derived from the optimization scheme are shown.

such an observable parameter as phonon frequency. Indeed, while our main target function, Δ, represents an integrate deviation from the Kohn-Sham result over the entire range of the phonon displacements, one can also use a less stringent criterion, namely the second derivative of the total energy with respect to the ionic displacement, i.e., the dynamic matrix. For the high-symmetry displacements this second derivative is an eigenmode, and is proportional to the square of the corresponding phonon,

$$\Omega = \frac{1}{M}\sqrt{\left(\frac{d^2E}{du^2}\right)\bigg|_{g=0}},$$

where u is the ionic displacement, and M the ionic mass. This is just the derivative of our energy deviation function Δ with respect to the perturbation parameter g. Then we can introduce the "phonon error" as

$$\delta_{\text{model}} = \left|\frac{\Omega_{\text{KS}} - \Omega\text{model}}{\Omega_{\text{KS}}}\right|,$$

where model $\in$ {OF, NL}. The results for representative cases are presented in Table 1.

Systems and Methods: Crystal structures and perturbations

We consider body-centered cubic (BCC), face-centered cubic (FCC), hexagonal close-packed (HCP), and body-centered tetragonal (BCT) crystal phases of magnesium (Mg), aluminum (Al), indium (In), and binary systems such as aluminium-magnesium (Al-Mg) and aluminium-indium (Al-In). These systems form a diverse set that includes simple metals and transition metals in a variety of geometries. Importantly, well-tested local pseudopotentials[57] are available for the chemical elements in question, i.e., Mg, Al, and In, allowing for a careful comparison of the results obtained from KS-DFT and OF-DFT calculations. The conventional cell consists of 2-atom cells, and we choose the ideal $\gamma = c/a$ ratios of 1.633 and 1.414 for HCP and BCT structure respectively.

We applied two types of perturbations to study the stability and structural response of Al, Mg, In, and their alloys (Al-Mg, Al-In) in various crystal configurations: BCC, FCC, BCT, and HCP.

1. **Frozen Phonon calculations** : To lower symmetry, we applied perturbation parameter g in the form of atomic displacement. These atomic displacement were applied along specific crystallographic directions as:
   - BCC/FCC: displacements along (110), (111), and (001)
   - BCT: displacements along (110), (100), and (001)
   - HCP: displacements along (100), (010), and (001)

   Each displacement magnitude is governed by the parameter g, where $—0.05 \leq g \leq 0.05$, modifying each atomic coordinate η to $\eta\prime = (1+g)\eta$.
2. **Uniaxial Distortion** (ϵ): We also simulated structural strain by stretching the crystal along the x-y plane while compressing it along the z-axis, ensuring constant volume. Here, ϵ quantifies the distortion, with values of 0% (no strain), 3%, and 5%.

For each uniaxial distortion (ϵ), we applied frozen-phonon atomic perturbations (g), enabling us to explore structural responses under combined symmetry-lowering displacements and uniaxial strain.

Discussion

In this work, we have introduced a novel class of KEDFs for OF-DFT that allows for the following decomposition: a local part in coordinate space and a local part in reciprocal space. As a demonstration of principle, we have selected the TFW KEDF for the former, and a Lindhard response function-based functional with two system-dependent parameters for the latter. We have machine-learned these parameters from Kohn-Sham data using Bayesian linear regression combined with the kernel method, while employing moments of the Fourier components of the electronic density as the descriptor. Through representative cases of metals and binary alloys, we have demonstrated that even for the rather simple functional form chosen, there is more than an order-of-magnitude improvement in the accuracy of the energy relative to the TFW KEDF, when the machine-learned model is applied non-self-consistently for frozen phonon calculations, with negligible increase in the cost.

The development of more sophisticated functionals, particularly for the part that is local in reciprocal space, along with self-consistent application, is likely to further increase the accuracy of OF-DFT calculations, making it a worthy subject of future research. The development of more accurate descriptors is expected to further increase the accuracy of the machine learned model and consequently OF-DFT calculations, making it another subject for future research.

