Editors' Suggestion

Investigating anharmonicities in polarization-orientation Raman spectra of acene crystals with machine learning

Paolo Lazzaroni, Shubham Sharma, and Mariana Rossi*
MPI for the Structure and Dynamics of Matter, Luruper Chaussee 149, 22761 Hamburg, Germany

(Received 2 October 2025; revised 17 December 2025; accepted 14 January 2026; published 20 February 2026; corrected 4 November 2026)

We present a first-principles machine-learning computational framework to investigate anharmonic effects in polarization-orientation (PO) Raman spectra of molecular crystals, focusing on anthracene and naphthalene. By combining machine learning models for interatomic potentials and polarizability tensors, we enable efficient, large-scale simulations that capture temperature-dependent vibrational dynamics beyond the harmonic approximation. Our approach reproduces key qualitative features observed experimentally. We show, systematically, what are the fingerprints of anharmonic lattice dynamics, thermal expansion, and Raman tensor symmetries on PO-Raman intensities. However, we find that the simulated polarization dependence of Raman intensities shows only subtle deviations from quasiharmonic predictions, failing to capture the pronounced temperature-dependent changes that have been reported experimentally in anthracene. We propose that part of these inconsistencies stem from the impossibility to deconvolute certain vibrational peaks when only experimental data is available. This work therefore provides a foundation to improve the interpretation of PO-Raman experiments in complex molecular crystals with the aid of theoretical simulations.

DOI: 10.1103/t4s2-45t8

I. INTRODUCTION

Molecular crystals are a class of solids where the molecular units, which constitute the basis of the crystal lattice, are kept together by intermolecular (noncovalent) interactions. The weak nature of these forces, when compared to those responsible in covalent bonds, allows one to identify two distinct energy scales in the nuclear lattice dynamics. At higher energies, intramolecular vibrations usually give rise to phonon branches with flat dispersion, as a result of the weakly perturbed internal motion of the individual molecules. In contrast, at low energies, the collective translations and hindered rotations of the molecular units are responsible for the unique spectral fingerprints of a given crystal structure.

These low-frequency modes, typically below 200 cm^{-1}, dictate numerous properties of organic molecular crystals. It has been discussed, for example, how this region determines polymorph ordering at finite temperatures, by strongly contributing to the vibrational free energy [1–3], or how it dominates heat [4,5] and charge [6,7] carrier mobility in organic semiconductors. Because the weak intermolecular interactions lead to large-amplitude motion, these collective lattice vibrations can exhibit pronounced anharmonic character. Taking into account deviations from the ideal harmonic crystal picture is especially relevant when modeling these systems.

Recently, it was shown [8,9] that polarization-orientation (PO) vibrational Raman scattering experiments conducted on several molecular crystals, including some linear oligoacenes, reveal an effect that has been attributed to anharmonic coupling between low-frequency intermolecular phonon modes. This effect consists of a deviation of the polarization dependence of the Raman intensity of some phonon modes from the prediction of the harmonic crystal approximation. To rationalize this finding, a two-mode model, allowing phonon-phonon coupling in the description of inelastic light scattering, was presented in Ref. [10]. Parameters that were fitted on this model, based on the experimental data, could successfully explain the pattern that was measured and led to the conclusion that anharmonic mode-coupling terms are necessary for a complete description of the Raman scattering process in molecular crystals.

In this paper, we provide an *ab initio* quality framework to reproduce, explain and provide quantitative insights into anharmonic effects directly probed by PO-Raman spectroscopy. As a case study, we focus on linear oligoacene molecular crystals, in particular naphthalene and anthracene. The standard approach of combining *ab initio* molecular dynamics (AIMD) with density functional perturbation theory (DFPT) to obtain Raman spectra guarantees that all orders of anharmonicity of the potential energy surface (PES) are included, but is prohibitively expensive for this task. This is due, for example, to the vast range of experimental conditions one needs to explore and the large cell sizes required to accurately describe the phonon manifold. Furthermore, to capture subtle variations in the angle-resolved Raman intensity, the spectra have to be thoroughly converged, and finite-size artifacts due to periodic boundary conditions need to be mitigated. While ML-based frameworks for computing Raman spectra without explicitly accounting for the crystal orientation

*Contact author: mariana.rossi@mpsd.mpg.de

Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI. Open access publication funded by Max Planck Society.

have been demonstrated [11–14], this work instead presents a ML framework to obtain accurate polarization-orientation Raman spectra across multiple temperatures. This extension is nontrivial: PO-Raman intensities provide a significantly more stringent test of ML accuracy than angle-integrated spectra, as subtle errors in polarizability tensor components or normal mode symmetries dramatically alter angular dependence while remaining invisible in unpolarized spectra. We thus combine state-of-the-art equivariant ML models for both interatomic potentials and polarizability tensors with the computationally efficient Raman-tensor weighted Γ-point density of states (RGDOS) approximation proposed in Refs. [15,16] to enable quantitative predictions of temperature-dependent anharmonic PO-Raman signals.

In this manuscript, we begin by introducing the theoretical background of PO-Raman scattering, emphasizing the limitations of the harmonic approximation and the need for anharmonic treatments. After benchmarking our ML models and validating the RGDOS approximation against full ML calculations, we simulate the temperature-dependent PO-Raman spectra of anthracene and naphthalene, capturing anharmonic effects. By comparing our results with experimental data, we analyze the polarization dependence of Raman intensities and discuss the subtle deviations from the harmonic predictions. Our findings highlight the interplay between anharmonic lattice dynamics and Raman tensor symmetries, providing insights into the origins of the temperature-dependent behavior of low-frequency Raman modes and suggesting guidelines for the interpretation of the rich amount of information contained in PO-dependent Raman spectra. This work demonstrates the power of combining machine learning with established theoretical frameworks to unravel complex anharmonic phenomena in molecular crystals.

II. THEORY OF POLARIZATION-ORIENTATION RAMAN SCATTERING

The Raman effect, first described in 1928 [17], is a process where photons undergo inelastic scattering on matter, resulting in a finite frequency (and polarization) shift between incident and scattered radiation. At a certain energy range, the spectrum of frequency differences is a manifestation of vibrational transitions occurring in the scattering medium, providing unique fingerprints of the vibrations of the material.

Single crystal polarization-orientation Raman (also referred to as angle-resolved polarized Raman) experiments are often conducted in a back- (or forward-) scattering geometry, where the incident ($\boldsymbol{e}_1$) and scattered ($\boldsymbol{e}_2$) fields are linearly polarized and propagate through the crystal, which is positioned at a certain orientation [8,18,19]. The scattered radiation is then collected with polarization either perpendicular ($\boldsymbol{e}_1 \perp \boldsymbol{e}_2$) or parallel ($\boldsymbol{e}_1 \parallel \boldsymbol{e}_2$) to that of the incident beam. Once the configuration is fixed, the experiment is repeated spanning an angle θ between $\boldsymbol{e}_1$ and an arbitrary reference, either by rotating the sample or the direction of $\boldsymbol{e}_1$.

Within a semiclassical treatment, where the light is treated classically and the matter quantum-mechanically, the expression for the Raman cross-section in the electric dipole approximation reads [20,21]

$$I(\omega,\beta,\theta) \propto \sum_{i,f} \rho_i^\beta \langle i|\boldsymbol{e}_1\cdot\boldsymbol{\alpha}\cdot\boldsymbol{e}_2|f\rangle^2 \delta(\omega-\omega_{fi}), \quad (1)$$

where $\boldsymbol{\alpha}$ is the polarizability tensor of the matter and, within the Born-Oppenheimer (BO) approximation, the summation runs over all vibrational states, with ρ_i^β being the thermal equilibrium population associated to the initial state at inverse temperature β. Here $\boldsymbol{e}_1$ and $\boldsymbol{e}_2$ are the polarization vectors defined above, either for the parallel or perpendicular setting. The unpolarized Raman spectrum can be obtained by integrating the Raman intensity over all angles in both configurations, and it is defined as

$$I^{\text{unpolarized}}(\omega,\beta) = \int_0^{2\pi} \frac{d\theta}{2\pi} [I^\perp(\omega,\beta,\theta) + I^\parallel(\omega,\beta,\theta)]. \quad (2)$$

If we restrict our description to nonresonant Raman, i.e., when the energy of the incident radiation is far below any electronic transition of the system, then only the static polarizability tensor of the electronic ground state is needed [21]. Within perturbation theory, its components take the form

$$\alpha_{qp} = 2\sum_{\sigma>0} \frac{\langle 0|\mu_p|\sigma\rangle\langle\sigma|\mu_q|0\rangle}{\varepsilon_\sigma - \varepsilon_0}, \quad (3)$$

where $|0\rangle$ is the electronic ground state, the summation runs over all the excited electronic states, $\boldsymbol{\mu}$ is the electric dipole operator, and q, p denote Cartesian components. With this definition of the polarizability tensor, Eq. (1) becomes the well-known Placzek approximation [21,22] to the full expression of the light scattering cross-section derived first by Kramers, Heisenberg and Dirac [23,24]. Note that, as recently highlighted in Ref. [25], by treating the quantity in Eq. (3) as the polarizability tensor (or dielectric susceptibility) entering Eq. (1), we are already excluding the effect of finite scattering wave-vectors, essentially assuming that only vibrations at the Γ point of the crystal phonon Brillouin zone are probed in the experiment, i.e., $\boldsymbol{q} = \hat{\boldsymbol{k}}_1 - \hat{\boldsymbol{k}}_2 \approx 0$, where the $\hat{\boldsymbol{k}}$ are the propagation directions of incoming and scattered fields. This tends to be a very good assumption for first-order Raman scattering, but breaks down in the context of second-order Raman scattering in polar materials [25].

At this stage, the polarizability tensor can be regarded as an operator for the nuclei, and, in the BO picture, it depends parametrically on the nuclear coordinates through the electronic states of Eq. (3). Thus, it is commonly expanded in a Taylor series around the nuclear equilibrium configuration $\boldsymbol{Q}_0$,

$$\alpha_{qp}(\boldsymbol{Q}) = \alpha(\boldsymbol{Q}_0) + \sum_k \left.\frac{\partial\alpha_{qp}(\boldsymbol{Q})}{\partial Q_k}\right|_{\boldsymbol{Q}=\boldsymbol{Q}_0} Q_k + \ldots, \quad (4)$$

where k runs over all nuclear degrees of freedom. While the first term in the expansion above is responsible for Rayleigh scattering, the higher-order terms, involving polarizability derivatives, enter the expression for the Raman cross-section. If the coordinates are chosen to be the mass-scaled normal modes of the system $\tilde{Q}_k$, then these derivatives are referred to as Raman tensors, $\boldsymbol{\mathcal{R}}_k$, i.e., at first order,

$$\boldsymbol{\mathcal{R}}_k^{(1)} = \frac{\partial\alpha_{qp}(\boldsymbol{Q})}{\partial\tilde{Q}_k}. \quad (5)$$

Note that $\tilde{Q}_k$ are Cartesian-space vectors, obtained by removing the mass-scaling that exists in the eigenvectors that come out of the diagonalization of the dynamical matrix. As such,

they form a nonorthonormal basis that needs to be handled carefully in common operations such as normal-mode projections of molecular dynamics trajectories, as will be relevant later in the manuscript (see Sec. S1 of the Supplemental Material (SM) [26]).

Retaining only the first-order term in the expansion of Eq. (4), the matrix element of Eq. (1) is approximated as

$$\langle i|\boldsymbol{e}_1\cdot\boldsymbol{\alpha}\cdot\boldsymbol{e}_2|f\rangle \approx \sum_k \langle i|\tilde{Q}_k|f\rangle \boldsymbol{e}_1\cdot\boldsymbol{\mathcal{R}}_k\cdot\boldsymbol{e}_2. \tag{6}$$

At this stage, if the vibrational Hamiltonian is assumed to be harmonic in the normal coordinates, then the initial and final states can be written as product states of harmonic eigenstates. It is easy to see, within these two approximations, that the matrix elements of the position operators in Eq. (6) connect only states that differ by a single vibrational excitation in the respective mode, giving rise to the common first-order Raman selection rule.

