ON-POLICY SELF-DISTILLATION FOR LEARNING KOHN–SHAM HAMILTONIANS

Anonymous authors
Paper under double-blind review

ABSTRACT

Neural networks that predict the Kohn–Sham Hamiltonian provide initial guesses that shorten self-consistent field (SCF) calculations. However, every training label is itself a converged SCF calculation. Extending a model to new molecules or functionals therefore requires a new labeled dataset. We observe that a few SCF iterations on the model's own prediction return a corrected Hamiltonian, which indicates how every matrix entry should change. Building on this, we propose on-policy self-distillation with the SCF map (OPSD), which uses the SCF solver as the teacher. A pretrained flow model predicts Hamiltonians, the solver corrects them, and the model learns from the corrections before generating the next round. The model thus amortizes the solver, using only geometries and standard SCF software. In eight transfer settings, including a new functional, implicit solvent, and larger molecules, OPSD has a lower mean error than fine-tuning on converged labels. After a change of functional, for example, the HOMO–LUMO gap error is 2.2 mHa with OPSD and 42 mHa with fine-tuning. On QH9, OPSD lowers the gap error of QHFlow2 models by up to 70%. On the 58-element HELM benchmark, it stabilizes the orbital spectrum of a pretrained QHFlow2 model. OPSD could thus serve as a general method for improving the transferability of Hamiltonian models.

1 INTRODUCTION

Density functional theory (DFT) is the most widely used electronic structure method in chemistry and materials science (Hohenberg & Kohn, 1964; Kohn & Sham, 1965). A DFT calculation is carried out by the self-consistent field (SCF) procedure (Roothaan, 1951). Starting from a guess of the Fock matrix, it computes the electron density, rebuilds the Fock matrix from it, and repeats until the two are consistent. Each iteration requires an expensive Fock build, so the cost depends heavily on the initial guess. Classical guesses use superpositions of atomic densities (van Lenthe et al., 2006; Lehtola, 2019) or extrapolation along trajectories (Pulay & Fogarasi, 2004; Niklasson, 2008). More recently, neural networks have been trained to predict the converged Fock matrix directly from the geometry (Schütt et al., 2019; Unke et al., 2021; Li et al., 2022; Yu et al., 2023b). The latest models, such as QHFlow, generate the matrix with conditional flow matching (Kim et al., 2025; 2026). The prediction gives orbital energies without any SCF iterations. As an initial guess, it also substantially reduces the iterations needed (Li et al., 2025; Kim et al., 2025).

As these models scale to hundreds of atoms, more elements, and diffuse triple-ζ bases (Luo et al., 2025a; Kaniselvan et al., 2025), training data becomes the limiting factor. Every label is a converged SCF calculation, and these calculations grow expensive with system size. Generating PubChemQH took about a month on 128 GPUs (Li et al., 2025), ∇^2DFT about 120 CPU-years (Khrabrov et al., 2024), and OMol25, the source of the HELM data, billions of CPU core-hours (Levine et al., 2025; Kaniselvan et al., 2025). Extending a model to larger molecules or off-equilibrium geometries therefore requires a new set of converged calculations each time.

In this paper we study how to extend a pretrained Hamiltonian model to such systems without converged SCF calculations. Self-consistency training (SCT) is one label-free alternative (Zhang et al., 2024). It minimizes the difference between a predicted Hamiltonian and the Fock matrix rebuilt from it. However, it must backpropagate through the eigensolver and the Fock build. This requires a differentiable electronic-structure code and is unstable when orbital energies are close. Reinforcement learning is another alternative. It is commonly used to post-train diffusion and flow models (Liu et al.,

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

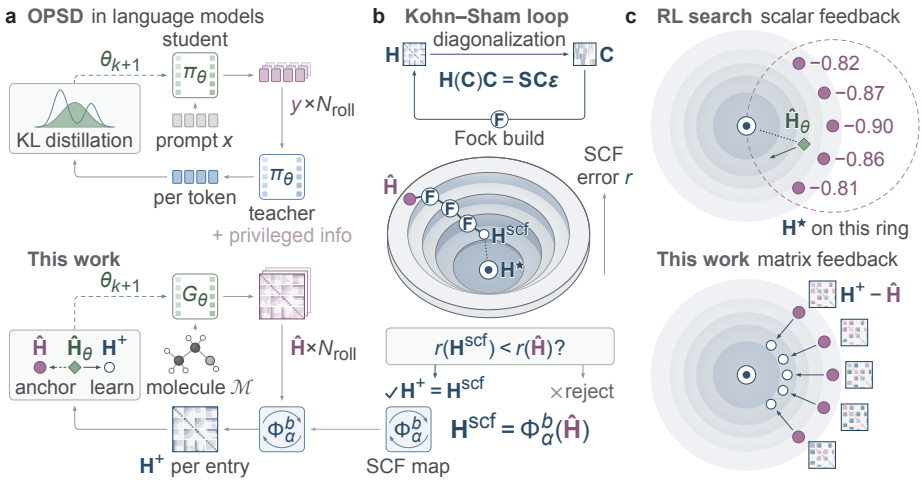


Figure 1: **Learning from SCF corrections.** (a) The actor's flow sampler G_θ produces clean Hamiltonian predictions, SCF supplies corrected targets, and the learner fits these targets before the actor is refreshed. The positive branch follows the correction, and the negative branch anchors the learner to the actor. The upper row shows LLM OPSD (Zhao et al., 2026): a privileged teacher supplies dense token-level feedback through next-token distributions along the student's own trajectories. (b) SCF diagonalizes a Hamiltonian and rebuilds it from its density. The panel idealizes the teacher as b damped iterations, Φ_α^b, without the trust-radius clip or the acceptance margin, and r is the residual r_{AO}. Section 3.1 gives the DIIS solver and the target used in training. (c) RL provides a sparse terminal reward, while SCF distillation provides dense supervision across the Hamiltonian matrix. Contours and trajectories are schematic.

2025; Li et al., 2026; Park & Walsh, 2026), and the self-consistency error of a prediction could serve as its reward. The drawback is sample efficiency. A scalar reward does not indicate how the matrix entries should change, so the model has to sample and compare many candidates. Each candidate costs a Fock build to score, and the number needed grows with the number of directions in which the model can change its prediction.

Robot learning and language modeling face the same limitation. Both address it with feedback on the model's own outputs. In human-in-the-loop reinforcement learning, a human operator briefly takes control when the robot goes wrong and supplies corrective actions (Luo et al., 2025b). In on-policy distillation for language models, a teacher scores every token of the student's own response (Agarwal et al., 2024; Lu & Thinking Machines Lab, 2025). The recipe works because the feedback acts on the policy's own mistakes. A Hamiltonian model already has such a source of corrections: the SCF solver. Its basic operation, which we call the *SCF map*, rebuilds the Fock matrix from the density of the current one. We apply it to the model's own prediction and obtain a corrected matrix. The corrected matrix gives a training target for every entry (§3.1). Computing it needs only a standard SCF code, with no gradients through the solver. We call the resulting method *on-policy self-distillation with the SCF map* (OPSD). A pretrained flow model (Kim et al., 2025; 2026) produces clean Hamiltonian predictions. A few SCF iterations improve each prediction. The model is then updated toward the improved matrices while staying close to the version of itself that produced them. The updated model generates the next round. Each geometry costs a few Fock builds per round, whereas a converged label costs a full SCF calculation. In this way, the model amortizes the SCF solver. SCT pursues the same goal by differentiating through the solver; OPSD instead learns from the solver's own corrections.

Experimentally, OPSD adapts a QHFlow2-s model pretrained on QH9 (Yu et al., 2023a) to eight new settings, such as a different functional, implicit solvent, and larger molecules, using only geometries (§5.1). Its gap error after a change of functional and its dipole error in implicit solvent are lower than those of fine-tuning on converged labels on nearly every molecule. On the four QH9 tasks, OPSD lowers the HOMO–LUMO gap error of QHFlow2 by up to 70%, while reinforcement learning and SCT improve it by 1.5–5.2%. On HELM, it stabilizes the orbital spectrum of a model pretrained across 58 elements. On MD17 (Chmiela et al., 2017), QHFlow2 already predicts forces on par with state-of-the-art machine-learned force fields, and OPSD improves them further (§5). Appendix B reviews related work.

2 PRELIMINARIES

This section reviews Kohn–Sham DFT, the SCF map, and the flow-matching model that OPSD post-trains. Appendix A collects the notation used throughout the paper.

Kohn–Sham DFT. Let $\mathcal{M}$ denote a molecule specified by its atomic numbers and nuclear positions, with N electrons. Within each experiment, the basis, functional, electron count, and environment fix the electronic-structure problem and its SCF map, and we suppress them in the notation. Settings that vary within a model's inputs, such as an applied field, are part of its conditioning. With effective core potentials, N counts the explicit electrons. We consider closed-shell molecules, for which the $n_{\mathrm{occ}} = N/2$ lowest orbitals are doubly occupied. In an atom-centered basis of size B with overlap matrix $\mathbf{S}$, the Kohn–Sham equations take the form (Kohn & Sham, 1965) $\mathbf{HC} = \mathbf{SC}\boldsymbol{\epsilon}$, $\mathbf{C}^\top\mathbf{SC} = \mathbf{I}$. Here $\mathbf{H}$ is a real symmetric $B \times B$ Hamiltonian, the columns of $\mathbf{C}$ are orbital coefficients, and $\boldsymbol{\epsilon}$ is diagonal with orbital energies $\epsilon_1, \ldots, \epsilon_B$ in increasing order. The complex-valued formulation is analogous, with Hermitian matrices and conjugate transposes. We assume $\mathbf{S} \succ 0$ in the full basis; Section 5 treats the near-singular HELM case. The HOMO and LUMO energies are $\epsilon_{n_{\mathrm{occ}}}$ and $\epsilon_{n_{\mathrm{occ}}+1}$, with gap $\epsilon_\Delta = \epsilon_{n_{\mathrm{occ}}+1} - \epsilon_{n_{\mathrm{occ}}}$. The first n_{occ} columns, collected in $\mathbf{C}_{\mathrm{occ}}$, determine the density matrix $\mathbf{P}(\mathbf{H}) := 2\,\mathbf{C}_{\mathrm{occ}}\mathbf{C}_{\mathrm{occ}}^\top$. A positive gap makes the density well defined. This density enters the Hamiltonian through $\mathbf{F}(\mathbf{P}) := \mathbf{H}_{\mathrm{core}} + \mathbf{J}(\mathbf{P}) + \mathbf{V}_{\mathrm{xc}}(\mathbf{P})$, where $\mathbf{H}_{\mathrm{core}}$, $\mathbf{J}$, and $\mathbf{V}_{\mathrm{xc}}$ are the core, Coulomb, and exchange–correlation matrices; $\mathbf{V}_{\mathrm{xc}}$ includes exact exchange for hybrid functionals. We call one evaluation of $\mathbf{F}$ a *Fock build* and use it as our unit of solver cost.

The SCF map. The density and Fock build define the SCF map $\Phi(\mathbf{H}) := \mathbf{F}\big(\mathbf{P}(\mathbf{H})\big)$. An application of Φ diagonalizes $\mathbf{H}$, forms the density, and rebuilds the Hamiltonian. At a self-consistent solution $\mathbf{H}^\star(\mathcal{M})$, this cycle returns the same matrix: $\mathbf{H}^\star = \Phi(\mathbf{H}^\star)$. Starting from an initial guess, an SCF solver approaches this fixed point through repeated updates. Each iteration costs one Fock build. Practical solvers accelerate convergence with damping, level shifts, and DIIS extrapolation over previous iterates (Saunders & Hillier, 1973; Pulay, 1980; Kudin et al., 2002). The residual $\Phi(\mathbf{H}) - \mathbf{H}$ vanishes at self-consistency, providing a criterion for assessing predictions.

Hamiltonian prediction. Hamiltonian models predict $\mathbf{H}^\star(\mathcal{M})$ from geometry using converged labels for training. Their predictions give orbital energies through the generalized eigenproblem and initial guesses for SCF. Equivariance ensures that these predictions transform consistently with the molecule. For example, a 3×5 block $\mathbf{H}^{pd}_{ij}$ couples the three $2p$ orbitals on atom i to the five $3d$ orbitals on atom j. Under a molecular rotation R, the p orbitals mix like Cartesian vectors, while the d orbitals mix through the 5×5 real Wigner matrix $D^{(2)}(R)$. Thus $\mathbf{H}^{pd}_{ij} \mapsto R\mathbf{H}^{pd}_{ij}D^{(2)}(R)^\top$, with a different rotation representation acting on each side. More generally, a rotation or atom permutation g acts through an orthogonal basis representation $D(g)$, giving $\mathbf{H} \mapsto D(g)\mathbf{H}D(g)^\top$ (Yu et al., 2023b; Li et al., 2025). In practice, equivariant networks commonly use Clebsch–Gordan tensor products to combine features while preserving these rotation rules.

Flow matching for Hamiltonian prediction. We use the clean-space formulation of QHFlow (Kim et al., 2025), a flow-matching model initialized from a MINAO guess with equivariant Gaussian noise. At time t, the network with parameters θ maps the noisy state $\mathbf{H}_t$ and geometry $\mathcal{M}$ to the clean Hamiltonian prediction $\hat{\mathbf{H}}_\theta(\mathbf{H}_t, \mathcal{M}, t)$. We call $(\mathbf{H}_t, \mathcal{M}, t)$ a *flow state*. Suppressing the common inputs, the velocity is $v_\theta = (\hat{\mathbf{H}}_\theta - \mathbf{H}_t)/(1-t)$ for $0 \le t < 1$, and the sampler follows $\mathrm{d}\mathbf{H}_t/\mathrm{d}t = v_\theta$. Equivalently, the clean Hamiltonian prediction is $\hat{\mathbf{H}}_\theta = \mathbf{H}_t + (1-t)v_\theta$. Thus the clean and velocity parameterizations describe the same flow for $t < 1$. Pretraining fits clean Hamiltonian predictions to $\mathbf{H}^\star$ along linear prior–target interpolations. Sampling takes a few Euler steps and returns the final clean Hamiltonian prediction. The network predicts all Hamiltonian entries at each flow state (Appendix E).

3 ON-POLICY SELF-DISTILLATION WITH THE SCF MAP

Starting from a pretrained Hamiltonian flow model, we turn the model's own predictions into training targets with a few SCF cycles. A frozen *actor* produces clean Hamiltonian predictions, the SCF solver corrects them, and a *learner* fits the corrected targets (Figure 1a). The actor is periodically refreshed from the learner, so later corrections act on the current model's predictions. This is on-policy distillation (Agarwal et al., 2024; Zhou et al., 2026) with the SCF solver as the teacher.

3.1 SCF CORRECTION TARGETS

For each training geometry, the actor samples N_{roll} flow trajectories of L steps, each from its own noisy MINAO guess. Along each trajectory, the actor visits a sequence of flow states $(\mathbf{H}_t, \mathcal{M}, t)$. We keep the actor-visited flow states at the low-noise end of each trajectory and store each with the actor's clean prediction $\hat{\mathbf{H}}$. The solver corrects $\hat{\mathbf{H}}$, and the learner is later evaluated at the same flow state, so both models see identical inputs. Starting from the density $\mathbf{P}(\hat{\mathbf{H}})$, we run b DIIS-accelerated SCF cycles with an empty DIIS history to obtain a candidate $\mathbf{H}^{\text{scf}}$. We accept or reject each candidate by whether it lowers the atomic-orbital (AO) residual, $\gamma = \mathbf{1}[r_{\text{AO}}(\hat{\mathbf{H}}) - r_{\text{AO}}(\mathbf{H}^{\text{scf}}) > \delta_r]$ with $r_{\text{AO}}(\mathbf{H}) = \|\Phi(\mathbf{H}) - \mathbf{H}\|_F / B$, where r_{AO} is the root-mean-square entry of the SCF residual and $\delta_r \geq 0$ is the required improvement. We further bound each accepted displacement by a trust radius $r_{\text{tr}} > 0$ and train on $\mathbf{H}^+ = \hat{\mathbf{H}} + d$, with $d = \min\left\{1, \frac{r_{\text{tr}}}{\|\mathbf{H}^{\text{scf}} - \hat{\mathbf{H}}\|_F}\right\}(\mathbf{H}^{\text{scf}} - \hat{\mathbf{H}})$. Each correction costs $b + 1$ Fock builds, b cycles plus one to evaluate the candidate's residual. Appendix E gives the solver settings for QH9 and the transfer studies.

3.2 LEARNING FROM THE CORRECTIONS

The corrected matrix $\mathbf{H}^+$ improves on the actor's prediction $\hat{\mathbf{H}}$ but is still inexact, so the learner should move only part of the way toward it. At a stored flow state, let $\hat{\mathbf{H}}_\theta$ be the learner's clean prediction, $\Delta_\theta = \hat{\mathbf{H}}_\theta - \hat{\mathbf{H}}$ its displacement from the actor, and $d = \mathbf{H}^+ - \hat{\mathbf{H}}$ the correction. As in negative-aware fine-tuning (DiffusionNFT; Zheng et al., 2026), two branches around the actor control this step. The positive branch $\mathbf{H}_\theta^+ = \hat{\mathbf{H}} + \beta\Delta_\theta$ fits the target $\mathbf{H}^+$, and the negative branch $\mathbf{H}_\theta^- = \hat{\mathbf{H}} - \beta\Delta_\theta$ fits the anchor $\hat{\mathbf{H}}$, both in clean Hamiltonian space: $\mathcal{L}_{\text{pair}}(\theta) = \mathbb{E}\big[\frac{\gamma}{2}\{\ell_{\text{mat}}(\beta\Delta_\theta - d) + \lambda_{\text{neg}}\ell_{\text{mat}}(-\beta\Delta_\theta)\}\big]$. In plainer words, the positive branch pulls the learner toward $\mathbf{H}^+$, and the negative branch holds it near the actor. Our analysis shows that, for the quadratic part of ℓ_{mat}, this loss regresses Δ_θ toward the damped target $d/(\beta(1 + \lambda_{\text{neg}}))$ (Appendix C.4), so β and λ_{neg} set the step size. The expectation in $\mathcal{L}_{\text{pair}}$ is over geometries, rollouts, and selected flow states. The matrix loss ℓ_{mat} combines MAE and scaled MSE and weights the atom-diagonal and interatomic blocks equally (Appendix E).

3.3 WAVEFUNCTION ALIGNMENT DISTILLATION

Matrix fitting weights entries regardless of their effect on the orbitals, whereas orbital energies and orbitals see the error through the orbital basis, so a small average Hamiltonian error can still leave inaccurate orbitals. We therefore adapt the wavefunction alignment (WA) loss of WANet (Li et al., 2025) to the corrections, using $\mathbf{H}^+$ in place of a converged label, and call the result the *WA distillation loss*. For each accepted correction, we solve $\mathbf{H}^+\mathbf{C}^+ = \mathbf{S}\mathbf{C}^+\boldsymbol{\epsilon}^+$ with $(\mathbf{C}^+)^\top\mathbf{S}\mathbf{C}^+ = \mathbf{I}$ and express the learner's error in the corrected orbital basis, $\mathbf{E}_\theta = (\mathbf{C}^+)^\top\hat{\mathbf{H}}_\theta\mathbf{C}^+ - (\mathbf{C}^+)^\top\mathbf{H}^+\mathbf{C}^+$, where the second term is the diagonal matrix $\boldsymbol{\epsilon}^+$. To first order in $\hat{\mathbf{H}}_\theta - \mathbf{H}^+$ and for nondegenerate orbitals, the learner's orbital energies and orbitals are $\epsilon_i(\hat{\mathbf{H}}_\theta) \approx \epsilon_i^+ + (\mathbf{E}_\theta)_{ii}$ and $\boldsymbol{c}_i(\hat{\mathbf{H}}_\theta) \approx \boldsymbol{c}_i^+ + \sum_{j\neq i}(\mathbf{E}_\theta)_{ji}\,\boldsymbol{c}_j^+/(\epsilon_i^+ - \epsilon_j^+)$. The diagonal of $\mathbf{E}_\theta$ thus carries the orbital-energy errors, and its off-diagonal entries carry the orbital mixing. The WA distillation loss penalizes both, averaging each over the B orbitals: $\mathcal{L}_{\text{WA}}(\theta) = \mathbb{E}\big[\gamma\big\{\frac{1}{B}\sum_i(\mathbf{E}_\theta)_{ii}^2 + \frac{1}{B(B-1)}\sum_{i\neq j}(\mathbf{E}_\theta)_{ij}^2\big\}\big]$. Its off-diagonal term weights all orbital pairs equally. Because $\mathbf{C}^+$ is computed once per flow state and held fixed, $\mathbf{E}_\theta$ is linear in $\hat{\mathbf{H}}_\theta$, and the loss requires no differentiation through an eigensolver.

