

Supplementary Information for “Relation between electro-optic coefficient magnitude and attenuation frequency in ferroelectrics”

Mouyi Weng^{1,*} and Lin-Wang Wang^{1,†}

¹*State Key Laboratory of Optoelectronic Materials and Devices,
Institute of Semiconductors, Chinese Academy of Sciences, Beijing, China*

(Dated: September 9, 2026)

S1. RAMAN-BORN PROJECTIONS, SUPERCELL NORMALIZATION, AND RESPONSE DERIVATION

A. Atomic vertices and equilibrium covariance

The main text defines the optical impermeability $\mathbf{B} = (\boldsymbol{\epsilon}^{\text{opt}})^{-1}$ and the electro-optic (EO) coefficient. Here we give the atomic vertices and normalization factors entering those expressions. For a primitive-cell atomic displacement $u_{a\alpha}$, let $R_{a\alpha,bc} = \partial \varepsilon_{bc}^{\text{opt}} / \partial u_{a\alpha}$ and $Z_{a\alpha,k}^* = e^{-1} \partial p_k^{\text{prim}} / \partial u_{a\alpha}$, where p_k^{prim} is the primitive-cell dipole and eZ^* is the physical Born effective charge. Differentiating the inverse dielectric tensor gives

$$\frac{\partial B_{ij}}{\partial u_{a\alpha}} = - \sum_{bc} B_{ib} R_{a\alpha,bc} B_{cj}. \quad (1)$$

For a field-independent observable A and the coupling $-E_k M_k$, the canonical derivative, with $\beta = (k_B T)^{-1}$, is

$$\frac{\partial \langle A \rangle}{\partial E_k} = \frac{\partial}{\partial E_k} \frac{\int dX A e^{-\beta(H_0 - E_k M_k)}}{\mathcal{Z}(E_k)} = \beta \langle \delta A \delta M_k \rangle. \quad (2)$$

An explicit electronic field derivative of A is a separate contribution, absent from the displacement-only estimator. Applying this equilibrium identity to stationary NVE trajectories assumes that their correlations represent the intended thermal state. Each trajectory is centered separately, with consistent domain alignment and polarization branch; real hopping within a trajectory is retained.

* mouyi.weng@semi.ac.cn

† lw.wang@semi.ac.cn

For N equivalent primitive cells and a diagonal reference dielectric tensor, the static transverse contribution is

$$r_{42}^{\text{ion}}(0) = -\frac{e}{Nk_{\text{B}}T\epsilon_{yy}^{\text{opt}}\epsilon_{zz}^{\text{opt}}}\sum_{l\alpha,m\beta}R_{a\alpha,yz}Z_{b\beta,y}^*\langle\delta u_{l\alpha}\delta u_{mb\beta}\rangle. \quad (3)$$

Here l, m index primitive cells and a, b index atoms within a cell. The $1/N$ normalizes the complete within-cell and between-cell sum; it is not a pair-count average. The optical indices are yz and the applied-field direction is y . With displacements in meters, R in m^{-1} , and eZ^* in coulombs, the result is in m/V ; multiplication by 10^{12} converts it to pm/V .

B. Connection to the harmonic mode sum

The static mode representation used for the comparison in main-text Fig. 1 is a specialization to an auxiliary harmonic model. For that comparison, the eigenvectors and restoring stiffnesses are taken from the mass-weighted SSCHA auxiliary force constants at 300 K. For normalized mass-weighted zone-center optical eigenvectors, $\sum_{a\alpha}e_{a\alpha,\nu}e_{a\alpha,\mu}=\delta_{\nu\mu}$, write

$$u_{l\alpha} = \sum_{\nu} \frac{e_{a\alpha,\nu}}{\sqrt{NM_a}} q_{\nu}, \quad \langle q_{\nu} q_{\mu} \rangle = \frac{k_{\text{B}}T}{\Omega_{\nu}^2} \delta_{\nu\mu}. \quad (4)$$

Here M_a is the atomic mass and Ω_{ν}^2 is a positive eigenvalue of the mass-weighted SSCHA auxiliary force-constant matrix. The associated Ω_{ν} is an auxiliary harmonic frequency, not necessarily a dynamical spectral peak. The covariance above follows from classical equipartition within the effective quadratic model. With $R_{ij,\nu}^{(q)} = \partial B_{ij}/\partial q_{\nu}$ and $Z_{k,\nu}^{(q)} = \partial M_k/\partial q_{\nu}$, substitution into the covariance gives

$$\begin{aligned} r_{ij,k}^{\text{ion}}(0) &= \sum_{\nu} \frac{R_{ij,\nu}^{(q)} Z_{k,\nu}^{(q)}}{\Omega_{\nu}^2} \\ &= -e \sum_{bc} B_{ib} \left[\sum_{\nu,a\alpha,d\delta} \frac{R_{a\alpha,bc} Z_{d\delta,k}^* e_{a\alpha,\nu} e_{d\delta,\nu}}{\sqrt{M_a M_d} \Omega_{\nu}^2} \right] B_{cj}. \end{aligned} \quad (5)$$

Rigid translations are excluded. The mode vertices scale as $N^{-1/2}$ and $N^{1/2}$, respectively, so the result is independent of supercell normalization. No additional volume or ϵ_0 factor is needed with relative-dielectric derivatives and physical dipoles in SI units. The $4\pi/\sqrt{v_0}$ prefactor in the convention of Veithen et al. [1] follows from their Gaussian-unit Raman susceptibility, $\partial\epsilon_{bc}/\partial q_{\nu} = 4\pi\alpha_{bc}^{\nu}/\sqrt{v_0}$; it must not be applied again to the derivatives used here. This harmonic specialization does not require the collective coordinates in the main-text effective model to be phonon eigenmodes.