Methods

DFT calculations

All calculations are performed using the M-SPARC code for rapid prototyping[56,58], which is a `Matlab` version of the large-scale parallel electronic structure code, SPARC[58,59]. This code implements the real-space finite-difference method, whose formulation and implementation in the context of KS-DFT and OF-DFT can be found in previous works[60–63]. We

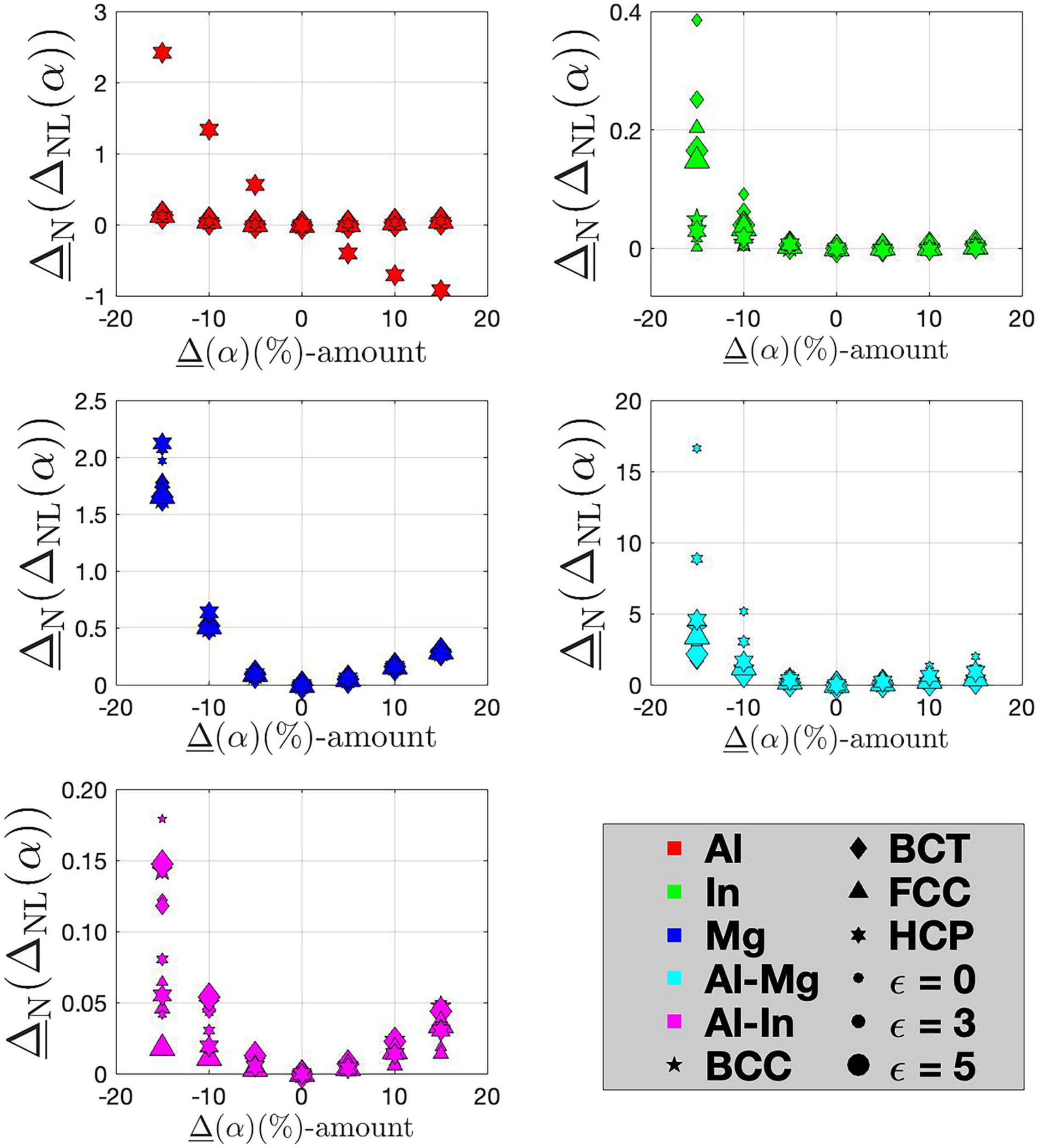


Fig. 5 | This figure illustrates the α-sensitivity to the normalized deviation Δ_N as a function of the nonlocal deviation (Δ_{NL}). The *x*-axis represents the deviation percentage of α_{opt} from 15% to 15%, while the y-axis shows the deviation in the nonlocal correction. The colors correspond to different materials, the marker shape represents different crystal structures, and the marker size indicates the introduced strain. The overlap of the data points suggests that there is no significant difference in the normalized deviation (Δ_N) for the cases shown in the legend. Positive values indicate that the nonlocal correction with α_{pert} is greater than the nonlocal correction with α_{opt} and vice-versa.

