Dissertation

Ab-initio modelling of material properties using elements of artificial intelligence

carried out for the purpose of obtaining the degree of Doctor of Natural Sciences, submitted at TU Wien, Institute of Materials Chemistry, by

Sebastian Bichelmaier
Mat.Nr.: 01325275

under the supervision of

Univ. Prof. Dr. Georg K. H. Madsen
Dr. Jesús Carrete Montaña
E165
Institute of Materials Chemistry

Vienna, March 2023

ii

Affidavit

I declare in lieu of oath, that I wrote this thesis and performed the associated research myself, using only literature cited in this volume. If text passages from sources are used literally, they are marked as such. I confirm that this work is original and has not been submitted elsewhere for any examination, nor is it currently under consideration for a thesis elsewhere.

Vienna, March 2023

Acknowledgements

Abstract

Density functional theory (DFT) has established itself as one of the staple tools for materials simulations. Methodological advances achieved in the past decades have alleviated one of the main weaknesses of DFT: the lack of temperature-dependent treatment. Arguably the most widespread method of including temperature is the harmonic approximation (HA). It is based on modelling the displacement-force relationship of the individual atoms by a harmonic potential. The nuclear Hamiltonian then consists of independent quantum harmonic oscillators which yield an expression for the free energy. Combining this with the electronic ground state energy obtained through DFT yields a temperature-dependent free energy for a system.

However, for systems unstable at $0\,\mathrm{K}$, in particular those stabilized by temperature, the HA is not applicable and other methods, such as molecular dynamics (MD) or effective harmonic potentials (EHP) need to be used. Both rely on importance sampling of the potential energy surface (PES) and thus require significantly more DFT evaluations than needed for the HA. While meeting these increased resource demands may be possible for high-symmetry configurations or short timescales, the computational cost quickly approaches prohibitive levels. It is thus natural that researchers aim at introducing surrogate models (SM) that calculate results much more rapidly than pure DFT. In particular, recent advances in machine learning (ML), provide access to exceptionally accurate SM and are thus in the focus of research across the field of computational materials science.

In this thesis the temperature-dependent behavior of HfO_2, a material commonly associated with a temperature-stabilized cubic (cI) phase, $Fm\bar{3}m$, is explored. For this high-symmetry structure, investigation by a DFT-backed EHP approach is made possible by including a reweighting procedure. It is shown, that said reweighting allows direct evaluation of the term responsible for describing anharmonicities in the EHP formalism, as well as the use of unregularized regression techniques. The results, such as thermal expansion and bulk moduli, appear to agree well with experimental data from literature.

For the lower-symmetry monoclinic (m) and tetragonal (t) phases, a DFT-based approach would incur unfeasible computational cost, thus the use of a SM in the form a neural network force field (NNFF) is explored for their EHP treatment. In the second manuscript, a NNFF data acquisition and training strategy are detailed, yielding a parametrization, with an accuracy comparable to ab-initio calculations at a fraction of the cost. The NNFF performs well on the m- and t-phase, the previously studied cI and a lower-symmetry cubic (cII) phase presenting the $P\bar{4}3m$ spacegroup. Excellent agreement of the thermal expansion of the m- and t-phase with experimental data is found. This is in contrast to the results obtained for both cubic phases, where lattice constants substantially lower than experiment are found. While it is shown that cII is favorable over cI, a phase transition to any cubic phase is not observed. It is thus hypothesized

vi

that cubic HfO_2 is only present in a defect-stabilized form.

The advantages of NNFFs over other SM, specifically Taylor expansions of the PES, become clear in the last manuscript included in this thesis. Automatic differentiation makes direct evaluation of said high-order Taylor potentials accessible. This approach is investigated in detail, based on three simple systems, a six-atom Lennard-Jones (LJ) cluster, an fcc-LJ-solid, as well as a silver cluster. The limited flexibility of polynomials, i.e. the fact that they can only tend to $\pm\infty$ for large arguments, leads to significant artifacts in free energies and derived quantities. Hence it is concluded that global interpolation strategies, such as NNFFs, are better suited as cost-effective SM and a critical look at power-series expansions and their applicability is recommended.

Kurzfassung

Dichtefunktionaltheorie (DFT) hat sich als eines der zentralen Werkzeuge für Materialsimulationen etabliert. In den vergangenen Jahrzehnten konnte durch methodische Weiterentwicklungen eine der größten Schwächen von DFT behoben und temperaturabhängige Studien ermöglicht werden. Eine der weitverbreitetsten Methoden zur Berücksichtigung von Temperatur ist die harmonische Approximation (HA). In dieser Näherung wird die Beziehung zwischen Auslenkung und Kraft einzelner Atome durch ein harmonisches Potential modelliert. Der dadurch bedingte Hamiltonoperator der Kerne beschreibt dann unabhängige quantenmechanische harmonische Oszillatoren, die einen Ausdruck für die freie Energie ergeben.

Allerdings ist die HA für Systeme, die bei $0\,$K nicht stabil sind, und auch für solche, die durch Temperatur stabilisiert werden, nicht anwendbar. In diesem Fall muss man auf andere Methoden wie Molekulardynamik (MD) oder effektive harmonische Potentiale (EHP) zurückgreifen. Beide dieser Methoden erfordern das Abtasten der potentiellen Energieoberfläche (PES) durch relevanzbasierte Stichprobenverfahren, wodurch die Anzahl an benötigten DFT Berechnungen signifikant ansteigt. Für Systeme mit hoher Symmetrie oder kurze Zeitspannen, kann der damit verbundene, gestiegene Ressourcenbedarf noch bewältigbar sein, allerdings nähert sich der Berechnungsaufwand schnell unerschwinglichen Größenordnungen. Daher versuchen Wissenschafter Surrogatmodelle (SM) einzuführen, die Resultate wesentlich ressourcenschonender als DFT erzielen können. Besonders die rasanten Entwicklungen im Bereich des maschinellen Lernens ermöglichen außergewöhnlich genaue SM, weswegen deren Verwendung in der computerunterstützten Materialchemie zunehmend weit verbreitet ist.

In dieser Arbeit wird das Temperaturverhalten von HfO_2 untersucht. Dieses Material ist bekannt dafür, eine temperaturstabilisierte kubische (cI) Phase, $Fm\bar{3}m$, zu haben. Da es sich hierbei um eine Hochsymmetriephase handelt, kann man mithilfe eines Gewichtungsverfahren eine DFT-basierte EHP Studie durchführen. Ebendieses Gewichtungsverfahren ermöglicht eine direkte Auswertung des Terms der Anharmonizitäten in EHP beschreibt, sowie die Verwendung nicht regularisierter Regressionstechniken. Thermisches Ausdehnungsverhalten, sowie Kompressionsmodul stehen in Einklang mit literaturbasierten experimentellen Ergebnissen.

Im Falle der monoklinischen (m) und tetragonalen (t) Phasen von HfO_2, welche geringere Symmetrien aufweisen, würde eine solche DFT-basierte EHP-Studie einen nicht durchführbaren rechnerischen Aufwand bedeuten, weshalb ein Kraftfeld basierend auf einem neuronalen Netzwerk (NNFF) als SM untersucht wird. Im zweiten Manuskript dieser Arbeit werden eine Datenakquise- und Trainingsstrategie beschrieben, mit deren Hilfe ein NNFF, mit einer Genauigkeit vergleichbar zu ab initio Methoden, aber deutlich vermindertem Rechenaufwand, erzielbar ist. Das NNFF kann sowohl die m- und t-Phase, sowie die bereits zuvor untersuchte cI-Phase, als auch eine kubische (cII) Phase niedrigerer Symmetrie, die der Raumgruppe $P\bar{4}3m$ entspricht, akkurat beschreiben.

viii

Für die m- und t-Phase ist die Übereinstimmung der thermischen Ausdehnung mit dem Experiment hervorragend, im Falle der beiden kubischen Phasen hingegen wird das Volumen deutlich unterschätzt. Während gezeigt werden konnte, dass die cII-Phase energetisch günstiger ist, als die cI-Phase, konnte kein Phasenübergang von der tetragonalen in eine der beiden kubischen Phasen festgestellt werden. Daher wird die Vermutung aufgestellt, dass kubisches HfO_2 ausschließlich in einer defektstabilisierten Form existiert.

Die Vorteile von NNFF gegenüber anderen SM, insbesondere Taylor-Expansionen der PES werden im letzten Manuskript dieser Arbeit herausgearbeitet. Automatische Differenzierung machen eine direkte Evaluierung besagter Taylor-Expansionen zugänglich. Diese Methodik wird basierend auf drei Modellsystemen untersucht: Einem Lennard-Jones (LJ) Cluster bestehend aus sechs Atomen, einem kfz-LJ Festkörper, sowie einem Silbercluster. Die Tatsache, dass Polynomialfunktionen bei großen Argument im Grenzwert lediglich die Werte $\pm\infty$ annehmen können, führt hierbei zu signifikanten Artefakten in den freien Energien und daraus abgeleiteten Größen. Die Verwendung von globalen Interpolationsstrategien, beispielsweise NNFF, als SM ist daher gegenüber Expansionen zu bevorzugen.

Contents

List of Figures

Chapter 1

Introduction

1.1 Motivation

Since they were conceived nearly six decades ago, density-functional-theory-based (DFT) electronic structure methods and the ecosystem surrounding them, have come a long way and are now an integral part of materials science research. This can partially be attributed to the enormous increase in computational power, but without methodological advances, these improvements would only allow obtaining a ground state energy faster, or for a bigger system. The developments that have pushed computational chemistry into "mainstream" research are methodological. Some of the key contributors of those methodological advances are those alleviating one of the main weaknesses of DFT - the lack of a treatment of temperature.

Temperature, however, is one of the most important quantities linking theoretical studies to real-life materials, products and processes. Not only are temperature-dependent phenomena responsible for thermal expansion, thus impacting e.g. strain engineering [1] or thermal management [2] in semiconductor technologies. But microscopic phase stability and hence the macroscopic behavior of materials, is also governed by it. Going beyond solids, predictive capabilities of theoretical studies describing some of the most important fields of our time, predicting the activity of catalysts [3], or the stability and potential uses of pharmaceutical compounds [4], benefit greatly from the inclusion of temperature. All of these problems, and the many more that are being explored in research groups all over the world, pose a serious challenge for purely experimental studies: The parameter space, e.g. chemical composition, arrangement or impurities, that needs to be searched is almost infinite. Due to monetary and temporal constraints enacting severely restricting boundary conditions upon the exploration, such research is often based on small, iterative improvements guided by experience. For the challenges ahead of us, be it shifting towards renewable energies, through novel solar and battery technologies, or hydrogen production and fuel cells, an iterative approach is likely not sufficient. Thus, computational methods will continue to be key enablers and perhaps even grow their importance as guides towards a solution. After all, "*bits are cheaper than atoms*" [5].

The idea of a "digital twin", acting as an exploratory model to rule out the most absurd ideas and close in on the most promising ones, is already widespread in an industrial setting, e.g. when constructing large machines. However, already before that in the 1990s the bioinformatics community coined the term of *in silico* [6] in the spirit of *in vivo* or *in situ*, to refer to biological

research done through computer simulations. Although this is a vogue expression, it describes the idea concisely and understandably, which could be why the terminology is starting to appear in an increasing number of DFT-based studies as well. In material science, however, a "digital materials twin" has not yet been achieved within an acceptable error tolerance or degree of usability and reaching this goal will likely still take years. Arguably though, the general idea behind it, is slowly becoming established in the minds of research and industry alike. From GUI-tools, like QuantumATK [7], over materials research consulting as done by Enthought[1], to the idea of a free database containing material parameters, such as the materials project [8]: DFT is approaching broad applicability in an industrial setting.

While DFT run on modern computers is quite useable already, the aforementioned vast search spaces can still create situations where resource limitations are met or exceeded, for example in low-symmetry systems or when predicting thermal transport properties. One of the most impactful recent concepts in all of modern computational chemistry was transferred from informatics into materials science: machine learning (ML). Neural-network force fields (NNFF), introduced in a 2007 publication by Behler [9], have experienced an enormous interest by the community. And justifiably so, as these approaches seem to alleviate computational issues with little to no downside and can thus be used as an appropriate surrogate model. Arguably, NNFFs could be the missing ingredient for true *in silico* materials studies.

ML models might even be the key to achieving the pinnacle of materials science: Materials by design. This is a common concept usually phrased in terms of a forward problem, i.e. predict the outcome based on inputs, and an inverse problem, i.e. predict the inputs based on desired outcomes. While it is certainly still a long road ahead until these problems are solved, there are many approaches being already explored today and the potential impact of the topic can hardly be overstated [10, 11].

1.2 Description of thesis

Timeline

The goal of this thesis is to develop and apply an above-described *in silico* approach in order to improve understanding of HfO_2, a material that has recently gathered much attention. It exhibits at large number of different phases, based on temperature, pressure, strain or dopants, each with different macroscopic properties. Initially, hopes were high to describe its ambient phases and their temperature-mediated transitions using well-established finite-difference- (FD) or perturbation-theory-based phonon methods, but it was soon realized that these descriptions are not suitable to provide an accurate model of the material. This prompted implementation of an idea already discussed in the 1950s, effective harmonic potentials (EHPs, [12, 13]), which essentially provide a temperature-dependent description of phonons and - more importantly - allow the study of temperature-stabilized phases, as some of HfO_2 are. This is described in the first article below.

Using the abundant data created for the cubic phase in the first article, as well as some additional data generated for the tetragonal and monoclinic phases, a systematic approach on

---

[1]https://www.enthought.com/materials-science-chemistry/

creating a proper training set, based on the iterative nature of EHPs, was devised. This approach is then used to construct a potential that transfers well between (at least) four phases of HfO_2. The NNFF is then used as a backend for EHP calculations of the phases and while achieving great agreement for the tetragonal and monoclinic phase, a large mismatch for the cubic phase is observed. There is, however, some ambiguity regarding its high-temperature phase and in article II an argument is presented that the spacegroup traditionally assumed is incorrect and that perhaps a stoichometric cubic phase does not exist at all.

The differences between FD- and EHP-based approaches were more significant than anticipated. To explore this idea further, we studied the applicability of such "simple" (essentially power-series-based) descriptions of the potential energy surface (PES) on a few simple toy models, a Lennard-Jones molecule and solid, and a silver cluster. While methodologically this was relatively well-trodden territory, the implications are quite striking and the work presented in article III hopefully convinces some of our colleagues to take care in evaluating the limitations of their approaches. Most of the work there was only made possible through automatic differentiability as implemented in Google's JAX [14, 15].

JAX is also what powers the NNFF used in our group (and beyond): NeuralIL. Much combined effort has been and continues to be put into developing augmentations and improvements for it, as it has become the working horse for almost all our projects from ionic liquids [16] and surfaces [17], to bulk solids (HfO_2 [18, 19], SiO_2, ...) and soon interfaces as well. Apart from numerous bug fixes, I worked on including a charge equilibration scheme and a repulsive short-range contribution.

Finally, an extension to the already-existing JAX-MD code [20], which integrates exceedingly well with our NNFF, is proposed. It supports not only isotropic cell fluctuations but also flexible changes of the lattice vectors. This allows studying phase transitions *in silico*. While thematically this study is part of this thesis and much progress has been achieved during it, the manuscript is still under construction and the results presented in chapter 5 preliminary.

Structure

The material HfO_2 is introduced in the second chapter. The basis and background for the attached articles is outlined in the third chapter, where we use examples and results from the already-published work to illustrate the theory. The fourth chapter contains brief and concise summaries of the most important publications written in the course of this thesis, while the fifth chapter presents the unpublished molecular dynamics study. Specifics of the published work are detailed within the corresponding manuscripts which can be found in the last chapter and more in-depth considerations are available at the references therein. The thesis concludes with a brief review of recent literature, combining the learnings and motivation into an outlook for what computational chemistry and materials science can bring to the table in the foreseeable future.

Chapter 2

Hafnia

Hafnia (HfO_2) has a multi-faceted phase diagram [21], with numerous industrially relevant applications: It is being used as a high-κ gate dielectric for semiconductors in its amorphous [22] and tetragonal phase [23]. Due to its high melting point of approximately 3100 K [24] and neutron absorption coefficient, hafnia-toughened ceramics are relevant for the nuclear industry [25]. Owing to its low thermal conductivity, HfO_2 lends itself to applications in thermal barriers [26]. There are reports of hafnia layers improving stability of ceria-based solid oxide fuel cells [27]. In its doping-stabilized [28] or pure [29] ferroelectric phase, hafnia is being considered for e.g. nonvolatile memory applications [30] or ferroelectric field-effect transistors [31]. The material has been researched and used for decades and industry has acquired a lot of expertise in its processing. Due to this and the potential broad technological applications, HfO_2 continues to be of high interest today. However, even though it has been studied for well over half a century now [24, 25, 32–44], its phase structure is still relatively poorly understood. An indication of this is that ferroelectricity was only discovered in 2011 [28], another that phase transition temperatures reported in literature span over 1000 K [24].

Crystal structure

At ambient pressures, it is believed that hafnia transitions from a monoclinic $P2_1c$ crystal structure to a tetragonal $P4_2nmc$ one between approximately 1800 K to 3000 K [24, 45]. With increasing temperature it is believed to transition to a high-symmetry cubic phase, presenting an $Fm\bar{3}m$ crystal structure, before melting [46]. With an applied pressure, one finds various nonpolar orthorhombic ($Pbca$, $Pnma$) phases, while strain or doping can lead to polar orthorhombic phases ($Pca2_1$, $Pmn2_1$) responsible for the ferroelectric properties. The discovery of these [28] has prompted renewed interest in the material's experimental and theoretical description, as well as potential applications [47]. Furthermore, there are reports of various different (meta)-stable phases discovered in theoretical investigations [38]. A non-exhaustive overview is given in Fig. 2.1.

However, similar to the isostructural material ZrO_2 [49–53] there are ambiguities with respect to the existence of a stable stoichiometric cubic phase. Not only do the reported tetragonal-to-cubic phase transition temperatures obtained from different authors span almost 500 K [24], but there is even some doubt regarding the precise space group of cubic HfO_2 [54]. While some researchers find a mixed phase of cubic and tetragonal symmetry suggesting metastability [55],
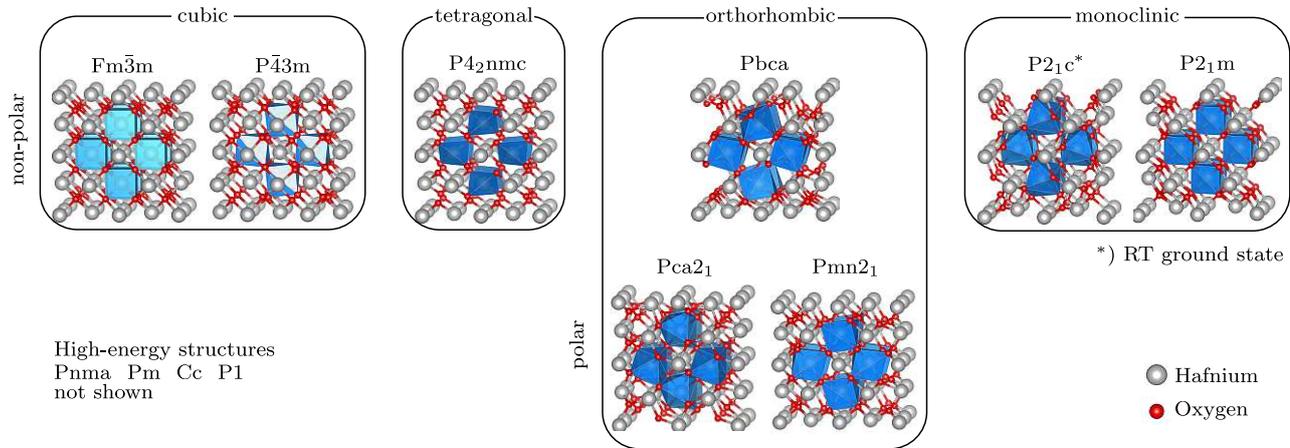
Figure 2.1: Phases of HfO_2, where $P2_1c$ constitutes the room temperature ground state structure. Several (very high energy) phases are not shown. The visualizations of the atomic structures were created using VESTA [48].

others see a clear second-order transition [42, 46]. In agreement with Barabash et al.'s [56] initial assessment, which disregarded temperature, we find the lower-symmetry $P\bar{4}3m$ cubic phase to be energetically favorable over the commonly-believed $Fm\bar{3}m$ one even with temperature included. To our knowledge there is, besides our work, no theoretical temperature-dependent analysis of the high-temperature phase stability of HfO_2 that would allow identification of the space group. Furthermore, there are only few experimental studies describing a tetragonal-to-cubic phase transition [42, 55, 57] and those are not in agreement.

Experimental studies at such extreme temperature conditions are difficult and, in particular for HfO_2, precise knowledge of the stoichiometry and potential impurities is key for an accurate interpretation of experimental results: Temperature-induced or processing-induced oxygen vacancies stabilize high-symmetry phases [58–61] and there are even reports of a room temperature cubic phase of (severely) oxygen deficient $HfO_{1.7}$ [59]. Similarly, numerous reports detail the effects a variety of dopants have on HfO_2 phase stability. These range from the ferroelectric orthorhombic phase (Si: [28], La: [62], Sr: [63], Gd: [64]), to the tetragonal (Y, Si, Gd: [65]) and even the cubic (Y: [66]). This implies the expressed phase has a high sensitivity to (e.g. processing-induced) impurities as well. Furthermore, several studies show the significant impact strain has on phase equilibria [3, 67–69]. Strain is an unavoidable consequence in the manufacturing of thin-layer technologies such as semiconductors, but it just as well occurs in powders commonly used for x-ray diffraction measurements in the form of surface strain.

All these factors combine in an intricate interplay which determines the exhibited crystal structure of HfO_2 and perhaps provide an explanation as to why over 60 years of research have not yet shed light on all phenomena present in this material. As such, even though there is a large number of studies published already, there is still a lot to learn about this material and I hope that this thesis offers a small contribution to this undertaking.

Previous computational work

A good starting point for the need of theoretical work supporting and verifying experimental studies is the article published by Wang et al. [24]. This publication is an exceptional source, containing not only their own differential thermal analyses of the ambient phase transitions of HfO_2, but also a detailed overview of previous experimental work covering over 20 years of publications. What becomes clear quickly, is that transition temperatures vary widely: Monoclinic to tetragonal ranges 1773 K to 2973 K with an average of 2200 K and Wang et al.'s own measurements amounting to 2066 K. The tetragonal-to-cubic transition shows an average of 2940 K and finally the average melting temperature measured is 3090 K.

After the somewhat surprising discovery of a ferroelectric orthorhombic phase in 2011 [28], several studies exploring the phase stability of HfO_2 were performed. In particular the work done by Huan et al. in 2014 [38], has established a nowadays widely accepted assortment of hafnia phases through a minima-hopping method. Based on the harmonic approximation - a simple model for incorporating temperature - a first attempt at a phase diagram was made there. However, methodologically, imaginary phonons occurring in some of the structures, render the free energy calculations unconvincing.

This work was extended by Barabash [56], who implemented a systematic search based on an enumeration of displacements and distortions of the fluorite structure. Apart from numerous relatively high energy structures, this work is to the best of my knowledge the first that deals with the lower-symmetry cubic $P\bar{4}3m$ phase. Antunes et al. [54] attempted the first thermodynamic assessment of this "new" cubic phase, again based on the harmonic approximation and - as they correctly point out - again producing "*unreliable*" results with "*well-known theoretical issues*" [54] that are raised by the presence of imaginary phonon frequencies.

A machine learning-based molecular dynamics study performed using DeepMD [70] was published by Wu et al. in 2021 [71]; however, temperatures larger than 2500 K are unfortunately not considered. They again show a monoclinic-to-tetragonal transition at approximately 2000 K confirming the results found by Fan et al., Antunes et al. and Huan et al. [38, 46, 54]. Haggerty et al. [39] used HTXRD to catalogue the thermal expansion behavior of monoclinic and tetragonal hafnia, highlighting in particular the anisotropic thermal expansion of monoclinic HfO_2.

One of the few experimental publications describing a measured tetragonal-to-cubic transition is provided by Tobase et al. [42]. There high-temperature x-ray diffraction (HTXRD) was performed on a levitating HfO_2 powder, measuring the lattice constants. Based on the c/a ratios, they conclude a transition towards cubic is taking place at around 2800 K. Similar results were obtained in an ab initio molecular dynamics study by Fan et al. [46]. There, a tetragonal-to-cubic phase transition was observed at 2600 K. Their study, however, has its limitations, many of which are stemming from the large computational cost of ab initio molecular dynamics. The small simulation box allows no long-wavelength deformations and the quick scans over a wide temperature range make relaxation into a steady state unlikely. Furthermore, even *if* they indeed observe a cubic phase, it would not be possible impossible to conclusively determine whether it corresponds to $Fm\bar{3}m$ or $P\bar{4}3m$.

There is still no clear road towards a theoretical description of a tetragonal-to-cubic phase transition and unambiguous determination of the space group. As is discussed above, literature knows many cases of non-stoichiometric hafnia stabilized in a cubic structure through oxygen

vacancies. The significant strains and extreme heat unavoidable in such experimental studies, can have a drastic impact and thus a detailed theoretical investigation of potential phenomena promoting cubicity in HfO_2 is warranted.

Chapter 3

Background

3.1 Density functional theory

While density functional theory (DFT) is essential for the work presented here, it is just a tool in the toolbox and not the focus of this thesis. Hence, only a brief overview of the most essential concepts is given here.

When Walter Kohn and Pierre Hohenberg first introduced DFT in their 1964 paper [72] they probably had not anticipated the impact and subsequent Nobel prize their findings would yield. In principle, their theory allows finding the correct ground-state energy, E_0, of a many-body (stationary) Schrödinger equation in the Born-Oppenheimer (BO) approximation [73], i.e. for "static" nuclei, located at $\mathbf{R}_k$ with charge Z_k

$$\hat{H} \ket{\Psi_n} = E_n \ket{\Psi_n}, \tag{3.1}$$

with the many-body electronic Hamiltonian

$$\hat{H} = -\frac{1}{2} \sum_i \hat{\nabla}_i^2 + \frac{1}{2} \sum_{i \neq j} \frac{1}{|\hat{\mathbf{r}}_i - \hat{\mathbf{r}}_j|} - \sum_{i,k} \frac{Z_k}{|\hat{\mathbf{r}}_i - \mathbf{R}_k|}, \tag{3.2}$$

based on the ground-state density, $n_0(\mathbf{r}) = \int d\mathbf{r}_2 \dots d\mathbf{r}_N |\Psi_0(\mathbf{r}, \mathbf{r_2}, \dots, \mathbf{r_N})|^2$ alone. At the core of the methodology lie two key points. First, a one-to-one correspondence of the density n_0 and the external potential $v_{\text{ext}} = -\sum_k \frac{Z_k}{|\hat{\mathbf{r}}_i - \mathbf{R}_k|}$, which parametrically depends on the nuclear coordinates, $\mathbf{R}_k$, and uniquely defines the potential energy surface and the Hamiltonian. Second, the variational principle, guaranteeing an energy minimum at the exact ground-state density (for a given set of $\mathbf{R}_k$, i.e. v_{ext}). Or, as Becke puts it in his 2014 perspective on the 50 year anniversary of DFT [74]

$$n_0 \rightarrow v_{\text{ext}} \rightarrow \Psi_0 \rightarrow \text{everything!}. \tag{3.3}$$

3.1.1 Kohn-Sham equations

A year later, Kohn and Sham published the Kohn-Sham (KS) equations [75], and by utilizing the Hohenberg-Kohn theorems, provided a self-consistency framework enabling a tractable and *practical* way forward. This is achieved by reformulating the many-body interacting-electron

Schrödinger equation (wavefunctions $|\Psi\rangle$) into a system describing non-interacting (NI) electrons (wavefunctions $|\phi\rangle$). The KS energy is given by

$$E(n) = T_0(n) + \int v_{\text{ext}}(\mathbf{r})n(\mathbf{r})d\mathbf{r} + J(n) + E_{\text{xc}}(n), \tag{3.4}$$

where

$$T_0 = \min_{\{\text{SD } n(\mathbf{r})\}} -\frac{1}{2}\sum_i \langle\phi_i|\nabla^2|\phi_i\rangle \qquad J = \frac{1}{2}\int \frac{n(\mathbf{r}_1)n(\mathbf{r}_2)}{r_{12}}d\mathbf{r}_1 d\mathbf{r}_2 \tag{3.5}$$

are the NI kinetic energy and Coulomb self-energy, respectively. Here $n(\mathbf{r})$ is constructed as

$$n(\mathbf{r}) = \sum_i |\langle\phi_i|\mathbf{r}\rangle|^2. \tag{3.6}$$

Finally, E_{xc} is the exchange and correlation energy, correcting for the error incurred by going from $|\Psi\rangle$ to $|\phi\rangle$, i.e. from interacting to non-interacting electrons by - at least formally - restoring the total Coulomb interaction energy V_{ee} and total kinetic energy T

$$E_{\text{xc}} = T + V_{\text{ee}} - T_0 - J. \tag{3.7}$$

Thus, the solutions of the Kohn-Sham-equations obtained through minimization of

$$\left(-\frac{1}{2}\nabla^2 + v_{\text{ext}}(\mathbf{r}) + \int d\mathbf{r}' \frac{n(\mathbf{r})}{|\mathbf{r} - \mathbf{r}'|} + \frac{\delta E_{\text{xc}}}{\delta\rho}(\mathbf{r})\right)\phi_i(\mathbf{r}) = \epsilon_i\phi_i(\mathbf{r}), \tag{3.8}$$

are in principle exact. The KS formalism hence allows studying real many-body systems in terms of NI single particle wavefunctions whose behavior is governed by an effective potential. However, approximations of E_{xc}, or rather, ϵ_{xc} defined through

$$E_{\text{xc}} = \int d\mathbf{r}\,\epsilon_{\text{xc}}(\mathbf{r})n(\mathbf{r}), \tag{3.9}$$

must be made.

3.1.2   XC functionals

Constructing and studying different xc-functionals is a scientific discipline in itself, hence I will limit myself to two well-known and widespread choices. The exchange contribution in the local density approximation (LDA) was already published in the original paper [75] and amounts to that of the homogeneous electron gas, $\epsilon_{\text{x}}^{\text{LDA}}(n) = -\frac{3}{4}\left(\frac{3}{\pi}\right)^{\frac{1}{3}}n^{\frac{1}{3}}$. The correlation contributions, $\epsilon_{\text{c}}^{\text{LDA}}$, can be obtained by interpolating quantum Monte Carlo results, as is done in [76]. Thus, the full xc-energy can be calculated through

$$E_{\text{xc}}^{\text{LDA}} = \int d\mathbf{r}\left(\epsilon_{\text{x}}^{\text{LDA}}(\mathbf{r}) + \epsilon_{\text{c}}^{\text{LDA}}(\mathbf{r})\right)n(\mathbf{r}). \tag{3.10}$$

A logical extension to this, is to not only include the density, but also its gradient, $\nabla n(\mathbf{r})$, [74]. This is typically done in the form of the reduced density gradient, $s = \frac{|\nabla n|}{2(3\pi^2)^{1/3}n^{4/3}}$, which is used
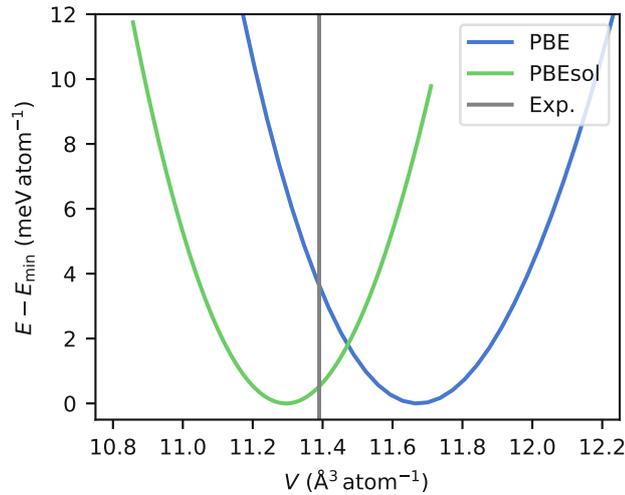
Figure 3.1: Energy-volume curves for $P2_1c$-HfO$_2$ obtained with PBE and PBEsol as compared to experiment [39].

to define a so-called enhancement factor, f_x. With this the generalized gradient approximation, or GGA, exchange can be introduced

$$\epsilon_\mathrm{x}^\mathrm{GGA} = \epsilon_\mathrm{x}^\mathrm{LDA} f_\mathrm{x}(s^2),\tag{3.11}$$

with one of the most ubiquitous flavors being the PBE functional [77]:

$$f_\mathrm{x}^\mathrm{PBE}(s^2) = 1 + \kappa - \frac{\kappa}{1 + \frac{\mu s^2}{\kappa}}.\tag{3.12}$$

Here the parameters $\kappa = 0.804$ and $\mu = 0.21951$ are determined by enforcing adherence to some selected theoretical constraints. Different choices of κ and μ lead to different behaviors but are just reparametrizations of the same functional. One such reparametrization is PBEsol [78].

Which functional is used depends on a virtually endless list of factors. If the results should be comparable to existing literature, likely the same functional will be chosen. LDA underestimates lattice constants, PBE overestimates them - a fact that is mostly corrected by PBEsol (see Fig. 3.1). But, even as Perdew himself says in [78], "*[a]t the GGA level, one must choose*", the choice is always a trade-off between accuracy in energy, bandgap or lattice constant and other factors, more specific to the planned study at hand. This work mostly employs PBE, as most literature data was obtained using this functional, where adequate and reasonable, we compare to PBEsol.

This is just a glimpse of the broad spectrum of functionals available, for a more comprehensive overview, the reader is referred to relevant literature, e.g. Ref. [79] for solids, or Ref. [80] for molecules.

3.2 Temperature-dependent phenomena

A typical shortcoming of DFT approaches when comparing the results with actual measurements, is the complete absence of any notion of temperature. Most DFT-backed calculations focus on

the electronic ground state energies (E_0) or differences thereof, which are purely mechanical quantities. To introduce temperature, actual thermodynamic properties need to be studied. E.g. the Helmholtz free energy

$$F = U - TS, \tag{3.13}$$

or the Gibbs free enthalpy,

$$G = F + PV, \tag{3.14}$$

that link together internal energy U, temperature, T, entropy, S, pressure, P and volume, V. For a wide-band-gap semiconductor, electronic excitations are unlikely, the internal energy can be approximated as the thermodynamic average of the potential energy

$$U = \langle E_{\mathrm{BO}} \rangle, \tag{3.15}$$

where

$$E_{\mathrm{BO}}(\{\mathbf{R}_i\}) = E_0(\{\mathbf{R}_i\}) + \frac{1}{2} \sum_{i \neq j} \frac{Z_i Z_j}{|\mathbf{R}_i - \mathbf{R}_j|} \tag{3.16}$$

is the BO potential energy.

The entropy itself has several constituents with the most important for us being the vibrational, or lattice contribution, S_{vib}. Approximations to S_{vib} and by extension, F and G, are ubiquitous, however, the most obvious and widely used one is the harmonic approximation.

3.2.1   Harmonic approximation

The harmonic approximation (HA) is, in essence, just an expansion of the BO potential energy in a power series as a function of the (mass-weighted) displacements, $\mathbf{u}_i = \sqrt{m_i}(\mathbf{R}_i - \mathbf{R}_{i,0})$ around the equilibrium atomic positions $\mathbf{R}_{i,0}$, truncated after the quadratic term:

$$E_{\mathrm{BO}}(\mathbf{u}_i) = E_{\mathrm{BO},0} + \sum_{i\alpha} \Phi_{i\alpha}^{(1)} u_{i\alpha} + \frac{1}{2} \underbrace{\sum_{i\alpha,j\beta} \Phi_{i\alpha,j\beta}^{(2)} u_{i,\alpha} u_{j\beta}}_{=E_{\mathrm{HA}}} + \dots, \tag{3.17}$$

where the indices i, j run over the number of atoms N_{atoms}, α, β over the Cartesian coordinates. The $\boldsymbol{\Phi}$ are the partial derivatives of the energy with respect to said displacements - the so-called force constants:

$$\Phi_{i\alpha,j\beta}^{(2)} = \Phi_{i\alpha,j\beta} = \frac{\partial^2 E_{\mathrm{BO}}}{\partial u_{i\alpha} \partial u_{j\beta}} \bigg|_{u_{i\alpha}=u_{i\beta}=0} = -\frac{1}{\sqrt{m_i}} \frac{\partial f_{i\alpha}}{\partial u_{j\beta}} \bigg|_{u_{i\alpha}=u_{i\beta}=0}. \tag{3.18}$$

The linear term, $\Phi_{i\alpha}^{(1)}$, vanishes for a structure in equilibrium, i.e. with no forces acting on it, which leaves a quadratic Hamiltonian. The quadratic form can be diagonalized by solving an eigenvalue problem[1]

$$\boldsymbol{\Phi} \boldsymbol{\epsilon}_n = \omega_n^2 \boldsymbol{\epsilon}_n. \tag{3.19}$$

---

[1]For a detailed derivation, please refer to [81]

Expressing the Hamiltonian in the basis of these eigenvectors, it takes the form of a system of uncoupled harmonic oscillators.

The link to DFT calculations is given through the Hellman-Feynman theorem [82], granting access to the forces of each atom, $\mathbf{f}_i$. The actual implementation within the PAW basis framework as used in VASP is detailed in [83, 84].

A compound is described on a periodic real-space lattice $\mathbf{R_I} = \sum_i^{n_{\dim}=3} I_i \boldsymbol{\alpha}_i$, indexing the periodic image $\mathbf{I} = (I_1, I_2, I_3) \in \mathbb{Z}^3$, with basis vectors $\boldsymbol{\alpha}_i$. On it, the coordinates of our atoms are expressed by placing a copy of a motif $\{\mathbf{R}_{i,0}\}$ at each node of the lattice:

$$\mathbf{R_{I},}_i = \mathbf{R_I} + \mathbf{R}_{i,0}. \tag{3.20}$$

Any such lattice is accompanied by a corresponding reciprocal-space lattice, made up of all vectors satisfying $e^{i\mathbf{Q}\mathbf{R_I}} = 1 \quad \forall \mathbf{I}$, or

$$\mathbf{Q}\mathbf{R_I} = 2\pi n, \tag{3.21}$$

for $n \in \mathbb{Z}$. Its respective reciprocal basis vectors $\boldsymbol{\beta}_i$ can hence be obtained from $\boldsymbol{\beta}_i \boldsymbol{\alpha}_j = 2\pi \delta_{ij}$. With this, we can directly evaluate Eq. (3.21) for an arbitrary reciprocal vector $\mathbf{Q} = \sum_i^3 J_i \beta_i$, to arrive at:

$$J_1 I_1 + J_2 I_2 + J_3 I_3 = n, \tag{3.22}$$

thus proving that the J_i are in fact also integers and $\mathbf{Q}$ forms a lattice as well. Now, for a reasonably well-behaved function, such as the eigenvectors of Eq. (3.19), the Fourier series

$$\tilde{\boldsymbol{\epsilon}}_n(\mathbf{q}) = \sum_{\mathbf{I}} e^{i\mathbf{q}\mathbf{R_I}} \boldsymbol{\epsilon}_n, \tag{3.23}$$

converges and the Fourier coefficients can be calculated by integrating over the first Brillouin zone (BZ), defined by the period of Eq. (3.23),

$$\boldsymbol{\epsilon}_n = \frac{1}{V_{\mathrm{BZ}}} \int_{\mathrm{BZ}} e^{-i\mathbf{q}\mathbf{R_I}} \tilde{\boldsymbol{\epsilon}}_n(\mathbf{q}) d\mathbf{q}. \tag{3.24}$$

After taking into account translational symmetry, combining Eq. (3.24) and Eq. (3.19), results in another eigenvalue problem

$$\mathbf{D}(\mathbf{q})\tilde{\boldsymbol{\epsilon}}_n(\mathbf{q}) = \omega_n^2(\mathbf{q})\tilde{\boldsymbol{\epsilon}}_n(\mathbf{q}), \tag{3.25}$$

centered around the dynamical matrix $\mathbf{D} = \sum_{\mathbf{I'}} e^{-i\mathbf{q}\mathbf{R_{I'}}} \boldsymbol{\Phi}_{0,\mathbf{I'}}$. Treating the problem in reciprocal space enables us to make a tractable approximation of an infinite number of real-space degrees of freedom using finite matrices.

The ability to calculate the forces of all reasonable structures allows construction of configurations with defined displacements and hence the calculation of force constants. However, most displacements can only be represented in so-called supercells, that is, $S_1 \times S_2 \times S_3$ repetitions of the unit cell. Periodic boundary conditions still apply; hence some aliasing effects remain and an additional term enters Eq. (3.18):

$$-\frac{1}{\sqrt{m_i}} \frac{\partial f_{i\alpha;0}}{\partial u_{j\beta}} = \Phi_{i\alpha,j\beta;0,\mathbf{I}} + \sum_{\text{images of } \mathbf{I}} \Phi_{i\alpha,j\beta;0,\mathbf{I'}} = \tilde{\Phi}_{i\alpha,j\beta;0,\mathbf{I}}, \tag{3.26}$$
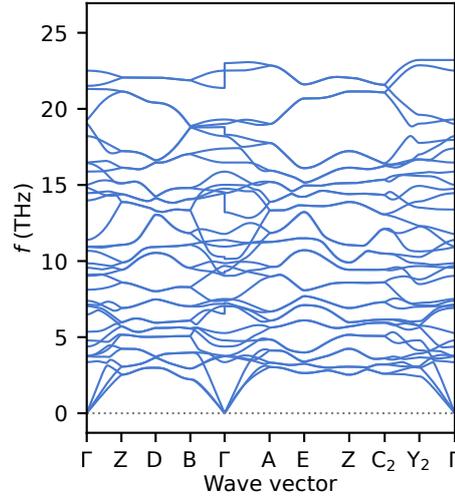
Figure 3.2: The phonon band structure of the $P2_1c$ phase of HfO_2 at $V = 11.67\,\text{Å}^3\,\text{atom}^{-1}$ obtained using DFT-backed FD calculations. The path in k-space was generated by SeeK-path [86].

resulting in the cumulant force constants [85], $\tilde{\boldsymbol{\Phi}}$ and finally the approximate dynamical matrix

$$\boldsymbol{\mathcal{D}}(\mathbf{q}) = \sum_{\mathbf{M}'} e^{-i\mathbf{q}\mathbf{R}_{\mathbf{M}'}} \tilde{\Phi}_{i\alpha,j\beta;0,\mathbf{M}'}, \tag{3.27}$$

where the $\mathbf{M}'$, now only include all the atoms within and on the surface of the supercell. For $\mathbf{q}$-points commensurate with the supercell, it describes the dynamical matrix exactly, while for incommensurate points it is an approximate Fourier interpolation, resulting in a frequency-, or band-plot as shown in Fig. 3.2.

