Machine Learning Prediction of Phonon Dispersion

Minjae Kwen[1], **Moritz Baumgarten**[1], **Tong Jiang**[1], **Jigyasa Nigam**[2], **and Joonho Lee**[1]

[1]Harvard University
[2]Massachusetts Institute of Technology

ABSTRACT

Accurate phonon dispersions are typically obtained from DFPT calculations. However, DFPT is computationally expensive and often becomes a bottleneck in simulations of phonon-related interactions in condensed-matter systems. In this project, we use a finetuned machine-learning interatomic potential (FT-MLIP) to predict phonon dispersions. We compute the Hessian (interatomic force constants) as the second derivative of the ML-predicted energy with respect to atomic positions. We then apply a long-range dipole–dipole correction to capture LO—TO splitting in polar crystals. Because Hessians can be obtained from the ML model at comparatively low cost, this approach suggests a promising route to accelerate phonon calculations. Nevertheless, achieving a robust and accurate long-range correction for ML-predicted force constants remains an open challenge.

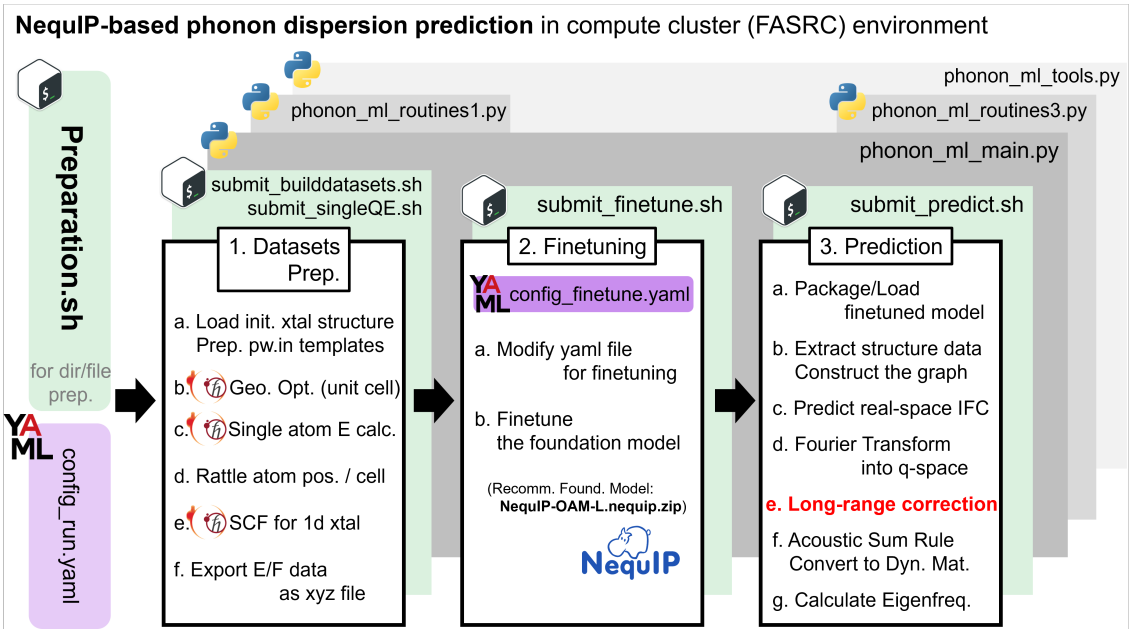

Figure 1. Flowchart of the phonon-dispersion predictor pipeline.

Background & Goal

Accurate phonon dispersions are typically obtained using density-functional perturbation theory (DFPT), but the computational cost can be prohibitive. We explore a machine-learning interatomic potential (MLIP) approach, where interatomic force constants (IFCs) are obtained from the second derivative of the predicted energy with respect to atomic displacements.

We therefore pursue the following overall workflow and objectives: (i) generate DFT-based training data, (ii) finetune a NequIP foundation model, (iii) predict real-space IFCs/Hessians and compute phonon dispersions with long-range correction.

- **Goal:** Predict phonon dispersion from structure with reduced computational cost.
- **Key challenge:** Stable enforcement of physical constraints (translation invariance, ASR) and long-range dipole interactions.

- **Deliverables:** End-to-end workflow (dataset → finetune → predict → dispersion), source codes and analysis of long-range correction strategies.