3.4 TRAINING AND ACTOR REFRESH

To limit drift across rounds, the term $\mathcal{L}_{\text{base}}(\theta) = \mathbb{E}\big[\big\|\hat{\mathbf{H}}_\theta - \hat{\mathbf{H}}_{\text{base}}\big\|_F^2\big]$ penalizes the squared deviation from $\hat{\mathbf{H}}_{\text{base}}$, the clean prediction of a frozen copy of the pretrained model at the same flow state. Each learner step minimizes $\mathcal{L}_{\text{total}} = \lambda_{\text{pair}}\mathcal{L}_{\text{pair}} + \lambda_{\text{WA}}\mathcal{L}_{\text{WA}} + \lambda_{\text{base}}\mathcal{L}_{\text{base}}$ with nonnegative weights. Appendix E gives the exact reductions and weights.

OPSD keeps two policies: an actor, which generates the predictions that the solver corrects, and a learner, which is trained on the corrections. We keep them separate so that each correction has a

fixed reference. The learner moves away from the actor's prediction toward the corrected matrix, and the negative branch keeps this move small. If one model both generated and fitted the targets, the targets would shift with every gradient step. The actor stays frozen for n_{fit} learner steps. To save SCF calls, these steps sample a replay buffer of recent corrections from earlier actor snapshots, as is standard in online RL (Piché et al., 2026). We then refresh the actor by parameter averaging, $\theta_{k+1} = m\theta_k + (1-m)\theta$ with $0 \le m < 1$, where θ_k and θ are the actor and learner parameters, and the refreshed actor generates the next round. Algorithm 1 summarizes the procedure.

4 Theoretical justification

The method of Section 3 raises two questions. Where does repeated distillation converge when the network cannot fit every correction? And why learn from corrected matrices rather than from a scalar reward such as the residual? We answer the first for an idealized version of the method: one linearized SCF step, accepted and unclipped corrections from the current actor, a fixed m, a learner that fits quadratically in the response metric (Definition 1), and no MAE, WA, or $\mathcal{L}_{\text{base}}$ terms (Appendix C.1; proofs in Appendix C).

Algorithm 1 OPSD

Require: Geometries; pretrained θ_0
Learner $\theta \leftarrow \theta_0$; actor starts at θ_0
for rounds $k = 0, 1, \ldots$ **do**
  for n_{fit} steps, actor θ_k frozen **do**
    Actor: generate
    Run ODE solver: N_{roll} samples, L steps
    Keep low-noise flow states $\mathbf{H}_t$
    $\hat{\mathbf{H}} \leftarrow$ actor's clean prediction at $\mathbf{H}_t$
    Solver: correct
    $\mathbf{H}^{\text{scf}} \leftarrow b$ SCF iterations from $\hat{\mathbf{H}}$
    $\gamma \leftarrow 1$ if residual drops by $> \delta_r$, else 0
    $\mathbf{H}^{+} \leftarrow \hat{\mathbf{H}} + \text{clip}_{r_{\text{tr}}}(\mathbf{H}^{\text{scf}} - \hat{\mathbf{H}})$
    Solve for $\mathbf{C}^{+}, \boldsymbol{\epsilon}^{+}$ if $\gamma = 1$; add to buffer
    Learner: fit
    Sample buffer; learner predicts $\hat{\mathbf{H}}_\theta$
    Step θ on weighted $\mathcal{L}_{\text{pair}} + \mathcal{L}_{\text{WA}} + \mathcal{L}_{\text{base}}$
  end for
  Refresh actor: $\theta_{k+1} \leftarrow m\theta_k + (1-m)\theta$
end for
return learner θ

The answer to the first question depends on how fitting errors are weighted. To first order, a Hamiltonian error moves the density only through its occupied–virtual entries, entry (a, i) in proportion to $2/(\epsilon_a^\star - \epsilon_i^\star)$, so errors across small gaps matter most. In the response metric $\|\cdot\|_\chi$ with these weights (Definition 1), one step of the SCF map linearized at $\mathbf{H}^\star$, written Φ_{lin}, is exactly a gradient step on the quadratic DFT energy error Q. Let $\boldsymbol{x}$ denote the network's predicted occupied–virtual entries at a fixed set of flow states, and $\boldsymbol{x}^\star$ their self-consistent values. To first order in its parameters, the network can reach only predictions in an affine set $\mathcal{A} = \boldsymbol{x}_0 + \mathcal{V}$, which need not contain $\boldsymbol{x}^\star$.

Theorem 1 (Distillation as projected gradient descent)**.** *Let the eigenvalues of $A = \mathrm{I} - \mathrm{D}\Phi(\mathbf{H}^\star)$ on the occupied–virtual entries lie in $[\delta_{\text{stab}}, 1+\Lambda]$ with $\delta_{\text{stab}} > 0$. If the learner fits each round's corrections in the response metric $\|\cdot\|_\chi$ as closely as $\mathcal{A}$ allows, the actor follows $\boldsymbol{x}_{k+1} = \boldsymbol{x}_k + \eta\,\Pi_{\mathcal{V}}\big(\Phi_{\text{lin}}(\boldsymbol{x}_k) - \boldsymbol{x}_k\big)$ with $\eta = (1-m)/\big(\beta(1+\lambda_{\text{neg}})\big)$, where $\Pi_{\mathcal{V}}$ is the response-metric projection onto $\mathcal{V}$. This is projected gradient descent on $Q(\boldsymbol{x}) = \frac{1}{2}\langle \boldsymbol{x} - \boldsymbol{x}^\star, A(\boldsymbol{x} - \boldsymbol{x}^\star)\rangle_\chi$, and for $0 < \eta < 2/(1+\Lambda)$ it converges linearly to $\boldsymbol{x}^\dagger = \arg\min_{\boldsymbol{x}\in\mathcal{A}} Q(\boldsymbol{x})$.*

Theorem 1 supports the amortization view of Section 1: each round of distillation is one damped SCF iteration with damping η, restricted to the predictions the network can represent. Under this idealization, distillation minimizes the DFT energy error without converged labels or differentiation through the eigensolver. Because the limit minimizes Q over $\mathcal{A}$, a network that cannot reproduce $\boldsymbol{x}^\star$ spends its capacity on the errors that matter most for the energy. An entrywise loss on converged labels instead weights every entry equally. This holds under response-metric fitting. The implemented WA loss shares its orbital basis but not its weights, and Appendix C.7 shows how the fitting metric affects the result. Finally, distillation inherits the conditioning of the SCF problem, so systems that are hard for SCF itself need a smaller η and more rounds.

We next compare the SCF call complexity of two kinds of feedback, that is, the number of Fock builds needed to reach a given accuracy. As a matrix target, one correction specifies a complete update. As a scalar reward, it gives one number, so the learner needs more calls to recover the update. Theorem 2 uses Euclidean fitting and assumes that the model can represent the self-consistent solution and that each correction shrinks the error by a fixed factor. The matrix cost then does not depend on the model,

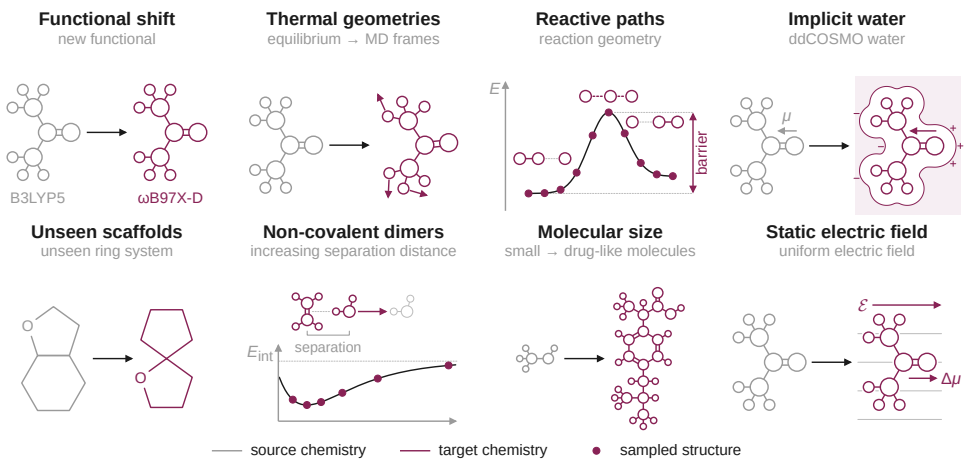


Figure 2: **Eight settings for adapting a pretrained Hamiltonian model.** The source model is trained on equilibrium QH9 molecules in the gas phase at B3LYP5/def2-SVP. The studies vary the electronic-structure functional, geometry, solvent, molecular scaffold, size, applied field, or intermolecular separation. Grey denotes the source setting and magenta the target; molecular drawings illustrate the shifts. Supervised fine-tuning (SFT) uses converged target labels, while OPSD uses molecular geometries and finite SCF corrections at the target conditions.

whereas the scalar cost can grow with the number of directions in which the model can change its prediction.

Theorem 2 (Solver calls for matrix and scalar feedback; informal)**.** *If the prediction family contains $\boldsymbol{x}^\star$ and a correction map costs c Fock builds and contracts errors by a factor $q_b < 1$, matrix feedback reaches relative mean-squared error ε^2 within $c\lceil \log(1/\varepsilon)/(\eta(1-q_b))\rceil$ Fock builds, whereas any residual-only algorithm needs at least $(1-\varepsilon^2)D$ in the worst case, where $D = \dim \mathcal{V}$ (full statement in Appendix C.1).*

It is worth noting that Theorem 2 counts Fock builds per geometry and leaves out optimization cost.

5 Experiments

QH9. QH9 (Yu et al., 2023a) contains B3LYP5/def2-SVP Hamiltonians for molecules with H, C, N, O, and F and up to nine heavy atoms. Its four tasks test random and size-based generalization on equilibrium structures (Stable-ID/OOD), and generalization to held-out geometries or molecules along molecular-dynamics trajectories (Dynamic-300k-Geo/Mol). Table 1 adds OPSD for QHFlow2-s, -l, and -xl on all four tasks to the published results. OPSD mainly improves the orbital spectrum. For the 12M-parameter QHFlow2-s, the Hamiltonian MAE changes by at most $1.8\,\mu E_h$ on any task, while the gap MAE falls by 27–70%. On Stable-ID, the LUMO MAE falls from 304.7 to $120.5\,\mu E_h$, and the HOMO MAE falls from 148.4 to $84.8\,\mu E_h$. After OPSD, QHFlow2-s has a lower gap MAE than the 43M-parameter QHFlow2-m on every task. RL and SCT start from the same backbone but lower the gap MAE by only 1.5–5.2%. OPSD also helps the larger models, lowering their gap MAE by 9–51% (QHFlow2-l) and 7–21% (QHFlow2-xl). With OPSD, QHFlow2-xl is best or tied for best in 23 of the 24 columns.

HELM. We use the OMol benchmark introduced with HELM (Kaniselvan et al., 2025), drawn from OMol25 (Levine et al., 2025). It provides ωB97M-V/def2-TZVPD Hamiltonians across 58 elements. We pretrain on `OMol_CSH_58k` and evaluate on `OMol_common_1k` and `OMol_all_5k` (common-1k and all-5k), covering eight common elements and all 58 elements, respectively. Post-training therefore starts from a model that already covers all 58 elements. Pretrained QHFlow2 achieves Hamiltonian MAEs of 705.10 and $1897.40\,\mu E_h$ on the two test sets, compared with HELM's published 961.64 and $2122.19\,\mu E_h$; HELM's all-5k value also includes 171 open-shell molecules. Post-training on HELM is a substantially harder engineering problem than on QH9, for two reasons.

First, OPSD corrections must converge toward the same electronic structure as the reference labels. We therefore generate them with ORCA (Neese et al., 2020), which computed OMol25, at the same ωB97M-V/def2-TZVPD level of theory and with the same effective core potentials. Each correction

Table 1: **Hamiltonian and orbital prediction on QH9.** Literature values and parameter counts were adapted from Table 2 of QHFlow2 (Kim et al., 2026) under their original training and inference settings. The RL and SCT rows are our runs from the same QHFlow2-s checkpoint. Hamiltonian and orbital-energy MAEs are in μE_h; ϵ_{occ} reports occupied-orbital energy MAE, and S_c is occupied-orbital coefficient cosine similarity in percent. Initialization settings and numerical precision follow the source. Bold marks the best value in each column. The -s/-m/-l/-xl suffixes indicate increasing QHFlow2 capacity. "–" denotes an unreported value.

		QH9-Stable-ID						QH9-Stable-OOD					
Method	Params	**H** ↓	ϵ_{occ} ↓	S_c ↑	HOMO↓	LUMO↓	Gap↓	**H** ↓	ϵ_{occ} ↓	S_c ↑	HOMO↓	LUMO↓	Gap↓
Regression													
QHNet (w/o init)	20M	77.7	963.5	94.8	1546.3	18257.3	17822.6	69.7	885.0	93.0	1046.0	25848.8	25370.1
QHNet	20M	46.0	459.6	98.3	470.6	2781.2	2967.9	38.1	170.5	97.8	255.6	1326.8	1359.9
WANet	29M	80.0	833.6	96.9	–	–	–	–	–	–	–	–	–
SPHNet	30M	45.5	334.3	97.8	–	–	–	43.3	186.4	98.2	–	–	–
QHNetV2 (w/o init)	–	31.5	417.9	98.6	–	–	–	23.0	165.9	97.7	–	–	–
Flow-based													
QHFlow	28M	23.0	119.7	99.5	179.5	438.0	553.9	20.0	84.5	99.0	130.7	321.2	395.8
QHFlow2-s	12M	16.3	94.9	99.6	148.4	304.7	411.4	12.5	54.6	99.4	72.8	182.9	200.2
QHFlow2-m	43M	9.2	59.2	99.7	77.7	175.7	216.0	7.8	42.1	99.5	49.5	134.8	147.3
QHFlow2-l	183M	5.9	45.7	**99.8**	50.5	103.3	119.6	5.2	31.9	**99.7**	33.7	99.0	105.9
QHFlow2-xl	990M	**4.6**	39.4	**99.8**	40.3	73.2	87.7	4.1	26.8	**99.7**	28.8	97.8	105.7
QHFlow2-s + RL (DiffusionNFT, SCF error reward)	12M	17.2	92.3	99.6	132.8	311.7	402.1	13.5	52.7	99.4	70.1	172.5	197.2
QHFlow2-s + SCT	12M	16.9	93.9	99.6	126.4	300.3	398.6	14.2	53.4	99.4	67.3	178.5	196.8
QHFlow2-s + OPSD	12M	16.5	76.0	99.7	84.8	120.5	147.8	12.5	53.2	99.4	54.6	141.1	145.4
QHFlow2-l + OPSD	183M	6.1	42.9	**99.8**	43.2	99.0	109.0	5.0	31.0	**99.7**	31.2	88.3	93.4
QHFlow2-xl + OPSD	990M	5.1	**37.1**	**99.8**	**37.2**	**55.4**	**69.1**	**4.0**	**26.3**	**99.7**	**26.7**	**87.3**	**88.0**

		QH9-Dynamic-300k-Geo						QH9-Dynamic-300k-Mol					
Method	Params	**H** ↓	ϵ_{occ} ↓	S_c ↑	HOMO↓	LUMO↓	Gap↓	**H** ↓	ϵ_{occ} ↓	S_c ↑	HOMO↓	LUMO↓	Gap↓
Regression													
QHNet (w/o init)	20M	88.4	1170.5	93.6	2040.1	23269.4	22408.0	121.4	5554.4	86.0	4352.8	53505.1	50424.9
WANet	29M	74.7	416.6	99.7	–	–	–	–	–	–	–	–	–
SPHNet	30M	52.2	100.9	99.1	–	–	–	108.2	1724.1	91.5	–	–	–
QHNetV2 (w/o init)	–	35.6	270.0	98.8	–	–	–	49.0	629.6	97.4	–	–	–
Flow-based													
QHFlow	28M	25.9	103.1	99.6	175.2	425.2	547.3	45.9	442.6	98.7	479.7	1344.7	1605.0
QHFlow2-s	12M	15.4	53.8	99.7	67.3	430.1	447.5	29.7	261.3	99.2	263.2	644.0	772.9
QHFlow2-m	43M	8.7	42.5	99.9	51.8	113.5	144.0	22.0	240.1	99.4	207.2	378.8	468.6
QHFlow2-l	183M	5.2	28.2	99.9	30.6	88.8	101.7	17.5	214.5	99.5	179.2	261.7	325.0
QHFlow2-xl	990M	3.7	22.8	**100.0**	21.0	44.4	50.6	15.9	202.3	**99.6**	167.4	219.5	278.1
QHFlow2-s + RL (DiffusionNFT, SCF error reward)	12M	15.9	62.8	99.6	65.3	420.2	434.2	31.4	251.2	99.3	252.2	632.1	732.4
QHFlow2-s + SCT	12M	14.9	63.3	99.6	70.6	422.4	438.8	33.5	257.3	99.3	259.5	633.3	740.3
QHFlow2-s + OPSD	12M	17.2	63.0	99.8	67.5	119.7	134.2	29.8	258.5	99.3	226.5	400.0	466.7
QHFlow2-l + OPSD	183M	5.1	26.0	99.9	24.1	45.0	49.8	17.5	212.5	99.5	169.2	214.4	265.2
QHFlow2-xl + OPSD	990M	**3.5**	**21.5**	**100.0**	**18.9**	**38.5**	**41.9**	**15.8**	**201.1**	**99.6**	**163.8**	**203.0**	**259.1**

is much more expensive than on QH9, so a CPU cluster coordinated by Ray (Moritz et al., 2018) serves the GPU learner asynchronously.

Second, the diffuse functions in def2-TZVPD make the overlap matrix $\mathbf{S}$ nearly singular; its smallest eigenvalues reach about 3×10^{-6}. Orbital energies solve $\mathbf{HC} = \mathbf{SC}\boldsymbol{\epsilon}$, so an atomic-orbital error $\delta\mathbf{H}$ reaches the spectrum through $\mathbf{S}^{-1/2}\,\delta\mathbf{H}\,\mathbf{S}^{-1/2}$ and is amplified by $1/s$ along each overlap eigenvector with eigenvalue s. Consequently, a prediction with errors of only $10^{-3}\,E_h$ per matrix element, which looks accurate under an element-wise loss, can yield a dozen or more spurious eigenvalues far below the core levels, and these fill the occupied space. This makes a stable learning signal on HELM nearly impossible to obtain from the raw atomic-orbital Hamiltonian.

To address this, we supervise orbital energies only in the well-conditioned part of the overlap space, while the model still predicts the full atomic-orbital Hamiltonian $\hat{\mathbf{H}}_\theta$. We diagonalize the overlap matrix, $\mathbf{S} = \mathbf{U}\,\mathrm{diag}(s)\mathbf{U}^\top$, where s collects its eigenvalues. We keep the eigenvectors $\mathbf{U}_{\text{kept}}$ whose eigenvalues s_{kept} exceed $\tau = 10^{-3}$ and set $\mathbf{Q} = \mathbf{U}_{\text{kept}}\,\mathrm{diag}(s_{\text{kept}})^{-1/2}$, so that $\mathbf{Q}^\top\mathbf{S}\mathbf{Q} = \mathbf{I}$. On the reference Hamiltonians, this cutoff changes orbital energies by only 6.4×10^{-4} eV on average. Let $\tilde{\boldsymbol{c}}_i$ be the eigenvectors of the accepted correction in this basis, $\mathbf{Q}^\top\mathbf{H}^+\mathbf{Q}$, in order of increasing energy. The projected WA loss penalizes the diagonal errors of the prediction in these orbitals, for all occupied orbitals and the LUMO: $\mathcal{L}_{\text{proj}} = \frac{1}{n_{\text{occ}}+1}\sum_{i=1}^{n_{\text{occ}}+1}\left|\tilde{\boldsymbol{c}}_i^\top\mathbf{Q}^\top(\hat{\mathbf{H}}_\theta - \mathbf{H}^+)\mathbf{Q}\tilde{\boldsymbol{c}}_i\right|$. This is the orbital-energy term of the WA distillation loss, restricted to the retained basis, the occupied orbitals, and the LUMO, with absolute errors.

