

Image super-resolution inspired electron density prediction

Received: 6 March 2024

Accepted: 7 May 2025

Published online: 23 May 2025

Chenghan Li¹, Or Sharir², Shunyue Yuan² & Garnet Kin-Lic Chan¹✉

Predicting ground-state electron densities of chemical systems has recently received growing attention in machine learning quantum chemistry, given their fundamental importance as highlighted by the Hohenberg-Kohn theorem. Drawing inspiration from the domain of image super-resolution, we view the electron density as a 3D grayscale image and use a convolutional residual network to transform a crude and trivially generated guess of the molecular density into an accurate ground-state quantum mechanical density. Here we show that this model produces more accurate predictions than all prior density prediction approaches. Due to its simplicity, the model is directly applicable to unseen molecular conformations and chemical elements. We show that fine-tuning on limited new data provides high accuracy even in challenging cases of exotic elements and charge states.

Computing electronic properties of the ground-state of molecules and materials is a central task in chemistry and materials science. Even within the common mean-field formulation known as Kohn-Sham density functional theory (DFT)¹, the computations can be prohibitively expensive for many applications, for example, in molecules with many atoms, or when many geometric configurations must be considered. To reduce this cost, machine learning techniques have risen to the fore. These substitute physics-based calculations with a machine learning model, turning the electronic structure problem into a regression task to map a chemical description to the electronic state and properties^{2,3}.

The most common machine learning models for electronic structure target the electronic energy. In one approach, known as machine learning force fields, the inputs are the molecular geometry and element types (and sometimes atomic energies)^{4–15}. Another common class of models requires first carrying out a simpler molecular electronic structure calculation (e.g., semi-empirical tight-binding density functional theory) to generate an input to a model that predicts the energy of a more sophisticated electronic structure method on the same molecule^{16–18}. In both approaches, much effort has focused on improving the featurization and model architectures, for example, to explicitly incorporate the equivariance and invariance that arise from physical symmetries¹³. These ML models have achieved considerable success, greatly reducing the time to obtain quantum mechanical quality energies. However, the energy is only a single facet

of the electronic structure, and machine learning other properties of the quantum state has been comparatively less explored¹⁹.

Within density functional theory, due to the Hohenberg-Kohn theorem²⁰, the ground-state electron density formally yields all properties of the associated quantum state. Machine learning the electron density has thus attracted recent attention^{17,21–42}. Like their counterparts in force fields, ML density models have been explored with different sources of input data, ranging from ones that just take in the geometry and element types^{21–25,27–33,35–42} to ones that first require a quantum mechanical calculation^{17,34}.

We will demonstrate that a powerful way to learn an accurate molecular electron density is to start from its crudest physical approximation, the superposition of atomic densities, represented on a spatial grid. In fact the atomic densities can be chosen to be very crude, and serve mainly to denote the size and location of the atoms, but importantly, once the choice is made, they do not require recalculation for any molecule. After featurizing the input superposed atomic density on a coarse grid, our model learns the accurate molecular electron density on a high-resolution grid. This learning task is the 3D analog of image super-resolution, a well-established domain in computer vision^{43,44}, where crude, lossy images at low resolution are enhanced to accurate, high-resolution images. We therefore adopt a standard image processing approach, using a simple convolutional residual neural net (ResNet)⁴⁵ that does not enforce any physical symmetries. We find that we achieve state-

¹Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena 91125 CA, USA. ²Division of Engineering and Applied Sciences, California Institute of Technology, Pasadena 91125 CA, USA. ✉e-mail: gkc1000@gmail.com

Table 1 | Density prediction mean absolute error on QM9 test set

ResNet	Charge3Net ⁴¹	DeepDFT (in-/equivariant) ³⁸	OrbNet-Equi ¹⁷	DF	SAD
0.14	0.196	0.36/0.27	0.21	0.26/ 0.27	13.9

ResNet (This work) starts from the superposition of spherical atomic density (SAD) guess and uses a spatial upscaling factor of 2. The density fitting (DF) errors were computed with both the Coulomb metric and the overlap metric (numbers are given in the order Coulomb/overlap). More details and other data are in the text.

Table 2 | Electronic property errors on QM9

Quantity	Unit	ResNet (This work)	DF
E^{elec}	meV	3.0	0.0083/0.014
ϵ^{HOMO}	meV	12.0	1.3/3.8
ϵ^{LUMO}	meV	9.5	2.0/6.1
$\Delta\epsilon$	meV	11.0	1.4/4.4

Quantities defined in the text. All errors are given as the mean absolute errors. ResNet (This work) starts from the superposition of spherical atomic density (SAD) guess and uses a spatial upscaling factor of 2. The density fitting (DF) errors were computed with both the Coulomb metric and the overlap metric (numbers are given in the order Coulomb/overlap). Note that access to the properties listed in this table (E^{elec} : total electronic energy, ϵ^{HOMO} : the highest occupied molecular orbital energy, ϵ^{LUMO} : the lowest unoccupied molecular orbital energy, and $\Delta\epsilon$: the HOMO-LUMO gap) from the electron density requires one diagonalization of the Kohn-Sham matrix built with the predicted density (note, this diagonalization yields a different density to the predicted density).

of-the-art accuracy in density prediction, improving on all previous models. Furthermore, because we do not use explicit element types, our architecture generalizes, and can be easily fine-tuned, across chemical space, which we demonstrate in molecules including elements we do not train against. From the density, we compute energies and orbitals through a single diagonalization of the Kohn-Sham Hamiltonian based on our predicted density. We find that even though we do not train against these quantities, we can obtain accurate predictions from this additional diagonalization step. We propose that the strategy in this work not only provides an important step forward in the density prediction problem but also opens up a distinct class of methods to predict real-space electronic structure properties based on well-established image processing techniques. We conclude with a discussion of the future of such models, including the current limitations and potential future solutions.

Results

Density learning on the test set

In Table 1, we summarize the accuracy of our density model trained on the QM9 dataset, which contains 134K small neutral organic molecules composed of the CHONF elements⁴⁶. As is usual, the dataset is split into training (used to optimize the model parameters), validation (used to determine the end-point of training), and test (used for independent assessment) datasets, and we follow the same number of training/validation/testing points used in Ref. 38.

We consider two types of input densities for our density prediction models (see Methods for model details). The first is the superposition of spherical neutral atomic density (SAD) (similar to the atomic guess approach in NWChem⁴⁷ but without building and diagonalizing a molecular Fock), based on pre-tabulated atomic natural orbitals expanded in a Gaussian atomic basis^{48–50}. The other is of even lower (atrocious) quality, and is generated by modifying the SAD density by multiplying the Gaussian basis exponents of the atomic natural orbitals to be filled by electrons by two. This results in a highly spatially contracted density that is very poor in terms of Err_ρ but which

still preserves qualitative information about the atom positions and the relative order of atom sizes.

