SAPIENZA
UNIVERSITÀ DI ROMA

Phase diagram, structure and spectroscopy of ordinary and high pressure ice: impact of quantum anharmonic nuclear motion

Scuola di dottorato Vito Volterra

Dottorato di Ricerca in Fisica – XXXIV Ciclo

Candidate

Marco Cherubini

ID number 1609542

Thesis Advisor

Prof. Francesco Mauri

February 2022

Thesis defended on 27 May 2022
in front of a Board of Examiners composed by:

Prof. Riccardo Mazzarello (chairman)

Prof. Andrea Perali

Prof. Marco Liscidini

Phase diagram, structure and spectroscopy of ordinary and high pressure ice: impact of quantum anharmonic nuclear motion

Ph.D. thesis. Sapienza – University of Rome

This thesis has been typeset by LaTeX and the Sapthesis class.

Version: June 14, 2022

Author's email: marco.cherubini@uniroma1.com

Abstract

Water ice is a unique material, presenting the most complex phase diagram known in the literature, ranging from low to high temperature and from low to high pressures. The low-pressure phases, like ordinary ice I_h and its proton-ordered counterpart ice XI, show intriguing physical properties, such as negative thermal expansion and anomalous volume isotope effect (VIE). In the opposite regime, at high pressure, the features of the phase transition of dense ices VII/VIII to the symmetric ice X are still open questions. The signatures of the phase transition are indirect (hydrogen atoms are invisible to X-ray diffraction and the limited data quality and uncertainties in the procedure of data correction in neutron scattering hampers an unambiguous interpretation) and come from vibrational spectroscopies that give contrasting results because of the strong anharmonic regime close to the phase transition. Experimental data need the support of theoretical simulations to understand the high-pressure phase diagram.

In this thesis, I explore the paramount importance of nuclear quantum fluctuations in the thermodynamic and vibrational properties of low- and high-pressure ice by employing the stochastic self-consistent harmonic approximation. For what concerns the VIE in low-pressure ices, I prove that quantum effects on hydrogen are so strong to be in a nonlinear regime: when progressively increasing the mass of hydrogen from protium to infinity (classical limit), the volume first expands and then contracts, with a maximum slightly above the mass of tritium. I manage to accurately reproduce, for the first time, the low-energy phonon dispersion, possible thanks to the correct treatment of nuclear quantum fluctuations, paving to way for the study of thermal transport in ice from first-principles.

I establish the second-order character of the high-pressure phase transition combining the results from classical and quantum simulations, where a continuous transformation of one phase into the other, the presence of soft modes, and the absence of an hysteresis cycle is proven. I show the importance of including quantum fluctuations that reduce the critical pressure of about 55 GPa at T=0 K, solving the problem of the strong underestimation of the critical pressure by the classical approximations. I simulate the Infrared absorption spectra sampling a fine pressure grid close to the transition, revealing the sudden collapse of the stretching mode toward the low-energy regime in less than 10 GPa. Simultaneously, in the same range, I show that the low-energy translational mode (situated in a region where there are no published data to date) increases its intensity by an order of magnitude. These two features can be regarded as unique signatures of the phase transition.

Contents

Chapter 1

Introduction

1.1 Main topic and motivations

Water is essential for life. It is ubiquitous on Earth in all states of aggregation, the cryosphere covers part of the Earth's atmosphere and surface, both land and sea in the form of ice or snow. Ice influences the refrigeration, transportation and the climate [1] system: The albedo and radiative balance on Earth are determined by the extent of ice cover, the geometry of the upper atmosphere waves is affected by the height of the ice sheets, melt water influences the ocean circulation [2]. The presence of ice is not limited to our planet, it is found in many celestial bodies in the Universe and interstellar space.

Ice is a molecular crystal composed of individual water molecules held to one another by hydrogen bonds, whose importance ranges from biology to astrophysics. Ice exhibits the polymorphism typical of molecular crystals. It has been found so far in 17 different crystalline structure phases (Fig. 1.1) embodying the most complex phase diagram known in literature [3, 4, 5, 6, 7, 8, 9, 10, 11, 12]. In each phase, the oxygen atoms are long-range ordered in a specific symmetry with the hydrogen atoms arranged around the oxygen according to the Bernal-Fowler ice rules [13]. The rules state that each oxygen atom forms two covalent bonds with two hydrogen atoms and it is hydrogen-bonded to two other water molecules so that there is precisely one hydrogen between each pair of oxygen atoms [14] (Fig. 1.2). Ice structures can be depicted as fully connected tetrahedral networks of water molecules satisfying the ice rules.

In the low-pressure phases (as ordinary ice I_h) the angle between oxygen atoms is close to that of an ideal tetrahedron, upon compression, initially, the structure undergoes an angle distortion and with a further pressure increase, the reduced volume cannot be filled by increasing the distortion and in the denser phases of ice (like ice VII-VIII) the water molecules form interpenetrated (but not interconnected) four-coordinated networks. This picture holds for the oxygen atoms of the water molecules, while for the hydrogen atoms the situation is more cumbersome: the central water molecule of the tetrahedral network can be oriented in six different ways in order to satisfy the Bernal-Fowler rules. As a consequence, the hydrogen atoms are not topologically ordered and do not relate to an underlying lattice as the oxygen ones, but they decorate the oxygen framework with the only constraint to

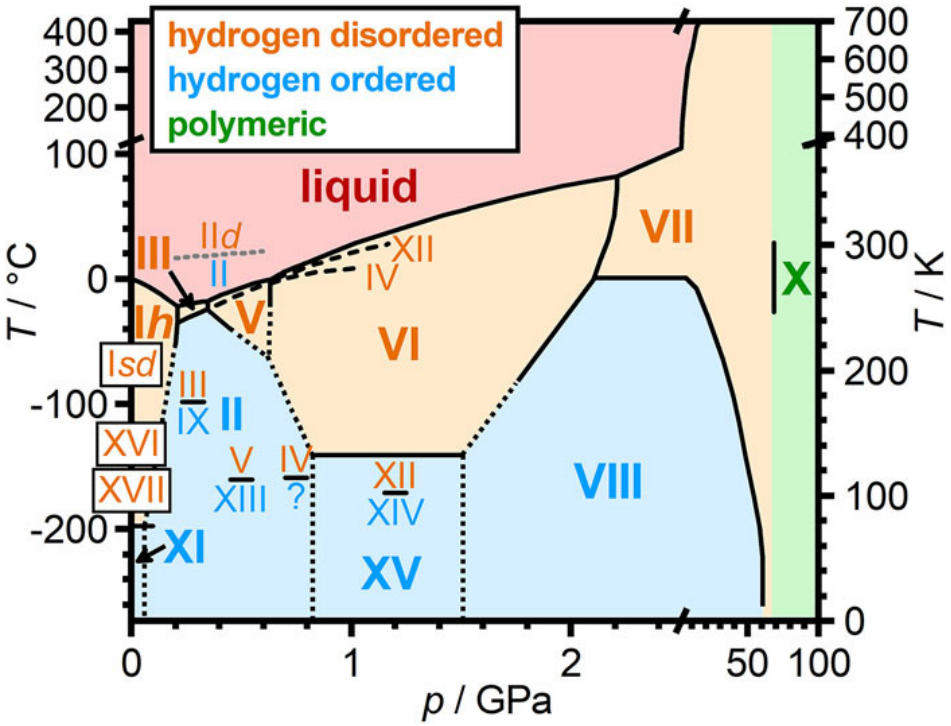

Figure 1.1. Phase diagram of ice. The phases shaded in orange, close to the liquid (red) are hydrogen disordered phases. Proton ordered structures are shaded in blue. The stable phases are indicated by bold Roman numbers, while metastable ones are highlighted by smaller font sizes. The green shaded area indicates the nonmolecular phase X, where the proton lies in the midpoint between two neighboring oxygen atoms. Figure taken from: *Advances in the experimental exploration of water's phase diagram, Salzmann, J. Chem. Phys. 150, 060901 (2019) [3]*

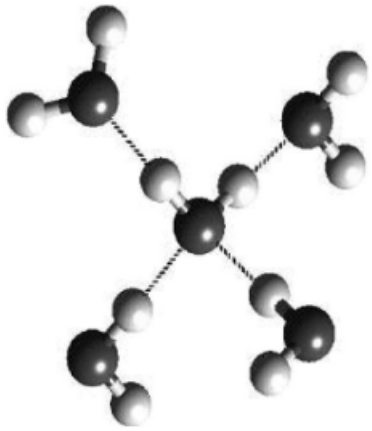

Figure 1.2. Local tetrahedron coordination of water molecules. Each oxygen atom forms two covalent bonds and two hydrogen bonds (dotted lines). Grey and white balls stand for oxygen and hydrogen atoms respectively. Figure taken from: *Ice structures, patterns, and processes: A view across the icefields, Bartels-Rausch et al. Rev. Mod. Phys. 84, 885* [2]

have one hydrogen atom along each hydrogen bond.

Several hydrogen arrangements are possible for proton disordered structures, depending on the number of molecules in the cell. In 1935, Pauling estimated the number of possible configurations as $M = W^N$. In the absence of the ice rules W = 4, since, for each N oxygen atoms in the piece of ice, the are 2N hydrogen atoms, each of them can sit in two different locations. Pauling recognizes that the Bernal-Fowler rules reduce the number of possible configurations so that $W \simeq 1.5$. The accuracy of his estimation was confirmed later in 1966 by Nagle's numerical calculations [15]. The ensemble of many possible hydrogen arrangements results in a residual entropy $S = k_b N \log W$, where k_b is the Boltzmann constant.

All the phases that share a phase boundary with the liquid (I_h,III,V,VI,VII) are orientationally disordered, while the others are proton-ordered.

The exploration of the full phase diagram of ice is a wonderful but challenging task and cannot be concluded in the relatively short time of a Ph.D. course. So, in the following, I will investigate the ordinary ice I_h and its proton-ordered counterpart ice XI [5] in the low-pressure region of the phase diagram, and I will explore in detail the high-pressure region, where the dense form of ice, phases VII and VIII transform into the symmetric ice X [16].

The proton ordered ice XI is thermodynamically stable below T = 72 K and it was discovered by calorimetric measurements on KOH doped ice I_h [17, 18, 19]. These phases show peculiar thermodynamic properties, like negative thermal expansion at low temperatures and anomalous volume isotope effect [20, 21] widely investigated which makes them a very interesting topic. Phases VII-VIII are another example of the ordered-disordered structures that fill the ice phase diagram, they occupy most of the high-pressure region up to $P \simeq 60GPa$, where the symmetrization of the hydrogen bond occurs [16, 22]: the proton lies midway the two oxygen atoms of neighboring water molecules. This new ice phase, named ice X, loses completely

the molecular character of ice. The phase transition toward this nonmolecular structure is very debated in literature [22, 23, 24, 25]; the critical pressure is hard to determine since the position of the hydrogen atoms cannot be identified by X-ray spectroscopy and in neutron scattering experiments, the limited quality of the data as well as the uncertainties in the procedure of data correction hamper an unambiguous interpretation of the results. All the evidence for the phase transition is indirect and comes mainly from spectroscopic experiments [26, 27] or from the study of the equation of state [16]. The full picture is still unclear, whether the transition is direct or if it exists an intermediate phase between ices VII/VIII and X, hypothesized because of the presence of broad spectra at the critical pressure that become narrower increasing the pressure [25, 16], is still under debate. Therefore, accurate numerical simulations can shed light on the problem.

More detailed descriptions of the ice phases investigated and their relative crystal structure are postponed to the dedicated chapters. In this work, I will perform numerical simulations on these ice phases. Through the years, the description of nuclear motion has undergone continuous improvement. The well-known failure of the harmonic approximation has been overcome by more sophisticated techniques. The quasi-harmonic approximation (QHA) is the commonly used technique to go beyond the harmonic approach. It has been shown to perform very well in the description of the low-pressure phases [28, 29, 30, 31], especially in the low-temperature regime, while at high temperatures, some discrepancies with respect to more accurate techniques arise. I will carry out the simulations of both low and high-pressure ice phases with the stochastic implementation of the Self Consistent Harmonic Approximation (SSCHA) [32] devised by the group and its time-dependent extension (TD-SSCHA) [33]. The SSCHA is based on a variable principle on the free energy and it allows me to include quantum anharmonicity in the nuclear motion at a non-perturbative level.

1.2 Structure of the thesis

This thesis is structured as follows, besides this brief introductory chapter in which I recall the main topic of the work and the motivations beyond it, I dedicated an entire chapter (Chap. 2) to the theoretical methodologies. There, I describe the SSCHA theory that we developed. Almost all the simulations done to achieve the results presented in this thesis are performed within the SSCHA framework and in the section dedicated to it, I describe also some improvements I applied to the code (Sec. 2.6) to have to possibility to simulate structures with many atoms (like the objects of this thesis) in a more efficient way. In Sec. 2.7, I report the time-dependent extension of the SSCHA, recently developed in Ref. [33] and I describe the force fields used for energies and forces calculations in Sec. 2.9.

The main results of the simulations on low- and high-pressure ices are in Chapter 3 and Chapter 4, respectively. The first one is completely dedicated to the investigation of ordinary ice I_h and its proton-ordered counterpart ice XI. First, I discuss their thermodynamic properties in Sec. 3.3.1, where, beyond the thermal expansion and the bulk modulus, I study in great detail the effect of quantum fluctuations on the anomalous isotopic effect. Vibrational spectroscopies, both Raman and IR, are

simulated in Sec. 3.3.5.

The high-pressure phase transition of ices VII/VIII to the symmetric phase X is reported in Chapter 4. There, I first investigate the order of the phase transition at a classical level at T=0 K for both the ordered and disordered phases in Sec. 4.4.1 and I confirm the result in the quantum regime in Sec. 4.4.2, where the effect of quantum fluctuations on the critical pressure and on the structural parameters of the ice phases investigated is underlined. The establishment of the phase diagram obtained through structure relaxation with and without quantum effects is then confirmed with the simulation of Infrared spectroscopy within the TD-SSCHA in Sec. 4.5. Since the experimental evidence of the phase transition are all indirect and come from the spectroscopy, I analyze in detail the comparison with the experiments (Sec. 4.5.2). In the end, Sec. 4.5.3 and 4.5.4 are dedicated to the analysis of quantum effects on the IR spectra and the comparison with MD simulations, respectively.

Chapter 2

Methods

2.1 Introduction

In this chapter, I report the theoretical methods and tools used to perform the numerical simulations necessary to reproduce all the results shown in this thesis. First, I describe the self-consistent harmonic approximation, a computational tool to study nuclear motion, including thermal and quantum anharmonic effects at a non-perturbative level (Sec. 2.2 - 2.6). We already published a full review of the SSCHA method [33] and released the code `https://github.com/SSCHAcode`.

I recall briefly the time-dependent extension of the SSCHA that consents the reproduction of vibrational spectra, infrared and Raman spectroscopy for instance, where atoms interact with a time-evolving external field. In the final part, I compare the SSCHA with other state-of-the-art approximations, mainly the QHA, highlighting the differences between the two approaches.

In the end, I describe the force fields used for the computation of energies and forces. The use of a force field based approach allows me to overcome the computational limitations coming from an *ab initio* method that would prevent the huge amount of calculations performed.

2.2 The Self-Consistent Harmonic Approximation

Ions fluctuate at any temperature in matter, also at zero kelvin due to the quantum zero-point motion. Even if the energy of ionic fluctuations is considerably smaller than the electronic one, many physical and chemical properties of materials and molecules cannot be understood without considering ionic vibrations. Since ionic vibrations are excited at much lower temperatures than electrons, ionic fluctuations are mainly responsible for the temperature dependence of thermodynamic properties of materials. They also determine heat and electrical transport through the electron-phonon and/or phonon-phonon interactions, as well as spectroscopic signatures detected in infrared, Raman, and inelastic x-ray or neutron scattering experiments. The large computational power available today has paved the way to material design and characterization, but advanced and reliable methods that accurately calculate vibrational properties of materials in the limit of strong quantum anharmonicity and that are easily interfaced with modern *ab initio* codes are required for accurately

describing materials' properties *in silico*.

Since electrons are faster than ions, the ionic motion is assumed to be described by the Born-Oppenheimer potential $V(\boldsymbol{R})$, which, at an ionic configuration $\boldsymbol{R}$, is given by the electronic ground state energy. In the standard harmonic approximation $V(\boldsymbol{R})$ is Taylor-expanded up to second-order around the $\boldsymbol{R}_0$ ionic positions that minimize $V(\boldsymbol{R})$. The resulting Hamiltonian is exactly diagonalizable in terms of phonons, the quanta of vibrations. Harmonic phonons are well-defined quasiparticles with an infinite lifetime, which energies do not depend on temperature. These two features are intrinsic failures of this approximation: phonons acquire a finite lifetime due to their anharmonic interaction with other phonons (also because of other types of interactions such as the electron-phonon coupling), and phonon energies do depend on temperature experimentally. When higher-order anharmonic terms are small compared to harmonic ones, anharmonicity can be treated within perturbation theory[34, 35, 36]. Even if within perturbative approaches phonons' temperature dependence and lifetimes can be understood, whenever anharmonic terms of the $V(\boldsymbol{R})$ potential are similar or larger than the harmonic terms in the range sampled by the ionic fluctuations, perturbative approaches collapse and are not valid[37]. This is often the case when light ions are present, as well as when the system is close to melting or a displacive phase transition, such as a ferroelectric or charge-density wave (CDW) instability.

In order to calculate from first principles vibrational properties of solids beyond perturbation theory and overcome these difficulties, several methods have been developed in the last years[38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 50, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63]. Many of them are based on extracting renormalized phonon frequencies from *ab initio* molecular dynamics (AIMD) through velocity autocorrelation functions[39, 40, 41, 42] or by extracting effective force constants from the AIMD trajectory[43, 44, 45]. In order to include quantum effects on the ionic motion, which are neglected on AIMD, the AIMD trajectory may be substituted by a path-integral molecular dynamics (PIMD) one[46]. Other methods are based on variational principles[49, 51, 55, 56, 57, 64], which are mainly inspired on the self-consistent harmonic approximation[65] or vibrational self-consistent field [66] theories, and yield free energies and/or phonon frequencies corrected by anharmonicity non-perturbatively.

Even if these methods have often successfully incorporated the effect of anharmonicity beyond perturbation theory in different materials, they usually lack a consistent procedure that prevents them from capturing properly both quantum effects and anharmonicity in the compound. For instance, many of them simply correct the free energy and/or the phonon frequencies assuming that the ions remain fixed at the $\boldsymbol{R}_0$ classical positions. However, as it has been shown recently in several compounds[48, 67, 68, 69], the ionic positions can be strongly altered by quantum effects and anharmonicity even at zero Kelvin. The structural changes are important for both internal degrees of freedom (the Wyckoff positions), and the lattice parameters themselves. Moreover, in many of the aforementioned methods, it is not clear what the meaning of the renormalized phonon frequencies is, i.e., whether they are auxiliary phonon frequencies intrinsic to the devised theoretical framework or if they really represent the physical vibrational excitations probed experimentally.

The stochastic self-consistent harmonic approximation (SSCHA)[55, 56, 57] is

a unique method that provides a full and complete way of incorporating ionic quantum and anharmonic effects on materials' properties without approximating the $V(\boldsymbol{R})$ potential. The SSCHA is defined from a rigorous variational method that directly yields the anharmonic free energy. It can optimize completely the crystal structure, including both internal and lattice degrees of freedom, accounting for the quantum nature of the ions at any target pressure or temperature. It computes thermal expansion even in highly anharmonic crystals. Furthermore, the SSCHA provides a well-defined approach to estimate at which thermodynamic conditions displacive second-order phase transitions occur. This is particularly challenging in *ab initio* molecular dynamics simulations, both for the dynamical slowing down that may hamper the thermalization close to the critical point, and for the difficulties in resolving the two distinct phases that continuously transform one into the other. Also, the rigorous theoretical approach of the SSCHA yields a clear distinction between auxiliary phonons of the theory and the phonon spectra probed experimentally, which can be accessed from a rigorous dynamical extension of the theory[33, 56, 70]. Lastly, the code provides non-perturbative third- and fourth-order phonon-phonon scattering matrices that can be fed in any external thermal transport code to compute thermal conductivity and lattice transport properties.

This section is organized to introduce the SSCHA algorithm and to review the recent developments in the SSCHA theory that lead to the SSCHA code. In Sec. 2.3, there is a simple overview of the method, presenting a simple picture of how it works with a model calculation on a highly anharmonic system with one particle in one dimension. Then, I review the full theory of the SSCHA in details, starting from the free energy calculation and structure optimization in Sec. 2.4. Then, I describe, in Sec. 2.5, the post-processing features of the code, which include calculations of the free energy Hessian for second-order phase transitions, as well as phonon spectral function and linewidth calculations. In Sec. 2.6, I provide some details of the Python code, including implementation of parallelized portion of the code that lead to an impressive speedup.

2.3 The variational free energy

The SSCHA is a theory that aims at describing the thermodynamics of a crystal, fully accounting for quantum, thermal, and anharmonic effects of nuclei within the Born-Oppenheimer approximation. The basis of all equilibrium thermodynamics is that a system in equilibrium at fixed volume, temperature, and number of particles is at the minimum of the free energy. The free energy is expressed by the sum of the internal energy E, which includes the energy of the interaction between the particles (kinetic and potential), and the product between the temperature T and the entropy S, which accounts for "disorder" and is related to the number of microstates corresponding to the same macrostate of the system:

$$F = E - TS. \tag{2.1}$$

In a classical picture, the free energy can be thus expressed in terms of the microscopic states of the system, which are determined by the classical probability distribution of atoms $\rho_{\text{cla}}(\boldsymbol{R})$. We remind that $\boldsymbol{R}$ is a vector of coordinates of all

atoms in the system (we will use bold symbols to denote vectors and tensors in component free notation). The same holds for a quantum system, but we need to account also for quantum interference. This is achieved by calculating the free energy with the many-body density matrix. As the system at equilibrium is at the minimum of the free energy, the Gibbs-Bogoliubov variational principle[71] states that between all possible trial density matrices $\tilde{\rho}$, the true free energy of the system is reached at the minimum of the functional $\mathcal{F}[\tilde{\rho}]$:

$$\mathcal{F}[\tilde{\rho}] = E[\tilde{\rho}] - TS[\tilde{\rho}] \geq F, \tag{2.2}$$

where

$$E[\tilde{\rho}] = \langle K + V(\boldsymbol{R}) \rangle_{\tilde{\rho}} \tag{2.3}$$

is the total energy (K is the kinetic energy operator and $V(\boldsymbol{R})$ the potential energy), and $S[\tilde{\rho}]$ the entropy calculated with the trial density matrix. $\langle \cdot | \cdot \rangle_{\tilde{\rho}} = Tr[\tilde{\rho}\ \cdot]$ indicates the quantum average of the the operator $\cdot$ taken with $\tilde{\rho}$.

If we pick any trial density matrix $\tilde{\rho}$, $\mathcal{F}[\tilde{\rho}]$ is an upper bound of the true free energy of the system. The SSCHA follows this principle: we optimize a trial density matrix $\tilde{\rho}$ to minimize the free energy functional $\mathcal{F}[\tilde{\rho}]$ of Eq. (2.2). Performing the optimization on any possible trial density matrix is, however, an unfeasible task due its many-body character that hinders an efficient parametrization. This is true also for a classical system: no exact parametrization of $\rho_{\text{cla}}(\boldsymbol{R})$ can be obtained in a computer with a finite memory.

The SSCHA solves the problem by imposing a constraint on the density matrix. In particular, the quantum probability distribution function that the SSCHA density matrix defines, $\tilde{\rho}_{\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}}(\boldsymbol{R}) = \langle \boldsymbol{R} | \, \tilde{\rho}_{\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}} \, | \boldsymbol{R} \rangle$, is a Gaussian. $\tilde{\rho}_{\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}}(\boldsymbol{R})$ is the quantum analogue of $\rho_{\text{cla}}(\boldsymbol{R})$, and determines the probability to find the atoms in the configuration $\boldsymbol{R}$. The trial SSCHA density matrix $\tilde{\rho}_{\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}}$ is uniquely identified by the average atomic positions (centroids) $\boldsymbol{\mathcal{R}}$ and the quantum-thermal fluctuations around them $\boldsymbol{\Phi}$ (we have explicitly expressed the dependence of $\tilde{\rho}$ on $\boldsymbol{\mathcal{R}}$ and $\boldsymbol{\Phi}$ by adding them as subindexes), just like any Gaussian is defined by the average and mean square displacements. Within the SSCHA, we optimize $\boldsymbol{\mathcal{R}}$ and $\boldsymbol{\Phi}$ to minimize the free energy of the system. In this way, we compress the memory requested to store $\tilde{\rho}_{\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}}$, as $\boldsymbol{\mathcal{R}}$ depends only on $3N_{\text{a}}$ numbers (the coordinates of the atoms), while the fluctuations $\boldsymbol{\Phi}$ are encoded in a symmetric square real matrix of $3N_{\text{a}} \times 3N_{\text{a}}$. N_{a} is the total number of atoms in the system. The free parameters in $\boldsymbol{\mathcal{R}}$ and $\boldsymbol{\Phi}$ can be further reduced by exploiting translation and point group symmetries of the crystal, resulting in an efficient and compact representation of the density matrix $\tilde{\rho}_{\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}}$.

The "harmonic" in the SSCHA name comes from the fact that any Gaussian density matrix that describes a physical system is the equilibrium solution of a particular harmonic Hamiltonian. Therefore, there is a one-to-one mapping between the trial density matrix $\tilde{\rho}_{\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}}$ and an auxiliary trial harmonic Hamiltonian $\mathcal{H}_{\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}}$:

$$\mathcal{H}_{\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}} = K + \frac{1}{2} \sum_{ab} (R_a - \mathcal{R}_a) \Phi_{ab} (R_b - \mathcal{R}_b). \tag{2.4}$$

Here, $\boldsymbol{\mathcal{R}}$ is a real vector and $\boldsymbol{\Phi}$ a real matrix that parametrize the trial Hamiltonian, while K and $\boldsymbol{R}$ are quantum operators that measure the kinetic energy and the

position of the state. For simplicity, unless otherwise specified, all indices a, b, etc. run over both atomic and Cartesian coordinates from 1 to $3N_\mathrm{a}$. Let us note here, that, inspired by the harmonic shape of $\mathcal{H}_{\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}}$, we will also refer to $\boldsymbol{\Phi}$ as the auxiliary force constants.

This mapping with a harmonic Hamiltonian is very useful, as both $\langle K|K\rangle_{\tilde{\rho}_{\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}}}$ and $S[\tilde{\rho}_{\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}}]$ become simply the kinetic energy and entropy of the auxiliary harmonic system $\mathcal{H}_{\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}}$, which are analytic functions of $\boldsymbol{\Phi}$. Hence, the only quantity that we really need to compute is the average over the interacting Born-Oppenheimer potential

$$\langle V(\boldsymbol{R})\rangle_{\tilde{\rho}} = \int d\boldsymbol{R}\, V(\boldsymbol{R})\tilde{\rho}(\boldsymbol{R}). \tag{2.5}$$

The potential $V(\boldsymbol{R})$ is the Born-Oppenheimer energy landscape, and can be easily computed *ab initio* by any DFT code (or by any energy and force engine).

The SSCHA algorithm starts with an initial guess on $\boldsymbol{\mathcal{R}}$ and $\boldsymbol{\Phi}$, and proceeds as follows:

- Use the trial Gaussian probability distribution function $\tilde{\rho}_{\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}}(\boldsymbol{R})$ to extract an ensemble of random nuclear configurations in a supercell.
- For each nuclear configuration in the ensemble, compute total energies and forces with an external code, either *ab initio* or via a force field.
- Use total energy and forces on the ensemble to compute the free energy functional and its derivatives with respect to the free parameters of our distribution $\boldsymbol{\mathcal{R}}$ and $\boldsymbol{\Phi}$.
- Update $\boldsymbol{\mathcal{R}}$ and $\boldsymbol{\Phi}$ to minimize the free energy.

These steps are repeated until the minimum of the free energy is found.

To illustrate better the philosophy of the method, we report in Figure 2.1 a simple application of the SSCHA to a one particle in one dimension at $T = 0\,\mathrm{K}$. In panel **a**, we plot the very anharmonic "Born-Oppenheimer" (BO) energy landscape $V(\boldsymbol{R})$ of our one-dimensional particle (of mass of an electron). In Hartree atomic units it is given by

$$V(R) = 3R^4 + \frac{1}{2}R^3 - 3R^2. \tag{2.6}$$

We first study the classical harmonic result, obtained by Taylor-expanding the potential in Eq. (2.6) to second order around the minimum R_0. Then, we use the harmonic solution to build our initial guess for the SSCHA density matrix $\tilde{\rho}_{\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}}$ and update the parameters ($\boldsymbol{\Phi}$ and $\boldsymbol{\mathcal{R}}$) until we reach the minimum of the free energy. In Figure 2.1(a), we compare the average atomic position and equilibrium free energy obtained with the harmonic approximation, with the SSCHA, and the result we obtained with the exact diagonalization of the potential. While the harmonic result clearly overestimates the energy and yields an average atomic position far from the exact result, the SSCHA energy and average position are very close to the exact solution. In Figure 2.1(b), we report the probability distribution functions of the particle for the different approximations compared with the exact result. By definition, both the harmonic and SSCHA results have Gaussian probability distributions. However, while the harmonic solution is centered in the minimum of

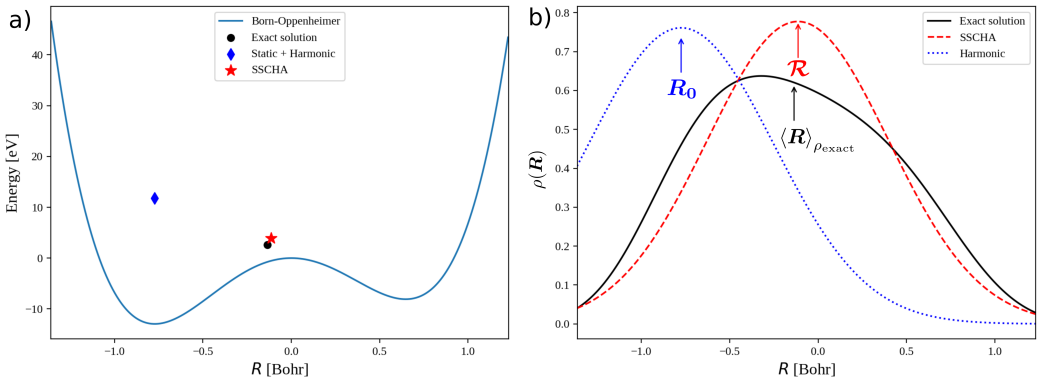

Figure 2.1. Illustration of the SSCHA method to a one dimensional particle problem at $T = 0\,\text{K}$. Panel **a**: The one dimensional Born-Oppenheimer energy landscape $V(R)$ as a function of the particle position R. The points represent the solution of the Harmonic approximation, the SSCHA, and the exact solution. The y coordinate of the points are the quantum total energy (including the zero-point motion), while the x axis coordinate is the average position of the particle. The SSCHA outperforms the harmonic approximation and it is very close to the exact solution. Panel **b**: Representation of the nuclear quantum distribution functions in the different approaches. The arrows point the average position of the particle in each distribution. Both harmonic and the SSCHA are Gaussians, while the exact solution is more complex. The harmonic solution is centered around the minimum of the energy landscape R_0, while the SSCHA centroid position $\mathcal{R}$ and width are optimized to satisfy the least energy principle. The average position in the exact case is however obtained as $\langle R \rangle_{\rho_{\text{exact}}}$.

the BO energy landscape (and the width is fixed by the harmonic frequencies), the SSCHA distribution is optimized to minimize the free energy. Notice how, even if the exact equilibrium distribution deviates from the Gaussian line-shape, the SSCHA energy and average nuclear position match almost perfectly the exact solution as stated above. The very good result on the free energy reflects that the SSCHA error is variational: the free energy of the exact density matrix is the minimum. This means that the free energy is stationary around the exact solution, assuring that even an approximate density matrix (like the SSCHA solution) describes very well the exact free energy. This is an excellent feature of the SSCHA, as the free energy and its derivatives fully characterize thermodynamic properties. Even if this simple calculation is performed at $T = 0\,\text{K}$, the SSCHA can simulate any finite temperature by mixing quantum and thermal fluctuations on the nuclear distribution.

The previously outlined straightforward implementation of the SSCHA becomes too cumbersome on a real system composed of many particles, especially if *ab initio* methods are used to extract $V(\boldsymbol{R})$. The reason is that at any minimization step we need to calculate total energies and forces for many ionic configurations with displaced atoms in a supercell. The bottleneck is the computational cost of the force engine adopted. In the next sections we will show how the number of force calculations can be minimized and how these issues can be overcome by the code implementation proposed here. The resulting SSCHA code is very efficient, and, in most of the core cases, much faster than standard AIMD, with the advantage of fully accounting for the quantum nature of nuclei.

Symbol	Meaning	First use
$\boldsymbol{R}$	Atomic position (canonical variable)	Eq. (2.3)
$V(\boldsymbol{R})$	Potential energy	Eq. (2.3)
$\boldsymbol{\mathcal{R}}$	Trial centroid positions (parameter)	Eq. (2.4)
$\boldsymbol{u}$	Displacement from the average atomic position $\boldsymbol{\mathcal{R}}$	Eq. (2.15)
$\boldsymbol{\Phi}$	Trial harmonic matrix (parameter)	Eq. (2.4)
$\boldsymbol{\Psi}$	Static displacement-displacement correlation matrix relative to $\boldsymbol{\Phi}$	Eq. (2.15)
$\mathcal{H}_{\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}}$	Trial harmonic Hamiltonian for given $\boldsymbol{\mathcal{R}}$, $\boldsymbol{\Phi}$	Eq. (2.4)
$\tilde{\rho}_{\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}}$	Density matrix of $\mathcal{H}_{\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}}$ (trial density matrix)	Pag. 10-1
$\tilde{\rho}_{\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}}(\boldsymbol{R})$	Gaussian positional probability density for $\tilde{\rho}_{\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}}$	Pag. 10-1
$\mathcal{F}[\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}]$	SSCHA Helmholtz free energy functional	Eq. (2.7)
$\mathcal{G}[\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}]$	SSCHA Gibbs free energy functional	Eq. (2.10)
$\mathbf{f}^{\mathcal{H}_{\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}}}(\boldsymbol{R})$	Forces for the Hamiltonian $\mathcal{H}_{\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}}$ acting on the ions when they are in the $\boldsymbol{R}$ positions	Eq. (2.12)
$\mathbf{f}^{(\mathrm{BO})}(\boldsymbol{R})$	Born-Oppenheimer forces acting on the ions when they are in the $\boldsymbol{R}$ positions	Eq. (2.12)
$\boldsymbol{P}^{(\mathrm{BO})}(\boldsymbol{R})$	Born-Oppenheimer stress tensor when the ions are in the $\boldsymbol{R}$ positions	Eq. (2.19)
$\boldsymbol{\Phi}_{\boldsymbol{\mathcal{R}}}$	2nd SSCHA force constants for a given $\boldsymbol{\mathcal{R}}$, it is the trial $\boldsymbol{\Phi}$ that minimizes $\mathcal{F}[\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}]$	Pag. 25-2
$\boldsymbol{D}_{\boldsymbol{\mathcal{R}}}$	$\boldsymbol{\Phi}_{\boldsymbol{\mathcal{R}}}$ divided by the square root of the masses	Eq. (2.52)
$\overset{(3)}{\boldsymbol{\Phi}}_{\boldsymbol{\mathcal{R}}}$, $\overset{(4)}{\boldsymbol{\Phi}}_{\boldsymbol{\mathcal{R}}}$	3rd and 4th order SSCHA force constants for a given $\boldsymbol{\mathcal{R}}$	Eqs. (2.53), (2.54)
$\overset{(3)}{\boldsymbol{D}}_{\boldsymbol{\mathcal{R}}}$, $\overset{(4)}{\boldsymbol{D}}_{\boldsymbol{\mathcal{R}}}$	$\overset{(3)}{\boldsymbol{\Phi}}_{\boldsymbol{\mathcal{R}}}$, $\overset{(4)}{\boldsymbol{\Phi}}_{\boldsymbol{\mathcal{R}}}$ divided by the square root of the masses	Eqs. (2.53), (2.54)
$F(\boldsymbol{\mathcal{R}})$	SSCHA positional Helmholtz free energy, given by $\mathcal{F}[\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}_{\boldsymbol{\mathcal{R}}}]$	Eq. (2.50)
$\boldsymbol{\mathcal{R}}_{\mathrm{eq}}$	SSCHA equilibrium centroids, trial centroids that minimizes $F(\boldsymbol{\mathcal{R}})$	Eq. (2.8)
$\boldsymbol{\Phi}_{\mathrm{eq}}$	SSCHA harmonic matrix $\boldsymbol{\Phi}_{\boldsymbol{\mathcal{R}}_{\mathrm{eq}}}$, the trial $\boldsymbol{\Phi}$ at the minimum of the free energy functional	Eq. (2.8)
$\mathcal{H}^{(\mathrm{S})}$	SSCHA effective harmonic Hamiltonian, given by $\mathcal{H}_{\boldsymbol{\mathcal{R}}_{\mathrm{eq}},\boldsymbol{\Phi}_{\mathrm{eq}}}$	Eq. (2.60)
$\boldsymbol{D}^{(\mathrm{S})}$	Dynamical matrix of $\mathcal{H}^{(\mathrm{S})}$, given by $\boldsymbol{D}_{\boldsymbol{\mathcal{R}}_{\mathrm{eq}}}$	Pag. 28-2
$\overset{(3)}{\boldsymbol{D}}_{\mathrm{eq}}$, $\overset{(4)}{\boldsymbol{D}}_{\mathrm{eq}}$	Symbols indicating $\overset{(3)}{\boldsymbol{D}}_{\boldsymbol{\mathcal{R}}_{\mathrm{eq}}}$ and $\overset{(4)}{\boldsymbol{D}}_{\boldsymbol{\mathcal{R}}_{\mathrm{eq}}}$, respectively	Eq. (2.62)
$\boldsymbol{D}^{(\mathrm{F})}$	Positional Helmholtz free energy Hessian divided by the square root of the masses	Pag. 25-1
$\boldsymbol{G}(z)$	One-phonon Green function	Eq. (2.67)
$\boldsymbol{\Pi}(0)$, $\boldsymbol{\Pi}(z)$	Static and dynamic SSCHA self-energy	Eqs. (2.62), (2.68)
$\overset{(\mathrm{B})}{\boldsymbol{\Pi}}(0)$, $\overset{(\mathrm{B})}{\boldsymbol{\Pi}}(z)$	Static and dynamic SSCHA bubble self-energy	Eqs. (2.64), (2.69)
$\sigma(\boldsymbol{q},\Omega)$	Phonon spectral function (with the reciprocal lattice vector made explicit)	eq.(2.70)
$\omega_\mu(\boldsymbol{q})$	Frequency of the $(\mu,\boldsymbol{q})$ SSCHA auxiliary phonon	Eq. (2.16)
$\Omega_\mu(\boldsymbol{q})$	Frequency of the $(\mu,\boldsymbol{q})$ static approximation phonon from $\boldsymbol{D}^{(\mathrm{F})}$	Pag. 29-1
$\Omega_\mu(\boldsymbol{q})$,$\Gamma_\mu(\boldsymbol{q})$	Frequency and linewidth of the $(\mu,\boldsymbol{q})$ anharmonic phonon in the Lorentzian approximation	Eq. (2.81)

Table 2.1. Collection of some symbols frequently used in the method section. First column, the symbol used. Second column, a short description. Third column, equation or page-column of the first occurrence.

2.4 Structure relaxation and free energy minimization

In this section, I explain the simplest and most common use of the code: the calculation of the free energy and the optimization of a structure by fully accounting for temperature and quantum effects. This enables the simulation of finite temperature and pressure phase-diagrams (with first order boundaries), as well as the calculation of the lattice thermal expansion. I start by briefly reviewing the theory of the SSCHA method and then the details of the implementation are explained.

2.4.1 The SSCHA free energy minimization

In the simplest and most standard usage, the SSCHA free energy functional that is minimized depends on the centroid positions $\boldsymbol{\mathcal{R}}$ and the auxiliary force constants $\boldsymbol{\Phi}$ as

$$\mathcal{F}[\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}] = \langle K + V(\boldsymbol{R})\rangle_{\tilde{\rho}_{\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}}} - T\, S_{\mathrm{ion}}\left[\tilde{\rho}_{\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}}\right]. \tag{2.7}$$

Here, we explicit that the entropy S_{ion} only accounts for ionic degrees of freedom (not electronic). After the SSCHA minimization, the final estimate of the equilibrium

free energy is given by

$$F = \min_{\boldsymbol{\mathcal{R}}, \boldsymbol{\Phi}} \mathcal{F}[\boldsymbol{\mathcal{R}}, \boldsymbol{\Phi}] = \mathcal{F}[\boldsymbol{\mathcal{R}}_{\text{eq}}, \boldsymbol{\Phi}_{\text{eq}}] \,. \tag{2.8}$$

Therefore, the final result of a SSCHA free energy calculation is in general given in terms of the equilibrium configuration $\boldsymbol{\mathcal{R}}_{\text{eq}}$, the free energy F, and the SSCHA auxiliary force constants $\boldsymbol{\Phi}_{\text{eq}}$. The final free energy accounts for quantum and thermal ionic fluctuations without approximating the BO energy surface, valid to study thermodynamic properties, and the $\boldsymbol{\mathcal{R}}_{\text{eq}}$ positions determine the most probable atomic positions also taken into account quantum/thermal flluctuations and anharmonicity. It is important to remark, however, that the Gaussian variance $\boldsymbol{\Phi}$ has, in principle, no relation with the experimentally observed phonon frequencies, as it is just a variable parametrizing the density matrix. The relation of it with the physical phonon frequencies is discussed in Sec. 2.5.3.

The SSCHA can also perform the free energy minimization at fixed pressure instead. In this case, the Gibbs-Bogoliubov inequality is satisfied by the Gibbs free energy G, defined as

$$G = F + P^* \Omega_{\text{Vol}}, \tag{2.9}$$

where P^* is the target pressure, Ω_{Vol} is the simulation box volume, and F is the Helmholtz free energy. In this case, the code optimizes

$$G \leq \mathcal{G}[\boldsymbol{\mathcal{R}}, \boldsymbol{\Phi}] = \mathcal{F}[\boldsymbol{\mathcal{R}}, \boldsymbol{\Phi}] + P^* \Omega_{\text{Vol}}, \tag{2.10}$$

which can be used, for instance, to estimate the structural changes imposed by pressure by fully accounting for fluctuations.

As made explicit in Eq. (2.7), only thermal effects on the ions are taken into account so far, whereas the electrons are considered at zero temperature. However, at very high temperatures the entropy associated to electrons may be important. Within the SSCHA, it is possible to explicitly include finite-temperature effects on the electrons too. The key is to replace in Eq. (2.7) the electronic ground state energy $V(\boldsymbol{R})$ with the finite-temperature electronic free energy $F_{\text{el}}(\boldsymbol{R}) = E_{\text{el}}(\boldsymbol{R}) - TS_{\text{el}}(\boldsymbol{R})$ (if electrons have finite temperature, in the adiabatic approximation forces and equilibrium position of the ions are ruled by the electronic free energy). In this case the SSCHA method minimizes the functional

$$\begin{aligned} \mathcal{F}[\boldsymbol{\mathcal{R}}, \boldsymbol{\Phi}] &= \langle K + F_{\text{el}}(\boldsymbol{R}) \rangle_{\tilde{\rho}_{\boldsymbol{\mathcal{R}}, \boldsymbol{\Phi}}} - T\, S_{\text{ion}}\left[\tilde{\rho}_{\boldsymbol{\mathcal{R}}, \boldsymbol{\Phi}}\right] = \\ &= \langle K + E_{\text{el}}(\boldsymbol{R}) \rangle_{\tilde{\rho}_{\boldsymbol{\mathcal{R}}, \boldsymbol{\Phi}}} - T\, S\left[\tilde{\rho}_{\boldsymbol{\mathcal{R}}, \boldsymbol{\Phi}}\right] \,, \end{aligned} \tag{2.11}$$

where $S\left[\tilde{\rho}_{\boldsymbol{\mathcal{R}}, \boldsymbol{\Phi}}\right] = \langle S_{\text{el}}(\boldsymbol{R}) \rangle_{\tilde{\rho}_{\boldsymbol{\mathcal{R}}, \boldsymbol{\Phi}}} + S_{\text{ion}}\left[\tilde{\rho}_{\boldsymbol{\mathcal{R}}, \boldsymbol{\Phi}}\right]$. The same trick can be applied to the Gibbs free energy minimization as well. Therefore, the SSCHA estimation of the system's entropy can also incorporate contributions from both electrons (averaged through the ionic distribution $\tilde{\rho}_{\boldsymbol{\mathcal{R}}, \boldsymbol{\Phi}}$) and ions. In a DFT framework, for example, this simply comes down to including the electronic temperature in the energy/forces/stress calculations for the ensemble elements through the Fermi-Dirac occupation of the Kohn-Sham states [72].

2.4.2 The implementation of the free energy minimization

In the SSCHA code, the minimization of $\mathcal{F}[\boldsymbol{\mathcal{R}}, \boldsymbol{\Phi}]$ is performed through a preconditioned gradient descent approach, which requires the calculation of the gradient of the free energy with respect to the centroid positions $\boldsymbol{\mathcal{R}}$ and the auxiliary force constants $\boldsymbol{\Phi}$. The partial derivatives are evaluated through the exact analytic formulas

$$\frac{\partial \mathcal{F}}{\partial \mathcal{R}_a} = -\left\langle \mathrm{f}_a^{(\mathrm{BO})}(\boldsymbol{R}) - \mathrm{f}_a^{\mathcal{H}_{\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}}}(\boldsymbol{R}) \right\rangle_{\tilde{\rho}_{\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}}} \tag{2.12}$$

and

$$\frac{\partial \mathcal{F}}{\partial \Phi_{cd}} = \frac{1}{2} \sum_{ab} \frac{\partial \Psi_{ab}}{\partial \Phi_{cd}} \times \left\langle \left(\mathrm{f}_b^{(BO)}(\boldsymbol{R}) - \mathrm{f}_b^{\mathcal{H}_{\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}}}(\boldsymbol{R}) \right) \sum_e \Psi_{ae}^{-1} \left(R_e - \mathcal{R}_e \right) \right\rangle_{\tilde{\rho}_{\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}}} . \tag{2.13}$$

Here $\mathbf{f}^{(\mathrm{BO})}(\boldsymbol{R})$ are the Born-Oppenheimer forces that act on the ions when they are in the $\boldsymbol{R}$ positions; $\mathbf{f}^{\mathcal{H}_{\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}}}(\boldsymbol{R})$ is the force given by the auxiliary harmonic Hamiltonian $\mathcal{H}_{\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}}$,

$$\mathrm{f}_a^{\mathcal{H}_{\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}}}(\boldsymbol{R}) = -\sum_b \Phi_{ab}(R_b - \mathcal{R}_b); \tag{2.14}$$

and $\boldsymbol{\Psi}$ is the displacement-displacement correlation matrix

$$\Psi_{ab} = \langle u_a u_b \rangle_{\tilde{\rho}_{\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}}} , \tag{2.15}$$

where with $\boldsymbol{u} = \boldsymbol{R} - \boldsymbol{\mathcal{R}}$ we indicate the displacement from the average atomic position. Explicitly,

$$\Psi_{ab} = \frac{1}{\sqrt{M_a M_b}} \sum_\mu \frac{\hbar(2n_\mu + 1)}{2\omega_\mu} e_\mu^a e_\mu^b. \tag{2.16}$$

In Eq. (2.16), ω_μ and $\boldsymbol{e}_\mu$ are the eigenvalues and eigenvectors of the mass rescaled auxiliary force constants $\Phi_{ab}/\sqrt{M_a M_b}$, and n_μ is the Bose-Einstein occupation number for the ω_μ frequency. We underline again here that ω_μ are not the phonon frequencies of the system, but just the frequencies of the auxiliary harmonic Hamiltonian $\mathcal{H}_{\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}}$. In other words, they are only used to define the trial density matrix $\tilde{\rho}_{\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}}$. We show how to compute the physical anharmonic phonon frequencies of the system in Sec. 2.5.

It is convenient to give an explicit expression for the gradient of the free energy with respect to the auxiliary force constants in terms of the ω_μ eigenvalues and $\boldsymbol{e}_\mu$ eigenvectors. As shown in Ref. [56] (see Appendix B), the gradient can be rewritten as

$$\frac{\partial \mathcal{F}}{\partial \Phi_{cd}} = \sum_{ab} \frac{\Lambda[0]^{abcd}}{\sqrt{M_a M_b M_c M_d}} \times \left\langle \left(\mathrm{f}_b^{(BO)}(\boldsymbol{R}) - \mathrm{f}_b^{\mathcal{H}_{\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}}}(\boldsymbol{R}) \right) \sum_e \Psi_{ae}^{-1} \left(R_e - \mathcal{R}_e \right) \right\rangle_{\tilde{\rho}_{\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}}} , \tag{2.17}$$

where

$$\begin{aligned}\Lambda[0]^{abcd} &= \sum_{\mu\nu} \frac{\hbar}{4\omega_\nu\omega_\mu} e_\nu^a e_\mu^b e_\nu^c e_\mu^d \\ &\times \begin{cases} \frac{dn_\mu}{d\omega_\mu} - \frac{2n_\mu+1}{2\omega_\mu} & ,\omega_\nu = \omega_\mu \\ \frac{n_\mu - n_\nu}{\omega_\mu - \omega_\nu} - \frac{1+n_\mu+n_\nu}{\omega_\mu+\omega_\nu} & ,\omega_\nu \neq \omega_\mu \end{cases} . \end{aligned} \tag{2.18}$$

Here, $n_\mu = 1/(e^{\beta\hbar\omega_\mu} - 1)$. The reason for the introduction of the $\mathbf{\Lambda}[0]$ tensor will be evident in Sec. 2.5. Even if Eq. (2.17) looks different to the gradient introduced in the original SSCHA work in Ref. [55], it can be demonstrated that both expressions are equivalent by simply playing with the permutation symmetry of $\left\langle \frac{\partial^2 V}{\partial R_a \partial R_b} \right\rangle_{\tilde{\rho}_{\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}}}$. However, Eq. (2.18) unambiguously determines the value taken by the $\mathbf{\Lambda}[0]$ tensor for the $\omega_\nu = \omega_\mu$ case, while the gradient in Ref. [55] did not describe explicitly what to do in this degenerate limit.

At the end of the SSCHA optimization, apart from the temperature-dependent $\boldsymbol{\mathcal{R}}_{\text{eq}}$ positions and the equilibrium auxiliary force constant matrix $\mathbf{\Phi}_{\text{eq}}$, the code also calculates the anharmonic stress tensor $\boldsymbol{P}$, which includes both quantum and thermal ionic fluctuations, as derivatives of the free energy with respect to a strain tensor $\boldsymbol{\varepsilon}$:

$$\begin{aligned} P_{\alpha\beta} = &- \frac{1}{\Omega_{\text{Vol}}} \frac{\partial \mathcal{F}}{\partial \varepsilon_{\alpha\beta}}\bigg|_{\boldsymbol{\varepsilon}=0} = \left\langle P_{\alpha\beta}^{(\text{BO})}(\boldsymbol{R}) \right\rangle_{\tilde{\rho}_{\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}}} \\ &- \frac{1}{2\Omega_{\text{Vol}}} \sum_{s=1}^{N_{\text{a}}} \left\langle u_s^\alpha \mathrm{f}^{(\text{BO})\beta}_s + u_s^\beta \mathrm{f}^{(\text{BO})\alpha}_s \right\rangle_{\tilde{\rho}_{\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}}} . \end{aligned} \tag{2.19}$$

Here, we have made explicit the atomic index s (lower index) and Cartesian α, β (upper index) of $\boldsymbol{u}$ and $\mathbf{f}^{(\text{BO})}$. $\boldsymbol{P}^{(\text{BO})}(\boldsymbol{R})$ is the Born-Oppenheimer stress tensor of the configuration with ions displaced in the $\boldsymbol{R}$ coordinates. This equation is slightly different from the stress tensor equation presented in [57]. The two equations coincide at equilibrium, but this is more general. The derivation of Eq. (2.19) is reported in Appendix A. Thanks to the temperature-dependent stress, the SSCHA code can optimize also the lattice parameters and the volume. Thus, by relaxing the lattice at different temperatures, we get the thermal expansion straightforwardly.

Remarkably, the stress tensor of Eq. (2.19) can be computed with a single SSCHA minimization at fixed volume. This is a huge advantage with respect to the standard quasi-harmonic approximation, not only because it includes quantum and anharmonic effects, but also because it is computationally much more efficient. In fact, the quasiharmonic approximation requires performing harmonic phonon calculations at different volumes (and/or internal lattice positions) to estimate the minimum of the quasi-harmonic free energy with finite differences. This process is extremely cumbersome for crystals with few symmetries and lots of internal degrees of freedom in the structure.

In the current implementation of the code, the symmetries of the space group are imposed *a posteriori* on the gradients of Eqs. (2.12) and (2.13), as well as on (2.19). This assures that the density matrix satisfies all the symmetries at each step of the minimization. Thus, during the geometry optimization, the system cannot lose any

symmetry, though it can gain them. The symmetries are imposed following the methodology explained in Appendix D, which is different to the method originally conceived[55]. The current SSCHA code can also work without imposing symmetries, allowing for symmetry loss, though the stochastic number of configurations needed to converge the minimization is larger (see Sec. 2.4.2).

The stochastic sampling

The stochastic nature of the SSCHA comes from the Monte Carlo evaluation of the averages in Eqs. (2.12), (2.13), and (2.19). A set of random ionic configurations are created in a chosen supercell according to the Gaussian ionic probability distribution

$$\tilde{\rho}_{\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}}(\boldsymbol{R}) = \sqrt{\det(\boldsymbol{\Psi}^{-1}/2\pi)} \times \exp\left[-\frac{1}{2}\sum_{ab}(R_a - \mathcal{R}_a){\Psi^{-1}}_{ab}(R_b - \mathcal{R}_b)\right]. \tag{2.20}$$

The Monte Carlo average of a generic observable $O(\boldsymbol{R})$, function only of the ionic position $\boldsymbol{R}$, is calculated then as weighted sum over the created ensemble:

$$\langle O(\boldsymbol{R})\rangle_{\tilde{\rho}_{\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}}} = \frac{1}{\sum_{j=1}^{N_\mathrm{c}}\rho_j}\sum_{j=1}^{N_\mathrm{c}}\rho_j O(\boldsymbol{R}_{\{j\}}). \tag{2.21}$$

Here, N_c is the total number of configurations in the ensemble, while $\boldsymbol{R}_{\{j\}}$ is the j-th ionic randomly displaced configuration. Each of the $\boldsymbol{R}_{\{j\}}$ configurations is generated according to the initial trial ionic distribution $\tilde{\rho}_{\boldsymbol{\mathcal{R}}^{(0)},\boldsymbol{\Phi}^{(0)}}(\boldsymbol{R})$ from which the minimization starts. To improve the stochastic accuracy, for each $\boldsymbol{R}_{\{j\}}$ configuration also $-\boldsymbol{R}_{\{j\}}$ is created, benefiting from $\tilde{\rho}_{\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}}(\boldsymbol{R}) = \tilde{\rho}_{\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}}(-\boldsymbol{R})$ property of the Gaussian distribution.

The ρ_j weights are computed and updated along the free energy minimization as the values of $\boldsymbol{\mathcal{R}}$ and $\boldsymbol{\Phi}$ change:

$$\rho_j = \frac{\tilde{\rho}_{\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}}\left(\boldsymbol{R}_{\{j\}}\right)}{\tilde{\rho}_{\boldsymbol{\mathcal{R}}^{(0)},\boldsymbol{\Phi}^{(0)}}\left(\boldsymbol{R}_{\{j\}}\right)}. \tag{2.22}$$

At the beginning, when the ensemble has just been generated and $\boldsymbol{\mathcal{R}} = \boldsymbol{\mathcal{R}}^{(0)}$ and $\boldsymbol{\Phi} = \boldsymbol{\Phi}^{(0)}$, all values of $\rho_j = 1$. However, as the $\boldsymbol{\mathcal{R}}$ and $\boldsymbol{\Phi}$ are updated during the minimization, the weights change. This reweighting technique is commonly used in Monte Carlo methods[73, 74] and takes the name of *importance sampling*. This allows avoiding generating a new ensemble and computing *ab initio* energies and forces at each step of the minimization, speeding up the SSCHA calculation.

Minimization algorithm

The minimization strategy implemented in the SSCHA code for the free energy is based on a preconditioned gradient descent. At each step, the $\boldsymbol{\mathcal{R}}$ and $\boldsymbol{\Phi}$ are updated as

$$\boldsymbol{\Phi}^{(n+1)} = \boldsymbol{\Phi}^{(n)} - \lambda_{\boldsymbol{\Phi}}\sum_{ab}\left(\frac{\partial^2\mathcal{F}}{\partial\boldsymbol{\Phi}\partial\Phi_{ab}}\right)^{-1}\frac{\partial\mathcal{F}}{\partial\Phi_{ab}} \tag{2.23}$$

$$\boldsymbol{\mathcal{R}}^{(n+1)} = \boldsymbol{\mathcal{R}}^{(n)} - \lambda_{\boldsymbol{\mathcal{R}}} \sum_a \left(\frac{\partial^2 \mathcal{F}}{\partial \boldsymbol{\mathcal{R}} \partial \mathcal{R}_a} \right)^{-1} \frac{\partial \mathcal{F}}{\partial \mathcal{R}_a}. \tag{2.24}$$

In a perfectly quadratic landscape, this algorithm assures the convergence in just one step if both $\lambda_{\boldsymbol{\Phi}}$ and $\lambda_{\boldsymbol{\mathcal{R}}}$ are set equal to one. However, in order to avoid too big steps in the minimization, often it is more convenient to chose $\lambda_{\boldsymbol{\Phi}|\boldsymbol{\mathcal{R}}} < 1$. This algorithm, with Hessian matrices that multiplies the gradient, is the *preconditioned* steepest descent. If the *preconditioning* option is set to false, a standard steepest descent minimization is followed instead, with $\lambda_{\boldsymbol{\Phi}}$ and $\lambda_{\boldsymbol{\mathcal{R}}}$ re-scaled to the maximum eigenvalue of the $\frac{\partial^2 \mathcal{F}}{\partial \boldsymbol{\Phi}^2}$ and $\frac{\partial^2 \mathcal{F}}{\partial \boldsymbol{\mathcal{R}}^2}$ Hessian matrices, respectively, in order to have adimensional values independent on the system.

The preconditioning Hessian matrices that multiplies the gradients in Eq. (2.23) and Eq. (2.24) are approximated by the code. Following the procedure introduced in Ref.[57], we use the exact Hessian in the minimum of a perfectly harmonic oscillator with the same frequencies as the SSCHA auxiliary Hamiltonian. In particular, they are:

$$\frac{\partial^2 \mathcal{F}}{\partial \Phi_{ab} \partial \Phi_{cd}} \approx \frac{1}{2} \frac{\partial \Psi_{ab}}{\partial \Phi_{cd}} \tag{2.25}$$

and

$$\frac{\partial^2 \mathcal{F}}{\partial \boldsymbol{\mathcal{R}} \partial \boldsymbol{\mathcal{R}}} \approx \boldsymbol{\Phi}. \tag{2.26}$$

Eq. (2.25) is presented differently from the original work in which it was derived[57]. We prove in Appendix B that they are exactly the same. Considering that the Hessian preconditioner cancels out the $\frac{1}{2} \frac{\partial \Psi_{ab}}{\partial \Phi_{cd}}$ term in Eq. (2.13), the resulting update of the variational parameters at each step in the minimization is performed as

$$\begin{aligned} \Phi_{ab}^{(n+1)} &= \Phi_{ab}^{(n)} + \\ &- \lambda_{\boldsymbol{\Phi}} \left\langle \left(\mathrm{f}_b^{(\mathrm{BO})}(\boldsymbol{R}) - \mathrm{f}_b^{\mathcal{H}_{\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}}}(\boldsymbol{R}) \right) \sum_c {\Psi^{-1}}_{ac} \left(R_c - \mathcal{R}_c \right) \right\rangle_{\tilde{\rho}_{\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}}} \end{aligned} \tag{2.27}$$

and

$$\mathcal{R}_a^{(n+1)} = \mathcal{R}_a^{(n)} + \lambda_{\boldsymbol{\mathcal{R}}} \sum_b \Phi_{ab}^{-1} \left\langle \mathrm{f}_b^{(\mathrm{BO})}(\boldsymbol{R}) - \mathrm{f}_b^{\mathcal{H}_{\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}}}(\boldsymbol{R}) \right\rangle_{\tilde{\rho}_{\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}}}. \tag{2.28}$$

This implementation is very efficient, especially for Eq. (2.27), as there is no need to calculate the $\Lambda_{\boldsymbol{\mathcal{R}}}[0]$ tensor. Therefore, computing directly Eq. (2.27) is much faster than calculating the gradient of Eq. (2.13).