Design & Implementation

0. Overall pipeline

Figure 1 summarizes the end-to-end workflow: dataset construction using QE, finetuning NequIP, predicting the phonon dispersion.
(Prediction Step: predict real-space Hessian (IFCs) → Fourier transforming to obtain $C(\mathbf{q})$ → applying long-range dipole correction → ASR and dynamical-matrix diagonalization)

1. Dataset preparation (QE-based)

The dataset is built from a reference crystal structure and a set of perturbed configurations:

a. Load the initial structure and prepare QE input templates (`pw.in`). Both Materials Project structures and local input files are supported.

b. Perform an initial geometric optimization (QE) at the unit-cell scale.

c. Compute single-atom reference energies (needed for the energy term used in XYZ file).

d. Generate perturbed structures by rattling atomic positions and cell parameters using Gaussian noise. From this step onward, calculations are performed in a **supercell**.

e. Run SCF calculations (QE) for each perturbed structure.

f. Export the dataset into XYZ format.

This workflow is automated via python method named `builddatasets(phonon_ml_main.py)` and supporting routines in `phonon_ml_routine1.py`.

2. Finetuning (NequIP)

Finetuning is configured through `config_finetune.yaml` and launched via `finetuning(phonon_ml_main.py)` with a train:valid split of 4:1. The objective is to adapt the foundation model to the target material dataset while maintaining stability for second-derivative (Hessian) predictions.
(`nequip-train --config-dir . --config-name config_finetune`)

3. Phonon prediction from ML Energy (IFCs)

a. Model packaging and inference setup

The finetuned checkpoint is packaged and loaded for inference:

- Create a model package: `nequip-package build "best.ckpt_path" "model_zip_path.nequip.zip"`.
- Load the packaged model
- Extract structural features
- Construct the graph of crystal.

b. Real-space Hessian construction and symmetrization

Let k,l index atoms in the unit cell, $a,b \in \{x,y,z\}$ index Cartesian components, and $\mathbf{R}_{ij}$ be the displacement between cells i and j in the supercell. Denote the atomic coordinate as τ^i_{ka} (component a of atom k in cell i), and define an energy function $E(\{\tau^i_{ka}\})$ through the ML model.

The real-space Hessian (IFC) is computed as

$$H_{kalb}(\mathbf{R}_{ij}) = \frac{\partial^2 E}{\partial \tau^i_{ka}\, \partial \tau^j_{lb}} \qquad \text{or} \qquad H_{kalb}(\mathbf{R}_{0j}) = \frac{\partial^2 E}{\partial \tau^0_{ka}\, \partial \tau^j_{lb}}. \tag{1}$$

In practice, we enforce translational consistency by working with $\mathbf{R}_{0j}$ and symmetrizing:

$$H_{kalb}(\mathbf{R}_{0j}) = H_{lbka}(-\mathbf{R}_{0j}) = \frac{1}{2}\Big[H_{kalb}(\mathbf{R}_{0j}) + H_{lbka}(-\mathbf{R}_{0j})\Big]. \tag{2}$$

This symmetrization improves numerical stability of the predicted IFCs.

c. Fourier transform to obtain the q-space force-constant matrix

The q-space force-constant matrix is formed by

$$C_{kalb}(\mathbf{q}) \;=\; \sum_{\text{cells } j} H_{kalb}(\mathbf{R}_{0j})\, e^{i\mathbf{q}\cdot\mathbf{R}_{0j}} \;=\; C^{\mathrm{ML}}(\mathbf{q}). \tag{3}$$

4. Long-range dipole correction (LO–TO splitting)

Long-range dipole interactions are incorporated using an Ewald-type dipole–dipole correction following the framework of Gonze and Lee (1997). We denote the dipole–dipole correction term by $C^{\mathrm{DD}}(\mathbf{q})$ and consider the corrected total matrix

$$C^{\mathrm{tot}}(\mathbf{q}) = C^{\mathrm{ML}}(\mathbf{q}) + C^{\mathrm{DD}}(\mathbf{q}). \tag{4}$$

Non-analytic term at Γ

The non-analytic contribution at Γ is applied as

$$C^{\mathrm{NA}}_{kalb}(\mathbf{q}\to\mathbf{0}) = \frac{4\pi}{\Omega_0}\frac{\left(\sum_c q_c Z^*_{kca}\right)\left(\sum_d q_d Z^*_{ldb}\right)}{\sum_{a'b'} q_{a'}\,\varepsilon_{a'b'}\,q_{b'}}, \tag{5}$$

where Z^* is the Born effective charge tensor, ε is the dielectric tensor, and Ω_0 is the unit-cell volume. This term should be added to $C^{\mathrm{DD}}(\mathbf{q})$ at Γ point. Because $\mathbf{q}=\mathbf{0}$ is not allowed, we use a normalized direction vector $\hat{\mathbf{q}}$ from the $\mathbf{q}$-path toward Γ and evaluate the expression for $\mathbf{q}\to\mathbf{0}$ along $\hat{\mathbf{q}}$.