For the same reason, when the square modulus of Eq. (6) is taken, the cross-products between different eigenmodes vanish, leaving us with the expression for the Raman cross-section,

$$I(\omega,\beta,\theta) \propto \sum_k \rho_k^\beta |\boldsymbol{e}_1\cdot\boldsymbol{\mathcal{R}}_k\cdot\boldsymbol{e}_2|^2\delta(\omega-\omega_k), \tag{7}$$

that will depend trivially on temperature through the population of the initial states. These derivations can also be found in the SM of Ref. [10]. We will often refer to the θ angular dependence, for a given k, as the "PO-pattern" of the corresponding mode in the following.

The result of Eq. (7) is the well-known expression for the (harmonic) polarized Raman scattering cross-section that is commonly employed to rationalize experimental results in terms of Raman tensors and factor group analysis [27]. As shown above and already clearly stressed in Ref. [10], this simple expression is the result of what is often termed a "double harmonic" approximation, as it relies on both harmonic normal modes and the truncation at first order of the expansion in Eq. (4). Clearly, this formula would fall short of accounting for any effect of anharmonic lattice dynamics in the Raman spectrum. Particularly relevant for this work is the fact that in this approximation, the dependence of the (normalized) scattered intensity on the polarization angle θ is fully determined by the shape, symmetry and relative values of the Raman tensor components of a given mode, and by the geometry of the experiment. Notably, the thermal populations enter only as a multiplicative factor. As the space group of the crystal and the molecular sites do not change with temperature (assuming no phase transitions), the shape and symmetry of $\mathcal{R}_k$ are fixed. A common interpretation of this observation is that besides a change in its absolute values, the angular dependence of the intensity calculated with Eq. (7) will be temperature independent in the harmonic picture. As we will see in this paper, though, keeping to Eq. (7) but including temperature-dependent lattice changes can lead to modifications in sign and relative magnitude of certain components of $\mathcal{R}_k$, which in turn lead to pronounced variations in the PO pattern that are not connected to dynamical anharmonic effects.

Beyond this, any further change with temperature of the PO-patterns must then be investigated taking into account the dynamical anharmonicity of molecular vibrations, which presents an added challenge to simulations. In particular, dynamical methods which are not based on perturbation theory for the nuclear degrees of freedom provide an attractive platform to include these effects in their totality. For practical purposes, it is convenient to switch from the energy representation of Eq. (1) to the Heisenberg representation in the time domain, resulting in

$$I(\omega,\beta,\theta) \propto \int_{-\infty}^{+\infty} dt\ e^{i\omega t}\ \langle\alpha(\theta,t)\alpha(\theta,0)\rangle_\beta, \tag{8}$$

where we have conveniently defined

$$\alpha(\theta) := \boldsymbol{e}_1\cdot\boldsymbol{\alpha}\cdot\boldsymbol{e}_2, \tag{9}$$

$\langle\cdot\rangle_\beta$ denotes a quantum thermal average, ω is frequency, and t is time. Note that, within the BO approximation, the time dependence of the polarizability is given by the time dependence of the nuclear positions. The expression of Eq. (8) can be rewritten as sum over component-wise cross-sections,

$$I(\omega,\beta,\theta) \propto \sum_{rsqp} e_1^r e_1^s I_{rsqp}(\omega,\beta) e_2^q e_2^p, \tag{10}$$

highlighting the fourth-rank tensorial nature of $I_{rsqp}(\omega,\beta)$, where r, s, q, p run over Cartesian components, as recently stressed in Refs. [10,25].

The time-correlation function appearing in Eq. (8) can be well approximated in many situations by considering a classical time-evolution for the nuclei, especially at higher temperatures or when focusing on low-energy vibrational modes. A full quantum-mechanical evaluation of these correlation functions for large and complex systems is too cumbersome with available methodologies, which require the use of explicit wave-function methods or real-time path integrals [28,29]. However, these expressions are also amenable to approximate methods that join quantum-statistics with classical nuclear dynamics, based on ring-polymer molecular dynamics [30–34] and centroid molecular dynamics [35–38]. These methods provide a way to incorporate zero-point-energy and nuclear delocalization effects in the time-correlation functions, but disregard dynamical quantum coherence of the nuclear wave packets. When coherence effects are small, for example at moderate to high temperatures and in the condensed phase [39], these methods become suitable for including nuclear quantum effects in molecular crystals.

III. SIMULATING PO-RAMAN SIGNALS

In this work, we focus on capturing, through first-principles simulations, anharmonic effects in the PO-Raman signal coming from low-frequency vibrational modes on polyacene molecular crystals. Going beyond the harmonic approximation without relying on perturbative expansions amounts to evaluating Eq. (8), where the time-dependence of the polarizability tensor through the nuclear coordinates is commonly taken into account by means of AIMD simulations.

Achieving first-principles accuracy in this evaluation relies on being able to obtain $O(10^6)$ calculations of forces and

polarizability tensors, through the realization of AIMD trajectories and perturbation theory calculations, resulting in a very high computational cost [40]. Previous studies [1,41,42] have used AIMD to successfully compute anharmonic unpolarized or powder Raman spectra of molecular crystals, albeit computing very few trajectories of the order of 10 ps, resulting on a high uncertainty in peak intensities. In the present study, converged peak intensities are necessary to capture subtle variations in polarization-orientation patterns at different temperatures.

A practical and efficient solution is achieved by evaluating the quantities needed in Eq. (8) with highly accurate ML models. In this paper we use two models: one model targets the prediction of energies and forces, needed for the evaluation of normal modes or nuclear time propagation, and the other model targets the prediction of polarizability tensors, needed for the evaluation of the Raman cross-sections. When combining ML models to calculate a particular physical observable, one needs to take into account that errors from both models will compound in the calculation. This calls for a stringent evaluation of the accuracy and performance of the framework.

Because the training and usage of machine-learning interatomic potentials (MLIPs) is nowadays common in the literature, we only cover the basic aspects of the models used in this work in Methods Sec. VI A. Below, we go into more detail on the models used for the polarizability tensors.

A. Machine learning models for polarizability tensors

The polarizability tensor is a second-rank tensor with an isotropic component that transforms as spherical harmonics of $l = 0$ and an anisotropic component that transforms as spherical harmonics of $l = 2$. As such, symmetry-adapted equivariant ML methods can deliver a noticeable boost in efficiency of training. All models we discuss in this paper account for the symmetry of the tensor in the learning problem, being thus equivariant. We train two such models here, which are based on an atomic decomposition of this quantity, i.e.,

$$\boldsymbol{\alpha} = \bar{\alpha}\mathbf{1} + \tilde{\boldsymbol{\alpha}} = \sum_I \left[\boldsymbol{\alpha}^{(l=0)}(\mathcal{A}_I) + \boldsymbol{\alpha}^{(l=2)}(\mathcal{A}_I)\right], \tag{11}$$

where $\bar{\alpha}$ is the isotropic, traceless part of the polarizability tensor, $\tilde{\boldsymbol{\alpha}}$ the anisotropic one and I runs over all atoms that define the center of the atomic environments $\mathcal{A}$.

The first method we employ to learn this tensor is a MACE model [43,44], which is an equivariant message-passing graph neural-network (NN). Due to its equivariant formulation, it can be seamlessly extended to predict vectorial and tensorial properties [45,46], allowing one to set the tensorial rank of the prediction and the symmetries of its components. We will refer to this model as MACE-α.

The other method is the symmetry-adapted Gaussian process regression (SA-GPR) proposed in Ref. [47]. As the name proposes, this method is not based on a NN architecture, but instead on a Bayesian Gaussian process regression [48]. This technique casts the problem of learning the target quantity as a problem of finding the coefficients of a linear expansion of the target quantity on a combination of similarity-kernel functions [49]. It is possible to define different kernels corresponding to the different rotational symmetries of each component of $\boldsymbol{\alpha}$, leading to an equivariant "symmetry-adapted" version of the learning problem. We will refer to this model as SA-GPR-α.

Both methods mentioned above have already been employed in the context of predicting polarizability tensors that were then used to calculate standard vibrational Raman spectra [11,12,14,45,46,50]. Here we instead assess them for the more stringent test of predicting the PO-pattern of each Raman peak.

B. RGDOS approximation to polarizability tensors

The RGDOS approximation to the polarizability tensor, as introduced in Ref. [15] and extended for molecular dynamics in Ref. [16], is an efficient approach to reconstruct such tensors of a large supercell as a weighted sum of polarizabilities of the multiple unit cells that constitute it. In Ref. [12], Berger *et al.* showed the good performance of this approximation in reproducing DFT-quality polarizabilities for large supercells and yielding good Raman spectra, including extensions to second-order and resonant contributions. The method relies on the existence of a unique equilibrium reference structure and its corresponding eigenmodes, being therefore unable to describe different phases of a material within a single model, but performing well when no phase transitions occur.

Following Ref. [12], one expands the polarizabiliy of the supercell as

$$\boldsymbol{\alpha}(\boldsymbol{u}) \approx \boldsymbol{\alpha}(\boldsymbol{u}_0) + \sum_{n,\nu,\boldsymbol{q}} \frac{1}{n!} \mathcal{R}^{(n)}_{\nu,\boldsymbol{q}} P^n_{\nu,\boldsymbol{q}}(\boldsymbol{u}), \tag{12}$$

where $\boldsymbol{u}$ is an atomic displacement vector in the supercell, n is the order of the expansion term, ν is the primitive-cell phonon branch index and $\boldsymbol{q}$ is the phonon wavevector, $\mathcal{R}^{(n)}_{\nu,\boldsymbol{q}}$ is the corresponding nth order Raman tensor and $P_{\nu,\boldsymbol{q}}(\boldsymbol{u})$ is the projection coefficient of the supercell displacement onto the primitive-cell eigenmode, as defined in Sec. S1 of the SM [26]. The most important approximation involved in the expression above is that nth order cross-derivative terms involving different phonon modes are disregarded, while they would be present in an exact expansion of the polarizability tensor for anharmonic systems.

The flexibility of the method stems from the fact that the expansion can be truncated at any desired order. If one is interested only in first-order Raman scattering processes, then the task is considerably simplified. In addition to keeping only the first order in the expansion of Eq. (12), only modes at the vicinity of the Γ point of the Brillouin zone contribute, due to photon momentum conservation. This is well approximated by keeping only contributions from Γ-point phonon modes, i.e.,

$$\boldsymbol{\alpha}(t) = \boldsymbol{\alpha}(\boldsymbol{u}(t)) \approx \boldsymbol{\alpha}(\boldsymbol{u}_0) + \sum_{\nu} \mathcal{R}_{\nu,\Gamma} P_{\nu,\Gamma}(t), \tag{13}$$

which is the RGDOS expression for the polarizability of the supercell. Once a phonon calculation is performed with the primitive cell and the corresponding Raman tensors are obtained, the Γ-eigenmodes of the supercell are reconstructed by tiling, i.e., concatenating primitive-cell eigenmodes up to the supercell dimensionality. The projection coefficients are then computed at every time-step to yield the polarizability time series for the supercell according to Eq. (13). In the SM [26] we provide more details about how we obtain these

projection coefficients in a nonorthogonal basis. Note that a different calculation (phonons + Raman tensors) has to be performed at every temperature if lattice thermal expansion is taken into account, as the equilibrium reference structure will be affected by the change in lattice vectors. The RGDOS PO-Raman spectrum is then obtained by plugging Eq. (13) into Eq. (8), yielding

$$I^{\mathrm{RGDOS}}(\theta) \propto \sum_{\nu\mu} (\boldsymbol{e}_1 \cdot \mathcal{R}_{\nu,\Gamma} \cdot \boldsymbol{e}_2)(\boldsymbol{e}_1 \cdot \mathcal{R}_{\mu,\Gamma} \cdot \boldsymbol{e}_2) \times \int_{-\infty}^{\infty} dt \, \langle P_{\nu,\Gamma}(0) P_{\mu,\Gamma}(t) \rangle_\beta \, e^{i\omega t}, \quad (14)$$

where anharmonic effects are included in the correlation function of the projected supercell displacements. A further possible simplification regards the truncation of Eq. (13) with respect to the number of Γ phonons considered. The evaluation of Eq. (14) then requires fewer Raman tensors and projection coefficients. We demonstrate in the SM [26] (Sec. S2) how retaining only the intermolecular modes in the expansion has no visible effect in the resulting low-frequency spectrum of the polyacene crystals studied here, making this method particularly attractive for the investigation of low THz regions of molecular crystals.

