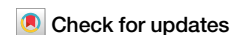


<https://doi.org/10.1038/s41524-026-02191-x>

Addressing hypersensitivity in density functional theory for reliable machine learning

L. Arnstein¹, J. Wetherell^{1,2}, R. Lawrence^{1,3}, P. J. Hasnip¹ & M. J. P. Hodgson^{1,2}✉

In principle, machine learning (ML) can be used within density functional theory to obtain any electronic property of a time-independent many-body system from its ground-state electron density alone. However, when the target quantity (the output) is highly sensitive to small variations in the density (the input), accurately learning the mapping between them is highly challenging in practice. We identify this ‘hypersensitivity’ in two classes of density functionals that contain observables central to most electronic structure calculations: (1) any observable that varies subject to a global constant shift in the external potential (such as the total energy); (2) any observable of the N -electron system that depends on the properties of the $(N + 1)$ -electron system (such as the electron affinity). We demonstrate that the accuracy with which widely used ML models approximate density functionals in these two classes is significantly limited by the pathological nature of the mapping from the density to the observable, even when the reference level of the external potential is fixed to prevent constant shifts. In addition, we find that providing the external potential as input to the ML model resolves this issue, in some cases reducing the error by up to three orders of magnitude when predicting quantities in these two classes.

Materials science relies on computational models to predict the properties of matter. This involves electronic structure calculations of many-electron systems, for which exact solutions to the Schrödinger equation are generally intractable.

Due to its balance of accuracy and computational efficiency, density functional theory (DFT) has become the most popular method for electronic structure calculations¹. Formally, DFT is an exact approach to quantum mechanics in which the ground-state electron density, $n(\mathbf{r})$, replaces the many-body wavefunction as the quantity from which observables are calculated. This formalism is rigorously justified by the Hohenberg-Kohn (HK) theorems², the first of which establishes that the ground-state electron density uniquely determines the external potential, $v_{\text{ext}}(\mathbf{r})$, up to an arbitrary additive constant. By extension, $n(\mathbf{r})$ determines the entire many-body Hamiltonian and, implicitly, all properties derivable from it³. While the first HK theorem is restricted to densities associated with an external potential (v -representable), the constrained search formalism⁴ extended DFT to all densities that can be derived from an N -electron antisymmetric wavefunction (N -representable). This ensures that all observables are well-defined functionals of the density^{5,6}, even in the case of degeneracy or when the density is not v -representable.

The second HK theorem states that there exists a total energy functional $E[n] = E_{\text{ext}}[n] + F[n]$ which is minimised when $n(\mathbf{r})$ is the

ground-state density. $E_{\text{ext}}[n] = \int v_{\text{ext}}(\mathbf{r})n(\mathbf{r})d^3r$ is the external potential energy. $F[n]$, which contains the electron-electron repulsion and the kinetic energy, is a universal functional of the density, meaning that it is the same for all systems, independent of $v_{\text{ext}}(\mathbf{r})$ ^{2,4}. As such, if the exact form of $F[n]$ was known, the ground-state total energy of any many-body system could be calculated via the variational principle⁴. In practice, expressing $F[n]$ as a pure density functional (particularly the kinetic energy component⁷) proved to be extremely challenging. This led the original, ‘orbital-free’ vision of DFT to be abandoned in favour of the Kohn-Sham (KS) approach⁸, which relies on single-particle orbitals in addition to the electron density. The orbitals describe an auxiliary system of non-interacting electrons that are constrained to have the same electron density as the real, many-body system. Their kinetic energy, T_s , is computed from the orbitals, side-stepping the problem of finding a density functional of the kinetic energy. As T_s tends to be close to the true, many-body kinetic energy and closed-form expressions for the Hartree energy and the external potential energy are known, this leaves only a (usually) relatively small unknown component of the total energy—termed the exchange-correlation energy, $E_{\text{xc}}[n]$ —to be approximated. Although many existing approximations perform well for specific system classes, the search for more reliable approximations to $E_{\text{xc}}[n]$ is ongoing⁹. Moreover, the introduction of orbitals came with a significant computational overhead, limiting KS-DFT to modelling systems of

¹School of Physics, Engineering and Technology, University of York, York, UK. ²The European Theoretical Spectroscopy Facility, Louvain-la-Neuve, Belgium.

³Physics Unit, Faculty of Engineering and Natural Sciences, Tampere University, Tampere, Finland. ✉e-mail: matt.hodgson@york.ac.uk

hundreds or at most a few thousand electrons¹⁰. In contrast, orbital-free DFT could, in theory, simulate millions of electrons.

Although the total energy is the central functional in DFT, *any* property of a time-independent many-body system can, in principle, be expressed as a pure density functional^{3,4}. However, owing to a lack of reliable explicit density functional approximations (DFAs) for many physical quantities, the electronic properties of the KS system are employed, despite known limitations. For instance, the KS band (fundamental) gap is often used as an approximation to the real, many-body band gap, even when they differ significantly^{11,12}. If accurate DFAs were available for properties such as the band gap, this could significantly accelerate progress in materials science¹³. Unfortunately, traditional methods for constructing DFAs rely on prior knowledge of the functional, which is lacking for almost all properties¹⁴. In addition, as the DFA for the band gap is likely to be highly non-linear and non-analytic, it may not be feasible to approximate it well with a simple analytic form.

Recently, machine learning (ML) has emerged as a promising approach for developing DFAs^{15–17}. ML models learn the relationship between the electron density, $n(\mathbf{r})$, and an observable, Q , via a collection of examples known as training data. A successfully trained model should then be able to generalise to unseen data, accurately predicting the values of $Q[n]$ corresponding to densities that were not used for training. These predictions can be made significantly faster than the original method used to generate the training data^{18,19}, even when non-local descriptors are used. Moreover, ML models such as kernel ridge regression (KRR) and neural networks (NNs) are universal approximators¹⁷; given sufficient parameters, they can, in theory, approximate any mapping to arbitrary precision. Crucially, these models do not require any prior knowledge of the functional¹⁴, enabling them to capture complex dependencies between $n(\mathbf{r})$ and Q that may evade human intuition²⁰. As such, ML offers an alternative route to attain DFAs for a variety of properties where traditional, analytical tools have fallen short.

A primary goal of ML-DFT is orbital-free DFT¹⁷, where the hope is that ML can eliminate the need for orbitals and thereby overcome the primary computational bottleneck of contemporary DFT calculations. At a minimum, this requires machine-learning the non-interacting (Kohn-Sham) kinetic energy as a density functional, $T_s[n]$ ^{20–24}. Alternatively, by explicitly modelling the universal functional $F[n]$ ²³, or equivalently the total energy $E[n]$ ^{19,25–29}, the auxiliary KS system can be bypassed altogether. Due to its flexibility, ML also opens up the possibility of modelling a variety of density functionals outside of the familiar energy functionals¹⁴, such as the electron affinity¹³ and the band gap^{13,14,28–31}.

A natural alternative to the electron density is to use the external potential as the ML input to predict these observables^{19,26,32,33}. This approach is based on potential functional theory (PFT)^{34–36}, a rigorous dual formalism to DFT¹⁷. Although not all potential $\rightarrow$ observable mappings are unique (e.g. in the case of ground-state degeneracy, the potential $\rightarrow$ density mapping is non-injective), many observables are well-defined functionals of the external potential. In these cases, the density and potential should represent equally valid ML inputs, as both contain *all* the information about a v -representable system with a given number of particles^{2–4}. However, they do not necessarily lead to equally accurate predictions in practice. Ref 26 found $n(\mathbf{r})$ to be the better input for predicting the total energy of molecules. On the other hand, ref. 19 found the opposite result for some system classes, with $v_{\text{ext}}(\mathbf{r})$ yielding more accurate predictions of the total energy. The results of ref. 19 also indicate that the relative effectiveness of using the potential or density depends on the observable being predicted. However, a systematic understanding of the origins of these differences and of the strengths and failure modes of each input is currently lacking. This understanding would facilitate the development of accurate and robust ML models for electronic structure calculations.