The code allows the user to select a different minimization algorithm specifically for the minimization with respect to $\boldsymbol{\Phi}$: the *root representation*. Since the minimization with respect to $\boldsymbol{\Phi}$ is the more challenging, this technique aims to further improving the $\boldsymbol{\Phi}$ optimization. In particular, the gradient has a stochastic error and the minimization is performed with a finite step size. For these reasons, $\boldsymbol{\Phi}$ could become non positive definite during the optimization (i.e. the dynamical matrix has imaginary frequencies). If this occurs, the minimization is halted raising an error, as the density matrix of Eq. (2.20) diverges. In such a case, the minimization must be manually restarted, either by taking a smaller step or by stopping the minimization before reaching imaginary frequencies (fixing the maximum number of steps). This kind of halts do not occur often when using the preconditioning in the

minimization. However, they may be encountered if few configurations are generated for each ensemble or the starting dynamical matrix is very far from equilibrium.

To solve these problems, we implement the root representation, in which, instead of updating $\boldsymbol{\Phi}$ as in Eq. (2.23), it updates a root of $\boldsymbol{\Phi}$:

$$\sqrt[n]{\boldsymbol{\Phi}}^{(i+1)} = \sqrt[n]{\boldsymbol{\Phi}}^{(i)} - \lambda_{\boldsymbol{\Phi}} \boldsymbol{\mathcal{G}}_{\boldsymbol{n}}. \tag{2.29}$$

The updating direction $\boldsymbol{\mathcal{G}_n}$ depends depends on the root order n:

$$\boldsymbol{\mathcal{G}_2} = \sqrt{\boldsymbol{\Phi}} \cdot \frac{\partial \mathcal{F}}{\partial \boldsymbol{\Phi}} + \frac{\partial \mathcal{F}}{\partial \boldsymbol{\Phi}} \cdot \sqrt{\boldsymbol{\Phi}}, \tag{2.30}$$

where with the $\cdot$ we indicate a matrix product. Similarly,

$$\boldsymbol{\mathcal{G}_4} = \sqrt[4]{\boldsymbol{\Phi}} \cdot \boldsymbol{\mathcal{G}_2} + \boldsymbol{\mathcal{G}_2} \cdot \sqrt[4]{\boldsymbol{\Phi}}. \tag{2.31}$$

We select the positive definite root matrix. Indeed, after the step of Eq. (2.29) the original force constant matrix is obtained as

$$\boldsymbol{\Phi}^{(i+1)} = \left(\sqrt[n]{\boldsymbol{\Phi}}^{(i+1)} \right)^n. \tag{2.32}$$

Thanks to the definition in Eq. (2.32), the dynamical matrix is always positive definite for any even value of n.

The *root representation* is independent of the preconditioning. With preconditioning, we replace the free energy gradient in Eq. (2.30) with the preconditioned direction in Eq. (2.27) (the gradient multiplied by the approximated Hessian). This is different from what was proposed in the original work[57], where the Hessian matrix was computed also for the $\sqrt{\boldsymbol{\Phi}}$ and $\sqrt[4]{\boldsymbol{\Phi}}$ cases. However, we noticed that in systems with many atoms, the Hessian matrix calculation becomes the bottleneck as it scales with N_{a}^6. The implementation here described allows for a much faster $\boldsymbol{\Phi}$ update and avoids calculating the Hessian matrix. The drawback is that the optimization step is not as optimal as it would be if the proposal in Ref. [57] was followed. The code offers six combinations for the minimization procedure: no root, square root ($n = 2$), and fourth-root ($n = 4$), all of them with or without the preconditioned direction. The optimal minimization step is $n = 1$ with preconditioning. If the square root is employed ($n = 2$), it is preferable to use the preconditioning. If fourth-root is employed ($n = 4$), the best performances are without *preconditioning*.

The lattice geometry optimization

The lattice degrees of freedom $\{\boldsymbol{a}_i\}$ are relaxed only after the minimization of the free energy with respect to $\boldsymbol{\mathcal{R}}$ and $\boldsymbol{\Phi}$ at a constant volume stops (see Sec. 2.4.2 for a detailed description of the stopping criteria). For this reason, the lattice geometry optimization is an "outer" optimization: at each step of the lattice geometry optimization, we perform a full free energy minimization with respect to the centroids $\boldsymbol{\mathcal{R}}$ and auxiliary force constants $\boldsymbol{\Phi}$. This means that each step of the lattice geometry optimization is performed with a different ensemble, whose configurations are all generated with the same lattice vectors.

To update the lattice, the code calculates the stress tensor with Eq. (2.19), and generates a strain for the lattice as

$$\varepsilon_{\alpha\beta} = \Omega_{\text{Vol}} \left(P_{\alpha\beta} - P^* \delta_{\alpha\beta} \right), \tag{2.33}$$

where P^* is the target pressure of the relaxation and $\delta_{\alpha\beta}$ is the Kronecker delta. The lattice parameters $\{\boldsymbol{a}_i\}$ are updated as

$$a'_{i\alpha} = a_{i\alpha} + \lambda_{\{\boldsymbol{a}_i\}} \sum_{\beta} \varepsilon_{\alpha\beta} a_{i\beta}, \tag{2.34}$$

where $\lambda_{\{\boldsymbol{a}_i\}}$ is the update step. Since each step requires a new ensemble, it is crucial to reduce the number of steps to reach convergence by properly picking the right value for $\lambda_{\{\boldsymbol{a}_i\}}$. In an isotropic material with a constant bulk modulus

$$B_0 = \Omega_{\text{Vol}} \left. \frac{d^2 V^{(BO)}}{d\Omega_{\text{Vol}}^2} \right|_{\boldsymbol{R}=\boldsymbol{\mathcal{R}}}, \tag{2.35}$$

the optimal value of the step is

$$\lambda_{\{\boldsymbol{a}_i\}} = \frac{1}{3\Omega_{\text{Vol}} B_0}. \tag{2.36}$$

B_0 is an input parameter given in GPa units. Good values of B_0 may range from 10 GPa for crystals at ambient conditions, like ice, up to 800 GPa in systems at Mbar pressures (or for diamond). Remember that increasing the value of B_0 produces smaller steps in the cell parameters. The optimal value of B_0 to assure the fastest convergence can be estimated by manually computing it from Eq. (2.35), by taking finite differences of the pressure obtained at two uniformly strained volumes, or by looking for the experimental value of similar compounds.

Alternatively to the fixed pressure optimization, it is also possible to perform the geometry lattice optimization at fixed volume. In this case P^* is recomputed at each step so that $\text{Tr}\{\boldsymbol{\varepsilon}\} = 0$. In this case, the final lattice parameters are also rescaled so that the final volume matches the one before the step. Since this algorithm has one less degree of freedom than the fixed pressure one, it usually converges faster.

The code flowchart

To start a SSCHA simulation, we need a starting guess on the trial positive definite force constants matrix $\boldsymbol{\Phi}^{(0)}$ and on the average atomic positions $\boldsymbol{\mathcal{R}}^{(0)}$. Even if in principle the starting point is arbitrary, the closer to the solution we begin, the faster the minimization converges. Thus, the coordinates at the minimum of the Born-Oppenheimer energy landscape and the harmonic force constants are usually good starting points, which can be obtained from any code that computes phonons. The supercell of the simulation is given by the dimension of the input force constants matrix, while the centroids $\boldsymbol{R}$ are defined in the unit cell (they satisfy translational symmetry). If the original dynamical matrix contains imaginary frequencies, it can be reverted to positive definite as

$$\Phi_{ab}^{(0)} = \sqrt{M_a M_b} \sum_{\mu} |\omega_\mu^2| e_\mu^a e_\mu^b. \tag{2.37}$$

Then, the first random ensemble (that we call population in the SSCHA language) can be generated. For each configuration inside the population, its total Born-Oppenheimer energy as well as its classical atomic forces $\mathbf{f}^{(\mathrm{BO})}$ and stress tensor $\boldsymbol{P}^{(\mathrm{BO})}$ must be computed. This is done with an external code, either manually (by computing externally the energies, forces, and stress tensor, and loading them back into the SSCHA code), or automatically (with an appropriate configuration discussed in Sec. 2.6). Once the Born-Oppenheimer energies, forces, and stress tensor of all the configurations have been computed, the minimization starts. The gradients of the free energy are computed as described in Sec. 2.4.2 and the minimization continues either until the stochastic sampling is not good or the algorithm converges.

If $\boldsymbol{\mathcal{R}}$ and $\boldsymbol{\Phi}$ change a lot during the minimization of the free energy, the original ensemble no longer describes well the new probability distribution $\tilde{\rho}_{\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}}(\boldsymbol{R})$, and the stochastic error increases. This occurrence is automatically checked by the SSCHA code calculating the Kong-Liu[75, 76] *effective sample size* N_{eff}:

$$N_{eff} = \frac{\sum_{j=1}^{N_\mathrm{c}} \rho_j^2}{\left(\sum_{j=1}^{N_\mathrm{c}} \rho_j\right)^2}. \tag{2.38}$$

The minimization is halted when the ratio between N_{eff} and the number of configurations N_c is lower than a parameter η defined by the user:

$$\frac{N_{eff}}{N_\mathrm{c}} < \eta. \tag{2.39}$$

A standard value of η that ensures a correct minimization is 0.5, but it can be convenient to lower it a bit to accelerate convergence in the first steps.

The convergence, on the contrary, is achieved only if the two gradients with respect to $\boldsymbol{\mathcal{R}}$ and $\boldsymbol{\Phi}$ are lower than a given threshold:

$$\left|\frac{\partial \mathcal{F}}{\partial \boldsymbol{\Phi}}\right| < \delta_{\boldsymbol{\Phi}} \tag{2.40}$$

$$\left|\frac{\partial \mathcal{F}}{\partial \boldsymbol{\mathcal{R}}}\right| < \delta_{\boldsymbol{\mathcal{R}}}. \tag{2.41}$$

The δ threshold is provided by the user and re-scaled at each step by the estimation of the stochastic error on the corresponding gradient (*meaningful_factor*). So, at each step, δ is

$$\delta_{\boldsymbol{\Phi}} = \text{meaningful\_factor} \cdot \left|\Delta \frac{\partial \mathcal{F}}{\partial \boldsymbol{\Phi}}\right| \tag{2.42}$$

$$\delta_{\boldsymbol{\mathcal{R}}} = \text{meaningful\_factor} \cdot \left|\Delta \frac{\partial \mathcal{F}}{\partial \boldsymbol{\mathcal{R}}}\right|. \tag{2.43}$$

In this way, the variable *meaningful_factor* is independent on the system size or the number of configurations used.

If the lattice parameters are free to move, then an additional condition must be fulfilled in order to end the minimization: each component of the strain per unit-cell volume Ω_Vol must be smaller than the stochastic error on the stress tensor:

$$\frac{\varepsilon_{\alpha\beta}}{\Omega_\mathrm{Vol}} \leq \Delta P_{\alpha\beta}. \tag{2.44}$$

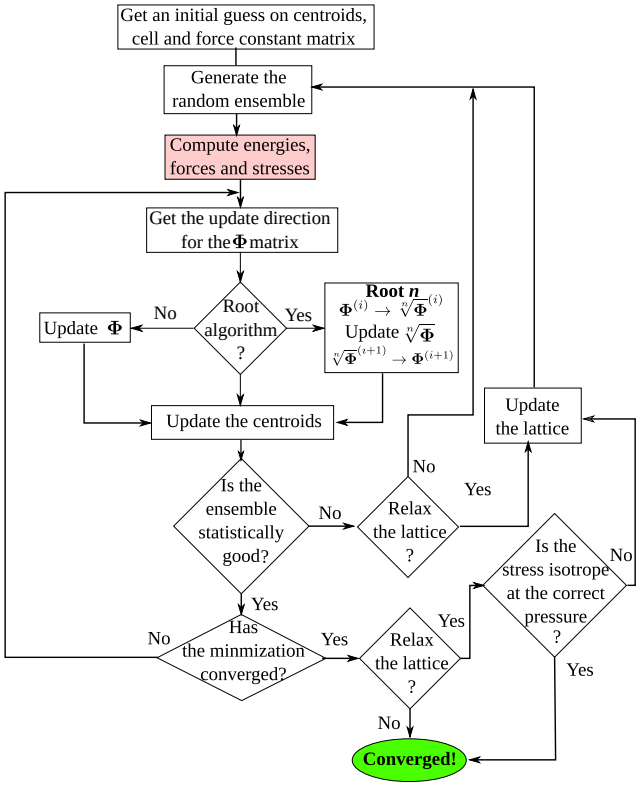

Figure 2.2. Flowchart of the SSCHA code. The most time consuming part of the diagram is the *ab initio* calculation of the Born-Oppenheimer forces, energies, and stress tensors for all the configurations inside the ensemble, and it is shaded in red. All the other steps usually take few seconds when executed on a standard workstation, even in systems that contain several hundreds of atoms.

If Eq. (2.44) is not fulfilled, even if all the gradients are lower than the chosen threshold, the code generates a new ensemble and continues (until both conditions are satisfied).

At the end of the minimization, the output of the SSCHA minimization gives the total free energy (with stochastic error), the average equilibrium ionic positions $\boldsymbol{\mathcal{R}}_{\text{eq}}$, the equilibrium auxiliary force constant matrix $\boldsymbol{\Phi}_{\text{eq}}$, and the stress tensor $\boldsymbol{P}$. All output quantities are temperature-dependent, and include quantum-thermal fluctuations and anharmonicity. A flowchart that represents the whole execution of a SSCHA run is presented in Figure 2.2. If the SSCHA code is coupled with an *ab initio* total-energy engine, the most expensive calculation in the flowchart is by far the calculation of Born-Oppenheimer energy, forces, and stress tensors on the whole ensemble, which may contain up to several hundreds or thousands of configurations. For this reason, the pretty complex workflow we set up is aimed to pass by the calculation of a new ensemble as few times as possible. Most materials studied are converged within 3 populations, and the CPU time required to minimize each population is few minutes on a single CPU of modern laptops.

2.4.3 The self-consistent equation and possible alternative implementations of the SSCHA

The preconditioned gradient descent approach sketched above offers a very efficient implementation of the SSCHA theory, in which the anharmonic free energy is optimized by all degrees of freedom in the crystal structure, including internal coordinates as well as lattice vectors. If the centroid positions $\boldsymbol{\mathcal{R}}$ are kept fixed in the minimization, the SSCHA self-consistent equation

$$\Phi_{ab}(\boldsymbol{\mathcal{R}}) = \left\langle \frac{\partial^2 V}{\partial R_a \partial R_b} \right\rangle_{\tilde{\rho}_{\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}(\boldsymbol{\mathcal{R}})}} \tag{2.45}$$

offers an alternative way of implementing the SSCHA theory (see Ref. [56] for a proof of Eq. (2.45)). It is important to underline the self-consistent condition required by the equation above, as the quantum statistical average is taken with a density matrix dependent on $\boldsymbol{\Phi}(\boldsymbol{\mathcal{R}})$, which must equal the result of the average. As in this approach the centroid positions are not optimized, the obtained auxiliary force constant matrix depends parametrically on $\boldsymbol{\mathcal{R}}$.

The self-consistent equation can be implemented stochastically, following the procedure outlined in Sec. 2.4.2. By using integration by parts [56], the right-hand-side of Eq. (2.45) can be rewritten in terms of forces and displacements. Thus, with the importance sampling technique and reweighting, the equation can be solved by calculating forces in supercells generated with the SSCHA density matrix. An equivalent approach [63] is to extract the auxiliary force constants by fitting the obtained forces in the supercells generated with the SSCHA density matrix to Eq. (2.14). This approach has been followed recently [63, 77, 78], where a least-squares technique is followed for the fitting.

The use of the self-consistent equation is valid, thus, only for fixed centroid positions. If the centroid positions want to be optimized as well within this approach, the self-consistent procedure should be repeated for different values of $\boldsymbol{\mathcal{R}}$, calculate the free energy for these positions, and see where its minimum is. Clearly this is a very cumbersome procedure unless centroid positions are fixed by symmetry. Moreover, solving the self-consistent equation fixing the centroid positions at the classical $\boldsymbol{R}_0$ positions, which it is usually the case [63, 77, 78], neglects all the effects of quantum and thermal fluctuations on the structure. Since within our approach based on the gradient descent we can optimize the free energy not only with respect to the auxiliary force constants but also all degrees of freedom in the crystal structure, the workflow outlined in Sec. 2.4.2 provides a full picture of the effect of quantum/thermal fluctuations as well as anharmonicity on crystals, much more efficient than the approaches based on Eq. (2.45).

2.5 Post-minimization tools: positional free energy Hessian, phonon spectral functions, and phonon linewidths

In the previous section we described how to compute the free energy of a system and fully optimize its structure by taking into account the anharmonicity that arises from both thermal and quantum fluctuations. After the free energy functional

minimization, additional information can be extracted from the results obtained, namely, the second derivative (Hessian) of the positional free energy with respect to the centroids, the anharmonic phonon spectral functions, and the anharmonic frequency linewidths and shifts. In the next subsections I will explain why these quantities are of physical interest and what is the strategy adopted by the code to computing them. The theory here reviewed was introduced in Ref. [56] and extensively applied for the first time in *ab initio* calculations to H_3S in Ref. [79].

2.5.1 Positional free energy Hessian

As shown in Sec. 2.3, for a given temperature the free energy at equilibrium F of a system with Hamiltonian H is obtained by minimizing the density-matrix functional $\mathcal{F}[\tilde{\rho}] = \langle K + V(\boldsymbol{R})\rangle_{\tilde{\rho}} - TS[\tilde{\rho}]$:

$$F = \min_{\tilde{\rho}} \mathcal{F}[\tilde{\rho}] = \mathcal{F}[\rho]\,, \tag{2.46}$$

where ρ is the equilibrium density matrix of the system obtained at the minimum. The average atomic positions at equilibrium are $\langle \boldsymbol{R}\rangle_{\rho} = \boldsymbol{\mathcal{R}}_{\mathrm{eq}}$. By minimizing the functional keeping fixed the average atomic positions in a generic configuration $\boldsymbol{\mathcal{R}}$, $\langle \boldsymbol{R}\rangle_{\tilde{\rho}} = \boldsymbol{\mathcal{R}}$, we define the positional free energy $F(\boldsymbol{\mathcal{R}})$:

$$F(\boldsymbol{\mathcal{R}}) = \min_{\substack{\tilde{\rho} \\ \langle \boldsymbol{R}\rangle_{\tilde{\rho}} = \boldsymbol{\mathcal{R}}}} \mathcal{F}[\tilde{\rho}] = \mathcal{F}[\rho_{\boldsymbol{\mathcal{R}}}]\,, \tag{2.47}$$

where $\rho_{\boldsymbol{\mathcal{R}}}$ is the density matrix giving the constrained minimum for the considered average position $\boldsymbol{\mathcal{R}}$. Since

$$F = F(\boldsymbol{\mathcal{R}}_{\mathrm{eq}}) = \min_{\boldsymbol{\mathcal{R}}} F(\boldsymbol{\mathcal{R}})\,, \tag{2.48}$$

$\boldsymbol{\mathcal{R}}$ and $F(\boldsymbol{\mathcal{R}})$ can be interpreted as a multidimensional order parameter and a thermodynamic potential, respectively, in the study of displacive phase transitions according to Landau's theory. Properly speaking, the "order" parameter would be $\boldsymbol{\mathcal{R}} - \boldsymbol{\mathcal{R}}_{\mathrm{hs}}$, where $\boldsymbol{\mathcal{R}}_{\mathrm{hs}}$ is the average position of the atoms when the system is in the high-symmetry phase. Therefore, the knowledge of the positional free energy landscape as a function of external parameters, like temperature or pressure, gives crucial information about the structural stability and evolution of a system, as it allows to determine the (meta-)stable configurations corresponding to (local) minima of the positional free energy.

The Hessian of the positional free energy $F(\boldsymbol{\mathcal{R}})$ in the equilibrium configuration is the inverse response function to a static perturbation on the nuclei (i.e. the inverse of the static susceptibility). In presence of a second order phase transition the static response function diverges, which results in one or more eigenvalues of the positional free energy Hessian going to zero. This means that the occurrence of displacive second-order phase transitions, like CDW or ferroelectric transitions [72, 80, 81, 82, 83], can be characterized by analyzing the evolution with temperature of the eigenvalues of the equilibrium positional free energy Hessian. Typically, in these cases at high temperature the minimum point of the free energy, i.e. the equilibrium configuration $\boldsymbol{\mathcal{R}}_{\mathrm{eq}}$, is a high-symmetry configuration $\boldsymbol{\mathcal{R}}_{\mathrm{hs}}$. Therefore, at high temperature the free

energy Hessian in $\boldsymbol{\mathcal{R}}_{\rm hs}$ is positive definite, i.e. its eigenvalues are positive. As the temperature decreases, the minimum in $\boldsymbol{\mathcal{R}}_{\rm hs}$ becomes less and less deep, until it becomes a saddle point at the transition temperature (i.e. at least one eigenvalue is zero), so that a second-order displacive phase transition occurs and the equilibrium configuration $\boldsymbol{\mathcal{R}}_{\rm eq}$ moves towards lower-symmetry configurations that reduce the free energy as the temperature decreases further (following the pattern indicated by the eigenvector of the vanishing eigenvalue). Using the same approach, it is possible to characterize second-order displacive phase transitions driven by other external parameters, like the pressure in high-pressure superconducting hydrides [54, 67, 68, 79, 84].

The role played by the eigenvalues and eigenvectors of the positional free energy Hessian in tracing the system's structural stability recalls the role played by the harmonic dynamical matrix in the standard harmonic approximation, but now including lattice thermal and quantum anharmonic effects in the dynamics of the nuclei. Therefore, the Hessian of the positional free energy, divided by the masses, $D^{(\rm F)}_{ab} = \partial^2 F/\partial\mathcal{R}^a\partial\mathcal{R}^b\big|_{\boldsymbol{\mathcal{R}}_{\rm eq}}/\sqrt{M_a M_b}$, can be considered a natural generalization of the harmonic dynamical matrix that, however, includes thermal and quantum effects.

What is explained hitherto about the role played by the positional free energy and its Hessian is general. In particular, the evaluation of the positional free energy within the SSCHA is pretty straightforward. Indeed, the average position for a trial harmonic density matrix $\tilde{\rho}_{\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}}$ coincides with the centroid parameter $\boldsymbol{\mathcal{R}}$,

$$\langle \boldsymbol{R} \rangle_{\tilde{\rho}_{\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}}} = \boldsymbol{\mathcal{R}}\,. \tag{2.49}$$

Thus, within the SSCHA the positional free energy is obtained by minimizing the SSCHA free energy functional $\mathcal{F}[\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}]$ with respect to the trial quadratic amplitude $\boldsymbol{\Phi}$ only:

$$F(\boldsymbol{\mathcal{R}}) = \min_{\boldsymbol{\Phi}} \mathcal{F}[\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}]\,. \tag{2.50}$$

The auxiliary force constants that minimize Eq. (2.50) for a given $\boldsymbol{\mathcal{R}}$ position of the centroids will be labeled in the following as $\boldsymbol{\Phi}_{\boldsymbol{\mathcal{R}}}$. Solving Eq. (2.50) allows to employ the SSCHA code to have direct access to $F(\boldsymbol{\mathcal{R}})$ for any $\boldsymbol{\mathcal{R}}$ and, in principle, to compute the Hessian by finite differences. However, as discussed above, such a finite-difference approach would be extremely expensive for two main reasons. First, it would require a large number of configurations in the ensemble to reduce the stochastic error and calculate the derivatives by finite differences. Second, because the large number of degrees of freedom in $\boldsymbol{\mathcal{R}}$ prevents any realistic finite-difference approach. Luckily the SSCHA code allows to avoid any cumbersome finite-difference approach by exploiting an analytic formula for the positional free energy Hessian.

Before describing the analytic formula, let us introduce a notation that will simplify the mathematical expressions. Given two tensors $\boldsymbol{X}$ and $\boldsymbol{Y}$, with the single dot product $\boldsymbol{X}\cdot\boldsymbol{Y}$ we will indicate the contraction of the last index of $\boldsymbol{X}$ with the first index of $\boldsymbol{Y}$, $\sum_c X_{\dots c} Y_{c\dots}$. Likewise, with the double-dot product $\boldsymbol{X}:\boldsymbol{Y}$ we will indicate the contraction of the last two indices of $\boldsymbol{X}$ with the first two indices of $\boldsymbol{Y}$, $\sum_{cd} X_{\dots cd} Y_{cd\dots}$. Moreover, any fourth-order tensor X_{pqlm} can be interpreted as a "super" matrix X_{AB}, with the composite indices $A = (pq)$ and $B = (lm)$, and vice versa (through this correspondence we can define, for example, the inverse of

a fourth-order tensor and the identity fourth-order tensor $\mathbb{1}$). Using this notation, we can express the positional free energy Hessian, $1/\sqrt{M_a M_b}\ \partial^2 F/\partial\mathcal{R}^a\partial\mathcal{R}^b$, in component-free form as

$$\begin{aligned}\frac{1}{\sqrt{\boldsymbol{M}}}\cdot\frac{\partial^2 F}{\partial\boldsymbol{\mathcal{R}}\partial\boldsymbol{\mathcal{R}}}\cdot\frac{1}{\sqrt{\boldsymbol{M}}} &= \\ &= \boldsymbol{D}_{\boldsymbol{\mathcal{R}}} + \overset{(3)}{\boldsymbol{D}}_{\boldsymbol{\mathcal{R}}} : \boldsymbol{\Lambda}_{\boldsymbol{\mathcal{R}}}[0] : \left[\mathbb{1} - \overset{(4)}{\boldsymbol{D}}_{\boldsymbol{\mathcal{R}}} : \boldsymbol{\Lambda}_{\boldsymbol{\mathcal{R}}}[0]\right]^{-1} : \overset{(3)}{\boldsymbol{D}}_{\boldsymbol{\mathcal{R}}}\,, \end{aligned} \tag{2.51}$$

where $M_{ab} = \delta_{ab} M_a$ is the mass matrix,

$$\begin{aligned}(D_{\boldsymbol{\mathcal{R}}})_{ab} &= \frac{1}{\sqrt{M_a M_b}}\left\langle\frac{\partial^2 V}{\partial R^a \partial R^b}\right\rangle_{\rho_{\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}_{\boldsymbol{\mathcal{R}}}}} \\ &= \frac{(\Phi_{\boldsymbol{\mathcal{R}}})_{ab}}{\sqrt{M_a M_b}}\,, \end{aligned} \tag{2.52}$$

$$\begin{aligned}(\overset{(3)}{D}_{\boldsymbol{\mathcal{R}}})_{abc} &= \frac{1}{\sqrt{M_a M_b M_c}}\left\langle\frac{\partial^3 V}{\partial R^a \partial R^b \partial R^c}\right\rangle_{\rho_{\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}_{\boldsymbol{\mathcal{R}}}}} \\ &= \frac{(\overset{(3)}{\Phi}_{\boldsymbol{\mathcal{R}}})_{abc}}{\sqrt{M_a M_b M_c}}\,, \end{aligned} \tag{2.53}$$

$$\begin{aligned}(\overset{(4)}{D}_{\boldsymbol{\mathcal{R}}})_{abcd} &= \frac{1}{\sqrt{M_a M_b M_c M_d}}\left\langle\frac{\partial^4 V}{\partial R^a \partial R^b \partial R^c \partial R^d}\right\rangle_{\rho_{\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}_{\boldsymbol{\mathcal{R}}}}} \\ &= \frac{(\overset{(4)}{\Phi}_{\boldsymbol{\mathcal{R}}})_{abcd}}{\sqrt{M_a M_b M_c M_d}}\,, \end{aligned} \tag{2.54}$$

and $\boldsymbol{\Lambda}_{\boldsymbol{\mathcal{R}}}[0]$ is the $z = 0$ value of the fourth-order tensor $\boldsymbol{\Lambda}_{\boldsymbol{\mathcal{R}}}[z]$, already introduced in Eq. (2.18). In the equations above the quantum statistical averages are taken with $\rho_{\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}_{\boldsymbol{\mathcal{R}}}}$, which for a given $\boldsymbol{\mathcal{R}}$ position of the centroids is taken with the $\boldsymbol{\Phi}_{\boldsymbol{\mathcal{R}}}$ auxiliary force constants that minimize the free energy. $\boldsymbol{\Lambda}_{\boldsymbol{\mathcal{R}}}[z]$ is given in components by

$$(\Lambda_{\boldsymbol{\mathcal{R}}}[z])^{abcd} = \sum_{\mu\nu}\mathscr{F}(z,\omega_\mu,\omega_\nu)\, e^a_\nu e^b_\mu e^c_\nu e^d_\mu, \tag{2.55}$$

where ω_μ^2 and e^a_ν are the eigenvalues and eigenvectors of $\boldsymbol{D}_{\boldsymbol{\mathcal{R}}}$, and

$$\begin{aligned}\mathscr{F}(z,\omega_\nu,\omega_\mu) = \frac{\hbar}{4\omega_\mu\omega_\nu}\Bigg[&\frac{(\omega_\mu-\omega_\nu)(n_\mu-n_\nu)}{(\omega_\mu-\omega_\nu)^2 - z^2} + \\ &- \frac{(\omega_\mu+\omega_\nu)(1+n_\mu+n_\nu)}{(\omega_\mu+\omega_\nu)^2 - z^2}\Bigg]. \end{aligned} \tag{2.56}$$

The only difference between $\boldsymbol{\Lambda}_{\boldsymbol{\mathcal{R}}}[0]$ and $\boldsymbol{\Lambda}[0]$ (introduced in Eq. (2.18)) is that in the former the eigenvalues and eigenvectors entering the equation are those associated to the $\boldsymbol{\Phi}_{\boldsymbol{\mathcal{R}}}$ auxiliary force constants at the centroid positions $\boldsymbol{\mathcal{R}}$, while in the latter this is not necessarily the case. The subindex $\boldsymbol{\mathcal{R}}$ in the equations above precisely indicates that the averages are calculated with a density matrix defined by $\boldsymbol{\mathcal{R}}$ and $\boldsymbol{\Phi}_{\boldsymbol{\mathcal{R}}}$ (after a full SSCHA relaxation at fixed nuclei position $\boldsymbol{\mathcal{R}}$). We will refer to

the $\overset{(n)}{\Phi}_{\boldsymbol{\mathcal{R}}}$ tensors as the nth-order SSCHA force constants (FCs). Note that for the second-order we drop the (2) upper index.

The SSCHA code computes the free energy positional Hessian through Eq. (2.51). At the end of a SSCHA free energy functional minimization, the SSCHA matrix Eq. (2.52), with its eigenvectors and eigenvalues, is available. Thus, $\boldsymbol{\Lambda}_{\boldsymbol{\mathcal{R}}}[0]$ is readily computable and the only quantities that need some effort to be calculated are the averages of Eqs. (2.53) and (2.54). The code computes them through these equivalent expressions (obtained by integrating by parts):

$$(\overset{(3)}{\Phi}_{\boldsymbol{\mathcal{R}}})_{abc} = -\sum_{pq} (\Psi^{-1}_{\boldsymbol{\mathcal{R}}})_{ap} (\Psi^{-1}_{\boldsymbol{\mathcal{R}}})_{bq} \left\langle u^p u^q \mathbb{f}_c \right\rangle_{\rho_{\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}_{\boldsymbol{\mathcal{R}}}}} \tag{2.57a}$$

$$(\overset{(4)}{\Phi}_{\boldsymbol{\mathcal{R}}})_{abcd} = -\sum_{pqr} (\Psi^{-1}_{\boldsymbol{\mathcal{R}}})_{ap} (\Psi^{-1}_{\boldsymbol{\mathcal{R}}})_{bq} (\Psi^{-1}_{\boldsymbol{\mathcal{R}}})_{cr} \left\langle u^p u^q u^r \mathbb{f}_d \right\rangle_{\rho_{\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}_{\boldsymbol{\mathcal{R}}}}}, \tag{2.57b}$$

where $\boldsymbol{\Psi}_{\boldsymbol{\mathcal{R}}}$ is the $\boldsymbol{\Psi}$ matrix with $\boldsymbol{\Phi} = \boldsymbol{\Phi}_{\boldsymbol{\mathcal{R}}}$ and

$$\mathbb{f}(\boldsymbol{R}) = \mathbf{f}^{(\mathrm{BO})}(\boldsymbol{R}) - \left\langle \mathbf{f}^{(\mathrm{BO})}(\boldsymbol{R}) \right\rangle_{\rho_{\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}_{\boldsymbol{\mathcal{R}}}}} - \mathbf{f}^{\mathcal{H}_{\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}_{\boldsymbol{\mathcal{R}}}}}(\boldsymbol{R}). \tag{2.58}$$

These averages are computed employing the stochastic approach already described in Sec. 2.4.2 (indeed, as explained in Ref. [56], the choice of Eq. (2.58), among other possible alternatives, aims at reducing the statistical noise). Note that if the calculation of the free energy Hessian is performed at $\boldsymbol{\mathcal{R}}_{\mathrm{eq}}$, $\left\langle \mathbf{f}^{(\mathrm{BO})}(\boldsymbol{R}) \right\rangle_{\rho_{\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}_{\boldsymbol{\mathcal{R}}}}}$ vanishes.

In order to minimize the number of energy-force calculations needed, it is advisable to compute these averages using the same ensemble used to minimize the free energy functional and obtain $\boldsymbol{\Phi}_{\boldsymbol{\mathcal{R}}}$ (at most adding new elements to reduce the statistical noise, if needed). Of course, since in this case the ensemble is not generated from $\boldsymbol{\Phi}_{\boldsymbol{\mathcal{R}}}$, an importance sampling reweighting has to be employed in order to evaluate the averages $\left\langle \cdot \right\rangle_{\rho_{\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}_{\boldsymbol{\mathcal{R}}}}}$. After computing the averages, the code symmetrizes the results with respect to the space group symmetries (including the lattice translation symmetries) and the index-permutation symmetry, following the approach described in Appendix D.

In order to reduce the computational cost, the SSCHA code can also compute the free energy Hessian discarding the contribution coming from the higher-order terms of the geometric-series expansion in Eq. (2.51), i.e. discarding the terms coming from $\overset{(4)}{\boldsymbol{D}}_{\boldsymbol{\mathcal{R}}}$. In many cases this approximation is extremely good, but it must be checked case by case. Within this so called "bubble" approximation, the free energy Hessian becomes

$$\frac{1}{\sqrt{\boldsymbol{M}}} \cdot \frac{\partial^2 F}{\partial \boldsymbol{\mathcal{R}} \partial \boldsymbol{\mathcal{R}}} \cdot \frac{1}{\sqrt{\boldsymbol{M}}} \simeq \boldsymbol{D}_{\boldsymbol{\mathcal{R}}} + \overset{(3)}{\boldsymbol{D}}_{\boldsymbol{\mathcal{R}}} : \boldsymbol{\Lambda}_{\boldsymbol{\mathcal{R}}}[0] : \overset{(3)}{\boldsymbol{D}}_{\boldsymbol{\mathcal{R}}}. \tag{2.59}$$

Using Eq. (2.51), or its approximated expression Eq. (2.59), the SSCHA code can compute the Hessian of the free energy at any $\boldsymbol{\mathcal{R}}$. However, as said, its most significant usage is in $\boldsymbol{\mathcal{R}}_{\mathrm{eq}}$, due to its relevance to characterize displacive second-order phase transitions. In this case, Eq. (2.51) can be written in a quite explanatory form. At the end of a full SSCHA minimization, the obtained $\boldsymbol{\mathcal{R}}_{\mathrm{eq}}$ and $\boldsymbol{\Phi}_{\mathrm{eq}}$ define the so-called SSCHA effective harmonic Hamiltonian

$$\mathcal{H}^{(\mathrm{S})} = K + \frac{1}{2}(\boldsymbol{R} - \boldsymbol{\mathcal{R}}_{\mathrm{eq}}) \cdot \boldsymbol{\Phi}_{\mathrm{eq}} \cdot (\boldsymbol{R} - \boldsymbol{\mathcal{R}}_{\mathrm{eq}}), \tag{2.60}$$

which replaces the conventional harmonic Hamiltonian to define non-interacting bosonic quasiparticles as a basis to describe the collective vibrational excitations in presence of strong anharmonic effects. In terms of the dynamical matrix $D^{(\mathrm{S})}_{ab} = (\Phi_{\mathrm{eq}})_{ab}/\sqrt{M_a M_b}$ of the SSCHA Hamiltonian $\mathcal{H}^{(\mathrm{S})}$, the anharmonic generalization of the dynamical matrix $\boldsymbol{D}^{(\mathrm{F})}$ can be written as

$$\boldsymbol{D}^{(\mathrm{F})} = \boldsymbol{D}^{(\mathrm{S})} + \boldsymbol{\Pi}(0)\,, \tag{2.61}$$

where

$$\boldsymbol{\Pi}(0) = \overset{(3)}{\boldsymbol{D}}_{\mathrm{eq}} : \boldsymbol{\Lambda}_{\mathrm{eq}}[0] : \left[\mathbb{1} - \overset{(4)}{\boldsymbol{D}}_{\mathrm{eq}} : \boldsymbol{\Lambda}_{\mathrm{eq}}[0]\right]^{-1} : \overset{(3)}{\boldsymbol{D}}_{\mathrm{eq}} \tag{2.62}$$

is the static SSCHA self-energy (the reason behind the use of this name will be clear in the Sec. 2.5.3). In particular, in the bubble approximation we have

$$\boldsymbol{D}^{(\mathrm{F})} = \boldsymbol{D}^{(\mathrm{S})} + \overset{(\mathrm{B})}{\boldsymbol{\Pi}}(0)\,, \tag{2.63}$$

where

$$\overset{(\mathrm{B})}{\boldsymbol{\Pi}}(0) = \overset{(3)}{\boldsymbol{D}}_{\mathrm{eq}} : \boldsymbol{\Lambda}_{\mathrm{eq}}[0] : \overset{(3)}{\boldsymbol{D}}_{\mathrm{eq}} \tag{2.64}$$

is the so called "bubble" static self-energy.

In conclusion, after the SSCHA minimization, the code allows to compute the high-order SSCHA force constants, Eqs. (2.57), and the free energy Hessian dynamical matrix

$$\boldsymbol{D}^{(\mathrm{F})}(\boldsymbol{q}) = \boldsymbol{D}^{(\mathrm{S})}(\boldsymbol{q}) + \begin{cases} \boldsymbol{\Pi}(\boldsymbol{q}, 0) & \text{(2.65a)} \\ \overset{(\mathrm{B})}{\boldsymbol{\Pi}}(\boldsymbol{q}, 0) & \text{(2.65b)} \end{cases}$$

(depending on whether the full or only the "bubble" static self-energy is computed) on the $\boldsymbol{q}$-points belonging to the reciprocal space grid commensurate with the real space supercell used to generate the ensemble. Here we are explicitly using the reciprocal-space formalism, i.e. we are Fourier transforming the quantities with respect to the lattice vector indices (see Appendix E.1 for more details). From the softening of the eigenvalues of $\boldsymbol{D}^{(\mathrm{F})}(\boldsymbol{q})$ as a function of external parameters (like temperature or pressure), it is possbile to observe the occurrence of second order displacive phase transitions, characterize the distortion patterns and compute the critical value of the external parameters driving it. Examples of the employ of this method are given for H_3S in Fig. 3 of Ref. [79], with the softening of an optical mode driven by pressure release, and for SnSe in Fig. 2 of Ref. [81], with the softening of the distortion mode obtained by decreasing the temperature.

2.5.2 Static bubble self-energy calculation: improved free energy Hessian calculation

The SSCHA code also allows to compute the free energy Hessian dynamical matrix $\boldsymbol{D}^{(\mathrm{F})}(\boldsymbol{q})$ on any reciprocal space $\boldsymbol{q}$-point, allowing to analyze the structural instabilities incommensurate with the used supercell. After the free energy evaluation and the subsequent free energy Hessian calculation, the real-space $\boldsymbol{D}^{(\mathrm{S})}(\boldsymbol{l}_1, \boldsymbol{l}_2)$ and

$\overset{(3)}{\boldsymbol{D}}_{\text{eq}}(\boldsymbol{l}_1, \boldsymbol{l}_2, \boldsymbol{l}_3)$ are available. Here, $D^{(\text{S})}_{ab}(\boldsymbol{l}_1, \boldsymbol{l}_2)$ is the real space $\boldsymbol{D}^{(\text{S})}$ matrix in which we made explicit the dependence of the lattice vectors $\boldsymbol{l}_1$ and $\boldsymbol{l}_2$ that identify the unit cells in which atom a and b are located, respectively. Using them, the code allows to compute the static bubble $\overset{(\text{B})}{\boldsymbol{\Pi}}(\boldsymbol{q}, 0)$ in any $\boldsymbol{q}$-point, through the formula

$$\begin{aligned}\overset{(\text{B})}{\Pi}_{\mu\nu}(\boldsymbol{q}, 0) = \frac{1}{N_{\boldsymbol{k}}} \sum_{\substack{\boldsymbol{k}_1 \boldsymbol{k}_2 \\ \rho_1 \rho_2}} \sum_{\boldsymbol{G}} \delta_{\boldsymbol{G}, \boldsymbol{q}+\boldsymbol{k}_1+\boldsymbol{k}_2} \mathscr{F}(0, \omega_{\rho_1}(\boldsymbol{k}_1), \omega_{\rho_2}(\boldsymbol{k}_2)) \\ \times \overset{(3)}{D}_{\mu\rho_1\rho_2}(-\boldsymbol{q}, -\boldsymbol{k}_1, -\boldsymbol{k}_2) \overset{(3)}{D}_{\rho_1\rho_2\nu}(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{q}) \ . \end{aligned} \tag{2.66}$$

This equation is Eq. (2.64) written in reciprocal space and SSCHA normal mode components, i.e. in components of the $\boldsymbol{D}^{(\text{S})}(\boldsymbol{q})$'s eigenvector basis. In Eq. (2.66), the $\boldsymbol{k}_i$ sums are performed on a Brillouin zone (BZ) mesh of $N_{\boldsymbol{k}}$ points; $\overset{(3)}{D}_{\mu\rho_1\rho_2}(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{q})$ are the SSCHA normal components of $\overset{(3)}{\boldsymbol{D}}_{\text{eq}}(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{q})$, the Fourier transform of $\overset{(3)}{\boldsymbol{D}}_{\text{eq}}(\boldsymbol{l}_1, \boldsymbol{l}_2, \boldsymbol{l}_3)$ in $(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{q})$; $\omega_\rho(\boldsymbol{k})$ are the frequencies of $\boldsymbol{D}^{(\text{S})}(\boldsymbol{k})$; the function $\mathscr{F}$ is defined by Eq. (2.56); $\boldsymbol{G}$ are reciprocal lattice vectors; and $\delta_{\boldsymbol{G}, \boldsymbol{q}+\boldsymbol{k}_1+\boldsymbol{k}_2}$ preserves crystal momentum. In this formula the $\boldsymbol{q}$ and the $\boldsymbol{k}_i$'s are not confined to the grid commensurate with the supercell used in the SSCHA minimization, as long as the ranges of the real space $\boldsymbol{D}^{(\text{S})}(\boldsymbol{l}_1, \boldsymbol{l}_2)$ and $\overset{(3)}{\boldsymbol{D}}(\boldsymbol{l}_1, \boldsymbol{l}_2, \boldsymbol{l}_3)$ are smaller than the supercell size, so as to be legitimately Fourier interpolated on any reciprocal space points (more about the Fourier interpolation in appendix E). This allows to obtain two results at once. First, the $\boldsymbol{k}$-mesh in Eq. (2.66) can be arbitrarly increased up to convergence, so as to reach the thermodynamic limit in the evaluation of the bubble static self-energy. Second, from $\overset{(\text{B})}{\boldsymbol{\Pi}}(\boldsymbol{q}, 0)$ and $\boldsymbol{D}^{(\text{S})}(\boldsymbol{q})$, through Eq. (2.65b) the code allows to compute the free energy Hessian dynamical matrix $\boldsymbol{D}^{(\text{F})}(\boldsymbol{q})$ (useful to detect and characterize the system instabilities) in $\boldsymbol{q}$ points not necessarily commensurate with the supercell (at variance with what is obtained with the simple Hessian calculation). This can be used, for instance, to study incommensurate second-order displacive phase transitions. In particular, this is the correct way to compute the frequencies $\Omega_\mu(\boldsymbol{q})$ along a reciprocal-space path, where $\Omega^2_\mu(\boldsymbol{q})$ are the eigenvalues of $\boldsymbol{D}^{(\text{F})}(\boldsymbol{q})$. This is, for example, the procedure followed to compute the (static) SSCHA phonon dispersions of NbS_2 shown in Figs. 2, 3 of Ref. [82], and to compute the interpolation-based convergence analysis shown in Fig. 3 of Ref. [85] for $TiSe_2$ monolayer.

2.5.3 Dynamic bubble self-energy calculation: spectral functions, phonon linewidth and shift

The anharmonic generalization of the harmonic dynamical matrix described in the previous sections is the starting point to build a quantum anharmonic ionic dynamical theory. As shown in Refs. [56, 79], in the context of the SSCHA it is possible to formulate an *ansatz* to give the expression of the one-phonon Green function $\boldsymbol{G}(z)$ for the variable $\sqrt{M_a}(R^a - \mathcal{R}^a_{\text{eq}})$. This *ansatz* has been rigorously proven within the Time Dependent Self-Consistent Harmonic Approximation (TD-SCHA)[33, 70]. In

this dynamical theory

$$\boldsymbol{G}^{-1}(z) = z^2\mathbb{1} - \left(\boldsymbol{D}^{(\mathrm{S})} + \boldsymbol{\Pi}(z)\right) , \tag{2.67}$$

where $\boldsymbol{D}^{(\mathrm{S})}$ is the dynamical matrix of the SSCHA effective harmonic Hamiltonian $\mathcal{H}^{(\mathrm{S})}$, and $\boldsymbol{\Pi}(z)$ is the SSCHA self-energy, in general given by

$$\boldsymbol{\Pi}(z) = \overset{(3)}{\boldsymbol{D}}_{\mathrm{eq}} \boldsymbol{:} \boldsymbol{\Lambda}_{\mathrm{eq}}(z) \boldsymbol{:} \left[\mathbb{1} - \overset{(4)}{\boldsymbol{D}}_{\mathrm{eq}} \boldsymbol{:} \boldsymbol{\Lambda}_{\mathrm{eq}}(z)\right]^{-1} \boldsymbol{:} \overset{(3)}{\boldsymbol{D}}_{\mathrm{eq}} , \tag{2.68}$$

and in the bubble approximation by

$$\overset{(\mathrm{B})}{\boldsymbol{\Pi}}(z) = \overset{(3)}{\boldsymbol{D}}_{\mathrm{eq}} \boldsymbol{:} \boldsymbol{\Lambda}_{\mathrm{eq}}(z) \boldsymbol{:} \overset{(3)}{\boldsymbol{D}}_{\mathrm{eq}} . \tag{2.69}$$

In the equations above we use the "eq" subindex to specify that the eigenvalues and eigenfunctions entering the equations are obtained from $\boldsymbol{\Phi}_{\mathrm{eq}}$ with the centroid positions at $\boldsymbol{\mathcal{R}}_{\mathrm{eq}}$.

With the Green function, we obtain the spectral function $\sigma(\Omega) = -2\,\mathrm{Im}\mathrm{Tr}\left[\boldsymbol{G}(\Omega + i0^+)\right]$, which provides the information that can be obtained with inelastic scattering experiments. Taking explicitly into account the lattice translational symmetry (i.e. Fourier transforming the quantities with respect to the lattice vector indices) we can write

$$\sigma(\boldsymbol{q}, \Omega) = -\frac{\Omega}{\pi}\,\mathrm{Im}\mathrm{Tr}\left[\boldsymbol{G}(\boldsymbol{q}, \Omega + i0^+)\right] \tag{2.70}$$

or

$$\begin{aligned}\sigma(\boldsymbol{q}, \Omega) = -\frac{\Omega}{\pi}\mathrm{Im}\mathrm{Tr}\Big[&\left(\Omega + i0^+\right)^2\mathbb{1} + \\ &- \left(\boldsymbol{D}^{(\mathrm{S})}(\boldsymbol{q}) + \boldsymbol{\Pi}(\boldsymbol{q}, \Omega + i0^+)\right)\Big]^{-1} ,\end{aligned} \tag{2.71}$$

where the multiplicative factor $\Omega/2\pi$ has been included to have, for each $\boldsymbol{q}$, a function that integrated on the real axis gives the total number of modes $3n_{\mathrm{a}}$ (n_{a} is the number of atoms in the unit cell, that may be different from the total number of atoms in the supercell N_{a}).

In the so-called "static approximation", we replace the full self-energy $\boldsymbol{\Pi}(z)$ with the static self-energy $\boldsymbol{\Pi}(0)$, where z is blocked at zero. In this case the spectral function is

$$\begin{aligned}\overset{(\mathrm{stat})}{\sigma}(\boldsymbol{q}, \Omega) &= -\frac{\Omega}{\pi}\mathrm{Im}\mathrm{Tr}\Big[\left(\Omega + i0^+\right)^2\mathbb{1} - \left(\boldsymbol{D}^{(\mathrm{S})}(\boldsymbol{q}) + \boldsymbol{\Pi}(\boldsymbol{q}, 0)\right)\Big]^{-1} \\ &= -\frac{\Omega}{\pi}\mathrm{Im}\mathrm{Tr}\left[\left(\Omega + i0^+\right)^2\mathbb{1} - \boldsymbol{D}^{(\mathrm{F})}(\boldsymbol{q})\right]^{-1} ,\end{aligned} \tag{2.72}$$

where in the last line we have used Eq. (2.65). Therefore,

$$\overset{(\mathrm{stat})}{\sigma}(\boldsymbol{q}, \Omega) = \sum_{\mu} \overset{(\mathrm{stat})}{\sigma_{\mu}}(\boldsymbol{q}, \Omega) \tag{2.73}$$

with

$$\overset{(\mathrm{stat})}{\sigma}_{\mu}(\boldsymbol{q},\Omega) = \frac{1}{2}\left[\delta(\Omega-\Omega_{\mu}(\boldsymbol{q})) + \delta(\Omega+\Omega_{\mu}(\boldsymbol{q}))\right] , \tag{2.74}$$

where $\Omega_{\mu}^2(\boldsymbol{q})$ are the eigenvalues of the free energy Hessian matrix $\boldsymbol{D}^{(\mathrm{F})}(\boldsymbol{q})$. In other words, the spectral function in the static limit is formed with delta peaks at the eigenvalues of $\boldsymbol{D}^{(\mathrm{F})}(\boldsymbol{q})$.

In the current version, the SSCHA code computes the full dynamical SSCHA self-energy ($z \neq 0$) only in the bubble approximation with the equation (see Eqs. (2.64), (2.66), and (2.69)))

$$\begin{aligned}\overset{(\mathrm{B})}{\Pi}_{\mu\nu}(\boldsymbol{q},\Omega+i\delta_{\mathrm{se}}) =& \frac{1}{N_c}\sum_{\substack{\boldsymbol{k}_1\boldsymbol{k}_2\\ \rho_1\rho_2}}\sum_{\boldsymbol{G}} \delta_{\boldsymbol{G},\boldsymbol{q}+\boldsymbol{k}_1+\boldsymbol{k}_2} \\ &\times \mathscr{F}(\Omega+i\delta_{\mathrm{se}},\omega_{\rho_1}(\boldsymbol{k}_1),\omega_{\rho_2}(\boldsymbol{k}_2)) \\ &\times \overset{(3)}{D}_{\mu\rho_1\rho_2}(-\boldsymbol{q},-\boldsymbol{k}_1,-\boldsymbol{k}_2)\,\overset{(3)}{D}_{\rho_1\rho_2\nu}(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{q}) ,\end{aligned} \tag{2.75}$$

where the summation $\boldsymbol{k}$-grid can be arbitrarily fine as long as the interpolation of the third-order SSCHA FCs can be performed (as in the static case Eq. (2.66)), and δ_{se} is an arbitrary small, but finite, positive smearing value used to obtain converged results in the computation. In fact, the exact result corresponds to the limiting value obtained with an infinite $\boldsymbol{k}$-grid and a zero δ_{se} smearing. In actual, finite-time calculations, the converged value of the dynamic self-energy is therefore estimated in this way. For a given summation $\boldsymbol{k}$-grid, the corresponding self-energy converged value is estimated by analyzing the result given by Eq. (2.75) for smaller and smaller δ_{se} values (for the used $\boldsymbol{k}$-grid, there will be a minimum value of δ_{se} under which the result shows numerical instability). This analysis is performed with finer and finer summation grids until the converged value in the thermodynamic limit is obtained. In principle, a dedicated convergence study of this kind has to be performed for all the specific observables of interest.

With the dynamical SSCHA bubble self-energy, the code allows to compute the spectral function by the equation (see Eq. (2.71))

$$\begin{aligned}\sigma(\boldsymbol{q},\Omega) = -\frac{\Omega}{\pi}\mathrm{Im}\mathrm{Tr}\Big[&\left(\Omega+i\delta_{\mathrm{id}}\right)^2\mathbb{1} + \\ &- \left(\boldsymbol{D}^{(\mathrm{S})}(\boldsymbol{q}) + \overset{(\mathrm{B})}{\boldsymbol{\Pi}}(\boldsymbol{q},\Omega+i\delta_{\mathrm{se}})\right)\Big]^{-1} ,\end{aligned} \tag{2.76}$$

with δ_{id} another arbitrary small, but finite, positive smearing value. The role of δ_{id} is significant when the imaginary part of the self-energy is small. A prominent example where this happens is when the spectral function is calculated in the static approximation, i.e. when the bubble self-energy is kept fixed at the static value $\overset{(\mathrm{B})}{\Pi}_{\mu\nu}(\boldsymbol{q},0)$ (see Eq. (2.72)). Indeed, in this case the self-energy is real (Hermitian) and the computed spectral function becomes

$$\begin{aligned}\overset{(\mathrm{stat})}{\sigma}(\boldsymbol{q},\Omega) = \sum_{\mu}\frac{1}{2}\mathrm{Im}\Big[&\frac{1}{\pi}\frac{1}{\Omega-\Omega_{\mu}(\boldsymbol{q})+i\delta_{\mathrm{id}}} + \\ &+ \frac{1}{\pi}\frac{1}{\Omega+\Omega_{\mu}(\boldsymbol{q})+i\delta_{\mathrm{id}}}\Big] ,\end{aligned} \tag{2.77}$$

where $\Omega_\mu^2(\boldsymbol{q})$ are the eigenvalues of $\boldsymbol{D}^{(\mathrm{F})}(\boldsymbol{q})$ in the bubble approximation. Therefore, for the numerical computation of the static spectral function, a finite δ_{id} value is necessary to recover the analytical result, Eq. (2.74), but with smeared Dirac delta functions. Actually, this is not just an extreme example, since the code really gives the opportunity to compute the spectral function in the static approximation, replacing in Eq. (2.76) the full bubble self-energy with its static value computed through Eq. (2.66). This can be used to double-check that, as expected from Eqs. (2.74) and (2.77), the obtained spectral function is given by spikes around the frequencies of the Hessian free energy matrix $\boldsymbol{D}^{(\mathrm{F})}$ (computed in the bubble approximation). However, the role played by δ_{id} is not as critical as δ_{se} since it is not typically system-dependent and it does not require a convergence study: in the code its default value is automatically set depending on the spacing of the energy Ω-grid used to compute the spectral function.

Given a $\boldsymbol{q}$, the calculation of the full spectral function $\sigma(\boldsymbol{q},\Omega)$ through Eq. (2.76) turns out to be quite a heavy task due to the inversion of a different $3n_{\mathrm{a}} \times 3n_{\mathrm{a}}$ matrix for each Ω value. The code also allows to employ a much less computational demanding approach by discarding the off-diagonal elements of the computed dynamical self-energy in the SSCHA normal modes components (i.e. the components in the $\boldsymbol{D}^{(\mathrm{S})}(\boldsymbol{q})$'s eigenvector basis). Within this "no mode-mixing" approximation, which usually proves to be extremely good, the SSCHA modes keep their individuality even after the renormalization due to anharmonic effects. Indeed in this case, as in the static approximation, Eq. (2.73), the total spectral function is given by the superposition of individual mode spectral functions:

$$\sigma(\boldsymbol{q},\Omega) = \sum_\mu \sigma_\mu(\boldsymbol{q},\Omega), \tag{2.78}$$

where now the $(\boldsymbol{q},\mu)$-mode spectral function $\sigma_\mu(\boldsymbol{q},\Omega)$ is computed with

$$\begin{aligned}\sigma_\mu(\boldsymbol{q},\Omega) = &\frac{1}{2}\Bigg[\frac{1}{\pi}\frac{-\mathrm{Im}\mathcal{Z}_\mu(\boldsymbol{q},\Omega)}{[\Omega - \mathrm{Re}\mathcal{Z}_\mu(\boldsymbol{q},\Omega)]^2 + [\mathrm{Im}\mathcal{Z}_\mu(\boldsymbol{q},\Omega)]^2} + \\ &+ \frac{1}{\pi}\frac{\mathrm{Im}\mathcal{Z}_\mu(\boldsymbol{q},\Omega)}{[\Omega + \mathrm{Re}\mathcal{Z}_\mu(\boldsymbol{q},\Omega)]^2 + [\mathrm{Im}\mathcal{Z}_\mu(\boldsymbol{q},\Omega)]^2}\Bigg]\end{aligned} \tag{2.79}$$

and

$$\mathcal{Z}_\mu(\boldsymbol{q},\Omega) = \sqrt{\omega_\mu^2(\boldsymbol{q}) + \Pi_{\mu\mu}(\boldsymbol{q},\Omega + i\delta_{\mathrm{se}})}\,. \tag{2.80}$$

Therefore, computing the spectral function in the "no mode-mixing" approximation, by measuring the deviation of $\sigma_\mu(\boldsymbol{q},\Omega)$ from a Dirac delta function around $\Omega_\mu(\boldsymbol{q})$, it is possible to asses the impact that anharmonicity has on the different SSCHA modes $(\boldsymbol{q},\mu)$, separately.

The form of the $(\boldsymbol{q},\mu)$-mode spectral function $\sigma_\mu(\boldsymbol{q},\Omega)$ in Eq. (2.79) resembles a Lorentzian, but with frequency-dependent center and width, meaning that the actual form of the spectral function $\sigma(\boldsymbol{q},\Omega)$ can be quite different from the superposition of true Lorentzian functions. However, in some cases the $\sigma_\mu(\boldsymbol{q},\Omega)$ can be expressed with good approximation as a true Lorentzian with a certain half width at half

maximum (HWHM) $\Gamma_\mu(\boldsymbol{q})$ and center $\Omega_\mu(\boldsymbol{q})$,

$$\sigma_\mu(\boldsymbol{q},\Omega) = \frac{1}{2}\left[\frac{1}{\pi}\frac{\Gamma_\mu(\boldsymbol{q})}{[\Omega-\Omega_\mu(\boldsymbol{q})]^2+[\Gamma_\mu(\boldsymbol{q})]^2}\right.$$

$$\left. + \frac{1}{\pi}\frac{\Gamma_\mu(\boldsymbol{q})}{[\Omega+\Omega_\mu(\boldsymbol{q})]^2+[\Gamma_\mu(\boldsymbol{q})]^2}\right], \tag{2.81}$$

meaning that the quasiparticle picture is still valid, even after the inclusion of anharmonicity, with the $(\mu,\boldsymbol{q})$ quasiparticle having frequency (energy) $\Omega_\mu(\boldsymbol{q})$ and lifetime $\tau_\mu(\boldsymbol{q}) = 1/(2\Gamma_\mu(\boldsymbol{q}))$. The difference between the renormalized and the "bare" SSCHA frequency, $\Delta_\mu(\boldsymbol{q}) = \Omega_\mu(\boldsymbol{q}) - \omega_\mu(\boldsymbol{q})$, is called the frequency shift of the $(\mu,\boldsymbol{q})$ mode.

