Evaluation of neutral network methodologies for understanding the interface phenomena in integrated materials in microelectronics : application to the case of solid phase epitaxy in silicon

Ruggero Lot

Université Fédérale
Toulouse Midi-Pyrénées

THÈSE

En vue de l'obtention du
DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE
Délivré par l'Université Toulouse 3 - Paul Sabatier

Présentée et soutenue par

Ruggero LOT

Le 12 décembre 2022

Évaluation des méthodologies de réseaux neuronaux pour la compréhension des phénomènes dinterface dans les matériaux intégrés dans les domaines de la microélectronique : application au cas de lépitaxie en phase solide dans le silicium

Ecole doctorale : **GEETS - Génie Electrique Electronique,Télécommunications et Santé : du système au nanosystème**

Spécialité : **MicroNano Systèmes**

Unité de recherche :
LAAS - Laboratoire d'Analyse et d'Architecture des Systèmes

Thèse dirigée par
Anne HEMERYCK et Stefano De Gironcoli

Jury

M. Normand MOUSSEAU, Rapporteur
Mme Francesca BALETTO, Rapporteure
Mme Milica TODOROVIC, Examinatrice
Mme Magali BENOIT, Examinatrice
Mme Anne HEMERYCK, Directrice de thèse
M. Stefano DE GIRONCOLI, Co-directeur de thèse

DOCTORAL THESIS

Evaluation of neural network methodologies for understanding the interface phenomena in integrated materials in microelectronics: application to the case of solid phase epitaxy in Silicon

ECOLE DOCTORALE:
GEETS - GÉNIE ÉLECTRIQUE, ÉLECRONIQUE, TÉLÉCOMMUNICATIONS ET SANTÉ: DU SYSTÈME AU NANOSYSTÈME
SPÉCIALITÉ: MICRONANO SYSTÈMES

Author:

Ruggero Lot
ruggero.lot@laas.fr

Advisor:

Anne Hémeryck
anne.hemeryck@laas.fr

Co-advisor:

Stefano Maria de Gironcoli
stefano.degironcoli@sissa.it

2022

LAAS-CNRS
7 avenue du colonel Roche, Toulouse - France

Contents

Acknowledgements

I am deeply indebted to my professors Stefano and Anne for their guidance during this project, and the time they dedicated to me in an uncountable number of discussions that lead to invaluable advice without which this work would not have been possible. Words cannot express my gratitude to Layla, who gave me the opportunity to start this journey that now is coming to an end.

I'm extremely grateful to Francesca and Normand for their precious comments and valuable suggestions on the manuscript. In particular, I would like to thank Normand for the precious lessons he taught me when our paths crossed at the beginning of my Ph.D.

I am also thankful to Franco, Emine, Yusuf, Gabriela, Miha, and Nico for their editing, moral support, and long walks accompanied by philosophical debates on the importance of science and how research should be conducted.

I would like to extend my sincere thanks to my beloved Stella, and my dearest friends Luca, Mirco, and Federico, for always being present through this part of my life.

Lastly, I would be remiss in not mentioning my mother and sister, I would not be where I am without you.

Introduction

Reducing chip dimensions has been one of the main challenges of the last twenty years in chip manufacturing. Smaller chips can reach higher frequencies, reduced latency, capacitance and inductance, and moving data for shorter distances significantly lower power consumption. All these advances reflect in better performance, which can improve our everyday life: from Internet of Things devices that last longer to high-performance facilities for calculations we only dreamed about one decade ago.

At the beginning of the 21st century, it was realized that reducing the size of planar transistors was no longer a viable solution to increase transistor density, and manufacturers began exploring the possibility of using the third dimension in making 3D transistors (an illustrative example is presented in fig. 1.). These 3D transistors were under development by the scientific community since the nineties, and thanks to their transposition from the laboratory to the production line they are now a standard in chip manufacturing.

Size reduction is reaching a point where quantum mechanical effects are not negligible and short channel effect and current leakage are now problems that can not be easily overcome anymore. As a result the industry is looking for new ways to take advantage of the third dimension.

As of today, the approaches taken by different firms are different flavors of chip stacking. In general, with this technology, chips, better defined as dies, implementing different functions, like CPU and storage, are stacked one on top of the other and soldered together. The electrical interconnections between dies are obtained through through-silicon-vias (TSV) etched into a silicon wafer and filled with conductive material. TSV employs a concept that has been known since the sixties, and the major limitation that this approach poses to future process scalability is the size of the vertical electrical connections (vias) since this quantity can not be arbitrarily reduced. The main limit is given by the manufacturing process that lacks the accuracy to align the dies within the micrometer scale required when soldering them together. [1]

The long-term solution to this problem is to move away from die stacking to

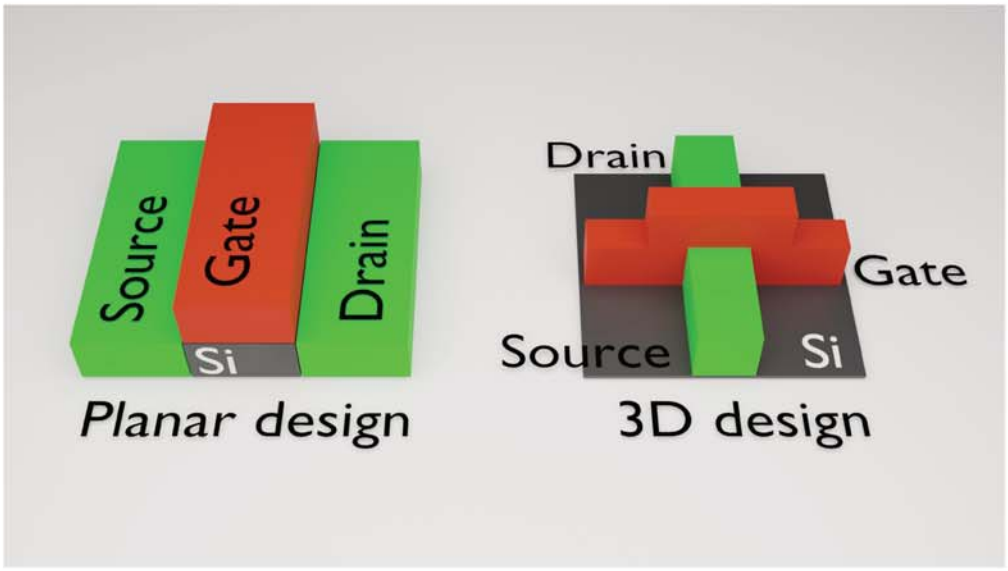


Figure 1: On the left side, we present a planar transistor design. In this design, we have the source/drain regions obtained by doping areas below the wafer's surface and the gate built on top. On the right, we can see a non-planar transistor, or "3D" transistor, referred to as fin field-effect transistor (FinFET). This name was given to this transistor geometry because the shape of the source/drain regions closely resembles "fins" on the silicon surface. While more complex to manufacture, this second geometry allows for faster switching times, higher current density, and higher transistor density. These, with other advantages, have made this geometry the basis for modern electronic semiconductor device fabrication.

a 3D monolithic chip integration. 3D monolithic integration starts from a regular die and instead of soldering a second die on top of it proceeds by covering it with an insulating material that allows for the direct construction of a new die on top of the buried one. The construction of the second die starts from the deposition of a new thin layer of silicon that acts as base to build the transistors. Once those are made, vias can be hatched directly in the insulator, and the process can be repeated [2]. Figure 2 shows a typical dies stack and the theoretical equivalent using 3D monolithic integration. This technique is still under active development and unlocks the third dimension in a way that is very similar to die stacking but solves the alignment issue to an unprecedented accuracy, reaching transistor densities that are two orders of magnitude larger than with conventional approaches [3].

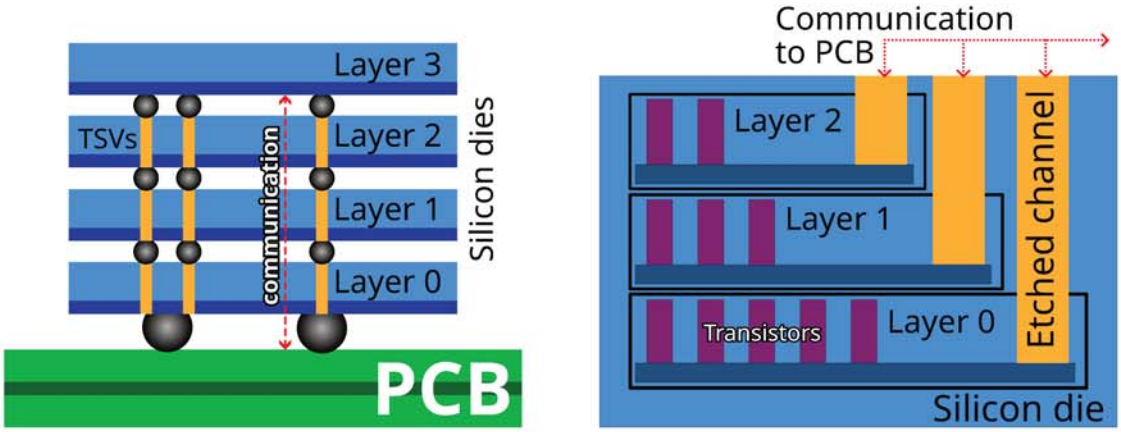


Figure 2: A typical stack of dies connected by TSV is presented on the left. Each layer is tenths of μm thick, and the size of the TSV has to be hundreds of nm wide because of the limited alignment precision available during the stacking process. On the right, an equivalent structure is embedded in a single silicon die where each layer is hundreds of nm thick. In this case, the communication between layers and the PCB is granted by etched channels less than 100nm. The alignment problems are directly solved while manufacturing the chip. This approach can increase transistor density by two orders of magnitude.

Functional prototypes have already been developed in the laboratory and the major technological issues are coming from the need to use low-temperature processing while making the transistors in the upper layers. This is necessary in order not to alter the devices present in the underlying layers, degrading their performances. Among all the processes involved during transistor fabrication, dopant activation is a key and delicate one that is usually performed at high

temperature. This result is commonly achieved through rapid thermal processing which heats silicon at temperatures exceeding 1300K for a few seconds, a temperature that in the case of 3D monolithic integration is sufficient to alter the properties of the device. A promising way to achieve excellent dopant activation levels in silicon is Solid Phase Epitaxy (SPE) at low temperature (900K) for several minutes after low-temperature implantation.

In this case, when dopants are added to silicon by high-energy implantation, the cascades resulting from atoms knocking add stress and point defects to the initial crystal structure, amorphizing it. The implantation energies used are however not enough to amorphize all the material and regions of perfect crystal remain. SPE involves the transition from the amorphous to the crystalline phases of the material. Raising the temperature of the sample triggers the epitaxial recrystallization of the amorphous layer, responsible for dopants activation in the implanted regions.

A detailed knowledge of the involved processes is desirable. Studying this problem in its whole complexity, using only quantum mechanics, is impossible even with modern computational capabilities and a multi-scale approach is needed, where different levels of theory are used to tackle different time scales.

In a multi scale approach more accurate but expensive theories are used to generate parameters for more approximate but computationally lighter theories that can tackle longer timescales. In the case of SPE simulations, a realistic timescale of the order of minutes can only be achieved with kinetic Monte-Carlo approaches. A typically employed approach is the on-lattice Monte-Carlo method. This approach starts considering the atoms on an oriented crystalline lattice and labels individual atoms crystalline or amorphous according to their position in the sample. Then, defining a catalog of events that can change the label of a given atom and the corresponding probability rates, the sample recrystallization can be simulated. This approach relies on the knowledge of the key events and of their probabilities and it falls short in describing phenomena that are characterized by events that are not included in the catalog. As of today, practically implemented on-lattice Monte-Carlo methods for SPE are based on empirically defined events that decide whether an atom at the crystal/amorphous interface can become crystalline with a probability that depends on the nature of its nearest neighbors and is directly fit to experimental results. This approach has limited predicting power in many respects and an active line of research is the search for these key events by employing techniques that explore the potential energy surface of the system when atoms are not confined to reside on a lattice.

Different attempts have been made in this direction in the last decades but

their success was limited by the accuracy of the employed potential energy surface. The investigations performed so far were based on energy models obtained from relatively simple empirical interatomic potentials, with a few parameters adjusted to reproduce specific lattice geometries and/or a few known experimental properties, but seldom on extensive accurate quantum mechanical calculations. In order to create a more accurate catalog of elemental events and improve existing simulation protocols a more accurate description of the potential energy surface is needed.

In this work we will pursue this program by exploring the ability of modern machine-learned interatomic potentials, based on neural networks trained on large datasets of accurate quantum-mechanical density functional theory calculations, to provide an overall accurate description of the properties of Silicon and a realistic description of the SPE.

In the first chapter, the SPE will be introduced in details from an experimental prospective covering properties such as the experimental setups and the experimental observables, that will be used to validate the final results. Then, Density Functional Theory, the quantum mechanical framework used to calculate the potential energy surface of interest, will be introduced discussing in particular the major approximations that may impact the obtained results, such as the exchange-correlation functional of choice and the electronic temperature. Finally, how artificial neural network can be used to generate an interatomic potential that retains the accuracy of DFT at a fraction of the cost will be discussed. The building steps of this approach will be explained in some details, such as the collection of the training dataset that defines where the potential energy surface is accurate, the specific description of the local environment on which the neural network acts on, and how the fit of this complex object is performed.

The second chapter will detail the creation of the potential developed in this work. First, it will cover the creation of the training dataset. This is done in an iterative way where first simple potentials are used to explore the phase space relevant for SPE that includes configuration such as the diamond bulk, the amorphous and the liquid phases, some surfaces. Energy and forces of the sampled configurations are then reevaluated with density functional theory and used to train a first neural network potential. This potential is then used to sample more representative configurations that are included in the training dataset, which is used to train newer, more realistic, potentials. Then the chapter will continue by showing the key aspects of the training procedure that lead to the final potential, and will finish by discussing its performance. The validation of the potential will be first done computing the standard machine-learning metrics over a set-aside fraction of data, not used during the training phase, and then checking a few selected physical properties that are valuable for SPE such

as the ability to produce good structure factors for the amorphous, an accurate melting temperature, stability of the main surface reconstructions, and realistic crystalline phonon dispersions.

Solid phase epitaxy will then be investigated in the third chapter. Here, a detour is first taken to discuss in detail the quality of the amorphous phase that the developed potential can generate with the procedure we follow to create it. This investigation is important because SPE rates seem to be sensitive to the level of relaxation of the amorphous phase present. During this analysis, we find that standard geometrical descriptors are insufficient to give a complete picture of the amorphous quality, and a more elaborate analysis is given in an attempt to obtain a more complete picture. Then a solid phase epitaxy simulation at a significant temperature is discussed in detail showing how homogeneity in the amorphous region is hard to reach, diffusion levels are higher than at low temperatures, and more than one phenomenon is present. Finally, the results of several solid phase epitaxy simulations at different temperatures are collected, and the barriers that characterize the physical phenomena are fit. The agreement between these fitted values and the barrier deduced from experiments proves the ability of the developed potential to provide a realistic SPE description, in spite of a severe overestimation of the measured recrystallization speed, and the comparison with results in the numerical simulation literature validates the proposed picture that two key events are involved in recrystallization: a defect-creation and a defect-migration event that together account for the single measured barrier.

In a final chapter a few directions for future research are discussed.

Chapter 1

Background

1.1 Solid Phase Epitaxy

Solid State Epitaxy (SPE) is the process by which a higher free-energy state of a given material progressively transforms into a more stable state, following the template that the substrate provides. This transformation takes place at temperatures below the melting point of the material and through atomic reordering and migration occurring at the interface between the two phases. In the case of silicon, the tetrahedral crystalline structure (c-Si) is the growing stable phase, while the meta-stable state is represented by the amorphous silicon phase (a-Si). Like the crystalline phase also a-Si is mainly tetrahedrally coordinated, but breaks the long-range order typical of the crystal. The increased strain caused by bond angle distortion and the presence of point defects are the main causes of the higher internal energy of the amorphous phase. When the temperature of the system is raised above a minimal kinetic energy needed for a-Si atoms to overcome the diffusion barriers and move, a-Si layers in the proximity of the crystalline template starts to re-order themselves. Such reordering can be achieved through a diverse set of structural relaxations with a relative relevance determined by the temperature [4, 5, 6, 7].

Unfortunately, SPE is not a stand-alone phenomenon, and other processes are compatible with the same experimental conditions, e.g. random nucleation growth (RNG), precipitation, phase separation, and defect-enhanced diffusion. All these processes are characterized by an activation barrier [8], hence their characteristic time τ_i follows an Arrhenius law:

$$1/\tau_i = 1/\tau_0 \ \exp(-\frac{E_i}{k_B T}) \tag{1.1}$$

where E_i is the activation energy of such processes, k_B is the Boltzmann constant, T is the sample temperature, and τ_0 is a process-dependent prefactor. This

exponential behavior is what allows for their identification during experiments and simulations.

Some of these phenomena are more compatible than others with SPE and their knowledge can give valid insight when interpreting numerical results. However, taking into account all these pathways can also increase the complexity of the analysis. To cite a few among the most relevant phenomena we have the creation of c-Si seeds in RNG, and explosive RNG. The first is a process with very high energy barrier, estimated to be in the range between 4 and 5.3 eV [9, 10], and it is often observed in the a-Si region of samples undergoing high-temperature SPE [4]. When the sample temperature approaches the melting point, explosive RNG is also observed. In this case, the exothermic a-Si/c-Si transformation provides enough energy to heat the interface above the melting point, increasing the mobility of the atoms at the border of the c-Si grain and hence allowing for a fast reconstruction. Such explosive behavior has been reported to occur also at the c-Si/a-Si interface in numerical simulations of SPE.

Comparing numerical simulations to experiments is made even harder by the fact that not all the a-Si are equal, and their properties strongly depend on the sample thermal history. Experimentally, when creating a sample for SPE the procedure starts from a perfect c-Si crystal, oriented in the direction of interest, that is either bombarded (implanted) with ions at different energies, flux, and angles, or over-grown by depositing atoms on the surface (sputtering/evaporation). While all these procedures result in the creation of a-Si, the samples obtained can have very different properties. For example, a-Si obtained by ion implantation is typically characterized by the presence of interstitial defects while, over-grown a-Si samples are characterized by more vacancies and voids. A commonly applied procedure to standardize the structure of different samples is to anneal them at low temperature (> 500 K). However, such a process has not been uniformly applied in experiments and so, for coherence, one should always specify the a-Si full thermal history. In the following, we will always indicate whether an a-Si sample is relaxed or non-relaxed, and further details will be added when relevant.

An obvious key role for the understanding, modeling, and interpretation of SPE results is played by temperature. While we can assume that the temperature in the material undergoing epitaxy is constant (thanks to the fact that silicon thermal diffusion length is greater than 1 μm), we do not know the properties of the specific a-Si under investigation at that temperature. This is due to two major issues. The first is the uncertainty in a-Si melting temperature: contrary to the well-defined melting temperature of the crystalline structure, the melting temperature of a-Si depends on its relaxation state. As the disorder of the sample increases its melting temperature can drop to values as low as 1000

K, while well relaxed a-Si samples show melting temperatures up to 1412 K. An example of the impossibility of defining a single melting temperature is shown in fig. 3 taken from [11] where the melting temperature of samples with different thermal history is modelled [12, 13, 11]. The second issue is the theorized existence of an a-Si state, often called low-density liquid (LDL), that shares more properties with liquid silicon (l-Si) than with glasses[14, 15].

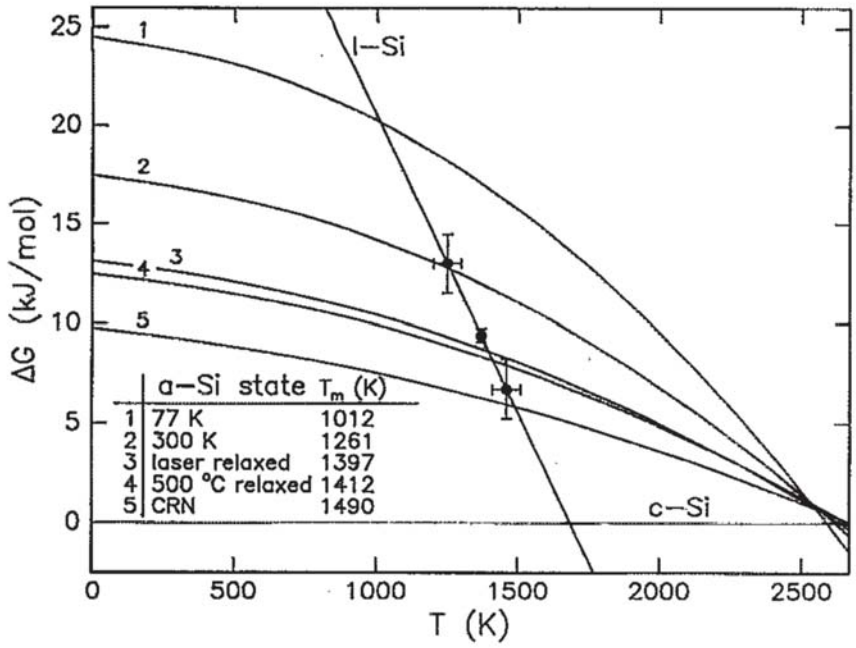


Figure 3: Reproduced from Ref. [11]. Calculation of the Gibbs free energy of l-Si and various states of a-Si relative to c-Si. The curves are theoretical estimates and the listed melting temperatures T are derived from the intersection of the free energy curves of a-Si with the one of l-Si. Parameters used in the theoretical model to obtain the curves have been obtained from samples with a thermal history specified in the inset. Continuous random network (CNR) is a theoretical defect-free a-Si structure. Points refer to experimental estimates performed on a-Si with different degrees of relaxation. Higher points are un-relaxed samples, while lower points are relaxed samples.

1.1.1 Measuring SPE velocity

Since SPE occurs progressively, it is possible to define a "velocity" of such process as the speed at which the interface between the two phases moves towards

the surface. This velocity can be inferred from experiments in different ways, and since it has an exponential dependence on the temperature, different techniques are needed to address different temperature ranges.

At low temperature (600–700,K) SPE rate ranges between 0.1 to 1 Å/min, and the regrowth of a typical 5000 Å slab requires days. The only experiments performed in this temperature range are indirect, based on differential scanning calorimetry (DSC) and isothermal calorimetry [16, 5]. In those experiments, the sample temperature is slowly raised following a 40 K/min ramp during which the excessive heat produced by internal relaxations is measured. The excessive heat can then be related to the change in enthalpy of the system and, by knowing the initial a-Si thickness, the SPE velocity is estimated as a function of the temperature [16].

In the range between 700 K and 900 K, crystal regrowth of the typical 5000 Å experimental slab happens in a matter of hours. Such time frame allows for the use of furnaces where the sample is maintained at a constant temperature for a given amount of time, after which the width of the regrown crystalline layer can be investigated through back-scattering spectra. Using this method, SPE velocity has been measured along different crystalline directions, leading to the hypothesis that the driving mechanism has preferential lattice directions [17, 18], a key point in modeling the phenomenon. Experimental speed as a function of interface crystallographic orientation is presented in Fig. 4.

The maximum temperature reachable with the furnace approach is upper bounded by practical issues. In particular, above $\approx$ 900 K crystalline regrowth becomes so fast that the reconstruction time becomes shorter than the time required to insert and extract the sample from the furnace. Indeed a typical 5000 Å sample would reconstruct in a matter of minutes, making accurate measurements impossible. A solution was put forward by Olson *et al.* [19] using time-resolved reflectivity (TRR) of samples maintained at some base temperature using a heating element from below and flash heated from above with a continuous wave (CW) laser. At high temperature, CW laser heating of the top portion of the sample becomes essential because the time otherwise required to rise the sample temperature with the bottom heating element would be comparable to the time required to crystallize the whole sample. In TRR, the crystalline/amorphous (c/a) interface position can be constantly probed by a second He-Ne laser: the higher refraction index in the visible range of a-Si w.r.t. c-Si X causes a reflection at the interface inside the material. Interference between the primary beam reflection from the sample surface and the beam component emerging after reflection by the c/a interface, causes oscillations in the reflectivity as a function of the interface position that can be monitored over time, as shown for instance in Fig. 5. The reconstruction velocity can be extracted from

the main frequency in the reflectivity oscillations and the laser wavelength. Using this technique, it is possible to observe SPE along the $\langle 100 \rangle$ direction up to 1600 K, where the reconstruction speed is of the order of cm/s.

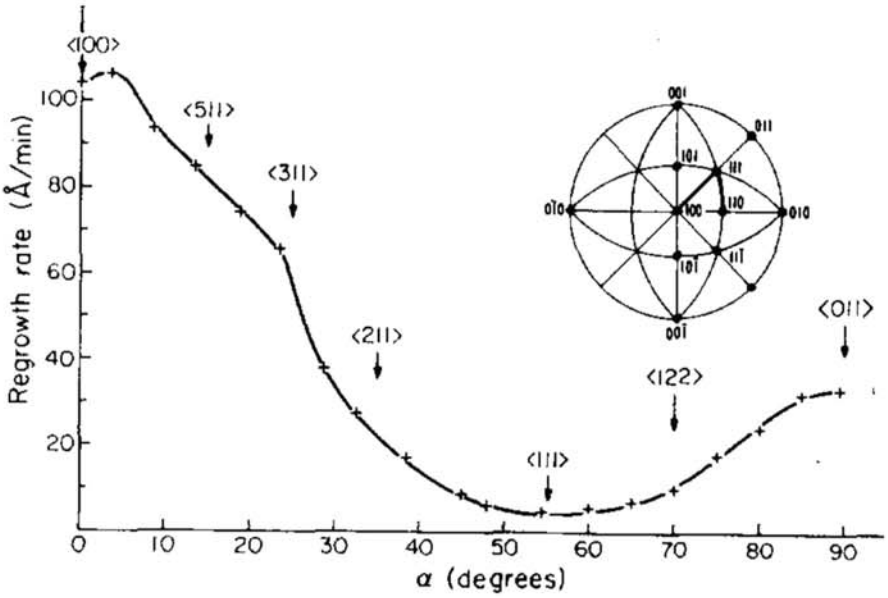


Figure 4: From Csepregi *et al.* [17] SPE velocity along different crystallographic axis measured at 820K. The sample was relaxed with furnace annealing and a-Si depth was characterized by back-scattering. More accurate selective measurements are available in [18], where the ratio between the three main directions of technological interests ($\langle 100 \rangle : \langle 110 \rangle : \langle 111 \rangle$) is established as 20:8.7:1, close to the 26:8:1 ratios reported by Csepergi.

The variability in measured SPE velocity can be readily seen by comparing different experiments: Fig. 6 reports the SPE velocity as a function of temperature for samples with different thermal history and compositions. The corresponding fits to parameters in an Arrhenius law $v = v_0 \exp -\frac{E_a}{k_b T}$, where v is the SPE rate, E_a is the activation energy, and v_0 is the process prefactor, are reported in table 1. From this table we can see that ion implanted layers reconstruct with a larger prefactor than the deposited films. Such behaviour has been attributed to their lack of voids in the a-Si. Moreover, the presence of boron and phosphorus impurities can drastically affect the barrier, suggesting a possible dependency between the speed and the Fermi level [20].

As we can see from fig. 6, no pure silicon investigation has been performed above 1200 K, but high-temperature points are available only for Si/As samples. Velocities extracted from these samples are consistent with the assumption of a

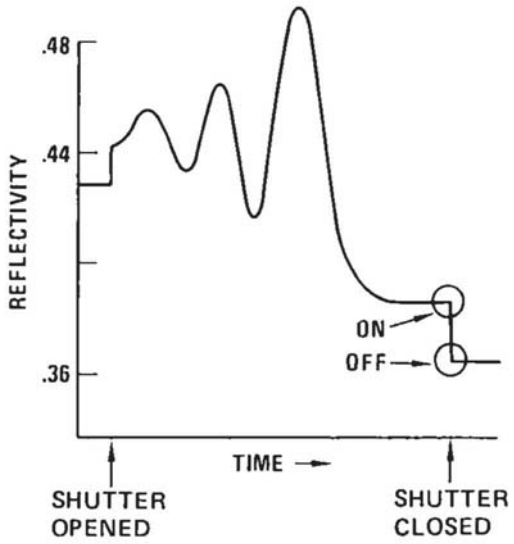


Figure 5: TRR procedure from Olson *et al.*[19]. Up to t=0 the sample is kept at a baseline temperature by the heated support and negligible SPE growth takes place. The observed reflectivity is the one of the amorphous layer combined with the bulk diamond layer at the temperature of the substrate. Once the heating CW laser shutter is open there is a sudden jump in reflectivity (solely due to the temperature change) and SPE growth starts. The reflectivity then oscillates depending on the depth of the amorphous. Moreover, as the depth decreases the light that remains trapped in the amorphous also decreases, leading to a larger oscillation amplitude. Once the whole system is crystallized the final reflectivity is that of the heated diamond structure. Closing the CW laser shutter causes a drop in the reflectivity proportional to the temperature drop of the sample surface.

single barrier process theorized for pure Si samples. The effect of arsenic implantation on SPE velocity has been investigated by Olson *et al.* [4] showing that the presence of low arsenic doses can not affect SPE velocity by more than approximately a factor of 2, a quantity that has a low impact on the energy barrier estimation. Although this is only an indication it is not unreasonable to assume that for pure Si the behavior remains the same over the whole temperature range studied.

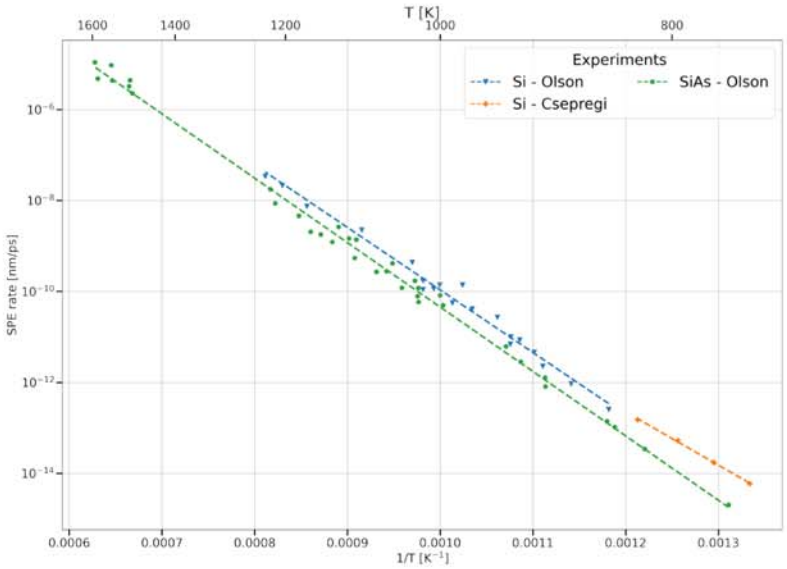


Figure 6: Selected SPE velocity as a function of temperature for different samples measured with different techniques. SiAs - Olson sample from [19] and Si - Olson sample from [4]. Both were obtained with TRR. Csepregi from [17] was obtained via furnace annealing. Different lines are exponential fit to the data, exact parameters are available in Table 1.

Differences between SPE velocities for "as implanted" samples and annealed samples have been investigated through both TRR and DSC. Rooda *et al.* [5] investigated the reconstruction speed at low temperatures for annealed and "as implanted" samples, measuring a higher speed for the latter (see Fig. 7). This higher speed has been attributed to defect-mediated diffusion that naturally leads to lower diffusion activation energies. At high temperature no speed differences are observed, thus supporting a mechanism change in favor of a native mechanism (therefore, independent of the sample history). Similar behaviors have also been observed for Ge [16].

Other experiments have, however, reported different behaviors. For instance, Lu *et al.* [22] found a speed increase lasting only tenths of a second and

Sample details	E_a (eV)	v_0 (cm/sec)
Implanted[a]	2.68 ± .05	3.1 $\times 10^8$
Implanted[b]	2.3 ± .1	2.8 $\times 10^6$
Implanted[c]	2.85	
Implanted[d]	2.9 ± .1	
Deposited[a]	2.71 ± .05	2.3 $\times 10^8$
Deposited[b]	2.7 ± .2	
Doped with Arsenic[a]	2.75 ± 0.05	3.68 $\times 10^8$
Doped with Boron[a]	2.52	
Doped with Phosphorus[a]	2.68	
Doped with Phosphorous[c]	2.5	

Table 1: Different prefactors, v_0, and energy barriers, E_a, experimentally measured on different samples. [a] [4], [b] [17], [c] [21], [d] [18].

postulated the presence of Hydrogen in the sample [23]. Aziz *et al.* [24], by performing low-temperature TRR, excluded any correlation between a-Si thermal history and SPE velocity, and suggested the existence of a native mechanism at the interface.