Table 2 compares the pretrained model with OPSD using the vanilla and the projected WA loss; both post-trained models use 240 updates on 64 training molecules. We evaluate on the common-1k and all-5k test sets, reading out predictions at the same τ against the released ORCA reference with a numerical overlap floor of 10^{-8}. Relative to the vanilla WA loss, the projected WA loss lowers occupied-orbital MAE from 88.375 to 8.344 eV on common-1k and from 119.637 to 31.509 eV

Table 2: **Post-training on HELM.** Molecule-averaged orbital-energy MAEs (eV) and occupied-subspace overlap (Subspace), $\|\mathbf{C}_{\mathrm{ref,occ}}^{\top}\mathbf{S}\mathbf{C}_{\mathrm{pred,occ}}\|_F^2/n_{\mathrm{occ}} \in [0,1]$, which is invariant to rotations within either occupied space. Post-trained models use 240 updates on 64 molecules. All methods use the prediction overlap cutoff $\tau = 10^{-3}$ and the released ORCA reference with a numerical overlap floor of 10^{-8}. The common-1k set includes all 1,008 molecules. The all-5k set includes 4,815 of 4,986 released molecules, excluding 171 open-shell records. Bold marks the best value in each column.

	common-1k ($n = 1{,}008$)					**all-5k** ($n = 4{,}815$)				
Method	$\epsilon_{\mathrm{occ}}\downarrow$	HOMO↓	LUMO↓	Gap↓	Subspace↑	$\epsilon_{\mathrm{occ}}\downarrow$	HOMO↓	LUMO↓	Gap↓	Subspace↑
Pretrained	157.327	29.856	30.446	1.582	0.327	178.982	33.766	35.529	2.804	0.408
OPSD (vanilla WA loss)	88.375	13.384	14.687	**1.466**	0.465	119.637	17.483	19.900	2.749	0.504
OPSD (projected WA loss)	**8.344**	**2.202**	**3.610**	1.471	**0.880**	**31.509**	**5.137**	**7.655**	**2.642**	**0.810**

Table 3: **MD17 and rMD17.** Energy (E, meV) and force (F, meV Å^{-1}) MAEs. Best values in bold.

	MD17						**rMD17**					
	Ethanol		Malonaldehyde		Uracil		Salicylic acid		Naphthalene		Aspirin	
Model	E	F	E	F	E	F	E	F	E	F	E	F
DimeNet (Gasteiger et al., 2020)	10.32	1.84	12.98	1.16	10.68	2.13	12.21	4.06	11.92	2.83	26.04	5.45
GemNet-T (Gasteiger et al., 2021)	7.44	3.77	8.12	5.73	9.37	5.51	15.64	4.85	19.62	2.87	10.27	5.28
NequIP (Batzner et al., 2022)	0.33	0.61	0.73	0.88	0.45	**0.60**	1.71	1.43	2.01	**0.76**	1.87	2.08
QHFlow2-s (12M)	0.005	0.36	0.058	0.57	0.234	1.01	0.110	1.48	0.063	1.17	0.124	1.42
QHFlow2-m (43M)	0.003	0.34	0.010	0.51	0.062	0.90	0.070	1.38	0.040	1.10	0.084	1.35
QHFlow2-s + OPSD	0.0048	0.347	0.0568	0.561	0.2324	1.008	0.1054	**1.179**	0.0628	1.154	0.1228	1.269
QHFlow2-m + OPSD	**0.0029**	**0.331**	**0.0097**	**0.504**	**0.0616**	0.894	**0.0688**	1.235	**0.0395**	1.079	**0.0829**	**1.242**

on all-5k, while occupied-subspace overlap rises from 0.465 to 0.880 and from 0.504 to 0.810, respectively. HOMO and LUMO errors decrease on both sets, and the gap MAE on all-5k decreases from 2.749 to 2.642 eV.

MD17. For molecular-dynamics geometries from MD17 (Chmiela et al., 2017) and rMD17 (Christensen & von Lilienfeld, 2020), we follow QHFlow2 (Kim et al., 2026); for MD17, this uses the PBE/def2-SVP Hamiltonians computed for the SchNOrb benchmark (Schütt et al., 2019). We calculate energies and forces from the predicted Hamiltonian and report their MAEs against converged DFT (Table 3). OPSD lowers or matches the energy and force MAEs of both QHFlow2 models on all six molecules, and its forces are lower than NequIP's (Batzner et al., 2022) on four of them.

5.1 Label-free post-training transfers to new chemistry

We evaluate adaptation without converged target labels across eight shifts from the QH9 pretraining distribution: a different functional, thermal geometries, reaction paths, implicit water, unseen scaffolds, larger molecules, static electric fields, and non-covalent dimers (Figure 2). Each study starts from the same QHFlow2-s checkpoint pretrained on QH9-Stable-ID (Kim et al., 2026) and compares the unadapted model with supervised fine-tuning (SFT) and OPSD. SFT fits converged target Hamiltonians; OPSD learns from four-cycle SCF corrections. SFT matches the WA loss, learning rate, updates, and geometry exposure. All arms use one training seed. We evaluate the fixed final checkpoints on held-out confirmation sets. Appendix E.1 gives the objectives, readouts, and paired comparisons. Appendices E.2 and E.3 add the Fock-build budget, water clusters, and SCF iterations.

Adapting the frontier spectrum and reaction energies. The largest spectral shift comes from changing the functional to ωB97X-D (Figure 3a). Mean HOMO–LUMO gap error falls from 155.4 mHa before adaptation to 42.48 mHa with SFT and 2.16 mHa with OPSD: a 94.9% reduction relative to SFT, with lower error on 196 of 200 molecules. On thermal geometries, OPSD reduces gap MAE from SFT's 4.16 to 0.93 mHa and improves on SFT for 26 of 32 parent molecules (panel b). On Transition1x reaction paths (Schreiner et al., 2022), the sampled-path barrier MAE falls from 200.9 with SFT to 11.35 kcal mol^{-1} with OPSD (panel c).

Generalization across molecular structures. On unseen scaffolds, OPSD lowers gap MAE from the pretrained model's 0.580 to 0.398 mHa, improving 134 of 200 molecules (Figure 3d). The seen-scaffold control tests what the models retain after adapting to unseen scaffolds: SFT produces 27 of 199 gap errors above 10 mHa, whereas every OPSD error is below 1.5 mHa (panel e). For 30–40-atom

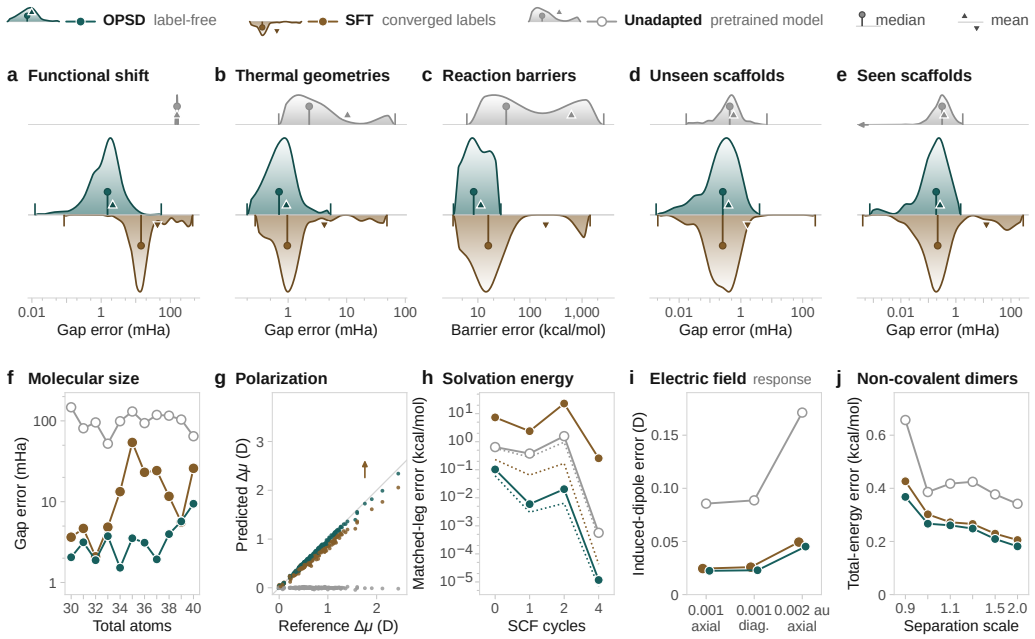


Figure 3: **Transfer after up to 2,000 adaptation updates.** Teal: OPSD; brown: SFT; grey: unadapted. **a–e**, Error distributions for functional transfer (200 molecules), thermal geometries (32 parent means, four frames each), sampled-path barriers (20 paths), and unseen/seen scaffolds (200/199 molecules). Densities are estimated in log-error space and normalized to each arm's peak; circles mark medians, triangles arithmetic means, and end caps the observed range. The arrow in **e** marks one unadapted error below the axis range. **f**, Mean gap error at each atom count (64 molecules total). **g**, Predicted versus reference gas-to-water dipole shift; the arrow marks one SFT prediction at 18.6 D. **h**, Electrostatic solvation-energy error, labeled matched-leg error, after 0, 1, 2, and 4 SCF cycles, using the water model for the water state and the unadapted gas model for the gas state; solid/dotted lines denote means/medians over 200 molecules. **i**, Mean induced-dipole vector error over 64 molecules at the training fields, unseen directions, and unseen magnitude. **j**, Mean total-energy error over ten dimer pairs at each separation, before SCF iteration. Table 6 reports the primary endpoints.

molecules, OPSD reduces mean gap error from 99.47 mHa before adaptation to 3.51 mHa, compared with 15.25 mHa for SFT. Its mean is lower at 10 of the 11 atom counts, with lower error on 38 of 64 individual molecules (panel f).

Polarization and energetics in new environments. In implicit water, OPSD reduces the dipole error from 0.182 to 0.022 D relative to SFT, improving 193 of 200 molecules (Figure 3g). Using each environment's own predicted electronic state, the mean electrostatic solvation-energy error is 0.103 kcal mol^{-1} before SCF iteration; all 200 molecules are within 1 kcal mol^{-1}, compared with 192 for SFT (panel h). Field-induced dipole error falls from SFT's 0.0245 to 0.0224 D at the six training fields, and the lower mean error extends to unseen field directions and twice the field strength (panel i). For non-covalent dimers, OPSD gives the lowest mean total-energy error at each of the six sampled separations (panel j); the corresponding interaction-energy errors are reported in Table 6.

6 Conclusion

OPSD lets a pretrained Hamiltonian model learn from SCF corrections of its own predictions, using only geometries and standard SCF software. In all eight transfer settings, each trained with a single seed, OPSD has a lower mean error than fine-tuning on converged labels. After a change of functional, its gap error is 2.16 mHa, compared with 42.48 mHa for fine-tuning. OPSD also lowers the QH9 gap error of QHFlow2-s, -l, and -xl by up to 70% and stabilizes the orbital spectrum of a QHFlow2 model pretrained on the 58-element HELM benchmark. OPSD could thus serve as a general method for improving the transferability of Hamiltonian models.

REFERENCES

Rishabh Agarwal, Nino Vieillard, Yongchao Zhou, Piotr Stanczyk, Sabela Ramos Garea, Matthieu Geist, and Olivier Bachem. On-Policy Distillation of Language Models: Learning from Self-Generated Mistakes. In *International Conference on Learning Representations*, 2024.

Simon Batzner, Albert Musaelian, Lixin Sun, Mario Geiger, Jonathan P Mailoa, Mordechai Kornbluth, Nicola Molinari, Tess E Smidt, and Boris Kozinsky. E(3)-Equivariant Graph Neural Networks for Data-Efficient and Accurate Interatomic Potentials. *Nature Communications*, 13(1):2453, 2022.

Lori A Burns, John C Faver, Zheng Zheng, Michael S Marshall, Daniel GA Smith, Kenno Vanommeslaeghe, Alexander D MacKerell, Kenneth M Merz, and C David Sherrill. The BioFragment Database (BFDb): An Open-Data Platform for Computational Chemistry Analysis of Noncovalent Interactions. *Journal of Chemical Physics*, 147(16), 2017.

Eric Cancès and Claude Le Bris. On the Convergence of SCF Algorithms for the Hartree-Fock Equations. *ESAIM: Mathematical Modelling and Numerical Analysis*, 34(4):749–774, 2000.

Eric Cancès, Gaspard Kemlin, and Antoine Levitt. Convergence Analysis of Direct Minimization and Self-Consistent Iterations. *SIAM Journal on Matrix Analysis and Applications*, 42(1):243–274, 2021.

Stefan Chmiela, Alexandre Tkatchenko, Huziel E Sauceda, Igor Poltavsky, Kristof T Schütt, and Klaus-Robert Müller. Machine Learning of Accurate Energy-Conserving Molecular Force Fields. *Science Advances*, 3(5):e1603015, 2017.

Anders S Christensen and O Anatole von Lilienfeld. On the Role of Gradients for Machine Learning of Molecular Energies and Forces. *Machine Learning: Science and Technology*, 1(4):045018, 2020.

Chandler Davis and William Morton Kahan. The Rotation of Eigenvectors by a Perturbation. III. *SIAM Journal on Numerical Analysis*, 7(1):1–46, 1970.

Peter Eastman, Pavan Kumar Behara, David L Dotson, Raimondas Galvelis, John E Herr, Josh T Horton, Yuezhi Mao, John D Chodera, Benjamin P Pritchard, Yuanqing Wang, et al. SPICE, a Dataset of Drug-Like Molecules and Peptides for Training Machine Learning Potentials. *Scientific Data*, 10(1):11, 2023.

John C Faver, Mark L Benson, Xiao He, Benjamin P Roberts, Bing Wang, Michael S Marshall, Matthew R Kennedy, C David Sherrill, and Kenneth M Merz Jr. Formal Estimation of Errors in Computed Absolute Interaction Energies of Protein-Ligand Complexes. *Journal of Chemical Theory and Computation*, 7(3):790, 2011.

Johannes Gasteiger, Janek Groß, and Stephan Günnemann. Directional Message Passing for Molecular Graphs. *arXiv preprint arXiv:2003.03123*, 2020.

Johannes Gasteiger, Florian Becker, and Stephan Günnemann. GemNet: Universal Directional Graph Neural Networks for Molecules. *Advances in Neural Information Processing Systems*, 34: 6790–6802, 2021.

Lucie Gráfová, Michal Pitoňák, Jan Řezáč, and Pavel Hobza. Comparative Study of Selected Wave Function and Density Functional Methods for Noncovalent Interaction Energy Calculations Using the Extended S22 Data Set. *Journal of Chemical Theory and Computation*, 6(8):2365–2376, 2010.

Jean-Bastien Grill, Florian Strub, Florent Altché, Corentin Tallec, Pierre Richemond, Elena Buchatskaya, Carl Doersch, Bernardo Avila Pires, Zhaohan Guo, Mohammad Gheshlaghi Azar, et al. Bootstrap Your Own Latent: A New Approach to Self-Supervised Learning. *Advances in Neural Information Processing Systems*, 33:21271–21284, 2020.

Pierre Hohenberg and Walter Kohn. Inhomogeneous Electron Gas. *Physical Review*, 136(3B):B864, 1964.

Roger A. Horn and Charles R. Johnson. *Matrix Analysis*. Cambridge University Press, 2nd edition, 2013.

Petr Jurečka, Jiří Šponer, Jiří Černý, and Pavel Hobza. Benchmark Database of Accurate (MP2 and CCSD(T) Complete Basis Set Limit) Interaction Energies of Small Model Complexes, DNA Base Pairs, and Amino Acid Pairs. *Physical Chemistry Chemical Physics*, 8(17):1985–1993, 2006.

Manasa Kaniselvan, Benjamin Kurt Miller, Meng Gao, Juno Nam, and Daniel S Levine. Learning from the Electronic Structure of Molecules across the Periodic Table. *arXiv preprint arXiv:2510.00224*, 2025.

GP Kerker. Efficient Iteration Scheme for Self-Consistent Pseudopotential Calculations. *Physical Review B*, 23(6):3082, 1981.

Kuzma Khrabrov, Anton Ber, Artem Tsypin, Konstantin Ushenin, Egor Rumiantsev, Alexander Telepov, Dmitry Protasov, Ilya Shenbin, Anton Alekseev, Mikhail Shirokikh, et al. ∇^2DFT: A Universal Quantum Chemistry Dataset of Drug-Like Molecules and a Benchmark for Neural Network Potentials. *Advances in Neural Information Processing Systems*, 37:36869–36889, 2024.

Seongsu Kim, Nayoung Kim, Dongwoo Kim, and Sungsoo Ahn. High-Order Equivariant Flow Matching for Density Functional Theory Hamiltonian Prediction. *Advances in Neural Information Processing Systems*, 38:13265–13307, 2025.

Seongsu Kim, Chanhui Lee, Yoonho Kim, Seongjun Yun, Honghui Kim, Nayoung Kim, Changyoung Park, Sehui Han, Sungbin Lim, and Sungsoo Ahn. Machine Learning Hamiltonians Are Accurate Energy-Force Predictors. *arXiv preprint arXiv:2602.16897*, 2026.

Walter Kohn and Lu Jeu Sham. Self-Consistent Equations including Exchange and Correlation Effects. *Physical Review*, 140(4A):A1133, 1965.

Konstantin N Kudin, Gustavo E Scuseria, and Eric Cancès. A Black-Box Self-Consistent Field Convergence Algorithm: One Step Closer. *Journal of Chemical Physics*, 116(19):8255–8261, 2002.

Susi Lehtola. Assessment of Initial Guesses for Self-Consistent Field Calculations. Superposition of Atomic Potentials: Simple yet Efficient. *Journal of Chemical Theory and Computation*, 15(3): 1593, 2019.

Daniel S Levine, Muhammed Shuaibi, Evan Walter Clark Spotte-Smith, Michael G Taylor, Muhammad R Hasyim, Kyle Michel, Ilyes Batatia, Gábor Csányi, Misko Dzamba, Peter Eastman, et al. The Open Molecules 2025 (OMol25) Dataset, Evaluations, and Models. *arXiv preprint arXiv:2505.08762*, 2025.

He Li, Zun Wang, Nianlong Zou, Meng Ye, Runzhang Xu, Xiaoxun Gong, Wenhui Duan, and Yong Xu. Deep-Learning Density Functional Theory Hamiltonian for Efficient Ab Initio Electronic-Structure Calculation. *Nature Computational Science*, 2(6):367–377, 2022.

Yunyang Li, Zaishuo Xia, Lin Huang, Xinran Wei, Samuel Harshe, Han Yang, Erpai Luo, Zun Wang, Jia Zhang, Chang Liu, et al. Enhancing the Scalability and Applicability of Kohn-Sham Hamiltonians for Molecular Systems. In *International Conference on Learning Representations*, 2025.

Yunyang Li, Lin Huang, Luojia Xia, Wenhe Zhang, and Mark Gerstein. Elign: Equivariant Diffusion Model Alignment from Foundational Machine Learning Force Fields. *arXiv preprint arXiv:2601.21985*, 2026.

Filippo Lipparini, Benjamin Stamm, Eric Cancès, Yvon Maday, and Benedetta Mennucci. Fast Domain Decomposition Algorithm for Continuum Solvation Models: Energy and First Derivatives. *Journal of Chemical Theory and Computation*, 9(8):3637–3648, 2013.

Jie Liu, Gongye Liu, Jiajun Liang, Yangguang Li, Jiaheng Liu, Xintao Wang, Pengfei Wan, Di Zhang, and Wanli Ouyang. Flow-GRPO: Training Flow Matching Models via Online RL. *arXiv preprint arXiv:2505.05470*, 2025.

