

## ARTICLE OPEN

## Theoretical analysis of spectral lineshapes from molecular dynamics

Andrew Cupo<sup>1</sup>, Damien Tristant<sup>1</sup>, Kyle Rego<sup>1</sup> and Vincent Meunier<sup>1</sup>

Conventional methods for calculating anharmonic phonon properties are computationally expensive. To address this issue, a theoretical approach was developed for the accelerated calculation of vibrational lineshapes for spectra obtained from finite-time molecular dynamics. The method gives access to the effect of anharmonicity-induced frequency shift and lifetime, as well as simulation broadening. For a toy model we demonstrate at least an order of magnitude reduction in the number of simulation steps needed to obtain converged vibrational properties in nearly all cases considered as compared to the standard extraction procedure. The theory is also illustrated for graphene, hexagonal boron nitride, and silicon at the density functional theory level, with up to nearly a factor of 9 reduction in the required simulation time to reach convergence in the vibrational frequencies and lifetimes. In general, we expect the newly developed method to outperform the standard procedure when the anharmonicity is sufficiently weak so that well-defined renormalized phonon quasiparticles emerge. Our extension of signal analysis to material vibrations represents a state-of-the-art advance in calculating temperature-dependent phonon properties and could be implemented in computational materials discovery packages that search for thermoelectric materials for instance, since the thermal conductivity contribution to  $ZT$  depends strongly on these characteristics.

*npj Computational Materials* (2019)5:82; <https://doi.org/10.1038/s41524-019-0220-1>

## INTRODUCTION

Phonons constitute an important topic in condensed matter physics and are required to explain phenomena ranging from thermal expansion to electrical resistivity to BCS superconductivity.<sup>1,2</sup> Due to advancements in supercomputing and density functional theory (DFT), first-principles calculations of phonon properties are now routine.<sup>3</sup> Within the harmonic approximation many phonon characteristics can already be calculated, including the phonon dispersion and density of states (DOS), specific heat, Helmholtz free energy, and thermal expansion.<sup>3</sup> These calculations are even permissible at the DFT level for systems with hundreds of atoms, one example being single-layer black phosphorus or blue phosphorus adsorbed onto Au (111).<sup>4</sup> However, some properties cannot be described within the quasi-harmonic approximation: Phonon anharmonicity results in phonon–phonon coupling and is required to describe frequency shift, lifetime, and thermal conductivity (all temperature-dependent). Anharmonic properties can be determined using many-body perturbation theory<sup>5–13</sup> or density functional perturbation theory.<sup>14–17</sup> For even the simplest materials the former method can require hundreds or thousands of accurate calculations on a large supercell using the finite displacement method. For example, with the PHONO3PY computer code<sup>13</sup> the four-atom primitive cell of single-layer black phosphorus<sup>18</sup> requires 4741 separate DFT calculations on a 256 atom supercell without using a cutoff distance to describe the three phonon processes. Quartic order treatment may also be necessary to obtain quantitative agreement with experiment.<sup>9,12</sup> Furthermore, at sufficiently high-temperature perturbation approaches break down.

On the other hand when molecular dynamics (MD) simulations are performed, phonon anharmonicity is treated exactly to all orders in principle<sup>19</sup> and is only limited by how accurately the exchange correlation functional (quantum)<sup>20</sup> or parameterized force-field<sup>21</sup> reproduce the actual energy surface. Regardless of the level of theory used to calculate the forces, the dynamics are still classical, but the quantum effect of the zero point motion can be included by a simple rescaling of the MD temperature.<sup>22</sup> The idea of obtaining vibrational properties from MD goes back to at least 1969: Temperature-dependent vibrational frequencies were extracted from the peak positions in spectra calculated from the Fourier transform of the velocity auto-correlation. It was also suggested early that the decay of the velocity auto-correlation of the normal mode coordinates is related to the phonon lifetime.<sup>19</sup> Later the phonon lifetime was extracted from the inverse of the full-width at half-maximum (FWHM) of the spectra peaks fitted with Lorentzians.<sup>22–24</sup> These methods have become popular and comprise fully dedicated analysis tools such as DynaPhoPy.<sup>25</sup> An extension of the auto-correlation method was proposed recently and involves the ratio of correlations of conjugate variables.<sup>26</sup> Furthermore, spectra obtained from the velocity auto-correlation without projections are related to the phonon DOS,<sup>27</sup> while spectra from the polarizability tensor auto-correlation give the Raman scattering cross section.<sup>28</sup> Another approach focuses on spectra derived from the kinetic energy and is termed the spectral energy density (SED) method.<sup>29</sup> In the temperature-dependent effective potential (TDEP) method, the harmonic, cubic, and quartic force constants can be obtained directly by minimizing an error function based on all forces from the trajectory, but the phonon lifetimes are still evaluated using perturbation theory.<sup>30–33</sup> On the other hand, self-consistent phonon methods offer

<sup>1</sup>Department of Physics, Applied Physics, and Astronomy, Rensselaer Polytechnic Institute, Troy, NY, USA  
Correspondence: Vincent Meunier (meuniv@rpi.edu)

Received: 14 February 2019 Accepted: 16 July 2019

Published online: 07 August 2019

non-perturbative results.<sup>34–41</sup> In these last two methods the benefit is that deviation of the dynamics lineshapes from what is expected in theory is not an issue, although determining the convergence properties with the simulation time analytically could be challenging.

In this paper, we derive exact analytical expressions for spectra obtained from MD starting with a simple perturbed normal mode expression for the velocities. The effect of anharmonicity-induced frequency shift and lifetime as well as the simulation time are included. The derived fitting functions are applied to a simple model, graphene, hexagonal boron nitride (hBN), and silicon to study the convergence of the vibrational frequency and lifetime with simulation time. It is expected that the proposed method could have far reaching impact on the determination of phononic properties from MD, in areas ranging from strongly correlated systems to biological materials.

## RESULTS

### Normal mode analysis

Consider a supercell subdivided into primitive cells each indexed by index  $\rho$  with corresponding atoms between primitive cells indexed by  $N$ . The atoms move about their equilibrium positions  $\mathbf{x}_{0,\rho N}$  as time  $t$  evolves:

$$\mathbf{x}_{\rho N}(t) = \mathbf{x}_{0,\rho N} + \mathbf{d}^{\rho N}(t). \quad (1)$$

Here  $\mathbf{x}_{\rho N}(t)$  are the time-dependent positions and  $\mathbf{d}^{\rho N}(t)$  are the time-dependent displacements from equilibrium. The potential energy  $V$  can be expanded in a Taylor series about the equilibrium configuration in the displacements of the atoms:

$$V = E_0 + \frac{1}{2} \sum_{\rho,\rho'} \sum_{N,N'} \sum_{\alpha,\alpha'} \left( \frac{\partial^2 V}{\partial x_{\rho N \alpha} \partial x_{\rho' N' \alpha'}} \right)_0 d^{\rho N \alpha} d^{\rho' N' \alpha'} + \dots \quad (2)$$

The ground state energy is  $E_0$ , the Cartesian components are represented by the index  $\alpha$ , and the second partial derivative is evaluated at equilibrium. In the harmonic approximation the higher order terms not shown in Eq. (2) are ignored and normal modes arise (indexed by  $n$ ).<sup>42</sup> The general solution for the time-dependent displacements can be written in terms of a superposition of the normal modes:<sup>42</sup>

$$\mathbf{d}^{\rho N}(t) = \sum_n \sum_{m=-1}^1 \int_{\text{BZ}} d\mathbf{q} e^{i\mathbf{q} \cdot \mathbf{x}_{0,\rho N}} a_{Nnm}(\mathbf{q}) e^{im\omega_n(\mathbf{q})t} \mathbf{u}_{Nn}(\mathbf{q}). \quad (3)$$

Here  $\mathbf{q}$  is a phonon wave vector in the first Brillouin zone (BZ), the coefficients  $a_{Nnm}(\mathbf{q})$  are determined by initial conditions,  $i^2 = -1$ ,  $\omega_n(\mathbf{q})$  is the angular frequency of the  $n$ th normal mode, and  $\mathbf{u}_{Nn}(\mathbf{q})$  is a normalized vector in the direction of the oscillation for the  $N$ th atom and the  $n$ th normal mode. The sum over  $m$  is introduced for compactness and  $a_{Nn0}(\mathbf{q}) = 0$ . It will be apparent in the discussion of the toy model that it is simpler to work with the velocities ( $b_{Nnm}(\mathbf{q}) \equiv im\omega_n(\mathbf{q})$ ):

$$\mathbf{v}^{\rho N}(t) = \frac{d}{dt} \mathbf{d}^{\rho N}(t) = \sum_n \sum_m \int d\mathbf{q} e^{i\mathbf{q} \cdot \mathbf{x}_{0,\rho N}} b_{Nnm}(\mathbf{q}) e^{im\omega_n(\mathbf{q})t} \mathbf{u}_{Nn}(\mathbf{q}). \quad (4)$$

Equation (4) implies that the velocities calculated from MD will contain information about all phonon wave vectors and normal modes. Taking the Fourier transform of these velocities directly will lead to complicated spectra, so it is best to select a single phonon wave vector and normal mode at a time. We can project onto a specific phonon wave vector  $\mathbf{Q}$ .<sup>19,43</sup>

$$\mathbf{v}_{\mathbf{Q}}^N(t) \equiv \sum_p \mathbf{v}^{\rho N}(t) e^{-i\mathbf{Q} \cdot \mathbf{x}_{0,\rho N}}. \quad (5)$$

Further, we can also project onto a phonon wave vector  $\mathbf{Q}$  and project onto normal mode  $n$ , provided the normal modes can be

calculated in advance:<sup>44</sup>

$$\mathbf{v}_{\eta\mathbf{Q}}^N(t) \equiv \mathbf{v}_{\mathbf{Q}}^N(t) \cdot \mathbf{u}_{N\eta}(\mathbf{Q}). \quad (6)$$

If the number of atoms per primitive cell is  $N_A$ , then we expect to find  $3N_A$  normal modes after the unfolding procedure. There are two general approaches to generating spectra: (i) projected velocities based on Eq. (6) and (ii) velocity auto-correlation based on Eqs. (5) or (6).