1.1.2 Atomistic models

As of yet, no agreement exists on the mechanism of the process. Simple atomistic models based on dangling bond migration at the interface have been proposed [25, 26, 27]. Starting from those models industrial on-lattice kinetic Monte Carlo (KMC) codes [28, 29] that try to describe the possible lattice faults during the crystallization of macroscopic samples, dopant activation and other phenomena have been developed. However, their ability to correctly describe the physics in configurations that are out of equilibrium is still limited because they are based on qualitative mechanisms. The ideal solution is to create a larger database of events by investigating SPE using other techniques and implementing these events in the KMC. When trying to do so, the first goal is always to obtain the correct barrier while simulating the reconstruction. Only then mechanisms can be isolated and analyzed.

Because of the length, time-scale and disorder involved in SPE a parameter-free quantum mechanical approach is not computationally affordable, even today with the new exascale High-Performance Computers. Several authors [30,

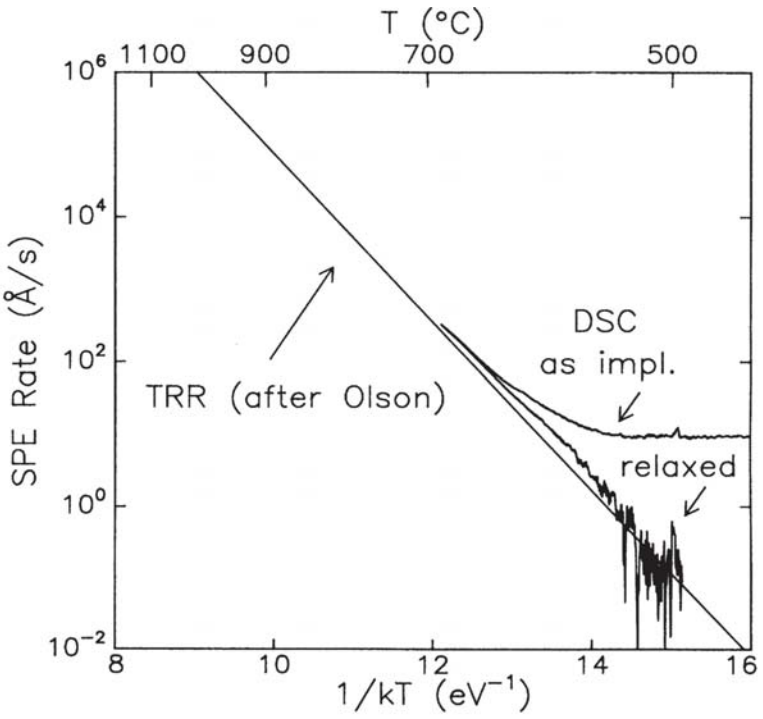


Figure 7: Crystallization velocity obtained by DSC measurements for as-implanted and relaxed structures, from [5]. Results from Olson *et al.* [4] (presented here in Fig. 5) are also present. The relaxed material differs from the implanted one because it has been subject to a 500 C furnace annealing. In the "as-implanted" sample a clear knee is present where the un-relaxed defects become the main driver of the reconstruction.

31, 32, 33, 34, 35, 36, 37] have, therefore, investigated the fundamental mechanisms using classical (and parametrized) interatomic potentials. The most successful studies were conducted using Tersoff [38] potential, Environment Dependent Interatomic Potential (EDIP) [39, 40] potential, and with a parametrization of the Stillinger-Weber [41, 42] (SW) potential specifically crafted for this phenomenon. Simulation with Tersoff potential succeeds in predicting a single barrier close to experiments, but the predicted bulk melting temperature is too high. Moreover, it does not reproduce the 26:8:1 experimental ratio between orientation dependent speeds shown in fig. 4 [35]. Therefore the mechanism events extrapolated from such simulations can be misleading. Similar issues are encountered with SW potential, but fewer studies have been performed. EDIP potential, on the other hand, predicts two different barriers similarly to what is inferred from fig. 7. These barriers have been associated with two separate phenomena that experiments have ruled as plausible, but it has a reconstruction speed prefactor v_0 that is orders of magnitude higher than the one observed in experiments [30, 22].

1.2 Computational approach

In order to overcome the limitations displayed in simulations using the available inter-atomic potentials, we propose to develop a new Machine Learned (Neural Network) inter-atomic potential fitted on accurate density functional theory (DFT) data. We briefly introduce here the main elements of these two topics.

1.2.1 Density functional theory

The DFT formalism reduces the complexity of the many body problem for the electrons by rephrasing it into a simpler one. In this section, we will first introduce the Hohenberg and Kohn (HK) theorem, showing how all properties of the electronic ground state depend on the electronic density. Then the Kohn and Sham (KS) ansatz will be discussed, allowing the reformulation of the interacting electron problem into a non-interacting electron one. This step allows to isolate the so-called exchange-correlation functional term, where the major practical approximations used in the theory can be introduced. Several other theoretical extensions that are relevant to some of the results of this work will be discussed later in the text.

The electronic Ground State Problem

The natural starting point of any quantum mechanical description is the time independent *Schrödinger equation* for a general electron and ion system:

$$\hat{H}(\{\mathbf{r}_i\},\{\mathbf{R}_I\})\ \ \Psi_s(\{\mathbf{r}_i\},\{\mathbf{R}_I\}) = E_s\ \ \Psi_s(\{\mathbf{r}_i\},\{\mathbf{R}_I\}), \tag{1.2}$$

where lower case letters ($\mathbf{r}_i$) are used to describe the coordinates of the N electrons of the system, and upper case letters ($\mathbf{R}_I$) for the coordinates of the M nuclei. Many body wave functions $\Psi_s(\{\mathbf{r}_i\},\{\mathbf{R}_I\})$ are the eigen-states of equation (1.2) corresponding to eigen-energies E_s, where $s = 1, 2, 3\ldots$ labels different solutions. Exact determination of such wave functions quickly becomes numerically intractable as the number of particles to describe increases. $\hat{H}$ is the Hamiltonian operator that can be expressed as:

$$\begin{aligned}\hat{H} = &-\sum_I \frac{\hbar^2}{2M_I}\nabla_I^2 - \frac{\hbar^2}{2m_e}\sum_i \nabla_i^2 \\ &-\sum_{i,I}\frac{Z_I e^2}{|\mathbf{r}_i - \mathbf{R}_I|} \\ &+\frac{1}{2}\sum_{i\neq j}\frac{e^2}{|\mathbf{r}_i - \mathbf{r}_j|} + \frac{1}{2}\sum_{I\neq J}\frac{Z_I Z_J e^2}{|\mathbf{R}_I - \mathbf{R}_J|}.\end{aligned} \tag{1.3}$$

Here $\hbar$ is the Plank constant, m_e is the electron mass and M_I is the mass of the I-th ion. These five terms can be analyzed separately: the first two terms represent the ion and electron kinetic energies respectively, where ∇ is the Laplace operator w.r.t. the ion/electron coordinates. The fourth and the last terms are the electron-electron and ion-ion Coulomb repulsion terms respectively, where e is the (modulus of the) electron charge and $Z_{I/J}$ is the atomic number of the ions. The third term is the electron-ion attractive interaction.

The problem of electrons coupled with ions is too hard to be tackled all at once and a first step which is almost invariably performed is the separation of the ionic and electronic degrees of freedom via the so called adiabatic or Born-Oppenheimer(BO) approximation [43]. Such approximation exploits the observation that the nuclear masses M_I are much larger than the electronic ones. The immediate consequence is that the nuclear kinetic energy term is negligible with respect to all the others or, in other words, the ions reaction time-scale is orders of magnitude longer than the electronic one. An "expansion" of the solution around the instantaneous positions of the ions can then be performed. First, the eigen-function solution to the Hamiltonian has to be decomposed in the ionic part and the electronic part. This decomposition can be achieved without loss

of generality by separating the ionic and electronic Hilbert sub-spaces:

$$\Psi_s(\{\mathbf{r}\},\{\mathbf{R}\}) = \sum_i \chi_{si}(\{\mathbf{R}\})\ \Phi_i(\{\mathbf{r}\}|\{\mathbf{R}\}), \tag{1.4}$$

where χ_{si} are purely ionic wave function components, and $\{\Phi_i\}$ is the complete basis set of the eigen-functions for the electronic problem (Eq. (1.3) without the ionic kinetic energy term, see also Eq. (1.6) below) computed at fixed ionic positions. Inserting Eq. (1.4) in Eq. (1.2) it is possible to write a matrix-form Schrödinger equation for the ions wavefunction components χ_{si}. Such equation correlates every χ_{si} with all the other $\{\chi_{si}\}$ with coefficients that represent quantum transitions between different electronic states. The adiabatic approximation, using the physical justification that electrons adjust much faster than the ions, neglects all these transitions, hence electrons remain in the same electronic state (the same potential energy surface (PES)), during time evolution.

Using the BO approximation it is possible to rewrite the main problem as two separate ones. The first one is an equation for the ions where the electronic potential energy surfaces as a function of ionic positions act as the external potentials for the ionic problem

$$\left[-\sum_j \frac{\nabla_j^2}{2M_j} + U_i(\{\mathbf{R}\})\right] \chi_{ni}(\{\mathbf{R}\}) = E_{ni}\ \chi_{ni}(\{\mathbf{R}\}), \tag{1.5}$$

where $U_i(\{\mathbf{R}\})$ is i-th eigenvalue of the effective Hamiltonian for the electronic wave function at fixed ionic positions:

$$\begin{aligned}\hat{H}_{BO} &= -\frac{\hbar^2}{2m_e}\sum_i \nabla_i^2 - \sum_{i,I}\frac{Z_I e^2}{|\mathbf{r}_i - \mathbf{R}_I|} + \frac{1}{2}\sum_{i\neq j}\frac{e^2}{|\mathbf{r}_i - \mathbf{r}_j|} + \frac{1}{2}\sum_{I\neq J}\frac{Z_I Z_J e^2}{|\mathbf{R}_I - \mathbf{R}_J|} \\ &= \hat{T} + \hat{V}_{\text{ext}} + \hat{V}_{\text{int}} + E_{II}\end{aligned} \tag{1.6}$$

where i and j label the electrons, while I labels the ions. The second line defines the operators that are going to be used hereafter: the kinetic energy $\hat{T}$, electron-electron Coulomb interaction $\hat{V}_{\text{int}}$, the classical ion-ion interaction E_{II}, and an external potential $\hat{V}_{\text{ext}}$ that here is generated by the nuclei but can also include the interaction with any external field.

The matrix elements connecting different electronic and ionic states that are neglected in the adiabatic approximation describe electron-phonon coupling. While these are fundamental to describe properties like electrical resistivity of material (and superconductivity), they typically have a minor effect for the ionic

dynamics when an electronic band gap is present and even in metals are typically introduced in a perturbative way when relevant.[1].

A further practical simplification resulting from the separation of the electronic and ionic problems is the fact that while the electronic problem requires a quantum mechanical treatment, the ionic problem can be addressed classically solving Newton's equation of motion for the ions

$$M_j \ddot{\mathbf{R}}_j = -\nabla_j U_i(\{\mathbf{R}\}), \tag{1.7}$$

where $U_i(\{\mathbf{R}\})$ is the interaction potential. The electronic state, i, considered in these simulations is almost invariably the ground state. Equation (1.7) is at the core of most atomic-scale simulations.

The ground state of the Born-Oppenheimer Hamiltonian (1.6) is in principle exactly solvable, but this is unfeasible in practice, except for systems with a very limited number of electrons. A first approximate approach proposed by Hartree and Fock can then be applied. This approach restricts the general solution to a single Slater determinant Φ of parametrized single particle spin orbitals $\phi_i^\sigma(\mathbf{r}_j) = \psi_i^\sigma(\mathbf{r}_j)\alpha_i(\sigma_j)$ where i and j retain their previous meaning, σ is the particle spin, ψ is the spatial part of the wave function, and α is the spin part of the wave function. By inserting such solution in (1.6) we obtain:

$$\begin{aligned}\left\langle \Phi \left| \hat{H} \right| \Phi \right\rangle = &\sum_{i,\sigma} \int d\mathbf{r} \psi_i^{\sigma\star}(\mathbf{r}) \left[-\frac{1}{2}\nabla^2 + V_{\text{ext}} \right] \psi_i^\sigma(\mathbf{r}) + E_{II} \\ &+ \frac{1}{2} \sum_{i,j,\sigma_i,\sigma_j} \int d\mathbf{r} d\mathbf{r}' \psi_i^{\sigma_i\star}(\mathbf{r}) \psi_j^{\sigma_j\star}(\mathbf{r}') \frac{1}{|\mathbf{r}-\mathbf{r}'|} \psi_i^{\sigma_i}(\mathbf{r}) \psi_j^{\sigma_j}(\mathbf{r}') \\ &- \frac{1}{2} \sum_{i,j,\sigma} \int d\mathbf{r} d\mathbf{r}' \psi_i^{\sigma\star}(\mathbf{r}) \psi_j^{\sigma\star}(\mathbf{r}') \frac{1}{|\mathbf{r}-\mathbf{r}'|} \psi_i^{\sigma}(\mathbf{r}) \psi_j^{\sigma}(\mathbf{r}').\end{aligned} \tag{1.8}$$

In this equation the second row describes the classical electrostatic energy of the electronic density, and is defined as the *Hartree* term, and the third row describes the *Exchange* term originating from the antisymmetry of the trial wave function and has no classical analog.

The total energy of the system, computed from Eq (1.8), can then be minimized with respect to the parameters in ψ_i to give the best approximate solution describable with the chosen set of single particle spin orbitals. This approach falls short as the complexity of the system increases and more accurate solutions require an expansion over a basis set of several Slater determinants with

[1] In this study, we have encountered several metallic configurations. Nevertheless we did not engage in addressing this problem because, *de facto*, we are not going to perform any molecular dynamics of the metallic phases. For all the calculations where the motion of the atoms is taken into account (e.g. phonon spectra), we restrict to configurations where there is a band gap.

increased memory (more wave functions to be memorized) and computational requirements (more integrals).

Two theorems, proven by Hohenberg and Kohn (HK), introduced a paradigm shift moving the focus of the problem from the wave function to the electron density, making this a density functional theory. The knowledge of the wave function is not the only way to access the properties of the system, but the electronic density of the ground state is what really matters, so if it is possible to find an alternative inexpensive way to compute such quantity, all observables can be derived from it. The two theorems state the following:

Theorem 1.1. *For any system of interacting particles in an external potential $V_{ext}(\mathbf{r})$, the potential $V_{ext}(\mathbf{r})$ is determined uniquely, up to an additive constant, by the ground state particle density $n_0(\mathbf{r})$.*

Theorem 1.2. *A universal functional for the energy $E[n]$ in terms of the density $n(\mathbf{r})$ can be defined, valid for any external potential $V_{ext}(\mathbf{r})$:*

$$E_{HK}[n] = F_{HK}[n] + \int d^3r V_{ext}(\mathbf{r}) n(\mathbf{r}) + E_{II} - \mu \left(\int n(\mathbf{r}) d^3r - N \right), \tag{1.9}$$

where the second term represents the interaction between the external fields (e.g. the electric field due to the ions positions) and the electrons, E_{II} is the repulsive interaction between nuclei, and the last term represents the chemical potential. The term $F_{HK} = T[n] + E_{int}[n]$ embeds all the internal quantities such as the kinetic energy and the electron-electron interaction.

For any particular $V_{ext}(\mathbf{r})$, the exact ground state energy of the system is the global minimum of this functional, and the density $n(\mathbf{r})$ that minimizes the functional is the exact ground state density $n_0(\mathbf{r})$.

Proof of these theorems can be found in a number of textbooks (see e.g. [44]).

If it were possible to find an explicit formula for the F_{HK} functional and minimize it with respect to the density, the final density would be a quantity that completely defines the properties of the system. Unfortunately the HK theorems do not give us any indication on how to do so.

The open question that remains is the definition of accurate approximations to such functional. A simple model, due to Thomas and Fermi, can provide an analytic approximation to it, but this misses some of the essential physics needed for the description of matter. A useful approach was proposed by Kohn and Sham (KS) who considered the ansatz that the ground state density of the original interacting system is equal to that of some auxiliary non-interacting system.

The construction of the auxiliary system can start from the definition of an independent-particle Hamiltonian:

$$\hat{H}^{\sigma}_{\text{aux}} = -\frac{\hbar}{2m_e}\nabla^2 + V^{\sigma}_{\text{KS}}(\mathbf{r}), \tag{1.10}$$

where the potential V^{σ}_{ks} is not yet specified. We will label the eigenfunctions of such Hamiltonian ψ^{σ}_i and the relative eigenvalues ϵ^{σ}_i.

The density of such simple system thus becomes:

$$n(\mathbf{r}) = \sum_{\sigma}\sum_{i=1}^{N^{\sigma}} \left|\psi^{\sigma}_i(\mathbf{r})\right|^2, \tag{1.11}$$

its kinetic energy

$$T_s = -\frac{\hbar}{2m_e}\sum_{\sigma}\sum_{i=1}^{N^{\sigma}} \left\langle \psi^{\sigma}_i \left| \nabla^2 \right| \psi^{\sigma}_i \right\rangle, \tag{1.12}$$

and its classical Coulomb interaction, the Hartree energy, can be written as:

$$E_{\text{Hartree}}[n] = \frac{e^2}{2}\int \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}d\mathbf{r}'. \tag{1.13}$$

We can apply the HK theorems to this non-interacting problem, and equation (1.9) will read:

$$E_{\text{KS}}[n] = T_s[n] + \int d^3r V_{\text{KS}}(\mathbf{r})n(\mathbf{r}) - \mu\left(\int n(\mathbf{r})d^3r - N\right). \tag{1.14}$$

The link between the two problems (HK and KS) comes from the observation that the two systems have the same ground state density. Taking the functional derivative with respect to the density at the minimum of equations (1.9) and (1.14) we obtain:

$$\frac{\delta F_{HK}[n]}{\delta n} + V_{\text{ext}} = \mu \tag{1.15}$$

$$\frac{\delta T_s[n]}{\delta n} + V_{\text{KS}} = \mu, \tag{1.16}$$

where in principle the two μ are the same up to a constant that can always be removed. We can formally single out the non interacting kinetic energy and the classical interaction energy in the $F_{HK}[n]$ functional, thus defining an *exchange-correlation* functional as the reminder:

$$F_{HK}[n] := T_s[n] + E_{\text{Hartree}}[n] + E_{\text{xc}}[n]. \tag{1.17}$$

Substituting (1.17) in (1.15) and making it equal to (1.16), after using the HK ansatz on the same density to simplify the two functional derivatives of the kinetic energy $T_s[n]$, we are left with:

$$\begin{aligned} V_{\text{KS}} &= V_{\text{ext}} + \frac{\delta E_{\text{Hartree}}}{\delta n} + \frac{\delta E_{\text{xc}}}{\delta n} \\ &= V_{\text{ext}} + V_{\text{Hartree}} + V_{\text{xc}}^{\sigma}. \end{aligned} \tag{1.18}$$

This is the formulation for the KS potential of the auxiliary non-interacting problem, that hides all the complexity behind the still unknown E_{xc} functional. The remaining problem requires the solution of a Schrödinger-like equation:

$$(H_{\text{KS}}^{\sigma} - \epsilon_i^{\sigma})\psi_i^{\sigma}(\mathbf{r}) = 0, \tag{1.19}$$

with the KS Hamiltonian (1.10), subject to the KS potential (1.18). These are known as the KS equations.

The KS equations (1.19) are typically solved self consistently: first an initial guess on the density is made and with it V_{KS} is and the corresponding eigenfunctions and eigenvalues of the Hamiltionian are computed. With the new quantities the energy is computed and if it is significantly different from the previous iteration a new charge density is built using the new eigenfunctions and the cycle is repeated until convergence.

Exchange and correlation functionals

Up to this point the KS equations are formally exact, but what has been done is nothing more than shuffling the cards and moving the interacting problem around, from the Slater determinant in the HF approximation, to the Exchange Correlation (xc) part of the potential in the KS non-interacting formulation. In fact, the problem still needs to be addressed: this requires to determine the functional form of the exchange-correlation term. The usefulness of DFT relies on the fact that practical approximations for XC energy can be defined. Historically, the first model proposed was the local density approximation (LDA), which was developed around the assumption that the exchange-correlation energy is equal to the one of a homogeneous electron gas with the same local density:

$$E_{\text{xc}}^{\text{LDA}}[n(\mathbf{r})] = \int d\mathbf{r} n(\mathbf{r}) \epsilon_{\text{xc}}^{\text{hom}}(n(\mathbf{r})), \tag{1.20}$$

where $\epsilon_{\text{xc}}^{\text{hom}}$, the xc-energy per particle of the homogeneous gas, can be numerically computed and is usually split in an analytic part [44] defined from the HF exchange ϵ_x, and a numerical part added to reach values predicted by more accurate quantum Monte Carlo calculations [45]. This second contribution is

defined as correlation ϵ_c and by construction the splitting is such that $\epsilon_x \gg \epsilon_c$. This simple formulation however has several shortcomings: it overestimates the atomization energy, and bonds tend to be too short and strong. A practical consequence is the prediction of lattice constants that are smaller than the experiments and bulk moduli that are too high.

Correcting equation (1.20) requires an extension of the theory. An attempt to pursue this extension systematically by performing a second-order gradient expansion in the density resulted in lower accuracy in the prediction because the resulting exchange-correlation energy functional displayed non-physical behaviors [46]. This led to the proposal of more general functional forms depending on the gradient of the density, such as the non empirical generalized-gradient approximation (GGA) formulation:

$$E_{\text{xc}}^{\text{GGA}}[n] = \int d\mathbf{r} f(n, \nabla n), \tag{1.21}$$

a form that includes many of the functionals that are commonly used today. The problem with these non empirical constructions is that there is no unique recipe to define the f function. The form of f can be subject to several different constraints, and different GGA approaches satisfy different subsets of them.

A standard GGA approach for the definition of f starts by rewriting (1.21) as:

$$\begin{aligned} E_{\text{xc}}^{\text{GGA}}[n] &= E_{\text{x}}^{\text{GGA}}[n] + E_{\text{c}}^{\text{GGA}}[n] \\ &= \int d\mathbf{r} n(\mathbf{r}) \epsilon_x^{\text{hom}}(n(\mathbf{r})) F_x(s(\mathbf{r})) \\ &+ \int d\mathbf{r} n(\mathbf{r}) \left\{ \epsilon_c^{\text{hom}}(n(\mathbf{r})) + H(r_s, s(\mathbf{r})) \right\}, \end{aligned} \tag{1.22}$$

where the separation of exchange and correlation energies defined in the LDA case has been taken as a starting point, and the gradient approximation has been introduced through the two empirical functions F_x and H. Here r_s is the local Seitz radius (such that $n = \frac{3}{4\pi r_s^3}$) and $s(\mathbf{r}) \propto \frac{|\nabla n|}{n^{4/3}}$.

The construction of the two functions F_x and H is not unique. A series of physically meaningful constraints for these functions have been identified [47], however the ultimate choice on which of those constraints needs to be fulfilled is always left to the authors.

Since this work focuses only on semiconductors, in particular on crystalline silicon and surfaces, the standard choice nowadys is the use of PBE [46] construction and its twin PBESol [48]. The latter uses the same functional form as the first one, but by prioritizing different constrains. PBEsol is believed to better describe solid phases.

In the PBE formulation the functionals F_x and H still have 2 parameters that need to be fixed, and to fix them one can use three physical constraints that however cannot be satisfied all together: *i)* reproduce the correct correlation for small values of s, *ii)* reproduce the correct exchange for small values of s, *iii)* have a global $E_{\text{xc}}^{\text{GGA}}$ that recovers LDA for small s. In the PBE case, one of these parameters is fit to reproduce the correct correlation value for slowly varying density (condition *i)*) while the second is chosen to respect the third condition, effectively violating the second condition and providing a worse approximation for the exchange. The PBESol functional instead, is fitted to the correct exchange expansion (condition *ii)*), and a compromise between the other two conditions. In particular condition *iii)* is almost fully satisfied, but corrected to better predict surface energies. Differences between different GGAs are typically reported through their enhancement factor F_{xc}, defined by rewriting equation (1.22) as:

$$E_{\text{xc}}^{\text{GGA}}[n] = \int d\mathbf{r} n(\mathbf{r}) \epsilon_x^{\text{hom}} F_{xc}(r_s, s), \tag{1.23}$$

where r_s and s have been previously defined and depend on $\mathbf{r}$. The plot for these two quantities for the two functionals under investigation is reported in Fig. 8. Unfortunately, until extensively tested on various properties, these graphs give little insight on how the functionals might behave outside of the fitted properties.

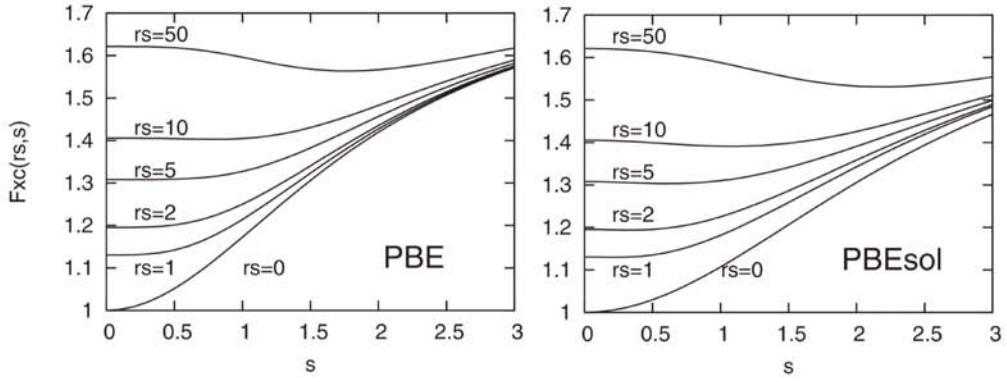


Figure 8: Enhancement factors of PBE and PBEsol, for spin-unpolarized systems, as a function of the reduced density gradient, for various values of r_s, from [48].

Plane waves

We will now describe an approach to practically solve the Schrödinger equation for the KS Hamiltionian (1.19) numerically. In a periodic system it is computationally convenient to expand the wave function over a complete set of Fourier components:

$$\psi_i(\mathbf{r}) = \sum_{\mathbf{q}} c_{i,\mathbf{q}} \frac{1}{\sqrt{\Omega}} \exp(i\mathbf{q}\mathbf{r}) \equiv \sum_{\mathbf{q}} c_{i,\mathbf{q}} \left|\mathbf{q}\right\rangle, \tag{1.24}$$

where $\mathbf{q}$ is the wave-vector, Ω is a normalization volume, and $c_{i,\mathbf{q}}$ are the coefficients of the wave function in the basis of orthonormal plane waves $|\mathbf{q}\rangle$ satisfying

$$\langle \mathbf{q}' | \mathbf{q} \rangle \equiv \frac{1}{\Omega} \int_{\Omega} d\mathbf{r} \, \exp(-i\mathbf{q}'\mathbf{r}) \, \exp(i\mathbf{q}\mathbf{r}) = \delta_{\mathbf{q}',\mathbf{q}}. \tag{1.25}$$

Inserting (1.25) into (1.19), multiplying from the left by $\langle \mathbf{q}' |$ and using the relation (1.25) leads to the Schrödinger equation in Fourier space:

$$\sum_{\mathbf{q}} \left\langle \mathbf{q}' \left| \hat{H}_{KS} \right| \mathbf{q} \right\rangle c_{i,\mathbf{q}} = \epsilon_i c_{i,\mathbf{q}}. \tag{1.26}$$

that must be solved to obtain the $c_{i,\mathbf{q}}$. Resolving the matrix element for the kinetic part is immediate

$$\left\langle \mathbf{q}' \left| -\frac{\hbar^2}{2m_e} \nabla^2 \right| \mathbf{q} \right\rangle = \frac{\hbar^2}{2m_e} \left|q\right|^2 \delta_{\mathbf{q}',\mathbf{q}}. \tag{1.27}$$

but things are more complicated for the periodic potential part such as those in periodic crystals. In the case of a periodic crystal of volume $\Omega_{\text{crystal}} = \prod_{i=1}^{3} N_i a_i$ (where N_i is the number of unit cells along the $\mathbf{a}_i$ basis vector)the potential must be periodic with the same periodicity as the crystal:

$$V_{\text{KS}}(\mathbf{r}) = V_{\text{KS}}(\mathbf{r} + \mathbf{T}_n), \tag{1.28}$$

where $\mathbf{T}_n \equiv \sum_{i=1}^{3} n_i \mathbf{a}_i$ is any lattice translational vector. Such a periodic function can be represented in terms of Fourier components that must satisfy the Born-Von Karman periodic boundary conditions (PBC) along each dimension:

$$\exp(i\mathbf{q}N_i\mathbf{a_i}) = 1 \Rightarrow \mathbf{q}_i\mathbf{a_i} = 2\pi \frac{\text{integer}}{N_i} \tag{1.29}$$

.

Now following the definition of the Fourier transform:

$$V_{\text{KS}}(\mathbf{q}) = \frac{1}{\Omega_{\text{crystal}}} \int_{\Omega_{\text{crystal}}} d\mathbf{r} \, V_{\text{KS}}(\mathbf{r}) \, e^{i\mathbf{q}\mathbf{r}}, \tag{1.30}$$

and by imposing the periodicity, we can reduce the integral to only the primitive cell

$$\begin{aligned} V_{\text{KS}}(\mathbf{q}) =& \frac{1}{\Omega_{\text{crystal}}} \sum_{\mathbf{n}} \int_{\Omega_{\text{cell}}} d\mathbf{r}\; V_{\text{KS}}(\mathbf{r})\; e^{i\mathbf{q}(\mathbf{r}+\mathbf{T}_n)} \\ =& \frac{1}{N_{\text{cell}}} \sum_{\mathbf{n}} e^{i\mathbf{q}\mathbf{T}_n} \frac{1}{\Omega_{\text{cell}}} \int_{\Omega_{\text{cell}}} d\mathbf{r}\; V_{\text{KS}}(\mathbf{r})\; e^{i\mathbf{q}\mathbf{r}}. \end{aligned} \tag{1.31}$$

Where N_{cell} is the total number of cells in the crystal.

The sum over all lattice points in the middle vanishes for all $\mathbf{q}$ except those for which $\mathbf{q}\mathbf{T}_n = 2\pi$ times an integer for all crystal lattice vectors $\mathbf{T}$. Since $\mathbf{T}_n$ is a sum of integer multiples of the primitive translations $\mathbf{a}_i$, it follows that $\mathbf{q}\mathbf{a}_i = 2\pi \times$ integer. The set of Fourier components $\mathbf{q}$ that satisfy this condition is the "reciprocal lattice". If we define the vectors $\mathbf{b}_i$ such that:

$$\mathbf{b_i a_j} = 2\pi\delta_{i,j}, \tag{1.32}$$

the nonzero components of the Fourier transform are only those that live on the lattice in the reciprocal space

$$q \equiv G(m_1, m_2, m_3) = \sum_{i=1}^{3} m_i \mathbf{b}_i. \tag{1.33}$$

With these considerations the potential can be written as

$$V_{\text{KS}}(\mathbf{r}) = \sum_m V_{\text{KS}}(\mathbf{G}_m) e^{i\mathbf{G_m r}}, \tag{1.34}$$

thus the matrix elements of the potential

$$\langle \mathbf{q}' \left| V_{\text{KS}} \right| \mathbf{q} \rangle = \sum_m V_{\text{KS}}(\mathbf{G}_m) \delta_{\mathbf{q}'-\mathbf{q},\mathbf{G}_m} \tag{1.35}$$

are nonzero only if the difference between $\mathbf{q}'$ and $\mathbf{q}$ is a reciprocal lattice vector $\mathbf{G}_m$.