Xin Liu, Xiao Wang, Zaiwen Wen, and Yaxiang Yuan. On the Convergence of the Self-Consistent Field Iteration in Kohn–Sham Density Functional Theory. *SIAM Journal on Matrix Analysis and Applications*, 35(2):546–558, 2014.

Yunhong Lou, Xihang Yue, Xinran Wei, Tianqi Deng, and Linchao Zhu. A Fixed-Point Neural Operator for Size- and Functional-Transferable Hamiltonian Prediction. *arXiv preprint arXiv:2606.14498*, 2026.

Kevin Lu and Thinking Machines Lab. On-Policy Distillation. *Thinking Machines Lab: Connectionism*, 2025. doi: 10.64434/tml.20251026. https://thinkingmachines.ai/blog/on-policy-distillation.

Erpai Luo, Xinran Wei, Lin Huang, Yunyang Li, Han Yang, Zaishuo Xia, Zun Wang, Chang Liu, Bin Shao, and Jia Zhang. Efficient and Scalable Density Functional Theory Hamiltonian Prediction through Adaptive Sparsity. *arXiv preprint arXiv:2502.01171*, 2025a.

Jianlan Luo, Charles Xu, Jeffrey Wu, and Sergey Levine. Precise and Dexterous Robotic Manipulation via Human-in-the-Loop Reinforcement Learning. *Science Robotics*, 10(105):eads5033, 2025b.

Roy Mathias. Quadratic Residual Bounds for the Hermitian Eigenvalue Problem. *SIAM Journal on Matrix Analysis and Applications*, 19(2):541–550, 1998.

Philipp Moritz, Robert Nishihara, Stephanie Wang, Alexey Tumanov, Richard Liaw, Eric Liang, Melih Elibol, Zongheng Yang, William Paul, Michael I Jordan, et al. Ray: A Distributed Framework for Emerging AI Applications. In *13th USENIX Symposium on Operating Systems Design and Implementation (OSDI 18)*, pp. 561–577, 2018.

Frank Neese, Frank Wennmohs, Ute Becker, and Christoph Riplinger. The ORCA Quantum Chemistry Program Package. *Journal of Chemical Physics*, 152(22), 2020.

Yurii Nesterov and Vladimir Spokoiny. Random Gradient-Free Minimization of Convex Functions. *Foundations of Computational Mathematics*, 17(2):527–566, 2017.

Anders MN Niklasson. Extended Born-Oppenheimer Molecular Dynamics. *Physical Review Letters*, 100(12):123004, 2008.

Hyunsoo Park and Aron Walsh. Guiding Generative Models to Uncover Diverse and Novel Crystals via Reinforcement Learning. *Nature Machine Intelligence*, 2026.

Alexandre Piché, Ehsan Kamalloo, Rafael Pardinas, Xiaoyin Chen, and Dzmitry Bahdanau. PipelineRL: Faster On-Policy Reinforcement Learning for Long Sequence Generation. *Transactions on Machine Learning Research*, 2026.

Péter Pulay. Convergence Acceleration of Iterative Sequences. The Case of SCF Iteration. *Chemical Physics Letters*, 73(2):393–398, 1980.

Peter Pulay. Improved SCF Convergence Acceleration. *Journal of Computational Chemistry*, 3(4): 556–560, 1982.

Peter Pulay and Géza Fogarasi. Fock Matrix Dynamics. *Chemical Physics Letters*, 386(4-6):272–278, 2004.

Jan Řezáč and Pavel Hobza. Describing Noncovalent Interactions beyond the Common Approximations: How Accurate Is the "Gold Standard," CCSD(T) at the Complete Basis Set Limit? *Journal of Chemical Theory and Computation*, 9(5):2151–2155, 2013.

Thorsten Rohwedder and Reinhold Schneider. An Analysis for the DIIS Acceleration Method Used in Quantum Chemistry Calculations. *Journal of Mathematical Chemistry*, 49(9):1889–1914, 2011.

Clemens Carel Johannes Roothaan. New Developments in Molecular Orbital Theory. *Reviews of Modern Physics*, 23(2):69, 1951.

Victor R Saunders and Ian H Hillier. A "Level–Shifting" Method for Converging Closed Shell Hartree–Fock Wave Functions. *International Journal of Quantum Chemistry*, 7(4):699–705, 1973.

Mathias Schreiner, Arghya Bhowmik, Tejs Vegge, Jonas Busk, and Ole Winther. Transition1x – a Dataset for Building Generalizable Reactive Machine Learning Potentials. *Scientific Data*, 9(1): 779, 2022.

Kristof T Schütt, Michael Gastegger, Alexandre Tkatchenko, Klaus-Robert Müller, and Reinhard J Maurer. Unifying Machine Learning and Quantum Chemistry with a Deep Neural Network for Molecular Wavefunctions. *Nature Communications*, 10(1):5024, 2019.

Daniel GA Smith, Lori A Burns, Andrew C Simmonett, Robert M Parrish, Matthew C Schieber, Raimondas Galvelis, Peter Kraus, Holger Kruse, Roberto Di Remigio, Asem Alenaizan, et al. PSI4 1.4: Open-Source Software for High-Throughput Quantum Chemistry. *Journal of Chemical Physics*, 152(18), 2020.

Qiming Sun, Xing Zhang, Samragni Banerjee, Peng Bao, Marc Barbry, Nick S Blunt, Nikolay A Bogdanov, George H Booth, Jia Chen, Zhi-Hao Cui, et al. Recent Developments in the PySCF Program Package. *Journal of Chemical Physics*, 153(2), 2020.

Antti Tarvainen and Harri Valpola. Mean Teachers Are Better Role Models: Weight-Averaged Consistency Targets Improve Semi-Supervised Deep Learning Results. *Advances in Neural Information Processing Systems*, 30, 2017.

Oliver Unke, Mihail Bogojeski, Michael Gastegger, Mario Geiger, Tess Smidt, and Klaus-Robert Müller. SE(3)-Equivariant Prediction of Molecular Wavefunctions and Electronic Densities. *Advances in Neural Information Processing Systems*, 34:14434–14447, 2021.

JH van Lenthe, Renate Zwaans, Huub JJ van Dam, and MF Guest. Starting SCF Calculations by Superposition of Atomic Densities. *Journal of Computational Chemistry*, 27(8):926–932, 2006.

Homer F Walker and Peng Ni. Anderson Acceleration for Fixed-Point Iterations. *SIAM Journal on Numerical Analysis*, 49(4):1715–1735, 2011.

Haiyang Yu, Meng Liu, Youzhi Luo, Alex Strasser, Xiaofeng Qian, Xiaoning Qian, and Shuiwang Ji. QH9: A Quantum Hamiltonian Prediction Benchmark for QM9 Molecules. *Advances in Neural Information Processing Systems*, 36:40487–40503, 2023a.

Haiyang Yu, Zhao Xu, Xiaofeng Qian, Xiaoning Qian, and Shuiwang Ji. Efficient and Equivariant Graph Networks for Predicting Quantum Hamiltonian. In *International Conference on Machine Learning*, pp. 40412–40424. PMLR, 2023b.

Haiyang Yu, Yuchao Lin, Xuan Zhang, Xiaofeng Qian, and Shuiwang Ji. Efficient Prediction of SO(3)-Equivariant Hamiltonian Matrices via SO(2) Local Frames. *arXiv preprint arXiv:2506.09398*, 2025.

Michael C Zerner and Michael Hehenberger. A Dynamical Damping Scheme for Converging Molecular SCF Calculations. *Chemical Physics Letters*, 62(3):550–554, 1979.

He Zhang, Chang Liu, Zun Wang, Xinran Wei, Siyuan Liu, Nanning Zheng, Bin Shao, and Tie-Yan Liu. Self-Consistency Training for Density-Functional-Theory Hamiltonian Prediction. *arXiv preprint arXiv:2403.09560*, 2024.

Siyan Zhao, Zhihui Xie, Mengchen Liu, Jing Huang, Guan Pang, Feiyu Chen, and Aditya Grover. Self-Distilled Reasoner: On-Policy Self-Distillation for Large Language Models. *arXiv preprint arXiv:2601.18734*, 2026.

Kaiwen Zheng, Huayu Chen, Haotian Ye, Haoxiang Wang, Qinsheng Zhang, Kai Jiang, Hang Su, Stefano Ermon, Jun Zhu, and Ming-Yu Liu. DiffusionNFT: Online Diffusion Reinforcement with Forward Process. In *International Conference on Learning Representations*, 2026.

Wei Zhou, Xiongwei Zhu, Lingdong Kong, Bo Chen, Lei Zhang, Yongyuan Liang, Xiaoxia Hou, Ye Tian, Xian Sun, Yingshuo Wang, et al. On-Policy Self-Distillation in Diffusion Models. *arXiv preprint arXiv:2608.24646*, 2026.

APPENDIX

Contents

A Notation **15**

B Related work **16**

C Theory and proofs **17**

C.1 Setting and interpretation of Section 4 . . . 17
C.2 Linear response at the fixed point . . . 18
C.3 Content of one SCF pass . . . 19
C.4 Branch geometry and the pointwise update . . . 20
C.5 Projected fitting in a shared prediction space . . . 20
C.6 The response operator and the proof of Theorem 1 . . . 21
C.7 A contracting SCF response can expand after fitting . . . 22
C.8 Contraction of the damped SCF map . . . 22
C.9 Finite-budget teachers . . . 23
C.10 Scalar residual feedback and the comparison of solver calls . . . 24
C.11 Numerical checks . . . 25

D Equivariance of the correction targets **25**

E Protocol details **26**

E.1 Transfer experiments . . . 27
E.2 Correction budget and water clusters . . . 29
E.3 SCF convergence from the predictions . . . 29

A NOTATION

Symbols used only within one proof are defined where they appear. The last column gives the section, appendix, or definition that introduces each symbol.

Symbol	Meaning	Defined
Molecules and Kohn–Sham DFT		
$\mathcal{M}$	Molecule: atomic numbers and nuclear positions	§2
N, n_{occ}	Electron count; number of doubly occupied orbitals, $n_{\mathrm{occ}} = N/2$	§2
B	Size of the atom-centered basis	§2
$\mathbf{S}$	Overlap matrix	§2
$\mathbf{H}$	Kohn–Sham Hamiltonian, a real symmetric $B \times B$ matrix	§2
$\mathbf{C}$, $\mathbf{C}_{\mathrm{occ}}$	Orbital coefficients, $\mathbf{HC} = \mathbf{SC}\boldsymbol{\epsilon}$ with $\mathbf{C}^\top\mathbf{SC} = \mathbf{I}$; their first n_{occ} columns	§2
$\boldsymbol{\epsilon}$, ϵ_p	Diagonal matrix of orbital energies; the pth orbital energy, in increasing order	§2
ϵ_Δ	HOMO–LUMO gap $\epsilon_{n_{\mathrm{occ}}+1} - \epsilon_{n_{\mathrm{occ}}}$	§2
$\mathbf{P}(\mathbf{H})$	Closed-shell density matrix $2\mathbf{C}_{\mathrm{occ}}\mathbf{C}_{\mathrm{occ}}^\top$	§2
$\mathbf{F}(\mathbf{P})$	Fock build $\mathbf{H}_{\mathrm{core}} + \mathbf{J}(\mathbf{P}) + \mathbf{V}_{\mathrm{xc}}(\mathbf{P})$ from the core, Coulomb, and exchange–correlation matrices; the unit of solver cost	§2
Φ	SCF map, $\Phi(\mathbf{H}) = \mathbf{F}(\mathbf{P}(\mathbf{H}))$	§2
$\mathbf{H}^\star$	Self-consistent Hamiltonian, $\mathbf{H}^\star = \Phi(\mathbf{H}^\star)$	§2
$D(g)$	Orthogonal basis representation of a rotation or atom permutation g	§2
Flow model		
θ	Network parameters; the learner's parameters during post-training	§2
t, $\mathbf{H}_t$	Flow time; noisy state	§2
$(\mathbf{H}_t, \mathcal{M}, t)$	Flow state	§2
$\hat{\mathbf{H}}_\theta$	Clean Hamiltonian prediction at a flow state	§2
v_θ	Velocity, $v_\theta = (\hat{\mathbf{H}}_\theta - \mathbf{H}_t)/(1-t)$	§2
$\mathbf{H}^{(0)}$, $\mathbf{H}_{\mathrm{prior}}$, σ	Fock build from a MINAO density; noisy prior draw around it; prior noise scale	App. E
L, N_{roll}	Integration steps per trajectory; trajectories per geometry	§3.1
SCF corrections		
$\hat{\mathbf{H}}$, $\hat{\mathbf{H}}_{\theta_k}$	Actor's clean prediction at a stored flow state, which the solver corrects ($\hat{\mathbf{H}}_{\theta_k}$ in Appendix C.4)	§3.1
b	DIIS-accelerated SCF cycles per correction	§3.1
$\mathbf{H}^{\mathrm{scf}}$	Candidate after b cycles from the density $\mathbf{P}(\hat{\mathbf{H}})$	§3.1
r_{AO}	Root-mean-square entry of the SCF residual, $\|\Phi(\mathbf{H}) - \mathbf{H}\|_F / B$	§3.1
δ_r, γ	Required residual improvement; acceptance indicator	§3.1
r_{tr}	Trust radius on the correction displacement	§3.1
d	Clipped correction displacement, $d = \mathbf{H}^+ - \hat{\mathbf{H}}$	§3.1
$\mathbf{H}^+$	Training target, the corrected Hamiltonian $\hat{\mathbf{H}} + d$	§3.1
Learning objective and actor refresh		
Δ_θ	Learner displacement from the actor, $\hat{\mathbf{H}}_\theta - \hat{\mathbf{H}}$	§3.2
β, $\mathbf{H}_\theta^\pm$	Branch step; positive and negative branch predictions $\hat{\mathbf{H}} \pm \beta\Delta_\theta$	§3.2
λ_{neg}	Weight of the negative branch	§3.2
ℓ_{mat}, $\mathcal{L}_{\mathrm{pair}}$	Matrix loss, MAE plus scaled MSE per block collection; paired loss	§3.2
s_a, sg	Detached scale of block collection a, shared by both branches; stop-gradient	App. E
$\mathbf{C}^+$, $\boldsymbol{\epsilon}^+$	Orbitals and orbital energies of $\mathbf{H}^+$	§3.3
$\mathbf{E}_\theta$	Learner error in the corrected orbital basis, $(\mathbf{C}^+)^\top\hat{\mathbf{H}}_\theta\mathbf{C}^+ - \boldsymbol{\epsilon}^+$	§3.3
$\mathcal{L}_{\mathrm{WA}}$	WA distillation loss	§3.3
$\hat{\mathbf{H}}_{\mathrm{base}}$, $\mathcal{L}_{\mathrm{base}}$	Prediction of the frozen pretrained model; penalty on the deviation from it	§3.4
λ_{pair}, λ_{WA}, λ_{base}	Loss weights in $\mathcal{L}_{\mathrm{total}}$	§3.4

Symbol	Meaning	Defined
θ_k, m	Actor parameters in round k; averaging coefficient of the refresh $\theta_{k+1} = m\theta_k + (1-m)\theta$	§3.4
n_{fit}	Learner steps per round	§3.4
Theory		
$\boldsymbol{x}$, $\boldsymbol{x}^\star$	Predicted occupied–virtual entries at fixed flow states (all represented entries in App. C.5 and C.9); their self-consistent values	§4
$\mathcal{A}$, $\mathcal{V}$	Predictions reachable to first order in the parameters, $\mathcal{A} = \boldsymbol{x}_0 + \mathcal{V}$; the reachable changes	§4
Φ_{lin}	SCF map linearized at $\mathbf{H}^\star$	§4
A, δ_{stab}, Λ	$\mathrm{I} - \mathrm{D}\Phi(\mathbf{H}^\star)$ on the occupied–virtual entries; bounds of its spectrum $[\delta_{\mathrm{stab}}, 1+\Lambda]$	§4
$\langle\cdot,\cdot\rangle_\chi$, $\|\cdot\|_\chi$	SCF response inner product and norm	Def. 1
$\Pi_{\mathcal{V}}$	Projection onto $\mathcal{V}$, in the response metric (Theorem 1) or the fitting metric W	§4
η	Step size of the learned iteration, $(1-m)/(\beta(1+\lambda_{\mathrm{neg}}))$	§4
Q, $\boldsymbol{x}^\dagger$	Quadratic DFT energy error; its minimizer over $\mathcal{A}$	§4
c, D, ε	Fock-build cost of a correction map; $\dim \mathcal{V}$; target relative error (Theorem 2)	§4
W	Fitting metric obtained by freezing the detached branch scales	App. C.4
Φ_α, α	Damped SCF map $(1-\alpha)\mathbf{H} + \alpha\Phi(\mathbf{H})$; damping factor	App. C.5
J_α	Derivative of Φ_α at $\mathbf{H}^\star$	App. C.2
$\mathcal{T}_b^{\mathrm{lin}}$, J_b	Linearized b-pass corrector $\boldsymbol{x}^\star + \mathrm{J}_b(\boldsymbol{x} - \boldsymbol{x}^\star)$; its linear part	App. C.5
$\mathcal{T}$	Correction map (teacher)	App. C.1
q_b	Error-reduction factor of $\mathcal{T}$ in the fitting norm	§4
q	Per-pass contraction factor	App. C.5
$e_{\mathcal{A}}$	Smallest error that $\mathcal{A}$ can represent	App. C.9
$\mathrm{Sym}(B)$, $\|\cdot\|_{\mathbf{S}}$	Real symmetric $B \times B$ matrices; Frobenius norm in the orthogonalized basis	App. C.2
δ, ω_{ia}	HOMO–LUMO gap of $\mathbf{H}^\star$; occupied–virtual energy differences $\epsilon_a^\star - \epsilon_i^\star \geq \delta$	App. C.2
$\mathcal{O}$, E_{ia}, Π_{ov}	Occupied–virtual subspace of $\mathrm{Sym}(B)$; its orthonormal basis; the orthogonal projection onto it	App. C.2
χ_0, χ_+	Non-interacting density response; its positive restriction $-\chi_0\|_{\mathcal{O}}$	App. C.2
$\mathcal{K}$, $\mathcal{K}_{\mathrm{ov}}$, $\mathcal{K}_{\perp,\mathrm{ov}}$	Hartree–exchange–correlation kernel; its changes within $\mathcal{O}$ and in $\mathcal{O}^\perp$	App. C.2
A_{full}	Jacobian $\mathrm{I} - \mathcal{K}\chi_0$ of the SCF residual $\mathbf{H} - \Phi(\mathbf{H})$ at $\mathbf{H}^\star$	App. C.6
A_{ov}	Occupied–virtual block of $\mathrm{I} - \mathcal{K}\chi_0$, the A of Section 4	App. C.6
$\|\cdot\|_\star$, r_{loc}	Local contraction norm and radius	App. C.8
Experiments		
ϵ_{occ}, S_c	Occupied-orbital energy MAE; occupied-orbital coefficient cosine similarity (%)	§5
τ, s, $\mathbf{Q}$	HELM overlap cutoff; overlap eigenvalues; orthonormalizing basis of the overlap directions above τ	§5
Subspace	HELM occupied-subspace overlap $\left\|\mathbf{C}_{\mathrm{ref,occ}}^\top \mathbf{S} \mathbf{C}_{\mathrm{pred,occ}}\right\|_F^2 / n_{\mathrm{occ}}$	§5
$\mathcal{L}_{\mathrm{proj}}$	Projected WA loss on HELM, in the retained overlap space	§5

B Related work

SCF convergence. The cost of a Kohn–Sham calculation is the number of Fock builds, set by the guess and by the convergence algorithm. Guesses come from atomic densities or potentials (van Lenthe et al., 2006; Lehtola, 2019) and, along trajectories, from extrapolation of previous steps (Pulay & Fogarasi, 2004; Niklasson, 2008). The iteration is stabilized by damping and level shifting (Zerner & Hehenberger, 1979; Saunders & Hillier, 1973), preconditioned against charge sloshing (Kerker, 1981), and accelerated by DIIS and its variants (Pulay, 1980; 1982; Kudin et al., 2002). Its convergence has been analyzed for damped and optimal-damping algorithms (Cancès & Le Bris,

2000) and under conditions relating the gap to the second derivative of the exchange–correlation functional (Liu et al., 2014). The contraction result in Appendix C.8 belongs to this line; it identifies a local contraction norm and a damping bound for the damped SCF map. Appendix C.7 then examines whether projection in the learning metric preserves that contraction.