### Spectral lineshapes for projected velocities

For the projected velocities in Eq. (6) the spectrum is defined by

$$S_{\eta\mathbf{Q}}(\omega) \equiv \hat{N}\hat{R} \sum_{N=1}^{N_A} M_N^{1/2} \int_{t_i}^{t_f} dt \mathbf{v}_{\eta\mathbf{Q}}^N(t) e^{i\omega t}. \quad (7)$$

The operator  $\hat{N}$  normalizes the spectrum so that the maximum deviation from zero is one and the operator  $\hat{R}$  takes the real part. The limits of the integral ( $t_i$  and  $t_f$ ) are the initial and final simulation times, respectively. The reason for not using the complex modulus will be apparent later. Contributions from all atoms are taken into account, and the weighting by  $M_N^{1/2}$  puts atoms with different masses on equal footing since  $b_{Nnm}(\mathbf{q}) \propto M_N^{-1/2}$ .<sup>42</sup> In the idealized case when  $t_i \rightarrow -\infty$  and  $t_f \rightarrow \infty$ , the resulting spectrum is

$$S_{\eta\mathbf{Q}}(\omega) = \hat{N}2\pi\hat{R} \sum_N M_N^{1/2} \sum_{\rho} \sum_n \sum_m \int d\mathbf{q} e^{i(\mathbf{q}-\mathbf{Q}) \cdot \mathbf{x}_{0,\rho N}} \dots \cdot b_{Nnm}(\mathbf{q}) (\mathbf{u}_{Nn}(\mathbf{q}) \cdot \mathbf{u}_{N\eta}(\mathbf{Q})) \delta(\omega + m\omega_n(\mathbf{q})). \quad (8)$$

From this expression we see that the spectrum will exhibit peaks at the normal mode frequencies. It is also apparent that there will be no direct comparison to Raman spectra, since the spectrum features all of the normal mode frequencies, i.e. there are no selection rules. For an actual simulation of duration  $T$  with  $t_i = -T/2$  and  $t_f = T/2$ :

$$S_{\eta\mathbf{Q}}(\omega) = \hat{N}T\hat{R} \sum_N M_N^{1/2} \sum_{\rho} \sum_n \sum_m \int d\mathbf{q} e^{i(\mathbf{q}-\mathbf{Q}) \cdot \mathbf{x}_{0,\rho N}} \dots \cdot b_{Nnm}(\mathbf{q}) (\mathbf{u}_{Nn}(\mathbf{q}) \cdot \mathbf{u}_{N\eta}(\mathbf{Q})) \text{sinc}[(\omega + m\omega_n(\mathbf{q}))T/2]. \quad (9)$$

Here  $\text{sinc}(x) = \sin(x)/x$ . If we make the approximation that both projections are complete, the appropriate fitting function is:

$$S_{\eta\mathbf{Q}}(\omega) = \sum_m C_{\eta m}(\mathbf{Q}) \text{sinc}[(\omega + m\omega_n(\mathbf{Q}))T/2]. \quad (10)$$

The contributions of a given normal mode at negative and positive frequencies need to be considered in particular when very few configurations are used and the two peaks overlap significantly. The fitting parameters are the normal mode frequency  $\omega_n(\mathbf{Q})$  and the  $C_{\eta m}(\mathbf{Q})$  amplitudes (all real). If the complex modulus were used instead of the real part, the amplitudes would be complex numbers and an artificial non-linearity between the peaks at negative and positive frequencies would be created. This would increase the computational cost of the fitting procedure. More importantly, it appears that the analytical manipulations would be more difficult when using the complex modulus. It makes no difference if what we have defined as the spectrum has negative values, since the precise fitting function is derived. For a sufficiently isolated peak, the full-width (FW) of the central peak as defined by the zeros of the sinc function is only dependent on the simulation time:

$$\Delta\omega_{\text{FW}} = \frac{4\pi}{T}. \quad (11)$$

So far the results are based on the harmonic approximation, which predicts no temperature dependence for the vibrational properties. Provided that the anharmonicity in the potential energy surface accessed by the atoms at a particular temperature is sufficiently small, we can assume that the superposition of normal

modes given in Eq. (4) is still valid with the harmonic vibrational frequencies  $\omega_n(\mathbf{q})$  replaced by their temperature-dependent renormalized values. This is analogous to perturbing eigenvalues on a fixed basis set in quantum theory. In the context of quantum field theory the phonon self-energy ( $PSE_n(\mathbf{q})$ ) describes how the properties of a given phonon are modified due to its anharmonicity-induced interaction with all other phonons. The real part of  $PSE_n(\mathbf{q})$  is the frequency shift  $\Delta\omega_n(\mathbf{q})$  and the imaginary part is the half-linewidth  $\Gamma_n(\mathbf{q})$  (both up to a factor of  $\hbar$ ).<sup>13</sup> The requirement for a weakly anharmonic system is that the magnitudes of the phonon self-energies are much smaller than the harmonic vibrational frequencies.<sup>24</sup> The lifetime  $\tau_n(\mathbf{q})$  is related to the linewidth  $2\Gamma_n(\mathbf{q})$  by

$$\tau_n(\mathbf{q}) = (2\Gamma_n(\mathbf{q}))^{-1}. \quad (12)$$

In the method developed here,  $\Gamma_n(\mathbf{q})$  becomes a fitting parameter in the spectral lineshape, which we achieve by making the following substitution in Eq. (4):<sup>45,46</sup>

$$e^{im\omega_n(\mathbf{q})t} \rightarrow e^{[im\omega_n(\mathbf{q}) - \Gamma_n(\mathbf{q})]t}. \quad (13)$$

A similar analysis results in the fitting function for spectra including the effect of phonon lifetime:

$$S_{\eta\mathbf{Q}}(\omega) = \frac{1}{m} \sum_n \frac{1}{(\omega + m\omega_n(\mathbf{Q}))^2 + (\frac{\Gamma_n(\mathbf{Q})}{2})^2} \dots \\ \dots \{ [C_{nm}(\mathbf{Q})(\omega + m\omega_n(\mathbf{Q})) + C'_{nm}(\mathbf{Q})\Gamma_n(\mathbf{Q})] \cosh(\Gamma_n(\mathbf{Q})T/2) \sin[(\omega + m\omega_n(\mathbf{Q}))T/2] \dots \\ \dots + [C_{nm}(\mathbf{Q})\Gamma_n(\mathbf{Q}) - C'_{nm}(\mathbf{Q})(\omega + m\omega_n(\mathbf{Q}))] \sinh(\Gamma_n(\mathbf{Q})T/2) \cos[(\omega + m\omega_n(\mathbf{Q}))T/2] \}. \quad (14)$$

$\Gamma_n(\mathbf{Q})$  and the amplitude  $C'_{nm}(\mathbf{Q})$  are additional fitting parameters (both also real). In the limit that  $\Gamma_n(\mathbf{Q})$  goes to zero, Eq. (14) reduces to Eq. (10) as expected.

#### Spectral lineshapes for velocity auto-correlation

In the previous section, it was assumed that the eigenvectors for the normal modes were already known. In cases where the system of interest is too large to determine eigenvectors, it is necessary to generate a full spectrum which contains information about all the normal modes. First, note that a full spectrum cannot be obtained from the magnitude of the velocity squared (see SI section I). However, this can be accomplished by manipulating the velocity auto-correlation as follows ( $\tau$  is the temporal offset):

$$A_{\mathbf{Q}}(\tau) \equiv \sum_N M_N \int dt \mathbf{v}_{\mathbf{Q}}^N(t) \cdot \mathbf{v}_{\mathbf{Q}}^N(t + \tau). \quad (15)$$

Note that the phonon DOS can also be obtained from the velocity auto-correlation.<sup>27</sup> In that case the sum over the unit cells ( $\rho$ ) occurs after the velocity auto-correlation is calculated, whereas in Eq. (15) the sum over the unit cells precedes the velocity auto-correlation calculation via Eq. (5). Similar to the previous section the spectrum is given by

$$S_{\mathbf{Q}}^A(\omega) \equiv \hat{N}\hat{R} \int d\tau A_{\mathbf{Q}}(\tau) e^{i\omega\tau}. \quad (16)$$

To obtain analytical expressions for the spectra it is useful to write:

$$\mathbf{v}_{\mathbf{Q}}^N(t) = \sum_{\rho'} \sum_{n'} \sum_{m'} \int d\mathbf{q}' e^{i(\mathbf{q}' - \mathbf{Q}) \cdot \mathbf{x}_{0,\rho'N}} b_{Nn'm'}(\mathbf{q}') e^{im'\omega_{n'}(\mathbf{q}')t} \mathbf{u}_{Nn'}(\mathbf{q}'), \\ \mathbf{v}_{\mathbf{Q}}^N(t + \tau) = \sum_{\rho} \sum_n \sum_m \int d\mathbf{q} e^{i(\mathbf{q} - \mathbf{Q}) \cdot \mathbf{x}_{0,\rho N}} b_{Nm}(\mathbf{q}) e^{im\omega_n(\mathbf{q})(t+\tau)} \mathbf{u}_{Nm}(\mathbf{q}).$$

The resulting spectrum is

$$S_{\mathbf{Q}}^A(\omega) = \hat{N}\hat{R} \sum_N M_N \sum_{\rho,\rho'} \sum_{n,n'} \sum_{m,m'} \int d\mathbf{q} \int d\mathbf{q}' e^{i(\mathbf{q} - \mathbf{Q}) \cdot \mathbf{x}_{0,\rho N} + i(\mathbf{q}' - \mathbf{Q}) \cdot \mathbf{x}_{0,\rho' N}} \dots \\ \dots b_{Nm}(\mathbf{q}) b_{Nn'm'}(\mathbf{q}') (\mathbf{u}_{Nm}(\mathbf{q}) \cdot \mathbf{u}_{Nn'}(\mathbf{q}')) I. \quad (17)$$

The integral  $I$  is defined as (the limits are not specified here):

$$I \equiv \int d\tau e^{i(\omega + m\omega_n(\mathbf{Q}))\tau} \int dt e^{i(m\omega_n(\mathbf{Q}) + m'\omega_{n'}(\mathbf{q}'))t}. \quad (18)$$

