Supplementary Information

Field theoretic atomistics: Learning thermodynamic and variational surrogate to density functional theory

Sambit Das[a, †, *], Bikash Kanungo[a, †, *], Arghadwip Paul[a], and Vikram Gavini[a,b]

[a]Department of Mechanical Engineering, University of Michigan, Ann Arbor, Michigan 48109, USA

[b]Department of Materials Science and Engineering, University of Michigan, Ann Arbor, Michigan 48109, USA

[†]These authors contributed equally

[*]Corresponding authors: {dsambit, bikash}@umich.edu

S1 Bijective Maps and Variational Relations

In this section, we derive the bijective maps and the variational relations presented in the main manuscript. We start with the saddle point reformulation of DFT

$$\widetilde{E}[v_{\text{aux}}, b_{\text{s}}, N_e] = \min_{\rho \in A_N} \max_{\phi} \left(\tilde{F}[\rho] + \int \rho(\mathbf{r}) v_{\text{aux}}(\mathbf{r})\, d\mathbf{r} + P[\phi, \rho, b_{\text{s}}] \right), \text{ s.t. } \int \rho(\mathbf{r})\, d\mathbf{r} = N_e\,, \tag{S1}$$

S1.1 Relation between (ρ_g, μ_g) and (v_{aux}, N_e)

Let us denote $\bar{P}[\rho, b_{\text{s}}] = \max_\phi P[\phi, \rho, b_{\text{s}}]$. Also, let us introduce a Lagrange multiplier μ for the constraint $\int \rho(\mathbf{r})\, d\mathbf{r} = N_e$. As will be shown later, the groundstate μ is the chemical potential of the system. Thus, using Eq. S1, the groundstate density (ρ_g) can be obtained by solving

$$\min_\rho \left(\tilde{F}[\rho] + \int \rho(\mathbf{r}) v_{\text{aux}}(\mathbf{r})\, d\mathbf{r} + \bar{P}[\rho, b_{\text{s}}] - \mu \left(\int \rho(\mathbf{r})\, d\mathbf{r} - N_e \right) \right) = 0\,, \tag{S2}$$

which is equivalent to solving

$$\frac{\delta(\tilde{F}[\rho] + \bar{P}[\rho, b_{\text{s}}])}{\delta\rho(\mathbf{r})} + v_{\text{aux}}(\mathbf{r}) = \mu\,. \tag{S3}$$

In the above, μ has to be updated until $\int \rho(\mathbf{r})\, d\mathbf{r} = N_e$. We now prove the bijective map between (ρ_g, μ_g) and (v_{aux}, N_e). The proof is by contradiction, akin to the proof of the original HK theorem. First of all, two different N_e's cannot lead to the same ρ_g, as the same density cannot integrate to different number of electrons. Thus, to show the bijectivity between (ρ_g, μ_g) and (v_{aux}, N_e), we need to show that, for a given N_e, two different v_{aux}'s cannot result in the same ρ_g and μ_g. To prove the above, let us assume the contrary: that, for a fixed N_e, two different potentials—$v^{(1)}_{\text{aux}}(\mathbf{r})$ and $v^{(2)}_{\text{aux}}(\mathbf{r})$—yield the same ρ_g and μ_g. Using Eq. S3, we have

$$\left.\frac{\delta(\tilde{F}[\rho] + \bar{P}[\rho, b_{\text{s}}])}{\delta\rho(\mathbf{r})}\right|_{\rho_g} + v^{(1)}_{\text{aux}}(\mathbf{r}) = \mu_g \tag{S4}$$

$$\left.\frac{\delta(\tilde{F}[\rho] + \bar{P}[\rho, b_{\text{s}}])}{\delta\rho(\mathbf{r})}\right|_{\rho_g} + v^{(2)}_{\text{aux}}(\mathbf{r}) = \mu_g \tag{S5}$$

Taking the difference of the above two equations leads to $v^{(1)}_{\text{aux}}(\mathbf{r}) - v^{(2)}_{\text{aux}}(\mathbf{r}) = 0$, which contradicts our original assumption. This proves the bijectivity between (ρ_g, μ_g) and (v_{aux}, N_e)