Practical strategies

From Gonze and Lee (1997), the full Ewald dipole–dipole correction follows:

$$C^{\mathrm{DD}}_{\mathrm{Ew},kalb}(\mathbf{q}) = \hat{C}^{\mathrm{DD}}_{\mathrm{Ew},kalb}(\mathbf{q}) - \delta_{kl}\sum_{l'}\hat{C}^{\mathrm{DD}}_{\mathrm{Ew},kal'b}(\mathbf{q}=\mathbf{0}), \tag{6}$$

$$\hat{C}^{\mathrm{DD}}_{\mathrm{Ew},kalb}(\mathbf{q}) = \sum_{a'b'} Z^*_{ka'a} Z^*_{lb'b}\,\overline{C}^{\mathrm{DD}}_{\mathrm{Ew},ka'lb'}(\mathbf{q}), \tag{7}$$

$$\begin{aligned}\overline{C}^{\mathrm{DD}}_{\mathrm{Ew},kalb}(\mathbf{q}) = &\sum_{\mathbf{K}=\mathbf{G}+\mathbf{q}} \frac{4\pi}{\Omega_0}\frac{K_a K_b}{\sum_{a'b'} K_{a'}\varepsilon_{a'b'}K_{b'}}\, e^{i\mathbf{K}\cdot(\tau^0_k-\tau^0_l)}\, \exp\Big(-\sum_{a'b'}\frac{K_{a'}\varepsilon_{a'b'}K_{b'}}{4\Lambda^2}\Big)\\ &-(\det\varepsilon)^{-1/2}\sum_{\text{cell } j}\Lambda^3\, e^{i\mathbf{q}\cdot\mathbf{R}_{0j}}\, H_{a'b'}(\Lambda\Delta,\Lambda D)\\ &-\frac{4}{3\sqrt{\pi}}\Lambda^3\delta_{kl}\,(\varepsilon^{-1})_{a'b'}(\det\varepsilon)^{-1/2},\end{aligned} \tag{8}$$

$$H_{ab}(x,y) = \frac{x_a x_b}{y^2}\left[\frac{3}{y^3}\mathrm{erfc}(y) + \frac{2}{\sqrt{\pi}}\, e^{-y^2}\left(\frac{3}{y^2}+2\right)\right] - (\varepsilon^{-\mathbf{1}})_{ab}\left[\frac{\mathrm{erfc}(y)}{y^3} + \frac{2}{\sqrt{\pi}}\frac{e^{-y^2}}{y^2}\right] \tag{9}$$

Here,
Λ: Ewald Coeffient (Ewald result should be Λ-independent)
Z^*: Born effective charge
ε: dielectric tensor
Ω_0: Volume of the unit cell
$\mathbf{d} = \mathbf{R}_{0j} + \tau^0_l - \tau^0_k, \quad \vec{\Delta} = \underline{\underline{\varepsilon}}^{-1}\vec{\mathbf{d}}, \quad D = \sqrt{\Delta\cdot\mathbf{d}}.$

Long-range correction proved challenging for ML-predicted phonon dispersion, so we tested the following approaches:

a. **Full Ewald dipole correction:** directly add the full $C^{\mathrm{DD}}_{\mathrm{Ew}}(\mathbf{q})$.