C. Derivation of the frequency-dependent collective-coordinate response

Here we derive main-text Eq. (7) from Eqs. (4) and (6) by a one-sided Fourier transform of the correlation equation. For a fixed tensor component and one collective coordinate, we temporarily suppress the index ν and write $R = R_{ij,\nu}$, $Z = Z_{k,\nu}$, and $C_0 = C(0)$. Define

$$\tilde{C}(\omega) = \int_0^\infty C(t)e^{i\omega t} dt. \quad (6)$$

The contribution of this coordinate to the Green-Kubo response is

$$r(\omega) = \frac{RZ}{k_B T} \left[C_0 + i\omega \tilde{C}(\omega) \right]. \quad (7)$$

Within the linear restoring-force and constant-friction model,

$$m\ddot{C}(t) + \gamma\dot{C}(t) + KC(t) = 0, \quad C(0) = C_0, \quad \dot{C}(0) = 0. \quad (8)$$

For a stable damped coordinate with $m, \gamma, K > 0$, the correlation and its derivative decay to zero as $t \rightarrow \infty$. Integration by parts therefore gives

$$\begin{aligned} \int_0^\infty \dot{C}(t)e^{i\omega t} dt &= [C(t)e^{i\omega t}]_0^\infty - i\omega \tilde{C}(\omega) \\ &= -C_0 - i\omega \tilde{C}(\omega), \end{aligned} \quad (9)$$

and

$$\begin{aligned} \int_0^\infty \ddot{C}(t)e^{i\omega t} dt &= -\dot{C}(0) - i\omega \int_0^\infty \dot{C}(t)e^{i\omega t} dt \\ &= i\omega C_0 - \omega^2 \tilde{C}(\omega). \end{aligned} \quad (10)$$

Transforming the correlation equation thus yields

$$m \left[i\omega C_0 - \omega^2 \tilde{C}(\omega) \right] + \gamma \left[-C_0 - i\omega \tilde{C}(\omega) \right] + K \tilde{C}(\omega) = 0, \quad (11)$$

so that

$$\tilde{C}(\omega) = C_0 \frac{\gamma - im\omega}{K - m\omega^2 - i\gamma\omega}. \quad (12)$$

Substitution into Eq. (7) gives

$$\begin{aligned} C_0 + i\omega \tilde{C}(\omega) &= C_0 \left[1 + \frac{i\gamma\omega + m\omega^2}{K - m\omega^2 - i\gamma\omega} \right] \\ &= C_0 \frac{K}{K - m\omega^2 - i\gamma\omega}, \end{aligned} \quad (13)$$

and hence

$$r(\omega) = \frac{RZC_0}{k_{\text{B}}T} \frac{K}{K - m\omega^2 - i\gamma\omega}. \quad (14)$$

Restoring the coordinate index and defining $r_\nu(0) = R_{ij,\nu}Z_{k,\nu}C_\nu(0)/(k_{\text{B}}T)$, $\omega_{0,\nu}^2 = K_\nu/m_\nu$, and $\Gamma_\nu = \gamma_\nu/m_\nu$, we obtain

$$r_\nu(\omega) = r_\nu(0) \frac{\omega_{0,\nu}^2}{\omega_{0,\nu}^2 - \omega^2 - i\Gamma_\nu\omega}, \quad (15)$$

which is main-text Eq. (7). This result is exact within the stated linear, constant-friction model and does not require an overdamped approximation.

S2. HSE EO SPECTRA FOR CUBIC SUPERCELLS

Figure S1 compares the effective ionic EO spectra for N^3 supercells with $N = 8, 12, 16, 20$, and 24 at 300 K. All five curves use 0.15 ns segments. The plotted quantity is the magnitude of the complex mean, $|\bar{r}_{\text{eff}}(f)|$, with $r_{\text{eff}} = r_{42} + r_{13}/2$. The complex real and imaginary parts are combined before taking the magnitude; the curves are not averages of segment magnitudes. No additional frequency smoothing or fitting is applied. At 10 GHz, the magnitudes are $471.3, 704.1, 706.9, 606.2$, and 590.2 pm/V for $N = 8, 12, 16, 20$, and 24 , respectively.

