

Pressure and stress tensor of complex anharmonic crystals within the stochastic self-consistent harmonic approximation

Lorenzo Monacelli,¹ Ion Errea,^{2,3} Matteo Calandra,⁴ and Francesco Mauri^{1,5}


¹*Department of Physics, University of “La Sapienza”, Rome, Italy*

²*Fisika Aplikatua I Saila, Bilboko Ingeniaritza Eskola, University of the Basque Country (UPV/EHU), Bilbao, Basque Country, Spain*

³*Donostia International Physics Center, Donostia-San Sebastian, Basque Country, Spain*

⁴*Sorbonne Université, Centre National de la Recherche Scientifique, Institut des Nanosciences de Paris, UMR7588, F-75252 Paris, France*

⁵*Graphene Labs, Fondazione Istituto Italiano di Tecnologia, Genova, Italy*

 (Received 18 April 2018; revised manuscript received 14 June 2018; published 26 July 2018)

The self-consistent harmonic approximation (SCHA) allows the computation of free energy of anharmonic crystals considering both quantum and thermal fluctuations. Recently, a stochastic implementation of the SCHA has been developed, tailored for applications that use total energy and forces computed from first principles. In this paper, we extend the applicability of the stochastic SCHA to complex crystals, i.e., systems in which symmetries do not fix the inner coordinates and require the optimization of both the lattice vectors and the atomic positions. To this goal, we provide an expression for the evaluation of the pressure and stress tensor within the stochastic SCHA formalism. Moreover, we develop a more robust free-energy minimization algorithm, which allows us to perform the SCHA variational minimization very efficiently in systems having a broad spectrum of phonon frequencies and many degrees of freedom. We test and illustrate the approach with an application to the phase XI of water ice using density-functional theory. We find that the SCHA reproduces extremely well the experimental thermal expansion of ice in the whole temperature range between 0 and 270 K, in contrast with the results obtained within the quasiharmonic approximation, that underestimates the effect by about 25%.

DOI: [10.1103/PhysRevB.98.024106](https://doi.org/10.1103/PhysRevB.98.024106)

I. INTRODUCTION

Atomic vibrations play a main role in many branches of physics and chemistry, as they are involved in thermodynamic, transport, and superconducting properties of materials and molecules. Many spectroscopic techniques, such as Raman and IR, measure how atoms vibrate. The standard approach to describe vibrations is the harmonic approximation, in which the Born-Oppenheimer (BO) energy surface is approximated as a $3N$ -dimensional paraboloid around the ionic positions. The solutions of the harmonic Hamiltonian are well-defined noninteracting vibrational quasiparticles, phonons with an infinite lifetime and temperature-independent spectrum. Anharmonic effects, due to higher orders in the BO energy surface, introduce interactions between phonons. As a result, phonons acquire a finite lifetime that is responsible for thermal transport. Furthermore, phonon spectra become temperature dependent.

Anharmonic effects are commonly accounted for by perturbation theory, the validity range of which is limited only when the harmonic contribution dominates in the range defined by the quantum zero-point motion (ZPM). This is not the case of many interesting phenomena, such as systems undergoing a displacive second-order structural phase transition in which a phonon branch softens as a function of temperature, e.g., charge-density waves and ferroelectrics [1–12], or in solids largely affected by the ZPM, for example in hydrides or in molecular crystals containing H, like water and high-pressure phases of hydrogen [13–18]. Classical molecular dynamics (MD) for ions or methods based on it can be used to extract the

nonperturbative anharmonic renormalized phonon dispersion [19–27]. However, within these approaches, quantum effects on nuclei are neglected. These methods are then inappropriate below the Debye temperature.

In order to correctly account for both quantum and anharmonic effects, the ideal technique is path-integral molecular dynamics (PIMD) [28–30], but its demanding computational cost limits its applicability to systems with few atoms or to the use of empirical potentials. To overcome these problems many self-consistent approximations have been developed [31,32]. Among them, the self-consistent harmonic approximation (SCHA) allows one to describe anharmonicity through a full-quantum variational theory. The stochastic implementation of the SCHA [16] (SSCHA) allows us to apply the powerful variational SCHA method to many systems with a lower numerical effort than MD and PIMD, making possible the calculation of nonperturbative anharmonic effects from first principles.

So far, the applications of the SSCHA method [15–17,33–35] have been limited to simple systems with high symmetry. The main reason is that the variational minimization as formulated in Ref. [16] can yield “runaway solutions” and become very inefficient in complex crystals that show a wide range of phonon frequencies and many degrees of freedom. Another limitation of the original SSCHA formulation is that it needs finite-difference approaches to estimate the effect of ionic fluctuations in the stress tensor, as it happens in the quasiharmonic approximation (QHA), which is extremely cumbersome for noncubic crystals. This hinders cell relaxation within the SSCHA.

In this paper, we efficiently overcome these difficulties by developing an equation for the stress tensor within the SCHA. Furthermore, we develop a more robust minimization algorithm based on an analytical preconditioner combined with a nonlinear change of variables that allows for efficient many-variables minimizations. Our developments pave the way to primitive cell relaxations including quantum and anharmonic effects avoiding finite difference approaches. Thanks to the improvements, we are able to drop the symmetry constraints and to study systems with little or no symmetry, as it is in molecular crystals.

We illustrate and benchmark the method with the phase XI of ice (H₂O), the perfect prototype of a complex molecular crystal. Ice XI is the ordered phase of common ice formed below 72 K in the presence of a small amount of an alkali-metal hydroxide [36]. It is commonly used to study quantum effects in water thanks to its great similarity to normal ice (Ih) [37,38]. Ice is characterized by the interplay between intramolecular covalent OH bonds and intermolecular hydrogen bonds. The great difference in strength of intermolecular and intramolecular forces makes ice phases acquire a very broad spectrum for their vibrational energies, from the very low-energy rotors to the large-energy vibrons. Moreover, this structure of ice experimentally exhibits at low temperature negative thermal expansion [39] and the ‘‘anomalous isotope volume effect’’ [37–40]: if hydrogen is replaced by deuterium, the crystal volume expands by about a 0.1%. This is the opposite behavior of what is usually observed when a heavier isotope is substituted in the crystal. These features make the XI phase of crystal ice a perfect benchmark for the here developed SSCHA algorithm (Secs. VI and VII).

This paper is organized as follows. We recall the basis of the SCHA algorithm in Sec. II. We introduce the stress tensor in the SCHA formalism in Sec. III. We discuss the stochastic implementation of the algorithm in Sec. IV. We face the issues of the SSCHA minimization algorithm in Sec. V: we get an ansatz on the condition number of the minimization process (Sec. VA), and provide two changes of variables that suppress it (Secs. VB and VC). Then, we benchmark the SCHA algorithm in ice XI in Sec. VI. Finally, Sec. VII reports the results computed with density functional theory (DFT) in the unit cell of ice XI, compared with the QHA. In Sec. VIII we summarize the main results of this paper. The paper is completed with three appendices, where the mathematical derivations of the presented equations are provided.

II. THE SELF-CONSISTENT HARMONIC APPROXIMATION

The SCHA is a variational principle on the BO free energy. The nuclear quantum Hamiltonian of a generic system can be defined in the BO approximation as

$$H = \sum_{n=1}^N \sum_{\alpha=1}^3 \frac{p_n^{\alpha 2}}{2M_n} + V(\vec{R}, \{\vec{a}_i\}), \quad (1)$$

where V is the BO energy surface, M_n is the mass of the n th atom, p_n^α and \vec{R} (R_n^α) are the momentum and position operators of the nuclei in the periodic cell (or supercell), N is the number of atoms, and $\{\vec{a}_i\}$ are the three unit-cell vectors. The α index

identifies the Cartesian coordinate. Fixing the temperature T and the volume (i.e., the cell vectors $\{\vec{a}_i\}$), the free energy of the ionic Hamiltonian H is

$$F(\{\vec{a}_i\}) = \langle H \rangle_{\rho_H} + k_b T \langle \ln \rho_H \rangle_{\rho_H}, \quad (2)$$

where ρ_H is the equilibrium density matrix

$$\rho_H = \frac{e^{-\beta H}}{\text{Tr} e^{-\beta H}}, \quad \beta = \frac{1}{k_b T}, \quad (3)$$

and the brackets $\langle O \rangle_{\rho_H}$ indicate the average of the observable O according to the ρ_H density matrix:

$$\langle O \rangle_{\rho_H} = \text{Tr} [\rho_H O]. \quad (4)$$

The equilibrium density matrix satisfies the free-energy least principle. Given a trial density matrix $\rho_{\mathcal{H}}$, we can define a free-energy functional the minimum of which is the free energy:

$$\mathcal{F}(\{\vec{a}_i\})[\rho_{\mathcal{H}}] = \langle H \rangle_{\rho_{\mathcal{H}}} + k_b T \langle \ln \rho_{\mathcal{H}} \rangle_{\rho_{\mathcal{H}}}, \quad (5)$$

$$F(\{\vec{a}_i\}) = \min_{\rho_{\mathcal{H}}} \mathcal{F}(\{\vec{a}_i\})[\rho_{\mathcal{H}}]. \quad (6)$$

The SCHA consists in the restriction of the possible trial density matrices to the equilibrium one obtained from a harmonic Hamiltonian:

$$\mathcal{H}_{\vec{R}, \Phi} = \sum_{n=1}^N \sum_{\alpha=1}^3 \frac{p_n^{\alpha 2}}{2M_n} + \mathcal{V}_{\Phi, \vec{R}}(\vec{R}), \quad \text{where} \quad (7a)$$

$$\mathcal{V}_{\Phi, \vec{R}}(\vec{R}) = \frac{1}{2} \sum_{n=1}^N \sum_{\alpha=1}^3 \sum_{m=1}^3 u_n^\alpha \Phi_{nm}^{\alpha\beta} u_m^\beta, \quad \text{and} \quad (7b)$$

$$u_n^\alpha = R_n^\alpha - \mathcal{R}_n^\alpha, \quad (7c)$$

$$\rho_{\mathcal{H}} = \rho_{\vec{R}, \Phi} = \frac{e^{-\beta \mathcal{H}_{\vec{R}, \Phi}}}{\text{Tr} e^{-\beta \mathcal{H}_{\vec{R}, \Phi}}}. \quad (7d)$$

Here u_n^α is the displacement of the n th atom along the α direction with respect to a central position \vec{R} , and $\Phi_{nm}^{\alpha\beta}$ is the matrix element of the real-space force constant matrix (we use bold font to indicate tensors and matrices). With the introduction of the auxiliary harmonic Hamiltonian it is possible to recast the free energy as

$$\mathcal{F}(\vec{R}, \Phi, \{\vec{a}_i\}) = F_\Phi + \langle V - \mathcal{V}_{\Phi, \vec{R}} \rangle_{\rho_{\vec{R}, \Phi}}, \quad (8)$$

where F_Φ is the exact free energy of the harmonic Hamiltonian:

$$F_\Phi(T) = \sum_{\mu=1}^{3N} \left[\frac{\hbar \omega_\mu}{2} + \frac{1}{\beta} \ln (1 - e^{-\beta \hbar \omega_\mu}) \right], \quad (9)$$

where ω_μ and \vec{e}_μ are, respectively, the eigenvalues and eigenvectors of the Φ matrix divided by the atomic masses:

$$\sum_{t=1}^N \sum_{\beta=1}^3 \frac{\Phi_{st}^{\alpha\beta}}{\sqrt{M_s M_t}} e_{\mu t}^\beta = \omega_\mu^2 e_{\mu s}^\alpha. \quad (10)$$

The real free energy can, therefore, be approximated as the minimum of the free-energy functional [Eq. (8)] with respect

to $\vec{\mathcal{R}}$ and Φ :

$$\mathcal{F}(\vec{\mathcal{R}}, \{\vec{a}_i\}) = \min_{\Phi} \mathcal{F}(\vec{\mathcal{R}}, \Phi, \{\vec{a}_i\}), \quad (11a)$$

$$\mathcal{F}(\{\vec{a}_i\}) = \min_{\Phi, \vec{\mathcal{R}}} \mathcal{F}(\vec{\mathcal{R}}, \Phi, \{\vec{a}_i\}). \quad (11b)$$

From now on, when we drop either the $\vec{\mathcal{R}}$ or Φ symbol, we mean the quantity computed in the value of that variable that minimizes the free energy. For example, the equilibrium SCHA density matrix is just ρ .

One of the advantages of using the harmonic Hamiltonian to restrict the $\rho_{\mathcal{H}}$ space is that we have a trivial physical interpretation of the minimization parameters. In fact $\vec{\mathcal{R}}$ represents the centroid positions of the atoms, i.e., the anharmonic average positions as measured by, e.g., diffraction experiments:

$$\vec{\mathcal{R}} = \langle \vec{R} \rangle_{\rho}. \quad (12)$$

In the same way, Φ is related to the thermal and quantum fluctuations and defines the real-space density matrix broadening. Within the harmonic auxiliary Hamiltonian, the probability distribution function defined by the real-space density matrix is a product of Gaussians:

$$\rho_{\vec{\mathcal{R}}, \Phi}(\vec{u}) = \langle \vec{u} | \rho_{\vec{\mathcal{R}}, \Phi} | \vec{u} \rangle, \quad (13a)$$

$$\rho_{\vec{\mathcal{R}}, \Phi}(\vec{u}) = \sqrt{\det(\Upsilon/2\pi)} \exp\left(-\frac{1}{2} \sum_{st\alpha\beta} \Upsilon_{st}^{\alpha\beta} u_s^{\alpha} u_t^{\beta}\right) \quad (13a)$$

where

$$\Upsilon_{st}^{\alpha\beta} = \sqrt{M_s M_t} \sum_{\mu} \frac{2\omega_{\mu}}{(1+2n_{\mu})\hbar} e_{\mu_s}^{\alpha} e_{\mu_t}^{\beta} \quad (13b)$$

and n_{μ} are the boson average occupation number for the μ mode. It is important to notice that ω_{μ} and \vec{e}_{μ} [Eq. (10)] are not directly equal to the physical phonons since they are constrained to be positive defined [35]. Instead, they are related to quantum and thermal fluctuations: they uniquely define the Υ tensor.

It is possible to define the SCHA force as the derivative of the free energy [Eq. (11a)] with respect to the nuclear average positions:

$$-\frac{\partial \mathcal{F}}{\partial \mathcal{R}_n^{\alpha}}(\vec{\mathcal{R}}, \{\vec{a}_i\}) = \langle f_n^{\alpha} - f_{\mathcal{H}_n}^{\alpha} \rangle_{\rho_{\vec{\mathcal{R}}}}, \quad (14)$$

where \vec{f} and $\vec{f}_{\mathcal{H}}$ are, respectively, the BO and harmonic forces:

$$f_n^{\alpha} = -\frac{\partial V}{\partial R_n^{\alpha}}(\vec{R}, \{\vec{a}_i\}), \quad (15)$$

$$f_{\mathcal{H}_n}^{\alpha} = -\frac{\partial \mathcal{V}_{\vec{\mathcal{R}}, \Phi}}{\partial R_n^{\alpha}} = -\sum_{m=1}^N \sum_{\beta=1}^3 \Phi_{nm}^{\alpha\beta} u_m^{\beta}. \quad (16)$$

It is interesting to notice how the harmonic potential $\mathcal{V}_{\vec{\mathcal{R}}, \Phi}$ does not depend explicitly on the unit-cell vectors $\{\vec{a}_i\}$, while the BO energy $V(\vec{R}, \{\vec{a}_i\})$ does.