In this paper, we identify a fundamental pitfall in using the electron density as the sole ML input to predict observables belonging to two classes (described below), such as the total energy, the electron affinity and the fundamental gap. We show that small variations in the density can correspond to dramatic changes in the value of these quantities. This

‘hypersensitivity’ is inherently challenging from an ML perspective^{17,37,38}, requiring complex models and large, high-quality datasets to accurately learn the steep density $\rightarrow$ observable mapping. We also show that this pitfall is avoided by using the external potential as input to the ML model. The hypersensitivity stems from significant changes in the potential that strongly affect the observable but leave the density largely unchanged; as such, the potential $\rightarrow$ observable mapping is not hypersensitive and hence more amenable to ML approximation. Our findings can therefore help to inform input selection for a variety of important observables. They also explain several aspects of the results of ref. 19, as well as their apparent contradiction with ref. 26.

The first class of density functionals in which we identify hypersensitivity corresponds to physical quantities that change value when a non-zero constant, c , is added to the external potential. Formally, these quantities are not invariant under a global $U(1)$ gauge transformation of the external potential, hereafter referred to as ‘constant shift dependence’ for simplicity:

Definition 1. Let the physical quantity Q correspond to the potential $v_{\text{ext}}(\mathbf{r})$, and, likewise, $\bar{Q}$ correspond to $\bar{v}_{\text{ext}}(\mathbf{r})$, where $\bar{v}_{\text{ext}}(\mathbf{r}) = v_{\text{ext}}(\mathbf{r}) + c$. Q is *constant shift dependent* if $Q \neq \bar{Q}$, and *constant shift invariant* if $\bar{Q} = Q$.

Examples of constant-shift dependent quantities include the total energy, E , ($\bar{E} = E + Nc$), the first ionisation energy, I , ($\bar{I} = I - c$) and the electron affinity, A , ($\bar{A} = A - c$). The electron density itself, on the other hand, is constant shift invariant, i.e. $\bar{n}(\mathbf{r}) = n(\mathbf{r})$. Therefore, a constant (spatially uniform) shift in the external potential will cause the same density to map to two different values of a constant-shift dependent quantity, Q , as depicted in Fig. 1a. However, this can be resolved by using a fixed reference level for the external potential, such as $v_{\text{ext}}(\mathbf{r} \rightarrow \infty) = 0$, which prevents any two potentials from differing exactly by a constant over all space. Under this constraint, each density maps to a single value of Q .

However, we note that even when the reference level is fixed, two potentials may differ approximately by a constant in the region where the density is concentrated (for example, see Fig. 2 in section ‘Dataset 1’). In this region, their difference can be described as $\Delta v_{\text{ext}}(\mathbf{r}) = c + \delta v_{\text{ext}}(\mathbf{r})$, where the spatially varying component, $\delta v_{\text{ext}}(\mathbf{r})$, is small. Note that $\delta v_{\text{ext}}(\mathbf{r})$ could even be 0 over a strict subset of space while still satisfying the fixed reference level. Since $n(\mathbf{r})$ is constant shift invariant, it only responds to the small, non-constant component of Δv_{ext} —therefore, the difference between the electron densities of the two systems, $\Delta n(\mathbf{r})$, is likely to be very small. On the other hand, the difference in a constant-shift dependent property, ΔQ , may be significant, e.g. $\Delta E \approx Nc$. $Q[n]$ is therefore hypersensitive, as similar inputs map to highly disparate outputs (this is illustrated in Figure 1b). In other words, $|\Delta Q/\Delta n|$ is very large, because $|\Delta Q| \gg |\Delta n|$. Conversely, $Q[v_{\text{ext}}]$ is not hypersensitive, because $\Delta v_{\text{ext}}(\mathbf{r})$ is not small.

In summary, the reference level of the external potential must be fixed to prevent global constant shifts and ensure a well-defined mapping from the electron density to a constant-shift dependent quantity. However, near-constant shifts, which can occur even with a fixed reference level, cause these quantities to vary rapidly with respect to the electron density, making learning highly challenging. Note that this problem does not affect all quantities that depend explicitly on the external potential, because some such quantities are not constant shift-dependent. For example, the fundamental gap, $E_g = I - A$, depends explicitly on the external potential but is constant shift invariant: $\bar{E}_g = \bar{I} - \bar{A} = I - c - (A - c) = I - A = E_g$. We emphasise that the relevant factor is constant shift dependence, not dependence on the external potential.

The second class of hypersensitive functionals that we highlight is functionals of the N -electron density that depend on the properties of the corresponding $(N + 1)$ -electron system. Examples include the electron affinity, A , and the fundamental gap, E_g , which both explicitly depend on the total energy of the $(N + 1)$ -electron system ($A = E_N - E_{N+1}$ and $E_g = E_{N-1} - 2E_N + E_{N+1}$). Since the density of the N -electron ground state uniquely determines the external potential, it by extension determines the Hamiltonian and its complete spectra—including the energies of the $N - 1$ and $N + 1$ states. Therefore, the fundamental gap and electron affinity are

Fig. 1 | Schematic of a constant-shift dependent density functional. **a** Under a constant shift in the external potential, the quantity jumps in value, but the electron density does not change. As the same input maps to two different outputs, the learning problem is intractable. This situation is avoided by fixing the reference level of the external potential, which prevents constant shifts from occurring. **b** Under a near-constant shift in the external potential, the functional is well-defined but changes so rapidly that it is challenging to machine-learn in this region of density space.

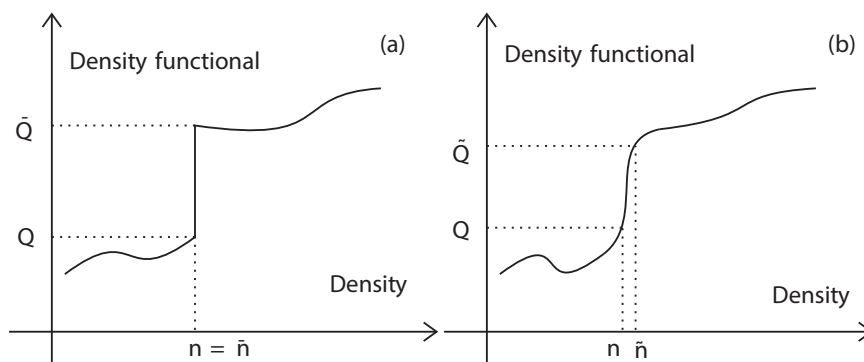
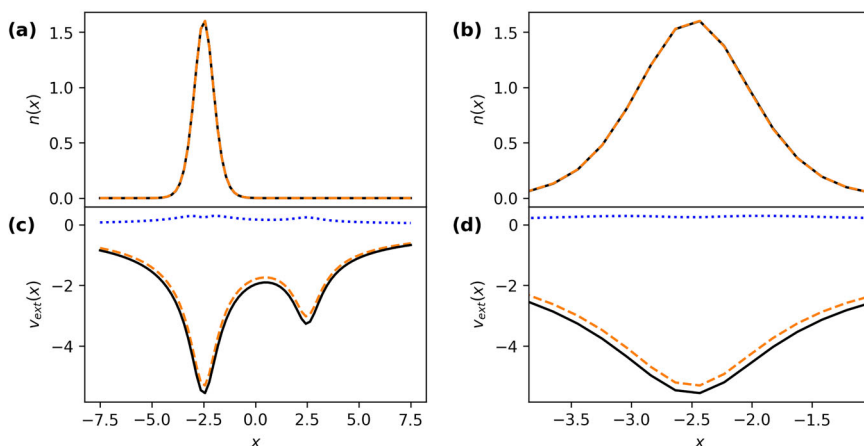