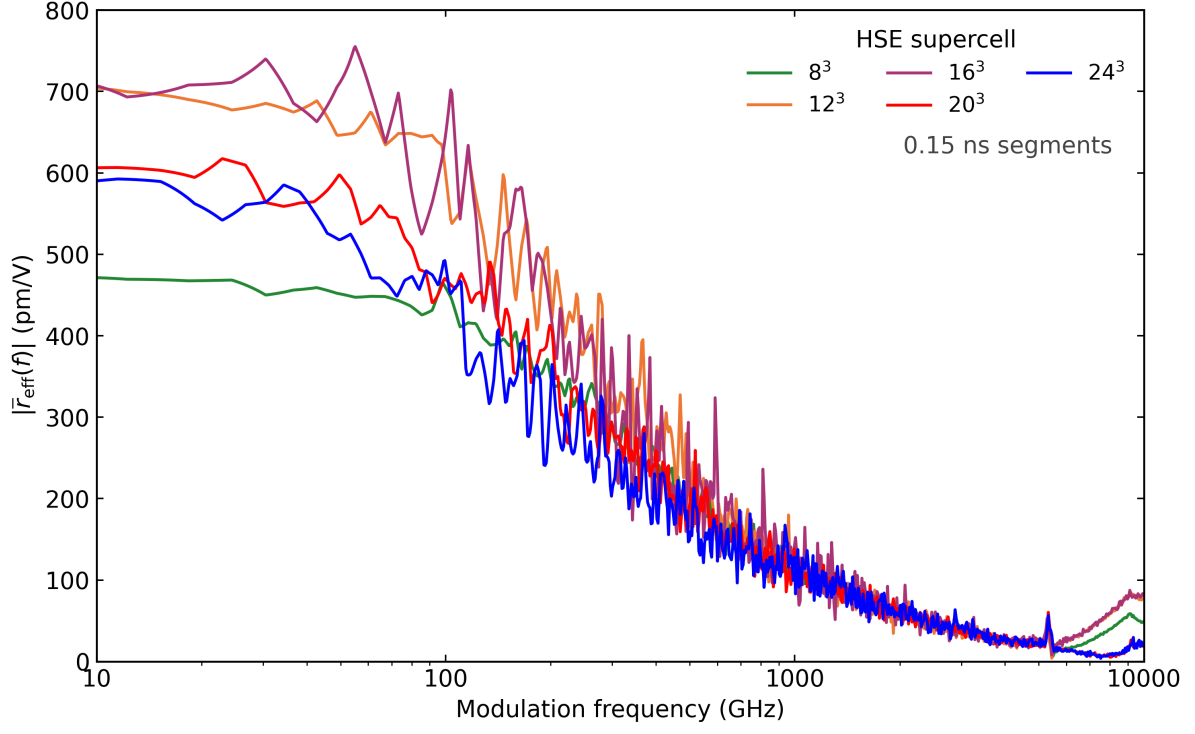


FIG. S1. **HSE EO spectra for different cubic supercell sizes.** Magnitude of the complex mean effective EO coefficient for 8^3 , 12^3 , 16^3 , 20^3 , and 24^3 supercells, using 0.15 ns segments. The full supplied frequency range from 10 GHz to 10 THz is shown. The 20^3 and 24^3 curves are the same data used in main-text Fig. 2(b), which displays frequencies up to 5 THz.

S3. SHELL-RESOLVED TI-TI FORCE RESPONSE

TABLE S1. Signed in-plane projected Ti-Ti response for twelve representative directions in the $24 \times 24 \times 24$ grid dataset. Values of $\beta_{\parallel}^{xy}$ are in eV Å⁻². The PBE SR and HSE SR columns correspond to the source labels **pbe_nep** and **hse_nep**, respectively. Both contributions use the same projection.

Direction	$R_{\text{Ti-Ti}}$ (Å)	PBE SR	HSE SR
[100]	4.04	5.1×10^0	5.5×10^0
[110]	5.82	1.0×10^0	1.1×10^0
[111]	7.06	2.5×10^{-1}	2.9×10^{-1}
[200]	8.02	7.3×10^{-1}	8.0×10^{-1}
[210]	9.05	2.7×10^{-1}	3.3×10^{-1}
[211]	9.93	1.1×10^{-1}	1.5×10^{-1}
[220]	11.54	6.5×10^{-2}	6.5×10^{-2}
[300]	12.08	3.2×10^{-2}	4.8×10^{-2}
[221]	12.08	1.7×10^{-2}	3.2×10^{-2}
[310]	12.61	-1.7×10^{-3}	8.2×10^{-3}
[311]	13.46	4.8×10^{-3}	4.2×10^{-3}
[222]	14.00	-8.6×10^{-4}	-4.7×10^{-5}

S4. CONTROLLED SHORT-RANGE POTENTIAL SEQUENCE

TABLE S2. EO coefficient and nearest-neighbor correlations for four potentials. Amplitudes are $|\bar{r}_{\text{eff}}(10 \text{ GHz})|$. The $\lambda = 0$ and 1 results use 0.15 ns segment means; $\lambda = 1.5$ and 2 use full-trajectory means. Brackets on amplitudes show 95% confidence intervals for the matching processing window. For $\lambda = 0$ and 1 these are archived hierarchical trajectory/window bootstrap intervals; the $\lambda = 1.5$ interval is retained from its full-current analysis, and the updated $\lambda = 2$ interval uses 2000 trajectory-bootstrap samples. All f_D values are reported as point estimates without uncertainty intervals. All reported fits use 25–5000 GHz with equal weights. The fitted $\lambda = 2$ frequency lies below the lower fit bound. Correlation uncertainties are SEM.

λ	$ r_{\text{eff}}(10 \text{ GHz}) \text{ (pm/V)}$	$f_D \text{ (GHz)}$	$C_{xy}(a\hat{x}, 0)$
0	196.3 [194.2, 199.0]	1211.59	0.4058 ± 0.0011
1	508.9 [493.3, 525.6]	220.05	0.4998 ± 0.0012
1.5	1055.6 [949.8, 1172.5]	56.53	0.5722 ± 0.0035
2	1634.6 [1457.0, 1860.6]	20.41	0.6328 ± 0.0025