To numerically minimize the SCHA free energy it is possible to use the steepest descent (SD) or conjugate gradient (CG) methods [41], both based on the knowledge of the gradient of

the function to minimize. This can be expressed as a function of the averages of the BO and harmonic forces [16]:

$$\nabla_{\Phi} \mathcal{F}(\vec{\mathcal{R}}, \Phi, \{\vec{a}_i\}) = -\sum_{st\alpha\beta\mu} \sqrt{\frac{M_t}{M_s}} (e_{\mu_s}^{\alpha} \nabla_{\Phi} \ln a_{\mu} + \nabla_{\Phi} e_{\mu_s}^{\alpha}) \times \langle [f_s^{\alpha}(\vec{u}) - f_{\mathcal{H}_s}^{\alpha}(\vec{u})] u_t^{\beta} \rangle_{\rho_{\vec{\mathcal{R}}, \Phi}} e_{\mu_t}^{\beta}, \quad (17a)$$

$$\nabla_{\mathcal{R}_s^{\alpha}} \mathcal{F}(\vec{\mathcal{R}}, \Phi, \{\vec{a}_i\}) = -\langle f_s^{\alpha} - f_{\mathcal{H}_s}^{\alpha} \rangle_{\rho_{\vec{\mathcal{R}}, \Phi}}. \quad (17b)$$

In the next section, we show how to implement the SCHA in an isobaric ensemble, allowing for the relaxation also of the unit cell. This is achieved thanks to the introduction of the stress tensor in the SCHA framework.

III. THE STRESS TENSOR IN THE SELF-CONSISTENT STOCHASTIC APPROXIMATION

To minimize the free energy with respect to the lattice parameters in a periodic system, knowledge of the stress tensor is crucial. The SCHA stress can be defined as

$$P_{\alpha\beta}(\vec{\mathcal{R}}, \{\vec{a}_i\}) = -\frac{1}{\Omega} \frac{\partial \mathcal{F}(\vec{\mathcal{R}}, \{\vec{a}_i\})}{\partial \varepsilon_{\alpha\beta}} \Bigg|_{\varepsilon=0}, \quad (18)$$

where Ω is the volume of the system and the strain tensor $\varepsilon_{\alpha\beta}$ identifies a generic deformation, where both the lattice parameters and the average central position are affected:

$$a_i^{\prime\alpha} = a_i^{\alpha} + \sum_{\beta=1}^3 \varepsilon_{\alpha\beta} a_i^{\beta}, \quad (19a)$$

$$\mathcal{R}_n^{\prime\alpha} = \mathcal{R}_n^{\alpha} + \sum_{\beta=1}^3 \varepsilon_{\alpha\beta} \mathcal{R}_n^{\beta}. \quad (19b)$$

This is equivalent to performing a strain keeping fixed the internal crystal coordinates of the system. The final result can be divided into three main contributions (see Appendix A for the proof):

$$P_{\alpha\beta}(\vec{\mathcal{R}}, \{\vec{a}_i\}) = P_{\alpha\beta}^H(\vec{\mathcal{R}}, \{\vec{a}_i\}) + P_{\alpha\beta}^{\text{FLC}}(\vec{\mathcal{R}}, \{\vec{a}_i\}) + P_{\alpha\beta}^{\text{FRC}}(\vec{\mathcal{R}}, \{\vec{a}_i\}), \quad (20)$$

where the $P_{\alpha\beta}^H(\vec{\mathcal{R}})$ is the static contribution, i.e., the stress tensor computed without quantum and thermal fluctuations (classical with $T = 0$), $P_{\alpha\beta}^{\text{FLC}}$ is the contribution of the fluctuations to the stress, and $P_{\alpha\beta}^{\text{FRC}}$ is an extra term that takes into account the work necessary to move the centroids according to the applied strain ε :

$$P_{\alpha\beta}^H(\vec{\mathcal{R}}, \{\vec{a}_i\}) = -\frac{1}{\Omega} \frac{\partial V(\vec{R}, \{\vec{a}_i\})}{\partial \varepsilon_{\alpha\beta}} \Bigg|_{\varepsilon=0}, \quad (21a)$$

$$P_{\alpha\beta}^{\text{FLC}}(\vec{\mathcal{R}}, \{\vec{a}_i\}) = \langle P_{\alpha\beta}^H(\vec{R}, \{\vec{a}_i\}) \rangle_{\rho_{\vec{\mathcal{R}}}} - P_{\alpha\beta}^H(\vec{\mathcal{R}}, \{\vec{a}_i\}) - \frac{1}{2\Omega} \sum_{s=1}^N \langle (f_{\mathcal{H}_s}^{\alpha} u_s^{\beta} + f_{\mathcal{H}_s}^{\beta} u_s^{\alpha}) \rangle_{\rho_{\vec{\mathcal{R}}}}, \quad (21b)$$

$$P_{\alpha\beta}^{\text{FRC}}(\vec{\mathcal{R}}, \{\vec{a}_i\}) = \frac{1}{2\Omega} \sum_{s=1}^N (\mathcal{R}_s^\beta \langle f_s^\alpha - f_{\mathcal{H}_s}^\alpha \rangle_{\rho_{\vec{\mathcal{R}}}} + \mathcal{R}_s^\alpha \langle f_s^\beta - f_{\mathcal{H}_s}^\beta \rangle_{\rho_{\vec{\mathcal{R}}}}). \quad (21c)$$

The last term in Eq. (21b) makes fluctuations on pressure disappear in a pure harmonic crystal (see Appendix A). The force term, i.e., Eq. (21c), is nonzero only if the SCHA centroids $\vec{\mathcal{R}}$ are not in the equilibrium configuration, i.e., the SCHA force Eq. (14) on each atom is not zero, and it is independent of the choice of the origin (the sum of the forces over atom indices is zero).

Equation (21) can be computed once we know the BO surface $V(\vec{R}, \{\vec{a}_i\})$, the atomic force $f_s^\alpha(\vec{R}, \{\vec{a}_i\})$, and the stress tensors $P_{\alpha\beta}^H(\vec{R}, \{\vec{a}_i\})$ for each ionic displacement \vec{R} in the ensemble of the configurations distributed according to $\rho_{\vec{\mathcal{R}}}(\vec{R})$. In Sec. IV we discuss an efficient stochastic implementation to numerically compute this average.

The computation of the SCHA stress tensor enables the complete unit-cell relaxation in isobaric conditions (fixing the external P^* pressure). This is done by minimizing the Gibbs free energy, that is obtained from the Helmholtz free energy through the Legendre transform:

$$\mathcal{G}(\vec{\mathcal{R}}, P^*) = \mathcal{F}(\vec{\mathcal{R}}, \{\vec{a}_i\}) + P^* \Omega(\{\vec{a}_i\}). \quad (22)$$

IV. THE STOCHASTIC IMPLEMENTATION

The SCHA algorithm can be implemented by performing the stochastic evaluation of all the averages. Thanks to the fact that the density matrix is a multidimensional Gaussian function [Eq. (13a)], it is possible to generate an ensemble distributed according to $\rho_{\vec{\mathcal{R}}, \Phi}$ without any Metropolis algorithm [16], and the average of a generic observable $O(\vec{R})$ can be computed through Monte Carlo integration:

$$\langle O(\vec{R}) \rangle_{\rho_{\vec{\mathcal{R}}, \Phi}} = \frac{1}{N_c} \sum_{\vec{R}_I} O(\vec{R}_I), \quad (23)$$

where N_c and \vec{R}_I are, respectively, the dimension and the configurations of the ensemble. To avoid regenerating the ensemble at each minimization step it is convenient to introduce the importance sampling reweighting [16]:

$$\rho_I = \frac{\rho_{\vec{\mathcal{R}}, \Phi}(\vec{R}_I)}{\rho_{\vec{\mathcal{R}}_{\text{SG}}, \Phi_{\text{SG}}}(\vec{R}_I)} \quad (24)$$

where $\rho_{\vec{\mathcal{R}}_{\text{SG}}, \Phi_{\text{SG}}}(\vec{R}_I)$ is the density matrix used to extract the ensemble configurations, i.e., computed with the starting guess for the centroid positions $\vec{\mathcal{R}}_{\text{SG}}$ and the auxiliary dynamical matrix Φ_{SG} . Then, the average of the observable O in a generic value of $\vec{\mathcal{R}}$ and Φ can be computed through

$$\langle O(\vec{R}) \rangle_{\rho_{\vec{\mathcal{R}}, \Phi}} = \frac{1}{N_c} \sum_{\vec{R}_I} \rho_I O(\vec{R}_I). \quad (25)$$

The reweighting procedure allows us to overtake the usually high computational effort required by the SSCHA minimization. In fact, the computation of the SCHA free-energy gradient [16], as well as the SCHA stress tensor [Eq. (21)], requires

only the knowledge of the first derivative of the BO energy in the ensemble, that can be obtained just in one total-energy calculation per configuration thanks to the Hellmann-Feynman theorem. Moreover, the total-energy calculation can be computed only one time in the starting ensemble of configurations \vec{R}_I , and then recycled on the whole minimization thanks to the reweighting. When the new variables $\vec{\mathcal{R}}$ and Φ are too distant from the initial ones, $\vec{\mathcal{R}}_{\text{SG}}$ and Φ_{SG} , the ensemble is no longer able to provide a good estimation of the stochastic averages and it must be reextracted. Thus, the overall computational effort to run an *ab initio* SSCHA calculation is given by the number of times the initial ensemble is regenerated.

It is crucial to improve the reliability of the ensemble, in order to minimize the number of times the initial ensemble is regenerated during the SSCHA free-energy optimization. To this purpose, we adopt both a *symmetrized sampling* and a stochastic threshold to evaluate the important sampling accuracy.

The real-space density matrix is a symmetric distribution, $\rho_{\vec{\mathcal{R}}, \Phi}(\vec{u}) = \rho_{\vec{\mathcal{R}}, \Phi}(-\vec{u})$, and all observables required in the SSCHA free-energy minimization are purely even or odd terms of the Taylor expansion of $V(\vec{R}, \{\vec{a}_i\})$ in $(\vec{R} - \vec{\mathcal{R}})$. To reduce the stochastic noise we implemented the symmetrized sampling [42]: for each displacement \vec{u} generated, also its opposite $-\vec{u}$ is included in the ensemble. This analytically cancels all the noncontributing terms in the Taylor expansion of the BO energy. It is important to notice that this advantage is lost when $\vec{\mathcal{R}} \neq \vec{\mathcal{R}}_{\text{SG}}$. However, we still find the symmetrized sampling to be convenient to reduce the stochastic noise even if the centroids do not match perfectly the starting guess.

The previous estimator of the importance sampling accuracy used by Errea *et al.* [16] was the check on the ρ_I normalization:

$$\left| \frac{1}{N_c} \sum_{I=1}^{N_c} \rho_I - 1 \right| < \eta. \quad (26)$$

However, this threshold can be exceeded if all the weight constantly drifts from the uniform value, or it can remain satisfied if they spread a lot. Thus, a much better estimator that considers the spreading of the different configuration weights can be implemented. In order to improve the reliability of the reweighting procedure we found more reliable the Kong-Liu effective sample size [43]:

$$N_{\text{eff}} = \frac{(\sum_I \rho_I)^2}{\sum_I \rho_I^2} < N_c. \quad (27)$$

A critical threshold η' can be defined as

$$\frac{N_{\text{eff}}}{N_c} > \eta'. \quad (28)$$

If the weight ρ_I of a configuration goes to zero, it does not contribute to the averages. The effective sample size counts how many configurations are actually contributing to the Monte Carlo average [Eq. (25)], even if the ρ_I are properly normalized.

We set $\eta' = 0.6$ in all the simulations reported in this paper. If the critical threshold is overcome, the minimization

is stopped and an ensemble is generated with the final trial density matrix $\rho_{\vec{R}, \Phi}$.

V. MINIMIZATION INSTABILITIES AND RUNAWAY SOLUTIONS

The SSCHA algorithm consists in minimizing the free energy through the stochastic evaluation of its gradient [Eq. (17)], employing the SD or CG algorithm, and taking advantage of reweighting to perform multiple SD or CG steps without recomputing energies and forces of the ensemble at each step. However, this minimization procedure was empirically found to be very difficult in some systems, especially in those near a structural instability, where a phonon mode frequency softens to zero, while all the other modes are substantially higher in energy, or molecular crystals, in which hard intermolecular vibrations coexist with low-energy intramolecular modes. In these cases, of very great physical interest, the stochastic free-energy minimization requires a large number of ensemble regenerations (and consequently a large number of first-principles force calculations) to converge. Moreover, the minimization can lead to runaway solutions: fake nonphysical solutions of the SCHA self-consistency where the auxiliary dynamical matrix Φ is not positive defined.

To understand the convergence properties we consider the SCHA free energy close to the minimum. It can be approximated as a quadratic form in the minimization variables (\vec{R} and Φ). Under this condition, the SCHA free energy is expressed by the Hessian matrix A with respect to those variables. From the Hessian matrix it is possible to define the condition number [44] C , as

$$C = \frac{\max \lambda_A}{\min \lambda_A}, \quad (29)$$

where λ_A is a generic eigenvalue of the A matrix. In the limit in which the number of degrees of freedom is much greater than the number of minimization steps, the SD and the CG algorithms converge into a fixed threshold with almost N steps proportional [41] to

$$N_{SD} \propto C, \quad (30a)$$

$$N_{CG} \propto \sqrt{C}. \quad (30b)$$

In the SCHA case, the number of minimization steps is proportional to the number of times the critical threshold η' is overcome. Then, this number must be carefully optimized, since each time the ensemble is reextracted the *ab initio* energies and forces for each configuration must be computed. This calculation is the overall computational cost of the algorithm. In the next sections, we provide an ansatz for the condition number, unveiling that it dramatically diverges in the aforesaid cases. We further provide two ways to prevent this divergence, paving the way for the application of SSCHA in these systems.