Given the fact that all that can be hoped for is a discrete number of $\mathbf{q}$-points, the frequencies and eigenvectors will now receive an index $\mathbf{q}$, instead of being a function thereof. The modes obtained through the above, $\omega_{n\mathbf{q}}$ and $\boldsymbol{\epsilon}_{n\mathbf{q}}$, are occupied by bosonic particles, phonons, of energy $\hbar\omega_{n\mathbf{q}}$ and momentum $\hbar\mathbf{q}$. The $3N_{\text{atoms}}$-dimensional real-space density function describing the distribution of displacements, ρ, can be expressed using a covariance matrix, $\mathbf{C}$ [87]. The elements of $\mathbf{C}$, the C_{ij}, can be obtained as

$$C_{ij} = \frac{\hbar}{2\sqrt{m_i m_j}} \sum_{n\mathbf{q}} \frac{1}{\omega_{n\mathbf{q}}^2 \tanh\left(\frac{\hbar\omega_{n\mathbf{q}}}{2k_B T}\right)} \epsilon_{n\mathbf{q},i} \otimes \epsilon_{n\mathbf{q},j}^*. \tag{3.28}$$

Hence, ρ amounts to

$$\rho(\mathbf{u}) = \frac{1}{\sqrt{(2\pi)^{3N_{\text{atoms}}} |\mathbf{C}|}} \exp\left(-\frac{1}{2}\mathbf{u}\mathbf{C}^{-1}\mathbf{u}\right). \tag{3.29}$$

Assuming all obtained frequencies correspond to oscillations, i.e. $\omega_{n\mathbf{q}}^2 > 0$, the partition function of the harmonic oscillators can be built:

$$Z_{\text{harm}} = \prod_{\mathbf{q}n} \frac{e^{-\frac{\hbar\omega_{n\mathbf{q}}}{2k_B T}}}{1 - e^{-\frac{\hbar\omega_{n\mathbf{q}}}{k_B T}}}, \tag{3.30}$$
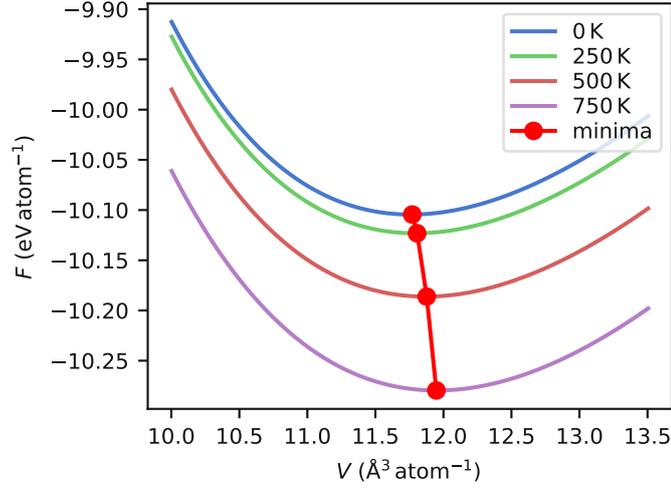
Figure 3.3: (F,V) curves at various temperatures for $P2_1c$ HfO$_2$, the red line indicating the corresponding minima.

and with it finally the Helmholtz free energy, $F = E_{\text{BO},0} + F_{\text{harm}}$, through,

$$F_{\text{harm}} = -k_B T \ln Z_{\text{harm}} = \sum_{n\mathbf{q}} \left(\frac{\hbar \omega_{n\mathbf{q}}}{2} + k_B T \ln \left[1 - e^{-\frac{\hbar \omega_{n\mathbf{q}}}{k_B T}} \right] \right). \quad (3.31)$$

Now, thermal expansion and the Gibbs free enthalpy are within reach: By performing phonon calculations at several volumes around the equilibrium volume of the unit cell and fitting an equation of state (EOS) a volume-dependency can be introduced. This broadly established extension of the HA is termed the *quasi*-harmonic approximation (QHA). With an EOS, such as Birch-Murnaghan's EOS (BMEOS, [88]) fitted onto a dataset of $(F(T), V)$, thermal expansion can be obtained by finding the minimum value of F for every value of T as indicated by the red line in Fig. 3.3. Furthermore as the BMEOS is an analytical expression, the pressure can be obtained as the derivative of the free energy with respect to volume:

$$\left. \frac{\partial F}{\partial V} \right|_{T=\text{const}} = -P. \quad (3.32)$$

Having obtained the pressure, G can also be computed and theoretically a phase diagram can be built. With PHONOPY, Atsushi Togo published a toolkit [89] which is nowadays widely used. Thanks to it, most of the steps described above are done automatically and the number of calculations is drastically reduced based on symmetry considerations. The user only has to care about properly converged force calculations, an adequate choice of the S_i and a mesh of $\mathbf{q}$-points for the Fourier interpolation.

However, there are a few shortcomings of the (Q)HA. As mentioned previously, the applicability of the statistical mechanics framework rests on $\omega_{n\mathbf{q}}^2 > 0$. If this does not hold the structure is dynamically unstable, the HA fails and the results are unreliable at best. Sometimes however, anharmonicities can alleviate this. To treat those phases or compounds one needs to resort

to different methods. Furthermore, scaling the volume does not necessarily lead to the lowest-energy thermal expansion and even relaxations for a given volume performed within DFT can be inaccurate at higher temperatures. At the absence of temperature the structure might behave differently than e.g. at $1000\,\mathrm{K}$. Finally, the limitations of the (Q)HA, which is after all a *local* Taylor-series-like expansion of the PES are rarely taken into consideration.

3.2.2   Effective harmonic potentials

Already in the 1950s Hooton et al. [12, 13] described an approach to treat anharmonicities using *effective* harmonic potentials. It was later reformulated using a variational Ansatz by Errea et al. [90, 91]: A trial density matrix, $\hat{\rho}$, is assumed, which exactly describes the statistics of a corresponding trial Hamiltonian, $\hat{\mathcal{H}}$, that only differs from the true Hamiltonian $\hat{H}$ through the form of the approximate potential energy operator, $\hat{\mathcal{U}}$, as opposed to the exact potential energy operator, $\hat{U}$. Minimizing the free energy with respect to the trial density matrix is guaranteed by the Gibbs-Bogoliubov inequality [92] to provide an upper bound on the free energy of the true system

$$F[\hat{U}, \hat{\rho}_0] \leq \mathcal{F}_{\mathrm{EHP}} = F[\hat{\mathcal{U}}, \hat{\rho}] + \mathrm{Tr}\left[\hat{\rho}(\hat{U} - \hat{\mathcal{U}})\right] = F_{\mathrm{harm}} + F_{\mathrm{corr}}, \tag{3.33}$$

As this is the effective *harmonic* approximation, $\hat{\mathcal{U}}$ corresponds to E_{HA} as defined in Eq. (3.17) and hence the trial density matrix is parameterized as Eq. (3.29). The term F_{harm} corresponds to Eq. (3.31). F_{harm} depends directly on the temperature T and indirectly on the harmonic trial potential through $\omega_{n\mathbf{q}}$. The optimal trial potential thus depends on the temperature. Ignoring F_{corr} [Eq. (3.33)] and the temperature dependence of the effective potential results in the harmonic approximation described in the previous section.

Determining the optimal EHP and calculating F_{corr} can be achieved through canonical importance sampling and building a self-consistency scheme around the aforementioned connection between $\mathcal{F}_{\mathrm{EHP}}$ and $\boldsymbol{\Phi}$, as is indicated in the inner-most box in Fig. 3.4. When self-consistency is reached, this corresponds to minimizing $\mathcal{F}_{\mathrm{EHP}}$ [93]. As a starting point, it is convenient to use the finite displacement (FD) second-order force constants and their corresponding potential. Then, using Eq. (3.29), a set, $\mathcal{S}$, of m displacements can be drawn, for which - by means of DFT or a surrogate model - potential energies and forces are calculated. These forces are then used to parameterize another set of force constants, leading to a new density, a new set of displacements and finally new forces and energies. Here, we use hiPhive [94] to efficiently determine force constants while also relying on their routines to simplify the significant bookkeeping involved.

This cycle continues until convergence is reached, which is aided by two augmentations:

1. Obtaining the new trial density matrix through a Pulay mixing scheme [95]. In the work presented throughout this thesis a memory of $n = 5$ steps and a mixing parameter of $\alpha = 0.1$ was chosen.

2. Reusing configurations from previous iterations (g) in the current iteration (k) through a reweighting procedure. As the probability densities are known, a reweighting factor is obtained easily:

$$w^{(g \to k)} = \frac{\rho^{(k)}(\mathbf{u}^{(g)})}{\rho^{(g)}(\mathbf{u}^{(g)})}. \tag{3.34}$$
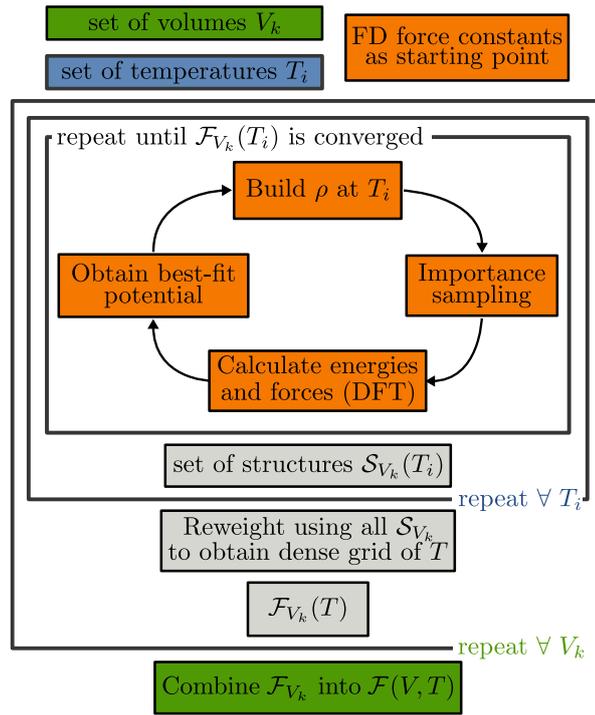
Figure 3.4: An illustration of the self-consistency workflow for effective harmonic potentials.

The latter improves the amount of available data and allows the use of unregularized fitting methods, i.e. those that do not reward or enforce (unphysical) sparseness of the force constants. In the presented articles the correction term, F_{corr} is obtained directly as a weighted average from the potential energies of the same samples used for determining the trial EHP

$$F_{\text{corr}} = \frac{1}{W} \sum_g \sum_m w_m^{(g \to k)} \left[E_{\text{BO}}(\mathbf{u}_m^{(g)}) - E_{\text{HA}}^{(k)}(\mathbf{u}_m^{(g)}) \right], \tag{3.35}$$

where W is the sum of all the weights. Furthermore, as described in the article below reweighting can be used to reuse samples drawn at temperature T_1 for a different temperature T_2, which can be used to obtain a dense grid of temperatures. Naturally, the extension to an effective *quasi* harmonic approximation, by performing the scheme at several volumes V_k, is possible as well, as indicated by the outermost box in Fig. 3.4.
The effective number of samples [96–99]

$$w_{\text{eff}}^{(k)} = \frac{\left[\sum_g \sum_m w_m^{(g \to k)} \right]^2}{\sum_g \sum_m \left[w_m^{(g \to k)} \right]^2}, \tag{3.36}$$

provides an additional guide to data completeness. With EHPs, temperature effects are considered not only in the frequencies and eigenvectors but also in the catch-all anharmonic correction term. Furthermore, it is possible to treat dynamically stabilized structures using Eq. (3.31), such as the fluorite phase of HfO_2, as the soft mode vanishes with temperature. This is shown in Fig. 3.5.
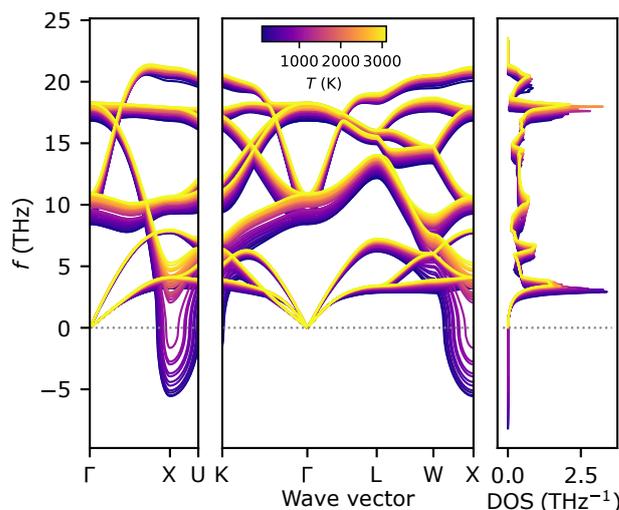
Figure 3.5: Stabilization of the soft mode at X exhibited by $Fm\bar{3}m$ HfO_2 at $V = 10.87\,\text{Å}^3/\text{atom}$ as temperature increases when using the EHP method. Adapted from [19].

3.3   Machine learning

Machine learning (ML) and "artificial intelligence" have become ubiquitous terms in business and science alike. From fault identification in semiconductor manufacturing [100], fraud detection and risk management in banking [101] to our application, the parametrization of the potential energy surface of various materials, there is hardly an area not benefitting from the recent surge in ML research. Regardless of the usecase, the archetypical design-process of an ML-based solution to a given problem can be broken down into a set of steps:

1. Identifying the problem type (e.g. regression vs. classification, supervised vs. unsupervised).

2. Choosing an algorithm or method and a corresponding framework to work with (tensorflow [102], scikit [103], PyTorch [104], JAX-based [14, 105], ...) and a corresponding problem-adapted and "computer-readable" representation of the data (a *descriptor*).

3. Generating or obtaining a *training*, a *validation* and a *test* set.

4. Obtaining model parameters using the *training set* and iteratively improving performance by tuning hyperparameters and model architecture based on the evaluation of the *validation* set.

5. Testing the performance and transferability on unseen *test* sets.

While this is a very abstract set of steps, it will be specified to the case of neural network force fields for computational chemistry in the following. It should be noted that the "recipe" provided above cannot and should not be stringently enforced, rather it will sometimes be necessary to backtrack and rethink previously-made decisions. For example, in our particular case, where new data can be generated as needed, the third and fourth step are combined into one.

3.3.1 Problem

ML applications in computational materials science can be roughly classified in three categories: (i) the direct prediction of hard-to-calculate physical quantities (e.g. thermal conductivity [106, 107]), based on several known or easy-to-calculate physical quantities (e.g. volume, bulk modulus, band gap, ...); (ii) the parametrization of the key DFT outputs (energy, forces) in terms of the Cartesian coordinates or quantities derived thereof [9, 108]; finally (iii) the full ML-based prediction of the electronic wavefunction [109, 110].

The work shown here is focused on obtaining a force field, i.e. achieving accurate and fast predictions of the energy and forces directly from the nuclear positions to facilitate application of the methods outlined in Section 3.4 and subsection 3.2.2, hence it belongs to the second category mentioned above. This naturally implies a supervised regression problem.

3.3.2 Framework

Since the implementation of a neural-network-based force field within our group was already underway, we decided to use it. To achieve a flexible and customizable implementation, while still relying on an efficient multi-purpose backend, Google's Deepmind JAX-ecosystem ([14, 15] and corresponding derived libraries (flax [105], optax [15]) were chosen. Using this ensures three key features:

- **Just-in-time compilation**: JIT drastically reduces the execution time of python code, making it competitive (and useable) for large-scale NNs.

- **Easy parallelization**: Regardless if the computing resources are CPUs, GPUs or TPUs, JAX and the XLA backend library are built to run efficiently and parallel on it, with no to minimal code adaptations required.

- **Automatic differentiability**: Adhering to a few coding practices further enables JAX code to be fully end-to-end differentiable. This is not only an obvious advantage when training NNs, but particularly when doing so for force fields as will become clear in the next subsection.

In the following a brief introduction to neural networks (NN) will be given, which we will then specify for the case of the neural network force field (NNFF).

Neural networks

Perhaps the simplest incarnation of an NN is the fully-connected NN, or perceptron. It consists of an input layer, several hidden layers, H_i, each with a number of so-called neurons and finally an output layer as shown in Fig. 3.6a. Each neuron receives all outputs of the previous layer as inputs, I, builds a weighted linear combination with weights, w, and a bias, b, and finally generates an output, O, by applying a nonlinear function, f, to this:

$$O = f(b + \sum_k w_k I_k). \tag{3.37}$$
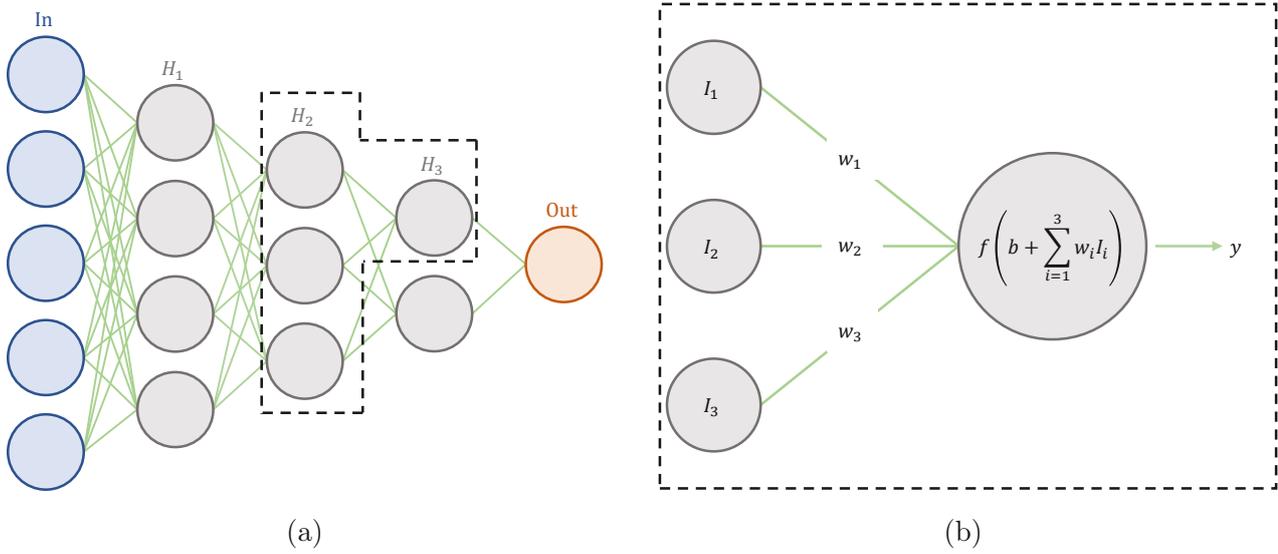
Figure 3.6: A schematic depiction of a simple fully-connected neural network (a) and the inner working of the outlined connections between H_2 and the first neuron of H_3 (b).

A schematic graphic is shown in Fig. 3.6b. The function f is a design choice, perhaps most commonly used is the rectified linear unit or ReLU [111]: $f(x) = \max(0, x)$. As can be easily seen, it is not smooth at $x = 0$, which would present an issue later on. Hence, an activation function found itself through an automated ML-based search, the Swish-1 activation function, [112, 113],

$$f(x) = \frac{x}{1 + e^{-x}},,$$ (3.38)

was chosen for our application. For some other examples, e.g. the tanh or sigmoid activation function, the gradient approaches 0 for large input values as is shown in the middle panel of 3.7. This results in a stalling training and is thus undesirable. The Swish-1 function is smooth and does not suffer from this so-called vanishing gradient problem. The successive application of non-linearities to linear combinations of input with a sufficient number of H_i has been shown to be a universal function approximator [114].

All weights and biases of all neurons, denoted as Ω, compose the set of trainable parameters. They are typically tuned by optimizing a so-called loss function through gradient descent or variations thereof. We denote the target value corresponding to a vector of input parameters $\mathbf{x}_i$ as y_i and the predicted output of the NN as $\tilde{y}_i$, where each entry i indicates a portion of the *training* set. With this, we can define a simple loss function, $\mathcal{L}$, based on the mean squared error (MSE):

$$\mathcal{L}_{\mathrm{MSE}}(\Omega) = \frac{1}{N_{\mathrm{train}}} \sum_{i=1}^{N_{\mathrm{train}}} (y_i - \tilde{y}_i)^2$$ (3.39)

The parameters are then adapted by building the gradient of Eq. (3.39) with respect to Ω and proceeding in the direction of $-\nabla_\Omega \mathcal{L}$ with a small stepsize (*learning rate*, LR). Repeating this for a number of iterations (*epochs*) would constitute the most primitive form of a NN training
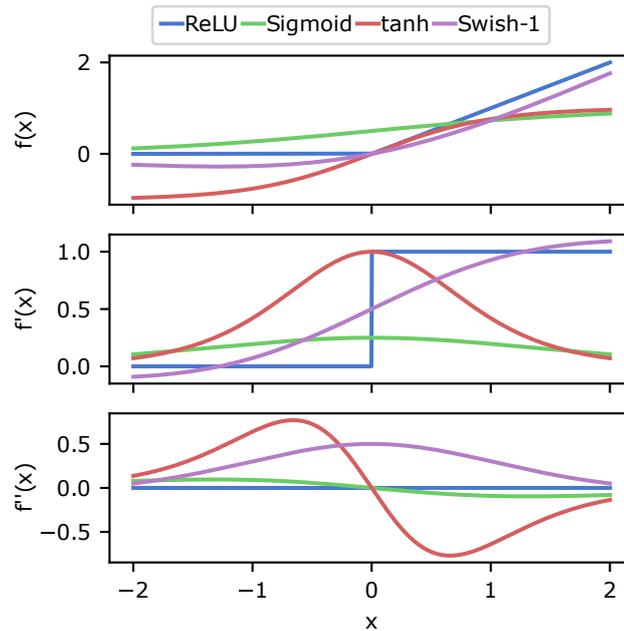
Figure 3.7: Several well-known activation functions and their first and second derivatives.

procedure. This simple picture is, however, only useful for illustrative purposes as it would be inefficient, unstable and slow.

In practice there are some improvements typically made when training a neural network:

- **Mini-batching**: Randomly splitting the training set into a number of mini-batches consisting of only a few samples. This makes mini-batching more memory efficient, decreases the risk of being stuck in a local minimum and in general provides a robust convergence behavior. We chose a batch-size of $N_{\text{batch}} = 4$.

- **ADAM optimizer**: The adaptive momentum-based optimizer [115] uses running averages and second momenta of the gradients to scale the loss gradients resulting in a faster, more stable training.

- **Normalization**: Normalization can substantially increase training speed of the network by centering and scaling inputs - either along batches (BatchNorm, [116]), or layers as we chose to do (LayerNorm [117]).

- **Learning rate scheduling**: A learning rate varying within an epoch has also been shown to significantly improve training time. The "schedule", i.e. the variation in the learning rate can be any of a number of functions, with the linear-onecycle schedule [118] being our choice. There, the LR is linearly increased for a fraction of the epoch, promoting exploration of the PES. Then it linearly decreases back to the original value, while finally decreasing even further improving exploitation.

- **Loss functions**: There is a host of loss functions available and the best choice depends on the problem at hand. For example, the MSE is by definition very sensitive to outliers
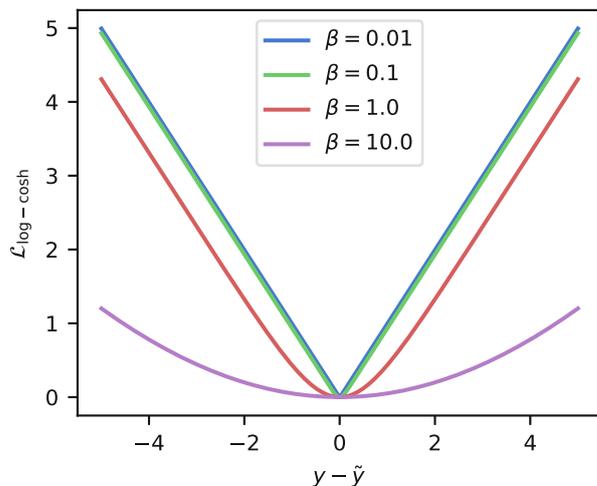
Figure 3.8: The log-cosh loss function as introduced in Eq. (3.40) for various values of β.

which can have detrimental impact on the overall model performance. The MAE on the other hand is not smooth. We thus use the log-cosh loss [119]:

$$\mathcal{L}_{\log-\cosh} = \frac{1}{N_{\text{batch}}} \sum_{i=1}^{N_{\text{batch}}} \beta \log \left[\cosh \left(\frac{y_i - \tilde{y}_i}{\beta} \right) \right], \tag{3.40}$$

which smoothly changes between the MSE and MAE for small and large values, respectively. This behaviour shown in Fig. 3.8 can be interpreted as a built-in gradient clipping, controlled by the hyperparameter β.

Neural Network Force Fields

To date one of the most fruitful ML methods to parameterize the PES as a function of the Cartesian coordinates are neural network force fields (NNFF, [120]). The approach Behler suggests is adapted throughout this work. The total energy of the system, E, is expressed as the sum over (not necessarily physical) individual contributions of each atom, e_k. Each of these contributions is obtained by passing a descriptor of the atom through the neural network - the sum then constitutes E.

A suitable choice of descriptor needs to take into account the symmetries and invariances inherent in the physics: translational, rotational and permutation invariance. In NeuralIL this is achieved through the recipe provided by Kocer et al. [121]. Starting from a local density within a pre-defined cutoff radius r_c for each chemical element, J, around each atom k

$$\rho_{kJ}(\mathbf{R}) = \sum_{j \in J; \mathbf{R}_{kj} < r_c; j \neq k} \delta(\mathbf{R} - \mathbf{R}_{kj}), \tag{3.41}$$

ensures translational invariance. This density is then projected onto an orthonormal basis set

$$c_{nlm;kJ} = \int_{\mathbb{R}^3} B^*(\mathbf{R} - \mathbf{R}_{kJ}) \rho_{kJ}(\mathbf{R}) d^3\mathbf{R}, \tag{3.42}$$
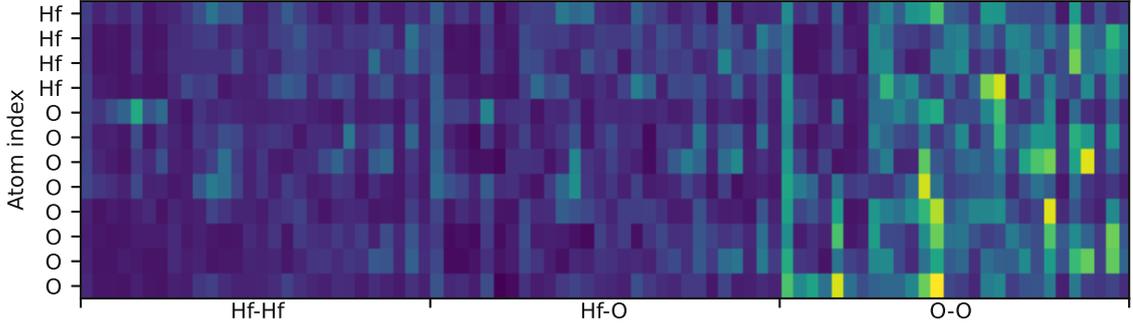
Figure 3.9: Descriptor of the 12-atom monoclinic HfO_2 unitcell.

where the $B_{nlm}(\mathbf{r}) = g_{n-l,l}(R)Y_l^m(\hat{\mathbf{R}})$ with the Y_l^m being the spherical harmonics and g_{nl} the radial basis functions as described in [121]. Subsequent construction of the power spectrum

$$\rho_{nl;kJJ'} = \sum_{m=-l}^{l} c_{nlm;kJ}c^*_{nlm;kJ'}, \tag{3.43}$$

ensures rotational invariance. The concatenation of the $\rho_{nl;kJJ'}$ for each J' in the set of chemical species, as well as an additional embedding vector encoding the species of atom k itself then constitutes the descriptor for one atom, $\mathcal{D}_k$. As an example we show the descriptors for the 12-atom monoclinic HfO_2 unit cell for $r_{\text{cut}} = 5\,\text{Å}$ and $n_{\text{max}} = 6$ in Fig. 3.9. Finally the summation of the atomic energy contributions described previously guarantees permutation invariance.

Thus the described framework, $\mathcal{E}$, accepts as input $(\mathbf{R}, \mathbf{J}, \mathbf{h})$, the positions and chemical types of all the atoms in the configurations, as well as the cell, constructs the descriptors $\mathcal{D}_k$ and passes them into a neural network, $\mathcal{N}(\cdot; \Omega)$ to output a prediction of the energy, $\tilde{E}$

$$\mathcal{E}(\mathbf{R}, \mathbf{J}, \mathbf{h}; \Omega) = \sum_{k=1}^{n_{\text{atom}}} \mathcal{N}(\mathcal{D}_k; \Omega) = \sum_{k=1}^{n_{\text{atom}}} e_k = \tilde{E}, \tag{3.44}$$

as depicted in Fig. 3.10. The corresponding log-cosh loss function, Eq. (3.40), can then be constructed as

$$\mathcal{L}_{\text{energy}} = \frac{1}{N_{\text{batch}}} \sum_{i=1}^{N_{\text{batch}}} \beta \log\left[\cosh\left(\frac{E_i - \mathcal{E}(\mathbf{R}_i, \mathbf{J}_i, \mathbf{h}_i; \Omega)}{\beta}\right)\right] \tag{3.45}$$

The advantage of a twice smoothly differentiable loss function such as Eq. (3.38) and a framework capable of automatic differentiation (AD), becomes clear when considering that the forces are the negative gradient of the energies with respect to the positions, or in terms of our NNFF

$$\nabla_{\mathbf{R}}\mathcal{E}(\mathbf{R}, \mathbf{J}, \mathbf{h}; \Omega) = -\tilde{\mathbf{f}}, \tag{3.46}$$

and the fact that one DFT calculation provides only a single energy, but $3 \times n_{\text{atoms}}$ forces, i.e. the amount of useable data is increased drastically when one trains on forces. With what amounts to essentially one line of PYTHON code, the energy-based loss can be formulated as a force-based
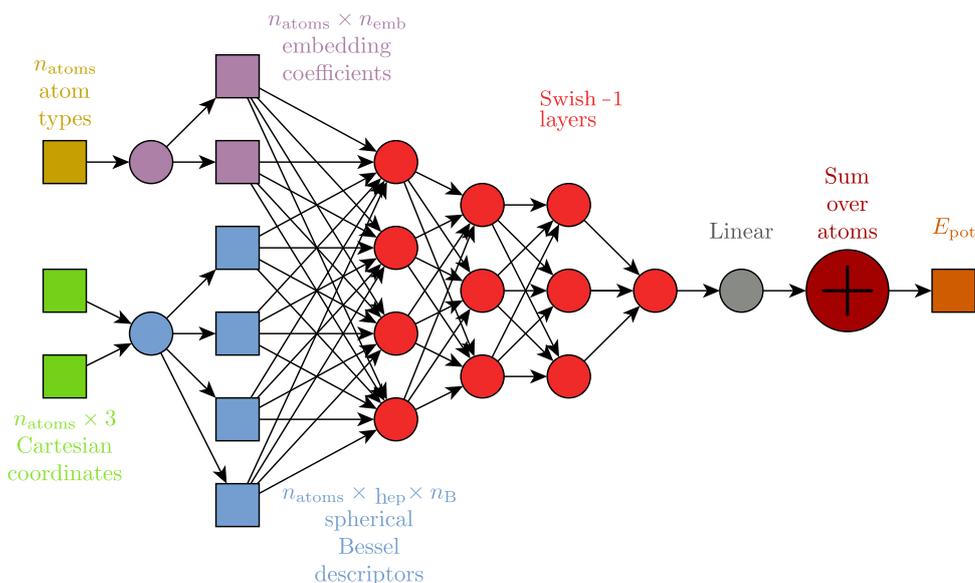
Figure 3.10: A schematic depiction of a full NeuralIL-type NNFF. Adapted from [16] with permission of the authors.

loss:

$$\mathcal{L}_{\text{forces}} = \frac{1}{N_{\text{batch}}} \sum_{i=1}^{N_{\text{batch}}} \frac{1}{n_{\text{atoms}}} \sum_{k=1}^{n_{\text{atoms}}} \beta \log \left[\cosh \left(\frac{\sqrt{\sum_{\alpha=1}^{3} \left(f_{ik\alpha} + \nabla_{\mathbf{R}} \mathcal{E}(\mathbf{R}, \mathbf{J}, \mathbf{h}; \Omega)_{ik\alpha} \right)^2}}{\beta} \right) \right] , \quad (3.47)$$

where the index i denotes the sample, the index k the atom and the index α the Cartesian coordinate, hence enabling force-based training. As the optimization is based on the gradient of the loss, $\nabla_{\Omega} \mathcal{L} \propto \nabla_{\Omega}(\nabla_{\mathbf{R}}(\mathcal{E}))$, an activation function that is at least twice continuously differentiable is required. In fact, the end-to-end differentiability can be exploited in numerous useful ways. It for example grants direct access to the stress tensor through

$$\boldsymbol{\sigma} = \frac{1}{V} \left. \frac{\partial \mathcal{E}(\mathbf{R}, \mathbf{J}, \mathbf{h}(\mathbf{I} + \epsilon))}{\partial \epsilon} \right|_{\epsilon=0} , \quad (3.48)$$

if $\mathcal{E}$ ensures correct transformation of the positions accounting for the changed cell. This is advantageous e.g. for evaluation of Eq. (3.78). Likewise, higher-order derivatives and even Taylor expansions can be obtained and their accuracy studied.

3.3.3   Data acquisition and training

In typical machine learning applications in industry, data acquisition and pre-processing come with several challenges. Data might be scant or originate from e.g. different measurement equipment, hence requiring more or less involved pre-processing. Further it might be difficult or impossible to acquire new data in a reasonable timespan. For a NNFF most of these hurdles can be eliminated already from the beginning and the amount of data is only limited by the
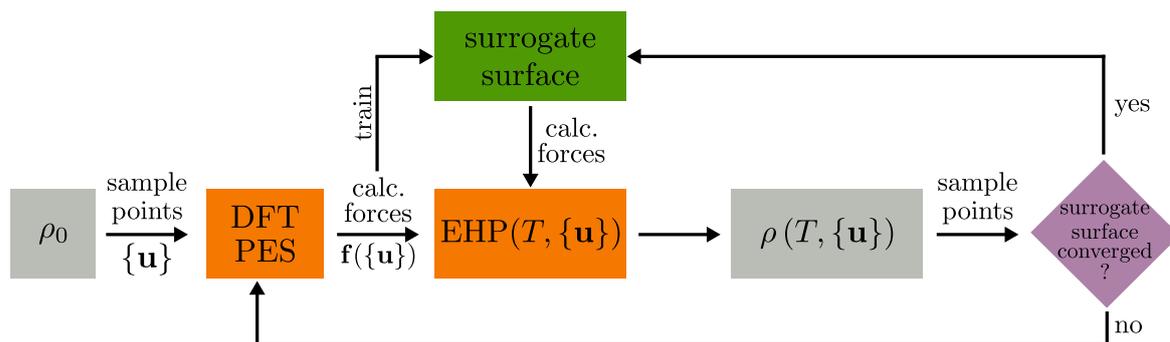
Figure 3.11: Flowchart describing the training of a surrogate model during EHP calculations.

resources allocatable to DFT calculations. However, just because it is possible, it does not necessarily mean that training a model on tens or even hundreds of thousands of configurations is desirable or efficient. In fact, arguably the less data required for satisfying accuracy, the better. More data still costs more, both in terms of time spent on DFT runs or training and, depending on the chosen architecture, might even have a detrimental impact on inference speed. All these factors would restrict scaling the model across different use cases. The most common approach of data acquisition for the purpose of building a NNFF is subsampling ab initio molcular dynamics (AIMD) trajectories. Subsampling is necessary as MD snapshots taken only a few fs apart are by definition strongly correlated, which would hinder training, but it also means that a significant portion of compute invested in obtaining the trajectory is not utilized for training.

In [19], a training strategy based on EHPs is proposed. Owing to the iterative nature of the method, the samples obtained during EHP calculations have an inherent hierarchy. Furthermore, the samples are uncorrelated. Starting with one iteration of $Fm\bar{3}m$ data, the amount of data is increased in an approach akin to active learning, until the test-set error converges. Subsequently, the surrogate model can be used to drive the EHP calculations. This *cascading* procedure is then repeated for the $P4_2nmc$ and finally the $P2_1c$ phase, thereby constructing a transferable surrogate model with a comparatively small number of DFT calculations. The framework is schematically shown as a flowchart in Fig. 3.11. Using it, we were able to obtain a parametrization with excellent accuracy for (at least) the three phases used in training and the additional $P\bar{4}3m$ cubic phase, requiring just ≈ 1500 structures.

A common way to visualize the performance of ML models are so-called parity plots, where the true values are plotted against the respective predictions. As such a 45° line indicates perfect agreement. An example of such a parity plot, showcasing the results of employing the described training strategy to HfO_2 is shown in Fig. 3.12. There, evaluations of the forces of an unseen *test set* consisting of 50 each of $Fm\bar{3}m$, $P4_2nmc$ and $P2_1c$, as well as an additional set of data points for the $P\bar{4}3m$ phase which is completely novel to the NNFF, are visualized. Indeed, the performance is comparable to other state-of-the-art NNFF for HfO_2, while requiring significantly less data [71].
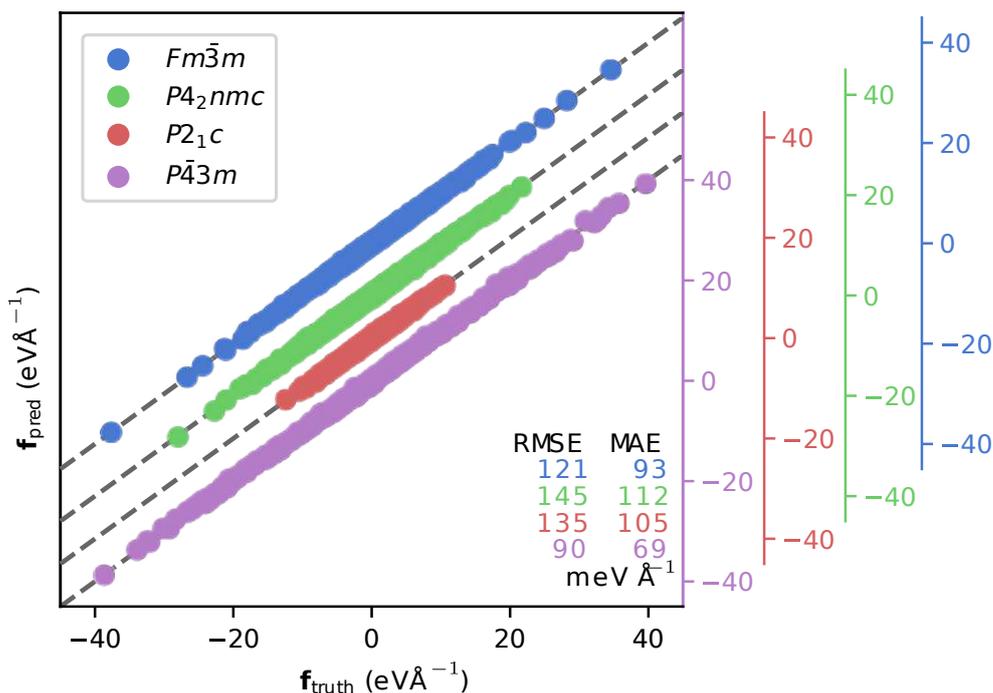
Figure 3.12: Parity plot and errors of the test set performance for the energies for the $Fm\bar{3}m$, $P\bar{4}3m$, $P2_1c$ and $P4_2nmc$ phases of HfO_2. Reproduced from [19].

3.3.4   Augmentations

Based on the training results and physical considerations we devised augmentations and modifications to the above-described procedure, several of which we introduce here.

Morse

Transferability typically refers to an ML model being applied to situations or problems it has not been explicitly trained on. As such, we have already shown a facet of transferability by achieving satisfactory performance on the low-symmetry cubic phase $P\bar{4}3m$. A different aspect is the performance of the NNFF for environments it has not previously seen. While under normal circumstances there is little that can be done to remedy this situation, aside from the generation of additional training data, in some cases physical intuition embedded into the architecture can alleviate the problem.

In particular, it is important that the NNFF properly reacts to two atoms coming into close proximity of one another - a situation typically resulting in large repulsive forces of several hundreds of eV Å^{-1} and high energies. While these events are usually rare, they do occur, for example when sampling from a high-dimensional Gaussian or doing MD simulations. It is undesirable to evaluate a sufficient number of those configurations, as their contribution to any final result is negligible. To test performance of the NNFF in these situations we constructed a training dataset where all configurations with (at least one) $|\mathbf{f}| > 50\,\text{eV}\,\text{Å}^{-1}$ were excluded and trained a NNFF on it. Evaluating its performance on the configurations exhibiting large forces, the predictions
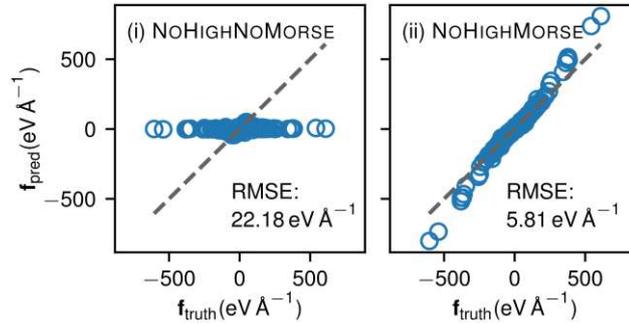
Figure 3.13: Comparison of NNFF performance on high-force configurations when not in- and excluding Eq. (3.49), in the left and right panel, respectively.

were found to be essentially $0\,\mathrm{eV\,\mathring{A}}^{-1}$ as indicated in the left panel in Fig. 3.13. This is of course unacceptable as it would result in an artifactual minimum state of bunched-together atoms and a resulting failure of the simulation. In contrast, the prediction of *any large force* irrespective of the accuracy will drive the system away from this situation thereby allowing the sampling process to continue and providing robustness in a sparsely sampled region. In the work presented here, this is achieved by including the repulsive part of a Morse potential [122], V_{rep} combined with a smooth bump function, h:

$$V_{\mathrm{rep}}(\mathbf{R}) = \frac{1}{2} \sum_{i \neq j} d_{ij} \times h(R_{ij}, R_s, R_c) e^{-2a_{ij}(R_{ij} - b_{ij})}$$

$$h(R, R_s, R_c) = 1 - g\left(\frac{R^2 - R_s^2}{R_c^2 - R_s^2}\right)$$

$$g(x) = \frac{f(x)}{f(x) + f(1-x)} \qquad (3.49)$$

$$f(x) = \begin{cases} 0 & x \leq 0 \\ \exp\left(-\frac{1}{x}\right) & x > 0 \end{cases},$$

where the a_{ij}, b_{ij} and c_{ij} are element-specific parameters, optimized during training. These are addressed by the corresponding atomic indices i and j. If $J_i \neq J_j$, i.e. the chemical species of the two atoms in question differ, the corresponding mixed parameters are obtained using the established mixing rules [123]. Using this, satisfactory performance for the high-forces is achieved even if they are not included in the training, as shown in the right panel of Fig. 3.13.