S5. LAYER CORRELATIONS AND FIXED-THICKNESS SIZE SERIES

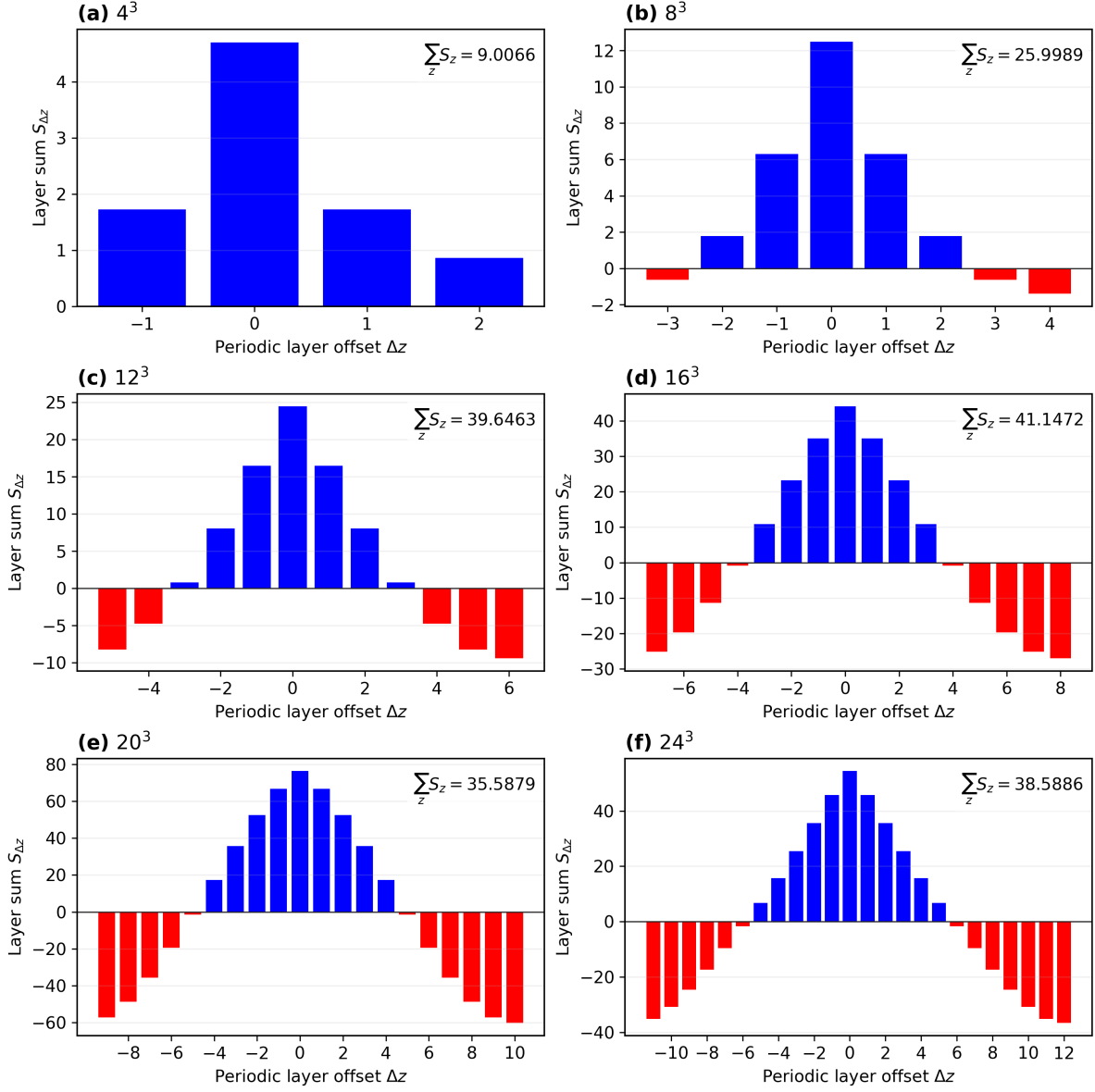


FIG. S2. **Layer-resolved HSE correlation sums.** (a-f) Results for the 4^3 , 8^3 , 12^3 , 16^3 , 20^3 , and 24^3 cells, respectively. At each periodic layer offset Δz , the layer sum is $S_{\Delta z} = \sum_{\Delta x, \Delta y} C_{xy}(\Delta x, \Delta y, \Delta z; 0)$, with the self-correlation at $(\Delta x, \Delta y, \Delta z) = (0, 0, 0)$ excluded. Blue and red bars denote positive and negative layer sums, respectively. Each panel reports the signed sum over all layers; the vertical scales differ between panels.

TABLE S3. HSE fixed- z size series. Every amplitude is the magnitude of the complex mean at 10 GHz. All rows use 0.15 ns segments except $24 \times 24 \times 4$, which retains the available full-current mean. The $8 \times 8 \times 4$ row reuses the $\lambda = 1$ result, including its fit; all reported fits use 25–5000 GHz with equal weights. Brackets on amplitudes are 95% sampling confidence intervals. Amplitude intervals for 0.15 ns means use the archived hierarchical trajectory/window bootstrap. For $24 \times 24 \times 4$, the amplitude interval uses 2000 trajectory-bootstrap replicates. All f_D values are reported as point estimates without uncertainty intervals. Correlation uncertainties are SEM.

Cell	$ r_{\text{eff}} $ (pm/V)	f_D (GHz)	$C_{xy}(a\hat{x}, 0)$
$6 \times 6 \times 4$	328.1 [316.1, 341.6]	313.73	0.5051 ± 0.0028
$8 \times 8 \times 4$	508.9 [493.3, 525.6]	220.05	0.4998 ± 0.0012
$12 \times 12 \times 4$	909.2 [858.2, 954.8]	143.98	0.4979 ± 0.0027
$16 \times 16 \times 4$	1510.2 [1431.5, 1596.3]	82.55	0.5016 ± 0.0016
$20 \times 20 \times 4$	2004.0 [1909.5, 2104.4]	60.10	0.4997 ± 0.0013
$24 \times 24 \times 4$	2465.5 [2308.9, 2652.7]	47.38	0.4981 ± 0.0015

S6. FORCE-MODEL VALIDATION

The HSE validation contains 59 independent 320-atom frames. The bare NEP force-component RMSE is 21.9 meV/Å; adding the fixed Born–LR correction gives 21.3 meV/Å, a 2.38% reduction. The corresponding MAE values are 16.7 and 16.4 meV/Å, and the absolute 95th-percentile errors are 44.2 and 42.9 meV/Å. The Born–LR force-component RMS is 6.43 meV/Å.

TABLE S4. Direct HSE force validation and the supplied MACE force-error summary.

Model	RMSE (meV/Å)	MAE (meV/Å)	abs. 95% (meV/Å)
NEP short range	21.9	16.7	44.2
NEP+Born–LR	21.3	16.4	42.9
MACE	11.3	8.7	23.3

S7. COMPLEX EO SPECTRA AND DEBYE FITS

Figure S3 shows the complex spectra and fits obtained using the protocol described in Methods, “Spectral analysis and relaxation fitting.”