C. Proposed framework

We assessed the quality of the MLIPs we trained for the frequencies of the crystal vibrations. A particularly challenging region, where we must guarantee that the errors are acceptable, is the low-THz part of the spectrum, which is determined by weak intermolecular interactions often poorly described in MLIPs that are inherently local. As detailed in Sec. VI A, we compared a second-generation Behler-Parrinello high-dimensional neural network potential [51], which we labeled high-dimensional neural-network potentials (HDNNP)-MLIP, and an equivariant graph neural-network MACE [43] interatomic potential, which we labeled MACE-MLIP.

Our benchmarks, detailed in Sec. S3 of the SM [26], show that MACE-MLIP outperforms the HDNNP-MLIP in the prediction of Γ-point harmonic phonon frequencies of the anthracene crystal, resulting in a lower average percentage error of $\approx 0.2\%$ against the $\approx 1.6\%$ obtained from HDNNP-MLIP. Moreover, MACE-MLIP exhibits a significantly tighter distribution of errors, with a low median and minimal variability, highlighting the consistency and accuracy of this model. In contrast, HDNNP-MLIP shows a broader error distribution, indicating less reliable predictions. When focusing specifically on the low THz region, the error on the harmonic frequency of Raman active modes is almost negligible for the MACE-MLIP model, while the HDNNP-MLIP largely underestimates the harmonic frequency of the three highest frequency librational modes, with errors up to 20%. These conclusions are consistent with the broader benchmark on the accuracy of vibrational properties obtained from MLIPs that we performed in Ref. [13], in which the reasons for the superior performance of MACE-MLIPs for vibrational frequencies of molecular crystals are discussed.

Next, we assessed the quality of our ML models for the prediction of the polarizability tensors. Our benchmarks, detailed in Sec. S3 of the SM [26], show that MACE-α predictions are almost a factor two more accurate than the SA-GPR predictions for diagonal and off-diagonal components of the tensor. We conclude that MACE is the most suitable solution for achieving higher prediction accuracy, especially when applied to larger datasets and system sizes.

Lastly, we analyzed the differences between (i) evaluating Eq. (8) with the MACE-MLIP potential for performing the nuclear time evolution combined with the MACE-α model for the polarizability evaluations and (ii) evaluating the RGDOS approximation in Eq. (14) with the same MACE-MLIP potential for the nuclear dynamics and fixed Raman tensors computed directly from DFPT. In the following, we name the first procedure "full-ML" and the second "RGDOS-ML."

In all the calculations performed in this work, the incident and scattered light are polarized in the xy plane and propagate along the z direction in the laboratory frame. This means that only the x and y components of the response are relevant. The PO-Raman spectra we show in the following figures are computed in the parallel scattering configuration, as defined in Sec. II.

In Figs. 1(a) and 1(b), we show the PO-Raman spectra of the anthracene crystal at 100 K in the range of 0 to 800 cm^{-1}, obtained with full-ML and RGDOS-ML, respectively. All the spectra were computed from $4 \times 4 \times 4$ supercell trajectories, which amounts to more than 3000 atoms in the case of anthracene. More details are given in Sec. VI. The overall visual comparison makes it clear that the differences between these two methods are almost imperceptible in this frequency range, with both yielding very similar polarization dependence of all peaks, and following a very similar relative-intensity pattern when regarding different peaks.

The small discrepancies that arise are shown in more detail in Fig. 1(c), where the difference between the PO-Raman spectra of anthracene at 100K obtained with each procedure is shown. Negative values correspond to features that are more intense in full-ML than in RGDOS-ML and vice-versa. In Fig. 1(d) we report the same difference spectrum, this time for the unpolarized Raman spectrum, as defined in Eq. (2). In general, RGDOS-ML demonstrates remarkable accuracy for PO-Raman analysis in the low-THz region of interest, with discrepancies in the order of ∼5–10% of the maximum amplitude. The larger discrepancy is found in the $B_g^{(1)}$ mode of anthracene located below 100 cm^{-1}, where RGDOS-ML overestimates the unpolarized peak amplitude $\sim$ 12%. Consequently, as our focus is limited to first-order Raman—which contains contributions almost solely from the vicinity of Γ-point phonons—we opted to proceed with the RGDOS-ML framework due to its efficiency and additional advantage of removing uncertainties in the prediction of polarizability tensors.

IV. RESULTS AND DISCUSSION

A. Quasiharmonic approximation to PO-Raman

As a first step toward the analysis of the impact of temperature on the polarization dependence of the Raman intensity, we compute the quasiharmonic PO-dependence

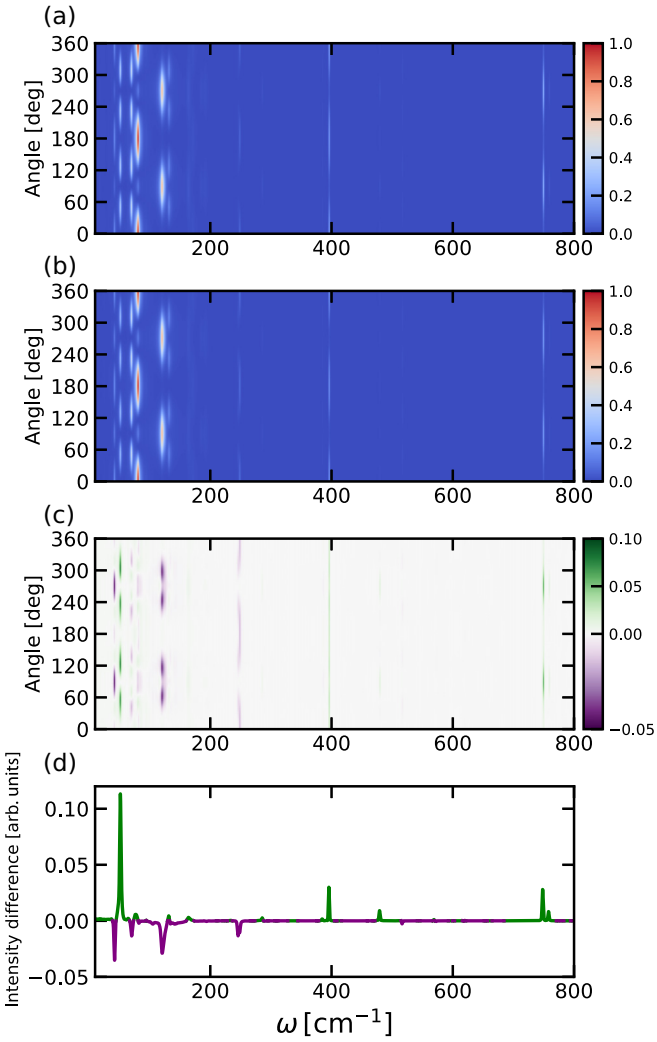


FIG. 1. (a) PO-Raman spectrum of anthracene at 100K obtained with the "full-ML" framework (4x4x4 supercell). (b) PO-Raman spectrum of anthracene at 100K obtained with the RGDOS-ML framework (4 × 4 × 4 supercell). Intensities are normalized to the highest peak amplitude in both cases. (c) Difference between the PO-Raman spectra obtained with the two frameworks. Negative values correspond to features that are more intense in full-ML than in RGDOS-ML and vice versa. (d) Difference between the unpolarized Raman spectra obtained with the two frameworks.

(parallel configuration) of the Raman intensity in anthracene and naphthalene using only the harmonic approximation of Eq. (7). This calculation is purely harmonic and nondynamical: the Raman tensors are obtained directly through DFPT at the optimized geometries corresponding to different lattice parameters, and no MD simulations are involved. Thermal lattice expansion is taken into account at 100K and 295K by using experimentally determined lattice constants. In this subsection, when we refer to different temperatures, we are simply referring to the corresponding different lattice constants. This approach is commonly referred to as quasiharmonic approximation.

Both materials studied here crystallize in the same space group, $P2_1/a$, with two molecules per primitive cell occupying the same sites in both crystals. Factor group analysis thus leads to the same symmetries for their Raman active modes. Namely, the A_g and B_g modes are respectively the three in-phase and the three out-of-phase librational modes of the two molecular units in the primitive cell, and are Raman active [52,53]. From the three out-of-phase translational modes, two bear A_u and one bears B_u symmetry, being instead IR active. As already mentioned, the crystal orientation in our simulations is along the crystallographic *ab* plane (001), meaning only the *xy* components of the Raman cross-section will be measured.

Within the quasiharmonic picture, the polarization pattern, i.e., the angular dependence of the Raman intensity for a given mode k in Eq. (7), is fundamentally determined by the shape, symmetry, and the relative magnitude and sign of the components of the corresponding $\mathcal{R}_k$. In Sec. S4 of the SM [26] we provide explicit expressions for the general angular dependencies of the Raman intensity patterns generated by A_g and B_g tensors. In the case of naphthalene and anthracene, with the crystals oriented in the crystallographic *ab* plane (001), we show how, once the intensity is normalized at each temperature, no variations with lattice expansion can occur in the polarization patterns of B_g modes, both in the parallel and perpendicular configuration, within the harmonic approximation of Eq. (7). The same is true for A_g modes in the perpendicular configuration. In contrast, we show how A_g modes in the parallel configuration are affected by changes in the relative value of their Raman tensor components, which can be induced by thermal lattice expansion.

We report in Fig. 2 the $A_g^{(1)}$ and $B_g^{(1)}$ modes of both crystals as interesting example cases. For completeness, we show the patterns for all modes in parallel and perpendicular configurations in Sec. S4 of the SM [26]. We start by discussing the PO-patterns of the anthracene crystal. All B_g modes show a fourfold oscillating pattern across polarization angles, as exemplified in Fig. 2(a), and the A_g modes show more variability (see also Fig. S4 of the SM [26]). Specifically, $A_g^{(1)}$ and $A_g^{(2)}$ have *xx* and *yy* tensor components of opposite sign, resulting in a fourfold pattern with two maxima of different intensities, whereas $A_g^{(3)}$ displays a twofold oscillation due to having these diagonal elements of the same sign. This is an important observation, as despite the fact that a square modulus appears in Eq. (7), the relative sign of the Raman tensor components is qualitatively relevant in the resulting PO-dependence of A_g modes, and can therefore be extracted from a fit of the experimental data. Variations of the PO-patterns at different temperatures are minimal. As a matter of fact, we only observe a change in the relative values of the diagonal components of the $A_g^{(1)}$ mode, resulting in an intensity variation of the relative maximum at $\theta = 270^\circ$ in Fig. 2(a).

In the case of naphthalene, the same analysis holds true for the B_g modes, which show a fourfold oscillating pattern, as expected from the fact that these crystals belong to the same space group. The A_g modes again show either a twofold or a fourfold pattern with maxima of different intensities, depending on the sign of the diagonal components of their Raman tensors (see Fig. S5 of the SM [26]). For example, the $A_g^{(2)}$ mode of naphthalene has a weak PO dependence in the parallel configuration, while having zero intensity in the perpendicular one. An inspection of its Raman tensor at 100K explains this, as the *xx* and *yy* components are almost

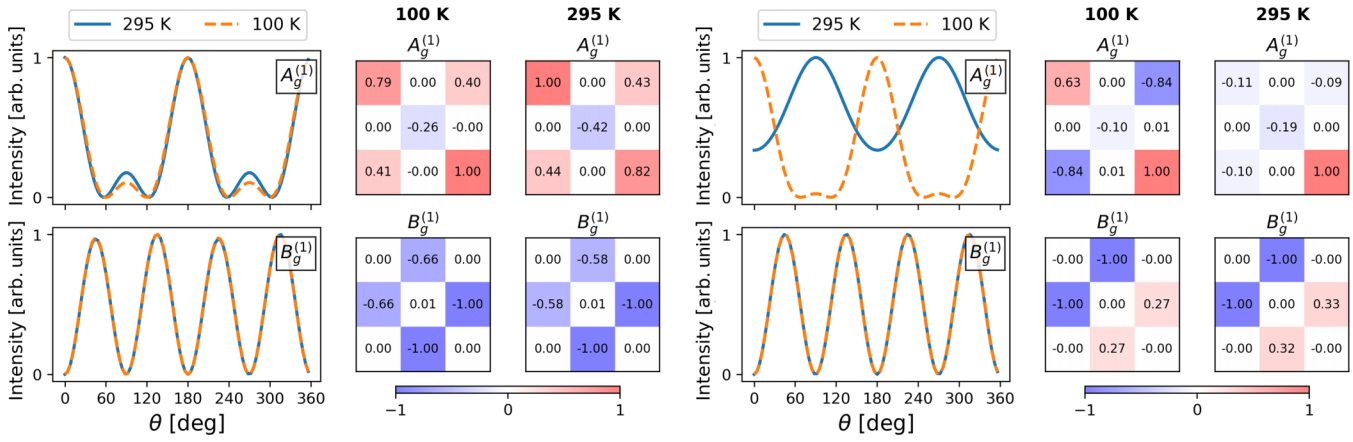


FIG. 2. Harmonic polarization dependence of the Raman intensity of modes $A_g^{(1)}$ and $B_g^{(1)}$ in anthracene (a) and naphthalene (b), obtained with the harmonic approximation of Eq. (7) in the parallel configuration. On the right, the corresponding Raman tensors for every mode computed with finite differences from DFPT polarizabilities along normal mode displacements of the MACE-MLIP at the 100K and 295K lattice vectors. Tensor components are normalized to their highest value for each mode. In the parallel configuration, only A_g modes can show a change in the polarization pattern with temperature (lattice expansion) within the quasiharmonic approximation. This effect is particularly pronounced in the case of naphthalene, where structural changes impact the $A_g^{(1)}$ mode. The full set of harmonic polarization patterns and a more detailed analysis can be found in Sec. S4 of the SM [26].

equal in magnitude and sign, acting like an identity matrix in Eq. (7).