Fig. 2 | An example pair of systems from our first dataset that illustrates a near-constant shift. In the left panels, they are plotted over the full domain of the 1D grid. The black solid lines and dashed orange lines each correspond to a system (see text for details), for which the external potentials and electron densities are plotted in (c) and (a), respectively. The blue dotted line shows the difference between the external potentials, $\Delta v_{\text{ext}}(x)$. In (c), it can be seen that $\Delta v_{\text{ext}}(x)$ varies over space. The right panels show a close-up to the region where most of the electron density is concentrated; in this region, $\Delta v_{\text{ext}}(x)$ is approximately constant (d). Consequently, the electron densities of the two systems are almost indistinguishable (b), but they differ substantially in their constant-shift dependent properties—for example, the system represented by solid black lines has a total energy of -7.72 Ha, whereas the other system has a total energy of -7.18 Ha.



mathematically well-defined functionals of the N -electron ground-state density. However, hypersensitivity arises in these functionals if some change to the external potential significantly affects the $(N + 1)$ -electron system, and hence the functional output, but not the N -electron system, and hence not the functional input. For example, consider two external potentials that only differ significantly in one region of space; see Fig. 3c in section ‘Dataset 1’. If the N -electron densities of both systems are low in this region, they are likely to be insensitive to the difference and hence similar overall. However, if the corresponding $(N + 1)$ -electron systems have significant density in the region where the external potentials differ appreciably, properties such as the affinity and fundamental gap will differ significantly for the two external potentials. These quantities are therefore hypersensitive when expressed as functionals of the N -electron density alone. In contrast, since the difference between the external potentials of the two systems is substantial, these same properties are *not* hypersensitive when expressed as potential functionals.

Below, we use one-dimensional (1D) model systems to test our hypotheses. We use the iDEA code³⁹ to model various systems of two fully interacting electrons subject to double-well molecular potentials by solving the many-body Schrödinger equation numerically. For each system, we calculate the exact many-body electron density $n(x)$, the total energy E , the first ionisation energy I , the electron affinity A , the fundamental gap E_g , and the HK universal functional F . We provide examples of hypersensitivity, wherein systems with near-identical electron densities differ substantially in their values of E (Class 1), I (Class 1), A (Classes 1 and 2) or E_g (Class 2). This is measured numerically as finite difference quotients $|\Delta Q/\Delta n|$ (see Eq. (3)) on the order of 10^2 . We contrast this with the universal functional, which falls in neither class, and subsequently, $|\Delta F/\Delta n|$ is consistently two orders of magnitude smaller. Similarly, we show that (within the dataset) $|\Delta Q/\Delta v_{\text{ext}}|$ does not exceed the order of 1 for all five observables, confirming our

hypothesis that the hypersensitivity is avoided by expressing these quantities as potential functionals. We note that while PFT avoids the ‘Class 1’ and ‘Class 2’ hypersensitivity we find in density functionals, potential functionals may still exhibit hypersensitivity in other contexts.

We then train KRR models to predict each physical quantity, first providing the entire (discretised) electron density as the input, and observe several orders of magnitude greater error for all hypersensitive quantities (compared to F). When we instead use the external potential as the input, this discrepancy in learning accuracy is not present, with similar accuracy for all observables. These results clearly demonstrate that the hypersensitivity in the density functionals we identified can be a major source of error that is avoided by using the external potential as input. This is further supported by ref. 19, who report results that we identify as having greater error for constant-shift dependent density functionals (such as $E[n]$) compared to constant-shift invariant density functionals, but more of a balance for potential functionals. Note that in the case of $E[n]$, the hypersensitivity stems from E_{ext} (the constant-shift dependent component). Since E_{ext} is exactly known, approximating it with ML accumulates unnecessary error.

We also construct a second dataset in which the hypersensitivity is intentionally avoided. This dataset is system-specific, containing a single molecule in which only the nuclear coordinates are varied. We show that within this dataset, all five quantities vary gradually with respect to the electron density, with $|\Delta Q/\Delta n|$ at or below the order of 1. Subsequently, all quantities are learned accurately regardless of the input. This highlights that the ‘learnability’ of a functional depends not only on the input and output variables, but on the steepness of the input-output mapping within the domain of the dataset. Our results suggest that the hypersensitivity we identify may not affect applications where system-specific datasets are used,

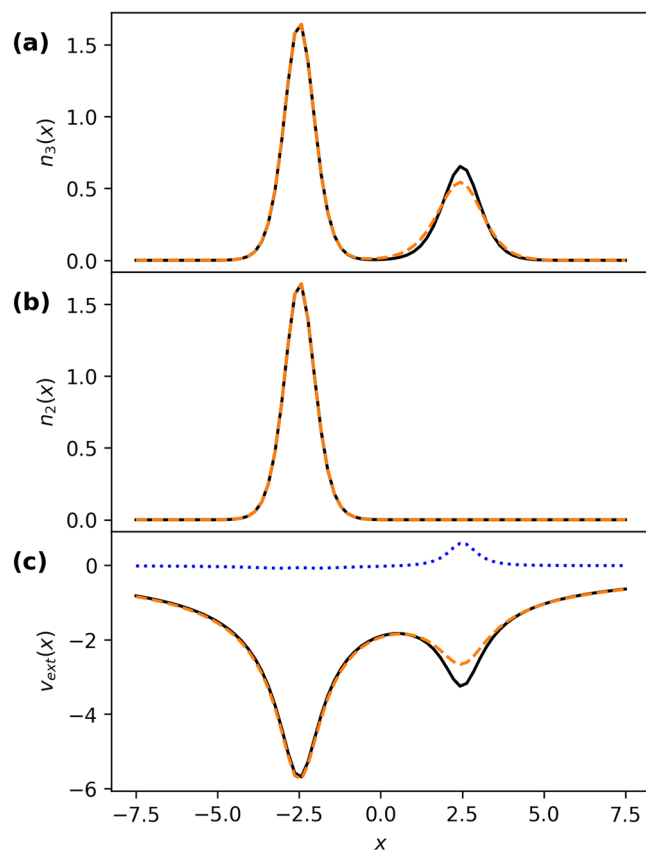


Fig. 3 | Two systems (one indicated by solid black lines, the other by dashed orange lines—details in the main text) that illustrate how dependence on the properties of the $(N + 1)$ -electron system can lead to hypersensitivity in our second class of density functionals. They are both two-electron molecular systems wherein both electrons are confined to the left well (b). In this region, the difference between the external potentials of the two systems (c) is close to 0, leading to similar electron densities. However, the corresponding three-electron densities (a) are spread between both wells, and the potentials differ significantly in the right well. This leads to large differences between the three-electron energies and quantities that depend on them, such as the affinity and fundamental gap. As such, the changes in these properties of the two-electron systems are disproportionate to the small change in the two-electron density.

such as molecular dynamics. System-specific datasets were used in ref. 26, which may partially explain why $E[n]$ was learned more accurately than $E[v_{\text{ext}}]$ —we expect that the hypersensitivity of $E[n]$ did not arise as the data did not explore the hypersensitive regions of the mapping. In contrast, datasets containing a wide variety of systems—such as those used in ref. 19—are more likely to cover the hypersensitive regions of the mapping. As a consequence, the effect of hypersensitivity is evident in the results of ref. 19, explaining a significant portion of the error in their ML approximations to constant-shift dependent density functionals ($E[n]$ and $E_{\text{ext}}[n]$).