A. Hessian matrix

In this section, we provide an analytical guess of the free-energy Hessian matrix A with respect to the minimization variable Φ . In general, this is not possible, since computing the real Hessian matrix corresponds to solving exactly the

problem. However, we can perform the computation in an analytical test case that, hopefully, will enclose all the physics of the minimization problems incurred so far. This is a purely harmonic system, described by a harmonic Hamiltonian. From now on we introduce a compact notation to describe both Cartesian and atomic indices ($v_a = v_s^a$):

$$H = \frac{1}{2} \sum_a \frac{(p_a)^2}{2M_a} + \frac{1}{2} \sum_{ab} u_a K_{ab} u_b. \quad (31)$$

The free-energy Hessian matrix with respect to the Φ variable can be computed analytically. The steps that lead to the following result are reported in Appendix B:

$$A_{\Phi}^{abcd} = \left. \frac{\partial^2 \mathcal{F}(\vec{R}, \Phi, \{\vec{a}_i\})}{\partial \Phi_{ab} \partial \Phi_{cd}} \right|_{\Phi=\mathbf{K}} = \frac{1}{2} \mathcal{P}_{ab} \mathcal{P}_{cd} (\Lambda^{abcd} + \Lambda^{abdc}), \quad (32)$$

where the Λ rank-4 tensor is the same as that introduced by Bianco *et al.* [35], and \mathcal{P} is a symmetrization factor:

$$\Lambda^{abcd} = -\frac{\hbar}{4} \sum_{\mu\nu} \frac{1}{\omega_{\mu} \omega_{\nu}} \frac{e_{\nu}^a e_{\mu}^b e_{\nu}^c e_{\mu}^d}{\sqrt{M_a M_b M_c M_d}} \times \begin{cases} \frac{2n_{\nu} + 1}{2\omega_{\nu}} - \frac{dn_{\nu}}{d\omega_{\mu}} & \omega_{\nu} = \omega_{\mu} \\ \frac{n_{\mu} + n_{\nu} + 1}{\omega_{\mu} + \omega_{\nu}} - \frac{n_{\mu} - n_{\nu}}{\omega_{\mu} - \omega_{\nu}} & \omega_{\nu} \neq \omega_{\mu} \end{cases}, \quad (33)$$

$$\mathcal{P}_{ab} = \sqrt{2}(1 - \delta_{ab}) + \delta_{ab}. \quad (34)$$

Here the ω_{μ} and \vec{e}_{μ} are the frequencies and polarization vectors of the \mathbf{K} matrix. These are, indeed, equal to the Φ matrix in the minimum of the SCHA free energy, and represent the real phonons of the system.

The Λ matrix can be diagonalized analytically if we consider the case of all equal masses:

$$\sum_{cd} \Lambda^{abcd} e_{\mu}^c e_{\nu}^d = \tilde{\lambda}_{\mu\nu} e_{\mu}^a e_{\nu}^b. \quad (35)$$

We can obtain an easy expression of the spectrum of the Hessian matrix in the pure quantum limit $T \rightarrow 0$ and the pure classical limit $T \rightarrow \infty$:

$$\lim_{T \rightarrow 0} \tilde{\lambda}_{\mu\nu} = -\frac{\hbar}{4M^2} \frac{1}{\omega_{\mu} \omega_{\nu} (\omega_{\mu} + \omega_{\nu})}, \quad (36a)$$

$$\lim_{T \rightarrow \infty} \tilde{\lambda}_{\mu\nu} = -\frac{1}{4\beta M^2} \frac{1}{\omega_{\mu}^2 \omega_{\nu}^2} \left[1 + \frac{\omega_{\mu} \omega_{\nu}}{(\omega_{\mu} + \omega_{\nu})^2} \right]. \quad (36b)$$

Therefore, the Hessian matrix spectrum goes as ω_{μ}^{-3} in the quantum limit and ω_{ν}^{-4} in the classical one. We can compute the condition numbers, as defined in Eq. (29):

$$C_{\Phi, T=0} \approx \left(\frac{\omega_{\max}}{\omega_{\min}} \right)^3, \quad (37a)$$

$$C_{\Phi, T \rightarrow \infty} \approx \left(\frac{\omega_{\max}}{\omega_{\min}} \right)^4. \quad (37b)$$

This unveils the pathology in the SSCHA minimization if the gradient is taken with respect to Φ as presented in

Ref. [16] for the mentioned systems: when we have a structural instability, there is a phonon mode that softens to zero ($\omega_{\min} \rightarrow 0$), producing a diverging condition number $C \rightarrow \infty$. In the same way, molecular crystals have a broad spectrum, with a very large difference between the highest vibron modes and the lowest intermolecular ones (for example, in common ice we have $\omega_{\max}/\omega_{\min} \sim 10^3$). This yields extremely high values of the condition numbers, that makes the minimization really difficult and requires lots of energy and force recalculations to achieve a good minimization. This obviously hinders the fully *ab initio* application of the SSCHA method in complex systems.

B. Nonlinear change of variable

The condition number is a function of the minimization variables. Therefore, a simple change of variables can result in a powerful improvement in the minimization algorithm. In this section, we show that it is possible to almost completely solve the divergences occurring in the condition numbers [Eq. (37)] with a simple nonlinear change in the Φ auxiliary matrix. Moreover, we can also completely cancel the aforesaid runaway solutions. The runaway solutions are fake nonphysical solutions of the SCHA that may arise during the minimization if the Φ matrix is not positive definite. In order to avoid this problem, one could perform a constrained minimization. It is difficult to implement this kind of constraint with the SD or CG algorithm. We find it much more convenient to introduce a nonlinear change of variables, where we replace the auxiliary dynamical matrix Φ with one of its even roots:

$$\Phi \rightarrow \sqrt[2n]{\Phi}. \quad (38)$$

This mathematically constrains the minimization to have only positive defined matrices Φ . Does this nonlinear change improve the condition number on the minimization?

We can compute the Hessian matrix $A_{\sqrt{\Phi}}$ with respect to the square root of Φ where Φ minimizes the free energy:

$$A_{\sqrt{\Phi}} = \frac{\partial^2 \mathcal{F}(\vec{\mathcal{R}}, \Phi, \{\vec{a}_i\})}{\partial \sqrt{\Phi} \partial \sqrt{\Phi}} = \Phi A_{\Phi} + 2\sqrt{\Phi} A_{\Phi} \sqrt{\Phi} + A_{\Phi} \Phi, \quad (39)$$

where A_{Φ} is the rank-4 Hessian with respect to Φ [Eq. (32)]. The procedure can be iterated to obtain any even root of Φ . Here we report also the $\sqrt[4]{\Phi}$ expression, since, as we will show, it has a very favorable condition number:

$$A_{\sqrt[4]{\Phi}} = \sqrt{\Phi} A_{\sqrt{\Phi}} + 2\sqrt[4]{\Phi} A_{\sqrt{\Phi}} \sqrt[4]{\Phi} + A_{\sqrt{\Phi}} \sqrt{\Phi}. \quad (40)$$

We can easily compute the condition numbers in the new variables if all the masses are equal substituting Eq. (32) into Eqs. (39) and (40) (recalling that $\Phi \sim \omega^2$):

$$C_{\sqrt[2]{\Phi}, T=0} \sim \left(\frac{\omega_{\max}}{\omega_{\min}} \right), \quad C_{\sqrt[2]{\Phi}, T \rightarrow \infty} \sim \left(\frac{\omega_{\max}}{\omega_{\min}} \right)^2, \quad (41)$$

$$C_{\sqrt[4]{\Phi}, T=0} \sim 1, \quad C_{\sqrt[4]{\Phi}, T \rightarrow \infty} \sim \left(\frac{\omega_{\max}}{\omega_{\min}} \right). \quad (42)$$

The nonlinear change of variable $\Phi \rightarrow \sqrt[4]{\Phi}$ both avoids the nonphysical runaway solutions constraining the minimization space to admit only positive defined matrices and strongly

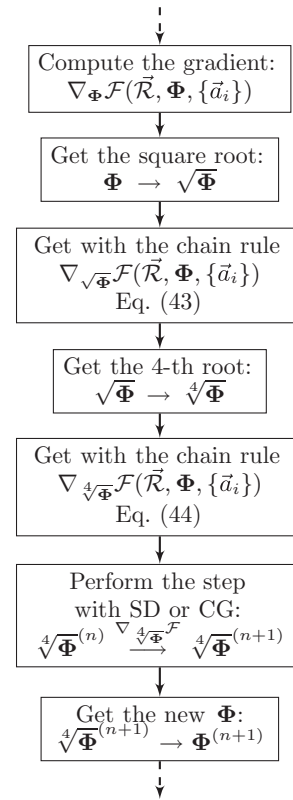


FIG. 1. Flowchart on a minimization step with the $\Phi \rightarrow \sqrt[4]{\Phi}$ change of variables.

suppress the condition number, making it independent on the phonon frequencies in the $T = 0$ case and suppressing it by a fourth root in the classical case.

In practice, the minimization in the $\sqrt[4]{\Phi}$ is performed by computing the free-energy gradient with respect to the new variable adopting the chain rule on the derivatives:

$$\begin{aligned} \nabla_{\sqrt{\Phi}} \mathcal{F}(\vec{\mathcal{R}}, \Phi, \{\vec{a}_i\}) &= \sqrt{\Phi} \nabla_{\Phi} \mathcal{F}(\vec{\mathcal{R}}, \Phi, \{\vec{a}_i\}) \\ &+ \nabla_{\Phi} \mathcal{F}(\vec{\mathcal{R}}, \Phi, \{\vec{a}_i\}) \sqrt{\Phi}, \end{aligned} \quad (43)$$

$$\begin{aligned} \nabla_{\sqrt[4]{\Phi}} \mathcal{F}(\vec{\mathcal{R}}, \Phi, \{\vec{a}_i\}) &= \sqrt[4]{\Phi} \nabla_{\sqrt{\Phi}} \mathcal{F}(\vec{\mathcal{R}}, \Phi, \{\vec{a}_i\}) \\ &+ \nabla_{\sqrt{\Phi}} \mathcal{F}(\vec{\mathcal{R}}, \Phi, \{\vec{a}_i\}) \sqrt[4]{\Phi}. \end{aligned} \quad (44)$$

The minimization step is updated as described by the flowchart reported in Fig. 1.

C. Preconditioning

Even if the fourth root change of variable considerably improves the condition number, for high-temperature calculations it still depends on the phonon frequencies linearly, which could be problematic when a phonon mode goes close to zero near a structural phase transition. The SSCHA minimization algorithm corresponds to finding the zeros of the free-energy gradient:

$$\nabla_{\Phi} \mathcal{F}(\vec{\mathcal{R}}, \Phi, \{\vec{a}_i\}) = 0. \quad (45)$$

From the above system, the SD and the CG algorithms are derived. However, since A_{Φ} is a positive defined matrix, the

solution of the SCHA equation coincides with the solution of the auxiliary problem:

$$\mathbf{A}_{\Phi}^{-1} \nabla_{\Phi} \mathcal{F}(\vec{\mathcal{R}}, \Phi, \{\vec{a}_i\}) = 0. \quad (46)$$

It can be shown [41,44] that the condition number on the problem defined by Eq. (46) is equal to 1 if \mathbf{A} is the exact Hessian matrix of $\mathcal{F}(\vec{\mathcal{R}}, \Phi, \{\vec{a}_i\})$. We can, therefore, use the analytic guess of the Hessian matrix \mathbf{A} provided in Eq. (32) to redefine the minimization algorithm. The SD algorithm on the problem of Eq. (46) becomes

$$\Phi^{(n+1)} = \Phi^{(n)} - \lambda \mathbf{A}_{\Phi}^{-1} \nabla_{\Phi} \mathcal{F}(\vec{\mathcal{R}}^{(n)}, \Phi^{(n)}, \{\vec{a}_i\}), \quad (47)$$

where λ is the minimization step. Another advantage of using the auxiliary problem is that, if \mathbf{A} is exact and $\mathcal{F}(\vec{\mathcal{R}}, \Phi, \{\vec{a}_i\})$ is quadratic, the minimization arrives in the minimum of the free energy in only one step with $\lambda = 1$. In a very similar way also the CG algorithm can be redefined for the auxiliary problem:

$$\mathbf{d}^{(0)} = 0, \quad (48a)$$

$$\mathbf{d}^{(n+1)} = \mathbf{A}_{\Phi}^{-1} \nabla_{\Phi} \mathcal{F}^{(n+1)} + \frac{\nabla_{\Phi} \mathcal{F}^{(n+1)} \mathbf{A}_{\Phi}^{-1} \nabla_{\Phi} \mathcal{F}^{(n+1)}}{\nabla_{\Phi} \mathcal{F}^{(n)} \mathbf{A}_{\Phi}^{-1} \nabla_{\Phi} \mathcal{F}^{(n)}} \mathbf{d}^{(n)}, \quad (48b)$$

$$\Phi^{(n+1)} = \Phi^{(n)} - \lambda \mathbf{d}^{(n)}. \quad (48c)$$

Here, we omit the explicit dependence of the free energy $\nabla_{\Phi} \mathcal{F}^{(n)} = \nabla_{\Phi} \mathcal{F}(\vec{\mathcal{R}}^{(n)}, \Phi^{(n)}, \{\vec{a}_i\})$ for simplicity.

Since we can compute the Hessian matrix even of the fourth root problem, we can combine the two approaches of the nonlinear change of variable and the preconditioner to achieve a minimization constrained only on the positive defined Φ with the smallest condition number.

D. Hessian in the $\vec{\mathcal{R}}$ vector

The analysis on the minimization conducted so far investigates only the minimization problems faced with the Φ parameter of the free energy. This is usually the most problematic part of the minimization, since, being a matrix, Φ has many more degrees of freedom than the centroid positions. Furthermore, the centroid positions are defined in the unit cell, while the force constant matrix is a supercell quantity. However, for generality, it is very easy to provide an approximation also for the Hessian matrix of the free energy with respect to the $\vec{\mathcal{R}}$ variables:

$$A_{\vec{\mathcal{R}}}^{ab} = \left. \frac{\partial^2 \mathcal{F}(\vec{\mathcal{R}}, \Phi, \{\vec{a}_i\})}{\partial \mathcal{R}_a \partial \mathcal{R}_b} \right|_{\Phi}. \quad (49)$$

Differences and similarities between this expression and the free-energy Hessian studied in Ref. [35] are discussed in Appendix C. Since we are both neglecting mixed terms in the Hessian, and we are taking an approximated Hessian also for the Φ minimization, we chose

$$\mathbf{A}_{\vec{\mathcal{R}}} = \Phi_{\text{SG}}. \quad (50)$$

This expression is correct when the $\vec{\mathcal{R}}$ and the Φ degrees of freedom are simultaneously minimized. Moreover, Eq. (50) provides a good preconditioner as it is always positive defined,

and it does not require any additional computational effort to the algorithm.

The eigenvalues of Φ are related to the square of the phonon frequencies for harmonic systems, therefore we can approximate the condition number on the $\vec{\mathcal{R}}$ variables as

$$C_{\vec{\mathcal{R}}} \sim \left(\frac{\omega_{\text{max}}}{\omega_{\text{min}}} \right)^2. \quad (51)$$

This is not as pathological as the condition number seen on the Φ minimization. However, we can introduce a preconditioner in the same way as described in Sec. VC to handle easier minimization in low-symmetry systems, as molecular crystals, where also many centroid degrees of freedom must be optimized, and the condition number (51) can be of the order of 10^6 . Preconditioning also the $\vec{\mathcal{R}}$ variables allows one to have a dimensionless step λ for the minimization algorithms [Eqs. (47) and (48c)], with a clear advantage of reducing the human time necessary to optimize the two λ steps for the Φ and $\vec{\mathcal{R}}$ minimizations. We remark that the terms in the Hessian matrix obtained by the mixed derivatives in $\vec{\mathcal{R}}$ and Φ are neglected.

