

---

# Equivariant Compression of Quantum Operator Representations

---

Jigyasa Nigam<sup>†\*</sup>, Utku Sirin<sup>‡\*</sup>, Tess Smidt<sup>†</sup>, Stratos Idreos<sup>‡</sup>

<sup>†</sup>Massachusetts Institute of Technology    <sup>‡</sup>Harvard University  
jnigam@mit.edu, utkusirin@seas.harvard.edu

## Abstract

Electronic structure is being increasingly prioritized in atomistic machine learning (ML), as it offers a direct path to predicting diverse properties beyond single-property ML surrogates. However, quantum mechanical (QM) data, such as matrix representations of the electronic Hamiltonian, are inherently high-dimensional. This makes standard operations, such as diagonalization, for obtaining eigenspectra and other observables prohibitively expensive, and also makes the ML of such data computationally demanding. In this work, we ask whether QM data can benefit from classical compression strategies that are widely used in other fields. Taking the example of effective single-particle Hamiltonians, we analyze how these techniques interact with the symmetric matrix structure and affect derived observables. We introduce an equivariant compression framework that preserves symmetry while reducing dimensionality, allowing quantum observables computed from the compressed representation to closely reproduce those from the full Hamiltonian. Our results underscore the need for physics-aware compression techniques and pave new directions for scalable, structure-preserving ML for electronic structure.

## 1 Introduction

Machine learning (ML) has transformed our approach to modeling complex physical systems, paving the way for newer computational pipelines, such as hybrid ML–quantum workflows for property prediction. Atomistic modeling, in particular, has progressed from learning interatomic potentials [1–9] and surrogate models for individual properties [10, 11] to directly targeting electronic structure itself. By targeting the fundamental ingredients of quantum mechanics, such as effective single-particle Hamiltonians, ML can simultaneously optimize multiple observables within a single framework, overcoming the limitations of separate property-specific surrogates.

Within these workflows, ML can take ingredients such as effective one-electron Hamiltonians as inputs to predict a wide range of molecular properties [12–14] or entirely replace the traditional quantum calculation [15–28]. The promise of unified modeling is, however, accompanied by the challenge of representing and learning quantum mechanical entities, which are inherently higher-dimensional than atomistic properties. For instance, matrix representations of the electronic Hamiltonian on atomic basis sets scale quadratically with their size. This rapid growth makes standard quantum mechanical operations computationally expensive and also poses a practical challenge for storing, sharing, and training ML models on these datasets. Consequently, there is growing interest in strategies to determine effective or downfolded representations of quantum mechanical operators without sacrificing their essential physical structure.

One strategy is to explicitly construct a low-rank symmetry-adapted projected Hamiltonian (SAPH) [20], which can then be used as a reduced target for ML. While this method accurately pre-

---

\*These authors contributed equally.

serves a subset of eigenvalues, it leads to unphysical predictions of other observables [29]. Refs. 29, 30 proposes an alternative strategy by exploiting differentiable pipelines to learn a compressed *effective* Hamiltonian by minimizing the error on several observables computed from it against their reference values from a larger basis set calculation.

In other domains routinely faced with high-dimensional data, classical compression techniques are reliable tools for handling data. For instance, image compression can reduce matrix dimensions by up to  $25\times$  by exploiting problem-specific features [31–33]. Techniques such as the Discrete Cosine Transform (DCT) enable efficient storage and reconstruction by using a fixed set of basis functions. While classical compression algorithms are typically optimized for storage rather than for dimensionality reduction, they have also been successfully applied in this context [31]. However, naive application of these methods on QM data can compromise physical symmetries, such as Hermiticity and rotational equivariance, leading to undesirable deviations in derived observables.

In this work, we analyze the interplay of several analytical compression techniques with the mathematically rich structure of QM data with the goal of designing compression strategies that lead to the most information-dense representations while preserving the underlying structure of the data. In doing so, we achieve a lower-dimensional representation of the operator matrix, from which chemical observables such as excitation energies, atomic charges, dipole moments, etc. may be derived. Although our focus here is on electronic Hamiltonians, the proposed approach naturally extends to other QM operators expressed on an AO basis and provides a strategy to balance cost–accuracy trade-offs in their modeling.

## 2 Background

**Electronic Hamiltonian Operator** The electronic Hamiltonian operator is central to QM as it encodes all ground-state properties of a given atomic configuration. In practice, many quantum-mechanical theories and software packages work with a matrix representation of the operator in a finite basis of atom-centered orbitals (AOs), each of which is characterized by radial ( $n$ ) and angular quantum numbers ( $l, m$ ) denoting the spherical harmonic centered on each atom. Note that these basis functions are independent of the molecular structure but may vary with the chemical species  $a$  of the atom on which they are centered. The matrix elements between  $\mu \equiv (anlm)$  and  $\nu \equiv (a'n'l'm')$  centered on atoms  $i$  and  $j$  in structure  $A$ , can then be represented as  $\mathbf{H}^{a\mu, a'\nu}(A_{ij}) = \langle a|nlm|\hat{H}|a'|n'l'm'\rangle$ . One-electron eigenvalues  $\varepsilon$  (molecular orbital (MO) energies) and MO coefficients ( $\mathbf{C}$ ) are obtained from the Hamiltonian by solving the generalized eigenvalue problem,