CENT

Owing to the fact that the original usecase for NeuralIL was ionic liquids [16], the question of the importance of long-range electrostatic interactions naturally arises. Even though the descriptors are purely local and can only describe what is within the respective cutoff, the performance was already very good.

Based on the charge equilibration via neural network technique (CENT, [124–126]), the total

energy was augmented by an additional term describing long-range electrostatic interactions

$$E_{\text{CENT}}(\mathbf{R}, \mathbf{J}, \mathbf{h}) = \sum_{i=1}^{n_{\text{atoms}}-1} \sum_{j=i+1}^{n_{\text{atoms}}} \frac{\text{erf}\left(\frac{R_{ij}}{\sqrt{2}\gamma_{ij}}\right)}{R_{ij}} Q_i Q_j + \sum_{i=1}^{n_{\text{atoms}}} \frac{Q_i^2}{2\sigma_i\sqrt{\pi}},$$ (3.50)

where the σ_i are taken to be the ionic radii as an approximation to the widths of the Gaussian charge density distributions and $\gamma_{ij} = \sqrt{\sigma_i^2 + \sigma_j^2}$. The charges themselves, Q_i are obtained from a *global* charge equilibration scheme, by minimizing

$$E_{\text{Q}} = E_{\text{CENT}} + \sum_{i=1}^{n_{\text{atoms}}} \chi(\mathcal{D}_i; \tilde{\Omega}) Q_i + \frac{1}{2} K_i Q_i^2,$$ (3.51)

where χ is another NN, conceptually identical to $\mathcal{N}$ but with a reduced neuron and layer count (16:16:16:1) described by the parameters $\tilde{\Omega}$. It is used to predict the electronegativity of the atom encoded by descriptor $\mathcal{D}_i$. The atomic hardnesses, K_i, are element-specific learnable parameters that - similarly to the a_{ij}, b_{ij}, d_{ij} for the Morse augmentation, are trained in conjunction with the parameters Ω and $\tilde{\Omega}$. Using a Lagrange multiplier charge conservation, $\sum_i Q_i = 0$, is ensured and the minimization problem Eq. (3.51) is solved using standard linear algebra routines. It is important to point out that, as opposed to [125], the $\tilde{\Omega}$ were not trained such that Eq. (3.51) leads to some partitioning of atomic charges, e.g. Hirshfeld charges [127], but rather the overall framework is redefined as

$$\mathcal{E} = \sum_{k=1}^{n_{\text{atom}}} \mathcal{N}(\mathcal{D}_k; \Omega) + E_{\text{CENT}},$$ (3.52)

i.e. both networks, $\mathcal{N}$ and χ were trained in one go to minimize the overall force loss introduced in Eq. (3.47). This has the additional benefit of only having to calculate $\mathcal{D}_i$ once. The automatic differentiability provided by JAX is key for a tractable implementation of this idea and the modular architecture of NeuralIL facilitated straightforward integration.

As opposed to what might be expected, the performance improvements were marginal at best ($\approx 8\,\%$ in force MAE). Similarly, the charges predicted by the scheme were unphysical with orders of magnitude in difference to the ionic charges that are typically considered for ionic liquids [128]. This points toward an efficient screening of long-range electrostatics, effectively allowing even the *local* descriptors to capture the most important contributions. Combined with the poor scaling behavior of such a global scheme, it was decided that the benefits of an inclusion of CENT do not outweigh the costs.

3.3.5   Alternative descriptors

While the choices described in the previous sections make up a capable architecture for a MLFF, they are still choices and other architectures are constantly being developed. In this final part of the background section, I want to point out some of those alternatives, briefly explain the fundamental concepts and provide references to the interested reader.

There is an abundance of choice regarding descriptors for representing atomic environments. I will present two archetypical approaches for atom-centered local atomic fingerprints, specifications and improvements of which constitute a majority of currently employed descriptors. For a more comprehensive overview, the reader is referred to [129, 130].

Atom-centered symmetry functions

In the original formulation [9] by Behler and Parinello (BP), these descriptors consist of a number of two- (G_i^1) and three-body (G_i^2) symmetry functions, which differ based on the choice of several parameters, $\eta, \zeta, R_s, \lambda$

$$G_i^1 = \sum_{j \neq i} e^{-\eta(R_{ij}-R_s)^2} f_c(R_{ij})$$

$$G_i^2 = 2^{1-\zeta} \sum_{j,k \neq i} (1 + \lambda \cos \theta_{ijk})^\zeta e^{-\eta\left(R_{ij}^2 + R_{ik}^2 + R_{jk}^2\right)} f_c(R_{ij}) f_c(R_{ik}) f_c(R_{jk}).$$

(3.53)

Here θ_{ijk} is the angle formed by atoms i, j and k, with i in the center and f_c is a cutoff function. A concatenation of a certain number of G_i^μ, for different values of the parameters, is then employed to construct the descriptor array $\mathcal{D}_i$. The precise formulation and notation has changed over the years [131], but the core concept and, with it, the key issue, has remained the same: The description cannot be systematically improved. Furthermore, the large number of "hyperparameters" results in a significant amount of physically informed manual fine tuning work. Nonetheless, impressive results can be achieved and the BP descriptors and flavors thereof are still widely used [132–135].

Smooth overlap of atomic positions

The second class of descriptors discussed here are the so-called smooth overlap of atomic positions (SOAP) descriptors. The descriptors used in NeuralIL and described earlier can be considered a subclass of these. The general idea is that an atomic density around an atom i is projected onto a (truncated) set of basis functions, each of which is typically split into an angular part described by the spherical harmonics, Y_{lm} [Eq. (3.42)], and a radial part g_n. Subsequently, the power spectrum [Eq. (3.43)], is constructed which constitutes the descriptor.

The differences between the different SOAP-like descriptors (e.g. [108, 121, 129, 136]) arise mainly from the radial functions used and the way the center atom's chemical species is included in the density. For example, in [129], the former are constructed as (orthonormalized) cubic and higher-order polynomials, while the latter is done through an explicit weighting of the densities, $\rho_{kJ} = \sum_{j \in J; j=k} w_{J_i} \delta(\mathbf{R} - \mathbf{R}_{kj})$.

3.3.6 Alternative architectures

Given the surge in ML-based computational chemistry studies[2], an exhaustive overview of methods lies beyond the scope of this section. Thus, I limit myself to the ones described below and refer the interested reader to available review articles e.g. [137, 138].

Gaussian approximation potentials

A widely used ML method in constructing force fields and arguably the most widespread alternative to NNs, is Gaussian process regression (GPR, or Kriging, [139]). The most well-known incarnation are Gaussian approximations potentials (GAP, [108, 140, 141]).

---

[2]A Web of Science query for (`machine learning AND density functional theory`) yields ≈ 500 results for 2022 alone

A Gaussian process is defined as a "*collection of random variables, any finite number of which have a joint Gaussian distribution*" [139], hence it is a generalization of a multidimensional Gaussian. It can thus be specified by a mean and any one of a number of covariance functions or kernels [103, 139], G, measuring the similarity of two inputs.

As opposed to NNFFs, GAPs does not require iterative training, it is an interpolation method where - in its simplest flavor - the coefficients can be obtained by matrix inversion. In the original formulation the truncated bispectrum of the projection of the density on the 4D-hyperspherical harmonics, $\tilde{\mathbf{b}}$, are used as descriptors. Using n_{train} tuples of descriptors and energies $(\mathbf{b}, E)$ as training data, Gaussian processes are used to interpolate the PES. The energy of an unknown structure is then obtained as [108]:

$$
\tilde{E}(\tilde{\mathbf{b}}_1, \ldots, \tilde{\mathbf{b}}_{n_{\text{atoms}}}) = \sum_{k}^{n_{\text{atoms}}} e_k = \sum_{k}^{n_{\text{atoms}}} \sum_{l}^{n_{\text{train}}} \alpha_l \exp\left[-\frac{1}{2} \sum_{m} \left(\frac{\tilde{b}_{km} - b_{lm}}{\theta_m} \right)^2 \right] =
$$
$$
= \sum_{k}^{n_{\text{atoms}}} \sum_{l}^{n_{\text{train}}} \alpha_l G(\tilde{\mathbf{b}}_k, \mathbf{b}_l),
\tag{3.54}
$$

where the index m runs over the components of the descriptors. The θ_l are hyperparameters and G is the squared-distance kernel, encoding a presumed knowledge about the covariance. The covariance matrix, $\mathbf{C}$, includes two further hyperparameters, δ and σ, and its matrix elements are given by:

$$
C_{mm'} = \delta^2 G(\mathbf{b}, \mathbf{b}') + \sigma^2 \mathbf{I}.
\tag{3.55}
$$

The interpolation parameters, $\boldsymbol{\alpha}$ are finally obtained as the product of the matrix inversion of $\mathbf{C}$ and the training data:

$$
\boldsymbol{\alpha} = \mathbf{C}^{-1}\mathbf{E}.
\tag{3.56}
$$

As the prediction is probabilistic, an uncertainty estimate is provided at no additional cost. Its usefulness, however, strongly depends on the appropriateness of the choice of G.

SchNet

Convolutional neural networks (CNN) have become the most important tool in computer vision applications. The name-giving convolutional layers are essentially filters, typically acting on discrete data points such as pixels. As such, it may seem surprising to find an application to a continuous dataset such as atomic positions.

With SchNet [142, 143], a continuous-filter convolutional layer was introduced, taking the $\mathbb{R}^{3n_{\text{atoms}}}$ positional vectors of the atoms to an $\mathbb{R}^F$-dimensional feature map X. This is achieved by expanding the distances between the atoms into radial basis functions (RBF), where the number of RBFs corresponds to the filter resolution in conventional CNNs. By passing the RBF representation through a number of dense layers, H, and combining the result with the feature representation X of the previous layer, SchNet (starting out from an embedding of the atomic charges in the first layer) essentially contains a "trainable" descriptor generator as part of its architecture.

NequIP

Neural equivariant interatomic potentials (NequIP, [144]) successfully combine a number of recent concepts, resulting in remarkable performance.

- It is built as a message passing neural network (MPNN, [145]). This type of NN is based on a graph representation of the atomic structure [146] where each atom i corresponds to a node and each node carries a state, $\mathbf{h}_i$. The edges of the graph are assigned a feature vector, $\mathbf{e}_{ij}$ as well - typically the distance between the two atoms i and j represented by the nodes it is connecting. A message $\mathbf{m}_i$ is then generated by a function M. Using an update function, U, the state $\tilde{\mathbf{h}}_i$ of the node i in the next layer, is obtained [144, 145, 147]

$$\mathbf{m}_i = \sum_{j; R_{ij} < r_c} M(\mathbf{h}_i, \mathbf{h}_j, \mathbf{e}_{ij}) \qquad \tilde{\mathbf{h}}_i = U(\mathbf{h}_i, \mathbf{m}_i), \tag{3.57}$$

  i.e. the messages are *passed* between the nodes.

- As the name suggests, it is equivariant, specifically with respect to E(3) [148]. This group comprises translations, rotations and inversions, the fundamental symmetries of free space. While, for example, the energy as a scalar is invariant with respect to rotations, the forces as vectors are not - they also need to rotate accordingly. Similarly all physical quantities, be it scalars, vectors or higher-order tensors, have well-defined transformation behaviors for the elements of E(3). An equivariant NN is constructed by ensuring only E(3)-equivariant operations are used. In contrast to invariant NNs, where only scalar inputs such as the power spectrum, Eq. (3.43), or the distance, Eq. (3.53), can be used, this enables including more information-rich quantities, such as relative displacements, $\mathbf{R}_i - \mathbf{R}_j$ [144].

- Finally, it also includes the idea of continuous-filter convolutions as introduced by SchNet. In NequIP, they consist of a radial part constructed from a NN acting on a radial expansion of the atomic distances and spherical harmonics acting on the corresponding unit vectors. This renders the convolution filters equivariant, as required.

While the results achieved using NequIP were best in-class across a variety of benchmarks, the message passing creates certain limitations. Every layer (i.e. every round of message passing) essentially increases the effective cutoff resulting in poor scalability and parallelizeability of the code. The authors of NequIP realized this and proceeded to introduce Allegro - an equivariant neural network based on local operations [147].

3.4 Molecular dynamics

In BO molecular dynamics [73] the interatomic forces are calculated quantum mechanically (typically through DFT), whereas the nuclei are treated as classical particles moving in the potential generated by the electrons. This means, the most simple ensemble, NVE, conserving the **N**umber of atoms, the **V**olume and the **E**nergy of the system, can be implemented by just solving Newton's

equations of motion (EOM):

$$\dot{x} = \frac{p}{M} \tag{3.58}$$

$$\dot{p} = f \tag{3.59}$$

using any symplectic integrator, such as velocity Verlet [149]

$$x(t + \Delta t) \approx x(t) + \dot{x}(t)\Delta t \tag{3.60}$$

$$\dot{x}(t + \Delta t) \approx \dot{x}(t) + \frac{\ddot{x}(t + \Delta t) + \ddot{x}(t)}{2}\Delta t, \tag{3.61}$$

with an error of $\mathcal{O}(\Delta t^2)$. After repeating this procedure for a number of timesteps Δt, a set of tuples of $(t, x(t), p(t))$, usually called a *trajectory*, is obtained. The goal of a typical MD simulation is to obtain a trajectory, and potentially additional quantities, by sampling the phase space according to a distribution, whose form is governed by the chosen ensemble. The NVE acts as a starting point for moving towards more involved systems, keeping e.g. the temperature (NVT) or the temperature and pressure (NPT) constant. Thus, a formal structure that enables systematic construction of numerical solvers will prove advantageous. One such formalisation is the Liouville-operator approach which in the following will be briefly explained on the example of the NVE ensemble and then extended to allow for NVT and finally NPT systems [150].

Liouville-operator approach: Microcanonical ensemble

Tuckerman et al. [150–153] have used the elegant Liouville-operator-based approach in their work. Here we will detail the methodology based on the simple NVE example, because it already contains the key insights. The following is hence adapted from [150], for an in-depth treatment please refer to it.

For any function $a(\mathbf{x}(t)) = a(q_i(t), \dots, q_n(t); p_i(t), \dots, p_n(t))$ acting on a phase space vector $\mathbf{x}(t)$ the time derivative in generalized coordinates $(p_i(t), q_i(t))$ is obtained through the chain rule

$$\frac{da}{dt} = \sum_i^n \left[\frac{\partial a}{\partial q_i}\dot{q}_i + \frac{\partial a}{\partial p_i}\dot{p}_i \right]. \tag{3.62}$$

Combined with Hamilton's EOM, $\dot{q}_i = \frac{\partial H}{\partial p_i}$ and $\dot{p}_i = -\frac{\partial H}{\partial q_i}$, we get the definition of the Poisson bracket, $\{\dots\}$

$$\frac{da}{dt} = \sum_i^n \left[\frac{\partial a}{\partial q_i}\frac{\partial H}{\partial p_i} - \frac{\partial a}{\partial p_i}\frac{\partial H}{\partial q_i} \right] = \{a, H\}. \tag{3.63}$$

With this, the Liouville operator, iL, can be defined

$$iLa = \{a, H\} = \frac{da}{dt}. \tag{3.64}$$

A formally exact solution of the equation above is given by $\mathbf{x}(t) = e^{iLt}\mathbf{x}_0$. For a numerical treatment, however, it is more convenient to split the operator described above into two contributions:

$$\underbrace{\sum_i^n \frac{\partial H}{\partial p_i}\frac{\partial}{\partial q_i}}_{iL_1} - \underbrace{\sum_i^n \frac{\partial H}{\partial q_i}\frac{\partial}{\partial p_i}}_{iL_2} = iL. \tag{3.65}$$

iL_1 and iL_2 do not commute, i.e. $[iL_1, iL_2] \neq 0$, hence $\exp{(iL_1 + iL_2)t} \neq \exp{(iL_1 t)}\exp{(iL_2 t)}$. Approximating the factorization using the Trotter therorem [154] yields an approximate time evolution operator

$$e^{iL\Delta t} \approx e^{iL_2\Delta t/2}e^{iL_1\Delta t}e^{iL_2\Delta t/2} + \mathcal{O}(3), \tag{3.66}$$

with an error proportionate to $\mathcal{O}(\Delta t^3)$ for a single step and an error $\mathcal{O}(\Delta t^2)$ for the whole trajectory. Considering now the special case of a particle of mass M, position x and momentum p whose one-dimensional motion is governed by a Hamiltonian $H = \frac{p^2}{2M} + U(x)$, we can explicitly write down Eq. (3.66)

$$e^{iL\Delta t} \approx e^{\Delta t/2 f(x)\frac{\partial}{\partial p}}e^{\Delta t \frac{p}{M}\frac{\partial}{\partial x}}e^{\Delta t/2 f(x)\frac{\partial}{\partial p}}, \tag{3.67}$$

where $f(x) = -\frac{dU}{dx}$. With the direct translation technique

$$\exp\left(c\frac{\partial}{\partial x}\right)g(x) = g(x + c), \tag{3.68}$$

as outlined in [151], the elegance of this approach will soon be clear. Application of the first "half" of iL_2, which acts on the momentum coordinate, onto the phase space vector $\mathbf{x}(t)$ yields:

$$e^{\Delta t/2 f(x)\frac{\partial}{\partial p}}\mathbf{x}(t) = e^{iL_2\Delta t/2}\begin{pmatrix} x \\ p \end{pmatrix} = \begin{pmatrix} x \\ p + \frac{\Delta t}{2}f(x) \end{pmatrix}.$$

The iL_1 acts on the position coordinate, yielding

$$e^{\Delta t \frac{p}{M}\frac{\partial}{\partial x}}\begin{pmatrix} x \\ p + \frac{\Delta t}{2}f(x) \end{pmatrix} = \begin{pmatrix} x + \Delta t\frac{p}{M} \\ p + \frac{\Delta t}{2}f(x + \Delta t\frac{p}{M}) \end{pmatrix}.$$

Finally, completing the momentum "translation", we arrive at

$$e^{\Delta t/2 f(x)\frac{\partial}{\partial p}}\begin{pmatrix} x + \Delta t\frac{p}{M} \\ p + \frac{\Delta t}{2}f(x + \Delta t\frac{p}{M}) \end{pmatrix} = \begin{pmatrix} x + \frac{\Delta t}{M}\left(p + \frac{\Delta t}{2}f(x)\right) \\ p + \frac{\Delta t}{2}f(x) + \frac{\Delta t}{2}f(x + \frac{\Delta t}{M}\left(p + \frac{\Delta t}{2}f(x)\right)). \end{pmatrix} \tag{3.69}$$

Hence

$$x(t + \Delta t) = x(t) + \Delta t v(t) + \frac{\Delta t^2}{2M}f(x(t)) \tag{3.70}$$

$$v(t + \Delta t) = v(t) + \frac{\Delta t}{2M}\left(f(x(t)) + f(x(t + \Delta t))\right), \tag{3.71}$$

which is precisely Eq. (3.61).

The beauty of this approach is that the problem is translated into simple, computer-interpretable instructions, which in the end recover the velocity-Verlet algorithm.

Canonical ensemble and Nosé-Hoover chains

The above-described formalism can be extended to describe the canonical (NVT) ensemble by connecting the system to an "external heat bath". Several methods exist to do so, from velocity rescaling, e.g. Berendsen thermostat [155], to the commonly-used Nosé-Hoover thermostat (NH, [156]). The NH belongs to the class of extended system thermostats, i.e. they treat temperature

through an additional heat bath variable, η, with its own mass-like parameter, Q, determining the coupling strength. It was shown, however, that the NH scheme is limited in its applications, as it fails to reproduce the proper distribution [152]. A possible solution in the form of Nosé-Hoover chains (NHC) was presented in [157]. Here, several heat bath variables, $\eta_1, \ldots, \eta_M$, with corresponding momenta, p_{η_i} and mass-like quantities Q_i are *chained* together. The first is linked to the particle momenta and ensures the correct average kinetic energy. The $i + 1$-th variable then ensures a proper canonical average for the i-th one, resulting finally in the following set of EOM for an N particle system of mass m_i in three dimensions [151, 153]:

$$\dot{\mathbf{R}}_i = \frac{\mathbf{p}_i}{m_i} \qquad \dot{\mathbf{p}}_i = \mathbf{f}_i - \frac{p_{\eta_1}}{Q_1}\mathbf{p}_i \qquad \dot{\eta}_k = \frac{p_{\eta_k}}{Q_k} \quad k = 1, \ldots, M$$

$$\dot{p}_{\eta_k} = G_k - \frac{p_{\eta_{k+1}}}{Q_{k+1}}p_{\eta_k} \qquad \dot{p}_{\eta_M} = G_M$$

$$G_1 = \sum_i^N \frac{\mathbf{p}_i^2}{m_i} - 3NkT \qquad G_k = \frac{p_{\eta_{k-1}}^2}{Q_{k-1}} - kT, \quad (3.72)$$

with the thermostat forces G_k. The corresponding Liouville operator obtains an additional term, iL_T

$$iL_T = \sum_{k=1}^M \left[\frac{p_{\eta_k}}{Q_k}\frac{\partial}{\partial \eta_k} + G_k\frac{\partial}{\partial p_{\eta_k}}\right] - \sum_{i=1}^N \frac{p_{\eta_1}}{Q_1}\mathbf{p}_i\frac{\partial}{\partial \mathbf{p}_i} - \sum_{k=1}^{M-1} \frac{p_{\eta_{k+1}}}{Q_{k+1}}p_{\eta_k}\frac{\partial}{\partial p_{\eta_k}}. \qquad (3.73)$$

To factorize this, Martyna and Tuckerman employ the Suzuki-Yoshida factorization scheme [158, 159] of order n_{sy} with a reduced timestep $\delta_j = \frac{w_j \Delta t}{n_c}$:

$$e^{i\frac{L_T \Delta t}{2}} = \prod_{i=1}^{n_c}\prod_{j=1}^{n_{\text{sy}}} e^{i\frac{L_t w_j \Delta t}{2n_c}} = \prod_{i=1}^{n_c}\prod_{j=1}^{n_{\text{sy}}} e^{i\frac{L_t \delta_j}{2}},$$

where the weights w_j are constructed to cancel out lower-order error terms conceptionally similar to e.g. the better-known Runge-Kutta methods [160, 161]. The resulting, now factorized, operator, then amounts to [151]

$$e^{i\frac{L_t \delta_j}{2}} = \overbrace{\exp\left[\frac{\delta_j}{4}G_M\frac{\partial}{\partial p_{\eta_M}}\right]}^{\mathcal{T}_1} \prod_{k=M-1}^1 \left\{\overbrace{\exp\left[-\frac{\delta_j}{8}\frac{p_{\eta_{k+1}}}{Q_{k+1}}p_{\eta_k}\frac{\partial}{\partial p_{\eta_k}}\right]}^{\mathcal{T}_2}\overbrace{\exp\left[\frac{\delta_j}{4}G_k\frac{\partial}{\partial p_{\eta_k}}\right]}^{\mathcal{T}_3}\right.$$

$$\times \exp\left[-\frac{\delta_j}{8}\frac{p_{\eta_{k+1}}}{Q_{k+1}}p_{\eta_k}\frac{\partial}{\partial p_{\eta_k}}\right]\Bigg\} \prod_{i=1}^N \overbrace{\exp\left[-\frac{\delta_j}{2}\frac{p_{\eta_1}}{Q_1}\mathbf{p}_i\frac{\partial}{\partial \mathbf{p}_i}\right]}^{\mathcal{T}_4} \qquad (3.74)$$

$$\times \prod_{k=1}^M \overbrace{\exp\left[-\frac{\delta}{2}\frac{p_{\eta_k}}{Q_k}\frac{\partial}{\partial \eta_k}\right]}^{\mathcal{T}_5} \prod_{k=1}^{M-1} \left\{\exp\left[-\frac{\delta_j}{8}\frac{p_{\eta_{k+1}}}{Q_{k+1}}p_{\eta_k}\frac{\partial}{\partial p_{\eta_k}}\right]\right.$$

$$\times \exp\left[\frac{\delta_j}{4}G_k\frac{\partial}{\partial p_{\eta_k}}\right]\exp\left[-\frac{\delta_j}{8}\frac{p_{\eta_{k+1}}}{Q_{k+1}}p_{\eta_k}\frac{\partial}{\partial p_{\eta_k}}\right]\Bigg\} \exp\left[\frac{\delta_j}{4}G_M\frac{\partial}{\partial p_{\eta_k}}\right].$$

While this result may seem intimidating at first, in reality there are only five unique operators, $\mathcal{T}_{1\ldots5}$. Of these, $\mathcal{T}_1$, $\mathcal{T}_3$ and $\mathcal{T}_5$ are simply translation operators as in Eq. (3.68). $\mathcal{T}_2$ and $\mathcal{T}_4$ are described as scaling operators, as their action is [150]:

$$\exp\left[cx\frac{\partial}{\partial x}\right]f(x) = f(xe^c). \tag{3.75}$$

Isobaric ensemble

Finally the Martyna-Tobias-Klein (MTK) equations [162] are considered. These describe an ensemble at constant particle number, pressure and temperature (NPT), i.e. allowing for volume and lattice constant fluctuations. In case of the former, yet another dynamical variable $\epsilon = \frac{1}{3}\ln\frac{V}{V_0}$, with a corresponding cell momentum, p_ϵ and mass W is introduced. With these, fluctuations in the internal pressure

$$P_{\text{int}} = \frac{1}{3V}\left[\sum_{i=1}^N \frac{\mathbf{p}_i^2}{m_i} + \sum_{i=1}^N \mathbf{R}_i \cdot \mathbf{f}_i - 3V\frac{\partial U}{\partial V}\right], \tag{3.76}$$

are controlled. Here $\mathbf{f}_i = -\nabla U$, are the forces on the atoms, resulting from the potential U. An additional NHC is coupled to this "volume" DOF to control the kinetic energy. An integrator for solving this system is presented in [153] and up to this point an implementation of the presented ensembles is given in JAX-MD [20]. However, isotropic volume fluctuations are only really suitable for fluid or gaseous systems, as in a solid, the cell shape is crucial for predicting materials properties and hence a non-isotropic variant is required.

In the course of this thesis, we implemented such a variant into JAX-MD. This was the natural choice as as interfaces well with our NNFF and can take full advantage of AD and JIT compilation. The required equations and a suitable integrator were presented by Yu et al. in [163] and amount to

$$\dot{\mathbf{R}}_i = \frac{\mathbf{p}_i}{m_i} + \frac{\mathbf{p}_g}{W_g}\mathbf{R}_i \qquad \dot{\mathbf{p}}_i = \mathbf{f}_i - \frac{\mathbf{p}_g}{W_g}\mathbf{p}_i - \frac{1}{N_f}\frac{\text{Tr}\left(\mathbf{p}_g\right)}{W_g}\mathbf{p}_i - \frac{p_{\eta_1}}{Q_1}\mathbf{p}_i$$

$$\dot{\mathbf{h}} = \frac{\mathbf{h}}{W_g}\mathbf{p}_g \qquad \dot{\mathbf{p}}_g = \det\left[\mathbf{h}\right]\left(\mathbf{P}^{(\text{int})} - \mathbf{I}P\right) + \frac{1}{N_f}\sum_{i=1}^N \frac{\mathbf{p}_i^2}{m_i}\mathbf{I} - \frac{p_{\zeta_1}}{Q_1'}\mathbf{p}_g$$

$$\dot{\eta}_k = \frac{p_{\eta_k}}{Q_k} \qquad \dot{p}_{\eta_k} = G_k - \frac{p_{\eta_{k+1}}}{Q_{k+1}}p_{\eta_k} \qquad \dot{p}_{\eta_M} = G_M$$

$$G_1 = \sum_i^N \frac{\mathbf{p}_i^2}{m_i} - 3NkT \qquad G_k = \frac{p_{\eta_{k-1}}^2}{Q_{k-1}} - kT \tag{3.77}$$

$$\dot{\zeta}_k = \frac{p_{\zeta_k}}{Q_k'} \qquad \dot{p}_{\zeta_k} = G_k' - \frac{p_{\zeta_{k+1}}}{Q_{k+1}'}p_{\zeta_k} \qquad \dot{p}_{\zeta_M} = G_M'$$

$$G_1' = \frac{\text{Tr}\left[\mathbf{p}_g^T\mathbf{p}_g\right]}{W_g} - d^2kT \qquad G_k' = \frac{p_{\zeta_{k-1}}}{Q_{k-1}'} - kT,$$

where P is the applied external pressure, d the dimensionality, $\mathbf{I}$ the $d \times d$ identity matrix, $\mathbf{h}$ the cell matrix, $\mathbf{p}_g$ and W_g the corresponding cell momentum and mass. The Q_k' are the masses of

the NHCs attached to the barostat. $\mathbf{P}^{\text{int}}$ is the pressure tensor whose elements are given by:

$$P_{\alpha\beta}^{(\text{int})}(\mathbf{p}, \mathbf{R}) = \frac{1}{\det(\mathbf{h})} \sum_{i=1}^{N} \left[\frac{(\mathbf{p}_i \cdot \hat{e}_\alpha)(\mathbf{p}_i \cdot \hat{e}_\beta)}{m_i} + (\mathbf{f}_i \cdot \hat{e}_\alpha)(\mathbf{R}_i \cdot \hat{e}_\beta) - \underbrace{\sum_{\gamma=1}^{d} \frac{\partial U}{\partial h_{\alpha\gamma}} h_{\gamma\beta}}_{=\boldsymbol{\sigma}}, \right] \tag{3.78}$$

with $\hat{e}_\alpha$ representing the unit vectors in the Cartesian direction α and $\boldsymbol{\sigma}$ representing the stress tensor. These EOMs result in yet another Liouville operator, $iL = iL_1 + iL_2 + iL_{g,1} + iL_{g,2} + iL_{T,\text{particles}} + iL_{T,\text{barostat}}$, with the terms

$$iL_1 = \sum_{i=1}^{N} \left(\frac{\mathbf{p}_i}{m_i} + \frac{\mathbf{p}_g}{W_g} \mathbf{R}_i \right) \nabla_{\mathbf{R}_i}$$

$$iL_2 = \sum_{i=1}^{N} \left[\mathbf{f}_i - \left(\frac{\mathbf{p}_g}{W_g} + \frac{1}{N_f} \frac{\text{Tr}[\mathbf{p}_g]}{W_g} \mathbf{I} \right) \mathbf{p}_i \right] \nabla_{\mathbf{p}_i} \tag{3.79}$$

$$iL_{g,1} = \frac{\mathbf{p}_g \mathbf{h}}{W_g} \nabla_{\mathbf{h}}$$

$$iL_{g,2} = \mathbf{G}_g \nabla_{\mathbf{p}_g},$$

where the "box forces" are given by

$$\mathbf{G}_g = \det[\mathbf{h}] \left(\mathbf{P}^{\text{int}} - \mathbf{I}P \right) + \frac{1}{N_f} \sum_{i=1}^{N} \frac{\mathbf{p}_i^2}{m_i} \mathbf{I}. \tag{3.80}$$

The particle and barostat thermostat, $iL_{T,\text{particles}}$ and $iL_{T,\text{barostat}}$ are again solved by using the factorization applied in Eq. (3.74). For iL_1 and iL_2 a matrix-vector differential equation [163] needs to be solved by decoupling the Cartesian coordinates of $\mathbf{v}_g = \frac{\mathbf{p}_g}{W_g}$ through an orthogonal transformation, $\mathbf{O}$

$$\boldsymbol{\lambda} = \mathbf{O}^T \mathbf{v}_g \mathbf{O}, \tag{3.81}$$

with $\boldsymbol{\lambda}$ being a diagonal matrix consisting of the eigenvalues, λ_α, of $\mathbf{v}_g$. At the same time a symmetric $\mathbf{P}^{\text{int}}$ needs to be ensured, in order to prevent cell rotations, which effectively reduces the DOF from d^2 to $d^2 - d$. Using this and four auxiliary quantities

$$D_{\alpha\beta} = e^{\lambda_\alpha \Delta t} \delta_{\alpha\beta}$$

$$\tilde{D}_{\alpha\beta} = e^{\lambda_\alpha \Delta t/2} \frac{\sinh(\lambda_\alpha \Delta t/2)}{\lambda_\alpha \Delta t/2} \delta_{\alpha\beta}$$

$$\Delta_{\alpha\beta} = \exp\left[-\left(\lambda_\alpha + \frac{1}{N_f} \text{Tr}[\mathbf{v}_g] \right) \frac{\Delta t}{2} \right] \delta_{\alpha\beta} \tag{3.82}$$

$$\tilde{\Delta}_{\alpha\beta} = \exp\left[-\left(\lambda + \frac{1}{N_f} \text{Tr}[\mathbf{v}_g] \right) \frac{\Delta t}{4} \right] \frac{\sinh\left[\left(\lambda_\alpha + \frac{1}{N_f} \text{Tr}[\mathbf{v}_g] \right) \frac{\Delta t}{4} \right]}{\left(\lambda_\alpha + \frac{1}{N_f} \text{Tr}[\mathbf{v}_g] \right) \frac{\Delta t}{4}} \delta_{\alpha\beta},$$

the effects of iL_1 and iL_2 can be expressed as [163]

$$\mathbf{R}_i(\Delta t) = \mathbf{R}_i(0)\mathbf{O}\mathbf{D}\mathbf{O}^T + \Delta t \mathbf{v}_i \mathbf{O}\tilde{\mathbf{D}}\mathbf{O}^T$$

$$\mathbf{v}_i(\Delta t/2) = \mathbf{v}_i(0)\mathbf{O}\boldsymbol{\Delta}\mathbf{O}^T + \frac{\Delta t}{2}\mathbf{f}_i\mathbf{O}\tilde{\boldsymbol{\Delta}}\mathbf{O}^T$$

(3.83)

and with that the solution of the flexible isotropic ensemble is complete.
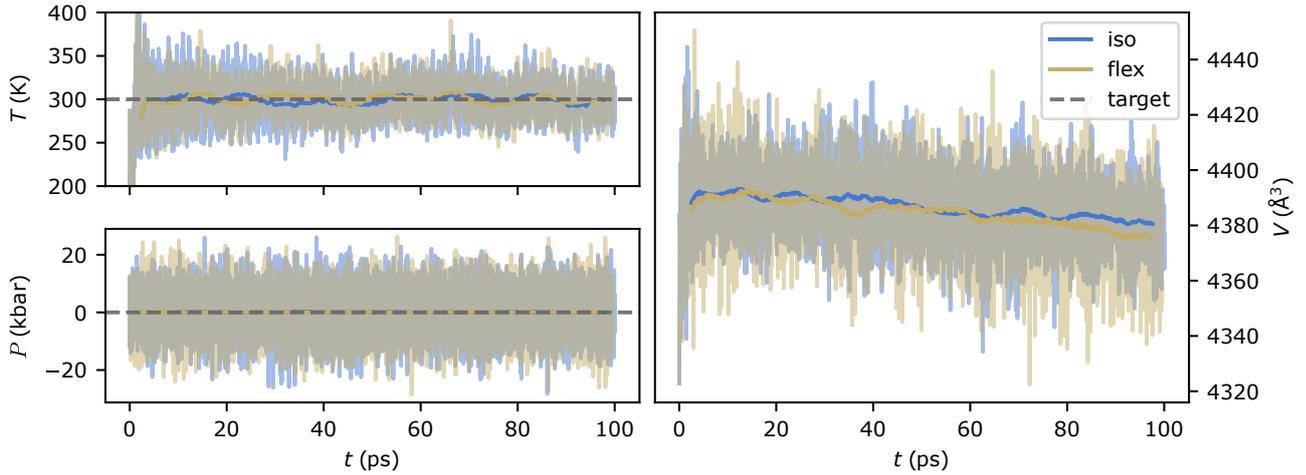
Figure 3.14: Temperature, pressure and volume trajectory of a 100 ps 216-atom Stillinger-Weber molecular dynamics run at 300 K as obtained with isotropic and flexible cell fluctuations.

In order to validate our implementation we chose a simple test system consisting of 216 Si atoms in the diamond structure, with the standard parametrization of the Stillinger-Weber potential as introduced in [164] and implemented in [20]. To ensure a representative sampling of the partition function, we choose a temperature of 300 K and pressure of 1 bar. In accordance with literature, we use a thermostat and barostat equilibration time of $\tau_T = 250$ fs and $\tau_p = 1000$ fs respectively, with a timestep of $\Delta t = 1$ fs for a total of 100 ps. We show the resulting trajectories comparing isotropic and flexible cell fluctuations in Fig. 3.14. The thermostat and barostat sufficiently control the temperature and pressure as indicated by the two panels on the left side. Similarly, the volume evolution which is shown on the right hand side of Fig. 3.14 is essentially indistinguishable between the two barostats and, perhaps even more importantly, the volume fluctuations are comparable as well.

Chapter 4

Summaries of publications
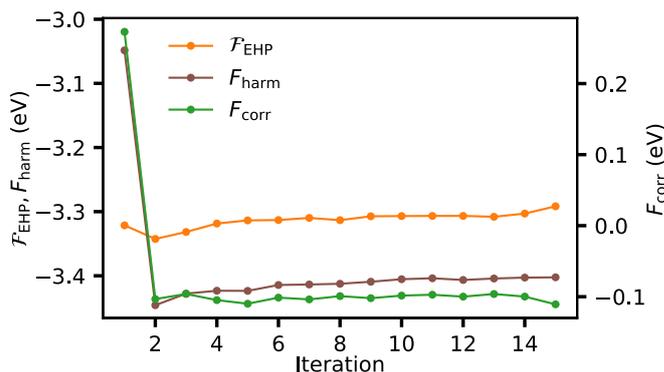
Figure 4.1: Convergence of F_{harm} and F_{corr} with iterations and samples.

4.1 Accurate first-principles treatment of the high-temperature cubic phase of hafnia

In this manuscript, we study the $Fm\bar{3}m$ phase of HfO_2 using EHPs. This phase of hafnia is unstable at $0\,\text{K}$, it is stabilized only through temperature effects (see Fig. 3.5) which in turn makes a phonon-based evaluation of the free energy possible. As this work is based on DFT calculations, data it costly and - particularly in the beginning - sparse. To overcome this limitation we evaluate two approaches:

- **Regularized regression methods**: Specifically, we include a L_1 regularization by employing the LASSO fitting procedure. It is found that convergence is not aided by this penalty. In fact by the time the free energy converges, the amount of data has increased to the point where a determination of the optimal regularization parameter through 5-fold cross validation results in $\alpha = 0$. We thus abandon this idea.

- **Reweighting procedure**: As described earlier, in Eq. (3.34), knowledge of the probability density at the time of sampling allows reusing samples from previous iterations in the current one. Additionally it is possible to use previously calculated samples at T_1 to obtain the EHP at T_2.

The reweighting procedure leads to a significant increase in data utilization and efficiency and based on the effective sample size, Eq. (3.36), we can determine when additional samples are required. We obtain the correction term introduced in Eq. (3.35) directly as a weighted average and in Fig. 4.1 we show that its convergence happens on a similar timescale as that of F_{harm}. Having obtained EHPs for several volumes of the $Fm\bar{3}m$ phase, a *quasi*-effective harmonic approximation is applied using a generalized form of the stabilized jellium equation of state. With this temperature-dependent physical properties are obtained that are in agreement with existing experimental data: (i) the thermal expansion coefficient and (ii) the bulk modulus. However, we *under*estimated the volume as compared to experiment which is atypical for a PBE based treatment.
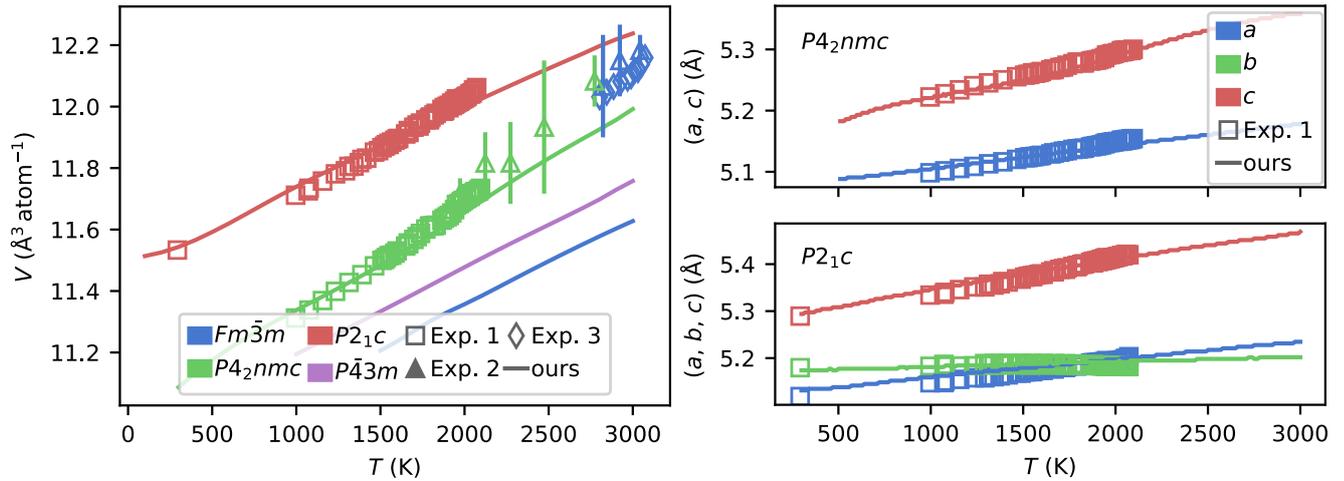
Figure 4.2: Volumetric thermal expansion of the four studied phases (left) and lattice thermal expansion of $P4_2nmc$ and $P2_1c$ (right) as compared to experimental data from Haggerty et al. (Exp. 1, [39]), Tobase et al. (Exp. 2, [42]) and Hong et al. (Exp. 3, [57]).

4.2 Neural-network-backed effective harmonic potential study of the ambient pressure phases of hafnia

As the first version of NeuralIL became available, just as the previous publication was finalized, a natural next step was to train a NNFF using the data created during the DFT-backed study. This enabled us to investigate not only the high-symmetry cubic phase, but also the other ambient pressure phases of hafnia - the lower symmetry monoclinic and tetragonal one. To properly and systematically build a potential transferable between those phases, we devised a framework integrating the EHP-sampling procedure with the neural network training (see Fig. 3.11). Data is added up to the point where the error is converged, from there on, the NNFF is used as a surrogate.