In the idealized case, all of the integration limits are infinite:

$$I = \left( \int_{-\infty}^{\infty} dt e^{i(m\omega_n(\mathbf{Q}) + m'\omega_{n'}(\mathbf{q}'))t} \right) \left( \int_{-\infty}^{\infty} d\tau e^{i(\omega + m\omega_n(\mathbf{Q}))\tau} \right), \\ I = (2\pi)^2 \delta(m\omega_n(\mathbf{Q}) + m'\omega_{n'}(\mathbf{q}')) \delta(\omega + m\omega_n(\mathbf{Q})). \quad (19)$$

The spectrum will then contain peaks at the normal mode frequencies, in similarity to the idealized result for the projected velocities (Eq. (8)). This is also similar to the result derived previously for the velocity auto-correlation.<sup>47</sup> For a finite simulation, the integrals are no longer independent:

$$I = \int_0^T dt e^{i(\omega + m\omega_n(\mathbf{Q}))t} \int_0^{T-t} dt' e^{i(m\omega_n(\mathbf{Q}) + m'\omega_{n'}(\mathbf{q}'))t'}. \quad (20)$$

Note that special consideration is required in evaluating this integral when  $m\omega_n(\mathbf{Q}) + m'\omega_{n'}(\mathbf{q}') = 0$ . Assuming the  $\mathbf{Q}$  projection is complete, the resulting fitting function is ( $C_{nm}^{(\lambda)}(\mathbf{Q})$  for  $\lambda = 1, 2, 3, 4$  are real fitting amplitudes):

$$S_{\mathbf{Q}}^A(\omega) = \sum_n \sum_m \left\{ C_{nm}^{(1)}(\mathbf{Q}) \text{sinc}[(\omega + m\omega_n(\mathbf{Q}))T] \dots \dots + C_{nm}^{(2)}(\mathbf{Q}) \text{sinc}[(\omega + m\omega_n(\mathbf{Q}))T/2] \text{sinc}[(\omega + m\omega_n(\mathbf{Q}))T/2] \dots \dots + C_{nm}^{(3)}(\mathbf{Q}) \text{sinc}^2[(\omega + m\omega_n(\mathbf{Q}))T/2] \dots \dots + C_{nm}^{(4)}(\mathbf{Q}) \frac{\partial \{\text{sinc}[(\omega + m\omega_n(\mathbf{Q}))T]\}}{\partial [(\omega + m\omega_n(\mathbf{Q}))T]} \right\}. \quad (21)$$

Equation (20) can be extended to include the effect of lifetime based on Eq. (13):

$$I = \int_0^T dt e^{[i(\omega + m\omega_n(\mathbf{Q}) - \Gamma_n(\mathbf{Q}))t]} \int_0^{T-t} dt' e^{[i(m\omega_n(\mathbf{Q}) + m'\omega_{n'}(\mathbf{q}')) - (\Gamma_n(\mathbf{Q}) + \Gamma_{n'}(\mathbf{q}'))t']}. \quad (22)$$

The appropriate fitting function can then be shown to be ( $C_{nm}^{(l)}(\mathbf{Q})$  and  $C'_{nm}(\mathbf{Q})$  are real fitting amplitudes, and  $l = \pm 1$ ):

$$S_{\mathbf{Q}}^A(\omega) = \frac{1}{T} \sum_n \sum_m \sum_{l'} \frac{1}{(\omega + m\omega_n(\mathbf{Q}))^2 + (\frac{\Gamma_n(\mathbf{Q})}{2})^2} \dots \dots \{ [C_{nm}^{(l)}(\mathbf{Q})(\omega + m\omega_n(\mathbf{Q})) + C'_{nm}(\mathbf{Q})\Gamma_n(\mathbf{Q})] e^{i\Gamma_n(\mathbf{Q})T} \text{sinc}[(\omega + m\omega_n(\mathbf{Q}))T] \dots \dots + [C_{nm}^{(l)}(\mathbf{Q})\Gamma_n(\mathbf{Q}) - C'_{nm}(\mathbf{Q})(\omega + m\omega_n(\mathbf{Q}))] \{ e^{i\Gamma_n(\mathbf{Q})T} \cos[(\omega + m\omega_n(\mathbf{Q}))T] - 1 \} \}. \quad (23)$$

Note that  $m\omega_n(\mathbf{Q}) + m'\omega_{n'}(\mathbf{q}') = 0$  does not cause any problem when the effect of lifetime is included. When  $\Gamma_n(\mathbf{Q}) = 0$  in Eq. (23), only the  $C_{nm}^{(1)}(\mathbf{Q})$  and  $C_{nm}^{(2)}(\mathbf{Q})$  terms in Eq. (21) will be recovered up to a constant factor. Equations (10), (14), (21) and (23) represent the key results of this paper. To summarize, the assumptions that are required to derive these lineshapes are quite simple: (i) Eq. (4) remains valid with the harmonic oscillation frequencies replaced by their renormalized values, (ii) the phonon lifetimes are included using a damped simple harmonic oscillator model (Eq. (13)), and (iii) the projections in Eqs. (5) and (6) perfectly select the desired phonon wave vector and normal mode. These assumptions are also required to obtain the standard fitting function used in the literature with the additional requirement that the simulation time be infinite (more discussion in the next section).

#### Comments on previous work

In the literature the standard procedure is to calculate spectra with the fast Fourier transform (FFT) and then fit with Lorentzian functions, however many configurations are required to obtain an accurate result.<sup>25</sup> Furthermore, a reasonable question is: Why has the sinc-like lineshape not been resolved for spectra resulting from low-temperature (nearly harmonic) simulations? The reason

is that the grid spacing for the FFT is too coarse:

$$\Delta\omega_{\text{FFT}} = \frac{2\pi}{T} = \frac{1}{2}\Delta\omega_{\text{FW}}. \quad (24)$$

That is, on average the FFT places two points on the central peak and one point on the side peaks of the sinc function. Another question is: Can the Lorentzian lineshape be rigorously justified? It is useful to see what happens when the effect of lifetime is included in the limit of an infinite simulation. This is most easily demonstrated for the projected velocities. To keep the real part of the exponential involving the half-linewidth finite, the time limits are changed to the asymmetric form  $t_i = 0$  and  $t_f = T$ , and only the peak at positive frequency is considered. The integral to evaluate is thus:

$$\lim_{T \rightarrow \infty} \int_0^T dt e^{[i(\omega - \omega_n(\mathbf{q})) - \Gamma_n(\mathbf{q})]t} = -\frac{1}{i(\omega - \omega_n(\mathbf{q})) - \Gamma_n(\mathbf{q})}.$$

If the complex modulus squared is used instead of the real part, then the spectral lineshape is

$$S_{\eta\mathbf{q}}(\omega) \propto \frac{1}{(\omega - \omega_n(\mathbf{Q}))^2 + (\frac{2\Gamma_n(\mathbf{Q})}{2})^2}. \quad (25)$$

This is indeed a Lorentzian peaked at  $\omega = \omega_n(\mathbf{Q})$  with  $\Delta\omega_{\text{FWHM}} = 2\Gamma_n(\mathbf{Q})$ . A similar result holds for the velocity auto-correlation approach. Therefore, the fitting function of the standard procedure is only valid for long simulation times. Also, part of the width of the Lorentzian will be due to the finite simulation time, which can be estimated by making a comparison between the Lorentzian and sinc functions using their limiting behavior to the Dirac delta function:

$$\delta(x) = \lim_{\epsilon \rightarrow 0} \frac{1}{\pi} \frac{\epsilon}{x^2 + \epsilon^2} = \lim_{\epsilon \rightarrow 0} \frac{1}{\pi\epsilon} \text{sinc}\left(\frac{x}{\epsilon}\right). \quad (26)$$

The FWHM of the Lorentzian (first term) is

$$\text{FWHM}_{\text{Lorentzian}} = 2\epsilon.$$

The full-width of the central peak as defined by the zeros of sinc (second term) is

$$\text{FW}_{\text{sinc}} = 2\pi\epsilon.$$

Combining these results implies:

$$\text{FWHM}_{\text{Lorentzian}} = \frac{\text{FW}_{\text{sinc}}}{\pi}. \quad (27)$$

Next, using Eq. (11) we obtain an important result with a quantitative approximation for the contribution to the FWHM of a Lorentzian due to the finite simulation time:

$$\Delta\omega_{\text{FWHM}} \approx \frac{4}{T}. \quad (28)$$

#### Illustration for a Toy model

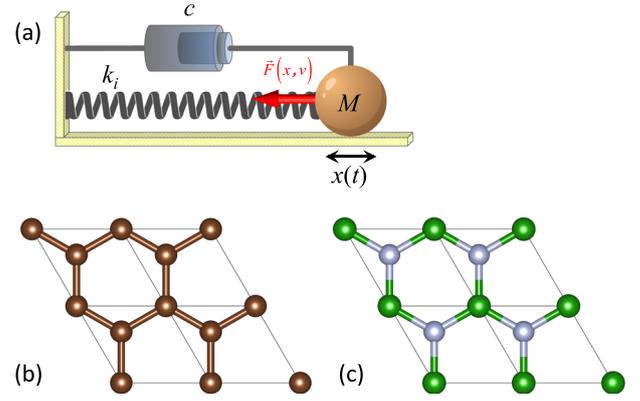
We now illustrate the formalism developed in this paper (specifically, Eqs. (10), (14), (21) and (23)) on a toy model, see Fig. 1a, which consists of a classical particle in one dimension subject to a force given by

$$F(x, v) = -k_1x - k_2x^2 - k_3x^3 - cv. \quad (29)$$

Here  $x$  is the position,  $v$  is the velocity, the  $k$ 's are spring constants, and  $c$  is a damping constant. Without the damping term this corresponds to a potential:

$$V(x) = \frac{1}{2}k_1x^2 + \frac{1}{3}k_2x^3 + \frac{1}{4}k_3x^4. \quad (30)$$

Explicitly, the acceleration  $a$  and jerk  $j$  are ( $M$  is the mass of the



**Fig. 1** Three systems considered as examples of the developed method. **a** Sketch of the toy model where a classical particle of mass  $M$  in one dimension is subjected to an external force.  $x$ ,  $k_i$  and  $c$  correspond to the position of the particle, spring constants, and damping constant, respectively. Top view of **b** graphene and **c** single-layer hexagonal boron nitride (hBN)<sup>59</sup>