Quite surprisingly, though, in naphthalene we observe pronounced changes in the overall polarization patterns of A_g modes with thermal lattice expansion. This is particularly evident in the case of mode $A_g^{(1)}$, where a significant change in the relative values of the diagonal components of the Raman tensor leads to a transition from a fourfold pattern with two maxima of different intensities at 100K to a twofold pattern at 295K [see Fig. 2(b)]. This observation highlights that even within the quasiharmonic approximation, temperature-induced lattice expansion can lead to substantial modifications in the polarization dependence of certain Raman active modes. In Sec. S4 of the SM [26], by comparing the normal modes displacement vectors at the two different temperatures, we show that the $A_g^{(1)}$ and $A_g^{(2)}$ modes of naphthalene present a measurable change in direction with the different lattice parameters, leading to significant variations in their Raman tensor components and consequently in their polarization patterns. This is not the case for anthracene, where the normal modes obtained for the structures at different temperatures overlap almost perfectly with each other.

As a further note, we demonstrate in Sec. S5 of the SM [26], that even small inaccuracies in the normal modes can lead to Raman tensors that yield incorrect polarization dependencies. These changes may be overlooked when the θ variable is integrated out, underlying again how PO-patterns are a much stricter test for the quality of the theoretical framework, compared to unpolarized Raman spectra.

In summary, once the experimental configuration is fixed (either $\boldsymbol{e}_1 \perp \boldsymbol{e}_2$ or $\boldsymbol{e}_1 \parallel \boldsymbol{e}_2$), only the PO-dependence of modes of a certain symmetry can be affected by thermal lattice expansion in the quasiharmonic approximation. In the case of anthracene and naphthalene, these are the A_g modes in the parallel configuration. The extent to which the PO-pattern of these modes is affected varies significantly between different materials, as exemplified by the comparison between anthracene and naphthalene. While we prove that not all temperature-dependent changes in PO-Raman patterns will stem from dynamical coupling between different vibrational modes, we also provide a clear baseline for the following analysis of temperature-dependent spectra. It is important to note that the changes we capture within the quasiharmonic approximation in Fig. 2, although sometimes pronounced, are still encoded in the form of second-rank Raman tensors of fixed symmetry. The quasiharmonic results thus establish a baseline for interpreting temperature effects that arise purely from lattice expansion. Any further changes in the polarization patterns beyond this baseline, as observed in experiment or in the dynamical simulations we present next, must originate from anharmonic dynamical effects involving the vibrational degrees of freedom, where the full, fourth-rank nature of the Raman response function [Eq. (10)] is revealed.

B. Temperature dependence of PO-Raman signals

We now turn our attention to the investigation of the temperature dependence of the PO-Raman spectra of anthracene and naphthalene crystals. In Fig. 3 we report the temperature evolution of the parallel configuration anharmonic PO-Raman spectra in the intermolecular motion range of 10 to 170 cm^{-1}, obtained with the RGDOS-ML method. The perpendicular configuration PO-spectra can be found in Sec. S6 of the SM [26] for completeness. Computational details on these production simulations are given in Sec. VI E.

In Figs. 3(c)–3(f), a clear redshift and broadening of all peaks is observed as the temperature increases in both crystals. The extracted peak positions are listed in Table I. The shift is similar across different modes, spanning a range of ∼3–9 cm^{-1}. This effect is a direct result of the incorporation of anharmonicity, in particular the temperature-dependent lattice expansion. As shown in Sec. S7 of the SM [26] for

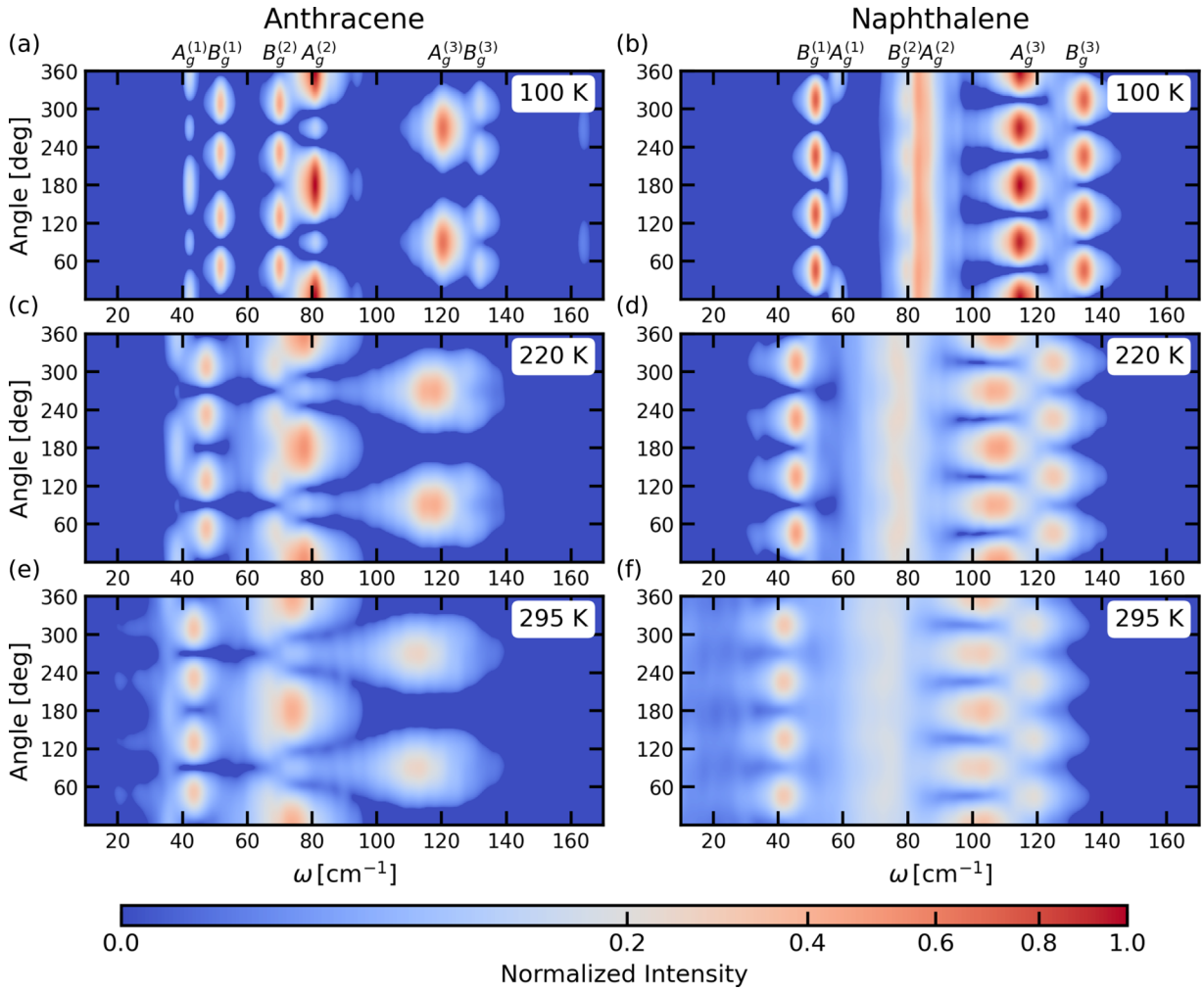


FIG. 3. Temperature dependence of the parallel configuration PO-Raman spectra of anthracene (a), (c), (e) and naphthalene (c), (d), (f) in the intermolecular motion range of 10 to 170 cm^{-1}, obtained with the RGDOS-ML method. The intensity is normalized to the maximum intensity of the 100K spectrum.

anthracene, the redshift is almost negligible when lattice expansion is not taken into account, despite the anharmonic dynamics. This highlights once more the strong influence of lattice parameter variations on the dynamics of low-frequency intermolecular modes. Notably, some modes even exhibit a blue shift in the absence of lattice expansion. Beyond these classical anharmonic effects, we also assessed the role of nuclear quantum effects (NQEs) on the PO-patterns. As shown in Sec. S8 of the SM [26] (see additional Refs. [54–57] therein), based on path-integral molecular dynamics simulations of the PO-Raman patterns of anthracene at 100K and 295K [38,58,59], NQEs are found to have a negligible impact on the polarization patterns of these modes below 400 cm^{-1}, confirming that classical-nuclei simulations are very accurate at the frequency range considered here.

In the following subsections, we will compare our simulated spectra with available experimental data and analyze in detail the signatures of anharmonic mode mixing and peak broadening that emerge at higher temperatures.

TABLE I. Evolution of the peak frequencies with temperature in anthracene and naphthalene from the PO-maps of Fig. 3. Peak positions are determined from the fitting procedure detailed in Sec. VI. For naphthalene, $B_g^{(2)}$ and $A_g^{(2)}$ peak positions were extracted at 220K and 295K from the VDOS decomposition of Fig. 5 due to their broadness and overlap.

Anthracene [cm^{-1}]				Naphthalene [cm^{-1}]			
Mode	100K	220K	295K	Mode	100K	220K	295K
$A_g^{(1)}$	42	38	36	$B_g^{(1)}$	51	45	41
$B_g^{(1)}$	52	47	44	$A_g^{(1)}$	58	50	48
$B_g^{(2)}$	70	68	67	$B_g^{(2)}$	79	75	69
$A_g^{(2)}$	81	77	74	$A_g^{(2)}$	84	79	72
$A_g^{(3)}$	121	117	112	$A_g^{(3)}$	115	108	103
$B_g^{(3)}$	132	128	124	$B_g^{(3)}$	134	125	119

1. Contrasting simulations with experimental data

The first observations of an anomaly in the polarization dependence of the Raman peaks in anthracene and pentacene were made by Asher *et al.* in 2020 [8]. For example, the

polarization dependence of some of the low-frequency Raman peaks of anthracene was found to gradually become less pronounced with increasing temperature. These effects were later reported for other classes of organic molecular crystals by the same authors, and extensive work has been performed in altering (quenching) this behavior through chemical substitutions [9].

This phenomenon, which was described as a dynamical symmetry breaking or mode-specific "liquidlike" transition, has been recently attributed by Benshalom *et al.* [10] to the manifestation of temperature activated, anharmonic coupling between the intermolecular modes. In their work, the authors tackled the problem from an analytical point of view. An expression for the Raman intensity similar to Eq. (8) was evaluated with fixed Raman tensors in a two-mode model, allowing for anharmonic coupling between modes of same and different symmetries. By relying on fourth-rank tensors for the Raman cross-sections like the ones in Eq. (10), this approach was successful in fitting the experimental PO-dependence of the Raman intensity when the harmonic expression of Eq. (7) failed. These results led the authors to conclude that the observed variation in the PO-dependence of the Raman peaks is a consequence of the anharmonic coupling between intermolecular modes, activated by temperature. Note that the harmonic approximation leading to Eq. (7), as explained in Sec. II, can only account for deviations due to thermal lattice expansion that conserve the symmetry of the Raman tensors. Wherever this approach fails, this can be regarded as a sign of a dynamical anharmonic effect.