As shown in Table 1, starting from the SAD guess, and using a spatial upscaling factor of 2 (i.e., the predicted density is output on a grid resolution that is uniformly improved from that of the input density by a factor of 2 along each axis) our model refines the density error of the input SAD density by two orders of magnitude. (As an immediate consequence of this, using the predicted density in self-consistent-field (SCF) iterations to obtain a converged DFT density required, on average, 5.5 fewer iterations than starting from the SAD guess, or roughly 35% of the iterations). The atrocious SAD density has a worse input density error by a factor of 2 from the SAD density ($\text{Err}_\rho = 27.7\%$). However, using this input, the model achieves a 173-fold factor of improvement, and the predicted density saved 10.2 SCF cycles (49% of the iterations) on average compared to the atrocious SAD guess. The predicted error in the density is then almost the same as when using the standard SAD input (model prediction $\text{Err}_\rho = 0.16\%$ on the test set), testifying to the robustness of the model. If we spatially upscale the resolution by a factor of 4, the result is only slightly less accurate than upscaling by a factor of 2, with $\text{Err}_\rho = 0.19\%$.

To place our results in a broader perspective, we compare to the density prediction error of some recent density learning approaches on the same datasets. Charge3Net⁴¹ and DeepDFT³⁸ are equivariant models that take the molecular structure and element types as inputs, while OrbNet-Equi¹⁷ uses input information from a molecular semi-empirical tight-binding DFT calculation. We reproduce their reported errors in Table 1 and we find that we achieve a better accuracy than all these prior models.

We further compare to the error from a direct numerical fit of the target density. For this, we use Gaussian density fitting, a common numerical representation in quantum chemistry which expands the density in an auxiliary Gaussian basis⁵¹. Here, we choose the auxiliary basis to be the default even-tempered Gaussian auxiliary basis in PYSCF for our DFT computational basis, and fit using both the Coulomb metric⁵¹ and the overlap metric⁵². Our density model outperforms even this direct numerical fit to the data by a factor of 2 in terms of density mean absolute error (MAE), although it falls behind according to some metrics, as discussed later.

Energies and properties from predicted densities

Within the framework of Kohn-Sham density functional theory, the total electronic energy and orbitals can be obtained from the exact electron density via a single diagonalization of the effective Kohn-Sham Hamiltonian (one-step DFT). Although this additional diagonalization means these quantities are not obtained within a pure data-driven approach, they provide a physically interpretable metric with which to understand the quality of a density prediction model, and one-step DFT metrics have been employed in several recent density prediction models^{28,38,41,53}. It is important to note that the energy obtained from the one-step DFT process is not the same as the energy computed directly from the Kohn-Sham energy functional using the model-predicted density. The latter is challenging to compute due to the unknown kinetic energy functional in the Kohn-Sham framework. This discrepancy arises because the diagonalization step also updates the density, typically improving it (e.g., by two-fold on average on the QM9 test set).

We list our MAE errors in Table 2. The one-step DFT electronic energy error is small (well within chemical accuracy of 1 kcal/mol \approx 43 meV). The one-step DFT highest-occupied-molecular-orbital (HOMO) and lowest-unoccupied-molecular-orbital (LUMO) energies (ϵ^{HOMO} , ϵ^{LUMO}), and HOMO-LUMO gap ($\Delta\epsilon$) are less accurate than the total energy (although still within chemical accuracy). We find larger errors in the individual energy components: for the Coulomb energy MAE = 167 meV, the exchange-correlation energy MAE = 40 meV, and the Kohn-Sham kinetic energy MAE = 303 meV. This is because near the

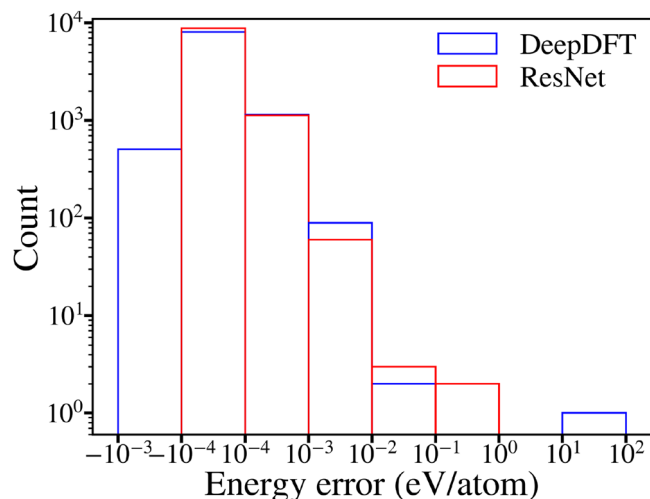


Fig. 1 | Energy error distribution from a one-step DFT procedure based on the predicted electron density. The equivariant DeepDFT result was extracted from Ref. 38. Source data are provided as a Source Data file.

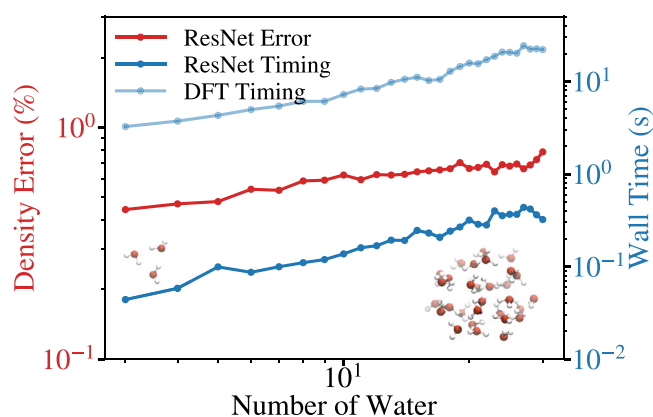


Fig. 2 | Zero-shot density prediction mean absolute error of the QM9 model and timings on a water cluster dataset. Representative molecular configurations of a three-water cluster and a thirty-water cluster are shown. The ResNet timing was performed using one A100 GPU, and the DFT timing was performed on one compute node with 28 cores [CPU model: Intel(R) Xeon(R) E5-2680 v4 @ 2.40GHz]. The computational cost of ResNet grows linearly with the system size due to the increasing box volume, where we maintain a constant margin around the actual extent of the atoms. Source data is provided as a Source Data file.

Table 3 | Transferability check of the ResNet model trained on QM9

Dataset		$\rho(\%)$	$E^{\text{elec}}(\text{meV})$
$\text{C}_7\text{O}_2\text{H}_{10}$	Zero-shot ResNet	0.65	24.5
	Fine-tuned ResNet	0.25	2.6
	From-scratch ResNet	1.5	
	SAD Guess	13.8	
	Zero-shot eqDeepDFT	1.8	
MT	Zero-shot ResNet	2.1	
	Fine-tuned ResNet	0.33	
	From-scratch ResNet	3.4	
	SAD Guess	11.9	

The errors are reported as mean absolute errors. Zero-shot means the model trained on QM9 is tested on the $\text{C}_7\text{O}_2\text{H}_{10}$ and MT datasets, fine-tuned means that after training on QM9, the ResNet model is further fine-tuned on a small portion of each dataset. From-scratch represents a ResNet model trained only on the $\text{C}_7\text{O}_2\text{H}_{10}$ or MT data, starting from randomly initialized parameters.

DFT electronic minimum, the energy error is quadratic in the density error (which is here small), while the relationship for other observables is linear⁵³.