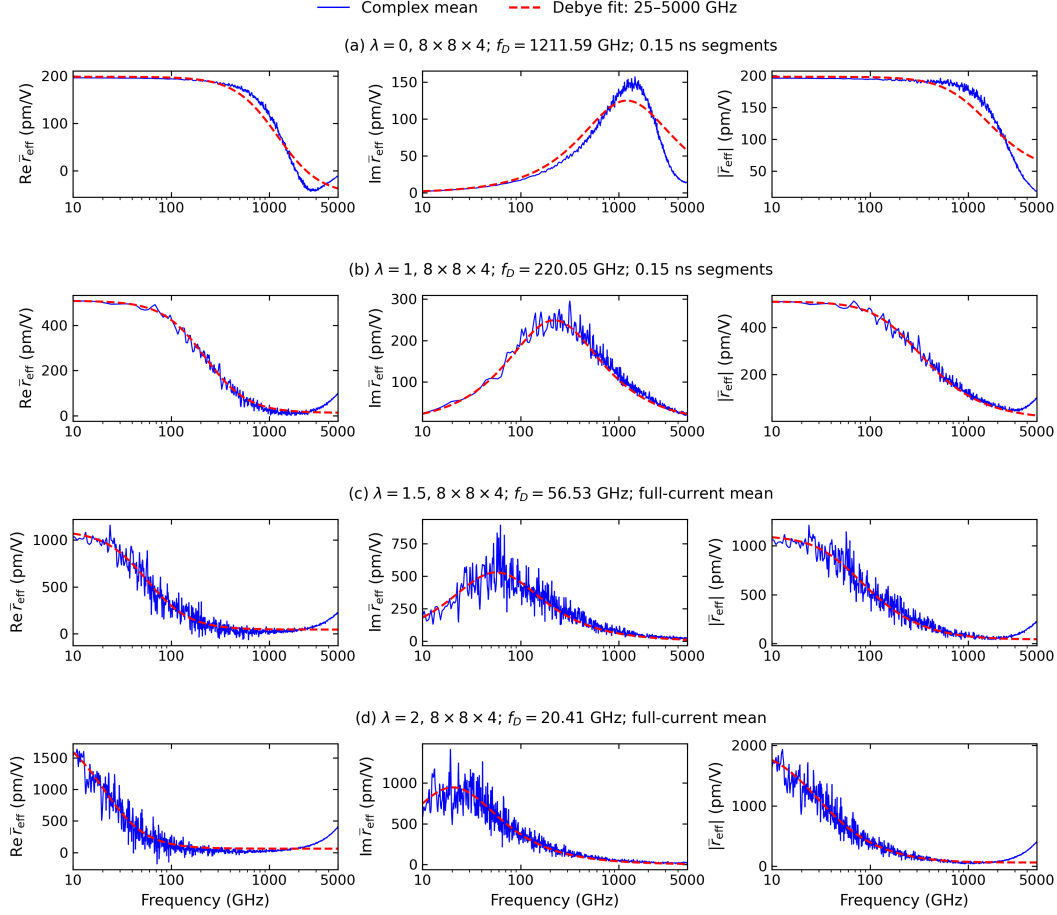


FIG. S3. **Complex EO coefficients and Debye fits.** Each row shows the real part, imaginary part, and magnitude of the same complex mean (pure blue), with the 25–5000 GHz equal-weight fit (dashed red). Rows (a–d) show the $8 \times 8 \times 4$ potential series, $\lambda = 0, 1, 1.5$, and 2 . The row headings specify the processing window and fitted f_D . Curves below 25 GHz are extrapolated outside the fit interval. For the PBE endpoint ($\lambda = 0$), systematic deviations in the THz region show that the single-Debye fit is an effective summary of the spectrum. For $\lambda = 2$, $f_D = 20.41$ GHz lies below the 25 GHz lower fit bound and is an extrapolated model parameter. The figure continues on the following two pages.