From our point of view, it is instructive to understand whether a first-principles procedure without fitting parameters can capture and provide deeper insight into the observed changes in PO-Raman intensities with increasing temperature. In this study, we are considering the full expression of the anharmonic intensity [i.e., Eq. (8)], thus retaining the generalization to the fourth-rank tensor of the scattering cross-section and the complete mode coupling given by the first-principles full-dimensional anharmonic PES. The expression Eq. (14) has the same form as the starting point of the analysis of Benshalom *et al.* [10], containing all the cross-correlations terms between phonon modes, i.e., the off-diagonal components of the phonon spectral function, albeit only at the Γ point. Moreover, our simulations give us the chance to directly compare harmonic and anharmonic results, with and without thermal lattice expansion, providing a tool to decouple and unravel the different effects that temperature has on the resulting Raman spectra.

Before focusing on the angle dependence of the integrated intensity, we highlight some general discrepancies between our simulated and the experimental spectra of anthracene. We observe good overall agreement with the experimental peak positions as reported in Fig. S7 of the SM of Ref. [8], with redshifts in the order of $\approx$ 5–10 cm^{-1} in the simulated spectra. Notably, the $B_g^{(2)}$ mode has negligible Raman intensity in the experiments, where no polarization dependence could be extracted. In contrast, the mode shows appreciable intensity in our results. This discrepancy is attributed to the level of theory employed, and is not affected by the subsequent dynamics. In fact, it is known to be already present when calculating a harmonic unpolarized Raman spectrum with the same level of theory (PBE+MBD), as reported in Ref. [8]. In Sec. S13 of the SM [26], we further show that using different functionals to calculate the crystal vibrations shift the position of the peaks in a nonuniform fashion in this frequency range, but do not alter the relative peak intensities.

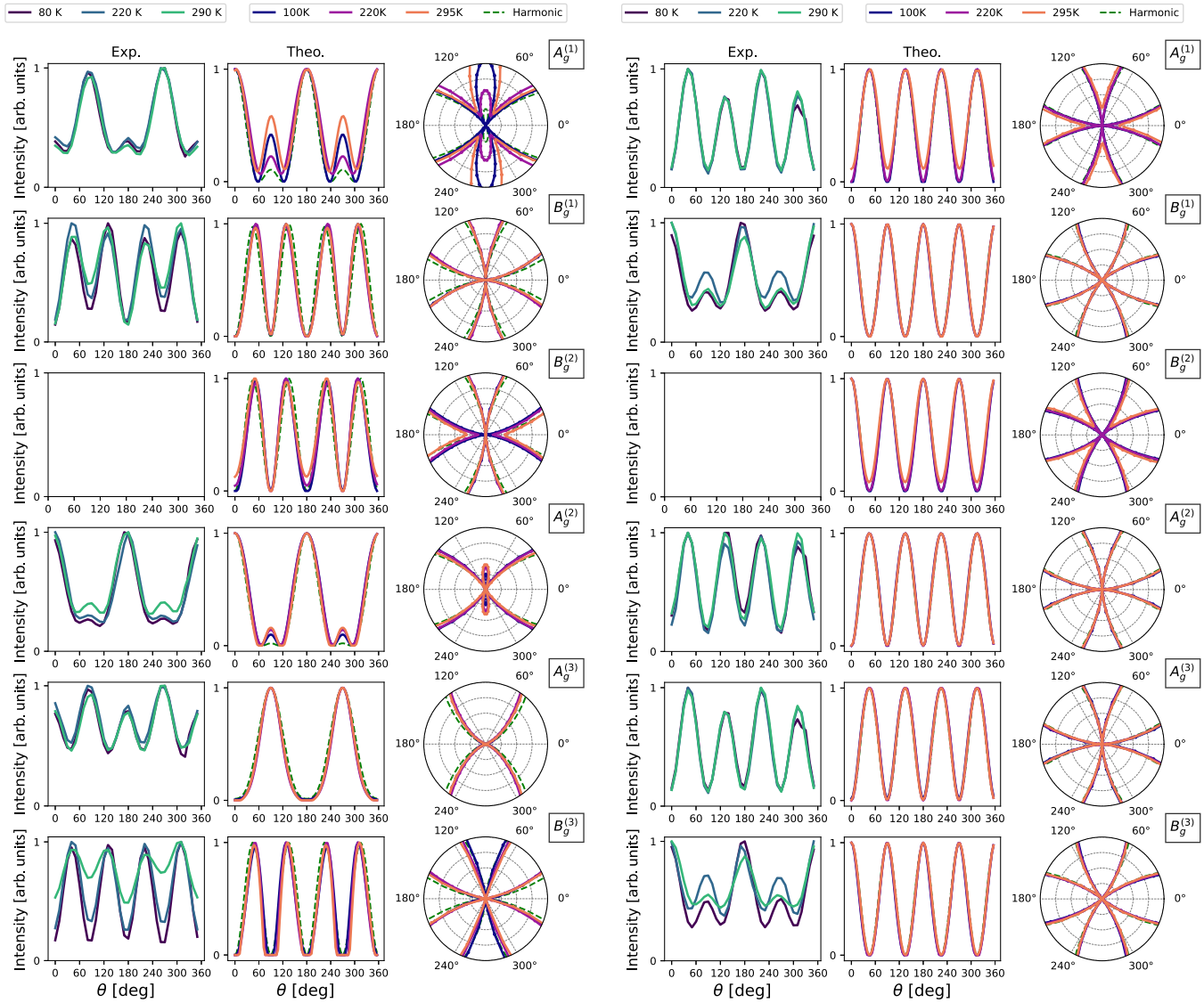
In this work, we attempt to extract the PO-dependence of the integrated intensity of every peak from the anharmonic PO-maps of Fig. 3 by fitting the Raman spectra of anthracene at 100K, 220K, and 295K to a multi-Lorentzian function at each angle. The fitting procedure mimics the one employed for the experiments, as detailed in the SM of Ref. [8], and reported in Sec. VI.

In Fig. 4 we show the fitted PO-dependence of the integrated intensity of the low-frequency Raman peaks of anthracene in both the parallel and perpendicular experimental configurations. The fitted data is normalized to the maximum intensity of each peak at the corresponding temperature. For comparison, we also include the harmonic PO-dependence and the experimental results from Ref. [8], at the closest available temperatures of 80K, 220K, and 290K. We remind the reader that for anthracene, the harmonic PO-pattern is insensitive to lattice expansion. In Fig. S13 of the SM [26], we show the result of attempting the same fit to the simulated spectra of naphthalene, which fails at higher temperatures due to excessive broadening of the peaks.

The fitted intensity oscillation patterns from the simulated data in the parallel configuration shown in Fig. 4(a) are in general agreement with the experimental results, with the notable exception of the $A_g^{(3)}$ mode. While both theory and experiment agree that the temperature variation has no impact on the PO-dependence of this mode, its polarization pattern differs. In experiment, it shows a fourfold pattern, with two relative maxima. In the simulations, only a twofold pattern is present. This discrepancy is entirely determined by the form of the Raman tensor predicted by the specific level of theory employed, and cannot attributed to the subsequent dynamics. Specifically, our PBE $A_g^{(3)}$ Raman tensor has xx and yy components of the same sign, which cannot give rise to a fourfold pattern. One must note that experimental artifacts, such as birefringence, can alter the polarization dependence of the Raman intensity in experiment, making the direct comparison even more challenging. Artificially including birefringence in the form of Jones matrices [60] can alter the simulated PO-patterns of A_g modes, but not to a sufficient extent to explain the observed discrepancies (see Sec. S9 of the SM [26]). In the perpendicular configuration, all modes present a fourfold pattern in our simulations. This is partially consistent with experiments, where indeed every mode appears as fourfold, but with B_g having two relative maxima of different intensities.

In our simulations, the $A_g^{(1)}$ mode shows variations in the PO pattern with temperature, although we attribute a portion of this variation to the change in the relative magnitudes of the diagonal components of the Raman tensor, discussed in Sec. IV A and shown in Fig. 2.

The experimental results in Fig. 4 show a clear change of the PO-dependence with temperature of the $B_g^{(1)}$ and $B_g^{(3)}$ modes in the parallel configuration. The largest change is seen in going from 220K to 290K. A similar effect is present for the $A_g^{(2)}$ mode in experiment, but less pronounced. In all cases, the effect can be described as a gradual loss of the



(a) Parallel configuration

(b) Perpendicular configuration

FIG. 4. Polarization dependence of the integrated intensity of low THz Raman peaks in anthracene for (a) Parallel configuration and (b) perpendicular configuration. Left columns are experimental results reproduced with permission from Ref. [8] at 80K, 220K, and 290K. Center and right columns fitted with Eq. (15) from the simulated PO-maps of Fig. 3 (parallel) and Fig. S8 of the SM [26] (perpendicular) obtained with the RGDOS-ML framework at 100K, 220K, and 295K. Intensity is normalized with respect to the highest intensity of every peak at the corresponding temperature. For comparison, the harmonic PO dependence computed with Eq. (7) is also shown. Polar plots zoomed in the region of intensities below 0.4, to highlight small differences arising with varying temperature.

oscillation pattern with increasing temperature. When comparing to our simulated, anharmonic PO-maps, we do not find a one-to-one, quantitative correspondence of this phenomenon. In some cases the trend is reproduced correctly, such as the $A_g^{(2)}$ mode, where the oscillation pattern is indeed less pronounced at higher temperatures in both experiment and theory in the parallel configuration. In the case of the $B_g^{(1)}$ and $B_g^{(3)}$ modes, which show a strong loss of polarization dependence in experiment, our simulated pattern is not affected by temperature. Nonetheless, we observe variations of the PO pattern in the $B_g^{(2)}$ mode, with a consistent trend of gradual attenuation of the oscillation pattern with increasing temperature. In the perpendicular configuration of Fig. 4(b), $B_g^{(2)}$ still shows subtle deviations in the simulated patterns. In contrast, the experimental data is substantially less clear regarding the temperature dependence of the PO patterns when the perpendicular geometry is employed. In this geometry, changes are still observed in experiment for $B_g^{(1)}$ and $B_g^{(3)}$, but no monotonic trend can be identified with the increasing temperature. We proceed with a deeper theoretical analysis of these observations.

2. Impact of mode mixing and mode broadening on anharmonic spectral weights

To address the role of anharmonicity in the vibrational dynamics of anthracene, we analyze the vibrational density of states (VDOS) and the cross-VDOS (CVDOS) computed in

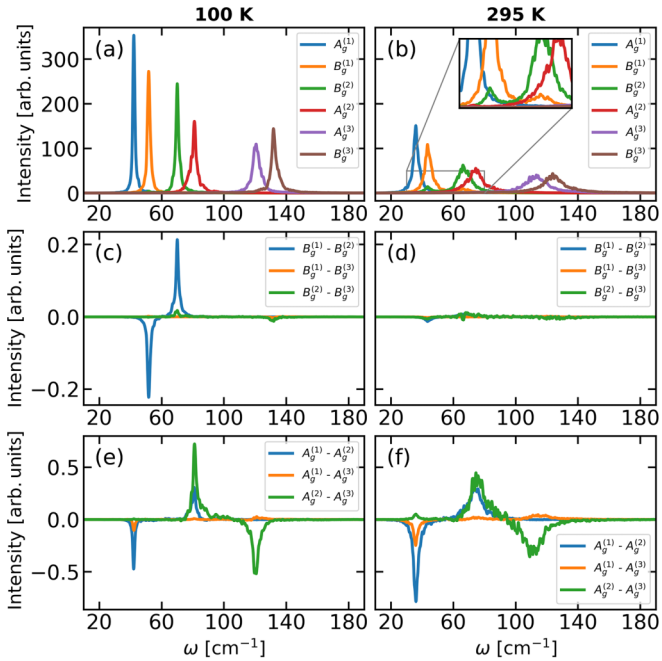


FIG. 5. Vibrational density of states (VDOS) and cross-VDOS (CVDOS) computed in the normal mode basis for the intermolecular Raman active modes of anthracene (Γ point only) at 100K (left column) and 295K (right column). Panels (a) and (b) show the VDOS projected on each mode, while panels (c–f) show the CVDOS between all pairs of B_g and A_g modes. Intensities are not normalized to give a better idea of the relative strength of the cross-correlations. The inset in panel b) zooms in on the spectral mixing between the $B_g^{(1)}$ and $B_g^{(2)}$ modes, as discussed in the text.

the normal mode basis for the Γ-point intermolecular modes. These quantities are defined in Eq. (16) of Sec. VI. The VDOS provides a measure of the distribution of vibrational frequencies, while the CVDOS captures correlations between different modes, revealing signatures of mode coupling.