b. **Leakage removal + full Ewald:** subtract a hypothesized "incomplete dipole-dipole leakage" term from the ML IFCs and then add full Ewald:

$$C^{\mathrm{tot}}(\mathbf{q}) = C(\mathbf{q}) + C^{\mathrm{DD}}_{\mathrm{Ew}}(\mathbf{q}) - C^{\text{DD-Leak}}(\mathbf{q}), \tag{10}$$

$$C^{\text{DD-Leak}}(\mathbf{q}) = \sum_{\text{cells } j} H^{\text{DD-Leak}}(\mathbf{R}_{0j})\, e^{i\mathbf{q}\cdot\mathbf{R}_{0j}}, \tag{11}$$

$$H_{kalb}^{DD-Leak}(\mathbf{R}_{0j}) = \sum_{a'b'} Z^*_{kaa'} Z^*_{lbb'} \left(\frac{(\varepsilon^{-1})_{a'b'}}{D^3} - \frac{3\Delta_{a'}\Delta_{b'}}{D^5} \right) \times (\det \varepsilon)^{-1/2}. \tag{12}$$

c. **Reciprocal-only Ewald (partial) correction:** add only the reciprocal-space term (from **the first term of Eqn 8**) of the Ewald expression, which converges rapidly but depends on the Ewald parameter Λ:

$$C^{\text{tot}}(\mathbf{q}) = C(\mathbf{q}) + C^{\text{DD}}_{\text{Ew-recip}}(\mathbf{q};\Lambda). \tag{13}$$

This approach produced the best qualitative result in our tests but requires calibration of Λ and may distort acoustic modes.

5. Post-processing: ASR enforcement and dynamical-matrix construction

After applying the long-range dipole correction to obtain the corrected force-constant matrix $C^{\text{tot}}_{kalb}(\mathbf{q})$, we enforce translational invariance via the acoustic sum rule (ASR). In practice, we correct the on-site blocks ($k = l$) so that the total force on a uniformly translated lattice vanishes:

$$C^{\text{ASR}}_{kalb}(\mathbf{q}) = C^{\text{tot}}_{kalb}(\mathbf{q}) - \delta_{kl} \sum_{l'} C^{\text{tot}}_{kalb}(\mathbf{q} = \mathbf{0}). \tag{14}$$

We then construct the mass-weighted dynamical matrix

$$D_{kalb}(\mathbf{q}) = \frac{1}{\sqrt{M_k M_l}} C^{\text{ASR}}_{kalb}(\mathbf{q}), \tag{15}$$

where M_k and M_l are the atomic masses of atoms k and l in the unit cell. Finally, phonon frequencies and eigenmodes are obtained by diagonalizing $D(\mathbf{q})$:

$$\sum_{l,b} D_{kalb}(\mathbf{q})\, U_{m\mathbf{q}}(l,b) = \omega^2_{m\mathbf{q}}\, U_{m\mathbf{q}}(k,a). \tag{16}$$

Here, m indexes the phonon branch and $U_{m\mathbf{q}}(k,a)$ is the polarization vector for atom k and Cartesian component a.

Results

The Foundation model used here is "NequIP-OAM-L". The target crystal is LiF, and datasets for finetuning are $3 \times 3 \times 3$ supercell with 250 samples.

1. Supercell size selection for ML prediction

Figure 2 compares phonon dispersions obtained without long-range correction for different supercell sizes. Based on this comparison, the $9 \times 9 \times 9$ supercell was selected for subsequent experiments.

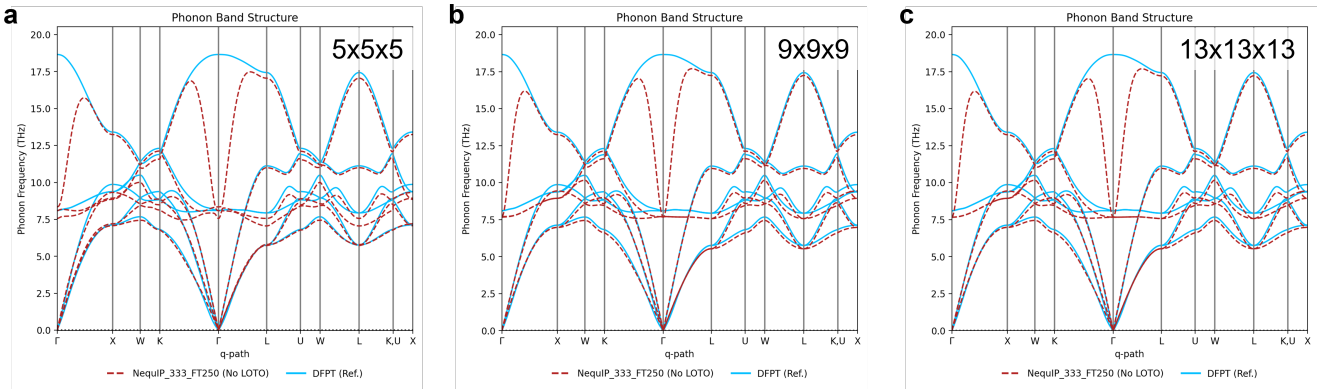

Figure 2. Phonon dispersion without long-range correction for a) 5x5x5, b) 9x9x9, c) 13x13x13 supercell

2. Long-range correction outcomes

Figure 3 shows the outcomes of (a) adding full Ewald correction, and (b,c) adding full Ewald after attempting leakage removal in different real-space regions. These attempts did not yield satisfactory agreement with DFPT results.