To increase the robustness of the neural network, particularly in the situation of two atoms getting close to each other, we augment its functionality with the repulsive part of a Morse potential and in Fig. 3.13 we show this augmentation has the desired effect. PBE typically overestimates lattice parameters, to calibrate for this, we extrapolate experimental results of [39] from room temperature to $0\,\mathrm{K}$ and extract an artificial isotropic pressure of $p_a = 4\,\mathrm{GPa}$. Application of this pressure results in excellent agreement of volume and lattice thermal expansion for monoclinic and tetragonal hafnia over a wide temperature range, as is shown in Fig. 4.2. At an additional pressure of $\approx 4\,\mathrm{GPa}$, we find a transition from monoclinic to tetragonal at $2600\,\mathrm{K}$.

Naturally the artificial pressure made the volume underestimation of the $Fm\bar{3}m$ phase even worse. This led to the inclusion of the lesser-known, lower-symmetry cubic $P\bar{4}3m$ phase into the study. While $P\bar{4}3m$ hafnia exhibits a larger volume, it is still lower than experiment suggests. Neither of the cubic phases is energetically favorable over the tetragonal phase, regardless of the temperature or the applied pressure. Even more so, both cubic phases seem to become *less* favorable with increasing temperature. We thus hypothesize that stoichometric cubic hafnia only exists in a defect-stabilized form.
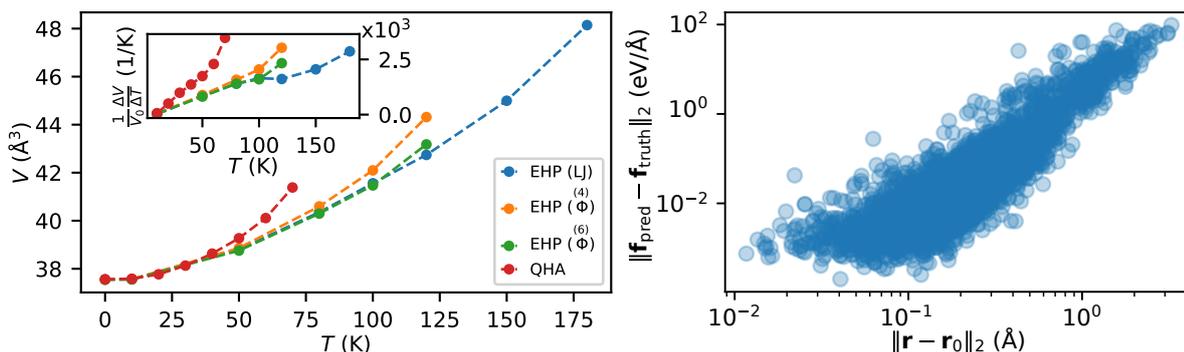
Figure 4.3: Left: Thermal expansion behavior of the fcc LJ solid as obtained from the QHA and EHPs obtained from $\overset{(4)}{\Phi}$ and $\overset{(6)}{\Phi}$ Taylor expansions as well as the exact potential. Right: Error of the force prediction of a $\overset{(4)}{\Phi}$ Taylor expansion of the silver cluster PES over the mean-square displacement.

4.3   Evaluating the efficiency of power-series expansions as model potentials for finite-temperature atomistic calculations

Due to the cost of studying temperature-dependent behavior, the use of inexpensive surrogate models is widespread. Many of these models rely on a local power-series expansion around the equilibrium. Equipped with automatic differentiation tools, we select three simple models - the Lennard-Jones cluster and fcc-solid and the five-atom silver cluster - as test benches to evaluate the efficiency of power-series models. To obtain a differentiable potential for the silver cluster a NNFF was parameterized. Based on the excellent agreement with DFT results, it is considered to be the ground truth. As such we arrive at three differentiable PES expressions, $V(\mathbf{R})$. Jacobian-vector product operators, allow evaluation of a function on an arbitrary input point $\mathbf{R}$ and tangent vector $\mathbf{t}$ without requiring the full Jacobian. With these, it is possible to recursively define a Taylor series operator, $\mathcal{T}_k$, and with it the order-n Taylor expansion, $\overset{(n)}{\Phi}$,

$$\mathcal{T}_k(\mathbf{R}, \mathbf{t}) = \frac{1}{k} \frac{\partial \mathcal{T}_{k-1}}{\partial \mathbf{R}} \mathbf{t} \qquad \overset{(n)}{\Phi}(\mathbf{R}) = \sum_{k=0}^{n} \mathcal{T}_k(\mathbf{R}_0, \Delta\mathbf{R}), \tag{4.1}$$

where $\mathcal{T}_0(\mathbf{R}, \mathbf{t}) = V(\mathbf{R})$, $\mathbf{R}_0$ are the equilibrium positions and $\Delta\mathbf{R} = \mathbf{R} - \mathbf{R}_0$. The performance of these expansions is then compared to the original potential for EHP generation. As polynomials can tend only towards $\pm\infty$ for large values of their argument, we see a significant decline in quality with increase in temperature, i.e. displacement distance, as shown e.g. in the left of Fig. 4.3 for the thermal expansion of the LJ solid. The root cause of this is shown in the right-hand-side of Fig. 4.3 on the example of the silver cluster: The error in the force obtained from a Taylor expansion increases by almost four order of magnitude, for one order in magnitude of displacement increase. Global interpolation strategies, such as NNFFs, thus are a much better alternative.

Chapter 5

Neural-network-backed molecular dynamics study of hafnia

Introduction

Effective harmonic potentials (EHP) are an efficient tool for studying the thermal expansion behavior of a single phase. In our previous work [18, 19], we have studied the ambient phase structure of HfO_2 in detail using EHPs. We have shown that application of an artificial pressure of $p_a = 4\,GPa$ can compensate the overestimation in lattice constants typically experienced by PBE [165] and results in excellent agreement of the anisotropic monoclinic and tetragonal lattice constants with experiment [39]. Furthermore, we observe a transition from monoclinic to tetragonal under the application of an additional moderate pressure ($p = 4\,GPa$), at $2600\,K$, a temperature comparable to literature results. However, as opposed to some experiments [42] or ab initio molecular dynamics [46], we were not able to observe a transition into a cubic phase, regardless of the temperature and pressure applied. Moreover, it was found that regardless of temperature or pressure, the commonly assumed cubic $Fm\bar{3}m$ phase is energetically less favorable than the lower-symmetry cubic $P\bar{4}3m$.

Extending on the study from [19] and using the neural network force field (NNFF) developed therein as a starting point, we decide to perform molecular dynamics (MD) simulations in the work presented here. In contrast to EHPs, MD simulations offer the possibility of studying dynamical phenomena and metastability. Furthermore, while EHPs require knowledge of the precise crystal structure, this restriction is lifted in MD. As such, MD offers much more flexibility to study phase transitions. Previously, however, MD studies for HfO_2 were limited. Existing classical force field parametrizations of hafnia are sparse [166–169], with the few existing being inaccurate, or tweaked towards the amorphous phase [169]. The computational demand of ab initio MD on the other hand enacts severe restrictions on system size and timescale [46]. Recently, a first machine-learning-based investigation of HfO_2 was enabled by DeepMD [71]. The NNFF developed in it was trained on over $21\,000$ DFT calculations, hence its construction requires a significant investment of resources.

Here, we describe a NNFF parametrization of HfO_2, obtained using a NeuralIL-type NNFF [16] and requiring less than 2000 DFT evaluations for construction. As NeuralIL is built on Google's highly efficient automatic differentiation framework JAX [15], it integrates exceedingly
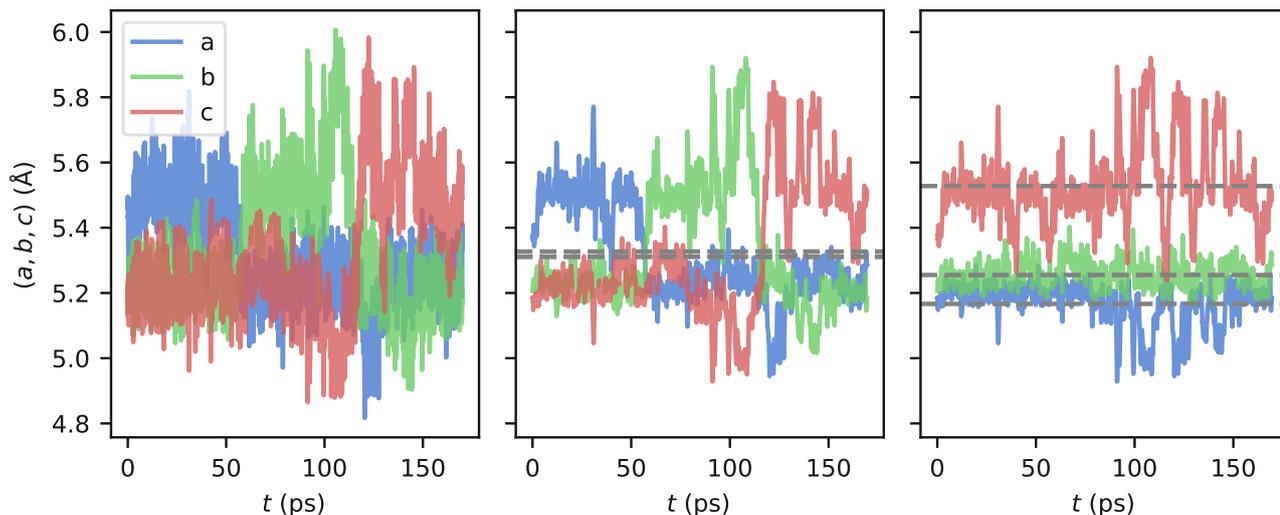
Figure 5.1: Analysis procedure of the trajectory of HfO$_2$ lattice parameters at 2000 K: Starting from the raw data (left), a moving average is applied (middle) before sorting the lattices at each timestep before obtaining the averages (right).

well with JAX-MD [20], a MD tool also built using JAX, and thus allows deployment of the code seamlessly on GPUs. We furthermore extend the NPT formalism implemented in JAX, which was previously limited to isotropic cell fluctuations, to include flexible cell shape changes, hence allowing the direct study of phase transitions.

Method

We reviewed the methodology of flexible-cell NpT MD in Section 3.4. Based on the implementation of isotropic cell fluctuations already covered in JAX-MD [20], we extend the code to include flexible cell changes as well. The implementation is validated by comparison to the isotropic case in Fig. 3.14. The NNFF was parameterized using the data obtained and methodology described in [19] (see Section 4.2), i.e. a cutoff radius of $r_c = 5\,\text{Å}$, an embedding dimension of 4 and a total of 128 basis functions. The architecture consists of 128 : 64 : 32 : 16 : 16 neurons.

To study HfO$_2$ we successfully ran flexible-cell MD calculations for a 96-atom supercell at temperatures ranging from 500 K to 2500 K at $P = 1$ bar, with a timestep of $\Delta t = 0.1$ fs and coupling constants of $\tau_T = 50$ fs and $\tau_p = 500$ fs used to construct the mass-like quantities, Q_k, Q'_k and W_g for the thermostat and barostat respectively. Of the 175 ps trajectory obtained for each temperature, the first 5 ps were considered relaxation time.

Preliminary results

In the left panel of Fig. 5.1, we show a trajectory as obtained using the above-described procedure for 2000 K. When averaging naively over the trajectory of cell parameters, as is indicated by the dashed grey lines in the middle panel of Fig. 5.1, one might conclude that the material is indeed

presenting a cubic phase. However, as can be clearly seen in Fig. 5.1, this is only an artifact as the elongated side is oscillating between a, b and c, resulting in essentially equal averages for the three sides. Hence, we apply a moving average of 500 fs to limit the impact of instantaneous oscillations, before sorting the thus obtained cell parameters at each timestep. From these sorted values we then obtain the averages as indicated for 2000 K by the dashed grey lines in Fig. 5.1 and shown for all the studied temperatures in Fig. 5.2.

From the bottom panel in Fig. 5.2, we can clearly see a transition from monoclinic to a cell consisting only of 90° angles, i.e. a collapse of β, that takes place from 1500 K to 2000 K. While the top panel could be seen as an indication for an orthorhombic phase, the fluctuations are quite large (see Fig. 5.1) and the supercell size is relatively small, which is known to lead to an extended phase transition window. More in-depth studies of the atomic arrangements and with a larger supercell could shed light on whether the material is indeed exhibiting an orthorhombic phase, or is approaching tetragonality as suggested by the EHP studies and experiment. The results shown here were obtained at ambient (i.e. 1 bar) pressure, which as already discussed in Section 4.2 results in a PBE-typical overestimation of the lattice parameters with regards to experiment. With the additional flexibility provided by MD it cannot be assumed that the application of pressure *just* results in a compression of the lengths and it does not impact e.g. the relative lengths too. In general, however, the agreement between MD and EHP can be interpreted as reasonable, especially if the unpredictable behavior during a phase transition is taken into account.

We do want to note, that these averages depend on a choice of order - had we not sorted the lattice parameters, the data point at 2000 K would suggest a cubic material. To better understand the behavior of the structure, a metric is required that does not depend on e.g. the identification of the "long" axis. We thus introduce a three-dimensional aspect ratio, $\aleph$, defined as

$$\aleph = \frac{p}{p_{\mathrm{cub}}} \qquad p = \frac{V_i}{V_e}, \tag{5.1}$$

where V_i (V_e) is the volume of the largest (smallest) sphere that can be inscribed (exscribed) in the cell (i.e. $p_{\mathrm{cub}} = \frac{1}{\sqrt{3}}$) and normalize it to a cubic phase. Thus $\aleph = 1$ if the structure is cubic. We show the result in Fig. 5.3 and judging from this metric, we can, on the one hand, confirm a phase transition taking place at 2000 K and, on the other hand, find an indication that the material seems to be indeed approaching a cubic phase. It is, however, unclear, if cubicity can be reached before melting.

Trying to go to temperatures higher than 2500 K originally led to failures in the MD run. At 2700 K, we experienced clustering of atoms of the same type, an environment not present in training at all. After these clusters started appearing, we saw a rapid expansion of the cell and a subsequent failure. To eliminate this behavior we augment the training set by randomly sampling an additional 200 structures from the database already curated, flipping the atomic species of each atom with a probability of 50 %, obtaining the respective energies and forces through DFT and subsequently retraining the NNFF. We then repeat the simulation at 2700 K using the new potential and an enlarged supercell of $n_{\mathrm{atoms}} = 768$. No clustering is observed and the results are shown in Figs. 5.2 and 5.3 in the area shaded in grey. Linearly extrapolating the four last data points of $\aleph$ would suggest a cubic transition at approximately 5800 K a point at which the material would have surely melted already.
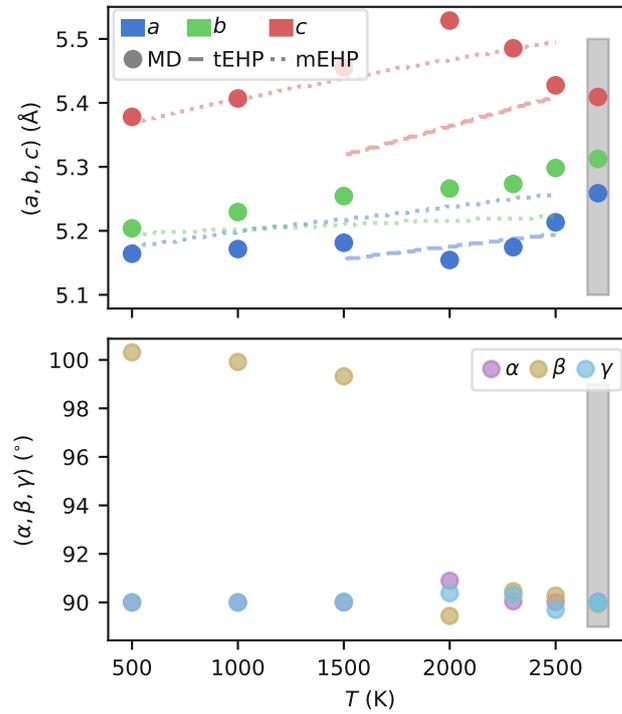
Figure 5.2: Lattice parameters as a function of temperature as obtained using MD (markers) compared to EHP results of the monoclinic (dotted lines) and tetragonal (dashed lines) phase. In the bottom panel the cell angles from MD are shown. The shaded area indicates a different potential and supercell size used (see text).
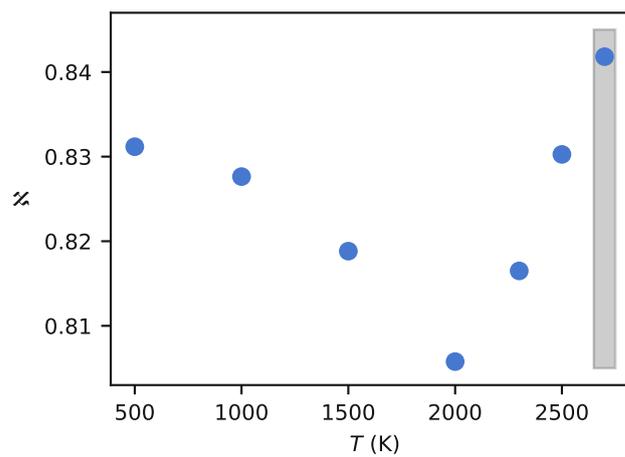
Figure 5.3: The 3D aspect ratio, $\aleph$, as a function of temperature. The shaded area indicates a different potential and supercell size used (see text).

Outlook

When extracting physical statements from the trajectories, it is key to assert when the NNFF is interpolating properly and when it is outputting nonsensical results due to a lack of training in an area. Introduction of an uncertainty metric, e.g. a committee-based approach, could alleviate situations similar to the observed clustering and provide guidance on the network's performance. Likewise, the impact of larger supercells on the temperature-dependent behavior and phase transitions temperature windows, as well as the application of an artificial pressure p_{corr}, as was done in the EHP-based manuscript in Section 4.2, need to be assessed. Furthermore, MD grants access to a rich selection of physical quantities through fluctuation formulas and Green-Kubo relations which are worth exploring. While promising, this work would go far beyond the scope of this thesis: We do, however, plan on targeting these goals and extending on the preliminary work shown here in the near future.

Chapter 6

Conclusion and Outlook

Conclusion

In this thesis methods to capture temperature-dependent behaviors of materials were developed, and employed on HfO_2. Based on an effective harmonic potential treatment in combination with a reweighting scheme, the temperature-aided stabilization of the soft mode in $Fm\bar{3}m$ hafnia was shown using density functional theory. Furthermore, the convergence behavior of the anharmonic correction term, as obtained by importance sampling, was found to be equivalent to the harmonic free energy contribution.

Building on the data obtained during this study, a enhanced NeuralIL-type neural network force field was trained and iteratively improved, by utilizing the inherent iterative nature of effective harmonic potentials. With this parametrization exploration of the phase stability and temperature-dependent behavior of four phases of HfO_2 was made possible. After accounting for the lattice overestimation induced by the PBE functional by introducing an artificial pressure, excellent agreement of the lattice thermal expansion of monoclinic and tetragonal hafnia with experiment was observed across a wide temperature range. Similarly, a free-energy-based phase transition from monoclinic to tetragonal was seen at 2600 K. In stark contrast, neither the $Fm\bar{3}m$ phase, nor the lower-energy and lower-symmetry $P\bar{4}3m$ phase were commensurable with experiment and irrespective of the applied pressure or temperature a tetragonal-to-cubic transition was not found.

To further explore the (lack of) cubicity in HfO_2, a preliminary molecular dynamics study was started. This required implementation of a flexible cell NpT-extension to JAX-MD. First results are in agreement with those obtained using effective harmonic potentials, although the work is far from finished.

Finally, the necessity of carefully assessing the validity and limitations of employed methods was made clear in a study exploring the anharmonicities in Lennard-Jones solids and clusters. In particular simple harmonic approaches and power-series-based surrogate models are shown to have severe shortcomings when used in inappropriate situations, constituted by e.g.high temperatures.

49

Outlook

Owing to the developments of the recent years and the challenges lying ahead of us, it can be expected that atomistic simulations will only continue to increase their pace of evolution. With machine learning as a catalyst, a whole new world of possibilities has been made available, the extent and impact of which I believe is not yet in sight. It truly is an excellent time to be working in this field as numerous competing ideas are appearing while a clear "winner" is yet to be determined. As such there is an abundance of methods, topics and fields one could dive into. Here I want to present a vision-board of topics that appear interesting to me.

Hafnia

Given the fact that not all our initial goals for HfO_2 were achieved, and the many more open questions that exist regarding this material, it is only natural to further pursue hafnia. Indeed, as pointed out in the preceding chapter, a future manuscript is already being planned. However, there is still much more to achieve: From the impact of defects or alloying HfO_2, the exploration of orthorhombic and ferroelectric phase stability, to interfaces and their induced strains - hafnia provides a multitude of scientifically interesting issues. Machine learned potentials equipped with an uncertainty metric that integrate well with molecular dynamics or effective harmonic potential codes can enable a more thorough and accurate investigation than ever before and in the future perhaps even the creation of a true *in silico* materials twin. The road to achieving this is not clear cut and there are several possibilities of advancing current methodology towards that goal.

Machine learned potentials

With the ideas introduced in NequIP and Allegro [144, 147], its successor in spirit, the group at Harvard university surrounding Boris Kozinsky have enabled ab initio levels of accuracy for calculations as large as 100 million atoms. Of course this still requires significant amount of computer time but with the large-scale GPU projects[1] currently being undertaken in the US and all over the world, those demands should not pose a significant obstacle in the near future. Thus, I think the usage of MLFFs will become even more widespread in the next years. Different methods should be explored, compared and benchmarked using a standardized framework and dataset. Owing to its architecture and functional programming style, NeuralIL lends itself to experimentation and continuous improvement. Already in the past we have shown that the modular style can be exploited to quickly implement and test novel concepts introduced in the much broader field of machine learning and continuing this work to further improve NeuralIL's capabilities is an exciting proposition.

One aspect I did not touch upon is uncertainty estimates. With GAPs those are essentially built into the model, but with NNs it is also possible to obtain them at a reasonable cost. There are several promising methods, ranging from seemingly trivial ensembling techniques based on bootstrapping or differently seeded initial weights, to more involved concepts like heteroscedastic loss functions [170]. The latter can, in a drastic simplification, be understood as a way of allowing

---

[1]e.g. https://www.nersc.gov/news-publications/nersc-news/science-news/2021/berkeley-lab-targets-exascale-with-perlmutter-and-nesap/

the NN to compensate for inaccurate predictions in the evaluation of the loss: an uncertainty. Implementing, validating and finally exploiting those uncertainty metrics would increase resilience and reliability of NNFFs across the board and should, in my opinion, be a priority.

Automatic differentiability

Automatic differentiability (AD) has so far only been sparsely used in atomistic simulations, aside from making machine learning implementations more tractable. This will likely change in the future. Combining uncertainty metrics with AD and an appropriate sampling scheme to perform active learning based on the uncertainties is a direction worth pursuing [171]. This approach would allow investing only in those DFT calculations that grant the largest improvements. In turn, this could open up efficient exploration and model construction of complex and hard-to-sample configurations, such as interfaces and grain boundaries, further closing the gap to classical MD use cases.

This is not the only area of potential impact of AD: With JAX-MD [20] it is possible to obtain essentially fully differentiable molecular dynamics trajectories. Some of the implications of this are explored in e.g. [172], but I am convinced that we have not yet grasped the potential impact of this development. Going even further, one could imagine a fully featured DFT code based on an AD framework. While there are some first exploratory studies [173], this idea is still in its infancy.

Effective harmonic potentials

Effective harmonic potentials are a versatile and useful tool for studying phases and their behaviors, and as we have shown, generate data for potential construction. As such it would prove to be a promising undertaking to integrate EHP generation into a JAX-based framework that interfaces well with NeuralIL. Combining this integrated framework with an active learning scheme, would allow not only obtaining a useable NNFF and temperature-dependent phonons in one shot, but also ionic relaxation based on free energies.

Moreover EHPs rely on sampling from a real-space probability distribution corresponding to a harmonic potential, a restriction that could be lifted by employing normalizing flows [174]. Conceptually this would even allow sampling directly from a canonical distribution and obtaining thermodynamic averages based on these samples alone, although an intermediate formulation could still rely on a surrogate potential and a Gibbs-Bogliubov-based free energy minimization.

Chapter 7

List of publications

First author

1. Sebastian Bichelmaier et al. "Accurate First-Principles Treatment of the High-Temperature Cubic Phase of Hafnia". In: *Phys. Status Solidi Rapid Res. Lett.* 16.10 (2022), p. 2100642. DOI: 10.1002/pssr.202100642

2. Sebastian Bichelmaier, Jesús Carrete, and Georg K. H. Madsen. "Evaluating the efficiency of power-series expansions as model potentials for finite-temperature atomistic calculations". In: *Int. J. Quantum Chem.* (2023). DOI: 10.1002/QUA.27095

3. Sebastian Bichelmaier et al. "A neural-network-backed effective harmonic potential study of the ambient pressure phases of hafnia". In: *Phys. Rev. B* accepted (2023)

Contributing author

1. Hadrián Montes-Campos et al. "A Differentiable Neural-Network Force Field for Ionic Liquids". In: *J. Chem. Inf. Model.* 62.1 (2022), pp. 88–101. DOI: 10.1021/acs.jcim.1c01380

2. Ralf Wanzenböck et al. "Neural-network-backed evolutionary search for SrTiO3(110) surface reconstructions". In: *Digital Discovery* 1 (5 2022), pp. 703–710. DOI: 10.1039/D2DD00072E

Bibliography

[1] Lixia Qin et al. "Hybrid density functional theory studies of AlN and GaN under uniaxial strain". In: *J. Phys. Condens. Matter* 25.4 (Dec. 2012), p. 045801. DOI: `10.1088/0953-8984/25/4/045801`.

[2] Jesús Carrete et al. "almaBTE : A solver of the space–time dependent Boltzmann transport equation for phonons in structured materials". In: *Comput. Phys. Commun.* 220 (2017), pp. 351–362. ISSN: 0010-4655. DOI: `10.1016/j.cpc.2017.06.023`.

[3] Yuke Zhang et al. "Effects of Strain and Film Thickness on the Stability of the Rhombohedral Phase of HfO_2". In: *Phys. Rev. Appl.* 14 (1 July 2020), p. 014068. DOI: `10.1103/PhysRevApplied.14.014068`.

[4] Nouman Rasool et al. "Biological perspective of thiazolide derivatives against Mpro and MTase of SARS-CoV-2: Molecular docking, DFT and MD simulation investigations". In: *Chem. Phys. Lett.* 771 (2021), p. 138463. ISSN: 0009-2614. DOI: `10.1016/j.cplett.2021.138463`.

[5] M Grieves. "Intelligent digital twins and the development and management of complex systems [version 1; peer review: 4 approved]". In: *Digital Twin* 2.8 (2022). DOI: `10.12688/digitaltwin.17574.1`.

[6] A. Danchin et al. "From data banks to data bases". In: *Res. Microbiol.* 142.7 (1991), pp. 913–916. ISSN: 0923-2508. DOI: `10.1016/0923-2508(91)90073-J`.

[7] Søren Smidstrup et al. "QuantumATK: an integrated platform of electronic and atomic-scale modelling tools". In: *J. Phys. Condens. Matter* 32.1 (Oct. 2019), p. 015901. DOI: `10.1088/1361-648X/ab4007`.

[8] Anubhav Jain et al. "Commentary: The Materials Project: A materials genome approach to accelerating materials innovation". In: *APL Materials* 1.1 (2013), p. 011002. DOI: `10.1063/1.4812323`.

[9] Jörg Behler and Michele Parrinello. "Generalized Neural-Network Representation of High-Dimensional Potential-Energy Surfaces". In: *Phys. Rev. Lett.* 98 (14 Apr. 2007), p. 146401. DOI: `10.1103/PhysRevLett.98.146401`.

[10] Kamal Choudhary et al. "Recent advances and applications of deep learning methods in materials science". In: *npj Comput. Mater.* 8.1 (Apr. 2022), p. 59. ISSN: 2057-3960. DOI: `10.1038/s41524-022-00734-6`.

[11] Ghanshyam Pilania. "Machine learning in materials science: From explainable predictions to autonomous design". In: *Comput. Mater. Sci.* 193 (2021), p. 110360. ISSN: 0927-0256. DOI: `10.1016/j.commatsci.2021.110360`.

[12] D.J. Hooton. "LI. A new treatment of anharmonicity in lattice thermodynamics: I". In: *Lond. Edinb. Dublin philos. mag. j. sci.* 46.375 (1955), pp. 422–432. DOI: `10.1080/14786440408520575`.

[13] D.J. Hooton. "LVIII. Some vibrational properties of solid helium". In: *Lond. Edinb. Dublin philos. mag. j. sci.* 46.376 (1955), pp. 485–498. DOI: `10.1080/14786440508520585`.

[14] James Bradbury et al. *JAX: composable transformations of Python+NumPy programs.* Version 0.3.13. 2018. URL: `http://github.com/google/jax`.

[15] Igor Babuschkin et al. *The DeepMind JAX Ecosystem.* 2020. URL: `http://github.com/deepmind`.

[16] Hadrián Montes-Campos et al. "A Differentiable Neural-Network Force Field for Ionic Liquids". In: *J. Chem. Inf. Model.* 62.1 (2022), pp. 88–101. DOI: `10.1021/acs.jcim.1c01380`.

[17] Ralf Wanzenböck et al. "Neural-network-backed evolutionary search for SrTiO3(110) surface reconstructions". In: *Digital Discovery* 1 (5 2022), pp. 703–710. DOI: `10.1039/D2DD00072E`.

[18] Sebastian Bichelmaier et al. "Accurate First-Principles Treatment of the High-Temperature Cubic Phase of Hafnia". In: *Phys. Status Solidi Rapid Res. Lett.* 16.10 (2022), p. 2100642. DOI: `10.1002/pssr.202100642`.

[19] Sebastian Bichelmaier et al. "A neural-network-backed effective harmonic potential study of the ambient pressure phases of hafnia". In: *Phys. Rev. B* accepted (2023).

[20] Samuel S. Schoenholz and Ekin D. Cubuk. "JAX M.D. A Framework for Differentiable Physics". In: *Advances in Neural Information Processing Systems.* Vol. 33. Curran Associates, Inc., 2020. URL: `bit.ly/jaxmdnips`.

[21] Brienne Johnson and Jacob L. Jones. "Chapter 2 - Structures, Phase Equilibria, and Properties of HfO_2". In: *Ferroelectricity in Doped Hafnium Oxide: Materials, Properties and Devices.* Ed. by Uwe Schroeder, Cheol Seong Hwang, and Hiroshi Funakubo. Woodhead Publishing Series in Electronic and Optical Materials. Woodhead Publishing, 2019, pp. 25–45. ISBN: 978-0-08-102430-0. DOI: `10.1016/B978-0-08-102430-0.00002-4`.

[22] G. D. Wilk, R. M. Wallace, and J. M. Anthony. "High-κ gate dielectrics: Current status and materials properties considerations". In: *J. Appl. Phys.* 89 (2001), pp. 5243–5275. DOI: `10.1063/1.1361065`.

[23] Dominik Fischer and Alfred Kersch. "Stabilization of the high-k tetragonal phase in HfO_2: The influence of dopants and temperature from ab initio simulations". In: *J. Appl. Phys.* 104 (2008), p. 084104. DOI: `10.1063/1.2999352`.

[24] Chong Wang, Matvei Zinkevich, and Fritz Aldinger. "The Zirconia–Hafnia System: DTA Measurements and Thermodynamic Calculations". In: *J. Am. Ceram. Soc.* 89.12 (2006), pp. 3751–3758. DOI: `10.1111/j.1551-2916.2006.01286.x`.

[25] J. Wang, H. P. Li, and R. Stevens. "Hafnia and hafnia-toughened ceramics". In: *J. Mater. Sci.* 27.20 (Oct. 1992), pp. 5397–5430. ISSN: 1573-4803. DOI: 10.1007/BF00541601.

[26] Kazuhide Matsumoto, Yoshiyasu Itoh, and Tsuneji Kameda. "EB-PVD process and thermal properties of hafnia-based thermal barrier coating". In: *Sci. Technol. Adv. Mater.* 4.2 (2003), pp. 153–158. ISSN: 1468-6996. DOI: 10.1016/S1468-6996(03)00009-3.

[27] You-Dong Kim et al. "Stable ceria-based electrolytes for intermediate temperature-solid oxide fuel cells via hafnium oxide blocking layer". In: *J. Alloys Compd.* 779 (2019), pp. 121–128. ISSN: 0925-8388. DOI: 10.1016/j.jallcom.2018.11.069.

[28] T. S. Böscke et al. "Ferroelectricity in hafnium oxide thin films". In: *Appl. Phys. Lett.* 99 (2011), p. 102903. DOI: 10.1063/1.3634052.

[29] Yongjian Luo et al. "Ferroelectricity in dopant-free HfO_2 thin films prepared by pulsed laser deposition". In: *J. Materiomics* 8.2 (2022), pp. 311–318. ISSN: 2352-8478. DOI: 10.1016/j.jmat.2021.09.005.

[30] J. Müller et al. "Ferroelectric Hafnium Oxide Based Materials and Devices: Assessment of Current Status and Future Prospects". In: *ECS J. Solid State Sci.* 4 (2015), N30–N35. DOI: 10.1149/2.0081505jss.

[31] Johannes Mueller, Stefan Slesazeck, and Thomas Mikolajick. "Chapter 10.4 - Ferroelectric Field Effect Transistor". In: *Ferroelectricity in Doped Hafnium Oxide: Materials, Properties and Devices.* Ed. by Uwe Schroeder, Cheol Seong Hwang, and Hiroshi Funakubo. Woodhead Publishing Series in Electronic and Optical Materials. Woodhead Publishing, 2019, pp. 451–471. ISBN: 978-0-08-102430-0. DOI: 10.1016/B978-0-08-102430-0.00022-X.

[32] C. E. Curtis, L. M. Doney, and J. R. Johnson. "Some Properties of Hafnium Oxide, Hafnium Silicate, Calcium Hafnate, and Hafnium Carbide". In: *J. Am. Ceram. Soc.* 37.10 (1954), pp. 458–465. DOI: 10.1111/j.1151-2916.1954.tb13977.x.

[33] R. N. Patil and E. C. Subbarao. "Axial thermal expansion of ZrO_2 and HfO_2 in the range room temperature to 1400°C". In: *J. Appl. Crystallogr.* 2.6 (1969). DOI: 10.1107/S0021889869007217.

[34] Robert Ruh and Peter W. R. Corfield. "Crystal Structure of Monoclinic Hafnia and Comparison with Monoclinic Zirconia". In: *J. Am. Ceram. Soc.* 53.3 (1970), pp. 126–129. DOI: 10.1111/j.1151-2916.1970.tb12052.x.

[35] Lin-Gun Liu. "New high pressure phases of ZrO_2 and HfO_2". In: *J. Phys. Chem. Solids* 41.4 (1980), pp. 331–334. ISSN: 0022-3697. DOI: 10.1016/0022-3697(80)90205-X.

[36] Osamu Ohtaka, Takamitsu Yamanaka, and Shoichi Kume. "Synthesis and X-Ray Structural Analysis by the Rietveld Method of Orthorhombic Hafnia". In: *J. Ceram. Soc. Jpn.* 99.1153 (1991), pp. 826–827. DOI: 10.2109/jcersj.99.826.

[37] J.M. Leger et al. "Pressure-induced phase transitions and volume changes in HfO_2 up to 50 GPa". In: *Phys. Rev. B* 48.1 (1993), pp. 93–98. DOI: 10.1103/PhysRevB.48.93.

[38] Tran Doan Huan et al. "Pathways towards ferroelectricity in hafnia". In: *Phys. Rev. B* 90 (6 Aug. 2014), p. 064111. DOI: 10.1103/PhysRevB.90.064111.

[39] Ryan P. Haggerty et al. "Thermal Expansion of HfO_2 and ZrO_2". In: *J. Am. Ceram. Soc.* 97.7 (2014), pp. 2213–2222. DOI: `10.1111/jace.12975`.

[40] Rohit Batra, Huan Doan Tran, and Rampi Ramprasad. "Stabilization of metastable phases in hafnia owing to surface energy effects". In: *Appl. Phys. Lett.* 108.17 (2016), p. 172902. DOI: `10.1063/1.4947490`.

[41] Rohit Batra et al. "Factors Favoring Ferroelectricity in Hafnia: A First-Principles Computational Study". In: *J. Phys. Chem. C* 121.8 (2017), pp. 4139–4145. DOI: `10.1021/acs.jpcc.6b11972`.

[42] Tsubasa Tobase et al. "Pre-Transitional Behavior in Tetragonal to Cubic Phase Transition in HfO_2 Revealed by High Temperature Diffraction Experiments". In: *Phys. Status Solidi (b)* 255.11 (2018), p. 1800090. DOI: `10.1002/pssb.201800090`.

[43] Weitong Ding et al. "The atomic-scale domain wall structure and motion in HfO_2-based ferroelectrics: A first-principle study". In: *Acta Mater.* 196 (2020), pp. 556–564. ISSN: 1359-6454. DOI: `10.1016/j.actamat.2020.07.012`.

[44] S. S. Behara and A. Van der Ven. "Ferroelectric HfO_2 and the importance of strain". In: *Phys. Rev. Materials* 6 (5 May 2022), p. 054403. DOI: `10.1103/PhysRevMaterials.6.054403`.

[45] Thaddeus B Massalski et al. *Binary alloy phase diagrams*. Vol. 1/2. American society for metals Metals Park, OH, 1986.

[46] P. Fan et al. "Origin of the Intrinsic Ferroelectricity of HfO_2 from ab Initio Molecular Dynamics". In: *J. Phys. Chem. C* 123.35 (2019), pp. 21743–21750. DOI: `10.1021/acs.jpcc.9b04106`.

[47] Uwe Schroeder et al. "The fundamentals and applications of ferroelectric HfO_2". In: *Nat. Rev. Mater.* 7.8 (Aug. 2022), pp. 653–669. ISSN: 2058-8437. DOI: `10.1038/s41578-022-00431-2`.

[48] Koichi Momma and Fujio Izumi. "*VESTA*: a three-dimensional visualization system for electronic and structural analysis". In: *J. Appl. Crystallogr.* 41.3 (June 2008), pp. 653–658. DOI: `10.1107/S0021889808012016`.

[49] K. J. McClellan et al. "Determination of the structure of the cubic phase in high-ZrO_2 Y_2O_3-ZrO_2 alloys by convergent-beam electron diffraction". In: *Philos. Mag. A* 70.1 (1994), pp. 185–200. DOI: `10.1080/01418619408242545`.

[50] C. M. Wang et al. "Distortion of the oxygen sublattice in pure cubic-ZrO_2". In: *J. Mater. Res.* 19.5 (May 2004), pp. 1315–1319. ISSN: 2044-5326. DOI: `10.1557/JMR.2004.0175`.

[51] M. Sternik and K. Parlinski. "Free-energy calculations for the cubic ZrO_2 crystal as an example of a system with a soft mode". In: *J. Chem. Phys.* 123.20 (2005), p. 204708. DOI: `10.1063/1.2124708`.

[52] Osamu Ohtaka et al. "Phase relations and equation of state of ZrO_2 to 100GPa". In: *J. Appl. Crystallogr.* 38.5 (Oct. 2005), pp. 727–733. DOI: `10.1107/S0021889805018145`.

[53] Peitao Liu et al. "Phase transitions of zirconia: Machine-learned force fields beyond density functional theory". In: *Phys. Rev. B* 105 (6 Feb. 2022), p. L060102. DOI: `10.1103/PhysRevB.105.L060102`.

[54] Luis Azevedo Antunes et al. "Characteristics of Low-Energy Phases of Hafnia and Zirconia from Density Functional Theory Calculations". In: *Phys. Status Solidi Rapid Res. Lett.* 16.10 (2022), p. 2100636. DOI: `10.1002/pssr.202100636`.

[55] Ganesh Sivaraman et al. "Experimentally Driven Automated Machine-Learned Interatomic Potential for a Refractory Oxide". In: *Phys. Rev. Lett.* 126 (15 Apr. 2021), p. 156002. DOI: `10.1103/PhysRevLett.126.156002`.

[56] S. V. Barabash. "Prediction of new metastable HfO_2 phases: toward understanding ferro- and antiferroelectric films". In: *J. Comput. Electron.* 16.4 (Dec. 2017), pp. 1227–1235. DOI: `10.1007/s10825-017-1077-5`.

[57] Qi-Jun Hong et al. "Combined computational and experimental investigation of high temperature thermodynamics and structure of cubic ZrO_2 and HfO_2". In: *Sci. Rep.* 8.1 (Oct. 2018), p. 14962. ISSN: 2045-2322. DOI: `10.1038/s41598-018-32848-7`.

[58] Terence Mittmann et al. "Origin of Ferroelectric Phase in Undoped HfO_2 Films Deposited by Sputtering". In: *Adv. Mater. Interfaces* 6.11 (2019), p. 1900042. DOI: `10.1002/admi.201900042`.

[59] Nico Kaiser et al. "Defect-Stabilized Substoichiometric Polymorphs of Hafnium Oxide with Semiconducting Properties". In: *ACS Appl. Mater. Interfaces* 14.1 (2022), pp. 1290–1303. DOI: `10.1021/acsami.1c09451`.

[60] P. Rauwel et al. "One step synthesis of pure cubic and monoclinic HfO_2 nanoparticles: Correlating the structure to the electronic properties of the two polymorphs". In: *J. Appl. Phys.* 112.10 (2012), p. 104107. DOI: `10.1063/1.4766272`.

[61] D. Islamov and T. Perevalov. "Effect of oxygen vacancies on the ferroelectric $Hf_{0.5}Zr_{0.5}O_2$ stabilization: DFT simulation". In: *Microelectron. Eng.* 216 (2019), p. 111041. ISSN: 0167-9317. DOI: `10.1016/j.mee.2019.111041`.

[62] Uwe Schroeder et al. "Lanthanum-Doped Hafnium Oxide: A Robust Ferroelectric Material". In: *Inorg. Chem.* 57.5 (2018), pp. 2752–2765. DOI: `10.1021/acs.inorgchem.7b03149`.

[63] T. Schenk et al. "Strontium doped hafnium oxide thin films: Wide process window for ferroelectric memories". In: *2013 Proceedings of the European Solid-State Device Research Conference (ESSDERC)*. 2013, pp. 260–263. DOI: `10.1109/ESSDERC.2013.6818868`.

[64] M. Hoffmann et al. "Stabilizing the ferroelectric phase in doped hafnium oxide". In: *J. Appl. Phys.* 118.7 (2015), p. 072006. DOI: `10.1063/1.4927805`.

[65] T. S. Boscke et al. "Tetragonal Phase Stabilization by Doping as an Enabler of Thermally Stable HfO_2 based MIM and MIS Capacitors for sub 50nm Deep Trench DRAM". In: *2006 International Electron Devices Meeting*. 2006, pp. 1–4. DOI: `10.1109/IEDM.2006.347011`.