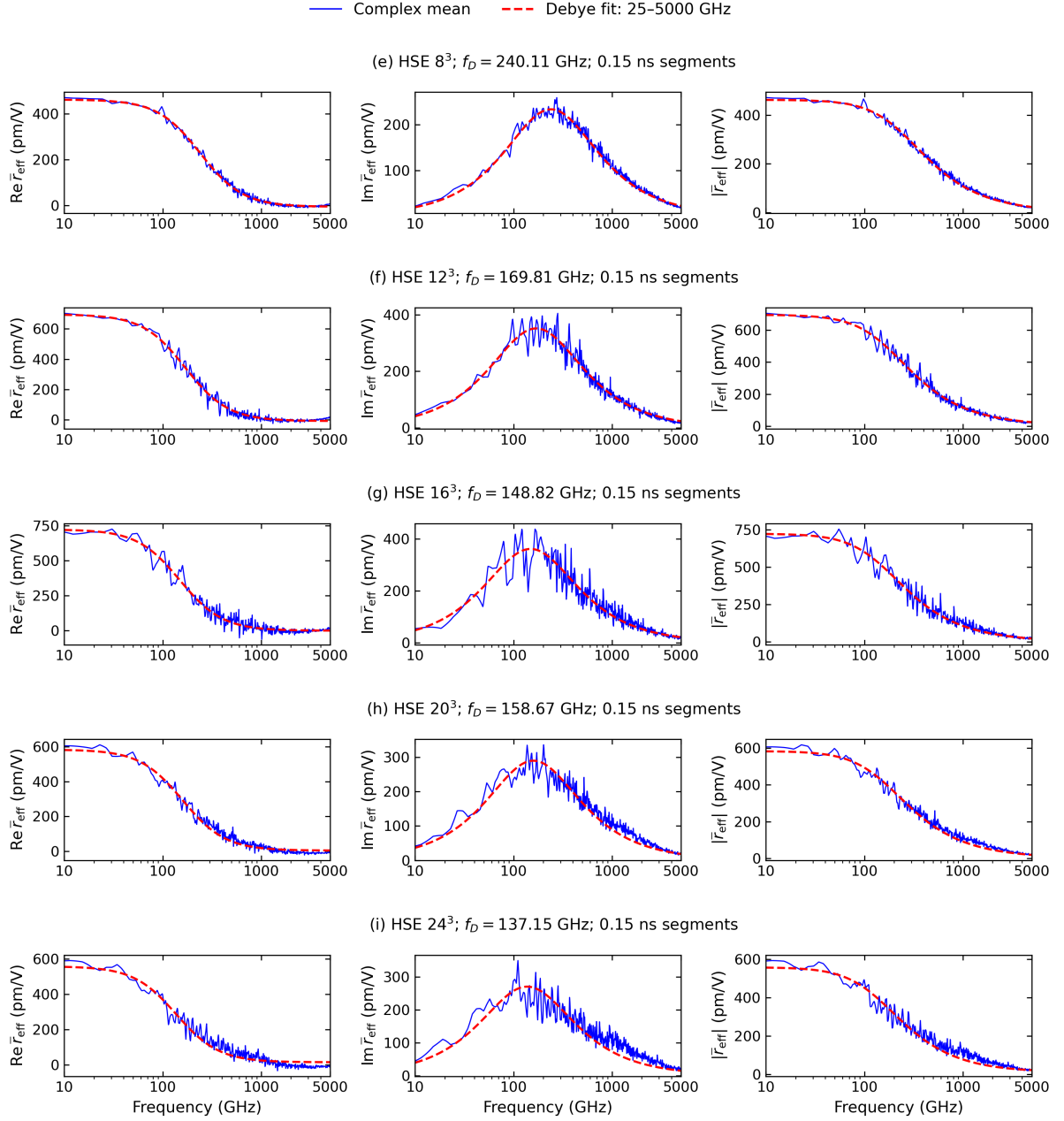


FIG. S3. **Continued: cubic HSE supercells.** Rows (e–i) show 8^3 , 12^3 , 16^3 , 20^3 , and 24^3 , respectively. All use 0.15 ns segment means. Colors and columns follow the first page.

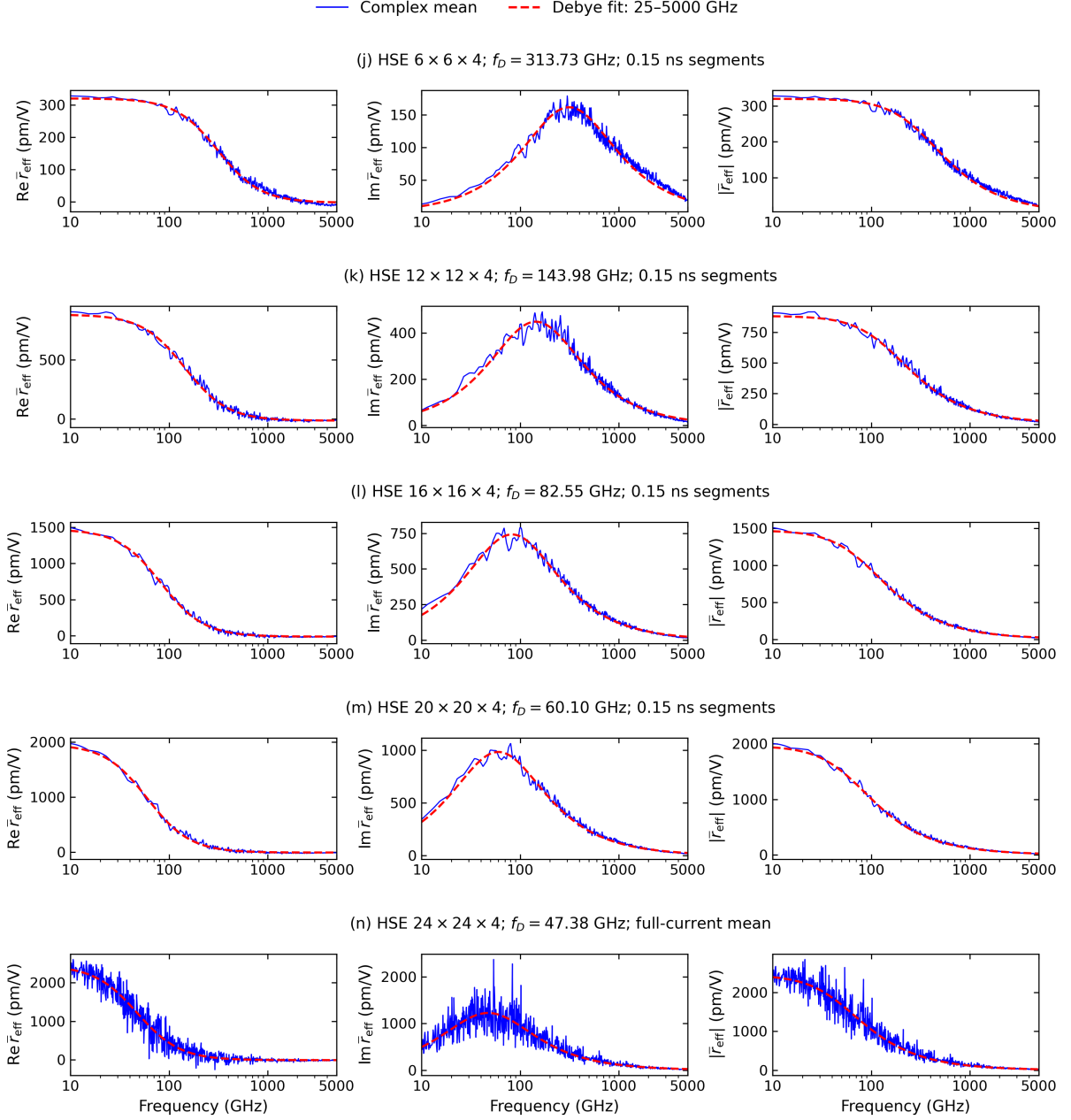


FIG. S3. **Continued: fixed-thickness HSE cells.** Rows (j–n) show $N_x = N_y = 6, 12, 16, 20$, and 24 with $N_z = 4$. The shared $8 \times 8 \times 4$ HSE result appears once, in row (b). The $24 \times 24 \times 4$ row uses a full-current mean; the others use 0.15 ns segments. Styling follows the first page.