employ the local density approximation (LDA-PZ)[4,64] for the exchange-correlation functional and use the bulk-derived local pseudopotentials (BLPS)[57]. In the OF-DFT calculations, we choose the TFW kinetic energy density functional with weight factor $\lambda = 1/9$. In the KS-DFT calculations, we perform Brillouin zone integration using a $15 \times 15 \times 15$ Monkhorst-Pack grid for the FCC, BCC lattices, and $15 \times 15 \times 10$ grid for the HCP and BCT lattices, which ensures that the energies are converged to within 10^{-7} ha/atom. In all calculations, we employ a 12-th order finite-difference approximation and a grid spacing of 0.3 bohr, which ensures that the computed energies are converged to within 10^{-6} ha/atom. Finally, the change in energy arising due to a perturbation is converged to within 10^{-7} ha/atom.

Machine learned model

The training and prediction of α are crucial steps, accomplished using a Laplacian-based activation kernel:

$$\mathrm{k_E} = \exp\left(-\parallel \mathbf{X}_i - \mathbf{X}_\tau \|_1\right),$$

where $\mathbf{X}_i$ and $\mathbf{X}_\tau$ correspond to the descriptor vectors and consist of the moments of the density:

$$n_G^m = \sum_{\mathbf{G}\setminus\{\mathbf{G}=0\}} |n_{\mathbf{G}}| \parallel \mathbf{G}\|_2^m,$$

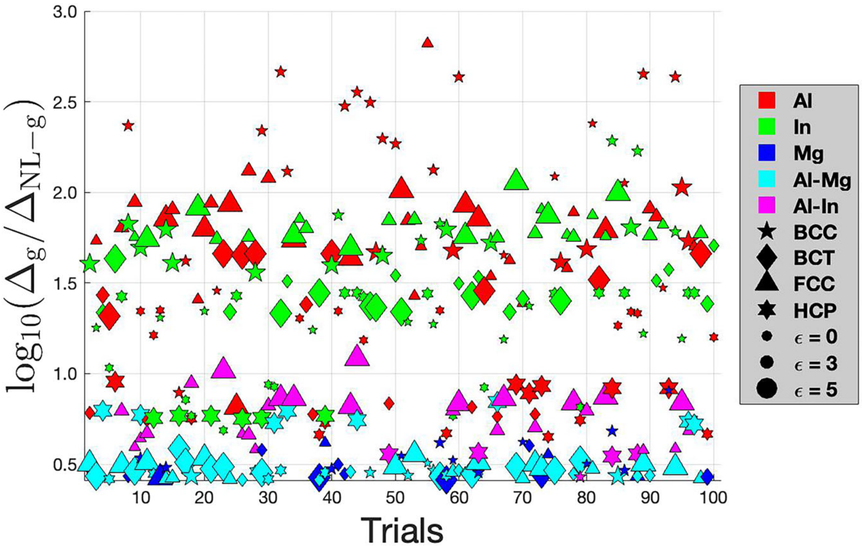


Fig. 6 | The figure demonstrates the enhancement in accuracy due to the inclusion of nonlocal correction, with no additional computational bottleneck compared to the T*F*λ*W* Orbital-Free model. Bayesian regression calculations were performed to predict α_{pred} for 200 different test sets obtained through cross-validation (Only 100 sample points are shown in the figure). Colors correspond to different material systems, and markers indicate different structures. The size of the markers represents the magnitude of strain (ϵ), with smaller markers corresponding to lower strain values and larger markers representing higher strain values.

Table 1 | Deviation of phonon frequencies from the KS ones for representative materials and $\epsilon = 0$

Material	$(\alpha_{\text{pred}}, \kappa_{\text{reg}})$	δ_{NL}	δ_{OF}
Al-FCC	(7.3595, 0.3319)	0.0040	0.2083
Mg-HCP	(5.7560, 0.1178)	0.0245	0.2261
In-BCT	(7.1430, 0.1673)	0.0434	2.3898
AlMg-HCP	(7.9616, 0.2118)	0.0044	0.0491
AlIn-FCC	(7.6429, 0.2543)	0.0103	0.2167

The second column shows the parameters obtained through machine learning regression, while the third and fourth columns present the deviation of phonon frequencies from "benchmark" KS-DFT calculations.