We now prove that $\frac{\delta \widetilde{E}[v_{\text{aux}}, b_{\text{s}}, N_e]}{\delta v_{\text{aux}}(\mathbf{r})} = \rho_g(\mathbf{r})$. At groundstate, $\widetilde{E}[v_{\text{aux}}, b_{\text{s}}, N_e] = \tilde{F}[\rho_g] +$

$\bar{P}[\rho_g, b_{\rm s}] + \int \rho_g(\mathbf{r}) v_{\rm aux}(\mathbf{r})$. Thus,

$$\begin{aligned}\frac{\delta \widetilde{E}[v_{\rm aux}, b_{\rm s}, N_e]}{\delta v_{\rm aux}(\mathbf{r})} &= \int \left(\frac{\delta(\tilde{F}[\rho] + \bar{P}[\rho, b_{\rm s}])}{\delta \rho_g(\mathbf{r}')} + v_{\rm aux}(\mathbf{r}') \right) \frac{\delta \rho_g(\mathbf{r}')}{\delta v_{\rm aux}(\mathbf{r})}\, d\mathbf{r}' + \rho_g(\mathbf{r}) \\ &= \mu_g \int \frac{\delta \rho_g(\mathbf{r}')}{\delta v_{\rm aux}(\mathbf{r})}\, d\mathbf{r}' + \rho_g(\mathbf{r}) \\ &= \rho_g(\mathbf{r})\,,\end{aligned} \tag{S6}$$

where in the second line we used Eq. S3. In the third line, we used the fact that $\int \rho_g(\mathbf{r}')\, d\mathbf{r}' = N_e$, which results in $\int \frac{\delta \rho_g(\mathbf{r}')}{\delta v_{\rm aux}(\mathbf{r})}\, d\mathbf{r}' = 0$.

We can, similarly, prove $\frac{\partial \widetilde{E}[v_{\rm aux}, b_{\rm s}, N_e]}{\partial N_e} = \mu_g$. We have

$$\begin{aligned}\frac{\partial \widetilde{E}[v_{\rm aux}, b_{\rm s}, N_e]}{\partial N_e} &= \int \left(\frac{\delta(\tilde{F}[\rho] + \bar{P}[\rho, b_{\rm s}])}{\delta \rho_g(\mathbf{r}')} + v_{\rm aux}(\mathbf{r}') \right) \frac{\partial \rho_g(\mathbf{r}')}{\partial N_e}\, d\mathbf{r}' \\ &= \mu_g \int \frac{\partial \rho_g(\mathbf{r}')}{\partial N_e}\, d\mathbf{r}' \\ &= \mu_g\,,\end{aligned} \tag{S7}$$

where in the second line we used Eq. S3. In the third line, we used the fact that $\int \rho_g(\mathbf{r}')\, d\mathbf{r}' = N_e$, which results in $\int \frac{\partial \rho_g(\mathbf{r}')}{\partial N_e}\, d\mathbf{r}' = 1$.

S1.2 Relation between ϕ_g and $b_{\rm s}$

We recall that $P[\phi, \rho, b_{\rm s}] = -\frac{1}{8\pi} \int |\nabla \phi(\mathbf{r})|^2\, d\mathbf{r} + \int \left(\rho(\mathbf{r}) + b_{\rm s}(\mathbf{r})\right) \phi(\mathbf{r})\, d\mathbf{r}$. Let us define $G[\phi, v_{\rm aux}, N_e]$ as

$$G[\phi, v_{\rm aux}, N_e] = \min_{\rho} \left(\tilde{F}[\rho] + \int (v_{\rm aux}(\mathbf{r}) + \phi(\mathbf{r})) \rho(\mathbf{r})\, d\mathbf{r} \right)\,, \quad \text{s.t.} \int \rho(\mathbf{r})\, d\mathbf{r} = N_e\,. \tag{S8}$$