$M = 104.5 \text{ eV}$ $\text{fs}^2 \text{ \AA}^{-2}$	$E_0 = 0.025 \text{ eV}$	$\nu_H = \sqrt{k_1/M}/(2\pi) =$ $100 \text{ cm}^{-1}$	$k_2 = 0.25 \text{ eV}$ $\text{ \AA}^{-3}$
$k_3 = -0.1 \text{ eV}$ $\text{ \AA}^{-4}$	$c = 0.01 \text{ eV fs}$ $\text{ \AA}^{-2}$	$\Gamma = c/(2M) = 1.60 \text{ cm}^{-1}$	$\Delta t = 0.5 \text{ fs}$

$M$  is the mass,  $E_0$  is the total initial energy,  $\nu_H$  is the harmonic vibrational frequency,  $k_2$  and  $k_3$  are the anharmonic constants in Eq. (29),  $c$  is the damping constant, and  $\Delta t$  is the time-step

particle):

$$a = -\frac{1}{M}(k_1x + k_2x^2 + k_3x^3 + cv), \quad (31)$$

$$j = \frac{da}{dt} = -\frac{1}{M}[(k_1 + 2k_2x + 3k_3x^2)v + ca]. \quad (32)$$

The system is propagated on the basis of a truncated Taylor expansion ( $\Delta t$  is the time-step and  $n = 1, 2, \dots$ ):

$$x(t_{n+1}) = x(t_n) + v(t_n)\Delta t + \frac{1}{2}a(t_n)(\Delta t)^2 + \frac{1}{6}j(t_n)(\Delta t)^3, \quad (33)$$

$$v(t_{n+1}) = v(t_n) + a(t_n)\Delta t + \frac{1}{2}j(t_n)(\Delta t)^2. \quad (34)$$

The total energy of the system is monitored and compared to its initial value to validate the choice of time-step ( $n \geq 2$ ):

$$E(t_n) = \frac{1}{2}M(v(t_n))^2 + V(x(t_n)) \dots \\ \dots + \sum_{n'=2}^n c \left( \frac{v(t_{n'-1}) + v(t_{n'})}{2} \right) (x(t_{n'}) - x(t_{n'-1})). \quad (35)$$

In order, these terms correspond to the kinetic energy, potential energy, and the cumulative drag energy. The simulation parameters are reported in Table 1. The mass is equivalent to that of a hydrogen atom, the total initial energy is the thermal energy  $k_B T$  at room temperature, and the time-step is the same as that used in our DFT calculations. For reference, for silicon and black phosphorus the vibrational frequencies are less than about  $600 \text{ cm}^{-1}$ , and the frequency shifts and linewidths are no more than about  $6 \text{ cm}^{-1}$  at room temperature.<sup>25,48</sup> The initial position is the amplitude for  $x < 0$  and is found by solving  $V(x) = E_0$ , with  $E_0 = E(t_1)$  as the specified initial energy. Therefore, the initial velocity is set to zero. Without the jerk term the percent deviation of the

total energy from its initial value is as much as  $\sim 10^3\%$ ; however, the inclusion of the jerk term reduced this to  $\sim 10^{-3}\%$ .

Combinations of harmonic/anharmonic and no damping/damping are considered for the derived projection and auto-correlation lineshapes, as well as the standard Lorentzians: (i) harmonic, no damping; (ii) harmonic, damping; and (iii) anharmonic, no damping. The anharmonic with damping case is not considered, since no general analytical solution exists to the equation of motion. The Lorentzians are fit to the modulus squared of the FFT ( $|\text{FFT}|^2$ ), while the derived lineshapes are fit to the real part of the Fourier transform ( $\text{Re}\{\text{FT}\}$ ) calculated on a fine non-uniform frequency grid. To construct the grid, the peak positions are estimated using the relative maxima of  $|\text{FFT}|^2$  and then a grid spacing of  $\Delta\omega_{\text{FW}}/50$  is set for  $\pm 5\Delta\omega_{\text{FW}}$  about these estimated positions (see Eq. (11)). Outside of these regions the grid spacing is increased by a factor of five. The convergence of the fitted vibrational frequencies and  $\Gamma$  values (damping only) with the number of configurations is considered. The configuration grid consists of values ranging from 11 to 10,001 (increment of 10) and 10,001 to 20,001 (increment of 100). The only exception is for harmonic with damping (auto-correlation only) for which the values range from 11 to 10,001 (increment of 10), 10,001–100,001 (increment of 100), and 100,001–200,001 (increment of 500). Only odd values are used to enable symmetry about zero time when necessary. The parameters extracted from the fitting procedure are averaged over 30 consecutive configuration grid points to prevent outliers from defining the convergence. Therefore, on the horizontal axis of the plots to follow in the main text and SI *maximum number of configurations* refers to the largest of the configuration grid values in a given average. Vibrational frequencies and  $\Gamma$  values are considered converged when the current and all subsequent averaged values are within the exact numbers by at least  $0.5 \text{ cm}^{-1}$ . Note that in the discussion and figures that follow *frequency* refers to the vibrational frequency  $\nu$ , not the angular frequency  $\omega$ .

**Toy model: harmonic, no damping**

Harmonic with no damping means that  $k_2$ ,  $k_3$ , and  $c$  are zero. The exact analytical solution takes the form:

$$v(t) = v_+ e^{i\omega_H t} + v_- e^{-i\omega_H t}. \quad (36)$$

This is a special case of Eq. (4) with one mode along one dimension, which implies that the derived harmonic with no damping lineshapes will apply here. The frequency convergence and sample spectra for the projection lineshape are illustrated in Figs. S1 and S2, respectively. The corresponding information for the auto-correlation lineshape is shown in Figs. S3 and S4. Our analysis indicates that using the proposed method reduces the required number of configurations by factors of about 10 (projection) and 29 (auto-correlation) for the vibrational frequency.

**Toy model: harmonic, damping**

Harmonic with damping means that  $k_2$  and  $k_3$  are set to zero. The exact analytical solution takes the form:

$$v(t) = v_+ e^{(i\omega_D - \Gamma)t} + v_- e^{(-i\omega_D - \Gamma)t}. \quad (37)$$

This is a special case of Eq. (4) with the inclusion of Eq. (13), which implies that the derived harmonic with damping lineshapes will apply here. The fitting procedure will extract a modified frequency:

$$\nu_D = \omega_D / (2\pi) = \sqrt{\omega_H^2 - \Gamma^2} / (2\pi) \approx 99.9997 \text{ cm}^{-1}.$$

Since  $\omega_H \gg \Gamma$ ,  $\nu_D \approx \nu_H$ . The frequency/damping constant convergence and sample spectra for the projection lineshape are illustrated in Figs. S5 and S6, respectively. The corresponding information for the auto-correlation lineshape is shown in Figs. S7

and S8. Using the new method reduces the required number of configurations by factors of about 10 (projection) and 26 (auto-correlation) for the vibrational frequency, and factors of about 12 (projection) and 99 (auto-correlation) for  $\Gamma$ .

**Toy model: anharmonic, no damping**

Anharmonic with no damping means that  $c$  is set to zero. The exact frequency for determining the convergence can be calculated starting from energy conservation:

$$\frac{1}{2} M \left( \frac{dx}{dt} \right)^2 + V(x) = E_0, \quad (38)$$

$$\nu_A = \left( \sqrt{2M} \int_{A_-}^{A_+} \frac{dx}{\sqrt{E_0 - V(x)}} \right)^{-1} \approx 86.8 \text{ cm}^{-1}.$$

The integration limits  $A_-$  and  $A_+$  are the amplitudes at negative and positive positions, respectively, as determined by solving  $V(x) = E_0$ . The anharmonicity causes a significant downshift in frequency ( $100 \text{ cm}^{-1} \rightarrow 86.8 \text{ cm}^{-1}$ ). Since the motion remains periodic, the solution for the position can be written as a Fourier series ( $\omega_A = 2\pi\nu_A$ ):

$$x(t) = \sum_{m=-\infty}^{\infty} a_m e^{im\omega_A t}. \quad (39)$$

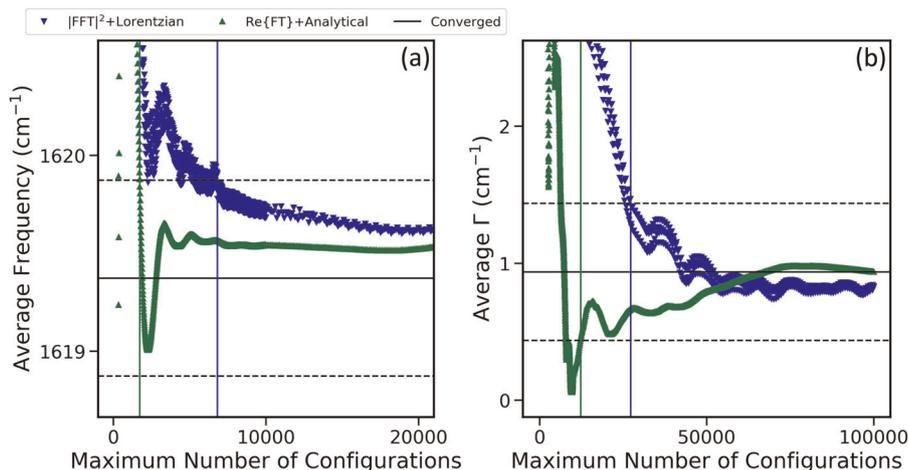
The constant term for  $m = 0$  in the Fourier series will not be zero here, since there is an odd term in the potential ( $k_3 \neq 0$ ). Taking a Fourier transform will then place a peak at zero as well as the vibrational frequency (and its other integral multiples), which complicates the analysis. Taking the derivative to get the velocity removes this effect and suggests why the velocities should be used instead of the displacements from equilibrium. The frequency convergence and sample spectra for the projection lineshape are illustrated in Figs. S9 and S10, respectively. The corresponding information for the auto-correlation lineshape is shown in Fig. S11. No additional interesting information is obtained from the spectra. Using the new method reduces the required number of configurations by factors of about 8 (projection) and 10 (auto-correlation) for the vibrational frequency.