Results

Data generation

To obtain our training and testing data, we start by defining 1D molecular potentials of the form

$$v_{\text{ext}}(x) = -\frac{Z_l}{\sqrt{(x + \frac{R}{2})^2 + \alpha_l^2}} - \frac{Z_r}{\sqrt{(x - \frac{R}{2})^2 + \alpha_r^2}}. \quad (1)$$

(Atomic units are used throughout.) These potentials have the soft-core Coulomb form, which uses a softening parameter, α (since the 1D bare-Coulomb form has a non-integrable singularity)^{40,41}. The reference level is

fixed, as $v_{\text{ext}}(|x| \rightarrow \infty) = 0$ for all potentials. We construct two datasets, each comprising one hundred potentials. For the first dataset, we use a fixed internuclear separation of $R = 5$ and sample the remaining parameters uniformly in the following ranges: $3 \leq Z_l \leq 3.5$, $1.25 \leq \frac{1}{\alpha_l} \leq 1.75$, $1.5 \leq Z_r \leq 2$, $1.25 \leq \frac{1}{\alpha_r} \leq 1.75$. The external potentials are represented on a spatial grid of 75 points from $x = -7.5a_0$ to $x = 7.5a_0$, ensuring that all relevant physical quantities are converged to within 10^{-5} Ha. For the second dataset, we fix both Z_l and Z_r at 2, and α_l and α_r at 1. We sample R uniformly in the range $2 \leq R \leq 7$. For this dataset, we use a spatial grid of 70 points from $x = -10a_0$ to $x = 10a_0$ (due to the greater internuclear separations and our infinite boundary conditions, a grid with a larger spatial extent is required to achieve the same level of convergence).

For each external potential, we model systems of two electrons in a spin singlet (spin up and spin down), interacting via the soft-core Coulomb potential with $\alpha = 1$, i.e. $U(x, x') = ((x - x')^2 + 1)^{-\frac{1}{2}}$. We numerically solve the many-body Schrödinger equation using the iDEA code³⁹ to find the exact total energy, E , and the many-body wavefunction, from which we obtain the exact many-body density, $n(x)$. We calculate the value of the HK universal functional for each system by subtracting the external potential energy from the total energy: $F = E - E_{\text{ext}}$ where $E_{\text{ext}} = \int_{-\infty}^{\infty} v_{\text{ext}}(x)n(x)dx$. To obtain the remaining quantities, we first solve the Schrödinger equation for the corresponding one- and three-electron systems. (For the three-electron systems, our electron spins are up, up, and down.) We then calculate the first ionisation energy, $I = E_1 - E_2$, the electron affinity, $A = E_2 - E_3$ and the fundamental gap, $E_g = I - A$.

Dataset 1

We now illustrate the hypersensitivity effects described above using examples from our first dataset, where systems with similar electron densities differ substantially in their electronic properties.

In this dataset, variations in the nuclear charges and softening parameters of the external potentials lead to near-constant differences between potentials. An example of this is shown in Figure 2. The molecular potentials in panel (c) are defined by $Z_l = 3.44$, $\alpha_l = 0.65$, $Z_r = 1.64$, $\alpha_r = 0.63$ (solid black line) and $Z_l = 3.10$, $\alpha_l = 0.62$, $Z_r = 1.52$, $\alpha_r = 0.63$ (dashed orange line). Their difference, $\Delta v_{\text{ext}}(x)$, is also shown as a blue dotted line. $\Delta v_{\text{ext}}(x)$ is non-constant, peaking at $x = \pm R/2$ and tending to 0 as $|x| \rightarrow \infty$ (as necessitated by the fixed reference level). However, over the domain $x = -3.9$ to $x = -1$ (right), which contains over 97% of the electron density, it is approximately constant. As a result, the electron densities corresponding to the two potentials, $n(x)$ and $\tilde{n}(x)$, are very similar. This is visually apparent in panel (b), but we also attain a numerical measure using the L^2 norm distance,

$$\|\tilde{n}(x) - n(x)\|_{L^2} = \sqrt{\int_{-\infty}^{\infty} [\tilde{n}(x) - n(x)]^2 dx}, \quad (2)$$

and find that $\|\tilde{n}(x) - n(x)\|_{L^2} = 0.004$. Despite their proximity in density space, the difference in total energy is significant at $|\tilde{E} - E| = 0.54$. As the total energy responds significantly to the near-constant shift but the electron density does not, the total energy functional, $E[n]$, is hypersensitive. In contrast, the value of the HK universal functional F differs by only 0.004 for the two systems, so $F[n]$ is not hypersensitive, as anticipated. Since $E = F + E_{\text{ext}}$, this analysis shows that the constant-shift dependent energy component, E_{ext} , is responsible for the significant difference in total energy, not F . The near-constant shift also leads to large changes in I (0.27) and A (0.21), because they are constant shift dependent.

The rate at which a functional $Q[n]$ changes with respect to the input can be approximated by the finite difference quotient

$$\left| \frac{\Delta Q}{\Delta n} \right| = \frac{|Q[\tilde{n}] - Q[n]|}{\|\tilde{n}(x) - n(x)\|_{L^2}}. \quad (3)$$

$|\Delta Q/\Delta n| \gg 1$ indicates hypersensitivity, as $|\Delta n|$ is small relative to $|\Delta Q|$. For the systems shown in Figure 2, $|\Delta E/\Delta n| = 122$ and $|\Delta F/\Delta n| = 0.96$.

Figure 3 shows a pair of systems from our dataset that exemplifies the hypersensitivity of our second class of functionals, i.e. any functional of the N -electron density that depends on the properties of the corresponding $(N + 1)$ -electron system, such as $A[n]$ and $E_g[n]$. The potentials (c) are defined by $Z_l = 3.49$, $\alpha_l = 0.64$, $Z_r = 1.52$, $\alpha_r = 0.77$ (solid black line) and $Z_l = 3.39$, $\alpha_l = 0.63$, $Z_r = 1.52$, $\alpha_r = 0.59$ (dashed orange line). The two-electron densities (b) are concentrated in the left wells, where the difference between the potentials (blue dotted line) is close to 0. As a result, they essentially experience the same potential and differ by only 0.001 in L^2 . However, the difference between the potentials is significant in the right well, a region in which both systems have significant three-electron density (a). Consequently, E_g and A for the two systems differ by 0.39 and 0.32, respectively, with finite difference quotients of $|\Delta E_g/\Delta n| = 318$ and $|\Delta A/\Delta n| = 259$. In contrast, $|\Delta F/\Delta n| = 0.83$.