Then, using Eq. S1, the groundstate electrostatic potential, ϕ_g, can be found by solving

$$\max_{\phi} \left(-\frac{1}{8\pi} \int |\nabla \phi(\mathbf{r})|^2\, d\mathbf{r} + \int b_{\rm s}(\mathbf{r}) \phi(\mathbf{r})\, d\mathbf{r} + G[\phi, v_{\rm aux}, N_e] \right)\,, \tag{S9}$$

which is equivalent to

$$\frac{1}{4\pi}\nabla^2\phi_g(\mathbf{r}) + b_{\mathrm{s}}(\mathbf{r}) + \frac{\delta G[\phi_g, v_{\mathrm{aux}}, N_e]}{\delta\phi_g(\mathbf{r})} = 0\,. \tag{S10}$$

To prove the bijectivity between ϕ_g and b_{s}, we use similar arguments of proof by contradiction used in Sec. S1.1. We assume there are two different nuclear charges, $b_{\mathrm{s}}^{(1)}$ and $b_{\mathrm{s}}^{(2)}$, which result in the same groundstate electrostatic potential (ϕ_g). Thus, both $b_{\mathrm{s}}^{(1)}$ and $b_{\mathrm{s}}^{(2)}$ should satisfy Eq. S10 at ϕ_g. Taking the difference of those two equations will lead to the condition: $b_{\mathrm{s}}^{(1)}(\mathbf{r}) - b_{\mathrm{s}}^{(2)}(\mathbf{r}) = 0$, which contradicts our original assumption.

Using Eq. S9, we can rewrite $\widetilde{E}$ as

$$\widetilde{E}[v_{\mathrm{aux}}, b_{\mathrm{s}}, N_e] = -\frac{1}{8\pi}\int |\nabla^2\phi_g(\mathbf{r})|^2 + \int b_{\mathrm{s}}(\mathbf{r})\phi_g(\mathbf{r})\,d\mathbf{r} + G[\phi_g, v_{\mathrm{aux}}, N_e]\,. \tag{S11}$$

Thus,

$$\begin{aligned}\frac{\delta\widetilde{E}[v_{\mathrm{aux}}, b_{\mathrm{s}}, N_e]}{\delta b_{\mathrm{s}}(\mathbf{r})} &= \int \left(\frac{1}{4\pi}\nabla^2\phi_g(\mathbf{r}') + b_{\mathrm{s}}(\mathbf{r}') + \frac{\delta G[\phi_g, v_{\mathrm{aux}}, N_e]}{\delta\phi_g(\mathbf{r}')}\right)\frac{\delta\phi_g(\mathbf{r}')}{\delta b_{\mathrm{s}}(\mathbf{r})}\,d\mathbf{r} + \phi_g(\mathbf{r}) \\ &= \phi_g(\mathbf{r})\,,\end{aligned} \tag{S12}$$

where the second line results from Eq. S10.

S1.3 Derivation of covariant transformation of ρ_g and ϕ_g

We now derive the covariant transformation of ρ_g and ϕ_g when the input fields to $\widetilde{E}[v_{\mathrm{aux}}, b_{\mathrm{s}}, N_e]$ transform as $v'_{\mathrm{aux}} = Tv_{\mathrm{aux}}$ and $b'_{\mathrm{s}} = Tb_{\mathrm{s}}$, where the linear operator T is a representation of the Euclidean symmetry group. Formally, the action of T of a function $f(\mathbf{r})$ is defined as: $f'(\mathbf{r}') = \int T(\mathbf{r}', \mathbf{r})f(\mathbf{r})\,d\mathbf{r}$. In other words, $T(\mathbf{r}', \mathbf{r}) = \frac{\delta f'(\mathbf{r}')}{\delta f(\mathbf{r})}$. Using the above definition, we derive the covariance of ρ_g with respect to $(v_{\mathrm{aux}}, b_{\mathrm{s}})$. The derivation for the ϕ_g covariance with $(v_{\mathrm{aux}}, b_{\mathrm{s}})$ follows along similar lines. Under Euclidean symmetry group transformations, $\widetilde{E}$ remains invariant:

$$\widetilde{E}[v'_{\mathrm{aux}}, b'_{\mathrm{s}}, N_e] = \widetilde{E}[v_{\mathrm{aux}}, b_{\mathrm{s}}, N_e]\,. \tag{S13}$$