**Comment on toy model**

For a given set of toy model parameters, the projection and auto-correlation approaches for the developed formalism converge in a similar number of configurations. However, for the standard procedure based on  $|\text{FFT}|^2$  and a Lorentzian fit the auto-correlation technique takes much longer to converge than the projection method. A possible explanation is that while the Lorentzian is the exact lineshape for both methodologies in the infinite simulation time limit, it is a worse approximation to the exact auto-correlation lineshape for finite simulation time as compared to the corresponding projection case.

**Application to real materials**

With the developed theory illustrated for a toy model, we now show its application to three materials: graphene, single-layer hBN (see Fig. 1b, c), and silicon. Ab initio molecular dynamics (AIMD) simulations were carried out for a supercell of each material and more than 100,000 velocity configurations were obtained (see “Methods” section for computational details). The effective velocities are calculated using Eq. (5) with projection at  $\mathbf{Q} = 0$  and the velocity auto-correlation is evaluated with Eq. (15). For illustration purposes only the BZ center is analyzed, which is relevant to single-phonon Raman spectroscopy.<sup>9,49,50</sup> For graphene and hBN only the in-plane components of the velocity auto-correlation are used to select the doubly degenerate in-plane optical modes at  $\mathbf{Q} = 0$ . For silicon the analysis is simple since all



**Fig. 2** Convergence properties for graphene treated using ab initio molecular dynamics with a temperature of around 50 K and the velocity auto-correlation method. The  $E_{2g}$  symmetry modes are isolated at the  $\Gamma$ -point by projections. The convergence is studied using two approaches, which are denoted by “|FFT|<sup>2</sup> + Lorentzian” and “Re{FT} + Analytical”. In both cases, the effective velocities at the  $\Gamma$ -point were obtained using Eq. (5) and the velocity auto-correlation was subsequently calculated with Eq. (15). In the “|FFT|<sup>2</sup> + Lorentzian” method, which is the standard procedure and assumes infinite simulation time, the spectrum is calculated as the modulus squared of the fast Fourier transform of the velocity auto-correlation and the resulting peaks are subsequently fit using Lorentzian functions (Eq. (25)). In the “Re{FT} + Analytical” method, which is the new approach and is valid at all simulation times based on our analytical derivations, the spectra are evaluated with Eq. (16) and the vibrational frequencies and half-linewidths are extracted by fitting with Eq. (23). **a** Convergence of the vibrational frequency with the number of configurations ( $=N_C$ ). Convergence is indicated by the vertical lines with  $N_C = 14,501$  for the Lorentzian fit and  $N_C = 2,461$  for the analytical model, with improvement by a factor of about 5.9. **b** Convergence of the half-linewidth  $\Gamma$  with the number of configurations. Convergence is indicated by the vertical lines with  $N_C = 27,301$  for the Lorentzian fit and  $N_C = 12,301$  for the analytical model, with improvement by a factor of about 2.2. The developed method results in a significant reduction in the required simulation time to obtain converged properties

optical modes are triply degenerate at  $\mathbf{Q} = 0$ . The spectra are calculated using both the standard |FFT|<sup>2</sup> approach and Eq. (16), and are correspondingly fit using Eqs. (25) and (23). To construct the non-uniform frequency grid for the new method, the peak positions are estimated using the relative maxima of |FFT|<sup>2</sup> and then a grid spacing of  $\Delta\omega_{FW}/50$  is set for  $\pm 5\Delta\omega_{FW}$  about these estimated positions (see Eq. (11)). The values of the configuration grid are from 11 to 10,001 (increment of 10) and 10,001 to 100,001 or 125,001 (increment of 100), and the convergence tolerance is still set to  $\pm 0.5 \text{ cm}^{-1}$ .

### Graphene

The vibrational frequency and  $\Gamma$  convergence rates are illustrated in Fig. 2 and sample spectra are shown in Fig. 3 for single-layer graphene at a temperature of 50 K. Details are provided in the captions of the figures.

Classical dynamics based on parameterized force-fields can be used as an alternative to AIMD, especially for very large systems.<sup>21</sup> In this case, it is easier to use larger supercells to include many more phonon wave vectors. To that end we carried out MD on a  $(20 \times 20)$  supercell of graphene at around 50 K using four common potentials (see “Methods” section for computational details) and plotted the spectra after 50,001 time steps (see Fig. 4). The fitted vibrational frequencies are close to the values obtained in a previous comprehensive study using very long simulation times and a detailed statistical analysis.<sup>51</sup> In general the analytical fitting function does not match the calculated spectra very well. If the potentials were producing accurate trajectories, the spectra in Fig. 4 would all match the spectrum in Fig. 3d obtained from AIMD for which the analytical fit exactly matches the calculated spectrum. This is clearly not the case, so we can conclude that the parameterized force-fields produce inaccurate trajectories for oscillations close to equilibrium despite their outstanding utility for other applications. Slightly inaccurate trajectories could also explain the asymmetries for the AIMD spectra in Fig. 3f, although

to a significantly smaller extent. In general the spectra in Fig. 4 are much broader than expected, which the analytical fitting function superficially attempts to recreate by using large  $\Gamma$  values. This observation is consistent with previous work, where the half-linewidth values predicted for silicon with the Tersoff potential disagreed with those obtained from AIMD.<sup>25</sup>

### Hexagonal boron nitride

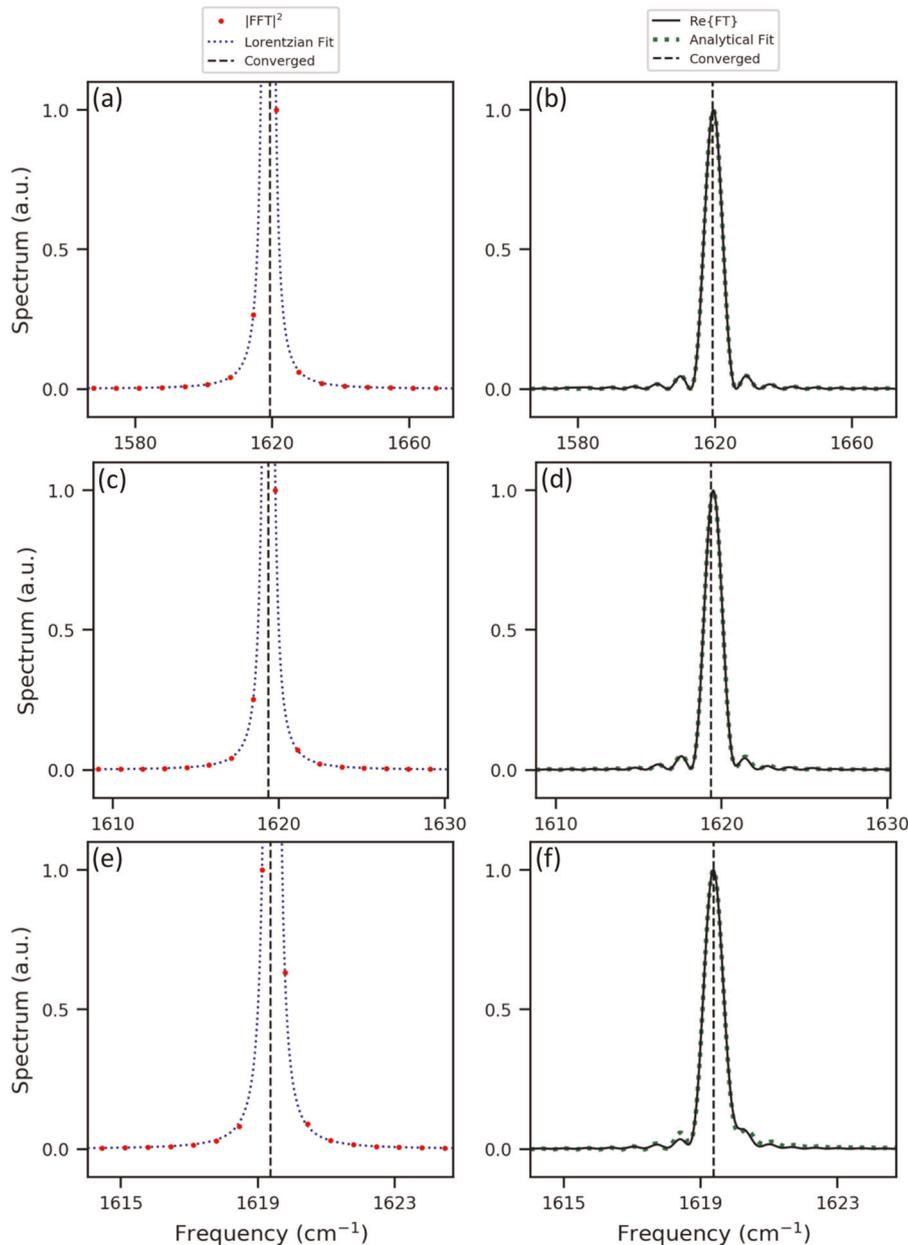
The vibrational frequency and  $\Gamma$  convergence rates are illustrated in Fig. 5. Details are provided in the caption of the figure. Sample spectra are not shown since they are very similar to those of graphene (Fig. 3).

### Silicon

The vibrational frequency and  $\Gamma$  convergence are illustrated in Fig. S12 and sample spectra are shown in Fig. S13 for silicon at a temperature of 50 K. There is almost no difference between the two methods for the convergence of the vibrational frequency; however, the convergence of the half-linewidth is improved by a factor of about 1.4 by using the new approach. Corresponding results are shown in Figs. 6 and 7 for room temperature (300 K) where the anharmonicity is significantly stronger.