The new SCHA algorithm flowchart is shown in Fig. 2.

VI. TESTS ON ICE XI (H₂O)

In order to present the impressive enhancement in the minimization procedure obtained thanks to the combination of the preconditioning with the root representation, we report the calculation on phase XI of ice. The difficulties of applying the SSCHA to this structure arise due to the presence of both hard covalent intramolecular bonds and soft intermolecular H bonds, resulting in a broad phonon spectrum.

Ice XI is the proton ordered phase of common ice [36] that is stable below 72 K. This is a typical prototype of a molecular crystal also for the low symmetry of the structure. It belongs to the $Cmc2_1$ group, with four symmetry operations. The unit cell contains four water molecules (12 atoms). The number of symmetry independent SCHA degrees of freedom is 11 for the inner coordinates (Wyckoff positions) and 159 for the unit-cell force-constant matrix. All the 11 inner coordinates, as well as the 159 parameters in the force-constant matrix, are allowed to move in the SSCHA.

We restricted the calculation to the unit cell, as we presented this example as a test case; however, all the methods developed here (both the minimization strategy and the stress tensor computation) are defined on an arbitrary large supercell.

In this section, we use a classical force field that explicitly includes anharmonicity of the water molecule to compute energies and forces. The model is q-SPC/FW+anh [45].

A. Stress tensor test

Here we test the anharmonic effects on the stress tensor with q-SPC/FW+anh. Equation (20) can be checked by performing the numerical derivative of the SCHA free energy at different volumes. In Fig. 3 we report the SCHA free energy as a function of the system volume, with a polynomial fit. The cell is deformed with an isotropic expansion of the volume so that the obtained pressure as the derivative of the free energy versus the volume can be compared with 1/3 of the stress tensor trace

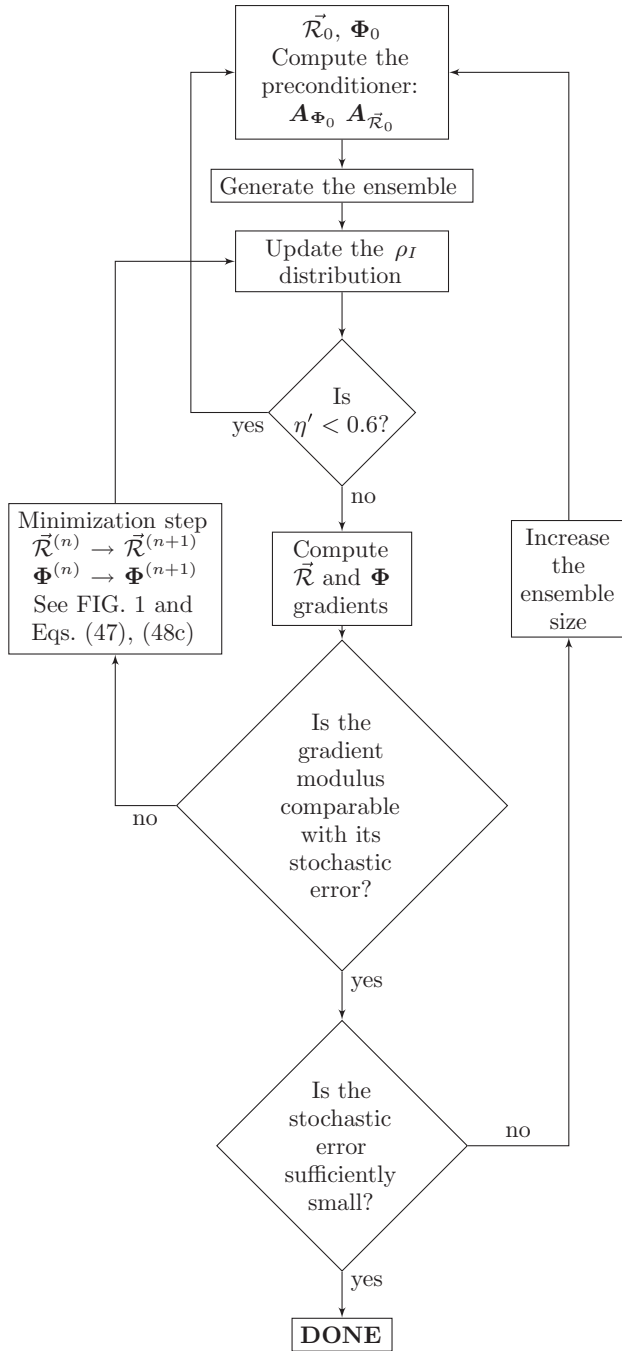


FIG. 2. Flowchart of the new SSCHA implementation. The minimization step can be expanded by using the root4 algorithm shown in Fig. 1. In this case, the preconditioner \mathbf{A}_{Φ_0} should be replaced with $\mathbf{A}_{\sqrt[4]{\Phi_0}}$ in the initial step. The minimization step is performed using the CG algorithm as long as the error is much greater than the stochastic noise, then the last steps are performed using SD. This prevents error propagation in the conjugation due to the correlated noise introduced by the importance sampling reweighting procedure.

of Eq. (20). The fit on the SCHA free energy is then used to evaluate the pressure as a function of the volume:

$$P = -\frac{d\mathcal{F}}{d\Omega} = \frac{1}{3} \sum_{\alpha=x,y,z} P_{\alpha\alpha}. \quad (52)$$

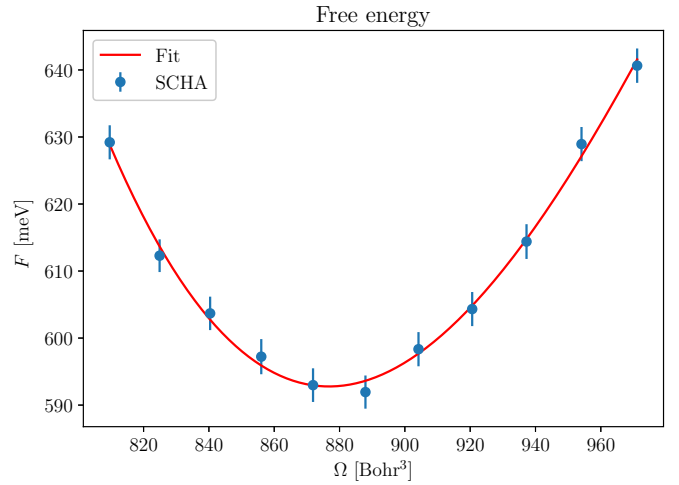


FIG. 3. SCHA free energy as a function of the volume. The unit cell is kept fixed, while only an isotropic scaling factor is considered. The solid line represents a cubic fit. The simulation is performed at $T = 100$ K.

In Fig. 4 we compare the SCHA pressure obtained both as indicated in Eq. (52) and as the opposite of the total derivative of the free energy. The stochastic average of the stress tensors $\langle P_{\alpha\beta}^H(\vec{R}, \{\vec{a}_i\}) \rangle_{\rho}$ is also reported, showing how the pressure cannot be considered as a physical observable to be computed in analogy to what is done for general operators: $P_{\alpha\beta} \neq \langle P_{\alpha\beta}^H \rangle_{\rho}$; in fact, this neglects the kinetic contribution of the vibrations. It is necessary to compute it as the derivative of the free energy, as done in Eq. (20). The pressure P_{cla} without quantum effects at $T = 0$ is also reported, and can be computed as 1/3 of the trace of the stress tensor in the classical equilibrium centroid

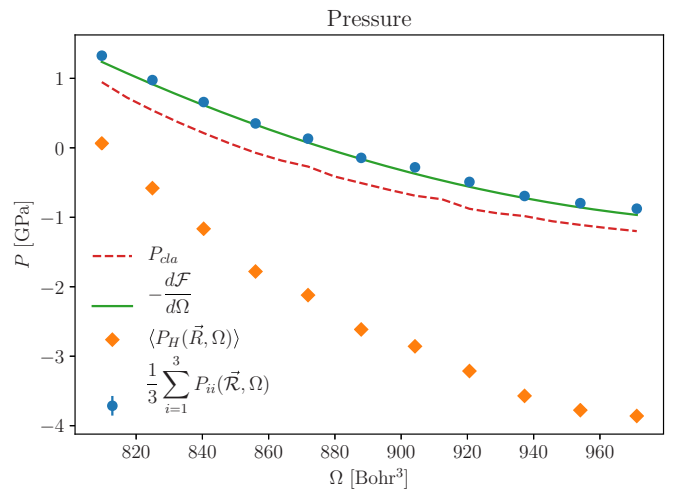


FIG. 4. The figure compares the pressure computed with Eq. (20) (blue circles), the classical pressure P_{cla} obtained neglecting thermal and quantum fluctuations (red dashed line), the average of the classical pressures over the SCHA ensemble (orange diamonds), and the analytical derivative of the free energy fit reported in Fig. 3 (solid green line). The simulation is performed at $T = 100$ K.

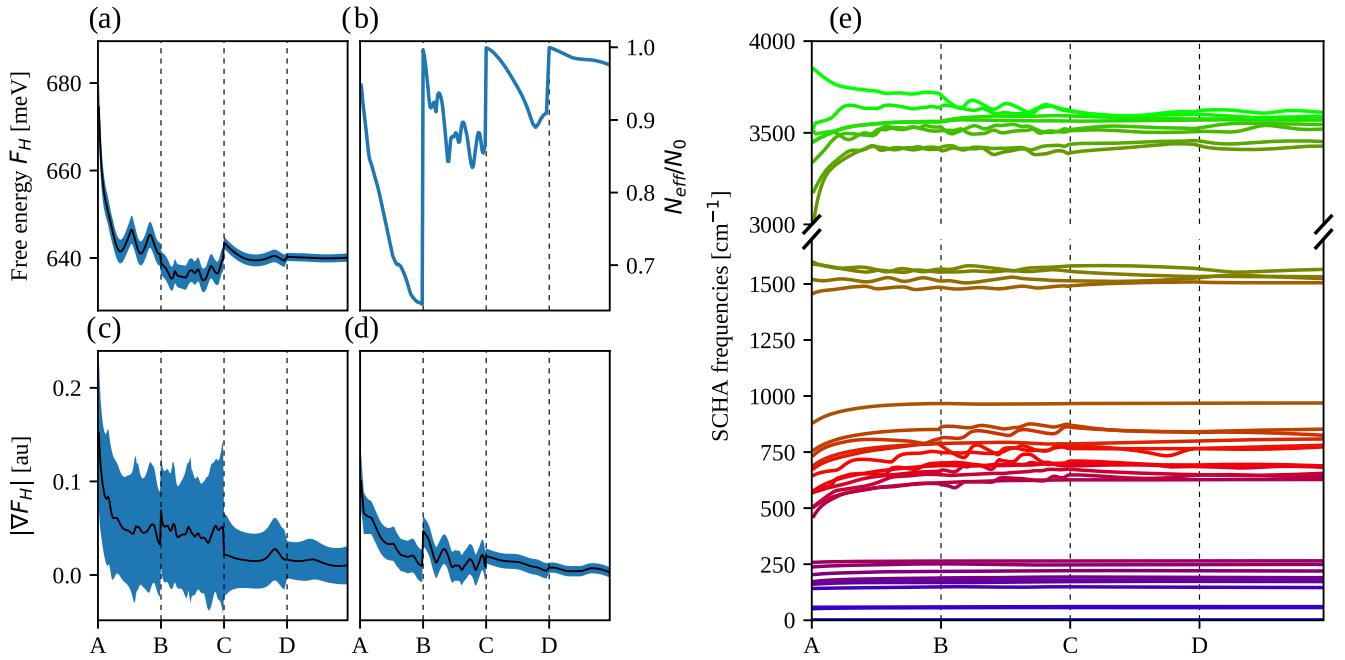


FIG. 5. Minimization progress starting from the harmonic result with the preconditioning linear change of variables. (a) The free energy. (b) The Kong-Liu effective sample size ratio, defined in Eq. (27). (c) Modulus of the free-energy gradient with respect to the dynamical matrix Φ . (d) Modulus of the free-energy gradient with respect to the centroids \vec{R} . (e) Frequencies obtained from the eigenvalues of the SCHA Φ matrix, as they evolve during the minimization. Preconditioning uniformly converges all the frequencies, achieving the final result much faster. Panels (a), (c), and (d) contain the stochastic error. For the two gradients, the error is computed as the norm of the error on each component of the gradient, to make it invariant with respect to the basis used to describe the Φ .

positions:

$$P_{\text{cla}} = -\frac{1}{3} \sum_{\alpha=x,y,z} \frac{1}{\Omega} \frac{\partial V(\vec{R}_0, \{\vec{a}_i\})}{\partial \varepsilon_{\alpha\alpha}}, \quad (53)$$

where \vec{R}_0 is defined as

$$\left. \frac{\partial V(\vec{R}, \{\vec{a}_i\})}{\partial \vec{R}} \right|_{\vec{R}=\vec{R}_0} = 0. \quad (54)$$

B. Tests on the new minimization algorithm

A typical SCHA run with the precondition is reported in Fig. 5. The $\rho_{\vec{r}, \Phi}$ ensemble is reextracted four times. The first two times (A and B) 2500 configurations were used, 10 000 were used in the C step, and 20 000 were used in D. As clearly reported, the frequencies of the dynamical matrix converge uniformly to the final result, as we expect from the preconditioning, and we achieve a converged good result after only two steps.

The comparison of the performances between the nonlinear change $\Phi \rightarrow \sqrt[4]{\Phi}$ and the preconditioning is reported in Fig. 6. As a reference, the SCHA run without the nonlinear change of variable and without preconditioning is also reported. The simulations are compared at $T = 100$ K. It is clear that both methods greatly outperform the standard algorithm. The harmonic dynamical matrix around the static equilibrium positions (neglecting quantum and thermal fluctuations) is used as a starting point, according to what is usually done in *ab initio* calculations [16,17,34]. The q-SPC/FW+anh harmonic

dynamical matrix is close to the SCHA result, as seen by the low value of the free-energy gradient with respect to Φ ,

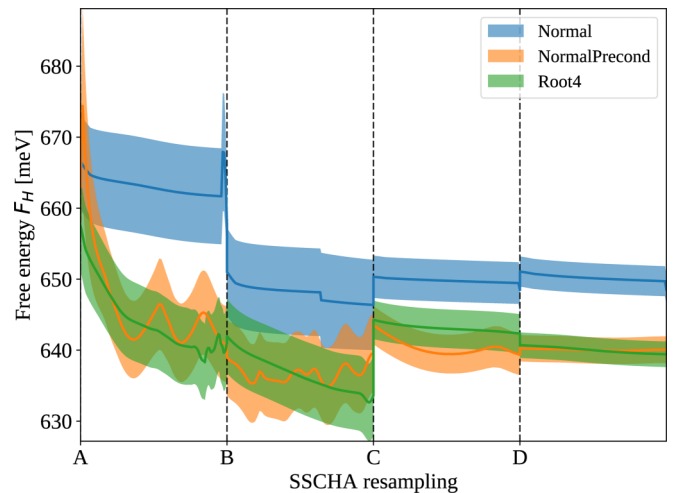


FIG. 6. Comparison between the different methods described here. The free energy is shown as a function of the number of configurations used for the stochastic evaluation together with its stochastic error. The first two calculations have 2500 configurations each. The third (C) is with 10 000. A final calculation is performed as a reference (D) with 20 000 configurations to check the convergence of the previous ones. As shown, the minimization without the preconditioning or the nonlinear change of variables is not able to get a converged result even using an overall of 35 000 configurations, preventing the old SCHA from being used with any *ab initio* technique in this kind of system.