[66] G. H. Chen et al. "Effects of Y doping on the structural stability and defect properties of cubic HfO2". In: *J. Appl. Phys.* 104.7 (2008), p. 074101. DOI: `10.1063/1.2985908`.

[67] Songsong Zhou, Jiahao Zhang, and Andrew M. Rappe. "Strain-induced antipolar phase in hafnia stabilizes robust thin-film ferroelectricity". In: *Sci. Adv.* 8.47 (2022), eadd5953. DOI: `10.1126/sciadv.add5953`.

[68] Shi Liu and Brendan M. Hanrahan. "Effects of growth orientations and epitaxial strains on phase stability of HfO_2 thin films". In: *Phys. Rev. Mater.* 3 (5 May 2019), p. 054404. DOI: `10.1103/PhysRevMaterials.3.054404`.

[69] Sheng-Ting Fan, Yun-Wen Chen, and C W Liu. "Strain effect on the stability in ferroelectric HfO_2 simulated by first-principles calculations". In: *J. Phys. D Appl. Phys.* 53.23 (Apr. 2020), 23LT01. DOI: `10.1088/1361-6463/ab7fd4`.

[70] Han Wang et al. "DeePMD-kit: A deep learning package for many-body potential energy representation and molecular dynamics". In: *Computer Physics Communications* 228 (2018), pp. 178–184. ISSN: 0010-4655. DOI: `10.1016/j.cpc.2018.03.016`.

[71] Jing Wu et al. "Deep learning of accurate force field of ferroelectric HfO_2". In: *Phys. Rev. B* 103 (2 Jan. 2021), p. 024108. DOI: `10.1103/PhysRevB.103.024108`.

[72] P. Hohenberg and W. Kohn. "Inhomogeneous Electron Gas". In: *Phys. Rev.* 136 (3B Nov. 1964), B864–B871. DOI: `10.1103/PhysRev.136.B864`.

[73] M. Born and R. Oppenheimer. "Zur Quantentheorie der Molekeln". In: *Ann. Phys.* 389.20 (1927), pp. 457–484. DOI: `10.1002/andp.19273892002`.

[74] Axel D. Becke. "Perspective: Fifty years of density-functional theory in chemical physics". In: *J. Chem. Phys.* 140.18 (2014), 18A301. DOI: `10.1063/1.4869598`.

[75] W. Kohn and L. J. Sham. "Self-Consistent Equations Including Exchange and Correlation Effects". In: *Phys. Rev.* 140 (4A Nov. 1965), A1133–A1138. DOI: `10.1103/PhysRev.140.A1133`.

[76] J. P. Perdew and Y. Wang. "Accurate and simple analytic representation of the electron-gas correlation energy". In: *Phys. Rev. B* 45 (23 June 1992), pp. 13244–13249. DOI: `10.1103/PhysRevB.45.13244`.

[77] J. P. Perdew, K. Burke, and M. Ernzerhof. "Generalized Gradient Approximation Made Simple". In: *Phys. Rev. Lett.* 77 (18 Oct. 1996), pp. 3865–3868. DOI: `10.1103/PhysRevLett.77.3865`.

[78] J. P. Perdew et al. "Restoring the Density-Gradient Expansion for Exchange in Solids and Surfaces". In: *Phys. Rev. Lett.* 100 (13 Apr. 2008), p. 136406. DOI: `10.1103/PhysRevLett.100.136406`.

[79] Pedro Borlido et al. "Exchange-correlation functionals for band gaps of solids: benchmark, reparametrization and machine learning". In: *npj Comput. Mater.* 6.1 (July 2020), p. 96. ISSN: 2057-3960. DOI: `10.1038/s41524-020-00360-0`.

[80] Narbe Mardirossian and Martin Head-Gordon. "Thirty years of density functional theory in computational chemistry: an overview and extensive assessment of 200 density functionals". In: *Mol. Phys.* 115.19 (2017), pp. 2315–2372. DOI: `10.1080/00268976.2017.1333644`.

[81]   C. Cohen-Tannoudji et al. *Quantum Mechanics*. A Wiley - Interscience publication Bd. 1. Wiley, 1977. ISBN: 9782705658335. URL: `https://books.google.at/books?id=iHcpAQAAMAAJ`.

[82]   R. P. Feynman. "Forces in Molecules". In: *Phys. Rev.* 56 (4 Aug. 1939), pp. 340–343. DOI: `10.1103/PhysRev.56.340`.

[83]   P. E. Blöchl. "Projector augmented-wave method". In: *Phys. Rev. B* 50 (24 Dec. 1994), pp. 17953–17979. DOI: `10.1103/PhysRevB.50.17953`.

[84]   G. Kresse and D. Joubert. "From ultrasoft pseudopotentials to the projector augmented-wave method". In: *Phys. Rev. B* 59 (3 Jan. 1999), pp. 1758–1775. DOI: `10.1103/PhysRevB.59.1758`.

[85]   K. Parlinski, Z. Q. Li, and Y. Kawazoe. "First-Principles Determination of the Soft Mode in Cubic ZrO_2". In: *Phys. Rev. Lett.* 78 (21 May 1997), pp. 4063–4066. DOI: `10.1103/PhysRevLett.78.4063`.

[86]   Yoyo Hinuma et al. "Band structure diagram paths based on crystallography". In: *Comput. Mater. Sci.* 128 (2017), pp. 140–184. ISSN: 0927-0256. DOI: `10.1016/j.commatsci.2016.10.015`.

[87]   L. E. Reichl. *A modern course in statistical physics; 1st ed.* London: Arnold, 1980. URL: `https://www.wiley.com/en-us/A+Modern+Course+in+Statistical+Physics%2C+4th+Edition-p-9783527413492`.

[88]   Francis Birch. "Finite Elastic Strain of Cubic Crystals". In: *Phys. Rev.* 71 (11 June 1947), pp. 809–824. DOI: `10.1103/PhysRev.71.809`.

[89]   Atsushi Togo and Isao Tanaka. "First principles phonon calculations in materials science". In: *Scr. Mater.* 108 (2015), pp. 1–5. ISSN: 1359-6462. DOI: `10.1016/j.scriptamat.2015.07.021`.

[90]   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". In: *Phys. Rev. B* 89 (6 Feb. 2014), p. 064302. DOI: `10.1103/PhysRevB.89.064302`.

[91]   Lorenzo Monacelli et al. "The stochastic self-consistent harmonic approximation: calculating vibrational properties of materials with full quantum and anharmonic effects". In: *J. Phys. Condens. Matter* 33 (July 2021), p. 363001. DOI: `10.1088/1361-648x/ac066b`.

[92]   A Isihara. "The Gibbs-Bogoliubov inequality dagger". In: *J. Phys. A* 1 (Sept. 1968), pp. 539–548. DOI: `10.1088/0305-4470/1/5/305`.

[93]   Ambroise van Roekeghem, Jesús Carrete, and Natalio Mingo. "Quantum Self-Consistent Ab-Initio Lattice Dynamics". In: *Comput. Phys. Commun.* 263 (2021), p. 107945. ISSN: 0010-4655. DOI: `10.1016/j.cpc.2021.107945`.

[94]   Fredrik Eriksson, Erik Fransson, and Paul Erhart. "The Hiphive Package for the Extraction of High-Order Force Constants by Machine Learning". In: *Adv. Theo. Simul.* 2.5 (2019), p. 1800184. DOI: `10.1002/adts.201800184`.

[95] P. Pulay. "Improved SCF convergence acceleration". In: *J. Comput. Chem* 3 (1982), pp. 556–560. DOI: 10.1002/jcc.540030413.

[96] Leslie Kish. "Sampling Organizations and Groups of Unequal Sizes". In: *Am. Sociol. Rev.* 30.4 (1965), pp. 564–572. DOI: 10.2307/2091346.

[97] Augustine Kong, Jun S. Liu, and Wing Hung Wong. "Sequential Imputations and Bayesian Missing Data Problems". In: *J. Am. Stat. Assoc.* 89.425 (1994), pp. 278–288. DOI: 10.1080/01621459.1994.10476469.

[98] N. Hansen. "The CMA Evolution Strategy: A Tutorial". Version 1. In: *arXiv:1604.00772* (2016). DOI: 10.48550/ARXIV.1604.00772.

[99] Michele Invernizzi, Pablo M. Piaggi, and Michele Parrinello. "Unified Approach to Enhanced Sampling". In: *Phys. Rev. X* 10 (4 Nov. 2020), p. 041034. DOI: 10.1103/PhysRevX.10.041034.

[100] Stefan Schrunner et al. "A Health Factor for Process Patterns Enhancing Semiconductor Manufacturing by Pattern Recognition in Analog Wafermaps". In: *2019 IEEE International Conference on Systems, Man and Cybernetics (SMC)*. Bari, Italy: IEEE Press, 2019, pp. 3555–3560. DOI: 10.1109/SMC.2019.8914309.

[101] Martin Leo, Suneel Sharma, and K. Maddulety. "Machine Learning in Banking Risk Management: A Literature Review". In: *Risks* 7.1 (2019). ISSN: 2227-9091. DOI: 10.3390/risks7010029.

[102] Martín Abadi et al. *TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems.* 2015. URL: https://www.tensorflow.org/.

[103] F. Pedregosa et al. "Scikit-learn: Machine Learning in Python". In: *J. Mach. Learn. Res.* 12 (2011), pp. 2825–2830. URL: http://jmlr.org/papers/v12/pedregosa11a.html.

[104] Adam Paszke et al. "PyTorch: An Imperative Style, High-Performance Deep Learning Library". In: *Advances in Neural Information Processing Systems 32*. Curran Associates, Inc., 2019, pp. 8024–8035. URL: http://papers.neurips.cc/paper/9015-pytorch-an-imperative-style-high-performance-deep-learning-library.pdf.

[105] Jonathan Heek et al. *Flax: A neural network library and ecosystem for JAX.* Version 0.6.3. 2020. URL: http://github.com/google/flax.

[106] Rasmus Tranås et al. "Lattice thermal conductivity of half-Heuslers with density functional theory and machine learning: Enhancing predictivity by active sampling with principal component analysis". In: *Comput. Mater. Sci.* 202 (2022), p. 110938. ISSN: 0927-0256. URL: https://www.sciencedirect.com/science/article/pii/S0927025621006376.

[107] Jesús Carrete et al. "Finding Unprecedentedly Low-Thermal-Conductivity Half-Heusler Semiconductors via High-Throughput Materials Modeling". In: *Phys. Rev. X* 4 (1 Feb. 2014), p. 011019. DOI: 10.1103/PhysRevX.4.011019.

[108] Albert P. Bartók et al. "Gaussian Approximation Potentials: The Accuracy of Quantum Mechanics, without the Electrons". In: *Phys. Rev. Lett.* 104 (13 Apr. 2010), p. 136403. DOI: 10.1103/PhysRevLett.104.136403.

[109] K. T. Schütt et al. "Unifying machine learning and quantum chemistry with a deep neural network for molecular wavefunctions". In: *Nat. Commun.* 10.1 (Nov. 2019), p. 5024. ISSN: 2041-1723. DOI: 10.1038/s41467-019-12875-2.

[110] M. Gastegger et al. "A deep neural network for molecular wave functions in quasi-atomic minimal basis representation". In: *J. Chem. Phys.* 153.4 (2020), p. 044123. DOI: 10.1063/5.0012911.

[111] Kunihiko Fukushima. "Cognitron: A self-organizing multilayered neural network". In: *Biol. Cybern.* 20.3 (Sept. 1975), pp. 121–136. ISSN: 1432-0770. DOI: 10.1007/BF00342633.

[112] Prajit Ramachandran, Barret Zoph, and Quoc V. Le. *Searching for Activation Functions.* 2017. DOI: 10.48550/ARXIV.1710.05941.

[113] Garrett B. Goh, Nathan O. Hodas, and Abhinav Vishnu. "Deep learning for computational chemistry". In: *J. Comput. Chem.* 38.16 (2017), pp. 1291–1307. DOI: 10.1002/jcc.24764.

[114] Kurt Hornik, Maxwell Stinchcombe, and Halbert White. "Multilayer feedforward networks are universal approximators". In: *Neural Netw.* 2.5 (1989), pp. 359–366. ISSN: 0893-6080. DOI: 10.1016/0893-6080(89)90020-8.

[115] Diederik P. Kingma and Jimmy Ba. *Adam: A Method for Stochastic Optimization.* 2014. DOI: 10.48550/ARXIV.1412.6980.

[116] Sergey Ioffe and Christian Szegedy. *Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift.* 2015. DOI: 10.48550/ARXIV.1502.03167.

[117] Jimmy Lei Ba, Jamie Ryan Kiros, and Geoffrey E. Hinton. *Layer Normalization.* 2016. DOI: 10.48550/ARXIV.1607.06450.

[118] Leslie N. Smith. *A disciplined approach to neural network hyper-parameters: Part 1 – learning rate, batch size, momentum, and weight decay.* 2018. DOI: 10.48550/ARXIV.1803.09820.

[119] Qi Wang et al. "A Comprehensive Survey of Loss Functions in Machine Learning". In: *Ann. Data Sci.* 9.2 (Apr. 2022), pp. 187–212. ISSN: 2198-5812. DOI: 10.1007/s40745-020-00253-5.

[120] Jörg Behler. "Constructing high-dimensional neural network potentials: A tutorial review". In: *Int. J. Quantum Chem.* 115.16 (2015), pp. 1032–1050. DOI: 10.1002/qua.24890.

[121] Emir Kocer, Jeremy K. Mason, and Hakan Erturk. "Continuous and optimally complete description of chemical environments using Spherical Bessel descriptors". In: *AIP Advances* 10.1 (2020), p. 015021. DOI: 10.1063/1.5111045.

[122] Albert P. Bartók et al. "Machine Learning a General-Purpose Interatomic Potential for Silicon". In: *Phys. Rev. X* 8 (4 Dec. 2018), p. 041048. DOI: 10.1103/PhysRevX.8.041048.

[123] Chang Lyoul Kong. "Combining rules for intermolecular potential parameters. II. Rules for the Lennard-Jones (12–6) potential and the Morse potential". In: *J. Chem. Phys.* 59.5 (1973), pp. 2464–2467. DOI: 10.1063/1.1680358.

[124] Somayeh Faraji et al. "High accuracy and transferability of a neural network potential through charge equilibration for calcium fluoride". In: *Phys. Rev. B* 95 (10 Mar. 2017), p. 104105. DOI: 10.1103/PhysRevB.95.104105.

[125] Tsz Wai Ko et al. "A fourth-generation high-dimensional neural network potential with accurate electrostatics including non-local charge transfer". In: *Nat. Commun.* 12.1 (Jan. 2021), p. 398. ISSN: 2041-1723. DOI: 10.1038/s41467-020-20427-2.

[126] S. Alireza Ghasemi et al. "Interatomic potentials for ionic systems with density functional accuracy based on charge densities obtained by a neural network". In: *Phys. Rev. B* 92 (4 July 2015), p. 045131. DOI: 10.1103/PhysRevB.92.045131.

[127] F. L. Hirshfeld. "Bonded-atom fragments for describing molecular charge densities". In: *Theor. Chim. Acta* 44.2 (June 1977), pp. 129–138. ISSN: 1432-2234. DOI: 10.1007/BF00549096.

[128] R M Lynden-Bell and T G A Youngs. "Simulations of imidazolium ionic liquids: when does the cation charge distribution matter?" In: *J. Phys. Condens. Matter* 21.42 (Sept. 2009), p. 424120. DOI: 10.1088/0953-8984/21/42/424120.

[129] Albert P. Bartók, Risi Kondor, and Gábor Csányi. "On representing chemical environments". In: *Phys. Rev. B* 87 (18 May 2013), p. 184115. DOI: 10.1103/PhysRevB.87.184115.

[130] Lauri Himanen et al. "DScribe: Library of descriptors for machine learning in materials science". In: *Comput. Phys. Commun.* 247 (2020), p. 106949. ISSN: 0010-4655. DOI: 10.1016/j.cpc.2019.106949.

[131] Nongnuch Artrith and Jörg Behler. "High-dimensional neural network potentials for metal surfaces: A prototype study for copper". In: *Phys. Rev. B* 85 (4 Jan. 2012), p. 045439. DOI: 10.1103/PhysRevB.85.045439.

[132] Mingjie Liu and John R. Kitchin. "SingleNN: Modified Behler–Parrinello Neural Network with Shared Weights for Atomistic Simulations with Transferability". In: *J. Phys. Chem. C* 124.32 (2020), pp. 17811–17818. DOI: 10.1021/acs.jpcc.0c04225.

[133] Kun Yao et al. "The TensorMol-0.1 model chemistry: a neural network augmented with long-range physics". In: *Chem. Sci.* 9 (8 2018), pp. 2261–2269. DOI: 10.1039/C7SC04934J.

[134] Samare Rostami, Maximilian Amsler, and S. Alireza Ghasemi. "Optimized symmetry functions for machine-learning interatomic potentials of multicomponent systems". In: *J. Chem. Phys.* 149.12 (2018), p. 124106. DOI: 10.1063/1.5040005.

[135] J. S. Smith, O. Isayev, and A. E. Roitberg. "ANI-1: an extensible neural network potential with DFT accuracy at force field computational cost". In: *Chem. Sci.* 8 (4 2017), pp. 3192–3203. DOI: 10.1039/C6SC05720A.

[136] Diego Milardovich et al. "Building robust machine learning force fields by composite Gaussian approximation potentials". In: *Solid-State Electron.* 200 (2023), p. 108529. ISSN: 0038-1101. DOI: 10.1016/j.sse.2022.108529.

[137] Emir Kocer, Tsz Wai Ko, and Jörg Behler. "Neural Network Potentials: A Concise Overview of Methods". In: *Annu. Rev. Phys. Chem.* 73.1 (2022), pp. 163–186. DOI: 10.1146/annurev-physchem-082720-034254.

[138] Oliver T. Unke et al. "Machine Learning Force Fields". In: *Chem. Rev.* 121.16 (2021), pp. 10142–10186. DOI: 10.1021/acs.chemrev.0c01111.

[139] Carl Edward Rasmussen and Christopher K. I. Williams. *Gaussian Processes for Machine Learning*. The MIT Press, Nov. 2005. ISBN: 9780262256834. DOI: `10.7551/mitpress/3206.001.0001`.

[140] Albert P. Bartók and Gábor Csányi. "Gaussian approximation potentials: A brief tutorial introduction". In: *Int. J. Quantum Chem.* 115.16 (2015), pp. 1051–1057. DOI: `10.1002/qua.24927`.

[141] Volker L. Deringer et al. "Gaussian Process Regression for Materials and Molecules". In: *Chem. Rev.* 121.16 (2021), pp. 10073–10141. DOI: `10.1021/acs.chemrev.1c00022`.

[142] Kristof T. Schütt et al. "SchNet: A continuous-filter convolutional neural network for modeling quantum interactions". In: *arXiv e-prints* (June 2017). arXiv: `1706.08566 [stat.ML]`.

[143] K. T. Schütt et al. "SchNet – A deep learning architecture for molecules and materials". In: *J. Chem. Phys.* 148.24 (2018), p. 241722. DOI: `10.1063/1.5019779`.

[144] Simon Batzner et al. "E(3)-equivariant graph neural networks for data-efficient and accurate interatomic potentials". In: *Nat. Commun.* 13.1 (May 2022), p. 2453. ISSN: 2041-1723. DOI: `10.1038/s41467-022-29939-5`.

[145] Justin Gilmer et al. "Neural Message Passing for Quantum Chemistry". In: *arXiv e-prints* (Apr. 2017). arXiv: `1704.01212 [cs.LG]`.

[146] David K. Duvenaud et al. "Convolutional Networks on Graphs for Learning Molecular Fingerprints". In: *NIPS*. 2015, pp. 2224–2232. URL: `http://papers.nips.cc/paper/5954-convolutional-networks-on-graphs-for-learning-molecular-fingerprints`.

[147] Albert Musaelian et al. "Learning Local Equivariant Representations for Large-Scale Atomistic Dynamics". In: *arXiv e-prints* (Apr. 2022). arXiv: `2204.05249 [physics.comp-ph]`.

[148] Mario Geiger and Tess Smidt. *e3nn: Euclidean Neural Networks*. 2022. DOI: `10.48550/ARXIV.2207.09453`.

[149] William C. Swope et al. "A computer simulation method for the calculation of equilibrium constants for the formation of physical clusters of molecules: Application to small water clusters". In: *J. Chem. Phys.* 76.1 (1982), pp. 637–649. DOI: `10.1063/1.442716`.

[150] M. Tuckerman. *Statistical Mechanics: Theory and Molecular Simulation*. Oxford Graduate Texts. OUP Oxford, 2010. ISBN: 9780191523465. URL: `https://books.google.at/books?id=Lo3Jqc0pgrcC`.

[151] Glenn J. Martyna et al. "Explicit reversible integrators for extended systems dynamics". In: *Mol. Phys.* 87.5 (1996), pp. 1117–1157. DOI: `10.1080/00268979600100761`.

[152] Mark E. Tuckerman et al. "Non-Hamiltonian molecular dynamics: Generalizing Hamiltonian phase space principles to non-Hamiltonian systems". In: *J. Chem. Phys.* 115.4 (2001), pp. 1678–1702. DOI: `10.1063/1.1378321`.

[153] Mark E Tuckerman et al. "A Liouville-operator derived measure-preserving integrator for molecular dynamics simulations in the isothermal–isobaric ensemble". In: *J. Phys. A Math. Gen.* 39.19 (Apr. 2006), p. 5629. DOI: `10.1088/0305-4470/39/19/S18`.

[154] H. F. Trotter. "On the product of semi-groups of operators". In: *Proc. Amer. Math. Soc.* 10 (1959), pp. 545–551. ISSN: 0002-9939. DOI: `10.2307/2033649`.

[155] H. J. C. Berendsen et al. "Molecular dynamics with coupling to an external bath". In: *J. Chem. Phys.* 81.8 (1984), pp. 3684–3690. DOI: `10.1063/1.448118`.

[156] William G. Hoover. "Canonical dynamics: Equilibrium phase-space distributions". In: *Phys. Rev. A* 31 (3 Mar. 1985), pp. 1695–1697. DOI: `10.1103/PhysRevA.31.1695`.

[157] Glenn J. Martyna, Michael L. Klein, and Mark Tuckerman. "Nosé–Hoover chains: The canonical ensemble via continuous dynamics". In: *J. Chem. Phys.* 97.4 (1992), pp. 2635–2643. DOI: `10.1063/1.463940`.

[158] Masuo Suzuki. "General theory of fractal path integrals with applications to many-body theories and statistical physics". In: *J. Math. Phys.* 32.2 (1991), pp. 400–407. DOI: `10.1063/1.529425`.

[159] Haruo Yoshida. "Construction of higher order symplectic integrators". In: *Phys. Lett. A* 150.5 (1990), pp. 262–268. ISSN: 0375-9601. DOI: `10.1016/0375-9601(90)90092-3`.

[160] C. Runge. "Über die numerische Auflösung von Differentialgleichungen." In: *Math. Ann.* 46 (1895), pp. 167–178. URL: `http://eudml.org/doc/157756`.

[161] W. Kutta. "Beitrag zur näherungsweisen Integration totaler Differentialgleichungen". In: *Zeit. Math. Phys.* 46 (1901), pp. 435–53.

[162] Glenn J. Martyna, Douglas J. Tobias, and Michael L. Klein. "Constant pressure molecular dynamics algorithms". In: *J. Chem. Phys.* 101.5 (1994), pp. 4177–4189. DOI: `10.1063/1.467468`.

[163] Tang-Qing Yu et al. "Measure-preserving integrators for molecular dynamics in the isothermal–isobaric ensemble derived from the Liouville operator". In: *Chem. Phys.* 370.1 (2010), pp. 294–305. ISSN: 0301-0104. DOI: `10.1016/j.chemphys.2010.02.014`.

[164] Frank H. Stillinger and Thomas A. Weber. "Computer simulation of local order in condensed phases of silicon". In: *Phys. Rev. B* 31 (8 Apr. 1985), pp. 5262–5271. DOI: `10.1103/PhysRevB.31.5262`.

[165] Pinchen Xie et al. "Ab initio multi-scale modeling of ferroelectrics: The case of PbTiO3". In: *arXiv e-prints* (May 2022). DOI: `10.48550/ARXIV.2205.11839`.

[166] A. K. Rappe et al. "UFF, a full periodic table force field for molecular mechanics and molecular dynamics simulations". In: *J. Am. Chem. Soc.* 114.25 (1992), pp. 10024–10035. DOI: `10.1021/ja00051a040`.

[167] Yin Wang et al. "Structure and dielectric properties of amorphous high-κ oxides: HfO_2, ZrO_2, and their alloys". In: *Phys. Rev. B* 85 (22 June 2012), p. 224110. DOI: `10.1103/PhysRevB.85.224110`.

[168] Tzu-Ray Shan et al. "Charge-optimized many-body potential for the hafnium/hafnium oxide system". In: *Phys. Rev. B* 81 (12 Mar. 2010), p. 125328. DOI: `10.1103/PhysRevB.81.125328`.

[169] G Broglia et al. "Molecular dynamics simulation of amorphous HfO_2 for resistive RAM applications". In: *Model. Simul. Mat. Sci. Eng.* 22.6 (July 2014), p. 065006. DOI: `10.1088/0965-0393/22/6/065006`.

[170] Romain Egele et al. "AutoDEUQ: Automated Deep Ensemble with Uncertainty Quantification". In: *arXiv e-prints* (Oct. 2021). arXiv: 2110.13511 [cs.LG].

[171] Daniel Schwalbe-Koda, Aik Rui Tan, and Rafael Gómez-Bombarelli. "Differentiable sampling of molecular geometries with uncertainty-based adversarial attacks". In: *Nat. Commun.* 12.1 (Aug. 2021), p. 5104. ISSN: 2041-1723. DOI: 10.1038/s41467-021-25342-8.

[172] Carl P. Goodrich et al. "Designing self-assembling kinetics with differentiable statistical physics models". In: *Proc. Natl. Acad. Sci.* 118.10 (2021), e2024083118. DOI: 10.1073/pnas.2024083118.

[173] Li Li et al. "Kohn-Sham Equations as Regularizer: Building Prior Knowledge into Machine-Learned Physics". In: *Phys. Rev. Lett.* 126 (3 Jan. 2021), p. 036401. DOI: 10.1103/PhysRevLett.126.036401.

[174] Peter Wirnsberger et al. "Normalizing flows for atomic solids". In: *Mach. Learn. Sci. Technol.* 3.2 (May 2022), p. 025009. DOI: 10.1088/2632-2153/ac6b16.

[175] Sebastian Bichelmaier, Jesús Carrete, and Georg K. H. Madsen. "Evaluating the efficiency of power-series expansions as model potentials for finite-temperature atomistic calculations". In: *Int. J. Quantum Chem.* (2023). DOI: 10.1002/QUA.27095.

Publication I

Accurate First-Principles Treatment of the High-Temperature Cubic Phase of Hafnia

Sebastian Bichelmaier, Jesús Carrete, Michael Nelhiebel, and Georg K. H. Madsen*

HfO_2 is an important high-k dielectric and ferroelectric, exhibiting a complex potential energy landscape with several phases close in energy. It is, however, a strongly anharmonic solid, and thus describing its temperature-dependent behavior is methodologically challenging. An approach based on self-consistent, effective harmonic potentials (EHP) to study the potential energy surface (PES) of anharmonic materials is proposed. The introduction of a reweighting procedure enables the usage of unregularized regression methods by efficiently utilizing the information contained in every data point obtained from density functional theory. The approach is detailed and tested on the example of the high-temperature cubic phase of HfO_2. It is demonstrated how the correction term for the deviation between the EHP and the true PES can be calculated directly from the same sampling used for determining the EHP. The calculated temperature-dependent physical properties are in agreement with existing experimental data, thereby opening for the predictive treatment of HfO_2 over a wide temperature range.

1. Introduction

One of the biggest drawbacks of density functional theory (DFT) calculations is the lack of temperature-induced effects. However, due to the exponential growth in computing power and continuous methodological developments, the previously prohibitively expensive calculations necessary to remedy that shortcoming are becoming viable for the investigation of new materials. Consequently, the inclusion of temperature is a prominent theme in many current computational efforts.[1–8]

In the present study we focus on hafnia, HfO_2, which has a multifaceted phase diagram[9] and numerous industrially relevant applications, ranging from a high-κ gate dielectric for semiconductors in its amorphous[10] and (more recently suggested) tetragonal phase,[11] to a ferroelectric in its orthorhombic states[12] for e. g. nonvolatile memory applications.[13]

An accurate DFT treatment of its temperature-dependent behavior has proven difficult to achieve in previous theoretical efforts.[14] In the simplest approach, the effect of temperature is included by means of the harmonic approximation (HA), where the second-order interatomic force constants (IFCs) are obtained by applying small displacements and mapping the corresponding forces induced by them. However, in the case of structures governed by anharmonic potential energy surfaces (PES), the HA might yield imaginary frequencies, thus indicating mechanical instability, even when experiments confirm the existence of those structures.

Ab initio molecular dynamics (AIMD) approaches[15] can, in principle, treat such temperature-stabilized structures, but obtaining the free energy of reasonably complex systems through thermodynamic integration proves to be a resource-intensive task and quickly becomes intractable. Moreover, AIMD treats the nuclear motion in a completely classical fashion, and therefore cannot capture effects such as zero-point motion, which can be relevant, e.g., accurately describing the vibrations of light and strongly bonded atoms.

An emerging category of alternatives to AIMD can be labeled as effective harmonic potentials (EHP). The idea goes back to 1955[16] and in essence, involves determining the best HA for the part of the PES which dominates nuclear motion. Temperature-dependent contributions to the free energy are then included using independent quantum harmonic oscillators based on these EHPs. Especially at lower temperatures, EHPs have proven to be a rich starting point for understanding temperature-dependent behavior using ab initio methods and have prompted various implementations and formulations of the underlying theory.[3,6,7,17–20] The various implementations differ in how the PES is sampled and how the HA is determined. The options include sampling along the HA eigenvectors, stochastic sampling, and molecular dynamics trajectories.

The term accounting for the deviation between the EHP and the true PES is likely the least well-determined in an actual EHP implementation. At lower temperatures, it can often be neglected but it is a key quantity at higher temperatures.[20] In the present study, we focus on the high-temperature cubic ($Fm\bar{3}m$) hafnia phase (c-HfO_2), which is an example of a temperature-stabilized

S. Bichelmaier, J. Carrete, G. K. H. Madsen
Institute of Materials Chemistry
TU Wien
A-1060 Vienna, Austria
E-mail: georg.madsen@tuwien.ac.at

M. Nelhiebel
KAI GmbH
Europastrasse 8, A-9524 Villach, Austria

The ORCID identification number(s) for the author(s) of this article can be found under https://doi.org/10.1002/pssr.202100642.

structure where the small displacement HA yields imaginary frequencies.[14] We show how the use of reweighting in combination with unregularized regression can be employed to obtain the temperature-dependent EHP. Special attention is given to the deviation between the EHP and the true DFT PES. We demonstrate how this correction term can be calculated directly from the same sampling used for determining the trial EHP. We find a good agreement with available AIMD calculations and experiments, thereby opening for the predictive treatment of HfO_2 over a wide temperature range.

2. Method

2.1. Background

A system described by a Hamiltonian $\widehat{H}$ is in a state of thermal equilibrium at constant volume, temperature T, and number of particles when its free energy

$$F[\widehat{\rho}_0] = \mathrm{Tr}(\widehat{\rho}_0\widehat{H}) + k_B T \mathrm{Tr}(\widehat{\rho}_0 \log \widehat{\rho}_0) \tag{1}$$

is at a minimum. This equilibrium state is described by a particular quantum mechanical density matrix, $\widehat{\rho}_0$, which, were if known, would provide access to the whole thermodynamics of the system. However, it is impossible to solve this problem exactly for all but trivial model systems.

The EHP can be formulated as a variational problem,[20,21] where a trial density matrix, $\widehat{\rho}$, which exactly describes the statistics of a corresponding trial Hamiltonian, $\widehat{\mathscr{H}}$, is introduced. $\widehat{\mathscr{H}}$ differs from the true Hamiltonian $\widehat{H}$ only in the form of the approximate potential energy operator, $\widehat{\mathcal{V}}$, as opposed to the exact potential energy operator, $\widehat{V}$. Minimizing the free energy with respect to the trial density matrix is guaranteed by the Gibbs–Bogoliubov inequality[22] to provide an upper bound on the free energy[20,21]

$$F[\widehat{V},\widehat{\rho}_0] \leq \mathscr{F}_{\mathrm{EHP}} = F[\widehat{\mathcal{V}},\widehat{\rho}] + \mathrm{Tr}[\widehat{\rho}(\widehat{V}-\widehat{\mathcal{V}})] = F_{\mathrm{harm}} + F_{\mathrm{corr}} \tag{2}$$

In the HA, the trial potential is parameterized as

$$\mathcal{V}(\mathbf{u}) = \frac{1}{2}\mathbf{u}^T\mathbf{\Phi}\mathbf{u} \tag{3}$$

in terms of the mass-weighted displacements, $\mathbf{u} = \mathbf{M}^{-1/2}(\mathbf{r}-\mathbf{r}_0)$, with ion masses $\mathbf{M}$, from the minimum-energy configuration and the second-order force constants, $\mathbf{\Phi}$, with eigenvalues and eigenvectors $\omega_\lambda^2\,\mathbf{\varepsilon}_\lambda$, so that

$$\mathbf{\Phi} = \sum_\lambda \omega_\lambda^2 \mathbf{\varepsilon}_\lambda \otimes \mathbf{\varepsilon}_\lambda^* \tag{4}$$

Within the HA, the projection onto real space of the trial density matrix can be expressed in closed form

$$\rho(\mathbf{u}) = \frac{1}{\sqrt{(2\pi)^{3N}|\mathbf{C}|}}\exp\left(-\frac{1}{2}\mathbf{u}\mathbf{C}^{-1}\mathbf{u}\right) \tag{5}$$

The covariance matrix $\mathbf{C}$ can be obtained from the aforementioned ω_λ and $\mathbf{\varepsilon}_\lambda$[23]

$$\mathbf{C} = \frac{\hbar}{2\sqrt{M_i M_j}}\sum_\lambda \frac{1}{\omega_\lambda \tanh\frac{\hbar\omega_\lambda}{2k_B T}}\mathbf{\varepsilon}_\lambda \otimes \mathbf{\varepsilon}_\lambda^* \tag{6}$$

Likewise, the expression for the harmonic contribution to the free energy, F_{harm}, is given by[23]

$$F_{\mathrm{harm}}(T) = \sum_\lambda \left(\frac{\hbar\omega_\lambda}{2} + k_B T \log\left[1 - \exp{-\frac{\hbar\omega_\lambda}{k_B T}}\right]\right) \tag{7}$$

F_{harm} depends directly on the temperature T and indirectly on the harmonic trial potential through ω_λ [Equation (4)]. The optimal trial potential thus depends on the temperature. Ignoring F_{corr} [Equation (2)] and the temperature dependence of the effective potential results in the well-known quasi-HA.

2.2. Temperature-Dependent Effective Potentials

We implement the search for the optimal EHP by approximating the real-space density matrix by means of canonical importance sampling and treating the interdependence of $\mathscr{F}_{\mathrm{EHP}}$ and Φ as a self-consistent problem. When self-consistency is reached, this corresponds to minimizing $\mathscr{F}_{\mathrm{EHP}}$.[7]

The starting point is the second-order force constants and corresponding potential, $\mathcal{V}^{(1)}$, obtained through small displacements. From the eigenvalues and eigenvectors and the temperature of interest the associated trial density, $\rho^{(1)}$, is obtained through Equation (5) and (6). We replace the imaginary square roots of possible negative eigenvalues from intermediate steps with their modulus. From this probability density, the first set of displacements, $\mathcal{S}^{(1)}$, is drawn. A new EHP is obtained by calculating the potential energies and forces corresponding to the displacements using DFT and finding the parametrization of the force constants in Equation (3), which best represent the relationship between forces and displacements, as will be discussed in the following. The iterative process then progresses by constructing a new density matrix using Equation (6) and (5). To aid convergence, the new trial density matrix, $\rho^{(k)}$, is obtained through a Pulay mixing scheme[24] with a memory of $n = 5$ steps and a mixing parameter of $\alpha = 0.1$, as commonly used. A new set of displacements is now drawn and the process continues until convergence is reached.

To efficiently use all the data obtained from DFT, the reweighting factor[17] is introduced

$$w_m^{(g\to k)} = \frac{\rho^{(k)}(\mathbf{u}_m^{(g)})}{\rho^{(g)}(\mathbf{u}_m^{(g)})} \tag{8}$$

Thereby displacement vectors, $\mathbf{u}_m^{(g)}$, belonging to a set drawn in a previous iteration, $\mathcal{S}^{(g)}$, and their corresponding forces and potential energies can be included as if they belong to the current set, $\mathcal{S}^{(k)}$. Using the reweighting factors significantly increases the amount of available data and allows using an unregularized fitting procedure to obtain the force constant matrix. To simplify the notation, we introduce the stiffness tensor, whose elements are defined as

$$\Psi_{ij}^{\alpha\beta} = \sqrt{M_i M_j}\,\Phi_{ij}^{\alpha\beta} \tag{9}$$

The trial potential for iteration k is found by finding the force constants which minimize the weighted sum of the least-squares deviations from the calculated forces $\mathbf{f}_m$, i. e.

$$\sum_g \sum_m w_m^{(g\to k)} \| \mathbf{f}_m^{(g)} + \Psi \mathbf{u}_m^{(g)} \mathrm{diag}(\mathbf{M})^{-\frac{1}{2}} \|_2^2 \tag{10}$$

In **Figure 1**, the free energy evaluated according to Equation (7) is shown as a function of the iterations until convergence for the 0 K DFT-relaxed structure using a temperature of $T = 2500$ K. Typically the convergence criterion of $\|\mathscr{F}_{\mathrm{EHP}}^{(g-1)} - \mathscr{F}_{\mathrm{EHP}}^{(g)}\| < 2.5\,\mathrm{meV\,fu^{-1}}$ for three successive iterations is reached in $10-15$ iterations when five structures are added per iteration for the initial temperature point, totaling 50–75 DFT runs.

Alternatively, a loss function involving a penalty

$$\left(\sum_g \sum_m w_m^{(g\to k)} \| \mathbf{f}_m^{(g)} + \Psi \mathbf{u}_m^{(g)} \mathrm{diag}(\mathbf{M})^{-\frac{1}{2}} \|_2^2 + \alpha \|\Psi\|_1 \right) \tag{11}$$

can be defined. This is known as the LASSO minimization target function. The α parameter determines the strength of the L_1 regularization and promotes sparsity of the force constant matrix. Regularization is common in machine learning tasks and can be a computational advantage.[25] However, there is no a priori reason that Ψ should be particularly sparse. Thus, a process called fivefold cross-validation [26], during which the available data is split into five complementary subsets, is performed at every step of the iteration. For each value of α in a predefined range, Equation (11) is then minimized for every unique choice of four of those subsets. The resulting models are then evaluated on the corresponding set left out in the minimization. Through this procedure, the regularization strength can be tuned to provide the model which generalizes best, i.e. has the best performance on said left-out set.

In Figure 1, we show the convergence behavior at $T = 2500$ K for the 0 K DFT-relaxed structure. Notably, the α parameter quickly approaches zero as the calculation progresses. This is understandable as regularization is typically applied when fitting

samples that insufficiently cover the sample space. Thus, as the amount of data points available for fitting increases, every fold in the cross-validation procedure will be more and more equally representative of the rest, so the L_1 penalty will actually hinder minimizing Equation (11) and will be forced toward zero by the algorithm itself, effectively resulting in Equation (10).

We choose the unregularized least-squares approach [Equation (10)] for two reasons: We are, for all but the first few iterations, confronted with an overdetermined system, as only a total of 52 independent force constants remain after considering symmetry and the cut-off and every sample provides 576 force–displacement pairs. Furthermore, if a nonzero L_1-penalty was indeed used throughout the calculation, the force constants obtained would not fulfill the property of minimizing the free energy once self-consistency is reached.[7] As argued previously, the regularization parameter must approach zero, because the coverage increases with every iteration. While we would presumably not have arrived at an artificially increased free energy, using LASSO only provides a minor speed up, while introducing additional uncertainty in the results.

The correction term, F_{corr}, has previously been calculated by representing the DFT PES in a simpler form[17,20] or by using the trajectory obtained from an AIMD run.[6,27] We calculate F_{corr} directly from the DFT potential energies obtained from the same sampling as used for determining the trial EHP, Equation (10), as a weighted average

$$F_{\mathrm{corr}} = \frac{1}{W} \sum_g \sum_m w_m^{(g\to k)} \left[V(\mathbf{u}_m^{(g)}) - \mathcal{V}^{(k)}(\mathbf{u}_m^{(g)}) \right] \tag{12}$$

where W is the sum of all the weights. Similar to the EHP the reweighting allows using all DFT calculations to obtain F_{corr} and Figure 1 illustrates that the convergence is also comparable, meaning that convergence of $\mathscr{F}_{\mathrm{EHP}}$ is reached within 10–15 iterations.

It is straightforward to extend the formalism described earlier to reuse samples drawn at temperature T_1 for a different temperature T_2, by building reweighting factors accounting for this. This provides a fast way to calculate force constants at temperatures near T_1. At temperatures more different from T_1 the sample set might not be adequate anymore, necessitating augmentation by additional DFT runs. As an example, we can
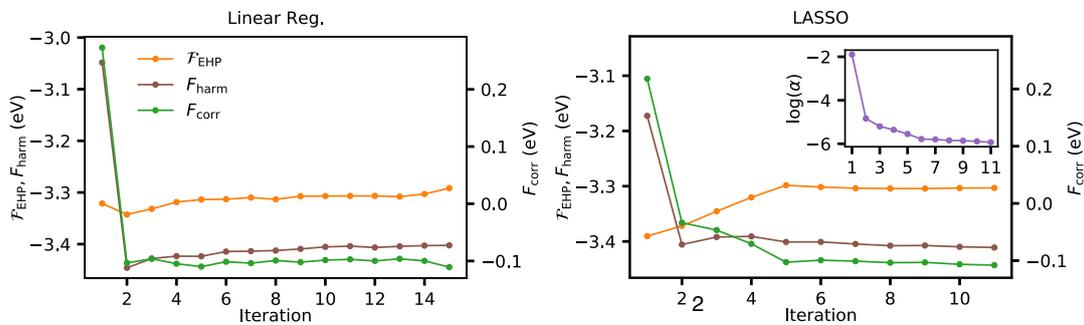

Figure 1. Convergence of the free energy contributions with iterations for the exemplary case of the 0 K density functional theory (DFT)-relaxed structure, $V_0 = 32.62\,\text{Å}^3$, at $T = 2500$ K. The upper plot is for the least-squares penalty function, Equation (10) and the lower for LASSO, Equation (11). The total free energy, $\mathscr{F}_{\mathrm{EHP}}$, and the harmonic contribution, F_{harm}, are plotted according to the scale on the left-hand side y-axis and F_{corr} with respect to the y-axis to the right. In every iteration, five new structures are added. In the inset, the decreasing optimal regularization strength, α, is shown.

mention that using the displacements and forces obtained for $T_1 = 2500$ K, shown in Figure 1, at $T_2 = 2100$ K, but reweighted according to Equation (8) results in convergence after adding only two additional iterations. Once convergence has been obtained for a mesh of temperatures, free energies can be obtained at intermediate temperatures without additional DFT calculations. As a measure of when those additional calculations are necessary, we use the effective number of samples,[28,29]

$$w_{\text{eff}}^{(k)} = \frac{\left[\sum_g \sum_m w_m^{(g \rightarrow k)}\right]^2}{\sum_g \sum_m \left[w_m^{(g \rightarrow k)}\right]^2} \quad (13)$$

which has also been widely used in other sampling strategies.[30,31] As can be observed in **Figure 2**, the number of effective samples is indeed a useful metric for the completeness of the data at a given temperature. However, it would be inefficient to augment the data if the poorly sampled regions are small and the surrounding points are described by the existing samples well enough, as is the case in the situation shown in the inset in Figure 2. To prevent this, we apply a smoothing spline, as implemented in ref. [26] weighted with the effective samples to $\mathscr{F}_{\text{EHP}}(T)$. This ensures smoothness of the free energy as a function of temperature, by a weighted interpolation according to the effective samples and hence avoids artifactual oscillations in later results.