The SSCHA code offers several tools to perform such a "Lorentzian analysis". In general, the best Lorentzian approximation is obtained with

$$\Omega_\mu(\boldsymbol{q}) = \mathrm{Re}\mathcal{Z}_\mu(\boldsymbol{q},\Omega_\mu(\boldsymbol{q})) \tag{2.82}$$

$$\Gamma_\mu(\boldsymbol{q}) = -\mathrm{Im}\mathcal{Z}_\mu(\boldsymbol{q},\Omega_\mu(\boldsymbol{q})). \tag{2.83}$$

Once the dynamical self-energy and the $\mathcal{Z}_\mu(\boldsymbol{q},\Omega)$ are computed, the SSCHA code allows to compute the single-mode spectral functions in the Lorentzian approximation, estimating the frequency $\Omega_\mu(\boldsymbol{q})$ and HWHMs $\Gamma_\mu(\boldsymbol{q})$ in different ways. One, optional, possibility is to solve self-consistently Eq. (2.82) to estimate $\Omega_\mu(\boldsymbol{q})$, and then $\Gamma_\mu(\boldsymbol{q})$, through Eq. (2.83). However, by default, the "one-shot" approximation is employed with

$$\overset{(\mathrm{os})}{\Omega}_\mu(\boldsymbol{q}) = \mathrm{Re}\mathcal{Z}_\mu(\boldsymbol{q},\omega_\mu(\boldsymbol{q})) \tag{2.84}$$

$$\overset{(\mathrm{os})}{\Gamma}_\mu(\boldsymbol{q}) = -\mathrm{Im}\mathcal{Z}_\mu(\boldsymbol{q},\omega_\mu(\boldsymbol{q}))\,. \tag{2.85}$$

If the SSCHA self-energy $\boldsymbol{\Pi}$ is a (small) perturbation on the SSCHA free propagator (not meaning that we are in a perturbative regime with respect to the harmonic approximation), then perturbation theory can be employed to evaluate the spectral function. If we keep the first order in the self-consistent equations Eq. (2.85), we get:

$$\overset{(\mathrm{pert})}{\Omega}_\mu(\boldsymbol{q}) = \frac{1}{2\omega_\mu(\boldsymbol{q})}\mathrm{Re}\Pi_{\mu\mu}(\boldsymbol{q},\omega_\mu(\boldsymbol{q})) \tag{2.86}$$

$$\overset{(\mathrm{pert})}{\Gamma}_\mu(\boldsymbol{q}) = -\frac{1}{2\omega_\mu(\boldsymbol{q})}\mathrm{Im}\Pi_{\mu\mu}(\boldsymbol{q},\omega_\mu(\boldsymbol{q})+i\delta_{\mathrm{se}})\,. \tag{2.87}$$

This perturbative approach is also employed by the SSCHA code to evaluate the quasiparticles' energies and lifetimes. Examples of spectral function calculations done with Eq. (2.76), Eqs. (2.78)- (2.80) and Eq. (2.81) can be found in Fig. 4 of Ref. [79]. In Fig. 5 of the same reference, the anaharmonic phonon frequencies and linewidths along a path, computed using the Lorenztian approximation through

Eqs. (2.82) and (2.83), are shown. The spectral function computed with Eq. (2.79) along a path is shown with a colorpolot in Fig. 3 of Ref. [80] for PbTe, and in Fig. 4 of Ref. [81] for SnSe.

In conclusion, with the SSCHA code we can calculate three frequencies for a mode $(\boldsymbol{q}, \mu)$: $\omega_\mu(\boldsymbol{q}), \Omega_\mu(\boldsymbol{q})$, and $\mathbb{\Omega}_\mu(\boldsymbol{q})$, which are the frequency of the SSCHA auxiliary boson, the frequency coming from the SSCHA free energy Hessian (i.e. from the static approximation), and the frequency of the SSCHA quasiparticle in the Lorentzian approximation. Only the last one is a true physical quantity as it can be measured in experiments. However, the static $\Omega_\mu(\boldsymbol{q})$ is also a physical meaningful quantity, as its zero value corresponds to a structural instability driving a second-order phase transition along the pattern characterized by the mode $(\boldsymbol{q}, \mu)$. The SSCHA provides a specific physical meaning of each of these frequencies, in contrast to other approaches used to estimate anharmonic phonons, where no distinction is usually done.

2.6 The Python code

During the last years, the code to implement all the SSCHA machinery has undergone substantial modification and improvements with respect to its original version. In the current version, two different Python libraries are provided with the SSCHA code: CellConstructor and Python-sscha. The latter is the library that performs the SSCHA minimization itself, while the former is a library that deals with the dynamical matrix, the crystal structure, the symmetrization, and performs the calculation of phonon spectral functions and linewidths as a post-processing tool.

After the generation of the ensemble the code requires the computation of atomic forces and energies and it is possible to implement interfaces with the most common *ab initio* codes like Quantum ESPRESSO[48, 86], VASP[87], SIESTA[88], CP2K[89], and many more. Force-field codes like LAMMPS[90] may also be used. The use of first-principles calculation of such energies and forces poses a serious limitation on the number of configuration to use because of the high computational cost.

My biggest contributions to the method development are related to the code optimization. I report in this section, some of the improvement I applied to allow a smart application of the SSCHA algorithm to systems with many atoms. In order to obtain accurate results with less configurations, I checked if one can use some shortcuts. For instance, for all the results in this work, one needs to relax crystal structures at a given pressure. Consequently, it was crucial to try to reduce the pressure fluctuations with the generated configurations. The Taylor expansion of the pressure for small displacements u gives a hint for a possible improvement of the code. We can generate couples of symmetric $\{-u, u\}$ configurations and consider the fluctuations of the even and odd part of stress tensor, where, for a generic $f(x)$ function, $f_{even}(x) = \frac{1}{2}(f(x) + f(-x))$ and $f_{odd}(x) = \frac{1}{2}(f(x) - f(-x))$. Now, the odd part has a linear dependence on the displacement (unless it vanishes for some reason) while the even part is quadratic in the displacement. So, in principle, for small displacement generation, I expect smaller fluctuations for the even term.

I report in Fig. 2.3 the even and odd contributions of the pressure extracted from an ensemble of 2000 symmetric configurations. The fluctuations around the

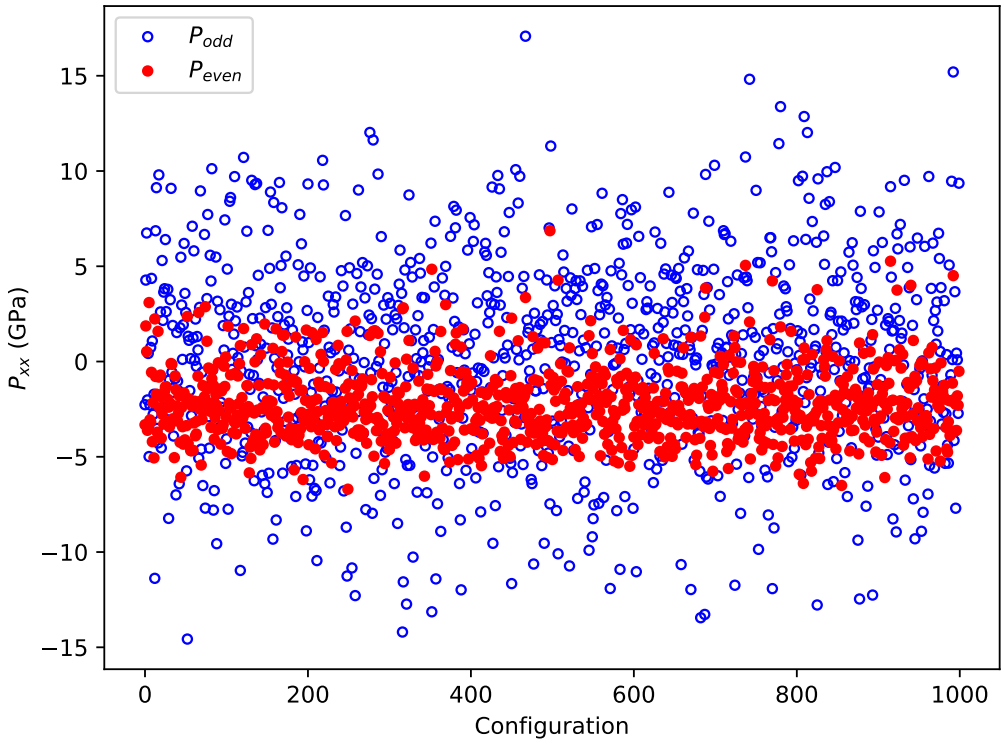

Figure 2.3. P_{xx} component of the stress tensor, computed within the SSCHA as in Eq. (2.19). The red filled points are the even part of the stress tensor $\frac{P_{xx}(u)+P_{xx}(-u)}{2}$ for each $\{u, -u\}$ configuration, while the blue empty ones indicate the odd part $\frac{P_{xx}(u)-P_{xx}(-u)}{2}$.

mean value of the odd contribution are three times bigger than that of the even one, as expected. This proves that the use of the even term of the pressure (that by construction has the same average value of the stress tensor computed with the original ensemble) grants a better accuracy with the same computational effort.

The SSCHA algorithm takes includes quantum anharmonicity in a nonperturbative way. In principle, in the self-energy term of the dynamical Green function in Eq. (2.69), all the phonon modes interact together. Sometimes, it is useful to isolate the interaction of given phonon bands in Eq. (2.56) to understand the origin of specific features in the spectral function as the combination modes (see Sec. 3.3.4). To achieve these results, I modified the original routines to give the possibility to consider the interaction of few selected phonon modes alone.

2.6.1 Parallelization

The nature of the algorithm makes it very simple to exploit massive parallelization strategies available in high performance computing (HPC) facilities. In particular, the most expensive part of the code is the calculation of Born-Oppenheimer energies, forces, and stress tensors of the generated ionic configurations in each population (the red shaded cell in the code flowchart in Figure 2.2). Each of this calculation is independent from the others, so they can be trivially run in parallel on different

computing nodes. This is a huge advantage with respect to other methods based on AIMD or PIMD, which mimic a time evolution of the system and thus require to calculate atomic forces on one configuration after the other.

Usually, all the other steps of the code are computationally very cheap compared to the energy and force calculation. This is not true when they are computed with force fields (like for the ice in this work) instead then from first-principles. In this case, when the supercell used contain many atoms, as it is for ices Ih/XI in Chapter 3, both the generation of the ensemble and the minimization itself can be challenging and their computational cost can be comparable or exceeding that necessary for the energies and forces computation. The SSCHA minimization cannot exploit so well the possibilities offered by parallelization, since each step of the main cycle depends on the previous one. Most of the computations executed in the cycle are linear algebra calculations but some of them can be speeded up with an explicit Fortran implementation. So, to study the convergence of the results with the supercell size, I completely rewrote some subroutines used for the generation of the ensembles and for the minimization. For the latter, in particular, I worked on the updating of the weights at each step that could be heavily slimmed down and some functions could be written in Fortran. Moreover, a massive parallelization of these subroutines gave me a huge speedup. Just to give a flavour of the improvement, the time elapsed to perform each minimization step in a structure with 768 atoms passed from $t \simeq 2500s$ to $t \simeq 200s$, with a speedup of an order of magnitude. I obtained even better results for the generation of the ensemble, that for structure with many atoms were stuck in very slow loops that could avoided by rewriting the code in a more efficient way, gaining almost two order of magnitudes in the elapsed time for the complete generation.

2.7 Time dependent SSCHA

I will use the dynamical *ansatz* postulated in Sec. 2.5.3 to define the one-phonon Green function in Chapter 3 to study the low-pressure ices I_h/XI. In a following work, Monacelli et al. [33] rigorously extended the SSCHA theory to account for the quantum time evolution of nuclear vibrations, proving the exactness of the dynamical *ansatz*. In the following, I will use the TD-SSCHA extension and in particular, the Lanczos algorithm (see below) to study high-pressure ice in Chapter 4. In this section, I report the main results and equation of the TD-SSCHA to make clear what is done in Chapter 4.

The idea is to describe how the nuclear physical properties change in time when the system in equilibrium is perturbed with an external probe. So, the system is set in equilibrium for $t < t_0$ and at $t = t_0$ it interacts with an external time-dependent perturbation $V^{ext}(\mathbf{R}, t)$. Then, the aim is to describe the evolution of the expectation value of a generic nuclear observable $\hat{\mathcal{A}}$ for $t > t_0$:

$$\mathcal{A}(t) = \langle \psi(t) | \hat{\mathcal{A}} | \psi(t) \rangle \,, \tag{2.88}$$

where, the time evolution of the wave function is governed by the time-dependent Schrödinger equation, through the time-dependent Hamiltonian

$$\hat{H}_{td}(t) = \hat{H} + V^{ext}(\hat{\mathbf{R}}, t) \tag{2.89}$$

At finite temperature, the wave function has to be replaced with the density matrix $\hat{\rho}$ that can describe also mixture of states. The equilibrium density matrix acquires time dependence for $t > t_0$ when the external perturbation is switched on

$$\hat{\rho}(t) = \sum_i p_i \left|\psi(t)\right\rangle \left\langle\psi(t)\right|, \tag{2.90}$$

where

$$p_i = \frac{e^{-\beta E_i}}{Z}, Z = \sum_i e^{-\beta E_i}, \beta = \frac{1}{k_b T} \tag{2.91}$$

In isolated systems, p_i does not depend on time, and the density matrix satisfy the Liouville-Von Neumann equation

$$i\hbar \frac{d}{dt}\hat{\rho}(t) = \hat{H}_{td}(t)\hat{\rho}(t) - \hat{\rho}(t)\hat{H}_{td}(t) \tag{2.92}$$

In this case, the average of the observable can be written as

$$\mathcal{A}(t) = \left\langle \hat{\mathcal{A}} \right\rangle_{\hat{\rho}(t)} = \mathrm{Tr}\left[\hat{\rho}(t)\hat{\mathcal{A}}\right] \tag{2.93}$$

This approach can be very useful to study linear response as Infrared absorption in Sec. 4.5, when the external perturbation does not change the status of the system, neither breaking chemical bonds, nor triggering a macroscopic rearrangement of the atoms, nor heating the sample. The small external time-dependent perturbation can be split in a term that depends only on the nuclear positions $\hat{B} = \mathcal{B}(\hat{\mathbf{R}})$ and the time envelop $\mathcal{V}(t)$

$$V^{ext}(\mathbf{R}, t) = V^{(1)}(\mathbf{R}, t) = \mathcal{B}(\hat{\mathbf{R}})\mathcal{V}(t) \tag{2.94}$$

Now, the time dependence of an observable can be expressed as

$$\mathcal{A}(t) = \mathcal{A}(t_0) + \int_{-\infty}^{+\infty} \chi_{AB}(t - t')\mathcal{V}(t')dt' \tag{2.95}$$

and the response function is directly related to time correlation functions through the Kubo formula

$$\chi_{AB}(t) = -\frac{i}{\hbar}\left\langle \left[e^{\frac{i}{\hbar}\hat{H}t}\mathcal{A}e^{-\frac{i}{\hbar}\hat{H}t}, \hat{\mathcal{B}}\right]\right\rangle_{\hat{\rho}^{(0)}} \theta(t) \tag{2.96}$$

In Eq. (2.96), $e^{-\frac{i}{\hbar}\hat{H}t}$ is the time evolution of an operator in absence of the external perturbation and $\theta(t)$ is the Heaviside function. In the frequency domain, the average of the observable $\mathcal{A}$ can be readily written as the product of the response function and the time envelop of the perturbation.

The SSCHA theory depicted in the first part of this Chapter is a static theory that cannot describe physical phonons. In the most general case, the nuclear time-dependent evolution cannot be described by the dynamics of pure states (suitable only at T=0 K), but mixed states have to be used. Shortly, one has to substitute

the time-dependent Schrödinger equation with the Liouville-Von Neumann equation of Eq. (2.92). As for the static SSCHA theory, the wave function is constrained to the most general Gaussian, but in this case, time dependence is introduced:

$$\begin{aligned}\langle \mathbf{R}'|\hat{\rho}(t)|\mathbf{R}\rangle =& \mathcal{N}(t) exp(i\sum_a Q_a(t)(R'_a - R_a) \\ &- \sum_{ab}\left[\frac{\Theta_{ab}(t)}{4} - iC_{ab}(t)\right](R_a - \mathcal{R}_a(t))(R_b - \mathcal{R}_b(t)) \\ &- \sum_{ab}\left[\frac{\Theta_{ab}(t)}{4} + iC_{ab}(t)\right](R'_a - \mathcal{R}_a(t))(R'_b - \mathcal{R}_b(t)) \\ &+ \sum_{ab} A_{ab}(t)(R_a - \mathcal{R}_a(t))(R'_b - \mathcal{R}_b(t)))\end{aligned} \tag{2.97}$$

Here, the $\mathbf{\Theta}$ matrix encodes pure quantum fluctuations, while the complex Hermitian $A(t)$ matrix, absent in pure states, has a real part ReA related to thermal fluctuations and an imaginary part ImA that plays the role of momentum of thermal fluctuations. The Q_i variables represent the momentum of the i-th atom. $\mathbf{C}$ is a quadratic phase that represents the chirp along the quantum fluctuations.

The equation of motion can be derived by minimizing the Dirac action

$$A = \frac{1}{t - t_0}\int_{t_0}^{t} \langle\psi(t')|\hat{H}_{td}(t') - i\hbar\frac{d}{dt'}|\psi(t')\rangle\, dt' \tag{2.98}$$

I show in Eq. (2.99) the final equation of motions, all the intermediate passages can be found in Ref. [33].

$$\begin{aligned}\frac{d\tilde{\mathcal{R}}}{dt} &= \hbar\tilde{\mathbf{Q}}, \frac{dQ_a}{dt} = \frac{\left\langle f_a^{(tot)}\right\rangle_{\rho(t)}}{\hbar} \\ \frac{d\tilde{\mathbf{\Upsilon}}}{dt} &= \hbar[\tilde{\mathbf{\Upsilon}}(2\tilde{\mathbf{C}} + \mathrm{Im}\tilde{\mathbf{A}}) + (2\tilde{\mathbf{C}} - \mathrm{Im}\tilde{\mathbf{A}})\tilde{\mathbf{\Upsilon}}] \\ \frac{d\mathrm{Re}\tilde{\mathbf{A}}}{dt} &= \frac{\hbar}{2}(4\tilde{\mathbf{C}}\mathrm{Re}\tilde{\mathbf{A}} + 4\mathrm{Re}\tilde{\mathbf{A}}\tilde{\mathbf{C}} - \tilde{\mathbf{\Theta}}\mathrm{Im}\tilde{\mathbf{A}} + \mathrm{Im}\tilde{\mathbf{A}}\tilde{\mathbf{\Theta}}) \\ \frac{d\mathrm{Im}\tilde{\mathbf{A}}}{dt} &= \frac{\hbar}{2}(4\tilde{\mathbf{C}}\mathrm{Im}\tilde{\mathbf{A}} + 4\mathrm{Im}\tilde{\mathbf{A}}\tilde{\mathbf{C}} + \tilde{\mathbf{\Theta}}\mathrm{Re}\tilde{\mathbf{A}} - \mathrm{Re}\tilde{\mathbf{A}}\tilde{\mathbf{\Theta}}) \\ \frac{d\tilde{\mathbf{C}}}{dt} &= \frac{1}{2\hbar}\left\langle\frac{\partial^2 V^{(tot)}}{\partial\tilde{\mathbf{R}}\partial\tilde{\mathbf{R}}}\right\rangle_{\rho(t)} + \frac{\hbar}{2}\left[\tilde{\mathbf{C}}^2 - \frac{1}{4}\tilde{\mathbf{\Theta}}^2 + \mathrm{Re}(\tilde{\mathbf{A}}\tilde{\mathbf{A}}^\dagger)\right] \\ \frac{d\mathcal{N}}{dt} &= 2\hbar\mathcal{N}\,\mathrm{Tr}\,\tilde{\mathbf{C}}\end{aligned} \tag{2.99}$$

where variables with a tilde are scaled by the square root of the masses.

2.7.1 Linear response theory

Experimental data are collected by probing the response of the system to an external time-dependent perturbation $V^{(1)}(\mathbf{R}, t)$. The perturbation can be either electromagnetic radiation or particles, like neutrons or electrons. When it is small with

respect to the nuclear potential $V(\mathbf{R})$, the density matrix can be separated into an equilibrium solution $\hat{\rho}^{(0)}$ and a small perturbation $\hat{\rho}^{(1)}(t)$:

$$\hat{\rho}(t) = \hat{\rho}^{(0)} + \hat{\rho}^{(1)}(t). \tag{2.100}$$

A linear expansion of the TD-SSCHA equation around equilibrium brings to Eq. (2.101) for the perturbed density matrix

$$i\hbar \frac{d}{dt}\hat{\rho}^{(1)}(t) = L_{sc}\hat{\rho}^{(1)}(t) + \left[\hat{V}_{sc}^{(1)}(t), \hat{\rho}^{(0)}\right], \tag{2.101}$$

where L_{sc} is a super-operator that describes the unperturbed anharmonic evolution according to the self-consistent Hamiltonian and $\hat{V}_{sc}^{(1)}$ is the interaction with the external potential:

$$L_{sc}\hat{\rho^{(1)}}(t) = \left[\hat{\mathcal{H}}[\rho^{(0)}], \hat{\rho}^{(1)}(t)\right] + \left[\hat{\mathcal{H}}[\rho^{(1)}(t)], \hat{\rho}^{(0)}\right] \tag{2.102}$$

$$\hat{V}_{sc}^{(1)} = \frac{1}{2}\sum_{ab}(\hat{R}_a - \mathcal{R}_a^{(0)}) \left\langle \frac{d^2V^{(1)}(t)}{dR_a dR_b} \right\rangle_{\rho^{(0)}} (\hat{R}_b - \mathcal{R}_b^{(0)}) + \sum_a \left\langle \frac{\partial V^{(1)}(t)}{dR_a} \right\rangle_{\rho^{(0)}} (\hat{R}_a - \mathcal{R}_a^{(0)}) \tag{2.103}$$

Eq. (2.101) can be solved passing to Fourier space. In the following, I indicate with $^{(0)}$ all the equilibrium quantities and with $^{(1)}$ a small perturbation around the SSCHA equilibrium. $\hat{\rho^{(1)}}(t)$ depends on the perturbation of all the variables appearing in the density matrix definition in Eq. (2.97). We can get rid of some of them by taking the time derivative of Eq. (2.101) leading to the general expression for the linear response of the system to any external perturbation on nuclei $V^{(1)}(\mathbf{R}, \omega)$

$$\begin{pmatrix} \tilde{\boldsymbol{\Upsilon}}^{(1)}(\omega) \\ \mathrm{Re}\tilde{\mathbf{A}}^{(1)}(\omega) \\ \tilde{\mathcal{R}}^{(1)}(\omega) \end{pmatrix} = \mathbf{G}(\omega) \begin{pmatrix} f_{\Upsilon}^{(1)}(\omega) \\ f_{\mathrm{Re}A}^{(1)}(\omega) \\ f_{\mathcal{R}}^{(1)}(\omega) \end{pmatrix} \tag{2.104}$$

In this way, we can introduce the Green function $\mathbf{G}(\omega)$

$$\mathbf{G}(\omega) = -(\omega^2 - \mathcal{L})^{-1}. \tag{2.105}$$

The $\mathcal{L}$ kernel and $\mathbf{f}$ vector are the free (anharmonic) evolution and the coupling of the phonons with the bare perturbation, respectively, in the space of the parameter of the Gaussian and their full expression can be found in Ref. [33].

Finally, a first-order expansion of the average of the observable $\hat{\mathcal{A}}$ around the equilibrium solution can be written as

$$\left\langle \hat{\mathcal{A}} \right\rangle_{\rho^{(1)}(\omega)} = \mathbf{p} \cdot \begin{pmatrix} \tilde{\boldsymbol{\Upsilon}}^{(1)}(\omega) \\ \mathrm{Re}\tilde{\mathbf{A}}^{(1)}(\omega) \\ \tilde{\mathcal{R}}^{(1)}(\omega) \end{pmatrix}, \tag{2.106}$$

where

$$\mathbf{p} = \begin{pmatrix} \frac{\partial < \hat{\mathcal{A}} >_{\rho(\omega)}}{\partial \tilde{\boldsymbol{\Upsilon}}} & \frac{\partial < \hat{\mathcal{A}} >_{\rho(\omega)}}{\partial \mathrm{Re}\tilde{\mathbf{A}}} & \frac{\partial < \hat{\mathcal{A}} >_{\rho(\omega)}}{\partial \tilde{\mathcal{R}}} \end{pmatrix} \tag{2.107}$$

We can also separate the frequency dependence of the $\mathbf{f}$ vector as $\mathbf{f}^{(1)}(\omega) = \mathbf{f}^{(1)}\mathcal{V}(\omega)$ [33], and define the $\mathbf{q}$ vector:

$$\mathbf{q} = \begin{pmatrix} \mathbf{f}_{\Upsilon}^{(1)} \\ \mathbf{f}_{\mathrm{Re}A}^{(1)} \\ \mathbf{f}_{\mathcal{R}}^{(1)} \end{pmatrix} \tag{2.108}$$

So that, at the end, the average of an operator can be written as in Eq. (2.109)

$$\left\langle \hat{\mathcal{A}} \right\rangle_{\rho^{(1)}(\omega)} = \mathbf{pG}(\omega)\mathbf{q}\mathcal{V}(\omega) \tag{2.109}$$

and the response function

$$\chi_{AB}(\omega) = \mathbf{pG}(\omega)\mathbf{q} \tag{2.110}$$

2.7.2 IR and Raman response

Infrared absorption is related to the dipole-dipole correlation function

$$\hat{\mathcal{A}} = M_\alpha(\hat{R}) \quad \hat{\mathcal{B}} = M_{\alpha'}(\hat{R}).$$

where $M_\alpha(\hat{R})$ is the α Cartesian component of the net dipole moment when the ions are displaced in the R position

$\mathbf{p}$ and $\mathbf{q}$ vector are related to first and second-order derivative of $\hat{\mathcal{A}}$ and $\hat{\mathcal{B}}$ operators, respectively, with respect to nuclear displacements [33].

Since the dipole can be written as $M_\alpha(\mathbf{R}) = |e| \sum_a Z_{\alpha a}(\mathbf{R}) u_a$, the averages of the dipole derivative are readily computed:

$$\left\langle \frac{\partial M_\alpha(\mathbf{R})}{\partial R_a} \right\rangle_{\rho^{(0)}} = \langle Z_{\alpha a}(\mathbf{R}) \rangle_{\rho^{(0)}} \tag{2.111}$$

The effective charges Z indicate how the atoms are displaced when an electric field is applied and they are related to the second derivative of the electronic energy:

$$Z_{\alpha a} = \frac{\partial^2 E_{el}}{\partial \mathcal{E}_\alpha \partial R_a} \tag{2.112}$$

The second order derivative can be easily written by integrating by parts:

$$\left\langle \frac{\partial^2 M_\alpha(\mathbf{R})}{\partial R_a \partial R_b} \right\rangle_{\rho^{(0)}} = \sum_c \Upsilon_{ac} \langle (R_c - \mathcal{R}_c) Z_{\alpha b}(\mathbf{R}) \rangle_{\rho^{(0)}} \tag{2.113}$$

In the case of linear coupling to the probe, the second order derivative in Eq. (2.113) vanishes, and if the effective charges do not depend on nuclear positions

$$\langle Z_{\alpha a}(\mathbf{R}) \rangle_{\rho^{(0)}} = Z_{\alpha a}(\mathbf{R})$$

and it can be shown that only the $\mathbf{R}$ block of the $\mathbf{p}$ and $\mathbf{q}$ vectors are different from zero [33]. In this case, the response function is

$$\chi_{M_\alpha M_{\alpha'}}(\omega) = \sum_{ab} \frac{Z_{\alpha a} Z_{\alpha' b}}{\sqrt{M_a M_b}} G_{ab}(\omega) \tag{2.114}$$

If the external perturbations is nonlinear in the ionic displacements, a contribution from the Υ and ReA blocks of the $\mathbf{p}$ and $\mathbf{q}$ vectors arises and in some systems it can be important [23]. The two-phonon propagation gives a non-zero response also in purely harmonic crystals. For example, the harmonic two-phonon IR response is

$$\chi^{(2ph)}_{M_\alpha M_{\alpha'}}(\omega) = \sum_{abcd\mu\nu} \frac{e^a_\mu e^b_\nu e^c_\mu e^d_\nu}{\sqrt{M_a M_b M_c M_d}} \frac{dZ_{\alpha a}}{dR_b} \chi_{\mu\nu\mu\nu}(\omega) \frac{dZ_{\alpha' c}}{dR_d} \tag{2.115}$$

To obtain the interacting response, it is sufficient to replace the harmonic two-phonon Green function $\chi_{\mu\nu\mu\nu}$ with the interacting one.

The same procedure applies to Raman scattering. Here, the $\hat{\mathcal{A}}$ and $\hat{\mathcal{B}}$ observable are the polarizability $\alpha_{\alpha\beta}(\hat{\mathbf{R}})$ of the sample along α (incoming) and β (outcoming) directions of the light polarization, when ions are displaced.

$$\hat{\mathcal{A}} = \alpha_{\alpha\beta}(\hat{R}) \quad \hat{\mathcal{B}} = \alpha_{\alpha'\beta'}(\hat{R}),$$

The polarizability of a molecule is its ability to respond to an electric field $\mathcal{E}$ and acquire a dipole moment, thus, it can be related to the electronic energy as

$$\alpha_{\alpha\beta} = \frac{\partial^2 E_{el}}{\partial \mathcal{E}_\alpha \partial \mathcal{E}_\beta}$$

where α and β are Cartesian indices. By assuming a linear dependence of the polarizability on the atomic displacement u, it can be recast as in Eq. (2.116)

$$\alpha_{\alpha\beta}(t) = \sum_a A_{\alpha\beta a} u_a(t) \tag{2.116}$$

where the Raman tensor $A_{\alpha\beta a}$ describes the coupling between the two electric fields and the sample

$$A_{\alpha\beta a} = \frac{\partial^3 E_{el}}{\partial \mathcal{E}_\alpha \partial \mathcal{E}_\beta \partial R_a} \tag{2.117}$$

The $\mathbf{p}$ and $\mathbf{q}$ vectors are derived in the same way as for Infrared absorption. It is sufficient to compute the first and second-order derivative of the polarizability as

$$\left\langle \frac{\partial^2 \alpha_{\alpha\beta}(\mathbf{R}}{\partial R_a} \right\rangle_{\rho^{(0)}} = \langle A_{\alpha\beta a} \rangle_{\rho^{(0)}} \tag{2.118}$$

$$\left\langle \frac{\partial^2 \alpha_{\alpha\beta}(\mathbf{R}}{\partial R_a \partial R_b} \right\rangle_{\rho^{(0)}} = \sum_c \Upsilon_{ac} \langle (R_c - \mathcal{R}_c) A_{\alpha\beta a} \rangle_{\rho^{(0)}} \tag{2.119}$$

Again, in the case of linear coupling to the probe and no dependence of the Raman tensor on the ionic displacement, the response function can be easily written

$$\chi_{\alpha_{\alpha\beta}\alpha_{\alpha'\beta'}} = \sum_{ab} \frac{A_{\alpha\beta a} A_{\alpha'\beta' b}}{\sqrt{M_a M_b}} G_{ab}(\omega) \tag{2.120}$$

The two-phonon form is analogous to the Infrared one in Eq. (2.115).

2.7.3 Lanczos algorithm

The computational cost of the one-phonon Green function as in Sec. 2.5.3 diverges quickly with the number of atoms and so the application of the full dynamical Green function has been performed in realistic systems under the assumption that the fourth-order term $\overset{(4)}{\boldsymbol{D}}$ is negligibly small. Moreover, the one phonon Green function does not provide the most general response to the experimental probe, but it is limited to probes interacting linearly with atomic displacements. In Ref. [33], an efficient algorithm that allows to compute the elements of the response function to any general external perturbation, computationally achievable in systems of hundreds of atoms in the full anharmonic regime, even with $\overset{(4)}{\boldsymbol{D}} \neq 0$, is derived.

Shortly, the biconjugate Lanczos algorithm computes in one shot the response function of Eq. (2.110) for any value of ω. This is done by building, with an iterative procedure, a basis-changing Q matrix so that

$$Q^{-1}\mathcal{L}Q = \mathcal{T} \tag{2.121}$$

is tridiagonal, with $\alpha_k, \beta_k, \gamma_k$ coefficients.

$$\mathcal{T} = \begin{pmatrix} \alpha_1 & \gamma_1 & & \cdots & 0 \\ \beta_1 & \alpha_2 & \ddots & & 0 \\ & \ddots & \ddots & \ddots & \\ \vdots & & \ddots & \ddots & \gamma_{n-1} \\ 0 & \cdots & & \beta_{n-1} & \alpha_n \end{pmatrix} \tag{2.122}$$

the basis in which the $\mathcal{L}$ is tridiagonal is represented by the column of the $\mathcal{Q}$ matrix:

$$\mathcal{Q} = \begin{pmatrix} \mathbf{q_1} & \cdots & \mathbf{q_n} \end{pmatrix} \tag{2.123}$$

and the rows of the inverse matrix define the conjugate vectors

$$\mathcal{Q}^{-1} = \begin{pmatrix} \mathbf{p_1} \\ \vdots \\ \mathbf{p_n} \end{pmatrix} \tag{2.124}$$

and both the coefficients of the tridiagonal matrix and the $\mathbf{q_k}$ and $\mathbf{p_k}$ vectors are found through the iterative biconjugate Lanczos algorithm; the algorithm formally ends when $\mathbf{q_k}$ and $\mathbf{p_k}$ are linear combinations of the previous vectors or when $\mathbf{p_k} \cdot \mathbf{q_k} = 0$. See Ref. [33] for further details. In the end, one can show that the response function can be written as

$$\chi_{AB}(\omega) = -(\mathbf{p} \cdot \mathbf{q})[(\mathcal{T} + \omega^2)^{-1}]_{11} \tag{2.125}$$

It depends only on the first element of the inverse matrix in the tridiagonal basis that can be written as a continued fraction.

$$\chi_{AB}(\omega) = \cfrac{-(\mathbf{p}\cdot\mathbf{q})}{\alpha_1 + \omega^2 - \cfrac{\gamma_1\beta 1}{\alpha 2 + \omega^2 - \cfrac{\gamma_2\beta_2}{\ddots}}} \tag{2.126}$$

This method is limited to response functions where the perturbation and the response are not orthogonal $\mathbf{p}\cdot\mathbf{q} \neq 0$

2.8 Comparison with other approximations

We work within the Born-Oppenheimer (BO) approximation [91] to separate electronic and nuclear degrees of freedom. In Sec. 2.2 and 2.7, I described the SSCHA approximation and its time-dependent extension, used to solve the nuclear BO Hamiltonian. It is a very sophisticated technique that allows to include nuclear quantum anharmonic effect in a non-perturbative way. In literature, there are a great number of theoretical approximations to treat the nuclear problem. The most used approach, which overcomes the limitations of the harmonic theory and gives very good results, is the so-called quasi-harmonic approximation (QHA) already mentioned in Sec. 2.2. In this section, I briefly describe the QHA theory and I highlight the differences and similarities with the SSCHA.

In the QHA, the BO energy surface is expanded as a quadratic function around its minimum at each volume. The free energy is the sum of the BO energy $V(\boldsymbol{\mathcal{R}}, \{\vec{a}_i\})$ at fixed nuclear position $\boldsymbol{\mathcal{R}}$ and cell parameters $\{\vec{a}_i\}$, and the harmonic vibrational contribution:

$$\mathcal{F}_{QHA}(\boldsymbol{\mathcal{R}}, \{\vec{a}_i\}) = V(\boldsymbol{\mathcal{R}}, \{\vec{a}_i\}) + F_{vib}(\boldsymbol{\mathcal{R}}, \{\vec{a}_i\}) \tag{2.127}$$

where

$$\begin{aligned} F_{vib}(\boldsymbol{\mathcal{R}}, \{\vec{a}_i\}) = \frac{1}{N_{\boldsymbol{q}}} \sum_{\boldsymbol{q}\in BZ} \sum_{\mu=1}^{3N} \Bigg[& \frac{\hbar\omega^{\mathcal{H}}_{\boldsymbol{q}\mu}(\boldsymbol{\mathcal{R}}, \{\vec{a}_i\})}{2} + \\ & + \frac{1}{\beta} \ln\Big(1 - e^{-\beta\hbar\omega^{\mathcal{H}}_{\boldsymbol{q}\mu}(\boldsymbol{\mathcal{R}}, \{\vec{a}_i\})}\Big)\Bigg]. \end{aligned} \tag{2.128}$$

Here, $N_{\boldsymbol{q}}$ is the number of $\boldsymbol{q}$ points in the Brillouin zone, $\beta = (k_B T)^{-1}$, and $\omega^{\mathcal{H}}_{\boldsymbol{q}\mu}$ are the volume-dependent harmonic frequencies in the $\boldsymbol{q}$-point for the μ mode. However, as in the harmonic model, the frequencies show no temperature dependence for a fixed volume.

In principle, the QHA free energy is obtained by minimizing the functional $\mathcal{F}_{\text{QHA}}$ in Eq. (2.127) at fixed volume and temperature. However, this minimization is computationally expensive for systems with many degrees of freedom like ice, since it requires the calculation of the harmonic phonon frequencies for each value of the nuclear positions.

As commonly done, in the study of low-pressure ice, the QHA free energy is computed in the minimum $\boldsymbol{\mathcal{R}}_0$ of the BO energy $V(\boldsymbol{\mathcal{R}}, \{\vec{a}_i\})$, obtained by relaxing both the internal coordinates and the cell vectors without vibrations at a fixed pressure.

A direct comparison with the SSCHA is available if we rewrite the SSCHA free energy of Eq. (2.7) in a more fashionable form by adding and subtracting the auxiliary trial harmonic Hamiltonian $\mathcal{H}_{\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}} = K + \mathcal{V}^{\mathcal{H}_{\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}}}$ (as in Eq. A.8):

$$\mathcal{F}[\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}] = \left\langle V - \mathcal{V}^{\mathcal{H}_{\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}}} \right\rangle_{\tilde{\rho}_{\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}}} + F_{vib}(\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}), \tag{2.129}$$

where the vibrational term has the same functional dependence as in Eq. (2.128). The average is computed in an ensemble of configurations generated according to the density matrix $\tilde{\rho}$.

The differences between the two approaches rely upon three main points:

a) The centroid position $\boldsymbol{\mathcal{R}}$

   As mentioned earlier, it is hard to optimize the centroids within the QHA at a reasonable computational cost, so I employ their equilibrium value without vibrations. In contrast, the minimization procedure in the SSCHA consents fully optimization of the geometry at any temperature, including the average nuclear position $\boldsymbol{\mathcal{R}}$ with thermal and quantum anharmonic effects included.

b) The frequencies:

   In the QHA approach, they are the harmonic frequencies. In the SSCHA framework, the frequencies are the eigenvalues of the dynamical matrix obtained through free energy minimization. Thus, they account for anharmonic quantum and thermal fluctuations. Moreover, the dynamical extension of the SSCHA and the methodology reported in Sec. 2.5.3 allows computing the physical phonons, probed in spectroscopic experiments.

c) The SSCHA explicitly accounts for deviation of the real ionic energy landscape from the Harmonic approximation: $\left\langle V - \mathcal{V}^{\mathcal{H}_{\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}}} \right\rangle_{\rho_{\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}}}$ in Eq. (2.129). The inclusion of this term makes the SSCHA free energy variational with respect to the exact one. This property is not shared by the QHA free energy.

Although an approximate technique, this self-consistent approach revealed to perform quite accurately in describing free energy differences, (See Ref. [92]), that determine the thermodynamic properties investigated in low- and high-pressure ice, for instance. Moreover, the large amount of simulations performed in this work prevents the use of more accurate path integral methods. In addition, most of the properties, like the VIE, are studied at zero temperature, a limit that is computationally very expensive and difficult to achieve through PI simulations. Furthermore, PI is a static theory and it requires alternative formulations to describe dynamical spectral functions, while the TD-SSCHA is a dynamical theory that grants direct access to the spectral functions.

2.9 Force Fields

The first-principle computation of energy and forces necessary to the SSCHA machinery is computational demanding. In systems with many atoms or that require

a huge number of configurations to converge or to achieve results with low uncertainties, the computational cost of an *ab initio* approach can be prohibitively expensive. Moreover, when one wants to use a hybrid functional where part of the exact exchange functional is used, the computational cost grows even more and it is almost mandatory to resort to alternative approaches.

Commonly, empirical force fields are used to overcome the great cost of DFT functionals. However, in recent years, the upcoming of machine learning techniques provided a new approach to obtain more accurate and versatile force fields. Shortly, the use of neural networks (NNs) allows learning the energy and forces output by the *ab initio* approach by training the network itself with a number of different atomic configurations that sample the phase space as much as possible.

By paying the price of lower accuracy, but with the advantage of a speedup of several orders of magnitudes, one can use the force fields fitted on DFT functionals by the NN. The computations performed to produce the data of this thesis exploit the availability of such force fields. In this section, I will describe the two adopted very briefly.

2.9.1 revPBE0-D3

I used this force field in the study of ordinary ice I_h and its ordered counterpart ice XI in Chapter 3. revPBE0-D3 is the hybrid counterpart [93, 94] of the Revised Perdew-Burke-Ernzerhof (revPBE) functional of the generalized gradient approximation (GGA) which incorporates 25 % of exact exchange with Grimme D3 dispersion correction [95]. It has been demonstrated to accurately predict structure, dynamics and spectroscopy of water systems in molecular dynamics (MD) and path integral MD simulations [96]. Since the sampling of the phase space of water at the revPBE0-D3 level of theory is prohibitively expensive, Cheng et al. [97] sampled the phase space using a surrogate machine learning (ML) potential energy surface (PES).

They built a flexible and fully dissociable neural network potential for bulk liquid water and ice, following the framework of Behler and Parrinello [98] using the RuNNer code [99], which was trained with revPBE0-D3 energies and forces computed on 1593 different reference structures of 64 water molecules, a $2 \times 2 \times 2$ supercell of the conventional 8 molecules orthorhombic unit cell of ordinary ice. Energy cutoffs for the plane-wave expansion of 1200 Ry have been used within the CP2K package [89].

The NN computes the total energy of the system as a sum of environment-dependent atomic energies. For each atom, the position of all neighboring atoms within a cutoff radius (to be converged), is described by a set of atom-centered many-body symmetry functions, that are used as an input for the NN, yielding the atomic energy contributions. Analytic total energy expression is available as the sum of all the individual atomic energies contribution and analytic gradients for the calculation of the forces are available too. The NN outputs also the stress tensor. Atomic environments within a cutoff radius of 12 Å are described using symmetry functions set (27 functions for hydrogens and 30 for oxygen atoms). The hydrogen and oxygen atomic NNs are made up of two hidden layers with 20 nodes each.

The 1593 reference structures were split into an 80% training set and a 20% test set to estimate the quality of the NN and avoid over-fitting. Structures were chosen

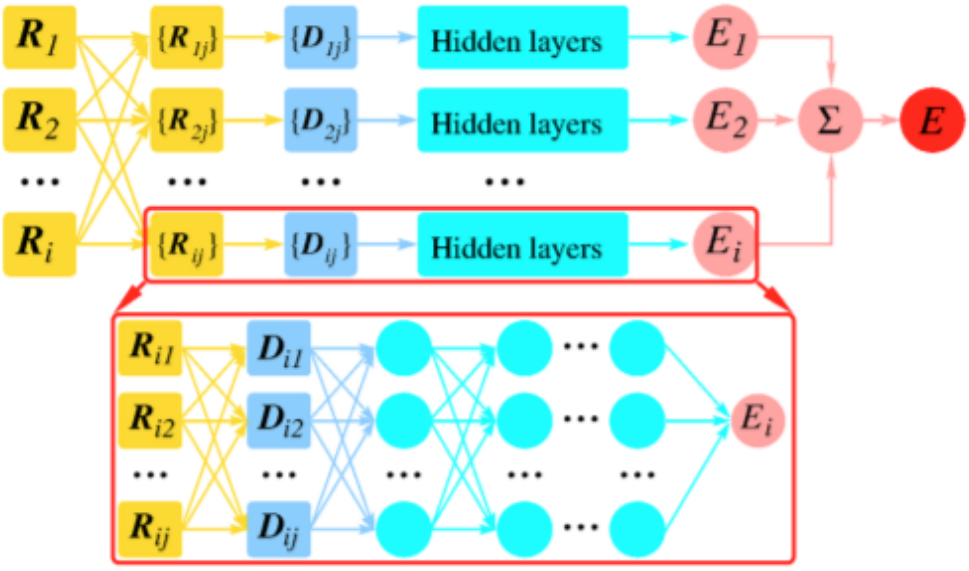

Figure 2.4. Schematic representation of the neural network of Ref. [100]. The enlarged box is the DNN itself. $\mathbf{R_i}$ are the input coordinate and a $\{\mathbf{R_{ij}}\}$ represent the relative positions of all neighboring atoms with respect to atom i. $\{\mathbf{D_{ij}}\}$ represent the local coordinate information and they serve as input for each hidden layer to compute the atomic energies E_i, whose sum gives the total energy. Figure taken from: *Deep Potential Molecular Dynamics: A Scalable Model with the Accuracy of Quantum Mechanics, Zhang et al. PRL 120, 143001 (2018) [101]*

to cover the phase space as much as possible since the accuracy of the predicted energy and forces for a given atomic environment is dictated by the structural similarity between this environment and the ones contained in the reference training set. Inaccurate energy predictions tend to lead the system to explore a non-physical part of the phase space.

2.9.2 PBE

Perdew–Burke–Ernzerhof (PBE) [102] functional is the most common choice in *ab initio* atomistic simulations of water [29, 103, 104]. The computational cost of performing DFT simulation using this kind of non-hybrid functional is smaller with respect to that expected if a hybrid functional, as revPBE0-D3 described in Sec. 2.9.1, is used. However, if the mole of numerical computations to carry out is important, the cumulative cost of using a first-principle approach can become prohibitive even with the PBE approximation of DFT.

This is the case, in the simulation of the phase transition of high-pressure ice, where several pressures, temperatures and structures are sampled. To overcome the computational limitation posed by the *ab initio* approach, I will exploit a Deep Neural Network (DNN) realization of the PBE approximation of DFT by Zhang et al. [24].

In the DNN scheme [101], a local reference frame and a local environment

are assigned to each atom. Each environment contains a finite number of atoms, whose local coordinates are arranged in a symmetry preserving way following the prescription of the deep potential method. The learning process is improved thanks to the use of a flexible family of loss functions. Despite being different from the NN used to produce the revPBE0-D3 force field of Sec. 2.9.1 it works in a very similar way.

In Fig. 2.4 there is a schematic diagram of the DNN used to fit the PBE force field. The atomic positions are rearranged to give the relative positions with respect to all neighbors $\{\mathbf{R_{ij}}\}$ for atom i and then transformed into local coordinate information (radial and angular) $\{\mathbf{D_{ij}}\}$ within a given cutoff radius R_c. $\{\mathbf{D_{ij}}\}$ serve as input for the DNN, made up of multiple hidden layers each with several nodes. Finally the atomic energies E_i are summed up to give the total energy (as in Sec. 2.9.1).

The training data are generated with an iterative learning procedure. The code for this work is integrated into the open-source software package DEEPMD-KIT [105] and the authors used the DP-GEN [100] package for the iterative scheme. The DNN model is an end-to-end scheme that does not use any *ad-hoc* construction in addition to the coordinate information and the network itself.

The model is size extensive, preserves translational, rotational and permutational symmetry, and yields also a polarization that varies continuously with the atomic coordinates. The authors built Deep Potentials (DP) for ice using 2248 single-shot DFT calculations with a 16 water molecules cell and 2400 single-shot DFT calculations with a 128 molecules cell. Recently, ML methods have been used to represent the polarization and its time derivatives [106] needed for IR or Raman calculations. Zhang et al, in their work, propose an approach based on DNN and maximally localized Wannier functions (MLWFs) to predict polarization and polarizability. It exploits the use of Wannier Centers (WC) in which the input atomic coordinates are mapped into output WC coordinates. For the moment, I have not yet exploited the availability of configuration-dependent polarization and polarizability (and so effective charges and Raman tensors) but I plan to do this in the following to analyze the spectral dependence on nonlinearity on the effective charges and Raman tensors.

Chapter 3

Low pressure ice: the I_h and XI phases

3.1 Introduction

In this chapter, I apply the theoretical methods developed in Chapter 2 to study the stable ice phases in the low-pressure region, that is the ordinary ice I_h and its proton-ordered counterpart ice XI. First, I report some information about the structure of the ice phases involved as well as a brief review of the experimental and theoretical literature (Sec. 3.2) that motivates the study of these low-pressure structures. I show that the thermodynamic properties of ice (Sec. 3.3.1), like its peculiar negative thermal expansion or the compressibility and internal geometry, are perfectly reproduced within the self-consistent harmonic approximation framework which provides some improvements with respect to other approximated theories. Ice exhibits an anomalous isotopic effect, and I show that quantum effects on hydrogen are in a strongly nonlinear regime (Sec. 3.3.2): when progressively increasing the mass of hydrogen from protium to infinity (classical limit), the volume first expands and then contracts, with a maximum slightly above the mass of tritium. The last part is dedicated to vibrational spectroscopies (Sec. 3.3.3 - 3.3.5). There, for the first time, I report an accurate comparison of the low-energy phonon dispersion with the experimental data, crucial for the study of thermal transport or to simulate ice under pressure. Finally, I describe the anharmonic renormalization of stretching and bending modes probed in Raman and IR spectroscopies.

The results reported in this chapter have been already published in Ref. [107]: *The microscopic origin of the anomalous isotopic properties of ice relies on the strong quantum anharmonic regime of atomic vibration*

3.2 Ice structures and motivations

Ice I_h [4] is the hexagonal crystal form of ordinary ice in the biosphere (frozen water). It has a completely proton disordered structure, with $P6_3/mmc$ space group and D_{6h} symmetry, the space group of any configuration in accord with the ice rules must be a subgroup of $P6_3/mmc$. In accordance with the third law of thermodynamics, one ordered configuration must have slightly lower energy than the others, resulting

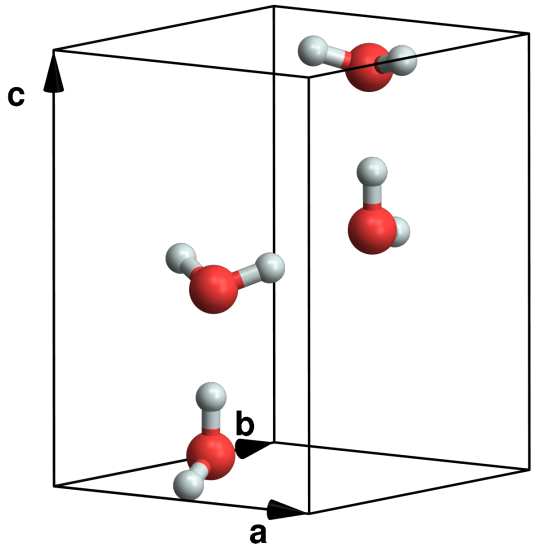

Figure 3.1. Orthorhombic unit cell of phase XI. The unit cell contains 4 water molecules (red and white balls in the figure are oxygen and hydrogen atoms respectively). The structure is ferroelectric, with a permanent electric dipole along the vertical axis.

in the equilibrium structure at low temperatures. The ordered structure, phase XI, is found to be thermodynamically stable below T = 72 K. However, at such low temperatures, the proton is essentially immobile so that on the laboratory time scale, in pure ice, the transition to proton ordered structure never occurs and the disordered ice I_h persists to low temperatures. The timescales to achieve proton order are expected to be of the order of 10^6 years, so that ice XI is the stable form of low-pressure ice on planets where the surface temperature is lower than 72 K, like Pluto.

However, the transformation can be catalyzed by doping I_h with KOH or other hydroxides that promote the migration of protons and the associated change in the proton configuration. Calorimetric measurements [17, 18, 19], in which the doped ice was kept some degrees below the critical temperature for several days to allow the ordering of the crystal and dielectric experiments [108] revealed the ordering transition at T=72 K.

Before the definitive establishment of the structure of ice XI, several proton-ordered structures, with different symmetries, were proposed. Neutron diffraction measurements [5, 109] determined the ferroelectric orthorhombic structure of phase XI. It belongs to $Cmc2_1$ space group with 4 water molecules per unit cell (Fig. 3.1), with oxygen and hydrogen atoms in the 4a e 8b Wyckoff position respectively (see Tab. 3.1)

Thermodynamic properties of the ordered and disordered structures present negligible differences [110], making ice XI a perfect prototype system to simulate ordinary ice I_h. This entails several advantages from a computational perspective:

Atom	Wyckoff position	site symmetry	Coordinates +(1/2,1/2,0)
O	4a	m..	(0,y,z)(0,-y,z+1/2)
H	8b	1	(x,y,z) (-x,-y,z+1/2) (x,-y,z+1/2) (-x,y,z)

Table 3.1. Space group ($Cmc2_1$), site symmetry and Wyckoff positions of ice XI. The orthorhombic unit cell contains 4 water molecules. The values of the free parameters in the SSCHA optimized structure at T=0 K are y=0.316 and z=0.060 for the oxygen atoms and x=0.089, y=0.4747, z=0.480 for the hydrogen atoms.

ordered ice has a smaller unit cell and it exists in only one configuration.

Liquid and crystalline water have been extensively studied in the last decades. The great effort that researchers devoted to the investigation of water is not attributable only to its pivotal role in biological processes, its anomalous properties attract scientific attention and make theoretical prediction challenging. The great difference in strength between the intermolecular hydrogen bonds and the intramolecular covalent OH bonds produces a vibrational spectrum with a wide energy range, from low energy rotons to high energy vibrons [111, 112, 113]. The stretching vibrational modes of the water molecule have an energy of about 3400 cm^{-1}, which requires a temperature of 4900 K to populate the first excited state. Thus, the nuclear motion at room temperature is completely quantum mechanical.

Anharmonic effects play a key role in ice, determining, for example, its anomalous thermal expansion [20, 21, 114] at low temperatures, the inverse volume isotope effect (VIE) [29, 115, 116] and the shifts in the vibrational spectra. The thermodynamic properties of ice have been investigated by using several approximations. Force-field and first-principles based path integral molecular dynamics (PIMD) and quasi-harmonic approximation (QHA) [28, 29, 30, 115, 116] were employed to describe the negative thermal expansion and VIE, enlightening the strong dependence of the results on the force field or the DFT functional [117] used as well as some discrepancies of the QHA at high temperatures and the improvement of the simulations of water and ice obtained by including long-range van der Waals interactions [31, 118]. The vibrational properties of ice have been widely investigated by using mainly Raman, Infrared and neutron spectroscopy [119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133]. Alongside experimental works, several theoretical studies focused on the librational modes [134, 135, 136, 137] and on the OH stretching bands [138, 139], where the authors were able to calculate linewidths in good agreement with experimental data but shifted peak positions. The low energy modes computed using standard *ab initio* techniques display a severe disagreement with experiments. This prevents the study of thermal transport properties in ice *ab initio* and its characterization under pressure, an area of extensive experimental research [140, 141, 142, 143, 144, 145, 146]. An accurate description of atomic vibrations is of paramount importance to reproduce thermodynamic and dynamical properties. In the study of low-pressure ice, I overcome the intrinsic limitations of other methodologies by using the self-consistent harmonic approximation (SSCHA) [32, 55, 56, 57] which exploits a full-quantum variational principle on the free energy to account for the effect of anharmonicity arising from thermal and quantum fluctuations.

3.3 Results

In this section, I report a detailed investigation of ordinary ice I_h and phase XI, its proton-ordered counterpart [17], stable below 72 K. Quantum anharmonic effects on the nuclei affect the properties of the hydrogen bonds, producing exotic behaviors, like negative thermal expansion or anomalous VIE. Soft intermolecular hydrogen bonds coexist with harder intra-molecular covalent OH bonds producing phonons with a very wide energy range, heavily impacted by anharmonicity.

We work within the Born-Oppenheimer (BO) approximation [91] to separate electronic and nuclear degrees of freedom. The total electronic energy at fixed nuclei is calculated using a Neural Network Potential (NNP) developed in Ref. [97], trained on the revPBE0 [93, 94, 147] functional with Grimme D3 dispersion correction [95, 147] to properly account for long-range van der Waals interactions. A detailed description of the force field can be found in Sec. 2.9.1.

It is known in the literature that the commonly used QHA produces accurate predictions of thermodynamic properties at low temperatures ($T \leq 100\,\mathrm{K}$), manifesting instead some inaccuracies at higher temperatures. Therefore, I solve the nuclear BO Hamiltonian by using both the QHA and the SSCHA. The description of the QHA method and the differences with respect to the SSCHA approach can be found in Sec. 2.8.

Both the harmonic and the SSCHA calculations are performed in supercells with periodic boundary conditions. Simulation details and a discussion about the convergence properties of the SSCHA and QHA calculations can be found in Appendix F. Results are converged for a 3x3x2 supercell in the SSCHA and a 14x14x14 for the QHA.

The following section is dedicated to the anomalous thermal expansion of ice, the bulk modulus, the temperature dependence of the crystalline properties and the anomalous isotope effect. For each temperature, I estimated the equilibrium volume $\Omega_{eq}(T)$ as the one where the pressure, defined as the derivative of the free energy with respect to a strain tensor $\boldsymbol{\varepsilon}$, $P = -1/\Omega\, \partial\mathcal{F}/\partial\boldsymbol{\varepsilon}$, vanishes.

$$\Omega_{eq}(T):\ P(\Omega_{eq}(T), T) = 0 \tag{3.1}$$

In the SSCHA framework, we have the analytical expression for the pressure in Eq. (2.19) [57], while in the QHA a finite difference approach is employed.

The application of external pressure to a solid produces a volume change; the bulk modulus is how a crystal withstands modifications of volume under pressure.

$$B(T) = -\Omega_{eq}(T)\frac{\partial P(\Omega, T)}{\partial \Omega}\bigg|_{\Omega_{eq}(T)} \tag{3.2}$$

The simulation of spectroscopic experiments, Raman and IR, can be found in the last part of this section. Throughout all this chapter and in the following one, I replace the volume denomination Ω_{vol} of Chapter 2 with the simple Ω to avoid a cumbersome and hard-to-read notation in the figures.