$$\mathbf{HC} = \mathbf{SC} \operatorname{diag} \varepsilon. \quad (1)$$

Other ground state properties, such as the Löwdin atomic charges  $q_i$ , can then be computed by first computing the density matrix,  $\rho = \mathbf{C} \operatorname{diag}(\mathbf{f}) \mathbf{C}^\dagger$ , with occupation of each MO in  $\mathbf{f}$ , followed by

$$q_i = Z_i - \sum_{\mu \in i} \sum_{\nu \in i} (\mathbf{S}^{1/2} \rho \mathbf{S}^{1/2})_{\mu\nu}. \quad (2)$$

In addition to satisfying Hermiticity,  $\mathbf{H} = \mathbf{H}^\dagger$ , the matrix representation of the Hamiltonian also transforms under rotations as a consequence of the angular nature of AOs. For each pair of angular quantum numbers  $(l, l')$ , associated with the radial functions  $(n, n')$  and atoms  $i$  and  $j$  in structure  $A$ , the matrix elements  $\mathbf{H}_{A_{ij}}^{bmm'}$ , are coupled using Clebsch–Gordan coefficients to obtain equivariant outputs indexed by  $L$ ,

$$\mathbf{H}_{nlnl''}^{\mathbf{b}_{\text{symm}}, M}(A_{ij}) = \sum_{mm'} \langle lm; l'm' | LM \rangle \mathbf{H}_{nn'}^{\mathbf{b}, mm'}(A_{ij}), \quad (3)$$

where we use the shorthand  $\mathbf{b} = (al, a'l')$  and  $\mathbf{b}_{\text{symm}} = (a, a', sL)$  to denote the combined set of indices for the angular and radial basis functions, as well as the chemical species of the atoms.  $\langle lm; l'm' | LM \rangle$  are the Clebsch–Gordan coefficients,  $s$  denotes inversion parity, i.e., whether the resulting irrep behaves as a polar or pseudotensor under inversions. The electronic Hamiltonian is, thus, a mathematically constrained object rather than an arbitrarily dense matrix. In the following, we review a class of matrix compression techniques, and analyze to what extent they preserve the structure of the matrix and the physical properties derived from it.

**Discrete Cosine Transform** Similar to the symmetry-adapted block decomposition in (3), popular data compression methods [34–36] also project matrices onto an orthogonal basis using transformations such as Discrete Cosine Transform (DCT) [37–39]. A two-dimensional DCT of an input matrix results in a matrix whose elements are coefficients of a data-independent basis of cosines of increasing frequency, such that coefficients on the top-left corner of the matrix correspond to low-frequency components, and coefficients in the bottom-right corner correspond to high-frequency components. This structure enables compression strategies such as high-frequency pruning (DCT-H), which removes the bottom-right frequency bands, and low-frequency pruning (DCT-L), which removes the top-left bands [31] (Figure 1). We collectively refer to these methods as 2D compression as they directly operate on two dimensional matrices<sup>2</sup>. However, when applied to the Hamiltonian, these methods are agnostic to the atom and orbital-resolved structure of the operator. Although compression reduces dimensionality, it obscures the chemically or physically meaningful structure associated with the resulting matrix elements, as they can no longer be ascribed to specific atoms. The compressed matrices, however, can still be diagonalized and their eigenvalues and eigenvectors used to obtain approximate physical observables. In the following, we propose a framework to leverage these frequency-based compression techniques while retaining the symmetry-adapted decomposition of the Hamiltonian.

### 3 Methods

As the physical symmetries of the Hamiltonian matrix are best captured by the symmetry-adapted block decomposition (3), we adopt a block-wise compression approach, treating each block separately and restricting compression along a *single* dimension that neither alters the atoms, nor compromises the O(3) symmetry of each block. Each block can then be compressed by a naive truncation, singular-value decomposition (SVD), or a one-dimensional version of the DCT method described above. We collectively refer to these methods as 1D compression as these approaches only act along the the ( $nl, n'l'$ ) labels corresponding to the columns of each block  $\mathbf{H}^{\mathbf{b}_{\text{symm}}}$  in Fig. 2. For example, consider a fictitious initial basis corresponding to the matrix in Fig. 2, containing 2p orbitals on oxygen and 1s, 2s orbitals on hydrogen. The matrix then features blocks involving the coupling 2p–2p, 1s–2s, and so on. In a target minimal basis where each oxygen and hydrogen is represented by a single s orbital, blocks corresponding to p orbitals or higher s orbitals are entirely removed, while remaining blocks may be compressed along the orbital indices to keep only the desired reduced dimensionality.