Now, consider the change (differential) in $\widetilde{E}$ due to a small variation $\delta v_{\text{aux}}(\mathbf{r})$ in $v_{\text{aux}}(\mathbf{r})$. The differentials on both sides of the above equation turn out to be

$$\int\int \frac{\delta \widetilde{E}[v'_{\text{aux}}, b'_{\text{s}}, N_e]}{\delta v'_{\text{aux}}(\mathbf{r}')} \frac{\delta v'_{\text{aux}}(\mathbf{r}')}{\delta v_{\text{aux}}(\mathbf{r})} \delta v_{\text{aux}}(\mathbf{r})\, d\mathbf{r}\, d\mathbf{r}' = \int \frac{\delta \widetilde{E}[v_{\text{aux}}, b_{\text{s}}, N_e]}{\delta v_{\text{aux}}(\mathbf{r})} \delta v_{\text{aux}}(\mathbf{r})\, \mathrm{d}\,\mathbf{r}\,. \tag{S14}$$

Note that for the left side in the above equation, the term $\int \frac{\delta v'_{\text{aux}}(\mathbf{r}')}{\delta v_{\text{aux}}(\mathbf{r})} \delta v_{\text{aux}}(\mathbf{r})\, d\mathbf{r}$ is nothing but $T\delta v_{\text{aux}}$. Further, we know that $\delta \widetilde{E}[v'_{\text{aux}}, b'_{\text{s}}, N_e]/\delta v'_{\text{aux}}(\mathbf{r}') = \rho_g[v'_{\text{aux}}, b'_{\text{s}}](\mathbf{r}')$ and $\delta \widetilde{E}[v_{\text{aux}}, b_{\text{s}}, N_e]/\delta v_{\text{aux}}(\mathbf{r}) = \rho_g[v_{\text{aux}}, b_{\text{s}}](\mathbf{r})$. Thus, using the bra-ket notation, we can rewrite the above equation as

$$\begin{aligned} &\langle T\delta v_{\text{aux}} | \rho_g[v'_{\text{aux}}, b'_{\text{s}}] \rangle = \langle \delta v_{\text{aux}} | \rho_g[v_{\text{aux}}, b_{\text{s}}] \rangle \ , \text{ or} \\ &\left\langle \delta v_{\text{aux}} \middle| T^\dagger \rho_g[v'_{\text{aux}}, b'_{\text{s}}] \right\rangle = \langle \delta v_{\text{aux}} | \rho_g[v_{\text{aux}}, b_{\text{s}}] \rangle \ , \end{aligned} \tag{S15}$$

where in the last step we have used the definition of the adjoint of a linear operator A ($\langle Af|g\rangle = \langle f|A^\dagger g\rangle$). Finally, as the differential relation in Eq. S15 is satisfied for any v_{aux} perturbation, we get

$$\begin{aligned} &T^\dagger \rho_g[v'_{\text{aux}}, b'_{\text{s}}] = \rho_g[v_{\text{aux}}, b_{\text{s}}] \,, \text{ or} \\ &T^\dagger \rho_g[Tv_{\text{aux}}, Tb_{\text{s}}] = \rho_g[v_{\text{aux}}, b_{\text{s}}] \,. \end{aligned} \tag{S16}$$

This proves the covariance of ρ_g with respect to input fields.

S2 Probe functions

Our goal is to use chemical species agnostic probe functions. To that end, we use a combination of short-ranged orthogonal Bessel (SR-OB) functions and long-ranged damped-multipole (LR-DM) functions for the radial probe functions $f_n(r)$ and $g_n(r)$ (see Eq. 14 in main text). The SR-OB are an extension of the standard spherical Bessel function of first kind [1]. For

a given radial cutoff (r_c), the n-th SR-OB function ($h_n(r)$) is defined as

$$h_n(r) = \frac{1}{\sqrt{d_n}} \left(j_n(r) + \frac{e_n}{d_{n-1}} j_{n-1}(r) \right) ; \; h_0(r) = j_0(r) \, , \tag{S17}$$

where

$$e_n = \frac{n^2(n+2)^2}{4(n+1)^4+1}; \quad d_n = 1 - \frac{e_n}{d_{n-1}}; \quad d_0 = 1 \, , \text{ and} \tag{S18}$$

$$j_n(r) = (-1)^n \frac{\sqrt{2}\pi}{r_c^{3/2}} \frac{(n+1)(n+2)}{\sqrt{(n+1)^2+(n+2)^2}} \left[\text{sinc}\left(\frac{r(n+1)\pi}{r_c} \right) + \text{sinc}\left(\frac{r(n+2)\pi}{r_c} \right) \right] . \tag{S19}$$