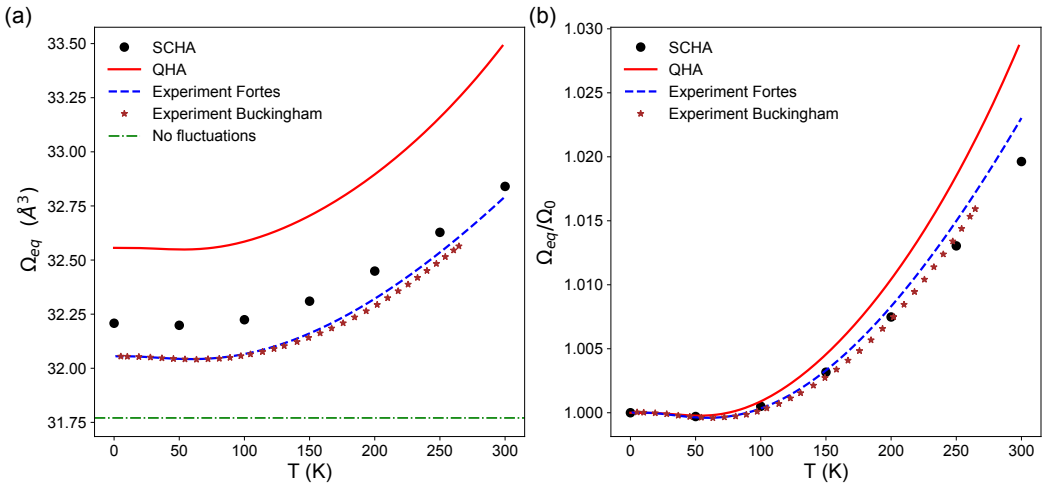

Figure 3.2. Equilibrium volume per H_2O molecule of ice XI as a function of temperature. **a** Comparison of the absolute value of the volume in the SSCHA (black circles) and QHA (solid red line) with the experimental measurements by Fortes [21] (blue dotted line) and Buckingham [148] (brown stars). The classical equilibrium volume, where thermal and quantum fluctuations are neglected, is the green dot-dashed line. **b** The normalized equilibrium volume $\Omega_{eq}(T)/\Omega_{eq}(T=0)$ in the SSCHA (black circles) and QHA (red solid line) is compared to the experimental data (blue dotted line and brown starts).

3.3.1 Thermodynamic properties

The absolute value of the equilibrium volume per H_2O molecule is in Fig. 3.2 (**a**). We notice a considerable effect of the zero point motion that shifts the curves one respect to the other. The zero-temperature equilibrium volumes predicted by the different theories and their percentage shift with respect to the classical limit are reported in the first part of Table 3.2. The agreement with the experiment is much better in the SSCHA than in the QHA picture. This data are discussed in detail in Sec. 3.3.2, concerning the isotope volume effect. Since $\Omega_{eq}(T=0)$ is theory dependent, to compare the thermal expansion, in Fig. 3.2 (**b**) I report the normalized volume $\Omega_{\mathrm{eq}}(T)/\Omega_{\mathrm{eq}}(T=0)$.

In the low-temperature regime, $T \leq 50K$, the predictions for the normalized equilibrium volume in the QHA and SSCHA are very similar and in a good match with the experimental data [21, 148]. For higher temperatures, essential differences between the two theories arise. The SSCHA agrees with the experimental measurements by Fortes [21] within a 0.1 % up to 200 K and the agreement with the Buckingham [148] data is almost perfect in all the temperature range explored. The QHA, otherwise, deviates from the experimental data above 100 K. This establishes the success of the SSCHA theory in describing the thermal expansion of ice, overcoming the limitations of commonly used techniques based on the harmonic approximation.

The temperature dependence of the internal geometry parameters is strictly correlated to the thermal expansion. Within the SSCHA framework we are able to extract information about the covalent and hydrogen bond lengths. The former is reported in Fig. 3.3.

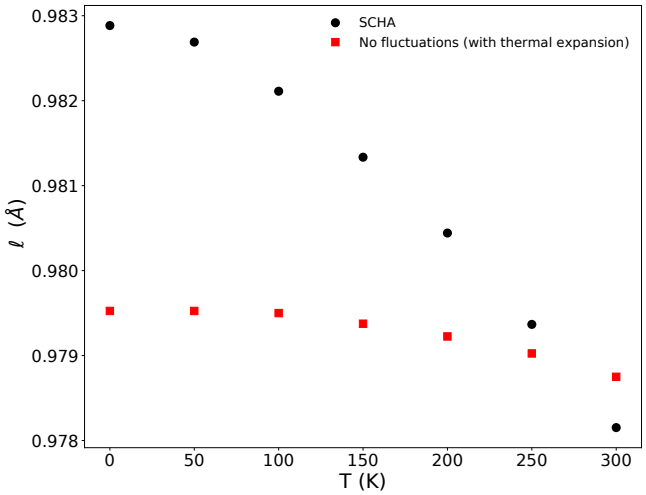

Figure 3.3. Temperature dependence of the covalent bond length. The SSCHA bond length (black circles) is compared with the classical result (red squares), computed from the minimization of the BO energy $V(\boldsymbol{\mathcal{R}}, \{\vec{a}_i\})$, where, the effect of the thermal expansion is introduced by fixing the volume to be the SSCHA equilibrium one at each temperature, in order to point out the contribution of thermal and quantum fluctuations.

Counterintuitively, the water molecules shrink upon heating at high temperatures. This is only marginally a consequence of intermolecular hydrogen bond weakening due to the increasing distance between molecules with temperature, but rather a complex effect of anharmonicity triggered by molecular vibration. If we relax the structure with static nuclei at the SSCHA equilibrium volume for each temperature (Fig. 3.3, red squares), we explain only the 15% of this effect.

Since the vibration that deforms the water molecule is a stretching mode, with $\omega \simeq 3400\text{cm}^{-1}$ and excitation temperature $T \simeq 4900K$, this relevant temperature dependence can only be explained by the anharmonic interaction between translational molecular modes, the only ones whose population changes in this temperature range, and the stretching mode, that affects the OH bond length.

Conversely, the volume expansion explains thoroughly the widening of the hydrogen bond with temperature (Fig. 3.4).

The bulk modulus

Within the thermodynamic properties of a solid, the bulk modulus is worth of a paragraph on its own. In geophysics, the correct description of the compressibility of a solid is of paramount importance for studying the inner composition of Earth and other planets [149, 150]. In fact, it is possible to put in a close relationship the velocity of propagation of sound waves, as those produced in an earthquake, with the bulk modulus of the media through which they propagate, having, in such a way, the possibility to identify its composition given the pressure and temperature conditions.

Eq. (3.2) suggests a direct dependence between equilibrium volume and bulk modulus. The absolute value and the normalized one $B(T)/B(T=0)$ of the bulk

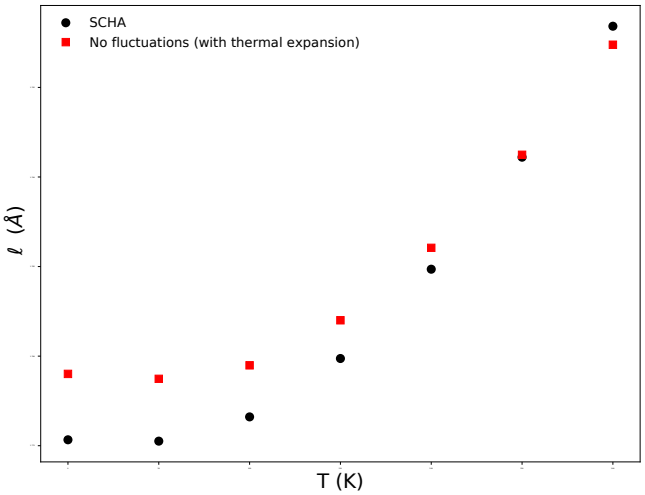

Figure 3.4. Temperature dependence of the hydrogen bond length. The SSCHA bond length (black circles) is compared with the classical result (red squares), computed from the minimization of the BO energy $V(\boldsymbol{\mathcal{R}}, \{\vec{a}_i\})$, where, the effect of the thermal expansion is introduced by fixing the volume to be the SSCHA equilibrium one at each temperature, in order to point out the contribution of thermal and quantum fluctuations.

modulus are reported in Fig. 3.5. Already at $T = 0\,\mathrm{K}$, the bulk modulus is strongly renormalized by ionic quantum fluctuations by a 22% and 15% in the QHA and SSCHA, respectively (Fig. 3.5 (**a**)).

The bulk modulus has an anomalous strong temperature dependence; experimental data [21] show a 20% reduction from 0 to 300 K. The SSCHA reproduces this behavior perfectly (Fig. 3.5 (**b**)), while the QHA overestimates the bulk modulus reduction of 10%.

In order to understand the origin of this strong temperature dependence it is possible to separate the bulk modulus, computed from Eq. (3.2) into a static and a vibrational contribution as for the free energy in Eq. (2.127). By restricting to the QHA picture, the static and vibrational bulk modulus can be expressed through Eq. (3.3):

$$B_{stat}(T) = \Omega_{eq}(T) \frac{\partial^2 V(\boldsymbol{\mathcal{R}}, \{\vec{a}_i\})}{\partial \Omega^2}\bigg|_{\Omega_{eq(T)}} \tag{3.3a}$$

$$B_{vib}(T) = \Omega_{eq}(T) \frac{\partial^2 F_{vib}(\boldsymbol{\mathcal{R}}, \{\vec{a}_i\})}{\partial \Omega^2}\bigg|_{\Omega_{eq(T)}} \tag{3.3b}$$

Fig. 3.6 (**a**) shows the static bulk modulus, where fluctuations are neglected.

I computed the curves at the QHA and SSCHA equilibrium volumes, in order to introduce the effect of thermal expansion. The continuous line is the classic result, where quantum and thermal effects are not included. The vibrational contribution of Eq. (3.3b), for the QHA and SSCHA volumes, is reported in Fig. 3.6 (**b**). It is always negative, meaning that it would increase the volume under compression.

Both the static and the vibrational bulk modulus have a non negligible temperature dependence, contributing to the total one for the 64% and 36% respectively.

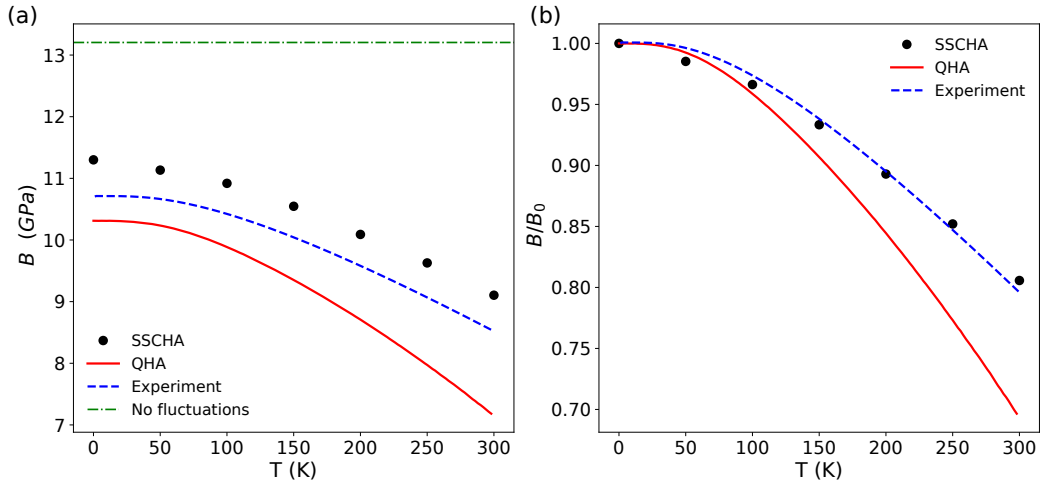

Figure 3.5. **a** Temperature dependence of the bulk modulus. The QHA (red solid line) and SSCHA (black circles) results are compared with the experiment [21] (blue dashed line). The green dot-dashed line is the classical value for the bulk modulus, obtained by neglecting thermal and quantum fluctuations. **b** Comparison of the normalized bulk modulus $^{B(T)}/_{B(T=0)}$ in QHA and SSCHA with the experiment. The bulk modulus is computed as in Eq. (3.2).

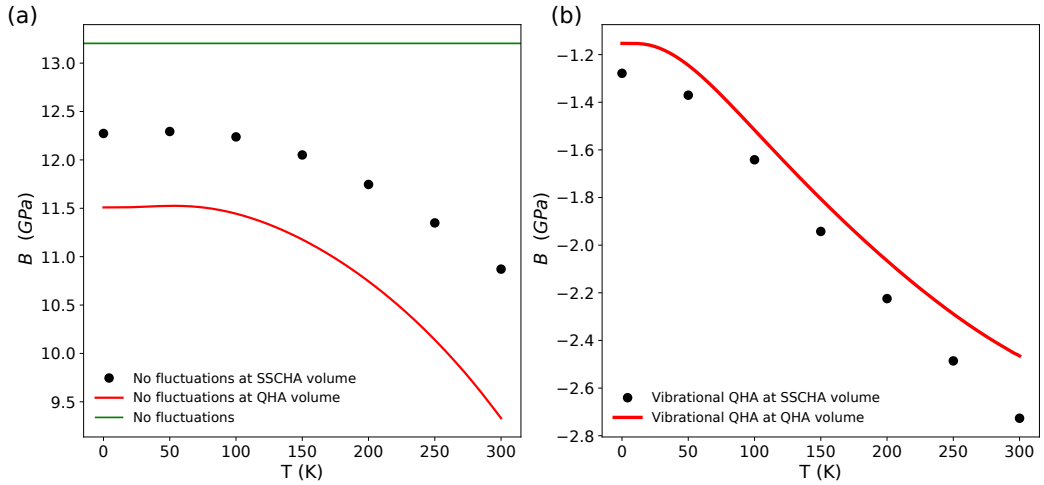

Figure 3.6. Analysis of the different contributions to the bulk modulus. **a** The static QHA bulk modulus is computed at the QHA equilibrium volumes (solid red line) and at the SSCHA equilibrium volumes (black circles). The green solid line is the classical bulk modulus where fluctuations are neglected. **b** Vibrational QHA bulk modulus computed at the SSCHA (black circles) and QHA (red solid line) equilibrium volumes.

Moreover, the effect of different equilibrium volumes is almost temperature independent in the vibrational term, being unable to explain the different thermal behaviour of the two theories, that can instead be addressed partially to the effect of volumes in the static bulk modulus as evident in Fig. 3.6 (**a**).

The combined employment of high accuracy in the electronic exchange-correlation and the accurate description of quantum nuclear motion provided by the SSCHA correctly describe the thermodynamics properties of ice beyond the accuracy of quasi harmonic simulations.

3.3.2 Volume Isotope Effect

According to classical mechanics, the equilibrium volume does not depend on the mass of the atoms, and, thus, it is isotope independent. Quantum effects overturn this simple situation. In most crystalline systems, the heavier the isotopes the smaller the equilibrium volume. In rare exceptions, like ice, by substituting hydrogen with deuterium the equilibrium volume increases [29, 115]. This is known as anomalous volume isotope effect.

In Table 3.2, I report the equilibrium volume per water molecule at zero temperature with protium mass of hydrogen, computed without thermal fluctuations, with quantum effects and harmonic phonons (QHA) and with full anharmonic quantum zero point motion (SSCHA) compared with the experiment in Ref. [21].

As seen, the quantum anharmonic theory is the closest match with the experiment, with an error smaller than 0.5%. The discrepancy between the SSCHA results and those obtained within the QHA, which account for volume effects but not for vibrational anharmonicity, is comparable in size to neglecting ionic motion altogether.

Anharmonicity affects the results in two ways: it changes the average position of nuclei (See Fig. 3.3) and modifies the vibrational frequencies. To shed light on which effect dominates the volume expansion, I repeated the harmonic calculation by fixing nuclear positions to the SSCHA result (QHA @ $\{\boldsymbol{\mathcal{R}}\}_{\mathrm{SSCHA}}$) and by employing also the frequencies shifted by the anharmonicity (QHA @ $\{\boldsymbol{\mathcal{R}}, \omega\}_{\mathrm{SSCHA}}$). The results are reported in the second part of Table 3.2. The last row indicates the volume difference with respect to the full anharmonic theory.

This analysis reveals the origin of the discrepancies between a quasi-harmonic approach and a full non perturbative anharmonic treatment of nuclear vibrations, unveiling how the key role played by anharmonicity is on the frequency renormalization rather than the significative structural changes.

Elucidated the crucial role of anharmonicity in describing the correct volume expansion, I systematically explored the volume effect by varying the mass of the isotopes, both hydrogen and oxygen.

At first, I investigate the dependence of the equilibrium volume at T=0 K on the hydrogen mass, numerical simulations give the possibility to modify this parameter without limitations, while experiments are available only for hydrogen and deuterium [20, 21].

In Fig. 3.7, the equilibrium volumes per water molecule obtained at different levels of the theory are compared with two experimental measurements [20, 21]. The solid line is the classical limit used as a reference. The numerical values for the volume difference at T = 0 K in the theoretical models and the experiments are

	Classic	QHA	SSCHA	Exp. [21]
$\Omega_{eq}(Å^3)$	31.771	32.555	32.207	32.055
%	0	2.47	1.38	

	QHA @ $\{\boldsymbol{\mathcal{R}}\}_{SSCHA}$	QHA @ $\{\boldsymbol{\mathcal{R}}, \omega\}_{SSCHA}$
$\Omega_{eq}(Å^3)$	32.613	32.068
$\Delta(Å^3)$	0.406	-0.139

Table 3.2. The first part of the table reports the equilibrium volume per water molecule of ice XI at T=0 K. In the first row there is the comparison of the volume computed in the QHA and SSCHA with the experiment [21] and with the classical result, where quantum and thermal fluctuations are neglected. In the second row, we report the percentage shift with respect to the classical volume. The second part of the table shows the equilibrium volume per water molecule in the QHA when free energy of Eq. (2.127) is computed using the SSCHA equilibrium positions $\{\boldsymbol{\mathcal{R}}\}_{SSCHA}$ or using both the SSCHA positions and frequencies and their error with respect to the SSCHA result.

	SSCHA	QHA	Exp. [21]	Exp. [20]	SSCHA RPBE-D3
$\Delta\Omega(Å^3)$	0.112	-0.075	0.017	0.029	0.156

Table 3.3. Equilibrium volume per water molecule difference $\Omega_{D_2O} - \Omega_{H_2O}$ at T=0 K. The SSCHA and QHA results computed in the converged meshes are compared with two experimental measurements [20, 21]. The last column shows the volume difference computed in the SSCHA by using the NNP RPBE-D3 functional devised in Ref. [97] in order to analyze the dependence of the VIE on the functional used.

reported in Table 3.3. The anharmonic theory (SSCHA) correctly predicts the sign of the VIE, while the quasi-harmonic approach fails, predicting a volume reduction. However, the SSCHA heavily overestimates the experiment, resulting in a volume difference from 3.8 to 6.7 larger than the measured data.

Experimental data are measured on hydrogen-disordered samples of ice I_h, while the simulation is performed on the hydrogen-ordered ice XI. To unveil the role of hydrogen ordering, I repeated the calculation of the VIE in a hydrogen-disordered structure for ice I_h. I generated disordered cells containing 24 atoms and I ran convergence tests (see App. F) using supercells containing up to 432 atoms (144 water molecules). I found converged properties for supercells with 192 atoms (64 molecules). This result is in agreement with previous theoretical works [97, 151], where 64 molecules simulation cells are used to attain convergence.

I obtained a result deviating by a 3.5 % from the hydrogen-ordered structure, unveiling how hydrogen ordering doesn't significantly affect the VIE, and it is not in origin beyond the discrepancy between theory and experiments.

Our calculation's most relevant source of error is in the electronic correlation: we repeated the simulation employing a different electronic energy engine (a neural network trained on RPBE with Grimme D3 dispersion correction [97]). I obtained a result deviating by a 38% from our simulation, giving a rough estimation of the error introduced by the DFT functional.

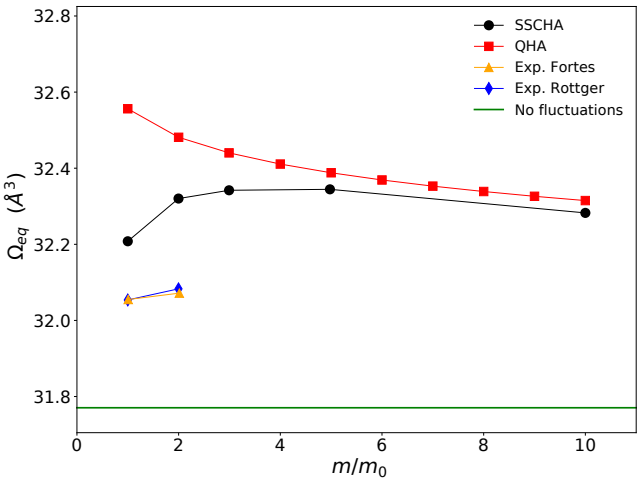

Figure 3.7. Dependence of the equilibrium volume per water molecule of ice on the hydrogen mass at T=0 K. The results obtained in the QHA (red squares) and in the SSCHA (black circles) are compared with two experimental results (yellow triangles and blue diamonds) [20, 21]. The mass independent classical equilibrium volume is reported as a reference (green solid line).

The QHA and SSCHA volumes are different for the physical isotopes of hydrogen (protium, deuterium, and tritium). The difference disappears as we increase the hydrogen mass. This derives from the reduced role of anharmonicity for higher mass where quantum fluctuations at zero temperature are smaller. The crossover above which the QHA correctly reproduces the VIE occurs for an isotope mass of five times the hydrogen's one, not a stable isotope. This means that the quantum regime of protium, deuterium, and tritium is anharmonic, beyond the range of validity of the quasi-harmonic theories.

In Fig. 3.8, I show the equilibrium volume of the solid when the mass of each atomic species (hydrogen and oxygen) is varied separately until the classical limit of infinite mass is reached, as well as the evolution of the equilibrium volume when the mass of the whole molecule is increased. Obviously, there are only few of these combinations in nature that are stable, but we can infer the quantum nature of each element from this plot.

By increasing the mass of the whole solid, we manage to approach the classical limit (the continuous line) for $m/m_{\mathrm{H_2O}} \gtrsim 10000$.

The largest natural atomic species weighs only 238 times the mass of protium. This unveils how quantum effects on nuclei, usually neglected in atomistic calculations, are of paramount importance even with "heavy" atoms. This is further proved by the isotope volume effect of oxygen that, scaled to its much lower mass ratio between its natural isotopes, is bigger than hydrogen.

Fig. 3.8 reveals a nonmonotonous volume expansion compared with quantum fluctuations for the hydrogen isotopes. First, the equilibrium volume expands when we increase the mass, reaching the maximum value for $m \simeq 5m_H$, then we observe a contraction to the classical value for bigger masses. This behavior explains that the VIE is due to a crucial nonlinear regime of quantum fluctuations in ice, overturning

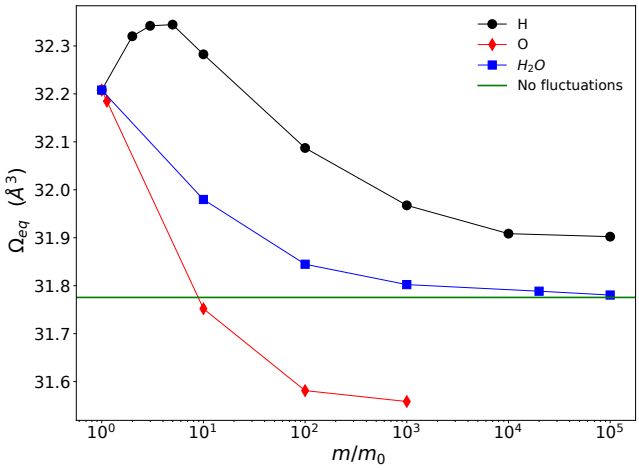

Figure 3.8. Volume isotope effect for all the atomic species in ice XI in the SSCHA framework at T=0 K. Black circles indicate the equilibrium volume per H_2O molecule when the hydrogen mass is increased by keeping the oxygen mass fixed to its ^{16}O isotope value. Red diamonds show the equilibrium volume when the oxygen mass is varied with fixed hydrogen mass. Blue squares stand for the equilibrium volumes when the mass of the entire water molecule is increased until reaching the classical limit shown as a reference (green solid line).

the hypothesis of a monotonous volume reduction due to quantum effects. This exotic behavior cannot be explained in a quasi-harmonic picture, as evident from Fig. 3.7.

To check the dependence of the nonmonotonous volume expansion for hydrogen isotopes on the DFT functional adopted, I repeated the simulation (see Fig. 3.9) with the NNP BLYP-D3 functional devised in Ref. [97]. The increase of the equilibrium volume with the mass of the hydrogen isotope is confirmed, as well as the maximum value reached for a hydrogen mass between three and five times the natural one and the successive decrease toward the classical limit. The use of a different approximation of the exchange and correlation DFT functional modifies the absolute value of the equilibrium volume and the relation between the quantum and classical results. BLYP-D3 predicts an anomalous VIE with a classical equilibrium volume that lies between the hydrogen and deuterium value, while revPBE0-D3 shows an anomalous VIE with volume expansion with respect to neglecting quantum fluctuations.

The correct inclusion of quantum anharmonic fluctuations within the SSCHA managed to explain the role of nuclear quantum effect (NQE) for what concerns the anomalous VIE in water ice. Indeed, NQE plays a major role in several properties of solid and liquid water. It is known that there are competing quantum effects. While NQE extends the covalent bond, delocalizing the protons and strengthening the hydrogen bond, if one takes into consideration the quantum fluctuations in other directions, along the bending modes, that lead to distortion we observe the weakening of the hydrogen bond and it would be interesting to separate the two contributions to understand which of them dominates. Moreover, the present analysis of VIE can

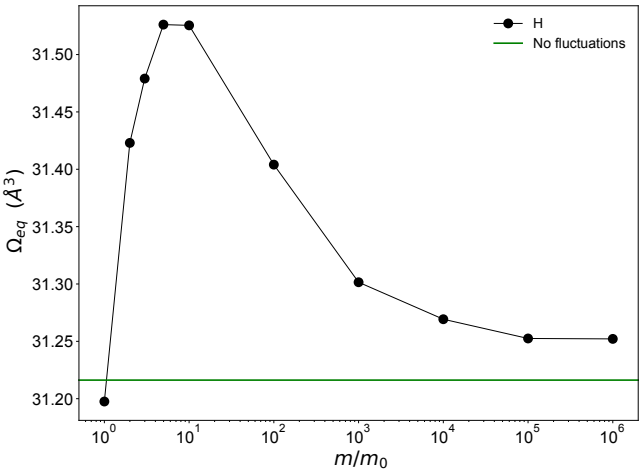

Figure 3.9. Volume isotope effect at T=0 K for the hydrogen atoms within the SSCHA framework by using the NNP BLYP-D3 functional devised in Ref. [97]. The black circles indicate the equilibrium volume per water molecule when the hydrogen mass is increased progressively, keeping the oxygen mass fixed to its ^{16}O isotope value. The green solid line is the classical reference for the particular force field adopted.

be easily extended to the finite temperature case, where the difference between the deuterated and hydrogenated ices equilibrium volumes is expected to increase.

Nuclear quantum effects play an important role also in the determination of the melting point of ice, in fact, D_2O ice forms a stronger hydrogen bond that increases the stability of ice over the liquid, increasing the melting temperature. In general, the effect of the NQE on the bond lengths has not been deeply investigated in this work, giving hints for future works. The impact of quantum fluctuations on hydrogen and covalent bonds is expected to give also further explanations to the VIE. Normal VIE is observed when the covalent bond length is small compared to the hydrogen bond one, as for ice VIII in the low-pressure regime (see Chap. 4.4) and vice-versa for the anomalous VIE.

3.3.3 Phonon dispersion

The coexistence of strong intra-molecular and weak intermolecular bonds in ice produces a vast vibrational spectrum. To compare with experimental results, I computed the real phonons from the dynamical interacting Green function within the time-dependent SSCHA [33, 56] (TD-SSCHA) to fully account for dynamical quantum anharmonic effects (see Sec. 2.5.3 for further details).

Acoustic modes play a major role in thermal transport. I compared the simulated phonon dispersion of deuterated ice at T=140 K and P=0.05 GPa in Fig. 3.10 with the experiment [143]. In the low energy regime, the static approximation of the self-energy of Eq. (2.63) is suitable to describe the real phonon dispersion. A comparison between the harmonic and the SSCHA auxiliary phonons and DOS is in App. G.

I consider the lowest order self-energy correction, the bubble of Eq. (2.64).

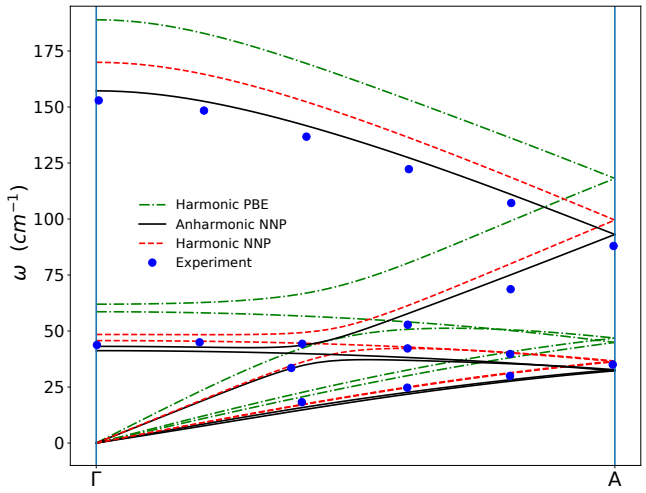

Figure 3.10. Comparison of the low-energy phonon dispersion (molecular translations) with the experiment [143] (blue circles) for deuterated ice at T = 140 K and P = 0.05 GPa. Harmonic (red dashed lines) and anharmonic phonons (black solid line), computed as the poles of the interacting one-phonon Green function in the static limit (using the hybrid revPBE0-D3 functional fitted with the NNP [97]) are shown. The effect of the DFT functional on the harmonic phonon dispersion is reported by showing the harmonic dispersion calculated with PBE [102] functional (green dot dashed lines).

This approximation is reasonable and routinely employed in many system with hydrogen[152, 153]. In the present case, I checked the approximation against the exact static self-energy, where phonon energies change less than $1\,\mathrm{cm}^{-1}$.

While the harmonic phonons can be computed without effort in the supercell described above, the computational cost of the bubble term is huge, and it is not possible to perform the calculations in the same supercell used for the SSCHA dynamical matrix $\boldsymbol{D}^{(\mathrm{S})}$. To overcome this problem, I first computed the bubble correction in a smaller supercell, namely a 3x3x2, and then I interpolated it to a finer supercell to obtain the Hessian matrix $\boldsymbol{D}^{(\mathrm{F})}$.

$$\boldsymbol{D}^{(\mathrm{F})}_{4x4x4} = \boldsymbol{D}^{(\mathrm{S})}_{4x4x4} + \left[\boldsymbol{D}^{(\mathrm{F})}_{3x3x2} - \boldsymbol{D}^{(\mathrm{S})}_{3x3x2}\right]^{4x4x4} \tag{3.4}$$

The harmonic acoustic phonons are in good agreement with the experiment and deviate significantly from experimental data at high frequency ($\omega \geq 120\,\mathrm{cm}^{-1}$). Anharmonic effects correct the deviation, obtaining a perfect match between theory and experimental data. The harmonic energy of the lowest acoustic branch overestimates the speed of sound of 20%, introducing a substantial error in the determination of thermal transport properties, further stressing the fundamental role of anharmonicity in thermal conductivity.

I report an excellent agreement between experimental data and a simulation of ice with *ab initio* accuracy for the first time. This result has a profound impact enabling the first principles simulation of thermal transport, where an accurate description of the acoustic phonons is required. Moreover, low-energy phonons are the only modes detectable at high pressure; our work paves the way to characterize

ice under pressure further.

The astonishingly good agreement I achieve is merit of the combined effect of the correct treatment of anharmonicity and of the electronic functional adopted for the calculation: the harmonic phonons with PBE, the common choice in *ab initio* atomistic simulations of water [29, 103, 104] are reported in Fig. 3.10 (green dot dashed lines), offering a comparison with phonons computed with NN-revPBE0 (employed in our work). We notice a considerable dependence on the DFT functional. The use of the NNP improves the PBE harmonic dispersion, where the error committed approximating the experimental points ranges from 30% to 36% e.g. in the A point.

In Fig. 3.10, I restricted the original path of the experiment (the same as in Fig. G.1) to the Γ-A direction. I report the phonon dispersion along the Γ-K direction in App. G.

3.3.4 Combination modes

The phonon spectral function $\sigma(q, \omega)$ gives access to the quasiparticles energies and lifetime. The spectral function is proportional to the signal probed in scattering experiments, as neutron or X-ray scattering, and it is computed from the diagonal elements of the dynamical one-phonon Green function $G(q, \omega)$ as in Eq. (2.70).

The phonon density of states (DOS) computed with the SSCHA dynamical matrix at equilibrium (without the self-energy correction) describes anharmonic non-interacting phonons, while the dynamical spectral function encapsulates all the effects of phonon-phonon interactions, where, the addition of the self-energy term Eq. (2.59) may produce combination of modes: the Λ tensor in the definition of the bubble correction of Eq. (2.55) is defined though the $\mathscr{F}$ function of Eq. (2.56) that has poles at the sum and differences of the SSCHA frequencies.

In Fig. 3.11, I report the comparison between the phonon DOS and the spectral function at Γ of H_2O ice XI to enlighten the presence of combination modes and anharmonic overtones. I reveal the presence of two structures in the spectral function that are absent in the DOS: one occurs at energies between the bending and stretching bands, the other at twice the frequency of the stretching modes.

The same results apply for the deuterated ice XI in Fig. 3.12, where the occurrence of features absent in the phonon DOS hints the presence of combination of modes due to phonon-phonon interaction.

The interaction between phonon bands can be dissected to unveil which modes originate these satellite peaks in the spectral function. I do this by restricting the sum over the modes μ, ν in Eq. (2.55) only to the phonon modes I want to make interact.

I report the results for H_2O ice at T = 200 K, but the same also holds in deuterated ice at T = 140 K and P = 0.05 GPa. I show in Fig. 3.13 that the first combination mode, occurring for $\omega \in [2250, 2800]\text{cm}^{-1}$, is originated from the interaction between the libration and bending bands, as the peak appears only if we account for their reciprocal interaction. If we account only for librational or bending modes in the calculations, the peak vanishes. The residual mismatch in the low-frequency tail with the full spectral function reveals a nonnegligible contribution of the other phonon branches (mainly translations).

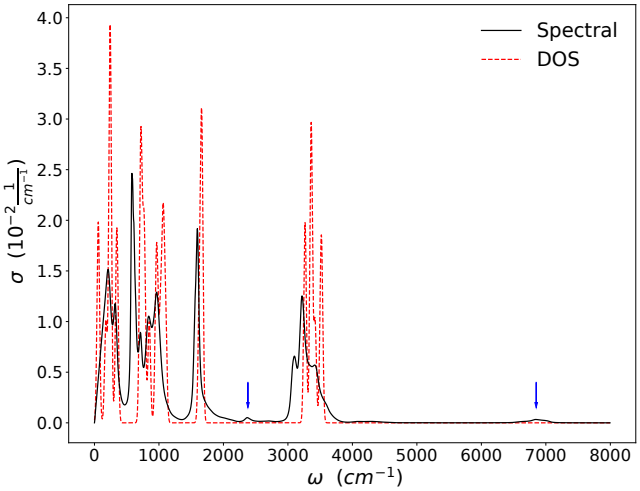

Figure 3.11. Comparison of the density of states and spectral function at Γ. H_2O ice XI at T=200 K and ambient pressure. Spectral function (black solid line) and DOS (red dashed line) are shown. The DOS is computed by using the SSCHA dynamical matrix at equilibrium without the inclusion of the self-energy term and by adding an artificial smearing factor of 20 cm^{-1}. Blue arrows indicate the combination modes and anharmonic overtones.

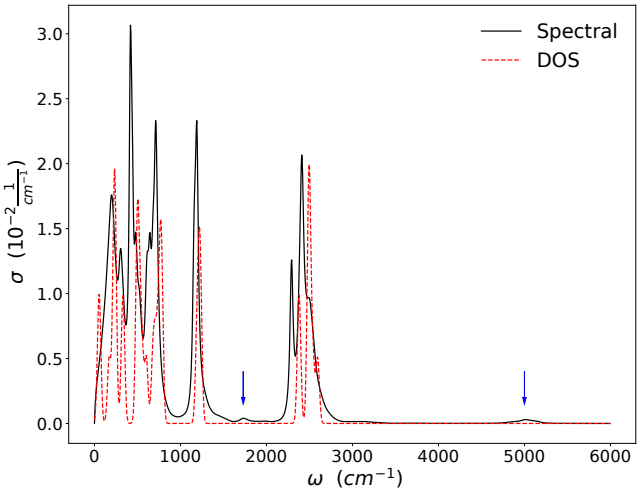

Figure 3.12. Comparison of the density of states and spectral function at Γ. D_2O ice XI at T=140 K and P = 0.05 GPa. Spectral function (black solid line) and DOS (red dashed line) are shown. The DOS is computed by using the SSCHA dynamical matrix at equilibrium without the inclusion of the self-energy term and by adding an artificial smearing factor of 20 cm^{-1}. Blue arrows indicate the combination modes and anharmonic overtones.

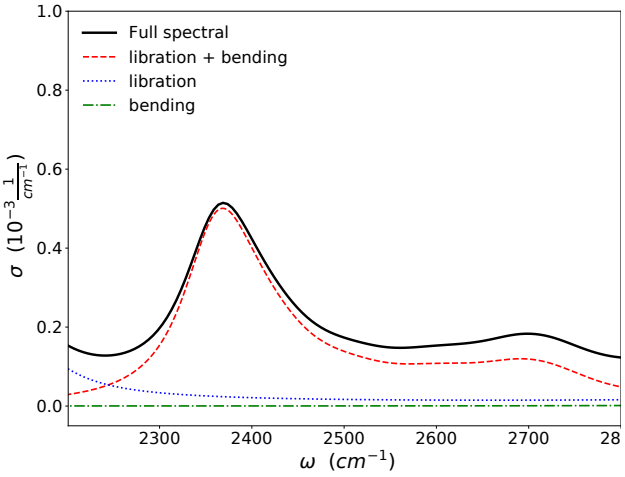

Figure 3.13. Spectral function of H_2O ice XI at T = 200 K in the energy range of the first combination mode at Γ. A comparison between the full spectral function (solid black line) and that obtained by selecting only the interaction between the librations and the bending band (red dashed line) is provided. Spectral functions obtained through the interaction of bending band with themselves (green dot dashed line) and librations with themselves (blue dotted line) are shown too.

The overtone at twice the stretching frequency is analyzed in Fig. 3.14. The perfect matching between the full spectral function and the same computed only considering stretching modes unveils how this peak is entirely generated by stretching modes interacting with themselves, without a significative contribution of other phonon branches. It is, in fact, the overtone of the stretching modes.

From the spectral function, we can extract the real phonon energies and their lifetimes in different ways as depicted in Sec. 2.5.3. In this case, I solve Eq. (2.82) and Eq. (2.83) self-consistently. I report in Table 3.4 the phonon energies and their linewidths for few selected intense modes at T = 0 K and T = 300 K in the Brillouin zone center.

We observe an essential reduction of the lifetime increasing the temperature. The phonon energies are less temperature-dependent than their lifetime. Stretching modes gain energy upon heating while all the others become softer.

3.3.5 Spectroscopy: Raman and IR

Anharmonicity shifts the frequencies of phonons and introduces a finite lifetime. Here, I simulate the Raman and IR vibrational spectroscopy on ice, adequately accounting for quantum and thermal anharmonic nuclear motion.

Raman scattering and infrared absorption are complementary tools to probe phonon energies at Γ. The former is based on an inelastic scattering process that detects modes due to changes in the polarizability, while the latter relies on the absorption process and the vibrations detected involve modifications of the dipole moment. Consequently, the selection rules for the two spectroscopies are different, and often active IR modes are Raman inactive (or vice versa).

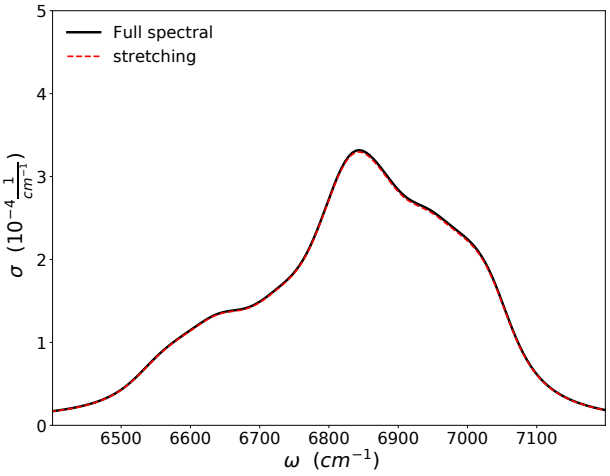

Figure 3.14. Spectral function of H_2O ice XI at T = 200 K in the energy range of the second combination mode at Γ. The full spectral function (solid black line) is compared to that obtained by selecting only the interaction of the stretching band with itself (red dashed line).

	(a) T=0 K		**(b)** T=300 K	
	Ω_μ [cm^{-1}]	Γ_μ [cm^{-1}]	Ω_μ [cm^{-1}]	Γ_μ [cm^{-1}]
	208	12	211	75
tr.	326	11	311	52
	602	4	557	20
	740	5	682	43
lib.	851	12	808	47
	968	16	940	54
	1565	16	1551	31
bend.	1606	20	1594	40
	3102	13	3080	68
str.	3207	16	3237	77
	3399	47	3439	97

Table 3.4. Peak frequency and linewidths of some selected intense modes in H_2O ice XI at ambient pressure for two values of temperature, T=0 K and 300 K at Γ. The first column indicates to which band the modes belong. See Sec. 2.5.3 for details about the calculations of frequencies and linewidths.

The Raman spectrum is proportional to the polarizability correlation function $\langle \alpha(t)\alpha(0)\rangle$. The Raman response function, when the Raman tensor does not depend on the ionic displacements can be easily expressed as a function of the one-phonon Green function as in Eq. (2.120). Details on the computation of the Raman tensor are in App. F

Considering the coupling of the polarizability with the incident and scattered electric field [125, 129, 130], we can define a "reduced" Raman tensor, depending only on the Cartesian indices, by contraction with the polarization vector of the incident ($\mathcal{E}^{in}$) and scattered ($\mathcal{E}^{out}$) light:

$$A'_c = \mathcal{E}^{\text{in}}_{\mathbf{a}} \cdot A_{abc} \cdot \mathcal{E}^{\text{out}}_{\mathbf{b}}$$

Finally, considering all the factors, the intensity of the Raman signal can be evaluated as:

$$I^{\text{Raman}}(\omega) \propto (\omega - \omega_L)^4 \sum_{a,b=1}^{3N} \frac{A'_a A'_b}{\sqrt{M_a M_b}} \text{Im} G_{ab}(\omega) \tag{3.5}$$

where ω_L is the frequency of the laser.

In most cases, the laser frequency is much bigger than the phonon frequencies, so that the approximation $|\omega - \omega_L| \simeq \omega_L$ holds. This is not completely true in ice, where phonons can be very energetic ($\omega \simeq 3400$ cm^{-1}). Here, that difference provides a slight reshaping of the spectra, and it is safer to avoid approximations.

In Fig. 3.15 (**a**), I compare the simulated Raman spectra with the experiment [130] for the same geometry c(a,*)b in H_2O ice XI at T=65 K. We report deuterated ice at T=269 K in the c(a,a)b geometry [125] in Fig. 3.15 (**b**) .

The result obtained in the anharmonic dynamical theory (TD-SSCHA) matches perfectly with the experimental results, correcting a shift of the harmonic phonon energy in the stretching modes of about 10% (7%) of the energy in H_2O (D_2O) ice. The theory can predict the presence of the combination mode (indicated by the blue arrow in Fig. 3.15 (**a**)) for H_2O ice in the considered geometry. Instead, this mode has very low intensity in the experimental geometry of D_2O ice.

Finite linewidths in the harmonic model are for presentation purposes only, as harmonic phonons have infinite lifetimes.

Phonon vibrations impact the low energy optical properties of any material. The physical quantity determining the Infrared absorption is the dielectric function and standard electromagnetism provides a simple relation between the dielectric tensor and the susceptibility.

$$\epsilon(\omega) = 1 + \frac{4\pi\chi^{(\text{tot})}}{\Omega}(\omega) = \epsilon^{\text{el}} + \frac{4\pi}{\Omega}\chi^{\text{ion}}(\omega) \tag{3.6}$$

The electronic part ϵ^{el} is computed *ab initio*. The absence of electronic transitions in the phonon energy range makes it real and it is frequency independent, $\epsilon^{\text{el}} = 1.65$. The ionic susceptibility of Eq. (3.6) is the Fourier transform of the dipole-dipole correlation function and it has been computed in Sec. 2.7.2, Eq. (2.110), in the case when the effective charges do not depend on the ionic atomic positions. Details on the computation of the effective charges are in App. F.

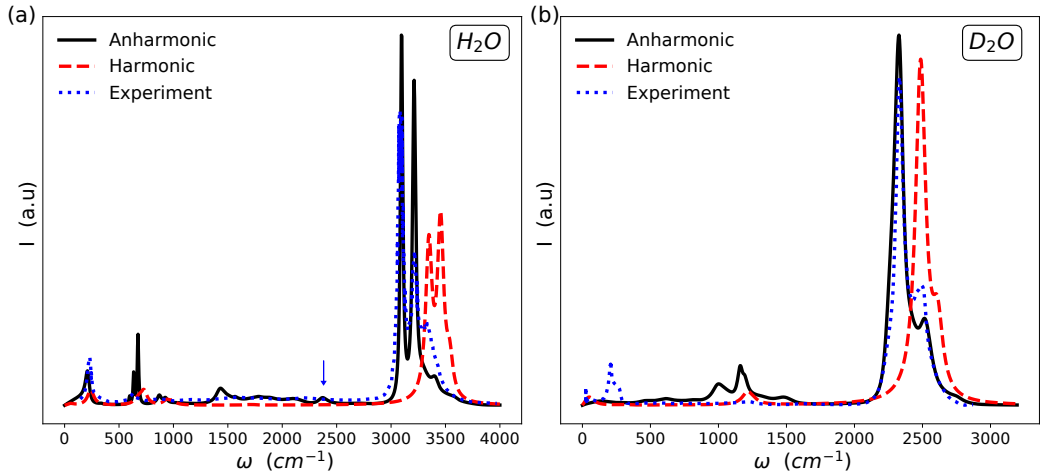

Figure 3.15. **a** Raman scattering spectra at T=65 K in H_2O ice XI in the a(c,*)b geometry. The experimental spectrum [130] (blue dotted line) is compared with the harmonic (red dashed line) and the anharmonic phonons (solid black line), computed in the SSCHA framework with the inclusion of the bubble term in Eq. (2.59). **b** Raman scattering spectra for deuterated ice at T=269 K in the c(a,a)b geometry. Comparison between harmonic (red dashed line), anharmonic SSCHA phonons with the bubble correction (solid black line) spectra, and the experiment [125] (blue dotted line). Blue arrows indicate the position of the combination mode. An artificial broadening is employed in the harmonic approximation to guide the eyes in the comparison of the spectrum with the experiment of $35\,\mathrm{cm}^{-1}$ ($45\,\mathrm{cm}^{-1}$) in D_2O (H_2O). Instead, the broadening of the anharmonic simulation is fully obtained *ab initio* from phonon-phonon scattering.

In analogy with the Raman case, the dielectric function can be expressed as a function of the one-phonon Green function:

$$\epsilon_{\alpha\beta}(\omega) = \epsilon^{\text{el}}_{\alpha\beta} + \frac{4\pi|e|^2}{\Omega} \sum_{ab} \frac{Z_{\alpha a} Z_{\beta b}}{\sqrt{M_a M_b}} G_{ab}(\omega) \tag{3.7}$$

and we averaged over all possible orientation of the crystal:

$$\epsilon(\omega) = \frac{1}{3} \sum_{\alpha=1}^{3} \epsilon_{\alpha\alpha}(\omega) \tag{3.8}$$

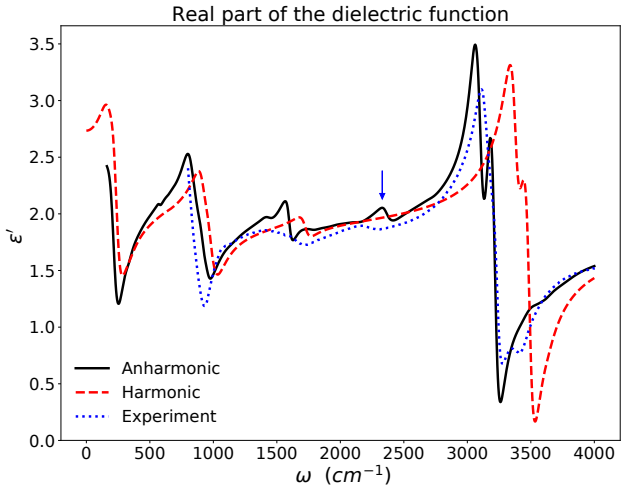

Figure 3.16. Real part of the dielectric function of H_2O ice at T=200 K. All the possible orientations of the crystal are taken into account by reporting the trace of the dielectric tensor as in Eq. (3.8). The results within the harmonic approximation (red dashed curve) and the anharmonic phonons (black solid line), computed in SSCHA framework with the inclusion of the bubble term of Eq. (2.59), are compared with the experiment [126] (blue dotted line). The blue arrow indicates a combination mode. I employed an artificial broadening of 50 cm^{-1} in the harmonic spectrum. Instead, the broadening of the anharmonic simulation is fully obtained *ab initio* from phonon-phonon scattering.

In Fig. 3.16 and Fig. 3.17, I report the real and imaginary part of the dielectric function at 200 K of H_2O ice. To include all the crystal orientations, I plot the trace of the dielectric tensor as in Eq. (3.8). As for the Raman, anharmonicity reduces by 10% the harmonic stretching band energy, providing perfect agreement with the experiment [126]. Also, the combination mode located at 2300 cm^{-1} (highlighted by the blue arrow in Fig. 3.17) observed experimentally is correctly reproduced by the anharmonic spectrum.

Both the dielectric function of Fig. 3.17 and the Raman scattering spectra in Fig. 3.15 confirm the importance of anharmonicity. They further demonstrate how it is fundamental to reproduce the experimental results and make the SSCHA (and its time-dependent extension) as the best tool for ice simulation.

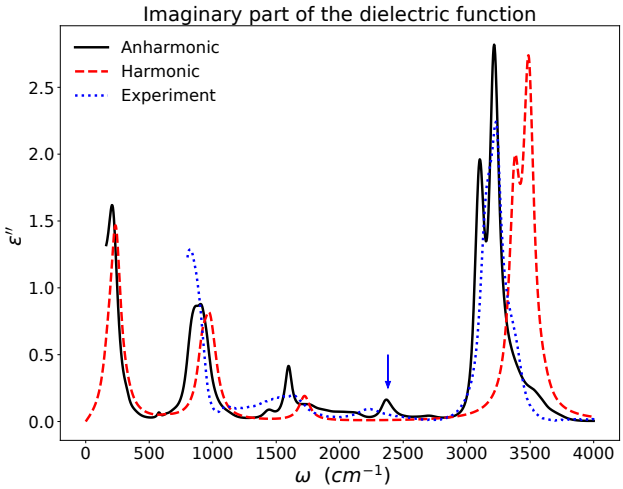

Figure 3.17. Imaginary part of the dielectric function of H_2O ice at T=200 K. The trace of the dielectric tensor as in Eq. (3.8) is reported to take into consideration all the possible orientation of the crystal. The experimental spectrum [126] (blue dotted line) is compared with the results within the harmonic approximation (red dashed curve) and the anharmonic phonons (black solid line), computed in SSCHA framework with the inclusion of the bubble term of Eq. (2.59). The blue arrow indicates a combination mode. An artificial broadening of $50\,\mathrm{cm}^{-1}$ is used in the harmonic spectrum. Instead, the broadening of the anharmonic simulation is fully obtained *ab initio* from phonon-phonon scattering.

3.4 Conclusions

Exploring the thermodynamic structural and vibrational properties of ice XI (hydrogen-ordered counterpart of ordinary ice I_h), I further elucidated the importance of quantum anharmonic effects. The anomalous strong temperature dependence of the bulk modulus, 20% variation from 0 to 300 K, is fully explained by thermal and quantum anharmonic fluctuations, revealing the combined effect of the vibrations (64%) and the thermal expansion (36%).

I highlight the inaccuracy of the quasi-harmonic approximation in reproducing the anomalous VIE. Only an anharmonic treatment of quantum nuclear motion enables the reproduction of the experimental results. In particular, I proved how the negative VIE originates from a nonmonotonous volume expansion due to quantum fluctuations. If we increase the mass of the hydrogen isotopes, the volume first expands, saturating slightly above the tritium mass, and then contracts to the classical value. This means that the VIE is due to a strongly nonlinear regime of quantum fluctuations in ice, which commonly employed approximate theories (such as the QHA) do not grasp. Notably, also oxygen is in a strong quantum mechanical regime, being responsible for a 2% volume reduction in the classical limit.

We observe an anharmonic renormalization of $8-10\%$ in the bending and stretching modes that grants a good prediction of Raman scattering spectra and dielectric function of H_2O and D_2O ice XI.

For the first time, the low-energy range of phonon dispersion of deuterated ice at T=140 K is excellently reproduced by the anharmonic renormalized phonons and

correct treatment of the electron exchange and correlation. This result paves the way for the study of thermal transport from first principles and the simulation of ice under pressure, where acoustic phonons are the only modes detectable.

The simulations deciphered the microscopic origin of many anomalous properties of ice, proving how anharmonicity and quantum fluctuations of ions are a mandatory ingredient to reproduce the thermodynamic structural and vibrational properties of ice. The work done does not cover all the roles of the nuclear quantum effects in ice but it can pave the way for further investigation. In particular, as discussed in Sec. 3.3.2, in solid and liquid water there are competing quantum effects that act as weakeners and strengtheners of hydrogen bonds and it would be interesting to separate their contribution within the SSCHA to understand which one dominates. The effect of the NQE on the bond lengths is also related to the presence or not of the anomalous VIE, so its analysis would give further strength to the explanation of the VIE given here.

Chapter 4

High pressure ice

4.1 Introduction

The symmetrization of the hydrogen bond length at high pressure is a long-debated issue. From the experiments, it is not clear if the transition from ice VII (or its proton-ordered counterpart, ice VIII, stable at low temperature) to ice X (the symmetric phase) is direct or if an intermediate phase exists. The experimental signatures of the phase transition are all indirect (hydrogen atoms are not visible by X-ray diffraction and the limited data quality and uncertainties in the procedure of data correction in neutron scattering hampers an unambiguous interpretation of the experiments) and come from vibrational spectroscopy. However, they give contrasting results because of the strong anharmonic regime close to the phase transition. Classical simulations fail in the prediction of the critical pressure, providing a value that strongly overestimates the physical one.

In this chapter, I solve the problem by simulating the structural and thermodynamic properties of the phase transition (Sec. 4.4), where I show that quantum effects play a major role, they reduce the classical critical pressure, producing a good agreement with the experimental value and reproducing perfectly the isotopic effect (the substitution of hydrogen with deuterium brings a 10 GPa increase of the critical pressure). Then, with the TD-SSCHA simulations of the vibrational spectra (Sec. 4.5), I show that the assignment of spectra to an intermediate phase is an error due to the strong anharmonicity. The analysis of the spectra with the clear underlying crystal structure allows me to identify uniquely the signatures of the phase transition as the strong intensity increase of the translational mode and the sudden collapse of the stretching mode in a small pressure range.

The results shown in this chapter have not been published yet.

4.2 Ice structures

The high-pressure region of the ice phase diagram has been widely investigated in recent years for many reasons. Besides the important role it plays in planetary physics, it offers the possibility to inspect the hydrogen bond interaction. The distance between the oxygens of two neighboring water molecules can be modified by acting on the pressure, leading to the symmetrization of the $O-H\cdots O$ bond. The

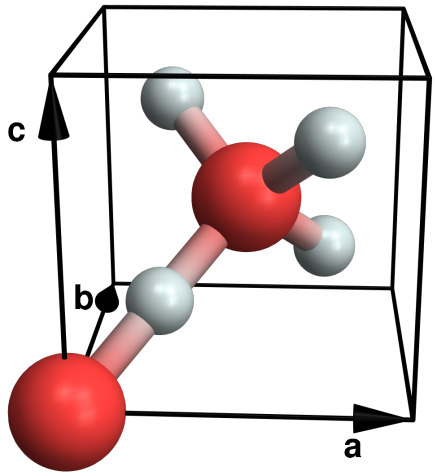

Figure 4.1. Unit cell of the nonmolecular ice X. Red and white atoms are oxygen and hydrogen atoms respectively. The cubic $\vec{a}, \vec{b}, \vec{c}$ lattice parameters are reported. Symmetrized H-bonds are evident from the shared bonds with two oxygens.

idea of H-bond symmetrization relies on the correlation between the covalent bond length d_{OH} and the $O \cdots O$ distance d_{OO} known at ambient pressure. A nonlinear covalent bond stretching is observed upon d_{OO} decreasing [154, 155]. This implies that during compression in hydrogen-bonded compounds, as the distance between oxygens decreases the covalent bond length increases, as supported by the observation of the softening of the stretching frequencies[26]. Upon further compression, at a critical value of the d_{OO} distance, the proton occupies the midpoint between the two neighboring oxygen atoms. In this phase (ice X) the arrangement of ice is isomorphic to the Cu_2O structure and the picture of ice as a molecular crystal breaks down: the water molecule cannot be identified anymore and the system is better described in terms of H^+ and O^{2-} ions. The resulting ionic arrangement of the atoms is characteristic of hydrogen-bonded compounds at high pressure. Ice X has a cubic structure, belonging to the $Pn\bar{3}m$ space group. Its unit cell contains two water molecules that occupy high symmetry positions. Structure and Wyckoff positions are in Fig. 4.1 and Tab. 4.1, respectively.

Atom	Wyckoff position	Coordinates
O	2a	(0,0,0) (1/2,1/2,1/2)
H	4b	(1/4,1/4,1/4) (3/4,1/4,3/4) (1/4,3/4,3/4) (3/4,3/4,1/4)

Table 4.1. Space group ($Pn\bar{3}m$) and Wyckoff positions of Ice X. The body centered cubic (bcc) unit cell contains 2 water molecules.

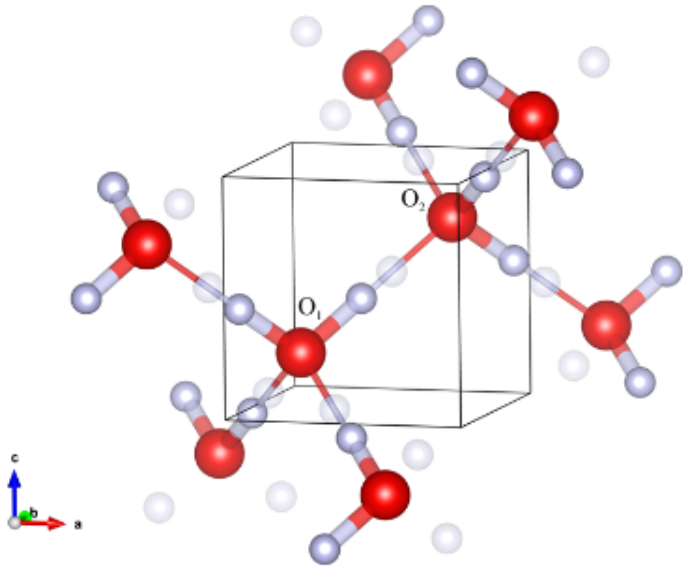

Figure 4.2. Fraction of the local structure of ice VII. Only one of the two hydrogen-bonded networks is shown. Oxygen and hydrogen atoms are in red and white colors respectively. Solid black lines indicate the cubic $Pn\bar{3}m$ unit cell. Partially transparent spheres are the two unoccupied hydrogen sites around each oxygen atom, in accordance with Bernal-Fowler ice rules. Figure taken from: *Structure and disorder in ice VII on the approach to hydrogen-bond symmetrization, Guthrie et al, Phys. Rev. B 99, 184112 (2019) [156]*

Before approaching the symmetrized phase X, the pressure-temperature phase space is occupied by the high-density ice phases VII and VIII. In all the dense ice phases, VII, VIII and X, each oxygen atom is surrounded by eight others, and four of them belong to the same sublattice that is interpenetrated but not interconnected with that of the remnant four water molecules. Each molecule can have six possible orientations in the unit cell that satisfy the ice rules.