In Fig. 5 we show the VDOS and CVDOS at the Γ point of anthracene at 100K and 295K. For completeness, we include the same figure for naphthalene in the SM [26] (Sec. S10). Figures 5(a) and 5(b) display the mode-projected VDOS, whereas Figs. 5(c)–5(f) report the CVDOS between all pairs of modes with the same symmetry.

At low temperatures, the VDOS of anthracene shows clear Lorentzian-like peaks for each mode. The corresponding CVDOS shows hints of mode coupling between modes $B_g^{(1)}$-$B_g^{(2)}$, $A_g^{(1)}$-$A_g^{(2)}$, and $A_g^{(2)}$-$A_g^{(3)}$. The CVDOS of modes with different symmetry is zero and not shown. As the temperature increases, we observe the emergence of spectral weight in the VDOS of a given mode stemming from other modes. This is particularly evident for the $B_g^{(1)}$-$B_g^{(2)}$ case at 295 K [inset of Fig. 5(b)] and provides a clear indication of anharmonic mode mixing. Note how in the CVDOS of Fig. 5(d), the cross-correlation intensity between these two modes vanishes going from 100K to 295K. Instead, the VDOS gains spectral weight for the $B_g^{(1)}$ and $B_g^{(2)}$ modes, signaling that the phonon basis is a less accurate representation of the vibrational dynamics at higher temperature. This observation can explain the attenuation of the PO-pattern of the $B_g^{(2)}$ mode observed in the simulated data of Fig. 4. The phenomenon of mode mixing has also been addressed in the analysis of Benshalom *et al.* [10] mentioned at the beginning of the previous section. They showed how, instead of isolated Lorentzian peaks, the spectral functions acquire contributions from multiple modes, resulting in nontrivial intensity redistribution and the possibility of temperature-dependent changes in the PO-patterns at a given frequency.

In addition, the broadening of the peaks in VDOS and CVDOS with increasing temperature, evident in Fig. 5, makes it increasingly difficult to assign spectral features to individual modes when one only has access to the PO-Raman maps such as those shown in Fig. 3. This delicate interplay of overlap and mixing between different modes stresses the need for robust deconvolution procedures when analyzing temperature-dependent PO-Raman spectra, as both true anharmonic coupling and spectral overlap can affect the PO-patterns extracted from the fitting procedure of the global maps of Fig. 3.

This analysis suggests that, while genuine mode coupling like the one that emerges in Fig. 5 can induce real changes in the polarization dependence, the additional overlap of non-Lorentzian, broadened peaks can also lead to apparent intensity sharing between modes, complicating the interpretation of PO-pattern evolution with temperature. Modes like $B_g^{(1)}$ and $B_g^{(3)}$, which show a pronounced change in the PO-pattern in experiment [Fig. 4(b)], both overlap substantially with an intense peak of a different symmetry: $A_g^{(2)}$ and $A_g^{(3)}$, respectively. Especially at polarization angles where one of the two overlapping features is in principle not present, any deviation of the other mode spectral feature from a perfect Lorentzian will lead to intensity sharing, thus potentially hiding real anharmonic coupling effects or giving rise to artificial contributions.

Finally, we note how the RGDOS method explained in Sec. III B amounts to a weighting of the Γ-point VDOS and CVDOS with Raman tensors. This operation adds an intensity modulation that is angle-resolved, and can highlight or suppress mode coupling at different polarization angles. Therefore, one can also perform a mode decomposition of the PO-maps of Fig. 3 in terms of self- and cross-mode contributions at each polarization angle. We present in Sec. S12 of the SM [26] examples of this decomposition for anthracene at selected angles, at both temperatures. While such analysis provides a powerful tool to aid the deconvolution of experimental PO maps, the conclusions drawn from the analysis of the VDOS and CVDOS remain unaltered.

V. CONCLUSIONS

In this work, we have established a computational framework to investigate anharmonic effects in polarization-orientation Raman spectra of molecular crystals, applying it to naphthalene and anthracene. By combining machine learning models for interatomic potentials with the RGDOS approximation, we achieve *ab initio* accuracy for very large supercells and systematically explore temperature-dependent effects. In particular, we showcased how focusing on the polarization dependence of Raman intensities provides a stringent test of the underlying theoretical framework.

Already at the quasiharmonic level, we observe that when lattice expansion affects the nature of the normal modes, as in naphthalene, significant changes in the polarization patterns can occur, which are still rationalized within a simplified, second-rank Raman tensor formalism. On top of this, our simulations capture clear anharmonic signatures, including temperature-driven redshifts and peak broadening, which are strongly influenced by lattice thermal expansion. We also confirm that we capture anharmonic mode mixing and that nuclear quantum effects have negligible impact on PO patterns of low-frequency modes. While the polarization dependence of Raman intensities remains largely consistent with harmonic predictions, subtle deviations emerge, reflecting the interplay between anharmonic lattice dynamics and Raman tensor symmetry, complementing the phenomenological approach of Ref. [10].

Notably, even though our approach captures the anharmonic effects on peak positions and intensities, the simulated PO-patterns show only modest temperature dependence compared to the pronounced changes observed experimentally in Ref. [8]. A careful assessment shows that the extraction of PO intensity variations from the full 2D maps that can be gathered from experiments or simulations can be tricky because of the fitting procedure involved. In particular, the simple multi-Lorentzian fitting becomes increasingly ambiguous as temperature-induced peak broadening leads to intensity overlaps between different spectral features. These factors can obscure genuine dynamical effects and complicate a direct comparison between theory and experiment.

Based on the analysis we have presented in this paper, we can propose a way forward. The ability of the RGDOS-ML framework to decompose the PO-Raman patterns into self- and cross-contributions from different vibrational modes could be used to direclty assign intensities to a given spectral feature in the simulated PO-Raman maps, avoiding the need to fit the entire pattern at once. While in this manuscript we chose to compare on equal grounds the global fit of both the simulation and experiment, this alternative first-principles decomposition could be used in the future to guide the experiments toward a well-informed fitting procedure that takes mode mixing and nontrivial spectral overlap into account [10].

It is important to emphasize that a perfect match between simulated and experimental PO-Raman patterns is not expected, as discrepancies already arise at the harmonic level due to the specific level of theory employed for the electronic structure and Raman tensor calculations. Other limitations include the use of finite supercell sizes which impose a finite sampling of the Brillouin zone (here a $4 \times 4 \times 4$ q grid), and exclude q points in the close vicinity of the Brillouin zone center. While the large supercell still mitigates finite-size effects and indirectly affects the dynamics at the Γ point, our analysis of the Raman signals is fundamentally restricted to the Γ point itself, as we stressed from the beginning. We note that allowing for contributions from the vicinity of the Γ point in Raman spectroscopy could strongly affect the mode-coupling picture, by breaking fundamental symmetry restrictions and allowing modes of different irreducible representations to couple [10]. These contributions are very challenging to assess by molecular dynamics alone, requiring very large system sizes.

Nevertheless, our framework demonstrates that, starting from a given theoretical baseline, we can systematically include temperature and dynamical (anharmonic) effects on top of it, providing a predictive route to assess their impact on the PO-Raman response. Overall, we showed that ML-driven simulations combined with established theoretical frameworks can unravel complex anharmonic phenomena in molecular crystals. The insights gained here advance the interpretation of PO-Raman spectra and provide a route for future improvements, including more robust deconvolution methods for analysis of two-dimensional polarization-orientation Raman patterns and larger-scale simulations to further bridge theory and experiment.

VI. METHODS AND COMPUTATIONAL DETAILS

A. Training machine learning models for interatomic interactions

In this paper we investigate two types of MLIPs based on NN architectures. One architecture is that of HDNNP [51,61], which uses two-body and three-body local atomic descriptors in the form of atom-centered symmetry functions. In this representation, there are only a few hidden layers and a few nodes on the NN attached to each atomic species are needed, and the construction of the potential is successful for quite complex systems [62–66]. This architecture does not explicitly encode the equivariance of the transformation of each component of the symmetry functions. The target of the optimization of the NN parameters is the total energy of the system, constructed from the sum of the individual contributions of the atomic energies predicted by the NN, and forces are readily obtained by back-propagation (i.e., the model is conservative). We have referred to this model as HDNNP-MLIP in the main text.

The other architecture is the equivariant message-passing graph NN employed in the MACE model [43,44]. In this architecture, atoms are represented as nodes of a graph and the interactions between pairs of atoms are the edges. The messages are constructed as many body expansions where each term is approximated by a linear combination of a complete basis of equivariant features [67]. These messages are thus of high-body order and can be understood as an efficient evaluation of a many-body expansion on ACE descriptors [68]. Forces are obtained by automatic differentiation and the model is also conservative. The use of effective high body-order messages delivers highly accurate predictions. The imposition of equivariance reduces the complexity of the problem by incorporating the rotational properties of the representation, resulting in the need of less training data points and removing the need of explicit data augmentation. We have referred to this model as MACE-MLIP in the main text.

To construct the training dataset for anthracene, we used the experimentally available lattice parameters at temperatures of 100K and 295K from Ref. [8] and performed a 2 ns NVT-MD run with the foundational model MACE-OFF [69] to generate trial structures. A total of 1200 configurations were selected through farthest point sampling and the corresponding total energies and forces were recomputed at the PBE [70] level of theory with many-body dispersion correction (MBD) [71] using the FHI-aims [72] electronic structure

TABLE II. Training and validation errors as RMSE for energies and forces of MACE-MLIP and HDNNP-MLIP for anthracene.

	Energy (meV/atom)		Force (meV/Å)	
	Training	Test	Training	Test
HDNNP-MLIP	0.3	2.5	45.6	47.7
MACE-MLIP	0.2	0.2	5.3	7.1

package, with the recommended *tight* basis set. Dispersion corrections are essential to capture the cohesive forces in molecular crystals and the MBD scheme has proven to be very accurate for structural properties of these systems [73]. Both of our models for anthracene were subsequently trained on this dataset, with 10% of the structures left as test set. The accuracy of the models on training and test set is reported in Table II.

For naphthalene, we used the MACE potential described in Ref. [13]. The underlying electronic-structure is also PBE+MBD with *tight* basis set. For any further detail on the training procedure, the reader is referred to Ref. [13]. The errors on energies and forces are reported here (Table III) for completeness.

B. Training procedure of the machine learning polarizability models

The dataset for both polarizability models consisted in 1000 structures sampled from NVT MD simulations generated by our MACE-MLIP. The structures were sampled at 100K, 150K, and 293K for anthracene and 30K, 80K, and 295K for naphthalene, employing available experimental lattice constants from Refs. [8,74]. The polarizability tensors were computed from DFPT as implemented in the FHI-aims code using the LDA functional, which has been shown in Ref. [75] to be accurate in capturing the polarizability variations that enter the Raman intensity calculation as long as the underlying dynamics is generated by a high quality potential. We also performed an explicit comparison of LDA versus PBE for Raman tensors in Sec. S5 of the SM [26].

The training set was split into 85% training and 15% validation sets. The MACE-α model was trained with the *mu_alpha* branch of the MACE code [76], while the SA-GPR model was trained with a modified version of the TENSOAP code [47] available at Ref. [77].

TABLE III. Training and validation errors as RMSE for energies and forces of MACE-MLIP for naphthalene, reproduced from Ref. [13].