Dataset 2

Figure 4 shows an example pair of systems from our second dataset, in which only the internuclear separation, R , is varied. The systems represented by solid black lines and dashed orange lines correspond to $R = 4$ and $R =$

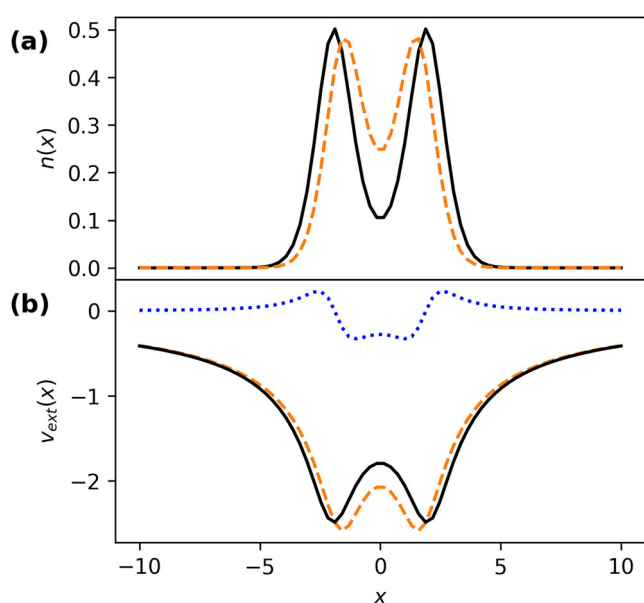
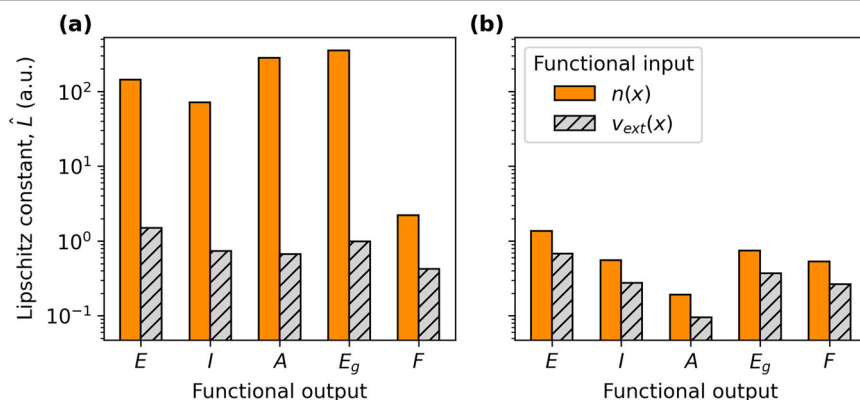


Fig. 4 | Two systems from our second dataset; the variation in the potentials in this dataset is such that our density functionals do not exhibit hypersensitivity. In this dataset, only the internuclear separation between the atomic wells, R , is varied—the external potential represented by a solid black line in (b) corresponds to $R = 4$, whereas the dashed orange line corresponds to $R = 3.31$. The blue dashed line shows the difference between the potentials, which varies significantly over the domain of the electron density (a). The electron density responds significantly to this change in potential, in contrast to the examples of hypersensitivity in Figs. 2 and 3.

Fig. 5 | The local Lipschitz constant, $\hat{L}$, is plotted for all density functionals (orange bars) and potential functionals (grey hatched bars), enabling a comparison of how rapidly each varies with respect to its input. As $\hat{L}$ represents (a lower bound on) the maximum rate of change within a restricted domain, it has different values in our two datasets. In our first dataset (a), the hypersensitivity effects we have identified manifest as values of $\hat{L}$ on the order of 10^2 for $E[n]$, $I[n]$, $A[n]$ and $E_g[n]$. For $F[n]$, which does not suffer from these effects, $\hat{L}$ is two orders of magnitude smaller. In our second dataset (b), none of the density functionals exhibit hypersensitivity, as expected. Moreover, in both datasets, $\hat{L}$ is small for all potential functionals, indicating the absence of hypersensitivity.



3.31, respectively. In contrast to Figs. 2 and 3, the difference between the two external potentials (blue dotted line) has significant spatial variation over the domain of the electron density. As a consequence, the two systems have significantly different electron densities ($L^2 = 0.36$), and the hypersensitivity issues do not arise: for example, $|\Delta E/\Delta n| = 0.56$ and $|\Delta A/\Delta n| = 0.08$.

Estimating the local Lipschitz constant

We now extend our analysis to each dataset as a whole, considering all pairs of systems i and j , $i \neq j$. For each pair, we compute the finite difference quotient $|\Delta Q/\Delta n|$ for all quantities. We then take the maximum finite difference quotient for each quantity as an empirical estimate of the local Lipschitz constant, $\hat{L}^{42}$.

$$\hat{L} = \max_{i \neq j} \frac{|Q[n^i] - Q[n^j]|}{\|n^i(x) - n^j(x)\|_{L^2}} \quad (4)$$

$\hat{L}$ is a lower bound on the maximum rate of change within the domain of the dataset. It is commonly used as a measure of learning complexity⁴³, e.g. to attain bounds for the worst-case prediction error⁴². We use it to characterise the degree of hypersensitivity of each functional in our datasets.

In Figure 5 we plot $\hat{L}$ for all density functionals (solid orange bars) in our datasets. In the first dataset (a), $\hat{L} \gg 1$ for $E[n]$, $I[n]$, $A[n]$ and $E_g[n]$, providing numerical confirmation of the hypersensitivity described above. (b) shows that, as expected, the same density functionals are *not* hypersensitive in our second dataset, as $\hat{L}$ is on the order of 10^{-1} – 10^0 . We also plot $\hat{L}$ for each quantity as a functional of the external potential (grey hatched bars), showing that the hypersensitivity does not arise for potential functionals in either dataset. This is because the small changes in electron density in general correspond to significant changes in the external potential (for example, see Figs. 2 and 3), so $|\Delta Q/\Delta v_{\text{ext}}|$ is consistently small.

Machine learning

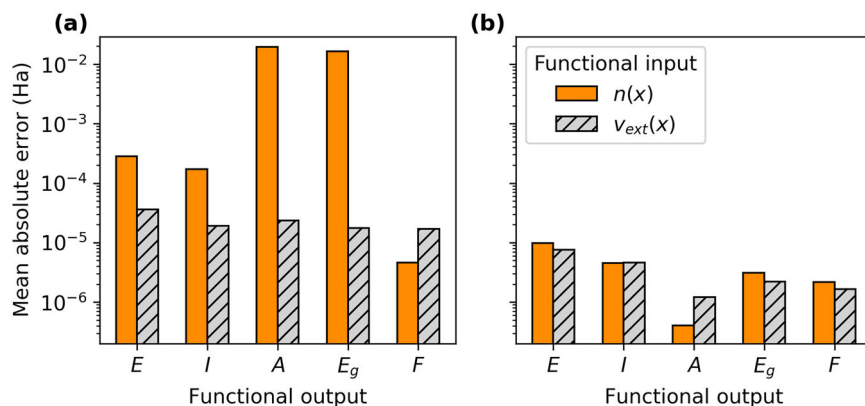
The ML model we use is KRR; see section ‘Kernel Ridge Regression’ for details. To assess the model’s accuracy, we use the mean absolute error (MAE)

$$\text{MAE} = \frac{1}{S} \sum_{i=1}^S |Q^{\text{ML}}[n^i] - Q[n^i]|, \quad (5)$$

where n^i are the inputs in the test set, of which there are $S = 10$ (note that since we use nested k-fold cross-validation, all 100 samples are used for testing at some point).