Ice VII-VIII are related by an order-disorder transition. Ice VII is characterized by orientational proton disorder and it has a cubic (space group $Pn\bar{3}m$) paraelectric structure. Indeed, due to proton disorder, the average unit cell, measured through crystallographic experiments, differs from the real (instantaneous and local) structure. The interest in ice VII is not purely academic, since scientists hypothesize that it may comprise the ocean floor of Europa as well as extrasolar planets that are largely made of water. Fig. 4.2 depicts a fraction of the local structure of phase VII, where only one of the two hydrogen-bonded networks is shown.

In the disordered phases, the ice configurations that arise from the combination of the six different orientations that a water molecule can take in the structure according to the Bernal-Fowler ice rules, are very close in energy and can be regarded as degenerated. Ice VII is a prototype of a disordered structure, and depending on the number of water molecules present in the cell, several hydrogen arrangements are possible, resulting in a residual entropy, estimated by Pauling. The number of

Atom	Wyckoff position	Coordinates +(1/2,1/2,1/2)
O	8e	(0,0,z+1/8) (0,1/2,z+3/8) (1/2,0,-z+5/8) (1/2,1/2,-z+3/8)
H	16h	(0,y+3/4,z+1/8)(1/2,-y+3/4,z+5/8)(-y+1/4,1/2,z+3/8) (y+1/4,0,z+7/8) (1/2,y+3/4,-z+5/8) (0,-y+3/4,-z+1/8) (y+1/4,1/2,-z+3/8) (-y+1/4,0,-z+7/8)

Table 4.2. Space group ($I4_1/amd$) and Wyckoff positions of ice VIII. The tetragonal unit cell contains 8 water molecules.

possible hydrogen configurations is $M = W^N$, where N is the number of oxygen atoms in the piece of ice. The ice rule reduces W from 4 (there are 2N hydrogen atoms with two possible locations) to $W \simeq 1.5$.

In the case of degenerate structures, the resulting residual entropy can be easily accounted for. If we call F the free energy of each hydrogen configuration (identical for each hydrogen atom arrangement in the degenerate case), we can express the free energy of phase VII as

$$\tilde{F} = F - k_b T N \log W \tag{4.1}$$

Conversely, in the most general case, where the ice configurations are not degenerate, it can be easily shown that the free energy can be written as in Eq. (4.2)

$$\tilde{F} = \sum_i \frac{e^{-\beta F_i}}{Z_F} F_i + \frac{1}{\beta} \sum_i \frac{e^{-\beta F_i}}{Z_F} \log \frac{e^{-\beta F_i}}{Z_F}, \tag{4.2}$$

where F_i is the free energy of the i-th configuration and $Z_F = \sum_i e^{\beta F_i}$. It is straightforward to demonstrate that this expression reduces to Eq. (4.1) in the degenerate case.

Ice VII orders under pressure, above 2.1 GPa and on cooling below 278 K. The ordered solid was named ice VIII, its lattice can be obtained by slightly distorting the cubic cell of phase VII into a tetragonal parallelepiped followed by a displacement of one of the two sublattices in the $\vec{c}$ axis direction. In this phase, the two sublattices of ice VIII have opposite dipole moments, resulting in an antiferroelectric structure that belongs to the $I4_1/amd$ space group.

The tetragonal unit cell of ice VIII with 8 water molecules is shown in Fig. 4.3 (**a**). Fig. 4.3 (**b**) contains a bigger supercell needed for computational reasons elucidated in Sec. 4.4.1.

Tab. 4.2 reports the symmetry properties of phases VIII. We can notice that while the atomic positions of ice X in Tab. 4.1 are fixed by symmetry, with the proton lying in the midpoint of the two neighboring oxygen atoms, the lower symmetry of phase VIII leaves free parameters in the atomic coordinates.

4.3 Experimental observations

The VII/VIII phase boundary has been studied extensively using Infrared and Raman spectroscopies [16, 25, 157] in a 10-300 K temperature range and in the pressure region of 2-60 GPa and by neutron scattering experiments [158, 159]. The

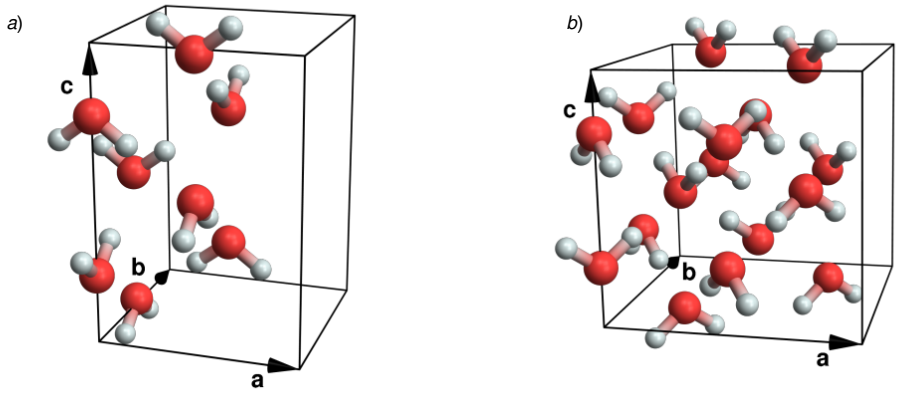

Figure 4.3. Phase VIII structure. Panel **a** shows the unit cell of ice VIII, a tetragonal cell with 8 water molecules. Panel **b**; a supercell of phase VIII containing 16 water molecules, commensurate with a supercell of phase X. The latter structure is the one used in the numerical simulations. White and red atoms are oxygen and hydrogen atoms respectively. The tetragonal lattice vectors $\vec{a}, \vec{b}, \vec{c}$ are reported.

experimental high-pressure phase diagram of ice taken from [16] is reported in Fig. 4.17 (**a**). The domain of stability of phases VII and VIII are very well established. By looking at the Clapeyron slope $\frac{dT_c}{dP}$, where T_c is the transition temperature, three different regimes can be identified. At low pressure, from 2 to 15 GPa the transition is driven by pure thermal effects where the transition temperature is almost pressure independent. In this region, the residual entropy of the disordered phase VII of Eq. (4.2) stabilizes this phase with respect to the ordered one. In the intermediate region, pressure and temperature are linearly dependent and here the system can be described with a pseudo-spin model [16]. Finally, in the third regime, the critical temperature drops abruptly to zero and the transition is dominated by quantum effects. At higher temperatures, a superionic VII" phase is observed [160].

While the VII/VIII transition is quite clear and well understood, the symmetrization of the hydrogen bond length is a long time debated issue. Theoretical studies [22, 23] predict that the energy barrier between a double-well potential description for the proton motion along the oxygen-oxygen ($O-O$) line is depressed with increasing pressure up to a critical pressure where the two potential minima eventually merge to a single minimum shape where the proton occupies the midpoint of the O-O distance so that the molecular solid becomes an ionic material at high pressure. From the experimental side, no definitive evidence of the transition is present at the moment. The optimal tool to probe atomic positions within a crystal would be X-ray diffraction (XRD). The intensity of the diffracted peaks is proportional to the arrangement of atoms in the entire crystal

$$I_{hkl} \propto |F_{hkl}|^2$$

where $\{h, k, l\}$ indicate the lattice planes and the structure factor

$$F_{hkl} = \sum_{j=1}^{m} N_j f_j e^{2\pi i(hx_j + ky_j + lz_j)}$$

sums the result of scattering from all of the m atoms in the unit cell. N_j is the fraction of equivalent positions that are occupied by atom j. So, the scattered intensity is determined by where and what the atoms are, in particular, the scattering factor f_j quantifies the efficiency of X-ray scattering and it is proportional to the electronic density of the given atomic species. Unfortunately, hydrogen atoms have only one electron and this results in a very little cross-section that makes them practically invisible to the X-ray diffraction experiments.

Conversely, XRD is suitable to sample the oxygen atoms sublattice in the unit cell thanks to their bigger cross-section. Different X-ray diffraction experiments [161, 162, 163] on phase VII under pressure revealed that the ice structure retains a body-centered close-packed oxygen sublattice from above 2.1 GPa up to over 100 GPa. XRD consents to analyze the equation of state of ice under pressure providing indirect evidence for the hydrogen bond symmetrization through abrupt changes in the EOS. However, direct evidence for the transition is still lacking. Neutron diffraction experiments [164], where the peak intensity is no more proportional to the electronic density, would be better suitable to probe the position of deuterium atoms [165]. Low-pressure neutron diffraction [166] confirmed the interpenetrated H-bond network structure and subsequent higher pressure measurements revealed little changes in the molecular geometry up to 20 GPa. However, the lack of neutron diffraction data at even higher pressure prevents the direct determination of the proton positions approaching the symmetrization.

The evolution of ice under pressure is even more complicated and "mysterious" because of the complication of the hydrogen-bond variation. Some quantum-mechanical calculations proposed the possibility of the existence of an intermediate phase prior to the VII/VIII to X transition [22, 167]. At the same time, several experimental anomalies, both in X-ray diffraction and spectroscopic experiments, were reported indicating a number of transitions occurring at moderate pressure ranges [157, 163, 168, 169, 170]. The hypothesis of the intermediate phase mainly rely on the presence of very broad peaks in the vibrational spectra at the critical pressure whose intensity enlarges (and they become narrower) increasing the pressure in the phase X [16]. Recently, neutron diffraction experiments [166] detected anomalies from 13 GPa that cannot be explained by conventional structure networks above 26 GPa, where structures with proton localization on the octahedral interstitial sites in the oxygen sublattice seem to fit well the data.

Hints for the symmetrization of the hydrogen bond can be detected from changes in spectroscopic observations. All the experimental and theoretical studies agree with a second-order character of the phase transition [16]. Hence, vibrational spectroscopies are suitable tools to identify the soft mode behavior of the transition. Unfortunately, experimental measurements are not easy to perform: Raman spectroscopy is a sensitive probe for phase transitions but Raman intensities for H_2O ices are very weak making it hard to measure signals under pressure. Opposite difficulties arise for infrared spectroscopy; the O-H stretching modes (the ones which soften) are so intense that the absorbance maxima saturate resulting in such a broad

band that the peak position cannot be identified with sufficient precision. Finally, the transmission through the diamond anvil cell, necessary to achieve high pressure, is around 0.05% in the frequency range from 1800 to 2400 cm^{-1}, a critical interval to study the softening of the stretching modes. The diamond absorption issue is overcome in [171], where a striking number of Fermi resonances between the soft antisymmetric stretching mode and the rotational mode, bending mode and their combination are observed during the softening of the mode.

Direct evidence of the hydrogen bond symmetrization is not available and the existence of an intermediate phase proposed in several works, both theoretical and experimental, is not clear. In the following, I elucidate on this problem by using the SSCHA theory [32] and its dynamical extension [33]. The phase diagram is investigated in Sec. 4.4, while in Sec. 4.5 I discuss spectroscopic simulations to give more reliability to the obtained results.

4.4 Phase Diagram

The SSCHA theory described in Chapter 2 provides the inclusion of quantum anharmonic effects in the description of nuclear vibrations. However, a qualitative analysis of some features, like the order of the phase transition can be done at a classical level at T=0 K, where thermal and quantum effects are completely neglected and numerical simulations are much faster. In this section, I will describe first the classical phase diagram and then the full quantum anharmonic case.

The Born-Oppenheimer energy landscape is represented by the deep neural network of Ref. [172] (fitted on the PBE functional approximation of DFT) described in Sec. 2.9.2.

4.4.1 The classical phase transition at zero temperature, namely neglecting quantum and thermal fluctuations

Since second-order phase transitions require a continuous transformation of one phase into the other, it is useful to have commensurated structures for the phases of interest. To this purpose, it is better not to use the conventional unit cell of ice VIII in Fig. 4.3 (**a**) but the supercell shown in Fig. 4.3 (**b**), that is a $\sqrt{2} \times \sqrt{2} \times 1$ of the unit cell obtained by summing and subtracting the lattice vectors in the basal plane. This unit cell, containing 16 water molecules, is commensurated to a $2 \times 2 \times 2$ supercell of the body-centered cubic unit cell of phase X, and a disordered structure for phase VII with the same number of molecules can be easily generated.

In this section, I analysis the phase transition in the classical approximation at T=0 K, where both quantum and thermal fluctuations are neglected, in the next sections, I will discuss the classical phase transition at finite temperature, where thermal fluctuations are included within the SSCHA but I keep neglecting quantum effects.

Ice structures for phases VIII and X have been relaxed at fixed external pressure and then I compared their Gibbs free energy. Notably, at T=0 K, the difference between the Gibbs free energy $G(P,T) = H - TS$ and the Hentalpy $H = U + P\Omega$ vanishes (also for the disordered phase VII, that has a non zero residual entropy S

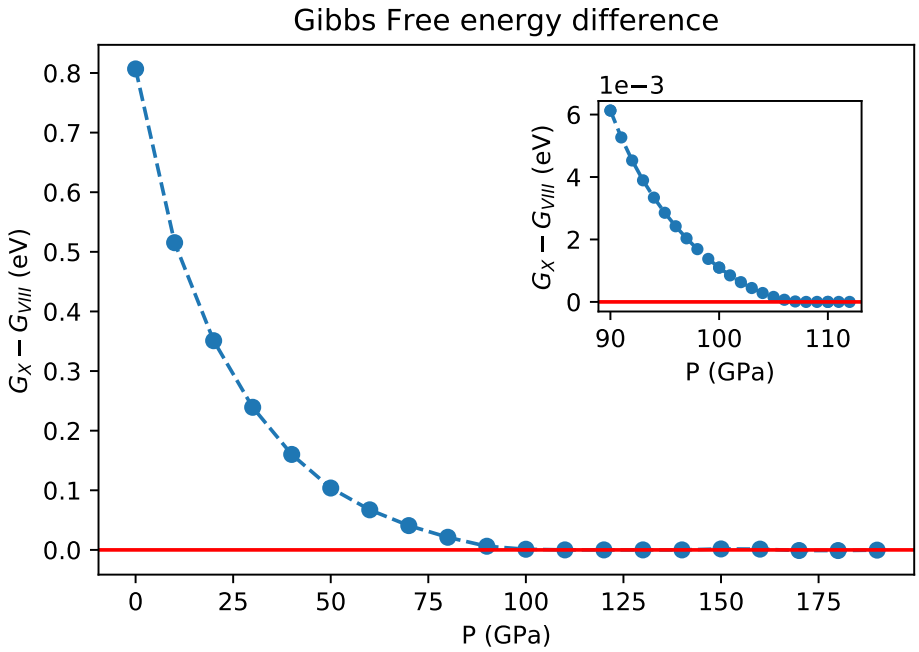

Figure 4.4. Gibbs free energy difference per molecule between phase VIII and phase X in the classical approximation at T=0 K. In the inset, a finer pressure sampling in the region close to the phase transition is shown. Ice structures containing 16 water molecules were used. The red line is a guide for the eye that indicates zero Gibbs free energy difference, the phase transition is located at $P \simeq 107$ GPa.

at zero Kelvin that gets anyway multiplied by T=0). Details on the simulations are in Appendix H.

I report the Gibbs free energy difference as a function of pressure in Fig. 4.4. The inset is a zoom around the region close to the phase transition, where a finer range of pressures is sampled. The behavior of the Gibbs free energy is that typical of a second-order phase transition, where one structure continuously transforms into the other so that the Gibbs free energy difference approaches zero smoothly. Looking at the inset of Fig. 4.4, I can locate the classical phase transition at $P_c \simeq 107$ GPa.

Experimentally, it is really hard to detect signatures of the phase transitions. X-ray diffraction cannot detect the position of the hydrogen atoms but it is well suited to identify the oxygen atoms that allows to determine the volume of the structure and consequently the equation of state. Changes in the EOS can be used to detect the phase transition as in Fig. 4.5 (**a**). The pressure as a function of volume for phases VIII and X (blue circles and red squares) are shown. While in phase X the positions are fixed by symmetry, in phase VIII they are free to evolve and for $P >= 107$, the EOS of the two phases are identical, indicating the occurrence of a second-order phase transition. To give further evidence of the phase transition, I report in Fig. 4.5 (**b**), the order parameter, $\Delta = \frac{l_{OO}}{2} - l_{OH}$, as function of pressure. The symmetrization of the hydrogen bond occurs when $\Delta = 0$ at P = 107 GPa, confirming the change in the EOS in Fig. 4.5 (**a**).

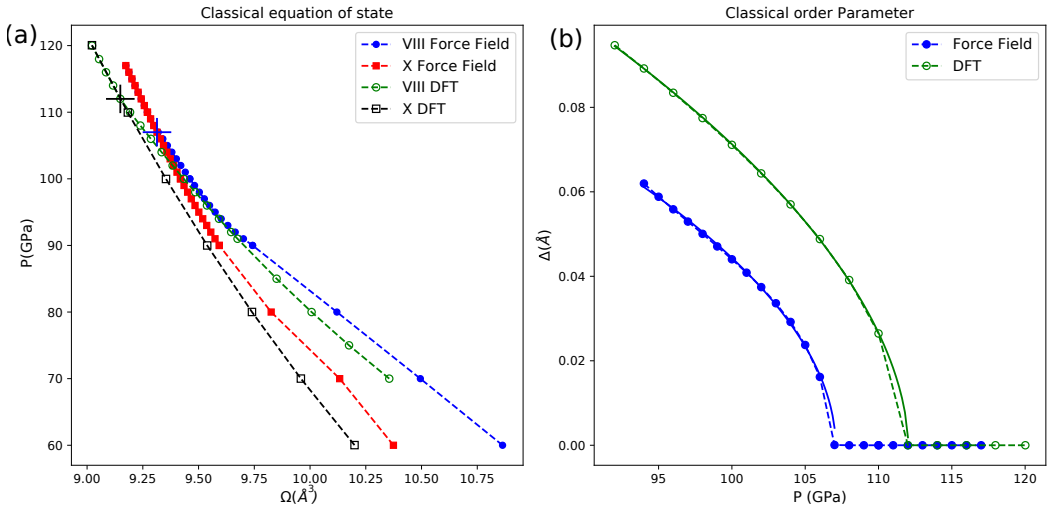

Figure 4.5. Structural parameters of the phase transition in the classical approximation at T=0 K. Panel **a** : Equation of state of phases VIII and X as a function of pressure. Pressure as a function of the equilibrium volume per molecule for phase VIII and X are shown as blue filled circles and red filled squares respectively. Green empty circles and black empty squares are data computed *ab initio* with the PBE approximation of DFT. The blue and black crosses indicate the phase transition at the force-field and *ab initio* level respectively. Panel **b**: Order parameter, defined as $\frac{l_{OO}}{2} - l_{OH}$ as a function of pressure when the force field (blue filled circles) or the DFT functional (green empty circles) are used.

The use of the force field necessarily introduces a source of error with respect to the pure *ab initio* approach. In order to quantify such an error, I repeat the same classical calculation at the *ab initio* level, using the PBE functional within Quantum ESPRESSO [86]. The green empty circles and the black empty squares in Fig. 4.5 (**a**) are the EOS for phases VIII and X, respectively computed *ab initio*. The corresponding order parameter is in Fig. 4.5 (**b**).

At the *ab initio* level, the phase transition occurs at $P_c = 112$ GPa and the use of the force field underestimates the critical pressure of about 5 GPa.

Another hint for the second-order character of the phase transition can be inferred from the energy profile. By starting from the relaxed ice X structure at a given pressure, I moved the position of all the hydrogen atoms simultaneously far from equilibrium, along the OO. The energy of each configuration as a function of the distance of the hydrogen atom from the equilibrium position for several pressures, inside and outside the stability domain of ice X is reported in Fig. 4.6. The resulting fourth-order potential is typical of second-order phase transitions, outside the stability region, for $P < 107$ GPa, the midpoint is unstable and the energy profile presents two well distinct minima. By increasing the pressure, approaching the phase transition, the double wells merge into a single minimum and at the critical pressure the energy profile is flat, since the second-order term in the potential expansion vanishes.

Actually, this cannot be considered the real energy profile since we are dealing with a multidimensional problem. The unstable modes, that lead the transition from

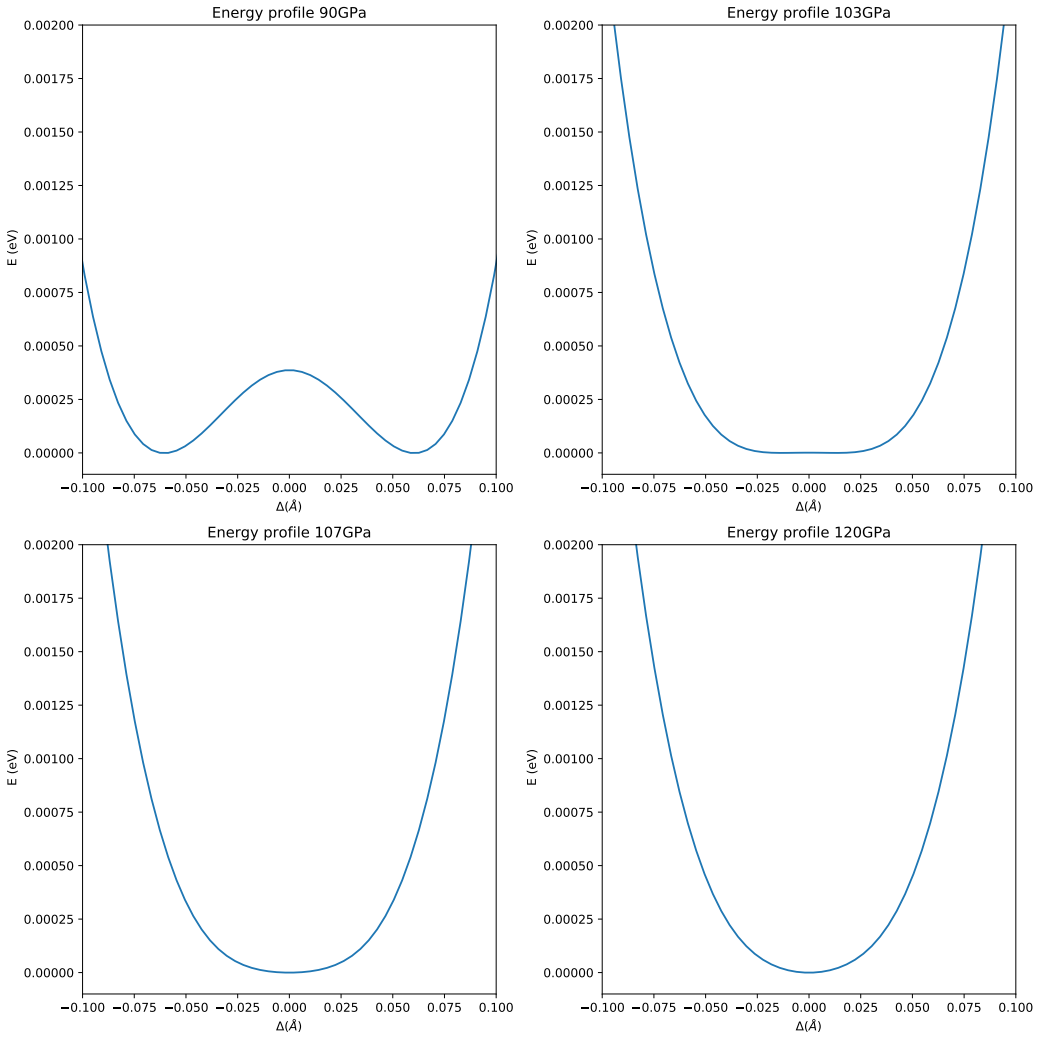

Figure 4.6. Energy profile at fixed pressure for phase X in the classical approximation at T=0 K. From left to right and top to bottom, the pressure is increased from below to above the critical pressure P_c=107 GPa. Starting from the midpoint, all the hydrogen atoms in the simulation cells have been moved along the OO direction, far from the equilibrium positions. The parameter Δ indicates the displacement with respect to the symmetric position. The energy profile changes from a single well to a double-well potential going through the region of instability of phase X.

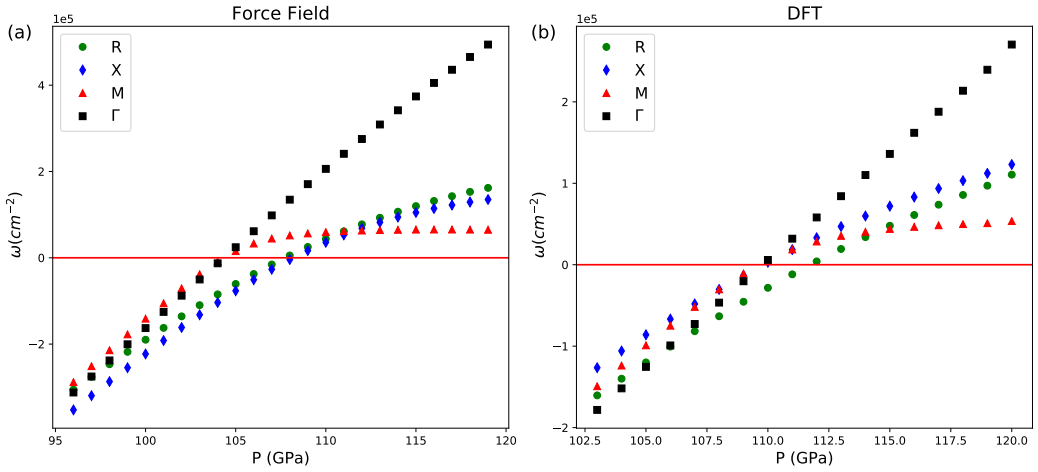

Figure 4.7. The figure shows the phonons in the **q**-points commensurated to the 2×2×2 supercell of phase X as a function of pressure in the classical approximation at T=0 K. For each pressure, the system is forced to stay in the symmetric phase X and out of all the $3N$ modes, I show the ones that are unstable below the critical pressure. Panel **a** describes the phonons computed with the force field (P_c =107 GPa) and in panel **b** there are the phonons at the DFT-PBE level (P_c=112 GPa). In both the approaches, phonons at $R = (\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$ are the green circles, blue diamonds indicate the $X = (\frac{1}{2}, 0, 0)$ point. Instead, the phonons in $M = (\frac{1}{2}, \frac{1}{2}, 0)$ are the red upper triangle and the black squares stand for the phonons at Γ. Phonons in R and Γ are triple degenerate while those in X and M are double degenerate.

one phase to the other will be stretching modes along the OO direction. So, Fig. 4.6 is a reasonable picture to grasp the qualitative feature of the phase transition. For instance, in the unstable region, the curvature is negative in the symmetric position of phase X leading to imaginary phonons. These are characteristic of second-order phase transitions, where the phonons that connect the two phases go soft at the critical pressure and then become imaginary in the unstable region.

The imaginary modes, that drive the phase transition, can be observed if we force the system to stay in the symmetric phase X for each pressure. In Fig. 4.7 (**a**), I report the frequency squared of the phonon modes observed in the **q**-point of the Brillouin zone commensurated to the 2×2×2 supercell of phase X as a function of pressure computed with the DNN of Sec. 2.9.2.

The $R = (\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$ point is commensurate to the transformation from phase X to VIII and we expect an instability there, with the zero-energy located exactly at the critical pressure $P_c = 107$ GPa. As further evidence, I check that the displacement of the atoms along this unstable mode followed by a structure relaxation leads exactly to phase VIII.

At the same pressure, the phonons in the $X = (\frac{1}{2}, 0, 0)$ point (not commensurate with the expected transformation to phase VIII) become imaginary. The relaxation of the structure displaced along that mode leads to a tetragonal structure belonging to $P4_12_12$ space group, which is similar to phase VIII. This raises the question about the existence of a competing phase. Energetically, phase VIII is slightly favorite with respect to the structure related to the instability in point X. By lowering again

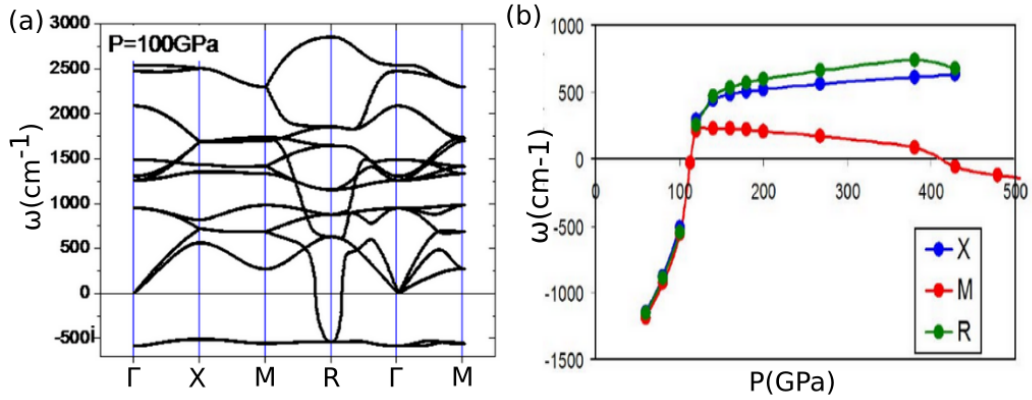

Figure 4.8. First-principles simulations of the structural instability in phase X taken from *Dynamical Instabilities of Ice X; Caracas, Phys. Rev. Lett. 101, 085502 (2008).* **a**: Dispersion of the phonon band in ice X at P=100 GPa. **b**: Pressure variation of the lowest-frequency modes in ice X in $R = (\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$, $M = (\frac{1}{2}, \frac{1}{2}, 0)$ and $X = (\frac{1}{2}, 0, 0)$. The study detected also an instability above 400 GPa in the M point not investigated in my work.

the pressure, in the region of instability of phase X, we can notice that also the modes in M and Γ become imaginary simultaneously at P=104 GPa. For instance, the relaxation of the structure along the M mode leads to tetragonal structures belonging to $I4_1$ space group. Actually, below the critical pressure, a whole phonon branch becomes imaginary and the instability for each **q** point is related to the existence of the disordered phase. This picture can be found also at the DFT-PBE level, where the phonon in the same **q** points are reported in Fig. 4.7 (**b**). The main difference is that in DFT the phonon at R is the only one that goes to zero at the critical pressure (112 GPa) while the others become imaginary simultaneously 2 GPa below. Also in the *ab initio* approach, a whole branch becomes imaginary by decreasing the pressure (not shown here). These results are in agreement with previous theoretical calculations in the classical approximation at T=0 K performed by Caracas in Ref. [167], reported in Fig. 4.8.

This concludes the classical analysis of the phase transition, where I have shown several elements to confirm its second-order character. The same kind of conclusions can be drawn about the disordered phase VII, where the transition toward the symmetric phase occurs exactly at the same pressure as for phase VIII. Details on the study of phase VII in the classical approximations are in App. I.

4.4.2 Quantum phase transition at T=0 K and T=300 K

The inclusion of quantum anharmonic effects within the SSCHA framework completely overturns the classical T=0 K picture. All the details about the SSCHA simulations are in App. H. Starting from phase VIII, I relaxed the crystal structures at zero Kelvin and at finite temperature, T=300 K, for increasing pressures until reaching the symmetric phase and beyond by using the SSCHA technique [32]. The equations of state are reported in Fig. 4.9, critical points for all the temperatures

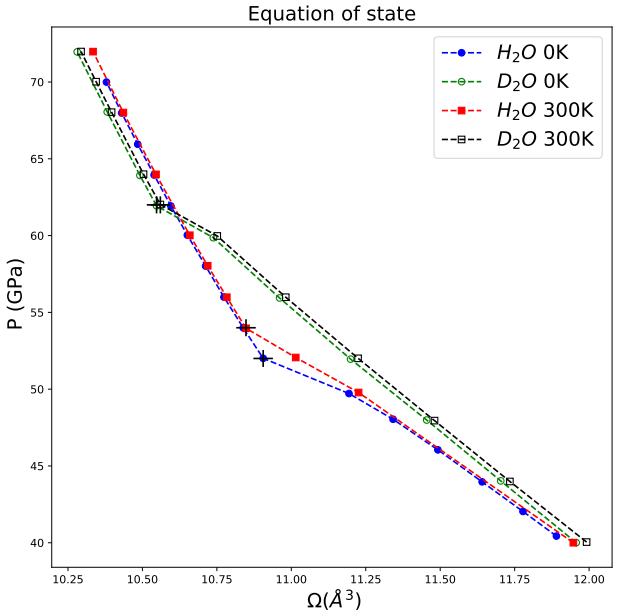

Figure 4.9. SSCHA Equation of state with quantum and thermal fluctuations included for phase VIII. Blue filled circles and red filled squares are hydrogenated ice at T=0 K and T=300 K respectively. Empty symbols, instead, describe deuterated ice at T=0 K (green circles) and T=300 K (black squares). The transformation to the symmetric phase can be spotted by the change in the EOS, where the black crosses are located.

and isotopes inspected are indicated with black crosses. The very tiny temperature dependence of both phases VIII and X stands out, notably, the only relevant difference for hydrogenated ice is for the critical pressure, where, a very tiny increase from $P_c = 52$ GPa to $P_c = 54$ GPa from T=0 K to T=300 K is detected. This difference in the critical pressure is probably an artifact of the discretization of the pressures calculated (see App. H).

Quantum effects heavily renormalize the critical pressure with respect to the purely classical case described in Sec. 4.4.1. A reduction of the critical pressure of about 30 GPa due to quantum effects is known from numerical simulations [22, 173]. However, within the SSCHA framework, an even bigger difference, of about 55 GPa at T=0 K, is observed.

Our data manage to well reproduce the isotopic shift of 10 GPa between hydrogenated and deuterated ice observed experimentally [16]. Empty symbols in Fig. 4.9 describe the EOS for D_2O ice VIII. Its critical pressure $P_c = 62$ GPa is temperature independent consistently with the experiment.

The error introduced by the SSCHA approximation is expected to lead to an overestimation of the transition pressure. The SSCHA theory is based on variational principle on the free energy of the system, so the Gaussian description of the SSCHA wavefunction describes well the minimum of the double-well potential of phase VIII, with hydrogen atoms well-localized close to the relative oxygen atom. Conversely, we expect a bigger error in the description of the single well potential in the symmetric phase. Such overestimation is instead harmless when dealing with pressure differences

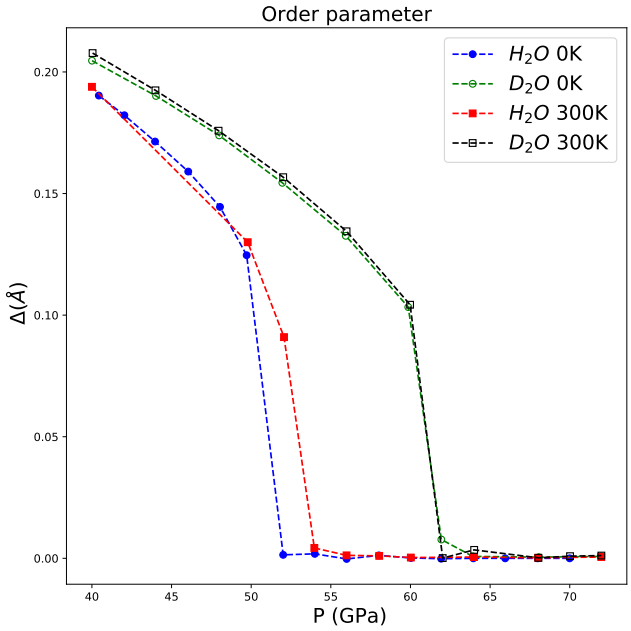

Figure 4.10. Order parameter, $\Delta = \frac{l_{OO}}{2} - l_{OH}$ computed within the SSCHA, with thermal and quantum fluctuations included. Filled blue circles and red squares are data for H_2O ice VIII at T=0 K and T=300 K, respectively. Instead, empty green circles and black squares refer to deuterated ice. The order parameter approaches zero at the phase transition, where phase VIII turns into the symmetric phase X.

like for the isotopic shift that is perfectly reproduced.

If we take into consideration the 5 GPa underestimation of the critical pressure by the DNN with respect to DFT-PBE, we recover an almost perfect agreement with the experimental data. The residual error with respect to the experiment is probably due to the exchange-correlation of the PBE functional used to fit the DNN.

As for the classical T=0 K case, the phase transition can be spotted more easily by looking directly at the order parameter in Fig. 4.10. The points where it vanishes for deuterated and hydrogenated ices at T=0 K and T=300 K correspond to the changes in the slope of the EOS of Fig. 4.9. For H_2O ice, a tiny temperature effect can be spotted as for the EOS, while in D_2O the order parameter seems really temperature insensitive. After the phase transition, the order parameter stays close to zero as expected for the symmetric phase X. The tiny and negligible oscillations are due to the fact that the phase X symmetries are not imposed above the critical pressure.

The critical pressure can be identified also by looking at displacement parameters, from diffraction experiments one can extract the anisotropic displacement parameters (ADP), related to the Debye-Waller factor introduced to keep into consideration thermal fluctuations. The anisotropic displacement parameters matrix is formally defined for each atom s as in Eq. (4.3),

$$ADP_s^{\alpha\beta} = < u_s^\alpha u_s^\beta > = \frac{\hbar}{N_{\mathbf{q}} M_s} \sum_{\mu \mathbf{q}} \frac{1 + 2n_\mu(\mathbf{q})}{2\omega_\mu(\mathbf{q})} \epsilon_\mu^{s\alpha}(\mathbf{q}) \epsilon_\mu^{s\beta}(\mathbf{q}), \tag{4.3}$$

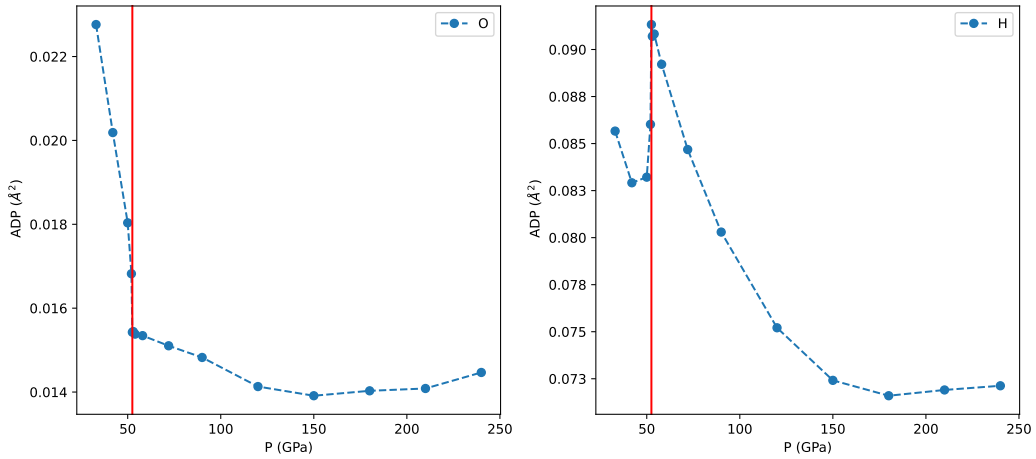

Figure 4.11. Trace of the anisotropic displacement matrix for oxygen (panel **a**) and hydrogen atoms (panel **b**) at T=300K in phase VIII within the SSCHA approximation. The red vertical line indicates the critical pressure at the given temperature. The ADPs in the disordered phase, not reported here, show the same behavior as in the ordered structure.

where $N_{\mathbf{q}}$ is the number of unit cells considered. The trace of the ADP matrices for hydrogen and oxygen atoms are shown in Fig. 4.11. Here, I report the results for the ordered phase at T=300 K. However, the same results can be drawn for the disordered structure, whose displacements factors are almost identical to those of phase VIII. In the latter, the ADPs along the x and y directions are degenerate, with a slightly different value along the vertical axis, conversely, disorder makes them different along the three Cartesian directions. Fig. 4.11 shows that these parameters can be adopted to spot the critical pressure from experiments. For the oxygen atoms, their value falls very rapidly approaching the transition pressure and then stabilizes in the symmetric phase, while, more interestingly, the displacement factors for the hydrogen atoms first decrease and then increase very rapidly close to the critical pressure, to decrease again in phase X.

First-order phase transition differentiates themselves from second-order ones for the existence of hysteresis cycles, that is if the state of the system has memory of what happened in the past, i.e if it reached that point upon compression or decompression. In the compression runs, the starting guess for each pressure is the minimized structure at the preceding lower pressure, instead in the decompression, the starting guess is that at the preceding higher pressure.

All the discussion in Sec. 4.4.1, as well as the continuous transformation between the two phases in the quantum regime in Fig. 4.9 and Fig. 4.10, hinted at a second-order character for the phase transition. Furthermore, I rule out the possibility of a first-order phase transition by excluding the presence of the hysteresis cycle in Fig. 4.12. There, the covalent and hydrogen bond lengths at T=0 K are reported as a function of pressure, in the region close to the VIII-X phase transition. The two bond lengths merge at the phase transition as expected from Fig. 4.10 and stay identical in phase X.

Sometimes, when studying phase transitions, a weekly first-order character can

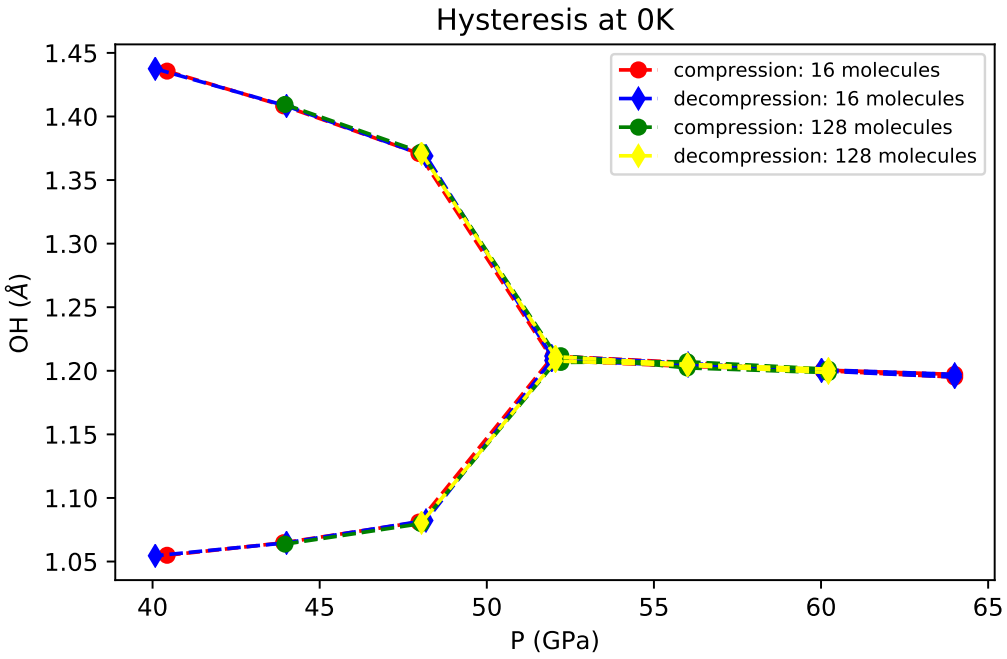

Figure 4.12. Hysteresis cycle for phase VIII close to the phase transition at T=0 K within the SSCHA framework. Hydrogen and covalent bond lengths are plotted as a function of pressure, above and below the transition to the symmetric ice, where they are identical. Red circle and blue diamonds are compression and decompression run in the unit cell with 16 water molecules. To check the convergence with a supercell, we report compression and decompression in a supercell with 128 water molecules: green circles and yellow diamond respectively.

appear because of the use of too small supercells. This is not the case, and moreover, to further check the convergence of the results, I include the calculation in a bigger supercell with 128 water molecules. Green circles and yellow diamonds in Fig. 4.12, show the hydrogen and covalent bond lengths in the compression and decompression run respectively in the supercell. The results are almost identical to those in the unit cell with 16 water molecules, where convergence is already granted. Consequently, in the following, all the results, when not specified, are in the cell with 16 molecules.

The study of the phase transition can be readily extended to the disordered phase VII. In principle, one should analyze all the possible nonequivalent hydrogen arrangements in the given simulation cell as depicted in the last part of Sec. 4.2. The large computational cost of the SSCHA method prevents the simulation of such a big number of structures, so I took 3 random configurations and I relaxed their structure for different pressures until reaching the symmetrized phase as for phase VIII.

In Fig. 4.13, I report the structural parameters, namely the EOS and the order parameter for phase VII at T=0 K. Sets 1-3 indicate the three different hydrogen arrangements. As we can see, the difference between the three structures is negligible far from the phase transition, while some discrepancies arise close to the critical pressure. Indeed, for two structures $P_c = 52$ GPa, while the third one reaches

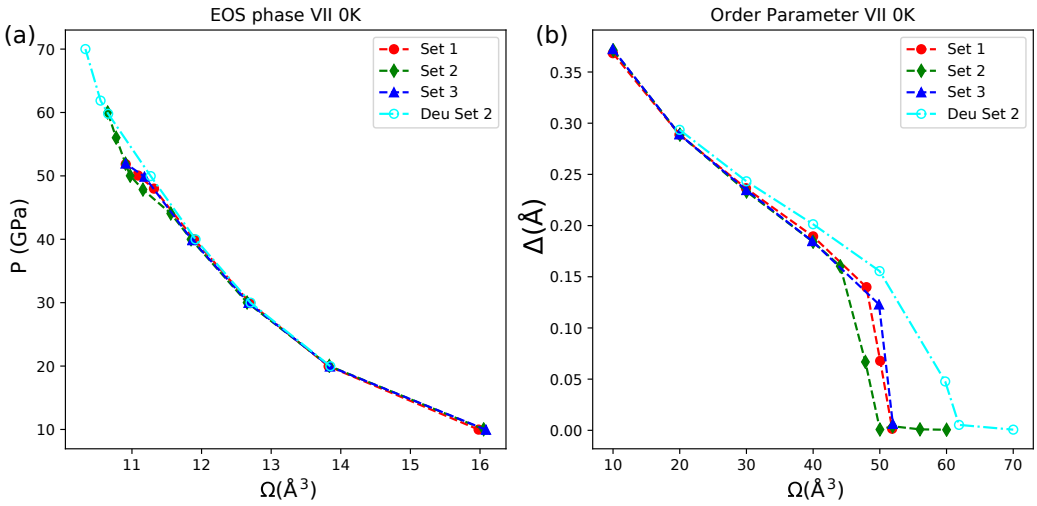

Figure 4.13. Structural parameters for ice VII at T=0 K within the SSCHA. Three different instances of proton decoration of the oxygen structure for H_2O ice are reported: red circles, green diamonds and blue upper triangles. The equation of states is in panel **a** and the order parameter, $\Delta = \frac{l_{OO}}{2} - l_{OH}$, is in panel **b**. Data for deuterated ice for the analysis of the isotopic shift are added as cyan empty circles.

symmetrization at P=50 GPa. The difference between the transition pressures is small compared to the usual experimental uncertainties in detecting the phase transition, so, we can safely take their average to give an estimate of the critical pressure $P_c = 51 \pm 2$ GPa.

The isotopic shift of 10 GPa is present also in the disordered phase at T=0 K. The EOS and order parameter of deuterated ice in a specific hydrogen arrangement (Set 2) is added for comparison in Fig. 4.13.

As for the ordered phase VIII, temperature affects negligibly the critical pressure of the disordered structures. Fig. 4.14 illustrates the EOS and ordered parameter of the same three disordered configurations described above at T=300 K. There, the critical pressures range from P=52 GPa to P=54 GPa, showing again, a little difference with respect to the experimental uncertainties.

Most experimental works adopt a linearization of the Vinet et al. EOS [174] to detect changes in its slope that suggest a phase transition. In Fig. 4.15, I show the natural logarithm $\ln(H)$ of the H function defined in Eq. 4.4,

$$H(x) = \frac{Px^2}{3(1-x)} = B_0 \exp\left\{\frac{3}{2}(B_0' - 1)(1-x)\right\}, \tag{4.4}$$

as a function of $1-x$, where $x = (\frac{V}{V_0})^{\frac{1}{3}}$. Its slope is related to the derivative of the bulk modulus with respect to pressure, B_0', and the offset is $\ln B_0$. The comparison with experimental results in Ref. [16], obtained through X-ray diffraction is quite good; the location of the change in the slope is in an optimal agreement, but the two curves are a shifted one with respect to the other, meaning a non-perfect description of the bulk modulus.

The structure of disordered ice shows some strange behavior during relaxation.

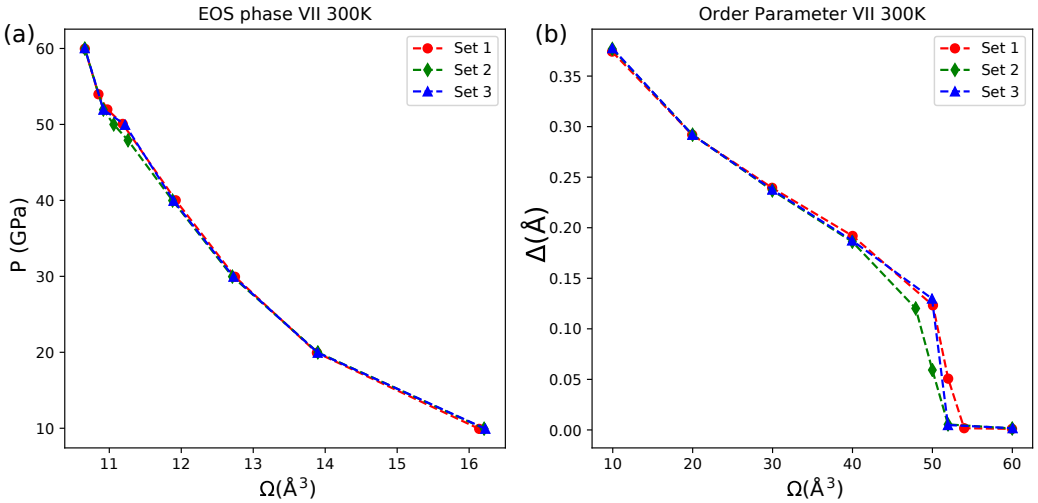

Figure 4.14. Structural parameters for ice VII at T=300 K within the SSCHA. Red circles, green diamonds and blue upper triangles are three different configuration of proton disordered ice H_2O structures. The equation of states is in panel **a** and the order parameter, $\Delta = \frac{l_{OO}}{2} - l_{OH}$, is in panel **b**.

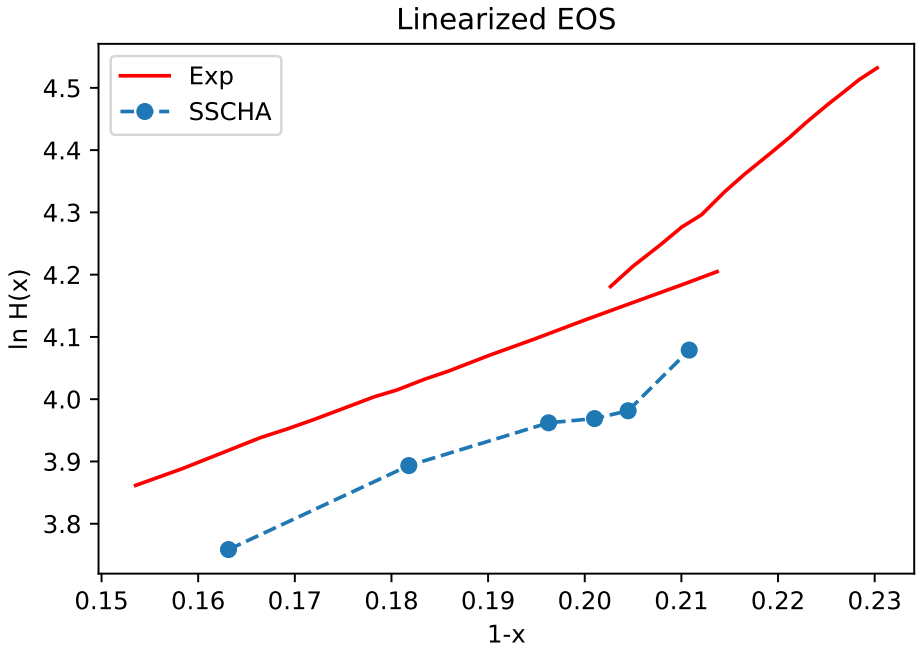

Figure 4.15. Linearized form of the Vinet et al. EOS [174]: $lnH(x)$ as a function of $1-x$, where $x = (\frac{V}{V_0})^{\frac{1}{3}}$. See text for further explanation. SSCHA simulations for H_2O ice VII at 300 K are compared to experimental data [16]. Changes in the slope in the linearized EOS indicate phase transitions.

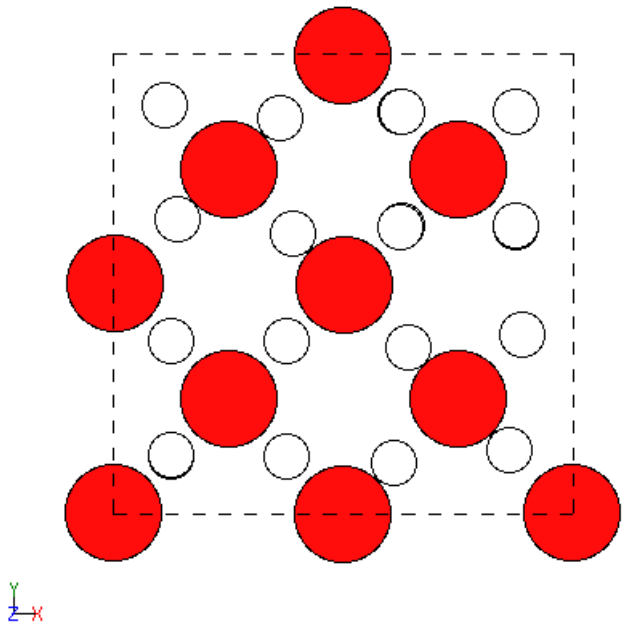

Figure 4.16. Front view of the relaxed structure of a disordered ice VII configuration at P=48GPa and T = 0 K within the SSCHA. The structure correspond to the hydrogen arrangements denominates Set 2 in Fig. 4.13. Some of the hydrogen bonds are symmetrized while others are not yet symmetric.

In a tiny pressure range close to the phase transition, some of the chosen hydrogen atoms arrangements display partial ordering. Fig. 4.16 shows a front view of the the relaxed structure at P = 48 GPa and T=0 K for the arrangement denoted Set 2 in Fig. 4.13. We can notice that some of the hydrogen bonds are fully symmetrized, while others are not. This situation persists at most in a range of 4 GPa before the phase transition, since at P=50 GPa phase X with all symmetrized bonds is reached. We argue if this structure, where not all the hydrogen bonds are symmetrized, can be related to the hypothesized disordered symmetric phase between phases VII and X. However, the very small pressure range in which this phase is observed probably rules out this possibility. Moreover, if we look at the Infrared absorption of Fig. 4.19 of Sec. 4.5, we see that the spectrum of this structure presents the same features as those of the purely disordered phase, with broader peaks but similar positions to those of ice VIII.

I have shown that quantum effects play a major role in the description of the phase transition. Within the SSCHA framework, it is possible to remove quantum effects, with a trick that consists in setting infinite masses for the nuclei (achieved by multiplying the physical ones by a factor 10^6), keeping instead only thermal fluctuations. The final results are indicated by the blue empty circles in Fig. 4.17 (**b**). Interestingly, now temperature plays an important role. The critical temperature decreases from $P_c \simeq 107$ GPa at T=0 K to $P_c \simeq 80$ GPa at T=300 K in a nonlinear way, where the stronger temperature dependence is in the low-temperature limit.

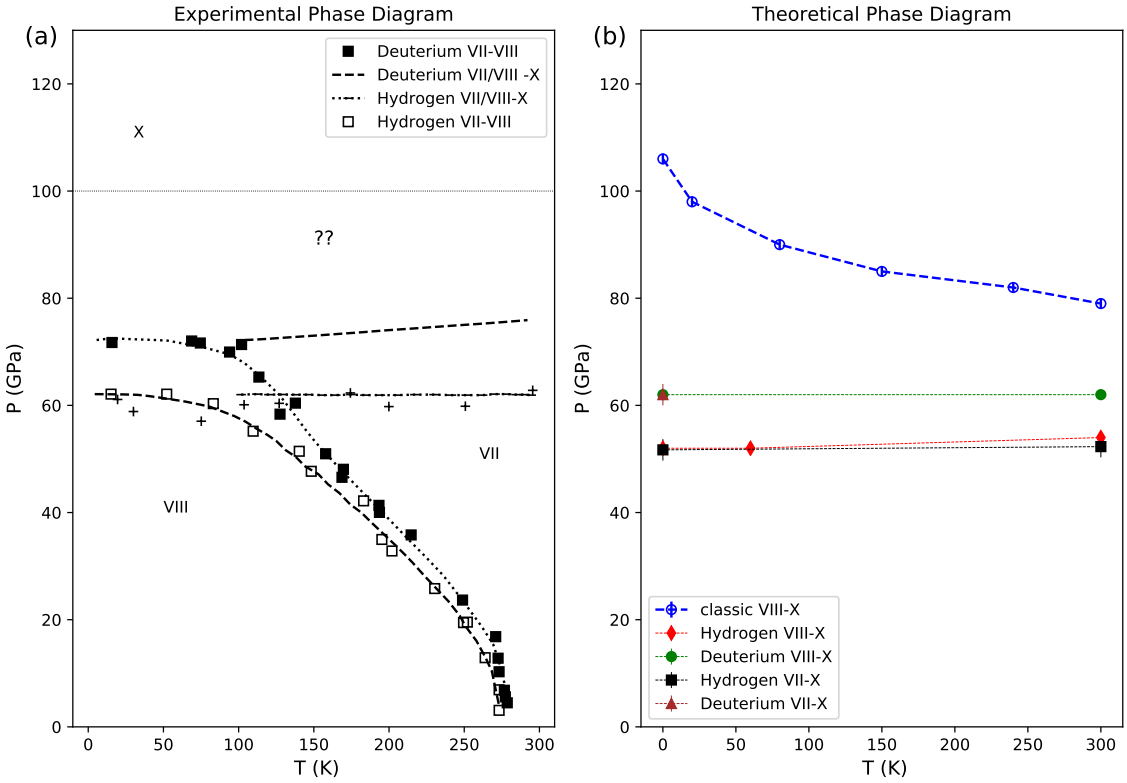

Figure 4.17. Panel **a**: Experimental phase diagram. Data are extracted from Pruzan et al. [16]. The stability domain of phases VIII and VII as a function of pressure and temperature is indicated. Open and closed squares are the transition lines for H_2O and D_2O ice, respectively [25], whereas dashed and dotted lines are fitted curves. Crosses are experimental data for the VII-X phase transition obtained through Raman scattering [26]. The horizontal dash-dotted and dashed lines are the expected VII-X transition lines for H_2O and D_2O ice, respectively. Experimentally, it is not clear if the transition to phase X (indicated above 100 GPa) is direct or if an intermediate state, indicated with ?? text in the figure, exists. Panel **b**: Theoretical phase diagram within the SSCHA framework. Only the phase boundaries between phase VIII/VII and X are shown. The transition VIII-VII is not reported. Red diamonds and black squares are the critical pressure for the VIII-X and VII-X transition respectively in hydrogenated ice. Deuterated VIII-X and VII-X transitions are indicated with green circles and brown upper triangles respectively. Empty blue circles stand for the critical pressures as a function of temperature in the SSCHA framework that includes thermal fluctuations but with removed quantum effects.

The critical pressure difference at T = 300 K is about 30 GPa as reported in some theoretical works [22, 173], justifying somehow the comparison of classical and quantum results with that pressure shift. Anyway, my results clearly show that it is not possible to adopt a constant shift but is it crucial to adopt a different one as a function of the temperature analyzed. Moreover, as I will show in Sec. 4.5 it is not even correct to use a constant pressure shift for fixed temperature. Convergence tests confirm that the unit cell with 16 water molecules is sufficient to give converged results for the thermodynamic case also in the classical approximation at finite temperature (I probed structures with up to 128 water molecules).

Direct comparison of Fig. 4.17 (**a**) and (**b**) confirms the good description of the high-pressure phase diagram within the SSCHA framework. The critical pressures are temperature independent in the quantum phase coherently with the experimental data, however, about 10 GPa reduction with respect to the experimental phase transition is observed, which can be explained by the different sources of error, like the employment of a force field description of the PBE functional as well as the adoption of the DFT methodology. However, we show that the DNN fitted potential shows a transition pressure lower of about 5 GPa with respect to PBE in the static case. If we correct our result with this constant shift, we get an almost perfect agreement between the SSCHA phase diagram and the experimental phase diagram. The isotopic shift of 10 GPa in Fig. 4.17 is well described within the SSCHA.