The h_n's have the following properties: (i) $h_n(r) = 0$, $\forall r \geq r_c$ (short-ranged); and (ii) $\int h_n(r) h_{n'}(r) r^2 \, d\mathbf{r} = \delta_{nn'}$ (orthogonal). Fig. S1 shows the first ten SR-OB functions. We use 20 and 15 SR-OB functions to probe v_{aux} and b_{s}, respectively. We use r_c values of 4 and 6.5 for v_{aux} and b_{s}, respectively.

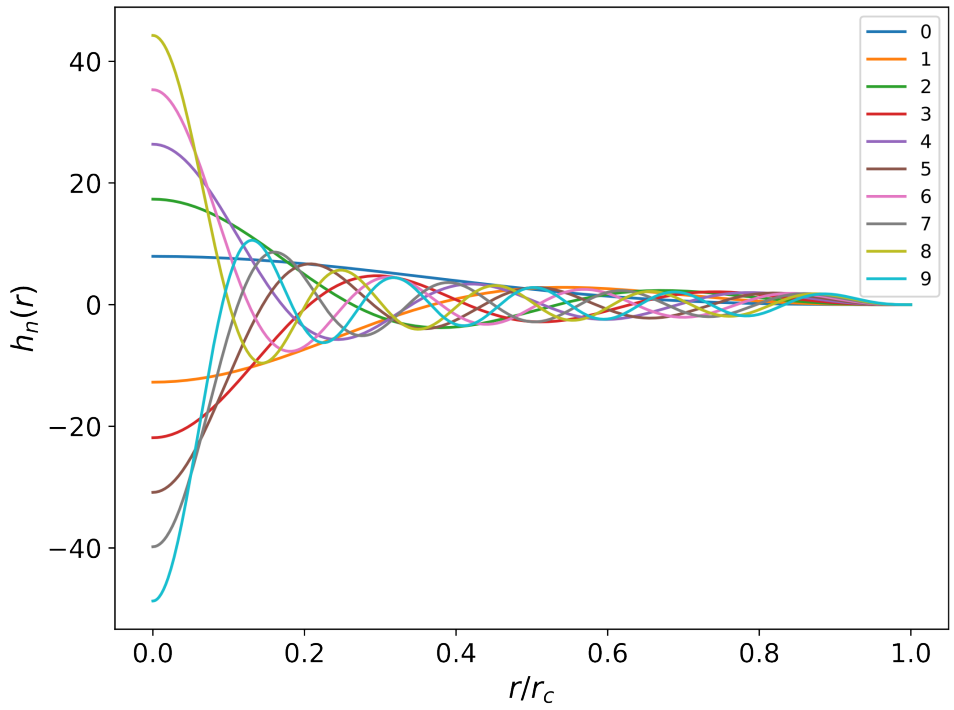

Figure S1: First ten short-ranged orthogonal Bessel (SR-OB) functions.