As shown in Fig. 2, this strategy corresponds to eliminating some blocks ( $\mathbf{H}^{\mathbf{b}_{\text{symm}}}$ ) or reducing the size of each block ( $\mathbf{b}$ ), based on the size and symmetries of the target reduced effective basis. In this work, we focus on SVD, as well as, a one-dimensional DCT, noting that these methods, unlike Refs. 29, 30, are not specifically optimized for the downstream properties (such as eigenvalues, charges derived from the matrix) and may only partially capture global properties due to the independent treatment of each block. This may be possibly remedied by applying similar compression techniques to functions of the matrix, such as algebraic powers, which we leave for future work. Another drawback of treating blocks separately is that even when the compressed basis has the same dimension as the original one, each block may undergo its own unitary transformation. This leads to a reassembled matrix that is not related to the original by a single global unitary (evidenced in Fig. 3c). One can avoid this issue by imposing an overall gauge that aligns the blocks in the original and compressed representations, ensuring consistency across the full matrix.

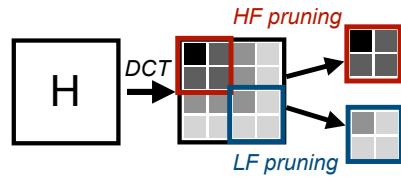


Figure 1: Discrete Cosine Transform (DCT) represents data in the frequency domain based on a range of low and high-frequency cosine functions. LF: Low-frequency. HF: High-frequency.

<sup>2</sup>Notice that for 2D compression, we perform a single DCT over entire matrix. This is unlike how popular storage formats such as JPEG operate on images, which decomposes the image into smaller spatial patches, corresponding to adjacent submatrices e.g., 8x8 patches, and apply DCT to each patch separately.

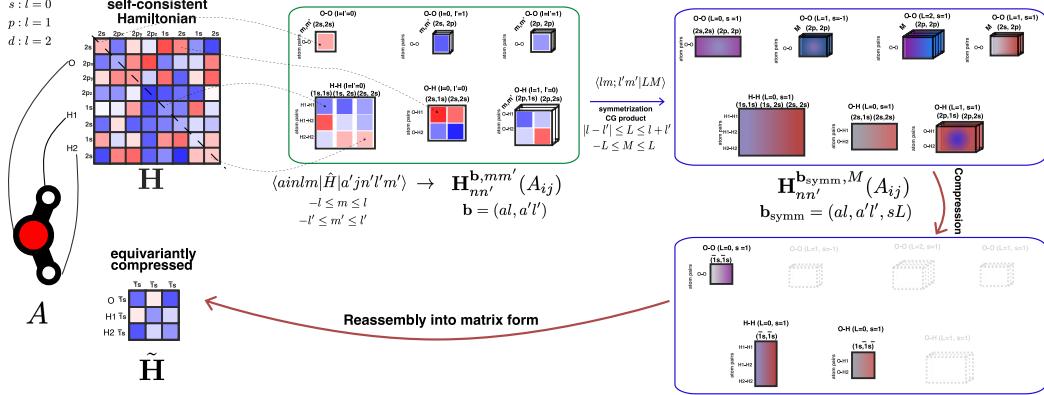


Figure 2: Schematic illustrating the decomposition of an input self-consistent Hamiltonian matrix into blocks  $\mathbf{H}_{A_{ij}}^{b,m,m'}$  labeled by angular quantum numbers [20], followed by symmetrization via a Clebsch-Gordan tensor product. Compression to an effective representation  $\tilde{\mathbf{H}}$  corresponds to retaining only the subset of matrix elements allowed by symmetry of the smaller effective basis. We follow the data structure proposed in Ref. 40 such that each block  $\mathbf{H}_{n,n'}^{b_{\text{symm}},M}(A_{ij})$  has the structure and atom labels along its rows, the symmetry index  $M$  along the depth, and the orbital label  $(nl, n'l')$  contributing to its symmetric label along its columns.