Despite the use of one-step DFT errors as a metric in density learning^{28,38,41,53}, reported numbers for the multiple quantities in Table 2 for the QM9 dataset are scarce. One comparable example is given by the DeepDFT model, which reported the distribution of energy errors on this dataset. We show the comparison with the distribution in Fig. 1 and we see that our errors are on average smaller (note the logarithmic axes), and we do not see large outliers. Removing the outliers of DeepDFT (defined as density MAE > 0.6%) cannot change its prediction accuracy more than 0.01% due to the relatively few number of outliers (15 data points, according to Fig. 2 of Ref. 38 compared to 10,000 total testing data). While our ResNet model achieves higher accuracy regarding the density MAE than a direct fitting to the SCF-converged density (Table 1), it does not in terms of energy errors (Table 2). One might prioritize energy-related metrics during training if energy errors are deemed more practically significant.

In addition to the above quantities, we can also assess the quality of prediction of properties that are directly available from the density without diagonalization, such as the multipole moments (non-traceless, using the center of electron charge density as the origin). The L^2 norm of our dipole, quadrupole and octupole errors are 68 mD, 0.14 a.u., and 1.0 a.u. respectively. Another model that reports these quantities on the QM9 data set is OrbNet¹⁷. Our results can be compared to the reported 185 mD, 0.856 a.u., and 8.68 a.u. errors for OrbNet¹⁷.

Transferability on unseen geometries and elements

We now assess the transferability of our model. As a first task, we consider the performance of our ResNet model trained on QM9 on a different dataset of the isomers $\text{C}_7\text{O}_2\text{H}_{10}$. Unlike QM9, which consists only of equilibrium molecular geometries, the $\text{C}_7\text{O}_2\text{H}_{10}$ dataset contains diverse conformers sampled at high temperature, and thus the geometries and associated electron densities extend outside of the training space of QM9. Specifically, the $\text{C}_7\text{O}_2\text{H}_{10}$ dataset comprises 5000 conformations for each of the 113 isomers of $\text{C}_7\text{O}_2\text{H}_{10}$. As these conformations were sampled at 1 fs intervals from molecular dynamics (MD) simulations and geometries at adjacent time steps are very correlated, we evenly downsampled by a factor of 100 for each isomer trajectory (keeping the 0-th, 99-th, ... geometries), resulting in a total of 5650 less correlated data points.

As shown in Table 3, our ResNet model trained on QM9 (denoted zero-shot) moderately decreases in accuracy when evaluated on the $\text{C}_7\text{O}_2\text{H}_{10}$ dataset, illustrating a qualitative difference between equilibrium and non-equilibrium structure densities. This hypothesis is supported by the high accuracy we observe when evaluating our model densities only for the initial conformation of each of the 113 isomer MD trajectories ($\text{Err}_\rho = 0.13\%$), which correspond to local energy minima. We also assess the zero-shot eqDeepDFT (trained on QM9) model. This gives a zero-shot density error of 1.8% on the full diluted dataset and 0.29% on the first conformation of each isomer. The zero-shot error of ResNet is not only smaller than that of eqDeepDFT, but the increase from the QM9 test error is also smaller than that of eqDeepDFT. This shows that our density model not only fits better but also generalizes better.

Despite the decrease in accuracy of our predicted densities when fed into a one-step diagonalization, our zero-shot ResNet model still predicts the total energy within chemical accuracy of 1 kcal/mol (24.5 meV = 0.56 kcal/mol). Since we did not train on energies and the density learning was trained without conformational diversity of this isomer dataset, this illustrates the surprising robustness of the model.

To obtain better performance, we further fine-tune our ResNet model after training on QM9, using a small dataset comprising the first

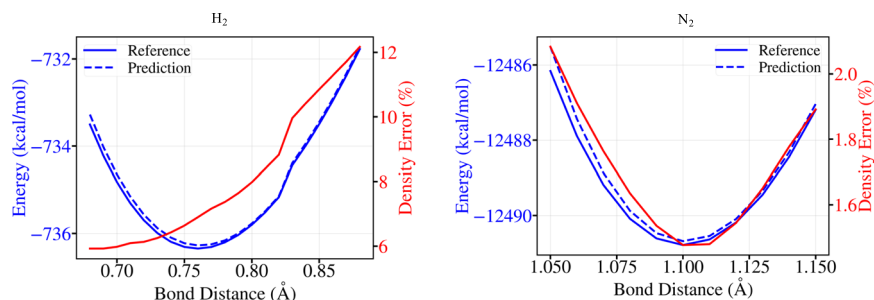


Fig. 3 | Density prediction mean absolute error and potential energy surfaces of a single H₂ and N₂ molecule as a function of bond distance. Source data is provided as a Source Data file.

two conformations from the downsampled dataset, including the one at the energy minimum—from each of the 113 trajectories. During training, we monitored the validation loss on the third confirmation of each trajectory. Since the conformations in the dataset are ordered by the MD simulation time, the use of the first 3 conformations for training and validation limits our model's fine-tuning to future conformations.

In Table 3, we show the accuracy of the fine-tuned model on the last 40 conformations of each downsampled trajectory. As such, the middle data points (4th–9th) of each trajectory are never used, and this gap between the training/validation set and test set helps minimize the correlation between the training and test data, creating a more stringent test. We note that using the first two conformations of each isomer corresponds to only 4% of our downsampled C₇O₂H₁₀ dataset (and 0.04% of the original dataset). Typically, such limited training data (113 × 2 = 226 data points) does not suffice to train a deep neural network, thus, a fine-tuning approach is essential. Training on this small dataset alone yields the from-scratch ResNet model, which is less accurate than the predictions from the zero-shot ResNet model. Fine-tuning leads to a (more than) two-fold decrease of the density and energy error on the C₇O₂H₁₀ set and a performance close to that achieved by ResNet in the QM9 testing set.

The isomer test evaluates the transferability of our model to moderately non-equilibrium geometries of a single gas-phase molecule. To assess the transferability of our model, which is trained on single-molecule data to aggregates and condensed-phase settings, we use a water cluster dataset⁵⁴. We tested the zero-shot prediction of the QM9 model on the energy-minimized geometries of water clusters ranging from 3 to 30 molecules. As shown in Fig. 2, our model trained on gas-phase single molecule data extrapolates reasonably well to the larger water clusters, with a computational cost that is a small fraction of that of the DFT calculations.

Comparing to literature models, the equivariant e3nn model trained on 7-water clusters achieved an Err_p of around 0.65% on 30-water clusters, and including clusters up to 14 waters in the training data brought the error down to ~0.4% (see Fig. 3 of Ref. 36). To test if similar behavior exists for our model, we fine-tuned the QM9 model on a few conformations containing up to 9 or 13 waters (see Methods for training and dataset details) and achieved a density error Err_p of 0.17% and 0.12%, respectively on the 30-water clusters. For a different comparison, we also trained from scratch on 4635 conformations of up to 13 waters and obtained Err_p = 0.11% on 30-water clusters.

The water cluster dataset further provides an opportunity to assess the quality of electrostatics associated with our predicted density. The Coulomb energy E^{coul} spans a large range in this dataset (its standard deviation on the 15–30 water data is 1.6×10^3 kcal/mol/molecule). Despite this, the Coulomb energy for the same 15–30 water data using the from-scratch model density prediction, followed by a single Fourier transform, gave an MAE of only 1.1 kcal/mol/molecule. As is well known, capturing long-range interactions is a challenge for

commonly employed ML approaches based on short-range potentials, and a hybrid approach⁵⁵ combining a density prediction model for the long-range electrostatic interactions with an ML potential for the remaining interactions, is a promising possibility.