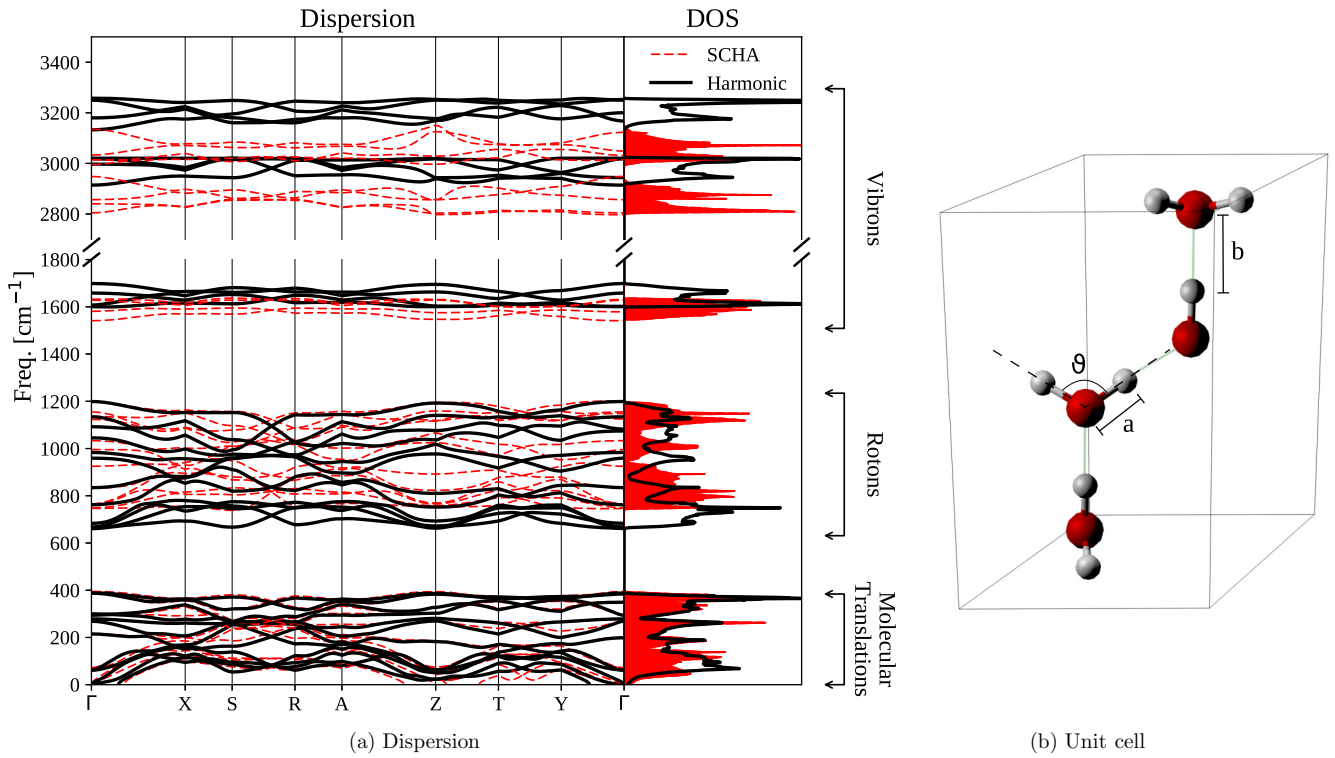


FIG. 7. (a) Comparison between harmonic (black solid lines) and SCHA (red dashed lines) dispersion and density of states. (b) Unit cell of the ice XI structure. The parameters a , b , and ϑ represent, respectively, the covalent OH bond, the hydrogen bond, and the molecule angle. Their average value as a function of the temperature is reported in Table I.

compared with its stochastic error, already in the first step of Fig. 5. However, the standard minimization is not able to further minimize the system.

The success of the SCHA implementation on this force field paves the way to its systematic utilization for the study of water and any other complex system with many degrees of freedom.

In the next section, we show the capabilities of our method in a more realistic first-principles potential.

VII. *AB INITIO* SIMULATION ON ICE XI

Encouraged by the success of the SCHA implementation on the q-SPC/FW+anh force field, we report also the SCHA results on a realistic DFT potential. The converged SCHA phonon dispersion ($T = 0$ K) is compared to the harmonic one in Fig. 7. The calculation of energies and forces required to minimize the SCHA free energy, as well as the computation of the harmonic dynamical matrix, are performed *ab initio* with DFT, Perdew-Burke-Ernzerhof (PBE) exchange-correlation functional [46], and ultrasoft pseudopotentials [47] from the *pslibrary* [48], as implemented in the QUANTUMESPRESSO suite [49,50]. The SCHA dispersion is computed in the unit cell with 13 000 overall configurations and a wave-function cutoff of 45 Ry (360 Ry being the charge density cutoff), then the difference between the harmonic and anharmonic dynamical matrices is extrapolated in a $3 \times 3 \times 3$ supercell, and the harmonic dispersion Φ_0 is added:

$$\Phi^{(3 \times 3 \times 3)} = \Phi_0^{(3 \times 3 \times 3)} + (\Phi^{(1 \times 1 \times 1)} - \Phi_0^{(1 \times 1 \times 1)})^{(3 \times 3 \times 3)}. \quad (55)$$

The harmonic phonon dispersion obtained interpolating the dynamical matrices converges already in a $2 \times 2 \times 2$ supercell, with a wave-function cutoff of 80 Ry (640 Ry being the charge density cutoff). The SCHA auxiliary dynamical matrix Φ is not directly related to the anharmonic phonon dispersion and, in general, a more sophisticated calculation is required to extract the real phonon frequencies in the SCHA approximation [35]. However, it is found [35] that the static phonon dispersion can be obtained as a perturbative series, the leading order of which is given by the Φ matrix itself plus a “bubble” correction. It has been found in many systems with hydrogen [34,51] that the bubble correction is much lower than the Φ contribution. As an explicative case, here we neglect this correction. It is, however, worth noticing that the developments presented in Sec. V A do not affect the bubble computation as reported in Ref. [35] since it depends only on the converged result and not on the particular minimization strategy. Therefore, we report the anharmonic phonon dispersion and density of states approximated by directly interpolating the Φ matrix after the SSCHA optimization in Fig. 7.

All phonon modes below 500 cm^{-1} (molecular translations) are almost unaffected by the anharmonicity. The two upper bands corresponding to symmetric and asymmetric stretching suffer a redshift, together with the band around 1600 cm^{-1} (molecular bending). These modes are well described by molecular vibrons, and the observed redshift is a general property of the water molecule [52]. Also, the lowest part of the molecular rotations (the bands between 600 and 1200 cm^{-1}) are blueshifted. This blueshift of the lowest modes is indeed very interesting since it involves intermolecular modes. Such

TABLE I. Anharmonic effects on the crystal structure predicted by the DFT-PBE at three temperatures. The average intramolecular OH distance (covalent bond), the average H-bond distance, and the water molecule mean angle, as reported in Fig. 7.

	Harmonic	0 K	150 K	300 K
a (OH-covalent)	1.00835 Å	1.0159 Å	1.0148 Å	1.0131 Å
b (H-bond)	1.6769 Å	1.673 Å	1.674 Å	1.6765 Å
ϑ (HOH angle)	106.715°	106.950°	106.951°	106.886°

an effect of anharmonicity is typical of this solid structure of ice, and cannot be predicted just studying anharmonicity in the isolated water molecule or the dimer. Moreover, the Debye temperature of these bands is far above room temperature, invalidating the dispersion obtained with classical molecular dynamics since zero-point motion has a predominating role in these lattice oscillations.

Also the average atomic positions are affected as reported in Table I. Here, the quantum fluctuations slightly stretch the water molecule, each covalent OH bond increases its length by almost 0.7%, and the molecular angle widens by 0.2%.

Even if the anharmonic molecular stretch can seem negligible compared to what it has been predicted to be for a high-pressure molecular phase of hydrogen [33], a difference of 1% in the OH covalent bond has a great contribution to the energy. As a test, the SCHA average structure can be used for a classical DFT calculation, where the classical pressure is found to be 1 GPa lower (negative) than its value in the equilibrium positions, suggesting that the anharmonic relaxation of the centroid positions may significantly affect the pressure and, consequently, the equilibrium volume.

The stress tensor calculation can be used to optimize the unit cell considering both thermal and quantum effects. The most advanced calculations to include these effects without involving PIMD in water have been performed within the QHA [37,38]. In this scheme the total pressure is obtained expanding the BO energy surface as a quadratic function around its minimum at each volume. Then the exact free energy of the approximated BO surface can be computed analytically:

$$\mathcal{F}_{\text{QHA}}(\vec{R}_c, \{\vec{a}_i\}) = V(\vec{R}_c, \{\vec{a}_i\}) + \sum_{\mu=1}^{3N} \left[\frac{\hbar \tilde{\omega}_{\mu}(\vec{R}_c, \{\vec{a}_i\})}{2} + \frac{1}{\beta} \ln(1 - e^{-\beta \hbar \tilde{\omega}_{\mu}(\vec{R}_c, \{\vec{a}_i\})}) \right], \quad (56)$$

where $\tilde{\omega}_{\mu}$ are the harmonic frequencies of the BO surface. The QHA free energy F_{QHA} is obtained minimizing the functional \mathcal{F}_{QHA} at fixed volume and temperature:

$$F_{\text{QHA}}(T, \{\vec{a}_i\}) = \min_{\vec{R}_c} \mathcal{F}_{\text{QHA}}(T, \vec{R}_c, \{\vec{a}_i\}). \quad (57)$$

The QHA pressure is obtained by differentiating the free energy with respect to a uniform volume deformation:

$$P_{\text{QHA}} = -\frac{dF_{\text{QHA}}}{d\Omega}. \quad (58)$$

In complex systems with many degrees of freedom, like in ice, the minimization in Eq. (57) is computationally very

expensive, since it requires the calculation of the gradient of the free energy (that depends on the harmonic dynamical matrix) with respect to any possible atomic displacement. This involves the calculation of a third-order derivative of the BO total energy for each minimization step [53]. Differences and analogies of QHA and SCHA approaches are discussed in Appendix D. The QHA implementation with the full atomic coordinates relaxation in H₂O system has never been performed, and usually the QHA free energy is approximated with $\vec{R}_c = \vec{R}_0$: the minimum of the BO energy. The pressure in Eq. (57) is computed numerically taking finite differences between the QHA free energies at several volumes. A more convenient way to compute the QHA pressure is to consider the harmonic frequencies as a linear function of the volume:

$$\tilde{\omega}_k(\Omega) = \tilde{\omega}_k(\Omega_0) \left[1 - \frac{\Omega - \Omega_0}{\Omega_0} \gamma_k \right], \quad (59)$$

where the γ_k are the Grüneisen parameters. Then the QHA pressure can be easily obtained at any temperature:

$$P_{\text{QHA}} = P_H(\Omega) - \sum_{\mu=1}^{3N} \frac{\hbar \omega_{\mu} \gamma_{\mu}}{2\Omega} \frac{1}{\tanh\left(\frac{\beta \hbar \omega_{\mu}}{2}\right)}. \quad (60)$$

The comparison between QHA and the SSCHA pressure calculations as a function of temperature is reported in Fig. 8(a). Both the calculations have been performed in the unit cell, allowing for a direct comparison between the SCHA and the QHA result. Moreover, we checked the QHA pressure convergence versus the supercell size, and found that the difference between the QHA in a $3 \times 3 \times 2$ supercell and in the unit cell was much lower than the SCHA stochastic error itself.

The experimental fit on the elastic bulk modulus and the volume expansivity have been used to compare the QHA and SSCHA pressures at a fixed volume. As clearly shown, the QHA pressure is shifted by about 4 kbars with respect to the SSCHA result. This is two times bigger than the whole pressure range between 0 and 300 K. In Fig. 8(b) the comparison between the QHA, the SSCHA, and the experimental results is reported. All the pressures are shifted with respect to their zero-temperature value. The SSCHA zero-temperature pressure has been obtained by fitting the SSCHA points with the experimental curve. The experimental data have been obtained from the fit reported in Ref. [54].

The QHA grasps the qualitative behavior of the pressure, including the low-temperature negative thermal expansion [55], but it deviates from the experiments at temperatures above 80 K. This effect has been associated with the entropy contribution of the proton disorder of the ice Ih with respect to the ice XI, that is not accounted for in the simulations [55]. However, the SSCHA result corrects the QHA estimation of the pressure by a significant amount, matching perfectly the experiments, suggesting that the underestimation of the pressure at high temperature can be simply explained as a failure of the QHA. This indicates that anharmonic effects beyond QHA play an important role in reproducing the physical properties of ice at temperatures above 80 K.