## DISCUSSION

Up until this point we have demonstrated that the developed method improves the convergence of vibrational frequencies and lifetimes with respect to the total simulation time. There are three simple points that explain this observation: (i) the derived lineshapes (Eqs. (10), (14), (21) and (23)) explicitly include the simulation time, (ii) the Lorentzian function used in the standard fitting procedure is only exact in the infinite simulation limit, and (iii) the FFT poorly samples each peak, only placing a few points on the dominant part of the peak, whereas in the developed method the Fourier transform is calculated on a much finer grid. The first

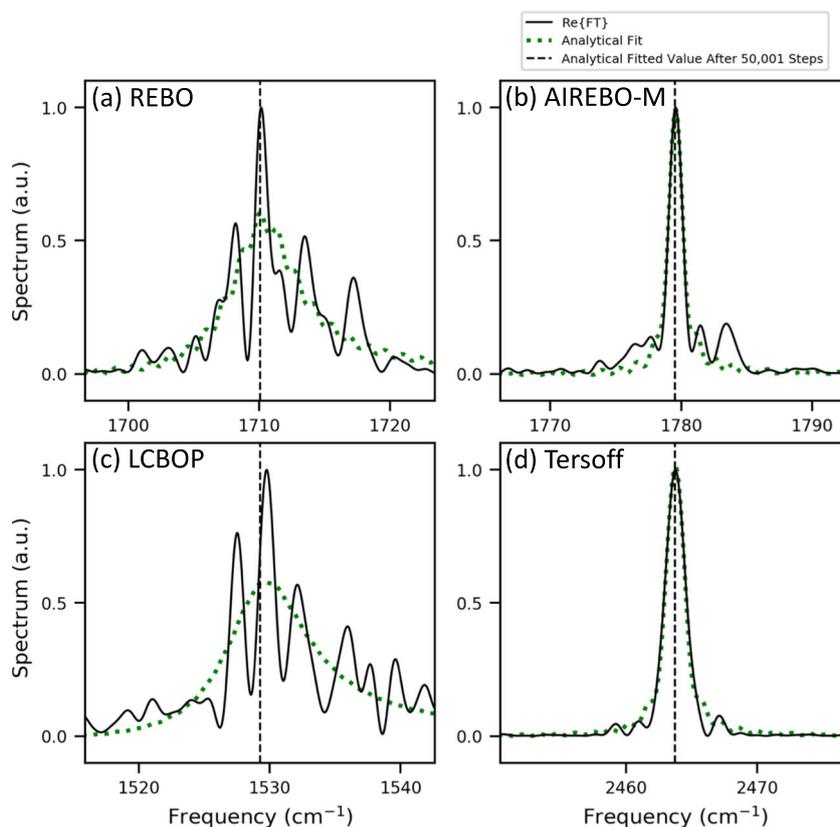


**Fig. 3** Sample spectra obtained from ab initio molecular dynamics applied to graphene (based on Fig. 2). **a** and **b** 10,001 configurations. **c** and **d** 50,001 configurations. **e** and **f** 100,001 configurations. The derived fitting function only deviates from the calculated spectrum slightly even after 100,001 configurations

two points are more important for the lifetime, since the simulation time and the lifetime both contribute to the width of a peak. The two methods can be evaluated based on more than the number of configurations to reach convergence. In general, the new method produces smooth curves, while the standard procedure results in rapid oscillations (this is also true for the toy model, see SI).

The developed method performs better for the toy model than the real materials, which can be explained by several points. Since there is only one vibrational mode and one phonon wavevector in the toy model, the system is initially thermalized, i.e. all phonon modes follow the equilibrium distribution. However, in the case of graphene, hBN, and silicon there are multiple modes and phonon wavevectors included, so the random initialization of the velocities does not result in the equilibrium distribution immediately. Since the anharmonicity is weak, the transfer of energy between the

phonon modes to reach the equilibrium distribution takes some critical time which is related to the phonon lifetimes. Furthermore, the distortion of the peaks in Figs. 3f, S13f, and 7f indicates at least a partial breakdown of the initial assumptions. These are to be contrasted with the corresponding toy model spectra based on the auto-correlation in the right column of Fig. S8, which are symmetric and have the fitting function matching the simulation data exactly. The anharmonicity-induced coupling between phonons partially invalidates the simplicity of the superposition in Eq. (4). Lastly, the inclusion of the phonon lifetime by a simple exponential decay (Eq. (13)) represents a continuous dissipation of energy from the mode during the simulation. This differs from the actual thermalization process, which is a continuous exchange of energy between different modes due to their interactions, with the total energy being constant. This simple assumption is



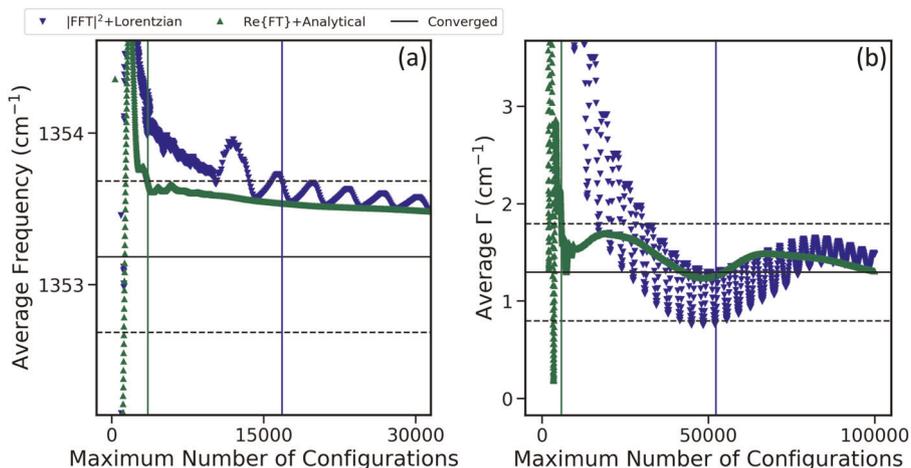
**Fig. 4** Sample spectra obtained for graphene from parameterized force-field molecular dynamics at 50 K and the velocity auto-correlation method. The  $E_{2g}$  symmetry modes are isolated at the  $\Gamma$ -point by projections. All spectra are calculated using 50,001 configurations. **a** REBO<sup>60</sup>, **b** AIREBO-M<sup>61</sup>, **c** LCBOP<sup>62</sup>, and **d** Tersoff<sup>63</sup>. Together, the deviation of the spectral lineshapes from that of the corresponding AIMD result (Fig. 3d) and the inability of the analytical model to fit the data suggest that the potentials create inaccurate trajectories for oscillations close to equilibrium

therefore perhaps only valid over short timescales, which may be difficult to treat due to the simulation broadening.

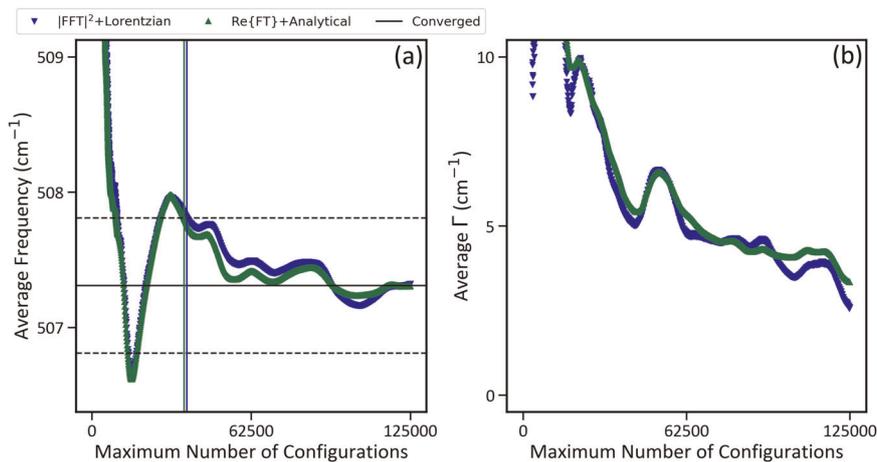
Although silicon is often featured in elementary textbooks as a standard example, it is evident that the phonon anharmonicity is strong even at 50 K (Fig. S13) in comparing to the spectra from graphene (Fig. 3) and hBN, which may in part be due to the difference in dimensionality. More specifically, the 50 K silicon spectrum broadens more than the fitting lineshape anticipates as the simulation evolves and a satellite peak at higher frequency than the central peak emerges, both of which contribute to the fact that the two fitting methods converge similarly for the vibrational frequency. At 300 K the phonon-phonon coupling is much stronger, causing the fitting functions to strongly deviate from the raw spectrum data. Therefore, the lifetime cannot be extracted reliably (Fig. 7). It is important to emphasize that when it is not possible to obtain accurate results from the developed method, the standard procedure will not produce meaningful results either since the underlying physical assumptions are the same except that the latter further imposes that the simulation time be infinite. Future work can address these issues, although the improvement with our new method is already striking for graphene and hBN.

To summarize, a new approach to understand vibrational lineshapes was derived for spectra calculated from MD and include the effect of anharmonicity-induced frequency shift and lifetime, as well as the simulation broadening (Eqs. (10), (14), (21) and (23)). For a toy model we demonstrated at least an order of magnitude reduction in the number of simulation steps to obtain converged vibrational properties in nearly all cases considered as compared to the standard extraction procedure. The theory was

also illustrated for graphene, hBN, and silicon with the dynamics carried out at the DFT level. For graphene the vibrational frequency and lifetime convergence were improved by factors of 5.9 and 2.2, and for hBN the corresponding factors were 4.7 and 8.9 (both materials at 50 K). For silicon the anharmonicity results in spectral lineshapes that deviate from the theoretically predicted ones, which is more extreme at higher-temperature. At 50 K the lifetime can be converged using both methods, with the new method reducing the required simulation time by a factor of about 1.4. However, at 300 K the deviation of the calculated spectra from the theoretically predicted lineshape is too extreme and claiming accurate values for the lifetimes from either method is not possible. Further analytical work will be required to accurately describe cases where the anharmonicity is moderate. Considering the simplicity of the initial assumptions, only resolving the spectrum on a fine grid and including the simulation time in the lineshapes has significantly reduced the computational effort to obtain temperature-dependent phonon properties from AIMD in the regime where the anharmonicity is sufficiently weak and results in well-defined renormalized phonon quasiparticles. In all cases, adoption of the developed method is encouraged, since, by design, it will never be outperformed by the conventional method. Our extension of signal analysis to material vibrations represents a state-of-the-art advance in calculating temperature-dependent phonon properties. Application of the proposed approach has the potential to be far reaching as the theory applies equally well to ab initio MD based on DFT, time-dependent DFT dynamics, and parameterized force-fields and is thus expected to have important impact on topics ranging from strongly correlated materials with sophisticated treatment of



**Fig. 5** Convergence properties for hexagonal boron nitride treated using ab initio molecular dynamics with a temperature of around 50 K and the velocity auto-correlation method. The  $E_{2g}$  symmetry modes are isolated at the  $\Gamma$ -point by projections. **a** Convergence of the vibrational frequency with the number of configurations ( $=N_C$ ). Convergence is indicated by the vertical lines with  $N_C = 16,801$  for the Lorentzian fit and  $N_C = 3561$  for the analytical model, with improvement by a factor of about 4.7. **b** Convergence of the half-linewidth  $\Gamma$  with the number of configurations. Convergence is indicated by the vertical lines with  $N_C = 52,301$  for the Lorentzian fit and  $N_C = 5851$  for the analytical model, with improvement by a factor of about 8.9. The developed method results in a significant reduction in the required simulation time to obtain converged properties



**Fig. 6** Convergence properties for silicon treated using ab initio molecular dynamics with a temperature of around 300 K and the velocity auto-correlation method. The triply degenerate optical modes are isolated at the  $\Gamma$ -point by projections. **a** Convergence of the vibrational frequency with the number of configurations ( $=N_C$ ), which is indicated by the vertical lines (almost no difference between the methods). **b** Convergence of the half-linewidth  $\Gamma$  with the number of configurations is not obtained due to the significant distortion of the spectral lineshapes from their theoretically predicted forms (see next figure)

electron–electron interactions to biological systems (large systems precluding long simulation times).