Figure 4 shows the reciprocal-only Ewald correction across different Λ values. Empirically, $\Lambda = 0.31$ showed the closest values with reference among the tested settings, but the correction can distort the longitudinal acoustic (LA) mode.

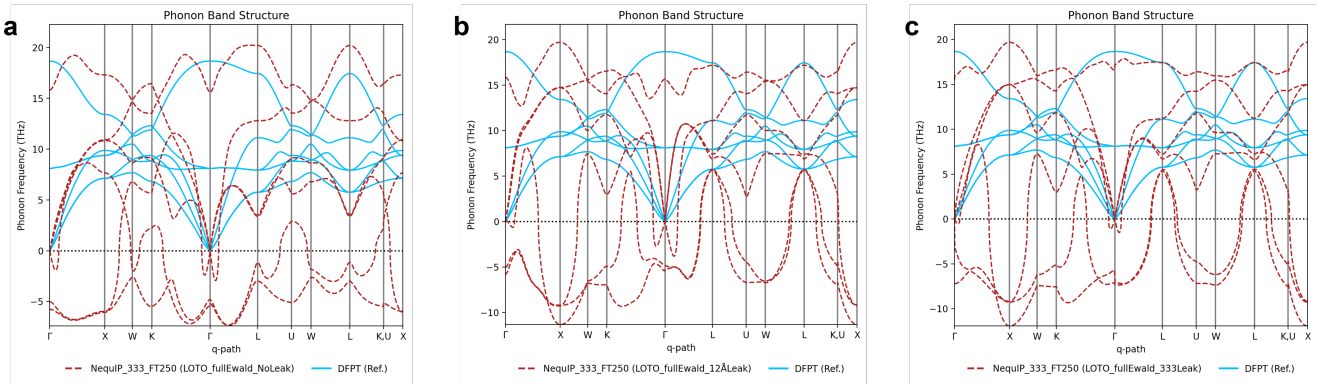

Figure 3. Phonon dispersion with adding Ewald C^{DD}: (a) only, (b) after removal of LR-Leakage in a 12 sphere, (c) after removal of LR-Leakage in a $3 \times 3 \times 3$ supercell box.