Moreover, if we define $\mathbf{q} = \mathbf{k} + \mathbf{G}_m$ and $\mathbf{q}' = \mathbf{k} + \mathbf{G}_{m'}$ (which differ by a reciprocal lattice vector), then the Shrödinger equation for any given $\mathbf{k}$ can be written as the matrix equation:

$$\sum_{m'} \hat{H}_{m,m'} c_{i,m'}(\mathbf{k}) = \epsilon_i(\mathbf{k}) c_{i,m}(\mathbf{k}), \tag{1.36}$$

where

$$\begin{aligned} H_{m,m'}(\mathbf{k}) &= \left\langle \mathbf{k} + \mathbf{G}_m \left| \hat{H}_{\text{KS}} \right| \mathbf{k} + \mathbf{G}_{m'} \right\rangle \\ &= \frac{\hbar^2}{2m_e} \left| \mathbf{k} + \mathbf{G}_m \right| \delta_{m,m'} + V_{\text{KS}}(\mathbf{G}_m - \mathbf{G}_{m'}). \end{aligned} \tag{1.37}$$

These last two equations are the core equation implemented in any DFT-plane wave code. Once these are solved, many properties can be obtained by integrating over the k-space. For example, the total energy, or the total electron density (1.11):

$$n(r) = \frac{1}{V} \sum_{\mathbf{k},i} \sum_{m,m'} f_{\mathbf{k},i} c^*_{i,m}(\mathbf{k}) c_{i,m'}(\mathbf{k}) e^{i(\mathbf{G}_m - \mathbf{G}_{m'})\mathbf{r}} \tag{1.38}$$

here we introduced $f_{\mathbf{k}i}$ as the occupation number which represents the occupation of a given state i at a given $\mathbf{k}$ by an electron.

Electronic temperature

The occupation number, $f_{\mathbf{k}i}$ introduced in equation (1.38), at zero temperature is a step-function that goes abruptly to zero above the Fermi energy. As this quantity is discontinuous, it may require a very dense sampling in the Brillouin Zone and lead to convergence issues during the self-consistent procedure as, for example, from one iteration to the next some states close to the Fermi energy can abruptly switch from being non-occupied to being occupied, and viceversa. Such jumps cause major changes in the charge distribution that reflects in instabilities in the total energy.

To prevent this from happening, an "electronic temperature" is added. Practically speaking, this corresponds to allowing the electronic states to have a partial occupation, an this can be done by substituting the step-function at the Fermi energy with smooth approximations such as the Fermi-Dirac distribution -in this case the electronic temperature could have a physical interpretation- or any other smooth approximate step-function. Unfortunately, using the Fermi-Dirac distribution the electronic temperature needed to sufficiently smooth the occupation, and hence to avoid convergence issue must be quite large (on the order of magnitude of 0.3 eV $\approx$ 4000K or more). Such large values cause partial occupation of electronic states at much higher energy than the Fermi energy with consequent nonphysical behaviors. To overcome this limitation, different occupation distributions have been developed, notably the Methfessel-Paxton [49] and the Marzari-Vanderbilt [50] ones. The latter will be adopted in this work as it increases the numerical stability when very large density gradients are expected, i.e. in the case of surfaces.

Pseudo-potentials

Equation (1.36) is the exact Shrödinger equation in Fourier space and its numerical solution is performed by truncating the (discrete infinite) set of plane

waves corresponding to a given k-vector to include only the plane waves with kinetic energy less than a given cutoff value. For sufficiently large cutoff values the total charge density and all related properties converge. Unfortunately the required cutoff may be extremely large, especially when the electronic wave functions are oscillating at very high frequency. This is the case when treating atomic core orbitals, and due to orthogonality, it also applies to valence wave functions in the core region. To overcome such problems, core electrons are removed from the problem and valence orbitals are modified so as to reproduce their correct behavior *outside* the core region. This is achieved by replacing V_{KS} in equation (1.35) with a V_{eff}. The theory of first principles pseudopotential and their optimization has been developed in a number of seminal papers to which we refer to [norm-conserving HSC, vanderbilt ultra soft , PAW] Using pseudopotentials materials, of any chemical composition can be computed with similar cutoff requirements.

1.2.2 From DFT to machine learning

SPE processes take hundreds of picosecond to be detectable at low temperature. In principle the classical dynamics of the ions in the system could be studied along the lines described previously solving the second order Newton equations

$$F_i(\{R_j\}) = -\nabla_i V(\{R_j\}), \tag{1.39}$$

where F_i is the force acting on the atom i and $\{R_i\}_{j\in 1..N}$ is the complete N atoms atomic configuration and V represents the ion-ion interaction energy that can in principle be obtained from DFT. However, even with advanced techniques such as the Car-Parrinello dynamics [51], due to the large samples and long simulation times needed to address the SPE, these simulations are too computationally demanding to be performed directly at the DFT level.

As a consequence researchers have been forced to trade accuracy for speed and all simulations so far have been performed using empirical interatomic potentials, constructed on empirical observations, that took into consideration the main chemical properties of the system, and were fit to reproduce specific experimental (where available), or some DFT properties that were considered relevant for the problem under study. The assumption being that for simple enough functional forms the parameters found in this way would extend beyond the region of the phase space used to fit them.

A common class of these functionals describes the interactions among particles by performing an expansion into different contributions depending on an

increasing numbers of atomic coordinates:

$$U(\{R_j\}_{j\in 1..N}) = \sum_i u_1(\mathbf{R}_i) + \sum_{\substack{i,j \\ i<j}} u_2(\mathbf{R}_i, \mathbf{R}_j) + \sum_{\substack{i,j,k \\ i<j<k}} u_3(\mathbf{R}_i, \mathbf{R}_j, \mathbf{R}_k) + \cdots + u_n(\mathbf{R}_1, \ldots, \mathbf{R}_n) \tag{1.40}$$

where u_1 is a one-body term describing the interaction with external fields, u_2 is a two-body term, e.g. Coulomb repulsion and other spherically symmetric interactions, u_3 is a three body term, and so on. All the terms along this expansion should be considered, but in practice they are limited by taking into consideration two factors.

First, the interaction often only has an effective radius r_c beyond which it becomes negligible. This limits the number of ions that need to be considered and, by consequence, the maximum order or the expansion. A second constraint comes from the observation, through theoretical modeling, that the contribution from high values of n rapidly converges to zero. In general, the energy splitting in different components is not uniquely defined, and may depend on the variety of the atomic configurations to be described. In the case of a periodic solid it can be addressed by a procedure that defines the one body term as the average energy per atom in the crystal, while the remaining error is then fit to the two-body term, and so on showing that expanding up to three-body terms is often enough.

On the basis of the locality considerations, a main ansatz that can be introduced is to write the total system energy as a sum of local atomic contributions:

$$U(\{R_j\}_{j\in 1..N}) = \sum_i U_i(\{\mathbf{R}_j \,|\, |R_j - R_i| < r_c\}) \tag{1.41}$$

where r_c is a chosen interaction radius and the function U depends on the central atom.

These different assumptions are not always valid, a common case being the presence of charge distributions that create long range electrostatic interactions that must be explicitly accounted for. Fortunately, silicon is neutral and does not need such corrections.

Silicon, being one of the most interesting materials of the last century, has seen a multitude of different empirical potentials developed for its simulation. The most common ones that will be used throughout this work have been developed by Stillinger and Weber (SW) [41], and Tersoff (in several parametrizations) [52, 53, 38, 54]. Also relevant is the EDIP [40] potential. SW potential uses a two- plus three-body description, where the latter is formulated to simulate the strong directional bonds that are present in the tetrahedral structure.

EDIP potential improves this picture by adding an environmental dependence to the SW three-body term. Tersoff potential instead adopts a different expansion: it starts from a two-body Morse potential [55] that has an attractive and a repulsive term and it expands the theory beyond the attractive term, effectively making it a three-body term.

The Newtons equations resulting from these potentials can then be integrated numerically using, for instance, the velocity Verlet integrator [56], as implemented in one of the available general molecular dynamics (MD) codes. In this work, we will use the LAMMPS [57] code, for its extendable code-base, its stability and well written documentation.

The main issue associated to the use of empirical potential is of course their limited accuracy. The last decade has seen a rapid evolution of new techniques based on the development of very flexible interatomic potentials trained by linear regression, kernel methods or neural networks (NN) [58] on datasets of atomic configurations whose energies and forces are evaluated at the DFT level. In contrast with standard approaches with their rigid, although physically motivated, functional form, these very general and often hyper-parametrized machine-learned potentials can describe much more complex potential energy surfaces (PES), thus delivering much better accuracy in regions of the material phase space well represented in the training dataset. In regions of the configurational space that the potential has not been trained on, the model can lead however, to an unphysical behavior, thus stressing the importance of an accurate configuration sampling.

In this section we are going to briefly review our adopted NN approach. This technique share with many other the approach: *i)* the generic atomic configuration is decomposed in atomic local environments, defined by a cutoff radius r_c, *ii)* each local environment is encoded in a descriptor which is used as an input to *iii)* optimize the parameters of the model (a NN in our case) by reproducing the result of a representative dataset of atomic configurations . In the production stage the trained model is implemented in the LAMMPS code and can be applied to the atomic configurations of interest to obtain energy and forces.

The local environment descriptor

One of the key ingredients of any potential, based on NNs or otherwise, is the descriptor i.e. how the local environment is represented before being processed by the numerical method of choice. In empirical potentials like SW, Tersoff the descriptor is comprised directly from distances and angles among the atoms within the cutoff. Distances from the atom central to the local environment are used to write the potential function two-body part, together with the angles for the three-body part. On potentials based on the fit of high dimensional

parametric functions instead, the descriptor is a multidimensional vector that tries to represent the atomic density in the local environment around the central atom.

The literature behind the descriptors is vast and multiple choices have proven valid in different contexts. For simple multilayer perceptrons, the descriptors proposed by Behler and Parrinello [59, 60] (BP) and successive variations [61], have shown to be particularly appropriate. At the same time, potentials based on Gaussian processes are often fitted with the SOAP descriptor [62], and many systems have been successfully studied through these Gaussian Approximation Potentials (GAP). Recently, Atomic Cluster Expansion (ACE) [63] descriptors have made their appearance, and preliminary results are pointing towards their effectiveness, also when used in combination with the NN approach. While these descriptors will be considered in future works, the following discussion will focus only on the BP-like descriptors, that are the ones used here.

When formulating an atomic descriptor, the effect of translational, rotational and permutational symmetries must be taken into consideration. In principle, the fact that symmetry related configurations have the same energy could be inferred from the data but the amount of data required would be enormous and it is much more convenient to enforce invariance between equivalent atomic configurations at the level of the descriptor.

BP *symmetry functions*, are discrete vectors $\mathbf{G}$ of features that represent a local environment and have identical representation for symmetrically equivalent configurations. These functions describe the local atomic configuration sampling the space on a carefully designed translationally and rotationally invariant grid. The concept of locality of a certain environment is enforced by using a cutoff function:

$$f_c(r) = \begin{cases} 0.5\left[\cos(\frac{\pi r}{r_c})+1\right], & \text{if } r \leq r_c \\ 0, & \text{if } r > r_c \end{cases}, \tag{1.42}$$

were r is the distance from a central atom, and r_c represents the cutoff distance. A first part of the symmetry function descriptor is purely radial:

$$G^R_{m,s;i} = \sum_{i \neq j}^{\text{All atoms of kind } s} e^{-\eta(R_{ij}-R_m)^2} f_c(R_{ij}), \tag{1.43}$$

where the distributions of atoms $\{j\}$ around the central atom i with distances $\{R_{ij}\}$ is filtered by a series of Gaussian functions. These Gaussian functions have a width controlled by a hyperparameter η, and are centered on a regular grid of radial distances R_m. The index s is then used to differentiate between different species.

This two-body term alone, however, is not sufficient to fully describe the configuration, and a higher body angular term is added. The detail of this 3 body term may vary slightly between different implementations. This work will use the formulation proposed by Smith *et al.* [61] as implemented in PANNA code [64]:

$$G^A_{n,m,s_1,s_2;i} = 2^{1-\xi} \sum_{j,k\neq i}^{j \text{ of kind } s_1, k \text{ of kind} s_2} (1+\cos(\Theta_{ijk}-\Theta_n))^{\xi}\ e^{-\eta\left(\frac{R_{ij}+R_{ik}}{2}-R_m\right)^2} f_c(R_{ij}) f_c(R_{ik}) \tag{1.44}$$

where the contribution of each pair of atoms j, k around the central atom i is resolved over a number of radial channels (with resolution defined by η) and angular channels (with angular resolution controlled by the number of probed angles Θ_n and the exponent ξ.[2] This part of the descriptor is also resolved by species trough the s_1 and s_2 indexes. A practical example for a case with a single specie is presented in Fig. 9.

To summarize: the descriptor is a vector of finite, predefined, length. The first part can be viewed as a concatenation of two-body features similar to the radial distribution function, one for each species. The second part is a concatenation of radio/angular distribution functions again divided by species.

Neural network architecture

The mathematical foundation of a NN is the artificial neuron developed by McCulloch-Pitts [65] at the beginning of the nineties to study signal processing in the brain. Despite the high biological complexity, some of the basic functionality of a neuron can be mathematically modeled as:

$$y = f(\mathbf{w}\mathbf{x} + b) \tag{1.45}$$

where y is the output of the neuron, $\mathbf{w}$ is a vector of weights, b is a bias that together with the weights defines the parameters of the neuron, and f is a non linear activation function. The variable $\mathbf{x}$ is the input vector, or the data that

[2] As written, Eq. 1.44 has a discontinuous derivative when a triplet of atoms is collinear. To resolve this problem the $\cos(\Theta_{ijk} - \Theta_n)$ is substituted with the expression:

$$2\frac{\cos(\Theta_{ijk})\cos(\Theta_n) + \sqrt{1-\cos(\Theta_{ijk})^2 + \epsilon}\sin(\Theta_n)^2\sin(\Theta_n)}{1+\sqrt{1+\epsilon\sin(\Theta_n)^2}}.$$

This recovers the original formulation in the $\epsilon \to 0$ limit. The value $\epsilon = 1e-3$ is found to be sufficient to not alter the accuracy of the model in any appreciable way and completely address the discontinuity problem for integration with a sufficiently small time step.

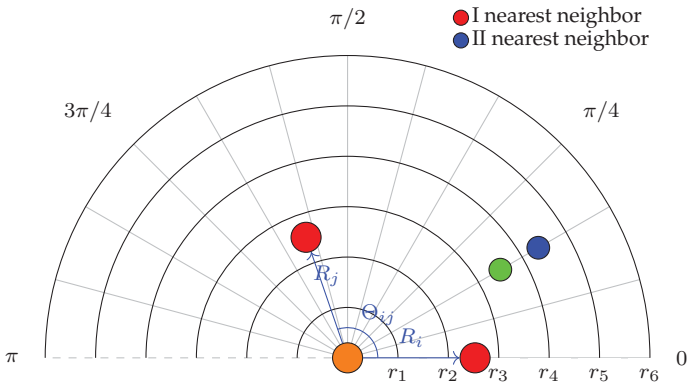


Figure 9: Pictorial view of the descriptors' three-body part. The radial grid shows the sampled radius between r_1 and r_6. The angular grid shows the sampled angles. The intersections between the two grids are the descriptor bins. The two parameters ξ and η of equation 1.44 are used to tune respectively the extension of the angle and radius probed by each bin. As an example the central atom of the environment is at the origin in orange,the two first nearest neighbors in a hypothetical c-Si structure are reported in red at positions R_i and R_j, and one second nearest neighbor is reported in blue. In this case, the triplet composed of the central atom and the first nearest neighbors populates the radial bins between r_2 and r_3 at an angle between $\frac{7}{12}\pi$ and $\frac{8}{12}\pi$. The triplet composed of the central atom, one first nearest neighbor (for convenience the one at a 0-degree angle), and the second nearest neighbor does not have $R_i = R_j$ and hence it will contribute to bins as if there is a virtual particle at an average distance(in green).

will be processed by the neuron. In different implementations f can have different shapes. The most common functional forms are the sigmoid $\frac{1}{1+e^{-x}}$ and the hyperbolic tangent, but also Gaussians and other*non linear* functions have been used.

The flexibility allowed by a single neuron is not enough to fully approximate the functions that we would describe. Theoretical results [66] have shown that networks of neurons, organized in layers of sufficiently large size, can implement functions of any complexity. We will consider here an all-to-all connected feed-forward neural network (FFNN), being among the simplest of many possible architectures. To create an all-to-all connected FFNN we will first collect several neurons creating a so-called layer:

$$y_i = f(\sum_j w_{ij} x_j + b_i), \tag{1.46}$$

where i is the index of the neuron inside the layer, and we have explicitly written the scalar product $\mathbf{wx}$. It is worth noting that the weights are now a matrix where the number of rows, i, is the size of the output of the layer while the number of columns, j, is the size of the input, which is shared by all neurons. Now that the concept of layer has been introduced, it is possible to stack a series of layers, using the output of one as the input of the next, obtaining the FFNN:

$$\begin{aligned} y_i^1 &= f^1(w_{ij}^1 x_j^0 + b_i^1) \\ y_i^2 &= f^2(w_{ij}^2 y_j^1 + b_i^2) \\ &\vdots \\ y_i^{\text{out}} &= f^{\text{out}}(w_j^{\text{out}} y_j^{\text{out}-1} + b_i^{\text{out}}) \end{aligned} \tag{1.47}$$

In this notation $\mathbf{x}^0$ is the input to the NN (the atomic descriptor in our specific case) and $\mathbf{y}^{\text{out}}$ is the neural network prediction (the predicted contribution of the given local environment to the total energy). In practice since we will only predict one energy for every descriptor the output vector is composed of a single element.

In the case of a NN potential (NNP) aimed at describing a system comprised of several species, the local contribution to the total energy depends on the nature of the central atom and one needs to train as many FFNN as the number of species present in the system. The descriptors are then labeled accordingly to the central atomic species and fed to the correct FFNN. This is presented schematically in fig. 10.

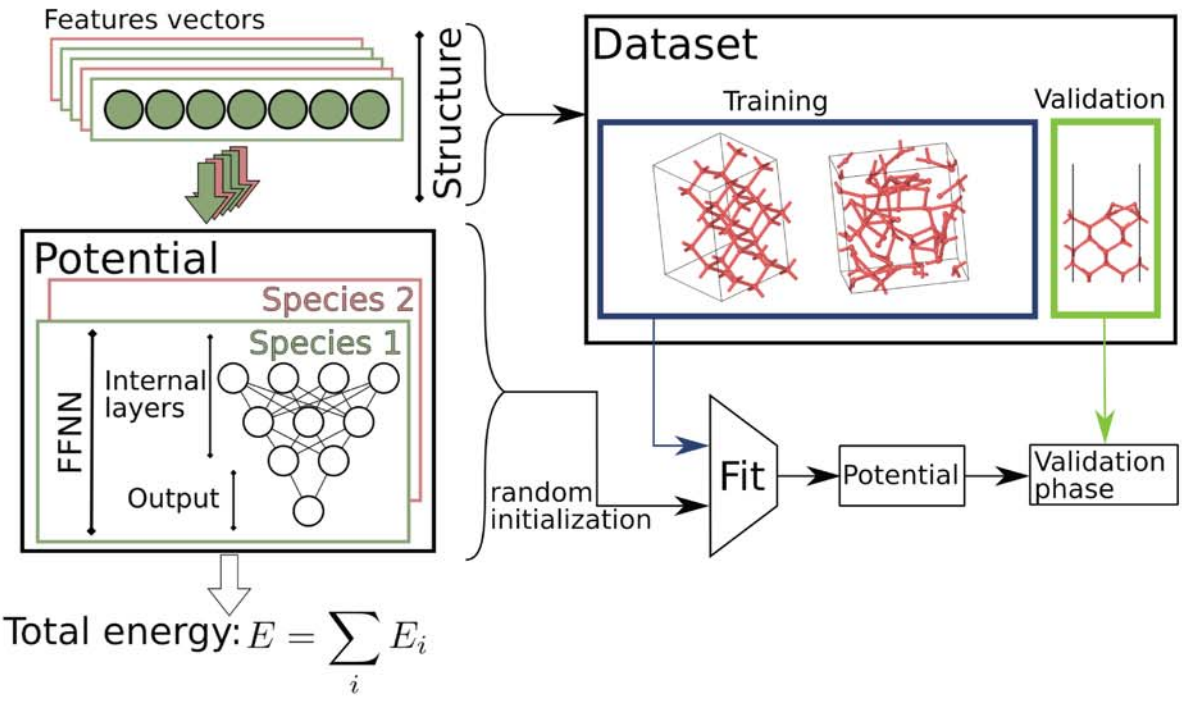


Figure 10: On the left side: a pictorial view of a FFNN can be realized as an ensemble of nodes arranged in layers. Each node or neuron is a uniquely parametrized mathematical function that takes as input the vector that is passed to the layer it is part of and returns as output the value of the function for the given input. Layers are chained together from the "input" to the "output" in a hierarchical manner such that the output of the above layer is given as input to the next layer. The vector passed to the input layer is commonly defined as the features vector and the output layer is composed of a single neuron, the output of which is the local contribution to the total energy. The extrapolation from one FFNN to a potential is trivial: each features vector is labeled with the species of the central atom and is processed accordingly by a different FFNN. The combination of all these networks (one per species) defines a full potential. On the right side: the main training/validation procedure described in the text is sketched.

Dataset, training and validation

At creation time, the NN parameters that compose the potential are randomly initialized. The weights are sampled from a uniform distribution, while the biases are set to zero. Starting from this point in the parameter space, an optimal model is found by evolving the parameters to minimize a "loss function" $\mathcal{L}$ with respect to a set of reference data called training dataset.

A dataset in this context is a set of structures resolved in total energy and forces by some reference method, DFT in our case. The creation of a dataset is a complex procedure that will be described in the next chapter for our application but the general idea consists in the generation, with different approaches, of atomic configurations that are representative of the phase space regions that will be explored by the problem under investigation. The dataset is then usually randomly split in 2 parts containing e.g. 80% and 20% of the data, respectively. The larger portion is defined as the training dataset, while the smaller one is the validation dataset. The training dataset is used in the first phase where the network is trained (fitted) by minimizing a cost function that is a measure of the distance between the neural network predictions and the true values. Usually, this part of the dataset is composed of thousands of structures and hundreds of thousands of atoms. The validation part of the dataset is instead used to verify the network accuracy over a set of points that have never been seen, in order to obtain a unbiased estimate of the error.

The loss function is a measure of the distance between the model prediction and its true value. We use the function:

$$\begin{aligned}\mathcal{L}(W) = \gamma_e \sum_{i \in \text{batch}} \frac{(E_i^{\text{DFT}} - E_i(W)^{\text{NN}})^2}{\#\text{atoms}_i^2} \\ + \gamma_f \sum_{i \in \text{batch}} \frac{\sum_{j \in \text{atoms}_i} \sum_{k=1}^{3} (F_{i,j,k}^{\text{DFT}} - F_{i,j,k}^{\text{NN}}(W))^2}{\#\text{atoms}_i} \\ + \gamma_{\text{norm}} (\frac{|W|^2}{2} + |W|).\end{aligned} \tag{1.48}$$

This function is composed of 3 weighted contributions. The first contribution is the squared error between the exact DFT energy E_i^{DFT} and the predicted values $E_i(W)^{\text{NN}}$ over a limited set of training configurations i in a batch. In this equation, W represents all the network parameters. Since different configurations can have a different number of atoms their contribution to the squared error is normalized with $\#^2_{\text{atoms}_i}$. In a standard gradient descent minimization gradients at each step are computed for an objective function summing over all the available points in the dataset. In this case, however, we restrict our sum over a batch of randomly extracted elements. This is done for two reasons: first, a practical one, as the dataset will often not fit in memory all at the same time. The second is a theoretical one, as the usage of batches continuously changes the shape of the target PES, introducing a noise term that helps avoid local minima. The second term is equivalent to the energetic one, but it works on forces. F_{ijk}^{DFT} is the exact force of the configuration i and atom j along the direction k and F_{ijk}^{NN} is the value predicted by the NN. This quantity is in principle unnecessary, since

energy data should naturally contain force information. Nevertheless, it would be very expensive to generate a dataset dense enough to describe its curvature around each point. It is instead convenient to use more information available from the configurations already computed. The third term consists of the L^1 and L^2 norms of all the parameters and is introduced to limit their growth since this seems to help in converging the model.

Different cost functions can be written to better reweigh outliers during training. However, we have found in this work that this simpler form leads to equivalent results.

Another hyper-parameter is the learning rate α. This enters in the equation that defines how weights are updated:

$$W^{t+1} = W^t - \alpha * \nabla_W \mathcal{L}(W^t), \tag{1.49}$$

where all the quantities have been previously defined and the subscript t is the minimization step. At each step t, the gradient of the loss function (1.48) with respect to the network parameters is computed and subtracted from the current set of network parameters. The result is a new set of optimized weights that has a lower loss function and hence is capable of better predictions. α regulates how big the step is in the weight space during the minimization. Since the gradient is a local property, a large value for α can overshoot its radius of trustfulness. This result in steps that do not minimize the loss function. On the other hand, small alpha can lead to very slow minimizations.

In practice, the real implemented equation is $W^{t+1} = W^t - \alpha * \text{Adam}(\nabla_W \mathcal{L}(W^t))$. Adam, or Adaptive Moment Estimation [67], is a protocol that uses running averages over the gradients to tune the learning rate at each minimization step. Regions where the gradients are large, tend to lower the learning rate while regions where gradients are small, tend to allow for steps as big as α. This greatly improves convergence and final accuracy helping to avoid being trapped in local minima.

Finally, to assess the fit quality, a validation is performed. In this phase, the NNP is evaluated against the validation portion of the dataset. Predicted energies and forces are compared to the DFT values, and root mean square errors(RMSE) are computed. If such errors are sufficiently small physical properties of the NNP are also tested against their DFT counterparts.

1.2.3 Codes

This work exploits several well established open source software tools and packages and ad-hoc libraries. The Machine Learning (Neural Network) part is based on PANNA [64]. PANNA is a code developed by our team and based

on TensorFlow [68]. Binary outputs are managed by a library (Molecular data base library) that I have developed. This library also provides several common utilities to work with atomic configurations. Jupyter, Numpy [69], Pandas [70], Matplotlib [71], and Seaborn [72] were used for the data analysis and visualization. Other codes used were: LAMMPS [73] with the PANNA plugin for the MD runs (part of the PANNA code); Ovito [74] for MDs visualization and post-processing. Quantum ESPRESSO (QE) [75], a software distribution for ab-initio calculations within the framework of Density Functional Theory, was used to perform all the required DFT calculations.

Chapter 2

Creation of the potential

The creation of the potential, as explained in the sec. 1.2.2 comprises two phases: a training phase and a validation phase. Both of these phases rely on a dataset of configurations, and this dataset physical expressivity is what determines which regions of the phase space are ultimately described by the potential. To create an optimal dataset an iterative procedure is often employed [76, 77]. It starts from a sampling of the phase space of interest. MDs or other techniques (eg. KART [78], NEB [79]) are adopted depending on the physical phenomena under investigation. The energy and forces of the atomic configurations thus collected are then evaluated with an accurate method of choice (DFT in our case) and these data are included in the dataset and a NNP is trained so as to reproduce them at best.

At the beginning, when the potential used to generate the atomic configurations is not very accurate, the atom dynamics will follow inaccurate PES, sampling configurations in the phase space far from the ones corresponding to the ground truth level of theory (i.e. DFT). As better potentials become available from the previous iterations more accurate PES will be followed, and more relevant configurations will be sampled and added to the dataset. The procedure can be iterated until a self-consistent solution is obtained without significantly new configurations to be added to the dataset.

During each cycle of this self-consistency, the potential performance is monitored by evaluating standard error metrics over the validation dataset (containing configurations not used during training) as well as monitoring a set of relevant (easily computable) physical properties. This quantitative analysis helps to locate whether and where more sampling is needed, and once a satisfactory accuracy is reached the potential creation procedure terminates and production runs can be undertaken.

The procedure mentioned above adapted to our work is shown in Figure 11,

and this chapter will elaborate on it in great detail. First, we are going to describe how to generate a meaningful dataset for each of the different explored phases (Sec. 2.2). This will be done while highlighting interesting results of surface properties related to the surface reconstruction phenomenon and optimization of the DFT related parameters. We are then going to discuss the training of the neural network potential that will be used in the next chapter to investigate SPE and how we validated it against an independent set of atomic configurations and additional physical properties (Sec. 2.5). Finally, a discussion on the ability of the ML approach to reflect DFT choices will be put into place in Sec.2.6.

2.1 Computational details of DFT calculations

To achieve accurate energy and forces for the configurations included in the training and validation datasets, we need accurate DFT calculations. This section discusses the main design choices and convergence criteria that we adopt to ensure accuracy.

The most important choice underlying the overall accuracy of the DFT calculations used as reference for our studies is the choice of the exchange-correlation functional. Among the plethora of possible functionals that have been proposed along the years, we selected two parametrizations of the PBE functional family as reported in sec. 1.2.1. First we considered PBESol parametrization due to its reported higher accuracy in solid-state applications [48] for which it was explicitly developed. As we will see in the following, we then switched back to the original PBE parametrization [81, 82] because of its better performance in the description of Si melting.

On the numerical side, the primary source of error is defined by the k-point density used when performing integrals in the reciprocal space to extract physical quantities (Sec. 1.2.1). To estimate this quantity we started from c-Si. In c-Si k-point sampling converges quite quickly, and a k-point density of 15.3 k-points $\cdot$ Å is sufficient to reproduce rather complex quantities such as phonon dispersions [83]. Disordered phases, such as a-Si and l-Si, are simulated with the single point Γ [84, 85, 86, 87] in rather large unit cells corresponding to a k-point density of the order of 16.3 k-points $\cdot$ Å. One of our goals is to accurately describe silicon surfaces and we therefore took this system as reference, obtaining a required density of 46.1 k-points $\cdot$ Å, in line with accurate surface studies in the literature [88]. See Sec. 2.2.4 for more details.

The presence of metallic configurations in the dataset is a source of numerical instabilities and convergence problems. To solve these issues a fictitious electronic temperature (smearing) is usually added. This induces partial occupation

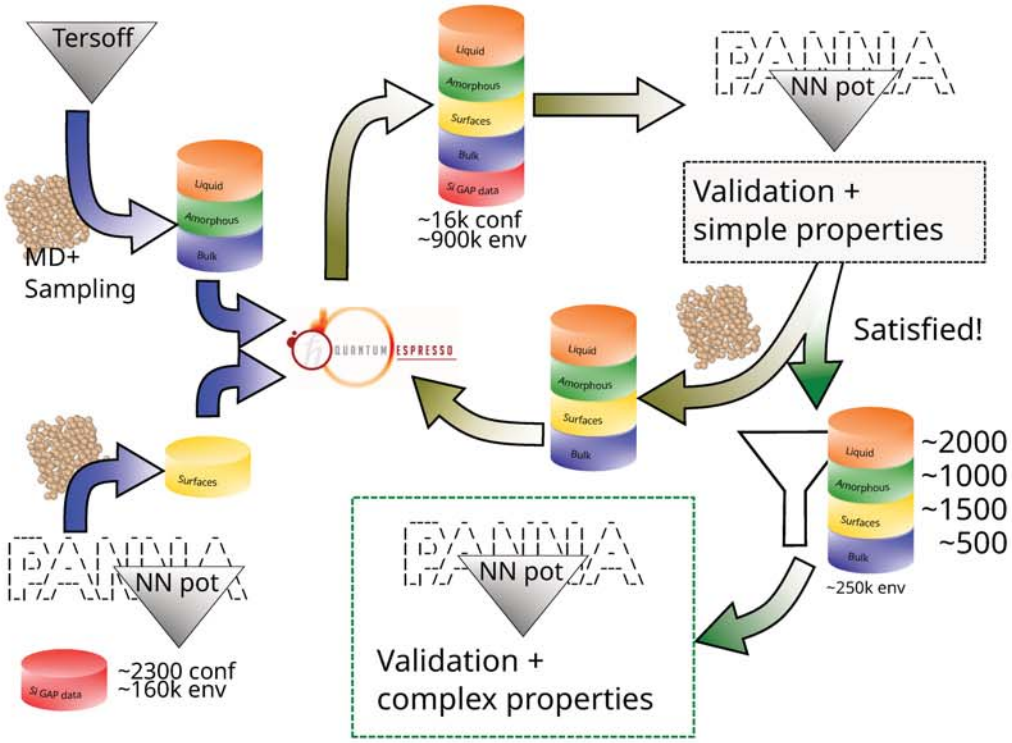


Figure 11: Data creation procedure we followed in this work. The first step is the choice of a simple potential or a series of potentials that are known to describe correctly macroscopical properties of the regions of interest. Then, a series of MDs covering the region of interest are performed with such potential, and from those simulations, we sampled the set of structures to create the initial database. In our case, we could not find an empirical potential capable of capturing the surface properties we were interested in. Hence we trained a first potential directly on a dataset[80] explicitly developed for gap potentials. We used it to sample surfaces using the same MD/sampling strategy described before. We then resolved all the collected structures in energy and forces at DFT accuracy using the Quantum espresso engine. With this operation, we obtained our first dataset. We then used this dataset to train a potential we validated against several simple properties. If the validation were satisfactory, we would have stopped here. If they were not, we must collect new data from the regions of interest to extend our database, resolve them at DFT accuracy, and train/validate a new potential. Once we were satisfied with the accuracy of our simple properties, we removed from the dataset structures that were not representative of our problem. Finally, we trained a potential and validated it on the same set of simple properties and more complex ones.

of the anti-bonding states and weakens the typically strong covalent bonds of c-Si. Different kinds of smearing are possible and in this work Marzari-Vanderbilt smearing was chosen because of its mild dependence on the smearing width and its reported good accuracy in surface applications [50]. In principle, electronic temperature could be optimized for each configuration and chosen as small as possible to minimize this side effect, but this is not possible in the high throughput approach adopted here. After a preliminary convergence study, a smearing width of 0.05 eV was fixed as a good compromise between accuracy of the calculation and numerical stability.