As another test for model transferability on unseen geometries, we examined the density MAE of the QM9 model on H₂ and N₂ molecules. The potential energy surface was scanned every 0.01 Å, and the density error as well as the one-step DFT energy were computed at the same resolution. H–H and N–N bonds are poorly represented in the training data, and we observe that Err_p is larger than in the transferability test on the C₇O₂H₁₀ dataset, but the one-step DFT energy remains accurate. Thus, the predicted equilibrium bond lengths are accurate (exact within the 0.01 Å spacing used in the PES scan). The harmonic frequencies at the minimum are 4153 cm^{−1} (one-step DFT) and 4172 cm^{−1} (reference) for H₂; 2425 cm^{−1} (one-step DFT) and 2330 cm^{−1} (reference) for N₂, computed from finite difference with $\delta x = 0.01$ Å.

We next evaluate the ResNet model transferability on unseen elements and molecular charges. These are non-trivial tasks for models that explicitly require the element types as input, and (indeed, without any modification) are impossible for certain models. However, our network can be applied without any modification in both these cases.

As a first test, we create a training and validation set for QM9 that deliberately omits the N atom, and train a model to this limited data. The model yields a density error of 1.4% when tested on the remaining nitrogen-containing QM9 molecules. This error is significantly higher than what we saw in the previous transferability tests on the isomer and water cluster dataset. However, the 1.4% error still represents a substantial improvement over the initial SAD guess, which has a density error of 13.6%. Further analysis reveals that the density error specifically associated with nitrogen is 3.3%, compared to a 7.8% error in the SAD guess around the nitrogen atoms. (The density error around the atoms was evaluated using grid points within the atomic radii as taken from Ref. 56). This indicates that although the pre-training itself is not sufficient for a quantitatively accurate prediction of the unseen nitrogen element, it enhances the model's ability to predict densities for the unseen element by learning about the density of other elements.

As a second, even harder test, we consider a large biological system, namely, a model of the GTP hydrolysis reaction in microtubules⁵⁷ (denoted here the MT dataset). This consists of, on average, 211 atoms and a wide range of conformations involving close interactions between GTP, water, and protein amino acid residues, and bond-breaking during GTP hydrolysis. From the density prediction perspective, the system is especially challenging as it carries a total charge of −4 and also includes two new elements compared to the QM9 dataset, namely phosphorus and magnesium.

The previous transferability task omitted nitrogen in the QM9 dataset but trained on data containing oxygen and fluorine elements. Thus testing on molecules with the unseen nitrogen element tests the

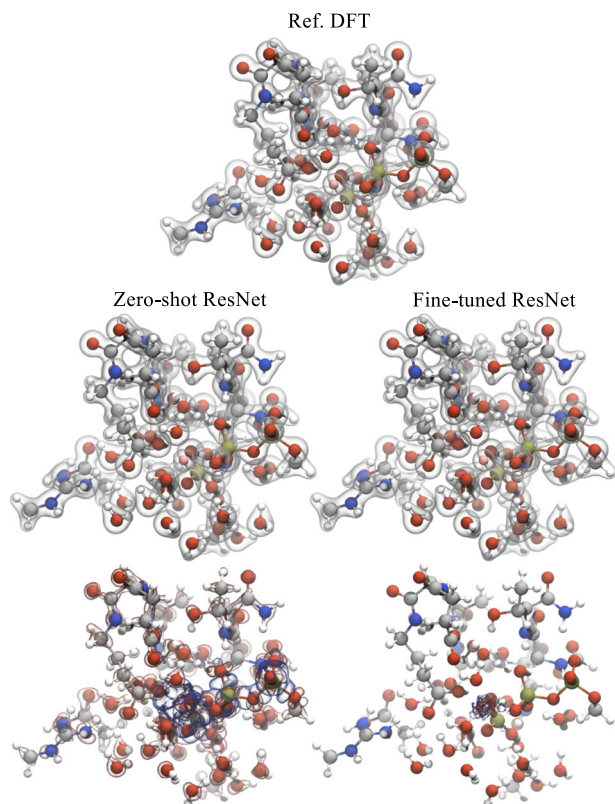


Fig. 4 | Density prediction on a microtubule-mediated GTP hydrolysis model. The first two rows show the electron density isosurface at 0.1 e/Bohr^3 . The third row shows the density prediction error ($\rho^{\text{pred}} - \rho^{\text{ref}}$) isosurfaces at 0.003 e/Bohr^3 (transparent red) and -0.003 e/Bohr^3 (transparent blue). Coloring of atoms: oxygen (red), nitrogen (blue), carbon (silver), hydrogen (white), phosphorus (gold), and magnesium (pink).

interpolation ability of the model between two adjacent elements in the periodic table. In contrast, moving from the QM9 dataset to the MT dataset involves extrapolating elemental and charge state behavior. As such, very accurate predictions are not anticipated in this scenario. Indeed, we find using the zero-shot ResNet model a relatively large average error (2.1%) with a large density error around the phosphorus and magnesium elements (7.8% and 33%, respectively). A visual check confirms that the density errors are predominantly around the unseen elements (Fig. 4).

To improve on the zero-shot model, we can use fine-tuning. Again, we limit ourselves to a small training set, containing a few initial conformations of the MD trajectories. Despite using limited training data (110 data points, comprising only 8% of the total data), the fine-tuned ResNet model achieves a 0.33% accuracy in density prediction on the test data (approximately 88% of the entire MT dataset), again approaching the accuracy of our state-of-the-art model on the simpler QM9 training dataset. In comparison, training from scratch with the same limited data results in a model that is much less accurate than the fine-tuned one. This highlights that pre-training our density model on a standard dataset such as QM9, in conjunction with a modest amount of fine-tuning, achieves quantitative accuracy even in the very challenging case of unseen systems with new charge states and unusual elements.

Discussion

We have demonstrated that a simple machine learning strategy, inspired by image super-resolution methods, is a promising approach for electron density prediction. The model is based on a convolutional residual network, and the input is a crude guess of the density

(superposition of atomic densities; SAD) represented on a coarse 3D real-space grid. We showed that even very poor input atomic guesses make virtually no difference to the accuracy, thus the input density mainly encodes the atomic positions and atom types (the latter is implicitly represented by the sizes of the atomic densities).

From a computer vision perspective, the above process mimics the well-studied image super-resolution problem. From a physical perspective, the model may be viewed as indirectly learning the Hohenberg-Kohn (HK) map from the nuclear electrostatic potential $v^{\text{ext}}(\mathbf{r})$ to the ground-state density, where the nuclear potential is encoded in the features of the input density. The latter may be seen as in the spirit of Bright-Wilson's argument^{58,59} that the nuclear potential of the Hamiltonian may be identified from the features of the density. It therefore provides an alternative target for ML approaches to the construction of the HK map from $v^{\text{ext}}(\mathbf{r})$ to $\rho(\mathbf{r})$ via minimization of ML density functionals^{60–70}.