4.4.3 Conclusions

The study of the phase transition in the classical approximation at T=0 K (without thermal and quantum effects) provided useful insights on the second-order character of the phase transition; a continuous transformation of one phase into the other has been observed, both looking at the structural parameters, like the EOS or the order parameter, and at the Gibbs free energy. A fourth-order energy profile projected on a single dimension with a relative soft mode behavior is detected. The inclusion of thermal fluctuations in the classical approximation within the SSCHA framework reveals a strongly nonlinear temperature dependence of the critical pressure that is not observed in the quantum regime. The critical pressure decreases from $P_c = 107$ GPa at T=0 K to $P_c = 80$ GPa at T=300 K demonstrating that the approximation of the quantum case (where the critical pressure is temperature independent) by the application of a rigid pressure shift to the classical results is wrong, as the shift should depend on the temperature considered. The SSCHA result for the classical approximation at finite temperature, T=300 K, overestimates the critical pressure computed from the spectral analysis of the MD simulation in Ref. [24] (P_c =70 GPa) with the same DNN. The origin of the discrepancy cannot be addressed to size effects, since convergence tests confirm that the results in the cell with 16 water molecules are converged. Partially, the difference can be explained by the discretization of the pressure sampled and by the fact that in the MD simulations the critical pressure is extracted from an indirect observable, namely the changes in the IR spectra; and the results in Sec. 4.5.4 reveal that the MD spectrum at the critical pressure is similar to the SSCHA one at about 7 GPa below (still in phase VII), questioning the location of the critical pressure from the MD simulations. Further details on this problem can be found in Sec. 4.5.4 that is fully dedicated to the comparison

with MD simulations. The remnant part of the error has to be addressed to the SSCHA approximation.

Quantum effects overturn the classical picture, inducing a 55 GPa reduction of the critical pressure at T=0 K, with a critical pressure that is temperature independent in agreement with the experiments. The analysis of the order and disordered phases has not reported remarkable differences in the symmetrization process, both phases VIII and VII symmetrize the hydrogen bond lengths at about 52 GPa turning into phase X. The unit cell with 16 water molecules is sufficient to produce converged results; a comparison with a 128 molecules supercell gave the same results and excluded the presence of hysteresis cycle and consequently a first-order phase transition.

Three of the many possible hydrogen arrangements for the disordered phase have been analyzed, with at most 2 GPa difference in the critical pressure, negligible with respect to typical experimental uncertainties. Fig. 4.17 confirms a good agreement with the experimental data, the critical pressure is shifted about 10 GPa below the experiment but the 10 GPa isotopic shift, between deuterated and hydrogenated ice, is perfectly described. If we correct the SSCHA critical pressure with the 5 GPa error found in the force field description with respect to DFT-PBE, we have an almost perfect agreement with the experimental transition pressure $P_c^{exp} \simeq 60$ GPa. The "mysterious" phase indicated by the question mark in Fig. 4.17 (**a**), hypothesized because of the shape of the absorption spectra with a broad peak at the critical pressure that becomes narrower above, is not detected but a direct transition between phase VIII/VII to X is observed. We did not show the VIII-VII phase boundary as it has been widely studied theoretically.

4.5 Spectroscopy

This section is dedicated to the analysis of infrared spectroscopy in high-pressure ices. As broadly explained in Sec. 4.3, all the signatures of the high-pressure phase transition are indirect. Apart from X-ray or neutron scattering, all of them are spectroscopy-based. First, I review the experimental observations in the field of IR absorption (Sec. 4.5.1) and then I show the simulated spectra within the TD-SSCHA framework in Sec. 4.5.2, offering a comparison with experimental measurements [26] and previous theoretical calculations [24] and the analysis of the role of quantum anharmonic effects and disorder. All the details of the simulations can be found in App. H.

4.5.1 Infrared spectroscopy: Experimental observations

Infrared absorption experiments can provide strong insights into the features of the phase transition. Group theory analysis of the vibrational modes predicts six Infrared active modes in phase VIII, namely symmetric ($\nu_1(A_{2u})$) and asymmetric ($\nu_3(E_u)$) stretching, a bending mode ($\nu_2(A_{2u})$), two rotational modes (ν_R and $\nu_{R'}$ with E_u symmetry) and a low energy translational mode ($\nu_T(E_u)$). Modes with E_u symmetry are double degenerate. The normal mode analysis is not available for ice VII, which has a cubic structure with proton disorder. However, the absorption bands observed for ice VII can be assigned on analogy with the mode assignments for ice

VIII [27, 175]. The fully centered ice X has only two infrared active modes, related to a deformational mode ν_D and a stretching mode ν_S, both with T_{1u} symmetry and triple degenerate.

Reflectivity measurements on ice VII by Goncharov et al. [26] allows us to follow the evolution of the IR-active modes as a function of pressure. The authors transformed the reflectivity data into the imaginary part of the dielectric function, through Kramers-Kronig transformations. The spectra for several pressures are reported in Fig. 4.18 (**a**). The strong absorption of the diamond anvil cell in the region between $1800\,\mathrm{cm}^{-1}$ and $2400\,\mathrm{cm}^{-1}$ prevents the acquisition of data and the dashed line in that region were obtained with a classical oscillator model fit.

In the low-pressure regime, the stretching mode near $3000\,\mathrm{cm}^{-1}$ softens with increasing pressure and splits in the pressure range of 25 to 45 GPa as can be seen from the pressure dependence of the vibrational modes in Fig. 4.18 (**b**). At still higher pressures, the stretching vibration continues to soften and intersect with the vibrational mode ν_R at $\omega \simeq 1250\,\mathrm{cm}^{-1}$ that hardens with pressure. The other IR-active rotational mode predicted by group theory in phase VII is not detected in the experiment because of its low intensity. During the softening, the stretching bands becomes always broader at it is hard to distinguish the peak position. At P=60 GPa the Infrared spectrum changes abruptly; the stretching mode becomes much weaker and a strong low-frequency band is observed. This abrupt change in the spectrum is considered the signature of the phase transition, the deformational mode ν_D is thought to be related to the transformation into the symmetric structure. It originated from the merging of the rotational mode ν_R and the bending mode, that in the experiment is not detected, probably because of its low intensity (as can be seen from the SSCHA simulations in Fig. 4.19) but predicted from the group theory analysis. For pressures above $P_c = 60$ GPa, the spectra are drastically different from those in phase VII. With increasing pressure, the two IR-active modes increase in frequency and gain intensity. Then, between P=150 GPa and P=160 GPa, a resonance between the stretching and the deformational mode occurs, characterized by the approach and separation of the two vibrations followed by an exchange of intensity, with the stretching modes that become dominant in the high-pressure region. Notably, there is a family of stretching modes that soften approaching the critical pressure, however, only one of these is the mode that goes exactly to zero driving phase transition. However, this is a silent mode in the symmetric phase: we saw in Sec. 4.4.1 that it is a mode in the **R**-point of the Brillouin zone, consequently, it can be neither Raman nor Infrared active.

4.5.2 TD-SSCHA simulations of Infrared spectroscopy

I compare the experimental measurements with the Infrared simulations of the ordered and disordered phases within TD-SSCHA (with thermal and quantum fluctuations included) at T=300 K in Fig. 4.19. Since at the phase transition the low-energy modes are much more intense than the high-energy ones, I plot the product of the imaginary part of the dielectric function ϵ_2 by the frequency ω to weigh more the high-frequency structures. Moreover, it can be proven that this quantity is related to the physical properties of the system that can be measured:

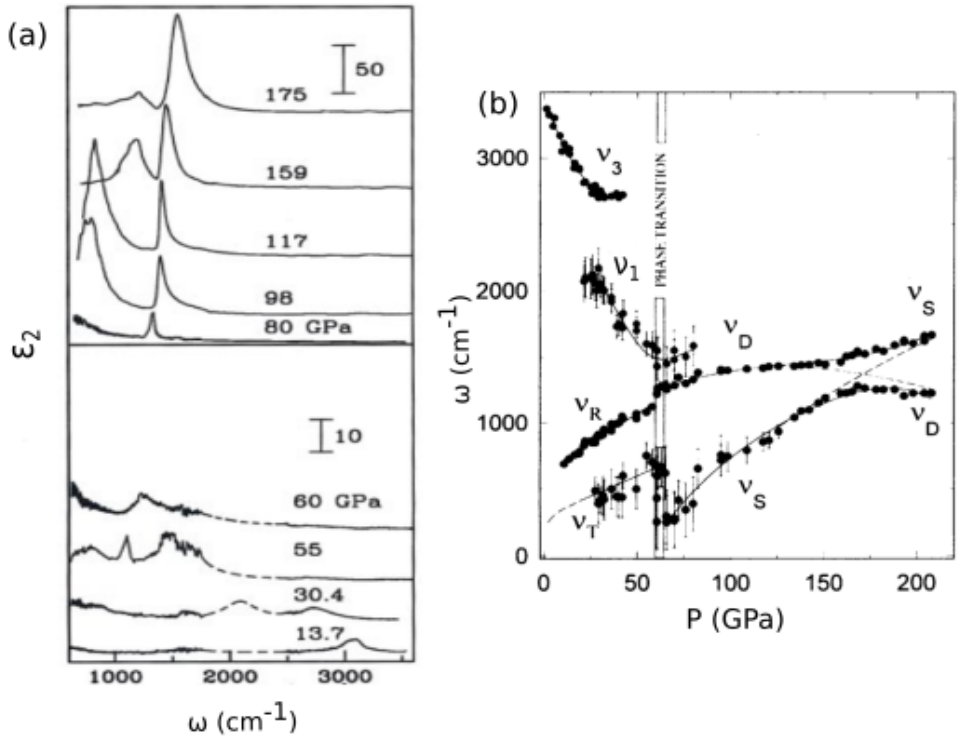

Figure 4.18. Experimental Infrared absorption data from: *Compression of ice to 210 gigapascals: Infrared evidence for a symmetric hydrogen-bonded phase. Goncharov et al. Science, 273(5272):218–220, July 1996 [26].* Panel **a**: imaginary part of the dielectric function ϵ_2 of H_2O ice VII at T=295 K obtained from a Kramers-Kronig analysis of reflectivity spectra. Continuous lines are the Kramers-Kronig transformation of the experimental data, while the dashed curves are derived from oscillator fit to fill the empty region due to strong diamond absorption. The critical pressure is identified at P=60 GPa. In the upper part of panel **a** the phase transition has already occurred and the system is in phase X. Panel **b**: pressure dependence of the IR-active vibrational frequencies. Stretching modes are labeled ν_1 and ν_3 before the phase transition and ν_s after. In phase VII, ν_R and ν_T are rotational and translational modes, respectively. ν_D is a deformational mode that arises at the phase transition.

$$\alpha(\omega)n(\omega) = 4\pi\omega\epsilon_2(\omega), \tag{4.5}$$

where, $\alpha(\omega)$ is the energy-dispersive absorption coefficient and $n(\omega)$ the frequency dependent real part of the refractive index.

The comparison with the experiment is done in such a way that the claimed transition pressures coincide, (the VIII phase is plotted at 2 GPa above phase VII: $P_c^{VII} = 52$ GPa, $P_c^{VIII} = 54$ GPa) so that, for each panel, the experimental spectra are about 10 GPa above the SSCHA ones according to the pressures available in the simulations. The pressure at which each spectrum is calculated can be read in the labels of Fig. 4.19. Notably, the experimental measurements are in phase VII which is the one stable at T=300 K.

The SSCHA manages to reproduce very well the infrared spectra at low pressure, P=30 GPa and P=40 GPa. The differences between the ordered and disordered phases are mostly in the line shape and become more remarkable close to the phase transition. Unfortunately, approaching the critical pressure, there are no experimental data in a small pressure range, where the SSCHA spectra reveal the most interesting changes, with the abrupt collapse of the soft stretching vibrations. The phonon frequencies of the IR-active modes as a function of pressure are reported in Fig. 4.20 (**a**). Notice that the analysis of the active modes in the SSCHA relaxed structure is in perfect agreement with group theory. In the SSCHA simulations, we barely detected the non-dispersive bending mode ν_2 at $\omega \simeq 1300\,\mathrm{cm}^{-1}$ invisible in the experiment. Its intensity is negligible as can be seen from the size of the dots that is a measure of the IR intensities (see App. H for further details). The same is for the IR activity of the second rotational mode ν_R' , explaining why it is not detected experimentally, in combination with the fact that has energy very close to the other rotation and probably cannot be distinguished. In a range of 10 GPa, from 40 to 50 GPa, the symmetric stretching vibration suffers 70% energy reduction. Within the TD-SSCHA, while softening, the stretching modes interact with the other modes, in a series of resonances followed by intensity exchanges, reproducing correctly the experimental trend. Increasing the pressure, the two rotational modes become indistinguishable and one of them is no more IR-active right before the phase transition. Following the abrupt collapse of the stretching modes toward the low energy regime, which can be understood very well by looking at the labeled peaks in Fig. 4.21 and from Fig. 4.20 (**a**), the spectra change abruptly. The intensity of the low-energy modes, in the region where there are no experimental data published to date, increases by an order of magnitude. The change in intensity of the modes cannot be grasped very well from Fig. 4.19 where I used panels of the same height, one should look at the scale of the ordinate axis that is different in each panel. Instead, it can be understood from Fig. 4.22, where the spectra are not rescaled.

The intensity increase of the translational mode, in combination with the birth of the deformational mode ν_D (coming from the merging of the rotational and bending modes), gives a clear sign of the transition to the symmetric phase. Experimentally, the signature of the phase transition is the birth of the asymmetric and broad deformational mode and the disappearing of all the other bands at $P \simeq 60$ GPa. The failure in the reproduction of the asymmetric shape of the ν_D in the TD-SSCHA simulations can be ascribed to the use of a small simulation cell and I'm

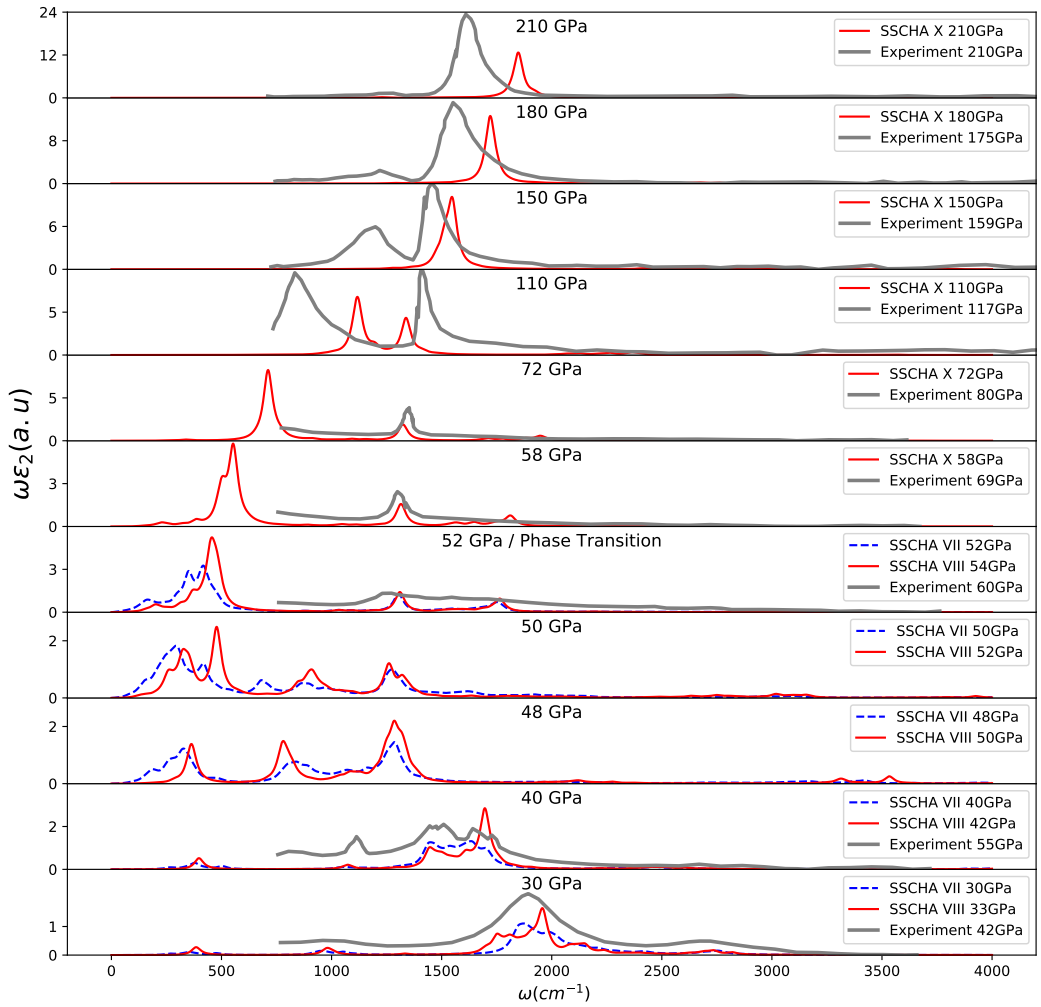

Figure 4.19. Comparison of the IR spectra of the ordered and disordered phases within TD-SSCHA framework at T=300 K with the experimental measurements in Ref. [26]. In the figure, the product of the $\epsilon_2(\omega)\omega \propto n(\omega)\alpha(\omega)$ is shown. The labeled pressure in the title of each panel refers to the SSCHA simulated disordered phase (blue dashed lines), and, in order to match the phase transition ($P_c^{VII} = 52$ GPa, $P_c^{VIII} = 54$ GPa), the spectra of phase VIII (red lines) are reported at 2 GPa above those of the disordered phase. The experimental spectra are in continuous thick grey lines and they are compared at around 10 GPa above to match the difference in the critical pressure $P_c^{exp} = 60$ GPa. When no experimental measurements are available at the given pressure only the SSCHA results are shown. Above 52 GPa (54 GPa) phase VII(VIII) transforms into phase X and so also the difference between the two vanishes. Details on the computation of the spectra are in App. H.

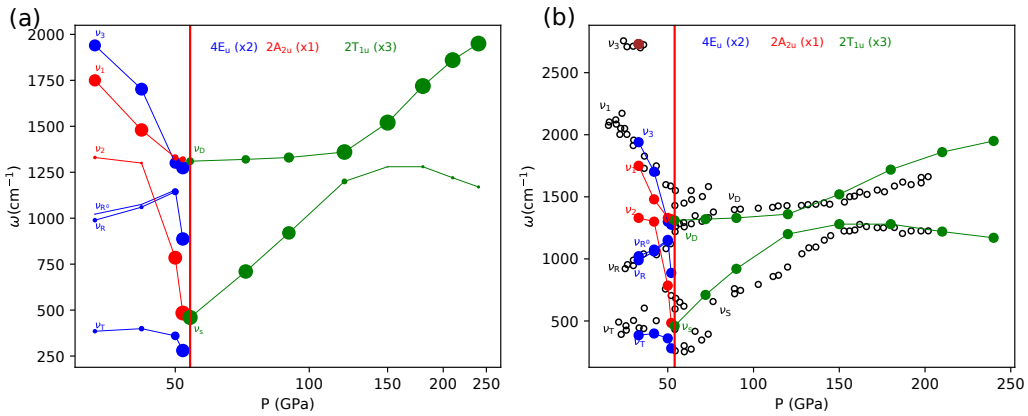

Figure 4.20. Panel **a**: Infrared active modes at T=300 K as a function of pressure. The six IR-active modes in phase VIII within the SSCHA framework are indicated with blue and red points for E_u and A_{2u} symmetry respectively. There are two stretching modes (asymmetric ν_3 and symmetric ν_1), a bending mode (ν_2), two rotational ones (ν_R and $\nu_{R'}$) and a translational mode ν_T. Green points indicate the two infrared active modes in phase X, namely a deformational (ν_D) and a stretching mode (ν_S). The size of the area of the points is a measure of the Infrared activity of each mode compared to that of the most intense one (See App. H for further details). The degeneracy of each representation is reported in the upper panel of the figure inside the parenthesis, while the number in front of the label indicates the number of different modes. The pressure scale is logarithmic to emphasize the pressure dispersion of the modes in phase VIII. Panel **b**: Comparison of the six IR-active modes with the experiment [26]. Here, the experimental frequencies are the black empty circles, and the denomination of the modes can be found written in black. In this case, the simulated spectra are reported with points of the same size in accordance with the experiment, where the intensity of each mode is not discussed. The experimental data are shifted to make the critical pressures (red vertical line) coincide ($P_c^{\text{exp}} \simeq$60 GPa). The ν_3 mode, addressed as the stretching mode in the experiment is probably a combination mode (see text for further details). In this panel the pressure scale is linear.

running simulating simulations in a cell with 128 water molecules to verify this hypothesis. Within the SSCHA, we observe the existence of a combination mode at 1800 cm^{-1} in phase X that blueshifts with increasing pressure and that is not detected experimentally.

In the experiment, the deformational mode assumes a symmetric shape as soon as the pressure is increased in phase X, and its non-dispersive character, as well the line shape are perfectly described within the SSCHA. Increasing again the pressure, the stretching mode that lies in the low-energy region at the critical pressure increases its frequency and loses intensity. At $P \simeq$ 110 GPa, we manage to reproduce the intensity exchange of the ν_D and ν_S modes as evident from Fig. 4.19 and Fig. 4.20 (**a**) and (**b**), where the size of the dots is a measure of the IR-activity with respect to the most intense mode. Already for $P \gtrsim$150 GPa, the stretching mode is the only one observable in the SSCHA, while in the experiment the deformational mode loses intensity slowlier and it is detectable even at P=210 GPa. The comparison of the TD-SSCHA IR-active modes with the experiment can be found in Fig. 4.20 (**b**), where I describe the SSCHA frequencies with dots of the same size, disregarding the

intensity of each mode, as did in the experiment. As discussed above, the SSCHA describes perfectly almost all of the IR-active modes in phase VIII. We can notice a visible error in the description of the ν_3 stretching mode, that in the experiment it is located around 2750 cm^{-1}. However, a deeper analysis of the asymmetric stretching mode with the Lanczos algorithm revealed that probably the mode indicated by ν_3 in the experiment is the combination mode indicated in Fig. 4.22 that crosses with the stretching as seen in Ref. [27]. Indeed, the asymmetric stretching mode is very close in energy with the symmetric one and the broad peak does not allow to distinguish the two of them. Anyway, a deeper analysis of the combinations is yet to be performed and it is scheduled for the future. The correct behavior of the two IR-active modes in the symmetric ice is correctly reproduced within the TD-SSCHA even if the evolution of the peak position is not perfectly grasped as evident in the region at the right of the red line in Fig. 4.20 (**b**).

In phase VII, the spectra are calculated on the disordered configuration denoted Set 2 in Sec. 4.4.2. The structure with partial symmetrization of Fig. 4.16 at T=0, arises at P=50 GPa at finite temperature, T=300 K. The comparison of the spectrum at P=50 GPa in Fig. 4.19 with those at 2 GPa below and above reveals a similar behavior. The peaks are very broad in the proximity of the phase transition, both in the purely disordered phase (P=48 GPa), in the partially ordered (50 GPa), and in phase X at the critical pressure. This definitely rules out the interpretation of the partially ordered structure as the hypothesized intermediate phase.

4.5.3 Quantum vs Classical SSCHA

Usually, theoreticians utilize the classical spectra to simulate the quantum absorption by applying a constant pressure shift according to the difference in the critical pressures. My simulations reveal that this procedure is not very accurate. In this section, I compare the quantum spectra computed within TD-SSCHA, including thermal and quantum effects, with the classical ones. The latter are again computed with the TD-SSCHA technique, but the quantum fluctuations are removed by substituting the physical masses of the atoms with "infinite" values (10^6 times the real ones), instead thermal fluctuations are considered to compute the IR absorption spectra at finite temperature.

In Fig. 4.21, I compare the quantum and classical spectra of phase VIII at T=300 K with classical simulations. The classical simulations are compared with the quantum ones at a pressure 26 GPa below to match the difference in the critical pressures ($P_c^{quantum} = 54$ GPa and $P_c^{cl} = 80$ GPa). When the classical spectra at the exact pressure difference required have not been calculated, I choose the simulation at the pressure value that matches the pressure difference the best.

There is an evident renormalization of the phonon frequencies due to quantum effects. The most relevant differences appear close to the phase transition, where the sudden collapse of the stretching modes toward the low energy region is faster in the quantum regime, albeit the same qualitative behavior can be observed in the classical simulations. The label over each mode indicates its name and symmetry and it allows to follow unambiguously their evolution as a function of pressure. In the region at the left of the black vertical line, there are no experimental data to date, and we can see there the importance of the quantum renormalization of the

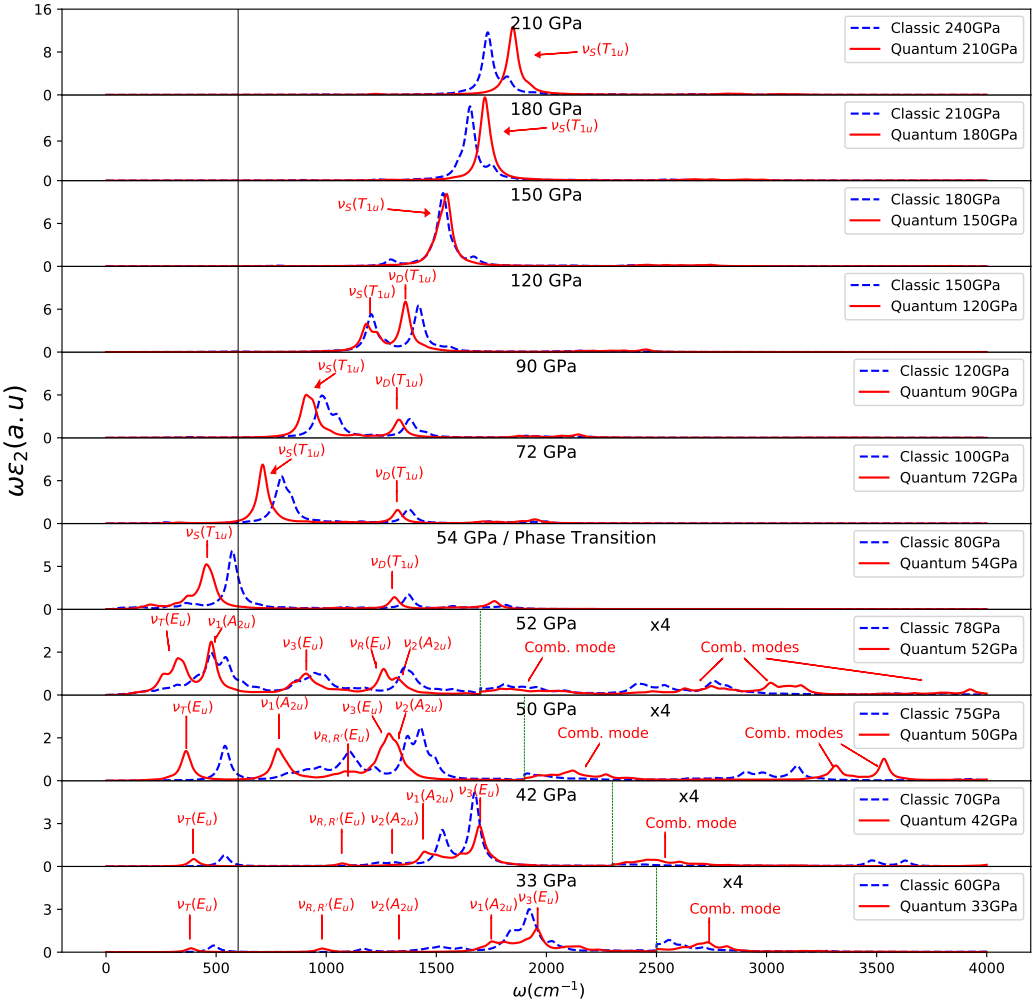

Figure 4.21. Direct comparison of the classical and quantum simulations of Infrared spectrum in phase VIII at T=300 K within the SSCHA. For a correct comparison, the 26 GPa difference of classical with respect to quantum critical pressure is taken into consideration. So, in each panel, the quantum IR spectra are compared with the classical ones computed at about 26 GPa above. When quantum spectra at exactly the correct pressure difference have not been computed, the closest ones are chosen. The black vertical line separates the region accessible experimentally. Above each peak, the name and the symmetry of the mode are reported, so that their evolution with pressure can be followed unambiguously. In phase VIII, the spectra at the right of the green dashed line are magnified by a factor 4 to make visible the combination modes, indicated by the arrows. Conversely, in phase X the scale is the same for each energy and the combination modes can be identified with the peaks without the label appended. Details on the computation of the spectra are in App. H.

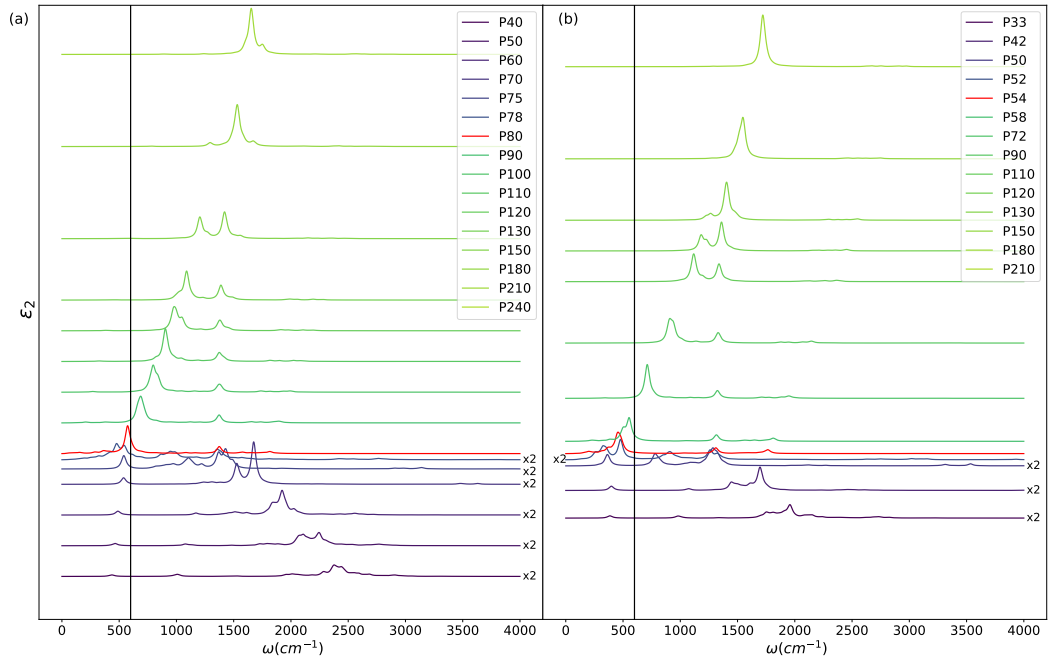

Figure 4.22. Imaginary part of the dielectric function in phase VIII at different pressures at T=300 K. **a** : Classical computations with critical pressure $P_c^{cl} = 80$ GPa. **b** : Quantum simulations with $P_c^{\text{quantum}} = 54$ GPa. The two panels are aligned in such a way that the critical pressures (red spectra) in the quantum and classical case are at the same height, so a 26 GPa difference shift is set. Within each panel, the plots are linearly shifted with the pressure along the vertical axis. Blue color gradient spectra indicate ice phase VIII structures and they are multiplied by a factor 2, as indicated in figure, to make them more visible against the more intense modes in the symmetric phase X (green color gradient spectra). Both quantum and classical spectra have been computed with the Lanczos algorithm of Sec. 2.7.3 and a smearing factor of $25\,\text{cm}^{-1}$ has been used (see also App. H). In the region at the left of the black vertical line, there are no experimental data up to now.

frequency: the translational mode is predicted to be partially visible at the critical pressure in the classical regime, while in the quantum simulations it is detectable only at higher pressures.

The increased intensity of the translational mode by approaching the critical pressure from below is common in both the picture, as well as the blueshift of the stretching mode in phase X and the intensity exchange with the deformational mode. However, the stretching mode is more dispersive in the quantum regime while the non-dispersive deformational mode is perfectly described in the classical simulations apart from the frequency renormalization. At high pressure, the classical spectra present a stretching peak with a shoulder at his right that is absent in the quantum simulations. The region at the right of the vertical green dashed line in phase VIII is magnified by a factor 4 to make more visible the combination modes, indicated by the arrows. This is not done in phase X, from P=54 GPa and above, where they are more easily identifiable.

In Fig. 4.21, the panels for different pressures are all equispaced and have the

same height, disregarding the different intensity value (visible from the different scale for each panel) and the nonconstant pressure shift between a panel and the following one. The absolute values of the intensity and the pressure evolution of the spectra with pressure in both quantum and classical simulation can be better understood from Fig. 4.22, where I show the imaginary part of the dielectric function for all the pressure analyzed, applying a shift in the vertical direction that is linear with the pressure. The spectra in phase VIII, below the red one that indicates the critical pressure, are magnified (x2) to compensate for their low intensity with respect to the high-pressure ones.

In conclusion, the direct comparison of quantum and classical infrared simulations reveals that is not correct to use a constant pressure shift to deduce the quantum regime from the classical one, but in principle, one should use a different pressure shift for each quantum pressure albeit classical simulations manage to reproduce qualitatively the evolution of the spectra under pressure.

4.5.4 Comparison with MD simulations

Most recent and accurate Infrared spectra simulations have been performed by Zhang et al. [24], from whom we borrowed the force field described in Sec. 2.9.2. The authors ran MD simulations on ice VII at T = 300 K as a function of pressure, crossing the phase transition toward symmetric ice X, using a simulation cell containing 128 water molecules.

In order to compare with the MD results, I computed the classical IR absorption spectra for the disordered phase too within the time-dependent extension of the SSCHA as in Sec. 4.5.3. Out of all the possible hydrogen atoms arrangements, I choose the one denoted Set 2 in Sec. 4.4.2 at T=300 K. All the details of the calculations are in App. H.

In principle, apart from quantum effects (that can be included with path integral methods), MD is exact as it can grasp anharmonicity at all orders. So, the comparison with MD in Fig. 4.23 is also useful to set a benchmark on the quality of the SSCHA simulations.

The agreement between MD and SSCHA spectra is quite good at low pressure, where the difference between ordered and disordered structures is also negligible. Instead, discrepancies both between MD and SSCHA and between ordered and disordered phases start to arise when approaching the phase transition. Already at P = 50 GPa, the position of the broad peak at about 2200 cm^{-1} is shifted with respect to that in the MD spectra, and also the peak shapes of ice VII and VIII differ. The abrupt change in the IR spectra at P = 70 GPa lead the authors of the MD simulations to position the phase transition at that pressure, while in the SSCHA simulations, the structure is still that of VIII/VII phases. The SSCHA phase transition occurs at P=82 GPa in the disordered phase and at P=80 GPa for the ordered one, where from the structure relaxations the symmetrization of the hydrogen bond is observed. The small difference in the critical pressure explains the visible differences between the infrared spectra of ice VII and VIII at P=70 GPa. In fact, as for the quantum case, the spectra suffer drastic and sudden variations in a small pressure range close to the phase transition, and a difference of 2 GPa implies a great difference in the spectra as can be better seen in Fig. 4.24, where I

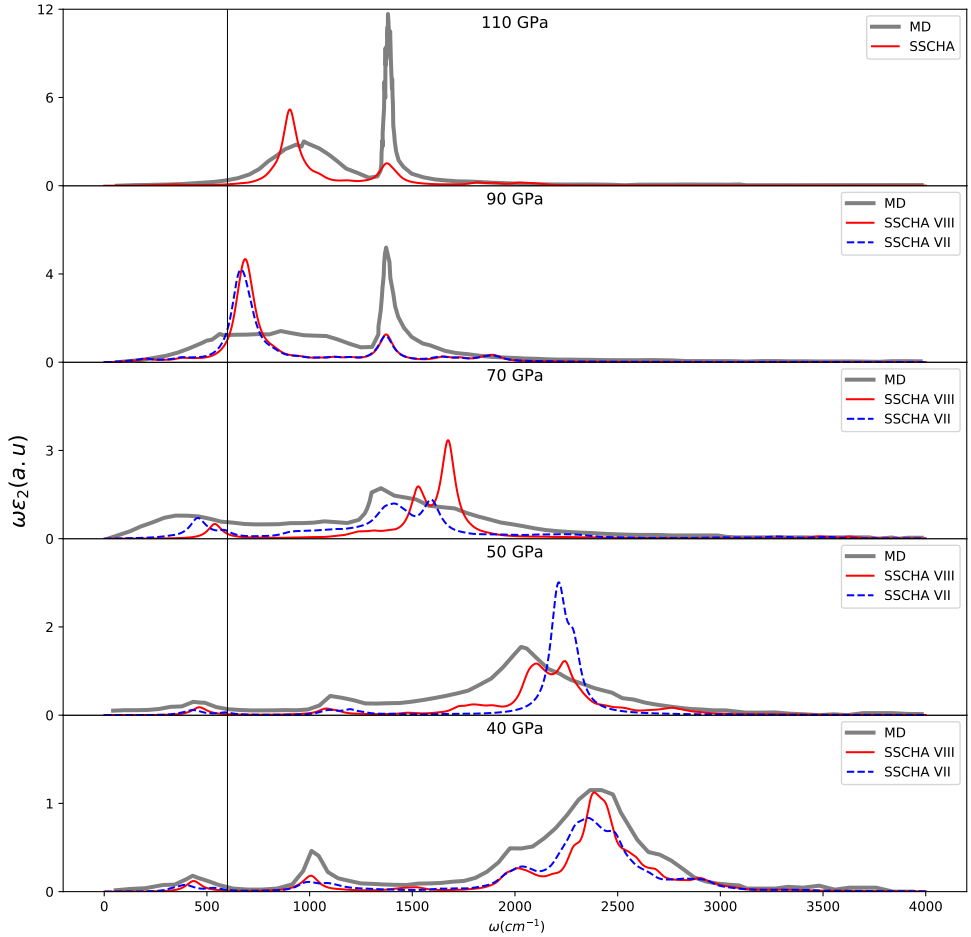

Figure 4.23. Comparison of MD simulations [24] with classical SSCHA results at T=300 K. The imaginary part of the dielectric function times the frequency for different pressure, below and above the phase transition are reported as a function or the frequency. The product $\omega\epsilon_2(\omega)$ is related to the absorption coefficient and the real part of the refractive index (see Eq. (4.5)) that are easily accessible experimentally. Changes in the MD spectra set the phase transition at P=70 GPa, while within the SSCHA $P_c^{cl,VIII} = 80$ GPa and $P_c^{cl,VII} = 82$ GPa. MD spectra are the thick grey lines. Spectra for phases VIII and VII within the SSCHA formalism are shown as solid red lines and dashed blue lines, respectively. The simulation cells contain 128 water molecules in MD and 16 water molecules within TD-SSCHA. Both the classical spectra of phases VIII and VII have been computed with the Lanczos algorithm of Sec. 2.7.3 and a smearing factor of $40\,\mathrm{cm}^{-1}$ is used (see also App. H). No experimental measurements are available in the region at the left of the black vertical line to date.

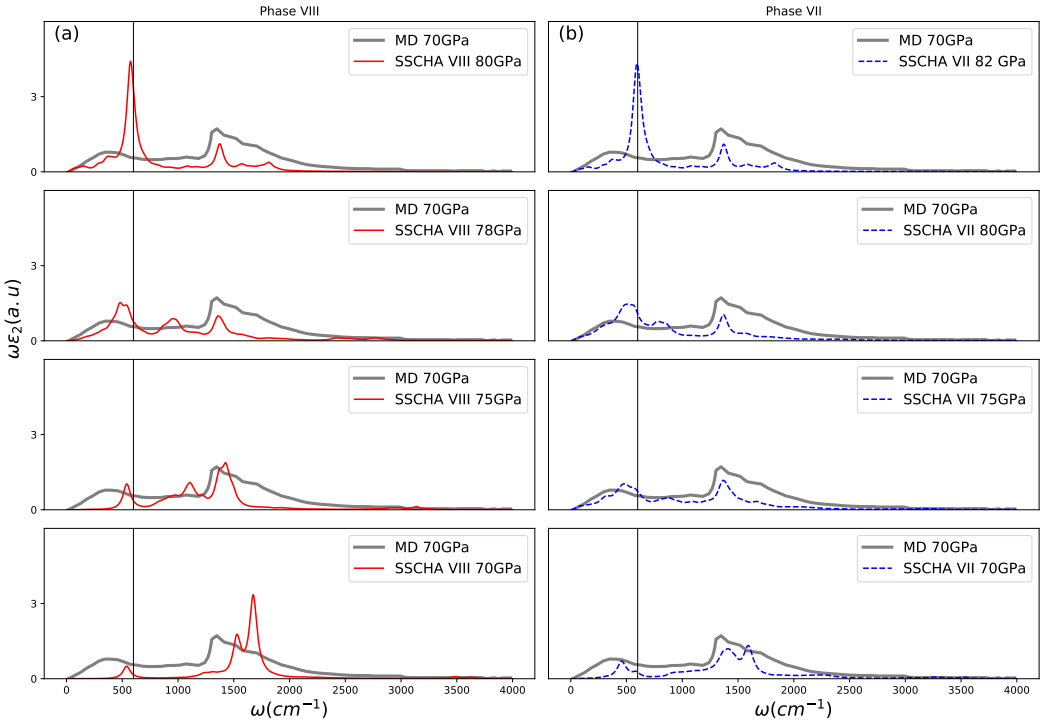

Figure 4.24. Zoom of the classical IR spectra close to the phase transition at T=300 K. The product of the imaginary part of the dielectric function times the frequency is reported $\omega\epsilon_2(\omega)$. Panel **a** and **b** represent the ordered and disordered phase respectively. The pressure are displayed vertically from P=70 GPa to the critical pressure $P_c^{VIII} = 80$ GPa ($P_c^{VII} = 82$ GPa) and each one is compared to the classical MD spectrum at the critical pressure $P_c^{MD} = 70$ GPa [24] (solid thick grey line). Solid red lines are the spectra in phase VIII, while the blue dashed ones represent the disordered phase VII. In the region at the left of the black vertical line, there are no published experimental data.

report the spectra of both the ordered and disordered phase from P=70 GPa up to the critical pressure. However, the different critical pressures can be explained by taking into consideration the results for the different hydrogen arrangements in phase VII of Sec. 4.4.2, since that 2 GPa difference is comparable to pressure shift in the phase transition for the different disordered configurations.

The discrepancies between the MD and phase VIII Infrared spectrum at P=70 GPa are evident, both in the peak position and shape. However, the phase VII spectrum shows some similarities with the MD simulation, and increasing the pressure we can see that both the disordered and the ordered phase spectra resemble more the MD one. In particular, for P=75 GPa, the IR spectrum of phase VII in Fig. 4.24 is very similar to that in MD and the same goes for the ordered phase in the pressure range 75-78 GPa, right before the phase transition.

At the critical pressure, the TD-SSCHA simulations manage to reproduce well the deformational mode of the MD but the low-energy stretching mode has a completely different shape: very broad in MD and sharp in SSCHA. This difference in the stretching band continuous to exist even at higher pressure, in the symmetric phase X (where only one plot for the SSCHA is reported since VII and VIII phase transform

both to ice X), and as discussed below I'm checking if it is due to size effects. The peak position is in good agreement with that predicted in MD.

The difference of about 10 GPa between SSCHA and MD critical pressures can be partially explained by a greater uncertainty in the MD value since spectra with 20 GPa steps are reported. Moreover, the similarity of the SSCHA spectra at P=75 GPa with the MD one at 70 GPa questions the occurrence of the phase transition at that pressure in MD (in the SSCHA we know for sure that this spectrum is still related to phase VII), reducing again the pressure difference.

Finally, the residual mismatch can be ascribed to size effects and to the SSCHA approximation. The MD simulation in cells with 128 molecules is surely converged, while in the SSCHA I'm checking the convergence of the classical simulations with the cell size. I'm currently running simulations in a supercell with 128 water molecules at P=110 GPa in phase X and at P=70 GPa in the disordered phase, to verify if the discrepancies in the lineshapes are entirely due to finite-size effects. The importance of including nonlinear effects in the effective charges for studying high-pressure ice discussed in previous works [23] is discussed in App. J and it seems to play a negligible role.

4.5.5 Conclusions and outlooks

Experimentally, the most relevant insights on the phase transition toward the symmetric ice come from spectroscopy. In this section, I have shown that the TD-SSCHA simulations of the IR absorption spectra shed light on the unique signature of the transition. The possibility to run simulations with small pressure intervals close to the critical pressure, allows me to notice sudden changes in the Infrared spectra, where the frequency of the stretching modes collapses rapidly to the low-energy regime in less than 10 GPa and I detected a strong intensity increase of the low-energy translational mode, that we retain as an important signature of the phase transition. In this low-energy region, where such important changes of spectra happen, there are no published experimental measurements to date. If our results were confirmed experimentally, they would set an indistinguishable signature for the phase transition, allowing us to detect unambiguously the critical pressure. The evolution of the modes with pressure predicted within the SSCHA framework is in good agreement with the experiment [26], from the strongly dispersive stretching modes to the non-dispersive bending and rotational ones, both in the high-pressure region and below the phase transition. Moreover, as for the static study of the hydrogen bond symmetrization in Sec. 4.4, the spectroscopic analysis does not reveal the presence of an intermediate disordered phase. The broad band close to the critical pressure that leads to the hypothesis of such a phase is detected also in the pure ice X right after the transition, meaning that in that region the structure is strongly anharmonic while increasing the pressure the anharmonicity is reduced.

The comparison of quantum and classical spectra clearly shows that despite the latter giving a qualitatively good representation of spectra evolution under pressure, except a quantum renormalization of the frequencies and the critical pressure shift, it is not correct to interpret the classical spectra as the quantum ones by just applying a rigid pressure shift as it is commonly done. Indeed, for quantitative good analysis, it is necessary to choose a different pressure shift for each quantum pressure.

The comparison of the classical SSCHA simulations at T=300 K with MD data [24], that are in principle exact (apart from quantum effects) shows good agreement in the low-pressure regime (in phase VIII/VII) with some differences arising in the proximity of the phase transition both in the peak shape and in the position of the critical pressure itself ($P_c^{\mathrm{MD}} = 70$ GPa and $P_c^{\mathrm{SSCHA}} = 80$ GPa). While size effects can explain completely the different peak shape, the discrepancy in the critical pressure is partially explained by the indirect observable coming from spectroscopy, where the MD spectrum at the phase transition is very similar to the TD-SSCHA one 7-8 GPa below; a fact that question the exact location of the critical pressure in the MD simulations. To check the importance of size effects, I'm running simulations in a bigger cell (containing 128 water molecules).

The effect of position-dependent effective charges is negligible if compared to the results obtained by using the average of the effective charges computed from an ensemble of 500 different configurations. A glance of the importance of nonlinear effects, that are considered to be important [23], can be seen from the comparison of the second derivative of the dipole moment with the effective charges, revealing a very small contribution.

The next goals are at first to check the origin of the discrepancy with MD simulations, completing the convergence tests in simulation cells with 128 water molecules. Moreover, I plan to investigate in detail the origin of the combination modes in the IR spectra and, to give further strength to the results obtained by inspecting the Infrared response of the dense high phases, I aim at simulating also the Raman scattering spectra, widely explored experimentally [176, 177].

Conclusions

This thesis is focused on the exploration of the phase diagram of ice, centering the attention on the low-pressure region, occupied by the ordinary ice I_h and its proton-ordered counterpart ice XI, and on the high-pressure regime, dominated by the high-density ices VII, VIII, and the symmetric ice X.

This is done through numerical simulations. I employed the stochastic self-consistent harmonic approximation (SSCHA), which we reviewed in Ref. [32], to fully account for the strong thermal and quantum anharmonic nuclear motion. In the first part of the thesis, I describe the main idea behind the SSCHA and I show how I improved the code to manage to simulate efficiently systems with many atoms in supercells, like the ice phases investigated in this thesis.

I elucidated the importance of quantum anharmonic effects in the thermodynamic, structural, and vibrational properties of low-pressure ices. The anomalous strong temperature dependence of the bulk modulus, 20% variation from 0 to 300 K, is fully explained by thermal and quantum anharmonic fluctuations, revealing the combined effect of the vibrations (64%) and the thermal expansion (36%). I proved that the anomalous volume isotope effect (VIE) in ice XI originates from a nonmonotonous volume expansion due to quantum effects. When the mass of the hydrogen isotopes is increased, the volume first expands, saturating slightly above the tritium mass, and then contracts to the classical value. This demonstrates that the VIE is due to a strongly nonlinear regime of quantum fluctuations in ice, which commonly employed approximated theories (such as the QHA) do not grasp. I show that also oxygen is in a strong quantum mechanical regime, being responsible for a 2% volume reduction in the classical limit.

In addition, for the first time, the low-energy range of phonon dispersion of deuterated ice at T=140 K under pressure [143] is excellently reproduced by the anharmonic renormalized phonons and correct treatment of the electron exchange and correlation. This result paves the way for the study of thermal transport from first principles and the simulation of ice under pressure, where acoustic phonons are the only detectable modes. I observe also an anharmonic renormalization of about 10% of the bending and stretching modes that grants a good prediction of the vibrational spectroscopies.

I solve the puzzle behind the long-debated phase transition [16, 22, 26] of the high-density ices, VIII/VII, to the symmetric phase X, where the proton lies in the midpoint of two neighboring oxygen atoms, by simulating the structural and thermodynamic properties of the transition, overcoming the absence of direct experimental signatures (the hydrogen atoms are not visible by X-ray diffraction and the limited data quality and uncertainties in the procedure of data correction in

neutron scattering hampers an unambiguous interpretation at high pressures).

I show that the simulations in the classical approximation at T=0 K manage to establish the second-order character of the phase transition, the structure of ice VIII/VII transforms continuously to the symmetric ice X by increasing the pressure and I detected the presence of soft modes typical of second-order phase transitions. I confirmed the result also in the quantum regime where I proved the absence of hysteresis too.

My simulations revealed that while in the classical approximation the critical pressure is strongly temperature-dependent, showing a nonlinear variation from $P_c(0\ \mathrm{K}) \simeq 107$ GPa to $P_c(300\ \mathrm{K}) \simeq 80$ GPa, quantum fluctuations overturn this picture, predicting a temperature-independent critical pressure for all the isotopes ($P_c^{\mathrm{H_2O}} \simeq 52$ GPa and $P_c^{\mathrm{D_2O}} \simeq 62$ GPa), in accordance with the experimental data. With this result, I proved that it is wrong to simulate the quantum regime by simply applying a pressure shift to the classical results as commonly done. Instead, one should apply a different pressure shift for each temperature. In fact, the 30 GPa reduction of the critical pressure due to quantum effects predicted in some theoretical works [22], is valid only at T=300 K, while at T=0 K, I detected a reduction even stronger, of about 55 GPa. Moreover, my simulations of the Infrared spectra revealed that the shift to apply is also pressure dependent.

I manage to describe perfectly the 10 GPa isotopic shift of the critical pressures between deuterated and hydrogenated ice. Instead, their absolute value underestimates by 10 GPa the experimental data, $P_c^{\mathrm{exp,H_2O}} \simeq 60$ GPa. However, I show that taking into consideration the 5 GPa underestimation of the critical pressure due to the use of a Deep Neural Network with respect to the *ab initio* calculation, I recover an almost perfect agreement with the experiment.

I shed light on the presence of the hypothesized intermediate phase between ice VIII/VII and X [16]: I show that neither the thermodynamic properties nor the simulated vibrational spectra detected its presence since the broad bands close the critical pressure that led to its hypothesis can be found also in pure ice X. The possibility to run simulations in a small pressure range close to the critical pressure allows me to notice sudden changes in the Infrared spectra in less than 10 GPa, where the stretching mode collapses rapidly toward the low-energy regime and the translational mode increases its intensity by about one order of magnitude. These two features can be regarded as the signature of the phase transition. Unfortunately, there are no published data in the low-energy regime, where the translational mode lies to confirm the theoretical predictions.

Acknowledgements

I want to thank all the people who supported me during the last three years. In particular, a special thanks goes to Lorenzo Monacelli for the scientific help.

Appendix A

Stress tensor

Here we derive the new equation for the stress tensor reported in the main text (Eq. (2.19)).

First we note that the quantum statistical averages taken with the trial density matrix can be written as

$$\langle O \rangle_{\tilde{\rho}_{\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}}} = \int O(\boldsymbol{\mathcal{R}} + \mathbf{J}\boldsymbol{y}, \{\boldsymbol{a}_i\})\,[dy]\,, \tag{A.1}$$

where

$$[dy] = \prod_{\mu} \frac{\exp\left(\frac{-y_\mu^2}{2}\right)}{\sqrt{2\pi}} dy_\mu. \tag{A.2}$$

This is obtained rewriting Eq. (2.20) applying the

$$u_a = \sum_{\mu} J_\mu^a y_\mu \tag{A.3}$$

change of variables, with

$$J_\mu^a = \frac{e_\mu^a}{\sqrt{M_a}} \sqrt{\frac{\hbar(1+n_\mu)}{2\omega_\mu}}. \tag{A.4}$$

Let's note that

$$\Psi_{ab} = \sum_{\mu} J_\mu^a J_\mu^b \tag{A.5}$$

after this change of variables. Note that in Eq. (A.1) we explicitly indicate the dependence of the operator O on the lattice parameters $\{\boldsymbol{a}_i\}$. Thus, in that equation, centroid positions $\boldsymbol{\mathcal{R}}$ refer only to the internal degrees of freedom of the crystal structure.

When calculating the stress tensor from Eq. (2.19) we are deriving the free energy functional in the minimum of the SSCHA free energy with respect to the auxiliary force constants $\boldsymbol{\Phi}$ for given centroid positions $\boldsymbol{\mathcal{R}}$. Thus, the stress tensor should be calculated considering the derivatives

$$\frac{\partial \mathcal{F}(\boldsymbol{\mathcal{R}})}{\partial \varepsilon_{\alpha\beta}} = \sum_{i=1}^{3} \left[\frac{\partial \mathcal{F}(\boldsymbol{\mathcal{R}})}{\partial \boldsymbol{a}_i} \cdot \frac{\partial \boldsymbol{a}_i}{\partial \varepsilon_{\alpha\beta}} + \frac{\partial \mathcal{F}(\boldsymbol{\mathcal{R}})}{\partial \boldsymbol{\mathcal{R}}} \cdot \frac{\partial \boldsymbol{\mathcal{R}}}{\partial \varepsilon_{\alpha\beta}} \right]. \tag{A.6}$$

The final equation of the strain should however be calculated for the minimum of the free energy with respect to the centroid positions, in which case the second addend above vanishes. Therefore we will just give the expression for the equilibrium situation:

$$\frac{\partial \mathcal{F}}{\partial \varepsilon_{\alpha\beta}} = \sum_{i=1}^{3} \frac{\partial \mathcal{F}}{\partial \boldsymbol{a}_i} \cdot \frac{\partial \boldsymbol{a}_i}{\partial \varepsilon_{\alpha\beta}}. \tag{A.7}$$

This expression coincides with the one usually employed to compute the stress tensor from the BO energy surface, but with the $V(\boldsymbol{R})$ potential substituted by the anharmonic free energy.

Let us write the free energy at the minimum as

$$\mathcal{F}(\boldsymbol{\mathcal{R}}) = F_{\boldsymbol{\Phi}_{\boldsymbol{\mathcal{R}}}} + \left\langle V(\boldsymbol{R}) - \mathcal{V}_{\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}_{\boldsymbol{\mathcal{R}}}} \right\rangle_{\rho_{\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}_{\boldsymbol{\mathcal{R}}}}}, \tag{A.8}$$

where $\boldsymbol{\Phi}_{\boldsymbol{\mathcal{R}}}$ is the dynamical matrix that minimizes $\mathcal{F}$ fixing the average atomic positions. $F_{\boldsymbol{\Phi}_{\boldsymbol{\mathcal{R}}}} = \left\langle K + \mathcal{V}_{\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}_{\boldsymbol{\mathcal{R}}}} \right\rangle_{\rho_{\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}_{\boldsymbol{\mathcal{R}}}}}$ and

$$\mathcal{V}_{\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}_{\boldsymbol{\mathcal{R}}}} = \frac{1}{2} (\boldsymbol{R} - \boldsymbol{\mathcal{R}}) \cdot \boldsymbol{\Phi}_{\boldsymbol{\mathcal{R}}} \cdot (\boldsymbol{R} - \boldsymbol{\mathcal{R}}). \tag{A.9}$$

The first term in Eq. (A.8) does not give any contribution to the derivative (as it depends on $\boldsymbol{\mathcal{R}}$ through $\boldsymbol{\Phi}_{\boldsymbol{\mathcal{R}}}$, which minimizes already the free energy). Therefore, the only therm that survives in the stress tensor is

$$\frac{\partial \mathcal{F}(\boldsymbol{\mathcal{R}})}{\partial \varepsilon_{\alpha\beta}} = \sum_{i=1}^{3} \frac{\partial \left\langle V(\boldsymbol{R}) - \mathcal{V}_{\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}_{\boldsymbol{\mathcal{R}}}} \right\rangle_{\rho_{\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}_{\boldsymbol{\mathcal{R}}}}}}{\partial \boldsymbol{a}_i} \cdot \frac{\partial \boldsymbol{a}_i}{\partial \varepsilon_{\alpha\beta}}. \tag{A.10}$$

Joining Eq. (A.10) with Eq. (A.2) we can compute the derivative of an average in the SSCHA ensemble with respect to the strain:

$$\begin{aligned} \frac{\partial \left\langle O \right\rangle_{\rho_{\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}_{\boldsymbol{\mathcal{R}}}}}}{\partial \varepsilon_{\alpha\beta}} &= \frac{\partial}{\partial \varepsilon_{\alpha\beta}} \int O\left(\boldsymbol{\mathcal{R}}(\boldsymbol{\varepsilon}) + \boldsymbol{J}\boldsymbol{y}, \{\boldsymbol{a}_i(\varepsilon)\}\right) [dy] = \\ &= \int \sum_{i=1}^{3} \frac{\partial O}{\partial \boldsymbol{a}_i} \cdot \frac{\partial \boldsymbol{a}_i}{\partial \varepsilon_{\alpha\beta}} [dy], \end{aligned} \tag{A.11}$$

$$\frac{\partial \left\langle O \right\rangle_{\rho_{\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}_{\boldsymbol{\mathcal{R}}}}}}{\partial \varepsilon_{\alpha\beta}} = \left\langle \sum_{i=1}^{3} \frac{\partial O}{\partial \boldsymbol{a}_i} \cdot \frac{\partial \boldsymbol{a}_i}{\partial \varepsilon_{\alpha\beta}} \right\rangle_{\rho_{\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}_{\boldsymbol{\mathcal{R}}}}}. \tag{A.12}$$

Replacing O by the BO energy landscape $V(\boldsymbol{R})$ we get

$$\begin{aligned} \frac{\partial \left\langle V \right\rangle_{\rho_{\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}_{\boldsymbol{\mathcal{R}}}}}(\boldsymbol{R})}{\partial \varepsilon_{\alpha\beta}} &= \left\langle \sum_{i=1}^{3} \frac{\partial V(\boldsymbol{R})}{\partial \boldsymbol{a}_i} \cdot \frac{\partial \boldsymbol{a}_i}{\partial \varepsilon_{\alpha\beta}} \right\rangle_{\rho_{\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}_{\boldsymbol{\mathcal{R}}}}} = \\ &= -\Omega_{\mathrm{Vol}} \left\langle P^{(\mathrm{BO})}_{\alpha\beta}(\boldsymbol{R}) \right\rangle_{\rho_{\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}_{\boldsymbol{\mathcal{R}}}}}. \end{aligned} \tag{A.13}$$

The term with the harmonic potential $\mathcal{V}$ can be derived writing its explicit dependence on the strain tensor $\boldsymbol{\varepsilon}$:

$$\mathcal{V}_{\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}_{\boldsymbol{\mathcal{R}}}}(\varepsilon) = \sum_{st} \frac{1}{2} \left[(\mathbb{1} + \boldsymbol{\varepsilon}) \cdot (\boldsymbol{R} - \boldsymbol{\mathcal{R}})_s\right] \cdot \boldsymbol{\Phi}_{\boldsymbol{\mathcal{R}}st} \left[(\mathbb{1} + \boldsymbol{\varepsilon}) \cdot (\boldsymbol{R} - \boldsymbol{\mathcal{R}})_t\right], \tag{A.14}$$

where the dot product is assumed in this equation only in the Cartesian indexes and st are atomic labels. From this equation we immediately can write the derivative

$$\left.\frac{\partial \mathcal{V}_{\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}_{\boldsymbol{\mathcal{R}}}}}{\partial \varepsilon_{\alpha\beta}}\right|_{\varepsilon=0} = -\frac{1}{2}\sum_s \left(u_s^\alpha (\mathbf{f}^{\mathcal{H}_{\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}_{\boldsymbol{\mathcal{R}}}}})_s^\beta + u_s^\beta (\mathbf{f}^{\mathcal{H}_{\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}_{\boldsymbol{\mathcal{R}}}}})_s^\alpha\right). \tag{A.15}$$

From which we obtain

$$\begin{aligned}\left\langle \left.\frac{\partial \mathcal{V}_{\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}_{\boldsymbol{\mathcal{R}}}}}{\partial \varepsilon_{\alpha\beta}}\right|_{\varepsilon=0} \right\rangle_{\rho_{\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}_{\boldsymbol{\mathcal{R}}}}} &= \\ -\frac{1}{2}\sum_s \left\langle \left(u_s^\alpha (\mathbf{f}^{\mathcal{H}_{\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}_{\boldsymbol{\mathcal{R}}}}})_s^\beta + u_s^\beta (\mathbf{f}^{\mathcal{H}_{\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}_{\boldsymbol{\mathcal{R}}}}})_s^\alpha\right)\right\rangle_{\rho_{\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}_{\boldsymbol{\mathcal{R}}}}}&.\end{aligned} \tag{A.16}$$

Combining Eqs. (A.13) and (A.16) with the definition of the stress tensor, it is trivial to get Eq. (2.19).