## METHODS

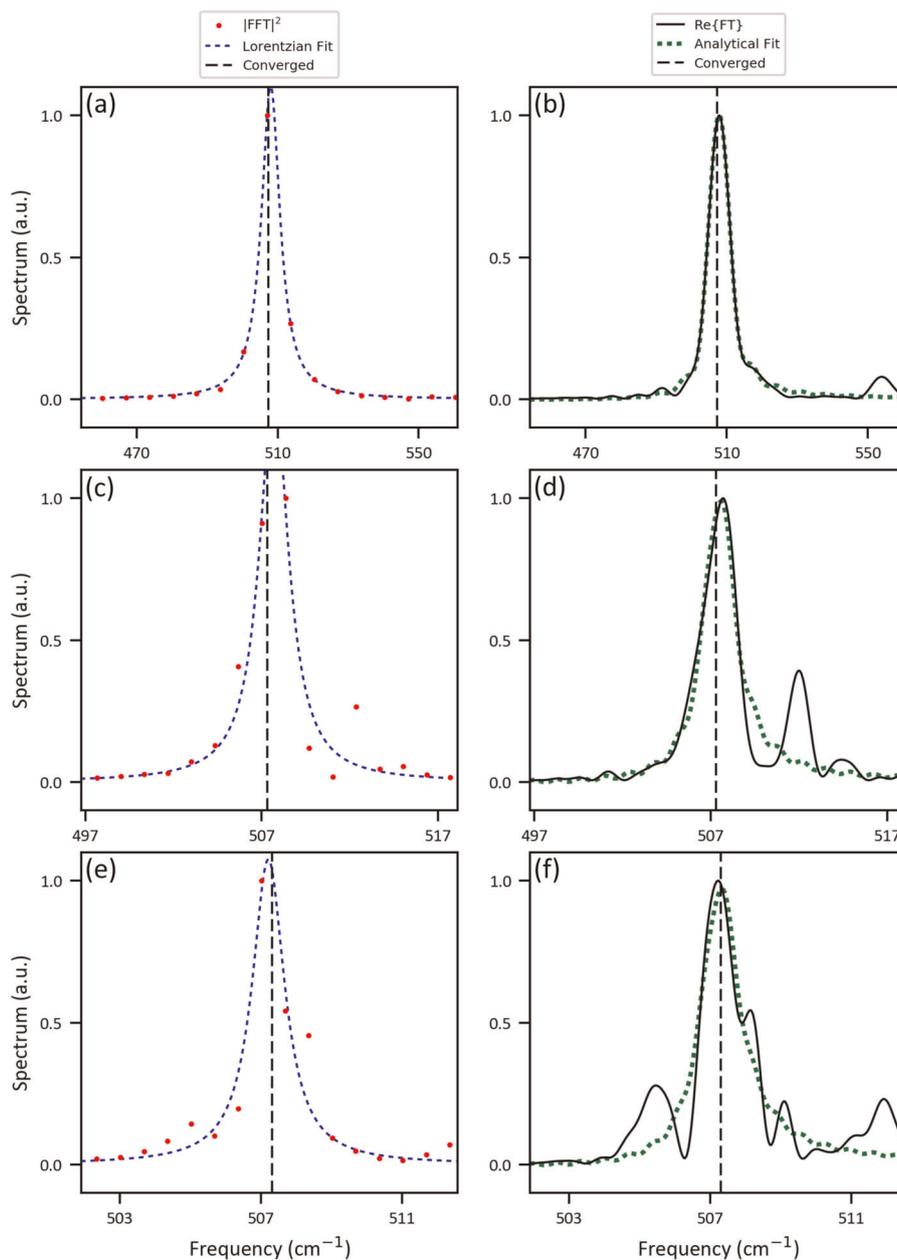
VASP was used for plane-wave DFT calculations<sup>52,53</sup> with the PBE exchange correlation functional.<sup>54</sup> The projector augmented wave pseudopotentials<sup>55</sup> were used with an energy cutoff of 500 eV. The tolerance for the electronic loop was set to  $10^{-5}$  eV. Single-layer graphene and hBN ( $4 \times 4$ ) supercells with a 15 Å vacuum region and a silicon ( $3 \times 3 \times 3$ ) supercell were each treated using a  $\Gamma$ -point  $k$ -point sampling and a Gaussian smearing of 0.01 eV. The lattice and ions were first relaxed to a force cutoff of 0.01 eV/Å. The ab initio MD<sup>56,57</sup> were carried out with a time-step of 0.5 fs. No thermostat was used to prevent systematic deviations from the desired Newtonian trajectories implicit in the initial assumptions for the analytical derivations. The dynamics were initialized with all atoms starting at their equilibrium positions and with velocities randomly assigned so that

the initial temperature was 100 K (600 K), which quickly stabilizes to an average temperature of around 50 K (300 K). No time-steps were discarded and more than 100,000 velocity configurations were obtained.

The MD with parameterized force-fields were carried out using LAMMPS.<sup>58</sup> A ( $20 \times 20$ ) supercell of graphene was used, and the lattice and ions were first relaxed to a force cutoff of 0.01 eV/Å. The dynamics were carried out with a time-step of 0.05 fs only saving the trajectory every tenth step. The smaller time-step is necessary since VASP uses a fourth-order predictor-corrector method, while LAMMPS uses a velocity-Verlet integrator. No thermostat was used and the dynamics were initialized in the same manner as VASP so that an average temperature of 50 K was maintained (although the fluctuations were smaller due to the larger supercell). No time-steps were discarded and more than 50,000 velocity configurations were obtained.

## DATA AVAILABILITY

Any data used to generate the results in this study can be obtained from the corresponding author upon reasonable request.



**Fig. 7** Sample spectra obtained from ab initio molecular dynamics applied to silicon at 300 K (based on Fig. 6). **a** and **b** 10,001 configurations. **c** and **d** 50,001 configurations. **e** and **f** 100,001 configurations. At high-temperature the anharmonicity causes strong deviations from the theoretically predicted lineshape. Therefore, the extracted half-linewidth values obtained from either method cannot be expected to be accurate

#### CODE AVAILABILITY

The spectral analysis code can be obtained from the corresponding author upon reasonable request.

#### ACKNOWLEDGEMENTS

Part of this work was performed using supercomputing resources provided by the Center for Computational Innovations (CCI) at Rensselaer Polytechnic Institute. A.C. was supported by NSF Grant EFRI 2-DARE (EFRI-1542707) and by a DOE SCGSR award. D.T. acknowledges support from the Office of Naval Research. We thank Bobby G. Sumpter for a critical reading of the project proposal. "This material is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Workforce Development for Teachers and Scientists, Office of Science Graduate Student Research (SCGSR) program. The SCGSR program is administered by the Oak

Ridge Institute for Science and Education (ORISE) for the DOE. ORISE is managed by ORAU under contract number DE-SC0014664. All opinions expressed in this paper are the authors and do not necessarily reflect the policies and views of DOE, ORAU, or ORISE."

#### AUTHOR CONTRIBUTIONS

A.C., D.T., and V.M. derived the spectral lineshapes, wrote the spectral analysis code, performed the DFT calculations, and produced the convergence studies for the toy model, graphene, hBN, and silicon. K.R. and V.M. carried out the MD calculations with parameterized force-fields for graphene. The manuscript was written by A.C. with contributions from all authors.

## ADDITIONAL INFORMATION

**Supplementary information** accompanies the paper on the *npj Computational Materials* website (<https://doi.org/10.1038/s41524-019-0220-1>).