where $-4 \le m \le 4$ and κ_{reg}. Each $\mathbf{X}$ vector is defined as

$$\mathbf{X} = \begin{bmatrix} n_G^m \\ \kappa_{\text{reg}} \end{bmatrix},$$

hence the descriptor vectors $X_{i=1}$ and X_τ have sizes $\mathbb{R}^{N_{\text{des}}\times 1}$ and $\mathbb{R}^{N_{\text{des}}\times N_{\text{st}}}$, respectively, where N_{des} is the dimension of the descriptor vectors and N_{st} is the exhaustive number of structures under study. Now, N_{st} is split into a test set (N_{Test}) and training sets (N_τ), ensuring that at least one instance of each material is included in the test set. The corresponding activation kernel matrix for the training sets is given as:

$$(\mathrm{k}_{\mathrm{E}})_\tau \in \left[\mathrm{k}_{\mathrm{E}}^1; \mathrm{k}_{\mathrm{E}}^2; \ldots; \mathrm{k}_{\mathrm{E}}^{N_{\text{st}}}\right], \qquad \mathbb{R}^{N_{\text{st}}\times N_\tau},$$

where the semicolon is used as a delimiter between the different rows. Note that $(\mathrm{k}_{\mathrm{E}})_\tau$ is a normalized, symmetric, positive definite matrix. The vectors α_{opt} (introduced in Eq. (10)) are obtained via a non-linear optimization scheme and have dimensions $\mathbb{R}^{N_{\text{st}}\times 1}$.

The loss function $\mathcal{L}$ corresponding to Bayesian regression technique is

$$\mathcal{L} = \frac{\Theta_2}{2}\left\|\alpha_{\text{opt}} - \mathrm{k}_{\mathrm{E}}\mathbf{w}\right\|_2^2 - \frac{\Theta_1}{2}\left\|\mathbf{w}\right\|_2^2, \qquad (14)$$

where Θ_1 and Θ_2 are hyper-parameters. The term involving Θ_2 ensures that the optimized coefficients α_{opt} remain close to the training values α_{Train}, as measured by the squared L^2 norm, $\|\alpha_{\text{opt}} - \mathrm{k}_{\mathrm{E}}\mathbf{w}\|_2^2$, between them. This encourages the model to closely approximate the training data. On the other hand, the term associated with Θ_1 acts as a regularization factor, penalizing large weight values $\mathbf{w}$ through the squared L^2 norm $\|\mathbf{w}\|_2^2$, which helps control model complexity and prevent overfitting. The optimal weights, $\mathbf{w}$, which minimize the loss function $\mathcal{L}$, defined in Eq. (14) are shown below:

$$\mathbf{w} = \Theta_2 C_{\mathbf{w}} (\mathrm{k}_{\mathrm{E}}^T)_\tau \alpha_{\text{opt}} \qquad \mathbb{R}^{N_\tau\times 1}$$
$$\text{with} \quad C_{\mathbf{w}} = \left(\Theta_1 \mathcal{I} + \Theta_2 (\mathrm{k}_{\mathrm{E}}^T \mathrm{k}_{\mathrm{E}})\right)^{-1} \qquad \mathbb{R}^{N_\tau\times N_\tau}.$$

The optimized weight vector ($\mathbf{w}$) is verified using k-fold cross-validation as referred in Fig. 6. Using the method defined above, the values of α for the test set are predicted as

$$\alpha_{\text{pred}} = (\mathrm{k}_{\mathrm{E}})_{\text{Test}}\mathbf{w}, \qquad \mathbb{R}^{N_{\text{Test}}\times 1},$$

where the size of $(\mathrm{k}_{\mathrm{E}})_{\text{Test}}$ is $\mathbb{R}^{N_{\text{Test}}\times N_\tau}$ and is briefly mentioned in Eq. (10).

Data availability

The data that support the findings of this study are available within the article and from the corresponding author upon reasonable request.

Received: 9 November 2024; Accepted: 2 May 2025;
Published online: 24 May 2025

References

1. Hohenberg, P. & Kohn, W. Inhomogeneous electron gas. *Phys. Rev.* **136**, B864–B871 (1964).
2. Lundqvist, S. & March, N. H. *Theory of the Inhomogeneous Electron Gas* 1st edn, Vol. 396 (Springer Science & Business Media, 2013).
3. Parr, R. G. Density functional theory of atoms and molecules. In *Horizons of Quantum Chemistry*, (eds. Fukui, K. & Bernard Pullman, B.) 310 (Springer, 1980).
4. Kohn, W. & Sham, L. J. Self-consistent equations including exchange and correlation effects. *Phys. Rev.* **140**, A1133 (1965).
5. Martin, R. M. *Electronic Structure: Basic Theory And Practical Methods* 2nd edn(Cambridge University Press, 2020).
6. Gavini, V. et al. Roadmap on electronic structure codes in the exascale era. *Model. Simul. Mater. Sci. Eng.* **31**, 063301 (2023).
7. Mi, W., Luo, K., Trickey, S. & Pavanello, M. Orbital-free density functional theory: an attractive electronic structure method for large-scale first-principles simulations. *Chem. Rev.* **123**, 12039–12104 (2023).