Figure 6 shows the MAE in the KRR predictions of each physical quantity when using the electron density (solid orange bars) and the external potential (hatched grey bars) as input. First, we will discuss the results for Dataset 1 (the dataset in which the density functionals exhibit hypersensitivity), which are shown in panel (a). Amongst the density functionals, the error is one to three orders of magnitude greater for the functionals we

Fig. 6 | The MAE in our ML approximations to density functionals (orange bars) and potential functionals (grey hatched bars). Comparing to Fig. 5, it is apparent that the ML accuracy is closely related to the Lipschitz constant, i.e. rapidly-varying functionals are approximated less accurately. Our hypotheses account for three key aspects of these results: for the first dataset (a), the ML approximations to the hypersensitive functionals $E[n]$, $I[n]$, $A[n]$ and $E_g[n]$ are less accurate than the $F[n]$ approximation, as $F[n]$ is neither constant-shift dependent nor does it depend on the properties of the $(N + 1)$ -electron system. In the second dataset (b), none of the functionals exhibit hypersensitivity, and hence the ML approximations are similarly accurate. Discrepancies in ML accuracy do not occur when the external potential is used as the input, because the potential functionals are not hypersensitive in either dataset.



identified as hypersensitive ($E[n]$, $I[n]$, $A[n]$ and $E_g[n]$). For example, although the error in $E[n]$ was very low in absolute terms (0.28 mHa or 0.18 kcal/mol) (due to the quality of data provided by our 1D systems), it was $61 \times$ greater than the error in $F[n]$ (4.65×10^{-3} mHa or 2.9×10^{-3} kcal/mol). This means that the vast majority of the error in $E[n]$ would be avoided by calculating $E_{\text{ext}}[n]$ via the closed-form expression rather than attempting to machine-learn it. The errors in $A[n]$ and $E_g[n]$ were the greatest by a significant margin at 19.4 and 16.4 mHa, respectively—around $4000 \times$ greater than the error in $F[n]$. This is likely because these quantities, due to their dependence on the $(N + 1)$ -electron energy, are the most sensitive in the dataset (note that they have the greatest values of $\hat{L}$ in Fig. 5).

E , I , A and E_g are all predicted with significantly greater accuracy from the external potential, because the hypersensitivity issue is avoided. Specifically, we observe about a 10-fold decrease in error for E and I , and a 1000-fold decrease for A and E_g , when using the external potential as input rather than the electron density. This confirms that the pathologies of the density $\rightarrow$ output mapping are the source of error, not the output quantities per se.

These results are in striking agreement with ref. 19, where ML was employed to predict the total energy, external potential energy, exchange, correlation and kinetic energy of electrons in random 2D potentials. When using the electron density as input to their ML models, the authors observed significantly greater error for the total energy and external potential energy relative to the exchange, correlation, and kinetic energy; the MAE was typically an order of magnitude greater. We posit that this is explained by hypersensitivity, as E and E_{ext} are constant shift-dependent and the other quantities are not. However, the authors attributed it to the fact that their dataset contained a larger range of values for E and E_{ext} compared to the other quantities. As this explanation relies solely on the target property, it implies that the discrepancy in predictive accuracy should not depend on the input to the ML model. However, when the authors used the external potential as the input, the error was more balanced between the different properties. Likewise, we find that when using the external potential as the input, our KRR model is equally capable of predicting constant-shift dependent quantities and constant-shift invariant quantities; see Fig. 6a. Our hypersensitivity hypothesis explains why the discrepancy occurs for density functionals and not potential functionals. In addition, as the authors of ref. 19 used NNs, while we used KRR, this implies the difficulties of learning these density functionals are independent of the ML method. Figure 6b shows the results for Dataset 2. This dataset is system-specific (only the internuclear separation is varied), and all observables vary gradually with respect to the electron density. As the hypersensitivity of $E[n]$, $I[n]$, $A[n]$ and $E_g[n]$ is intentionally avoided in this dataset, the resulting ML approximations are far more accurate. This highlights that the difficulty of learning a functional depends critically on its local behaviour within the domain of the data. Ref. 44 similarly found that ML was more

challenging in the region of a quantum phase transition, where the functional is non-analytic. Our results suggest that the hypersensitivity effects we identified may not arise for applications where system-specific datasets are used, such as molecular dynamics simulations. This also explains why in ref. 26, in which only system-specific datasets were used, the external potential was not a better input than the electron density for predicting the total energy; since each dataset only consisted of a single molecule, we hypothesise that the hypersensitivity of $E[n]$ did not arise. However, while system-specific functionals are useful for certain applications, the central aim of ML-DFT is to train a model that can reliably generalise across diverse physical systems¹⁷. This requires a broad and varied dataset, so the hypersensitivity is much more likely to arise. As ref. 19 used a very broad dataset, it is not surprising that these effects were encountered. This is why an alternative solution, such as using the external potential as input, is required.

Although our findings reveal a distinct advantage to using the potential rather than the density in specific contexts, we note that there may be scenarios in which the potential is the inferior input. Like density functionals, potential functionals may also exhibit hypersensitivity, non-analyticity, and other pathologies that make ML challenging. Moreover, for some systems, the potential to observable mapping is one-to-many, such as an open-shell atom or molecule with orbital degeneracy, for which the external potential maps to several distinct, degenerate densities. Hence, when selecting which input to use, the uniqueness and numerical behaviour of the input-output mapping must be considered. In this paper, we find that for a variety of important observables with well-defined mappings in both DFT and PFT, hypersensitivity affecting the density functionals is absent in the corresponding potential functionals, demonstrating a clear advantage to the potential-based approach in the contexts we outline.

Discussion

Although in principle, ML models such as KRR and NNs can learn the mapping from the electron density to any electronic property, some physical quantities can be highly sensitive to small variations in the electron density, which makes accurate ML far more challenging. In this paper, we identified two sources of this 'hypersensitivity' within DFT, each affecting an entire class of observables: (1) quantities that vary when a constant shift is applied to the external potential and (2) functionals of the N -electron density that depend on the properties of the $(N + 1)$ -electron system.

The hypersensitivity is caused by changes to the external potential of a system, $\Delta v_{\text{ext}}(\mathbf{r})$, that have only a negligible impact on its electron density but a disproportionately large impact on the target quantity. This occurs in the first class of functionals when $\Delta v_{\text{ext}}(\mathbf{r})$ is near-constant over the region where the electron density is concentrated (which is possible even when the reference level of the external potential is fixed). Hypersensitivity arises in the second class of quantities when $\Delta v_{\text{ext}}(\mathbf{r})$ has

small spatial variation over the domain of the N -electron density but is significant in the domain of the $(N + 1)$ -electron density. In both cases, $\Delta v_{\text{ext}}(\mathbf{r})$ itself is not small, so the observables are not hypersensitive to these changes in the potential.

Using exactly solvable one-dimensional model systems that form the basis of our ML datasets, we provide examples of these effects, where pairs of systems with near-identical electron densities differ substantially in their total energy (Class 1), first ionisation energy (Class 1), electron affinity (Classes 1 and 2), or fundamental gap (Class 2). We find that with a fixed dataset size and training procedure, KRR models exhibit significantly greater error when applied to these functionals compared to the HK universal functional, which does not suffer from the sources of hypersensitivity we identified. Both in this research and previous work which employed NNs¹⁹, the external potential energy, $E_{\text{ext}}[n]$, has proven to be the most challenging component of the total energy to learn as a density functional. Our analysis explains this surprising result; although $E_{\text{ext}}[n]$ has a simple, closed form, it is hypersensitive when expressed as a pure density functional and hence highly complex from an ML standpoint. Since E_{ext} can be calculated via a closed-form expression, our findings highlight that it is preferable to machine-learn $F[n]$ and then add the external potential energy contribution, rather than attempting to learn the mapping from the density to the total energy directly.