Learned Hamiltonians. SchNOrb, PhiSNet, and DeepH established neural Hamiltonian prediction (Schütt et al., 2019; Unke et al., 2021; Li et al., 2022), and QHNet and QHNetV2 made it efficient (Yu et al., 2023b; 2025). WANet introduced the wavefunction alignment loss and the PubChemQH dataset, which exposed the divergence between matrix error and usefulness at scale (Li et al., 2025). SPHNet added adaptive sparsity (Luo et al., 2025a), and HELM scaled the task to 58 elements and diffuse triple-zeta bases (Kaniselvan et al., 2025). QHFlow replaced regression by conditional flow matching and showed that sampled matrices reduce SCF iterations as guesses (Kim et al., 2025); its successor evaluates predicted matrices as energy and force predictors (Kim et al., 2026). All of these train on converged labels.

Training with the SCF equations. Self-consistency training minimizes the residual by differentiating through the Fock build, which is label-free but requires a differentiable electronic-structure program inside the training loop (Zhang et al., 2024). HamEvo learns a neural operator from single SCF updates recorded along solver trajectories and recovers the converged matrix as its fixed point (Lou et al., 2026). The loop here treats the solver as a black box and takes its passes at the model's own recent samples. Fitting fixed targets exactly would reproduce the error that their finite correction leaves.

On-policy self-distillation. Generalized knowledge distillation trains a language model on its own samples under a teacher's per-token feedback (Agarwal et al., 2024), and Lu & Thinking Machines Lab (2025) quantify the efficiency of dense on-policy supervision over reinforcement learning. In on-policy self-distillation, the teacher is the same model given privileged information, such as a reference solution (Zhao et al., 2026). DiffusionNFT post-trains a diffusion model through two implicit branches of one network, attracting it toward well-scored samples and repelling it from poorly scored ones (Zheng et al., 2026), and DiffusionOPSD turns this into self-distillation with a frozen behavior copy, low-noise flow states, and targets built from reward gradients (Zhou et al., 2026). The loop here shares the on-policy sampling, the frozen model, and the two-branch fit; its teacher is the SCF map, so its targets are a corrected sample and the sample itself rather than steps along a reward gradient, and its local convergence depends jointly on the teacher and the projection induced by fitting. Mean Teacher and BYOL likewise fit targets produced by a moving-average teacher (Tarvainen & Valpola, 2017; Grill et al., 2020). Our RL baseline adapts DiffusionNFT with the SCF error as a scalar reward (Appendix E).

C Theory and proofs

C.1 Setting and interpretation of Section 4

A shared network may not fit every SCF correction exactly. The fitting metric determines how it approximates these corrections and whether the resulting updates preserve the SCF map's local contraction (Appendix C.7).

The SCF map depends on the prediction through its density, $\Phi(\mathbf{H}) = \mathbf{F}(\mathbf{P}(\mathbf{H}))$. Around a gapped, energy-stable solution $\mathbf{H}^\star$, express a perturbation $\delta\mathbf{H}$ in its orthonormal orbital basis. The first-order density response is $\delta P_{ai} = -\frac{2\,\delta H_{ai}}{\epsilon_a^\star - \epsilon_i^\star}, \quad i$ occupied, a virtual. The factor 2 comes from double occupancy. To first order, only Hamiltonian perturbations that couple occupied and virtual orbitals change the density.

We choose the metric so that the linearized SCF correction is the negative gradient of the quadratic DFT energy error Q introduced below (Appendix C.6).

Definition 1 (SCF response metric)**.** At $\mathbf{H}^\star$, let χ_+ weight each occupied–virtual Hamiltonian entry by $2/(\epsilon_a^\star - \epsilon_i^\star)$, where $\epsilon_i^\star$ are the orbital energies of $\mathbf{H}^\star$. For Hamiltonian perturbations u, v in this subspace, define the response inner product $\langle u, v\rangle_\chi := \langle u, \chi_+ v\rangle_F$ and norm $\|u\|_\chi := \sqrt{\langle u, u\rangle_\chi}$, where $\langle\cdot,\cdot\rangle_F$ is the Frobenius inner product.

The metric gives greater weight to orbital pairs with smaller energy gaps, reflecting their larger density response.

Convergence within the prediction family. At fixed flow states, collect the predicted occupied–virtual entries into $\boldsymbol{x}$, with self-consistent values $\boldsymbol{x}^\star$. We approximate the network locally by $\mathcal{A} = \boldsymbol{x}_0 + \mathcal{V}$: $\boldsymbol{x}_0$ is its starting prediction, and $\mathcal{V}$ contains the output changes achievable through parameter updates. Linearizing one SCF update at the self-consistent solution with this density response gives Φ_{lin}. We assume the learner finds the best fit within $\mathcal{A}$. For accepted, unclipped corrections, the quadratic paired objective and actor averaging then give

$$\boldsymbol{x}_{k+1} = \boldsymbol{x}_k + \eta\, \Pi_{\mathcal{V}}\big(\Phi_{\mathrm{lin}}(\boldsymbol{x}_k) - \boldsymbol{x}_k\big), \qquad \eta = \frac{1-m}{\beta(1+\lambda_{\mathrm{neg}})}, \tag{1}$$

where $\Pi_{\mathcal{V}}$ is the response-metric projection onto $\mathcal{V}$ (derivation in Appendix C.5). Let $A = \mathrm{I} - \mathrm{D}\Phi(\mathbf{H}^\star)$ on these components, where $\mathrm{D}\Phi(\mathbf{H}^\star)$ is the SCF map's Jacobian at $\mathbf{H}^\star$. Theorem 1 states that this update converges to

$$\boldsymbol{x}^\dagger = \operatorname*{arg\,min}_{\boldsymbol{x}\in\mathcal{A}} Q(\boldsymbol{x}), \qquad Q(\boldsymbol{x}) = \tfrac{1}{2}\langle \boldsymbol{x} - \boldsymbol{x}^\star, A(\boldsymbol{x} - \boldsymbol{x}^\star)\rangle_\chi. \tag{2}$$

The objective Q sums over flow states. Its gradient in the response metric is $\nabla_\chi Q(\boldsymbol{x}) = \boldsymbol{x} - \Phi_{\mathrm{lin}}(\boldsymbol{x})$. Energy stability supplies the lower spectral bound δ_{stab} in Theorem 1; the upper bound $1 + \Lambda$ limits the step size (Cancès et al., 2021). This result assumes quadratic response-metric fitting; the MAE term, blockwise losses, $\mathcal{L}_{\mathrm{WA}}$, $\mathcal{L}_{\mathrm{base}}$, DIIS, replay of corrections from earlier actor snapshots, a round-dependent m, and finite optimization lie outside its scope.

Matrix versus scalar feedback. One Fock build supplies a corrected matrix; a residual-based reward keeps only its distance from the prediction. We compare their information content using Euclidean fitting. The full assumptions of Theorem 2 are as follows. For one geometry, let $\mathcal{A} = \boldsymbol{x}_0 + \mathcal{V}$ contain $\boldsymbol{x}^\star$, with $D = \dim \mathcal{V} \geq 1$. Assume $\boldsymbol{e} = \boldsymbol{x}^\star - \boldsymbol{x}_0$ is rotation-invariant in $\mathcal{V}$, with $0 < \mathbb{E}\,\|\boldsymbol{e}\|_2^2 < \infty$, so no correction direction is preferred. Let a correction map $\mathcal{T}$ cost c Fock builds and satisfy $\|\mathcal{T}(\boldsymbol{x}) - \boldsymbol{x}^\star\|_2 \leq q_b \|\boldsymbol{x} - \boldsymbol{x}^\star\|_2$ with $0 \leq q_b < 1$. For $0 < \eta \leq 1$ and $0 < \varepsilon < 1$, let $\hat{\boldsymbol{x}}$ be an output satisfying $\mathbb{E}\,\|\hat{\boldsymbol{x}} - \boldsymbol{x}^\star\|_2^2 \leq \varepsilon^2 \mathbb{E}\,\|\boldsymbol{e}\|_2^2$. The Fock-build counts with matrix and scalar feedback then satisfy

$$N_{\mathrm{matrix}} \leq c \left\lceil \frac{\log(1/\varepsilon)}{\eta(1-q_b)} \right\rceil, \qquad N_{\mathrm{scalar}} \geq (1-\varepsilon^2)D.$$

The upper bound holds for Equation (1) with $\mathcal{T}$ and Euclidean projection, for every such teacher. The lower bound holds in the worst case for every residual-only algorithm on $\Phi(\boldsymbol{x}) \equiv \boldsymbol{x}^\star$, where one matrix-valued call returns the solution.

The comparison counts solver calls only; it excludes optimization cost and generalization across molecules.

C.2 Linear response at the fixed point

We fix a geometry and work in the orthogonalized basis: the Hamiltonian is $\mathbf{S}^{-1/2}\mathbf{H}\mathbf{S}^{-1/2}$ and the density is $\mathbf{S}^{1/2}\mathbf{P}\mathbf{S}^{1/2}$. For perturbations, define $\|\Delta\mathbf{H}\|_{\mathbf{S}} := \left\|\mathbf{S}^{-1/2}\Delta\mathbf{H}\mathbf{S}^{-1/2}\right\|_F$, where $\|\cdot\|_F$ is the Frobenius norm. We suppress these basis transformations below, so $\|\cdot\|_{\mathbf{S}}$ becomes the Frobenius norm, and $\langle \mathbf{U}, \mathbf{V}\rangle = \mathrm{Tr}(\mathbf{U}^\top\mathbf{V})$ is the inner product on $\mathrm{Sym}(B)$, the space of real symmetric $B \times B$ matrices. Norms of linear operators are induced by this Frobenius norm unless specified otherwise. We assume $1 \leq n_{\mathrm{occ}} < B$ and that the electronic energy is twice continuously differentiable near the fixed-point density.

Let $\mathbf{H}^\star = \sum_{p=1}^{B} \epsilon_p^\star |p\rangle\langle p|$, where the eigenvectors are orthonormal and $\epsilon_1^\star \leq \cdots \leq \epsilon_B^\star$. Occupied indices satisfy $i \leq n_{\mathrm{occ}}$ and virtual indices satisfy $a > n_{\mathrm{occ}}$. Assume the HOMO–LUMO gap $\delta = \epsilon^\star_{n_{\mathrm{occ}}+1} - \epsilon^\star_{n_{\mathrm{occ}}}$ is positive. The closed-shell density is $\mathbf{P}^\star = 2\sum_i |i\rangle\langle i|$. Differentiating its occupied projector gives the non-interacting response

$$\chi_0(\Delta) = 2\sum_{i,a} \frac{\Delta_{ai}}{\epsilon_i^\star - \epsilon_a^\star}\big(|a\rangle\langle i| + |i\rangle\langle a|\big), \qquad \Delta_{ai} = \langle a|\Delta|i\rangle. \tag{3}$$

Only occupied–virtual energy differences enter, so degeneracies within either group do not affect the expression. The matrices $E_{ia} = (|a\rangle\langle i| + |i\rangle\langle a|)/\sqrt{2}$ form an orthonormal basis of the occupied–virtual subspace $\mathcal{O} \subset \mathrm{Sym}(B)$. On this basis, $\chi_0(E_{ia}) = -2E_{ia}/\omega_{ia}$ with $\omega_{ia} = \epsilon_a^\star - \epsilon_i^\star \geq \delta$; on $\mathcal{O}^\perp$, χ_0 vanishes. Consequently χ_0 is self-adjoint and negative semidefinite, with $\|\chi_0\| = 2/\delta$.

Write $\chi_+ = -\chi_0|_{\mathcal{O}} \succ 0$, and let Π_{ov} be the orthogonal projection onto $\mathcal{O}$. The Fock build is the density derivative of the electronic energy, $\mathcal{E}(\mathbf{P}) = \mathrm{Tr}(\mathbf{H}_{\mathrm{core}}\mathbf{P}) + \mathcal{E}_{\mathrm{Hxc}}(\mathbf{P})$. Its derivative at $\mathbf{P}^\star$ is the self-adjoint kernel $\mathcal{K} = \partial^2\mathcal{E}_{\mathrm{Hxc}}/\partial\mathbf{P}^2$; write $\mathcal{K}_{\mathrm{ov}} = \Pi_{\mathrm{ov}}\mathcal{K}|_{\mathcal{O}}$ and $\mathcal{K}_{\perp,\mathrm{ov}} = (\mathrm{I} - \Pi_{\mathrm{ov}})\mathcal{K}|_{\mathcal{O}}$ for its changes within $\mathcal{O}$ and in the orthogonal complement. The chain rule gives $\mathrm{D}\Phi(\mathbf{H}^\star) = \mathcal{K}\chi_0$ and $\mathrm{J}_\alpha = \mathrm{D}\Phi_\alpha(\mathbf{H}^\star) = (1-\alpha)\mathrm{I} + \alpha\mathcal{K}\chi_0$, where I denotes the identity operator on $\mathrm{Sym}(B)$ and Φ_α is the damped SCF map defined in Appendix C.5.

Orbital stability. For a skew-symmetric occupied–virtual rotation $\kappa = \sum_{ia}\kappa_{ai}(|a\rangle\langle i| - |i\rangle\langle a|)$, let $\mathbf{P}(\kappa) = e^\kappa\mathbf{P}^\star e^{-\kappa}$ and $v = \sum_{ia}\kappa_{ai}E_{ia} \in \mathcal{O}$. With $[\mathbf{U},\mathbf{V}] = \mathbf{UV} - \mathbf{VU}$, the density expansion is $\mathbf{P}(\kappa) = \mathbf{P}^\star + [\kappa,\mathbf{P}^\star] + \frac{1}{2}[\kappa,[\kappa,\mathbf{P}^\star]] + O(\|\kappa\|^3)$, and $[\kappa,\mathbf{P}^\star] = 2\sqrt{2}\,v$. The quadratic energy change at the fixed point is

$$\begin{aligned} Q(\kappa) &= \tfrac{1}{2}\,\mathrm{Tr}\big(\mathbf{H}^\star[\kappa,[\kappa,\mathbf{P}^\star]]\big) + \tfrac{1}{2}\langle[\kappa,\mathbf{P}^\star],\mathcal{K}[\kappa,\mathbf{P}^\star]\rangle \\ &= 4\langle v,(\chi_+^{-1} + \mathcal{K}_{\mathrm{ov}})v\rangle. \end{aligned} \tag{4}$$

Indeed, $\mathrm{Tr}(\mathbf{H}^\star[\kappa,[\kappa,\mathbf{P}^\star]]) = 4\sum_{ia}\kappa_{ai}^2(\epsilon_a^\star - \epsilon_i^\star)$, and χ_+^{-1} has eigenvalue $(\epsilon_a^\star - \epsilon_i^\star)/2$ on E_{ia}. Positive orbital curvature means $\chi_+^{-1} + \mathcal{K}_{\mathrm{ov}} \succ 0$; it is a sufficient condition for a strict local minimum on the fixed-rank density manifold. Equivalently, every eigenvalue of $-\chi_+^{1/2}\mathcal{K}_{\mathrm{ov}}\chi_+^{1/2}$ is below one. This supplies the stability margin used below.

C.3 Content of one SCF pass

An SCF pass responds differently to errors that mix occupied and virtual orbitals and errors confined to either group. Express $\delta\mathbf{H} = \hat{\mathbf{H}} - \mathbf{H}^\star$ in the orbitals of $\mathbf{H}^\star$. The rectangular block $\delta\mathbf{H}_{\mathrm{vo}}$ couples the two groups; $\delta\mathbf{H}_{\mathrm{ov}} \in \mathrm{Sym}(B)$ contains this block and its transpose. The remainder $\delta\mathbf{H}_\perp = \delta\mathbf{H} - \delta\mathbf{H}_{\mathrm{ov}}$ contains the within-group blocks $\delta\mathbf{H}_{\mathrm{oo}}$ and $\delta\mathbf{H}_{\mathrm{vv}}$. These blocks leave a gap of at least $\delta' = \delta - \|\delta\mathbf{H}_{\mathrm{oo}}\|_2 - \|\delta\mathbf{H}_{\mathrm{vv}}\|_2$.

Proposition 1 (Content of one SCF pass)**.** *Suppose $\delta' > 0$, and let $L_{\mathbf{F}}$ be a Lipschitz constant of the Fock build $\mathbf{F}$ near $\mathbf{P}^\star$ in the Frobenius norm. Then*

(i) $\Phi(\mathbf{H}^\star + \delta\mathbf{H}_\perp) = \mathbf{H}^\star$*, so one pass removes the error $\delta\mathbf{H}_\perp$ exactly;*

(ii) $\left\|\mathbf{P}(\hat{\mathbf{H}}) - \mathbf{P}^\star\right\|_F \leq 2\,\|\delta\mathbf{H}_{\mathrm{ov}}\|_F/\delta'$*;*

(iii) every ordered orbital energy satisfies $|\epsilon_p(\hat{\mathbf{H}}) - \epsilon_p(\mathbf{H}^\star + \delta\mathbf{H}_\perp)| \leq \|\delta\mathbf{H}_{\mathrm{vo}}\|_2^2/\delta'$ *and* $|\epsilon_p(\Phi(\hat{\mathbf{H}})) - \epsilon_p^\star| \leq 2L_{\mathbf{F}}\,\|\delta\mathbf{H}_{\mathrm{ov}}\|_F/\delta'$.

Proof. In the orbitals of $\mathbf{H}^\star$ the clean Hamiltonian prediction reads

$$\hat{\mathbf{H}} = \begin{pmatrix} \boldsymbol{\epsilon}_{\mathrm{o}}^\star + \delta\mathbf{H}_{\mathrm{oo}} & \delta\mathbf{H}_{\mathrm{vo}}^\top \\ \delta\mathbf{H}_{\mathrm{vo}} & \boldsymbol{\epsilon}_{\mathrm{v}}^\star + \delta\mathbf{H}_{\mathrm{vv}} \end{pmatrix},$$

with $\boldsymbol{\epsilon}_{\mathrm{o}}^\star$ and $\boldsymbol{\epsilon}_{\mathrm{v}}^\star$ the occupied and virtual orbital energies. By Weyl's inequality (Horn & Johnson, 2013) the spectrum of the first diagonal block lies below $\epsilon_{n_{\mathrm{occ}}}^\star + \|\delta\mathbf{H}_{\mathrm{oo}}\|_2$ and that of the second above $\epsilon_{n_{\mathrm{occ}}+1}^\star - \|\delta\mathbf{H}_{\mathrm{vv}}\|_2$, so the two are separated by at least $\delta' > 0$.

(i) Errors within either orbital group. When $\delta\mathbf{H}_{\mathrm{vo}} = 0$, the matrix is block diagonal. Its n_{occ} lowest eigenvalues remain in the first block, so its occupied subspace is unchanged. Hence $\mathbf{P}(\mathbf{H}^\star + \delta\mathbf{H}_\perp) = \mathbf{P}^\star$ and $\Phi(\mathbf{H}^\star + \delta\mathbf{H}_\perp) = \mathbf{F}(\mathbf{P}^\star) = \mathbf{H}^\star$.

(ii) Density error from mixing the groups. $\hat{\mathbf{H}}$ adds an off-diagonal perturbation to $\mathbf{H}^\star + \delta\mathbf{H}_\perp$, whose diagonal-block spectra are separated by δ'. The $\tan 2\Theta$ theorem of Davis and Kahan (Davis & Kahan, 1970) bounds the canonical angles Θ between the two occupied subspaces by $\|\tan 2\Theta\|_F \leq 2\,\|\delta\mathbf{H}_{\mathrm{vo}}\|_F/\delta'$, and the angles lie in $[0,\pi/4)$, where $\sin\Theta \leq \frac{1}{2}\tan 2\Theta$. The occupied projectors

differ by $\sqrt{2}\,\|\sin\Theta\|_F$ in Frobenius norm and the closed-shell density is twice the projector, so $\left\|\mathbf{P}(\hat{\mathbf{H}}) - \mathbf{P}^\star\right\|_F = 2\sqrt{2}\,\|\sin\Theta\|_F \le 2\sqrt{2}\,\|\delta\mathbf{H}_{\mathrm{vo}}\|_F / \delta'$. Finally $\|\delta\mathbf{H}_{\mathrm{ov}}\|_F = \sqrt{2}\,\|\delta\mathbf{H}_{\mathrm{vo}}\|_F$.