	Energy (meV/atom)		Force (meV/Å)	
	Training	Test	Training	Test
MACE-MLIP	0.1	0.1	2.6	3.4

C. Phonons calculations and geometry optimizations

For the comparison of Sec. S3 of the SM [26] we computed the harmonic phonon frequencies for anthracene on the minimum energy atomic configuration with lattice vectors fixed to the experimental structure at 295K from Ref. [8] with both our ML models and DFT. The geometry optimizations were separately performed with FHI-aims (PBE + MBD, *tight* basis set) for the DFT case and with our two ML potentials. Phonon calculations were subsequently performed using the phonopy code [78,79].

For the production simulations, energy minimization and phonon calculations needed to obtain the normal modes for the RGDOS projection coefficients of Sec. III B were performed with the i-PI [80] code, using the same MLIP as used in the dynamics as the force driver.

D. Raman tensors for RGDOS

The Raman tensors needed for the RGDOS expansion were obtained from finite differentiation of DFPT polarizabilities of displaced minimum energy structures along the normal coordinates. The procedure is repeated for every material at every different lattice parameters considered. The LDA functional was used for the benchmarks of Sec. III C for a direct comparison with the LDA-trained SA-GPR model, while the PBE functional was employed in the production calculations of the RGDOS spectra. All DFPT calculations employed *tight* basis sets. While our comparison of LDA versus PBE Raman tensors in Sec. S5 of the SM [26] shows nearly identical results, we opted for PBE in production calculations due to reduced computational cost of the RGDOS methodology, and to ensure full consistency with the PBE level of theory of the MLIPs employed. Section S5 of the SM [26] (see also additional Ref. [81] therein) also includes a brief discussion on *a posteriori* symmetrization of numerically noisy Raman tensors.

A further benchmark unpolarized harmonic Raman spectra calculated at different levels of theory is provided for completeness in Sec. S13 of the SM [26]. We show how different DFT functionals (including hybrid-GGA) and dispersion corrections do not significantly alter Raman intensities in anthracene at the harmonic level.

E. Production simulations for Raman spectra

For the production runs, we performed 48 independent NVE-MD simulations of 100 ps each with a time step of 1 fs for both anthracene and naphthalene at temperatures of 100K, 220K, and 295K. For both crystals, we employed $4 \times 4 \times 4$ supercells to mitigate finite-size effects. This amounts to a total of 3072 atoms for anthracene and 2304 atoms for the naphthalene crystal. The convergence of the PO-Raman spectra with the supercell dimension is reported in Sec. S11 of the SM [26]. The NVE trajectories were generated from a 2 ns parent NVT trajectory where temperature was controlled with a stochastic velocity-rescaling thermostat [82] with a time constant of 100 fs. The polarizability tensors were computed on-the-fly every 10 fs from the RGDOS expansion Eq. (13), with the procedure to obtain projections coefficients described in Sec. S1 of the SM [26]. The Raman spectra were

subsequently computed through numerical Fourier transformation with a Hanning windowing function applied to the time series. No additional broadening was introduced by the windowing process.

F. Fitting of PO-dependence

Following Ref. [8] we fitted the PO-dependence of the integrated intensity of every peak in the PO-maps, like the ones in Fig. 3, to a multi-Lorentzian function of the form

$$I(\omega) = \sum_i \frac{A_i \Gamma_i^2 \omega_{0,i} \omega}{\omega^2 \Gamma_i^2 + \left(\omega^2 - \omega_{0,i}^2\right)^2}, \tag{15}$$

where A_i is the amplitude, $\omega_{0,i}$ the position, and Γ_i the width of the ith peak. The position and the width for each peak were determined from a first round of fitting at fixed polarization angles at which peaks were as clearly separated as possible. Once determined, under the well-founded assumption that the polarization angle does not influence the peak position and width but only the amplitude, we fitted the oscillating intensities.

G. VDOS and CVDOS calculations

The VDOS and CVDOS of Fig. 5 were computed from Fourier transformation of the velocity autocorrelation function of the Γ-point projections $\tilde{v}_k$ with k being a Γ-point normal mode, following the same procedure as Ref. [83]. The final expression reads as

$$C_{ij}(\omega) = \mathcal{F}\{\langle \tilde{v}_i(t)\tilde{v}_j(0)\rangle\}, \tag{16}$$

where $i = j$ gives the VDOS of mode i and $i \neq j$ gives the CVDOS between modes i and j. When defined like this, the VDOS is normalized to twice the vibrational kinetic energy of the system, $2\langle K\rangle$. The projection velocities $\tilde{v}_k$ are obtained through finite differentiation of normal mode projection coefficients $\tilde{q}_k$ along the MD trajectory. These are defined as

$$\boldsymbol{u}(t) = \sum_k \tilde{Q}_k \tilde{q}_k(t), \tag{17}$$

where $\boldsymbol{u}(t)$ are the atomic displacements at time t and $\tilde{Q}_k$ is the cartesian normal mode k. To extract only the Γ-point projection coefficients, we follow the procedure we described in Sec. S1 of the SM [26]. No windowing was applied to the time series before Fourier transformation.

ACKNOWLEDGMENTS

M.R. acknowledges funding by the European Research Council (ERC) Consolidator Grant (CoG) under the European Union's Horizon Europe Research and Innovation Programme (Grant Agreement No. 101169761, Project Acronym QUADYMM). P.L. and S.S. acknowledge support from the UFAST International Max Planck Research School. The authors are grateful to Nimrod Benshalom, Noam Pinsk, and Omer Yaffe for insightful discussions, valuable input, and for inspiring this work through their earlier experiments. We further acknowledge Nimrod Benshalom for his critical reading of the manuscript. We also thank Burak Gurlek, Elia Stocco, and George Trenins for helpful discussions.

P.L. and M.R. designed and conceptualized the research. S.S. and P.L. trained machine-learning models. P.L. ran simulations, conducted theoretical derivations, analyzed data and produced all figures in the manuscript. M.R. contributed to the data analysis the discussion. P.L. and M.R. wrote the first draft of the paper. All authors edited the draft.

DATA AVAILABILITY

The data that support the findings of this article are openly available [84]. This includes the trained models, training datasets and example input files for every calculation mentioned in Sec VI. Further details are available upon request.

[1] N. Raimbault, V. Athavale, and M. Rossi, Anharmonic effects in the low-frequency vibrational modes of aspirin and paracetamol crystals, Phys. Rev. Mater. **3**, 053605 (2019).

[2] M. Rossi, P. Gasparotto, and M. Ceriotti, Anharmonic and quantum fluctuations in molecular crystals: A first-principles study of the stability of paracetamol, Phys. Rev. Lett. **117**, 115702 (2016).

[3] M. Krynski and M. Rossi, Efficient Gaussian process regression for prediction of molecular crystals harmonic free energies, npj Comput. Mater. **7**, 169 (2021).

[4] V. A. Konstantinov, Heat transfer by low-frequency phonons and "diffusive" modes in molecular crystals, Low Temp. Phys. **29**, 422 (2003).

[5] L. Legenstein, L. Reicht, S. Wieser, M. Simoncelli, and E. Zojer, Heat transport in crystalline organic semiconductors: Coexistence of phonon propagation and tunneling, npj Comput. Mater. **11**, 29 (2025).

[6] S. Fratini, S. Ciuchi, D. Mayou, G. T. de Laissardière, and A. Troisi, A map of high-mobility molecular semiconductors, Nat. Mater. **16**, 998 (2017).

[7] V. Coropceanu, J. Cornil, D. A. da Silva Filho, Y. Olivier, R. Silbey, and J.-L. Brédas, Charge transport in organic semiconductors, Chem. Rev. **107**, 926 (2007).

[8] M. Asher, D. Angerer, R. Korobko, Y. Diskin-Posner, D. A. Egger, and O. Yaffe, Anharmonic lattice vibrations in small-molecule organic semiconductors, Adv. Mater. **32**, 1908028 (2020).

[9] M. Asher, R. Jouclas, M. Bardini, Y. Diskin-Posner, N. Kahn, R. Korobko, A. R. Kennedy, L. Silva de Moraes, G. Schweicher, J. Liu, D. Beljonne, Y. Geerts, and O. Yaffe, Chemical modifications suppress anharmonic effects in the lattice dynamics of organic semiconductors, ACS Materials Au **2**, 699 (2022).

[10] N. Benshalom, M. Asher, R. Jouclas, R. Korobko, G. Schweicher, J. Liu, Y. Geerts, O. Hellman, and O. Yaffe, Phonon–phonon interactions in the polarization dependence of Raman scattering, J. Phys. Chem. C **127**, 18099 (2023).

[11] N. Raimbault, A. Grisafi, M. Ceriotti, and M. Rossi, Using Gaussian process regression to simulate the vibrational Raman spectra of molecular crystals, New J. Phys. **21**, 105001 (2019).

[12] E. Berger and H.-P. Komsa, Polarizability models for simulations of finite temperature Raman spectra from machine learning molecular dynamics, Phys. Rev. Mater. **8**, 043802 (2024).
[13] B. Gurlek, S. Sharma, P. Lazzaroni, A. Rubio, and M. Rossi, Accurate machine learning interatomic potentials for polyacene molecular crystals: Application to single molecule host-guest systems, npj Comput. Mater. **11**, 318 (2025).
[14] M. Grumet, C. Von Scarpatetti, T. Bučko, and D. A. Egger, Delta machine learning for predicting dielectric properties and Raman spectra, J. Phys. Chem. C **128**, 6464 (2024).
[15] A. Hashemi, A. V. Krasheninnikov, M. Puska, and H.-P. Komsa, Efficient method for calculating Raman spectra of solids with impurities and alloys and its application to two-dimensional transition metal dichalcogenides, Phys. Rev. Mater. **3**, 023806 (2019).
[16] E. Berger, Z.-P. Lv, and H.-P. Komsa, Raman spectra of 2D titanium carbide MXene from machine-learning force field molecular dynamics, J. Mater. Chem. C **11**, 1311 (2023).
[17] C. V. Raman and K. S. Krishnan, A new type of secondary radiation, Nature (London) **121**, 501 (1928).
[18] J. Kim, J.-U. Lee, J. Lee, H. Ju Park, Z. Lee, C. Lee, and H. Cheong, Anomalous polarization dependence of Raman scattering and crystallographic orientation of black phosphorus, Nanoscale **7**, 18708 (2015).
[19] B. Xu, N. Mao, Y. Zhao, L. Tong, and J. Zhang, Polarized Raman spectroscopy for determining crystallographic orientation of low-dimensional materials, J. Phys. Chem. Lett. **12**, 7442 (2021).
[20] D. A. McQuarrie, *Statistical Mechanics* (University Scince Books, Mill Valley, CA, 2000).
[21] D. A. Long, *The Raman Effect: A Unified Treatment of the Theory of Raman Scattering by Molecules* (Wiley, New York, NY, 2002).
[22] M. Walter and M. Moseler, *Ab initio* wavelength-dependent Raman spectra: Placzek approximation and beyond, J. Chem. Theory Comput. **16**, 576 (2020).
[23] H. A. Kramers and W. Heisenberg, Über die Streuung von Strahlung durch Atome, Z. Phys. **31**, 681 (1925).
[24] P. A. M. Dirac and R. H. Fowler, The quantum theory of dispersion, Proc. R. Soc. London, Ser. A **114**, 710 (1927).
[25] F. Knoop, N. Benshalom, M. Menahem, P. Gartner, T. Salzillo, O. Yaffe, and O. Hellman, *Ab initio* theory of the non-resonant Raman effect in crystals at finite temperature in comparison to experiment: The examples of GaN and BaZrS3, arXiv:2412.17711.
[26] See Supplemental Material at http://link.aps.org/supplemental/10.1103/t4s2-45t8 for details on normal modes projections, supercell convergence, perpendicular configuration PO-maps, Raman tensors benchmarking, and full harmonic PO-Raman results.
[27] B. A. DeAngelis, R. E. Newnham, and W. B. White, Factor group analysis of the vibrational spectra of crystals: A review and consolidation, Am. Mineral. **57**, 255 (1972).
[28] M. H. Beck, A. Jäckle, G. A. Worth, and H. D. Meyer, The multiconfiguration time-dependent Hartree (MCTDH) method: A highly efficient algorithm for propagating wavepackets, Phys. Rep. **324**, 1 (2000).
[29] S. Chatterjee and N. Makri, Real-time path integral methods, quantum master equations, and classical vs quantum memory, J. Phys. Chem. B **123**, 10470 (2019).
[30] I. R. Craig and D. E. Manolopoulos, Quantum statistics and classical mechanics: Real time correlation functions from ring polymer molecular dynamics, J. Chem. Phys. **121**, 3368 (2004).
[31] B. J. Braams and D. E. Manolopoulos, On the short-time limit of ring polymer molecular dynamics, J. Chem. Phys. **125**, 124105 (2006).
[32] S. Habershon, D. E. Manolopoulos, T. E. Markland, and T. F. Miller, Ring-polymer molecular dynamics: Quantum effects in chemical dynamics from classical trajectories in an extended phase space, Annu. Rev. Phys. Chem. **64**, 387 (2013).
[33] M. Rossi, M. Ceriotti, and D. E. Manolopoulos, How to remove the spurious resonances from ring polymer molecular dynamics, J. Chem. Phys. **140**, 234116 (2014).
[34] S. C. Althorpe, Path integral simulations of condensed-phase vibrational spectroscopy, Annu. Rev. Phys. Chem. **75**, 397 (2024).
[35] S. Jang and G. A. Voth, A derivation of centroid molecular dynamics and other approximate time evolution methods for path integral centroid variables, J. Chem. Phys. **111**, 2371 (1999).
[36] G. Trenins, M. J. Willatt, and S. C. Althorpe, Path-integral dynamics of water using curvilinear centroids, J. Chem. Phys. **151**, 054109 (2019).
[37] J. E. Lawrence, A. Z. Lieberherr, T. Fletcher, and D. E. Manolopoulos, Fast quasi-centroid molecular dynamics for water and ice, J. Phys. Chem. B **127**, 9172 (2023).
[38] F. Musil, I. Zaporozhets, F. Noé, C. Clementi, and V. Kapil, Quantum dynamics using path integral coarse-graining, J. Chem. Phys. **157**, 181102 (2022).
[39] T. E. Markland and M. Ceriotti, Nuclear quantum effects enter the mainstream, Nat. Rev. Chem. **2**, 0109 (2018).
[40] J. Hoja, A. M. Reilly, and A. Tkatchenko, First-principles modeling of molecular crystals: Structures and stabilities, temperature and pressure, WIREs Comput. Mol. Sci. **7**, e1294 (2017).
[41] A. Putrino and M. Parrinello, Anharmonic Raman spectra in high-pressure ice from *ab initio* simulations, Phys. Rev. Lett. **88**, 176401 (2002).
[42] M. Pagliai, C. Cavazzoni, G. Cardini, G. Erbacci, M. Parrinello, and V. Schettino, Anharmonic infrared and Raman spectra in Car–Parrinello molecular dynamics simulations, J. Chem. Phys. **128**, 224514 (2008).
[43] I. Batatia, D. P. Kovacs, G. Simm, C. Ortner, and G. Csanyi, MACE: Higher order equivariant message passing neural networks for fast and accurate force fields, in *Advances in Neural Information Processing Systems*, edited by S. Koyejo, S. Mohamed, A. Agarwal, D. Belgrave, K. Cho, and A. Oh (MIT Press, Cambridge, MA, 2022), Vol. 35, pp. 11423–11436.
[44] I. Batatia, S. Batzner, D. P. Kovács, A. Musaelian, G. N. C. Simm, R. Drautz, C. Ortner, B. Kozinsky, and G. Csányi, The design space of E(3)-equivariant atom-centred interatomic potentials, Nat. Mach. Intell. **7**, 56 (2025).
[45] V. Kapil, D. P. Kovács, G. Csányi, and A. Michaelides, First-principles spectroscopy of aqueous interfaces using machine-learned electronic and quantum nuclear effects, Faraday Discuss. **249**, 50 (2024).
[46] B. A. A. Martin, A. M. Ganose, V. Kapil, and K. T. Butler, General learning of the electric response of inorganic materials, arXiv:2508.17870.