8. Thapa, B., Jing, X., Pask, J. E., Suryanarayana, P. & Mazin, I. I. Assessing the source of error in the Thomas–Fermi–von weizsäcker density functional. *J. Chem. Phys.* **158**, 214112 (2023).
9. Ullmo, D., Nagano, T., Tomsovic, S. & Baranger, H. U. Semiclassical density functional theory: Strutinsky energy corrections in quantum dots. *Phys. Rev. B* **63**, 125339 (2001).
10. Yannouleas, C., Bogachek, E. N. & Landman, U. Energetics, forces, and quantized conductance in jellium-modeled metallic nanowires. *Phys. Rev. B* **57**, 4872 (1998).
11. Zhou, B. & Wang, Y. A. Orbital-corrected orbital-free density functional theory. *J. Chem. Phys.* **124**, 081107 (2006).
12. Zhou, B. & Wang, Y. A. Accelerating the convergence of the total energy evaluation in density functional theory calculations. *J. Chem. Phys.* **128**, 084101 (2008).
13. Gross, E. K. & Dreizler, R. M. *Density Functional Theory*, Vol. 337 (Springer Science & Business Media, 2013).
14. Kirzhnits, D. A. Quantum corrections to the Thomas-Fermi equation. *Sov. Phys. JETP* **5**, 64 (1957).
15. Mazin, I. I. One method for building non local density functionals for energy. *arXiv* https://arxiv.org/abs/2209.02807 (1988).
16. Mazin, I. & Singh, D. Nonlocal density functionals and the linear response of the homogeneous electron gas. *Phys. Rev. B* **57**, 6879 (1998).
17. Tomishima, Y. & Yonei, K. Solution of the thomas-fermi-dirac equation with a modified weizsäcker correction. *J. Phys. Soc. Jpn.* **21**, 142–153 (1966).
18. Constantin, L. A., Fabiano, E. & Della Sala, F. Performance of semilocal kinetic energy functionals for orbital-free density functional theory. *J. Chem. Theory Comput.* **15**, 3044–3055 (2019).
19. Luo, K., Karasiev, V. V. & Trickey, S. A simple generalized gradient approximation for the noninteracting kinetic energy density functional. *Phys. Rev. B* **98**, 041111 (2018).
20. Constantin, L. A., Fabiano, E. & Della Sala, F. Semilocal Pauli–Gaussian kinetic functionals for orbital-free density functional theory calculations of solids. *J. Phys. Chem. Lett.* **9**, 4385–4390 (2018).
21. Francisco, H. I., Carmona-Espíndola, J. & Gázquez, J. L. Analysis of the kinetic energy functional in the generalized gradient approximation. *J. Chem. Phys.* **154**, 084107 (2021).
22. Perdew, J. P. & Constantin, L. A. Laplacian-level density functionals for the kinetic energy density and exchange-correlation energy. *Phys. Rev. B* **75**, 155109 (2007).
23. Constantin, L. A., Fabiano, E. & Della Sala, F. Modified fourth-order kinetic energy gradient expansion with Hartree potential-dependent coefficients. *J. Chem. theory Comput.* **13**, 4228–4239 (2017).
24. Ziman, J. M. *Principles of the Theory of Solids* 2nd edn, Vol 452 (Cambridge University Press, 1972).
25. Huang, C. & Carter, E. A. Nonlocal orbital-free kinetic energy density functional for semiconductors. *Phys. Rev. B* **81**, 045206 (2010).
26. Constantin, L. A., Fabiano, E. & Della Sala, F. Nonlocal kinetic energy functional from the jellium-with-gap model: applications to orbital-free density functional theory. *Phys. Rev. B* **97**, 205137 (2018).
27. Ho, G. S., Lignères, V. L. & Carter, E. A. Analytic form for a nonlocal kinetic energy functional with a density-dependent kernel for orbital-free density functional theory under periodic and Dirichlet boundary conditions. *Phys. Rev. B* **78**, 045105 (2008).
28. Karasiev, V. V. & Trickey, S. B. Issues and challenges in orbital-free density functional calculations. *Comput. Phys. Commun.* **183**, 2519–2527 (2012).
29. Chacón, E., Alvarellos, J. & Tarazona, P. Nonlocal kinetic energy functional for nonhomogeneous electron systems. *Phys. Rev. B* **32**, 7868 (1985).