We computed also the whole SCHA anharmonic stress tensor. The effect of fluctuations on pressure anisotropy is much smaller than that on the isotropic pressure. Indeed, the

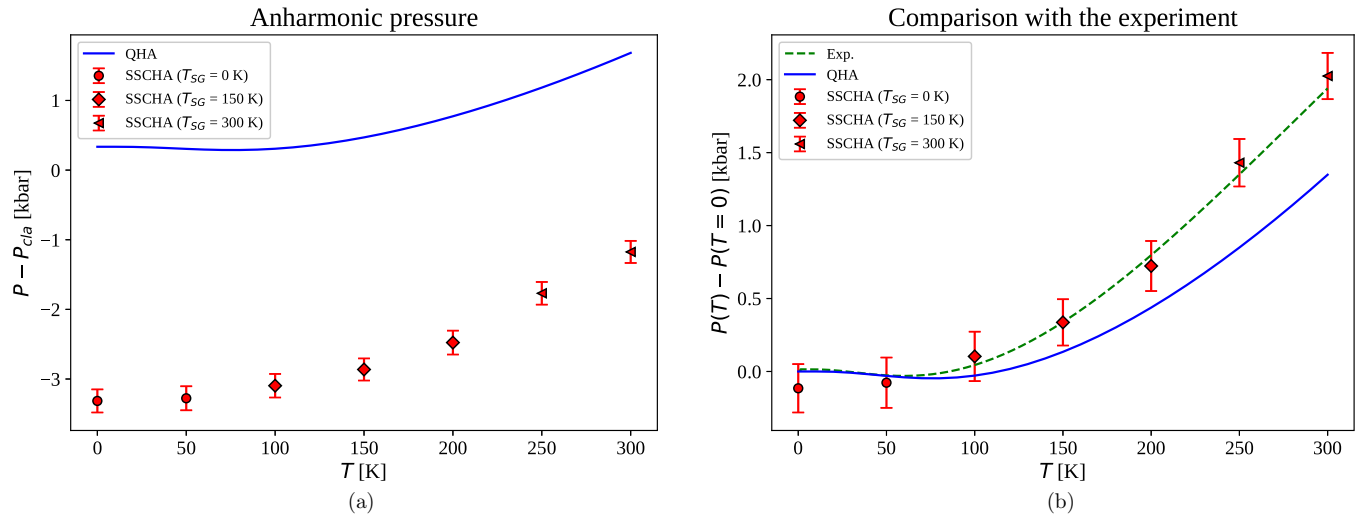


FIG. 8. (a) Comparison between the QHA and the SSCHA pressure as a function of the temperature at fixed volume. The zero value of the pressure is P_{cla} , obtained as $1/3$ of the stress tensor trace at $T = 0$ neglecting quantum fluctuations. (b) Comparison between QHA and SSCHA simulations and the fit of the experimental results from Ref. [54] (ice Ih). The SSCHA results are computed at temperatures of 0, 150, and 300 K with 40 000 stochastic configurations. The other temperatures reported have been obtained through reweighting [16], therefore their stochastic error is correlated.

computed pressure anisotropy is comparable with the statistical error bars.

Another interesting feature of ice at low temperature is the anomalous volume isotope effect: the D_2O equilibrium volume is bigger than the H_2O one. This effect has been recently studied within the QHA [37,38]. In particular, Pamuk *et al.* [37] showed how the QHA result systematically overestimates this effect with several DFT functionals but it depends slightly on the chosen functional, on the difference between ice Ih and XI, and on the q -grid interpolation. The experimental [39] difference between the two volumes at 10 K is about 0.09%, while the difference between the QHA equilibrium volume and the SSCHA one is 1.8%. Therefore, the isotope volume effect is a tiny correction with respect to the ZPE contribution on the equilibrium volume and the difference between the SSCHA and the QHA.

VIII. CONCLUSIONS

The study of quantum anharmonic effects in complex crystals with lots of degrees of freedom, e.g., molecular crystals, is a major challenge that impacts many domains of physics and chemistry, including high-pressure phases of hydrogen, water anomalies, thermoelectric materials, charge-density waves, ferroelectrics, multiferroics, and so on. In this paper, we derive an expression for the anharmonic contribution to the stress tensor in the SCHA theory. This correction is very important for accurate pressure estimations and phase-diagram computations and paves the way for isobaric unit-cell relaxation. We further improved the stochastic implementation of the SCHA theory to apply it in complex crystals with a large number of degrees of freedom. This aim has been achieved thanks to a preconditioning on the free-energy minimization algorithm, based on an analytical guess of the Hessian matrix of both the force-constant matrix and the central nucleus positions, and with a nonlinear change of variables that restricts the space

of allowed dynamical matrices only to the positive defined ones.

The algorithm is benchmarked with the phase XI of ice, the proton-ordered phase of common ice, a prototype molecular crystal. The quantum ZPM and anharmonicity are proven to affect the phonon dispersion both in the molecular and in the intermolecular modes. Also, the O-H and H-H bond distances are slightly affected by anharmonicity. The importance of the nonperturbative SCHA contribution to the pressure in this system has been benchmarked in q-SPC/FW+anh and calculated with *ab initio* DFT PBE, where the quantum fluctuations at 0 K are shown to affect the equilibrium volume by 1.8%.

The thermal expansion of the system has been computed within both the QHA and the SSCHA. The QHA is found to miscalculate both the ZPM contribution to the equilibrium volume (with the wrong sign) and the effect of the thermal fluctuations at temperatures above 80 K. The latter discrepancy was associated with the proton disorder of phase Ih of ice. However, we found the SSCHA to correct this effect, and to exhibit an excellent agreement with the experiments, unveiling that anharmonic effects behind QHA are crucial to correctly describe the thermodynamic properties of ice.

The cell-relaxation and stress calculation here developed paves the way to a refreshed quantitative and accurate study of anharmonic effects on water, like the anomalous isotope volume, the equilibrium isotope fraction, the negative thermal expansion, and the high-pressure phase diagram. More generally, the developed stress tensor derivation and the improved minimization algorithm make the SSCHA an efficient method to calculate quantum and thermal anharmonic effects on complex systems with many degrees of freedom.

ACKNOWLEDGMENTS

I.E. acknowledges financial support from the Spanish Ministry of Economy and Competitiveness (Grant No. FIS2016-

76617-P). We acknowledge a CINECA award under the IS-CRA initiative (Grant No. HP10BLTB9A); PRACE; IDRIS; CINES; and TGCC under EDARI Project No. A0030901202 for the use of high-performance computing resources and support.

APPENDIX A: STRESS TENSOR DERIVATION

To easily compute the derivative of the SCHA free-energy functional with respect to the strain tensor it is convenient to use the formalism introduced by Bianco *et al.* [35]. The average of a generic observable can be written as

$$\langle O \rangle_{\rho_{\vec{R}, \Phi}} = \sqrt{\det \left(\frac{\Upsilon}{2\pi} \right)} \int O(\vec{R} + \vec{u}, \{\vec{a}_i\}) \times \exp \left(-\frac{1}{2} \vec{u} \Upsilon \vec{u} \right) d^{3N} u. \quad (\text{A1})$$

In order to normalize the Gaussian integral a change of variable can be applied, so that

$$u_s^\alpha = \sum_{\mu} J_{\mu_s}^\alpha y_{\mu}, \quad J_{\mu_s}^\alpha = \frac{e_{\mu_s}^\alpha}{\sqrt{M_s}} \sqrt{\frac{\hbar(1+n_{\mu})}{2\omega_{\mu}}},$$

$$(\Upsilon^{-1})_{st}^{\alpha\beta} = \sum_{\mu} J_{\mu_s}^\alpha J_{\mu_t}^\beta. \quad (\text{A2})$$

Then we have

$$\langle O \rangle_{\rho_{\vec{R}, \Phi}} = \int O(\vec{R} + \mathbf{J}\vec{y}, \{\vec{a}_i\}) [dy],$$

$$[dy] = \prod_{\mu=1}^{3N} \frac{\exp \left(-\frac{y_{\mu}^2}{2} \right)}{\sqrt{2\pi}} dy_{\mu}. \quad (\text{A3})$$

Since we are deriving the $\mathcal{F}_{\vec{R}}$ functional [Eq. (9)], the Hellman-Feynman theorem allows us to neglect the changes introduced by the strain on the dynamical matrix. Only \vec{R} is affected by the deformation, according to Eq. (19). Therefore we have

$$\frac{d \langle O(\vec{R}, \{\vec{a}_i\}) \rangle_{\rho_{\vec{R}}} }{d\varepsilon_{\alpha\beta}} = \left. \frac{\partial \langle O(\vec{R}, \{\vec{a}_i\}) \rangle_{\rho_{\vec{R}}} }{\partial \varepsilon_{\alpha\beta}} \right|_{\Phi=\Phi(\vec{R})}, \quad (\text{A4})$$

$$\frac{\partial \langle O \rangle_{\rho_{\vec{R}}} }{\partial \varepsilon_{\alpha\beta}} = \frac{\partial}{\partial \varepsilon_{\alpha\beta}} \int O(\vec{R}(\boldsymbol{\varepsilon}) + \mathbf{J}\vec{y}, \{\vec{a}_i(\boldsymbol{\varepsilon})\}) [dy]$$

$$= \int \left(\sum_{s\gamma} \frac{\partial O}{\partial R_s^\gamma} \frac{\partial R_s^\gamma}{\partial \varepsilon_{\alpha\beta}} + \sum_{i\gamma} \frac{\partial O}{\partial a_i^\gamma} \frac{\partial a_i^\gamma}{\partial \varepsilon_{\alpha\beta}} \right) [dy]. \quad (\text{A5})$$

Note that the observable $O(\vec{R})$ is derived with respect to its argument, i.e., the atom positions in the ensemble configuration \vec{R} , not the centroid position \vec{R} . This happens because the $\vec{R}(\boldsymbol{\varepsilon})$ appears linearly in the configuration position of O after the change of variable:

$$\frac{\partial \langle O \rangle_{\rho_{\vec{R}}} }{\partial \varepsilon_{\alpha\beta}} = \frac{1}{2} \sum_s \left(\mathcal{R}_s^\beta \left\langle \frac{\partial O}{\partial R_s^\alpha} \right\rangle_{\rho_{\vec{R}}} + \mathcal{R}_s^\alpha \left\langle \frac{\partial O}{\partial R_s^\beta} \right\rangle_{\rho_{\vec{R}}} \right)$$

$$+ \left\langle \sum_{i\gamma} \frac{\partial O}{\partial a_i^\gamma} \frac{\partial a_i^\gamma}{\partial \varepsilon_{\alpha\beta}} \right\rangle_{\rho_{\vec{R}}}. \quad (\text{A6})$$

The free-energy functional is

$$\mathcal{F}_{\vec{R}} = F_{\Phi(\vec{R})} + \langle V - \mathcal{V}_{\vec{R}, \Phi(\vec{R})} \rangle_{\rho_{\vec{R}}}, \quad (\text{A7})$$

where $\Phi(\vec{R})$ is the dynamical matrix that minimizes $\mathcal{F}(\vec{R}, \Phi, \{\vec{a}_i\})$ fixing the average atomic positions. The first term, $F_{\Phi(\vec{R})}$ is an explicit function only of the SCHA dynamical matrix, and therefore does not contribute to the derivative. The latter average can be derived thanks to Eq. (A6):

$$\frac{\partial \langle V \rangle_{\rho_{\vec{R}}} }{\partial \varepsilon_{\alpha\beta}} = -\frac{1}{2} \sum_{s=1}^N \left(\mathcal{R}_s^\beta \langle f_s^\alpha \rangle_{\rho_{\vec{R}}} + \mathcal{R}_s^\alpha \langle f_s^\beta \rangle_{\rho_{\vec{R}}} \right) - \Omega \langle P_{\alpha\beta}^H \rangle_{\rho_{\vec{R}}}, \quad (\text{A8})$$

where $P_{\alpha\beta}^H$ is the BO stress tensor. In fact the last term of Eq. (A6) is the average of the derivatives of the BO energy when the strain is applied to the unit cell. The ‘‘harmonic’’ term can be computed in a similar way:

$$\frac{\partial \langle \mathcal{V}_{\vec{R}, \Phi} \rangle_{\rho_{\vec{R}}} }{\partial \varepsilon_{\alpha\beta}} \Big|_{\Phi} = \frac{1}{2} \sum_s \left(\mathcal{R}_s^\beta \left\langle \frac{\partial \mathcal{V}_{\vec{R}, \Phi}}{\partial R_s^\alpha} \right\rangle_{\rho_{\vec{R}}} + \mathcal{R}_s^\alpha \left\langle \frac{\partial \mathcal{V}_{\vec{R}, \Phi}}{\partial R_s^\beta} \right\rangle_{\rho_{\vec{R}}} \right)$$

$$+ \left\langle \frac{\partial \mathcal{V}_{\vec{R}, \Phi}}{\partial \varepsilon_{\alpha\beta}} \right\rangle. \quad (\text{A9})$$

In the same way as done for the BO energy surface, it is possible to introduce the harmonic stress tensor as

$$P_{\alpha\beta}^H = -\frac{1}{\Omega} \left\langle \frac{\partial \mathcal{V}_{\vec{R}, \Phi}}{\partial \varepsilon_{\alpha\beta}} \right\rangle_{\rho_{\vec{R}}} = \frac{1}{2\Omega} \sum_s \langle f_{\mathcal{H}_s}^\alpha u_s^\beta + f_{\mathcal{H}_s}^\beta u_s^\alpha \rangle_{\rho_{\vec{R}}}$$

$$= -\frac{1}{\Omega} \sum_{\mu=1}^{3N} \sum_{s=1}^N \frac{\hbar\omega_{\mu}}{2 \tanh \left(\frac{\beta\hbar\omega_{\mu}}{2} \right)} e_{\mu_s}^\alpha e_{\mu_s}^\beta, \quad (\text{A10})$$

$$\frac{\partial \langle \mathcal{V} \rangle_{\rho_{\vec{R}}} }{\partial \varepsilon_{\alpha\beta}} = -\frac{1}{2} \sum_{s=1}^N \left(\mathcal{R}_s^\beta \langle f_{\mathcal{H}_s}^\alpha \rangle_{\rho_{\vec{R}}} + \mathcal{R}_s^\alpha \langle f_{\mathcal{H}_s}^\beta \rangle_{\rho_{\vec{R}}} \right) - \Omega P_{\alpha\beta}^H. \quad (\text{A11})$$

The first term is zero (the harmonic forces $f_{\mathcal{H}}$ are odd, while the probability distribution $\rho_{\vec{R}}$ is even). However, we keep it as it helps to increase the numerical accuracy [16], as we can combine it with Eq. (A8) to exploit the correlation between f_s^α and $f_{\mathcal{H}_s}^\alpha$ to reduce the statistical noise on the average. In a pure harmonic crystal also the quantities $P_{\alpha\beta}^H$ and $P_{\alpha\beta}^H$ are correlated. Therefore, the final expression of the pressure can be written as follows:

$$P_{\alpha\beta} = \langle P_{\alpha\beta}^H - \frac{1}{2\Omega} \sum_{s=1}^N (f_{\mathcal{H}_s}^\alpha u_s^\beta + f_{\mathcal{H}_s}^\beta u_s^\alpha) \rangle_{\rho_{\vec{R}}}$$

$$+ \frac{1}{2\Omega} \sum_{s=1}^N \left(\mathcal{R}_s^\beta \langle f_s^\alpha - f_{\mathcal{H}_s}^\alpha \rangle_{\rho_{\vec{R}}} + \mathcal{R}_s^\alpha \langle f_s^\beta - f_{\mathcal{H}_s}^\beta \rangle_{\rho_{\vec{R}}} \right). \quad (\text{A12})$$

The last term is zero if the free energy has been minimized also with respect to the \vec{R} variables (as the average of the BO forces is the SCHA force acting on each atom, it is zero in the equilibrium).