[47] A. Grisafi, D. M. Wilkins, G. Csányi, and M. Ceriotti, Symmetry-adapted machine learning for tensorial properties of atomistic systems, Phys. Rev. Lett. **120**, 036002 (2018).
[48] V. L. Deringer, A. P. Bartók, N. Bernstein, D. M. Wilkins, M. Ceriotti, and G. Csányi, Gaussian process regression for materials and molecules, Chem. Rev. **121**, 10073 (2021).
[49] M. J. Willatt, F. Musil, and M. Ceriotti, Atom-density representations for machine learning, J. Chem. Phys. **150**, 154110 (2019).
[50] A. Jana, S. Shepherd, Y. Litman, and D. M. Wilkins, Learning electronic polarizations in aqueous systems, J. Chem. Inf. Model. **64**, 4426 (2024).
[51] J. Behler and M. Parrinello, Generalized neural-network representation of high-dimensional potential-energy surfaces, Phys. Rev. Lett. **98**, 146401 (2007).
[52] M. Suzuki, T. Yokoyama, and M. Ito, Polarized Raman spectra of naphthalene and anthracene single crystals, Spectrochim. Acta A: Mol. Spectrosc. **24**, 1091 (1968).
[53] G. N. Zhizhin and E. Mukhtarov, Vibrational spectra and structure, in *Optical Spectra and Lattice Dynamics of Molecular Crystals* (Elsevier Science, Amsterdam, 1995), Vol. 21, pp. 245–267.
[54] H. Seiler, M. Krynski, D. Zahn, S. Hammer, Y. W. Windsor, T. Vasileiadis, J. Pflaum, R. Ernstorfer, M. Rossi, and H. Schwoerer, Nuclear dynamics of singlet exciton fission in pentacene single crystals, Sci. Adv. **7**, eabg0869 (2021).
[55] H.-H. Kowalski, Theory of thermal conductivity, Ph.D. thesis, TU Berlin, 2023.
[56] Y. Litman, J. Behler, and M. Rossi, Temperature dependence of the vibrational spectrum of porphycene: A qualitative failure of classical-nuclei molecular dynamics, Faraday Discuss. **221**, 526 (2020).
[57] M. Rossi, Progress and challenges in *ab initio* simulations of quantum nuclei in weakly bonded systems, J. Chem. Phys. **154**, 170902 (2021).
[58] J. Wang, S. Olsson, C. Wehmeyer, A. Pérez, N. E. Charron, G. de Fabritiis, F. Noé, and C. Clementi, Machine learning of coarse-grained molecular dynamics force fields, ACS Cent. Sci. **5**, 755 (2019).
[59] J. Castro, G. Trenins, V. Kapil, and M. Rossi, Vibrational spectra of materials and molecules from partially-adiabatic elevated-temperature centroid molecular dynamics, J. Chem. Phys. **163**, 204102 (2025).
[60] C. Kranert, C. Sturm, R. Schmidt-Grund, and M. Grundmann, Raman tensor formalism for optically anisotropic crystals, Phys. Rev. Lett. **116**, 127401 (2016).
[61] J. Behler, Four generations of high-dimensional neural network potentials, Chem. Rev. **121**, 10037 (2021).
[62] A. Omranpour and J. Behler, Insights into the structure and dynamics of water at Co_3O_4 (001) using a high-dimensional neural network potential, Chem. Cat. Chem. **18**, e01321 (2026).
[63] N. Artrith, T. Morawietz, and J. Behler, High-dimensional neural-network potentials for multicomponent systems: Applications to zinc oxide, Phys. Rev. B **83**, 153101 (2011).
[64] N. Artrith and J. Behler, High-dimensional neural network potentials for metal surfaces: A prototype study for copper, Phys. Rev. B **85**, 045439 (2012).
[65] G. C. Sosso, G. Miceli, S. Caravati, J. Behler, and M. Bernasconi, Neural network interatomic potential for the phase change material gete, Phys. Rev. B **85**, 174103 (2012).
[66] S. K. Natarajan and J. Behler, Neural network molecular dynamics simulations of solid–liquid interfaces: Water at low-index copper surfaces, Phys. Chem. Chem. Phys. **18**, 28704 (2016).
[67] S. Batzner, A. Musaelian, L. Sun, M. Geiger, J. P. Mailoa, M. Kornbluth, N. Molinari, T. E. Smidt, and B. Kozinsky, E(3)-equivariant graph neural networks for data-efficient and accurate interatomic potentials, Nat. Commun. **13**, 2453 (2022).
[68] R. Drautz, Atomic cluster expansion for accurate and transferable interatomic potentials, Phys. Rev. B **99**, 014104 (2019).
[69] D. P. Kovács, J. H. Moore, N. J. Browning, I. Batatia, J. T. Horton, Y. Pu, V. Kapil, W. C. Witt, I.-B. Magdău, D. J. Cole, and G. Csányi, MACE-OFF: Short-range transferable machine learning force fields for organic molecules, J. Am. Chem. Soc. **147**, 17598 (2025).
[70] J. P. Perdew, K. Burke, and M. Ernzerhof, Generalized gradient approximation made simple, Phys. Rev. Lett. **77**, 3865 (1996).
[71] A. Tkatchenko and M. Scheffler, Accurate molecular Van der Waals interactions from ground-state electron density and free-atom reference data, Phys. Rev. Lett. **102**, 073005 (2009).
[72] V. Blum, R. Gehrke, F. Hanke, P. Havu, V. Havu, X. Ren, K. Reuter, and M. Scheffler, *Ab initio* molecular simulations with numeric atom-centered orbitals, Comput. Phys. Commun. **180**, 2175 (2009).
[73] L. M. Hunnisett, N. Francia, J. Nyman, N. S. Abraham, S. Aitipamula, T. Alkhidir, M. Almehairbi, A. Anelli, D. M. Anstine, J. E. Anthony, J. E. Arnold, F. Bahrami, M. A. Bellucci, G. J. O. Beran, R. M. Bhardwaj, R. Bianco, J. A. Bis, A. D. Boese, J. Bramley, D. E. Braun, *et al.*, The seventh blind test of crystal structure prediction: Structure ranking methods, Acta Crystallogr. B: Struct. Sci. Cryst. Eng. Mater **80**, 548 (2024).
[74] S. C. Capelli, A. Albinati, S. A. Mason, and B. T. M. Willis, Molecular motion in crystalline naphthalene: Analysis of multi-temperature X-ray and neutron diffraction data, J. Phys. Chem. A **110**, 11695 (2006).
[75] H. Shang, N. Raimbault, P. Rinke, M. Scheffler, M. Rossi, and C. Carbogno, All-electron, real-space perturbation theory for homogeneous electric fields: Theory, implementation, and application within DFT, New J. Phys. **20**, 073040 (2018).
[76] GitHub—ACEsuit/mace: MACE—Fast and accurate machine learning interatomic potentials with higher order equivariant message passing—github.com, https://github.com/ACEsuit/mace.
[77] GitHub—alanmlewis/TENSOAP: Public repository for symmetry-adapted Gaussian process regression (SA-GPR)—github.com, https://github.com/alanmlewis/TENSOAP.
[78] A. Togo, First-principles phonon calculations with Phonopy and Phono3py, J. Phys. Soc. Jpn. **92**, 012001 (2023).
[79] A. Togo, L. Chaput, T. Tadano, and I. Tanaka, Implementation strategies in phonopy and phono3py, J. Phys.: Condens. Matter **35**, 353001 (2023).
[80] Y. Litman, V. Kapil, Y. M. Y. Feldman, D. Tisi, T. Begušić, K. Fidanyan, G. Fraux, J. Higer, M. Kellner, T. E. Li, E. S. Pós, E. Stocco, G. Trenins, B. Hirshberg, M. Rossi, and M. Ceriotti,

I-PI 3.0: A flexible and efficient framework for advanced atomistic simulations, J. Chem. Phys. **161**, 062504 (2024).

[81] A. Togo, K. Shinohara, and I. Tanaka, Spglib: A software library for crystal symmetry search, Sci. Technol. Adv. Mater.: Methods **4**, 2384822 (2024).

[82] G. Bussi, D. Donadio, and M. Parrinello, Canonical sampling through velocity rescaling, J. Chem. Phys. **126**, 014101 (2007).

[83] A. Carreras, A. Togo, and I. Tanaka, DynaPhoPy: A code for extracting phonon quasiparticles from molecular dynamics simulations, Comput. Phys. Commun. **221**, 221 (2017).

[84] https://doi.org/10.5281/zenodo.18445141.

Correction: A support statement in the first sentence of the Acknowledgments has been corrected.