30. Mazin, I. Constructing nonlocal density functionals for kinetic energy. *arXiv* https://doi.org/10.48550/arXiv.2209.02807 (1988).
31. Wang, L.-W. & Teter, M. P. Kinetic-energy functional of the electron density. *Phys. Rev. B* **45**, 13196 (1992).
32. Wang, Y. A., Govind, N. & Carter, E. A. Orbital-free kinetic-energy density functionals with a density-dependent kernel. *Phys. Rev. B* **60**, 16350 (1999).
33. Carling, K. M. & Carter, E. A. 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).
34. Zhou, B., Ligneres, V. L. & Carter, E. A. Improving the orbital-free density functional theory description of covalent materials. *J. Chem. Phys.* **122**, 044103 (2005).
35. Ho, G., Ong, M. T., Caspersen, K. J. & Carter, E. A. Energetics and kinetics of vacancy diffusion and aggregation in shocked aluminium via orbital-free density functional theory. *Phys. Chem. Chem. Phys.* **9**, 4951–4966 (2007).
36. Shao, X., Mi, W. & Pavanello, M. Revised Huang-Carter nonlocal kinetic energy functional for semiconductors and their surfaces. *Phys. Rev. B* **104**, 045118 (2021).
37. Mi, W., Genova, A. & Pavanello, M. Nonlocal kinetic energy functionals by functional integration. *J. Chem. Phys.* **148**, 184107 (2018).
38. Xu, Q., Ma, C., Mi, W., Wang, Y. & Ma, Y. Nonlocal pseudopotential energy density functional for orbital-free density functional theory. *Nat. Commun.* **13**, 1–7 (2022).
39. Shin, I. & Carter, E. A. Enhanced von weizsäcker wang-govind-carter kinetic energy density functional for semiconductors. *J. Chem. Phys.* **140**, 18A531 (2014).
40. Golub, P. & Manzhos, S. Kinetic energy densities based on the fourth order gradient expansion: performance in different classes of materials and improvement via machine learning. *Phys. Chem. Chem. Phys.* **21**, 378–395 (2019).
41. Alghadeer, M., Al-Aswad, A. & Alharbi, F. H. Highly accurate machine learning model for kinetic energy density functional. *Phys. Lett. A* **414**, 127621 (2021).
42. Fujinami, M., Kageyama, R., Seino, J., Ikabata, Y. & Nakai, H. Orbital-free density functional theory calculation applying semi-local machine-learned kinetic energy density functional and kinetic potential. *Chem. Phys. Lett.* **748**, 137358 (2020).
43. del Mazo-Sevillano, P., & Hermann, J. Variational principle to regularize machine-learned density functionals: The non-interacting kinetic-energy functional. *J. Chem. Phys.* **159**, 194107 (2023).
44. Remme, R., Kaczun, T., Scheurer, M., Dreuw, A. & Hamprecht, F. A. Kineticnet: Deep learning a transferable kinetic energy functional for orbital-free density functional theory. *J. Chem. Phys.* **159**, 144113 (2023).
45. Ryczko, K., Wetzel, S. J., Melko, R. G. & Tamblyn, I. Toward orbital-free density functional theory with small data sets and deep learning. *J. Chem. Theory Comput.* **18**, 1122–1128 (2022).
46. Kumar, S. et al. Accurate parameterization of the kinetic energy functional for calculations using exact-exchange. *J. Chem. Phys.* **156**, 024107 (2022).
47. Meyer, R., Weichselbaum, M. & Hauser, A. W. 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**, 5685–5694 (2020).
48. Snyder, J. C. et al. Orbital-free bond breaking via machine learning. *J. Chem. Phys.* **139**, 224104 (2013).
49. Benguria, R., Brézis, H. & Lieb, E. H. The Thomas-Fermi-von Weizsäcker theory of atoms and molecules. *Commun. Math. Phys.* **79**, 167–180 (1981).
50. Jones, W. & Young, W. Density functional theory and the von Weizsacker method. *J. Phys. C: Solid State Phys.* **4**, 1322 (1971).
51. Wang, L.-W. & Teter, M. P. Kinetic-energy functional of the electron density. *Phys. Rev. B* **45**, 13196–13220 (1992).
52. Perrot, F. Hydrogen-hydrogen interaction in an electron gas. *J. Phys.: Condens. Matter* **6**, 431 (1994).