Another important convergence parameter is the kinetic energy cutoff of the plane wave expansion that defines the completeness of the basis set used. Cutoff energies of 680 eV for the expansion of the wave functions, and 3400 eV for the charge density expansion, were selected as sufficient for an accurate treatment of the pseudo-potentials used that were taken from the standard PSlibrary of the Quantum ESPRESSO distribution. Convergence tests were performed for c-Si and, since both considered functionals performed very similarly, we used the same parameters for both types of calculations.

2.2 Dataset creation

The creation of a potential is based on the knowledge of the configurations that will be explored during the simulation of the target physical phenomenon. It has been shown [76] that the creation of potentials trained on narrow regions of the phase space can be very accurate where trained on, but suffer from poor generalization in unexplored regions; on the other hand potentials trained on wider regions of the phase space, while in principle more robust, may be forced to make compromises in any specific phase space region, thus resulting in a less accurate potential in the region of interest. A balanced choice must then be made.

As a consequence, for this study, we restricted our attention only to those regions that are likely explored during SPE. As introduced in Sec. 1.1, the most relevant regions can be identified as the crystalline solid, the amorphous phase, the liquid phase, and some of the surfaces of the crystalline phase. These last ones are relevant because of the limited size of our sample, and for future interest in low Miller index surfaces (and their *vicinal* surfaces) [89] that are known to be a key feature in SPE regrown samples [90].

In the first dataset iteration different potentials were used to explore different regions of the phase space. Starting with potentials that follow PES closer to the real one implies that sampled configurations are already a good representative of those in the phase space region under investigation. Such optimization

results in the need for fewer iterations to converge the dataset. At each iteration the collected data was obtained by combining MDs with different sampling methodologies, depending on the region under study as explained in the following sections. This allowed us to capture relevant features that led to a better final potential.

2.2.1 Crystalline silicon

Starting with the structure with the lowest entropy, c-Si has been sampled in its harmonic basin by using a Tersoff potential [53]. For this particular step, we used a tetragonal cell with 32 atoms. This choice was made in order to have fast DFT calculations with local environments that do not have atom periodic replicas inside the local environment cutoff radius. The phase space spanned by c-Si thermal vibrations is sufficiently well explored by any empirical potential. Hence, a few thermal and pressure ramps have been deemed sufficient to correctly capture all the needed physics.

Sampling structures that average at different per atom volumes (eg different densities) is very important in this case, as generating a dataset with a single density would only succeed in reproducing the local energy volume curvature but not its linear term. This can be seen by expanding the internal energy with respect to the volume around the equilibrium value. Using the notation introduced in section 1.2.2:

$$\begin{aligned} U(\{r_j\}_{j\in 1..N}, V) &= \sum_i u_i(\{\mathbf{r}_j| \, |r_j - r_i| < r_c\}, v_i) \\ &= \sum_i u_i(v_{\text{eq}}) + \left(\frac{\partial u_i}{\partial v}\right)_{v=v_{\text{eq}}} \sum_i (v_i - v_{\text{eq}}) \\ &+ \left(\frac{\partial^2 u_i}{\partial v^2}\right)_{v=v_{\text{eq}}} \sum_i \frac{(v_i - v_{\text{eq}})^2}{2} + \dots \end{aligned} \tag{2.1}$$

where V is the total cell volume, v_i is the volume occupied by the atom i, and v_{eq} is the per atom equilibrium volume. We can see that the linear term is proportional to $\sum_i (v_i - v_{\text{eq}})$ but if the overall cell volume is not allowed to change during our sampling this term vanishes. The result is that during the fit no constraint is applied to the linear term. This is not the case for higher-order terms.

The simulations were performed at 300K and explored a range of pressures from -15 GPa to +65 GPa. From the simulations, we collected 1162 configurations by sampling them from the dynamics at equally spaced intervals. The presence of high-pressure regions is necessary to sample the strong short-range repulsion. Such repulsion term prevents atoms from collapsing on each other.

In standard empirical potentials, the repulsive term is built-in by construction in the choice of the functional, but for NNPs each desired feature needs to be represented in the training set. A sketch of a potential with such short-range issues is shown in Fig. 12. Above 14 GPa, we observed a phase transition and the system goes through a state that seems compatible with the β-tin [91] transformation predicted by the potential [92]. The reached state is not a real physical state but rather an idiosyncrasy of the Tersoff potential. Since we do not plan to reach such high pressure in any of our simulations, we didn't further investigate this region. We kept the few collected points in that region only to provide a strong interatomic repulsion.

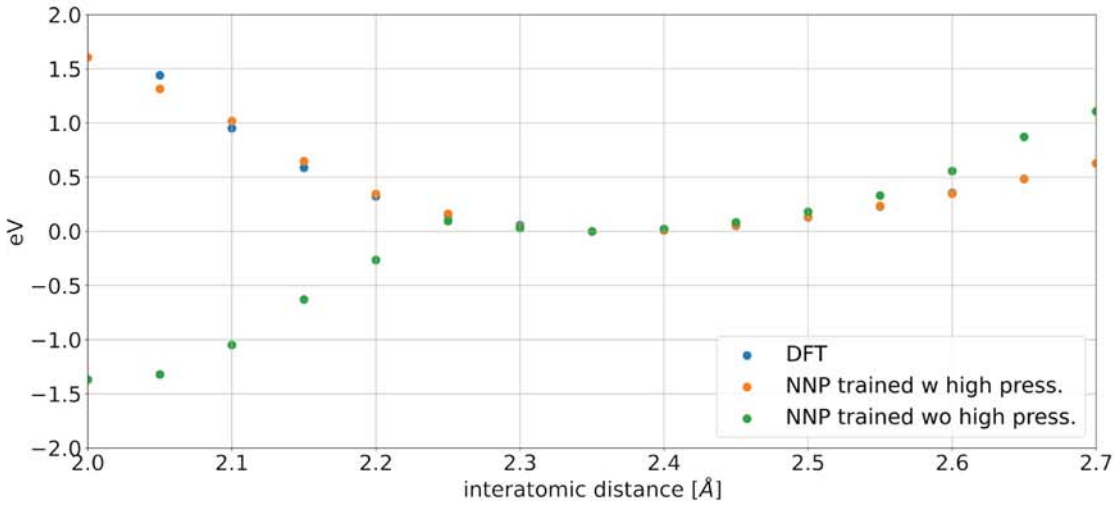


Figure 12: Equation of states curves for the diamond structure with different potential realizations. In Blue DFT references. In orange a potential that has been correctly trained with diamond structures with an interatomic distance as small as 2 Åand high pressure structures. In green a potential trained without high pressure configurations and without structures with an interatomic distance below ≈ 2.25 Å. The potential depicted in green shows a spurious attraction at short distances. This is reflecting the lack of high pressure structures in the training set. The major issue of this spurious minimum and the existence of a barrier-less path to reach it is the existence of un-physical bonding states.

2.2.2 Amorphous

A full microscopic knowledge of the amorphous structure has yet to be determined, and there likely is not only one amorphous but rather a large variety

of them. This lack of knowledge, however, has not limited the amount of research efforts to identify possible microscopic atomic arrangements. To date, there are mainly two different numerical approaches to obtaining a-Si structures. The cheaper WWW [93, 94, 95] approach is based on the iterative distortion of local bonds while keeping the tetrahedral coordination and relaxing with a potential. This approach preserves the local order while changing bond-angle and length distributions, generating a structure that has a radial distribution function compatible with experimental measurements. More recently, thanks to the increase in computational power, the golden standard for the generation of amorphous samples has become the melting-quenching technique. This technique tries to replicate what is done in some experiments when obtaining the amorphous phase: the sample is first melted and then it is slowly cooled to the amorphous state. Metastable glassy phases could be reached if the quench is too rapid, but this can be avoided by reducing the cooling-ramp steepness. Upper bounds on extreme experimental ramps are estimated to be of the order of $10^{10} \sim 10^{12}$K/s [96].

Two potential forms are commonly adopted in a-Si investigations: SW and Tersoff potential types. The original SW potential [41] can not be be employed directly in the melting-quenching technique. This is due to SW's inabilities to correctly describe the l-Si to c-Si phase transition [97], that instead of generating a structure mostly composed of fourfold coordinated atoms generates a glass with an excess of 5-fold coordinated atoms. To overcome this issue, if the l-Si/a-Si phase transition is of interest, it is possible to tune SW parameters and the simulation details so as to force the appearance of fourfold structures [98] during a-Si phase formation. We excluded this approach because it is rather hand-wavy, and does not guarantee good properties, other than the coordination number, during the phase transition.

Tersoff potentials on the other hand, despite having a severe problem with the liquid phase, where it underestimates the density [99], undergo a realistic phase transition providing a-Si of good quality [100]. Since the error in the liquid phase can be easily corrected by performing dynamics at the desired constant volume we started our exploration of the amorphous phase with this functional form, in the original Tersoff [53] parametrization.

A sample of 64 atom/cell, a number chosen to allow a fast DFT computation, was selected. The cell was prepared at a density that is compatible with the experimental crystal density. The box was then melted in a canonical (NVT) simulation at 3500K using a Nosé-Hoover thermostat. This temperature was chosen because it is well above the Tersoff melting temperature of approximately 2500K. The liquid state of the final configuration was assessed by computing

the radial distribution function structure and comparing it with known l-Si references [99]. From the melted system, another NVT MD with a target temperature of 2500K of 110 ps was performed saving a snapshot every 10 ps. After discarding the first 10 ps of the simulation due to equilibration, the remaining 10 configurations were used as a starting point for the liquid annealing. The annealing was performed by changing the bath temperature with a cooling ramp of $6 \cdot 10^{12}$K/s down to a temperature of 500K.

During annealing, the pressure was monitored showing a positive value typical of Tersoff potentials at this volume in the liquid phase. This is in contrast with the experimental negative value that would lead to liquid contraction. As the temperature drops, the pressure decreases reaching zero at about 800K.

The limited size of the samples poses serious limitation to the configurational variability that is in principle present in the a-Si. In particular at very low temperatures, the system freezes and can remain stuck in a metastable state. To investigate this phenomenon, we employ a simple description based on the measured mean-squared displacement (MSD):

$$R^2(t) = \frac{1}{N} \sum_{i=0}^{N} \left| \mathbf{r}_i(t) - \mathbf{r}_i(t_{\text{final}}) \right|^2 \tag{2.2}$$

that is an ensemble average over the particles in the cell. In this expression, $\mathbf{r}_i(t_{\text{final}})$ is the reference position of the particle i at the end of the simulation, and $\mathbf{r}_i(t)$ is the position of the particle i at time t. An example of the evolution of this quantity on a single MD run is reported in Fig. 13.

Fig. 13 shows that the exploration of these metastable states starts approximately at 1200K, and for each one of them the soft modes are probed. Although it could be argued that the identified clusters might not correspond to a single cluster but rather multiple clusters sharing similar distance from the final configuration, this was dismissed by analyzing a few configurations and verifying that the clusters do not undergo major changes over a long period. Furthermore, the change from one cluster to the next is usually due to the change in a single local environment. This happens through the crossing of an energy barrier that separates the two local minima. This behaviour was also observed later with the NNP at different temperatures and with a different frequency. Such a scenario would be very unlikely for larger systems (more than 64 atoms).

It is important to note that this approach, that will be repeated in the iterative part of the potential generation, will not sample any voids in the a-Si. It is possible that the final potential may not be well suited for a-Si with pockets that could be realized by evaporation [4]. Thankfully, experimental SPE samples are usually prepared by implantation processes which avoids this problem.

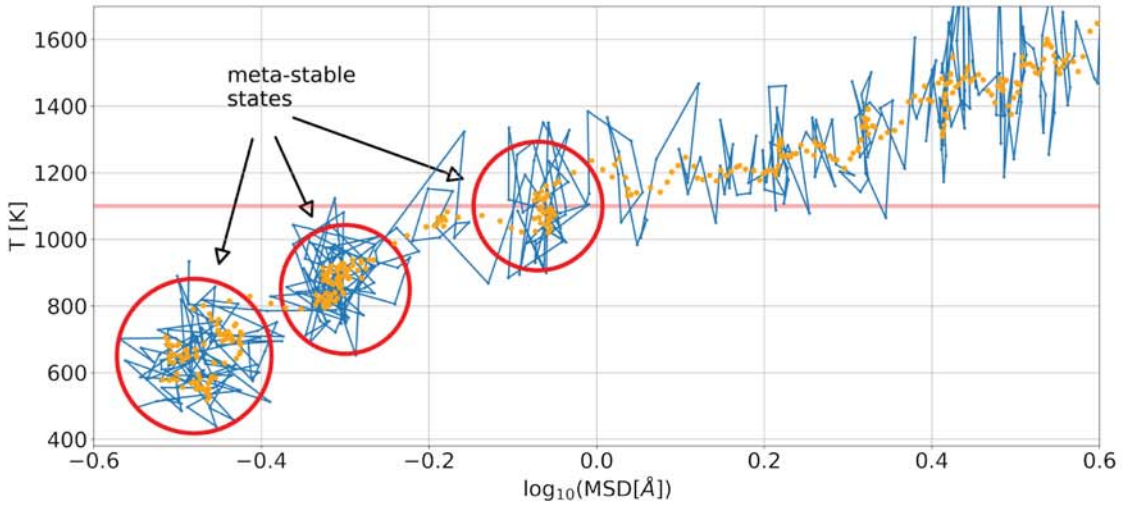


Figure 13: MSD in Å with respect to the last configuration as a function of the temperature. Blue lines represent successive sampled time steps in the simulation. Orange points are running averages over 7 blue points to aid the understanding. In the presented MD simulation the system visits 3 meta-stable states circled in red and recognizable from the clustering of the orange dots. The red line at 1100 K represent the arbitrary temperature below which configurations where sampled.

This procedure was repeated six times, first with Tersoff and then with different iterations of the NNP. Configurations were randomly sampled from the dynamics below the temperature where metastable states start to be explored. A total of 4485 amorphous configurations (accounting for 287040 local atomic environments) were collected.

2.2.3 Liquid

Choosing an initial potential for the liquid phase is easy because any functional form, if the correct density is imposed, produces reasonable geometrical properties (e.g. the radial distribution function). Given that we used the Tersoff potential for the previous two phases, we also adopted it for the liquid phase.

Knowing that the equilibrium density of Tersoff potential l-Si phase is larger than the equilibrium density of experimental l-Si, we decided to address this problem by adding an external pressure. For this reason, a preliminary NVT run at the experimental l-Si density (0.0522 atom/Å^{-3} [41]) was performed to determine the pressure needed to keep the system stable at this density. This resulted in a pressure of approximately 8.9 GPa.

For the first sampling, the positions for the initial l-Si structure were taken from an l-Si configuration encountered in the a-Si investigation and scaled to the desired density. At this point, both the NVT and isothermal-isobaric (NPT) simulations were run at 2500K for 500 ps and uncorrelated samples were taken every 50 ps. These simulations were performed on 64 atom cells following what it was done in the a-Si exploration.

We explored the possibility to use larger 216-atom cells but did not include these simulations in the final dataset as pathological cases that were difficult to converge in DFT calculations were rather common, making this size impractical for the high throughput approach. Selected structures were investigated and we found their metallic behavior and the Marzari-Vanderbilt smearing chosen to deal with it (Sec. 2.1 and 1.2.1) to be the main issue.

This procedure was repeated eight times, first with Tersoff and then with different iterations of the NNP. Simulation temperatures and pressures were adjusted accordingly to the potential performances. A total of 4612 configurations (accounting for 194248 atomic environments) were collected.

2.2.4 Surfaces

The sample size that we will be able to afford when simulating SPE poses an additional problem: its thickness is limited. The natural consequence is that the

a-Si/vacuum interface will be close to the c-Si/a-Si interface. This limited thickness forces the creation of a potential able to stabilize the surface and provide realistic dynamics. Behaviors to be avoided are, for example, fictitious sublimation, or unrealistic reconstructions. Another advantage of having a model that contains this information comes from its ability to reconstruct steps like those that are present in the vicinal surfaces. This is not immediately useful in our case but will leave doors open for future investigations of these surface orientations.

To create a model capable of describing surfaces, configurations including them must be sampled as done with the other phases. Since complex reconstructions are possible, before proceeding with the sampling, a detailed investigation is required.

As illustrated in the background (Sec. 1.1.1) SPE proceeds along different crystalline directions with different speeds. The main direction of SPE regrowth that we will target is the $\langle 100 \rangle$, because of its high technological utilization. As a consequence, the surfaces of focus will be the (100). Another valid reason to start from this surface is the limited number and cell periodicity of the possible reconstructions: the (100) surface is known to have a set of five possible reconstructions [101]. Three of these are known to be stable and the other two are not experimentally observed and are needed to define reference quantities and/or compare with other works.

The p(1x1) or ideal reconstruction is a structure with a surface periodicity of 1 single atom. In this case, during the relaxation, no symmetry is broken, and the only relaxation that takes place is a contraction along the $\langle 100 \rangle$ direction. This results in the reconstruction with the highest energy (per surface atom). The p(2x1) symmetric (s-p(2x1)) reconstruction is the first real reconstruction: composed of two p(1x1) surface unit cells along the $\langle 001 \rangle$ direction it is obtained by moving the two atoms in the top layer closer together (dimerization) to remove one of the two dangling bonds pointing outwards from the surface. The net result is the formation of rows of dimers along the $\langle 010 \rangle$ direction on the surface with a decrease in energy of ≈ 2 eV/dimer w.r.t. the ideal 1x1 case. Due to the Jahn–Teller effect, the dimer symmetry can be broken by displacing one atom of the dimer further and moving the other atom closer to the surface. This symmetry-breaking lowers the dimer energy by a further ≈ 0.1 eV/dimer, and the obtained configuration is referred to as p(2x1) asymmetric (a-p(2x1)) reconstruction. Alternating the direction of symmetry-breaking of the dimers along the $\langle 010 \rangle$ direction can lower the energy by another ≈ 0.02 eV/dimer, leading to what is known as p(2x2) reconstruction. If two consecutive rows of dimers are shifted by one cell along the $\langle 010 \rangle$ direction the slightly more stable reconstruction c(4x2) can be obtained, further lowering the energy by 2 meV/dimer.

A top view of these different surface reconstructions is shown schematically in Fig. 14.

We realized the surface calculations by creating periodic (super)cells containing a finite slab of c-Si exposing two (100) surfaces. The slab must be thick enough to accommodate elastic distortions propagating from the two surfaces in the Si bulk. In addition, the cell must include a vacuum layer sufficiently wide to avoid surface-surface interaction. This makes the calculation expensive. A common trick to reduce the cell size and the cost of the calculation is to consider only half of the slab, freezing the atoms on one side in their bulk positions, passivating the dangling bonds with hydrogen atoms, and studying the remaining single surface. Unfortunately, while this approach can be used to determine the relative stability of the different reconstructions, as the contribution from the fixed artificial passivated surface cancel out in the comparison, this cannot be done in our case where the absolute energy of the different reconstructions is required and the presence of the hydrogen atoms would need to be modelled as well, making the potential design and training needlessly more complex.

Elastic distortions caused by surface reconstruction extend for several layers within the bulk of c-Si, and the minimization of large slabs may be expensive. Therefore, we developed a procedure to speed up these calculations. First, we considered a crystalline slab of 8 atomic layers with the surface periodicity of the desired reconstructions and we added a 10 Å vacuum layer in the $\langle 001 \rangle$ direction to create the two surfaces. The exponential decay of the wave functions in the direction perpendicular to the surface makes the chosen distance sufficient to grant decoupling between the surfaces. The surface atom positions were then perturbed by a few tenths of an Angstrom in the direction of the desired relaxation and the sample was relaxed with the Broyden–Fletcher–Goldfarb–Shanno (BFGS) algorithm as implemented in QE. The DFT computational details used were those fixed in section 2.1 together with a first low-accuracy grid with 15.4 k-points $\cdot$ Å. Atoms located in the middle of the sample were kept frozen during the relaxation. This was done for a slab with a width of a unit cell. If at the end of the relaxation the forces on the frozen atoms exceeded the BFGS tolerance of 26 meV/Å, an additional crystalline bilayer was inserted at the center of the slab and the procedure repeated freezing only the central atoms. Once the cell was large enough to accommodate all the elastic distortions making the forces acting over the frozen atoms negligible, the k-points grid was increased and the cell relaxed again. If the forces remained below the threshold convergence was reached, alternatively more layers were added and a new iteration was performed.

The procedure just described was applied to the five surface reconstructions,

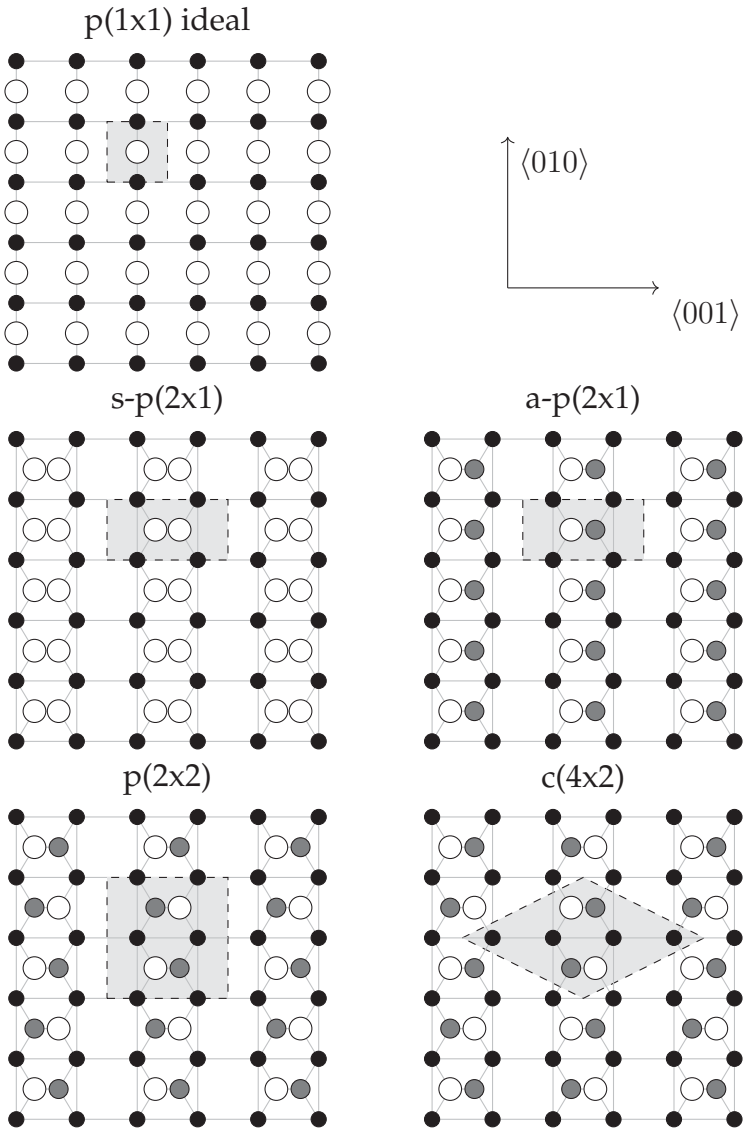


Figure 14: Top view of the unreconstructed (ideal) surface and the four reconstructions to be considered in this study. Smaller and darker circles represent deeper atoms. Thus the small black circles are second layer atoms and the larger grey and white circles are surface atoms. In the reconstructions with buckled dimers, the large white circles protrude further from the surface than the grey circles. The dashed lines and shaded areas show the surface unit cells used in the calculations.

and the two different functionals, PBE and PBESol, were tested to investigate their geometrical and energy differences. The final energy differences between the relaxed reconstructions are presented in Table 2 where the per-dimer energy is defined as:

$$E_{\text{dimer}} = \frac{E_{\text{cell}} - E_{\text{equivalent ideal surface}}}{n_{\text{dimers}}}, \tag{2.3}$$

and the surface energy as:

$$E_{\text{surface}} = \frac{E_{\text{cell}} - E_{\text{bulk}} * n_{\text{atoms}}}{A} \tag{2.4}$$

where E_{bulk} is the per-atom energy of the ideal crystal and A is the total surface area of the considered unit cell (in our case twice the surface exposed by the cut because of the PBC). For these calculations a final k-points density of 46.1 k-points · Å was deemed sufficiently accurate for both functionals.

The error due to the finite vacuum size was investigated on the a-p(2x1) reconstruction for the PBESol functional. A cell with 10 Å of vacuum was compared to one with 15 Å of vacuum, and an error of only 1.8 10^{-3} meV/atom was found. Adding extra vacuum didn't cause any significant further relaxation.

The results collected in Table 2 show that the energetic ordering for the two functionals is identical and that the c(4x2) reconstruction is always the most stable. In Table 3, a comparison with LDA is also presented showing that both PBE-type functionals are in agreement with the commonly accepted energy order. Unlike previous findings [101] (and references therein), the s-p(2x1) configuration hereby reported is not a stable state. It was found that such symmetry is a local minimum only if the BFGS relaxation is performed at an insufficiently accurate precision (in k-point grid density). The reported values are obtained from a calculation relaxed at a k-points density of 15 k-points · Å and then recomputed with a single point self-consistent calculation at the same higher accuracy as the other calculations.

Regarding geometrical (schematic views in Fig. 15 and 16 and energy analysis (Table 3), the description of c-Si (100) surfaces is very similar for both PBE and PBESol functionals. Almost equivalent geometries are found and elastic distortions propagate similarly through the sample (both require similar width to converge). From these analyses, it is hard to ascertain which functional best describes the surfaces, also due to the lack of consistent high-resolution experiments at low temperatures. Furthermore, this suggests that the dynamics performed with the two functionals should be very similar, allowing a simple switch from one to the other without the need for new sampling.

In the literature, different electronic quantities, such as charge distribution and band structures are also discussed. But since the potential we are aiming to

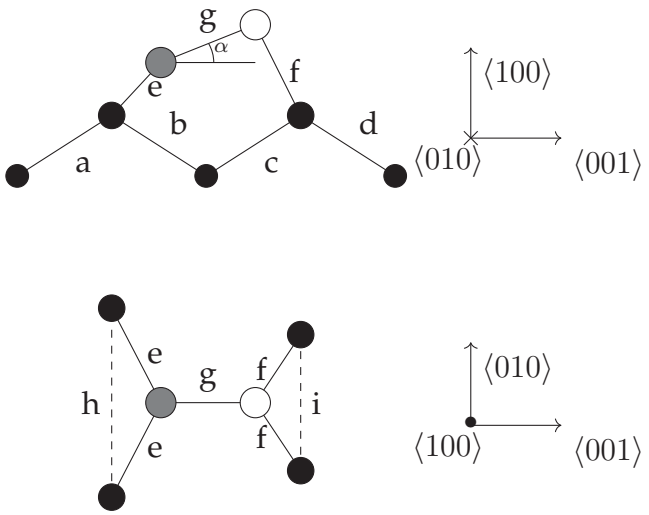
surface	dimer	PBESol			PBE		
		slab-size	E_{dimer}	$E_{surface}$	slab-size	E_{dimer}	$E_{surface}$
ideal	-	20	0	143.607	20	0	134.703
s-p(2x1)	2	32	-1.533	91.668	32	-1.456	86.028
a-p(2x1)	2	36	-1.702	85.955	36	-1.614	80.746
p(2x2)	4	36	-1.768	83.706	40	-1.684	78.431
c(4x2)	4	40	-1.775	83.486	40	-1.686	78.359

surface	dimer	NNP	
		E_{dimer}	$E_{surface}$
ideal	-	0	110.117
s-p(2x1)	2	-	-
a-p(2x1)	2	-0.894	79.836
p(2x2)	4	-0.983	76.812
c(4x2)	4	-0.981	76.880

Table 2: Energy difference per dimer and surface energy for the two functionals. The reported slab-size is the number of layers of the final slab and dimer is the number of dimer. In the NNP case a 40 layers slab-size was used for all the surfaces since computational time was not an issue. E_{dimer} in eV/dimer and $E_{surface}$ in meV/Å^2

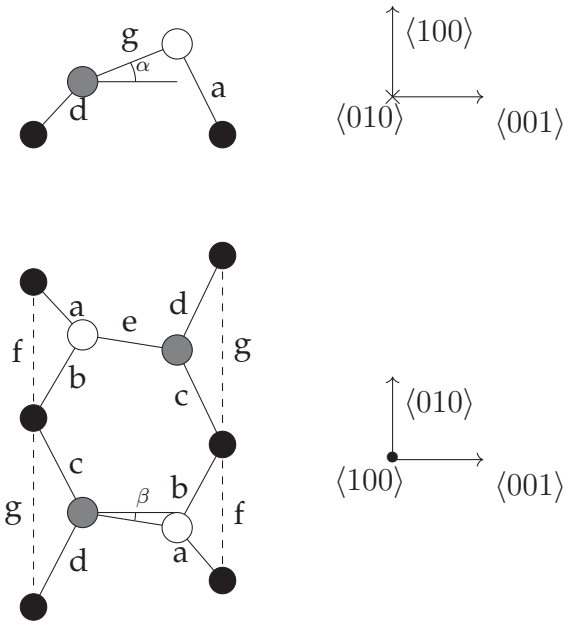
surface	PBESol	PBE	LDA[101]	NNP
ideal -> a-p(2x1)	1.70	1.61	1.9	.894
a-p(2x1) -> p(2x2)	0.066	0.069	0.048	0.089
p(2x2) -> c(4x2)	$6.68\ 10^{-3}$	$2.15\ 10^{-3}$	$3\ 10^{-3}$	$-2\ 10^{-3}$

Table 3: Difference in energy in meV/dimer between different functionals at relaxed reconstructions (different geometries). PBE and PBESol share the same accuracy.



conf			a-p(2x1)			p(2x2)
	PBEsol	PBE	NN	PBEsol	PBE	NN
a	2.38	2.38	2.39	2.38	2.40	2.39
b	2.42	2.42	2.43	2.38	2.39	2.39
c	2.35	2.36	2.36	2.38	2.39	2.39
d	2.38	2.39	2.40	2.38	2.40	2.39
e	2.32	2.32	2.33	2.33	2.34	2.34
f	2.38	2.39	2.40	2.38	2.40	2.40
g	2.29	2.31	2.31	2.34	2.35	2.35
α	18.81	16.3	18.34	19.52	19.04	19.44
h	3.84	3.88	3.87	4.08	4.09	4.02
i	3.84	3.88	3.87	3.60	3.64	3.62

Figure 15: Top figure: Schematic Top and side view of the a-p(2x1) and p(2x2) reconstructions. Bottom table: Geometric characterization of configurations reported in the top figure for PBESol and PBE. Distances are reported in Å and angles in degrees. A comparison with our final potential is also available. No major difference is noticeable.



conf			c(4x2)
	PBEsol	PBE	NN
a	2.38	2.34	2.39
b	2.38	2.34	2.40
c	2.33	2.34	2.34
d	2.33	2.34	2.34
e	2.34	2.35	2.35
f	3.58	3.62	3.60
g	4.10	4.10	4.12
α	19.90	19.00	19.90
β	0.86	0.60	0.86

Figure 16: Top figure:Schematic Top and side view of the c(4x2) reconstruction. In this case, differently in the p(2x2) case, more cells are presented in the top view. This has been done to show more clearly the presence the small angle β with respect to the $\langle 001 \rangle$ direction that was not present in other reconstructions. Bottom table: Geometric characterization of the configuration for the c(4x2) reconstruction. In this case, we characterized only the first two layers, and they are very similar. Our potential trained on PBE in the last column. Distances are reported in Å and angles in degrees.

create will not contain any of this information, these quantities have not been investigated.