Following the density prediction, a one-step DFT calculation enables the extraction of various DFT electronic properties, such as the total energies and molecular orbitals. We found that compared to other recent density prediction models that have used this one-step diagonalization metric, we obtained smaller average errors in the energies with fewer outliers. Our model also displayed significant transferability to previously unseen molecular geometries but only moderate transferability for interpolated chemical elements (with atomic numbers in between elements seen during training). Although we do not achieve quantitative accuracy in the zero-shot predictions for new elements, our model is easily fine-tuned on new data with unseen chemical elements, because element types are not part of the input. After fine-tuning our pre-trained models on limited data corresponding to the new geometries and elements, we found that we could achieve accurate predictions in all the cases we examined. Such transferability and fine-tunability suggests our approach is well-suited to an ensemble-based active learning setup^{71,72} on large unlabeled datasets. In particular, given its transferability, we expect a significant reduction in the need for expensive DFT data labeling.

Our model features a real-space convolution that necessitates a uniform grid. Such a requirement prevents us from directly working with angular-radial grids (such as atom-centered grids) that are more economical for small gas-phase molecules when all-electron calculations are conducted. Currently, the model predicts the valence density as the result of employing pseudo-potentials in the periodic DFT calculations used for data generation. The prediction of all-electron densities may be facilitated by a hybrid approach that employs models that predict density fitting coefficients for the core part of the density, while maintaining the ResNet approach for the valence part.

The success of our real-space approach to learning densities also suggests applications to other objects that are naturally expressed on real-space grids. In addition, while the electron density provides access to the electronic observables such as the energy after a single diagonalization, as performed in this work, it is desirable to use a data-driven approach to bypass this expensive computational step. These directions will be explored in the future.

Methods

Density learning and the superposition of atomic densities

The main objective of image super-resolution is to establish a mapping from a low-resolution and potentially inaccurate image (e.g., generated by lossy compression), denoted \mathbf{x} , to its high-resolution counterpart, \mathbf{X} . We view density learning through the same lens. Namely, given an inaccurate and low-resolution molecular density, we wish to predict the accurate molecular density, here defined as the self-consistent pseudopotential DFT density, on a finer grid. The density learning essentially performs two tasks: it transforms the inaccurate density to an accurate one, and it upscales the spatial resolution of the density. The latter is similar to interpolation and can be achieved

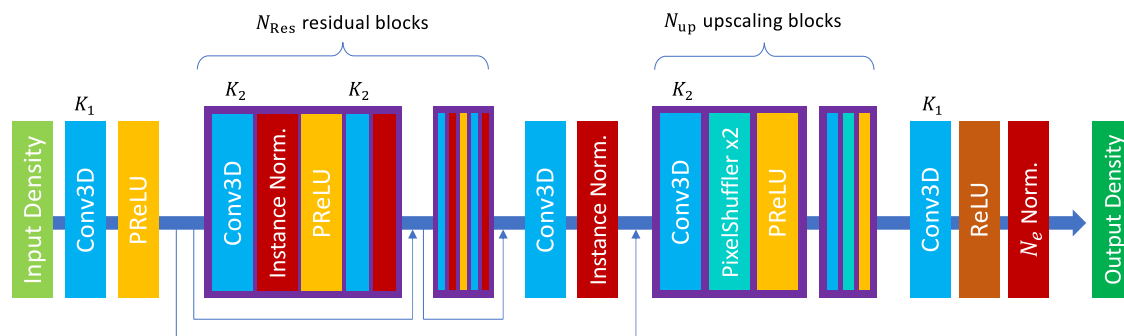


Fig. 5 | Model architecture. K indicates the kernel size used in the convolution.

without machine-learning techniques. The former is a more complex map, and is where data-driven approaches are most useful.

Since we wish to learn accurate densities for large molecules, it is essential for our input data to be cheap to generate. This suggests we should avoid approaches where we must first perform a low-level molecular quantum mechanical calculation, as these will be prohibitive in large problems. On the other hand, the list of atomic types and positions alone seems more parsimonious than necessary. We will define our input data to be the physically motivated approximation of the superposition of (spherical) neutral atomic densities, normalized to the total number of electrons, expressed on a low-resolution uniform grid. The spherical neutral atomic densities are obtained from filling electrons into atomic natural orbitals according to the Aufbau occupations, which need to be tabulated once for the periodic table, while the cost of assembling the SAD guess on the molecular grid requires no additional quantum calculation, is linear in the molecular size, and scalable to large molecules. The SAD guess preserves important physical constraints (e.g., it is positive) and provides qualitative information about the size and location of the atoms, but otherwise does not provide an accurate density, as quantified below.

DFT data generation

Our training data consisted of DFT densities computed using periodic boundary conditions in the Gaussian-Plane-Wave approach^{73–75} with the Perdew-Burke-Ernzerhof (PBE) functional⁷⁶, the Goedecker-Teter-Hutter (GTH) pseudopotential⁷⁷, and the GTH-TZV2P basis set⁷³, computed on a fine uniform grid (plane wave cutoff = 400 Ry), as implemented in the PYSCF^{74,75} package. The relation between the coarse grid for the input density and the predicted fine uniform grid is a factor of 2 or 4 in each dimension as specified below (an overall factor of 8 or 64 in the number of density points; note, the high-resolution grid is fixed, so the coarse-grid is scaled).

For all the datasets, the box size was determined based on the closest even-number multiple of the grid spacing while ensuring a minimum of 4 Å of space around the molecule in every dimension. This box size occasionally led to SCF convergence issues for some long conjugated chains in QM9, and a 5 Å margin was used instead for these cases. A SCF convergence threshold of 10^{-11} Hartree was set for the QM9, C₇O₂H₁₀, and water cluster calculations. For the water cluster dataset, we consider three subsets of the full data optimized by the TTM2.1-F potential (<https://sites.uw.edu/wdbase/database-of-water-clusters/>): (1) 1904 structures that are within 1 kJ/mol of the minimum (denoted as Wmin/ n) (2) the first few structures of each of the n -water clusters within 1 kcal/mol of the minimum (in the order of geometries listed in the xyz file) such that n times the number of structures does not exceed 1500 (denoted as W1500/ n ; 692 geometries in total), and (3) structures of each of the n -water clusters within 1 kcal/mol of the minimum such that n times the number of geometries does not exceed 50,000 (denoted as W50,000/ n ; 5137 geometries in total). The GTP hydrolysis in microtubules (MT) dataset contains 1442

configurations extracted from 11 metadynamics runs⁵⁷. The total simulation time for the metadynamics trajectories was about 750 ps. Configurations spaced every 12 hours of wall time, corresponding to a 0.5-1 ps simulation interval, were included in the dataset. For the MT dataset, the SCF convergence criteria were 10^{-7} Hartree for the energy and 10^{-5} Hartree for the orbital gradient. To aid SCF convergence, we used an electronic smearing with an electronic temperature of 0.01 Hartree.