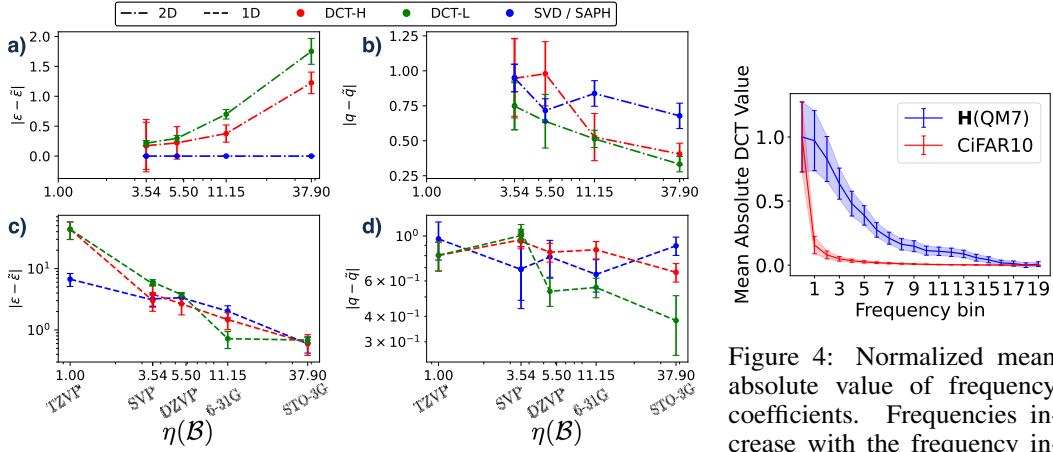


Figure 3: MAE on frontier MO eigenvalues and Löwdin charges (atomic units) obtained through compressed matrices obtained by 2D (matrix) (dash-dot lines) and 1D (block-wise) (dashed lines) compression. Compression ratios  $\eta(\mathcal{B})$  in Eq. (10) are defined as the average ratio of the matrix dimension in the TZVP basis to that in a smaller basis ( $\mathcal{B}$ ) across all molecules in the dataset, corresponding to 1, 3.54, 5.50, 11.15, and 37.90 for TZVP, SVP, DZVP, 6-31G, and STO-3G, respectively. We also compute SAPH [41], which is an equivariant 2D compression.

## 4 Results and Discussion

We evaluate the performance of these compression techniques on a dataset of 1000 molecules, containing C, H, N, and O atoms, from the QM7b dataset [42], comprising Fock and overlap matrices computed at the B3LYP/def2-TZVP level in PySCF [43]. We benchmark the compression across a range of resolutions corresponding to smaller basis sets, namely, def2SVP, cc-PVDZ, DZVP, 6-31G, and STO-3G, noting that the compressed matrices resulting from 2D DCT may be comparable in size but cannot be truly interpreted in terms of specific orbital labels. For all experiments, we report the

Figure 4: Normalized mean absolute value of frequency coefficients. Frequencies increase with the frequency index. For the  $\mathbf{H}$  dataset, frequencies decay more slowly than for natural image data, motivating a re-examination of image-inspired compression and ML techniques and their adaptation modeling QM data.

mean absolute error (MAE) on a set of frontier MO eigenvalues and Löwdin atomic charges computed from the compressed matrices against the reference calculations (def2-TZVP basis). In computing (2), we use the exact overlaps computed on the corresponding reduced basis, since compressing the def2-TZVP overlap matrix directly can easily compromise its positive definiteness. In the following, we report the mean absolute error (MAE) on frontier eigenvalues and Löwdin charges across the dataset, computed as

$$\text{MAE} = \sum_{A=1}^{N_{\text{struct}}} \sum_{j=\varepsilon_{\text{start}}(A)}^{\varepsilon_{\text{end}}(A)} |\varepsilon_j(A) - \tilde{\varepsilon}_j(A)|, \quad (4)$$

where  $A$  enumerates the molecular structures in the dataset,  $\varepsilon$  denotes the reference eigenvalue,  $\tilde{\varepsilon}$  the eigenvalue of the compressed matrix. The range of eigenvalues considered for computing the MAE [ $\varepsilon_{\text{start}}(A)$ ,  $\varepsilon_{\text{end}}(A)$ ] is a restricted subset of five eigenstates on either side of the HOMO-LUMO gap for each structure. The MAE for atomic charges, on the other hand, is simply an average of the difference between predicted and reference charges across all atoms in each molecule.

Fig. 3a) aligns most closely with intuition, DCT-H, which preserves low frequencies, better approximates global functions of the matrix (eigenvalues) than DCT-L, which prunes low frequencies. SAPH corresponds to zero errors as it preserves a subset of eigenvalues by design. In comparison, 3b) suffers from the gauge fixing problem and shows a systematic improvement on frontier MO eigenvalues with compression across all methods. However, 1D DCTs are more accurate than their corresponding 2D counterparts at the highest compression ratio. Frequency-compressed matrices approximate charges much better than SAPH, especially at the highest compression ratio using both 2D and 1D equivariant frequency compression (3c), d)). The decreasing trend of the MAE with basis set size warrants further investigation, especially as the compressed Hamiltonian interacts with the exact overlap matrix in the computation of properties, causing the effects of compression and basis truncation to be intrinsically intertwined.

Finally, we use the 2D frequency analysis to examine the structure of the Hamiltonian matrix and compare it with image data using the CIFAR10 dataset [44] in Fig. 4. As expected, the frequency distribution of QM matrices shows a distinct pattern relative to standard images. This difference becomes important in view of the growing interest in applying image based machine learning tools to electronic structure problems [28, 45–48] and may be used to adapt these techniques to the structure of the data.

While we demonstrated a strategy for compressing electronic Hamiltonians that preserves symmetry, this work highlights the opportunity to link molecular properties and eigenvalues with matrix frequencies and symmetries, guiding the selection of Hamiltonian functions to compress and combining classical and ML techniques for interpretable, efficient representations of quantum data. A promising extension could be to learn frequency cutoffs or adaptive filters that depend on angular momentum channels or local atomic environments, revealing data driven spectral priors for quantum matrices.

## Data Availability

The dataset comprising molecular configurations and reference electronic-structure observables can be obtained from the archive at Ref. 49.