Similarly to the other regions of the phase space, also for the surface performing dynamics closer to real PES will significantly reduce the number of iterations needed for an exhaustive sampling. To achieve this, after ruling out the Car-Parrinello approach because of its numerical cost, we explored different empirical potentials: Tersoff [53], Purja Pun [102], SW [41], MEAM with spline [103], and ReaxFF [104] potentials were tested. All these potentials failed to stabilize any of the experimentally observed reconstructions and all of them relaxed back to the s-p(2x1). Noticeably the ReaxFF parametrization also showed particularly poor results also for the lattice constant near the surface and it will not be considered.

To overcome this issue we decided to train a first preliminary NN potential, using the data provided in [80]. After checking the ability of the network to relax in the three stable geometries, we adopted it as a starting point for surface sampling. The first sampling consisted of a dynamic for each reconstruction and a few snapshots taken from high-temperature dynamics before evaporation started. The sample sizes were standardized to slabs containing 64 atoms to simplify the DFT computations.

It is interesting to observe that already with this early iteration of the potential, during the molecular dynamics runs an oscillation from the p(2x2) to the c(4x2) reconstruction takes place. This resembles what is observed experimentally at room temperature [105] and would be an interesting topic for further studies.

This procedure was repeated eight times, first with the preliminary potential and then with different iterations of the NN potential. Different slab thicknesses, surface sizes, and temperatures were sampled. A total of 3175 configurations (accounting for 203200 local environments) were collected.

2.3 Dataset analysis

In the first iteration with the Tersoff potential and our preliminary NNP for surfaces, about 2 thousand points were collected. The data were then used to train the first NN potential. This potential was then used to generate new data following procedures closely related to those just detailed. A selection of these new data was added to the global pool of data to train a new NNP in a self-consistent cycle that led to the final potential. For this procedure, PBESol was chosen as the DFT flavor in the first-principle calculations.

In each iteration, different properties were monitored. Among the most difficult to converge were the melting temperature and the stability of the surfaces,

especially the minute energy difference between the c(4x2) and the p(2x2) reconstructions. To converge this quantity, we added specific configurations to the dataset, notably larger surfaces (at the cost of sample thickness), and low-temperature simulation to sample the reconstructions' basins. For the melting temperature, we added a few solid/liquid interface samples to include realistic intermediate local environments.

During the evolution, the dataset variability was monitored qualitatively using a t-SNE bidimensional projection of the explored dataset as presented in Fig. 17. This showed how the configurations generated by Tersoff and the early generation NN potentials explored different geometries than those found self-consistently by the later generations of NN potentials.

About 16 thousand points were collected, for a total of 890 thousand local environments. After removing the data from [80], the structures that were no longer visited during the latest iterations of the self-consistent procedure, and applying a random uniform sub-sampling on the remaining dataset, we were left with about 5 thousand rapresentative data points, and 250 thousand local environments. An exact summary is available in Table 4. A standard splitting of the dataset in a training part (containing 80% of the data) and a validation part (containing 20% of the data) was performed, ensuring that the same fraction of training and validation data were present in each subset to allow for an homogeneous testing in different regions of the phase space.

Table 4: Data distribution before and after pruning. Pruned points are available in both DFT flavors, PBE and PBESol. Before pruning only PBESol is available.

	All atomic configurations	All environments	selected configurations	Selected environments
amorphous	4485	287040	1077	68928
bulk	1162	37184	504	16128
interfaces	98	12544	42	5376
liquid	4612	194248	2000	83904
surfaces	3175	203200	1376	88064
dataset [80]	2312	161048	0	0
total	15844	895264	4999	262400

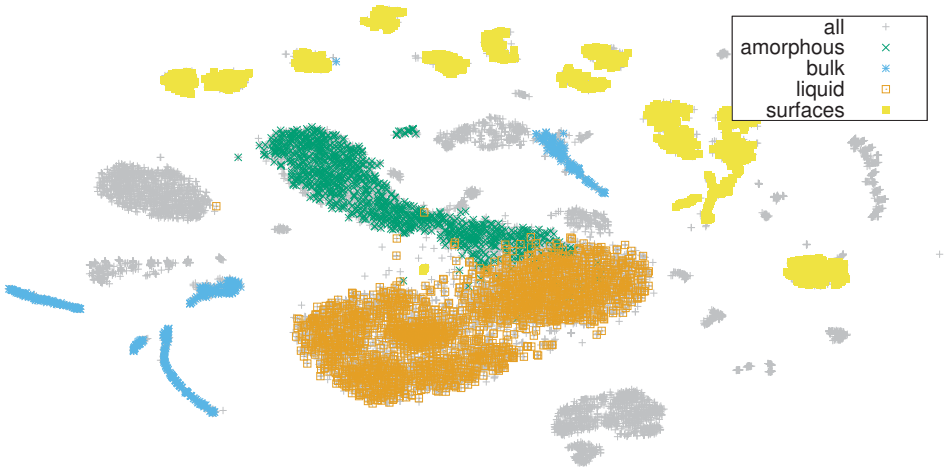


Figure 17: The figure shows a t-sne bi-dimensional map of the atomic configurations in the dataset using the structural fingerprint defined in Valle et al. [106] and the cosine distance. Other distances were tested. Both Cartesian and Wasserstein distances showed no appreciable difference. This picture has shown also a strong resilience towards t-sne parameters. In gray, all the available structures are presented while colored points refer to the final dataset. Different colors highlight the macro-regions of the phase space. The gray macroscopic islands are points explored by early potential iterations and data collected with Tersoff and discarded in the final self-consistent potential iteration (e.g. the gray island in the top left is the first iteration of a-Si structures; the small gray islands are from the environments in the [80] dataset). The yellow points correspond to surface configurations ordered in an arch from right to left for increasing temperature. For each temperature two clusters are visible: one for the c(4x2)/p(2x2) points and one for the a-p(2x1) points. The blue points in the center are configurations describing vibrations in c-Si and on the left are high-pressure structures. Configurations for a-Si and l-Si are the two large green and orange islands in the center of the figure. The contact points between the two islands are configurations containing l-Si/a-Si interfaces.

2.4 Parametrization of local environment descriptors

As explained in section 1.2.2, to train a NN potential it is first necessary to encode in a descriptor the local environment of each atom in the dataset. Descriptors, described in section 1.2.2, are vectors of a fixed length that contain the information needed by the NN to predict the local energy. Larger descriptors increase the computational time while providing more details on the local environment. On the other hand, smaller descriptors are faster but less accurate.

Since we are concerned with a single atomic type, all the species indexes in equations (1.43) and (1.44) are dropped and we selected the following hyper-parameters to achieve a reasonable tradeoff between size and expressivity: for the radial part η = 16 Å^{-2}, r_c = 4.6 Å, and thirty two values of R_m equally spaced between 0.5 Å and r_c. For the angular part instead: η =6 Å, $r_c = 4.6$ Å^{-2}, $\xi = 50$, with six values of R_m equally spaced between 1.5 Å and r_c and twelve values of Θ_n in the 0-π range. The short-range cutoffs were chosen to take into account the short-range interactions and are mostly populated by high pressure structures. The long-range cutoffs instead were chosen similar to the one in other empirical potentials such as SW and Tersoff, extending beyond the second nearest-neighbour distance in c-Si. These hyper-parameter set results in local environment feature vectors with 104 components, *i.e.* 32 in the radial part and 72 in the angular one.

To understand which are the most important descriptor features appearing in our training dataset we plotted their distribution. This data is presented as box-plots in Fig. 18 for the radial part and Fig. 19 for the angular part, where the outliers are automatically identified using the Tukey definition as detailed in [107].

Focusing on those outliers, we can observe that they are distributed among configurations that contain liquid-solid interfaces and high-pressure solids. The sampling over each one of these regions could have been increased to provide a better description, but we reached satisfactory physical accuracy without having to do so. High-pressure configurations are seldom explored in our simulations. The solid-liquid interface data is present to stabilize the melting temperature estimate which, as we will see, is already reliably determined.

Unlike similar previous studies [64, 76, 61], in this work we employ a $\tanh$ activation function for the hidden layers of the NN, instead of a Gaussian one. Such activation was chosen because it has the advantage of a linear behavior near zero and a stronger theoretical background to understand and optimize the training phase. On the other hand, this choice requires a preconditioning of the input that must be transformed from the positive defined descriptor to a

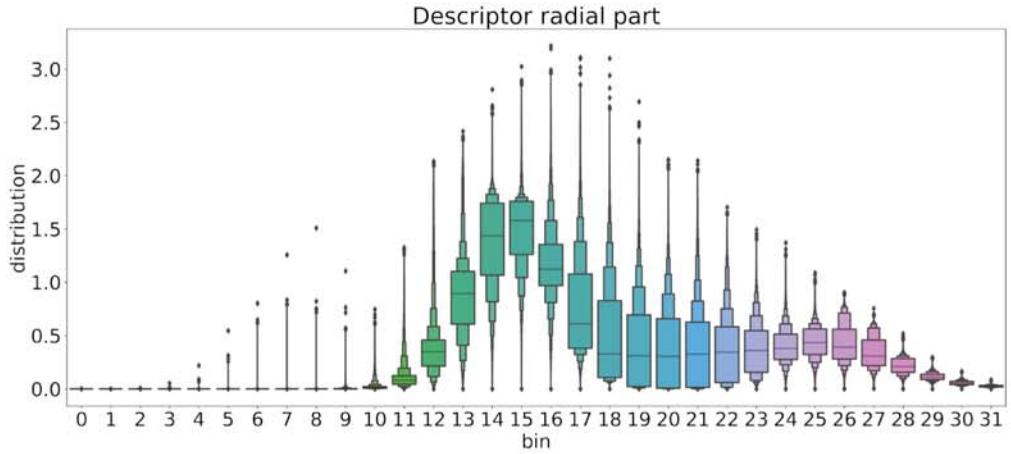


Figure 18: Distribution of the values assumed by each bin of the radial descriptor in the training dataset. Blocks are percentiles, 14 boxes are present, the larger one contains 50% of the data values assumed by the bin, the next ones 75% and 87.5% respectively. The line in the middle of the main block is the average bin value. Points at the margin of the distributions are outliers identified using the Tukey definition. Color is present to help the eye. The shape of the radial part descriptor closely resembles the radial distribution function, with a first peak approximately at bin 14 (2.4 Å). The tail does not converge to 1 because of the dumping function (see Eq. 1.42).

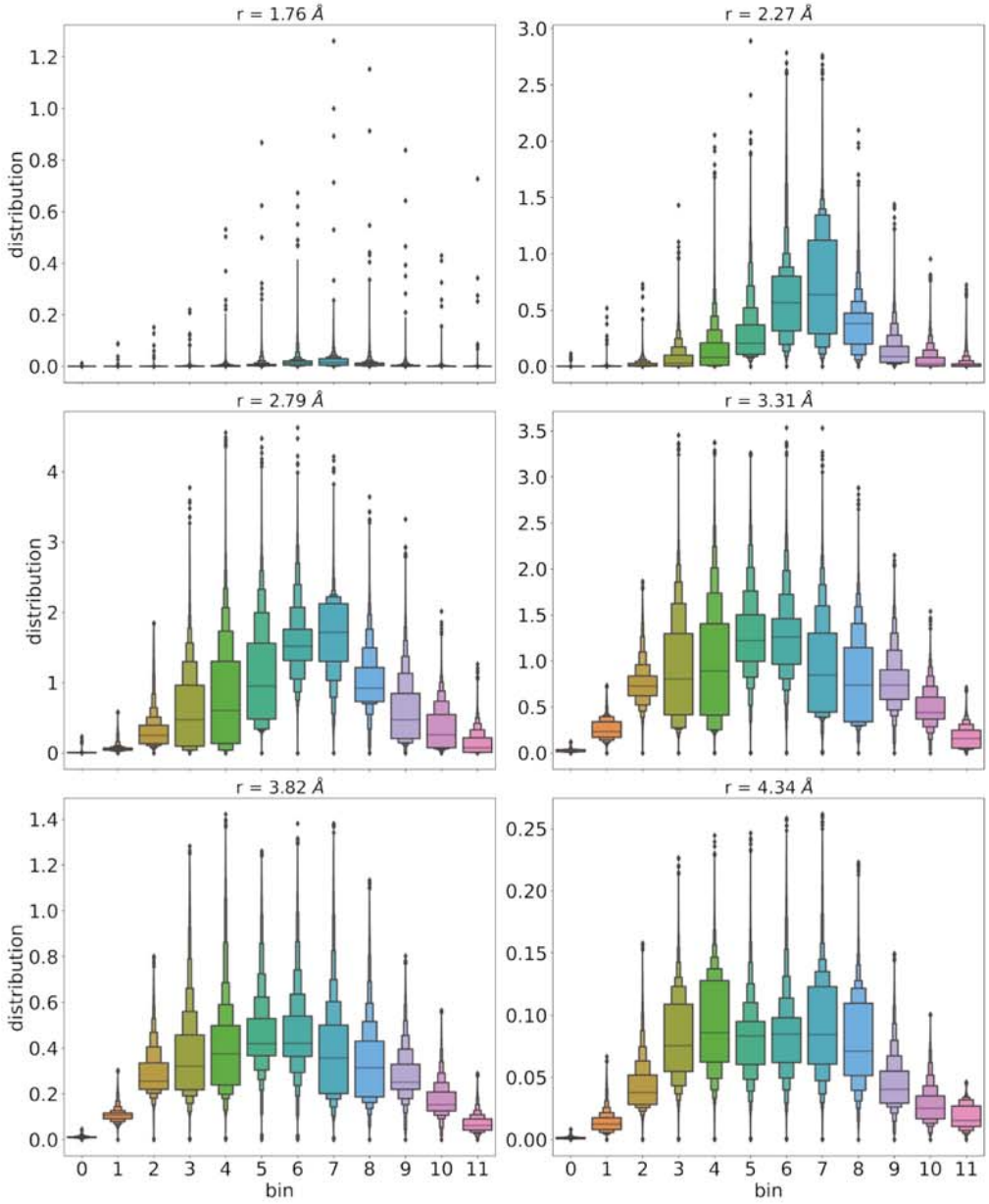


Figure 19: Angular average descriptor, same representation as in Figure 18. Each panel represents a probed radius. Bins inside each panel sample the angle between 0 and π; all the angles are homogeneously sampled from the dataset.

normalized one with zero mean and unit standard deviation. For this normalization step, two possibilities have been identified: *i)* a simple bin re-scaling or *ii)* adoption of principal components analysis (PCA). The re-scaling is obtained by subtracting from each bin its mean value and dividing by the standard deviation over the training dataset. A problem that arises with the chosen descriptor parametrization and dataset is that a few bins have little to no standard deviation, forcing the introduction of a small regularization constant to avoid extreme re-scaling. We found PCA a more systematic approach. From our training dataset, the directions of maximum variance of the descriptor (or principal components) are computed. As a result, it is possible to define projectors of the old descriptor coordinates on this new base. The variance in each new direction can be normalized to one.

We found that 8 out of the 104 components are linearly dependent on the others, hence having no variance at all. This means that they can be discarded with no information loss limiting the principal components under consideration to 96.

Furthermore, directions of very low variance can also be discarded to avoid division by very small numbers in the normalization step. A possible metric to understand how much information is lost in the PCA step is to measure the "total explained variance". In our case, we discarded the principal components that explain less than 10^{-5} % of the total variance, thus reducing the number of bins from 96 to 86. We do not expect this small reduction to have any measurable impact on the training results.

2.5 Neural network training and validation

Different rounds of training were performed to explore the landscape of NN architecture hyper-parameters. For every NN architecture, at least two trainings were performed to monitor differences due to the stochastic nature of the process. Among all these trainings, we analyze here the ones that leads to the potential that will be used in the rest of this thesis. This NNP was trained on the 5000 points dataset described earlier, with PBE-DFT flavor. In this case, the network was structured as follows: two hidden layers (hence L=3) of 256 and 128 nodes, respectively, with hyperbolic tangent activation functions and a single output node with linear activation. The PCA input normalization step was implemented adding a linear, non trainable, layer to the network. We are left with a model with 59905 degrees of freedom.

The training was performed with a manually tuned schedule for the learning rate parameter. The learning rate was changed from 10^{-3} to 10^{-5} in three steps. The batch size was composed of 150 elements. Regularization terms on

both L1 and L2 norms on weights and biases were included in the cost function with a multiplicative factor of 10^{-8} , while energy and force loss terms were both weighted by a unit factor. Training curves and detailed information on the schedule are presented in Fig. 20. Validation curves are presented in Fig. 21.

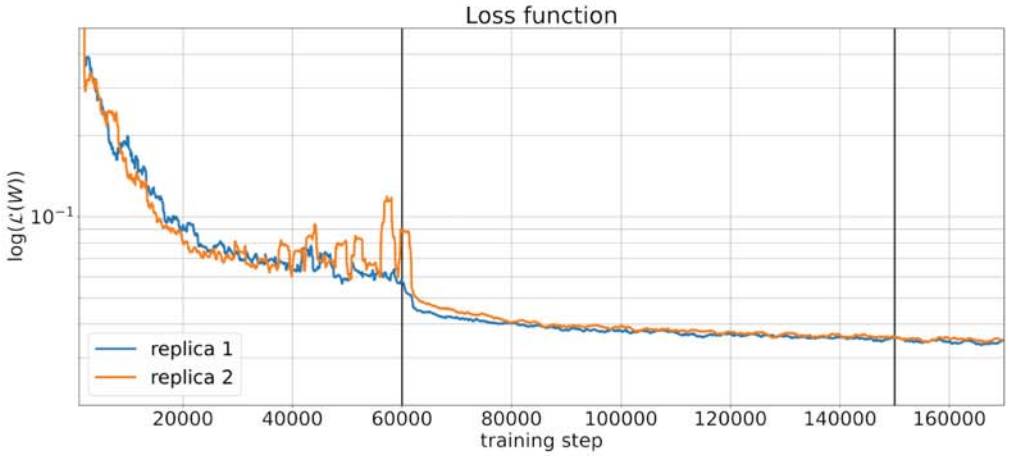


Figure 20: Loss function curves during training for the two replicas, in this case the learning rate was manually tuned. At the beginning a large value of 10^{-3} was used for the learning rate and this results in a rapid decrease of the cost function while allowing for larger oscillations when outliers are present in the batch. As the learning rate is decreased to 10^{-4} at 50K steps also the oscillations are dumped and the system freezes near a local minimum. At 150K steps the learning rate is further reduced by an order of magnitude. The training was stopped when no more improvement was seen on the validation side, Fig. 21.During the training 90 to 95% of the contribution to the cost function was coming from the force term and the remaining 5 $\sim$10% from the energy term. The normalization contribution is negligible.

The overall accuracy on the validation set for the NN potential corresponds to a root mean square error (RMSE) of 5.6 meV/atom for the prediction of the energy and a RMSE of 153 meV/Å per force component. This accuracy is comparable to the one obtained by others with similar approaches [108]. To give a more detailed picture of the validation accuracy, we computed the error over the different sampled macro-regions of the phase space described in the previous section, obtaining 5.8 meV - 170 meV/Å for the amorphous phase, 0.9 meV - 66 meV/Å for the crystalline bulk, 7.5 meV - 196 meV/Å for the liquid phase and 3.4 meV - 103 meV/Å for the surfaces. The distributions of these errors are reported in Fig. 22, showing that the accuracy is rather uniform across the phase space. Fig. 23 shows the force error distribution across the different parts

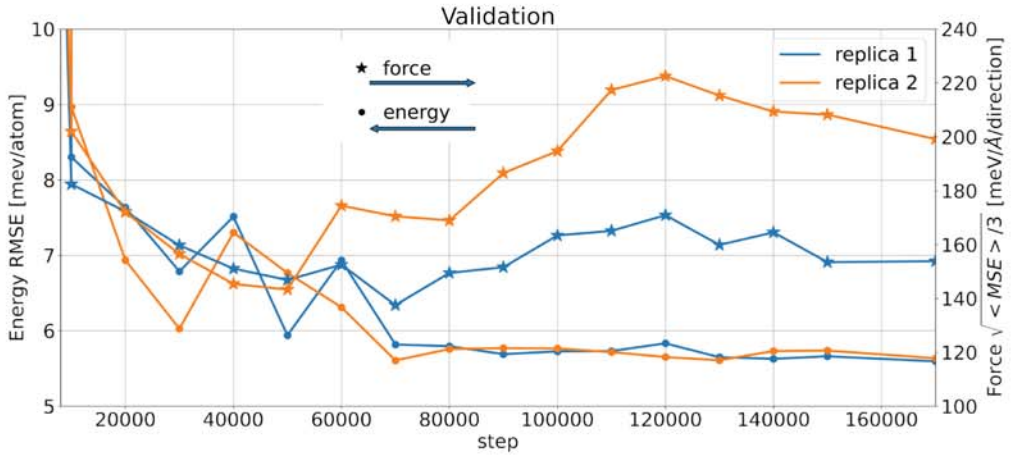


Figure 21: Validation error of selected points along the trajectory. The curves follow the typical shape of a training. Initially a fast increase in accuracy is experienced followed by a plateau where only small increases are measured. In this case however there is an interesting deviation in the validation of the forces (star points that refer to the right axis) where an apparent over-fitting is present. A detailed analysis of the force distribution showed that this is not the case and instead we are only looking at a change in the distribution of the ouliers.

of the dataset: observe that none of these distributions is a Gaussian. The presence of heavy tails in the error distributions indicates that the RMSE may not be very relevant in the comparison of different trainings as it is very sensitive to the presence of outliers.

A second potential was similarly trained on the same 5000 configuration dataset, evaluated with PBESol functional. Also in this case, different RMSE can be computed for energies and forces depending on the subset of points: 6.3 meV - 165 meV/Å for the amorphous phase, 1.0 meV - 61 meV/Å for the crystalline bulk, 8.4 meV - 183 meV/Å for the liquid phase, and 6.9 meV - 256 meV/Å for the surfaces. This shows that the accuracy of the potentials developed for the two functional variants is indeed very similar.

2.6 Testing the potential on additional physical properties

In order to get a better feeling of the quantitative behavior of our PBE NNP associated to the numerical errors reported above, we tested it by investigating a few

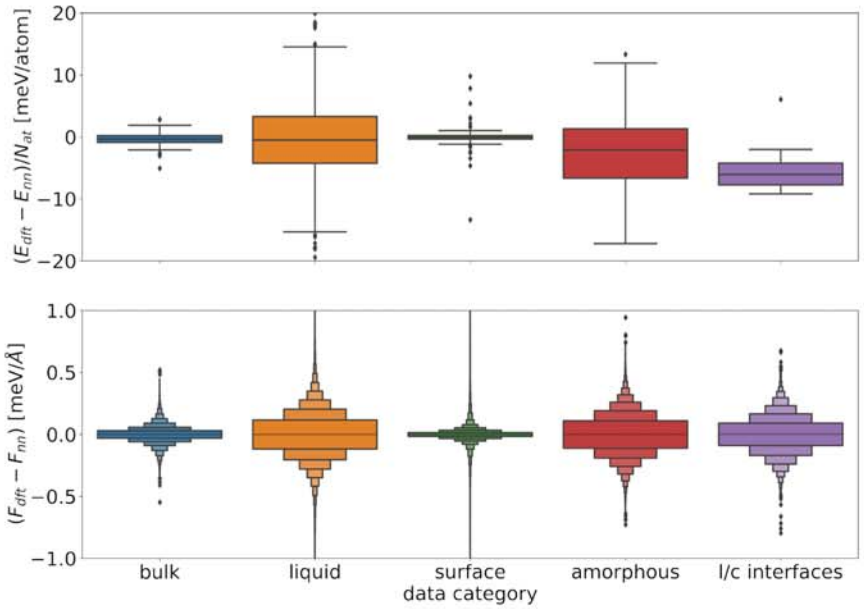


Figure 22: Distributions of the errors per dataset in Energies (top panel), box corresponds to the first quartile, and forces (bottom panel), different boxes represent different 14 different percentiles are presented. It can be seen how phases with the highest variability (liquid and amorphous) show the highest errors.

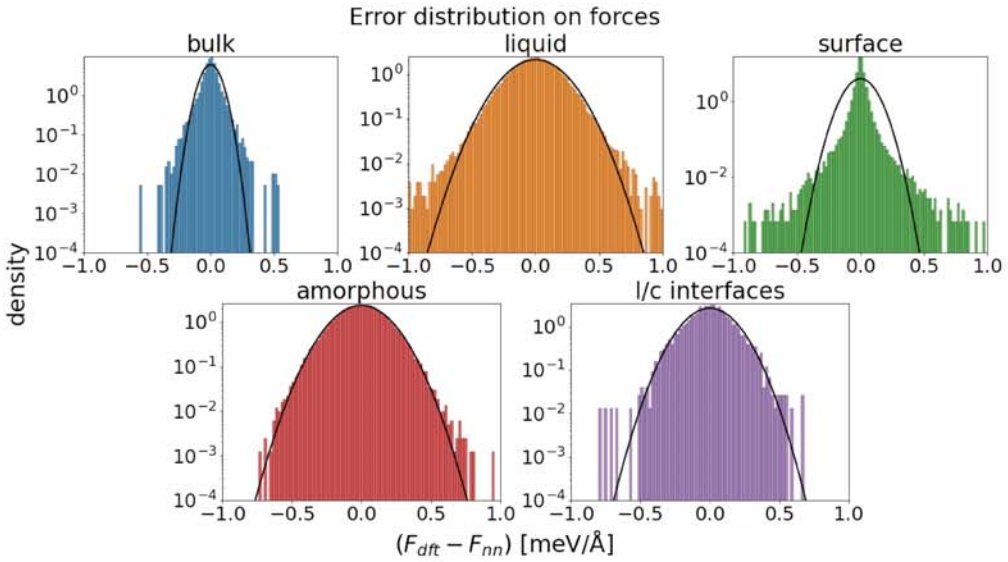


Figure 23: Error distribution of force components over the validation set. y axis is in logarithmic scale and black lines correspond to Gaussians with mean and standard deviation computed on the presented data. If the data were to be Gaussianly distributed they should stay below the black curves. Different regions present very different accuracy in the forces over the validation dataset. Bulk and surfaces have the highest accuracy thanks to their simple structures. In the case of surface (panel with the green distribution), atoms that are close to the vacuum often are badly predicted and appears in the large tails far form the Gaussian. Liquid (orange) and amorphous (red) instead are complex to describe and follow large Gaussian with large tails. Strictly speaking none of these distributions has a Gaussian shape with a negligible number of outlayers.

physical properties that can be compared with experimental and/or DFT reference data. A comprehensive test has been carried out for the Si-diamond structure, where we checked and computed the bulk modulus, obtaining a value of 87.2 GPa close to the experimental value of 97.6 GPa [109]. Table 5 reports a full comparison of bulk elastic constants. Full phonon dispersions were also investigated obtaining results very close to the DFT reference results (Fig. 24). The acoustic dispersion around the X and L points shows some oscillations which suggest that the long-range interaction along the bond-chains in the [110] directions, responsible for the flatness of the dispersion in the DFT reference, is not perfectly captured by our NNP. We were able to improve the latter result by increasing the weighting factor of the force term in the loss function (1.48) by two orders of magnitude. The ability to approach the DFT long range dispersion curve with high precision suggests that this piece of information is present in the training dataset, but in order to obtain such an improvement in the dispersion we would have to compromise on the energy accuracy. The optical modes are overall slightly softer than the reference DFT results and it is possible that this softening is due to the presence in the training set of a number of metallic or small-gap configurations where the partial occupation of the conduction bands due to the use of Marzari-Vanderbilt smearing makes the high frequency vibrations less stiff.

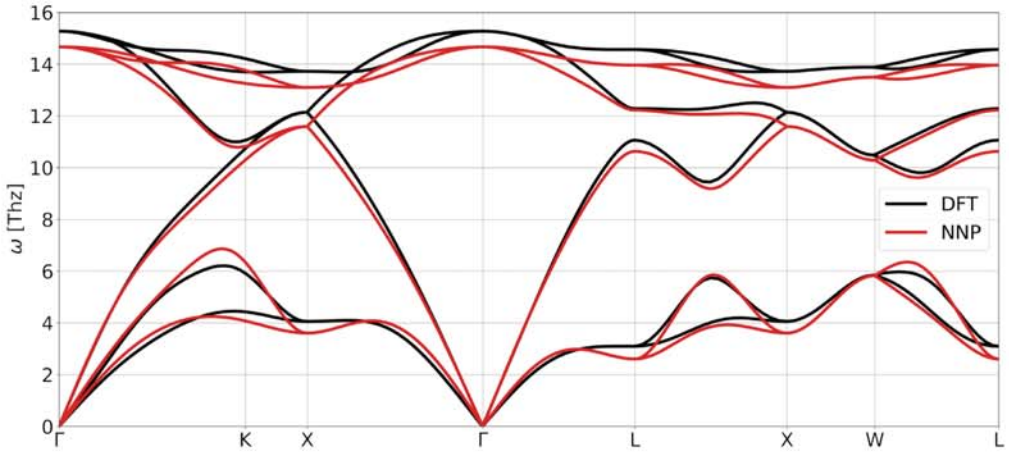


Figure 24: Diamond phonons of silicon. DFT prediction with PBE (black line) and NNP (red line).

The dynamics of SPE is critically dependent on temperature, and the commonly adopted empirical potentials either fail in describing the thermal stability of the different phases or, if they succeed, fail in other aspects. For example, the standard SW potential has a melting temperature of 1691 ± 20 K [97], very close

	NN		DFT [110]			Experiment
	PBE	PBESol	PBE	PBESol	LDA	(4.2 K)[111]
C_{11}	150.1	156.2	153	156	161	167.5
C_{12}	56.4	61.9	57	62	65	64.9
C_{44}	72.0	73.9	74	74	76	80.02
bulk mod.	87.2	93.3	89	94	97	97.6 [109]

Table 5: Elastic constants in GPa for Silicon in diamond phase at 0K calculated in this work for different NN potential, and compared to theoretical and experimental results from the literature.

to the experimental one, but generates amorphous configurations of very poor quality. This is due to the limited strength of the three-body term that is not able to force enough tetrahedrally coordinated environments. The Tersoff potential, on the other hand, is believed to generate realistic amorphous geometries, but it predicts a very high melting temperature of 2547 ± 22 K [99]. Errors in the predicted melting temperature are likely to have an impact on defect formation and diffusion dynamics, relevant for the SPE. It is therefore interesting to study the bulk melting temperature associated to our NN potential and we determined it from numerical simulation of the coexistence of the crystalline and liquid phases [112, 113].

Disappointingly, the NN potential developed from data computed with the PBE-sol exchange-correlation functional, originally adopted on the basis of its reported better description of solid phases [48], resulted in a bulk melting temperature as low as 1194 ± 29 K. The unsuitability of the PBE-sol functional to describe the melting behavior of materials has later been independently confirmed by simulations based on GAP ML potentials [114]. For this reason, we reverted our exchange-correlation choice to the original PBE functional for which a fully ab-initio estimate of the bulk melting temperature of 1540 ± 50 K [85] is available, still slightly off from the experimental value of 1685 ± 2 K but reasonably close. With the NN potential trained on our dataset with PBE energies and forces a melting temperature of 1468 ± 11 K, in line with the above literature value, was obtained showing that the main source of discrepancy with the experiment is not in the numerical errors associated to the ML potential accuracy or in the limited size of the cells used in the DFT calculations [85] but in the quality of the adopted DFT functional form. A more detailed description of the method used to compute the melting temperature, and its dependence on simulation cell size is given in Appendix A.