2.3. Computational Details

We use the plane-augmented-wave (PAW) formalism[32] in the frozen-core implementation of VASP[33] with the valence configurations 5p 6s 5d for Hf and s2p4 for O and the recommended cutoff of 600 eV. The force calculations for the phonons are performed on a supercell, consisting of a $4 \times 4 \times 4$ repetition of the primitive cubic unit cell, using just the Γ-point. We use phonopy[34] with the nonanalytical correction described in ref. [35] for obtaining the initial small displacements force constants. The fitting of the force constants [Equation (10) and (11)] is done
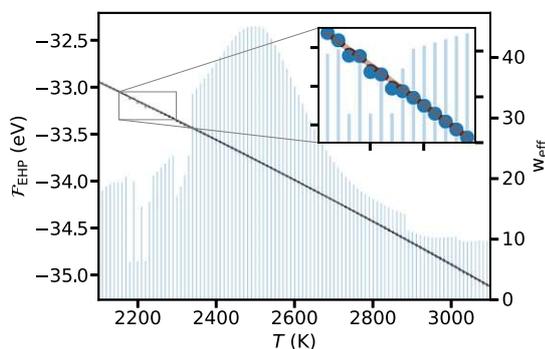
using scikit-learn[26] and the cluster formalism established in ref. [36] given a pre-defined cutoff $r_{\text{cut}} = 7$ Å.

We perform the steps outlined earlier for various deformations of the 0 K DFT-relaxed structure; we include volumes from $V = 30.70$ to $37.77 \text{Å}^3 \text{fu}^{-1}$, as well as various temperatures ranging from 2100 to 3100 K. To arrive at a simple but general analytical expression we then fit the free energies using

$$\mathscr{F}_{\text{EHP}}(V; T) = c_0 + \frac{c_1}{V^{1/3}} + \frac{c_2}{V^{2/3}} + \frac{c_3}{V} \quad (14)$$

achieving an average deviation between the actual and fitted free energy of less than 7 meV fu^{-1}, which corresponds to less than 3 meV atom^{-1}. This allows us to find the equilibrium lattice parameter and volume at every temperature point of interest with a high accuracy.

3. Results and Discussion

We settle on a temperature range from 2100 to 3100 K to ensure full coverage of the stable range of $\text{c} - \text{HfO}_2$, 2800–3100 K,[37] with the upper bound close to the melting temperature. To accurately treat this temperature range, we perform an initial self-consistent run at $T_1 = 2500$ K, augment it with samples at $T_2 = 2100$ K and, guided by the effective sample size, Equation (13), includes $T_3 = 3000$ K for some deformations.

As an example, we show the phonon band structure of the 0 K DFT-relaxed structure in **Figure 3**, as obtained using small displacements and using the EHP at elevated temperatures. As can be seen $\text{c} - \text{HfO}_2$ shows instability in the small-displacements (0 K) phonon spectrum at $X = (0, 1/2, 1/2)$ in the Brillouin zone, which in, the structurally similar, ZrO_2 has been linked to the cubic-to-tetragonal phase transition.[38] Within the studied temperature range, we see a continuous hardening (shown in the inset) of said mode indicating that temperature-induced anharmonic effects are stabilizing this phase. For the additional volumes that are studied (-6 to 16% volume changes), we find a similar behavior and can report stable phonon spectra for the whole volume and temperature range.
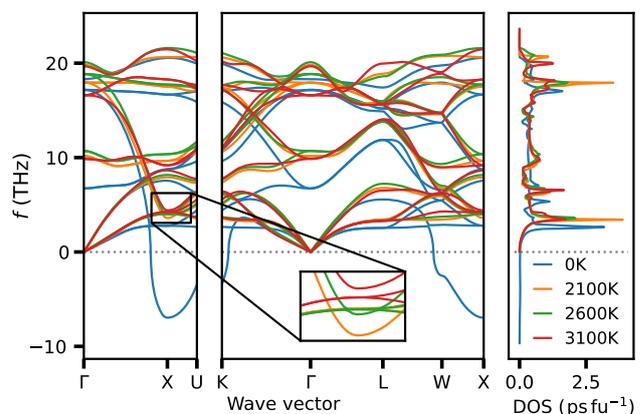

Figure 2. Free energy and effective samples as a function of temperature at $+2\%$ deformation, i.e., a volume of 34.62 Å^3. In regions where the effective samples, as shown by the bars in the background, are low, the free energy shows discontinuous behavior. The brown line is a smoothing spline using the effective samples as weights, whereas the blue points are the data points.

Figure 3. Phonon band structure and density of states of $\text{c} - \text{HfO}_2$ of the 0 K DFT-relaxed structure. The soft mode at 0 K indicated by the negative frequencies at X in blue continuously hardens as temperature increases and the structure becomes stable.
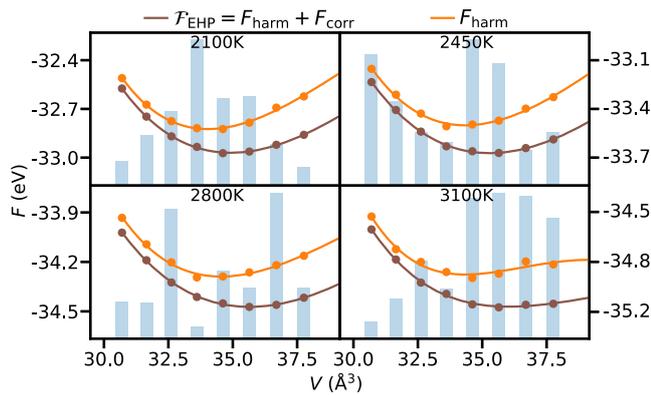
Figure 4. Comparison of the free energies obtained for various volumes and temperatures with and without the correction term in Equation (2). The solid lines are the fitted SJEOS as described in Equation (14), while the bars in the background are representing the effective sample size from Equation (13). The correction shifts the free energy downwards and towards larger volumes.
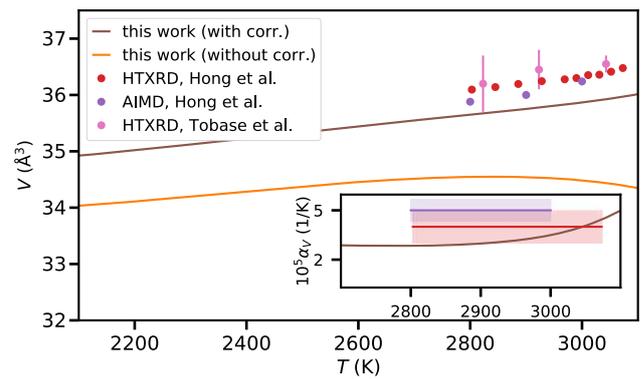

Figure 5. The unit-cell volume expansion as a function of temperature calculated with (without) the correction term plotted in brown (orange). The results are compared to experimental data[9,39] and AIMD calculations.[39] The thermal expansion coefficient, $\alpha_V = \frac{1}{V} \left(\frac{\partial V}{\partial T} \right)_P$, is shown in the inset. The shaded area in the inset indicates the experimental uncertainty.

The absence of imaginary phonon frequencies makes it possible to calculate the vibrational contribution to the free energy according to Equation (7). **Figure 4** depicts the volume dependence of the free energy at four different temperatures. Equation (14) captures the behavior of $c - HfO_2$ across the studied temperature range and as expected the equilibrium volume increases with temperature. Interestingly, F_{corr} not only shifts the curve to lower energies, $\approx 100\,\text{meV fu}^{-1}$, but it also changes the positions of the minima. As expected, F_{corr} gets larger with temperature, i.e., with increasing anharmonic contributions to the relevant parts of the PES. As a result, the contribution generally favors larger volumes and can be expected to be important for a correct prediction of thermal expansion.

The role of the correction term becomes even more apparent when looking at the thermal expansion in Equation (5). Comparing the unit-cell volume of $c - HfO_2$ with experimental and theoretical results from the literature[9,39] illustrates how neglecting F_{corr} will provide underestimated unit-cell volume in situations where anharmonicity contributes a significant portion of the total energy. Such situations can arise when describing materials that are inherently anharmonic, or generally for materials at elevated temperatures. HfO_2 is in our study subject to both of these circumstances and thus any description not taking the anharmonic correction into account is bound to lead to inaccurate conclusions.

When F_{corr} is included, the offset in the thermal expansion coefficient is reduced to about $0.5\,\text{Å}^3$, or 1.2%, **Figure 5**. The remaining offset can be due to limiting the trial density matrices, ρ, in Equation (2) to those which can result from a harmonic model Equation (3). However, as the offset is constant with temperature it is likely not the result of the neglected higher order force constants as this would entail additional temperature dependence. We argue instead that it can be attributed to inherent approximations of the chosen DFT functional. It is worth noting that the AIMD study reported in ref. [39] (employing the same PBE functional) also finds slightly lower volumes than the experiment. The calculated thermal expansion coefficient [see inset in Figure 5] is approximately constant, $3.3 \times 10^{-5}\,\text{K}^{-1}$, until an

increase in thermal expansion can be observed as the melting temperature, $T_m \approx 3100\,\text{K}$, is approached. The thermal expansion coefficient is not impacted by a small constant offset in the volume, and our results are within or very close to the uncertainty, indicated by a red bar, of the experimental average over the range from 2800 to 3100 K, $4(1) \times 10^{-5}\,\text{K}^{-1}$, obtained by Hong. et al.[39] The volume data from Tobase et al.[9] would result in $\alpha_V = 4.39 \times 10^{-5}\,\text{K}^{-1}$ over the region of ≈ 2823–3043 K. However, due to the large uncertainty in the volume measurements, the error bar would span about $8.3 \times 10^{-5}\,\text{K}^{-1}$, and hence it is not shown in the graph.

Finally, we report the temperature dependence of the $c - HfO_2$ bulk modulus. Due to the fact that the second derivative of Equation (14) with respect to the volume is not linear, the bulk modulus obtained by our method is naturally dependent on the volume at a given pressure at which it is evaluated. Evaluated at $0\,\text{GPa}$ we find a bulk modulus of $B_0 = 180\,\text{GPa}$, which decreases to $120\,\text{GPa}$ over the range of 2100 to 3100 K. This drastic softening is expected as we are approaching the melting point of the material. In their recent study, Irshad et al.[40] measured the bulk modulus of pressure-stabilized, nanocrystalline $c - HfO_2$ at ambient temperature, finding $B_0 = 242(16)\,\text{GPa}$. Applying a small pressure of $4\,\text{GPa}$ (corresponding to a volume change of 1.5%) to our result, yields a comparable bulk modulus of about $210\,\text{GPa}$ at 2100 K, decreasing to $170\,\text{GPa}$ at 3100 K.

4. Conclusion

The behavior of $c - HfO_2$ in the high-temperature regime was studied using EHP. At elevated temperatures, the unstable mode exhibited by the cubic structure hardens, resulting in a stable phonon spectrum. It was shown that, without consideration of the anharmonic correction term [Equation (2)], an accurate description of this phase is not possible and it is conjectured that this term is crucial throughout a broad spectrum of high-temperature materials studies. The thermal expansion behavior reported, $\alpha_V = 3.3 \times 10^{-5}\,\text{K}^{-1}$, is in good agreement with the existing experimental and theoretical data, if averaged over

the same temperature range. In the range of 2100 to 3100 K, the bulk modulus of $c - HfO_2$ exhibits a softening from 180 to 120 GPa.

The lack of quantum effects in molecular dynamics and the inadequacy of a purely harmonic treatment at high temperatures have often made it necessary to mix the two methods for a predictive treatment of phase diagrams over a wide temperature range (see ref. [41] for a recent study of HfO_2). The present results establish EHP as a viable alternative to molecular dynamics at high temperatures, thereby presenting the opportunity to use them to study temperature-induced effects over a broad temperature range.

Machine learned force fields have recently appeared as a viable surrogate for DFT in the research of HfO_2.[42,43] With the recent improvements to the accuracy of the predicted forces[44] of these force fields, it is conceivable that they could be used for obtaining EHP, thereby enabling large-scale studies of phase space.

Acknowledgements

Conflict of Interest

The authors declare no conflict of interest.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request. The code employed in this study is available on the BitBucket repository https://bitbucket.org/sbichelm/scphonons/src/master/.

Keywords

[1] J. Xie, S. de Gironcoli, S. Baroni, M. Scheffler, *Phys. Rev. B* **1999**, *59*, 970.

[2] B. Grabowski, T. Hickel, J. Neugebauer, *Phys. Rev. B* **2007**, *76*, 024309.

[3] P. Souvatzis, O. Eriksson, M. I. Katsnelson, S. P. Rudin, *Phys. Rev. Lett.* **2008**, *100*, 095901.

[4] B. Grabowski, L. Ismer, T. Hickel, J. Neugebauer, *Phys. Rev. B* **2009**, *79*, 134106.

[5] O. Hellman, I. A. Abrikosov, S. I. Simak, *Phys. Rev. B* **2011**, *84*, 180301.

[6] O. Hellman, P. Steneteg, I. A. Abrikosov, S. I. Simak, *Phys. Rev. B* **2013**, *87*, 104111.

[7] A. van Roekeghem, J. Carrete, N. Mingo, *Comput. Phys. Commun.* **2021**, *263*, 107945.

[8] S. Ehsan, M. Arrigoni, G. K. H. Madsen, P. Blaha, A. Tröster, *Phys. Rev. B* **2021**, *103*, 094108.

[9] T. Tobase, A. Yoshiasa, H. Arima, K. Sugiyama, O. Ohtaka, T. Nakatani, K.-i. Funakoshi, S. Kohara, *Phys. Status Solidi B* **2018**, *255*, 1800090.

[10] G. D. Wilk, R. M. Wallace, J. M. Anthony, *J. Appl. Phys.* **2001**, *89*, 5243.

[11] D. Fischer, A. Kersch, *J. Appl. Phys.* **2008**, *104*, 084104.

[12] T. S. Böscke, J. Müller, D. Bräuhaus, U. Schröder, U. Böttger, *Appl. Phys. Lett.* **2011**, *99*, 102903.

[13] J. Müller, P. Polakowski, S. Mueller, T. Mikolajick, *ECS J. Solid State Sci.* **2015**, *4*, N30.

[14] T. D. Huan, V. Sharma, G. A. Rossetti, R. Ramprasad, *Phys. Rev. B* **2014**, *90*, 064111.

[15] R. Iftimie, P. Minary, M. E. Tuckerman, *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 6654.

[16] D. Hooton, *Lond. Edinb. Dublin Philos. Mag. J. Sci.* **1955**, *46*, 422.

[17] I. Errea, M. Calandra, F. Mauri, *Phys. Rev. Lett.* **2013**, *111*, 177002.

[18] T. Tadano, S. Tsuneyuki, *Phys. Rev. B* **2015**, *92*, 054301.

[19] R. Stern, G. K. H. Madsen, *Phys. Rev. B* **2016**, *94*, 144304.

[20] L. Monacelli, R. Bianco, M. Cherubini, M. Calandra, I. Errea, F. Mauri, *J. Phys. Condens. Matter* **2021**, *33*, 363001.

[21] I. Errea, M. Calandra, F. Mauri, *Phys. Rev. B* **2014**, *89*, 064302.

[22] A. Isihara, *J. Phys. A* **1968**, *1*, 539.

[23] L. E. Reichl, in *A Modern Course in Statistical Physics*, 1st ed., Arnold, London **1980**.

[24] P. Pulay, *J. Comput. Chem.* **1982**, *3*, 556.

[25] E. Fransson, F. Eriksson, P. Erhart, *NPJ Comput. Mater.* **2020**, *6*, 135.

[26] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, E. Duchesnay, *J. Mach. Learn. Res.* **2011**, *12*, 2825.

[27] E. Metsanurk, M. Klintenberg, *Phys. Rev. B* **2019**, *99*, 184304.

[28] L. Kish, *Am. Sociol. Rev.* **1965**, *30*, 564.

[29] A. Kong, J. S. Liu, W. H. Wong, *J. Am. Stat. Assoc.* **1994**, *89*, 425.

[30] N. Hansen, arXiv:1604.00772, **2016**.

[31] M. Invernizzi, P. M. Piaggi, M. Parrinello, *Phys. Rev. X* **2020**, *10*, 041034.

[32] P. E. Blöchl, *Phys. Rev. B* **1994**, *50*, 17953.

[33] G. Kresse, J. Furthmüller, *Phys. Rev. B* **1996**, *54*, 11169.

[34] A. Togo, I. Tanaka, *Scr. Mater.* **2015**, *108*, 1.

[35] Y. Wang, J. J. Wang, W. Y. Wang, Z. G. Mei, S. L. Shang, L. Q. Chen, Z. K. Liu, *J. Phys. Condens. Matter* **2010**, *22*, 202201.

[36] F. Eriksson, E. Fransson, P. Erhart, *Adv. Theory Simul.* **2019**, *2*, 1800184.

[37] C. Wang, M. Zinkevich, F. Aldinger, *J. Am. Chem. Soc.* **2006**, *89*, 3751.

[38] A. Kuwabara, T. Tohei, T. Yamamoto, I. Tanaka, *Phys. Rev. B* **2005**, *71*, 064301.

[39] Q.-J. Hong, S. V. Ushakov, D. Kapush, C. J. Benmore, R. J. K. Weber, A. van de Walle, A. Navrotsky, *Sci. Rep.* **2018**, *8*, 14962.

[40] K. A. Irshad, V. Srihari, D. S. Kumar, K. Ananthasivan, H. Jena, *J. Am. Chem. Soc.* **2020**, *103*, 5374.

[41] J. J. Low, N. H. Paulson, M. D'Mello, M. Stan, *Calphad* **2021**, *72*, 102210.

[42] G. Sivaraman, A. N. Krishnamoorthy, M. Baur, C. Holm, M. Stan, G. Csányi, C. Benmore, Á. Vázquez-Mayagoitia, *NPJ Comput. Mater.* **2020**, *6*, 104.

[43] J. Wu, Y. Zhang, L. Zhang, S. Liu, *Phys. Rev. B* **2021**, *103*, 024108.

[44] H. Montes-Campos, J. Carrete, S. Bichelmaier, L. M. Varela, G. K. H. Madsen, *J. Chem. Inf. Model.* **2022**, *62*, 88.

Publication II

A neural-network-backed effective harmonic potential study of the ambient pressure phases of hafnia

Sebastian Bichelmaier,[1, 2] Jesús Carrete,[1] Ralf Wanzenböck,[1] Florian Buchner,[1] and Georg K. H. Madsen[1, *]

[1]*Institute of Materials Chemistry, TU Wien, A-1060 Vienna, Austria*
[2]*KAI GmbH, Europastrasse 8, A-9524 Villach, Austria*
(Dated: March 9, 2023)

Phonon-based approaches and molecular dynamics are widely established methods for gaining access to a temperature-dependent description of material properties. However, when a compound's phase space is vast, density-functional-theory-backed studies quickly reach prohibitive levels of computational expense. Here, we explore the complex phase structure of HfO_2 using effective harmonic potentials based on a neural-network force field (NNFF) as a surrogate model. We detail the data acquisition and training strategy that enable the NNFF to provide almost ab-initio accuracy at a significantly reduced cost and present a recipe for automation. We demonstrate how the NNFF can generalize beyond its training data and that it is transferable between several phases of hafnia. We find that the thermal expansion of the low-symmetry phases agrees well with experimental results and we determine the $P\bar{4}3m$ phase to be the favorable (stoichiometric) cubic phase over the established $Fm\bar{3}m$. In contrast, the experimental lattice constants of the cubic phases are substantially larger than what is calculated for the corresponding stoichiometric phases. Furthermore, we show that the stoichiometric cubic phases are unlikely to be thermodynamically stable compared to the tetragonal and monoclinic phases, and hypothesize that they only exist in defect-stabilized forms.

I. INTRODUCTION

The intricate link between microscopic crystal structures and macroscopic materials properties necessitates understanding the former to predict the latter. Temperature is an important contributor to phase stability. However, computational studies of its impact generally require extensive sampling and with that a large number of energy and force evaluations. Hence, temperature dependence has been notoriously difficult to combine with accurate computational methods like density functional theory (DFT) [1].

Effective harmonic potentials (EHPs) are a widely used method to include temperature dependency into ab-initio studies [2–8]. An EHP is constructed by sampling the potential energy surface (PES) using temperature-dependent displacement distribution functions and fitting effective force constants. Performing the described procedure with a DFT backend is computationally demanding, especially for low-symmetry structures. Thus, while the method gives access to structure- and pressure-dependent free energies, its application to the calculation of phase transitions has been somewhat limited [6]. It would be an obvious advantage if the sampling of the PES inherent to the EHP could be used to simultaneously train a surrogate model.

Recently, machine-learning (ML) force fields (FFs) are becoming increasingly widespread in the field of computational materials science [9–11], with applications ranging from an accurate description of the martensitic phase transition in the memory alloy NiTi [12] or grain boundaries in copper systems [13] to structural phase transitions [14] and surface reconstructions [15] in $SrTiO_3$. Given sufficient training, MLFFs can replace DFT calculations at a vastly reduced cost, and combining them with EHPs can lead to significantly lower computational demands.

Hafnia (HfO_2) has recently attracted particular research interest [8, 16–21]. Similar to the isostructural material zirconia (ZrO_2) it boasts a rich phase diagram with various internal (doping, alloying) and external (strain, temperature) factors contributing to phase stability and it is thus an ideal use case for MLFFs [22, 23]. At ambient pressure and temperatures up to approximately $2050\,K$ the material presents a monoclinic ($P2_1c$) crystal structure (m-phase), which, as temperature increases, undergoes a first-order transition into a tetragonal ($P4_2nmc$) one (t-phase) [24, 25]. Furthermore, hafnia is believed to transition to a high-symmetry cubic ($Fm\bar{3}m$) phase (cI-phase) in a narrow temperature window immediately below the melting point [26].

There is, however, still quite some ambiguity regarding the transition to a high-symmetry cubic phase. The reported t-to-c phase transition temperatures for HfO_2 span almost $1000\,K$ [25], and there is even some doubt regarding the existence of a pure cubic phase or its precise space group [27]. Indeed, some researchers find the lower-symmetry $P\bar{4}3m$ (cII-phase) phase to be energetically favorable [28], although to our knowledge no theoretical studies, that properly take temperature into account, exist. For HfO_2, few experimental studies describing a t-to-c phase transition exist, potentially due

* Correspondence email address: georg.madsen@tuwien.ac.at

2

to the extreme temperature requirements and the difficulties these conditions present for accurate experimental analysis. Nonetheless, some researchers find a mixed phase of cubic and tetragonal symmetry suggesting metastability [29], while others see a clear second-order transition [19]. It should be pointed out that similar ambiguities have been reported with respect to the existence of a stable stoichiometric cubic phase in the related ZrO_2 [30–34]. Thus, additional theoretical efforts are warranted, not only to clarify if a stable stoichiometric cubic phase exists, but also to shed light on its crystal structure.

In this work, we build an MLFF using a neural network (NNFF) with a methodology corresponding to an evolved version of the one presented in Ref. 35. We construct a NNFF that can describe the part of the potential energy landscape studied well, and then employ it as the backend calculator for an EHP-based treatment of HfO_2. We analyze the m-to-t phase transition and establish a case for $P\bar{4}3m$ being the more likely candidate for a high-T cubic phase in HfO_2 over the traditionally assumed $Fm\bar{3}m$. The full thermodynamical picture however suggests that neither of the cubic phases is completely stable in purely stoichiometric conditions, but may be stabilized by oxygen vacancies.

We furthermore discuss EHPs for training set generation. Generation of a proper training set for MLFFs is far from trivial [36]. Molecular dynamics (MD) trajectories typically consist of strongly correlated structures thus subsampling very long trajectories is required, resulting in hundreds of DFT calculations not used in training. Some authors combine this with active learning, using a subsampled trajectory as a baseline and an uncertainty metric, such as a committee of networks, to guide additional calculations [37–39]. Continuously retraining NNFFs with often only marginal augmentations to the data is computationally demanding and the accuracy of committee-based error estimates is still under debate [40]. Here, we show a training strategy based on EHPs, whereby the inherent physically meaningful and iterative sampling of the PES can be exploited. Thereby an NNFF is trained systematically on the cumulative data obtained up to each iteration, mimicking an active learning approach.

II. METHOD

A. Effective harmonic potentials

As presented in detail elsewhere [8], an effective harmonic potential $\mathcal{V}$ and its associated density matrix $\hat{\rho}$ can be obtained through a variational approach [2, 6, 41], resulting in a parametrization of the optimal harmonic potential in terms of the second-order force

constants, $\boldsymbol{\Phi}$:

$$\mathcal{V}(\mathbf{u}) = \frac{1}{2}\mathbf{u}^{\mathrm{T}}\boldsymbol{\Phi}\mathbf{u} \qquad (1)$$

out of which a corresponding density matrix

$$\rho(\mathbf{u}) = \frac{1}{\sqrt{(2\pi)^{3N}\,|\mathbf{C}|}}\exp\left(-\frac{1}{2}\mathbf{u}\mathbf{C}^{-1}\mathbf{u}\right), \qquad (2)$$

can be built. Here $\mathbf{u}$ refers to the mass-weighted displacements of the atoms, $\mathbf{u} = \sqrt{M_i}\left(\mathbf{R}_i - \mathbf{R}_{i,0}\right)$. The elements of the covariance matrix, C_{ij} are obtained as

$$C_{ij} = \frac{\hbar}{2\sqrt{M_i M_j}}\sum_\lambda \frac{1}{\omega_\lambda \tanh\frac{\hbar\omega_\lambda}{2k_B T}}\epsilon_{\lambda,i}\otimes\epsilon^*_{\lambda,j}, \qquad (3)$$

where ω_λ and $\boldsymbol{\epsilon}_\lambda$ are the eigenvalues and -vectors of the second-order force constants [Eq. (1)] and M_i is the mass of atom i. The free energy, $\mathcal{F} = F_{\mathrm{harm}} + F_{\mathrm{corr}}$, consisting of a harmonic contribution and an anharmonic correction term, can then be expressed as

$$F_{\mathrm{harm}}(T) = \sum_\lambda \left(\frac{\hbar\omega_\lambda}{2} + k_{\mathrm{B}}T\log\left[1 - \exp\left\{-\frac{\hbar\omega_\lambda}{k_{\mathrm{B}}T}\right\}\right]\right) \qquad (4)$$

$$F_{\mathrm{corr}}(T) = \frac{1}{N}\sum_n \left[V(\mathbf{u}_n) - \mathcal{V}(\mathbf{u}_n)\right], \qquad (5)$$

with $\mathbf{u}_n$ corresponding to samples drawn from the real-space distribution defined by the density matrix, Eq. (2), and V to the true energy as obtained from DFT.

B. Neural-network force field

1. Architecture and loss

We use a NeuralIL architecture [35], with a cutoff-radius of 5 Å, an embedding dimension of 4 and a total of 28 basis functions in a perceptron consisting of $128 : 64 : 32 : 16 : 16 : 16$ neurons. We employ a log-cosh based loss [42] for the forces

$$\mathcal{L}_f = \frac{0.1\,\mathrm{eV\mathring{A}}^{-1}}{n_{\mathrm{atoms}}}\sum_i^{n_{\mathrm{atoms}}}\log\left[\cosh\left(\frac{\sqrt{\frac{1}{3}\sum_\alpha \Delta f_{i,\alpha}^2}}{0.1\,\mathrm{eV\mathring{A}}^{-1}}\right)\right], \qquad (6)$$

where Δf is the difference between the actual force as obtained from DFT and the NNFF-predicted force, α indexes the three Cartesian axes and i the atoms.

While the forces in principle contain enough information to model the potential energy surface, including an

energy loss can enable more efficient learning of multiple local minima. Furthermore, once unit cell volume is included as a variable, a continuum of structures exhibiting zero forces, but with different energies, needs to be accurately described. Hence, an energy loss can be included in the overall loss

$$\mathcal{L}_E = \frac{0.1\,\text{eV}}{n_{\text{atoms}}} \log \left[\cosh \left(\frac{\Delta E}{0.1\,\text{eV}} \right) \right], \quad (7)$$

where ΔE corresponds to the error of the predicted energy as compared to the DFT result. The obvious approach of combining Eqs. (6) and (7) is to introduce a new hyperparameter, w_E, tuning the fraction of the energy loss included in the overall loss:

$$\mathcal{L}_{\text{tot,manual}} = (1 - w_E) \mathcal{L}_f + w_E \mathcal{L}_E. \quad (8)$$

However, as the loss measures the quality of the NNs approximation of the PES, the optimal choice of weight might change or depend on the composition or the dataset of the compound studied. Furthermore, manually tuning a hyperparameter is time- and resource-intensive.

Several approaches have been developed to overcome these obstacles. We implemented and tested the inverse certainty weighting [43] and the inverse Dirichlet weighting [44], but both resulted in instabilities during training, which manifested themselves as very large energy weights and essentially halted any progress. The most stable and efficient approach tested for the present datasets was Mitrevski's [45] implementation of the stochastic multi-gradient descent algorithm [46]. Here the gradients of Eqs. (6) and (7), $\nabla \mathcal{L}_f$ and $\nabla \mathcal{L}_E$, are corrected using the first and second momenta (*adamizing*, [45]), resulting in smoothed contributions of the force ($\tilde{\mathcal{L}}_f$) and energy loss ($\tilde{\mathcal{L}}_E$). Subsequently the common descent vector is calculated

$$\nabla \mathcal{L} = (1 - \omega_E) \nabla \tilde{\mathcal{L}}_f + \omega_E \nabla \tilde{\mathcal{L}}_E \quad (9)$$

where ω_E is obtained through a quadratic optimization

$$\min_{\omega_E} \left[\left\| \omega_E \nabla \tilde{\mathcal{L}}_E + (1 - \omega_E) \nabla \tilde{\mathcal{L}}_f \right\|^2 \quad 0 \le \omega_E \le 1 \right]. \quad (10)$$

The use of a common descent vector has been shown to result in a stable and efficient multi-objective optimization and enables exploration of the Pareto border [45]. To validate the method in the context of our study, we trained a neural network and stored the weights over all epochs and batches. The resulting NN is compared to two NNs trained with the average of these gathered weights $\omega_E \approx 0.273$, one using just the adamized loss, Eq. (9), and ignoring Eq. (10) and one using Eq. (8). Indeed, training NNs with the fixed energy weight results in a parametrization with similar performance as

applying Eqs. (9) and (10). The advantage thus lies in the elimination of the ω_E hyperparameter, which otherwise would have to be manually tuned.

The loss described above is minimized over a total of 1000 epochs using a one-cycle scheduler, varying the learning rate linearly from 1.5×10^{-4} to 1.5×10^{-3} and back, with the final 10 % decreasing further to 1.5×10^{-5}.

2. Augmentation

Due to the high dimensionality of the sampling space described earlier, we find forces in the training set ranging from a few meV Å^{-1} to several hundreds of eV Å^{-1}, caused by atoms being close to each other. As it is impossible and undesirable to enumerate a sufficient number of atomic environments resulting in those high forces and the NNFF behaving potentially erratically in untrained areas we include a repulsive contribution as suggested in [47]. We choose the repulsive part of the Morse potential to provide some robustness in those sparsely sampled regions:

$$V_{\text{rep}}(\mathbf{r}) = \frac{1}{2} \sum_{i \neq j} d_{ij} \times h(r_{ij}, r_s, r_c) e^{-2a_{ij}(r_{ij} - b_{ij})}, \quad (11)$$

where r_{ij} is the distance between atoms i and j, r_s is a trainable switching radius and r_c the cutoff radius, set equal to the cutoff radius of the NNFF, i.e. $r_c = 5$ Å. The a_{ij}, b_{ij}, d_{ij} represent element-specific parameters which are optimized during the training process if atom i and j are of the same type. For interactions between two different chemical species, the corresponding "mixed" parameters obtained through the rules established in Ref. 48 are used. The bump function, h, is defined by

$$f(x) = \begin{cases} 0 & x \le 0 \\ \exp\left(-\frac{1}{x}\right) & x > 0 \end{cases} \quad (12)$$

$$g(x) = \frac{f(x)}{f(x) + f(1-x)} \quad (13)$$

$$h(r, r_s, r_c) = 1 - g\left(\frac{r^2 - r_s^2}{r_c^2 - r_s^2}\right). \quad (14)$$

To test how efficiently a NNFF can be built to be useful as a surrogate model for EHPs, we decided to analyse a subset of the DFT data generated for our investigation of cI-HfO$_2$ from [8]. As the cubic phase shows an instability at 0 K, the first iteration data contains numerous high-energy configurations with large forces. We use a subset of the first 5 iterations, resulting in 172 structures, to validate this approach and train four NNFFs: using (i) only those configurations where the norms of all forces are below 50 eV Å^{-1} and
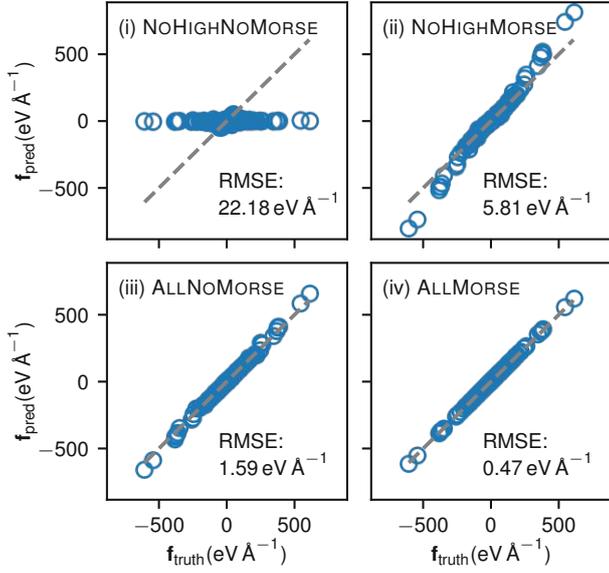
Figure 1. Comparison of the NNFF with and without the Morse model evaluated on high-force configurations, with training data including and excluding high-force configurations.

without Eq. (11) (NoHighNoMorse), (ii) only those configurations where the norms of all forces are below $50\,\mathrm{eV\,\AA^{-1}}$ and with Eq. (11) (NoHighMorse), (iii) all configurations without Morse (AllNoMorse) and (iv) all configurations with Morse (AllMorse). The resulting parity plots on the high-force configurations are shown in Fig. 1. As expected, the NoHighNoMorse NNFF performs poorly; in fact, it predicts a negligible force where large forces should be expected. This would lead to quickly-failing sampling as an artifactual minimum state with clusters of atoms would be considered advantageous. The NoHighMorse potential, on the other hand, predicts at least *some* high force. As the exact value in this case is not necessarily important the addition of the term Eq. (11) can indeed be understood as fulfilling its intended purpose: providing robustness in sparsely sampled regions. Naturally, for production runs we include all data, where the addition of the Morse potential proves superior as well.

C. Computational details

The DFT data used throughout this manuscript is obtained using the projector-augmented-wave (PAW) formalism [49] as implemented in VASP [50] using the PBE functional [51], an energy cutoff of $600\,\mathrm{eV}$ and the 5p6s5d and 2s2p as valence states for Hf and O respectively. Supercells of size $4\times4\times4$, $4\times4\times3$, $2\times2\times2$ and
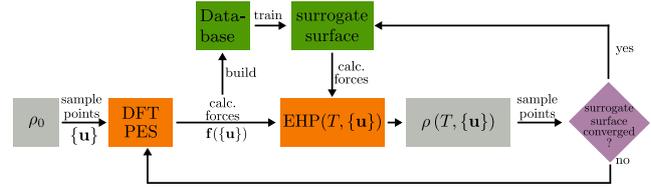

Figure 2. Flowchart of the EHP sampling procedure and gradual construction of a data base used to parametrize the surrogate NNFF.
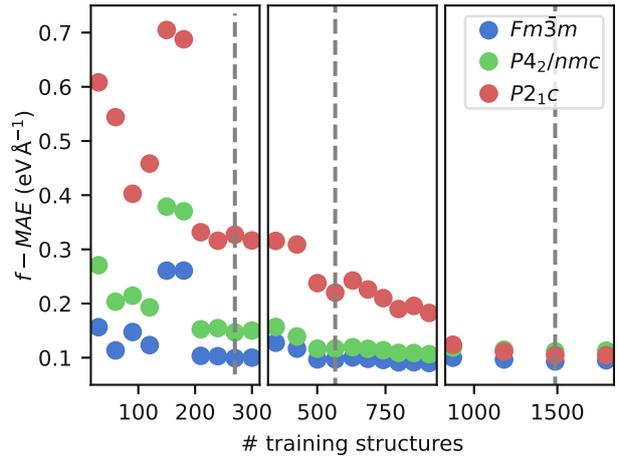

Figure 3. Decrease of force and energy MAE for increasing number of structures. The first panel shows convergence of the error as cubic data is added; the vertical dashed line indicates the chosen number of cubic training structures. In panels two and three this is repeated for tetragonal and monoclinic structures, respectively.

$2\times2\times2$ were used for the $Fm\bar{3}m$, $P4_2nmc$, $P\bar{4}3m$ and $P2_1c$ phases, respectively. A Γ-only k-point mesh was used for all calculations. Doubling the k-point mesh in each direction only had minimal impact on the energies ($\approx 3\,\mathrm{meV\,atom^{-1}}$) and forces ($< 0.1\,\mathrm{meV\,\AA^{-1}}$), comparable to the error incurred by the NNFF. The lattice constants of both cubic phases cover $-2\,\%$ to $5\,\%$ of their respective equilibrium values. For the tetragonal phase, deformations in a and c span $-1\,\%$ to $4\,\%$ and $-1.5\,\%$ to $3\,\%$, respectively. The monoclinic lattice parameters vary individually in the range of $-5\,\%$ to $5\,\%$.

III. RESULTS

A. Neural network validation and transferability

To construct a NNFF as a surrogate model for the EHPs we started with a subset, the $T = 2500\,\mathrm{K}$ samples, of the DFT data generated for our investigation

of cI-HfO$_2$ from Ref. 8. As this data was generated 30 points at a time through an iterative DFT-based EHP scheme, it has a naturally implied hierarchy. Starting out from a density matrix (ρ_0) constructed from the FD force constants, we move along this EHP hierarchy as indicated in Fig. 2, adding data points to the electronic structure database after each iteration. Based on this database an NNFF is parameterized at each iteration. The performance is assesed on a test set consisting of 50 randomly chosen and unseen structures from each of the phases, $Fm\bar{3}m$, $P4_2nmc$ and $P2_1c$ resulting in the first panel of Fig. 3. This process continues, mimicking active learning, until the error has converged for the cubic phase, as is indicated by the bottom arrow in Fig. 2. After convergence, the number of cubic structures contained in the database is fixed at 270 structures (the dashed line in the first panel of Fig. 3). Now, using the NNFF the force evaluations for treatment of the cubic phase with the EHP algorithm can be performed, following the top-right arrow in the flowchart (Fig. 2).

Next, we start from this database and repeat the procedure for each (a, c) pair defined for the tetragonal phase, adding ≈ 75 tetragonal structures sampled at $2500\,\mathrm{K}$ to the database at each iteration. In the middle panel of Fig. 3 it can already be seen that the local environments in the tetragonal and cubic phase have a high degree of similarity, requiring only an additional ≈ 300 structures until the prediction error stops improving. Finally the same procedure is repeated for all the (a, b, c) combinations of the monoclinic phase. A major improvement in performance for the monoclinic test set can be seen already after the first addition of monoclinic data to the training set (corresponding to the first set of points in the third panel of Fig. 3). For convergence a total of 921 additional monoclinic data points is required.
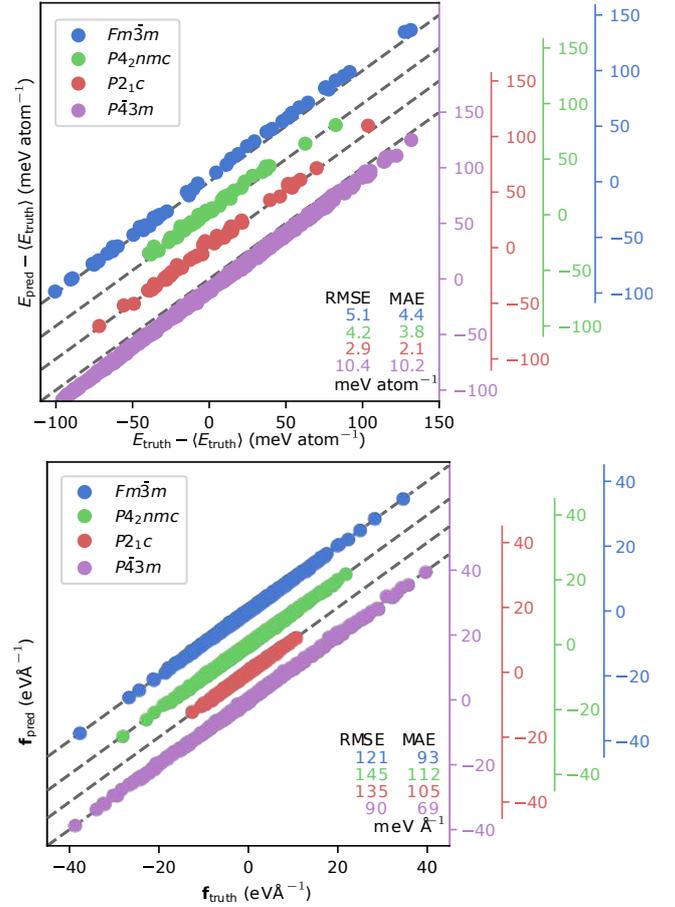

Figure 4. Parity plots and errors of the test set performance for the energies (top) and forces (bottom) for the $Fm\bar{3}m$, $P\bar{4}3m$, $P2_1c$ and $P4_2nmc$ phases of HfO$_2$.