The LR-DM functions, $v_{a,b,c,k,r_c,t}(r)$, are given as

$$v_{a,b,c,k,r_c,t}(r) = \xi_{a,b,c,k}(r) \; s_{r_c,t}(r) \, , \tag{S20}$$

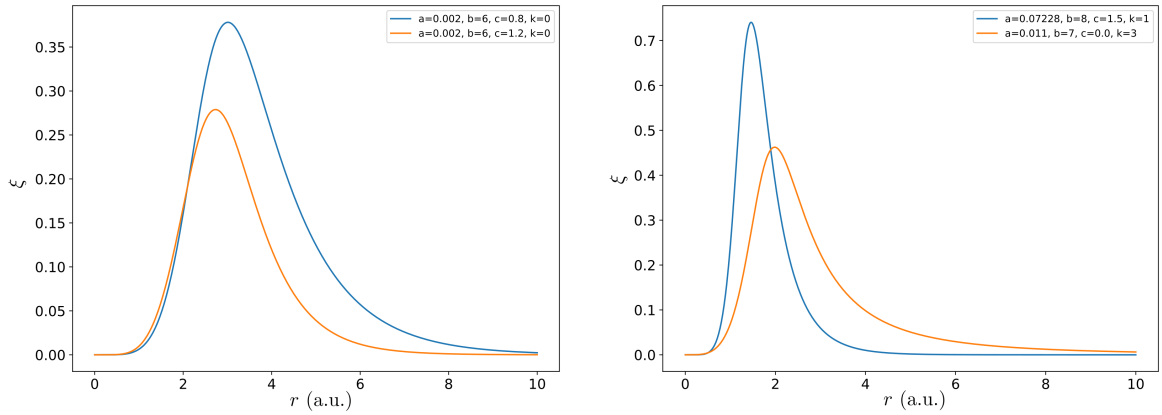

Figure S2: Long-ranged damped multipole (LR-DM) functions used to probe v_{aux} (left) and b_{s} (right).

where

$$\xi_{a,b,c,k}(r) = \left(1 - \frac{1}{1+ar^b}\right)\frac{\exp(-cr)}{r^k}\,, \tag{S21}$$

with $a, b, c, k \geq 0$. $s_{r_c,t}(r)$, with $t > 0$, is a smooth cutoff function with the following properties:

$$\begin{cases} s_{r_c,t}(r) = 1, & r < r_c\,\frac{t}{t+1} \\ 0 < s_{r_c,t}(r) < 1, & r_c\,\frac{t}{t+1} \leq r < r_c \\ s_{r_c,t}(r) = 0, & r \geq r_c\,. \end{cases} \tag{S22}$$

In our work, we use $r_c = 24$ and $t = 2$ for $s_{r_c,t}$. For the LR-DM probe functions for $v_{\text{aux}}(\mathbf{r})$, we set $k = 0$ (in Eq. S21). This is motivated by the fact that the probe functions for v_{aux} serves the purpose of being a basis for $\Delta\rho_g(\mathbf{r})$ and hence should have an exponential decay. Similarly, the probe functions for $b_{\text{s}}(\mathbf{r})$ should serve as a good basis for $\Delta\phi_g(\mathbf{r})$, which for a charge neutral system has screened multipole decay of the form $\exp(-cr)/r^k$. Accordingly, for the LR-DM probe functions for $b_{\text{s}}(\mathbf{r})$, we use $k = \{1, 3\}$.

S3 Form of nuclear charge distribution

We recap that our auxiliary nuclear charge takes the form $b_{\text{s}}(\mathbf{r}) = \sum_I b_I(\mathbf{r})$, where $b_I(\mathbf{r}) = -Z_I b_{\text{u}}(|\mathbf{r}-\mathbf{R}_I|; r_c)$. Here, $b_{\text{u}}(r; r_c)$ is a compactly supported positive unit charge distribution

with a cutoff radius of r_c. The form of $b_{\mathrm{u}}(r;r_c)$ we employ is borrowed from Ref. [2] and is given as

$$b_{\mathrm{u}}(r;r_c) = \begin{cases} \frac{-21(r-r_c)^3(6r^2+3rr_c+r_c^2)}{5\pi r_c^8}, & 0 \le r \le r_c, \\ 0, & r > r_c\,. \end{cases} \tag{S23}$$

The potential corresponding to $b_{\mathrm{u}}(r,r_c)$, is given as

$$v_{\mathrm{u}}(r;r_c) = \begin{cases} \frac{9r^7-30r^6r_c+28r^5r_c^2-14r^2r_c^5+12r_c^7}{5r_c^8}, & 0 \le r \le r_c \\ \frac{1}{r}, & r > r_c\,. \end{cases} \tag{S24}$$

Using the above, we can simplify $v_{\mathrm{S}}(\mathbf{r}) = \int \frac{b_{\mathrm{S}}(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}dr'$ as

$$v_{\mathrm{S}}(\mathbf{r}) = -\sum_I Z_I v_{\mathrm{u}}(|\mathbf{r}-\mathbf{R}_I|;r_c)\,. \tag{S25}$$

An $r_c < \frac{1}{2}\min\left(|\mathbf{R}_I - \mathbf{R}_J|\right)$ (i.e., less than half the minimum pairwise distance between atoms) ensures non-overlapping nuclear charges. In this work, we use $r_c = 0.5$ a.u.