Model details

The model is a convolutional residual neural network, adapted from the ResNet part of the SRGAN superresolution model⁷⁸ for electron density prediction. Our modifications were (1) changing the 2D convolution to a 3D one, (2) substituting batch normalization with instance normalization, (3) appending a ReLU layer to the network output to guarantee positive density values, and (4) normalizing the ReLU output to match the correct electron number, i.e., $\rho \leftarrow N_e \rho / \int \rho d\mathbf{r}$. All convolutions maintained a stride of 1, with padding applied to preserve the original size. Circular padding was employed to respect the periodic boundary conditions. The resulting architecture is depicted in Fig. 5. Note that, as we do not target human perception of the visual quality of the image, we do not use an adversarial network for our loss function (as used in image-processing tasks to produce images that are visually realistic to humans⁷⁸) but instead use a standard mean absolute error loss

$$\text{Err}_\rho = \frac{1}{N_e} \int d\mathbf{r} |\rho^{\text{pred}}(\mathbf{r}) - \rho^{\text{ref}}(\mathbf{r})|. \quad (1)$$

The reported MAE over a dataset is computed as the average of the MAE of individual molecules.

Supplementary Table 1 summarizes the hyperparameters and density errors for all models in this study. Each model underwent 50 epochs of training using the density MAE as the training loss, with the best model chosen based on the lowest validation error. The first epoch used a warm-up with a 10^5 times reduced learning rate, followed by a cosine scheduler from the second epoch starting with a learning rate specified in Supplementary Table 1. We employed the Adam optimizer⁷⁹ with $\beta_1 = 0.9$ and $\beta_2 = 0.99$. Training data was augmented with a random rotation along the x, y, or z axes. A batch size of 1, combined with instance normalization and data augmentation, proved sufficient for regularization across all cases, eliminating the need for weight decay. The data splitting for QM9 was random, and the number of training, testing, and validation points followed Ref. 38. The QM9 model was tested on the Wmin/3-30 datasets to produce Fig. 2. Fine-tuning on water clusters was performed on W1500/10 (59 geometries), and the validation error on W1500/10 was monitored. Another fine-tuning trial was performed on W1500/13 (226 geometries) and validated on W1500/14. Training from scratch on water clusters was performed on W50,000/3-13 (4635 geometries) and validated on

W50,000/14. As mentioned in the main text, the original $C_7O_2H_{10}$ dataset, comprising 5000 conformations for each of the 113 isomers, was downsampled by a factor of 100 to yield less correlated data. The first two conformations of each isomer in the diluted dataset served as the training data, with the third used for validation. The entire (diluted) dataset was used for testing if the model was not trained on the dataset, and otherwise the final 40 conformations were used for testing. For the MT dataset, the first 10 conformations from each of the 11 metadynamics trajectories were used for training, the 11th for validation, and the 17th to the last for testing. Fine-tuning on $C_7O_2H_{10}$ and MT followed exactly the same training protocol, but the models were initialized from the best QM9-trained model.

Recent learning approaches for molecular systems have favored the incorporation of physical symmetries into the model, such as equivariance under the coordinate transformations corresponding to translations, rotations, and reflections. Equivariant models for density prediction have also been formulated, for example, in Refs. 17,24,27,28,35,36,38,40,41. Our approach satisfies a transformation property that the predictions are unchanged if the molecule is rotated and the grid is simultaneously rotated with it. This is because the input SAD density is invariant under the simultaneous coordinate transformations of the molecule and grid points, and the model uses only the values of the density (not the coordinates of the grid points) thus, the predictions are unchanged. However, the set of density values is not invariant under molecular symmetry transformations if the grid points are not also transformed, i.e., the set $\{(\mathbf{r}|\rho(\mathbf{M}))\} \neq \{(\mathbf{r}|\rho(\mathbf{R}\mathbf{M}))\}$. A simple way to improve this is to augment the density data with that from rotated molecules on the same fixed grid, and we thus performed this data augmentation in our model training.

Computation of electronic properties

Some electronic properties were obtained from a one-step DFT procedure, which requires one Fock build and one diagonalization. Explicitly, a Fock matrix was built using the predicted electron density $\rho(\mathbf{r})$:

$$f_{\mu\nu} = h_{\mu\nu}^{\text{core}} + J_{\mu\nu}[\rho] + \int d\mathbf{r} \phi_{\mu}(\mathbf{r}) v_{\text{XC}}[\rho](\mathbf{r}) \phi_{\nu}(\mathbf{r}) \quad (2)$$

with μ and ν indexing atom orbitals, h^{core} , J , and v_{XC} adopting their usual quantum chemical meanings of the core Hamiltonian, Coulomb matrix, and exchange-correlation potential. The one-body reduced density matrix $\gamma_{\mu\nu}$ and orbital energies were obtained from diagonalizing the Fock matrix. The total energy and the Kohn-Sham kinetic energy were evaluated using the resulting $\gamma_{\mu\nu}$. The dipole moment, the Coulomb energy, and the exchange-correlation energy were obtained directly from the predicted density without the need for one step of DFT.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The DFT density data and molecular geometries are available at Refs. 80–82. Source data for the Figures are provided with this paper. Source data are provided with this paper.

Code availability

Implementation and minimal examples can be found at Ref. 83. Pre-trained models can be found in <https://doi.org/10.6084/m9.figshare.25365508>.

References

- Parr, R. & Weitao, Y. *Density-Functional Theory of Atoms and Molecules*. International Series of Monographs on Chemistry (Oxford University Press, New York, 1994).
- Kulik, H. J. et al. Roadmap on machine learning in electronic structure. *Electron. Struct.* **4**, 023004 (2022).
- Cerioti, M., Clementi, C. & Anatole von Lilienfeld, O. Machine learning meets chemical physics. *J. Chem. Phys.* **154**, 160401 (2021).
- Behler, J. & Parrinello, M. Generalized neural-network representation of high-dimensional potential-energy surfaces. *Phys. Rev. Lett.* **98**, 146401 (2007).
- Schütt, K. et al. Schnet: A continuous-filter convolutional neural network for modeling quantum interactions. *Adv. Neural Inf. Process. Syst.* **30** (2017).
- Bartók, A. P., Payne, M. C., Kondor, R. & Csányi, G. Gaussian approximation potentials: The accuracy of quantum mechanics, without the electrons. *Phys. Rev. Lett.* **104**, 136403 (2010).
- Unke, O. T. & Meuwly, M. Physnet: A neural network for predicting energies, forces, dipole moments, and partial charges. *J. Chem. Theory Comput.* **15**, 3678–3693 (2019).
- Zhang, L., Han, J., Wang, H., Car, R. & Weinan, E. Deep potential molecular dynamics: a scalable model with the accuracy of quantum mechanics. *Phys. Rev. Lett.* **120**, 143001 (2018).
- Smith, J. S. et al. Approaching coupled cluster accuracy with a general-purpose neural network potential through transfer learning. *Nat. Commun.* **10**, 2903 (2019).
- Smith, J. S., Isayev, O. & Roitberg, A. E. ANI-1: an extensible neural network potential with DFT accuracy at force field computational cost. *Chem. Sci.* **8**, 3192–3203 (2017).
- Gasteiger, J., Giri, S., Margraf, J. T. & Günnemann, S. Fast and uncertainty-aware directional message passing for non-equilibrium molecules. arXiv preprint arXiv:2011.14115 (2020).
- Thölke, P. & De Fabritiis, G. Equivariant transformers for neural network based molecular potentials. In *International Conference on Learning Representations* (2021).
- Batzner, S. et al. E (3)-equivariant graph neural networks for data-efficient and accurate interatomic potentials. *Nat. Commun.* **13**, 2453 (2022).
- Musaelian, A. et al. Learning local equivariant representations for large-scale atomistic dynamics. *Nat. Commun.* **14**, 579 (2023).
- Batatia, I., Kovacs, D. P., Simm, G., Ortner, C. & Csányi, G. Mace: Higher order equivariant message passing neural networks for fast and accurate force fields. *Adv. Neural Inf. Process. Syst.* **35**, 11423–11436 (2022).
- Welborn, M., Cheng, L. & Miller III, T. F. Transferability in machine learning for electronic structure via the molecular orbital basis. *J. Chem. Theory Comput.* **14**, 4772–4779 (2018).
- Qiao, Z. et al. Informing geometric deep learning with electronic interactions to accelerate quantum chemistry. *Proc. Natl Acad. Sci.* **119**, e2205221119 (2022).
- Ramakrishnan, R., Dral, P. O., Rupp, M. & Von Lilienfeld, O. A. Big data meets quantum chemistry approximations: the Δ -machine learning approach. *J. Chem. Theory Comput.* **11**, 2087–2096 (2015).
- Cerioti, M. Beyond potentials: Integrated machine learning models for materials. *Mrs Bull.* **47**, 1045–1053 (2022).
- Hohenberg, P. & Kohn, W. Inhomogeneous electron gas. *Phys. Rev.* **136**, B864 (1964).
- Brockherde, F. et al. Bypassing the Kohn-Sham equations with machine learning. *Nat. Commun.* **8**, 872 (2017).
- Schütt, K. T., Gastegger, M., Tkatchenko, A., Müller, K.-R. & Maurer, R. J. Unifying machine learning and quantum chemistry with a deep neural network for molecular wavefunctions. *Nat. Commun.* **10**, 5024 (2019).