APPENDIX B: DETAILED CALCULATION FOR THE HESSIAN MATRIX

The real and trial classical forces acting on each configuration identified by the displacements \vec{u} are

$$f_s^\alpha = -\frac{\partial V}{\partial u_s^\alpha} = -\sum_{t\beta} K_{st}^{\alpha\beta} u_t^\beta, \quad (\text{B1a})$$

$$f_{\mathcal{H}_s}^\alpha = -\frac{\partial \mathcal{V}_{\vec{\mathcal{R}}, \Phi}}{\partial u_s^\alpha} = -\sum_{t\beta} \Phi_{st}^{\alpha\beta} u_t^\beta. \quad (\text{B1b})$$

Defining $\vec{\delta f} = \vec{f} - \vec{f}_{\mathcal{H}}$ we have

$$\langle \delta f_s^\alpha u_t^\beta \rangle_{\rho_{\mathcal{H}}} = -\sum_{n\eta} (K_{sn}^{\alpha\eta} - \Phi_{sn}^{\alpha\eta}) \langle u_n^\eta u_t^\beta \rangle_{\rho_{\vec{\mathcal{R}}, \Phi}}. \quad (\text{B2})$$

From now on, we drop the subscript $\rho_{\vec{\mathcal{R}}, \Phi}$ for each average, and consider all the averages computed with respect to the trial density matrix. We further simplify the notation, introducing one index for each Cartesian and atomic coordinate, so $v_s^\alpha \rightarrow v_a$. In this new notation Eq. (B2) reads

$$\langle \delta f_a u_b \rangle = -\sum_{c=1}^{3N} (K_{ac} - \Phi_{ac}) \langle u_c u_b \rangle. \quad (\text{B3})$$

The average of the product between two displacements of a Gaussian distributed variable is the covariance between the two displacements [Eq. (13a)]:

$$\langle u_c u_b \rangle = (\Upsilon^{-1})_{cb} = \frac{1}{\sqrt{M_c M_b}} \sum_{v=1}^{3N} e_v^c e_v^b a_v^2, \quad (\text{B4})$$

where we introduce the mode length a_μ :

$$a_\mu = \sqrt{\frac{\hbar}{2\omega_\mu}} (1 + 2n_\mu). \quad (\text{B5})$$

The gradient of the SCHA free-energy functional with respect to Φ is [16]

$$\nabla_\Phi \mathcal{F}_{\vec{\mathcal{R}}, \Phi} = -\sum_{ab\mu} \sqrt{\frac{M_a}{M_b}} (e_\mu^b \nabla_\Phi \ln a_\mu + \nabla_\Phi e_\mu^b) e_\mu^a \langle \delta f_b u_a \rangle. \quad (\text{B6})$$

Substituting the explicit expression of the forces we have

$$\begin{aligned} \nabla_\Phi \mathcal{F}(\vec{\mathcal{R}}, \Phi, \{\vec{a}_i\}) &= \sum_{abc\mu\nu} (K_{ac} - \Phi_{ac}) (e_\mu^a \nabla_\Phi \ln a_\mu + \nabla_\Phi e_\mu^a) \\ &\quad \times \frac{e_\mu^b e_\nu^c e_\nu^b a_\nu^2}{\sqrt{M_c M_a}}. \end{aligned} \quad (\text{B7})$$

It is clear from Eq. (B7) that in the minimum $\Phi = \mathbf{K}$. Therefore, it is convenient to compact all the other terms into a symbol:

$$\frac{\partial \mathcal{F}(\vec{\mathcal{R}}, \Phi, \{\vec{a}_i\})}{\partial \Phi_{cd}} = \sum_{ab} (K_{ab} - \Phi_{ab}) \mathcal{L}_{abcd}. \quad (\text{B8})$$

Here \mathcal{L} is a rank-4 tensor. Since we sum on all a and b indices and the \mathcal{L} rank-4 tensor multiplies a symmetrical matrix, it is convenient to recast it into a symmetrical form:

$$L_{abcd} = \sum_{k,\mu\nu} \left(e_\mu^a \frac{\partial \ln a_\mu}{\partial \Phi_{cd}} + \frac{\partial e_\mu^a}{\partial \Phi_{cd}} \right) e_\mu^k e_\nu^b e_\nu^k a_\nu^2, \quad (\text{B9})$$

$$\mathcal{L}_{abcd} = \frac{\mathcal{P}_{ab}}{\sqrt{M_a M_b}} \frac{L_{abcd} + L_{bacd}}{2}, \quad (\text{B10})$$

$$\mathcal{P}_{ab} = \sqrt{2}(1 - \delta_{ab}) + \delta_{ab}, \quad (\text{B11})$$

$$\mathcal{L}_{abcd} = \frac{\mathcal{P}_{ab}}{\sqrt{M_a M_b}} \sum_{\mu} \left[e_\mu^a e_\mu^b \frac{\partial \ln a_\mu}{\partial \Phi_{cd}} + \frac{1}{2} \frac{\partial (e_\mu^a e_\mu^b)}{\partial \Phi_{cd}} \right] a_\mu^2. \quad (\text{B12})$$

In the minimum the only nonzero term of the Hessian matrix is given by

$$\left. \frac{\partial^2 \mathcal{F}(\vec{\mathcal{R}}, \Phi, \{\vec{a}_i\})}{\partial \Phi_{ab} \partial \Phi_{cd}} \right|_{\Phi=\mathbf{K}} = -\mathcal{L}_{abcd}, \quad (\text{B13})$$

$$\begin{aligned} \frac{\partial^2 \mathcal{F}(\vec{\mathcal{R}}, \Phi, \{\vec{a}_i\})}{\partial \Phi_{ab} \partial \Phi_{cd}} &= -\frac{\mathcal{P}_{ab}}{\sqrt{M_a M_b}} \\ &\quad \times \sum_{\mu} \left[a_\mu e_\mu^a e_\mu^b \frac{\partial a_\mu}{\partial \Phi_{cd}} + \frac{1}{2} a_\mu^2 \frac{\partial (e_\mu^a e_\mu^b)}{\partial \Phi_{cd}} \right]. \end{aligned} \quad (\text{B14})$$

Let us start with the term inside the square brackets. The derivative of a_μ can be obtained with the chain rule:

$$\frac{\partial a_\mu}{\partial \Phi_{cd}} = \frac{\partial a_\mu}{\partial \omega_\mu} \frac{\partial \omega_\mu}{\partial \Phi_{cd}} = \frac{\mathcal{P}_{cd}}{2\omega_\mu} \frac{e_\mu^c e_\mu^d}{\sqrt{M_c M_d}} \frac{\partial a_\mu}{\partial \omega_\mu}. \quad (\text{B15})$$

The derivative of the polarization versors can be computed with first-order perturbation theory:

$$\begin{aligned} \frac{\partial (e_\mu^a e_\mu^b)}{\partial \Phi_{cd}} &= e_\mu^a \frac{\partial e_\mu^b}{\partial \Phi_{cd}} + e_\mu^b \frac{\partial e_\mu^a}{\partial \Phi_{cd}} \\ &= \frac{\mathcal{P}_{cd}}{\sqrt{M_c M_d}} \sum_{v \neq \mu} \frac{(e_\mu^a e_\nu^b + e_\mu^b e_\nu^a) (e_\nu^c e_\mu^d + e_\nu^d e_\mu^c)}{2(\omega_\mu^2 - \omega_\nu^2)}. \end{aligned} \quad (\text{B16})$$

We have a complete expression for the Hessian matrix:

$$\frac{\partial^2 \mathcal{F}(\vec{\mathcal{R}}, \Phi, \{\vec{a}_i\})}{\partial \Phi_{ab} \partial \Phi_{cd}} = -\frac{\mathcal{P}_{ab} \mathcal{P}_{cd}}{\sqrt{M_a M_b M_c M_d}} \left[\sum_{\mu} \frac{e_\mu^a e_\mu^b e_\mu^c e_\mu^d}{4\omega_\mu} \frac{\partial a_\mu^2}{\partial \omega_\mu} + \sum_{\mu\nu}^{\mu \neq \nu} \frac{e_\mu^a e_\nu^b (e_\mu^c e_\nu^d + e_\nu^c e_\mu^d)}{4} \left(\frac{a_\mu^2}{\omega_\mu^2 - \omega_\nu^2} + \frac{a_\nu^2}{\omega_\nu^2 - \omega_\mu^2} \right) \right]. \quad (\text{B17})$$

We can use the bosonic occupation number and write a_μ as a function of n_μ :

$$a_\mu = \sqrt{\frac{\hbar}{\omega_\mu} \left[n_\mu(\beta) + \frac{1}{2} \right]}, \quad (\text{B18a})$$

$$\frac{a_\mu}{2\omega_\mu} \frac{\partial a_\mu}{\partial \omega_\mu} = -\frac{\hbar}{8\omega_\mu^3} (2n_\mu + 1 + 2\beta\hbar\omega_\mu n_\mu^2 + 2\beta\hbar\omega_\mu n_\mu). \quad (\text{B18b})$$

Therefore we have

$$\begin{aligned} \frac{\partial^2 \mathcal{F}(\vec{\mathcal{R}}, \Phi, \{\vec{a}_i\})}{\partial \Phi_{ab} \partial \Phi_{cd}} &= \frac{\hbar \mathcal{P}_{ab} \mathcal{P}_{cd}}{\sqrt{M_a M_b M_c M_d}} \left[\sum_\mu e_\mu^a e_\mu^b e_\mu^c e_\mu^d \frac{2n_\mu + 1 + 2\beta\hbar\omega_\mu n_\mu^2 + 2\beta\hbar\omega_\mu n_\mu}{8\omega_\mu^3} \right. \\ &\quad \left. - \sum_{\mu \neq \nu} \frac{e_\mu^a e_\nu^b (e_\mu^c e_\nu^d + e_\nu^c e_\mu^d)}{8(\omega_\mu^2 - \omega_\nu^2)} \left(\frac{2n_\mu + 1}{\omega_\mu} - \frac{2\omega_\nu + 1}{\omega_\nu} \right) \right]. \end{aligned} \quad (\text{B19})$$

It is clear from Eq. (B19) that a Λ matrix can be introduced so that

$$\frac{\partial^2 \mathcal{F}(\vec{\mathcal{R}}, \Phi, \{\vec{a}_i\})}{\partial \Phi_{ab} \partial \Phi_{cd}} = \frac{1}{2} \mathcal{P}_{ab} \mathcal{P}_{cd} \sum_{\mu\nu} (\Lambda_{\mu\nu}^{abcd} + \Lambda_{\mu\nu}^{abdc}), \quad (\text{B20})$$

where

$$\Lambda_{\mu\mu}^{abcd} = \frac{\hbar e_\mu^a e_\mu^b e_\mu^c e_\mu^d}{\sqrt{M_a M_b M_c M_d}} \frac{2n_\mu + 1 + 2\beta\hbar\omega_\mu n_\mu^2 + 2\beta\hbar\omega_\mu n_\mu}{8\omega_\mu^3}, \quad (\text{B21a})$$

$$\Lambda_{\mu\nu}^{abcd} = -\frac{\hbar}{\sqrt{M_a M_b M_c M_d}} \frac{e_\mu^a e_\nu^b e_\mu^c e_\nu^d}{(\omega_\mu - \omega_\nu)(\omega_\mu + \omega_\nu)} \frac{2n_\mu \omega_\nu - 2\omega_\mu n_\nu + \omega_\nu - \omega_\mu}{4\omega_\mu \omega_\nu}. \quad (\text{B21b})$$

To conclude the proof it is sufficient to show that the Λ matrix of Eq. (32) is equal to

$$\Lambda^{abcd} = \sum_{\mu\nu} \Lambda_{\mu\nu}^{abcd}. \quad (\text{B22})$$

First, we introduce an auxiliary function $f(\omega_\mu, \omega_\nu)$ as

$$f(\omega_\mu, \omega_\nu) = \frac{2\omega_\nu n_\mu - 2\omega_\mu n_\nu + \omega_\nu - \omega_\mu}{4\omega_\mu \omega_\nu (\omega_\mu + \omega_\nu)(\omega_\mu - \omega_\nu)} = -\frac{1}{4\omega_\mu \omega_\nu} \left[\frac{n_\mu + n_\nu + 1}{\omega_\mu + \omega_\nu} - \frac{n_\mu - n_\nu}{\omega_\mu - \omega_\nu} \right]. \quad (\text{B23})$$

In the limit $\omega_\nu \rightarrow \omega_\mu$ we get

$$f(\omega_\mu) = \lim_{\omega_\nu \rightarrow \omega_\mu} f(\omega_\mu, \omega_\nu) = -\frac{2n_\mu + 1 + 2\hbar\beta\omega_\mu n_\mu^2 + 2\hbar\beta n_\mu \omega_\mu}{8\omega_\mu^3}, \quad (\text{B24})$$

$$f(\omega_\mu) = -\frac{1}{4\omega_\mu^2} \left[\frac{2n_\mu + 1}{2\omega_\mu} - \frac{\partial n}{\partial \omega} \right]. \quad (\text{B25})$$

So $\Lambda_{\mu\mu}^{abcd}$ is obtained as the continuous limit of $\Lambda_{\mu\nu}^{abcd}$ when $\mu \rightarrow \nu$:

$$\Lambda_{\mu\mu}^{abcd} = -\frac{\hbar e_\mu^a e_\mu^b e_\mu^c e_\mu^d}{\sqrt{M_a M_b M_c M_d}} f(\omega_\mu, \omega_\nu), \quad \Lambda_{\mu\mu}^{abcd} = -\frac{\hbar e_\mu^a e_\mu^b e_\mu^c e_\mu^d}{\sqrt{M_a M_b M_c M_d}} f(\omega_\mu). \quad (\text{B26})$$

Substituting Eqs. (B23) and (B25) we finally get

$$\Lambda_{\mu\nu}^{abcd} = \frac{\hbar}{4\omega_\mu \omega_\nu} \frac{e_\mu^a e_\nu^b e_\mu^c e_\nu^d}{\sqrt{M_a M_b M_c M_d}} \times \begin{cases} \frac{n_\mu + n_\nu + 1}{\omega_\mu + \omega_\nu} - \frac{n_\mu - n_\nu}{\omega_\mu - \omega_\nu} & \omega_\mu \neq \omega_\nu \\ \frac{2n_\mu + 1}{2\omega_\mu} - \frac{\partial n_\mu}{\partial \omega_\mu} & \omega_\mu = \omega_\nu \end{cases}. \quad (\text{B27})$$

APPENDIX C: HESSIAN IN THE CENTROIDS

The Hessian matrix approximation that we provide for the centroids is compared with the free-energy Hessian calculated

in Ref. [35]. In particular, the correct preconditioner should be chosen according to the minimization strategy. If the inner degrees of freedom are optimized simultaneously with the force constant matrix, then the correct preconditioner should be

the complete Hessian matrix between any couple of degrees of freedom while the others are fixed. Therefore we are neglecting the mixed derivatives between the force constant and the centroids. However, the centroids preconditioner we provided is the correct one, since the second derivative of the free energy is computed at a fixed force constant matrix.