S4 Reference field data

In this section, we discuss how the reference field, $\Delta\rho_g(\mathbf{r})$ and $\Delta\phi_g(\mathbf{r})$, used in our training and testing are represented. For numerical efficiency and compact representation, we represent the fields in the corresponding probe function basis. Let's denote $\eta(\mathbf{r})$ to be a generic representation for $\Delta\rho_g(\mathbf{r})$ and $\Delta\phi_g(\mathbf{r})$. Similarly, let's denote $p_i(\mathbf{r})$ as a generic representation of the probe function basis (i.e., for $u_i(\mathbf{r})$ and $w_i(\mathbf{r})$ defined in Eq. 13 of main text). We define an approximation of $\eta(\mathbf{r})$ in the $p_i(\mathbf{r})$ basis, denoted by $\widetilde{\eta}(\mathbf{r})$, as

$$\widetilde{\eta}(\mathbf{r}) = \sum_I \sum_i c_{I,i} p_i(\mathbf{r}-\mathbf{r}_I)\,. \tag{S26}$$

The linear coefficients $c_{I,i}$ can be obtained through a minimization of the difference between $\eta(\mathbf{r})$ and $\widetilde{\eta}(\mathbf{r})$ in an appropriate norm. We evaluate them through the following minimization

$$\min_{\{c_{I,i}\}} \int\int \left(\eta(\mathbf{r}) - \widetilde{\eta}(\mathbf{r})\right) \left(a\delta(\mathbf{r}-\mathbf{r}') + \frac{b}{|\mathbf{r}-\mathbf{r}'|}\right) \left(\eta(\mathbf{r}') - \widetilde{\eta}(\mathbf{r}')\right)\, d\mathbf{r}\, d\mathbf{r}' \,. \tag{S27}$$

The solution to the above takes the form of the following linear system of equations

$$(a\mathbf{S} + b\mathbf{J})\boldsymbol{c} = a\mathbf{t}_S + b\mathbf{t}_J \,, \tag{S28}$$

where $\mathbf{S}$ and $\mathbf{J}$ are the overlap and Coulomb matrix for the p_i basis, given as

$$S_{I,i}^{J,j} = \int p_i(\mathbf{r}-\mathbf{R}_I) p_j(\mathbf{r}-\mathbf{R}_J)\,, \quad J_{I,i}^{J,j} = \int\int \frac{p_i\left(\mathbf{r}-\mathbf{R}_I\right) p_j\left(\mathbf{r}'-\mathbf{R}_J\right)}{|\mathbf{r}-\mathbf{r}'|}\, d\mathbf{r}\, d\mathbf{r}' \,. \tag{S29}$$

The vectors $\mathbf{t}_S$ and $\mathbf{t}_J$ are defined as

$$b_S^{I,i} = \int \eta(\mathbf{r}) p_i(\mathbf{r}-\mathbf{R}_I)\, d\mathbf{r}\,, \quad b_J^{I,i} = \int\int \frac{\eta(\mathbf{r}) p_i\left(\mathbf{r}'-\mathbf{R}_I\right)}{|\mathbf{r}-\mathbf{r}'|}\, d\mathbf{r}\, d\mathbf{r}' \,. \tag{S30}$$

Using $b = 0$ results in the familiar L_2 (or least square) projection. Using $a = 0$ leads to minimization in the Coulomb norm. In this work, we use $a = 1, b = 1$ for the projection of $\Delta\rho_g(\mathbf{r})$ and $a = 1, b = 0$ for the projection of $\Delta\phi_g(\mathbf{r})$.