53. Chai, J.-D., Ligneres, V. L., Ho, G., Carter, E. A. & Weeks, J. D. Orbital-free density functional theory: linear scaling methods for kinetic potentials, and applications to solid al and si. *Chem. Phys. Lett.* **473**, 263–267 (2009).
54. Kirzhnits, D. Quantum corrections to the thomas-fermi equation. *Soviet Phys. JETP* https://doi.org/10.1139/p73-189 (1957).
55. Nabney, I. *NETLAB: Algorithms for Pattern Recognition* (Springer Science & Business Media, 2002).
56. Xu, Q., Sharma, A. & Suryanarayana, P. M-SPARC: Matlab-simulation package for ab-initio real-space calculations. *SoftwareX* **11**, 100423 (2020).
57. Zhou, B., Wang, Y. A. & Carter, E. A. Transferable local pseudopotentials derived via inversion of the kohn-sham equations in a bulk environment. *Phys. Rev. B* **69**, 125109 (2004).
58. Zhang, B. et al. SPARC v2.0.0: Spin-orbit coupling, dispersion interactions, and advanced exchange–correlation functionals. *Softw. Impacts* **20**, 100649 (2024).
59. Xu, Q. et al. SPARC: Simulation package for ab-initio real-space calculations. *SoftwareX* **15**, 100709 (2021).
60. Ghosh, S. & Suryanarayana, P. SPARC: Accurate and efficient finite-difference formulation and parallel implementation of density functional theory: extended systems. *Comput. Phys. Commun.* **216**, 109–125 (2017).
61. Ghosh, S. & Suryanarayana, P. SPARC: Accurate and efficient finite-difference formulation and parallel implementation of density functional theory: isolated clusters. *Comput. Phys. Commun.* **212**, 189–204 (2017).
62. Ghosh, S. & Suryanarayana, P. Higher-order finite-difference formulation of periodic orbital-free density functional theory. *J. Comput. Phys.* **307**, 634–652 (2016).
63. Suryanarayana, P. & Phanish, D. Augmented Lagrangian formulation of orbital-free density functional theory. *J. Comput. Phys.* **275**, 524–538 (2014).
64. Perdew, J. P. & Zunger, A. Self-interaction correction to density-functional approximations for many-electron systems. *Phys. Rev. B* **23**, 5048 (1981).