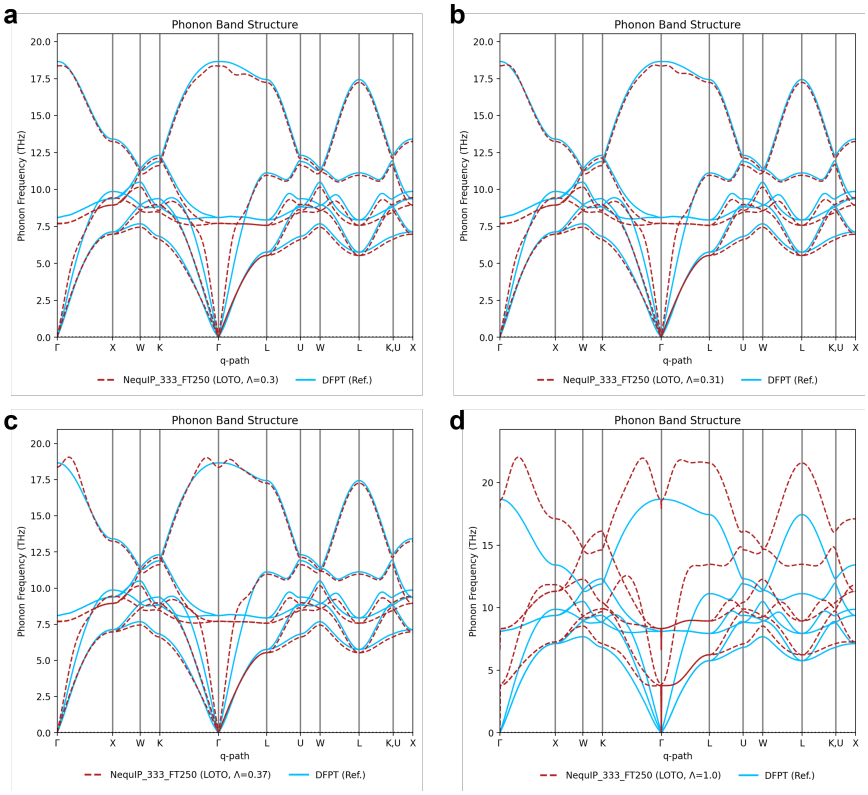

Figure 4. Phonon dispersion with adding reciprocal-only Ewald $C^{DD}_{\text{Ew-recip}}$ with (a) $\Lambda = 0.30$, (b) $\Lambda = 0.31$, (c) $\Lambda = 0.37$, (d) $\Lambda = 1.0$.

3. Computational costs

Representative wall-clock times (serial job with parallelized DFT workloads where applicable):

- Dataset construction ($3 \times 3 \times 3$ supercell, 250 samples): 56m27s (20-way parallel).
- finetuning: up to $\sim$3 days depending on convergence settings (main bottleneck).
- Prediction ($9 \times 9 \times 9$ supercell): $<$ 98m19s.

4. finetuning dataset supercell size (additional study)

In addition to selecting the prediction supercell size used to construct the real-space IFCs, we also investigated the supercell size used to generate the finetuning dataset. Since training configurations are created by rattling structures in a supercell, the dataset supercell determines the range of local environments seen during finetuning and strongly affects both accuracy and data-generation cost.

We therefore performed an additional study comparing different dataset supercell sizes (and a step-wise finetuning scheme from smaller to larger supercells) to assess the trade-off between computational cost and model performance. Figure 5 summarizes the comparison: (a) a model finetuned only on the $3 \times 3 \times 3$ dataset without LO–TO correction, (b) a step-wise finetuned model ($3 \times 3 \times 3 \rightarrow 4 \times 4 \times 4$) without LO–TO correction, (c) the $3 \times 3 \times 3$ model with LO–TO correction, and (d) the step-wise model with LO–TO correction.

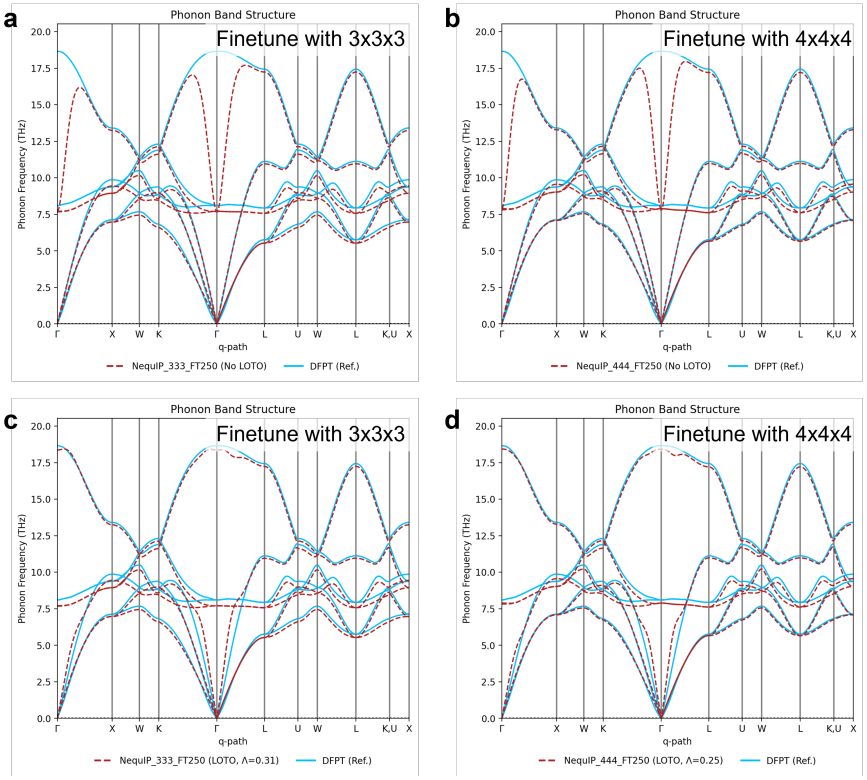

Figure 5. Stepwise Finetuning 3x3x3(250) - 4x4x4(125): Phonon dispersion of (a) 3x3x3 FT model, (b) 4x4x4 FT model without LR correction, (c) 3x3x3 FT model ($\Lambda = 0.31$), (d) 4x4x4 FT model with LR correction ($\Lambda = 0.25$).