## Acknowledgments

JN is grateful for funding from the MIT Postdoctoral Fellowship for Excellence in Engineering and thanks Genevi  e Dusson for insightful discussions. JN and TS acknowledge support from the National Science Foundation under Cooperative Agreement PHY-2019786 (The NSF AI Institute for Artificial Intelligence and Fundamental Interactions). US acknowledges funding from the Swiss National Science Foundation Postdoc Mobility scholarship P500PT\_217934.

## A An Alternative View of Compression

The eigenfunctions of the one-dimensional Laplacian operator provide a natural basis for representing smooth functions. Upon discretization on a uniform grid, the Laplacian can be written as a tridiagonal

matrix using a three-point finite-difference stencil,

$$\mathcal{L} = \frac{1}{\Delta x^2} \begin{bmatrix} -2 & 1 & 0 & \cdots & 0 \\ 1 & -2 & 1 & \cdots & 0 \\ 0 & 1 & -2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & 1 \\ 0 & 0 & 0 & 1 & -2 \end{bmatrix}. \quad (5)$$

The entries of the first and last row may be modified to reflect a change in boundary conditions. We refer the interested reader to Ref. 50, 51 for more detailed derivations of the various variants of discrete Fourier, cosine transform, and sine transforms. For instance, the 2D discrete cosine transform of a matrix is computed as

$$\mathbf{H}_{p,q}^{\text{DCT}} \propto \sum_{k=0}^{N-1} \sum_{l=0}^{N-1} \mathbf{H}_{s,t} \cos\left(\frac{(2s+1)\pi p}{2N}\right) \cos\left(\frac{(2t+1)\pi q}{2N}\right) \quad (6)$$

where  $N$  is the block size of a square matrix,  $\mathbf{H}_{s,t}$  the matrix elements of the 2D Hamiltonian matrix,  $\mathbf{H}_{p,q}^{\text{DCT}}$  the matrix element of the DCT compressed matrix. The eigenvectors  $\mathbf{D}$  of  $\mathcal{L}$  form an orthonormal set and correspond to discrete cosine modes, with eigenvalues proportional to the squared spatial frequencies. These eigenvectors are the smoothest functions on a real grid in the sense that they minimize the Rayleigh quotient, making them ideal for representing low-energy modes. The 1D discrete cosine transform (DCT) can be viewed as a projection onto this basis. A finite sequence or vector  $\mathbf{M}$  can be transformed via

$$\mathbf{M}' = \mathbf{MD}^\top. \quad (7)$$

If all eigenvectors are retained, this is a complete change of basis. However, if only a subset is selected, the procedure leads to a compressed representation of  $\mathbf{M}$ . In particular, selecting the eigenvectors corresponding to the smallest eigenvalues (lowest frequencies) defines the DCT-L variant, which emphasizes the smooth components of  $\mathbf{M}$ , while choosing the eigenvectors with the largest eigenvalues (highest frequencies) defines DCT-H.

On the other hand, for two-dimensional compression, the Laplacian eigenfunctions are separable along each dimension. The 2D DCT of a matrix  $\mathbf{M}$  is then given by

$$\mathbf{M}' = \mathbf{D}_{\text{row}} \mathbf{MD}_{\text{col}}^\top, \quad (8)$$

where  $\mathbf{D}_{\text{row}}$ , and  $\mathbf{D}_{\text{col}}$  are the 1D DCT eigenvectors along rows and columns, respectively. In the context of a quantum operator matrix, the 2D DCT thus mixes up correlations across both atom and orbital indices. This frequency space analysis provides a measure of smoothness for the matrix elements when arranged by radial or angular quantum numbers. As matrix elements between nearby orbitals vary smoothly with these indices, one may assume that the Hamiltonian is band-limited in this index space, allowing a select set of frequencies to capture most of its physically relevant structure.

## B Compression across basis sets

We define the compression ratio,  $\eta$ , reported in Fig. 3 as the ratio of the number of elements in the upper triangle of the matrix in the reference basis set (def2-TZVP) and those in the smaller basis. Let the dimension of the matrix in a basis  $\mathcal{B}$  be

$$D(\mathcal{B}) = \sum_a N_{\text{at}}(a) N_{\text{basis}}(\mathcal{B}(a)), \quad (9)$$

where the sum runs over all chemical species  $a$ ,  $N_{\text{at}}(a)$  is the number of atoms of type  $a$ , and  $N_{\text{basis}}(\mathcal{B}(a))$  is the number of basis functions associated with that element in  $\mathcal{B}$ . The compression ratio can then be obtained as

$$\eta(\mathcal{B}) = \frac{D(\text{TZVP})(D(\text{TZVP}) - 1)}{D(\mathcal{B})(D(\mathcal{B}) - 1)} \quad (10)$$

For example, in the def2-TZVP basis, there are 31 orbitals per C, N, and O atom and 6 per H atom, while in the STO-3G basis, there are 5 orbitals per C, N, and O atom and 1 per H atom. For a molecule such as CH<sub>4</sub>, this results in  $\eta \approx 41.25$ , whereas for a larger molecule like C<sub>6</sub>H<sub>12</sub>O<sub>6</sub>, we obtain  $\eta = 38.47$ . We therefore report the average of the compression ratio per molecule across the dataset in the main text.