Appendix B

Gradient equation

The gradient equation presented here in Eq. (2.13) can be obtained starting from the

$$\frac{\partial \mathcal{F}}{\partial \Phi_{cd}} = \frac{1}{2} \sum_{ab} \frac{\partial \Psi_{ab}}{\partial \Phi_{cd}} \left[\left\langle \frac{\partial^2 V(\boldsymbol{R})}{\partial R_a \partial R_b} \right\rangle_{\tilde{\rho}_{\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}}} - \Phi_{ab} \right] \tag{B.1}$$

equation obtained in Ref. [56]. By using the

$$\left\langle \frac{\partial O(\boldsymbol{R})}{\partial R_a} \right\rangle_{\tilde{\rho}_{\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}}} = \sum_b \Psi^{-1}_{ab} \left\langle u^b O(\boldsymbol{R}) \right\rangle_{\tilde{\rho}_{\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}}} \tag{B.2}$$

result proved in the same reference, we have

$$\begin{aligned} \left\langle \frac{\partial^2 V(\boldsymbol{R})}{\partial R_a \partial R_b} \right\rangle_{\tilde{\rho}_{\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}}} &= \sum_e \Psi^{-1}_{ae} \left\langle u^e \frac{\partial V(\boldsymbol{R})}{\partial R_b}(\boldsymbol{R}) \right\rangle_{\tilde{\rho}_{\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}}} = \\ &= -\sum_e \Psi^{-1}_{ae} \left\langle u^e \mathrm{f}^{(\mathrm{BO})}_b(\boldsymbol{R}) \right\rangle_{\tilde{\rho}_{\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}}} . \end{aligned} \tag{B.3}$$

Anologously,

$$\Phi_{ab} = -\sum_e \Psi^{-1}_{ae} \left\langle u^e \mathrm{f}_b^{\mathcal{H}_{\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}}}(\boldsymbol{R}) \right\rangle_{\tilde{\rho}_{\boldsymbol{\mathcal{R}},\boldsymbol{\Phi}}} . \tag{B.4}$$

Substituting Eq. (B.3) and (B.4) into Eq. (B.1) we get Eq. (2.13).

In Ref. [56] it was also shown that

$$\frac{1}{2} \sum_{ab} \frac{\partial \Psi_{ab}}{\partial \Phi_{cd}} A_{ab} = \sum_{ab} \frac{\Lambda_{abcd}[0]}{\sqrt{M_a M_b M_c M_d}} A_{ab}, \tag{B.5}$$

where $\mathbf{A}$ is a symmetric matrix. This also proves Eq. (2.17).

Appendix C

The Hessian preconditioner

In Ref. [57] it was shown that

$$\frac{\partial^2 \mathcal{F}}{\partial \Phi_{ab} \partial \Phi_{cd}} = \frac{\Lambda_{abcd}[0]}{\sqrt{M_a M_b M_c M_d}}. \tag{C.1}$$

However, due to the relationship in Eq. (B.5), we can see that we can effectively extend this equality to

$$\frac{\partial^2 \mathcal{F}}{\partial \Phi_{ab} \partial \Phi_{cd}} = \frac{\Lambda_{abcd}[0]}{\sqrt{M_a M_b M_c M_d}} = \frac{1}{2} \frac{\partial \Psi_{ab}}{\partial \Phi_{cd}}. \tag{C.2}$$

With the latter result, it is trivial to see how the preconditioned gradient that is used along the minimization can be written as in Eq. (2.23).

Appendix D

Symmetries

The original algorithm proposed to account for symmetry in Ref. [55] was based on the Gram-Schmidt orthonormalization of the symmetry generators. This algorithm is suited for systems with a reduced number of atoms in the unit cell, but scales as n_a^6, with n_a the number of atoms in the unit cell. This symmetrization procedure becomes thus a real bottleneck of the SSCHA code for systems with more than 30 atoms in the unit cell. In the version of the code we describe here, the orthonormalized generators are not calculated and, instead, the starting dynamical matrix and the gradient are directly symmetrized. The symmetrization of the dynamical matrix, or its gradient, is made in $\mathbf{q}$ space, which allows for a very fast implementation even for big supercells.

The code enforces all the symmetries in the auxiliary force constant matrix as

$$\boldsymbol{\Phi}(\mathbf{q}) = \frac{1}{N_S} \sum_{i=1}^{N_S} \mathbf{T}_{\hat{S}_i}(S_i^{-1}\mathbf{q}) \boldsymbol{\Phi}(S_i^{-1}\mathbf{q}) \mathbf{T}_{\hat{S}_i}^{\dagger}(S_i^{-1}\mathbf{q}), \tag{D.1}$$

where S_i are the 3×3 point group matrices of the space group, N_S the number of symmetries of the crystal, $\mathbf{T}_{\hat{S}}(\mathbf{q})$ are unitary matrices that represent the S symmetry in the $\mathbf{q}$ point. These matrices are reported in Refs. [178, 179, 180]. To find the symmetries given the structure, we wrapped into the SSCHA the symmetry module of Quantum ESPRESSO[86, 48].

This operation is performed also on the gradient of the dynamical matrix each time it is computed. Since the dynamical matrices satisfying the symmetries define a linear subspace, if both the gradient and the original dynamical matrix belong to this subspace, any linear combination of them will also satisfy the symmetry constrains. Thus, it is necessary to symmetrize the dynamical matrix once at the beginning, and then apply the symmetry constrains only to the gradient to preserve the symmetries in the whole simulation.

The symmetry module from Quantum ESPRESSO only recognizes symmetries when the unit cell is the primitive one. Sometimes, it could be convenient to choose a different unit cell. Therefore, we also interfaced the SSCHA code with the *spglib* package[181] to improve the identification of symmetries. Instead of working in the unit cell in $\mathbf{q}$ space, spglib provides the symmetry operations in real space. In this case, the SSCHA code divides the symmetry matrices identified by spglib into pure translations and point group operations. Then, symmetries are enforced in real

space by first imposing pure translations, followed by point group operations as

$$\boldsymbol{\Phi} = \frac{1}{N_S} \sum_{i=1}^{N_S} \mathbf{T}_{\hat{S}_i} \boldsymbol{\Phi} \mathbf{T}_{\hat{S}_i}^{\dagger}. \tag{D.2}$$

Then, the permutation symmetry on the indices is imposed. Finally, the code transforms back the real space dynamical matrix (or the gradient) in $\mathbf{q}$ space.

This operation takes more time than the symmetrization in $\mathbf{q}$ space, as it is performed in the supercell. However, due to its simplicity and to avoid the cumbersome $\mathbf{q}$-space symmetrization of higher-order force constants, the same supercell approach is used to symmetrize the third- and fourth-order force constant matrices, namely $\overset{(3)}{\Phi}$ and $\overset{(4)}{\Phi}$ introduced in Sec. 2.4.2.

Symmetries are also enforced on the average positions of the nuclei (and the forces). After computing the SSCHA forces on atom t along direction α, f_t^{α}, we impose symmetry as

$$\mathrm{f}_t^{\alpha} = \frac{1}{N_S} \sum_{i=1}^{N_S} \sum_{t=1}^{n_a} \sum_{\beta=1}^{3} S_{i\alpha\beta} \mathrm{f}_{S^{-1}(t)}^{\beta}, \tag{D.3}$$

where $S^{-1}(t)$ is the atom in which the S symmetry maps the t one to. In this way, the forces are correctly directed only along the Wyckoff coordinates, and the atomic positions relax subsequently keeping the correct Wyckoff positions.

D.1 Acoustic sum rule on the auxiliary force constants

Besides space group symmetries, also the acoustic sum rule (ASR) must be imposed. The acoustic sum rule is a condition that arises from the momentum conservation (the center of mass of the system is fixed). The energy must not change after a rigid translation of the whole system. This can be translated in a trivial condition for the force constant matrix in the supercell:

$$\sum_t \Phi_{st}^{\alpha\beta} = \sum_s \Phi_{st}^{\alpha\beta} = 0. \tag{D.4}$$

In general, the SSCHA gradient computed from a finite ensemble violates this condition due to the stochastic noise. We enforce the sum rule on the gradient at each step. As for the symmetries, also matrices that satisfy the acoustic sum rule define a linear subspace. Thus we define the orthogonal projector operator that enforces the acoustic sum rule as

$$\boldsymbol{\Phi}^{(asr)} = \boldsymbol{P} \boldsymbol{\Phi} \boldsymbol{P}^{\dagger}. \tag{D.5}$$

The projection matrix in real space is

$$P_{st}^{\alpha\beta} = \delta_{st}\delta_{\alpha\beta} - \frac{\delta_{\alpha\beta}}{n_a} \sum_{u=1}^{n_a} \delta_{tu}. \tag{D.6}$$

This operation only affects the dynamical matrix at Γ. The same projector is employed to impose the ASR on the forces:

$$\mathrm{f}^{(asr)}{}_s^{\alpha} = (\boldsymbol{P} f)_s^{\alpha} = \mathrm{f}_s^{\alpha} - \frac{1}{n_a} \sum_{k=1}^{n_a} \mathrm{f}_k^{\alpha} \tag{D.7}$$

Notably, it can be proved that this procedure does not spoil the symmetrization described above.

This ASR imposition procedure analytically cancels out the frequencies of acoustic modes at Γ and any rigid translation of the atomic positions, thus it is the most indicated for the SSCHA minimization. A different approach, implemented for the Fourier interpolation, is described in Appendix E.2. The latter affects not just phonons at Γ, and, thus, it is more suited for interpolating dynamical matrices close to the Brillouin zone center.

Appendix E

Reciprocal space formalism and Fourier interpolation

E.1 Reciprocal space formalism

The SSCHA code is designed to be used with crystals, thus it takes advantage of lattice periodicity and Fourier transforms the relevant quantities with respect to the lattice vectors. That allows to make independent analysis for each $\boldsymbol{q}$ point in reciprocal space. When we need to stress this aspect, we will modify the notation adopted, partitioning the supercell atomic index into a unit-cell index plus a lattice index $(s, \boldsymbol{l})$, with s now ranging from 1 to n_a (the number of atoms in the unit cell), and $\boldsymbol{l}$ being a 3 dimensional integer vector assuming N_c total values (the numer of unit cells forming the supercell). Thus, in this notation, in general we will have n-th order tensors in a $3n_\mathrm{a}$ dimensional space (indicated with bold symbols, in free-component notation), which in real space depend on n lattice-vector parameters, $\overset{(n)}{\boldsymbol{D}}(\boldsymbol{l}_1, \dots \boldsymbol{l}_n)$ (to be precise, due to the translation symmetry, this real-space tensor actually depends only on $n-1$ independent values $\boldsymbol{l}_i$). The reciprocal-space expression of such a tensor, $\overset{(n)}{\boldsymbol{D}}(\boldsymbol{q}_1, \dots, \boldsymbol{q}_n)$, is obtained through the Fourier transform

$$\overset{(n)}{\boldsymbol{D}}(\boldsymbol{q}_1, \dots, \boldsymbol{q}_n) = \frac{1}{N_\mathrm{c}} \sum_{\boldsymbol{l}_1 \dots \boldsymbol{l}_n} e^{i \sum_h \boldsymbol{q}_h \cdot \boldsymbol{l}_h} \overset{(n)}{\boldsymbol{D}}(\boldsymbol{l}_1, \dots, \boldsymbol{l}_n). \tag{E.1}$$

Notice that, due to the lattice translation symmetry, $\boldsymbol{D}^{(\mathrm{S})}(\boldsymbol{q}_1, \dots, \boldsymbol{q}_n)$ is zero unless $\sum_h \boldsymbol{q}_h$ is a reciprocal lattice vector, thus we have again only $n-1$ independent parameters $\boldsymbol{q}_i$. In particular, after the calculation performed on a real-space supercell, for each $\boldsymbol{q}$ point of the commensurate grid of the reciprocal-space unit cell, the SSCHA code computes the Fourier transformed matrices $\boldsymbol{D}^{(\mathrm{S})}(-\boldsymbol{q}, \boldsymbol{q})$, which we will shortly indicate as $\boldsymbol{D}^{(\mathrm{S})}(\boldsymbol{q})$, and the relative eigenvalues $\omega_\mu(\boldsymbol{q})$ and eigenvectors $\boldsymbol{e}_\mu(\boldsymbol{q})$. Similarly, the Hessian calculation provides the matrix $\boldsymbol{D}^{(\mathrm{F})}(-\boldsymbol{q}, \boldsymbol{q})$, which we indicate as $\boldsymbol{D}^{(\mathrm{F})}(\boldsymbol{q})$, where (see Eq. (2.61))

$$\boldsymbol{D}^{(\mathrm{F})}(\boldsymbol{q}) = \boldsymbol{D}^{(\mathrm{S})}(\boldsymbol{q}) + \boldsymbol{\Pi}(\boldsymbol{q}, 0)\,, \tag{E.2}$$

and its eigenvalues $\Omega_\mu(\boldsymbol{q})$ and eigenvectors $\boldsymbol{f}_\mu(\boldsymbol{q})$.

E.2 Fourier interpolation: centering and acoustic sum rule

The SSCHA code computes the FCs in real space supercells with periodic boundary conditions (PBCs). As shown in the previous section, a crucial feature of the SSCHA code is the use of the Fourier interpolation technique in order to extrapolate the results to the thermodynamic limit (infinite supercell results) without recurring to expensive large supercell calculations. In order to Fourier interpolate the computed FCs on arbitrary points of the reciprocal space, as a first thing it is necessary to reconstruct the real-space infinite-crystal FCs from them. Roughly speaking, this is done by removing the PBCs, i.e. superlattice equivalent atoms are not considered identical anymore, and assuming that only the FCs between atoms in the same supercell are different from zero. Of course, this gives correct results as long as the supercell used in the calculations is large enough to consider negligible the FCs between atoms at distances comparable with the distances between the periodic boundary replica. However, an intrinsic arbtrariness is present in this recipe, due to the fact that the supercell is not univocally defined and the choice of different supercells leads to different interpolation results (i.e. as long as the reciprocal-space point in which we are interpolating does not belong to the original commensurate grid, different - yet superlattice equivalent - lattice points give different contributions to the Fourier transform). This problem is solved by wisely selecting the supercell according to a prescription based on a physical principle: among equivalent superlattice points, the ones closest to each other must be selected. This procedure defines the so called "centering" of the FCs and, as explained, it is a necessary step to be done before Fourier interpolating the real space FCs. The SSCHA code centers 2nd and 3rd order FCs (with a procedure that can be generalized to any nth-order FCs. In particular, the next release of the code will apply the same procedure to center and interpolate the 4th order FCs). Here we explictly describe the 3rd FCs centering algorithm [182].

The PBCs are defined on a superlattice $\mathscr{R}_{\mathrm{lat}}^{(\mathrm{S})}$ of the original lattice $\mathscr{R}_{\mathrm{lat}}$. The lattice vectors set $\mathscr{R}_{\mathrm{lat}}$ can be equivalently described as the superlattice $\mathscr{R}_{\mathrm{lat}}^{(\mathrm{S})}$ plus the basis given by the lattice vectors in a superlattice unit cell SC. In other words, a lattice vector $\boldsymbol{l} \in \mathscr{R}_{\mathrm{lat}}$ indentifies a set of superlattice-equivalent lattice vectors $\{\boldsymbol{l} + \boldsymbol{T} \text{ with } \boldsymbol{T} \in \mathscr{R}_{\mathrm{lat}}^{(\mathrm{S})}\}$, and we have $\mathscr{R}_{\mathrm{lat}} = \{\boldsymbol{l} + \boldsymbol{T} \text{ with } \boldsymbol{l} \in SC\,, \boldsymbol{T} \in \mathscr{R}_{\mathrm{lat}}^{(\mathrm{S})}\}$. Given three atoms s_1, s_2, s_3 in the unit cells $\mathbf{0}, \boldsymbol{l}_2, \boldsymbol{l}_3$, respectively (due to the lattice translation symmetry we can confine the first atom to the origin unit cell), they indentify a triangle with vertices in $\boldsymbol{\tau}_{s_1}, \boldsymbol{\tau}_{s_2} + \boldsymbol{l}_2, \boldsymbol{\tau}_{s_3} + \boldsymbol{l}_3$ ($\boldsymbol{\tau}_{s_i}$ is the position vector of atom s_i in the original unit cell). For these three points we define the weight $\mathcal{W}_{s_1 s_2 s_3}(\mathbf{0}, \boldsymbol{l}_2, \boldsymbol{l}_3)$ in this way: it is zero if there is at least another "equivalent-vertices" triangle having as vertices points $\boldsymbol{\tau}_{s_1}$, $\boldsymbol{\tau}_{s_2} + \boldsymbol{l}_2 + \boldsymbol{T}_2$, $\boldsymbol{\tau}_{s_3} + \boldsymbol{l}_3 + \boldsymbol{T}_3$ with $\boldsymbol{T}_2, \boldsymbol{T}_3 \in \mathscr{R}_{\mathrm{lat}}^{(\mathrm{S})}$ (i.e. points that are superlattice-equivalent to $\boldsymbol{\tau}_{s_1}, \boldsymbol{\tau}_{s_2} + \boldsymbol{l}_2, \boldsymbol{\tau}_{s_3} + \boldsymbol{l}_3$) with smaller perimeter, otherwise it is the inverse of the number of equivalent-vertices triangles having the same (minimal) perimeter. In formulas, indicated with $\mathcal{P}_{s_1 s_2 s_3}(\mathbf{0}, \boldsymbol{l}_2 + \boldsymbol{T}_2, \boldsymbol{l}_3 + \boldsymbol{T}_3)$ the perimeter of the triangle with vertices $\boldsymbol{\tau}_{s_1}$, $\boldsymbol{\tau}_{s_2} + \boldsymbol{l}_2 + \boldsymbol{T}_2$, $\boldsymbol{\tau}_{s_3} + \boldsymbol{l}_3 + \boldsymbol{T}_3$, this amounts to

$$\mathcal{W}_{s_1s_2s_3}(\mathbf{0},\boldsymbol{l}_2,\boldsymbol{l}_3) = \begin{cases} 0 & \text{if } \begin{array}{c}\exists\, \boldsymbol{T}_2,\boldsymbol{T}_3 \in \mathscr{R}_{\text{lat}}^{(\text{S})}: \\ \mathcal{P}_{s_1s_2s_3}(\mathbf{0},\boldsymbol{l}_2+\boldsymbol{T}_2,\boldsymbol{l}_3+\boldsymbol{T}_3) < \mathcal{P}_{s_1s_2s_3}(\mathbf{0},\boldsymbol{l}_2,\boldsymbol{l}_3)\end{array} \\ \left[\begin{array}{c}\#(\boldsymbol{T}_2,\boldsymbol{T}_3) \in \mathscr{R}_{\text{lat}}^{(\text{S})}: \\ \mathcal{P}_{s_1s_2s_3}(\mathbf{0},\boldsymbol{l}_2+\boldsymbol{T}_2,\boldsymbol{l}_3+\boldsymbol{T}_3) = \mathcal{P}_{s_1s_2s_3}(\mathbf{0},\boldsymbol{l}_2,\boldsymbol{l}_3)\end{array}\right]^{-1} & \text{if } \begin{array}{c}\nexists\, \boldsymbol{T}_2,\boldsymbol{T}_3 \in \mathscr{R}_{\text{lat}}^{(\text{S})}: \\ \mathcal{P}_{s_1s_2s_3}(\mathbf{0},\boldsymbol{l}_2+\boldsymbol{T}_2,\boldsymbol{l}_3+\boldsymbol{T}_3) < \mathcal{P}_{s_1s_2s_3}(\mathbf{0},\boldsymbol{l}_2,\boldsymbol{l}_3)\end{array} \end{cases} \tag{E.3}$$

The weights $\mathcal{W}_{s_1s_2s_3}(\mathbf{0},\boldsymbol{l}_2,\boldsymbol{l}_3)$ are pure geometrical factors, different from zero for "compact" three-atom clusters, and they satisfy the normalization

$$\sum_{\boldsymbol{T}_2,\boldsymbol{T}_3\in\mathscr{R}_{\text{lat}}^{(\text{S})}} \mathcal{W}_{s_1s_2s_3}(\mathbf{0},\boldsymbol{l}_2+\boldsymbol{T}_2,\boldsymbol{l}_3+\boldsymbol{T}_3) = 1 \qquad \begin{array}{l}\forall\, \boldsymbol{l}_2,\boldsymbol{l}_3 \in \mathscr{R}_{\text{lat}} \\ \forall\, s_1,s_2,s_3 \in \{1,\dots,n_{\text{a}}\}\,.\end{array} \tag{E.4}$$

The weights are used to define the centering. Given a 3rd-order FCs, $\Phi^{\alpha_1\alpha_2\alpha_3}_{s_1,s_2,s_3}(\mathbf{0},\boldsymbol{l}_2,\boldsymbol{l}_3)$, its "centered" version $\overset{(\text{cent})}{\Phi}{}^{\alpha_1\alpha_2\alpha_3}_{s_1s_2s_3}(\mathbf{0},\boldsymbol{l}_2,\boldsymbol{l}_3)$ is given by

$$\overset{(\text{cent})}{\Phi}{}^{\alpha_1\alpha_2\alpha_3}_{s_1s_2s_3}(\mathbf{0},\boldsymbol{l}_2,\boldsymbol{l}_3) = \Phi^{\alpha_1\alpha_2\alpha_3}_{s_1s_2s_3}(\mathbf{0},\boldsymbol{l}_2,\boldsymbol{l}_3) \,\times\, \mathcal{W}_{s_1s_2s_3}(\mathbf{0},\boldsymbol{l}_2,\boldsymbol{l}_3)\,, \tag{E.5}$$

where we have separately indicated cartesian (α_h) and atomic (s_h) indices. The idea behind this definition is pretty simple: once the PBCs are discarded, of the infinite set of superlattice-equivalent atoms only the"closest one" are characterized by a force constant different from zero. If there are several equivalent triplets at the minimal reciprocal distance, all of them are considered (to preserve the symmetry) and the force constants are consequently scaled (to avoid a wrong multiple counting effect). The centering definition has some degree of arbitrariness, though, due to the arbitrariness of the criterion employed to evaluate the "size" of a three atoms cluster. We took the perimeter of the triangle, a criterion that is a direct generalization of the distance between atoms, which is the one used in the 2nd order FCs centering. More in general, for an n-atoms cluster this size measure is readily generalized as the sum of the distances between all the $n(n-1)/2$ couples of atoms. However, even if in principle other choices could be done, this arbitrariness is immaterial as long as the supercell calculation is large enough (in the thermodynamic limit all the possible choices, if reasonable, are expected to be equivalent).

A delicate issue is associated with the centering: the spoiling of the acoustic sum rule (ASR). For a nth-order FC, the acoustic sum rule is

$$\sum_{\boldsymbol{l}_i}\sum_{s_i} \Phi^{\alpha_1\dots\alpha_i\dots\alpha_n}_{s_1\dots s_i\dots s_n}(\boldsymbol{l}_1,\dots,\boldsymbol{l}_i,\dots,\boldsymbol{l}_n) = 0 \qquad \begin{array}{ll}\forall\, \alpha_h \in \{x,y,z\} & \\ \forall\, s_h \in \{1\dots n_{\text{a}}\} & \text{with } h\neq i \\ \forall\, \boldsymbol{l}_h \in \mathscr{R}_{\text{lat}} & \text{with } h\neq i\end{array}\,. \tag{E.6}$$

The ASR comes is crucial, among other things, to have the correct acoustic phonon dispersion at and close to Γ. The FCs computed with SSCHA (in supercells with PBCs) fulfill the acoustic sum rule but, in general, the centering spoils it (except for

the $n = 2$ case). In fact, in general the centered FCs fulfill a "weaker" version of the ASR in Eq. (E.6), since only the simultaneous sum on $n-1$ indices is zero:

$$\sum_{\boldsymbol{l}_{i_1},\ldots,\boldsymbol{l}_{i_{n-1}}} \sum_{s_{i_1},\ldots,s_{i_{n-1}}} \overset{(\mathrm{cent})}{\Phi}{}^{\alpha_1\ldots\alpha_n}_{s_1\ldots s_n}(\boldsymbol{l}_1,\ldots,\boldsymbol{l}_n) = 0\,. \tag{E.7}$$

In particular, this explains why the centering of 2nd-order FCs does not spoil the ASR (for $n = 2$ the weak ASR is nothing but the proper ASR, as the sum over $n-1$ indices coincides with the sum over one index).

In order to see why this happens let us consider, as an example, the 3rd-order FCs case and the sum over the third index. It is:

$$\sum_{s_3} \sum_{\boldsymbol{l}_3\in\mathscr{R}_{\mathrm{lat}}} \overset{(\mathrm{cent})}{\Phi}{}^{\alpha_1\alpha_2\alpha_3}_{s_1s_2s_3}(\mathbf{0},\boldsymbol{l}_2,\boldsymbol{l}_3) = \tag{E.8}$$

$$= \sum_{s_3} \sum_{\boldsymbol{l}_3\in SC} \sum_{\boldsymbol{T}_3\in\mathscr{R}^{(\mathrm{S})}_{\mathrm{lat}}} \overset{(\mathrm{cent})}{\Phi}{}^{\alpha_1\alpha_2\alpha_3}_{s_1s_2s_3}(\mathbf{0},\boldsymbol{l}_2,\boldsymbol{l}_3+\boldsymbol{T}_3) \tag{E.9}$$

$$= \sum_{s_3} \sum_{\boldsymbol{l}_3\in SC} \sum_{\boldsymbol{T}_3\in\mathscr{R}^{(\mathrm{S})}_{\mathrm{lat}}} \Phi^{\alpha_1\alpha_2\alpha_3}_{s_1s_2s_3}(\mathbf{0},\boldsymbol{l}_2,\boldsymbol{l}_3+\boldsymbol{T}_3)\ \mathcal{W}_{s_1s_2s_3}(\mathbf{0},\boldsymbol{l}_2,\boldsymbol{l}_3+\boldsymbol{T}_3) \tag{E.10}$$

$$= \sum_{s_3} \sum_{\boldsymbol{l}_3\in SC} \sum_{\boldsymbol{T}_3\in\mathscr{R}^{(\mathrm{S})}_{\mathrm{lat}}} \Phi^{\alpha_1\alpha_2\alpha_3}_{s_1s_2s_3}(\mathbf{0},\boldsymbol{l}_2,\boldsymbol{l}_3)\ \mathcal{W}_{s_1s_2s_3}(\mathbf{0},\boldsymbol{l}_2,\boldsymbol{l}_3+\boldsymbol{T}_3) \tag{E.11}$$

$$= \sum_{s_3} \sum_{\boldsymbol{l}_3\in SC} \Phi^{\alpha_1\alpha_2\alpha_3}_{s_1s_2s_3}(\mathbf{0},\boldsymbol{l}_2,\boldsymbol{l}_3) \underbrace{\sum_{\boldsymbol{T}_3\in\mathscr{R}^{(\mathrm{S})}_{\mathrm{lat}}} \mathcal{W}_{s_1s_2s_3}(\mathbf{0},\boldsymbol{l}_2,\boldsymbol{l}_3+\boldsymbol{T}_3)}_{\text{nonconstant w.r.t. } s_3 \text{ and } \boldsymbol{l}_3}\,. \tag{E.12}$$

Since the last factor, highlighted with a brace under, in general is not constant with respect to s_3 and $\boldsymbol{l}_3$, it cannot be factored out from the sums, so that the ASR for the original $\Phi^{\alpha_1\alpha_2\alpha_3}_{s_1s_2s_3}(\mathbf{0},\boldsymbol{l}_2,\boldsymbol{l}_3)$

$$\sum_{s_3} \sum_{\boldsymbol{l}_3\in SC} \Phi^{\alpha_1\alpha_2\alpha_3}_{s_1s_2s_3}(\mathbf{0},\boldsymbol{l}_2,\boldsymbol{l}_3) = 0 \tag{E.13}$$

cannot be used to obtain the ASR for the centered $\overset{(\mathrm{cent})}{\Phi}{}^{\alpha_1\alpha_2\alpha_3}_{s_1s_2s_3}(\mathbf{0},\boldsymbol{l}_2,\boldsymbol{l}_3)$. However, using Eq. (E.4), with similar passages we can show that the sum over the last two

indices is zero:

$$\sum_{s_2,s_3}\sum_{\boldsymbol{l}_2,\boldsymbol{l}_3\in\mathscr{R}_{\mathrm{lat}}} \overset{(\mathrm{cent})}{\Phi}{}^{\alpha_1\alpha_2\alpha_3}_{s_1s_2s_3}(\boldsymbol{0},\boldsymbol{l}_2,\boldsymbol{l}_3) = \tag{E.14}$$

$$= \sum_{s_2,s_3}\sum_{\boldsymbol{l}_2,\boldsymbol{l}_3\in SC} \Phi^{\alpha_1\alpha_2\alpha_3}_{s_1s_2s_3}(\boldsymbol{0},\boldsymbol{l}_2,\boldsymbol{l}_3) \underbrace{\sum_{\boldsymbol{T}_2,\boldsymbol{T}_3\in\mathscr{R}^{(\mathrm{S})}_{\mathrm{lat}}} \mathcal{W}_{s_1s_2s_3}(\boldsymbol{0},\boldsymbol{l}_2+\boldsymbol{T}_2,\boldsymbol{l}_3+\boldsymbol{T}_3)}_{=1} \tag{E.15}$$

$$= \sum_{s_2,s_3}\sum_{\boldsymbol{l}_2,\boldsymbol{l}_3\in SC} \Phi^{\alpha_1\alpha_2\alpha_3}_{s_1s_2s_3}(\boldsymbol{0},\boldsymbol{l}_2,\boldsymbol{l}_3) \tag{E.16}$$

$$= 0\,, \tag{E.17}$$

thus the weak ASR is fulfilled.

The spoling of the ASR after the centering dictates to impose it. In principle, there is not a unique way of doing it, as imposing the ASR on FCs simply consists in finding new FCs that fullfil the ASR and differ the least from the original FCs (according to some reasonable but arbitrary metric). In this release of the SSCHA code we impose the ASR by employing an iterative procedure, consisting of two steps [182]. First, the ASR is imposed on one index (the last one, for example). This spoils the permutation symmetry, which is consequently imposed. In general, the resulting permutation-symmetric FCs do not fulfill the ASR yet, thus this procedure is repeatedly applied until the permutation-symmetric FCs fulfill the ASR within a certain tolerance. The imposition of the permutation symmetry is a straightforward task. The ASR is imposed on the third index of a centered 3rd-order FCs by updating its values on the compact three-atom clusters that defined the centering (in order to preserve the short-sightdness of the centered FCs even after the ASR imposition). Given a centered $\Phi^{\alpha_1\alpha_2\alpha_3}_{s_1s_2s_3}(\boldsymbol{0},\boldsymbol{l}_2,\boldsymbol{l}_3)$, the $\widetilde{\Phi}^{\alpha_1\alpha_2\alpha_3}_{s_1s_2s_3}(\boldsymbol{0},\boldsymbol{l}_2,\boldsymbol{l}_3)$ that fulfills the ASR on the third index is computed with

$$\widetilde{\Phi}^{\alpha_1\alpha_2\alpha_3}_{s_1s_2s_3}(\boldsymbol{0},\boldsymbol{l}_2,\boldsymbol{l}_3) = \Phi^{\alpha_1\alpha_2\alpha_3}_{s_1s_2s_3}(\boldsymbol{0},\boldsymbol{l}_2,\boldsymbol{l}_3) - \mathcal{K}^{\alpha_1\alpha_2\alpha_3}_{s_1s_2s_3}(\boldsymbol{0},\boldsymbol{l}_2,\boldsymbol{l}_3|p) \times \sum_{\bar{s}_3,\bar{\boldsymbol{l}}_3} \Phi^{\alpha_1\alpha_2\alpha_3}_{s_1s_2\bar{s}_3}(\boldsymbol{0},\boldsymbol{l}_2,\bar{\boldsymbol{l}}_3)\,, \tag{E.18}$$

where $\mathcal{K}^{\alpha_1\alpha_2\alpha_3}_{s_1s_2s_3}(\boldsymbol{0},\boldsymbol{l}_2,\boldsymbol{l}_3|p)$ is the scaling factor

$$\mathcal{K}^{\alpha_1\alpha_2\alpha_3}_{s_1s_2s_3}(\boldsymbol{0},\boldsymbol{l}_2,\boldsymbol{l}_3|p) = \begin{cases} \dfrac{\left|\Phi^{\alpha_1\alpha_2\alpha_3}_{s_1s_2s_3}(\boldsymbol{0},\boldsymbol{l}_2,\boldsymbol{l}_3)\right|^p}{\sum_{\bar{s}_3,\bar{\boldsymbol{l}}_3}\left|\Phi^{\alpha_1\alpha_2\alpha_3}_{s_1s_2\bar{s}_3}(\boldsymbol{0},\boldsymbol{l}_2,\bar{\boldsymbol{l}}_3)\right|^p} & \text{if} \quad p=0 \quad \text{or} \quad p>0 \quad \text{and} \quad \displaystyle\sum_{\bar{s}_3,\bar{\boldsymbol{l}}_3}\left|\Phi^{\alpha_1\alpha_2\alpha_3}_{s_1s_2\bar{s}_3}(\boldsymbol{0},\boldsymbol{l}_2,\bar{\boldsymbol{l}}_3)\right| \neq 0 \\ 0 & \text{if} \quad p>0 \quad \text{and} \quad \displaystyle\sum_{\bar{s}_3,\bar{\boldsymbol{l}}_3}\left|\Phi^{\alpha_1\alpha_2\alpha_3}_{s_1s_2\bar{s}_3}(\boldsymbol{0},\boldsymbol{l}_2,\bar{\boldsymbol{l}}_3)\right| = 0 \end{cases}, \tag{E.19}$$

with p a non-negative real number which can be arbitrarily fixed to optimize the calculation performances (in the equation above the convention $0^0 = 1$ has been adopted). The $\widetilde{\Phi}^{\alpha_1\alpha_2\alpha_3}_{s_1s_2s_3}(\boldsymbol{0},\boldsymbol{l}_2,\boldsymbol{l}_3)$ defined through Eq. (E.18) fulfills the ASR on the third index, since for any $p \geq 0$ the scaling factor fulfills the normalization condition

$$\sum_{s_3,\boldsymbol{l}_3} \mathcal{K}^{\alpha_1\alpha_2\alpha_3}_{s_1s_2s_3}(\boldsymbol{0},\boldsymbol{l}_2,\boldsymbol{l}_3|p) = 1\,. \tag{E.20}$$

The value of p has effects on the way the different terms of $\Phi^{\alpha_1\alpha_2\alpha_3}_{s_1s_2s_3}(\mathbf{0},\boldsymbol{l}_2,\boldsymbol{l}_3)$ are scaled. For $p=0$ the scaling factor is a pure geometric quantity related to the three atoms clusters. Indeed, given $s_1, s_2, \boldsymbol{l}_2$, the scaling factor $\mathcal{K}^{\alpha_1\alpha_2\alpha_3}_{s_1s_2s_3}(\mathbf{0},\boldsymbol{l}_2,\boldsymbol{l}_3|p=0)$ is fully determined (it is the same for all the $\alpha_h, s_3, \boldsymbol{l}_3$) and, in particular, it does not depend on the FCs value. On the contrary, for $p\neq 0$, given $\alpha_h, s_1, s_2, \boldsymbol{l}_2$ we have

$$\frac{\mathcal{K}^{\alpha_1\alpha_2\alpha_3}_{s_1s_2s_3'}(\mathbf{0},\boldsymbol{l}_2,\boldsymbol{l}_3'|p)}{\mathcal{K}^{\alpha_1\alpha_2\alpha_3}_{s_1s_2s_3''}(\mathbf{0},\boldsymbol{l}_2,\boldsymbol{l}_3''|p)} = \left|\frac{\Phi^{\alpha_1\alpha_2\alpha_3}_{s_1s_2s_3''}(\mathbf{0},\boldsymbol{l}_2,\boldsymbol{l}_3'')}{\Phi^{\alpha_1\alpha_2\alpha_3}_{s_1s_2s_3'}(\mathbf{0},\boldsymbol{l}_2,\boldsymbol{l}_3')}\right|^p \tag{E.21}$$

so that if $p>1$ the scaling factor is higher (lower) for FCs have lower (higher) absolute value, otherwise the opposite.

E.3 Effective charges

In ionic crystals the nuclei displacement induces dipoles (proportional to the Born effective charge tensors), and this adds a dipole-dipole interaction term to the interatomic forces. This contribution, because of its long-range character (it goes as the inverse of the third power of the nuclei distances), is not suited to be Fourier interpolated and it is at the origin of the nonanalytic behavior of the dynamical matrix at Γ, with (in general anisotropic) LO-TO splitting of the phonon frequencies at BZ center. The long-range dipole-dipole contribution to the force constants can be calculated analytically since it is fully determined by the Born effective charges $(Z^*_s)^{\alpha\beta}$ (effective charge tensor of atom s) and the electronic dielectric permittivity tensor $(\epsilon^\infty)^{\alpha\beta}$, which can both be calculated from first principles. For a given $\boldsymbol{q}\in BZ$, this dipole-dipole contribution is given by [183, 184]

$$\boldsymbol{\Phi}^{(\mathrm{dd})}_{st}(\boldsymbol{q}) = \widehat{\boldsymbol{\Phi}}^{(\mathrm{dd})}_{st}(\boldsymbol{q}) - \delta_{st}\sum_{\bar{t}}\widehat{\boldsymbol{\Phi}}^{(\mathrm{dd})}_{s\bar{t}}(\boldsymbol{q}=\mathbf{0}) \tag{E.22}$$

with

$$\widehat{\boldsymbol{\Phi}}^{(\mathrm{dd})}_{st}(\boldsymbol{q}) = \frac{4\pi}{\Omega_{\mathrm{Vol}}}{\sum_{\boldsymbol{G}}}' \frac{[(\boldsymbol{G}+\boldsymbol{q})\cdot\boldsymbol{Z}^*_s]\otimes[(\boldsymbol{G}+\boldsymbol{q})\cdot\boldsymbol{Z}^*_t]}{(\boldsymbol{G}+\boldsymbol{q})\cdot\boldsymbol{\epsilon}^\infty\cdot(\boldsymbol{G}+\boldsymbol{q})}\, e^{-\frac{(\boldsymbol{G}+\boldsymbol{q})\cdot\boldsymbol{\epsilon}^\infty\cdot(\boldsymbol{G}+\boldsymbol{q})}{4\eta^2}}\, e^{i(\boldsymbol{G}+\boldsymbol{q})\cdot(\boldsymbol{\tau}_s-\boldsymbol{\tau}_t)}, \tag{E.23}$$

where we have explicitly indicated only the atomic indices (i.e. we are using component-free notation for the cartesian indices), η is a parameter whose value has to be large enough to allow to include only the reciprocal space terms in the Ewald sum, and $\sum'_{\boldsymbol{G}}$ is the sum over reciprocal lattice vectors such that $\boldsymbol{G}+\boldsymbol{q}\neq\mathbf{0}$ (the sum includes as many $\boldsymbol{G}$'s as it is necessary to reach the convergence for the considered η) [86].

Once $\boldsymbol{Z}^*_s$ and $\boldsymbol{\epsilon}^\infty$ are available, the problem caused to the Fourier interpolation by the long-range dipole-dipole interaction is thus bypassed in the SSCHA code in two steps. First, from the $\boldsymbol{\Phi}(\boldsymbol{q})$ calculated on a (coarse) grid of $\boldsymbol{q}$ point of the Brillouin zone, the corresponding dipole-dipole terms $\boldsymbol{\Phi}^{(\mathrm{dd})}(\boldsymbol{q})$ are subtracted and the resulting short range FCs is Fourier transformed to the real space. Subsequently, this real space short-range FCs, $\boldsymbol{\Phi}^{(\mathrm{sr})}(\boldsymbol{l})$, can be Fourier transformed back to any

$\boldsymbol{k} \in BZ$ and the corresponding long-range dipole-dipole analytical contribution $\boldsymbol{\Phi}^{(\mathrm{dd})}(\boldsymbol{k})$ is added [86]:

$$\underset{\text{on } BZ\ \boldsymbol{q}\text{-grid}}{\boldsymbol{\Phi}(\boldsymbol{q})} \xrightarrow[\text{Fourier transform to real space}]{\substack{\text{Subtract dipole-dipole interaction terms} \\ \boldsymbol{\Phi}^{(dd)}(\boldsymbol{q}) \\ + \\ \text{Fourier transform} \\ \text{to real space}}} \boldsymbol{\Phi}^{(\mathrm{sr})}(\boldsymbol{l}) \xrightarrow[\boldsymbol{\Phi}^{(dd)}(\boldsymbol{k})]{\substack{\text{Fourier transform} \\ \text{back to } \boldsymbol{k} \in BZ \\ + \\ \text{Add dipole-dipole interaction term}}} \boldsymbol{\Phi}(\boldsymbol{k}) \tag{E.24}$$

The dipole-dipole correction to the FCs given by Eqs. (E.22), (E.23) is nonanalytic at zone center and its $\boldsymbol{q} \to \boldsymbol{0}$ limit depends on the direction $\widehat{\boldsymbol{q}} = \boldsymbol{q}/\|\boldsymbol{q}\|$ along which the limit is performed:

$$\lim_{\delta \to 0^+} \boldsymbol{\Phi}_{st}^{(\mathrm{dd})}(\delta \widehat{\boldsymbol{q}}) = \boldsymbol{\Phi}_{st}^{(\mathrm{dd})}(\boldsymbol{0}) + \boldsymbol{\Phi}_{st}^{(\mathrm{dd\text{-}na})}(\widehat{\boldsymbol{q}})\,, \tag{E.25}$$

where

$$\boldsymbol{\Phi}_{st}^{(\mathrm{dd\text{-}na})}(\widehat{\boldsymbol{q}}) = \frac{4\pi}{\Omega_{\mathrm{Vol}}} \frac{[\widehat{\boldsymbol{q}} \cdot \boldsymbol{Z}_s^*] \otimes [\widehat{\boldsymbol{q}} \cdot \boldsymbol{Z}_t^*]}{\widehat{\boldsymbol{q}} \cdot \boldsymbol{\epsilon}^{\infty} \cdot \widehat{\boldsymbol{q}}} \tag{E.26}$$

is the nonanalytic zone-center correction term. When a phonon dispersion through Γ is calculated, the SSCHA code includes the nonanalytic correction term in the zone center, with the direction given by the followed path [86]. When the SSCHA code calculate the spectral properties (static or dynamic), it adds the nonanalytic correction term in the zone center dynamical matrix (necessary for the integral over the BZ) from a random direction.

Appendix F

Simulation details phase XI

Whatever the theoretical framework employed to describe the system, the free energy, Eq. (2.127) and Eq. (2.129) for the QHA and SSCHA respectively, clearly depend on the number N_q of q points sampled in the Brillouin zone. We employ the supercell method, which consists in replicating an integer number of times $N_1 \times N_2 \times N_3$ the unit cell along the three Cartesian directions and imposing periodic boundary conditions. In Sec. 3.3.1, we pointed out that the thermodynamic properties in the QHA are derived from the free energy, while in the SSCHA the pressure has been used, thanks to the availability of a handy analytical formula (see Eq. 2.19) [57].

Consequently, the converged supercells are chosen by looking at the free energy in the QHA and at the pressure in the SSCHA. We analyzed the convergence for the two extremal temperatures, T=0 K and T=300 K, in Fig. F.1. As a general trend, we notice that thermal fluctuations slow down the convergence of both the free energy and the pressure. The limited computational cost of the QHA doesn't place any limitation on the mesh to use.

Moreover, we have the possibility to interpolate the dynamical matrix to a finer mesh, as it has been done for the two bigger meshes, namely the $10 \times 10 \times 10$ and the $14 \times 14 \times 14$. So, by looking at Fig. F.1 (**b**) we decided to sample phonons in the biggest mesh of the Brillouin zone analyzed.

Conversely, the challenging computational cost of the SSCHA prevents the use of a big supercell. Fig. F.1 (**a**) reveals as the $3 \times 3 \times 2$ supercell grants at most 3% error with respect to the converged mesh at high temperature and a discrepancy smaller than 1% at zero temperature, where most of the results in this work are presented. We are satisfied with this accuracy level and we compute the thermodynamic properties in this $\boldsymbol{q}$-point mesh. Instead, to compute the phonon dispersion of Sec. 3.3.3, a 4x4x4 supercell has been used.

It is well known in literature [97, 151, 185] that very large simulation cells can be necessary to reproduce the thermodynamic behavior of proton disordered ice, according to the property of interest. In this perspective, we did a convergence test for ice I_h considering simulation cells containing up to 432 atoms (64 molecules). The tests for pressure and free energy are in Fig. F.2. A satisfactory accuracy is achieved when a simulation cell contains 64 water molecules, in accordance with the theoretical works cited above.

To simulate the thermodynamic properties, we computed the QHA harmonic

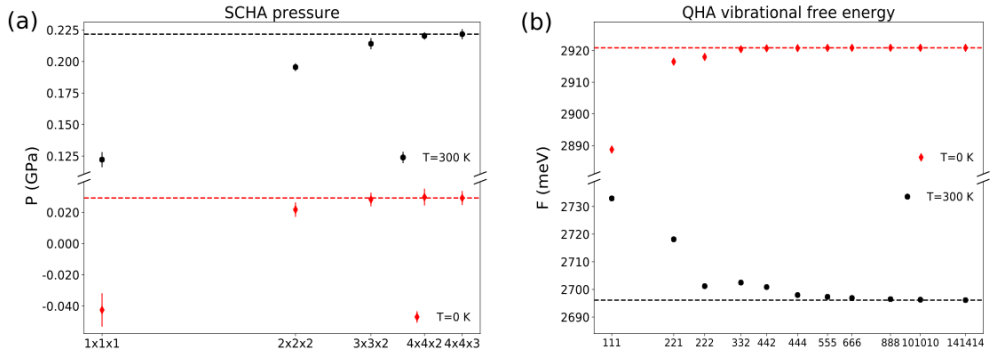

Figure F.1. Convergence tests in the QHA and SSCHA. **a** : Pressure within the SSCHA framework computed as in Eq. (2.19) as a function of the supercell dimension for T=0 K and T = 300 K **b** : Vibrational term of the QHA free energy in Eq. (2.127) as a function of the supercell dimension for T=0 K and T = 300 K. Here we use the short hard notation $N_1N_2N_3$ to indicate the $N_1 \times N_2 \times N_3$ supercell. The dashed lines in panel **a** and **b** are the converged results for the pressure or the free energy at the given temperature.

free energy (relaxing the atomic position at fixed cell) at 120 volumes. The free energy is fitted with the Vinet equation of state (EOS) [174] in Eq. (F.1), where $\eta = (V/V_0)^{1/3}$, B_0 and B'_o are the bulk modulus and its derivative with respect to pressure respectively. In this way, the equilibrium volume and the bulk modulus for any temperature are readily available.

$$P(V) = 3B_0 \frac{1-\eta}{\eta^2} \exp\left\{\frac{3}{2}(B'_0 - 1)(1-\eta)\right\} \tag{F.1}$$

In contrast, for the SSCHA, we computed the pressure as a function of temperature for six volumes. For each volume, we relaxed the atomic positions accounting for quantum and thermal anharmonic effects. We employed ensembles with as many as 100000 configurations in the converged supercell to reduce the statistical noise. We evaluated the equilibrium volume and the bulk modulus fitting the $P(\Omega)$ curve with the Vinet EOS at each temperature.

The spectral properties are computed from the dynamical one-phonon interacting Green function. After checking that high order corrections to the self-energy are negligible, we restrict to the bubble approximation of Eq. (2.59) and it is computed through Eq. (2.75). We ran convergence tests on the $\boldsymbol{k}$ integration grid and smearing factor δ_{se} that required a 14x14x14 $\boldsymbol{k}$-grid and $\delta_{se} = 45\text{cm}^{-1}$ to achieve convergence. The same values hold for H_2O and D_2O ice. Consequently, all the SSCHA spectral functions are computed with those values of smearing and integration $\boldsymbol{k}$-grid. Finite linewidths in the DOS and harmonic model are for presentation purposes only.

The Raman tensor A_{abc} of Eq. (2.117) is computed *ab initio* with Quantum ESPRESSO [86] in the SSCHA average centroid positions within the LDA approximation. The electronic wave functions were expanded in plane waves up to a cutoff of 120 Ry (960 Ry for the electronic density), with a uniform grid of 4x4x4 for the

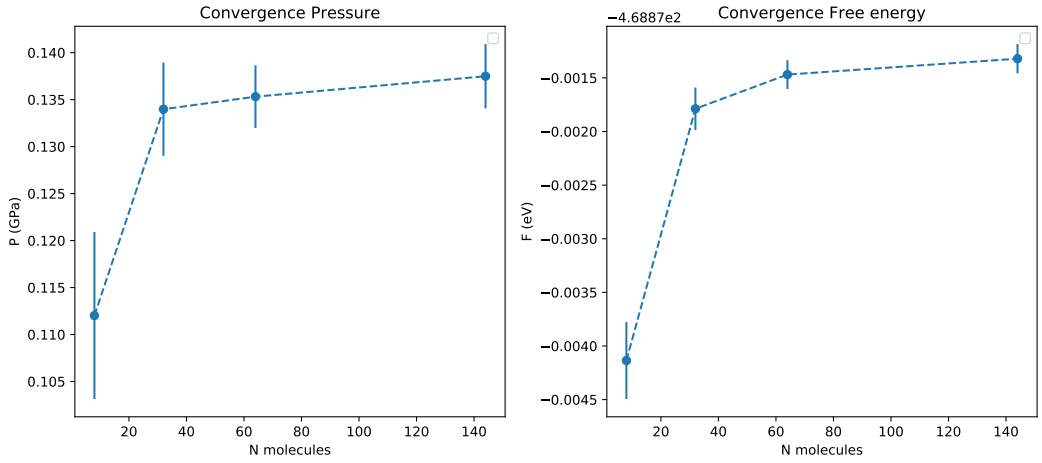

Figure F.2. Pressure (left) and free energy per molecule (right) at T=0 K for a proton-disordered realisation of ice I_h as a function of the number of molecules in the simulation cell. The dashed lines stand for the converged results.

Brillouin zone integrals. The effective charges Z_{ab} of Eq. (2.112) are computed *ab initio* in Quantum ESPRESSO [86] in LDA approximation. The electronic wave functions were expanded in plane waves up to a cutoff of 120 Ry (960 Ry for the electronic density), with a uniform grid of 4x4x4 for the Brillouin zone integrals.

Appendix G

Phonon dispersion in ice XI

Real phonon dispersion is calculated from the dynamical interacting Green function of Eq. (2.67). Here, we show the density of states (DOS) and dispersion for H_2O ice at T=200 K and D_2O ice at T=140 K and P=0.05 GPa, comparing the harmonic phonons and the SSCHA auxiliary phonons in Fig. G.1.

The harmonic dynamical matrix is computed with a 5x5x5 $\boldsymbol{q}$ mesh, while for the SSCHA I employed the 4x4x4 supercell.

As a molecular crystal, the phonon branches are well separated in translational modes, librations, narrow bending and stretching from low to high energy. In Fig. G.1 (**a**), we report the hydrogen and in Fig. G.1 (**b**) the deuterium.

In the SSCHA, the harmonic translational and rotational modes suffer a blue shift of the order of 4.2% (3.5%) and 8.2% (7.5%) respectively for H_2O (D_2O) ice. Instead, anharmonicity reduces the frequencies of the most energetic modes. Both the harmonic bending and stretching band are red-shifted of about 3.1% $(2.2-2.5\%)$ in H_2O (D_2O) ice.

The dispersion of deuterated ice under pressure in Fig. 3.10 was limited to the Γ-A direction. Here, I include the Γ-K path dispersion in Fig. G.2. In ice XI, we have 2 inequivalent Γ-K paths originated by the presence of the hydrogen sublattice, that overturns the perfect equivalence we would obtain in the presence of oxygen atoms alone.

The projection of the first Brillouin zone of ice XI in the plane is reported in the inset of Fig. G.2 (**a**). I show the planar reciprocal vectors, $\{\vec{b}_1, \vec{b}_2\}$ and the three K points. Notably, the three K points are actually equivalent (the phonon frequencies are the same) but the modes along the path connecting Γ and K aren't.

We average the three dispersion and compared the average with the experiment in Fig. G.2 (**b**). The matching between theory and experiment is very good.

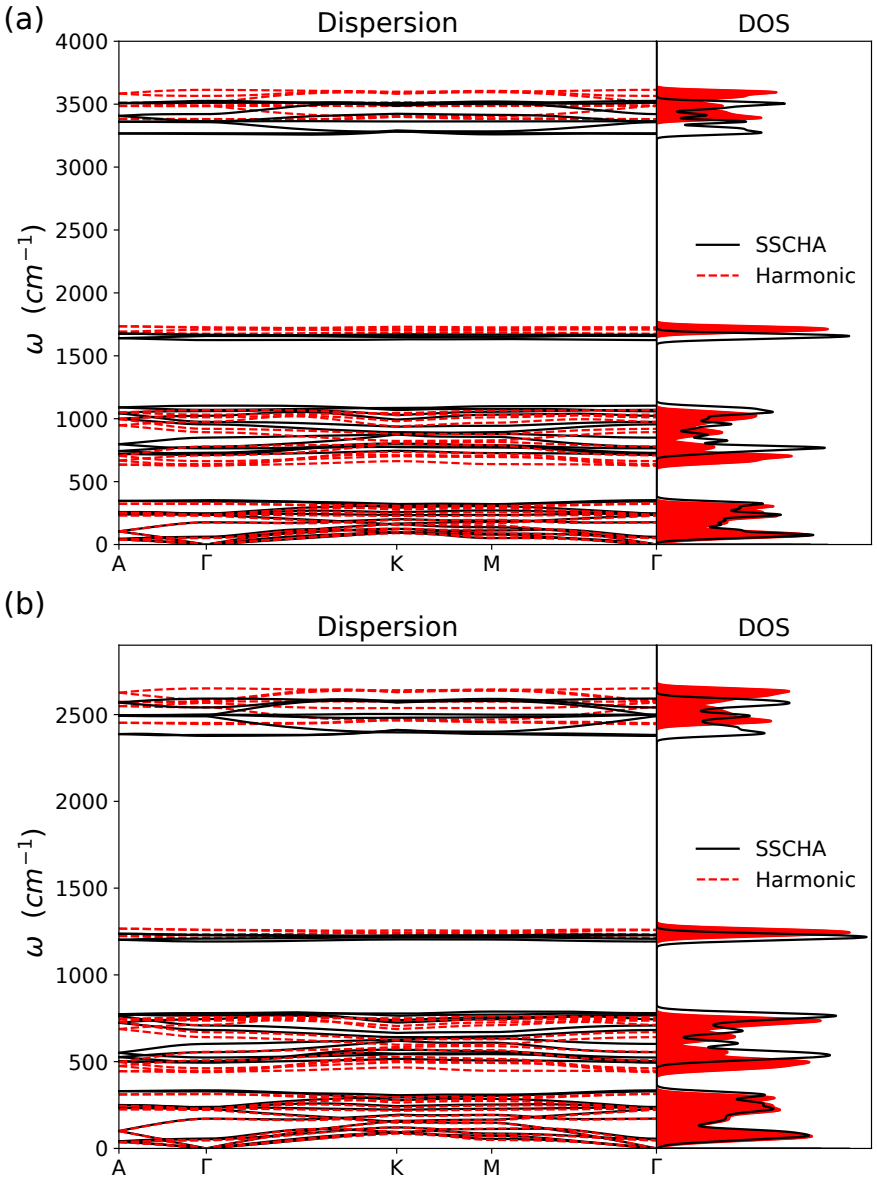

Figure G.1. Phonon dispersion and density of states. **a** H_2O ice XI at T=200 K at ambient pressure.**b** D_2O ice XI at T=140 K and P=0.05 GPa. In both panel **a** and **b**, the harmonic results (red dashed line) are compared with the dispersion and DOS in the SSCHA framework (black solid lines).

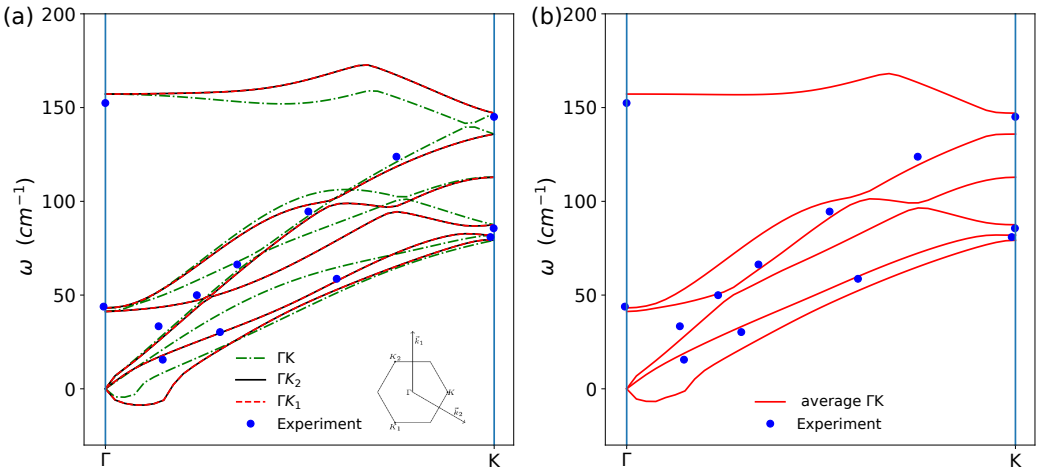

Figure G.2. Phonon dispersion in the SSCHA framework with the inclusion of the bubble term in the self-energy correction in the static limit for D_2O ice at T=140 K and P = 0.05 GPa. **a** Phonon dispersion along the three nonequivalent ΓK directions compared to the experimental measurement (blue circles) [143]. The three K points and the projection in the plane of the first Brillouin Zone are shown in the inset. **b** Comparison of the experiment with the average of the phonon dispersion along the three ΓK directions (red solid lines).