S8. OPTICAL-MODE RELAXATION AND COMPARISON WITH THE EO RESPONSE

The extraction of the collective coordinates and the spectral fitting protocol are described in Methods, “Spectral analysis and relaxation fitting.” Here we provide the mode correspondence, the relation between displacement spectra and EO response, and the detailed comparison for the HSE 24^3 supercell at 300 K.

The two leading MD covariance coordinates account for 93.2% of the covariance trace. Their subspace has a mean squared overlap of 0.99754 with the PBE SSCHA reference optical modes 4 and 5 (the two principal angles are approximately 2.84°). This establishes the correspondence of the collective displacement patterns; it does not assume that covariance eigenvectors diagonalize the dynamics. The SSCHA vectors are used only for this subspace comparison and do not define the two MD covariance coordinates.

For the PSD convention defined in Methods, $S_\nu(f) = 2 \operatorname{Re} \int_0^\infty C_\nu(t) e^{i2\pi f t} dt$. For a single exponential, $C_\nu(t) = C_\nu(0) e^{-t/\tau_\nu}$ for $t \geq 0$, the Fourier transform gives $S_\nu(f) = 2C_\nu(0)\tau_\nu/[1 + (2\pi f\tau_\nu)^2]$. Combining this expression with the single-coordinate Green–Kubo response gives

$$\frac{S_\nu(f)}{S_\nu(0)} = \frac{1}{1 + (f/f_{D,\nu})^2} = \left| \frac{r_\nu(f)}{r_\nu(0)} \right|^2, \quad f_{D,\nu} = \frac{1}{2\pi\tau_\nu}. \quad (16)$$

This equality applies to an isolated relaxing contribution without a background. It motivates the square-root display in panel (b), but is not an identity between an unweighted average of modal spectra and the total EO response, which contains optical/electrical weights, other modes, and possible cross correlations.

Using the Lorentzian-plus-background fitting protocol specified in Methods, fits to the two coordinate spectra and their mean give 91.47, 119.30, and 104.09 GHz, respectively. The mean-spectrum result corresponds to an effective relaxation time of approximately 1.53 ps. All values are point estimates without uncertainty intervals.

The EO curve is the same 24^3 complex mean used in the main text. The joint complex fit specified in Methods gives $f_D = 137.15$ GHz, corresponding to approximately 1.16 ps. The different spectral estimators and fit intervals should be kept in mind when comparing this value with the 104.09 GHz mean displacement-spectrum result. For panel (b), the raw curves and their corresponding fits are divided by the same raw values at 100 GHz, using linear interpolation where needed; only the PSD curves are then square-root transformed. No

fit is performed on the square-root spectrum, and the fitted curves are not forced through unity at the normalization frequency. The scalar PSD background and the complex EO background are fitted separately and are not identified with each other. The comparison tests the scale and trend of the modal relaxation; it does not extract a unique microscopic friction coefficient or implement a Raman–Born-weighted sum over all optical modes.