## C Transforming Non-orthogonal vs Orthogonal Matrices

Both  $\mathbf{H}$  and  $\mathbf{S}$  appear as components of the generalized eigenvalue equation in Eq. (1). In principle, we can choose to compress only the Hamiltonian  $\mathbf{H}$  while keeping the overlap matrix  $\mathbf{S}$  from the basis set corresponding to the chosen compression level. This approach is less physically consistent, since the resulting overlap matrix no longer reflects the true overlaps between basis functions. We do not attempt to compress  $\mathbf{S}$  from the larger basis, as doing so is numerically unstable and can violate its required positive semidefinite structure.

An alternative is to work in the Löwdin-orthogonalized representation, where the eigenvalue problem takes the simpler form

$$\mathbf{H}_{\text{Löwdin}} = \mathbf{S}^{-1/2} \mathbf{H} \mathbf{S}^{-1/2} \quad (11)$$

$$\mathbf{H}_{\text{Löwdin}} \mathbf{C} = \mathbf{C} \text{ diag } \boldsymbol{\varepsilon} \quad (12)$$

where  $\mathbf{C}$  and  $\boldsymbol{\varepsilon}$  are the same as in Eq. (1)

In the main text, we report results obtained by compressing the nonorthogonal Hamiltonian  $\mathbf{H}$ , noting however the inconsistency that would arise from combining atomic orbital overlaps with the compressed matrix, though similar techniques may also be applied to the orthogonalized Hamiltonian  $\mathbf{H}_{\text{Löwdin}}$ .

## References

- [1] Jörg Behler and Michele Parrinello. Generalized Neural-Network Representation of High-Dimensional Potential-Energy Surfaces. *Phys. Rev. Lett.*, 98:146401, 2007.
- [2] Albert P. Bartók, Mike C. Payne, Risi Kondor, and Gábor Csányi. Gaussian Approximation Potentials: The Accuracy of Quantum Mechanics, without the Electrons. *Phys. Rev. Lett.*, 104: 136–403, 2010.
- [3] S. Batzner, A. Musaelian, L. Sun, M. Geiger, J. P. Mailoa, M. Kornbluth, N. Molinari, T. Schmidt, and B. Kozinsky. E(3)-Equivariant Graph Neural Networks for Data-efficient and Accurate Interatomic Potentials. *Nature Communications*, 13(2453), 2022.
- [4] Ilyes Batatia, David P Kovacs, Gregor Simm, Christoph Ortner, and Gabor Csanyi. MACE: Higher Order Equivariant Message Passing Neural Networks for Fast and Accurate Force Fields. In S. Koyejo, S. Mohamed, A. Agarwal, D. Belgrave, K. Cho, and A. Oh, editors, *NeurIPS*, volume 35, pages 11423–11436. Curran Associates, Inc., 2022.
- [5] Matthias Rupp, Alexandre Tkatchenko, Klaus-Robert Müller, and O. Anatole von Lilienfeld. Fast and Accurate Modeling of Molecular Atomization Energies with Machine Learning. *Phys. Rev. Lett.*, 108:058301, 2012.
- [6] Alexander V. Shapeev. Moment Tensor Potentials: A Class of Systematically Improvable Interatomic Potentials. *Multiscale Modeling & Simulation*, 14:1153–1173, 2016.
- [7] Ralf Drautz. Atomic Cluster Expansion for Accurate and Transferable Interatomic Potentials. *Phys. Rev. B*, 99:014104, 2019.
- [8] Jorg Behler. Four generations of high-dimensional neural network potentials. *Chemical Reviews*, 121(16):10037–10072, 2021.