Appendix H

Simulation details for high-pressure ice

In this section, I report all the details of the numerical simulations. The simulation cell used for the high-pressure phases have been already described in Sec. 4.2, they all contain 16 water molecules, and the convergence with respect to the number of molecules for what concerns the thermodynamic properties have been tested and shown in Fig. 4.12.

The relaxation of the structures in the classical approximation at T=0 K of Sec. 4.4.1 has been performed with LAMMPS [90], by fixing the external pressure and allowing for a triclinic relaxation. The pressures sample have been discretized with steps of 10 GPa far from the critical pressure and with a refined sampling close to the phase transition, where I relaxed the structures each 1 GPa. I used strict cutoffs(10^{-18}) on the energies and forces. Energies and forces are calculated within LAMMPS with the use of the DNN described in Sec. 2.9.2, fitted on DFT-PBE. The relaxation at the *ab initio* level, using the PBE approximation of DFT are done within Quantum ESPRESSO [86]. The electronic wave functions were expanded in plane waves up to a cutoff of 120 Ry (960 Ry for the electronic density), with a uniform grid of 4×4×4 for the Brillouin zone integrals.

Quantum and thermal effects are included performing the structure relaxation within the SSCHA framework. Relaxations are done every 2 GPa in the proximity of the phase transition and steps of 10 GPa are chosen far for the critical pressure. The choice of such a fine grid allows us to achieve an accuracy that is better than the typical experimental one and we can assume that the error on the critical pressures extracted from the simulations is at most 2 GPa. The SSCHA minimizations are performed by generating ensembles with 25000 configurations to reduce the statistical noise and the pressure is calculated with Eq. (2.19). The hysteresis cycle is computed by simulating compression and decompression runs. In the former, the starting guess for each structure is the result for the previous pressure, while in the latter, it is the relaxed structure for the immediately higher pressure.

Within the SSCHA simulations, we manage to simulate the classical picture but with included thermal effects. This is done by substituting the physical masses of the atomic species with fictitious infinite ones. In practice, this is done by multiplying the real masses times a 10^6 factor. The number of configurations generated is the

same as for the quantum case.

The infrared spectra are simulated within the TD-SSCHA framework by using the Lanczos algorithm of Sec. 2.7.3. The effective charges have been computed *ab initio* with Quantum ESPRESSO using the PBE approximation of DFT. After checking the convergence, the electronic wave functions were expanded in plane waves up to a cutoff of 80 Ry (640 Ry for the electronic density), with a uniform grid of 4×4×4 for the Brillouin zone integrals. The absorption spectra are calculated by tracing over the dielectric function as in Eq. 3.8 to take into consideration all the possible orientations of the crystal. Most of the spectra are calculated neglecting the position dependence of the effective charges, whose importance is partially explored in Sec. 2.7.2, where the simulations obtained using the average of the effective charges (extracted from the calculations on 500 configurations) and the position-independent ones are compared.

Regarding the Lanczos algorithm. A sub-ensemble with 15000 configurations of the 25000 generated has been used to compute the Infrared response after checking the convergence with the number of configurations. In principle, the Lanczos algorithm gives converged results in the limit of infinite steps and infinitesimal smearing factor. I found that about 350 Lanczos steps are sufficient to achieve convergence with a broadening factor of about $15\,\mathrm{cm}^{-1}$. To have smooth curves, we used a smearing factor of $25\,\mathrm{cm}^{-1}$ for the simulation of classical and quantum Infrared spectra. As a unique exception, we choose a smearing factor of $40\,\mathrm{cm}^{-1}$ to compare the TD-SSCHA simulations with the Molecular dynamics spectra and let the algorithm run for 500 steps.

The IR intensity of each mode described in Fig. 4.20 (**a**) is computed by using the polarization vectors e_μ^a of the Hessian matrix of Eq. 2.59 in the bubble approximation with the relation in Eq. (H.1).

$$I_\mu = 2 \sum_{\alpha=\{x,y,z\}} |\sum_{a=1}^{3N} Z_a^\alpha \frac{e_\mu^a}{\sqrt{M_a}}|^2, \tag{H.1}$$

Then, the area of the dots is the ratio of the IR-intensity of the given mode with respect to the most intense one.

Appendix I

Classical phase VII

The phase transition at a classical level at T=0 K, studied through structure relaxation of the disordered structures at different pressure with LAMMPS [90], shows the same features as for the ordered phase, as for the quantum case of Sec. 4.4.2. Out of the many possible hydrogen arrangements, I study only five of them to check if relevant differences arise in the critical pressure according to the configuration chosen. I report in Fig. I.1 (**a**) the equation of states for the 5 arrangements, denoted as Set 1-5, and the order parameter $\frac{l_{\mathrm{OO}}}{2} - l_{\mathrm{OH}}$ is in Fig. I.1 (**b**). Both the EOS and order parameter of the five structures show some discrepancies below the phase transition at about P=90 GPa, but, as the pressure increases, they all merge and reach zero at the same critical pressure, namely $P_c = 109$ GPa indicated by the red cross in the figure. The critical pressure is 2 GPa greater than for the ordered phase, a difference that is anyway negligible compared to the experimental uncertainties.

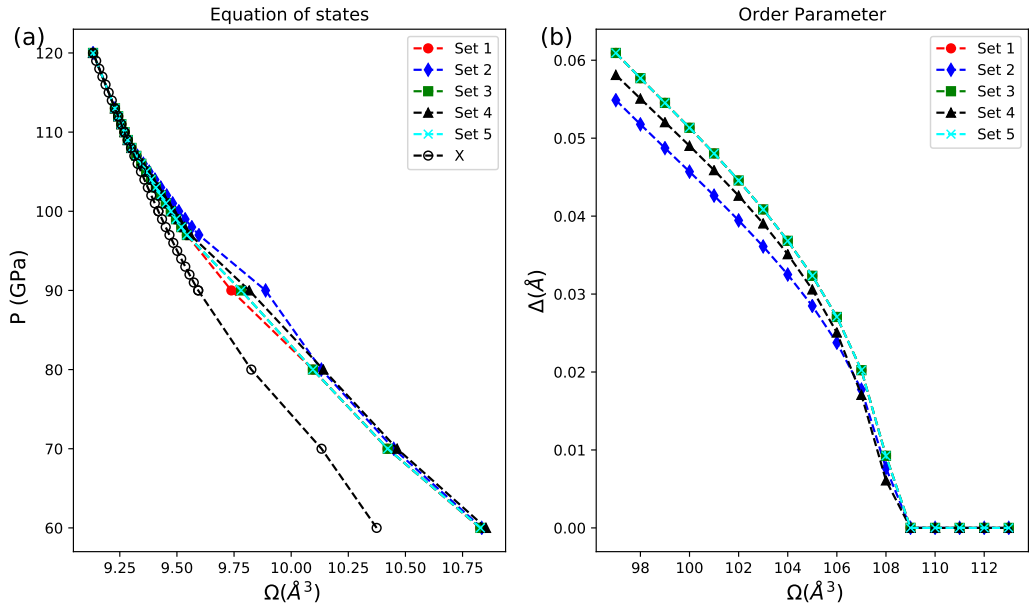

Figure I.1. Structure parameters of the disordered phase in the classical approximation at T=0 K. Five different hydrogen arrangements compatible with the VII phase structure are shown, denoted as Set1 to Set5. Structure relaxation are performed with LAMMPS [90] for different pressures, above and below the phase transition. Panel **a**: Equation of state for the disordered sets compared to that of the symmetric phase X. The red cross indicates the phase transition at P=109 GPa. Panel **b** Order parameter $\frac{l_{\mathrm{OO}}}{2} - l_{\mathrm{OH}}$ as a function of pressure.

Appendix J

Infrared absorption in high-pressure ice

The critical pressure in the quantum regime is almost temperature independent. In Fig. J.1, I report the product of the imaginary part of the dielectric function times the frequency of the ordered phase that according to Eq. (4.5) is related to the absorption coefficient. Two temperatures, T=60 K and T=300 K are compared. The IR spectra are almost unaffected by temperature too. The most relevant differences can be seen in the low-pressure regime, in phase VIII and, mostly, in the region close to the phase transition. The slight difference in the critical pressures, $P_c(60K) = 52$ GPa and $P_c(300K) = 54$ GPa, is reflected in the simulated spectra, as the strong intensity increase of the low energy peak is 2 GPa delayed in the high-temperature spectra.

In some works, [23], it is discussed the importance of including nonlinear effects in the effective charges for studying high-pressure ice. In all the spectra computed above, I computed the effective charges in the equilibrium position and the response function of Eq. (2.114) within the TD-SSCHA framework with the Lanczos algorithm of Sec. 2.7.3. Here, I include the effect of position dependence of the effective charges. I compute at the DFT level, within Quantum ESPRESSO [86] with PBE functional [102] the effective charges for a set of 500 configurations out of the 15000 ones in the SSCHA ensemble. Then, their average is used in Eq. (2.114) to computed the Infrared response at T=300 K.

The comparison of the IR spectra with position-dependent and independent effective charges is in Fig. J.2. I analyzed their effect in different situations, below the phase transition, in phase VIII, at P=33 GPa, close to critical pressure at P=52 GPa and in phase X at P=62 GPa. Fig. J.2 clearly reveals the use of the position-independent effective charges or their average does not affect the IR spectra.

The effect of the nonlinearity in the ionic displacements of the effective charges can be seen from Eq. (2.113). Unfortunately, a code that includes nonlinearity in the perturbation is not yet available, but we can make an idea of their importance by comparing the position-independent effective charges with the second derivative of the dipole moment, opportunely multiplied by the mean square displacement to match the effective charges dimension. By doing this, I can see that, on the diagonal elements, the correction never exceeds 1-2% percent of the leading order.

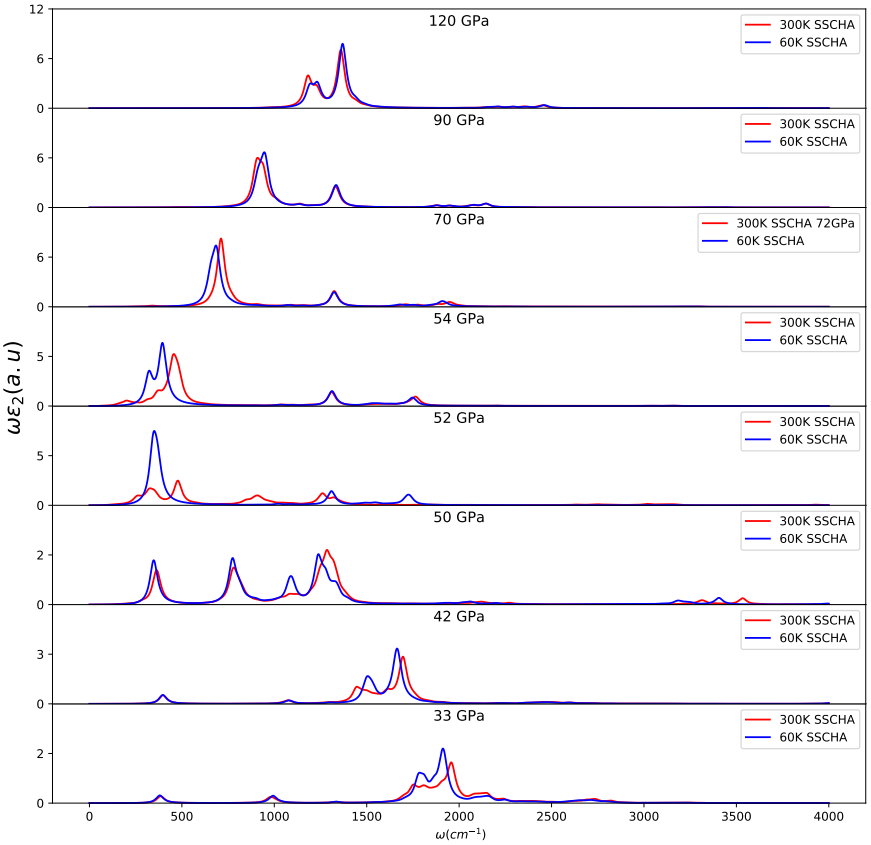

Figure J.1. Comparison of the SSCHA spectra for phase VIII at different temperatures and pressures. The figure reports the product of the imaginary part of the dielectric function and the frequency $\omega\epsilon_2(\omega)$ that is proportional to the absorption coefficient and the real part of the refractive index as $\alpha(\omega)n(\omega)$ as in Eq. (4.5). The spectra are computed within TD-SSCHA with a 16 water molecules simulation cell and a 25 cm^{-1} smearing factor is applied. Red lines are spectra at 300 K, while for the blue ones T=60 K. Between the two temperatures there is a 2 GPa shift in the critical pressure, $P_c(60K) = 52$ GPa and $P_c(300K) = 54$ GPa.

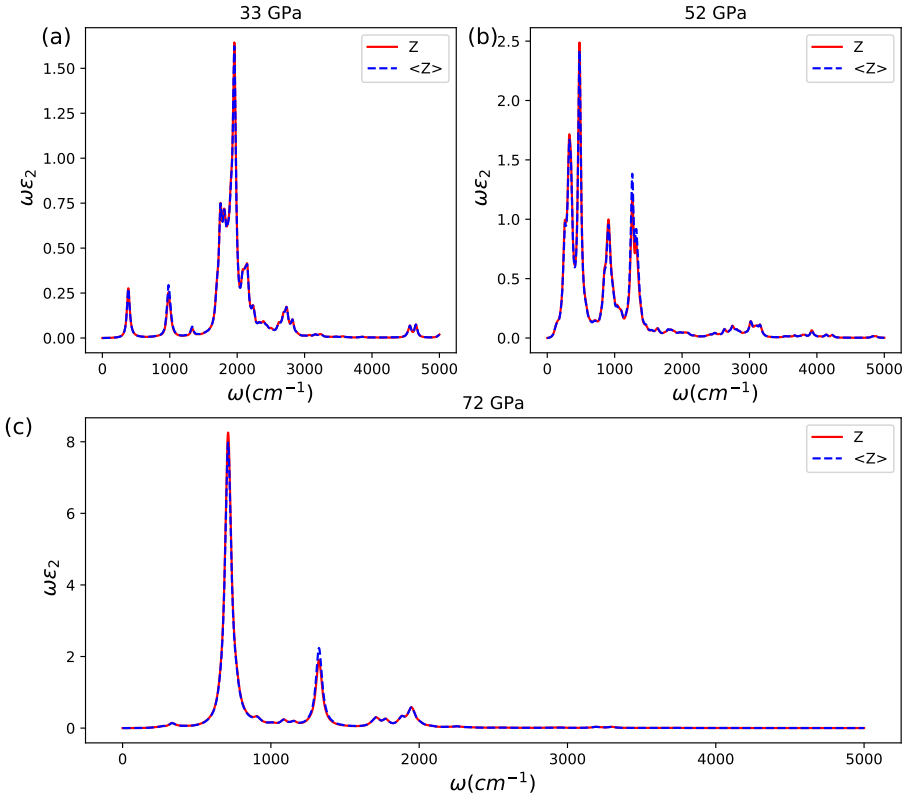

Figure J.2. Effect of the position dependence on the effective charges at T=300 K within TD-SSCHA. Three sample pressure have been chosen, below the phase transition, in phase VIII (33 GPa), close to the phase transition (52 GPa) and above, in phase X (72 GPa). The spectra computed with position independent effective charges computed as in Eq. (2.114) are compared with those where the average of the effective charges is used.

Bibliography

[1] Thorsten Bartels-Rausch, Vance Bergeron, Julyan H. E. Cartwright, Rafael Escribano, John L. Finney, Hinrich Grothe, Pedro J. Gutiérrez, Jari Haapala, Werner F. Kuhs, Jan B. C. Pettersson, Stephen D. Price, C. Ignacio Sainz-Díaz, Debbie J. Stokes, Giovanni Strazzulla, Erik S. Thomson, Hauke Trinks, and Nevin Uras-Aytemiz. Ice structures, patterns, and processes: A view across the icefields. *Reviews of Modern Physics*, 84(2):885–944, May 2012.

[2] John T. Andrews. Glaciers, oceans, atmosphere and climate. In *Glaciers, Oceans, Atmosphere and Climate*, pages 95–113. Blackwell Publishing, 2006.

[3] Christoph G. Salzmann. Advances in the experimental exploration of water's phase diagram. *The Journal of Chemical Physics*, 150(6):060901, February 2019.

[4] A. D. Fortes, I. G. Wood, D. Grigoriev, M. Alfredsson, S. Kipfstuhl, K. S. Knight, and R. I. Smith. No evidence for large-scale proton ordering in antarctic ice from powder neutron diffraction. *The Journal of Chemical Physics*, 120(24):11376–11379, June 2004.

[5] A. J. Leadbetter, R. C. Ward, J. W. Clark, P. A. Tucker, T. Matsuo, and H. Suga. The equilibrium low-temperature structure of ice. *The Journal of Chemical Physics*, 82(1):424–428, January 1985.

[6] C. Lobban, J. L. Finney, and W. F. Kuhs. The structure of a new phase of ice. *Nature*, 391(6664):268–270, January 1998.

[7] B. Kamb. Ice. II. a proton-ordered form of ice. *Acta Crystallographica*, 17(11):1437–1449, November 1964.

[8] B. Kamb, A. Prakash, and C. Knobler. Structure of ice. v. *Acta Crystallographica*, 22(5):706–715, May 1967.

[9] B. Kamb and A. Prakash. Structure of ice III. *Acta Crystallographica Section B Structural Crystallography and Crystal Chemistry*, 24(10):1317–1327, October 1968.

[10] B. Kamb and B. L. Davis. ICE VII, THE DENSEST FORM OF ICE. *Proceedings of the National Academy of Sciences*, 52(6):1433–1439, December 1964.

[11] Hermann Engelhardt and Barclay Kamb. Structure of ice IV, a metastable high-pressure phase. *The Journal of Chemical Physics*, 75(12):5887–5899, December 1981.

[12] L. G. DOWELL and A. P. RINFRET. Low-temperature forms of ice as studied by x-ray diffraction. *Nature*, 188(4757):1144–1148, December 1960.

[13] J. D. Bernal and R. H. Fowler. A theory of water and ionic solution, with particular reference to hydrogen and hydroxyl ions. *The Journal of Chemical Physics*, 1(8):515–548, August 1933.

[14] Sherwin J. Singer, Jer-Lai Kuo, Tomas K. Hirsch, Chris Knight, Lars Ojamäe, and Michael L. Klein. Hydrogen-bond topology and the Ice-VII/VIIIand IceIh/XIProton-ordering phase transitions. *Physical Review Letters*, 94(13):135701, April 2005.

[15] J. F. Nagle. Lattice statistics of hydrogen bonded crystals. i. the residual entropy of ice. *Journal of Mathematical Physics*, 7(8):1484–1491, August 1966.

[16] Ph. Pruzan, J. C. Chervin, E. Wolanin, B. Canny, M. Gauthier, and M. Hanfland. Phase diagram of ice in the VII-VIII-x domain. vibrational and structural data for strongly compressed ice VIII. *Journal of Raman Spectroscopy*, 34(7-8):591–610, 2003.

[17] Y. Tajima, T. Matsuo, and H. Suga. Phase transition in KOH-doped hexagonal ice. *Nature*, 299(5886):810–812, October 1982.

[18] Yoshimitsu Tajima, Takasuke Matsuo, and Hiroshi Suga. Calorimetric study of phase transition in hexagonal ice doped with alkali hydroxides. *Journal of Physics and Chemistry of Solids*, 45(11-12):1135–1144, January 1984.

[19] Syuji Kwada. Acceleration of dielectric relaxation by KOH-doping and phase transition in ice ih. *Journal of Physics and Chemistry of Solids*, 50(11):1177–1184, January 1989.

[20] K. Röttger, A. Endriss, J. Ihringer, S. Doyle, and W. F. Kuhs. Lattice constants and thermal expansion of H_2O and D_2O ice ih between 10 and 265 k. *Acta Crystallographica Section B Structural Science*, 50(6):644–648, December 1994.

[21] A. Dominic Fortes. Accurate and precise lattice parameters of H_2O and D_2O ice ihbetween 1.6 and 270 k from high-resolution time-of-flight neutron powder diffraction data. *Acta Crystallographica Section B Structural Science, Crystal Engineering and Materials*, 74(2):196–216, March 2018.

[22] Magali Benoit, Dominik Marx, and Michele Parrinello. Tunnelling and zero-point motion in high-pressure ice. *Nature*, 392(6673):258–261, Mar 1998.

[23] M. Bernasconi, P. L. Silvestrelli, and M. Parrinello. Ab InitioInfrared absorption study of the hydrogen-bond symmetrization in ice. *Physical Review Letters*, 81(6):1235–1238, August 1998.

[24] Linfeng Zhang, Mohan Chen, Xifan Wu, Han Wang, Weinan E, and Roberto Car. Deep neural network for the dielectric response of insulators. *Physical Review B*, 102(4):041121, July 2020.

[25] Ph. Pruzan, J. C. Chervin, and B. Canny. Stability domain of the ice VIII proton-ordered phase at very high pressure and low temperature. *The Journal of Chemical Physics*, 99(12):9842–9846, December 1993.

[26] A. F. Goncharov, V. V. Struzhkin, M. S. Somayazulu, R. J. Hemley, and H. K. Mao. Compression of ice to 210 gigapascals: Infrared evidence for a symmetric hydrogen-bonded phase. *Science*, 273(5272):218–220, July 1996.

[27] M. Song, H. Yamawaki, H. Fujihisa, M. Sakashita, and K. Aoki. Infrared investigation on ice VIII and the phase diagram of dense ices. *Physical Review B*, 68(1):014106, jul 2003.

[28] R. Ramírez, N. Neuerburg, M.-V. Fernández-Serra, and C. P. Herrero. Quasi-harmonic approximation of thermodynamic properties of ice ih, II, and III. *The Journal of Chemical Physics*, 137(4):044502, July 2012.

[29] B. Pamuk, J. M. Soler, R. Ramírez, C. P. Herrero, P. W. Stephens, P. B. Allen, and M.-V. Fernández-Serra. Anomalous nuclear quantum effects in ice. *Physical Review Letters*, 108(19):193003, May 2012.

[30] Carlos P. Herrero and Rafael Ramírez. Isotope effects in ice ih: A path-integral simulation. *The Journal of Chemical Physics*, 134(9):094510, March 2011.

[31] Betül Pamuk, Philip B. Allen, and M.-V. Fernández-Serra. Electronic and nuclear quantum effects on the ice XI/ice ih phase transition. *Physical Review B*, 92(13):134105, October 2015.

[32] Lorenzo Monacelli, Raffaello Bianco, Marco Cherubini, Matteo Calandra, Ion Errea, and Francesco Mauri. The stochastic self-consistent harmonic approximation: calculating vibrational properties of materials with full quantum and anharmonic effects. *Journal of Physics: Condensed Matter*, 33(36):363001, July 2021.

[33] Lorenzo Monacelli and Francesco Mauri. Time-dependent self-consistent harmonic approximation: Anharmonic nuclear quantum dynamics and time correlation functions. *Physical Review B*, 103(10):104305, March 2021.

[34] A. A. Maradudin and A. E. Fein. Scattering of neutrons by an anharmonic crystal. *Phys. Rev.*, 128(6):2589–2608, 1962.

[35] R A Cowley. Anharmonic crystals. *Reports on Progress in Physics*, 31(1):123–166, 1968.

[36] Matteo Calandra, Michele Lazzeri, and Francesco Mauri. Anharmonic and non-adiabatic effects in mgb2: Implications for the isotope effect and interpretation of raman spectra. *Physica C: Superconductivity*, 456(1-2):38 – 44, 2007.

[37] Ion Errea. Approaching the strongly anharmonic limit with ab initio calculations of materials' vibrational properties – a colloquium. *The European Physical Journal B*, 89(11):237, nov 2016.

[38] R. Car and M. Parrinello. Unified approach for molecular dynamics and density-functional theory. *Phys. Rev. Lett.*, 55:2471–2474, Nov 1985.

[39] C. Z. Wang, C. T. Chan, and K. M. Ho. Tight-binding molecular-dynamics study of phonon anharmonic effects in silicon and diamond. *Phys. Rev. B*, 42:11276–11283, Dec 1990.

[40] Mathias P. Ljungberg and Jorge Íñiguez. Temperature-dependent classical phonons from efficient nondynamical simulations. *Phys. Rev. Lett.*, 110:105503, Mar 2013.

[41] Ioan B. Magdău and Graeme J. Ackland. Identification of high-pressure phases iii and iv in hydrogen: Simulating raman spectra using molecular dynamics. *Phys. Rev. B*, 87:174110, May 2013.

[42] Dong-Bo Zhang, Tao Sun, and Renata M. Wentzcovitch. Phonon quasiparticles and anharmonic free energy in complex systems. *Phys. Rev. Lett.*, 112:058501, Feb 2014.

[43] O. Hellman, I. A. Abrikosov, and S. I. Simak. Lattice dynamics of anharmonic solids from first principles. *Phys. Rev. B*, 84:180301, Nov 2011.

[44] Olle Hellman, Peter Steneteg, I. A. Abrikosov, and S. I. Simak. Temperature dependent effective potential method for accurate free energy calculations of solids. *Phys. Rev. B*, 87:104111, Mar 2013.

[45] Olle Hellman and I. A. Abrikosov. Temperature-dependent effective third-order interatomic force constants from first principles. *Phys. Rev. B*, 88:144301, Oct 2013.

[46] D. M. Ceperley. Path integrals in the theory of condensed helium. *Rev. Mod. Phys.*, 67:279–355, Apr 1995.

[47] W. Zhong, David Vanderbilt, and K. M. Rabe. First-principles theory of ferroelectric phase transitions for perovskites: The case of batio$_3$. *Phys. Rev. B*, 52:6301–6312, Sep 1995.

[48] P Giannozzi, O Andreussi, T Brumme, O Bunau, M Buongiorno Nardelli, M Calandra, R Car, C Cavazzoni, D Ceresoli, M Cococcioni, N Colonna, I Carnimeo, A Dal 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 Otero 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. Advanced capabilities for materials modelling with quantum ESPRESSO. *Journal of Physics: Condensed Matter*, 29(46):465901, oct 2017.

[49] Ion Errea, Bruno Rousseau, and Aitor Bergara. Anharmonic stabilization of the high-pressure simple cubic phase of calcium. *Phys. Rev. Lett.*, 106(16):165501, 2011.

[50] Fei Zhou, Weston Nielson, Yi Xia, and Vidvuds Ozoliņš. Lattice anharmonicity and thermal conductivity from compressive sensing of first-principles calculations. *Phys. Rev. Lett.*, 113:185501, Oct 2014.

[51] Terumasa Tadano and Shinji Tsuneyuki. Self-consistent phonon calculations of lattice dynamical properties in cubic srtio$_3$ with first-principles anharmonic force constants. *Phys. Rev. B*, 92:054301, Aug 2015.

[52] T Tadano, Y Gohda, and S Tsuneyuki. Anharmonic force constants extracted from first-principles molecular dynamics: applications to heat transfer simulations. *Journal of Physics: Condensed Matter*, 26(22):225402, may 2014.

[53] Sam Azadi, Bartomeu Monserrat, W. M. C. Foulkes, and R. J. Needs. Dissociation of high-pressure solid molecular hydrogen: A quantum monte carlo and anharmonic vibrational study. *Phys. Rev. Lett.*, 112:165501, Apr 2014.

[54] Ion Errea, Matteo Calandra, and Francesco Mauri. First-principles theory of anharmonicity and the inverse isotope effect in superconducting palladium-hydride compounds. *Phys. Rev. Lett.*, 111:177002, Oct 2013.

[55] Ion Errea, Matteo Calandra, and Francesco Mauri. Anharmonic free energies and phonon dispersions from the stochastic self-consistent harmonic approximation: Application to platinum and palladium hydrides. *Phys. Rev. B*, 89:064302, Feb 2014.

[56] Raffaello Bianco, Ion Errea, Lorenzo Paulatto, Matteo Calandra, and Francesco Mauri. Second-order structural phase transitions, free energy curvature, and temperature-dependent anharmonic phonons in the self-consistent harmonic approximation: Theory and stochastic implementation. *Phys. Rev. B*, 96:014111, Jul 2017.

[57] Lorenzo Monacelli, Ion Errea, Matteo Calandra, and Francesco Mauri. Pressure and stress tensor of complex anharmonic crystals within the stochastic self-consistent harmonic approximation. *Phys. Rev. B*, 98:024106, Jul 2018.

[58] P. Souvatzis, O. Eriksson, M. I. Katsnelson, and S. P. Rudin. Entropy driven stabilization of energetically unstable crystal structures explained from first principles theory. *Phys. Rev. Lett.*, 100(9):095901, 2008.

[59] Xiaoli Tang, Chen W. Li, and B. Fultz. Anharmonicity-induced phonon broadening in aluminum at high temperatures. *Phys. Rev. B*, 82:184301, Nov 2010.

[60] Nikolas Antolin, Oscar D. Restrepo, and Wolfgang Windl. Fast free-energy calculations for unstable high-temperature phases. *Phys. Rev. B*, 86:054119, Aug 2012.

[61] Krzysztof Parlinski. Ab initio determination of anharmonic phonon peaks. *Phys. Rev. B*, 98:054305, Aug 2018.

[62] Fredrik Eriksson, Erik Fransson, and Paul Erhart. The hiphive package for the extraction of high-order force constants by machine learning. *Advanced Theory and Simulations*, 2(5):1800184, 2019.

[63] Ambroise van Roekeghem, Jesús Carrete, and Natalio Mingo. Quantum self-consistent ab-initio lattice dynamics. *Computer Physics Communications*, 263:107945, June 2021.

[64] Bartomeu Monserrat, N. D. Drummond, and R. J. Needs. Anharmonic vibrational properties in periodic systems: energy, electron-phonon coupling, and stress. *Phys. Rev. B*, 87:144302, Apr 2013.

[65] D. J. Hooton. Li. a new treatment of anharmonicity in lattice thermodynamics: I. *Philosophical Magazine Series 7*, 46:422–432, 1955.

[66] Joel M. Bowman. Self-consistent field energies and wavefunctions for coupled oscillators. *The Journal of Chemical Physics*, 68(2):608–610, January 1978.

[67] Ion Errea, Matteo Calandra, Chris J. Pickard, Joseph R. Nelson, Richard J. Needs, Yinwei Li, Hanyu Liu, Yunwei Zhang, Yanming Ma, and Francesco Mauri. Quantum hydrogen-bond symmetrization in the superconducting hydrogen sulfide system. *Nature*, 532(7597):81–84, Apr 2016. Letter.

[68] Ion Errea, Francesco Belli, Lorenzo Monacelli, Antonio Sanna, Takashi Koretsune, Terumasa Tadano, Raffaello Bianco, Matteo Calandra, Ryotaro Arita, Francesco Mauri, and José A. Flores-Livas. Quantum crystal structure in the 250-kelvin superconducting lanthanum hydride. *Nature*, 578(7793):66–69, Feb 2020.

[69] Lorenzo Monacelli, Ion Errea, Matteo Calandra, and Francesco Mauri. Black metal hydrogen above 360 gpa driven by proton quantum fluctuations. *Nature Physics*, 17(1):63–67, Sep 2020.

[70] Jae-Mo Lihm and Cheol-Hwan Park. Gaussian time-dependent variational principle for the finite-temperature anharmonic lattice dynamics. *Phys. Rev. Research*, 3:L032017, Jul 2021.

[71] G. Q. Huang, L. F. Chen, M. Liu, and D. Y. Xing. Electronic structure and electron-phonon interaction in the ternary silicides *malsi* ($m = ca$, sr, and ba). *Phys. Rev. B*, 69(6):064509, Feb 2004.

[72] Raffaello Bianco, Lorenzo Monacelli, Matteo Calandra, Francesco Mauri, and Ion Errea. Weak dimensionality dependence and dominant role of ionic fluctuations in the charge-density-wave transition of $nbse_2$. *Phys. Rev. Lett.*, 125:106101, Sep 2020.

[73] Radford M. Neal. Annealed importance sampling. *Statistics and Computing*, 11(2):125–139, 2001.

[74] Mattia Miotto and Lorenzo Monacelli. Entropy evaluation sheds light on ecosystem complexity. *Physical Review E*, 98(4):042402, October 2018.

[75] Augustine Kong, Jun S. Liu, and Wing Hung Wong. Sequential imputations and bayesian missing data problems. *Journal of the American Statistical Association*, 89(425):278–288, mar 1994.

[76] R. H. Liu, T. Wu, G. Wu, H. Chen, X. F. Wang, Y. L. Xie, J. J. Ying, Y. J. Yan, Q. J. Li, B. C. Shi, W. S. Chu, Z. Y. Wu, and X. H. Chen. A large iron isotope effect in smfeaso$_{1-x}$f$_x$ and ba$_{1-x}$k$_x$fe$_2$as$_2$. *Nature*, 459:64–67, 2009.

[77] Nina Shulumba, Olle Hellman, and Austin J. Minnich. Intrinsic localized mode and low thermal conductivity of pbse. *Phys. Rev. B*, 95:014302, Jan 2017.

[78] Nina Shulumba, Olle Hellman, and Austin J. Minnich. Lattice thermal conductivity of polyethylene molecular crystals from first-principles including nuclear quantum effects. *Phys. Rev. Lett.*, 119:185901, Oct 2017.

[79] Raffaello Bianco, Ion Errea, Matteo Calandra, and Francesco Mauri. High-pressure phase diagram of hydrogen and deuterium sulfides from first principles: Structural and vibrational properties including quantum and anharmonic effects. *Physical Review B*, 97(21):214101, June 2018.

[80] Guilherme A. S. Ribeiro, Lorenzo Paulatto, Raffaello Bianco, Ion Errea, Francesco Mauri, and Matteo Calandra. Strong anharmonicity in the phonon spectra of pbte and snte from first principles. *Phys. Rev. B*, 97:014306, Jan 2018.

[81] Unai Aseginolaza, Raffaello Bianco, Lorenzo Monacelli, Lorenzo Paulatto, Matteo Calandra, Francesco Mauri, Aitor Bergara, and Ion Errea. Phonon collapse and second-order phase transition in thermoelectric snse. *Phys. Rev. Lett.*, 122:075901, Feb 2019.

[82] Raffaello Bianco, Ion Errea, Lorenzo Monacelli, Matteo Calandra, and Francesco Mauri. Quantum enhancement of charge density wave in nbs2 in the two-dimensional limit. *Nano Letters*, 19(5):3098–3103, 2019. PMID: 30932501.

[83] Unai Aseginolaza, Raffaello Bianco, Lorenzo Monacelli, Lorenzo Paulatto, Matteo Calandra, Francesco Mauri, Aitor Bergara, and Ion Errea. Strong anharmonicity and high thermoelectric efficiency in high-temperature sns from first principles. *Phys. Rev. B*, 100:214307, Dec 2019.

[84] Ion Errea, Matteo Calandra, Chris J. Pickard, Joseph Nelson, Richard J. Needs, Yinwei Li, Hanyu Liu, Yunwei Zhang, Yanming Ma, and Francesco Mauri. High-pressure hydrogen sulfide from first principles: A strongly anharmonic phonon-mediated superconductor. *Phys. Rev. Lett.*, 114:157004, Apr 2015.

[85] Jianqiang Sky Zhou, Lorenzo Monacelli, Raffaello Bianco, Ion Errea, Francesco Mauri, and Matteo Calandra. Anharmonicity and doping melt the charge density wave in single-layer tise2. *Nano Letters*, 20(7):4809–4815, 2020. PMID: 32496779.

[86] Paolo Giannozzi, Stefano Baroni, Nicola Bonini, Matteo Calandra, Roberto Car, Carlo Cavazzoni, Davide Ceresoli, Guido L Chiarotti, Matteo Cococcioni, Ismaila Dabo, Andrea Dal Corso, Stefano de Gironcoli, Stefano Fabris, Guido Fratesi, Ralph Gebauer, Uwe Gerstmann, Christos Gougoussis, Anton Kokalj, Michele Lazzeri, Layla Martin-Samos, Nicola Marzari, Francesco Mauri, Riccardo Mazzarello, Stefano Paolini, Alfredo Pasquarello, Lorenzo Paulatto, Carlo Sbraccia, Sandro Scandolo, Gabriele Sclauzero, Ari P Seitsonen, Alexander Smogunov, Paolo Umari, and Renata M Wentzcovitch. QUANTUM ESPRESSO: a modular and open-source software project for quantum simulations of materials. *Journal of Physics: Condensed Matter*, 21(39):395502, sep 2009.

[87] G. Kresse and J. Furthmüller. Efficient iterative schemes forab initiototal-energy calculations using a plane-wave basis set. *Physical Review B*, 54(16):11169–11186, October 1996.

[88] José M Soler, Emilio Artacho, Julian D Gale, Alberto García, Javier Junquera, Pablo Ordejón, and Daniel Sánchez-Portal. The SIESTA method forab initioorder-nmaterials simulation. *Journal of Physics: Condensed Matter*, 14(11):2745–2779, March 2002.

[89] Thomas D. Kühne, Marcella Iannuzzi, Mauro Del Ben, Vladimir V. Rybkin, Patrick Seewald, Frederick Stein, Teodoro Laino, Rustam Z. Khaliullin, Ole Schütt, Florian Schiffmann, Dorothea Golze, Jan Wilhelm, Sergey Chulkov, Mohammad Hossein Bani-Hashemian, Valéry Weber, Urban Borštnik, Mathieu Taillefumier, Alice Shoshana Jakobovits, Alfio Lazzaro, Hans Pabst, Tiziano Müller, Robert Schade, Manuel Guidon, Samuel Andermatt, Nico Holmberg, Gregory K. Schenter, Anna Hehn, Augustin Bussy, Fabian Belleflamme, Gloria Tabacchi, Andreas Glöss, Michael Lass, Iain Bethune, Christopher J. Mundy, Christian Plessl, Matt Watkins, Joost VandeVondele, Matthias Krack, and Jürg Hutter. CP2k: An electronic structure and molecular dynamics software package - quickstep: Efficient and accurate electronic structure calculations. *The Journal of Chemical Physics*, 152(19):194103, May 2020.

[90] Steve Plimpton. Fast parallel algorithms for short-range molecular dynamics. *Journal of Computational Physics*, 117(1):1–19, March 1995.

[91] M. Born and R. Oppenheimer. Zur quantentheorie der molekeln. *Annalen der Physik*, 389(20):457–484, 1927.

[92] Venkat Kapil, Edgar Engel, Mariana Rossi, and Michele Ceriotti. Assessment of approximate methods for anharmonic free energies. *Journal of Chemical Theory and Computation*, 15(11):5845–5857, September 2019.

[93] Yingkai Zhang and Weitao Yang. Comment on "generalized gradient approximation made simple". *Physical Review Letters*, 80(4):890–890, January 1998.

[94] Carlo Adamo and Vincenzo Barone. Toward reliable density functional methods without adjustable parameters: The PBE0 model. *The Journal of Chemical Physics*, 110(13):6158–6170, April 1999.

[95] Stefan Grimme, Jens Antony, Stephan Ehrlich, and Helge Krieg. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-d) for the 94 elements h-pu. *The Journal of Chemical Physics*, 132(15):154104, April 2010.

[96] Ondrej Marsalek and Thomas E. Markland. Quantum dynamics and spectroscopy of ab initio liquid water: The interplay of nuclear and electronic quantum effects. *The Journal of Physical Chemistry Letters*, 8(7):1545–1551, mar 2017.

[97] Bingqing Cheng, Edgar A. Engel, Jörg Behler, Christoph Dellago, and Michele Ceriotti. Ab initio thermodynamics of liquid and solid water. *Proceedings of the National Academy of Sciences*, 116(4):1110–1115, January 2019.

[98] Jörg Behler and Michele Parrinello. Generalized neural-network representation of high-dimensional potential-energy surfaces. *Physical Review Letters*, 98(14):146401, April 2007.

[99] J Behler. Runner-a neural network code for high-dimensional potential-energy surfaces, universität göttingen, 2018.

[100] Yuzhi Zhang, Haidi Wang, Weijie Chen, Jinzhe Zeng, Linfeng Zhang, Han Wang, and Weinan E. DP-GEN: A concurrent learning platform for the generation of reliable deep learning based potential energy models. *Computer Physics Communications*, 253:107206, August 2020.

[101] Linfeng Zhang, Jiequn Han, Han Wang, Roberto Car, and Weinan E. Deep potential molecular dynamics: A scalable model with the accuracy of quantum mechanics. *Physical Review Letters*, 120(14):143001, April 2018.

[102] John P. Perdew, Kieron Burke, and Matthias Ernzerhof. Generalized gradient approximation made simple. *Physical Review Letters*, 77(18):3865–3868, October 1996.

[103] Soohaeng Yoo, Xiao Cheng Zeng, and Sotiris S. Xantheas. On the phase diagram of water with density functional theory potentials: The melting temperature of ice ih with the perdew–burke–ernzerhof and becke–lee–yang–parr functionals. *The Journal of Chemical Physics*, 130(22):221102, June 2009.

[104] Andrew J. Rusnak, Eric R. Pinnick, Camilo E. Calderon, and Feng Wang. Static dielectric constants and molecular dipole distributions of liquid water and ice-ih investigated by the PAW-PBE exchange-correlation functional. *The Journal of Chemical Physics*, 137(3):034510, July 2012.

[105] Han Wang, Linfeng Zhang, Jiequn Han, and Weinan E. DeePMD-kit: A deep learning package for many-body potential energy representation and molecular dynamics. *Computer Physics Communications*, 228:178–184, July 2018.

[106] Andrea Grisafi, David M. Wilkins, Gábor Csányi, and Michele Ceriotti. Symmetry-adapted machine learning for tensorial properties of atomistic systems. *Physical Review Letters*, 120(3):036002, January 2018.

[107] Marco Cherubini, Lorenzo Monacelli, and Francesco Mauri. The microscopic origin of the anomalous isotopic properties of ice relies on the strong quantum anharmonic regime of atomic vibration. *The Journal of Chemical Physics*, 155(18):184502, November 2021.

[108] Shuji Kawada. Dielectric dispersion and phase transition of KOH doped ice. *Journal of the Physical Society of Japan*, 32(5):1442–1442, May 1972.

[109] S. M. Jackson, V. M. Nield, R. W. Whitworth, M. Oguro, and C. C. Wilson. Single-crystal neutron diffraction studies of the structure of ice XI. *The Journal of Physical Chemistry B*, 101(32):6142–6145, August 1997.

[110] Steve Vance, Thomas Loerting, Josef Stern, Matt Kropf, Baptiste Journaux, Corey Jamieson, Morgan L. Cable, and Olivier Bollengier. Chapter 3 solids and fluids at low temperatures. In *Low Temperature Materials and Mechanisms*, pages 27–54. CRC Press, July 2016.

[111] Massimo Marchi, John S. Tse, and Michael L. Klein. Lattice vibrations and infrared absorption of ice ih. *The Journal of Chemical Physics*, 85(5):2414–2418, September 1986.

[112] John S. Tse, Michael L. Klein, and Ian R. McDonald. Lattice vibrations of ices ih, VIII, and IX. *The Journal of Chemical Physics*, 81(12):6124–6129, December 1984.

[113] Paola Bosi, Riccardo Tubino, and Giuseppe Zerbi. On the problem of the vibrational spectrum and structure of ice ih: Lattice dynamical calculations. *The Journal of Chemical Physics*, 59(9):4578–4586, November 1973.

[114] Hideki Tanaka. Thermodynamic stability and negative thermal expansion of hexagonal and cubic ices. *The Journal of Chemical Physics*, 108(12):4887–4893, March 1998.

[115] Koichiro Umemoto, Emiko Sugimura, Stefano de Gironcoli, Yoichi Nakajima, Kei Hirose, Yasuo Ohishi, and Renata M. Wentzcovitch. Nature of the volume isotope effect in ice. *Physical Review Letters*, 115(17):173005, October 2015.

[116] Michael A. Salim, Soohaeng Yoo Willow, and So Hirata. Ice ih anomalies: Thermal contraction, anomalous volume isotope effect, and pressure-induced amorphization. *The Journal of Chemical Physics*, 144(20):204503, May 2016.

[117] M. V. Fernández-Serra, G. Ferlat, and E. Artacho. Two exchange-correlation functionals compared for first-principles liquid water. *Molecular Simulation*, 31(5):361–366, April 2005.

[118] Jue Wang, G. Román-Pérez, Jose M. Soler, Emilio Artacho, and M.-V. Fernández-Serra. Density, structure, and dynamics of water: The effect

of van der waals interactions. *The Journal of Chemical Physics*, 134(2):024516, January 2011.

[119] F. Li and J. L. Skinner. Infrared and raman line shapes for ice ih. i. dilute HOD in H_2O and D_2O. *The Journal of Chemical Physics*, 132(20):204505, May 2010.

[120] F. Li and J. L. Skinner. Infrared and raman line shapes for ice ih. II. H_2O and D_2O. *The Journal of Chemical Physics*, 133(24):244504, December 2010.

[121] Marek J. Wójcik, Krzysztof Szczeponek, and Susumu Ikeda. Theoretical study of the OH/OD stretching regions of the vibrational spectra of ice ih. *The Journal of Chemical Physics*, 117(21):9850–9857, December 2002.

[122] Michael S. Bergren and Stuart A. Rice. An improved analysis of the OH stretching region of the vibrational spectrum of ice ih. *The Journal of Chemical Physics*, 77(2):583–602, July 1982.

[123] Stuart A. Rice, Michael S. Bergren, Alan C. Belch, and Glenn Nielsen. A theoretical analysis of the hydroxyl stretching spectra of ice ih, liquid water, and amorphous solid water. *The Journal of Physical Chemistry*, 87(21):4295–4308, October 1983.

[124] J. E. Bertie and E. Whalley. Infrared spectra of ices ih and ic in the range 4000 to 350 cm—1. *The Journal of Chemical Physics*, 40(6):1637–1645, March 1964.

[125] James R. Scherer and Robert G. Snyder. Raman intensities of single crystal ice ih. *The Journal of Chemical Physics*, 67(11):4794–4811, December 1977.

[126] M. L. Clapp, D. R. Worsnop, and R. E. Miller. Frequency-dependent optical constants of water ice obtained directly from aerosol extinction spectra. *The Journal of Physical Chemistry*, 99(17):6317–6326, April 1995.

[127] A. Y. Zasetsky, A. F. Khalizov, M. E. Earle, and J. J. Sloan. Frequency dependent complex refractive indices of supercooled liquid water and ice determined from aerosol extinction spectra. *The Journal of Physical Chemistry A*, 109(12):2760–2764, March 2005.

[128] Kohji Abe, Kazumasa Ishii, Masaki Nakajima, Hideyuki Fukuda, and Takeshi Shigenari. Raman scattering at the proton ordering phase transition in ice crystal. *Ferroelectrics*, 239(1):1–8, February 2000.

[129] Kohji Abe and Takeshi Shigenari. Raman spectra of proton ordered phase XI of ICE i. translational vibrations below 350 cm-1. *The Journal of Chemical Physics*, 134(10):104506, March 2011.

[130] Takeshi Shigenari and Kohji Abe. Vibrational modes of hydrogens in the proton ordered phase XI of ice: Raman spectra above 400 cm-1. *The Journal of Chemical Physics*, 136(17):174504, May 2012.

[131] J.-C. Li, V.M. Nield, and S.M. Jackson. Spectroscopic measurements of ice XI. *Chemical Physics Letters*, 241(4):290–294, July 1995.

[132] M. Arakawa, H. Kagi, and H. Fukazawa. LABORATORY MEASUREMENTS OF INFRARED ABSORPTION SPECTRA OF HYDROGEN-ORDERED ICE: A STEP TO THE EXPLORATION OF ICE XI IN SPACE. *The Astrophysical Journal Supplement Series*, 184(2):361–365, September 2009.

[133] Hiroshi Fukazawa, Susumu Ikeda, and Shinji Mae. Incoherent inelastic neutron scattering measurements on ice XI; the proton-ordered phase of ice ih doped with KOH. *Chemical Physics Letters*, 282(2):215–218, January 1998.

[134] Hidenosuke Itoh, Katsuyuki Kawamura, Takeo Hondoh, and Sinji Mae. Molecular dynamics studies of proton ordering effects on lattice vibrations in ice ih. *Physica B: Condensed Matter*, 219-220:469–472, April 1996.

[135] Hidenosuke Itoh, Katsuyuki Kawamura, Takeo Hondoh, and Shinji Mae. Polarized librational spectra of proton-ordered ice XI by molecular dynamics simulations. *The Journal of Chemical Physics*, 109(12):4894–4899, September 1998.

[136] Alessandro Erba, Silvia Casassa, Roberto Dovesi, Lorenzo Maschio, and Cesare Pisani. Periodic density functional theory and local-MP2 study of the librational modes of ice XI. *The Journal of Chemical Physics*, 130(7):074505, February 2009.

[137] Maciej Gług, Marek Boczar, Łukasz Boda, and Marek J. Wójcik. Analysis of librational modes of ice XI studied by car–parrinello molecular dynamics. *Chemical Physics*, 459:102–111, September 2015.

[138] Radu Iftimie and Mark E. Tuckerman. Decomposing total IR spectra of aqueous systems into solute and solvent contributions: A computational approach using maximally localized wannier orbitals. *The Journal of Chemical Physics*, 122(21):214508, June 2005.

[139] Wei Chen, Manu Sharma, Raffaele Resta, Giulia Galli, and Roberto Car. Role of dipolar correlations in the infrared spectra of water and ice. *Physical Review B*, 77(24):245114, June 2008.

[140] Stefan Klotz, Th Strässle, A M Saitta, G Rousse, G Hamel, R J Nelmes, J S Loveday, and M Guthrie. In situ neutron diffraction studies of high density amorphous ice under pressure. *Journal of Physics: Condensed Matter*, 17(11):S967–S974, mar 2005.

[141] S. Klotz, G. Hamel, J. S. Loveday, R. J. Nelmes, M. Guthrie, and A. K. Soper. Structure of high-density amorphous ice under pressure. *Physical Review Letters*, 89(28):285502, December 2002.

[142] S. Klotz, Th. Strässle, R. J. Nelmes, J. S. Loveday, G. Hamel, G. Rousse, B. Canny, J. C. Chervin, and A. M. Saitta. Nature of the polyamorphic transition in ice under pressure. *Physical Review Letters*, 94(2):025506, January 2005.

[143] Th. Strässle, A. M. Saitta, S. Klotz, and M. Braden. Phonon dispersion of ice under pressure. *Physical Review Letters*, 93(22):225901, November 2004.

[144] Thierry Strässle, Stefan Klotz, Gérard Hamel, Michael M. Koza, and Helmut Schober. Experimental evidence for a crossover between two distinct mechanisms of amorphization in ice ih under pressure. *Physical Review Letters*, 99(17):175501, October 2007.

[145] J. M. Besson, S. Klotz, G. Hamel, W. G. Marshall, R. J. Nelmes, and J. S. Loveday. Structural instability in ice VIII under pressure. *Physical Review Letters*, 78(16):3141–3144, April 1997.

[146] Richard J. Nelmes, John S. Loveday, Thierry Strässle, Craig L. Bull, Malcolm Guthrie, Gérard Hamel, and Stefan Klotz. Annealed high-density amorphous ice under pressure. *Nature Physics*, 2(6):414–418, May 2006.

[147] Lars Goerigk and Stefan Grimme. A thorough benchmark of density functional methods for general main group thermochemistry, kinetics, and noncovalent interactions. *Physical Chemistry Chemical Physics*, 13(14):6670, 2011.

[148] David T. W. Buckingham, J. J. Neumeier, Sueli H. Masunaga, and Yi-Kuo Yu. Thermal expansion of single-crystal h_2O and d_2O ice ih. *Phys. Rev. Lett.*, 121:185505, Nov 2018.

[149] J. Matas, J. Bass, Y. Ricard, E. Mattern, and M. S. T. Bukowinski. On the bulk composition of the lower mantle: predictions and limitations from generalized inversion of radial seismic profiles. *Geophysical Journal International*, 170(2):764–780, August 2007.

[150] Jessica C. E. Irving, Sanne Cottaar, and Vedran Lekić. Seismically determined elastic parameters for earth's outer core. *Science Advances*, 4(6):eaar2538, June 2018.

[151] Edgar A. Engel, Bartomeu Monserrat, and Richard J. Needs. Anharmonic nuclear motion and the relative stability of hexagonal and cubic ice. *Physical Review X*, 5(2):021033, June 2015.

[152] Ion Errea, Matteo Calandra, Chris J. Pickard, Joseph R. Nelson, Richard J. Needs, Yinwei Li, Hanyu Liu, Yunwei Zhang, Yanming Ma, and Francesco Mauri. Quantum hydrogen-bond symmetrization in the superconducting hydrogen sulfide system. *Nature*, 532(7597):81–84, March 2016.

[153] Lorenzo Paulatto, Ion Errea, Matteo Calandra, and Francesco Mauri. 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, Feb 2015.

[154] A. Novak. Hydrogen bonding in solids correlation of spectroscopic and crystallographic data. In *Large Molecules*, pages 177–216. Springer Berlin Heidelberg, 1976.

[155] M. Ichikawa. The o–h vs o.o distance correlation, the geometric isotope effect in OHO bonds, and its application to symmetric bonds. *Acta Crystallographica Section B Structural Crystallography and Crystal Chemistry*, 34(7):2074–2080, July 1978.

[156] M. Guthrie, R. Boehler, J. J. Molaison, B. Haberl, A. M. dos Santos, and C. Tulk. Structure and disorder in ice VII on the approach to hydrogen-bond symmetrization. *Physical Review B*, 99(18):184112, May 2019.

[157] Yukihiro Yoshimura, Sarah T. Stewart, Maddury Somayazulu, Ho kwang Mao, and Russell J. Hemley. High-pressure x-ray diffraction and raman spectroscopy of ice VIII. *The Journal of Chemical Physics*, 124(2):024502, January 2006.

[158] S. Klotz, K. Komatsu, H. Kagi, K. Kunc, A. Sano-Furukawa, S. Machida, and T. Hattori. Bulk moduli and equations of state of ice VII and ice VIII. *Physical Review B*, 95(17):174111, May 2017.

[159] Kazuki Komatsu, Stefan Klotz, Shinichi Machida, Asami Sano-Furukawa, Takanori Hattori, and Hiroyuki Kagi. Anomalous hydrogen dynamics of the ice VII–VIII transition revealed by high-pressure neutron diffraction. *Proceedings of the National Academy of Sciences*, 117(12):6356–6361, mar 2020.

[160] Linfeng Zhang, Han Wang, Roberto Car, and Weinan E. Phase diagram of a deep potential water model. *Physical Review Letters*, 126(23):236001, June 2021.

[161] R. J. Hemley, A. P. Jephcoat, H. K. Mao, C. S. Zha, L. W. Finger, and D. E. Cox. Static compression of H_2O-ice to 128 GPa (1.28 mbar). *Nature*, 330(6150):737–740, December 1987.

[162] E. Wolanin, Ph. Pruzan, J. C. Chervin, B. Canny, M. Gauthier, D. Häusermann, and M. Hanfland. Equation of state of ice VII up to 106 GPa. *Physical Review B*, 56(10):5781–5785, September 1997.

[163] Paul Loubeyre, René LeToullec, Elodie Wolanin, Michel Hanfland, and Daniel Hausermann. Modulated phases and proton centring in ice observed by x-ray diffraction up to 170 GPa. *Nature*, 397(6719):503–506, February 1999.

[164] Stefan Klotz. *Techniques in High Pressure Neutron Scattering*. CRC Press, dec 2012.

[165] R. J. Nelmes, J. S. Loveday, R. M. Wilson, J. M. Besson, Ph. Pruzan, S. Klotz, G. Hamel, and S. Hull. Neutron diffraction study of the structure of deuterated ice VIII to 10 GPa. *Physical Review Letters*, 71(8):1192–1195, August 1993.

[166] M. Guthrie, R. Boehler, C. A. Tulk, J. J. Molaison, A. M. dos Santos, K. Li, and R. J. Hemley. Neutron diffraction observations of interstitial protons in dense ice. *Proceedings of the National Academy of Sciences*, 110(26):10552–10556, June 2013.

[167] Razvan Caracas. Dynamical instabilities of ice x. *Physical Review Letters*, 101(8):085502, August 2008.

[168] A. Polian and M. Grimsditch. New high-pressure phase ofH2o: Ice x. *Physical Review Letters*, 52(15):1312–1314, April 1984.

[169] Chang-Sheng Zha, Russell J. Hemley, Stephen A. Gramsch, Ho kwang Mao, and William A. Bassett. Optical study of H_2O ice to 120gpa: Dielectric function, molecular polarizability, and equation of state. *The Journal of Chemical Physics*, 126(7):074506, February 2007.

[170] Maddury Somayazulu, Jinfu Shu, Chang sheng Zha, Alexander F. Goncharov, Oliver Tschauner, Ho kwang Mao, and Russell J. Hemley. In situ high-pressure x-ray diffraction study of H_2O ice VII. *The Journal of Chemical Physics*, 128(6):064510, February 2008.

[171] Viktor V. Struzhkin, Alexander F. Goncharov, Russell J. Hemley, and Ho kwang Mao. Cascading fermi resonances and the soft mode in dense ice. *Physical Review Letters*, 78(23):4446–4449, June 1997.

[172] Chunyi Zhang, Cui Zhang, Mohan Chen, Wei Kang, Zhuowei Gu, Jianheng Zhao, Cangli Liu, Chengwei Sun, and Ping Zhang. Finite-temperature infrared and raman spectra of high-pressure hydrogen from first-principles molecular dynamics. *Physical Review B*, 98(14):144301, oct 2018.

[173] Yael Bronstein, Philippe Depondt, Fabio Finocchi, and Antonino Marco Saitta. Quantum-driven phase transition in ice described via an efficient langevin approach. *Physical Review B*, 89(21):214101, June 2014.

[174] Pascal Vinet, John R. Smith, John Ferrante, and James H. Rose. Temperature effects on the universal equation of state of solids. *Physical Review B*, 35(4):1945–1953, February 1987.

[175] W. B. Holzapfel, B. Seiler, and M. Nicol. Effect of pressure on infrared-spectra of ice VII. *Journal of Geophysical Research*, 89(S02):B707, 1984.

[176] K. R. Hirsch and W. B. Holzapfel. Effect of high pressure on the raman spectra of ice VIII and evidence for ice x. *The Journal of Chemical Physics*, 84(5):2771–2775, March 1986.

[177] Chang-Sheng Zha, John S. Tse, and William A. Bassett. New raman measurements for H_2O ice VII in the range of 300 cm-1 to 4000 cm-1 at pressures up to 120 GPa. *The Journal of Chemical Physics*, 145(12):124315, September 2016.

[178] A. A. Maradudin and S. H. Vosko. Symmetry properties of the normal vibrations of a crystal. *Rev. Mod. Phys.*, 40(1):1–37, Jan 1968.

[179] J. L. Warren. Further considerations on the symmetry properties of the normal vibrations of a crystal. *Rev. Mod. Phys.*, 40:38–76, Jan 1968.

[180] Z.W. Hendrikse, M.O. Elout, and W.J.A. Maaskant. Computation of the independent elements of the dynamical matrix. *Computer Physics Communications*, 86(3):297 – 311, 1995.

[181] Atsushi Togo and Isao Tanaka. Spglib: a software library for crystal symmetry search. *arXiv preprint arXiv:1808.01590*, 2018.

[182] Lorenzo Paulatto, Francesco Mauri, and Michele Lazzeri. Anharmonic properties from a generalized third-order *ab initio* approach: Theory and applications to graphite and graphene. *Phys. Rev. B*, 87:214303, Jun 2013.

[183] Xavier Gonze and Changyol Lee. Dynamical matrices, born effective charges, dielectric permittivity tensors, and interatomic force constants from density-functional perturbation theory. *Phys. Rev. B*, 55:10355–10368, Apr 1997.

[184] Paolo Giannozzi, Stefano de Gironcoli, Pasquale Pavone, and Stefano Baroni. Ab initio calculation of phonon dispersions in semiconductors. *Phys. Rev. B*, 43:7231–7242, Mar 1991.

[185] Jer-Lai Kuo, Michael L. Klein, and Werner F. Kuhs. The effect of proton disorder on the structure of ice-ih: A theoretical study. *The Journal of Chemical Physics*, 123(13):134505, October 2005.