Having obtained a reasonable description of the liquid/solid transition, we proceeded to study the amorphous phase. We followed the procedure described

in Ref. [115, 35] where the liquid samples were quenched with a 10K/ps temperature ramp. Amorphous samples of very good quality were thus obtained and in Fig. 25 the structure factor from the largest simulated cell (4096 atoms) is compared with experimental results.

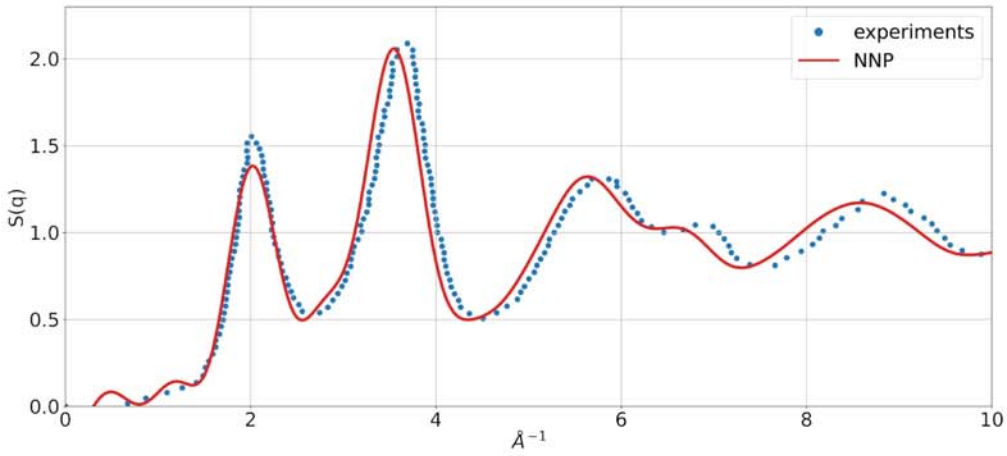


Figure 25: Amorphous silicon structure factor $S(q)$. Dots are experimental points [116], red line from PBE NNP.

We examined in detail the (001) surfaces that play a very important role in microelectronic devices [101]. As mentioned previously, several reconstructions have been proposed in the literature [101]: the ideal unreconstructed surface with p(1x1) surface periodicity, the symmetric-dimer reconstruction with p(2x1) periodicity, the asymmetric-dimer p(2x1) reconstruction, and the larger reconstructions with p(2x2) and c(4x2) periodicity. Of these five structures, only the last three have proven to be stable with sufficiently accurate calculations, while the ideal p(1x1) and the symmetric p(2x1) reconstructions are only meta-stable. The energy ordering of the various structures according to different methods can be seen in Fig. 26, and the corresponding values have been reported earlier in Table 3. All reconstructions are predicted with differences in energy in good agreement with DFT reference values from the literature. In the the second to last column of Fig. 26 the geometry of the dimers is taken from the DFT calculations, scaled so as to reproduce the lattice constant of the NNP. The last column is obtained by relaxing the structures with the NNP. It shows that the symmetric dimer reconstruction is not a local minimum and spontaneously relax in the asymmetric dimer structure as expected from accurate DFT calculations. It turns out the predicted energy ordering between the c(4x2) and the p(2x2) reconstructions obtained with the NNP is opposite to the DFT result, corresponding to very close energies for the two configurations (within 2-3 meV per dimer)

in both schemes.

The atomic geometries of the different reconstructions were also compared and the NNP calculations were able to correctly predict the dimer tilting angle of 18° for the p(2x1) asymmetric phase as well as the 19° angle in the p(2x2) and c(4x2) ones, in agreement with DFT results. More details on the geometrical characterization are reported earlier in Figures 15 and 16.

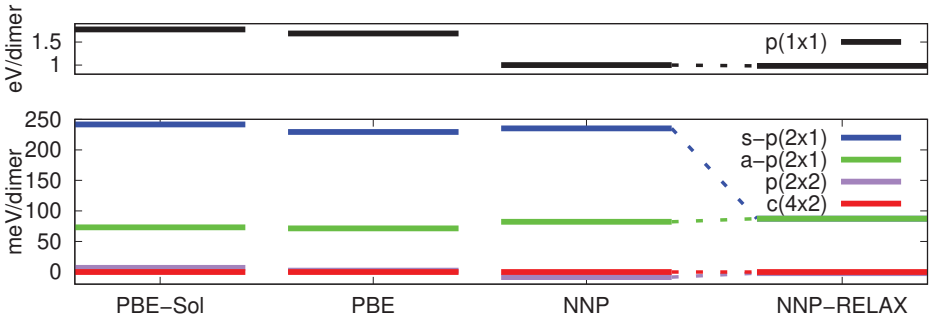


Figure 26: Surface energies. From left to right, PBE-Sol energies, PBE energies, NN energies for PBE geometries, NN energies after relaxation. The c(4x2) structure has always been taken as zero reference. DFT geometries have been obtained by relaxation with the BFGS algorithm, with the default 25 meV/Å as a force threshold. Minimizations with NNPs were performed with the same threshold. Both DFT calculations predicted the s-p(2x1) structure as stable only with a low k-point grid, the reported energy is a successive SCF calculation with the high accuracy grid. In the last column the s p(2x1) geometry has relaxed to the a p(2x1) structure. The p(2x2) and the c(4x2) are incorrectly ordered but their difference in energy is below our error threshold.

Chapter 3

Solid Phase Epitaxy

The major issue of the current on-lattice KMC approaches that simulate SPE is their lack in microscopical knowledge of the key events 1.1.2. The end goal of all the developed ML potentials should always finally be their ability to predict new physics and extend our knowledge. Hence, in this chapter, we are going to use the predictive power of the NN potential previously created to investigate SPE from an atomistic prospective, computing its characteristic barrier and isolating the time sequence of the simulations that are most likely to contain the key events for further analysis. This will be done using MDs, that have been proven to be able to describe with a reasonable degree of accuracy structural and dynamical properties of simple and complex materials. Such results will pave the way for the search for new microscopic mechanisms for the KMC, and possibly validate or disprove the results obtained with well-known empirical potentials.

3.1 The l-Si/a-Si phase transition

In the section 1.1 we reported that experimentally the SPE speed follows an Arrheniuhs' law, Eq. 1.1. This exponential relation between reconstruction velocity and temperature means that even small changes in the simulated temperature result in large changes in the computational time necessary to observe a reconstruction. It is then mandatory to find a temperature upper bound at which the SPE can be observed, so that we can simulate the fastest possible phenomenon, resulting in a reasonable MD simulation time.

Two phase transitions are involved in this process: the c-Si/a-Si phase transition and the a-Si/l-Si phase transition. The first one is the phase transition we are interested in, the second phase transition is a problem we have to deal with when investigating high-temperature regimes.

We lack SPE experiments on pure silicon samples so we don't know precisely at which temperature SPE ceases to happen, but Olson *et al.* [4] performed several high temperature experiments in samples with arsenic. After excluding the a-Si/l-Si phase transition in the high-temperature range they measured a barrier very similar to the pure silicon one. Further analysis showed that the presence of low As concentrations does not induce major differences with respect to pure samples. Such result offers solid empirical evidence that the microscopic mechanisms, whatever they are, should not change significantly up to the melting point. With this in mind, we can assume that an extrapolation to pure silicon samples from experimentally available data at high temperatures should hold.

At this point, the high-temperature regime of interest lies between the melting temperature of c-Si (1685K) and the melting temperature of a-Si (1260K-1450K depending on its level of relaxation [117, 118] or section 1.1). The melting temperature of c-Si for the NNP, as explained in section 2 is lower than the experimental one by a factor of about 1.14. Experimental values estimated for the melting temperature of a-Si can not be used directly in our simulations. The approach often adopted in the literature to address this issue is to linearly scale these temperatures with the aforementioned factor. This approach fixes our temperature boundaries with a maximum theoretical temperature that can be adopted in simulations of 1400K, and a melting temperature of a-Si of approximately 1100-1240 K.

For the sake of simplicity, all reported temperatures in the following text will refer to these scaled simulated temperatures unless otherwise specified.

The 1100 to 1240 K temperature range is only a qualitative estimate, and there is no reason to trust it blindly. Quantitative proof of the presence of an l-Si/c-Si phase transition and more accurate values for its temperature can be given by looking for the presence of a van der Walls loop (vdW) in isothermal lines in the P-V space, below the critical temperature. In realistic systems, isotherms in P-V space near a first-order phase transition display two distinct behaviors separated by the isotherm at the critical temperature. Above the critical temperature, P-V isotherms have no stationary points, and no phase transition can be defined. At the critical temperature, the P-V isotherm displays a stationary point that separates the pressure/volume ranges assigned to the two phases. Phase coexistence is possible at the stationary point. Below the critical temperature isotherms are characterized by a intermediate range of volumes accessible to the system in equilibrium only as a mixture of the two phases, coexisting at the same pressure and different volumes. In NVT numerical simulations performed in simulation cells of limited size, at these intermediate volumes the sample can not undergo the phase separation process and is trapped

in either an undercooled or an overheated homegeneous state of one or the other of the two phases. The result is that below the critical temperature, what was supposed to be a straight line connecting the two phases becomes an oscillation. From the oscillation, it is possible to recover the equilibrium pressure and the transition volumes of the two phases by applying the Maxwell construction. In our case, the phase transition occurs around zero pressure so, instead of using the Maxwell construction to obtain the pressure from a given temperature, we will use it to look for a range of temperatures that satisfy it at vanishing pressure.

To investigate the vdW loop we proceeded as follows. First, a set of liquid configurations with different densities and a size that does not allow for the coexistence of the two phases was generated. The density range goes from 10% less dense than a-Si to 10% denser than l-Si. From each configuration, different NVT relaxations with initial random velocities compatible with a temperature of 2000K were performed. During the relaxations, a 10K/ps cooling ramp was applied, and each relaxation ended at the temperature of a target isotherm. This procedure leads to each density having as many different configurations as target isotherms. Then, for each isotherm the pressure at a given density can be estimated by running long NVT dynamics. The main limitation of this procedure is that if the initial configuration is trapped in a local minimum the sampling for that isotherm will be very limited. To overcome this we applied the replica exchange method [119] that regularly attempts to swap configurations between nearby isotherms.

For the vdW study we used samples composed of 216 atom cells and 15 isotherms in the range between 880 and 1280K were targeted. The NVT sampling with the replica exchange method was run for 40 ps. Results are presented in fig. 27. Assuming that the Maxwell equal-area construction must hold at zero pressure, we can estimate a temperature on the order of $1100 \approx 1050$ as the maximum temperature at which the generated amorphous can coexists with the liquid phase. This result is approximately in line with our initial estimate. Furthermore, its density is approximately between 18.2 and 20.7 Å^3/atom as expected.

An interesting result, never considered in SPE literature, regarding the a-Si is that there are numerical [14, 15, 120] as well as experimental [121] observations that indicate the existence of an intermediate phase between l-Si and a-Si, that shares geometrical structure and electrical characteristics with a-Si while diffusing as a liquid. This state has been defined as low-density liquid (LDL) in contrast with the usual l-Si phase, called high-density liquid (HDL). The existence of this state is not unchallenged [122] but, as we will see, some of our high-temperature simulations have properties that are close to this state.

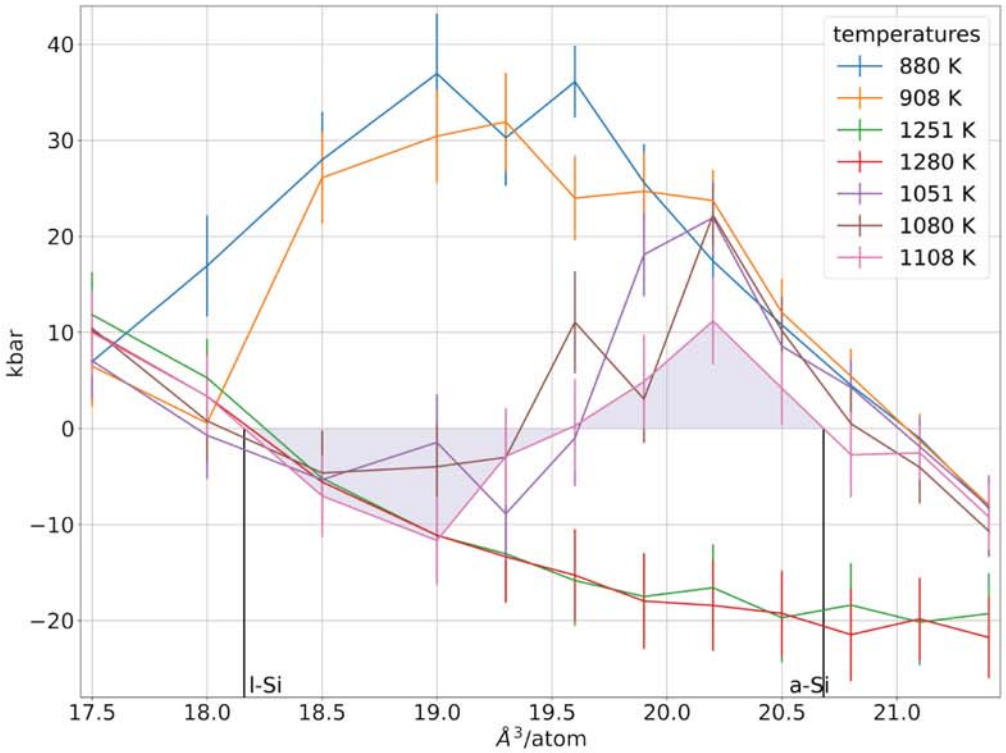


Figure 27: Isotherms in the pressure-volume plane near the l-Si/a-Si phase transition. 15 temperatures were simulated in the 880-1280 K range but only 7 are reported for clarity. A possible Maxwell construction is highlighted and the liquid , In the figure, a positive P(V) slope indicates that the system is in a thermodynamically unstable state, and, if larger samples were simulated, it would separate in the two phases. The three temperatures presented in the middle region (1051, 1080, and 1108 K) are those that show a clear vdW loop at zero pressure, and give a solid upper and lower bounds for the hypothetical T_m value. From the figure, the critical temperature, which is the one corresponding to the lowest isotherm that does not show a stationary point, can also be estimated in the 1220-1250K range, a value very close to other DFT theoretical results [15] performed with GGA functionals.

3.2 a-Si characterization

During the generation of all our amorphous structures, we quenched the samples from the liquid by decreasing the temperature with a linear ramp. The linear ramp was chosen with a 10K/ps rate. This value is similar to fast quenching experiments but is still several orders of magnitude faster than the quenching rates used in SPE experiments. Therefore, it is important to understand the quality of the obtained a-Si samples.

When talking about comparisons, the properties accessible with numerical simulations can be roughly divided into two categories: those that require the knowledge of the charge distribution and of the wave function, and those that do not. Electronic properties belong to the first category and are notoriously expensive to compute, their knowledge would allow a precise characterization of point defects. This can be achieved, for example, by smartly combining the electron localization function and the charge distribution to classify all the types of bonds [123] or by associating them to localized states within the bandgap [124]. The second category includes instead all the geometrical and vibrational properties that can be investigated by limiting the domain to classical mechanics. Among the geometrical properties, we find the density, the structure factor, the angular/dihedral distributions, the ring counting distribution [125, 126], the excess energy w.r.t. c-Si, and the vibrational density of states (v-DOS). Since the nature of our potential is classical we will restrict our investigation to this second category.

Volume and structure factor metrics have the advantage of being directly comparable with experiments but are known to be very little informative since different local structures can easily give rise to similar results [127]. a-Si excess energy w.r.t. c-Si is also a property that can be directly compared to calorimetric experiments [128]. Since such extra energy is stored in bond angle distortions and point defects it should be a good overall indication of the a-Si quality. Angles and other geometric quantities, on the other hand, are not directly measured but are accessible to atomistic simulations and allow comparison between atomic configurations reported in the literature. v-DOS is sensitive to the local structure and can be accessed at a reasonable cost from the simulation.

To numerically create a-Si samples a variety of techniques have been proposed and combined with all sorts of potentials [129]. These methods produced the most disparate results, and comparing our a-Si with all of them is beyond the scope of this work. However, defining a baseline to compare with is necessary. We opted for two comparisons, one with structures obtained from classical approaches and one with a DFT "optimal" structure proposed in [86].

For the comparison with classical approaches, we focus on SPE: Krzeminski

et al. [34] started from a WWW a-Si, and relaxed it with MDs using different potentials. After extensive tests, they concluded that Tersoff [38] potential, thanks to its correct density prediction and correct SPE barriers, is the best choice for this application, in spite of the large overestimation of Si melting temperature. Lampin *et al.* [35] expanded the previous study showing that results can be improved if instead of using the WWW approach, the a-Si was prepared through a melting quenching technique. Hence we adopted the latter approach to generate all the structures used for comparison.

In this characterization, we are not interested in creating an excellent amorphous but rather in understanding the properties of the average amorphous we deal with during SPE simulations. This is the reason why we followed an approach that produced many cheap samples instead of a few large cells. To create samples that share the same thermal history as the a-Si used in SPE, the sample creation for both the neural network potential and the Tersoff potential was performed following the melting-quenching technique with a 10K/ps ramp. These ramps were started from configurations extracted by initial liquid dynamics at temperatures adapted to the potential (3500K for Tersoff potential and 2000K for the neural network potential); the liquid presence was assessed by qualitatively analyzing the radial distribution function. For the Tersoff potential the liquid dynamics and the first part of the annealing (at 2500K) were NVT MD runs constrained to have the correct l-Si density. For both potentials, the procedure was repeated with cells of 216 and 1728 atoms to probe for finite-size effects. Ten samples from each potential and cell size were collected, providing a dataset of 40 amorphous configurations.

For the comparison with DFT structures instead, we considered the best configuration generated by Pedersen *et al.* [86] in a recent work. This structure was generated with a novel procedure that aims at minimizing the excess energy of the amorphous sample w.r.t. c-Si. Starting from a structure generated with the melting-quenching technique (using the Tersoff [38] potential) and then relaxed with PBE-DFT, for each atom in the cell the corresponding relaxed removal energy was computed. The removal energy is defined, similarly to the vacancy formation energy as how much cohesive energy is needed to remove the atom from a particular a-Si realization. Among the atoms, those with negative removal energy (that is atoms not strongly bound to their neighbors whose removal lowers the sample energy [87]) were of particular interest. The atom with the lowest (more negative) removal energy was indeed removed, the system relaxed, and the procedure repeated until no more atoms with negative removal energy were present. The best structure obtained (with the lowest excess energy w.r.t. c-Si) is used as a reference and referred to as y, following the notation used in [86]. The y structure relaxed with our NNP will also be the

object of discussion and will be labeled as y_{NN}.

3.2.1 Geometrical characterization

The simplest quantities that can be used to compare different samples generated with numerical procedures are geometric quantities such as bond angles, dihedral angles, the ring counting distribution, the per atom volume, and the radial distribution function. The radial distribution function was already investigated indirectly during the potential creation in chapter 2 (see Fig. 25), where we showed that the structure factor, that is the Fourier transform of the radial distribution function, obtained with our NNP is in good agreement with experiments. Therefore we are not showing it here again. Bond angle and dihedral angle distributions are reported in Fig. 28, and a summary of the angle distribution properties are reported in Table 6. Fig. 28 shows that amorphous samples obtained from Tersoff potential have noticeable features in the angular distribution around 60 and 80 degrees. Those featutres are not present in the samples generated with the NNP. This difference easily explains the standard deviation differences reported in Table 6 between the two approaches. However, the origin of these two peaks is still unclear. An investigation over the molecular database created by Allen *et al.* [130] showed that both 60 and 80 degrees bond angles are reasonable if one considers the possible hybridizations of silicon atoms [131]. This leaves open the question whether such a result is an idiosyncrasy of Tersoff potential or is a real feature of a-Si not described by PBE and consequently by any MLP trained of DFT-PBE.

Fig. 28 and Table 6 show also that statistically no major difference in angles statistics is due to the sample size, and all the features seem to be determined by the potential. To further prove the goodness of our NNP, Table 6 reports angle means and standard deviations also for the y structure and its sibling structure relaxed with the NNP, y_{NN}. As expected the distribution proprieties are very similar to the ones obtained from the melting-quenching technique with our NNP.

Another quantity commonly analyzed in amorphous samples is the ring-size distribution. Results for this metric are presented in Table 7, for samples generated from our NNP and with Tersoff potential, as well as for the DFT optimal structure y and its NNP relaxed variant, y_{NN}. From the reported results we stress the similarity between the samples generated from the NNP and the one optimized with DFT, and the larger deviation noticeable for the samples generated with Tersoff potential. The agreement with the DFT optimized sample could likely be further improved if we adopted the same optimization protocol.

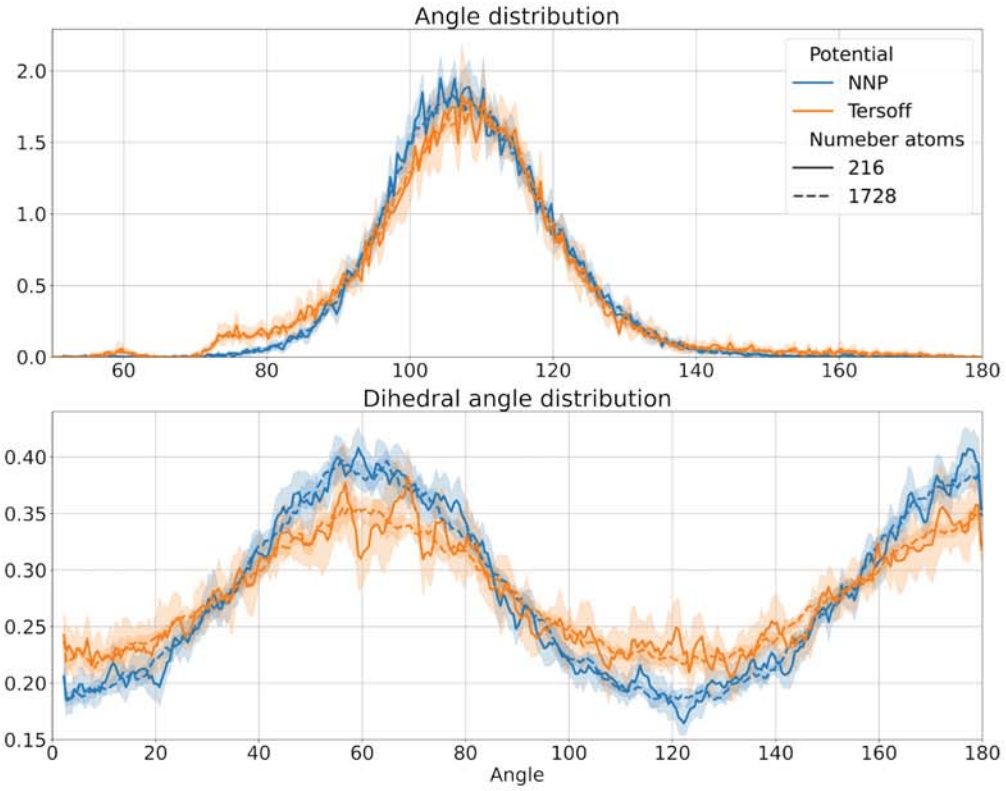


Figure 28: Angular distribution (top panel), and dihedral distribution (bottom panel) for the NN potential and and Tersoff [38] potentials, and different numbers of atoms. For each potential increasing the number of atoms only smoothes the curves thanks to the increased sampling. The features in the dihedral angle distribution shows the existence of a medium-range order, up to second nearest neighbors, more pronounced for the NN potential.

potential	n atoms	mean	std
NNP	216	109.08(4)	11.9(4)
NNP	1728	109.03(3)	12.0(2)
Tersoff	216	108.64(18)	14.3(9)
Tersoff	1728	108.62(5)	14.4(3)
y	216	108.92	11.6
y_{NN}	216	109.02	11.2

Table 6: Average angle distribution and standard deviation obtained with different potentials, and sample size. Ten samples were collected for each row. The last two rows are the statistics for the y structure provided by [86] relaxed with DFT and with our NN potential. The only significant difference in the results seems to originate from the potential choice.

Ring length	NN potential	Tersoff	y	y_{NN}
3	0.006(2)	0.012(5)	0.005	0.0
4	0.031(10)	0.087(28)	0.042	0.037
5	0.429(27)	0.520(57)	0.391	0.395
6	0.717(73)	0.668(112)	0.772	0.767
7	0.046(30)	0.042(23)	0.023	0.023

Table 7: Ring counting distribution, obtained with the ring definition from [132] and the code [125]. Results averaged over 10 structures with 216 atoms for NN potential and Tersoff potential. Also shown results for y and y_{NN} structures.

In conclusion, all these geometric metrics do not indicate any critical statistical difference between well-relaxed DFT configurations and the average configuration obtained with our NN potential. The same can not be said for structures produced with simulations using Tersoff potential.

3.2.2 Excess energy distribution

The excess energy with respect to c-Si is defined as the per-atom difference in energy between c-Si and a-Si and is stored in either point defects and/or lattice distortions. This value can be experimentally estimated with calorimetry by measuring the excess heat during a-Si/c-Si recrystallization and is estimated to be between 0.07 and 0.15 eV per atom [133, 128] depending on sample preparation method and thermal history.

Another quantity that we will investigate in parallel to the excess energy is the density of a-Si with respect to the one of c-Si. Experimentally, a-Si density is expected to be lower than c-Si density by 1.8±0.1% [134, 135].

These two quantities are presented together in Fig. 29. Looking at the relative density axis in Fig. 29, we can conclude that the a-Si per atom equilibrium volume, in general, is a property of the potential that is rather independent from the system size. The same cannot be said about the excess energy. Tersoff potential gives generally bad results and always predicts very high excess energy values that seem to be influenced by the sample size. The NNP always shows a consistent behavior, and the average configuration has the same excess energy regardless of the number of atoms in the cell. Unfortunately, the applied melting-quenching generation procedure seems to fail to generate amorphous structures as relaxed as the y_{NN} structure, the black dot in Fig. 29, so it must be concluded that the amorphous structures produced with our melting-quenching protocol still have a significant margin of excess energy that could

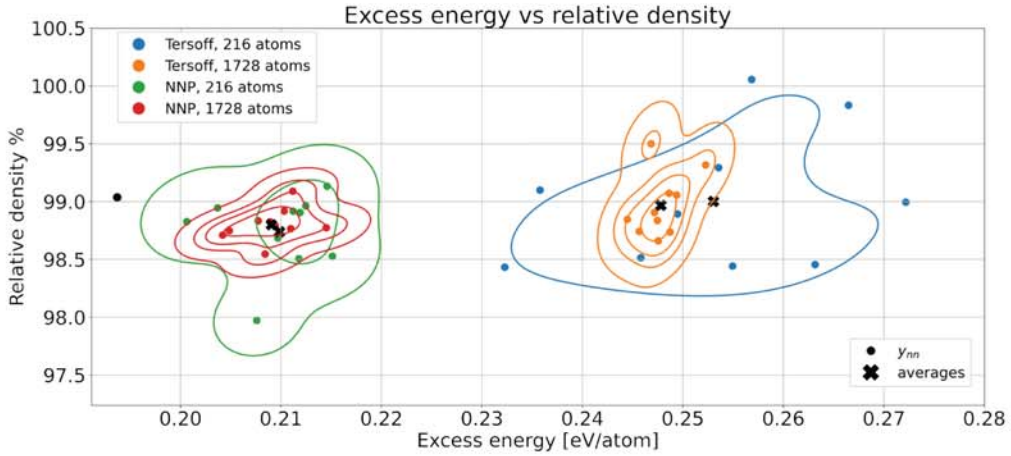


Figure 29: Two dimensional map of the excess energy and relative density with respect to crystalline diamond. For each potential, the corresponding diamond structure has been taken as a reference. The data are grouped by potential and number of atoms. The black dot corresponds to the reference structure y_{NN}. Solid lines are kernel density estimates of the distributions, due to the small number of data points they should be considered as an aid to the eye to estimate the reachable regions. Increasing the number of atoms only reduces the energy fluctuation while keeping the same average.

be further relaxed to bring them closer to the DFT optimized reference, for instance adopting the same optimization protocol. Unfortunately this important energetic difference is however not captured by the geometrical structural analysis described earlier.

3.2.3 Vibrational density of states

Another material quantity that can be used to characterize the amorphous phase is the vibrational density of states (v-DOS) which is sensitive to the local environment of the vibrating atoms. To compute v-DOS we adopted the frozen phonon approach as implemented in `phon` [136] as well as a code of our own that fixes several performance issues that `phon` experiences for big cells. Results for selected configurations are shown in Fig. 30. So far we only made the statistical comparison for cells containing 216 atoms.

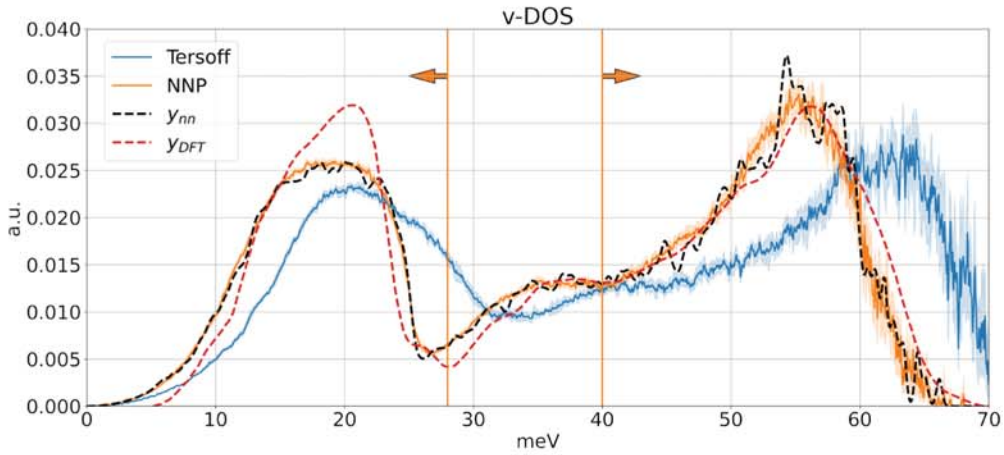


Figure 30: Vibrational density of states for 216 atoms cells for the NNP and Tersoff potential with error bars. Tersoff potential results are presented for comparison but shares little with our potential or DFT. DFT results are taken from Pedersen *et al.*[86] and refer to the y structure. Vertical orange lines represent the extrema used to perform the table's 8 integrals.

Starting by comparing the DFT spectrum of the y configuration from [86] with the v-DOS spectrum computed for the y_{NN} configuration, the main features are: 1) the difference in shape of the low frequency peak, 2) the presence of a double peak in the in the optical region of y_{NN}, 3) an overall underestimation of the frequency in the optical peak. For the first difference, we were not able to find its origin. For the second, it is very likely that our NNP relaxed some local configuration to something that locally closely resembles c-Si

making a c-Si-like peak appear at approximately 59 meV. This explanation is supported by the fact that a double peak is indeed often seen in experiments during implantation when the whole c-Si structure has not yet been completely amorphized [137, 138], and could be also the reason why the sample has the low excess energy observed in the previous section. As for the third difference, it is very likely related to the overall underestimation of the optical branches in c-Si phonon dispersion observed in section 2.6. Despite all of these differences the agreement between the NN potential results and the DFT ones is rather good and the number of modes encompassed by the two main low/high frequency peaks, reported in table 8, is very similar. The average v-DOS of the samples we created behaves very similarly to the v-DOS of y_{NN} except for the less structured optical peak. This, as commented before, could be a sign of the absence of c-Si-like environments that would give sharp c-Si-like modes in the optical region. More details on the v-DOS features are available in Appendix B

	low	high
DFT	36.14%	51.25%
y_{NN}	36.64%	49.77%
NNP	37.0(2) %	49.3(5)%

Table 8: Average Integrals for the low (below 28 meV) and high (above 40 meV) peaks in the v-DOS.