**Competing interests:** The authors declare no competing interests.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## REFERENCES

- Kittel, C. *Introduction to Solid State Physics*. 8 edn (John Wiley & Sons, Hoboken, NJ, USA, 2005).
- Ashcroft, N. W. & Mermin, N. D. *Solid State Physics*. (Cengage Learning, Boston, MA, USA, 1976).
- Togo, A. & Tanaka, I. First principles phonon calculations in materials science. *Scr. Mater.* **108**, 1 (2015).
- Tristant, D., Cupo, A. & Meunier, V. Finite temperature stability of single-layer black and blue phosphorus adsorbed on Au (111): a first-principles study. *2D Mater.* **5**, 035044 (2018).
- Maradudin, A. & Fein, A. Scattering of neutrons by an anharmonic crystal. *Phys. Rev.* **128**, 2589 (1962).
- Cowley, R. Raman scattering from crystals of the diamond structure. *J. Phys.* **26**, 659 (1965).
- Klemens, P. Anharmonic decay of optical phonons. *Phys. Rev.* **148**, 845 (1966).
- Hart, T., Aggarwal, R. & Lax, B. Temperature dependence of Raman scattering in silicon. *Phys. Rev. B* **1**, 638 (1970).
- Balkanski, M., Wallis, R. & Haro, E. Anharmonic effects in light scattering due to optical phonons in silicon. *Phys. Rev. B* **28**, 1928 (1983).
- Vanderbilt, D., Louie, S. G. & Cohen, M. L. Calculation of anharmonic phonon couplings in C, Si, and Ge. *Phys. Rev. B* **33**, 8740 (1986).
- Haro, E., Balkanski, M., Wallis, R. & Wanser, K. Theory of the anharmonic damping and shift of the Raman mode in silicon. *Phys. Rev. B* **34**, 5358 (1986).
- Narasimhan, S. & Vanderbilt, D. Anharmonic self-energies of phonons in silicon. *Phys. Rev. B* **43**, 4541 (1991).
- Togo, A., Chaput, L. & Tanaka, I. Distributions of phonon lifetimes in Brillouin zones. *Phys. Rev. B* **91**, 094306 (2015).
- Gonze, X. Adiabatic density-functional perturbation theory. *Phys. Rev. A* **52**, 1096 (1995).
- Debernardi, A., Baroni, S. & Molinari, E. Anharmonic phonon lifetimes in semiconductors from density-functional perturbation theory. *Phys. Rev. Lett.* **75**, 1819 (1995).
- Baroni, S., de Gironcoli, S., Dal Corso, A. & Giannozzi, P. Phonons and related crystal properties from density-functional perturbation theory. *Rev. Mod. Phys.* **73**, 515 (2001).
- Paulatto, L., Mauri, F. & Lazzeri, M. Anharmonic properties from a generalized third-order *ab initio* approach: Theory and applications to graphite and graphene. *Phys. Rev. B* **87**, 214303 (2013).
- Liu, H. et al. Phosphorene: an unexplored 2D semiconductor with a high hole mobility. *ACS Nano* **8**, 4033 (2014).
- Dickey, J. & Paskin, A. Computer simulation of the lattice dynamics of solids. *Phys. Rev.* **188**, 1407 (1969).
- Car, R. & Parrinello, M. Unified approach for molecular dynamics and density-functional theory. *Phys. Rev. Lett.* **55**, 2471 (1985).
- Senftle, T. P. et al. The ReaxFF reactive force-field: development, applications and future directions. *npj Comput. Mater.* **2**, 15011 (2016).
- Wang, C., Chan, C. & Ho, K. Tight-binding molecular-dynamics study of phonon anharmonic effects in silicon and diamond. *Phys. Rev. B* **42**, 11276 (1990).
- Lan, T., Tang, X. & Fultz, B. Phonon anharmonicity of rutile TiO<sub>2</sub> studied by Raman spectrometry and molecular dynamics simulations. *Phys. Rev. B* **85**, 094305 (2012).
- Zhang, D.-B., Sun, T. & Wentzcovitch, R. M. Phonon quasiparticles and anharmonic free energy in complex systems. *Phys. Rev. Lett.* **112**, 058501 (2014).
- Carreras, A., Togo, A. & Tanaka, I. DynaPhoPy: a code for extracting phonon quasiparticles from molecular dynamics simulations. *Comput. Phys. Commun.* **221**, 221 (2017).
- Raj, A. & Eapen, J. Computing phonon dispersion using fast zero-point correlations of conjugate variables. *MRS Adv.* **3**, 531 (2018).
- Lee, C., Vanderbilt, D., Laasonen, K., Car, R. & Parrinello, M. *Ab initio* studies on the structural and dynamical properties of ice. *Phys. Rev. B* **47**, 4863 (1993).
- Putrino, A. & Parrinello, M. Anharmonic Raman spectra in high-pressure ice from *ab initio* simulations. *Phys. Rev. Lett.* **88**, 176401 (2002).
- Thomas, J. A., Turney, J. E., Iutzi, R. M., Amon, C. H. & McGaughey, A. J. H. Predicting phonon dispersion relations and lifetimes from the spectral energy density. *Phys. Rev. B* **81**, 081411 (2010).
- Hellman, O., Abrikosov, I. & Simak, S. Lattice dynamics of anharmonic solids from first principles. *Phys. Rev. B* **84**, 180301 (2011).
- Hellman, O., Steneteg, P., Abrikosov, I. A. & Simak, S. I. Temperature dependent effective potential method for accurate free energy calculations of solids. *Phys. Rev. B* **87**, 104111 (2013).
- Hellman, O. & Abrikosov, I. A. Temperature-dependent effective third-order interatomic force constants from first principles. *Phys. Rev. B* **88**, 144301 (2013).
- Tadano, T., Gohda, Y. & Tsuneyuki, S. Anharmonic force constants extracted from first-principles molecular dynamics: applications to heat transfer simulations. *J. Phys. Condens. Matter* **26**, 225402 (2014).
- Hooton, D. L. I. A new treatment of anharmonicity in lattice thermodynamics: I. *Philos. Mag.* **46**, 422 (1955).
- Goldman, V., Horton, G. & Klein, M. An improved self-consistent phonon approximation. *Phys. Rev. Lett.* **21**, 1527 (1968).
- Werthamer, N. Self-consistent phonon formulation of anharmonic lattice dynamics. *Phys. Rev. B* **1**, 572 (1970).
- Souvatzis, P., Eriksson, O., Katsnelson, M. & Rudin, S. Entropy driven stabilization of energetically unstable crystal structures explained from first principles theory. *Phys. Rev. Lett.* **100**, 095901 (2008).
- Errea, I., Calandra, M. & Mauri, F. First-principles theory of anharmonicity and the inverse isotope effect in superconducting palladium-hydride compounds. *Phys. Rev. Lett.* **111**, 177002 (2013).
- Errea, I., Calandra, M. & Mauri, F. Anharmonic free energies and phonon dispersions from the stochastic self-consistent harmonic approximation: application to platinum and palladium hydrides. *Phys. Rev. B* **89**, 064302 (2014).
- Paulatto, L., Errea, I., Calandra, M. & Mauri, F. First-principles calculations of phonon frequencies, lifetimes, and spectral functions from weak to strong anharmonicity: the example of palladium hydrides. *Phys. Rev. B* **91**, 054304 (2015).
- Tadano, T. & Tsuneyuki, S. Self-consistent phonon calculations of lattice dynamical properties in cubic SrTiO<sub>3</sub> with first-principles anharmonic force constants. *Phys. Rev. B* **92**, 054301 (2015).
- Giustino, F. *Materials Modelling using Density Functional Theory: Properties and Predictions*. (Oxford University Press, Oxford, UK, 2014).
- Wang, C., Fasolino, A. & Tosatti, E. Molecular-dynamics theory of the temperature-dependent surface phonons of W(001). *Phys. Rev. B* **37**, 2116 (1988).
- Meyer, R. & Comtesse, D. Vibrational density of states of silicon nanoparticles. *Phys. Rev. B* **83**, 014301 (2011).
- Ladd, A. J. C., Moran, B. & Hoover, W. G. Lattice thermal conductivity: a comparison of molecular dynamics and anharmonic lattice dynamics. *Phys. Rev. B* **34**, 5058 (1986).
- Turney, J., Landry, E., McGaughey, A. & Amon, C. Predicting phonon properties and thermal conductivity from anharmonic lattice dynamics calculations and molecular dynamics simulations. *Phys. Rev. B* **79**, 064301 (2009).
- Lan, T. *Studies of Phonon Anharmonicity in Solids*. Ph.D. thesis, California Institute of Technology (2014).
- Łapińska, A., Taube, A., Judek, J. & Zdrojek, M. Temperature evolution of phonon properties in few-layer black phosphorus. *J. Phys. Chem. C* **120**, 5265 (2016).
- Ferrari, A. C. & Basko, D. M. Raman spectroscopy as a versatile tool for studying the properties of graphene. *Nat. Nanotechnol.* **8**, 235 (2013).
- Reich, S. et al. Resonant Raman scattering in cubic and hexagonal boron nitride. *Phys. Rev. B* **71**, 205201 (2005).
- Koukaras, E. N., Kalosakas, G., Galiotis, C. & Papagelis, K. Phonon properties of graphene derived from molecular dynamics simulations. *Sci. Rep.* **5**, 12923 (2015).
- Kresse, G. & Furthmüller, J. Efficiency of *ab-initio* total energy calculations for metals and semiconductors using a plane-wave basis set. *Comput. Mater. Sci.* **6**, 15 (1996).
- Kresse, G. & Furthmüller, J. Efficient iterative schemes for *ab initio* total-energy calculations using a plane-wave basis set. *Phys. Rev. B* **54**, 11169 (1996).
- Perdew, J. P., Burke, K. & Ernzerhof, M. Generalized gradient approximation made simple. *Phys. Rev. Lett.* **77**, 3865 (1996).
- Blöchl, P. E. Projector augmented-wave method. *Phys. Rev. B* **50**, 17953 (1994).
- Kresse, G. & Hafner, J. *Ab initio* molecular dynamics for liquid metals. *Phys. Rev. B* **47**, 558 (1993).
- Kresse, G. & Hafner, J. *Ab initio* molecular-dynamics simulation of the liquid-metal-amorphous-semiconductor transition in germanium. *Phys. Rev. B* **49**, 14251 (1994).
- Plimpton, S. Fast parallel algorithms for short-range molecular dynamics. *J. Comput. Phys.* **117**, 1 (1995).
- Momma, K. & Izumi, F. VESTA 3 for three-dimensional visualization of crystal, volumetric and morphology data. *J. Appl. Crystallogr.* **44**, 1272 (2011).

60. Brenner, D. W. et al. A second-generation reactive empirical bond order (REBO) potential energy expression for hydrocarbons. *J. Phys. Condens. Matter* **14**, 783 (2002).
61. O'Connor, T. C., Andzelm, J. & Robbins, M. O. AIREBO-M: a reactive model for hydrocarbons at extreme pressures. *J. Chem. Phys.* **142**, 024903 (2015).
62. Los, J. & Fasolino, A. Intrinsic long-range bond-order potential for carbon: performance in Monte Carlo simulations of graphitization. *Phys. Rev. B* **68**, 024107 (2003).
63. Tersoff, J. Modeling solid-state chemistry: interatomic potentials for multi-component systems. *Phys. Rev. B* **39**, 5566 (1989).



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing,

adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019