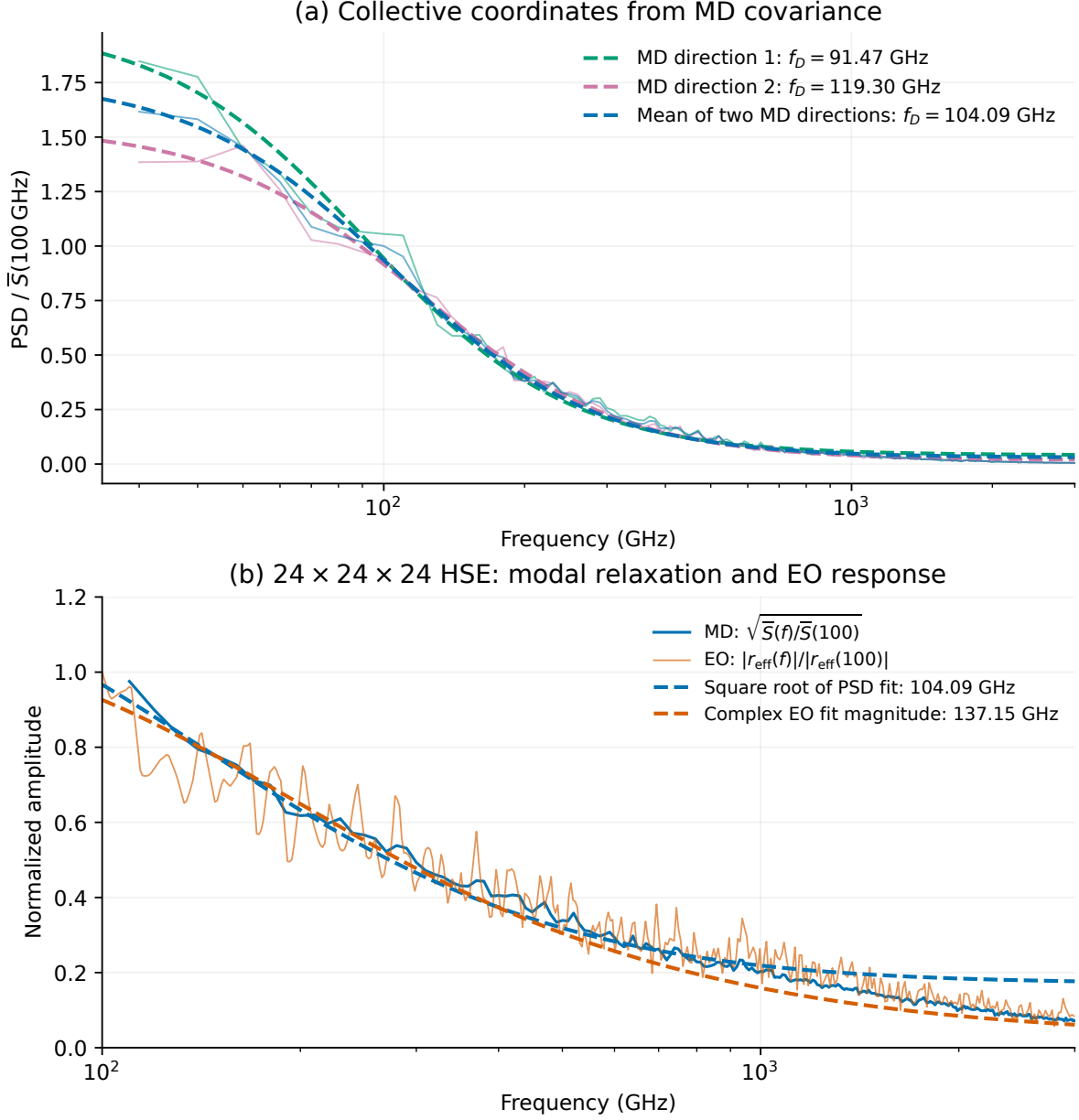


FIG. S4. **Optical-mode displacement spectra and EO attenuation in HSE 24^3 .** (a) PSDs of the two leading MD covariance coordinates and their arithmetic mean (solid curves), with free-background Lorentzian fits (dashed curves). All three spectra share the raw mean-spectrum normalization at 100 GHz. (b) Square root of the normalized mean PSD compared with the normalized Green-Kubo EO magnitude. Dashed curves are the square root of the PSD fit and the magnitude of the joint complex EO fit, respectively. The mean PSD gives $f_D = 104.09 \text{ GHz}$ and the complex EO response gives $f_D = 137.15 \text{ GHz}$. PSD and EO window lengths are 100 and 150 ps, respectively. Normalization at 100 GHz is for display and does not constrain the fits.

S9. TI DISPLACEMENT DISTRIBUTIONS AT THE TRAINING SUPERCELL SIZE

The PBE trajectories used to select and expand the training configurations were generated in $4 \times 4 \times 4$ supercells (320 atoms). Figure S5 shows substantial overlap of the main PBE and HSE displacement regions at this size, although HSE retains more pronounced four-fold maxima. The 4^3 PBE distribution is broader and flatter near its center than the 12^3 distribution in main-text Fig. 2(c). Thus, that large-cell comparison does not represent the PBE training pool. This local comparison does not establish complete many-atom configuration coverage; force validation is reported in Section S6.

The histograms use a 240×240 grid over $[-0.34, 0.34]$ Å per component. We plot the supplied Gaussian-smoothed grids (standard deviation of two bins), normalized by their respective maxima, without further smoothing.

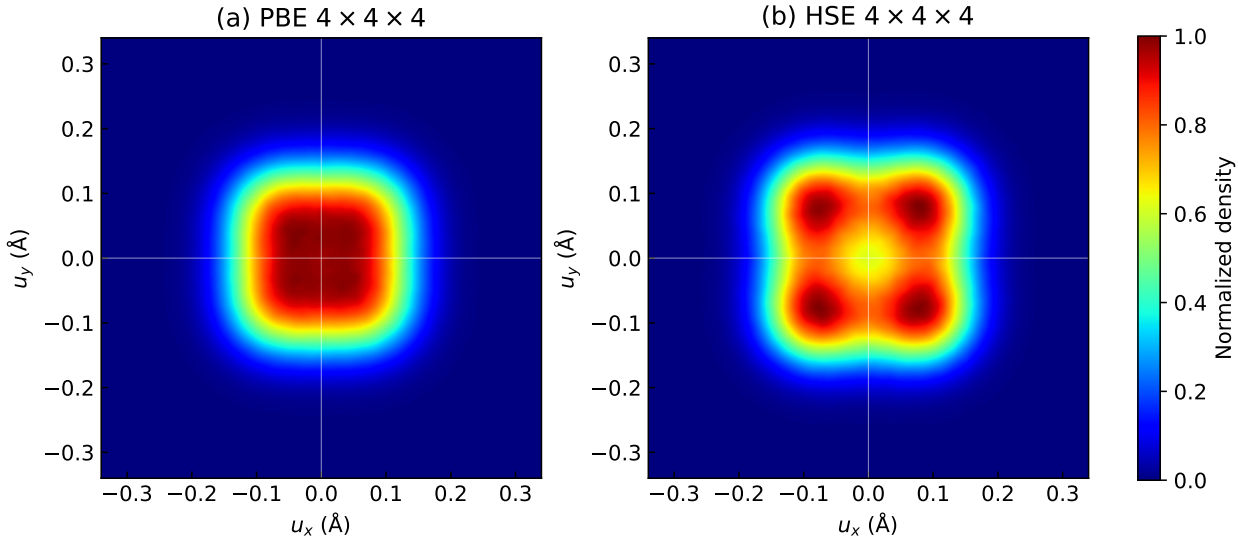


FIG. S5. **Ti off-centering distributions in $4 \times 4 \times 4$ supercells.** In-plane Ti displacements relative to the oxygen cage for (a) PBE and (b) HSE. Each smoothed distribution is normalized by its own maximum; the common color scale compares distribution shapes rather than absolute probability densities. These cells have the same size as those used to generate the PBE training configurations.

-
- [1] M. Veithen, X. Gonze, and P. Ghosez, Nonlinear optical susceptibilities, raman efficiencies, and electro-optic tensors from first-principles density functional perturbation theory, [Physical Review B](#) **71**, 125107 (2005).