3.3 Creation of the SPE samples

From an estimate of the T at which a-Si/l-Si phase transitions occur of approximately 1100K, and a characterization of a-Si quality, we can proceed with SPE investigation. First, we create the sample as follows(a pictorial view with the key events is also presented in fig. 31):

1. An initial set of boxes of c-Si is created by replicating the 3.84×3.84×5.43 Å^3 ($[01\bar{1}] \times [011] \times [100]$) 4-atom unit cell, and then thermalized at 900K in an NPT ensemble. The cells used for this procedure were composed of 512 atoms (4x4x8), 2560 atoms (8x8x10), and 10240 atoms (16x16x10).

2. Before starting the next phase 4 (100) contiguous layers are frozen in their c-Si equilibrium configurations to ensure the final presence of a c-Si/l-Si interface.

3. The next step, the model of the l-Si phase, is achieved through an initial NVT run at a high temperature (2300K), followed by a cooling phase

down to 1800K, and a subsequent thermalization at 1700K using an NPT ensemble to allow the volume of the box to shrink and accommodate the correct liquid density.

4. At temperatures sufficiently low to create a l-Si/vacuum interface without any risk of material evaporation. This is done by cutting open one of the two l-Si/c-Si interfaces.

5. Two of the four c-Si frozen layers in contact with the l-Si are freed and set to interact with a reservoir at 900K, while the rest of the l-Si is slowly cooled at a 10K/ps rate as done for the a-Si creation in section 3.2

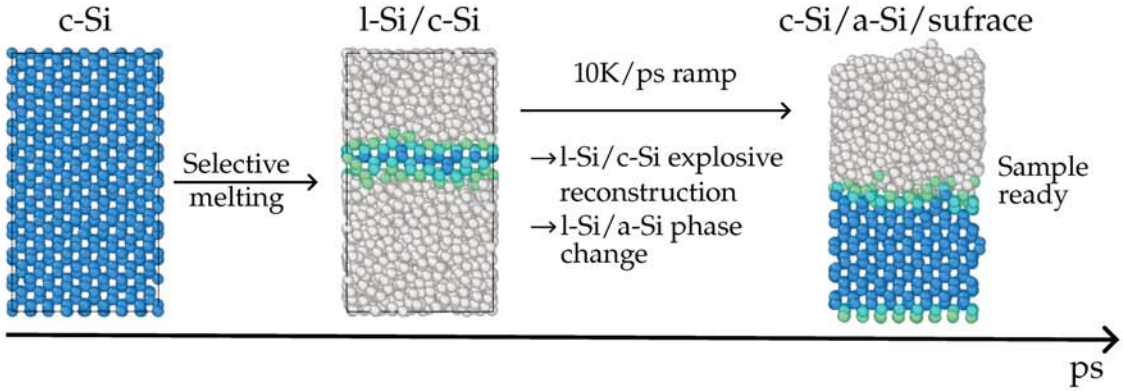


Figure 31: To create the sample needed to investigate SPE, we start with a perfect c-Si crystal. Then, by abruptly raising the temperature and freezing two layers, we achieve a selective melting of all the material except for the frozen layers. Then, after an equilibration, we open the sample on one side (this is achieved by removing the PBE and creating a surface) and slowly cool down the sample following a 10K/ps thermal ramp. During this thermal ramp, we observe an explosive reconstruction at the c-Si/l-Si interface that promptly stops leaving a more regular l-Si/a-Si phase transition. Once the ramp reaches a sufficiently low temperature, we are left with a sample ready to simulate SPE.

The arbitrary choice of adding 2 crystalline layers thermalized at 900K was made in order to mimic a realistic sample mounted upon a heated support, where a temperature gradient must be present [4], and to act as an intermediate stage towards the fictitious frozen bulk. Except for the smallest cell, which does not give sensible results, for the other two cell sizes a similar evolution can be observed during the last cooling ramp. At the beginning of the cooling ramp, the c-Si/a-Si interface grows homogeneously and without defects parallel to the

(100) surface, and does not show any specific feature or pattern. A rough estimate of the velocity with the method explained later in section 3.4.1 results in a value of approximately 13 10^{-3} nm/ps suggesting this to be an explosive regrowth (a value between 0.01 and 20 10^{-3} nm/ps is experimentally estimated for this phenomenon [139]). The measured explosive reconstruction put a further bound on the speed of our cooling ramp because in order to be able to spend more time at each temperature we would need to start with a much thicker sample.

This first mechanism lasts until a temperature of approximately 1400K is reached when some features start to appear at the c-Si/l-Si interface. Such features can be described as small "hills and valleys" of atoms with Voronoi volume larger than the liquid one. At 1050K the formation of a geometrical pattern exposing the {111} faces on the sides is always present. This pattern is compatible with what was experimentally observed in [17], where different reconstruction speeds were reported in different directions. An example can be seen in Fig. 32. This is a positive and easily reproducible result, not always obtained with other empirical potentials, that instead naturally occurs with ours.

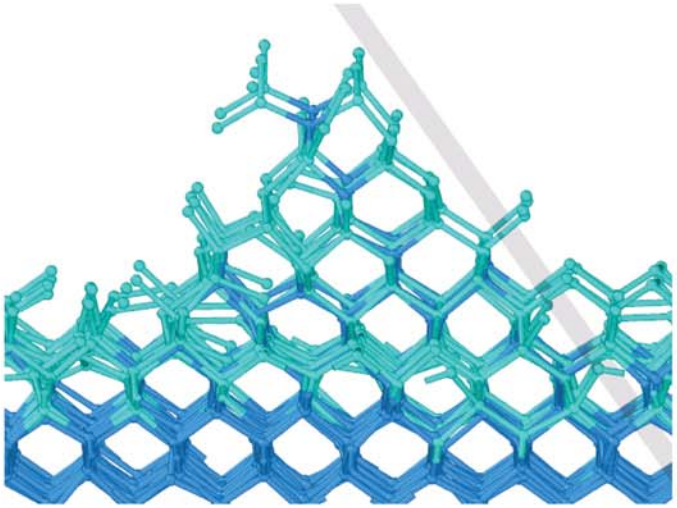


Figure 32: 3D wedge structure appearing during the solid phase epitaxy simulation at 1050K. Only tetrahedrally coordinated atoms with angles compatible with the diamond structure are shown, for simplicity. A (111) is drawn in gray.

We know from section 3.1 that the l-Si/a-Si phase transition at ambient pressure is possible between 1050 and 1100K. This, however, does not grant that the phase transition is taking place since coexistence between the two phases is still possible. To assess if our sample displays coexistence between the two states of matter the simplest discriminant to investigate is the density or its counterpart,

the per atom local volume. By defining the volume occupied by each atom as the Voronoi volume that encloses that atom it is possible to plot the distribution of this quantity in the sample as a function of the temperature. If the distribution is centered around the average liquid value no phase changes have happened yet; if instead it is distributed around the a-Si average value the phase change has already occurred. a Since there is a significant overlap between the a-Si per atom volume distribution and the c-Si per atom volume distribution, we need to exclude the atoms in the sample that are in the c-Si state from this analysis. This is done by performing a common neighbor analysis [140, 141], with an adaptive cutoff, modified to detect the diamond structure [142, 74]. This metric can detect also atoms in structures with symmetries different from the diamond one, but we restricted our investigation to two categories one of atoms detected as embedded in a perfect diamond structure and their first nearest neighbors, and one composed by of all the other atoms.

The distributions shown in Fig. 33 show, as expected, that the phase transition causes a shift in the volume distribution that starts from one compatible with the liquid phase and evolves to one closer to the silicon diamond phase, but broader. At 1100 K the presence of residual l-Si can not be excluded, suggesting that a fully glassy regime has not yet been achieved. Further analysis of such distribution as a function of the position along the [100] growth direction highlights that atoms closer to the a-Si/vacuum surface tend to have lower per atom volumes and therefore to be more liquid-like. At 1050 K the system is almost fully transformed to a-Si even if some sporadic l-Si-like environments are possible. And finally, below 1000K the phase transition is completed.

From this analysis, we can identify which temperatures are good candidates for our SPE simulation study: 4 configurations were therefore selected between 1100K and 900K every 50K. Finally, another sample was taken at 800K to investigate a low-temperature regime. For consistency: during the relaxation below 900K the bottom layer thermostat was attached to the top one and cooled with the same ramp.

To measure the barrier that regulates SPE one major problem is left, to identify where the c-Si/l-Si interface is located. This will be discussed in the next section.

3.4 c-Si/l-Si interface growth

From the previous analysis regarding the phase transition, it is possible to estimate that the fastest and most significant simulation should be the one at 1050K. We used this as a reference temperature to define a reliable way to identify the

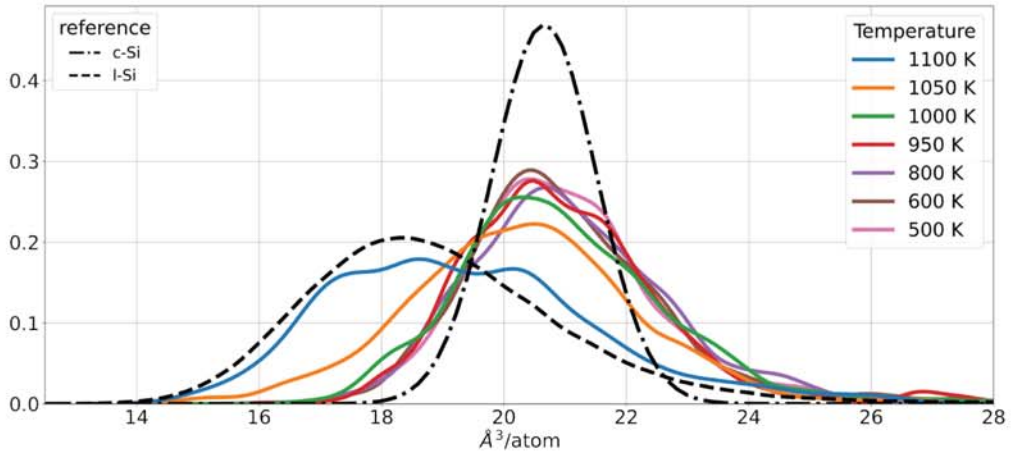


Figure 33: Distribution of the per atom volume measured using Voronoi tessellation and presented as a kernel density estimate fit for better visualization. In black are the two references: dot-dash line for distribution of c-Si at 900K, and dashed line for the distribution of l-Si at 1700K ; a-Si distribution is not reported for clarity as it would essentially overlap with the one of the sample at 500K. Other lines describe distributions after a 100 ps equilibration at different temperatures (see the figure legend for the corresponding color scheme) in a 2560 atom simulation cell. To obtain such distributions we took configurations at selected temperatures along the final thermal ramp (step 5) performed to create the sample. Those samples were then evolved, similarly to what was done to simulate SPE, in an NVT MD at a constant temperature.

position of the interface and investigate size effects. These latter will be discussed using the 10240 atoms and the 2560 atoms cell sizes previously mentioned in section 3.3.

3.4.1 Interface position

The measure of the reconstruction speed along the [100] direction requires a reliable definition of the interface location, and different approaches have been proposed to achieve this.

Mattoni *et al.* [33] proposed a possible definition based on a directional structure factor $S(z) = \left| \frac{1}{N_z} \sum_{z<z_i<z+dz} e^{i\mathbf{k}\mathbf{r_i}} \right|$ where $\mathbf{r_i}$ are the atomic coordinates of the atoms in a slab of width dz, $\mathbf{k}$ is a fixed vector in the reciprocal space parallel to the a-Si/c-Si interface chosen to match the crystalline lattice periodicity in the plane, and N_z is the number of atoms in the slab. The width d_z is chosen such that the slab contains at least two c-Si (100) lattice planes to always have a value different from zero. Once computed, at each z the function $S(z)$ will assume high values, lower than one, where the structure is ordered with the probed periodicity or small values where there are unordered structures. At the interface, a sharp (or smooth, depending on the local order) transition from a high value to a low value would be present, and an average position can be estimated from the decay.

This technique has been successfully applied in other studies [35] and gives good results in the case of sharp interfaces and large reconstructions. We have tried to use this method but it does not perform well in cases, like ours, where a non planar geometrical pattern is present: instead of observing a sharp degradation in $S(z)$ at the interface, a smooth transition is observed from the value associated with the crystal to the value associated with the amorphous. It is reasonable to try to extend this method by testing different reciprocal vectors other than the c-Si periodicity $\frac{2\pi}{a_0}(2, 2, 0)$. For this reason, we tried the orthogonal vector $\frac{2\pi}{a_0}(2, -2, 0)$, and the same vectors by doubling the frequency. In our experiments, adding the orthogonal vector did indeed help give sharper changes in the $S(z)$ at the interfaces, while sampling higher frequency vectors degraded the whole metric to the point of making it a nonviable option. Overall, the best results we were able to obtain were not sufficient to uniquely identify the position of the interface lacking a clear interface motion as a function of simulation time.

Following these considerations, we opted for a different method relying on the atom classification used in section 3.3, and performed via the OVITO [74] software. We restricted our analysis only to atoms detected as perfect diamonds

(d) or as nearest neighbors of perfect diamonds (d1). Once these atoms are identified, we can multiply their number by either an average c-Si volume compatible with the simulated temperature or their exact Voronoi volume. It is, then, possible to divide this number by the interface area in order to obtain the c-Si layer width $\overline{S}$, and the evolution of this quantity represents an approximation to the interface speed. Since this way of counting the atoms is a somehow arbitrary definition, we evaluated two different options to appreciate its robustness: using only the number of the perfect diamond atoms, d atoms, or using both the perfect diamond and the nearest neighbor to diamond atoms, d and d1 atoms, combined.

Since we have found that a change in the method used to estimate the atomic volume can impact the speed estimate up to 5% of its magnitude, while a change in the atoms categorization can impact it up to 40%, we will report only the discussion over the second quantity. As for the volume, we will only consider the use of local Voronoi volumes. We believe that obtaining consistent results with the two crystalline descriptors (d vs d and d1) for counting crystal atoms will give a sensible feeling over the errors incurred by the two methods. Results are shown in Fig. 34, and detailed speed estimates for the various regions of the simulation, see later, are available in Table 9.

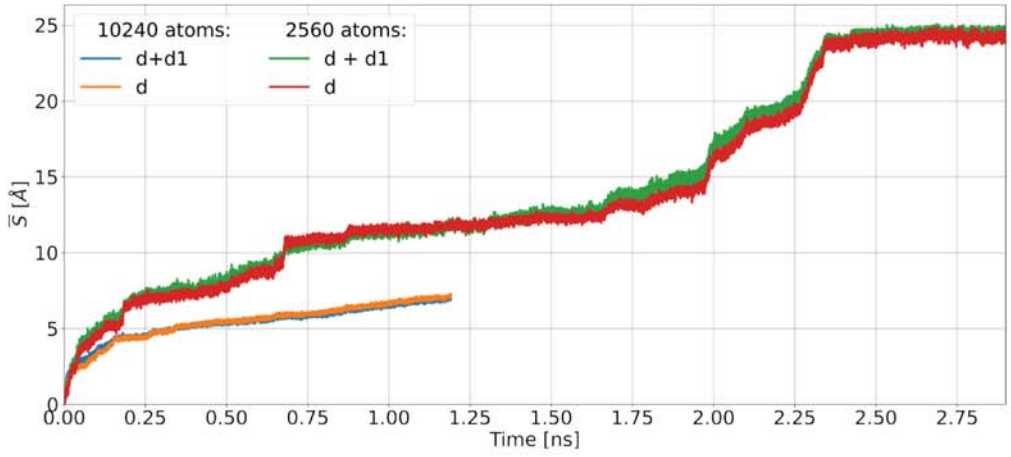


Figure 34: Position along the [100] axis of the a-Si/c-Si interface as a function of time, computed with the two different metrics and for simulations of different sizes at 1050K. The large cell simulation lasted 1.2 ns and consisted of 10240 atoms, while the small cell one consisted of 2560 atoms and lasted 3 ns. A detailed analysis on the physics is available in section 3.4.2.

3.4.2 Size effects

Since simulations are expensive, it would be desirable to use the smallest possible cell that can capture the physics of interest, although this might come at the price of more complex analysis. We will here present a comparison of the simulations performed with the two cell sizes anticipated in section 3.3 of 2560 atoms and 10240 atoms.

As shown in Fig. 34, when comparing the two cell sizes there are different mechanisms that come into play during the recrystallization, and not all of them are common to the two cell sizes. In the larger cell it is possible to see that, after an initial equilibration phase of about 165 ps, the system evolves at a roughly constant speed with rare small jumps. We have isolated these jumps and determined that they are due to a rapid phase change of small clusters (approximately 25/35 atoms) embedded in the interface, that are kept under stress by defects. Once the defect is removed, the whole cluster becomes crystalline. This crystallization is roughly one order of magnitude faster than the regular reconstruction, and it is thus registered as a jump in the interface position, but it does not impact significantly the average velocity.

Considering the smaller cell, a similar initial 150 ps equilibration period is present. After it, we can identify at least two rather large jumps, one at 180 ps and one at 660 ps, connected by what could be another equilibration phase or a first steady recrystallization phase. However, since the measured speed seemed to be accelerating, we decided to consider the system not yet equilibrated and excluded this region from further analysis. The range between 680 ps and 1600 ps is a region with small jumps that shows the same behavior as in the large cell. As such, we chose this time interval as the prototype to estimate the steady-state SPE velocity. Fig. 35 shows how the atomic environments changed during this time. After 1600 ps, the amount of material left to undergo a-Si/c-Si phase transformation is not sufficient to observe a clean SPE reconstruction, due to the closeness of the a-Si vacuum surface the quality of the simulation degrades. The final reconstructed material is less than the maximum available amount because, in this last part of the dynamics, defects start to embed in the crystalline lattice, and some atoms that compose the crystal are no more recognized as diamond atoms by the descriptor.

The two jumps at 180 ps and 660 ps in the 2560 atom simulation are of particular interest, because similar phenomena have been reported in the literature [36, 32, 37, 143], but they are absent in our simulations with the larger cell. In our cases, we have approximately 50 atoms that change in a way close to what happens in the small jumps, but on a much larger scale. Snapshots before and after the 680 ps jump are presented in Fig. 36. The presence of jumps in the numerical simulations of the SPE with various potentials led previous authors

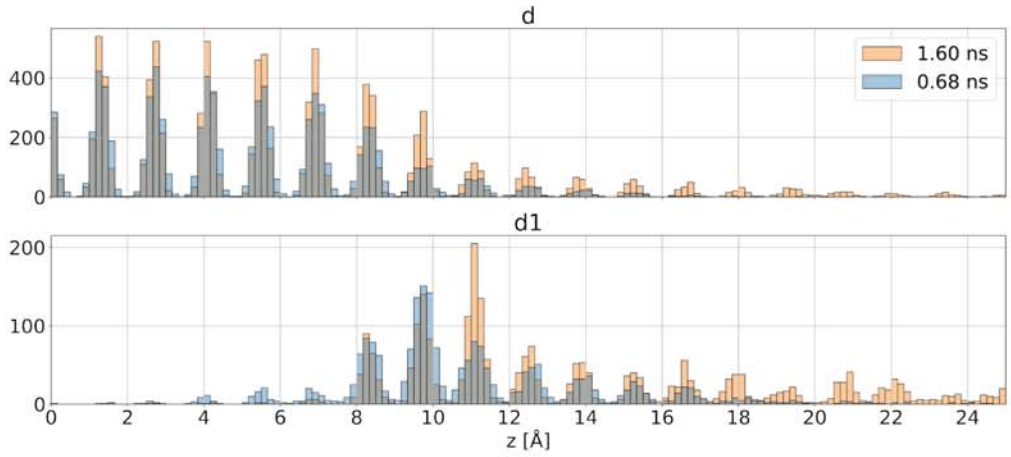
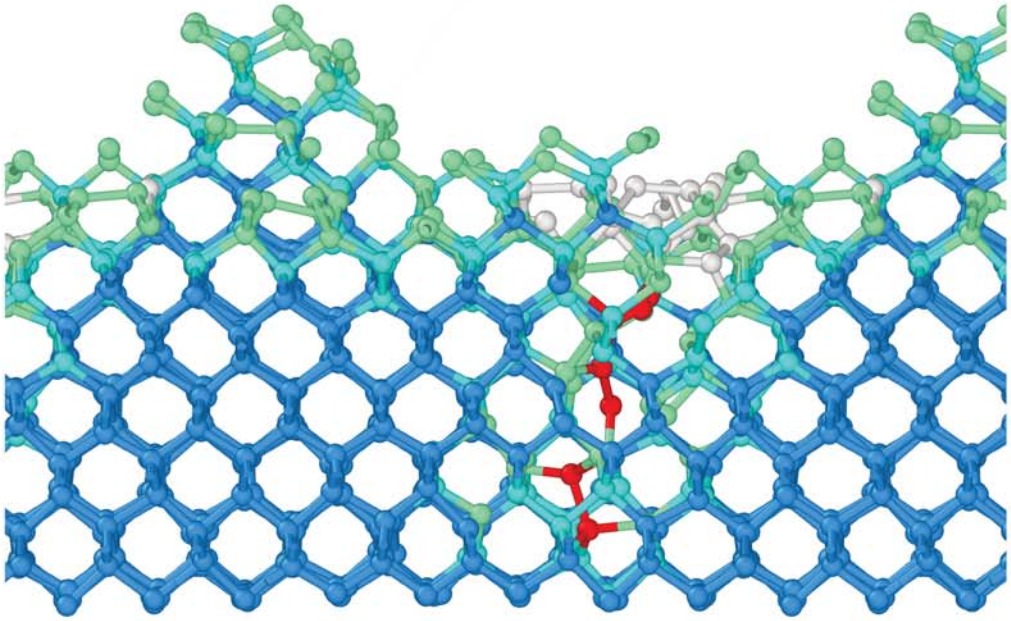


Figure 35: Distribution of the atoms category average over 2 ps at the interface during a normal reconstruction in the 2560 atoms case. The x axis represents atoms positions in the [100] direction. The first row shows the number of atoms in the crystalline configuration at a given height: we see that they are organized according to the lattice, and their number decreases as we get closer to the interface. The color represents time steps: 680 ps and 1600 ps. As time passes, the atoms deeper in to the sample (left) only reorganize under thermal vibrations, while those closer to the interface (right) tend to present more c-Si atoms. The second row is the same graph for the d1 atoms: we can see that they are distributed around the interface and follow, as a function of time, a pattern similar to the one of c-Si.

Before the jump, at 650 ps



After the jump, at 680 ps

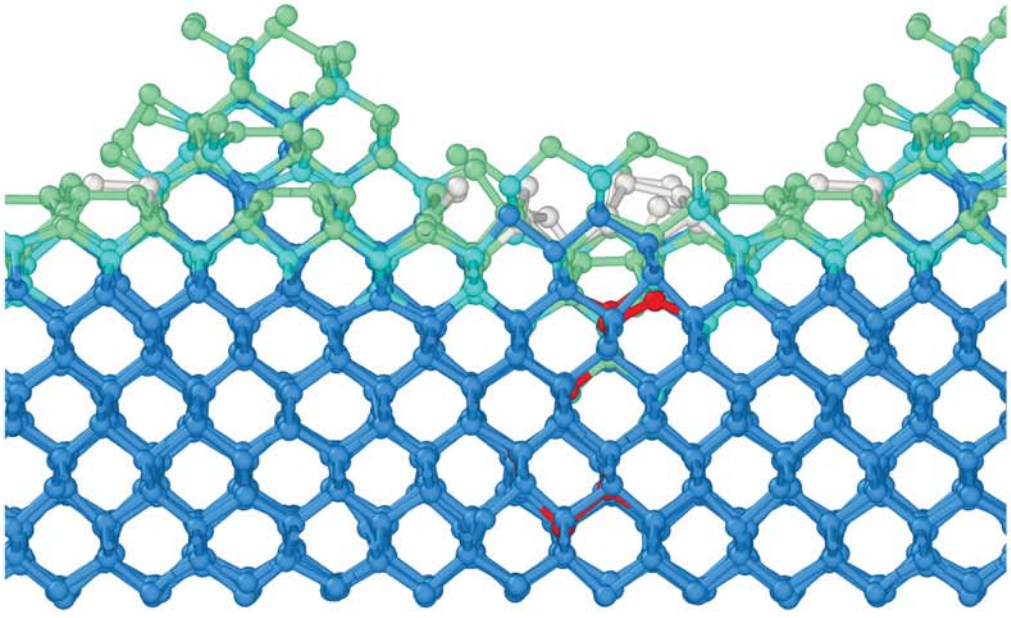


Figure 36: Snapshot of the atomic configurations during the jump registered between 650 and 680 ps in the simulation with 2560 atoms. Atoms are colored accordingly to the classification provided by Maras *et al.* descriptor.

Figure 36: (continuation) Blue atoms (d) are in perfect crystalline positions. Light blue atoms (d1) are first nearest neighbor of crystalline atoms and greens (d2) are second nearest neighbor. White atoms are in an amorphous state. Red atoms where originally white and are colored to show where the event is happening. It can be seen how this red atoms are trapped between two 111 wedges that are artificially close due to the presence of the periodic replica of the main feature.

to consider this as the main mechanism of the reconstruction, but our analysis would suggest otherwise. We hypothesize that these events are only possible due to the limited cell size that does not allow for the $\{111\}$ facets that compose the wedges at the c-Si/a-Si interface, see section 3.3 and Fig. 32, to grow beyond the periodic cell size. Indeed similar jumps are observed, with somewhat greater frequency, in SPE simulations performed in smaller cells with a potential reported to be better than SW [36], but the effect of cell size was not studied. Simulations with both SW [32] and Tersoff [37] potentials show rare jumps on time-scales and cell sizes similar to ours, and report no jumps in larger cell, as in our case. Tersoff potential displays a similar behavior also for germanium [143].

Our analysis allows us to split every simulation in different regions, each displaying its own regime. The average velocities computed for these regimes are presented for different cell sizes and descriptor definitions in table 9. It can be observed that the velocity estimates are stable with respect to the change in the descriptor threshold. After eliminating the regions that we deemed not representative of steady SPE (equilibration periods, large jumps, etc.) we can compute the average SPE velocity in the steady regime, presented in bold in the table 9. We can notice that these SPE velocity estimates are rather insensitive to system size, thus justifying our choice of restricting the more systematic study to the small system only.

The analysis described was applied to all simulations performed at different temperatures and only the average velocity corresponding to steady SPE growth simulation region will be reported in the following.

3.4.3 Evolution of the sample profile

Previous studies neglected an analysis of the material properties as a function of the positions along the [100] direction. We believe that there are at least two relevant quantities to monitor that give powerful insights into any further discussion. These are temperature and diffusion. In order to do this analysis, samples were sliced into three parts. The top, or surface, region was composed exclusively of a-Si atoms. The central, or interface, region was composed of a mixture

Event category	Time interval [ps] begin	end	Speed [10^{-3} nm / ps] d	d+d1	Temperature NVT sim. [K]
2560 atom cell					
equilibration	30	150	2.18	2.24	1038
	200	650	.495	.539	1044
jump	650	680	2.68	1.93	1041
	860	880	2.38	2.77	1050
	1600	1700	.785	1.06	1046
	1950	2050	3.02	3.15	1043
avg	**680**	**1600**	**.175**	**.243**	**1043**
end region	1700	1950	.526	.588	1044
	2050	2300	1.25	1.22	1049
10240 atom cell					
equilibration	10	30	3.34	3.80	1030
	30	165	1.56	1.35	1024
jump	250	275	1.52	1.28	1030
	325	350	1.28	.804	1026
avg	**165**	**1191**	**.265**	**.246**	**1027**

Table 9: Speed estimates during different regimes (time regions) of the trajectories. The upper part of the table refers to the 2560 atom cell simulation, the lower part refers to the 10240 atom cell simulation. Different qualitative categories have been defined based on the sample behavior in that time interval and refer to the discussion in section 3.4.2. Speeds are reported according the two considered ways for identifying bulk atoms in the interface definition (d vs d + d1, as defined in section 3.4.1). The reported temperatures are the average temperatures in at the interface in the specified time interval. The average SPE speeds after equilibration (excluding the surface-affected final phase in the 2560 atom cell) are marked in bold and are our best estimates for the SPE speeds.

of a-Si and c-Si atoms. And the bottom, or bulk, region was composed only by c-Si atoms. During the analysis, the boundaries of each region have been kept fixed. This choice was taken because the interface motion during the simulation was small compared to the 15–20 Å typical region width.

Starting the analysis with the temperature, we observe that the system is never at equilibrium with the thermal bath of Nosé-Hoover thermostat, and a temperature gradient is present. Fig. 37 reports the temperature evolution as a function of time for the three selected regions. During the crystallization process, the interface releases energy, and this tends to rise the sample temperature. The excess kinetic energy migrates towards the surface and the bulk. On the surface, the only mechanism that can remove the energy is the interaction with the thermostat, which is not happening fast enough. Therefore, the surface becomes hotter than the connected thermostat. This could, in principle, be fixed by increasing the frequency at which the atoms interact with the thermal bath, but we preferred not to do so to sample more closely the more realistic microcanonical ensemble. On the bulk side, our choice of adding a two-layer buffer zone interacting with a colder bath at 900K between the frozen atoms and the rest of the sample prevented the temperature from rising significantly above the bath temperature and kept the c-Si cool enough to have an interface at the desired temperature.

Another quantity that is not constant across the three sections of the sample is the diffusion coefficient. This quantity is derived from the mean square displacement of atoms per unit time :

$$D(t) = \lim_{T\to\infty} \frac{1}{6\ T\ N_{\text{at}}} \sum_{\mathbf{r}_i(t)\in\text{region}} \left|\mathbf{r}_i(t+T) - \mathbf{r}_i(t)\right|^2 \tag{3.1}$$

where the sum, generally extended over all the atoms in the sample, is restricted to those (N_{at}) that are in the region of interest at time t. We found that for the surface and interface regions the value of D is converged for $T = 100$ ps. It is not the same for the bulk region where diffusion is basically not present. Results are reported in figure 38.

Starting from the bulk region the lack of convergence in the c-Si region means that if we were able to increase T further in equation 3.1 the net effect would be only a rigid shift of the curve, while other curves would not change. This is important because it gives us a reference value for the non-diffusive behavior with the chosen value of T. The interface region instead displays a low but clearly non-zero diffusion coefficient as expected from the regular a-Si/c-Si mixture. The most interesting information comes from the surface region. This region has an average density compatible with a-Si (see Fig. 39) but it has a diffusion coefficient that is in the range between the one of regular l-Si (HDL) and

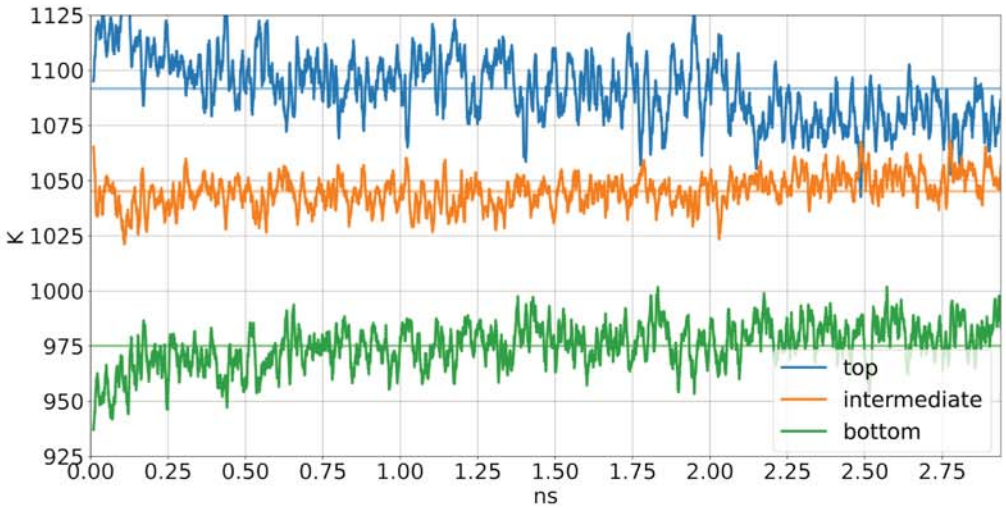


Figure 37: Temperature in the different regions for the case of the 2560k atoms as a function of time in the simulation targeted at 1050K previously presented in figure 34. In the beginning, we can see that the 180 ps equilibration phase is common to all regions. In the middle, the period of time deemed valid for the reconstruction has constant temperatures in all the regions, especially at the interface where relevant phenomena are happening. In the end, after 1.7 ns once the sample has crystallized no more energy is released by the phase transition, and the surface cools down.