Phase	Force (meVÅ^{-1})			Energy (meV atom^{-1})		
	RMSE	MAE	σ	RMSE	MAE	σ
$Fm\bar{3}m$	174	108	2989	4	3	185
$P4_2/nmc$	201	141	4474	5	4	227
$P2_1c$	155	120	2327	3	2	44

Table I. Comparison of the root-mean-square error (RMSE), the mean-absolute error (MAE) and the standard deviation (σ) of the data for the three phases contained in the validation set.

This cascading algorithm results in a training set with 1487 structures and a validation set with 599 structures. The corresponding validation statistics and the standard deviation of the dataset are shown in Table I. The parity plots and statistics of the final model applied to said test set, as well as to an additional set of data points for $P\bar{4}3m$, are shown in Fig. 4. Clearly, the model performs well on the three phases used to

generate configurations constituting the training data set, but it also provides excellent predictions for the completely unknown cII phase. This indicates the local environments sampled throughout the training provide enough diversity to result in a transferable model. As an additional qualitative validation metric we show the excellent agreement of the NNFF and DFT-backed finite displacement (FD) and EHP phonon spectrum of the $Fm\bar{3}m$ phase in the supplemental materials. Notably, none of the small-displacement FD configurations were part of the training, yet their prediction is sufficient to reconstruct a phonon spectrum.

B. Influence of pressure

For a first look at the stability of the four phases considered, energy-volume curves with all atoms at their equilibrium positions are shown in Fig. 5. For the tetragonal ($P4_2nmc$) and monoclinic phase ($P2_1c$) we
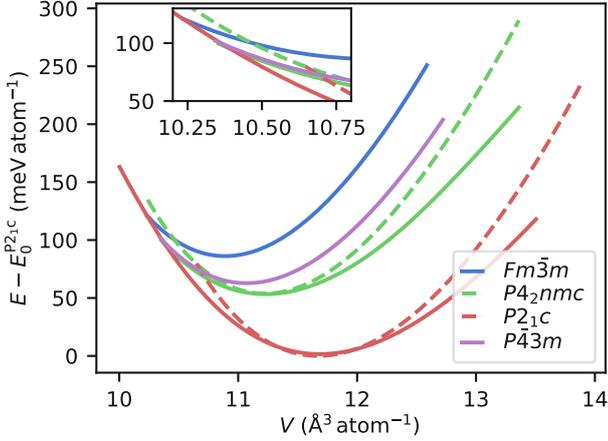
6

Figure 5. The (E,V)-curves of the various phases of HfO_2 under consideration. The solid lines show the actual (E,V) curves, i.e., taking variations of the lattice constants into account, while the dashed lines are only volume-scaled.

have calculated curves both by fully relaxing all lattice constants at each volume (full lines) and by simply scaling the volumes and keeping the aspect ratios fixed at the optimal volume values (dashed lines). For the $P2_1c$ phase, the effect of the unit cell can be split into the impact of changing the cell lengths and of changing β. Within the studied volume region $10.5\,\text{Å}^3\,\text{atom}^{-1}$ to $12.5\,\text{Å}^3\,\text{atom}^{-1}$, the impact of β relaxation is negligible; hence, β is kept constant at its optimal equilibrium angle $\beta_0 = 99.69°$. Energy-volume curves in literature (e.g. Fig. 4 in [22]) often suggest a first-order phase transition between the $P2_1c$, $P4_2nmc$ and $Fm\bar{3}m$ induced by a pressure of $\approx 8.7\,\text{GPa}$ when the effect of vibrations is neglected. This, however, only holds if the lattice constants are volume-scaled. With variable cell lengths, the tetragonal and monoclinic phases are favorable over the $Fm\bar{3}m$ until their axes are compressed to the point that they become essentially cubic at $10.3\,\text{Å}^3\,\text{atom}^{-1}$. Fig. 5 also shows that the lower-symmetry cubic $P\bar{4}3m$ phase is lower in energy (by $23\,\text{meV}\,\text{atom}^{-1}$) than $Fm\bar{3}m$ at ambient pressure. Again, no first-order pressure induced phase transition to $Fm\bar{3}m$ is found and the two phases gradually become energetically indistinguishable at volumes around $10.3\,\text{Å}^3\,\text{atom}^{-1}$.

The optimal PBE volumes are given in Table II. Extrapolating the experimental data of the monoclinic and tetragonal phase by Haggerty et al. [52] to $0\,\text{K}$, we arrive at equilibrium volumes of $V_0^{\text{exp}} = 11.39\,\text{Å}^3\,\text{atom}^{-1}$ and $V_0^{\text{exp}} = 10.91\,\text{Å}^3\,\text{atom}^{-1}$ for the monoclinic phase and tetragonal phases respectively. The PBE results in Table II thus overestimate the volume by $\approx 2.5 - 3.0\%$.

	0 GPa		4 GPa
	PBE	**PBEsol**	**PBE**
$Fm\bar{3}m$	10.89	10.54	10.73
$P\bar{4}3m$	11.07	10.68	10.89
$P4_2nmc$	11.23	10.79	11
$P2_1c$	11.68	11.28	11.44
	V ($\text{Å}^3\text{atom}^{-1}$)		

Table II. Volumes of the phases as obtained using PBE and PBEsol functionals at $0\,\text{GPa}$ in comparison to those obtained by application of an artificial pressure $p_a = 4\,\text{GPa}$

The tendency to overestimate unit cell volumes can substantially influence the calculated thermodynamic properties [53–55] and is a well-known shortcoming of the PBE functional [53, 56]. It has led to the development of more specialized functionals like the Wu-Cohen [57] and PBEsol [58] functionals where a lower exchange enhancement leads to volumes in better agreement with experiment. In Table II we show that PBEsol underestimates the volumes by approximately 1 %. The most straightforward strategy, however, is the addition of a small artificial isotropic pressure, p_a [53–55]. Here we choose to perform the calculations with PBE and fix the artificial pressure $p_a = 4\,\text{GPa}$, resulting in the volumes shown in Table II. The thus obtained bulk modulus for the m-phase, $B_0 = 184\,\text{GPa}$ is in agreement with experimental results ranging from $187\,\text{GPa}$ to $197\,\text{GPa}$ [59].

C. Influence of temperature

Besides pressure, temperature is an important contributor to phase stability. Conceptually, the EHP treats the influence of temperature in a manner similar to the quasi-harmonic approach, by calculating phonons on a grid of lattice constants, but with the important difference that a self-consistent phonon spectrum must be calculated for each set of lattice constants and temperatures. Furthermore the free energy contains the correction term accounting for the difference between the ensemble-averaged potential energy and its effective harmonic approximation.

For the investigation of the cubic phase in Ref. 8 eight volume points were considered. Using an effective sampling technique to reuse data points, approximately 1200 single-point calculations were required to achieve convergence for a rather narrow temperature range. Achieving convergence for the full temperature range for a low symmetry structure like the $P2_1c$ can become computationally prohibitive. Using a conservative estimate, the total number of calculations required for all four phases at the current temperature range would exceed 80000. While possible, this would have

been a major undertaking and the advantage of using a comparatively cheap surrogate model becomes apparent. The NNFF evaluates a configuration in negligible time and there is no need to be frugal when it comes to the number of samples drawn at each iteration. Consider for example, the $Fm\bar{3}m$ phase, where we evaluated 150 000 structures. For the tetragonal phase approximately 500 000 structures were evaluated, and for the monoclinic phase that number was almost 2.3 million. This allows obtaining good effective sample sizes and, with them, a much more accurate evaluation of particularly F_{corr}.

The finite-displacement phonon band structure of $Fm\bar{3}m$ exhibits imaginary frequencies around the X point [8, 16, 27]. This instability at $X = (\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$ was linked to a cubic-to-tetragonal deformation in zirconia which is structurally similar [60]. This would indicate that the phase is mechanically unstable and would make the application of Eq. (4) invalid. We have already shown that the EHP makes $Fm\bar{3}m$ become mechanically stable at temperatures of 2100 K and above [8]. Using the NNFF as the backend we can employ a much finer temperature mesh and provide a clearer picture. As shown in Fig. 6, the lowest X-point frequency increases continuously with temperature until it abruptly changes sign at around 1200 K. Moreover, we find a small instability in the phonon band structure close to the Γ point, at $(\frac{1}{30}, \frac{1}{30}, 0)$ and permutations thereof, in the $P\bar{4}3m$ phase when using the finite displacement method (Fig. 7). Clearly, this phase is also mechanically stabilized through temperature, and using the EHP $P\bar{4}3m$ becomes stable at temperatures lower than 1000 K.

Having achieved real phonon spectra, the free energies of the phases can be calculated using Eq. (4) and the correction term Eq. (5). We use the EHP approach for all phases, including those that are mechanically stable in the ordinary harmonic approximation (namely $P2_1c$ and $P4_2nmc$), and obtain free energy-volume curves as a function of temperature and phase and hence the thermal expansion. As shown in Fig. 8, under the artificial pressure $p_a = 4\,\text{GPa}$, the thermal expansion of the monoclinic and tetragonal phases very closely matches the results measured by Haggerty et al. [52] and Tobase et al. [19]. The monoclinic result at room temperature, $V = 11.54\,\text{Å}^3/\text{atom}$, is further corroborated by measurements performed by Ruh et al. ($V = 11.52\,\text{Å}^3/\text{atom}$, [61]) and Akahama et al. ($V = 11.53\,\text{Å}^3/\text{atom}$, [59]). Also individual lattice parameters agree well with experiment, as shown in the supplementary information.

Considering the good agreement found for the tetragonal and monoclinic phases, it is somewhat surprising that we see a distinct mismatch with experiment for the presumed cubic HfO_2 phases [19, 62]. Nonetheless, we
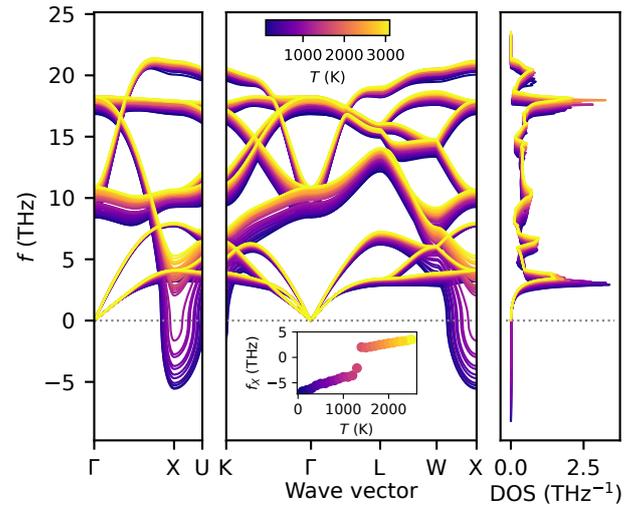

Figure 6. Continuous hardening with temperature of the soft mode exhibited by the $Fm\bar{3}m$ phase at $V = 10.87\,\text{Å}^3\text{atom}^{-1}$. The inset shows the value of the lowest-lying acoustic mode at X as a function of temperature.
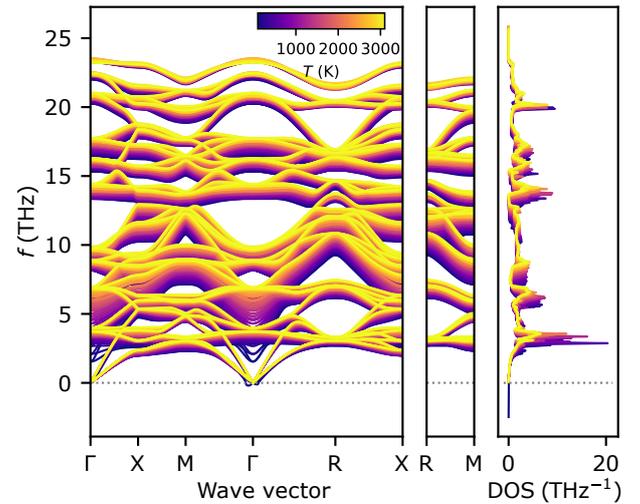

Figure 7. Continuous hardening with temperature of the soft mode exhibited by the $P\bar{4}3m$ phase at $V = 11\,\text{Å}^3\text{atom}^{-1}$.

can draw some important conclusions: First, the rarely studied $P\bar{4}3m$, while still differing by 3 %, is closer in volume to those measurements than the $Fm\bar{3}m$ phase. Secondly, the phases measured in [19, 62] might not correspond to stoichiometric HfO_2. Previous investigations indicate a stabilization of high-symmetry phases over the monoclinic phase particularly with decreasing oxygen content [63–66]. Specifically a low-temperature cubic phase of $HfO_{1.7}$ is reported in Ref. 64. Further-
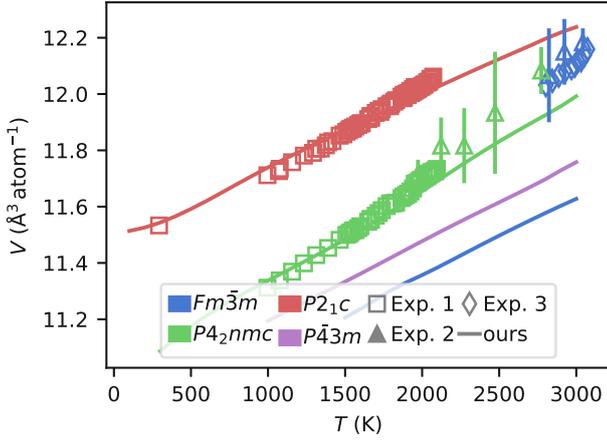
Figure 8. Thermal expansion of the two cubic, the tetragonal and the monoclinic phase obtained using the NNFF as compared to Haggerty et al. (Exp. 1, [52]), Tobase et al. (Exp. 2, [19]) and Hong et al. (Exp. 3, [62]).

Figure 9. The differences in free enthalpy of the monoclinic and tetragonal phase of HfO_2. The x-axis includes an artificial pressure of $p_a = 4\,\mathrm{GPa}$ as discussed in the text.

more, oxygen-deficient, but stoichiometrically unspecified, nanoparticles of cubic hafnia have been obtained through synthesis in a reductive solvent [65]. These oxygen-deficient structures could in turn be affected by vacancy-mediated chemical expansion, as discussed extensively in literature [67–69], which would explain the discrepancy visible in Fig. 8.

D. Thermodynamic phase stability

Using the free energy-volumes curves we furthermore find that temperature stabilizes the tetragonal $P4_2nmc$ phase compared to the $P2_1c$ phase, in agreement with experiment [24, 25]. However, as seen in Fig. 9, the application of additional pressure is essential to stabilize the $P4_2nmc$ phase over the $P2_1c$ within the temperature window studied in this work. For example, at 2600 K, a pressure of roughly 4 GPa in addition to p_a is needed to stabilize the tetragonal phase. This can again potentially be attributed to the chosen functional, PBE, which tends to produce slightly enlarged cells. A further indication supporting this is that the pressure needed to take the PBE equilibrium volume of the $Fm\bar{3}m$ phase to the PBEsol equilibrium volume, is of the same order of magnitude (7.7 GPa).

Interestingly, we do not find that temperature stabilizes the two studied cubic phases, $Fm\bar{3}m$ and $P\bar{4}3m$, compared to the tetragonal phase, Fig. 10. While the EHP methodology, due to the harmonic ansatz used in the construction of the density matrix, is not able to capture all anharmonic contributions, the effect of further corrections would have to be very significant
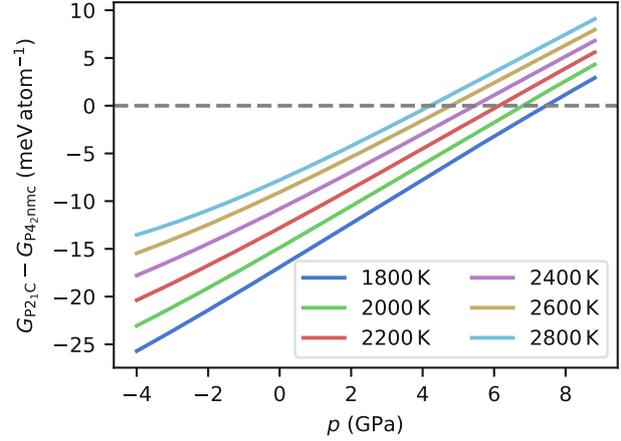
to overcome the energy differences required for a transition to a cubic lattice. Not even the studied pressure ranges are able to stabilize the cubic lattices, as is shown in Fig. 10. On the other hand, the energy differences are small in comparison with the characteristic scale of thermal fluctuations per degree of freedom (e.g. the $P4_2nmc$-to-$P\bar{4}3m$ enthalpy difference at 4 GPa is approximately 12 meV atom^{-1}, but k_BT at 2000 K is $\approx 172\,\mathrm{meV\,atom^{-1}}$). Processing-induced oxygen vacancies, or, at temperatures this high, even temperature-induced oxygen vacancies [62, 70, 71] might stabilize the cubic phase over its tetragonal counterpart, in particular in combination with interface and surface strain of the fine powders typically used in XRD measurements. Future research should more closely investigate how the energy balance between $Fm\bar{3}m$ and $P\bar{4}3m$ shifts when those parameters are considered. Likewise, the impact potential contributions to this picture arising from configurational entropy in the presence of defects should be probed. As a simple estimate $TS_{\mathrm{conf}} = Tk_B\left[(1-c)\log{(1-c)} + c\log{c}\right] \approx 28\,\mathrm{meV\,atom^{-1}}$ at an assumed defect concentration of $c = 10\,\%$.

These results do not preclude metastability as might be achieved in experiment: We emphasize that for a phase to be stable within the EHP formalism, its free energy needs to be the lowest. This might even be the case in molecular dynamics under certain conditions, such as in the work by Fan et al. [72] using ab-initio molecular dynamics (AIMD), in which they claim to observe a transition to $Fm\bar{3}m$. However, the computational restrictions inherent to AIMD might limit the interpretability of these results. First of all, the use of a small simulation box precludes relaxation through long-wavelength deformations. Secondly, the
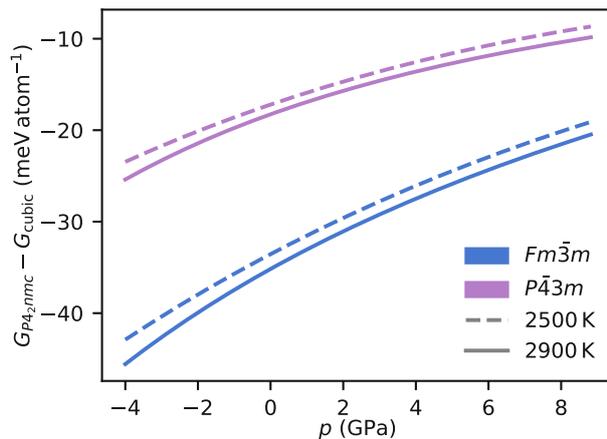
Figure 10. Impact of temperature, 2500 K (dashed) and 2900 K (solid), and pressure on the free enthalpy difference between $Fm\bar{3}m$ (blue), $P\bar{4}3m$ (violet) and $P4_2nmc$.

molecular-dynamics trajectory quickly scans through temperatures or pressures without a guarantee that the system is allowed to fully relax. To the extent a cubic phase is observed in Ref. 72, the methodology is furthermore insufficient to conclusively establish whether it corresponds to $Fm\bar{3}m$ or $P\bar{4}3m$. Those limitations, imposed by the high computational cost of direct ab-initio molecular dynamics, highlight the value of an accelerated surrogate model.

IV. CONCLUSION

We have described a framework to construct neural network force fields (NNFF) using effective harmonic potentials (EHP) and applied it towards the parametrization of a transferable NNFF for HfO_2. The potential performs comparably to state-of-the-art, with an out-of-sample force-MAE of $\approx 100 \, \mathrm{meV} \, \mathrm{\AA}^{-1}$ and an energy-MAE of $\approx 10 \, \mathrm{meV} \, \mathrm{atom}^{-1}$, despite being trained on only ≈ 1500 structures. The NNFF is then applied in conjunction with EHPs to study phase stability of the ambient phases of HfO_2. The methodology leads to agreement with experimental literature regarding the thermal expansion of the $P2_1c$ and $P4_2nmc$ phases. At roughly 4 GPa, which is comparable to an effective pressure linking PBE and PBEsol, we find a monoclinic-to-tetragonal phase transition at a temperature comparable to experiment (2500 K vs. 2050 K). We furthermore conclude that the most likely stoichiometric high-temperature space group of cubic bulk HfO_2 is $P\bar{4}3m$, as opposed to the traditionally assumed $Fm\bar{3}m$. Nonetheless, we do not see a tetragonal-to-cubic phase transition within the temperature and pressure range studied in this work and find that the observed lattice constants of the cubic phases are substantially larger than what is calculated for the corresponding stoichiometric phases. Future studies should explore the potential of metastability (e.g. through molecular dynamics) and the impact of oxygen vacancies on the various phases.

[1] A. Glensk, B. Grabowski, T. Hickel, and J. Neugebauer, Phys. Rev. X **4**, 011018 (2014).

[2] I. Errea, M. Calandra, and F. Mauri, Phys. Rev. Lett. **111**, 177002 (2013).

[3] O. Hellman, P. Steneteg, I. A. Abrikosov, and S. I. Simak, Phys. Rev. B **87**, 104111 (2013).

[4] R. Stern and G. K. H. Madsen, Phys. Rev. B **94**, 144304 (2016).

[5] A. van Roekeghem, J. Carrete, C. Oses, S. Curtarolo, and N. Mingo, Phys. Rev. X **6**, 041061 (2016).

[6] L. Monacelli, R. Bianco, M. Cherubini, M. Calandra, I. Errea, and F. Mauri, J. Phys. Condens. Matter **33**, 363001 (2021).

[7] A. van Roekeghem, J. Carrete, and N. Mingo, Comput. Phys. Commun. **263**, 107945 (2021).

[8] S. Bichelmaier, J. Carrete, M. Nelhiebel, and G. K. H. Madsen, Phys. Status Solidi Rapid Res. Lett. **n/a**, 2100642 (2022).

[9] O. T. Unke, S. Chmiela, H. E. Sauceda, M. Gastegger, I. Poltavsky, K. T. Schütt, A. Tkatchenko, and K.-R. Müller, Chem. Rev. **121**, 10142 (2021).

[10] T. W. Ko, J. A. Finkler, S. Goedecker, and J. Behler, Nature Communications **12**, 398 (2021).

[11] S. Faraji, S. A. Ghasemi, S. Rostami, R. Rasoulkhani, B. Schaefer, S. Goedecker, and M. Amsler, Phys. Rev. B **95**, 104105 (2017).

[12] H. Tang, Y. Zhang, Q.-J. Li, H. Xu, Y. Wang, Y. Wang, and J. Li, Acta Mater. **238**, 118217 (2022).

[13] Y. Du, Z. Meng, Q. Yan, C. Wang, Y. Tian, W. Duan, S. Zhang, and P. Lin, Phys. Chem. Chem. Phys. **24**, 18361 (2022).

10

[14] R. He, H. Wu, L. Zhang, X. Wang, F. Fu, S. Liu, and Z. Zhong, Phys. Rev. B **105**, 064104 (2022).

[15] R. Wanzenböck, M. Arrigoni, S. Bichelmaier, F. Buchner, J. Carrete, and G. K. H. Madsen, Digit. Discov. **1**, 703 (2022).

[16] T. D. Huan, V. Sharma, G. A. Rossetti, and R. Ramprasad, Phys. Rev. B **90**, 064111 (2014).

[17] R. Batra, H. D. Tran, and R. Ramprasad, Appl. Phys. Lett. **108**, 172902 (2016).

[18] R. Batra, T. D. Huan, J. L. Jones, G. Rossetti, and R. Ramprasad, J. Phys. Chem. C **121**, 4139 (2017).

[19] T. Tobase, A. Yoshiasa, H. Arima, K. Sugiyama, O. Ohtaka, T. Nakatani, K.-i. Funakoshi, and S. Kohara, Phys. Status Solidi (b) **255**, 1800090 (2018).

[20] W. Ding, Y. Zhang, L. Tao, Q. Yang, and Y. Zhou, Acta Mater. **196**, 556 (2020).

[21] S. S. Behara and A. Van der Ven, Phys. Rev. Materials **6**, 054403 (2022).

[22] J. Wu, Y. Zhang, L. Zhang, and S. Liu, Phys. Rev. B **103**, 024108 (2021).

[23] G. Sivaraman, A. N. Krishnamoorthy, M. Baur, C. Holm, M. Stan, G. Csányi, C. Benmore, and Á. Vázquez-Mayagoitia, npj Comput. Mater. **6**, 104 (2020).

[24] T. B. Massalski, H. Okamoto, P. Subramanian, L. Kacprzak, and W. W. Scott, *Binary alloy phase diagrams*, Vol. 1/2 (American society for metals Metals Park, OH, 1986).

[25] C. Wang, M. Zinkevich, and F. Aldinger, J. Am. Ceram. Soc. **89**, 3751 (2006).

[26] P. Fan, Y. K. Zhang, Q. Yang, J. Jiang, L. M. Jiang, M. Liao, and Y. C. Zhou, J. Phys. Chem. C **123**, 21743 (2019).

[27] L. Azevedo Antunes, R. Ganser, C. Kuenneth, and A. Kersch, Phys. Status Solidi Rapid Res. Lett. **n/a**, 2100636 (2022).

[28] S. V. Barabash, J. Comput. Electron. **16**, 1227 (2017).

[29] G. Sivaraman, L. Gallington, A. N. Krishnamoorthy, M. Stan, G. Csányi, A. Vázquez-Mayagoitia, and C. J. Benmore, Phys. Rev. Lett. **126**, 156002 (2021).

[30] K. J. McClellan, S.-Q. Xiao, K. P. D. Lagerlof, and A. H. Heuer, Philos. Mag. A **70**, 185 (1994).

[31] C. M. Wang, S. Azad, S. Thevuthasan, V. Shutthanandan, D. E. McCready, and C. H. F. Peden, J. Mater. Res. **19**, 1315 (2004).

[32] M. Sternik and K. Parlinski, J. Chem. Phys. **123**, 204708 (2005).

[33] O. Ohtaka, D. Andrault, P. Bouvier, E. Schultz, and M. Mezouar, J. Appl. Crystallogr. **38**, 727 (2005).

[34] K. Tolborg and A. Walsh, ChemRxiv e-prints (2022), 10.26434/chemrxiv-2022-lkzm9.

[35] H. Montes-Campos, J. Carrete, S. Bichelmaier, L. M. Varela, and G. K. H. Madsen, J. Chem. Inf. Model. **62**, 88 (2022).

[36] V. Zaverkin, D. Holzmüller, I. Steinwart, and J. Kästner, Digit. Discov. **1**, 605 (2022).

[37] L. Zhang, J. Han, H. Wang, R. Car, and W. E, Phys. Rev. Lett. **120**, 143001 (2018).

[38] R. Jinnouchi, K. Miwa, F. Karsai, G. Kresse, and R. Asahi, J. Phys. Chem. Lett. **11**, 6946 (2020).

[39] D. Schwalbe-Koda, A. R. Tan, and R. Gómez-Bombarelli, Nat. Comm. **12**, 5104 (2021).

[40] G. Scalia, C. A. Grambow, B. Pernici, Y.-P. Li, and W. H. Green, J. Chem. Inf. Model. **60**, 2697 (2020).

[41] I. Errea, M. Calandra, and F. Mauri, Phys. Rev. B **89**, 064302 (2014).

[42] Q. Wang, Y. Ma, K. Zhao, and Y. Tian, Ann. Data Sci. **9**, 187 (2022).

[43] S. Wang, Y. Teng, and P. Perdikaris, SIAM J. Sci. Comput. **43**, A3055 (2021), arXiv:2001.04536 [cs.LG].

[44] S. Maddu, D. Sturm, C. L. Müller, and I. F. Sbalzarini, Mach. learn.: sci. technol. **3**, 015026 (2022), arXiv:2107.00940 [cs.LG].

[45] B. Mitrevski, M. Filipovic, D. Antognini, E. Lejal Glaude, B. Faltings, and C. Musat, arXiv e-prints , arXiv:2009.04695 (2020), arXiv:2009.04695 [cs.LG].

[46] S. Liu and L. N. Vicente, arXiv e-prints (2019), 10.48550/ARXIV.1907.04472.

[47] A. P. Bartók, J. Kermode, N. Bernstein, and G. Csányi, Phys. Rev. X **8**, 041048 (2018).

[48] C. L. Kong, J. Chem. Phys. **59**, 2464 (1973).

[49] P. E. Blöchl, Phys. Rev. B **50**, 17953 (1994).

[50] G. Kresse and J. Furthmüller, Phys. Rev. B **54**, 11169 (1996).

[51] J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996).

[52] R. P. Haggerty, P. Sarin, Z. D. Apostolov, P. E. Driemeyer, and W. M. Kriven, J. Am. Ceram. Soc. **97**, 2213 (2014).

[53] Z. Wu, R. E. Cohen, and D. J. Singh, Phys. Rev. B **70**, 104112 (2004).

[54] W. Zhong, D. Vanderbilt, and K. M. Rabe, Phys. Rev. Lett. **73**, 1861 (1994).

[55] P. Xie, Y. Chen, E. Weinan, and R. Car, arXiv e-prints (2022), 10.48550/ARXIV.2205.11839.

[56] F. Tran, J. Stelzl, and P. Blaha, J. Phys. Chem. **144**, 204120 (2016), https://doi.org/10.1063/1.4948636.

[57] Z. Wu and R. E. Cohen, Phys. Rev. B **73**, 235116 (2006).

[58] J. P. Perdew, A. Ruzsinszky, G. I. Csonka, O. A. Vydrov, G. E. Scuseria, L. A. Constantin, X. Zhou, and K. Burke, Phys. Rev. Lett. **100**, 136406 (2008).

[59] Y. Akahama, S. Kawaguchi, N. Hirao, and Y. Ohishi, Appl. Phys. Lett. **117**, 182903 (2020).

[60] A. Kuwabara, T. Tohei, T. Yamamoto, and I. Tanaka, Phys. Rev. B **71**, 064301 (2005).

[61] R. Ruh and P. W. R. Corfield, J. Am. Ceram. Soc. **53**, 126 (1970).

[62] Q.-J. Hong, S. V. Ushakov, D. Kapush, C. J. Benmore, R. J. K. Weber, A. van de Walle, and A. Navrotsky, Sci. Rep. **8**, 14962 (2018).

[63] T. Mittmann, M. Materano, P. D. Lomenzo, M. H. Park, I. Stolichnov, M. Cavalieri, C. Zhou, C.-C. Chung, J. L. Jones, T. Szyjka, M. Müller, A. Kersch, T. Mikolajick, and U. Schroeder, Adv. Mater. Interfaces **6**, 1900042 (2019).

[64] N. Kaiser, T. Vogel, A. Zintler, S. Petzold, A. Arzumanov, E. Piros, R. Eilhardt, L. Molina-Luna, and L. Alff, ACS Appl. Mater. Interfaces **14**, 1290 (2022), https://doi.org/10.1021/acsami.1c09451.

[65] P. Rauwel, E. Rauwel, C. Persson, M. F. Sunding, and A. Galeckas, J. Appl. Phys. **112**, 104107 (2012), https://doi.org/10.1063/1.4766272.

[66] D. Islamov and T. Perevalov, Microelectron. Eng. **216**, 111041 (2019).

[67] D. Marrocchelli, S. R. Bishop, H. L. Tuller, G. W. Watson, and B. Yildiz, Phys. Chem. Chem. Phys. **14**, 12070 (2012).

[68] D. S. Aidhy, B. Liu, Y. Zhang, and W. J. Weber, Comput. Mater. Sci. **99**, 298 (2015).

[69] U. Aschauer, R. Pfenninger, S. M. Selbach, T. Grande, and N. A. Spaldin, Phys. Rev. B **88**, 054111 (2013).

[70] R. Kumar, R. Kumar, M. Singh, D. Meena, and A. Vij, J. Phys. D **55**, 495302 (2022).

[71] N. Hadacek, A. Nosov, L. Ranno, P. Strobel, and R.-M. Galéra, J. Phys. Condens. Mater. **19**, 486206 (2007).

[72] P. Fan, Y. K. Zhang, Q. Yang, J. Jiang, L. M. Jiang, M. Liao, and Y. C. Zhou, J. Phys. Chem. C **123**, 21743 (2019).

Supplementary information

January 17, 2023

Comparison DFT and NN-backed phonon spectra
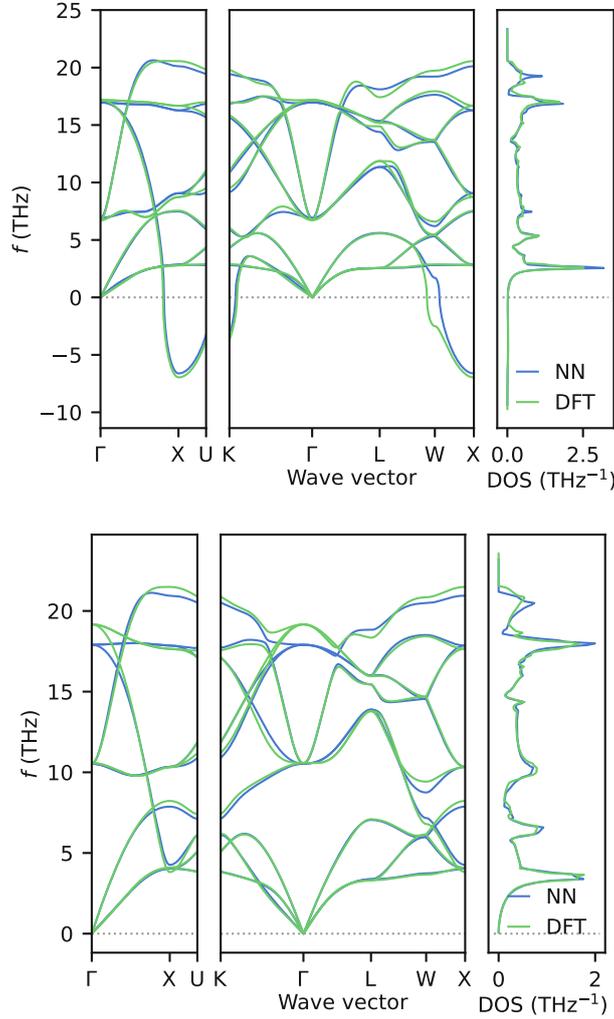
Figure 1: Comparison of finite displacement (top) and 2500 K EHP (bottom) phonon spectra for $Fm\bar{3}m$ phase of HfO$_2$ at $V = V_0 = 32.62\,\text{Å}^3$

As shown in the top panel of Figure 1, the finite displacement phonons agree exceptionally well. Even more so, when considering that the NNFF is not trained to reproduce the very carefully chosen and very small (0.01 Å) displacements used in obtaining these spectra. Comparing the spectra obtained at 2500 K using EHP shown in the bottom panel of Figure 1, we find a good agreement too. The slight deviations in the optical modes can likely be attributed to the fact that the convergence criterion is based on the free energy and the fact that the NN-obtained spectra are obtained from a much higher sampling, i.e. they can be expected to be more accurate.

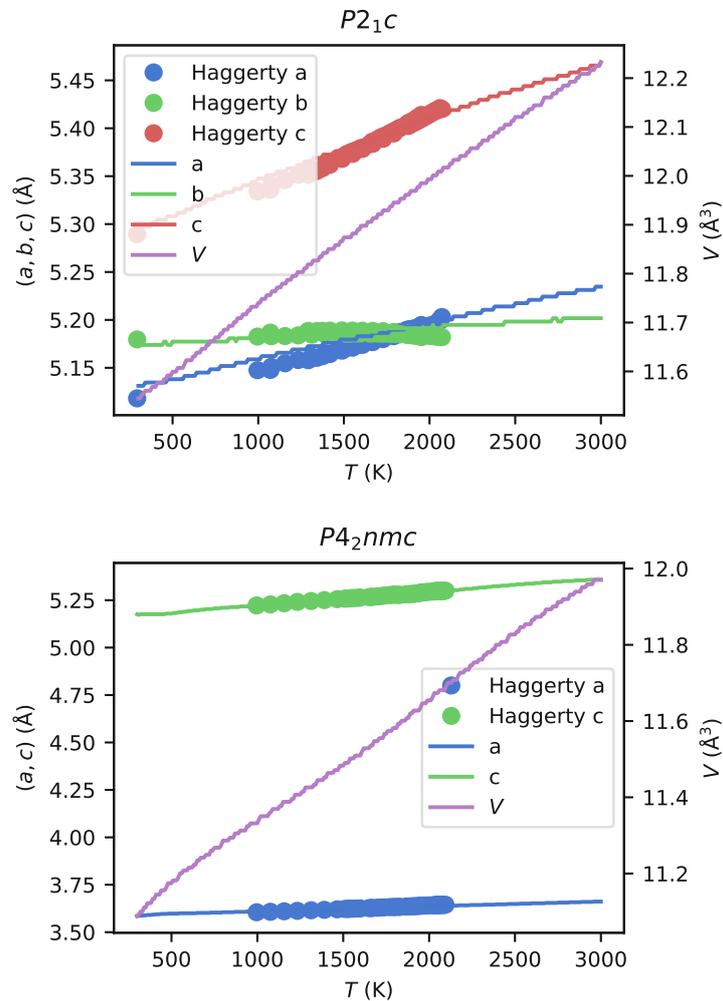# Comparison experimental and NN-backed EHP lattice constants

Figure 2: Lattice constants for $P2_1c$ (top) and $P4_2nmc$ (bottom) as obtained with the NN-backed EHP approach as compared to Haggerty et al.[1].

References

[1] R. P. Haggerty, P. Sarin, Z. D. Apostolov, P. E. Driemeyer, and W. M. Kriven, J. Am. Ceram. Soc. **97**, 2213 (2014).

Publication III

International Journal of **QUANTUM CHEMISTRY** **WILEY**

Evaluating the efficiency of power-series expansions as model potentials for finite-temperature atomistic calculations

Sebastian Bichelmaier[1,2] | Jesús Carrete[1] | Georg K. H. Madsen[1]

[1]Institute of Materials Chemistry, TU Wien, Vienna, Austria

[2]KAI GmbH, Villach, Austria

Correspondence
Georg K. H. Madsen, Institute of Materials Chemistry, TU Wien, A-1060 Vienna, Austria.
Email: georg.madsen@tuwien.ac.at

Funding information
Horizon 2020 Framework Programme, Grant/Award Number: 826060

Abstract

Given the cost of ab-initio calculations, predictive studies of temperature-dependent phenomena in strongly anharmonic systems pose a serious challenge. Using a relatively inexpensive surrogate model for the potential energy surface to build temperature-dependent effective harmonic potentials is a possible solution. Automatic differentiation makes high-order Taylor potentials as surrogate models a relatively accessible possibility. Here Lennard-Jones clusters and solids are used as a test bench on which to perform a detailed analysis of such a procedure. It is found that results might only be valid in a narrow temperature regime, outside of which the local nature of Taylor expansions leads to drastic artifacts in the free energies and derived quantities such as the thermal expansion. Those shortcomings are traced to the limited flexibility of polynomials as approximants and are therefore fundamental. The observed behavior is confirmed using density functional theory on a five-atom silver cluster. A global interpolation strategy, in the form of a neural-network force field, is suggested as a better path to cost-effective surrogate models.

KEYWORDS
automatic differentiation, Lennard-Jones clusters, neural-network force field, Taylor potential

1 | INTRODUCTION

An ultimate goal of computational materials science is to deliver predictions about the suitability of a given material for a specific practical application. Progress toward that end requires calculations to be performed under increasingly realistic conditions. In this context, what can be generally classified as effective harmonic potentials (EHPs) have emerged to study temperature-dependent effects using density functional theory (DFT). EHPs date back to Hooton in 1955 [1] and retain the analytical and computational convenience of the well-known harmonic approximation while avoiding some of its worst limitations. They have recently been successfully used for the calculation of the phase diagram of $Ti_{1-x}Al_xN$ [2], the anomalous thermal behavior of ScF_3 [3] and FeSi [4], the structural phase transitions of $BaTiO_3$ [5], $SrTiO_3$ [6] and SnSe [7, 8] as well as the stability of the high-temperature cubic phase of HfO_2 [9].

EHPs aim at approximating the actual dynamics of the nuclei in the relevant parts of the potential energy surface (PES) at a given temperature and offer an attractive balance of accuracy and efficiency. This has given rise to a number of frameworks differing mostly in the criteria used to find the optimal EHP, and in the way, the PES is sampled [10–16]. However, common to all implementations of EHPs is that they rely on importance sampling of the PES for a given temperature and set of structural parameters. As a result, when a large range of different conditions, for example, temperatures and structural phases, are of interest, the required number of DFT calculations can be prohibitive.

The use of a relatively inexpensive surrogate model for the PES then becomes advantageous. One possibility in this respect is the use of higher-order Taylor expansions of the PES. This has been done in a number of molecular dynamics simulations [17–19] and recent EHP studies [6, 8, 15, 20]. At temperatures nearing absolute zero and in particular contexts (e.g., Reference 21 compared to [22]) a remarkable agreement of such methods with experiment can be obtained. The open-source hiPhive package [15] makes these high-order expansions of the PES accessible and eliminates the need to reimplement the significant bookkeeping involved. An alternative for efficiently deriving Taylor expansions is available through automatic differentiation (autodiff) [23, 24]. While autodiff is a well-established method [25], it has only risen to prominence with the current, rapid development of machine learning (ML) methodologies and frameworks such as JAX [26]. Recently, the use of machine-learned force fields (MLFF) as surrogate models has surged in the study of solids [27, 28] and molecules [29, 30] alike. At the same time, the underlying autodiff is also having an impact on more traditional chemistry tools like PES exploration [31], molecular dynamics [32] and electronic structure calculation [33]. Applying autodiff to PES allows the setup of higher-order Taylor expansions with machine precision. However, due to the intrinsic local nature of Taylor expansions, it is important to take a critical look at their use as surrogate models for finite-temperature calculations beyond the small-displacement regime.

In this paper, we employ the widely used Lennard-Jones potential (LJ) as a test case to assess the use of Taylor expansions as surrogate PES models when constructing EHPs. Despite its simplicity, the LJ model offers all the required ingredients for this kind of study, including a significant degree of anharmonicity. Furthermore, it has been shown that the model, though used only as an illustrative example in this work, can provide valuable insight into the physics particularly of solid noble gases and the impact vibrational effects have on their bulk properties [34]. We analyze both the case of finite LJ clusters using autodiff as well as that of the fcc LJ solid using hiPhive [15], with a focus on normal modes and thermal expansion. Finally, to emphasize the transferability of the analysis described before to more practical situations, we study a five-atom silver cluster, comparing exact calculations using DFT and a machine-learned global approximant [35] to higher-order expansions thereof.