(iii) Orbital-energy error. The first inequality follows from the quadratic residual bound for an off-diagonal perturbation of a block-diagonal Hermitian matrix with separation δ' (Mathias, 1998). For the second, $\Phi(\hat{\mathbf{H}}) - \mathbf{H}^\star = \mathbf{F}(\mathbf{P}(\hat{\mathbf{H}})) - \mathbf{F}(\mathbf{P}^\star)$ has Frobenius norm at most $L_{\mathbf{F}} \left\|\mathbf{P}(\hat{\mathbf{H}}) - \mathbf{P}^\star\right\|_F$; apply (ii) and Weyl's inequality. □

To first order, $\epsilon_p(\hat{\mathbf{H}}) - \epsilon_p^\star = \delta\mathbf{H}_{pp}$, $\mathbf{P}(\hat{\mathbf{H}}) - \mathbf{P}^\star = \chi_0\delta\mathbf{H}_{\mathrm{ov}}$ and $\Phi(\hat{\mathbf{H}}) - \mathbf{H}^\star = \mathcal{K}\chi_0\delta\mathbf{H}_{\mathrm{ov}}$. Thus $\mathrm{D}\Phi(\mathbf{H}^\star)$ vanishes on $\mathcal{O}^\perp$. These are standard response facts: rotations within either orbital group preserve the density, and Equation (3) gives the density change from mixing the groups.

Orbital energies and entrywise Hamiltonian error measure different aspects of a prediction. The former can improve while the latter barely changes (Section 5), since each orbital energy is one functional of $B(B+1)/2$ entries. The b-cycle teacher also receives its input through $\mathbf{P}(\hat{\mathbf{H}})$, so (i) applies to it as well.

C.4 Branch geometry and the pointwise update

Fix an actor-visited flow state and hold its actor prediction $\hat{\mathbf{H}}_{\theta_k}$ and targets constant. We suppress the state arguments. The corrected proposal is the actor's own clean Hamiltonian prediction at that flow state, $\hat{\mathbf{H}} = \hat{\mathbf{H}}_{\theta_k}$. For its correction $\mathbf{H}^+$, let $\Delta = \hat{\mathbf{H}}_\theta - \hat{\mathbf{H}}_{\theta_k}$ and $d = \mathbf{H}^+ - \hat{\mathbf{H}}$. The positive and negative branch predictions are $\hat{\mathbf{H}}_{\theta_k} + \beta\Delta$ and $\hat{\mathbf{H}}_{\theta_k} - \beta\Delta$. Negative-aware fine-tuning uses these same affine mixtures of an actor and learner (Zheng et al., 2026).

In a shared positive metric W, held fixed during differentiation, the quadratic component of the pointwise branch objective is

$$\mathcal{L}_{\mathrm{pair}}^{\mathrm{quad,pt}}(\Delta) = \frac{1}{2}\,\|\beta\Delta - d\|_W^2 + \frac{\lambda_{\mathrm{neg}}}{2}\beta^2\,\|\Delta\|_W^2\,.$$

Averaging over flow states and completing the square gives

$$\mathcal{L}_{\mathrm{pair}}^{\mathrm{quad}} = \mathbb{E}\left[\frac{\gamma(1+\lambda_{\mathrm{neg}})\beta^2}{2}\left\|\Delta_\theta - \frac{d}{(1+\lambda_{\mathrm{neg}})\beta}\right\|_W^2\right] + \mathrm{const.} \tag{5}$$

The second branch penalizes departure from the actor and shrinks the displacement target to $d/((1+\lambda_{\mathrm{neg}})\beta)$. A flow state whose correction is rejected ($\gamma = 0$) supplies no update.

Both branches share a detached scale per molecule, flow state, and block collection (Appendix E). The quadratic identities apply to the MSE component with these scales held fixed; they do not cover the MAE term.

C.5 Projected fitting in a shared prediction space

Collect the represented Hamiltonian entries at a fixed set of flow states, and equip this space with a fixed positive metric W, with $\langle u, v\rangle_W = u^\top W v$ and $\|u\|_W^2 = \langle u, u\rangle_W$. Padding entries are omitted. Let $\mathcal{A} = \boldsymbol{x}_0 + \mathcal{V}$ be the affine prediction family and let $\Pi_{\mathcal{V}}$ and $\Pi_{\mathcal{A}}$ be the W-orthogonal projections onto $\mathcal{V}$ and $\mathcal{A}$. The set and metric remain fixed in this local model. In training, W comes from freezing the detached branch scales; Theorem 1 is the case of the occupied–virtual entries in the response metric.

For the linearized corrector, write $\mathcal{T}_b^{\mathrm{lin}}(\boldsymbol{x}) = \boldsymbol{x}^\star + \mathrm{J}_b(\boldsymbol{x} - \boldsymbol{x}^\star)$. The operator J_b applies the linearized correction independently to each flow state. Define the damped SCF map $\Phi_\alpha(\mathbf{H}) = (1-\alpha)\mathbf{H} + \alpha\Phi(\mathbf{H})$, with $0 < \alpha \le 1$. For b passes with fixed damping, J_b consists of blocks J_α^b, where $\mathrm{J}_\alpha = \mathrm{D}\Phi_\alpha(\mathbf{H}^\star)$. A per-pass bound $\|\mathrm{J}_\alpha\|_W \le q < 1$ gives $\|\mathrm{J}_b\|_W \le q^b$. Appendix C.8 relates the SCF response to the orbital gap and stability.

Deriving the learned iteration. At actor prediction $\boldsymbol{x} \in \mathcal{A}$, write $\boldsymbol{d} = \mathcal{T}_b^{\mathrm{lin}}(\boldsymbol{x}) - \boldsymbol{x}$. The accepted, unclipped quadratic paired objective over representable displacements is

$$\min_{\Delta \in \mathcal{V}} \frac{1}{2} \|\beta\Delta - \boldsymbol{d}\|_W^2 + \frac{\lambda_{\mathrm{neg}}}{2}\beta^2 \|\Delta\|_W^2 = \min_{\Delta \in \mathcal{V}} \frac{\beta^2(1+\lambda_{\mathrm{neg}})}{2} \left\|\Delta - \frac{\boldsymbol{d}}{\beta(1+\lambda_{\mathrm{neg}})}\right\|_W^2 + \text{constant}.$$

Orthogonal projection gives $\Delta^{\mathrm{fit}} = \Pi_{\mathcal{V}}\boldsymbol{d}/[\beta(1+\lambda_{\mathrm{neg}})]$. If a fraction $f \in (0,1]$ of this fitted displacement is transferred to the actor in output space, $\eta = f/[\beta(1+\lambda_{\mathrm{neg}})]$. Because $\boldsymbol{x} \in \mathcal{A}$, $\Pi_{\mathcal{A}}\mathcal{T}_b^{\mathrm{lin}}(\boldsymbol{x}) = \boldsymbol{x} + \Pi_{\mathcal{V}}(\mathcal{T}_b^{\mathrm{lin}}(\boldsymbol{x}) - \boldsymbol{x})$, which proves Equation (1). For output averaging, $f = 1 - m$. If predictions are affine in parameters, parameter averaging gives the same identity. For a smooth network the identity holds to first order, with a remainder of order the squared parameter displacement. A uniform fixed clip factor $\lambda_{\mathrm{clip}} \in (0,1]$ replaces η by $\lambda_{\mathrm{clip}}\eta$. State-dependent clipping and rejection do not reduce to such a factor; they change the operator.

C.6 The response operator and the proof of Theorem 1

In the splitting $\mathcal{O} \oplus \mathcal{O}^{\perp}$ the Jacobian $\mathrm{D}\Phi(\mathbf{H}^\star) = \mathcal{K}\chi_0$ and $A_{\mathrm{full}} = \mathrm{I} - \mathcal{K}\chi_0$ have the block forms

$$\mathcal{K}\chi_0 = \begin{pmatrix} -\mathcal{K}_{\mathrm{ov}}\chi_+ & 0 \\ -\mathcal{K}_{\perp,\mathrm{ov}}\chi_+ & 0 \end{pmatrix}, \qquad A_{\mathrm{full}} = \begin{pmatrix} \mathrm{I} + \mathcal{K}_{\mathrm{ov}}\chi_+ & 0 \\ \mathcal{K}_{\perp,\mathrm{ov}}\chi_+ & \mathrm{I} \end{pmatrix}. \tag{6}$$

The upper-right zero block shows that occupied–virtual errors evolve independently. Write $A_{\mathrm{ov}} = \mathrm{I} + \mathcal{K}_{\mathrm{ov}}\chi_+$ and $\mathcal{B} = \chi_+ A_{\mathrm{ov}} = \chi_+ + \chi_+\mathcal{K}_{\mathrm{ov}}\chi_+$. The main text denotes this occupied–virtual block by A. Although A_{ov} need not be Frobenius-symmetric, $\mathcal{B}$ is. This motivates the response inner product $\langle u, v\rangle_\chi = \langle u, \chi_+ v\rangle_F$: it weights each transition by $2/\omega_{ia}$ and makes A_{ov} self-adjoint.

The weighting also connects the correction to energy. For the quadratic energy $Q(\boldsymbol{x}) = \frac{1}{2}\langle \boldsymbol{x} - \boldsymbol{x}^\star, A_{\mathrm{ov}}(\boldsymbol{x} - \boldsymbol{x}^\star)\rangle_\chi$, the metric gradient is $\nabla_\chi Q(\boldsymbol{x}) = A_{\mathrm{ov}}(\boldsymbol{x} - \boldsymbol{x}^\star) = -(\Phi_{\mathrm{lin}}(\boldsymbol{x}) - \boldsymbol{x})$ in the linearized model. Thus fitting the SCF correction projects an energy-descent direction onto the updates the network can realize.

For convergence, A_{ov} is similar to $\mathrm{I} + \chi_+^{1/2}\mathcal{K}_{\mathrm{ov}}\chi_+^{1/2}$, with spectrum in $[\delta_{\mathrm{stab}}, 1+\Lambda]$. Across flow states, use the smallest δ_{stab} and largest Λ. The orbital Hessian is $\chi_+^{-1} + \mathcal{K}_{\mathrm{ov}} = \chi_+^{-1}\mathcal{B}\chi_+^{-1}$ by Equation (4); its positivity gives $\delta_{\mathrm{stab}} > 0$. These response facts and the unprojected rate $(\kappa_{\mathrm{cond}} - 1)/(\kappa_{\mathrm{cond}} + 1)$ follow Cancès et al. (2021), where κ_{cond} is the condition number of the orbital Hessian preconditioned by orbital-energy differences. Theorem 1 extends the energy interpretation to shared fitting.

Proof of Theorem 1. (a) Convergence within the prediction family. The linearized teacher displacement is $\Phi_{\mathrm{lin}}(\boldsymbol{x}) - \boldsymbol{x} = -A_{\mathrm{ov}}(\boldsymbol{x} - \boldsymbol{x}^\star)$ on the stacked occupied–virtual components, with A_{ov} acting state by state. Projecting A_{ov} onto $\mathcal{V}$ preserves χ-self-adjointness. For vectors in $\mathcal{V}$, projection also preserves the Rayleigh quotients, so the spectrum of the restricted operator $\Pi_{\mathcal{V}}A_{\mathrm{ov}}$ stays in $[\delta_{\mathrm{stab}}, 1+\Lambda]$. It is therefore invertible on $\mathcal{V}$, and the equation $\Pi_{\mathcal{V}}A_{\mathrm{ov}}(\boldsymbol{x}^\dagger - \boldsymbol{x}^\star) = 0$ has a unique solution $\boldsymbol{x}^\dagger \in \mathcal{A}$. The error $\boldsymbol{v}_k = \boldsymbol{x}_k - \boldsymbol{x}^\dagger \in \mathcal{V}$ obeys $\boldsymbol{v}_{k+1} = (\mathrm{I} - \eta\Pi_{\mathcal{V}}A_{\mathrm{ov}})\boldsymbol{v}_k$, so its χ-norm contracts by at most $\max\{|1 - \eta\delta_{\mathrm{stab}}|, |1 - \eta(1+\Lambda)|\}$ per round. This is below one for $0 < \eta < 2/(1+\Lambda)$ and is minimized at $\eta^\star = 2/(1 + \Lambda + \delta_{\mathrm{stab}})$. Applying this certificate to a numerical step size requires corresponding bounds on the response spectrum.

(b) Why the limit minimizes the local energy. The fixed-point condition is $\langle u, \mathcal{B}(\boldsymbol{x}^\dagger - \boldsymbol{x}^\star)\rangle_F = 0$ for all $u \in \mathcal{V}$. These are the normal equations of the strictly convex quadratic in Equation (2). For the energy, the first-order density change is $\delta\mathbf{P} = -\chi_+\delta\mathbf{H}_{\mathrm{ov}}$, and the second-order energy change of a density variation on the fixed-rank manifold is $\frac{1}{2}\langle \delta\mathbf{P}, (\chi_+^{-1} + \mathcal{K}_{\mathrm{ov}})\delta\mathbf{P}\rangle_F$, which is Equation (4) with $[\kappa, \mathbf{P}^\star] = \delta\mathbf{P}$. Substitution gives $Q = \frac{1}{2}\langle \delta\mathbf{H}_{\mathrm{ov}}, \mathcal{B}\delta\mathbf{H}_{\mathrm{ov}}\rangle_F$. Stationarity removes the first-order term, so twice-continuous differentiability gives $\mathcal{E}(\mathbf{P}(\boldsymbol{x})) - \mathcal{E}(\mathbf{P}^\star) = Q(\boldsymbol{x}) + o(\|\boldsymbol{x} - \boldsymbol{x}^\star\|_F^2)$.

To bound the remaining error, let $\boldsymbol{z}$ be the closest point to $\boldsymbol{x}^\star$ in $\mathcal{A}$ under the χ-norm. Then $\delta_{\mathrm{stab}} \left\|\boldsymbol{x}^\dagger - \boldsymbol{x}^\star\right\|_\chi^2 \le \langle \boldsymbol{x}^\dagger - \boldsymbol{x}^\star, A_{\mathrm{ov}}(\boldsymbol{x}^\dagger - \boldsymbol{x}^\star)\rangle_\chi \le \langle \boldsymbol{z} - \boldsymbol{x}^\star, A_{\mathrm{ov}}(\boldsymbol{z} - \boldsymbol{x}^\star)\rangle_\chi \le (1+\Lambda)\|\boldsymbol{z} - \boldsymbol{x}^\star\|_\chi^2$. This is Céa's lemma for A_{ov}.

(c) Bounds from the orbital gap. Let $k_+ = \max\{\lambda_{\max}(\mathcal{K}_{\mathrm{ov}}), 0\}$, the largest eigenvalue of the kernel's positive part. Since Λ is the largest eigenvalue of $\chi_+^{1/2}\mathcal{K}_{\mathrm{ov}}\chi_+^{1/2}$, it is at least the largest diagonal entry

$2\mathcal{K}_{ia,ia}/\omega_{ia}$ and at most $\|\chi_+\|\, k_+ = 2k_+/\delta$. With k_- the largest eigenvalue of the kernel's negative part, the smallest eigenvalue is at least $-2k_-/\delta$, giving $\delta_{\mathrm{stab}} \geq 1 - 2k_-/\delta$. The bounds depend on both the gap and the interaction kernel; a small gap alone does not determine the response. □

The energy-minimization result applies to the occupied–virtual space $\mathcal{O}$, with response-weighted fitting. On the full matrix space, A_{full} need not be self-adjoint, and entrywise fitting need not follow the energy gradient. The exact weights require the unavailable orbitals of $\mathbf{H}^\star$; using those of $\mathbf{H}^+$ would be an approximation.

C.7 A CONTRACTING SCF RESPONSE CAN EXPAND AFTER FITTING

Consider the linear teacher $\boldsymbol{x} \mapsto \mathrm{J}_{\mathrm{ex}}\boldsymbol{x}$ with

$$\mathrm{J}_{\mathrm{ex}} = \begin{pmatrix} 1.2 & 1 \\ -0.7 & -0.5 \end{pmatrix}, \qquad W_\star = \begin{pmatrix} 43 & 50 \\ 50 & 60 \end{pmatrix}. \tag{7}$$

The leading principal minors of $W_\star$ are 43 and 80, so it is positive definite. Direct multiplication gives $W_\star \mathrm{J}_{\mathrm{ex}} = \mathrm{J}_{\mathrm{ex}}^\top W_\star$, and the eigenvalues of J_{ex} are $1/2$ and $1/5$. Thus J_{ex} is self-adjoint in the $W_\star$ inner product and $\|\mathrm{J}_{\mathrm{ex}}\|_{W_\star} = 1/2$.

Now fit in the Euclidean metric on $\mathcal{A} = \mathcal{V} = \mathrm{span}\{e_1\}$, where $e_1 = (1, 0)^\top$, and let Π_2 denote that projection. For $\boldsymbol{x} = ue_1$, $\Pi_2 \mathrm{J}_{\mathrm{ex}}\boldsymbol{x} = (6/5)\boldsymbol{x}$. The relaxed learned update is $(1-\eta)\boldsymbol{x} + \eta\Pi_2\mathrm{J}_{\mathrm{ex}}\boldsymbol{x} = (1+\eta/5)\boldsymbol{x}$. Every positive η expands the error; at $\eta = 1/2$ the factor is 1.1. A smaller positive common clip or slower output averaging reduces the expansion but does not change its sign. By contrast, the $W_\star$-orthogonal projection has $\Pi_{W_\star}\mathrm{J}_{\mathrm{ex}}e_1 = (83/215)e_1$, so $\eta = 1/2$ gives the contracting multiplier $149/215$.

Residual acceptance does not remove this example. For the raw Euclidean residual $r(\boldsymbol{x}) = \|(\mathrm{J}_{\mathrm{ex}} - I)\boldsymbol{x}\|_2$, direct calculation gives $r(e_1)^2 = 53/100$ and $r(\mathrm{J}_{\mathrm{ex}}e_1)^2 = 2557/10000 < 53/100$. The teacher's proposed correction therefore passes a zero-margin residual-decrease test at every nonzero ue_1, while the learned update increases the residual by a factor of 1.1 when $\eta = 1/2$. Replacing the Euclidean norm by RMS divides both residuals by the same constant. The example can be scaled arbitrarily close to zero, where a positive trust radius leaves clipping inactive.

Compatibility with a positive SCF kernel. The matrix also has the damped-response form

$$\mathrm{J}_{\mathrm{ex}} = (1-\alpha)I - \alpha\mathcal{K}_{\mathrm{ex}}\chi_{+,\mathrm{ex}}, \qquad \alpha = \frac{2}{5}, \quad \chi_{+,\mathrm{ex}} = W_\star, \quad \mathcal{K}_{\mathrm{ex}} = \frac{1}{320}\begin{pmatrix} 140 & -130 \\ -130 & 123 \end{pmatrix}. \tag{8}$$

Both $\mathcal{K}_{\mathrm{ex}}$ and $\chi_{+,\mathrm{ex}}$ are positive definite. With $\chi_0 = -\chi_{+,\mathrm{ex}}$ this has the same structure as J_α on an occupied–virtual response subspace. The raw SCF linearization is $\mathrm{J}_{\mathrm{raw}} = -\mathcal{K}_{\mathrm{ex}}\chi_{+,\mathrm{ex}}$, so $\mathrm{J}_{\mathrm{ex}} = (1-\alpha)I + \alpha\mathrm{J}_{\mathrm{raw}}$ and $\|(\mathrm{J}_{\mathrm{raw}} - I)\boldsymbol{x}\|_2 = \|(\mathrm{J}_{\mathrm{ex}} - I)\boldsymbol{x}\|_2/\alpha$. The acceptance comparison therefore also holds for the raw SCF residual of Section 3.1. An orthogonal eigenbasis of $\chi_{+,\mathrm{ex}}$ gives positive occupied–virtual gaps $2/p_i$, where p_i are its eigenvalues; extending by zero response on the complementary subspace recovers the projector-response structure. Thus the obstruction comes from fitting in a metric that fails to preserve the contraction, even for a stable positive-kernel response.