The free-energy Hessian provided by Ref. [35] is, instead, the total derivative of the free energy:

$$\begin{aligned} \frac{d^2 \mathcal{F}(\vec{\mathcal{R}}, \Phi(\vec{\mathcal{R}}), \{\vec{a}_i\})}{d\mathcal{R}_a d\mathcal{R}_b} &= \frac{\partial^2 \mathcal{F}(\vec{\mathcal{R}}, \Phi, \{\vec{a}_i\})}{\partial \mathcal{R}_a \partial \mathcal{R}_b} \\ &+ \frac{\partial^2 \mathcal{F}(\vec{\mathcal{R}}, \Phi, \{\vec{a}_i\})}{\partial \mathcal{R}_a \partial \Phi} \frac{\partial \Phi}{\partial \mathcal{R}_b} \\ &+ \frac{\partial^2 \Phi}{\partial \mathcal{R}_a \partial \mathcal{R}_b} \frac{\partial \mathcal{F}(\vec{\mathcal{R}}, \Phi, \{\vec{a}_i\})}{\partial \Phi}. \quad (\text{C1}) \end{aligned}$$

The last term is zero in the minimum of the free energy, due to the Hellman-Feynman theorem. This is, indeed, the correct free-energy Hessian to study structural instabilities. It is also the correct preconditioner if the centroids are moved only after the full relaxation of the force constant is performed at each step. This is unpractical: the so-defined minimization algorithm converges slower, as it needs a full force constant minimization, the most expensive one, before starting to move the inner degrees of freedom. Moreover, the computation of the Hessian in Eq. (C1) is more expensive than the one provided in this paper, and it is not always positive defined.

APPENDIX D: QHA IN THE SCHA FRAMEWORK

The QHA can be reformulated in the SCHA framework in order to understand differences between the two approaches. The SCHA free energy is

$$\mathcal{F}(\vec{\mathcal{R}}, \Phi, \{\vec{a}_i\}) = F_\Phi + \langle V - \mathcal{V}_{\vec{\mathcal{R}}, \Phi} \rangle_{\rho_{\vec{\mathcal{R}}, \Phi}}. \quad (\text{D1})$$

If the system is perfectly harmonic, then the minimum of the free energy is found when $\mathcal{V}_{\vec{\mathcal{R}}_0, \Phi_0} = V$, and we get the QHA

free energy:

$$\begin{aligned} F_{\text{QHA}} &= \mathcal{F} = F_{\Phi_0} + V(\vec{\mathcal{R}}_0), \\ \Phi_{0\alpha\beta} &= \left. \frac{\partial^2 V}{\partial \mathcal{R}_\alpha \partial \mathcal{R}_\beta} \right|_{\vec{\mathcal{R}}=\vec{\mathcal{R}}_0}, \quad (\text{D2}) \end{aligned}$$

where $\vec{\mathcal{R}}_0$ is the minimum of the BO energy surface. So QHA is equivalent to SCHA for any harmonic potential. If the system is anharmonic, QHA approximates the potential as the second-order Taylor expansion around the equilibrium position. This makes the QHA theory not a self-consistent approach but a series expansion of the real potential.

If the atomic position coordinates relaxation is allowed, as introduced by Lazzeri and de Gironcoli [53,56], then the QHA free energy becomes

$$\begin{aligned} \mathcal{F}_{\text{QHA}}(\vec{\mathcal{R}}_c) &= F_{\Phi(\vec{\mathcal{R}}_c)} + V(\vec{\mathcal{R}}_c), \\ \tilde{\Phi}_{\alpha\beta}(\vec{\mathcal{R}}_c) &= \left. \frac{\partial^2 V}{\partial \mathcal{R}_\alpha \partial \mathcal{R}_\beta} \right|_{\vec{\mathcal{R}}=\vec{\mathcal{R}}_c}. \quad (\text{D3}) \end{aligned}$$

This is equivalent to SCHA [Eq. (D1)] keeping Φ fixed to the harmonic dynamical matrix and neglecting the contribution arising from $\langle V - \mathcal{V} \rangle_{\rho_{\vec{\mathcal{R}}_c, \Phi(\vec{\mathcal{R}}_c)}}$. The anharmonicity is taken into account by the fact that the harmonic dynamical matrix is a function of the atomic positions. This approximation is equivalent to neglecting all the even (from the fourth order) contribution in the BO surface Taylor expansion around the $\vec{\mathcal{R}}_c$ that minimizes \mathcal{F}_{QHA} . In this case, the average $\langle V - \mathcal{V} \rangle_{\rho_{\Phi(\vec{\mathcal{R}}_c)}}$ is equal to zero, and the harmonic dynamical matrix is the one that minimizes the SCHA free energy [Eq. (17) is exactly zero]. If only odd anharmonicities are present in the system (i.e., they dominate in the region of the quantum and thermal fluctuations), the QHA relaxed free energy coincides with the SCHA. The SCHA, therefore, is a natural extension to the relaxed QHA that assures the self-consistency of the theory for any kind of anharmonicity by explicitly including the average $\langle V - \mathcal{V} \rangle_{\rho_{\vec{\mathcal{R}}, \Phi}}$ in the free energy. Indeed, the SSCHA algorithm is much more efficient than the QHA relaxation, since it requires us only to compute energies and forces, while the QHA relaxation requires the third-order derivatives of the energy, and the application of the $2n + 1$ theorem for each minimization step [53].

[1] A. A. Volkov, Y. G. Goncharov, G. V. Kozlov, S. P. Lebedev, A. M. Prokhorov, R. A. Aliev, and K. R. Allakhverdiev, *Pis'ma Zh. Eksp. Teor. Fiz.* **37**, 517 (1983) [*JETP Lett.* **37**, 615 (1983)].
[2] J. Tsang, J. Smith, Jr., and M. Shafer, *Phys. Rev. Lett.* **37**, 1407 (1976).
[3] G. Pawley, W. Cochran, R. Cowley, and G. Dolling, *Phys. Rev. Lett.* **17**, 753 (1966).
[4] O. Delaire, J. Ma, K. Marty, A. F. May, M. A. McGuire, M.-H. Du, D. J. Singh, A. Podlesnyak, G. Ehlers, M. Lumsden *et al.*, *Nat. Mater.* **10**, 614 (2011).
[5] Y. Luspin, J. Servoin, and F. Gervais, *J. Phys. C* **13**, 3761 (1980).
[6] F. Weber, S. Rosenkranz, J.-P. Castellán, R. Osborn, G. Karapetrov, R. Hott, R. Heid, K.-P. Bohnen, and A. Alatas, *Phys. Rev. Lett.* **107**, 266401 (2011).

[7] M. Leroux, I. Errea, M. Le Tacon, S.-M. Souliou, G. Garbarino, L. Cario, A. Bosak, F. Mauri, M. Calandra, and P. Rodière, *Phys. Rev. B* **92**, 140303 (2015).
[8] M. Holt, P. Zschack, H. Hong, M. Y. Chou, and T.-C. Chiang, *Phys. Rev. Lett.* **86**, 3799 (2001).
[9] M. Leroux, M. Le Tacon, M. Calandra, L. Cario, M. A. Measson, P. Diener, E. Borrisenko, A. Bosak, and P. Rodiere, *Phys. Rev. B* **86**, 155125 (2012).
[10] P. S. H. Ghosez, X. Gonze, and J.-P. Michenaud, *Ferroelectrics* **206**, 205 (1998).
[11] M. Calandra, I. I. Mazin, and F. Mauri, *Phys. Rev. B* **80**, 241108 (2009).
[12] M. Calandra and F. Mauri, *Phys. Rev. Lett.* **106**, 196406 (2011).
[13] I. Errea, *Euro. Phys. J. B* **89**, 237 (2016).

- [14] B. Stritzker and W. Buckel, *Z. Phys. A* **257**, 1 (1972).
- [15] I. Errea, M. Calandra, and F. Mauri, *Phys. Rev. Lett.* **111**, 177002 (2013).
- [16] I. Errea, M. Calandra, and F. Mauri, *Phys. Rev. B* **89**, 064302 (2014).
- [17] I. Errea, M. Calandra, C. J. Pickard, J. Nelson, R. J. Needs, Y. Li, H. Liu, Y. Zhang, Y. Ma, and F. Mauri, *Phys. Rev. Lett.* **114**, 157004 (2015).
- [18] J. Schirber and C. Northrup, Jr., *Phys. Rev. B* **10**, 3818 (1974).
- [19] R. Car and M. Parrinello, *Phys. Rev. Lett.* **55**, 2471 (1985).
- [20] C. Z. Wang, C. T. Chan, and K. M. Ho, *Phys. Rev. B* **42**, 11276 (1990).
- [21] I. B. Magdău and G. J. Ackland, *Phys. Rev. B* **87**, 174110 (2013).
- [22] M. P. Ljungberg and J. Íñiguez, *Phys. Rev. Lett.* **110**, 105503 (2013).
- [23] A. M. Teweldeberhan, J. L. Dubois, and S. A. Bonev, *Phys. Rev. Lett.* **105**, 235503 (2010).
- [24] D.-B. Zhang, T. Sun, and R. M. Wentzcovitch, *Phys. Rev. Lett.* **112**, 058501 (2014).
- [25] O. Hellman, I. A. Abrikosov, and S. I. Simak, *Phys. Rev. B* **84**, 180301 (2011).
- [26] O. Hellman, P. Steneteg, I. A. Abrikosov, and S. I. Simak, *Phys. Rev. B* **87**, 104111 (2013).
- [27] O. Hellman and I. A. Abrikosov, *Phys. Rev. B* **88**, 144301 (2013).
- [28] D. Chandler and P. G. Wolynes, *J. Chem. Phys.* **74**, 4078 (1981).
- [29] J. A. Barker, *J. Chem. Phys.* **70**, 2914 (1979).
- [30] M. Ceriotti, M. Parrinello, T. E. Markland, and D. E. Manolopoulos, *J. Chem. Phys.* **133**, 124104 (2010).
- [31] T. Tadano and S. Tsuneyuki, *Phys. Rev. B* **92**, 054301 (2015).
- [32] B. Monserrat, N. D. Drummond, and R. J. Needs, *Phys. Rev. B* **87**, 144302 (2013).
- [33] M. Borinaga, P. Riego, A. Leonardo, M. Calandra, F. Mauri, A. Bergara, and I. Errea, *J. Phys.: Condens. Matter* **28**, 494001 (2016).
- [34] I. Errea, M. Calandra, C. J. Pickard, J. R. Nelson, R. J. Needs, Y. Li, H. Liu, Y. Zhang, Y. Ma, and F. Mauri, *Nature (London)* **532**, 81 (2016).
- [35] R. Bianco, I. Errea, L. Paulatto, M. Calandra, and F. Mauri, *Phys. Rev. B* **96**, 014111 (2017).
- [36] Y. Tajima, T. Matsuo, and H. Suga, *Nature (London)* **299**, 810 (1982).
- [37] B. Pamuk, J. M. Soler, R. Ramírez, C. P. Herrero, P. W. Stephens, P. B. Allen, and M.-V. Fernández-Serra, *Phys. Rev. Lett.* **108**, 193003 (2012).
- [38] K. Umamoto, E. Sugimura, S. de Gironcoli, Y. Nakajima, K. Hirose, Y. Ohishi, and R. M. Wentzcovitch, *Phys. Rev. Lett.* **115**, 173005 (2015).
- [39] K. Röttger, A. Endriss, J. Ihringer, S. Doyle, and W. F. Kuhs, *Acta Crystallogr. Sect. B* **50**, 644 (1994).
- [40] K. Röttger, A. Endriss, J. Ihringer, S. Doyle, and W. F. Kuhs, *Acta Crystallogr. Sect. B* **68**, 91 (2012).
- [41] W. H. Press, *Numerical Recipes: The Art of Scientific Computing*, 3rd ed. (Cambridge University Press, Cambridge, England, 2007).
- [42] P. López Ríos, B. Monserrat, and R. J. Needs, *Phys. Rev. B* **97**, 054104 (2018).
- [43] A. Kong, J. S. Liu, and W. H. Wong, *J. Am. Stat. Assoc.* **89**, 278 (1994).
- [44] F. Tassone, F. Mauri, and R. Car, *Phys. Rev. B* **50**, 10561 (1994).
- [45] C. Pinilla, M. Blanchard, E. Balan, G. Ferlat, R. Vuilleumier, and F. Mauri, *Geochim. Cosmochim. Acta* **135**, 203 (2014).
- [46] J. P. Perdew, K. Burke, and M. Ernzerhof, *Phys. Rev. Lett.* **77**, 3865 (1996).
- [47] D. Vanderbilt, *Phys. Rev. B* **41**, 7892 (1990).
- [48] A. D. Corso, *Comput. Mater. Sci.* **95**, 337 (2014).
- [49] P. Giannozzi, S. Baroni, N. Bonini, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, G. L. Chiarotti, M. Cococcioni, I. Dabo, A. D. Corso, S. de Gironcoli, S. Fabris, G. Fratesi, R. Gebauer, U. Gerstmann, C. Gougoussis, A. Kokalj, M. Lazzeri, L. Martin-Samos, N. Marzari, F. Mauri, R. Mazzarello, S. Paolini, A. Pasquarello, L. Paulatto, C. Sbraccia, S. Scandolo, G. Sclauzero, A. P. Seitsonen, A. Smogunov, P. Umari, and R. M. Wentzcovitch, *J. Phys.: Condens. Matter* **21**, 395502 (2009).
- [50] P. Giannozzi, O. Andreussi, T. Brumme, O. Bunau, M. B. Nardelli, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, M. Cococcioni, N. Colonna, I. Carnimeo, A. D. Corso, S. de Gironcoli, P. Delugas, R. A. DiStasio, A. Ferretti, A. Floris, G. Fratesi, G. Fugallo, R. Gebauer, U. Gerstmann, F. Giustino, T. Gorni, J. Jia, M. Kawamura, H.-Y. Ko, A. Kokalj, E. Küçükbenli, M. Lazzeri, M. Marsili, N. Marzari, F. Mauri, N. L. Nguyen, H.-V. Nguyen, A. O. de-la Roza, L. Paulatto, S. Poncé, D. Rocca, R. Sabatini, B. Santra, M. Schlipf, A. P. Seitsonen, A. Smogunov, I. Timrov, T. Thonhauser, P. Umari, N. Vast, X. Wu, and S. Baroni, *J. Phys.: Condens. Matter* **29**, 465901 (2017).
- [51] L. Paulatto, I. Errea, M. Calandra, and F. Mauri, *Phys. Rev. B* **91**, 054304 (2015).
- [52] M. J. Gillan, D. Alfè, and A. Michaelides, *J. Chem. Phys.* **144**, 130901 (2016).
- [53] M. Lazzeri and S. de Gironcoli, *Phys. Rev. B* **65**, 245402 (2002).
- [54] A. D. Fortes, *Acta Crystallogr. Sect. B* **74**, 196 (2018).
- [55] M. Gupta, R. Mittal, B. Singh, S. Mishra, D. Adroja, A. Fortes, and S. Chiplot, [arXiv:1803.08769](https://arxiv.org/abs/1803.08769).
- [56] M. Lazzeri and S. de Gironcoli, *Phys. Rev. Lett.* **81**, 2096 (1998).