S5 Evaluation of field loss

Recall that our loss function takes the form (see Eq. 19 in main text),

$$\begin{aligned}\mathcal{L} = &\frac{t_E}{M}\sum_{\alpha=1}^{M}\left(\Delta\widetilde{E}_{\text{ref}}^{(\alpha)} - \Delta\widetilde{E}^{(\alpha)}\right)^2 + \frac{t_F}{M}\sum_{\alpha=1}^{M}\left[\sum_I \frac{1}{N_a}\left(\left\|\mathbf{f}_{I,\text{ref}}^{(\alpha)} - \mathbf{f}_I^{(\alpha)}\right\|^2\right)\right] \\ &+ \frac{t_\rho}{M}\sum_{\alpha=1}^{M}\left\|\Delta\rho_{g,\text{ref}}^{(\alpha)}(\mathbf{r}) - \Delta\rho_g^{(\alpha)}(\mathbf{r})\right\|_{\text{J}}^2 + \frac{t_\phi}{M}\sum_{\alpha=1}^{M}\left\|\Delta\phi_{g,\text{ref}}^{(\alpha)}(\mathbf{r}) - \Delta\phi_g^{(\alpha)}(\mathbf{r})\right\|_{\text{S}}^2,\end{aligned} \tag{S31}$$

where α indexes the M training samples; t_E, t_F, t_ρ, and t_ϕ are the weights assigned to the energy, force, density, and potential loss terms; $\|\cdot\|$ is the Eucledian norm, $\|f(\mathbf{r})\|_{\mathrm{S}} = \sqrt{\int (f(\mathbf{r}))^2 \, d\mathbf{r}}$ is the L_2 norm for a continuous function; and $\|f(\mathbf{r})\|_{\mathrm{J}} = \sqrt{\int\int \frac{f(\mathbf{r})f(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} \, d\mathbf{r} \, d\mathbf{r}'}$ is the Coulomb norm. In the above, the third and fourth term are the field loss terms corresponding to the density and potential, respectively. Using a 3D quadrature grid to evaluate the field loss terms can be computationally expensive, especially when it is done in every epoch in the training. This cost can be alleviated by using the fact that both the reference and the predicted fields are represented in the same basis. To elaborate, consider the $\boldsymbol{c}_\alpha$ and $\widetilde{\boldsymbol{c}}_\alpha$ be the linear coefficients for $\Delta\rho^{(\alpha)}_{g,\mathrm{ref}}(\mathbf{r})$ and $\Delta\rho^{(\alpha)}_{g}(\mathbf{r})$, respectively, in the $u_i(\mathbf{r})$ basis. Thus, we can rewrite the density loss $\left\|\Delta\rho^{(\alpha)}_{g,\mathrm{ref}}(\mathbf{r}) - \Delta\rho^{(\alpha)}_{g}(\mathbf{r})\right\|^2_{\mathrm{J}} = \boldsymbol{c}^T_\alpha \mathbf{J} \widetilde{\boldsymbol{c}}_\alpha$. Similarly, denoting $\mathbf{d}_\alpha$ and $\widetilde{\mathbf{d}}_\alpha$ as the linear coefficients for $\Delta\phi^{(\alpha)}_{g,\mathrm{ref}}(\mathbf{r})$ and $\Delta\phi^{(\alpha)}_{g}(\mathbf{r})$, respectively, in the $w_i(\mathbf{r})$ basis, the potential loss can be written as $\left\|\Delta\phi^{(\alpha)}_{g,\mathrm{ref}}(\mathbf{r}) - \Delta\phi^{(\alpha)}_{g}(\mathbf{r})\right\|^2_{\mathrm{S}} = \mathbf{d}^T_\alpha \mathbf{S} \widetilde{\mathbf{d}}_\alpha$. In other words, the field loss terms can be efficiently recast into matrix-vector products.

References

[1] Emir Kocer, Jeremy K. Mason, and Hakan Erturk. A novel approach to describe chemical environments in high-dimensional neural network potentials. The Journal of Chemical Physics, 150(15):154102, 04 2019.

[2] JE Pask, N Sukumar, and SE Mousavi. Linear scaling solution of the all-electron coulomb problem insolids. International Journal for Multiscale Computational Engineering, 10(1), 2012.