A condition restricted to representable directions. Full-space contraction in the fitting norm is sufficient but stronger than necessary. If constants $\nu > 0$ and $c_{\mathrm{Lip}} > 0$ satisfy $\langle v, (I - \mathrm{J}_b)v\rangle_W \geq \nu\|v\|_W^2$ and $\|\Pi_{\mathcal{V}}(I - \mathrm{J}_b)v\|_W \leq c_{\mathrm{Lip}}\|v\|_W$ for every $v \in \mathcal{V}$, then

$$\|v - \eta\Pi_{\mathcal{V}}(I - \mathrm{J}_b)v\|_W^2 \leq (1 - 2\eta\nu + \eta^2 c_{\mathrm{Lip}}^2)\|v\|_W^2 .$$

Hence $0 < \eta < 2\nu/c_{\mathrm{Lip}}^2$ contracts differences of model predictions. This follows by expanding the square and using $\langle v, \Pi_{\mathcal{V}}(I - \mathrm{J}_b)v\rangle_W = \langle v, (I - \mathrm{J}_b)v\rangle_W$. In the counterexample, $\langle e_1, (I - \mathrm{J}_{\mathrm{ex}})e_1\rangle_2 = -1/5$, so the condition fails along e_1.

C.8 CONTRACTION OF THE DAMPED SCF MAP

We seek a ball $\Omega_\star(\mathbf{H}^\star, r_{\mathrm{loc}}) = \{\mathbf{H} : \|\mathbf{H} - \mathbf{H}^\star\|_\star \leq r_{\mathrm{loc}}\}$ and a norm equivalent to $\|\cdot\|_{\mathbf{S}}$ such that

$$\|\Phi_\alpha(\mathbf{H}) - \Phi_\alpha(\mathbf{H}')\|_\star \leq q\|\mathbf{H} - \mathbf{H}'\|_\star , \qquad \mathbf{H}, \mathbf{H}' \in \Omega_\star(\mathbf{H}^\star, r_{\mathrm{loc}}), \quad 0 \leq q < 1. \tag{9}$$

Theorem 3 (Local contraction from gap and stability)**.** *Under the regularity and gap assumptions above:* (i) *Every eigenvalue of $\mathcal{K}\chi_0$ is real and has magnitude at most $2\|\mathcal{K}\|/\delta$.* (ii) *Suppose all eigenvalues of $\mathcal{K}\chi_0$ lie in $[-\Lambda, 1-\delta_{\mathrm{stab}}]$, with $\Lambda \geq 0$ and $0 < \delta_{\mathrm{stab}} \leq 1$. For $0 < \alpha \leq 1$ and $\alpha < 2/(1+\Lambda)$, the spectral radius of J_α is at most*

$$\bar{q}(\alpha) = \max\{1 - \alpha\delta_{\mathrm{stab}},\ \alpha(1+\Lambda) - 1\} < 1, \qquad \alpha^\star = \min\left\{1, \frac{2}{1+\Lambda+\delta_{\mathrm{stab}}}\right\}. \tag{10}$$

The value $\alpha^\star$ minimizes this upper bound among admissible damping factors. Its minimum is $(1+\Lambda-\delta_{\mathrm{stab}})/(1+\Lambda+\delta_{\mathrm{stab}})$ when $\Lambda + \delta_{\mathrm{stab}} \geq 1$, and $1-\delta_{\mathrm{stab}}$ otherwise. (iii) *For every $q \in (\bar{q}(\alpha), 1)$ there are a norm $\|\cdot\|_\star$ equivalent to $\|\cdot\|_{\mathbf{S}}$ and a radius $r_{\mathrm{loc}} > 0$ for which Equation* (9) *holds.*

For a positive semidefinite kernel, one may take $\delta_{\mathrm{stab}} = 1$ and $\Lambda = 2\|\mathcal{K}\|/\delta$, obtaining $\bar{q}(\alpha^\star) = \frac{\|\mathcal{K}\|}{\delta+\|\mathcal{K}\|}$. The bound improves as the gap grows relative to the kernel. More generally, positive orbital curvature ensures a positive stability margin, and sufficiently small damping yields local convergence. The sufficient condition $\Lambda < 1$ also permits undamped iteration. These dependencies agree with gap-based and orbital-curvature analyses of SCF (Liu et al., 2014; Cancès et al., 2021).

Proof. (i) The nonzero spectra of $\mathcal{K}\chi_0$ and $\chi_0\mathcal{K}$ coincide. An eigenvector of $\chi_0\mathcal{K}$ with nonzero eigenvalue belongs to $\mathcal{O}$, on which the operator is $-\chi_+\mathcal{K}_{\mathrm{ov}}$. This operator is similar to the symmetric operator $-\chi_+^{1/2}\mathcal{K}_{\mathrm{ov}}\chi_+^{1/2}$. Its eigenvalues are real and satisfy $|\lambda| \leq \|\chi_+\|\,\|\mathcal{K}_{\mathrm{ov}}\| \leq 2\|\mathcal{K}\|/\delta$. All remaining eigenvalues are zero. Positive orbital curvature makes these nonzero eigenvalues strictly smaller than one, so the spectral enclosure in part *(ii)* is available with some $\delta_{\mathrm{stab}} > 0$.

(ii) Each eigenvalue λ of $\mathcal{K}\chi_0$ gives an eigenvalue $1 - \alpha + \alpha\lambda$ of J_α. The enclosure for λ maps to the interval $[1-\alpha(1+\Lambda), 1-\alpha\delta_{\mathrm{stab}}]$. Since $0 < \alpha \leq 1$ and $\delta_{\mathrm{stab}} \leq 1$, its right endpoint is nonnegative. Taking the largest absolute value on this interval gives the upper bound $\bar{q}(\alpha)$; equality with the spectral radius holds if both spectral endpoints are attained. The two affine terms in $\bar{q}(\alpha)$ meet at $2/(1+\Lambda+\delta_{\mathrm{stab}})$. Constraining this value to $\alpha \leq 1$ gives $\alpha^\star$ and the stated minimum.

(iii) Choose q_{aux} with $\bar{q}(\alpha) < q_{\mathrm{aux}} < q$. Define

$$\|\Delta\|_\star^2 = \sum_{j=0}^{\infty} q_{\mathrm{aux}}^{-2j} \left\|\mathrm{J}_\alpha^j \Delta\right\|_F^2. \tag{11}$$

The series converges because the spectral radius of J_α is smaller than q_{aux}. In finite dimension,

$$\|\Delta\|_F \leq \|\Delta\|_\star \leq c_\star \|\Delta\|_F, \qquad c_\star^2 = \sum_{j=0}^{\infty} q_{\mathrm{aux}}^{-2j} \left\|\mathrm{J}_\alpha^j\right\|^2 < \infty. \tag{12}$$

Moreover, $\|\mathrm{J}_\alpha\Delta\|_\star^2 = q_{\mathrm{aux}}^2(\|\Delta\|_\star^2 - \|\Delta\|_F^2) \leq q_{\mathrm{aux}}^2\|\Delta\|_\star^2$. The positive gap makes the occupied projector smooth near $\mathbf{H}^\star$, and the energy regularity makes the Fock build continuously differentiable. Continuity therefore gives $r_{\mathrm{loc}} > 0$ such that the derivative of Φ_α has induced $\star$-norm at most q throughout $\Omega_\star(\mathbf{H}^\star, r_{\mathrm{loc}})$. The mean value inequality on this convex ball proves Equation (9). □

C.9 Finite-budget teachers

The teacher $\mathcal{T}$ used in training runs b DIIS cycles with input-dependent coefficients. The result below assumes only that the teacher reduces error by a factor $q_b < 1$ in the fitting norm at each visited prediction; no linear response is required.

The next result uses Equation (1) with $\mathcal{T}$ replacing Φ_{lin}. It allows any fitting norm $\|\cdot\|_W$ and its orthogonal projection $\Pi_{\mathcal{V}}$. The quantity $e_{\mathcal{A}} := \inf_{\boldsymbol{y}\in\mathcal{A}} \|\boldsymbol{y} - \boldsymbol{x}^\star\|_W$ is the smallest error that the prediction family can represent.

Theorem 4 (SCF improvement under shared fitting)**.** *Suppose $\|\mathcal{T}(\boldsymbol{x}) - \boldsymbol{x}^\star\|_W \leq q_b\|\boldsymbol{x} - \boldsymbol{x}^\star\|_W$ with $0 \leq q_b < 1$ at every prediction the iteration visits. If $\boldsymbol{x}_0 \in \mathcal{A}$ and $0 < \eta \leq 1$, then with $r_\eta := 1 - \eta(1-q_b)$,*

$$\|\boldsymbol{x}_k - \boldsymbol{x}^\star\|_W \leq \underbrace{r_\eta^k \|\boldsymbol{x}_0 - \boldsymbol{x}^\star\|_W}_{\text{error reduced across rounds}} + \underbrace{\frac{e_{\mathcal{A}}}{\sqrt{1-q_b^2}}}_{\text{bound on limiting error}}. \tag{13}$$

If $e_{\mathcal{A}} = 0$, repeated fitting reaches self-consistency. Otherwise, learning may stop while SCF still supplies a correction, because that correction is orthogonal to every attainable update.

Proof. Let $\boldsymbol{x}_{\mathcal{A}}$ be the W-orthogonal projection of $\boldsymbol{x}^\star$ onto $\mathcal{A}$, so $\|\boldsymbol{x}_{\mathcal{A}} - \boldsymbol{x}^\star\|_W = e_{\mathcal{A}}$, and set $t_k = \|\boldsymbol{x}_k - \boldsymbol{x}_{\mathcal{A}}\|_W$. Since $\boldsymbol{x}_k \in \mathcal{A}$, the update is $\boldsymbol{x}_{k+1} = (1-\eta)\boldsymbol{x}_k + \eta\,\Pi_{\mathcal{A}}\mathcal{T}(\boldsymbol{x}_k)$, and $\Pi_{\mathcal{A}}\mathcal{T}(\boldsymbol{x}_k) - \boldsymbol{x}_{\mathcal{A}} = \Pi_{\mathcal{V}}(\mathcal{T}(\boldsymbol{x}_k) - \boldsymbol{x}^\star)$. Projection does not increase the norm, and $\boldsymbol{x}_k - \boldsymbol{x}_{\mathcal{A}} \in \mathcal{V}$ is orthogonal to $\boldsymbol{x}_{\mathcal{A}} - \boldsymbol{x}^\star$, so

$$t_{k+1} \le (1-\eta)t_k + \eta q_b \|\boldsymbol{x}_k - \boldsymbol{x}^\star\|_W = \varphi(t_k), \qquad \varphi(t) = (1-\eta)t + \eta q_b \sqrt{t^2 + e_{\mathcal{A}}^2}.$$

The increasing map φ has Lipschitz constant $r_\eta = 1 - \eta(1-q_b)$ and fixed point $t_\infty = q_b e_{\mathcal{A}}/\sqrt{1-q_b^2}$. Induction gives $t_k \le t_\infty + r_\eta^k (t_0 - t_\infty)_+$. Returning to the full error, $\|\boldsymbol{x}_k - \boldsymbol{x}^\star\|_W = \sqrt{t_k^2 + e_{\mathcal{A}}^2} \le \sqrt{t_\infty^2 + e_{\mathcal{A}}^2} + r_\eta^k (t_0 - t_\infty)_+$, the first term equals $e_{\mathcal{A}}/\sqrt{1-q_b^2}$, and $t_0 \le \|\boldsymbol{x}_0 - \boldsymbol{x}^\star\|_W$. □

The argument above requires neither linearity nor differentiability of the teacher. It covers gating and clipping when they preserve the assumed error reduction. For a linear contractive teacher, the iteration has a unique fixed point.

For DIIS, we measure q_b. On the linearized problem, DIIS coincides with a Krylov method on A_{ov} whose polynomial depends on the input (Rohwedder & Schneider, 2011; Walker & Ni, 2011). With our start cycle, the first two cycles are undamped passes. Table 4 shows why the budget matters: for B3LYP water, $b = 2$ enlarges the error at small gaps by up to 15 times, whereas $b = 4$ reduces it by a factor between 24 and 370.

C.10 Scalar residual feedback and the comparison of solver calls

Proposition 2 (Scalar residual feedback needs as many calls as directions)**.** *Fix a linear map $\Psi : \mathbb{R}^D \to \mathbb{R}^p$ and initial coefficients $\theta_0 \in \mathbb{R}^D$, with $\boldsymbol{x}_0 = \Psi\theta_0$. Let the residual at a prediction $\boldsymbol{x} \in \mathbb{R}^p$ be $-(\mathrm{I} - \mathrm{J})(\boldsymbol{x} - \boldsymbol{x}^\star)$ with J known, let $\boldsymbol{x}^\star = \Psi\theta^\star$ with $(\mathrm{I} - \mathrm{J})\Psi$ injective on $\mathbb{R}^D$, and let $\theta^\star$ have a prior for which the initial residual $(\mathrm{I} - \mathrm{J})\Psi(\theta_0 - \theta^\star)$ is rotation invariant in the range of $(\mathrm{I} - \mathrm{J})\Psi$. An algorithm chooses predictions $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{n_{\mathrm{q}}}$ adaptively, possibly at random, observes at each one any function of the residual norm and of independent noise, and outputs $\hat{\boldsymbol{x}}$. Then*

$$\mathbb{E}\,\|(\mathrm{I} - \mathrm{J})(\hat{\boldsymbol{x}} - \boldsymbol{x}^\star)\|_2^2 \;\ge\; \Big(1 - \frac{n_{\mathrm{q}}}{D}\Big)_+ \mathbb{E}\,\|(\mathrm{I} - \mathrm{J})(\Psi\theta_0 - \boldsymbol{x}^\star)\|_2^2 . \tag{14}$$

A scalar residual norm gives a distance, leaving the algorithm to infer the correction direction. The bound is sharp and covers the reward $-\frac{1}{2}r_{\mathrm{AO}}^2$, its standardized versions and the acceptance bit. It holds even for $\mathrm{J} = 0$, when one matrix-valued call returns $\boldsymbol{x}^\star$ exactly. Dimension-dependent complexity also appears in gradient-free optimization (Nesterov & Spokoiny, 2017). Methods that differentiate through the solver (Zhang et al., 2024) receive a full direction per build and fall outside the bound.

Proof. Reduce residual norms to distances. Write $\mathcal{M} = (\mathrm{I} - \mathrm{J})\Psi = U(\mathcal{M}^\top\mathcal{M})^{1/2}$ with U an isometry of $\mathbb{R}^D$ onto the range of $\mathcal{M}$, and $\phi = (\mathcal{M}^\top\mathcal{M})^{1/2}\theta$. At a prediction $\boldsymbol{x}_j$ the squared residual norm is $\big\|U^\top(\mathrm{I} - \mathrm{J})\boldsymbol{x}_j - \phi^\star\big\|_2^2$ plus a term the algorithm knows, so we may regard queries as points $\phi_j \in \mathbb{R}^D$ and answers as functions of $\|\phi_j - \phi^\star\|_2$ and independent noise. Translate so that $\phi_0 = 0$; by assumption $\phi^\star = \rho\,\zeta$ with ζ uniform on the unit sphere and independent of $\rho \ge 0$. The residual of any output satisfies $\|(\mathrm{I} - \mathrm{J})(\hat{\boldsymbol{x}} - \boldsymbol{x}^\star)\|_2 \ge \big\|\hat{\phi} - \phi^\star\big\|_2$ for $\hat{\phi} = U^\top(\mathrm{I} - \mathrm{J})\hat{\boldsymbol{x}}$, so it suffices to bound $\mathbb{E}\,\big\|\hat{\phi} - \phi^\star\big\|_2^2$.

Give the algorithm at most one exact coordinate per call. Reveal ρ and, at step j, the exact inner product $\langle u_j, \phi^\star\rangle$, where u_j is the unit vector obtained by orthogonalizing ϕ_j against $S_{j-1} = \mathrm{span}\{\phi_1, \ldots, \phi_{j-1}\}$; if $\phi_j \in S_{j-1}$ it learns nothing new. Since $\|\phi_j - \phi^\star\|_2^2 = \|\phi_j\|_2^2 - 2\langle\phi_j, \phi^\star\rangle + \rho^2$ and $\phi_j \in S_j$, every answer of the original oracle is a function of this information and of the noise, so a lower bound for the informed algorithm applies to the original one.

Bound the error in the remaining directions. Let $\mathcal{F}_j$ contain the internal randomness, noise, ρ and inner products through step j. Conditional on this information, the component of $\phi^\star$ in $S_j^\perp$ remains uniform on a sphere of radius ρ_j. This holds for $j = 0$. If it holds for $j-1$, the direction $u_j \in S_{j-1}^\perp$ is $\mathcal{F}_{j-1}$-measurable, and for a vector uniform on a sphere, conditioning on its coordinate along a fixed direction leaves the orthogonal component uniform on the sphere of the remaining radius; applying this conditionally on $\mathcal{F}_{j-1}$ proves the claim for j. With $d_{j-1} = \dim S_{j-1}^\perp \geq D - j + 1$, uniformity gives $\mathbb{E}[\langle u_j, \phi^\star\rangle^2 \mid \mathcal{F}_{j-1}] = \rho_{j-1}^2/d_{j-1}$, hence $\mathbb{E}\rho_j^2 \geq \frac{D-j}{D-j+1}\mathbb{E}\rho_{j-1}^2$ and $\mathbb{E}\rho_{n_\mathrm{q}}^2 \geq (1-n_\mathrm{q}/D)_+\mathbb{E}\rho^2$. Finally, given $\mathcal{F}_{n_\mathrm{q}}$ the component of $\phi^\star$ in $S_{n_\mathrm{q}}^\perp$ has mean zero, so every $\mathcal{F}_{n_\mathrm{q}}$-measurable output has $\mathbb{E}\left\|\hat{\phi} - \phi^\star\right\|_2^2 \geq \mathbb{E}\rho_{n_\mathrm{q}}^2$. For sharpness, the differences of the answers at 0 and at te_j reveal one coordinate of $\phi^\star$ per call. □

Proof of Theorem 2. (a) Matrix feedback. Apply Theorem 4 in the Euclidean norm. Since $\boldsymbol{x}^\star \in \mathcal{A}$ we have $e_\mathcal{A} = 0$, so $\|\boldsymbol{x}_k - \boldsymbol{x}^\star\|_2 \leq r_\eta^k \|\boldsymbol{e}\|_2$ for every realization of $\boldsymbol{e}$, with $r_\eta = 1 - \eta(1-q_b) \leq e^{-\eta(1-q_b)}$. Hence $k \geq \log(1/\varepsilon)/(\eta(1-q_b))$ rounds give $\mathbb{E}\|\boldsymbol{x}_k - \boldsymbol{x}^\star\|_2^2 \leq \varepsilon^2\mathbb{E}\|\boldsymbol{e}\|_2^2$, and each round costs c Fock builds.