For all observables in Classes 1 and 2, we found that the hypersensitivity can be avoided by using the external potential as input to the ML model. With this approach, we attained considerably more accurate predictions of the total energy, first ionisation energy, electron affinity and fundamental gap. This also explains why, in ref. 19, the error was more balanced between constant-shift dependent and constant-shift invariant quantities when using the external potential as input than when using the electron density.

We also provide a second example dataset which restricts the variations between systems so that density functionals in Classes 1 and 2 do *not* exhibit hypersensitivity, and are consequently learned with far greater accuracy. Note that it is not that these functionals are not hypersensitive on the whole, but rather the dataset does not explore the hypersensitive regions of the mapping. This highlights the dataset-dependent nature of the problem and raises the question: in which contexts is it likely to arise? As our second dataset is system-specific (containing only a single 1D molecule), we hypothesise that the hypersensitivity is unlikely to arise in system-specific datasets; this would also explain why it does not seem to have arisen in ref. 26. On the other hand, in a typical ML-DFT dataset that contains a broad range of physical systems, there are more likely to be near-constant differences between potentials (for example), and hence hypersensitivity.

Although our numerical examples are one-dimensional, our arguments readily extend to two- and three-dimensional systems. Moreover, while we have focused on two specific sources of hypersensitivity, there are likely to be other cases that follow the same pattern, i.e. the observable is sensitive to some change in the potential but the density is not. As such, our solution of using the external potential as input to the ML model could also circumvent other possible sources of hypersensitivity in density functionals that we have not identified in this paper.

Future research could apply our methodology to identify hypersensitivity in realistic ML-DFT datasets and to determine whether our proposed potential-based approach could be beneficial. As such, our findings offer the prospect of reliable ML models within DFT and PFT for a variety of important physical quantities that may, in turn, lead to more robust computational modelling within materials science.

Methods

Kernel ridge regression

KRR is a non-linear regression model with regularisation²². To approximate a functional $Q[n]$, KRR uses a kernel function $k[n^i, n^j]$ that measures the similarity of two inputs n^i and n^j . We use the Gaussian kernel, $k[n^i, n^j] = \exp(-||n^i - n^j||_{L^2}^2 / (2\sigma^2))$, a standard choice in ML-

DFT^{20,22,26,45}. For a given input, n , the predicted value, $Q^{\text{ML}}[n]$, is calculated as a weighted sum of the similarity scores of n with each n^j in the training set⁴⁵, of which there are M , i.e.

$$Q^{\text{ML}}[n] = \sum_{j=1}^M \alpha^j k[n, n^j]. \quad (6)$$

The weights $\alpha = (\alpha^1, \dots, \alpha^M)^T$ are found by minimising the cost function

$$C(\alpha) = \sum_{j=1}^M (Q^{\text{ML}}[n^j] - Q[n^j])^2 + \lambda ||Q||_{\mathcal{H}_K}^2, \quad (7)$$

where $||Q||_{\mathcal{H}_K}^2 = \alpha \cdot \mathbb{K} \alpha$ is the reproducing kernel Hilbert space (RKHS) norm⁴⁶ and $\mathbb{K}$ is a matrix with elements $K_{ij} = k[n^i, n^j]$. The greater the value of the regularisation parameter λ , the more emphasis is placed on minimising the RKHS norm, which can prevent the model from overfitting the training data.

The cost function is minimised by solving $\alpha = (\mathbb{K} + \lambda \mathbb{I})^{-1} \mathbf{Q}$, where $\mathbb{I}$ is the identity matrix and $\mathbf{Q} = (Q^1, \dots, Q^M)^T$ are the outputs (or ‘labels’) in the training set⁴⁵. Note that both λ and σ are hyperparameters, meaning their values are chosen explicitly before training. The small number of hyperparameters in KRR makes optimising (or ‘tuning’) them simpler than in more complex models such as NNs. Training is also easier for KRR due to the closed-form solution, which finds the global minimum of the loss function. In contrast, there is no such guarantee for NNs, which are typically trained via stochastic optimisation methods. As such, we employ KRR in this paper as the technique is less dependent upon user input than NNs, offering a fairer demonstration of our hypotheses. However, we note that our findings are not specific to KRR, as similar results have been reported in previous work using NNs; see ref. 19 (discussed in section ‘Results’).

To tune and test our KRR models, we use nested k-fold cross-validation in combination with grid searches⁴⁷. The nested validation comprises 10 outer folds and 9 inner folds with an 80-10-10 train-validation-test split, providing robust measures of model performance both at the validation and testing stages. Previous studies have shown that KRR can learn highly accurate DFAs from a small number of training samples²¹, and may outperform NNs in the small-data regime⁴⁷.

Data availability

The datasets generated and analysed during the current study are openly available in the research data repository of the University of York, <https://doi.org/10.15124/be174a35-5ae0-4758-85f1-d21d7cffdeba>.

Code availability

The underlying code for this study is available in the iDEA-org GitHub repository and can be accessed via this link <https://github.com/iDEA-org/iDEA>.

Received: 18 February 2026; Accepted: 4 June 2026;

Published online: 19 June 2026

References

1. Becke, A. D. Perspective: fifty years of density-functional theory in chemical physics. *J. Chem. Phys.* **140**, 18A301 (2014).
2. Hohenberg, P. & Kohn, W. Inhomogeneous electron gas. *Phys. Rev.* **136**, B864–B871 (1964).
3. Kohn, W. Nobel lecture: electronic structure of matter-wave functions and density functionals. *Rev. Mod. Phys.* **71**, 1253 (1999).
4. Levy, M. Universal variational functionals of electron densities, first-order density matrices, and natural spin-orbitals and solution of the v-representability problem. *Proc. Natl. Acad. Sci.* **76**, 6062–6065 (1979).
5. Lieb, E. H. Density functionals for coulomb systems. *Int. J. quantum Chem.* **24**, 243–277 (1983).