23. Unke, O. et al. SE (3)-equivariant prediction of molecular wave-functions and electronic densities. *Adv. Neural Inf. Process. Syst.* **34**, 14434–14447 (2021).
24. Grisafi, A. et al. Transferable machine-learning model of the electron density. *ACS Cent. Sci.* **5**, 57–64 (2018).
25. Ryczko, K., Strubbe, D. A. & Tamblyn, I. Deep learning and density-functional theory. *Phys. Rev. A* **100**, 022512 (2019).
26. Zepeda-Núñez, L. et al. Deep density: circumventing the Kohn-Sham equations via symmetry preserving neural networks. *J. Comput. Phys.* **443**, 110523 (2021).
27. Lewis, A. M., Grisafi, A., Ceriotti, M. & Rossi, M. Learning electron densities in the condensed phase. *J. Chem. Theory Comput.* **17**, 7203–7214 (2021).
28. Fabrizio, A., Grisafi, A., Meyer, B., Ceriotti, M. & Corminboeuf, C. Electron density learning of non-covalent systems. *Chem. Sci.* **10**, 9424–9432 (2019).
29. Kamal, D., Chandrasekaran, A., Batra, R. & Ramprasad, R. A charge density prediction model for hydrocarbons using deep neural networks. *Mach. Learn.: Sci. Technol.* **1**, 025003 (2020).
30. Gong, S. et al. Predicting charge density distribution of materials using a local-environment-based graph convolutional network. *Phys. Rev. B* **100**, 184103 (2019).
31. del Rio, B. G., Phan, B. & Ramprasad, R. A deep learning framework to emulate density functional theory. *npj Comput. Mater.* **9**, 158 (2023).
32. Chandrasekaran, A. et al. Solving the electronic structure problem with machine learning. *npj Comput. Mater.* **5**, 22 (2019).
33. Fiedler, L. et al. Predicting electronic structures at any length scale with machine learning. *npj Comput. Mater.* **9**, 115 (2023).
34. Sinitskiy, A. V. & Pande, V. S. Deep neural network computes electron densities and energies of a large set of organic molecules faster than density functional theory (DFT). arXiv preprint arXiv:1809.02723 (2018).
35. Lee, A. J., Rackers, J. A. & Bricker, W. P. Predicting accurate ab initio dna electron densities with equivariant neural networks. *Biophys. J.* **121**, 3883–3895 (2022).
36. Rackers, J. A., Tecot, L., Geiger, M. & Smidt, T. E. A recipe for cracking the quantum scaling limit with machine learned electron densities. *Mach. Learn.: Sci. Technol.* **4**, 015027 (2023).
37. Jørgensen, P. B. & Bhowmik, A. DeepDFT: Neural message passing network for accurate charge density prediction. arXiv preprint arXiv:2011.03346 (2020).
38. Jørgensen, P. B. & Bhowmik, A. Equivariant graph neural networks for fast electron density estimation of molecules, liquids, and solids. *npj Comput. Mater.* **8**, 183 (2022).
39. Pathrudkar, S., Thiagarajan, P., Agarwal, S., Banerjee, A. S. & Ghosh, S. Electronic structure prediction of multi-million atom systems through uncertainty quantification enabled transfer learning. *npj Comput. Mater.* **10**, 175 (2024).
40. Lee, A. J., Rackers, J. A. & Bricker, W. P. Machine-learned electron densities of nucleic acids. *Biophys. J.* **123**, 499a (2024).
41. Koker, T., Quigley, K., Taw, E., Tibbetts, K. & Li, L. Higher-order equivariant neural networks for charge density prediction in materials. *npj Comput. Mater.* **10**, 161 (2024).
42. Fu, X. et al. A recipe for charge density prediction. *Adv. Neural Inf. Process. Syst.* **37**, 9727–9752 (2024).
43. Wang, Z., Chen, J. & Hoi, S. C. Deep learning for image super-resolution: A survey. *IEEE Trans. Pattern Anal. Mach. Intell.* **43**, 3365–3387 (2020).
44. Chen, H. et al. Real-world single image super-resolution: A brief review. *Inf. Fusion* **79**, 124–145 (2022).
45. He, K., Zhang, X., Ren, S. & Sun, J. Deep residual learning for image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 770–778 (2016).
46. Ramakrishnan, R., Dral, P. O., Rupp, M. & Von Lilienfeld, O. A. Quantum chemistry structures and properties of 134 kilo molecules. *Sci. Data* **1**, 1–7 (2014).
47. Apra, E. et al. NWChem: Past, present, and future. *J. Chem. Phys.* **152**, 184102 (2020).
48. Widmark, P.-O., Malmqvist, P.-Å & Roos, B. O. Density matrix averaged atomic natural orbital (ANO) basis sets for correlated molecular wave functions: I. first row atoms. *Theor. Chim. Acta* **77**, 291–306 (1990).
49. Roos, B. O., Veryazov, V. & Widmark, P.-O. Relativistic atomic natural orbital type basis sets for the alkaline and alkaline-earth atoms applied to the ground-state potentials for the corresponding dimers. *Theor. Chem. Acc.* **111**, 345–351 (2004).
50. Roos, B. O., Lindh, R., Malmqvist, P.-Å, Veryazov, V. & Widmark, P.-O. Main group atoms and dimers studied with a new relativistic basis set. *J. Phys. Chem. A* **108**, 2851–2858 (2004).
51. Whitten, J. L. Coulombic potential energy integrals and approximations. *J. Chem. Phys.* **58**, 4496–4501 (1973).
52. Mintmire, J. & Dunlap, B. Fitting the coulomb potential variationally in linear-combination-of-atomic-orbitals density-functional calculations. *Phys. Rev. A* **25**, 88 (1982).
53. Grisafi, A., Lewis, A. M., Rossi, M. & Ceriotti, M. Electronic-structure properties from atom-centered predictions of the electron density. *J. Chem. Theory Comput.* **19**, 4451–4460 (2022).
54. Rakshit, A., Bandyopadhyay, P., Heindel, J. P. & Xantheas, S. S. Atlas of putative minima and low-lying energy networks of water clusters $n=3-25$. *J. Chem. Phys.* **151**, 214307 (2019).
55. Yue, S. et al. When do short-range atomistic machine-learning models fall short? *J. Chem. Phys.* **154**, 034111 (2021).
56. Pyykkö, P. & Atsumi, M. Molecular single-bond covalent radii for elements 1–118. *Chem. – A Eur. J.* **15**, 186–197 (2009).
57. Beckett, D. & Voth, G. A. Unveiling the catalytic mechanism of GTP hydrolysis in microtubules. *Proc. Natl Acad. Sci.* **120**, e2305899120 (2023).
58. Tozer, D. J., Ingamells, V. E. & Handy, N. C. Exchange-correlation potentials. *J. Chem. Phys.* **105**, 9200–9213 (1996).
59. Rich, A., Davidson, N. & Pauling, L. *Structural Chemistry and Molecular Biology* (W. H. Freeman, San Francisco, 1968).
60. Snyder, J. C., Rupp, M., Hansen, K., Müller, K.-R. & Burke, K. Finding density functionals with machine learning. *Phys. Rev. Lett.* **108**, 253002 (2012).
61. Nagai, R., Akashi, R., Sasaki, S. & Tsuneyuki, S. Neural-network Kohn-Sham exchange-correlation potential and its out-of-training transferability. *J. Chem. Phys.* **148**, 241737 (2018).
62. Schmidt, J., Benavides-Riveros, C. L. & Marques, M. A. Machine learning the physical nonlocal exchange–correlation functional of density-functional theory. *J. Phys. Chem. Lett.* **10**, 6425–6431 (2019).
63. Zhou, Y., Wu, J., Chen, S. & Chen, G. Toward the exact exchange–correlation potential: A three-dimensional convolutional neural network construct. *J. Phys. Chem. Lett.* **10**, 7264–7269 (2019).
64. Nagai, R., Akashi, R. & Sugino, O. Completing density functional theory by machine learning hidden messages from molecules. *npj Comput. Mater.* **6**, 43 (2020).
65. Dick, S. & Fernandez-Serra, M. Machine learning accurate exchange and correlation functionals of the electronic density. *Nat. Commun.* **11**, 3509 (2020).
66. Li, L. et al. Kohn-Sham equations as regularizer: Building prior knowledge into machine-learned physics. *Phys. Rev. Lett.* **126**, 036401 (2021).
67. Kasim, M. F. & Vinko, S. M. Learning the exchange-correlation functional from nature with fully differentiable density functional theory. *Phys. Rev. Lett.* **127**, 126403 (2021).