- [9] Albert Musaelian, Simon Batzner, Anders Johansson, Lixin Sun, Cameron J Owen, Mordechai Kornbluth, and Boris Kozinsky. Learning Local Equivariant Representations for Large-Scale Atomistic Dynamics. *Nature Communications*, 14(1):579, 2023.
- [10] Max Veit, David M. Wilkins, Yang Yang, Robert A. DiStasio, and Michele Ceriotti. Predicting Molecular Dipole Moments by Combining Atomic Partial Charges and Atomic Dipoles. *J. Chem. Phys.*, 153(2):024–113, 2020.
- [11] Jiace Sun, Lixue Cheng, and III Miller, Thomas F. Molecular Dipole Moment Learning via Rotationally Equivariant Derivative Kernels in Molecular-orbital-based Machine Learning. *The Journal of Chemical Physics*, 157(10):104–109, 2022.
- [12] Alberto Fabrizio, Ksenia R Briling, and Clemence Corminboeuf. SPA $\hat{H}$ M: The Spectrum of Approximated Hamiltonian Matrices representations. *Digital Discovery*, (1):286–294, 2022.
- [13] Matthew Welborn, Lixue Cheng, and Thomas F Miller III. Transferability in Machine Learning for Electronic Structure via the Molecular Orbital Basis. *Journal of Chemical Theory and Computation*, 14(9):4772–4779, 2018.
- [14] Zhuoran Qiao, Matthew Welborn, Animashree Anandkumar, Frederick R Manby, and Thomas F Miller III. OrbNet: Deep Learning for Quantum Chemistry Using Symmetry-adapted Atomic-orbital Features. *The Journal of Chemical Physics*, 153(12):124–111, 2020.
- [15] Andrea Grisafi, Alberto Fabrizio, Benjamin Meyer, David M Wilkins, Clemence Corminboeuf, and Michele Ceriotti. Transferable machine-learning model of the electron density. *ACS Central Science*, 5(1):57–64, 2018.
- [16] Joshua A Rackers, Lucas Tecot, Mario Geiger, and Tess E Smidt. A Recipe for Cracking the Quantum Scaling Limit with Machine Learned Electron Densities. *Machine Learning: Science and Technology*, 4(1):015–027, 2023.
- [17] Teddy Koker, Keegan Quigley, Eric Taw, Kevin Tibbetts, and Lin Li. Higher-order equivariant neural networks for charge density prediction in materials. *npj Computational Materials*, 10(1):161, 2024.
- [18] Ganesh Hegde and R Chris Bowen. Machine-learned Approximations to Density Functional Theory Hamiltonians. *Scientific reports*, 7(1):1–11, 2017.
- [19] Kristoff 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. *Nat Commun*, 10(1):50–24, 2019.
- [20] Jigyasa Nigam, Michael J. Willatt, and Michele Ceriotti. Equivariant Representations for Molecular Hamiltonians and  $N$ -Center Atomic-Scale Properties. *J. Chem. Phys.*, 156(1):014–115, 2022.
- [21] Oliver T. Unke, Mihail Bogojeski, Michael Gastegger, Mario Geiger, Tess Smidt, and Klaus-Robert Müller. SE(3)-equivariant Prediction of Molecular Wavefunctions and Electronic Densities. *NeurIPS*, (1106):14434–14447, 2021.
- [22] Liwei Zhang, Berk Onat, Geneviève Dusson, Adam McSloy, Gautam Anand, Reinhard J Maurer, Christoph Ortner, and James R Kermode. Equivariant Analytical Mapping of First Principles Hamiltonians to Accurate and Transferable Materials Models. *npj Computational Materials*, 8(1):158, 2022.
- [23] Shi Yin, Xudong Zhu, Tianyu Gao, Haochong Zhang, Feng Wu, and Lixin He. Harmonizing Covariance and Expressiveness for Deep Hamiltonian Regression in Crystalline Material Research: a Hybrid Cascaded Regression Framework. *arXiv:2401.00744*, 2024.
- [24] Yang Zhong, Jihui Yang, Hongjun Xiang, and Xingao Gong. Universal Machine Learning Kohn-Sham Hamiltonian for Materials. *arXiv:2402.09251*, 2024.