A model finetuned on larger-supercell datasets shows improved predictions, and the dip of the LO mode near the Γ point becomes narrower. This suggests that using sufficiently large supercells for finetuning may help mitigate issues in the long-range correction. However, generating such datasets can be computationally expensive.

Discussion

Overall, the end-to-end workflow—dataset generation, NequIP finetuning, Hessian/IFC prediction—was implemented successfully and produced stable phonon dispersions in the absence of long-range effects. However, Long-range dipole correction was the main difficulty when combining ML-predicted IFCs with LO–TO splitting. Directly applying the full Ewald correction

led to poor agreement with DFPT, and "leakage removal" heuristics did not sufficiently improve results. Using only the reciprocal-space Ewald contribution yielded qualitatively improved dispersions but introduced dependence on Λ and noticeable acoustic-mode distortion (LA-mode contamination).

From a computational-cost perspective, the main bottleneck was the finetuning stage. In our setup, the runtime may be influenced by factors such as stringent convergence criteria and the size of the finetuning dataset. This suggests that there is room to reduce the overall cost by revisiting the convergence settings, optimizing the dataset sample size, and exploring more efficient training strategies (e.g., step-wise fine-tuning).

Conclusion

We presented an end-to-end NequIP-based workflow for predicting phonon dispersion from structure by (i) generating QE-based datasets, (ii) finetuning a foundation model, and (iii) predicting real-space Hessians that are Fourier transformed to obtain phonon dispersions. The long-range dipole correction remains the most delicate component: reciprocal-only Ewald correction can improve agreement with DFPT results but requires calibration and may pollute acoustic modes.

Suggested follow-ups.

a. **Improve the ML model:** reduce finetuning burden via stronger foundations or training directly with Hessian targets (e.g., Shoot from the HIP, 2025 and Koker *et al.*, 2026).

b. **Step-wise finetuning:** train progressively from smaller to larger supercells to reduce data-generation cost.

c. **Systematic calibration:** determine supercell size and Ewald parameter Λ from the model cutoff/message passing range.

d. **Mitigate acoustic-mode pollution:** project long-range correction onto optical (or LO) subspace to avoid LA-mode contamination.

Additional Information

Prepared by Minjae Kwen for personal documentation. Not intended for publication in its current form. Some text was generated with assistance from an LLM (ChatGPT) and reviewed by Minjae Kwen. The source code (configured for the Harvard FASRC cluster) is available in a private GitHub repository; please contact minjaekwen@gmail.com for access.

References

1. Batzner, S. *et al.* E(3)-equivariant graph neural networks for data-efficient and accurate interatomic potentials. *Nat. Commun.* **13**, 2453, DOI: https://doi.org/10.1038/s41467-022-29939-5 (2022).
2. Kavanagh, S. R. & MIR Group @ Harvard. mir-group/NequIP-OAM-L:0.1 (foundation potential model card and artifacts). *nequip.net* https://www.nequip.net/models/mir-group/NequIP-OAM-L:0.1; *Zenodo* https://doi.org/10.5281/zenodo.16980200 (2025).
3. Gonze, X. & Lee, C. Dynamical matrices, born effective charges, dielectric permittivity tensors, and interatomic force constants from density-functional perturbation theory. *Phys. Rev. B* **55**, 10355–10368, DOI: https://doi.org/10.1103/PhysRevB.55.10355 (1997).
4. Fang, S., Geiger, M., Checkelsky, J. G. & Smidt, T. Phonon predictions with e(3)-equivariant graph neural networks. *arXiv* https://arxiv.org/abs/2403.11347 (2024).
5. Burger, A. *et al.* Shoot from the hip: Hessian interatomic potentials without derivatives. *arXiv* https://arxiv.org/abs/2509.21624; *OpenReview (ICLR 2026 submission)* https://openreview.net/forum?id=CNLC4ZkLmW (2025).
6. Koker, T., Gangan, A., Kotak, M., Marian, J. & Smidt, T. Pft: Phonon fine-tuning for machine learned interatomic potentials. *arXiv* https://arxiv.org/abs/2601.07742 (2026).