68. Kirkpatrick, J. et al. Pushing the frontiers of density functionals by solving the fractional electron problem. *Science* **374**, 1385–1389 (2021).
69. 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).
70. Ma, H., Narayanaswamy, A., Riley, P. & Li, L. Evolving symbolic density functionals. *Sci. Adv.* **8**, eabq0279 (2022).
71. Beluch, W. H., Genewein, T., Nürnberger, A. & Köhler, J. M. The power of ensembles for active learning in image classification. In *Proceedings of the IEEE Conference On Computer Vision and Pattern Recognition*, 9368–9377 (2018).
72. Smith, J. S., Nebgen, B., Lubbers, N., Isayev, O. & Roitberg, A. E. Less is more: Sampling chemical space with active learning. *J. Chem. Phys.* **148**, 241733 (2018).
73. VandeVondele, J. et al. Quickstep: Fast and accurate density functional calculations using a mixed Gaussian and plane waves approach. *Comput. Phys. Commun.* **167**, 103–128 (2005).
74. Sun, Q. et al. PySCF: the Python-based simulations of chemistry framework. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **8**, e1340 (2018).
75. Sun, Q. et al. Recent developments in the PySCF program package. *J. Chem. Phys.* **153**, 024109 (2020).
76. Perdew, J. P., Burke, K. & Ernzerhof, M. Generalized gradient approximation made simple. *Phys. Rev. Lett.* **77**, 3865 (1996).
77. Krack, M. Pseudopotentials for H to Kr optimized for gradient-corrected exchange-correlation functionals. *Theor. Chem. Acc.* **114**, 145–152 (2005).
78. Ledig, C. et al. Photo-realistic single image super-resolution using a generative adversarial network. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 4681–4690 (2017).
79. Kingma, D. P. & Ba, J. Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980 (2014).
80. Li, C., Sharir, O., Yuan, S. & Chan, G. K.-L. QM9 DFT Electron Density [Data set]. CaltechDATA <https://doi.org/10.22002/7vr2f-Or732> (2024).
81. Li, C., Sharir, O., Yuan, S. & Chan, G. K.-L. C7O2H10 Isomer DFT Electron Density [Data set]. CaltechDATA <https://doi.org/10.22002/04d1z-1d608> (2024).
82. Li, C., Sharir, O., Yuan, S. & Chan, G. K.-L. GTP-in-Microtubule DFT Electron Density [Data set]. CaltechDATA <https://doi.org/10.22002/v5tec-g6p43> (2024).
83. Li, C., Sharir, O., Yuan, S. & Chan, G. K.-L. Image super-resolution inspired electron density prediction. *resnet_density* <https://doi.org/10.5281/zenodo.15226766> (2025).

Acknowledgements

This work was primarily supported by the United States Department of Energy, Office of Science, Basic Energy Sciences, Chemical Sciences, Geosciences, and Biosciences Division, FWP LANLE3F2 awarded to Los Alamos National Laboratory under Triad National Security, LLC (Triad) contract grant no. 89233218CNA000001, subaward C2448 to the California Institute of Technology for C.L., O.S., S.Y., and G.K.C. Additional support for G.K.C. was provided by the Camille and Henry Dreyfus

Foundation via a grant from the program “Machine Learning in the Chemical Sciences and Engineering”. G.K.C. is a Simons Investigator in Physics. The authors disclose the use of the GPT-4 (OpenAI) model during the writing of the first draft of the article. The artificial intelligence model was used to polish the language, and the generated text was carefully inspected, validated, and edited by the authors.

Author contributions

C.L., O.S., and G.K.C. conceived the original study. C.L., S.Y. carried out DFT calculations to generate the training and testing data. C.L. developed the density prediction model and conducted the research. All authors discussed the results of the manuscript, and all authors contributed to the writing of the manuscript.

Competing interests

G.K.C. is a part owner of QSimulate Inc. The remaining authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-025-60095-8>.

Correspondence and requests for materials should be addressed to Garnet Kin-Lic Chan.

Peer review information *Nature Communications* thanks the anonymous reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

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-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, 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 you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. 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-nc-nd/4.0/>.

© The Author(s) 2025