- [25] Anubhab Haldar, Ali K Hamze, Nikhil Sivadas, and Yongwoo Shin. GEARS H: Accurate Machine-learned Hamiltonians for Next-Generation Device-scale Modeling. *arXiv:2506.10298*, 2025.
- [26] Pol Febrero, Peter Jørgensen, Miguel Pruneda, Alberto Garcia, Pablo Ordejon, and Arghya Bhowmik. Graph2Mat: Universal Graph to Matrix Conversion for Electron Density Prediction. 2024.
- [27] Liwei Zhang, Patrizia Mazzeo, Michele Nottoli, Edoardo Cignoni, Lorenzo Cupellini, and Benjamin Stamm. A Symmetry-preserving and Transferable Representation for Learning the Kohn-Sham Density Matrix. *arXiv:2503.08400*, 2025.
- [28] Chenghan Li, Or Sharir, Shunyue Yuan, and Garnet K Chan. Image Super-resolution Inspired Electron Density Prediction. *arXiv:2402.12335*, 2024.
- [29] Edoardo Cignoni, Divya Suman, Jigyasa Nigam, Lorenzo Cupellini, Benedetta Mennucci, and Michele Ceriotti. Electronic excited states from physically constrained machine learning. *ACS Central Science*, 10(3):637–648, 2024.
- [30] Divya Suman, Jigyasa Nigam, Sandra Saade, Paolo Pegolo, Hanna Türk, Xing Zhang, Garnet Kin-Lic Chan, and Michele Ceriotti. Exploring the Design Space of Machine Learning Models for Quantum Chemistry with a Fully Differentiable Framework. *Journal of Chemical Theory and Computation*, 2025.
- [31] Utku Sirin and Stratos Idreos. The Image Calculator: 10x Faster Image-AI Inference by Replacing JPEG with Self-designing Storage Format. *Proc. ACM Manag. Data*, 2(1):52:1–52:31, 2024.
- [32] Utku Sirin, Victoria Kauffman, Aadit Saluja, Florian Klein, Jeremy Hsu, and Stratos Idreos. Frequency-Store: Scaling Image AI by A Column-Store for Images. In *CIDR*, 2025.
- [33] Kai Xu, Minghai Qin, Fei Sun, Yuhao Wang, Yen-Kuang Chen, and Fengbo Ren. Learning in the Frequency Domain. In *CVPR*, pages 1737–1746, 2020.
- [34] G.K. Wallace. The JPEG Still Picture Compression Standard. *IEEE Transactions on Consumer Electronics*, 38(1):xviii–xxxiv, 1992.
- [35] JPEG. Jpeg 2000, 2023. URL <https://jpeg.org/jpeg2000/>. Accessed on Jan. 02, 2023.
- [36] Gary J. Sullivan, Jens-Rainer Ohm, Woo-Jin Han, and Thomas Wiegand. Overview of the High Efficiency Video Coding (HEVC) Standard. *IEEE Transactions on Circuits and Systems for Video Technology*, 22(12):1649–1668, 2012.
- [37] R. David Evans, Lufei Liu, and Tor M. Aamodt. JPEG-ACT: Accelerating Deep Learning via Transform-based Lossy Compression. In *ISCA*, pages 860–873, 2020.
- [38] Zachi Karni and Craig Gotsman. Spectral Compression of Mesh Geometry. In *SIGGRAPH*, page 279–286, 2000.
- [39] N. Ahmed, T. Natarajan, and K. R. Rao. Discrete Cosine Transfom. *IEEE Trans. Comput.*, 23(1):90–93, 1974.
- [40] Filippo Bigi, Joseph W Abbott, Philip Loche, Arslan Mazitov, Davide Tisi, Marcel F Langer, Alexander Goscinski, Paolo Pegolo, Sanggyu Chong, Rohit Goswami, et al. Metaten-sor and Metatomic: Foundational Libraries for Interoperable Atomistic Machine Learning. *arXiv:2508.15704*, 2025.
- [41] Jigyasa Nigam, Sergey Pozdnyakov, and Michele Ceriotti. Recursive evaluation and iterative contraction of  $N$ -body equivariant features. *J. Chem. Phys.*, 153(12):121101, September 2020. ISSN 0021-9606, 1089-7690. doi: 10.1063/5.0021116.
- [42] Grégoire Montavon, Matthias Rupp, Vivekanand Gobre, Alvaro Vazquez-Mayagoitia, Katja Hansen, Alexandre Tkatchenko, Klaus-Robert Müller, and O Anatole Von Lilienfeld. Machine Learning of Molecular Electronic Properties in Chemical Compound Space. *New Journal of Physics*, 15(9):095–003, 2013.

- [43] Qiming Sun, Timothy C Berkelbach, Nick S Blunt, George H Booth, Sheng Guo, Zhendong Li, Junzi Liu, James D McClain, Elvira R Sayfutyarova, Sandeep Sharma, et al. PySCF: the Python-based Simulations of Chemistry Framework. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 8(1):e1340, 2018.
- [44] Alex Krizhevsky. Learning Multiple Layers of Features from Tiny Images. 2009.
- [45] Murat Isik, Mandeep Kaur Saggi, Humaira Gowher, and Sabre Kais. Multimodal Quantum Vision Transformer for Enzyme Commission Classification from Biochemical Representations. *arXiv:2508.14844*, 2025.
- [46] Ryong-Gyu Lee and Yong-Hoon Kim. Convolutional Network Learning of Self-consistent Electron Density via Grid-projected Atomic Fingerprints. *npj Computational Materials*, 10(1):248, 2024.
- [47] Yong Zhao, Kunpeng Yuan, Yinqiao Liu, Steph-Yves Louis, Ming Hu, and Jianjun Hu. Predicting Elastic Properties of Materials from Electronic Charge Density Using 3D Deep Convolutional Neural Networks. *The Journal of Physical Chemistry C*, 124(31):17262–17273, 2020.
- [48] Yilong Ju, Shah Saad Alam, Jonathan Minoff, Fabio Anselmi, Han Pu, and Ankit Patel. Interpreting convolutional neural networks' low-dimensional approximation to quantum spin systems. *Physical Review Research*, 7(1):013–094, 2025.
- [49] Divya Suman, Jigyasa Nigam, Sandra Saade, Paolo Pegolo, Hanna Tuerk, Xing Zhang, Garnet Kin-Lic Chan, and Michele Ceriotti. Exploring the design space of machine-learning models for quantum chemistry with a fully differentiable framework. <https://doi.org/10.24435/materialscloud:mg-8f>, 2025. URL <https://doi.org/10.24435/materialscloud:mg-8f>.
- [50] Vladimir Britanak, Patrick C. Yip, and K.R. Rao. *Discrete Cosine and Sine Transforms: General Properties, Fast Algorithms and Integer Approximations*. Academic Press, 2006. ISBN 978-0-12-373624-6.
- [51] N. Ahmed, T. Natarajan, and K.R. Rao. Discrete Cosine Transform. *IEEE Transactions on Computers*, C-23(1):90–93, 1974.