Acknowledgements

B.T. acknowledges the support of the Quantum Science and Engineering Center (QSEC) at George Mason University. P.S. gratefully acknowledges funding from the U.S. Department of Energy, Office of Science, under Grant No. DE-SC0023445. T.O. acknowledges support provided by the George Mason University Provost Dissertation Completion Grant. M.E. has been partially supported by the Simons Foundation.

Author contributions

B.T. performed the calculations and coding, O.T. and B.T. did machine learning, I.M. and P.S. provided DFT guidance and M.E. and P.S. provided machine learning guidance. All authors participated in writing the manuscript.

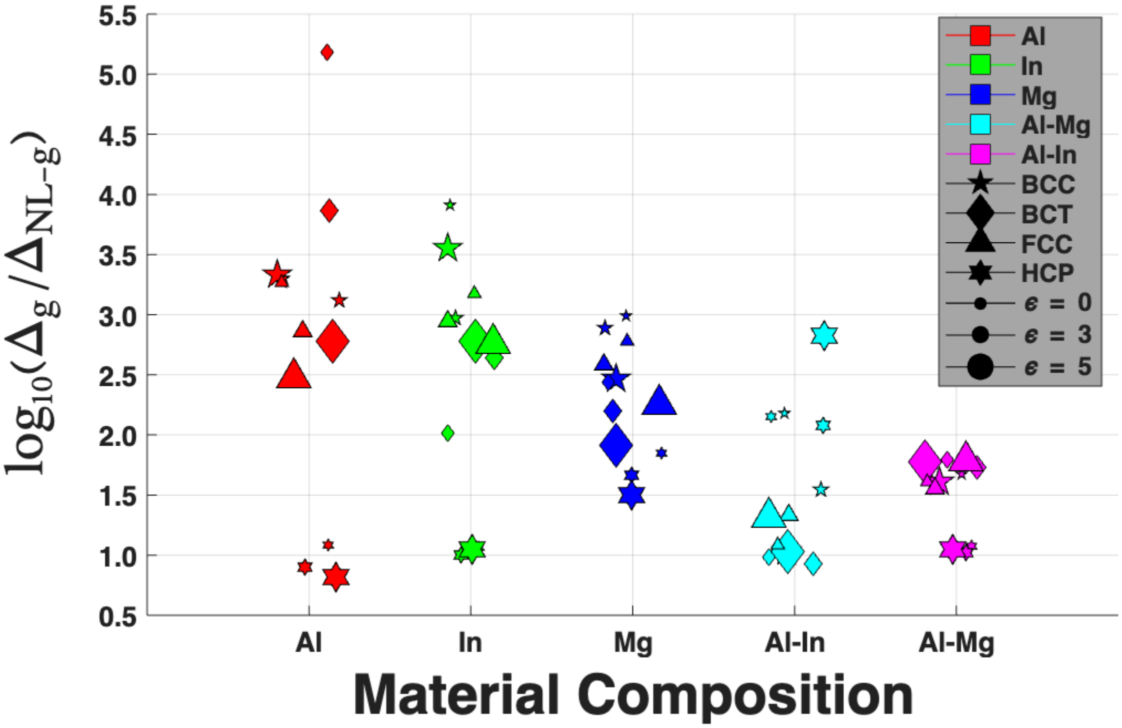
Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41524-025-01643-0.

Correspondence and requests for materials should be addressed to Igor I. Mazin.



Supplementary Figure 1: Ideal improvement averages at $\log_{10}\left(\frac{\Delta_g}{\Delta_{NL-g}}\right) = 2.25$

II. OPTIMIZED ENHANCEMENT:

The section reports the value of κ, obtained through regression (κ_{reg}), and the corresponding non-linear optimization search for α_{opt}, for various structures of the Al, In, Mg, AlIn, and AlMg systems, and the average improvement ratio is 1.77 in $\log_{10}$ units.

cases	κ_{reg}	α_{opt}	Improvement Factor
Al-BCC$-\epsilon = 0$	0.3205	6.6521	3.3033
Al-BCC$-\epsilon = 3$	0.3205	7.0629	2.6759
Al-BCC$-\epsilon = 5$	0.3205	6.6291	3.3193
Al-BCT$-\epsilon = 0$	0.3318	7.7202	0.8697
Al-BCT$-\epsilon = 3$	0.3318	6.8034	1.7682
Al-BCT$-\epsilon = 5$	0.3318	7.1830	1.6611
Al-FCC$-\epsilon = 0$	0.3319	6.6289	2.8902
Al-FCC$-\epsilon = 3$	0.3319	6.9591	2.1708
Al-FCC$-\epsilon = 5$	0.3319	7.4064	2.0201
Al-HCP$-\epsilon = 0$	0.3269	6.6692	1.0807
Al-HCP$-\epsilon = 3$	0.3269	7.7819	0.8980
Al-HCP$-\epsilon = 5$	0.3269	8.9994	0.7903

Supplementary Table 2: Table showing the values of κ_{reg}, α_{opt}, and Improvement Factor for different Al structure

cases	κ_{reg}	α_{opt}	Improvement Factor
In-BCC$-\epsilon = 0$	0.1649	6.4026	1.5328
In-BCC$-\epsilon = 3$	0.1649	6.4034	2.3855
In-BCC$-\epsilon = 5$	0.1649	6.4733	2.0313
In-BCT$-\epsilon = 0$	0.1673	6.2370	3.0341
In-BCT$-\epsilon = 3$	0.1673	6.0451	1.9007
In-BCT$-\epsilon = 5$	0.1673	6.0345	1.8371
In-FCC$-\epsilon = 0$	0.1671	7.9507	1.9038
In-FCC$-\epsilon = 3$	0.1671	6.1322	2.2844
In-FCC$-\epsilon = 5$	0.1671	6.0975	2.2420
In-HCP$-\epsilon = 0$	0.1654	8.3252	0.9609
In-HCP$-\epsilon = 3$	0.1654	8.2519	0.9582
In-HCP$-\epsilon = 5$	0.1654	8.9994	1.0173

Supplementary Table 3: Table showing the values of κ_{reg}, α_{opt}, and Improvement Factor for different In structure