6. Levy, M. & Perdew, J. P. The constrained search formulation of density functional theory. In *Density Functional Methods in Physics* 11–30 (Springer, 1985).
7. Alharbi, F. H. Kinetic energy density functional: advances, challenges, and future directions. *Electron. Struct.* **8**, 023001 (2026).
8. Kohn, W. & Sham, L. J. Self-consistent equations including exchange and correlation effects. *Phys. Rev.* **140**, A1133–A1138 (1965).
9. Wu, J., Pun, S.-M., Zheng, X. & Chen, G. Construct exchange-correlation functional via machine learning. *J. Chem. Phys.* **159**, 090901 (2023).
10. Kumar, S. et al. Accurate parameterization of the kinetic energy functional. *J. Chem. Phys.* **156**, 024110 (2022).
11. Perdew, J. P. Density functional theory and the band gap problem. *Int. J. Quantum Chem.* **28**, 497–523 (1985).
12. Godby, R. W., Schlüter, M. & Sham, L. J. Self-energy operators and exchange-correlation potentials in semiconductors. *Phys. Rev. B* **37**, 10159–10175 (1988).
13. Pilania, G., Wang, C., Jiang, X., Rajasekaran, S. & Ramprasad, R. Accelerating materials property predictions using machine learning. *Sci. Rep.* **3**, 2810 (2013).
14. Kolb, B., Lentz, L. C. & Kolpak, A. M. Discovering charge density functionals and structure-property relationships with PROPhet: a general framework for coupling machine learning and first-principles methods. *Sci. Rep.* **7**, 1192 (2017).
15. Nagai, R., Akashi, R. & Sugino, O. Completing density functional theory by machine learning hidden messages from molecules. *npj Comput. Mater.* **6**, 43 (2020).
16. Kirkpatrick, J. et al. Pushing the frontiers of density functionals by solving the fractional electron problem. *Science* **374**, 1385–1389 (2021).
17. Akashi, R., Sogal, M. & Burke, K. Can machines learn density functionals? Past, present, and future of ML in DFT. <https://arxiv.org/abs/2503.01709> (2025).
18. Chandrasekaran, A. et al. Solving the electronic structure problem with machine learning. *npj Comput. Mater.* **5**, 22 (2019).
19. Ryczko, K., Strubbe, D. A. & Tamblyn, I. Deep learning and density-functional theory. *Phys. Rev. A* **100**, 022512 (2019).
20. Snyder, J. C., Rupp, M., Hansen, K., Müller, K.-R. & Burke, K. Finding density functionals with machine learning. *Phys. Rev. Lett.* **108**, <https://doi.org/10.1103/PhysRevLett.108.253002> (2012).
21. Snyder, J. C. et al. Orbital-free bond breaking via machine learning. *J. Chem. Phys.* **139**, <https://doi.org/10.1063/1.4834075> (2013).
22. Li, L. et al. Understanding machine-learned density functionals. *Int. J. Quantum Chem.* **116**, 819–833 (2016).
23. Li, L., Baker, T. E., White, S. R. & Burke, K. Pure density functional for strong correlation and the thermodynamic limit from machine learning. *Phys. Rev. B* **94**, <https://doi.org/10.1103/PhysRevB.94.245129> (2016).
24. 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**, <https://doi.org/10.1021/acs.jctc.0c00580> (2020).
25. Bogojeski, M., Vogt-Maranto, L., Tuckerman, M. E., Müller, K.-R. & Burke, K. Quantum chemical accuracy from density functional approximations via machine learning. *Nat. Commun.* **11**, 5223 (2020).
26. Brockherde, F. et al. Bypassing the Kohn-Sham equations with machine learning. *Nat. Commun.* **8**, 872 (2017).
27. Ray, P., Choudhary, K. & Kalidindi, S. R. Lean CNNs for mapping electron charge density fields to material properties. *Integrating Mater. Manuf. Innov.* **14**, 1–13 (2025).
28. del Rio, B. G., Phan, B. & Ramprasad, R. A deep learning framework to emulate density functional theory. *npj Comput. Mater.* **9**, 158 (2023).
29. Chen, F. et al. Towards universal material property prediction with deep learning and single-descriptor electronic density. <https://arxiv.org/abs/2510.13207> (2025).
30. Moreno, J. R., Flick, J. & Georges, A. Machine learning band gaps from the electron density. *Phys. Rev. Mater.* **5**, 083802 (2021).
31. Gedeon, J. et al. Machine learning the derivative discontinuity of density-functional theory. *Mach. Learn. Sci. Technol.* <https://doi.org/10.1088/2632-2153/ac3149> (2021).
32. Mills, K., Spanner, M. & Tamblyn, I. Deep learning and the schrödinger equation. *Phys. Rev. A* **96**, <https://doi.org/10.1103/PhysRevA.96.042113> (2017).
33. Nigam, J., Smidt, T. & Dusson, G. Machine learning electronic structure and atomistic properties from the external potential. <https://arxiv.org/abs/2602.15345> (2026).
34. Englert, B.-G. & Schwinger, J. Statistical atom: some quantum improvements. *Phys. Rev. A* **29**, 2339 (1984).
35. Yang, W., Ayers, P. W. & Wu, Q. Potential functionals: dual to density functionals and solution to the v -representability problem. *Phys. Rev. Lett.* **92**, 146404 (2004).
36. Cangi, A., Gross, E. & Burke, K. Potential functionals versus density functionals. *Phys. Rev. A At. Mol. Opt. Phys.* **88**, 062505 (2013).
37. Kornowski, G., Hanneke, S. & Kontorovich, A. Near-optimal learning with average hölder smoothness. *Adv. Neural Inf. Process. Syst.* **36**, 21135–21157 (2023).
38. Martinetto, V., Shah, K., Cangi, A. & Pribram-Jones, A. Inverting the Kohn-Sham equations with physics-informed machine learning. *Mach. Learn.: Sci. Technol.* **5**, 015050 (2024).
39. Hodgson, M. J. P., Ramsden, J. D., Chapman, J. B. J., Lillystone, P. & Godby, R. W. Exact time-dependent density-functional potentials for strongly correlated tunneling electrons. *Phys. Rev. B* **88**, 241102 (2013).
40. Gordon, A., Santra, R. & Kärtner, F. X. Role of the coulomb singularity in high-order harmonic generation. *Phys. Rev. A* **72**, 063411 (2005).
41. Loudon, R. One-dimensional hydrogen atom. *Proc. R. Soc. A Math., Phys. Eng. Sci.* **472**, Article-ID-20150534 (2016).
42. Huang, J. W., Roberts, S. J. & Callies, J.-P. On the sample complexity of Lipschitz constant estimation. *Trans. Mach. Learn. Res.* <https://openreview.net/forum?id=UlaIYAHdBH> (2023).
43. Ashlagi, Y., Gottlieb, L.-A. & Kontorovich, A. Functions with average smoothness: structure, algorithms, and learning. *J. Mach. Learn. Res.* **25**, 117:1–117:54 (2024).
44. Moreno, J. R., Carleo, G. & Georges, A. Deep learning the Hohenberg-Kohn maps of density functional theory. *Phys. Rev. Lett.* **125**, <https://doi.org/10.1103/PhysRevLett.125.076402> (2020).
45. Vu, K. et al. Understanding kernel ridge regression: common behaviors from simple functions to density functionals. *Int. J. Quantum Chem.* **115**, 1115–1128 (2015).
46. Kanagawa, M., Hennig, P., Sejdinovic, D. & Sriperumbudur, B. K. Gaussian processes and kernel methods: a review on connections and equivalences. <https://arxiv.org/abs/1807.02582> (2018).
47. Hansen, K. et al. Assessment and validation of machine learning methods for predicting molecular atomization energies. *J. Chem. Theory Comput.* **9**, 3404–3419 (2013).

Acknowledgements

This work made use of the facilities of the N8 Centre of Excellence in Computationally Intensive Research (N8 CIR) provided and funded by the N8 research partnership and EPSRC (Grant No. EP/T022167/1). The Centre is coordinated by the Universities of Durham, Manchester, and York. The funder played no role in study design, data collection, analysis and interpretation of data, or the writing of this manuscript. P.J.H. acknowledges funding from UKRI (grant References UKRI2710, EP/W030489/1 and EP/X035891/1).

Author contributions

L.A. performed all numerical calculations, including the 1D simulations, Lipschitz analysis and the full machine learning pipeline. Under the supervision of M.J.P.H. and R.L., L.A. identified the sources of numerical hypersensitivity. M.J.P.H. and J.W. conceptualised the project, with

M.J.P.H. directing the research and proposing the focus on constant shifts in the potential. M.J.P.H. and J.W. authored the iDEA software, implementing substantial improvements that directly enabled this work; J.W. further extended iDEA with GPU acceleration for many-body interacting systems. R.L. and P.J.H. provided insights into DFT and ML, proposed conceptual ideas, and contributed to discussions throughout the project. L.A. and M.J.P.H. drafted the manuscript, while all authors contributed to the review and editing of the manuscript and approved the final version.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to M. J. P. Hodgson.

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

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

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

© The Author(s) 2026