2 | METHODOLOGY

2.1 | Effective harmonic potentials

Let $\widehat{V}$ and $\widehat{\rho}_0$ be the ionic PES and density matrix of the system, respectively, and let us introduce a trial harmonic potential $\widehat{\mathcal{V}}$, with an associated trial density matrix $\widehat{\rho}$. Any EHP fulfills the following variational property [8, 13, 14] based on the Gibbs-Bogoliubov inequality [36] for the free energy F:

$$F\left[\widehat{V},\widehat{\rho}_0\right] \leq \mathcal{F}_{\text{EHP}} = F\left[\widehat{\mathcal{V}},\widehat{\rho}\right] + \text{Tr}\left\{\widehat{\rho}\left(\widehat{V}-\widehat{\mathcal{V}}\right)\right\} = F_{\text{harm}} + F_{\text{corr}}. \tag{1}$$

All harmonic quantities, the potential $\mathcal{V}$, the density matrix $\widehat{\rho}$ and the expression for the free energy, F_{harm}, can be evaluated explicitly in terms of the mass-weighted displacements ($\mathbf{u} = \mathbf{M}^{1/2}(\mathbf{r}-\mathbf{r}_0)$, expressed using the diagonal matrix of ionic masses $\mathbf{M}$) of the atoms from their equilibrium positions and the eigenvalues and eigenvectors (ω_λ^2 and ϵ_λ, respectively) of the EHP:

$$\mathcal{V}(\mathbf{u}) = \sum_\lambda \frac{1}{2}\omega_\lambda^2 \mathbf{u} \cdot \left(\epsilon_\lambda \otimes \epsilon_\lambda^*\right)\mathbf{u} \tag{2a}$$

$$\langle\mathbf{u}|\widehat{\rho}|\mathbf{u}\rangle = \rho(\mathbf{u}) = \frac{1}{\sqrt{(2\pi)^{3N}|\mathbf{C}|}} \exp{-\frac{1}{2}\mathbf{u} \cdot \mathbf{C}^{-1}\mathbf{u}} \tag{2b}$$

Here, $\hbar$ is the reduced Planck constant, k_B the Boltzmann constant and T the temperature. The matrix $\mathbf{C}$ expresses the covariance of the atomic displacements in the harmonic approximation, with elements C_{ij} given by

$$C_{ij} = \frac{\hbar}{2\sqrt{M_i M_j}} \sum_\lambda \frac{1}{\omega_\lambda \tanh\frac{\hbar\omega_\lambda}{2k_\text{B}T}} \epsilon_{\lambda,i}\epsilon_{\lambda,j}^*. \tag{3}$$

The EHP free energy is then given by the contributions

$$F_{\text{harm}}(T) = \sum_\lambda \left(\frac{\hbar\omega_\lambda}{2} + k_\text{B}T\log\left[1-\exp\left\{-\frac{\hbar\omega_\lambda}{k_\text{B}T}\right\}\right]\right) \tag{4}$$

$$F_{\text{corr}}(T) = \frac{1}{N} \sum_n [V(\mathbf{u}_n) - \mathcal{V}(\mathbf{u}_n)],\tag{5}$$

where $\mathbf{u}_n$ are the samples from the real-space distribution defined by the density matrix, Equation (2b). If no structural parameters are relaxed, finding the effective force constants which minimize the weighted sum of the least-squares deviations from the calculated forces for those samples is equivalent to a direct minimization of $\mathcal{F}_{\text{EHP}}$ [16] and leads to an EHP providing a upper bound of Equation (1).

The dependence of the density matrix, Equations (2b) and (3), on the eigenvectors defining the EHP, Equation (2a), makes sampling the PES by an iterative algorithm natural. The closed-form expression of the density matrix in Equation (2b) allows an efficient reuse of data points sampled using different covariance matrices for obtaining both the EHP and the correction term [9]. Still, when constructing EHPs for a range of different temperatures and structural phases it would be an obvious advantage to train a relatively inexpensive and transferable surrogate PES on which the iterative sampling can be performed.

2.2 | High-order Taylor expansions using automatic differentiation

One possible surrogate PES explored is the use of high-order Taylor expansions. We write the Taylor expansion of the PES around the equilibrium positions, $\mathbf{r}_0$, as

$$\overset{(n)}{\Phi}(\mathbf{r}) = V(\mathbf{r}_0) + \frac{1}{2!} \sum \frac{\partial^2 V}{\partial r_i \partial r_j}\Big|_{\mathbf{r}_0} \Delta r_i \Delta r_j + \frac{1}{3!} \sum \frac{\partial^3 V}{\partial r_i \partial r_j \partial r_k}\Big|_{\mathbf{r}_0} \Delta r_i \Delta r_j \Delta r_k \ldots\tag{6}$$

where the indices i, j, and k are composite indices labeling a specific ion and Cartesian direction, with $\Delta r_i = r_i - r_{i,0}$ being the displacement from equilibrium.

To construct the expansions, we rely on the automatically differentiable python framework JAX [24, 26]. The JAX framework creates the Jacobian-vector product (JVP) operator, which for a given function, $f(\mathbf{x}): \mathbb{R}^n \mapsto \mathbb{R}^m$, provides a map $\mathbf{J}_f(\mathbf{x})\mathbf{t}: \mathbb{R}^n \mapsto \mathbb{R}^m$ that can be evaluated at an arbitrary input point $\mathbf{x}$ and tangent vector $\mathbf{t}$. The JVP of any composite function, $f(\mathbf{x}) = g(h(\mathbf{x}))$ for which the JVPs of its constituents are known, can be computed via the chain rule

$$\mathbf{J}_f(\mathbf{x})\mathbf{t} = \mathbf{J}_g(h(\mathbf{x}))\mathbf{J}_h(\mathbf{x})\mathbf{t},\tag{7}$$

By starting from $\mathcal{T}_0(\mathbf{r},\mathbf{t}) = V(\mathbf{r}): \mathbb{R}^{3n_{\text{atoms}}} \mapsto \mathbb{R}$, we can hence recursively define a Taylor-series operator

$$\mathcal{T}_k(\mathbf{r},\mathbf{t}) = \frac{1}{k} \frac{\partial \mathcal{T}_{k-1}(\mathbf{r},\mathbf{t})}{\partial \mathbf{r}} \cdot \mathbf{t}\tag{8}$$

and build the order-n Taylor expansion as

$$\overset{(n)}{\Phi}(\mathbf{r}) = \sum_{k=0}^{n} \mathcal{T}_k(\mathbf{r}_0, \Delta\mathbf{r}).\tag{9}$$

The JAX framework creates the JVP operator so that the JVP of a composite function can be evaluated without ever requiring the full Jacobian. The recursion in Equation (9) can thus be performed without evaluating the derivatives, which grow in dimensionality with the order of the expansion. This allows us to build a Taylor expansion of arbitrary order accurate to within machine precision through successive use of the JVP.

2.3 | Lennard-Jonesium

Our first model system is the six-atom LJ Ar cluster. The potential contains a standard LJ contribution from each pair of atoms:

$$V(\mathbf{r}) = 2\varepsilon \sum_{i \neq j} \left[\left(\frac{\sigma}{r_{ij}}\right)^{12} - \left(\frac{\sigma}{r_{ij}}\right)^{6} \right].\tag{10}$$

We take the LJ parameters for Ar ($M = 39.9$ amu) introduced in Reference [37], that is, $\sigma = 3.4$ Å and $\epsilon = 10.3$ meV. All Taylor potentials as expressed through Equation (9) are obtained through the autodiff procedure described above.

We also study thermal expansion in an infinite fcc LJ crystal, which involves deformations not spanned by the basis of normal modes. Due to the long-range nature of the LJ potential, we introduce a smooth cutoff at 12 Å using the bump function bump function [38] defined through a composition of functions that ensure a differentiable, smooth change from zero to one for arguments ranging from the cutoff radius, r_c to the switching radius, r_s:

$$f(x) = \begin{cases} 0 & x \leq 0 \\ \exp{-\dfrac{1}{x}} & x > 0 \end{cases} \tag{11}$$

$$g(x) = \frac{f(x)}{f(x) + f(1-x)} \tag{12}$$

$$h(r, r_s, r_c) = 1 - g\left(\frac{r^2 - r_s^2}{r_c^2 - r_s^2}\right) \tag{13}$$

The chosen cutoff fits well in an $8 \times 8 \times 8$ supercell with the equilibrium lattice parameter $a = 5.272$ Å, corresponding to an equilibrium volume of $V_0 = 36.64$ Å^3. We obtain the force constants for this case using hiPhive [15], which greatly simplifies the handling of space-group symmetries and thus drastically reduces the number of force constants that need to be calculated. We obtain a dataset of 76 800 configurations from supercells with random displacements drawn from a multidimensional normal distribution $\mathcal{N}(\mu = 0$ Å, $\sigma = 0.03$ Å$)$, which we split into training (90%) and test (10%) subsets. All Taylor potentials are fitted to the training set using linear regression and evaluated on the test set. For the fourth-order potential, we chose [12 Å, 7 Å, 6 Å] as the cutoffs for the respective orders; for the sixth order we add [5.5 Å, 5 Å] to that list. Therefore, in contrast to the case of the cluster, the Taylor expansions do not contain all possible terms, which would be infinitely many. The use of short cutoffs for the higher orders is standard practice in force-constant calculations for crystalline solids and both the $\overset{(4)}{\Phi}$ and the $\overset{(6)}{\Phi}$ potentials achieve a coefficient of determination $R^2 \approx 1$ for both the test and training forces, indicating that the models describe the data well. We repeat the fit for 10 volumes from 35.9 to 51.3 Å^3 to obtain free energy versus volume curves for each expansion order.

2.4 | Machine-learned global surrogate

As a first surrogate PES, we employ NeuralIL [35], a neural network force field (NNFF) loosely following the Behler-Parinello template [39]. A cutoff radius of 6 Å, an embedding dimension of 4 and a total number of 28 basis functions are used [35]. Additionally, we include the repulsive term of a Morse-potential

$$V_{\text{rep}}(\mathbf{r}) = \frac{1}{2} \sum_{i \neq j} d \times h(r_{ij}, r_s, r_c) e^{-2a(r_{ij} - b)} \tag{14}$$

where r_{ij} is the distance between atoms i and j, while a, b, d represent parameters optimized during the training process and h is the bump function defined by Equation (13).

We use a fully connected pyramidal NN architecture built from six hidden layers consisting of 128:64:32:16:16:16 neurons with Swish-1 [40] nonlinearities. The learning rate is varied linearly [41] using a one-cycle scheduler that sweeps from 5×10^{-4} to 5×10^{-3} and back, ending with 5×10^{-5} for the last 10% of each epoch. The target of the AdamW-based [42] optimization is a log-cosh loss function [43]. It consists mainly of a force-based loss, but we also include a small portion of the energy-loss to ensure a correct origin of energies

$$\mathcal{L} = \left\langle \frac{0.1 \text{eV\AA}^{-1}}{3 n_{\text{atoms}}} \sum_i^{3 n_{\text{atoms}}} \log\left[\cosh\left(\frac{f_{i,\text{pred}} - f_{i,\text{truth}}}{0.1 \text{eV\AA}^{-1}}\right)\right] + \frac{0.1 \text{eV}}{n_{\text{atoms}}} \log\left[\cosh\left(\frac{E_{\text{pred}} - E_{\text{truth}}}{1 \text{eV}}\right)\right] \right\rangle. \tag{15}$$

The DFT calculations are performed using NWChem [44], with the B3PW91 hybrid approximation to the exchange and correlation terms of the Hamiltonian [45]. Silver atoms are described using a double-ζ LANL2DZ basis set [46] with the valence configuration $4s^2 4p^6 4d^{10} 5s^1$ and the remaining electrons were treated with an effective core potential. The iterative calculations require between 20 and 45 iterations of the procedure outlined in the methods section. In each iteration, 100 structures are added. The convergence of $\mathcal{F}_{EHP}$, as well as of the frequencies, f_i is shown for the example of $T = 1000$ K in Figure 1. A run is considered converged if none of the frequencies change by more than 5×10^{-3} for
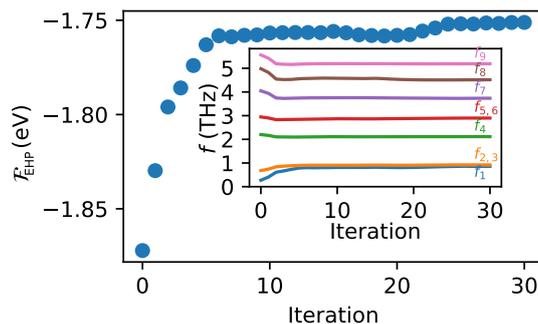
FIGURE 1   Convergence of $\mathcal{F}_{EHP}$ and of the frequencies (inset), as a function of iterations using the NWChem backend. In every iteration, 100 new structures are added.
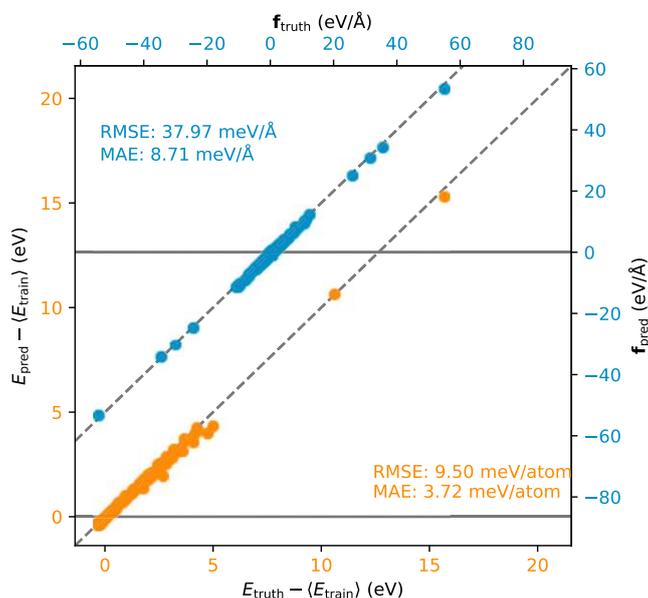

FIGURE 2   Truth versus prediction plot of the energy (orange) and forces (blue) of the 1000 validation data points produced by the NNFF trained on 2000 points. The gray dashed lines indicate perfect agreement.

three successive iterations. In the first iteration we draw 100 structures from a normal distribution, $\mathcal{N}\left(\mu = 0\,\text{Å}, \sigma = 10^{-3}\,\text{Å}\right)$, to start the calculation off.

The potential for the high-symmetry singlet five-atom orthogonal ditrigonal cationic silver cluster [47] is trained for 500 epochs using 2000 training and 1000 validation data points. The data points are randomly chosen out of over 20 000 generated during the NWChem-backed EHP calculations performed for temperatures 10, 50, 70, 100, 500, 750 and 1000 K. The validation statistics suggest a high-performance parametrization through the NNFF with a mean average error (MAE) in the forces of $\text{MAE} = \left\langle \left\| \mathbf{f}_{pred} - \mathbf{f}_{truth} \right\| \right\rangle = 8.7\,\text{meV}\,\text{Å}^{-1}$ and a per-atom energy MAE of $\frac{1}{n_{atoms}} \left\langle \left\| E_{pred} - E_{truth} \right\| \right\rangle = 3.7\,\text{meV}\,\text{atom}^{-1}$ on the validation data-set. This compares favorably to the standard deviations of the forces ($1\,\text{eV}\,\text{Å}^{-1}$) and energies ($180\,\text{meV}\,\text{atom}^{-1}$). The corresponding parity plot is shown in Figure 2.

3 | RESULTS AND DISCUSSION

3.1 | Lennard-Jones cluster

In its ground state, the six-atom LJ Ar cluster adopts the structure of a regular octahedron, with all six atoms at symmetrically equivalent positions [48]. Keeping those positions constant, we obtain Taylor approximants, Equation (9), from second up to eighth order. We then scan a range of
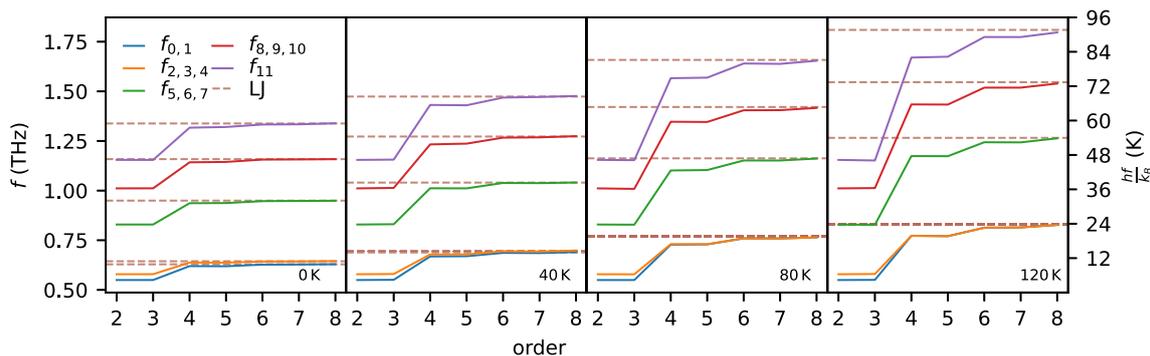
FIGURE 3 Comparison of frequencies obtained for different orders and temperatures. The dashed lines in brown indicate the results obtained using the exact potential.

temperatures up to 120 K, corresponding to the reduced temperature $k_B T/\epsilon \simeq 1$. At each temperature and for each Taylor expansion we obtain an EHP following the procedure described above for the exact potential.

Figure 3 shows the frequencies extracted from those EHPs at four representative temperatures. Due to the high symmetry of the structure, only five unique non-zero eigenvalues exist. This can be understood in terms of the irreducible representations of the point group O_h. The vibrational representation can be decomposed in terms of one singly degenerate (A_{1g}), one doubly degenerate (E_g) and three triply degenerate (T_{2g}, T_{1u}, T_{2u}) irreducible representations

$$\Gamma_{\text{vib}} = A_{1g} + 2E_g + 3T_{2g} + 3T_{1u} + 3T_{2u},$$

The breathing mode of the cluster, with the highest frequency, is nondegenerate and the two lowest-lying modes are doubly degenerate. The remaining nine modes appear in groups of three.

The second-order Taylor expansion is a plain harmonic approximation to the PES and naturally leads to temperature-independent frequencies. Moreover, from perturbation theory it is known that the term of odd order $2n+1$ does not introduce changes in the frequencies with respect to the even term $2n$. The results of our numerical procedure are consistent with both facts, which serve as a test of the method.

The errors of the frequencies obtained from the EHPs, corresponding to each of the Taylor expansions, as compared to those of the exact potential EHP, increase with increasing temperature and decrease with increasing order. They range from 13.7% to 36% for the $\Phi^{(2)}$ and 0.01% to 0.7% for the $\Phi^{(8)}$ potential over the investigated temperature interval from 0 to 120 K. Strictly speaking, the 0 K frequencies do not include any temperature effects, but still take into account the dynamics of the ground state of each normal mode considered as a quantum harmonic oscillator, reflected in the first term of Equation (4). Even this ground-state effect is grossly underestimated when considering only the harmonic approximation of the PES. The fourth-order expansion has an acceptable error for intermediate temperatures up to 50 K, particularly for the low-frequency modes, which have the largest impact on the free energy. For higher temperatures, at least the sixth order has to be taken into account as well. Obtaining such an expansion for actual solids is a daunting task, given the increase in the number of derivatives with their order and the number of force calculations required by each of them.

It is noticeable how all frequencies increase monotonically with the order of the expansion. The Taylor expansions thus systematically lead to a softer EHP than the actual potential. That phenomenon points to a dominant effect of the strongly repulsive $\propto r^{-12}$ short-range component of the pair potential, which grows more steeply toward $r \to 0$ than any power law with a positive exponent can capture.

Looking at the convergence of the approximation to the free energy $\mathcal{F}_{\text{EHP}}$ and its components, F_{corr} and F_{harm}, in Figure 4, we find a similar situation as for the frequencies. Due to the failure of the second-order expansion to reproduce the low frequencies even at low temperatures, it performs an order of magnitude worse than $\Phi^{(4)}$ and we thus omit it from the figure. The error in the harmonic portion, F_{harm} (top panel) is negative in accordance with the observation of the Taylor approximations yielding too soft EHPs. F_{corr} partially compensates that negative error (see middle panel) but the negative part of the error still dominates (bottom panel). As expected, the quality of the approximation decreases with increasing temperature for all orders.

This behavior with respect to temperature is rooted in the intrinsically local character of power-series expansions. In particular, the highest-order term always dominates as the distance from the center of the approximation grows, and the model potential can only tend toward $\pm\infty$. The effect is apparent even for the simple problem of the LJ Ar dimer, with a minimum at an interatomic distance of $2^{1/6}\sigma = 3.82$ Å. The intervals within which a Taylor model incurs an error of 25% or less are 3.53–4.22 Å for $\Phi^{(2)}$, 3.43–4.37 Å for $\Phi^{(4)}$, 3.40–4.51 Å for $\Phi^{(6)}$, 2.51–4.65 Å for $\Phi^{(8)}$, and 2.28–4.77 Å for $\Phi^{(10)}$. The gains with each order become subsequently more and more modest and extremely high orders would be required to get a significant improvement with respect to the basic harmonic approximation.
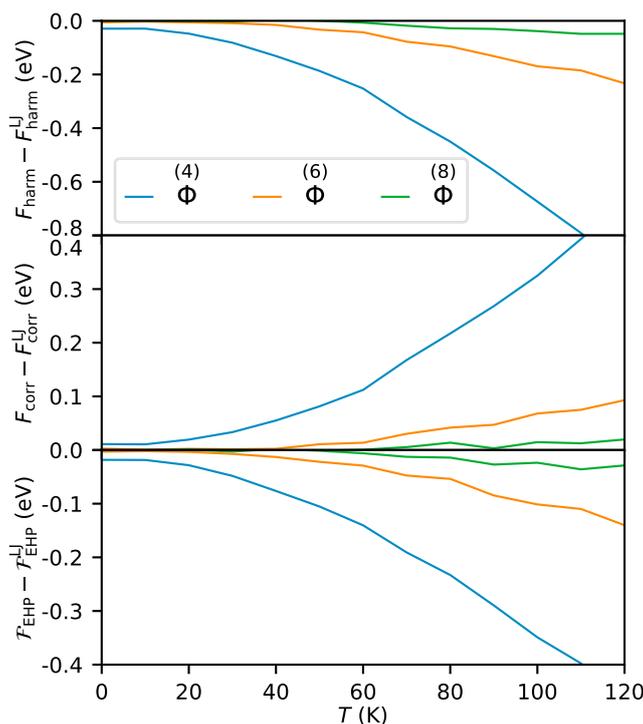
FIGURE 4 Comparison between the differences of the contributions to the free energy obtained with Taylor potentials of different orders and those obtained using the exact potential, as a function of temperature. Top: harmonic contribution; middle: correction term; bottom: full free energy.

To highlight the practical relevance of the shortcomings of power series as global models of the energy landscape, we consider the probability that a point sampled from a high-dimensional Gaussian distribution falls outside of the core region of that distribution along at least one axis. In one dimension most of the mass of a Gaussian distribution lies in that core region: for instance, there is a 95% probability that a random deviate falls within 2σ of the mean. However, for an isotropic D-dimensional Gaussian, the maximum of the density function is located $\sigma\sqrt{D-1}$ away from the origin and the probability that a random deviate falls in a hyperspherical shell of half-width 2σ around this value is at least 95%. As a consequence, even for rather moderate D the probability of sampling regions of space where the Taylor expansion is a poor approximant is overwhelming. This applies both to the sampling procedure used in this article and to the trajectory of a system simulated using molecular dynamics, for which series expansions are sometimes also used as model potentials. This observation is not tied to the choice of potential; it rather stems from the intrinsic inadequacy of evaluating samples drawn from a particular high-dimensional probability distribution with local expansions. It thus applies in an identical fashion to real solids treated with DFT or any other higher-level theory, necessitating at least a careful analysis of the applicability of such an expansion for the particular situation in question.

3.2 | Lennard-Jones solid

Considering now the fcc LJ crystal, the temperature-dependent behavior of the vibrational spectrum at a given volume shows some features reminiscent of the finite systems. As shown in Figure 5, the frequencies evaluated at a fixed volume are significantly hardened by temperature, and the effect of the ground-state harmonic motion is already very significant.

An expansion of the solid leads to a softening of the LJ interactions, and therefore, to lower effective frequencies. The harmonic portion of the free energy decreases accordingly, resulting in a temperature dependent minimum of $\mathcal{F}_{EHP}$. To study the impact of temperature on the equilibrium volumes in the framework of the Taylor-series surrogate PESs we compute Taylor-series for each volume. We then obtain the EHPs as described in the methodology section for each combination of volumes and Taylor-series expansions. Finally, the contribution to the free energies are obtained according to Equations (4) and (5) and the corresponding free-energy-volume curves, as well as a breakdown of $\mathcal{F}$ into its contributions F_{harm} and F_{corr} can be found in the Supporting Information.

We perform this procedure for $\overset{(2)}{\Phi}$, $\overset{(4)}{\Phi}$, and $\overset{(6)}{\Phi}$. The simplest approximation is a volume-dependent $\overset{(2)}{\Phi}$, which corresponds to the widely used quasi-harmonic approximation (QHA). We depict the thermal expansion behavior calculated with three EHPs in Figure 6, as well as the thermal
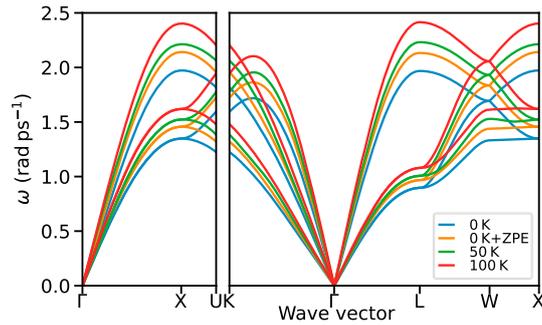
FIGURE 5 The three phonon bands of the Lennard-Jones fcc solid at the 0 K volume as a function of temperature in the regular harmonic approximation, at 0 K including zero-point motion, at 50 K, and at 100 K.
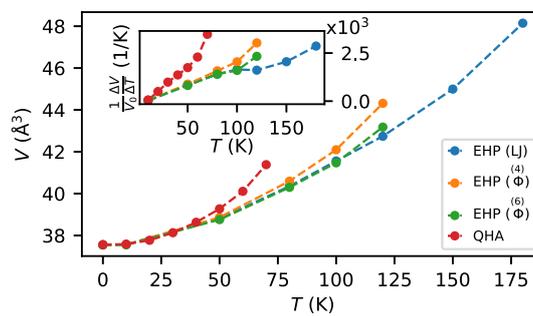
FIGURE 6 Thermal expansion as a function of temperature of the Lennard-Jones fcc solid as obtained from the QHA and the EHPs constructed from $\Phi^{(4)}$, $\Phi^{(6)}$ and the exact potential. The inset shows the thermal expansion coefficient.
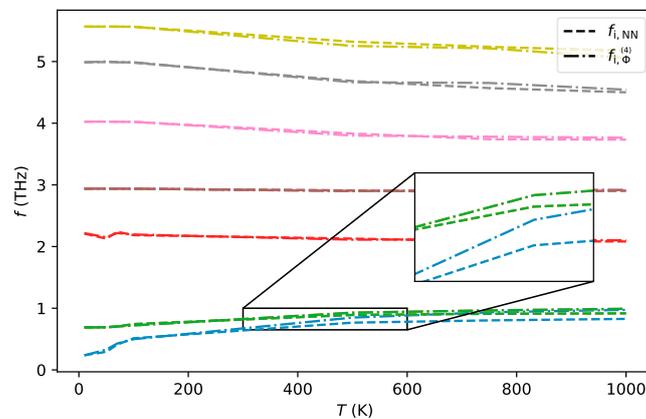
FIGURE 7 Normal-mode frequencies of the 5-atom ditrigonal orthogonal silver cluster as a function of temperature, obtained by constructing EHPs using a NNFF (dashed lines) and the fourth-order Taylor expansion thereof, $\Phi^{(4)}$ (dash-dotted lines).

expansion coefficient in the inset. Clearly, the QHA performs poorly even for low temperatures and grossly overestimates thermal expansion. Moreover it only predicts a minimum up to $T > 70$ K, after which the free energy decreases monotonically with increasing volume. The fourth order improves on this slightly, yielding minima up to $T = 120$ K, but performance is comparable to the exact calculation only up to about 80 K. The results of the $\Phi^{(6)}$ potential are reasonably close to those obtained using the exact LJ energy function. However, it yields instabilities for large volumes and temperatures. If the thermodynamics of this solid were studied on the basis of, one could wrongly conclude from this that a phase transition will happen at this temperature and volume. None of the approximations yielded results above the characteristic temperature of the system, $T = 120$ K, where the depth of the well, ϵ, is comparable to $k_B T$.
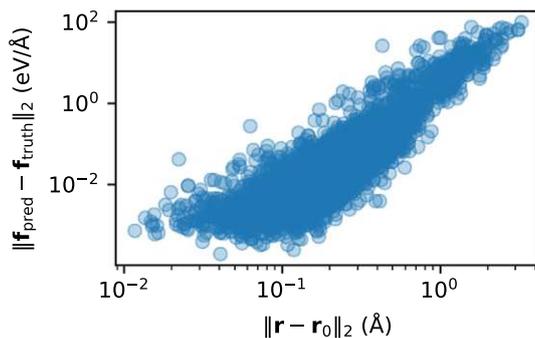
FIGURE 8 RMSE of the force prediction of $\overset{(4)}{\Phi}$ as a function of displacement magnitude for the validation data. Points not obtained through Equation (2b) (i.e., the first iterations) were omitted.

3.3 | The silver cluster

To validate these studies for a system studied using DFT calculations we use the high-symmetry singlet five-atom orthogonal ditrigonal cationic silver cluster [47]. Comparing the results of the NWChem-backed EHP runs to those obtained using the NNFF as described in the methods section, we find an excellent agreement of the frequencies. Furthermore, a relative mean squared error in the Hessians of less than 1% is achieved, indicating that the NNFF has captured the small displacement regime as well. Thus, the assumption that the NNFF can indeed be used as a global surrogate is warranted.

This enables us to study power-series expansions of the PES using the NNFF instead of DFT. Notably, Equation (9) does not impose any requirements on the function aside from smoothness and differentiability, both of which are ensured by the construction of the NNFF as described in detail in Reference [35]. For brevity, we limit ourselves to the fourth-order Taylor expansion, $\overset{(4)}{\Phi}$, which is obtained through Equation (9). On first glance at Figure 7, the frequencies obtained using $\overset{(4)}{\Phi}$ the potential satisfactorily reproduce those obtained through the NNFF. Nonetheless, despite the system being much less anharmonic than the LJ systems, the $\overset{(4)}{\Phi}$ prediction of the lowest-frequency mode deviates significantly from the NNFF- and NWChem-backed result starting at about 300 K, as shown in the inset. Specifically, it approaches the doubly degenerate second-lowest mode, suggesting that the $\overset{(4)}{\Phi}$ potential induces a spurious symmetry.

A different perspective on this can be gathered by looking at the relationship between the root-mean-squared error (RMSE) of the forces and the L_2-norm of the displacements in the validation dataset, as depicted in Figure 8. Here, the local nature of the expansion and the fact that polynomials can only tend to for large values of their argument become visible—the further away from the equilibrium positions $\overset{(4)}{\Phi}$ is evaluated, the worse the accuracy. Using the inverse relationship of normal-mode frequency and thermal amplitude implied by the equipartition theorem, this inaccuracy can give an approximate but intuitive explanation for the deviation of the lowest-frequency mode.

4 | CONCLUSION

Using finite and periodic Lennard-Jones systems as easy-to-analyze benchmarks with significant anharmonicity, we explore the usefulness of power-series expansions as relatively inexpensive surrogate models of the potential energy surface for the evaluation of thermodynamical properties at finite temperatures via effective harmonic potentials.

A general trend is that local expansions of the potential energy surface fail to capture important behavior at elevated temperatures and even corrections due to zero-point motions of the ions. In particular, the very commonly used quasi-harmonic approximation performs very poorly when it comes to predicting thermodynamic properties. High-order (sextic or octic) expansions are required to obtain reasonable results at intermediate temperature, but even those can introduce significant artifacts like instabilities and nonexistent phase transitions. Furthermore, in real solids, the number of parameters contained in those high-order Taylor approximants renders them unfeasible or only possible with very short cutoffs.

The root cause of such deficiencies is the behavior of polynomials for large values of their arguments, together with the fact that a many-body system at finite temperature has a significant probability of visiting areas far from the minimum of the potential energy surface. One possible solution to this issue is the parametrization of the potential energy surface using a global description, such as the neural network employed in this manuscript. Not only are such global approximants more flexible and transferable to different numbers of atoms or, in the case of solids, different cell shapes, they also require less computational effort than, for example, a full third-order, or partial fourth-order correction.

AUTHOR CONTRIBUTIONS

FUNDING INFORMATION

DATA AVAILABILITY STATEMENT

Data available on request from the authors.

ORCID

Sebastian Bichelmaier https://orcid.org/0000-0003-0997-8788

Georg K. H. Madsen https://orcid.org/0000-0001-9844-9145

REFERENCES

[1] D. Hooton, _Lond. Edinb. Dublin Philos. Mag. J. Sci._ **1955**, _46_, 422.

[2] N. Shulumba, O. Hellman, Z. Raza, B. Alling, J. Barrirero, F. Mücklich, I. A. Abrikosov, M. Odén, _Phys. Rev. Lett._ **2016**, _117_, 205502.

[3] A. van Roekeghem, J. Carrete, N. Mingo, _Phys. Rev. B_ **2016**, _94_, 20303.

[4] R. Stern, G. K. H. Madsen, _Phys. Rev. B_ **2016**, _94_, 144304.

[5] S. Ehsan, M. Arrigoni, G. K. H. Madsen, P. Blaha, A. Tröster, _Phys. Rev. B_ **2021**, _103_, 94108.

[6] T. Tadano, S. Tsuneyuki, _J. Ceram. Soc. Jpn._ **2019**, _127_, 404.

[7] A. Dewandre, O. Hellman, S. Bhattacharya, A. H. Romero, G. K. H. Madsen, M. J. Verstraete, _Phys. Rev. Lett._ **2016**, _117_, 276601.

[8] L. Monacelli, R. Bianco, M. Cherubini, M. Calandra, I. Errea, F. Mauri, _J. Phys. Condens. Matter_ **2021**, _33_, 363001.

[9] S. Bichelmaier, J. Carrete, M. Nelhiebel, G. K. H. Madsen, _Phys. Status Sol. RRL_ **2022**, _16_, 2100642.

[10] P. Souvatzis, O. Eriksson, M. I. Katsnelson, S. P. Rudin, _Phys. Rev. Lett._ **2008**, _100_, 95901.

[11] O. Hellman, I. A. Abrikosov, S. I. Simak, _Phys. Rev. B_ **2011**, _84_, 180301.

[12] O. Hellman, P. Steneteg, I. A. Abrikosov, S. I. Simak, _Phys. Rev. B_ **2013**, _87_, 104111.

[13] I. Errea, M. Calandra, F. Mauri, _Phys. Rev. Lett._ **2013**, _111_, 177002.

[14] I. Errea, M. Calandra, F. Mauri, _Phys. Rev. B_ **2014**, _89_, 64302.

[15] F. Eriksson, E. Fransson, P. Erhart, _Adv. Theory Simul._ **2019**, _2_, 1800184.

[16] A. van Roekeghem, J. Carrete, N. Mingo, _Comput. Phys. Commun._ **2021**, _263_, 107945.

[17] K. Esfarjani, G. Chen, H. T. Stokes, _Phys. Rev. B_ **2011**, _84_, 85204.

[18] F. Zhou, W. Nielson, Y. Xia, V. Ozoliņš, _Phys. Rev. Lett._ **2014**, _113_, 185501.

[19] T. Tadano, S. Tsuneyuki, _Phys. Rev. B_ **2015**, _92_, 54301.

[20] T. Tadano, S. Tsuneyuki, _J. Phys. Soc. Jpn._ **2018**, _87_, 41015.

[21] P. Schwerdtfeger, R. Tonner, G. E. Moyano, E. Pahl, _Angew. Chem., Int. Ed._ **2016**, _55_, 12200.

[22] C. S. Barrett, L. Meyer, _J. Chem. Phys._ **1964**, _41_, 1078.

[23] B. A. Pearlmutter and J. M. Siskind, in _Proceedings of the 2007 Symposium on Principles of Programming Languages_ (Nice, France, 2007), pp. 155–60.

[24] J. Bettencourt, M. J. Johnson, and D. Duvenaud, in _Program Transformations for ML Workshop at NeurIPS 2019_ (**2019**), https://openreview.net/forum?id=SkxEF3FNPH.

[25] A. G. Baydin, B. A. Pearlmutter, A. A. Radul, J. M. Siskind, _J. Mach. Learn. Res._ **2018**, _18_, 1.

[26] J. Bradbury, R. Frostig, P. Hawkins, M. J. Johnson, C. Leary, D. Maclaurin, G. Necula, A. Paszke, J. VanderPlas, S. Wanderman-Milne, Q. Zhang, JAX: Composable Transformations of Python+NumPy Programs (**2018**), http://github.com/google/jax.

[27] C. Verdi, F. Karsai, P. Liu, R. Jinnouchi, G. Kresse, _NPJ Comput. Mater._ **2021**, _7_, 156.

[28] J. Wu, Y. Zhang, L. Zhang, S. Liu, _Phys. Rev. B_ **2021**, _103_, 24108.

[29] F. Noé, A. Tkatchenko, K.-R. Müller, C. Clementi, _Annu. Rev. Phys. Chem._ **2020**, _71_, 361.

[30] C. Zeni, K. Rossi, T. Pavloudis, J. Kioseoglou, S. de Gironcoli, R. E. Palmer, F. Baletto, _Nat. Commun._ **2021**, _12_, 6056.

[31] D. Schwalbe-Koda, A. R. Tan, R. Gómez-Bombarelli, _Nat. Commun._ **2021**, _12_, 5104.

[32] S. S. Schoenholz, E. D. Cubuk, _Advances in Neural Information Processing Systems_, Vol. 33, Curran Associates, Inc., Red Hook, NY **2020**. https://papers.nips.cc/paper/2020/file/83d3d4b6c9579515e1679aca8cbc8033-Paper.pdf

[33] M. F. Kasim, S. Lehtola, S. M. Vinko, _J. Chem. Phys._ **2022**, _156_, 84801.

[34] P. Schwerdtfeger, A. Burrows, O. R. Smits, _J. Phys. Chem. A_ **2021**, _125_, 3037.

[35] H. Montes-Campos, J. Carrete, S. Bichelmaier, L. M. Varela, G. K. H. Madsen, _J. Chem. Inf. Model._ **2022**, _62_, 88.

[36] A. Isihara, _J. Phys. A_ **1968**, _1_, 539.

[37] A. Rahman, *Phys. Rev.* **1964**, *136*, A405.

[38] L. W. Tu, *An Introduction to Manifolds*, Springer-Verlag, New York, NY **2011**.

[39] J. Behler, *J. Chem. Phys.* **2016**, *145*, 170901.

[40] P. Ramachandran, B. Zoph, Q. V. Le, *arXiv e-Prints* arXiv:1710.05941 **2017**.

[41] L. N. Smith, arXiv:1803.09820 [cs.LG], 1803.09820, **2018**.

[42] I. Loshchilov and F. Hutter, in International Conference on Learning Representations, **2019**, https://openreview.net/forum?id=Bkg6RiCqY7.

[43] Q. Wang, Y. Ma, K. Zhao, Y. Tian, *Ann. Data Sci.* **2022**, *9*, 187.

[44] E. Aprà, E. J. Bylaska, W. A. de Jong, N. Govind, K. Kowalski, T. P. Straatsma, M. Valiev, H. J. J. van Dam, Y. Alexeev, J. Anchell, V. Anisimov, F. W. Aquino, R. Atta-Fynn, J. Autschbach, N. P. Bauman, J. C. Becca, D. E. Bernholdt, K. Bhaskaran-Nair, S. Bogatko, P. Borowski, J. Boschen, J. Brabec, A. Bruner, E. Cauët, Y. Chen, G. N. Chuev, C. J. Cramer, J. Daily, M. J. O. Deegan, T. H. Dunning Jr., M. Dupuis, K. G. Dyall, G. I. Fann, S. A. Fischer, A. Fonari, H. Früchtl, L. Gagliardi, J. Garza, N. Gawande, S. Ghosh, K. Glaesemann, A. W. Götz, J. Hammond, V. Helms, E. D. Hermes, K. Hirao, S. Hirata, M. Jacquelin, L. Jensen, B. G. Johnson, H. Jónsson, R. A. Kendall, M. Klemm, R. Kobayashi, V. Konkov, S. Krishnamoorthy, M. Krishnan, Z. Lin, R. D. Lins, R. J. Littlefield, A. J. Logsdail, K. Lopata, W. Ma, A. V. Marenich, J. Martin del Campo, D. Mejia-Rodriguez, J. E. Moore, J. M. Mullin, T. Nakajima, D. R. Nascimento, J. A. Nichols, P. J. Nichols, J. Nieplocha, A. Otero-de-la-Roza, B. Palmer, A. Panyala, T. Pirojsirikul, B. Peng, R. Peverati, J. Pittner, L. Pollack, R. M. Richard, P. Sadayappan, G. C. Schatz, W. A. Shelton, D. W. Silverstein, D. M. A. Smith, T. A. Soares, D. Song, M. Swart, H. L. Taylor, G. S. Thomas, V. Tipparaju, D. G. Truhlar, K. Tsemekhman, T. van Voorhis, Á. Vázquez-Mayagoitia, P. Verma, O. Villa, A. Vishnu, K. D. Vogiatzis, D. Wang, J. H. Weare, M. J. Williamson, T. L. Windus, K. Woliński, A. T. Wong, Q. Wu, C. Yang, Q. Yu, M. Zacharias, Z. Zhang, Y. Zhao, R. J. Harrison, *J. Chem. Phys.* **2020**, *152*, 184102.

[45] A. D. Becke, *J. Chem. Phys.* **1993**, *98*, 5648.

[46] P. J. Hay, W. R. Wadt, *J. Chem. Phys.* **1985**, *82*, 299.

[47] T. Yumura, M. Kumondai, Y. Kuroda, T. Wakasugi, H. Kobayashi, *RSC Adv.* **2017**, *7*, 4950.

[48] M. Hoare, P. Pal, *Adv. Phys.* **1971**, *20*, 161.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.