(b) Scalar feedback. The solver $\Phi(\boldsymbol{x}) \equiv \boldsymbol{x}^\star$ satisfies the hypothesis for every q_b, with the teacher $\mathcal{T} = \Phi$ and $c = 1$. For it the residual norm at $\boldsymbol{x}_j$ is $\|\boldsymbol{x}_j - \boldsymbol{x}^\star\|_2$, and Proposition 2 with $\mathrm{J} = 0$ and Ψ an isometry onto $\mathcal{V}$ gives $\mathbb{E}\|\hat{\boldsymbol{x}} - \boldsymbol{x}^\star\|_2^2 \geq (1 - n_\mathrm{q}/D)_+\mathbb{E}\|\boldsymbol{e}\|_2^2$. The target error therefore requires $n_\mathrm{q} \geq (1-\varepsilon^2)D$. □

C.11 Numerical checks

Table 4: Water, B3LYP/6-31G, symmetric O–H stretch relative to equilibrium. The response is obtained by finite differences of the SCF map at the converged solution. $\|\mathrm{D}\Phi\|$ is the Frobenius operator norm of its occupied–virtual block on $\mathcal{O}$. The energy column compares the quadratic Q in Equation (2) to the DFT energy change, taking their ratio over four random perturbations. The divergence column counts, among 300 random ten-dimensional prediction families in $\mathcal{O}$, those for which the relaxed iteration with the undamped one-pass teacher and $\eta = 1/(1+\Lambda)$ fails to contract, for the response projection and for the Frobenius projection. The last three columns are worst-case Frobenius error ratios over five random starts of size 10^{-3}: per cycle for optimal fixed damping, and in total for fresh-history DIIS with $b = 2$ and $b = 4$.

stretch	gap (Ha)	$\|\mathrm{D}\Phi\|$	δ_stab	Λ	energy	diverged	damped	DIIS, 2	DIIS, 4
1.0	0.350	1.57	0.84	1.42	1.000	0 / 0	0.43	0.44	0.0027
1.4	0.225	2.93	0.75	2.44	1.000	0 / 0	0.56	1.46	0.0081
1.8	0.146	5.30	0.61	4.34	0.999	0 / 0	0.70	5.06	0.020
2.2	0.104	8.96	0.46	7.36	0.998	0 / 0	0.80	14.9	0.041

The numerical response has the predicted structure: for Hartree–Fock, LDA and B3LYP, Jacobian columns on $\mathcal{O}^\perp$ vanish to 10^{-10}, the spectrum of $\mathcal{K}\chi_0$ is real to machine precision, and $\chi_+\mathcal{K}_\mathrm{ov}\chi_+$ is symmetric to 10^{-7}.

The smallest eigenvalue of the Frobenius-symmetric part of A_ov is 0.82, 0.72, 0.58 and 0.38 along the series, so by the representable-direction condition of Appendix C.7, the Frobenius metric in the orthogonalized basis also yields convergence for every family on the occupied–virtual block. For the full operator A_full on $\mathrm{Sym}(B)$, it is -0.28 at equilibrium and -3.1 at stretch 1.8, because of the coupling $\mathcal{K}_{\perp,\mathrm{ov}}\chi_+$ of norm 3.6 and 11.4; thus some prediction families make the entrywise projected iteration expand for this molecule.

D Equivariance of the correction targets

Lemma 1 (Equivariant SCF corrections)**.** *Let $D(g)$ be the orthogonal representation of a molecular rotation or a permutation of identical atoms in the atom-centered basis. Assume the electronic energy and basis operations transform consistently and the occupied–virtual gap is positive. Then Φ and Φ_α^b commute with the action $\mathbf{H} \mapsto D(g)\mathbf{H}D(g)^\top$, and $\|\cdot\|_\mathbf{S}$ is invariant. A correction displacement*

transforms by the same action, independently of the choice of eigenvector signs or bases within occupied or virtual degenerate eigenspaces.

Proof. Write $D = D(g)$. The overlap and core Hamiltonian obey $\mathbf{S}(g\mathcal{M}) = D\mathbf{S}(\mathcal{M})D^\top$ and $\mathbf{H}_{\mathrm{core}}(g\mathcal{M}) = D\mathbf{H}_{\mathrm{core}}(\mathcal{M})D^\top$. Covariance of the energy derivative gives $\mathbf{F}(D\mathbf{P}D^\top; g\mathcal{M}) = D\mathbf{F}(\mathbf{P}; \mathcal{M})D^\top$. If $\mathbf{HC} = \mathbf{SC}\boldsymbol{\epsilon}$ and $\mathbf{C}^\top\mathbf{SC} = \mathbf{I}$, the transformed eigenproblem has eigenvectors $D\mathbf{C}$ and the same ordered eigenvalues. The positive gap makes its occupied subspace unique, giving $\mathbf{P}(D\mathbf{H}D^\top) = D\mathbf{P}(\mathbf{H})D^\top$. Composing the density and Fock maps proves covariance of Φ and hence Φ^b_α.

The identity $\mathbf{S}(g\mathcal{M})^{-1/2} = D\mathbf{S}(\mathcal{M})^{-1/2}D^\top$ and orthogonal invariance of the Frobenius norm prove $\left\|\mathbf{S}(g\mathcal{M})^{-1/2}D\Delta D^\top\mathbf{S}(g\mathcal{M})^{-1/2}\right\|_F = \|\Delta\|_{\mathbf{S}}$. A difference of two covariant Hamiltonians is covariant. The density projector is unchanged by eigenvector signs and by rotations within a fully occupied or fully virtual eigenspace, so the correction displacement requires no eigenvector alignment. □

The Frobenius displacement clip and the full-matrix AO RMS residual are also invariant under this action. A finite-budget solver that respects the action therefore yields covariant clipped targets and invariant acceptance decisions. Parameter updates preserve the architectural equivariance of the actor and learner. A squared blockwise loss with invariant block weights is rotation invariant; the entrywise MAE normalization is not. The predictions remain equivariant in either case, because their equivariance comes from the architecture.

E PROTOCOL DETAILS

Flow states. The deterministic guess $\mathbf{H}^{(0)}$ is a Fock build from a MINAO density. The prior draw $\mathbf{H}_{\mathrm{prior}}$ adds perturbations obtained by an equivariant expansion of isotropic Gaussian coefficients with scale $\sigma = 0.1$; its interatomic entries equal those of $\mathbf{H}^{(0)}$. With L integration steps, the flow times are $t_j = 0.01 + 0.99j/L$, $j = 0, \ldots, L-1$, and the first state is initialized to $\mathbf{H}_{\mathrm{prior}}$. The network predicts all entries, and its final clean Hamiltonian prediction supplies the sampled Fock matrix. The default $L = 3$ gives flow states at $t = 0.01$, 0.34, and 0.67. Post-training selects flow states from the low-noise portion of the rollout and corrects their clean Hamiltonian predictions.

Finite SCF correction and acceptance. For QH9, the corrector uses restricted Kohn–Sham calculations with B3LYP5, def2-SVP, and integration-grid level three in PySCF (Sun et al., 2020). DIIS uses a history of eight and first extrapolates in the third cycle. Each call starts a new solver history from the density of the actor prediction and runs exactly b cycles; early convergence termination is disabled. It returns the last matrix whose occupied orbitals were used in a cycle, together with the next raw Fock build that evaluates its residual.

Each proposal is accepted or rejected by its raw AO residual improvement (Section 3.1). Validity checks require finite matrices with the correct shape and symmetry, finite occupations, and the correct electron count. The accepted total displacement is clipped once by the trust radius of Section 3.1. Rejected corrections have zero weight in both paired branches and in the WA distillation loss.

Exact paired-loss reduction. For one molecule and flow state, let $\mathcal{I}_{\mathrm{d}}$ and $\mathcal{I}_{\mathrm{o}}$ be the valid entries in its atom-diagonal and off-diagonal block collections. Padded entries are excluded. For an error matrix E and $a \in \{\mathrm{d}, \mathrm{o}\}$, define

$$\mathrm{MSE}_a(E) = \frac{1}{|\mathcal{I}_a|}\sum_{u\in\mathcal{I}_a} E_u^2, \qquad \mathrm{MAE}_a(E) = \frac{1}{|\mathcal{I}_a|}\sum_{u\in\mathcal{I}_a} |E_u|.$$

An empty block collection contributes zero. Write the branch residuals as $E^+ = \beta\Delta_\theta - d$ and $E^- = -\beta\Delta_\theta$. The matrix loss in $\mathcal{L}_{\mathrm{pair}}$ (Section 3.2) uses the shared correction scale $s_a = \frac{1}{2}\,\mathrm{sg}[\mathrm{MAE}_a(\mathbf{H}^+ - \hat{\mathbf{H}})] + \epsilon_{\mathrm{num}}$ and

$$\ell_{\mathrm{mat}}(E) = \frac{1}{2}\sum_{a\in\{\mathrm{d},\mathrm{o}\}}\left[\mathrm{MAE}_a(E) + \frac{\mathrm{MSE}_a(E)}{2s_a}\right]. \tag{15}$$

Here sg holds its argument fixed during differentiation and $\epsilon_{\mathrm{num}} > 0$ is a numerical floor. Each scale is detached and shared by both branches for that molecule, flow state, and block collection.

WA distillation loss. The WA distillation loss (Section 3.3) is evaluated per molecule and flow state on the learner's clean Hamiltonian prediction in double precision. The correction's eigenproblem $\mathbf{H}^{+}\mathbf{C}^{+} = \mathbf{S}\mathbf{C}^{+}\boldsymbol{\epsilon}^{+}$ is solved once per flow state through the Cholesky factor of $\mathbf{S}$, and $\mathbf{C}^{+}$ and $\boldsymbol{\epsilon}^{+}$ are held fixed in differentiation. Rejected corrections are skipped, keeping their place in the batch mean.

Actor refresh. The learner and actor start from the pretrained checkpoint. The actor remains fixed between refreshes, while a prefetcher adds new rollouts and SCF corrections to the replay buffer. Every n_{fit} optimizer steps, the actor is updated by parameter averaging as described in Section 3.4. The learner retains its parameters and optimizer state. With asynchronous prefetching, stored flow states, anchors, and corrections remain associated with the actor snapshot that generated them.

Hyperparameters. Table 5 lists the OPSD settings for each experiment.

Table 5: **OPSD settings.** Loss weights are those of $\mathcal{L}_{\mathrm{total}}$ (Section 3.4), and k is the refresh round. In the projected HELM run, $\mathcal{L}_{\mathrm{proj}}$ replaces $\mathcal{L}_{\mathrm{WA}}$ with the same weight. The transfer trust radius is 2.0 for the functional study, 1.0 for the size and water-cluster studies, 0.25 for the thermal, reactive, dimer, electric-field, solvent, and scaffold studies; "none" means the displacement is not clipped. The transfer studies keep the learning rate constant.

	QH9	**HELM**	**MD17**	**Transfer**
SCF correction targets				
Rollouts per geometry N_{roll}	1			
SCF cycles per correction b	3–4	3	4	4
Required improvement δ_r	0			
Trust radius r_{tr}	0.0054	none	0.0054	0.25–2.0
Loss				
λ_{pair}	1.76	1	1.76	1
λ_{WA}	0.0045	3×10^{-4}	0.018	0.0225
λ_{base}	10^{-4}	0	10^{-4}	10^{-4}
Branch step β	1			
Negative-branch weight λ_{neg}	0.25	1	0.25	1
Actor refresh				
Learner steps per round n_{fit}	1			
Refresh coefficient m	$\min(10^{-3}k, 0.5)$	0.99	$\min(10^{-3}k, 0.5)$	$\min(10^{-3}k, 0.5)$
Replay buffer (minibatches)	32			
Optimization				
Peak learning rate	10^{-6}	10^{-5}	5×10^{-7}	10^{-5}
Global batch size	32 or 64	16	64	8
Updates at evaluated checkpoint	3,400–7,650	240	600	up to 2,000

Baselines. The RL and SCT baselines start from the same QHFlow2-s checkpoints as OPSD. The RL baseline adapts DiffusionNFT (Zheng et al., 2026) with the SCF error reward $-\frac{1}{2}r_{\mathrm{AO}}^2$, standardized within each group of samples. SCT (Zhang et al., 2024) minimizes the difference between the predicted Hamiltonian and the Fock matrix rebuilt from it, differentiating through the eigensolver and the Fock build.

E.1 Transfer experiments

Adaptation and evaluation. All studies start from the same QHFlow2-s checkpoint pretrained on QH9-Stable-ID (Kim et al., 2026). OPSD and SFT share the adaptation geometries and three-step flow sampler, with global batch size eight, constant learning rate 10^{-5}, and up to 2,000 optimizer updates. OPSD corrects the prediction at the final flow state of each trajectory with four SCF cycles and trains on the accepted corrected targets. SFT fits converged target Hamiltonians with the Hamiltonian loss

Table 6: **Primary transfer endpoints on the confirmation sets.** Entries are mean errors. Gap errors are in mHa, barrier and interaction-energy errors in kcal mol^{-1}, and dipole errors in D.

Setting / endpoint	Unadapted	SFT	OPSD
Functional / gap	155.387	42.483	2.158
Thermal / gap	10.323	4.162	0.934
Solvent / dipole	0.696	0.182	0.022
Unseen scaffolds / gap	0.580	1.648	0.398
Size / gap	99.473	15.247	3.505
Reactive / barrier	625.643	200.921	11.350
Dimers / interaction	0.10753	0.08758	0.08226
Electric (training fields) / induced dipole	0.08559	0.02450	0.02239

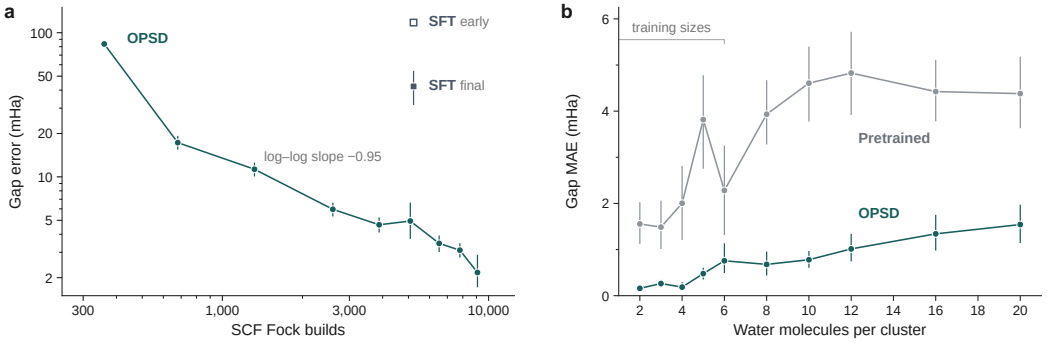


Figure 4: **Correction budget and water clusters. a**, Gap error of the functional study against the SCF Fock builds spent on training targets. **b**, Gap MAE of water clusters before and after OPSD; the bracket marks the training sizes. Error bars are 95% bootstrap confidence intervals over molecules (a) and clusters (b).

and the same WA loss, computed against the converged labels. Optimizer updates and geometry exposure were matched.

Electronic-structure targets and readouts. The functional study uses the exchange–correlation component of ωB97X-D with def2-SVP and grid level three; the empirical dispersion energy is omitted consistently. Thermal, scaffold, size, reactive, and dimer references use B3LYP5/def2-SVP. Thermal errors average four frames within each parent. Reaction barriers are the maximum energy over ten sampled path images minus the reactant energy, evaluated separately for each method. The size study uses 30–40-atom SPICE (Eastman et al., 2023) PubChem molecules, including hydrogens, with disjoint source identities across the splits. Dimer geometries are a subset of the A24 (Řezáč & Hobza, 2013), HSG (Faver et al., 2011), JSCH (Jurečka et al., 2006), S22by5 (Gráfová et al., 2010), and SSI (Burns et al., 2017) databases distributed with Psi4 (Smith et al., 2020). Dimer interaction energies subtract each arm's own isolated-monomer energies from its dimer energy, using common non-counterpoise references; curve MAEs average six separations within each pair. Energy readouts before SCF iteration evaluate the predicted density and include the Fock build needed to evaluate that energy.

Solvent and electric-field response. Implicit-water calculations use ddCOSMO (Lipparini et al., 2013) with dielectric constant 78.3553. The electrostatic solvation energy is $E_{\mathrm{water}}(\mathbf{P}_{\mathrm{water}}) - E_{\mathrm{gas}}(\mathbf{P}_{\mathrm{gas}})$: the water-adapted model supplies the water density, and the unadapted gas model supplies the gas density for both adapted arms. This definition gives the results in Figure 3h. For the static-field study, all arms use the same field-conditioned architecture, initialized by adding a zero-initialized field-mixing channel to the pretrained model. Adaptation uses 64 parents at zero field and six signed axial fields of magnitude 0.001 atomic units. The 64 confirmation parents are evaluated at those fields, six face-diagonal fields at the same magnitude, and six axial fields at magnitude 0.002. Each arm predicts its own zero-field dipole for the subtraction $\boldsymbol{\mu}(\boldsymbol{\mathcal{E}}) - \boldsymbol{\mu}(\mathbf{0})$; vector-error norms are averaged over each field family within a parent.

Distribution of the gains. Median paired gap improvements over SFT are 12.60 mHa for functional transfer, 0.177 mHa for thermal geometries, and 0.689 mHa for larger molecules; the median paired barrier improvement is 3.29 kcal mol^{-1}. In these four studies, the three largest paired gains contribute 15.2%, 94.9%, 58.1%, and 97.8% of the net mean improvement, respectively.

E.2 Correction budget and water clusters

Correction budget. Figure 4a plots the gap error of the functional study against the SCF Fock builds spent on training targets. For OPSD, these are the builds actually run up to each checkpoint; for SFT, they are the builds needed to converge its labels. The prefetcher adds about one minibatch of corrections every eight updates, so each OPSD correction is replayed in about eight updates. The OPSD error falls with a log–log slope of -0.95 over the plotted range. OPSD reaches the final SFT error with at least 7.7 times fewer builds.

Water clusters. We also adapt the QH9-pretrained QHFlow2-s to water clusters at B3LYP5/def2-SVP, with the settings of the transfer studies. OPSD trains on clusters of up to six molecules and is evaluated on clusters of 2 to 20 molecules (Figure 4b). It lowers the gap MAE at every cluster size, including sizes larger than those seen in training.

E.3 SCF convergence from the predictions

A prediction can also serve as the starting point of a full SCF calculation. Figure 5 counts the SCF iterations needed to converge from the SFT and OPSD predictions in four transfer studies. OPSD needs fewer iterations in all four: 13.8% fewer for the functional study, 9.7% for implicit water, 3.8% for thermal geometries, and 4.4% for dimers. It improves on SFT for 186 of 200, 130 of 200, 20 of 32, and 7 of 10 systems, respectively.

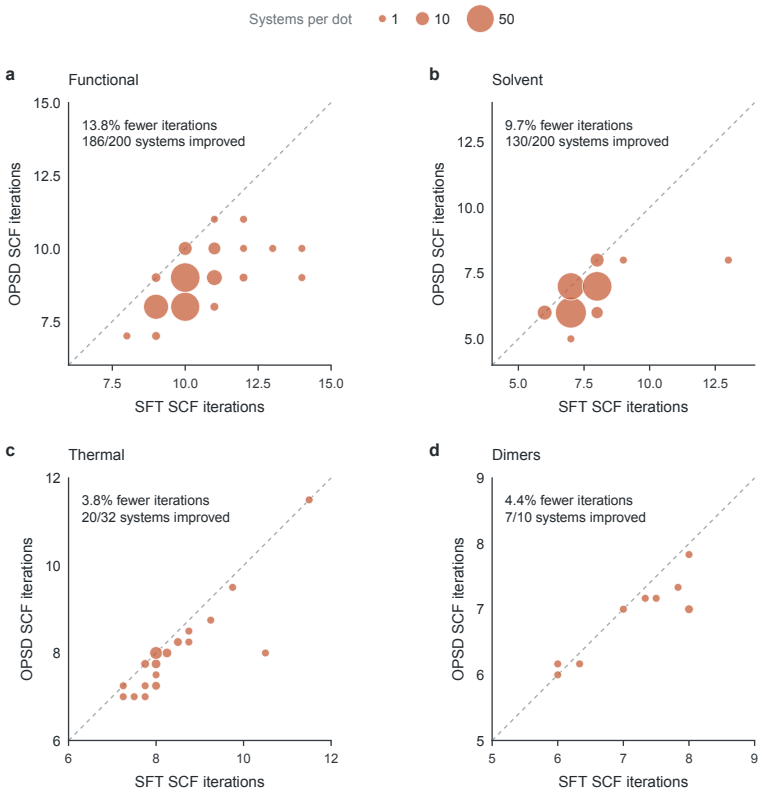


Figure 5: **SCF iterations from the adapted models.** Iterations needed to converge SCF from the SFT (horizontal axis) and OPSD (vertical axis) predictions for **a**, the functional study, **b**, implicit water, **c**, thermal geometries, and **d**, dimers. Points below the dashed diagonal converge faster from OPSD, and dot area grows with the number of systems at that point. Thermal counts average four frames per parent, and dimer counts average six separations per pair. Both models use the same stopping test, an energy change below 10^{-9} Ha and an orbital-gradient norm below 3.16×10^{-5}, and every plotted calculation converged within 50 iterations.