LDL [144]. Even if part of this effect can be attributed to surface diffusion, and more targeted simulations are needed to obtain an exact value for the diffusivity of LDL, we can not exclude that the surface region of our sample is in the LDL state and not in a classical a-Si phase.

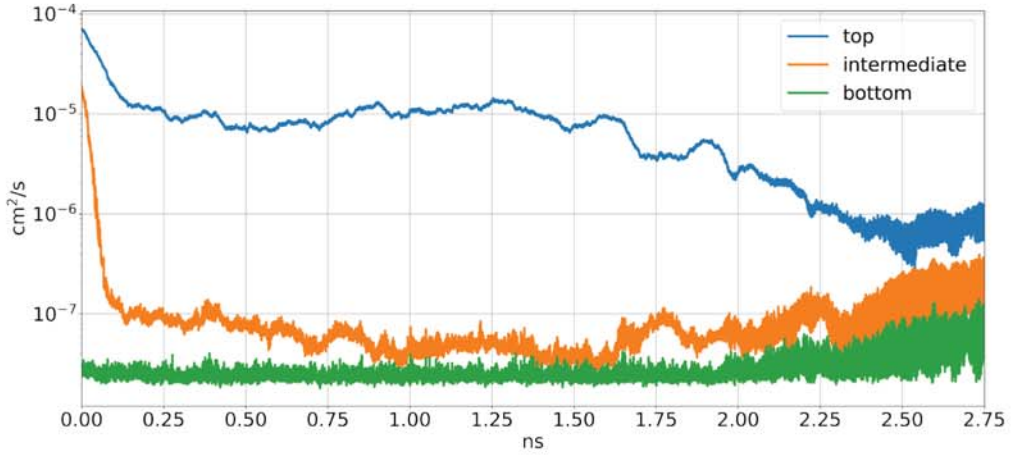


Figure 38: Diffusion coefficient in the different regions of the sample for the case of the 2560 atom cell as a function of time in the simulation targeted at 1050K previously presented in figure 34. In the first 180 ps during the equilibration, we can see a very high diffusion phase is common to both the top and the intermediate regions. In the middle part of the simulation diffusion is constant in all the regions. Toward the end, after 1.7 ns, the sample has crystallized, and diffusion reduces to defect migration and other unknown spurious phenomena we did not investigate.

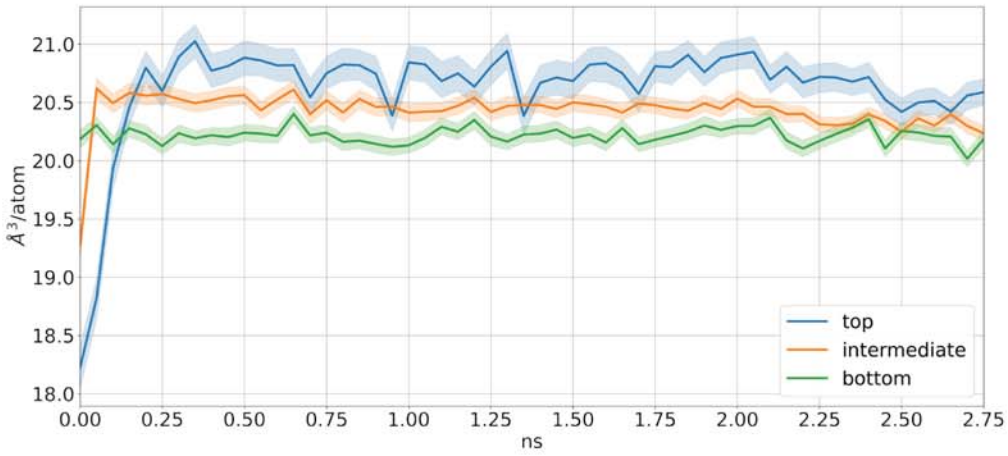


Figure 39: Average and standard deviation of the per-atom volume in the different regions of the 2560 atom cell as a function of time in the simulation targeted at 1050K previously presented in figure 34. This shows us exactly what is happening in the two relaxation regime below 180 ps, during this period a transformation of the material from the l-Si state to the a-Si state can be seen. The first relaxation regime terminates when the interface region becomes amorphous. The lower asymptotic value is due to the presence of both c-Si and a-Si in this region. The second relaxation regime terminates when the whole sample has become a-Si. In the end, after 1.7 ns all the densities converge towards the c-Si value.

3.5 The physical phenomenon

With the definitions given in the previous sections, we performed different simulations of SPE at five different temperatures between 1100 and 900 K and the collected data are presented in Table 10. The two velocity values reported at each temperature, obtained from the two atomic descriptors used to identify bulk atoms, crf. section 3.4.1, give an estimate of the uncertainty due to this choice, which, we believe, is the largest source of uncertainty in the determination of the velocity. On the obtained velocities we fitted the previously introduced Arrhenius law (sec 1.1):

$$v = v_0 \exp \frac{-E_a}{k_b T} \tag{3.2}$$

where E_a is the activation energy of the process and v_0 is the velocity prefactor. By excluding the low temperature point at 800K that is an outlier for reason that will be discussed later, the measured barrier E_a, is between 2.72 eV and

2.82 eV, depending on the chosen descriptor. Either value is in very good agreement with the experimental accepted value of approximately 2.7eV (table 1 in section 1.1.1). The velocity prefactor v_0 instead is estimated to be between 4 10^9 and 1.1 10^{10} nm/ps, and this is considerably overestimated with respect to experimental estimates. This overestimate is a direct consequence of the velocities obtained in our simulations being roughly eight orders of magnitude faster than the experimental observed values. To explain this large discrepancy as well as the results obtained at the lowest simulated temperature, 800K, we need to put these data in the context of the literature.

temp [K]	d [10^{-3}nm/ps]	d + d1 [10^{-3}nm/ps]	sampling time [ps]
1098	1.93	1.71	290
1043	.175	.243	920
1027	.265	.246	1026
995	.0209	.0303	2628
943	.0175	.0184	2214
799	.0171	.00981	3156

Table 10: SPE velocities with different descriptor thresholds (second and third columns) at different temperatures (first column) averaged over different time-length. Data were all extracted from steady state periods in the cell with 2560 atoms except for the point at 1027 K that was obtained from the simulation with 10240 atoms. At each temperature, the available time length depends is due to several factors. In the high-temperature regime, the reconstruction speed and the limited amount of material available to reconstruct are binding us to only consider a small time window as a valid representative of the real SPE. At low temperatures, on the other hand, to produce very small reconstructions (hundred of atoms) longer runs are needed and computational time is the only boundary. Temperatures reefers to average temperatures during the sampled time at the interface slab. The last raw has not been used in the fit of the Arrhenius law because because we believe it describe a different phenomena, more details are available in the text

We collect in Fig. 40 SPE speed results from experimental and numerical simulation literature together with the results of our simulations. As a general comment, it is clear that all theoretical estimates severely overestimate SPE speed, with the exception of simulations using Tersoff potential which however severely overestimate the temperature range in which the phenomenon takes place.

Bernstein *et al.* [30] performed a study with the EDIP potential [40, 39] (pink line and points in Fig. 40). Their velocities were obtained with a procedure

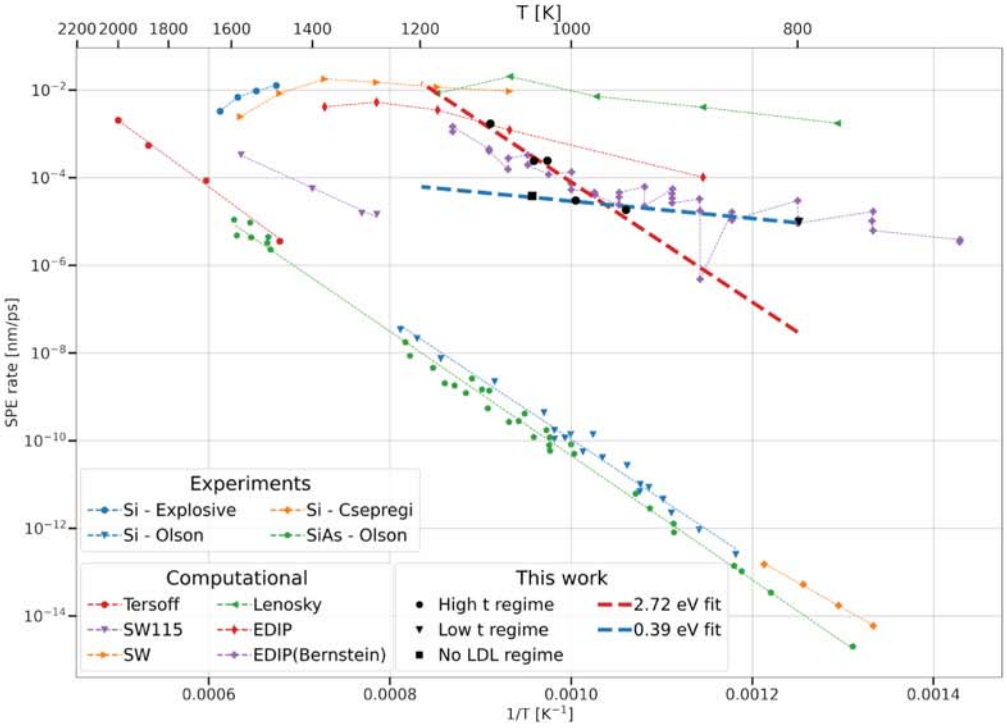


Figure 40: Arrhenius plot of experiments, numerical experiments, and our calculations. Experiments: pure Si from Olson *et al.* [19], Csepregi *et al.* [17], SiAs from Olson *et. al* [145], and explosive reconstruction from Galvin*et. al* [146]. Numerical works: EDIP, Lenosky, SW (no barrier was fit for those tree) SW115(with a barrier of 1.87 eV), Tersoff (with a barrier of 2.99 eV) from Krzeminski*at al.* [34], Bernstein *et al.* [30] (2.0 and 0.4 EV). This work: black dots as indicated in the legend, red fit (2.72 eV) and blue (0.39 eV) as discussed in the text.

close to the one used in our approach and with similar sample sizes. In these simulations, they obtained points that show a change of regime as a function of temperature, similar to the change of behavior we observe when comparing simulations above or below 900K. To justify their finding in a way compatible with the observed single barrier experimental evidence, they postulated the existence in their sample of two mechanisms with different barriers, of which only one would be relevant in the experimental conditions. Of these two mechanisms, one mechanism with a barrier of 0.4 ± 0.2 eV is the dominant one below $950K$ while at higher temperature the dominant mechanism has an activation energy of 2.0 ± 0.5 eV. This higher activation energy event is in fair agreement with the experimentally observed barrier of 2.7 eV [4] which has then been interpreted to represent a combination of a defect-formation event with an energy of 2.3 eV and a defect-migration event with an energy barrier of 0.4 eV [22] resulting in the experimentally observed barrier. They attributed their observation of a second mechanism at lower temperature to an excess of defects in a-Si in their sample (of the order of 5 % according to their estimates) giving a significant contribution to the reconstruction due to migration of pre-existing defects. This contribution is absent in well relaxed experimental samples while an enhanced recrystallization speed is reported in non-relaxed ion-implanted samples by DSC experiments (see Fig. 7) in 1.1.1).

This interpretation is not shared by Krezeminski *et al.* [34] that performed a comparative study with different potentials and deemed EDIP (blu points and line in Fig. 40) not to be able to correctly describe the phenomena that are at the base of SPE. In this work, an a-Si generated with the WWW technique combined with the Keating potential [147] is placed in contact with a substrate of c-Si while the other side of a-Si is exposed to vacuum. This initial configuration is then relaxed for 4 ps at different temperatures and with different potentials. After the relaxation, similarly to what we have done, a Nosé-Hoover thermostat at the target temperature is used to sample the reconstruction speed for each temperature/potential pair. Once results were collected and a fit to Arrhenius law was performed (Fig. 40), they argued that potentials such as MEAM (in the Lenosky parametrization [103]), SW [41], and EDIP (all potentials featuring in the topmost region of Fig. 40) are not capable of reproducing any SPE, while this is not the case for SW115 [42] (a SW parametrization that favours tetrahedral coordination) and Tersoff [38] potentials which correctly extrapolate the barrier while incorrectly describing other important physical properties, such as melting temperatures, phonons and local geometries that we think are important for this phenomena.

We believe that Krezeminski *et al.* [34] conclusions, at least as far as EDIP

potential is concerned, are not compelling. Recent calculations place the melting temperature corresponding to EDIP potential around 1520K [148], close to the 1350K predicted by LDA/DFT [84]. For this potential the amorphous melting temperature is therefore expected to be around 1175-1200K [42, 149]. This means that three of the five temperatures simulated by Krezeminski *et al.* are in the liquid regime and, therefore, can not be regarded as representative of SPE. As for the two simulation points at lower temperatures, analyzing the geometrical data provided in the paper, it is possible to conclude that the a-Si relaxation reached in these cells is not as good as the one obtained by Bernstein *et al.*. If this is the case, then, using Bernstein *et al.* argument, one would expect to measure a faster SPE velocity with respect to the one reported in their study, which is indeed the case.

Our simulation points endorse Bernstein *et al.* "two barriers" interpretation where defects can migrate with a 0.4 eV barrier, and the migration/formation happens with another barrier that we estimated to be about 2.7 eV, in excellent agreement with the experimental determination. What we obtain is also in line with DSC experiments (shown in section) where SPE speed for defective samples at low temperatures is driven by a second mechanism characterized by a lower barrier. The elephant in the room that both our and Bernstein *et al.* models suffer is the six orders of magnitude difference with respect to experimentally measured recrystallization velocities.

We believe that no single source can be responsible for this error. Accepting the two barrier interpretation the equation that governs the SPE can be written as:

$$\left[\rho_0 + \rho_1 \exp\left(-\frac{\Delta E_0}{k_b T}\right)\right] \exp\left(-\frac{\Delta E_1}{k_b T}\right) \tag{3.3}$$

where ΔE_0 is the defect formation energy, ΔE_1 is the activation energy of the defect migration, ρ_0 is the pre-formed defect concentration and ρ_1 is the concentration of sites where defects can form.

In the low temperature regime $\frac{\Delta E_0}{k_b T}$ is negligible and the only important exponential that remains is $\frac{\Delta E_1}{k_b T}$ associated to defect migration. As a consequence, in order to bring our results in better agreement with experiments the needed correction is a reduction in the pre-formed defect concentration ρ_0. This observation is compatible with our a-Si sample characterization showing highly defective a-Si.

Raising the temperature triggers the second mechanism. If we believe ρ_1 in our samples to be in line with experimental values, the estimated value of ρ_0 from our low temperature simulations is so high that higher barrier mechanism would become observable only around $\approx$1900 K. At such a temperature, however, our a-Si would be completely melted. This is clearly not the case since

in our simulations already at approximately 950 K we see a behavior that is explainable only by combining the two exponential.

Since the estimated 2.7 eV barrier is in line with experiments, and hence the underlying mechanisms should be correct, we believe that the source of the discrepancy is mainly imputable to a very different value for ρ_1. Indeed it is reasonable to assume that the density of locations where a new defect can form is not simply proportional to the number of atoms in the a-Si state but could be also influenced by the concentration of neighboring atoms already in a defective state. As previously discussed above 950 K our analysis shows a complex slab where the interface is in contact with a region of amorphous material, itself in contact with another region displaying liquid-like diffusion like LDL [120]. We think that in this region the blurry and rather large a-Si/LDL interface increases the number of atoms in defective states effectively making ρ_1 much larger than what it should be.

We verified the relevance of the LDL/a-Si interface by artificially removing the LDL. The removal of this state was obtained by freezing the c-Si/a-Si interface region and evolving only the portion of atoms in the LDL state for 1 ns in an NVT environment. The final result is that, since no heat is produced by the SPE, an equilibrium with the bath temperature can be achieved. This was done for the simulation at 1050K. After such a procedure the frozen atoms were freed and SPE restarted. The behavior is similar to the one previously observed with an initial equilibration that then saturates in a stationary SPE that shows a velocity of 4.35 10^{-5} nm/ps for the d metric and 3.81 10^{-5} nm/ps for the d + d1 metric. These values, combined with the ones presented in table 10 for 1000 and 950 K gives a barrier of 0.2 eV and 0.39 eV respectively that is also presented in figure 40, and is in good agreement with the 0.4 eV value estimated in experiments. This result shows how the presence of LDL or of a region of a-Si with higher diffusivity is key to reproduce our results.

We believe that in order to observe the 2.7 eV barrier in our samples without the aid of an LDL region a much less defective a-Si is needed.

Achieving such high quality a-Si samples by the melting-quenching approach would require thermal ramps similar to the experimental ones which are several orders of magnitude slower than the ones possible in MD simulations, hence other techniques are required, for example by applying methods such as off-lattice KMC [150], or by optimizing the excess energy as done by Pedersen *et al* [151, 86]. Another possibility that we are pursuing is the characterization of the amorphous trough quantities such as phonon localization [152, 153], and soft spots [154, 155] that will allow for surgically act on regions with very localized low frequency modes. The result of such defect removal will corresponds to a lower ρ_0 in equation 3.3, moving the corresponding contribution to SPE to

lower values, and offering the possibility of observing the mechanism characterized by the 2.7 eV barrier already at lower temperatures. The reduction of ρ_0, and the removal of LDL-like regions in the a-Si region, will probably also affect ρ_1, lowering the SPE simulation points characterized by an high barrier toward experimental values.

Another, independent, improvement to the simulation quality would derive from the use of a better DFT functional, able to predicts a more accurate melting temperature. Adoption of a different DFT functional would not have a simple effect like a mere temperature scaling as discussed briefly in Appendix A (that will have the inevitable side effect of scaling all the barriers), but will likely change the equilibrium between a-Si and LDL. The more likely scenario is that LDL will appear only at higher temperatures with the result of shifting the 2.7 eV portion of the simulated SPE in Fig. 40) to the left thus improving agreement with experiments.

Applying all these steps in an attempt to make our simulations closer to reality will simultaneously lower the simulated SPE speed to a point where direct simulation by MD would not be a viable option anymore. However, since the end goal of this study is the identification of the correct microscopic mechanism underlying the phenomenon we think that the current level of accuracy of the simulation is already sufficient to start an event characterization, that will lead to improved KMC approaches. This can be achieved for example, in a statistical fashion, by following the evolution during the MD runs of the local environment of every single atom and by clustering and labeling them so as to allow to compute the transition probability between states from the simulation itself. The resulting models could then be run on a larger system to validate their predictions against known quantities, such as the different reconstruction speeds along different directions, and to investigate the formation of $\{111\}$ wedges at the c-Si/a-Si interface during SPE, a feature that, if large enough and sufficiently periodic, should also be accessible to TEM investigations.

Chapter 4

Summary and Conclusion

The work presented in this Ph.D. thesis tries to elucidate the physics behind SPE, an essential process in the design of chips using the 3D monolithic integration technique as a way to leverage the third dimension. We achieved so by performing in silico MD simulations with a novel potential developed within this work. This potential is based on an artificial neural network (NN) that allows the scaling from a quantum mechanical approach -numerically unfeasible- to a classical one.

In the first part of the project, addressed in chapter 2, we developed a NN potential and the associated data set needed for its training. To create the dataset we sampled, by means of MD, all the regions of the phase space that we deemed relevant for the SPE investigation. Namely c-Si, a-Si, l-Si, and some surfaces. This sapling was performed using an iterative approach as it is customary when generating ML potential. The energy and forces for the final set of configurations have then been computed with DFT and two different functionals: PBE and PBESol. To construct the interatomic potential, we trained an all-to-all fully connected feed-forward ANN network, and as a descriptor, we used a modified version of the Behler and Parrinello atomic symmetry functions as we implemented in the "Properties from Artificial Neural Network Architectures" (PANNA) package.

The final accuracy reached by both potentials expressed by standard ML metrics (e.g. RMSE) is in line with results published in the literature, but we showed how this metric varies among different regions of the data and can not be trusted as the only way to characterize the goodness of a potential. Accuracy in regions with lower entropy such as c-Si systematically outperforms accuracy in higher entropy regions such as l-Si. On top of standard characterization, we investigated several physical properties and compared them with DFT ones. For c-Si, we investigated the energy-volume curve around the minimum, phonons, and elastic constants. For all of them, we obtained an agreement with

the reference. It stands out that we were never able to recover the stiffness of the TO modes, and we think that this is due to the presence of smearing in the reference DFT. For a-Si we investigated the structure factor, obtaining sufficient agreement with experiments, and for the surface, we showed how the potential is always capable of recovering the correct reconstruction geometries. However, for the surface, we have identified a limitation: the small difference in energy between the p(2x2) and c(4x2) is too small ($\approx$0.2 meV/Å^2) to be correctly recovered.

One of the main characteristics needed to study SPE is a good prediction of the c-Si melting temperature. Our first functional tentative (PBESol) systematically gave us a very low melting temperature of 1194$\pm$29 K. We were only able to solve this problem by chaining the functional to the original PBE which gave us a melting temperature of 1468$\pm$11 K. This result is still below the experimental value of 1685$\pm$2 K but usable for our purposes. This discrepancy among the two functionals however showed that with the accuracy reached by our potential the difference with the experiment can be traced back to the quality of the DFT functional, giving a solid starting point for future improvements.

In the second part of this work, presented in chapter 3, we have applied the generated PBE version of the NN potential to the investigation of SPE. Since experimentally SPE microscopical mechanism should hold up to the melting point of c-Si and the speed follows an Arrhenius law as a function of the temperature we have searched, within the limit of our potential, the highest temperature we could simulate. In order to obtain this temperature, we observed at which temperature a vdW loop at zero pressure is present. This limited the temperature for the phase transition in the 1050$\approx$1100 K range. Knowing that the a-Si produced by our procedure is poorly relaxed this value should fall in the upper part of the experimentally measured range (1250-1450 K) of a-Si melting temperatures, and we can confidently assert that the major contribution to the difference is due to the quality of the DFT functional.

Another central aspect of SPE is the relaxation reached while producing a-Si, that in our case, is generated through a melting quenching technique with a 10 K/ps ramp. Geometrical metrics were used at first to compare a-Si obtained with our NN potential and with Tersoff potential. In addition to a good agreement with experimental density, this analysis showed that both potentials predict a medium-range order, but our potential lacks the presence of triplets of atoms with angles in the 60-80 degree range. The lack of these angles in a-Si is also confirmed by a realistic a-Si configuration obtained by a carefully crafted procedure in DFT. Since no geometrical difference is clear between our average a-Si structure and the reference one, we investigated the excess energy with respect to c-Si. This investigation showed that our average structure is not as

relaxed as the reference, and our samples display systematically higher excess energies that, to be lowered, would require much slower quenching rates or specifically crafted procedures.

Once we understood the capabilities of our potential we finally addressed the SPE simulation. In order to measure the SPE barriers, we simulated the reconstruction process at different temperatures in the 800-1100 K range and fit the Arrhenius law. The collected points support a two-barrier interpretation of the simulated phenomenon, a lower barrier for defects migration around 0.4 eV and a second experimentally measured one that characterizes the combined creation and migration of defects around 2.7 eV. The 0.4 eV barrier, not present in experiments with relaxed samples, is predominant in our low-temperature simulations because of the poor relaxation reached in our amorphous samples with the implemented thermal ramp. Furthermore, we noticed that the number of defects that are present in our samples should not allow for the investigation of the 2.7 eV barrier, but our analysis shows a picture of the phases involved during high-temperature SPE more complex than previously thought. In particular, at temperatures where the 2.7 eV barrier is measurable, we found the presence of highly diffusive a-Si (probably LDL) only a few layers apart from our interface where the reconstruction is happening. After showing that the removal of this artificially high diffusivity region recovers a reconstruction speed in line with the 0.4 eV barrier, we concluded that its presence must alter the exponential prefactors of the Arrhenius law, allowing the appearance of the events with the barrier of our interest.

Our work can now be extended in several directions.

On one hand, an improvement in the quality of a-Si by increasing its level of relaxation obtainable by designing and applying ad-hoc procedures will lower the prefactors in the Arrhenius law leading to slower but more realistic SPEs simulations. This can be performed alongside research for a better DFT functional capable of better describing key properties such as melting temperatures for c-Si and a-Si. These more realistic simulations can then be used to investigate new and more accurate microscopical events and elucidate the relevance and existence of the LDL phase.

On the other hand, one can move forward and, assuming the accuracy of the present simulations to be already sufficient, one can proceed in the analysis of the MD simulations to identify a library of local environments involved in reconstruction events and estimate the corresponding transition probabilities to be used in a KMC approach in order to address the challenge of the longer time scales involved in a realistic SPE simulation.

Appendix A

Melting temperature

The melting temperature of a material described by a potential can be calculated numerically either by simulating the melting process through molecular dynamics or by more rigorous - but complicated to apply - free energy methods [156]. The interface-based method[112, 113] adopted here is part of the first category. This method is based on the observation that a system whose simulation cell contains both the liquid and the crystalline phases will evolve toward the equilibrium melting point. This approach has limitations when dealing with complex systems with slow dynamics but we don't think this to be our case and, as it will be shown, the accuracy obtained is higher than, for instance the error due to the choice of exchange-correlation functional in the potential generation.

To measure the melting temperature in crystalline silicon the procedure starts from a solid diamond bulk cell with PBC where the size in one direction (for simplicity z) is larger than the other two. The minimal cell size for the shorter directions can be defined based on the potential cutoff radius, and later incremented to investigate size effects. The longer direction instead must be tuned to allow the system to contain enough l-Si and c-Si material to avoid the interfaces separating the two to interact with each other. Further space along this direction must be provided to have enough crystal to melt or liquid to crystallize during the thermal equilibration. Once sizes are defined, using the NPT ensemble, a three-step process is applied. First, the velocities are initialized accordingly to a Maxwell Boltzmann distribution at a temperature known to be below the melting point. After an equilibration phase of 5 ps, the part of the cell meant to remain in the bulk structure (usually half of it) is frozen, and the velocities of the atoms meant to undergo the phase transition are scaled to a temperature close to the melting one. Then another 5 ps thermal ramp with a final temperature above the expected melting temperature is applied to the free atoms. This ramp must provide enough kinetic energy to escape the superheated crystal state and grant the fusion of the structure. Since in the NPT

ensemble used in this procedure the target pressure is always $0\,\mathrm{GPa}$ and l-Si has a higher density than c-Si the occurrence of phase transition can be easily detected by monitoring the expansion of the cell along the z-axis. The system is then brought back to a temperature close to the melting one by a final cooling ramp. An example of final configuration obtained with this procedure is reported in Fig. 41

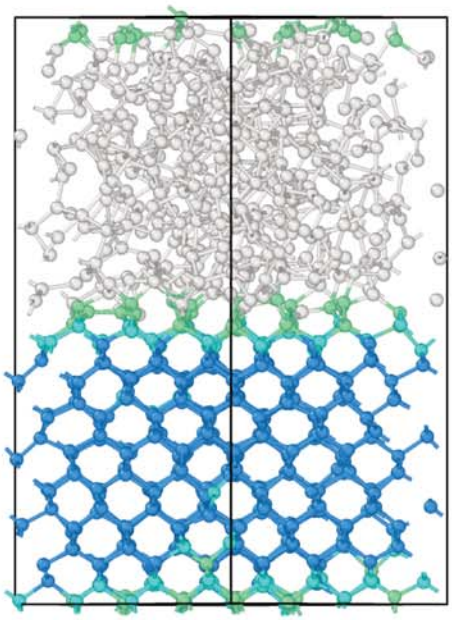


Figure 41: Example of a representative sample to study the melting temperature. Colors are obtained with the Maras *et al.* descriptor[142] and highlight the perfect diamond structure in blue, the first and second order coordinated diamond structure in light blue, and the only first neighbor tetrahedrally coordinated in green. White atoms are randomly arranged in a liquid phase. This particular configuration is a 4x4x8 cell.

The created cell can now be initialized with different temperatures below and above the expected melting value, and the system is equilibrated in an isobaric-isenthalpic (NPH) dynamics. The expected behavior is that when the temperature is slightly higher than the melting temperature some material should melt, and the temperature stabilizes at the melting value, due to the adsorption of the corresponding latent heat. When the temperature is lower than the melting one some material should freeze, and again the temperature stabilizes at the melting value due to release of latent heat. In the latter case however we saw that defects are easily embedded in the recrystallized part of the sample

creating more of an amorphous layer, hence the determination of the melting temperature approaching from below is less precise.

The final results obtained for the melting temperature of bulk silicon were 1194 $\pm$ 29 K for the PBESol and 1468 $\pm$11 K for the PBE derived potential as reported in the main text. We also examined the dependence of the calculated temperature on the cell size and results are reported in table 11. We can see that the estimated melting temperatures are rather stable with respect system size, the main difference being the larger statistical uncertainty in smaller systems due to more pronounced temperature oscillations during the simulation due to the limited amount of atoms present.

Table 11: Size effects in melting temperature.

atoms	replicas	mean [K]	std[K]
128	2x2x8	1465	107
256	2x2x16	1462	75
1024	4x4x8	1460	37
8192	8x8x16	1468	11

Fixing the melting temperature

We observed that our potential, based on PBE functional, underestimates the melting temperature by about 14%. Since we know that the melting temperature T_m is defined by the condition that the free energy of the liquid F_l is equal to the free energy of the crystal F_c, we can obtain the thermodynamic relations:

$$\begin{aligned} F_l &= F_c \\ U_l - T_m S_l &= U_c - T_m S_c \\ T_m &= \frac{\Delta U}{\Delta S} \end{aligned} \tag{A.1}$$

where $U_{l/c}$ is the internal energy and $S_{l/c}$ is the entropy of the two phases respectively.

Starting from these relations we can see whether there are simple ways to modify the potential so as to fix the melting temperature. Indeed if one scales the potential by a constant factor of 1.14 the atomic configurations visited during the dynamics remain unchanged if the time-step is scaled down by the corresponding factor $\sqrt{1.14}$. In the above thermodynamic relations the internal energy estimates would therefore be scaled, while the entropy estimates would remain unchanged thus giving the desired scaling of the melting temperature.

We did so and computed the melting temperature as described previously using a scaled potential and we determined a melting temperature of 1673 ± 40

K (this test was limited to the 1024 atom cell) confirming the soundness of the analysis.

This procedure, besides the conceptual flaw of not being ab-initio, has however serious consequences on all the other physical properties of the material: the elastic constants of the bulk are made stiffer by a factor 1.14, similarly the phonon frequencies are scaled by a factor $\sqrt{1.14}$ breaking the obtained good agreement with the DFT results. More critically, defect formation and migration energies would be similarly affected.

Appendix B

v-DOS

Acoustic modes

We know that a-Si, as all glassy systems, is characterized by an excess of specific heat in the range between 15-30K [157, 158]. This property has been linked to an excess of states, wrt the Debye theory, in the low spectrum region of the Raman spectra and consequently in the v-Dos in the region between 2 and 15 meV. Commonly refereed as the boson peak, this anomaly has been found by some authors to be due to the existence of localized soft modes [159, 155].

To test the presence of the boson peak and the capacity of our potential in reproducing it comes in handy in two ways. First because it can be used to determine if the atoms on the top part of our SPE sample are indeed in a glassy state. Such comparison can be done by computing the phonons of those atoms using their self correlation velocity as for example is done in [14]. Second atoms participation ratio[153, 152] in those soft modes can be used as a proxy to locate regions that can undergo irreversible changes that are expected to sensibly effect dynamics [160]. Furthermore such regions can be taken as first candidate to investigate glassy dynamics and to search for defects.

We have tested our 216 atoms models and found that all of them show such anomaly as reported in fig. 42 that present an ω_0 at approximately 12meV. Value that must be compared with experimental expectations that in the range 14.13-30.37 meV in the Raman spectrum[161, 162]. The same pattern is also visible even if less marked in the Tersoff samples. SW is also able to show it but larger cells are needed because of it very low frequency [163, 164].

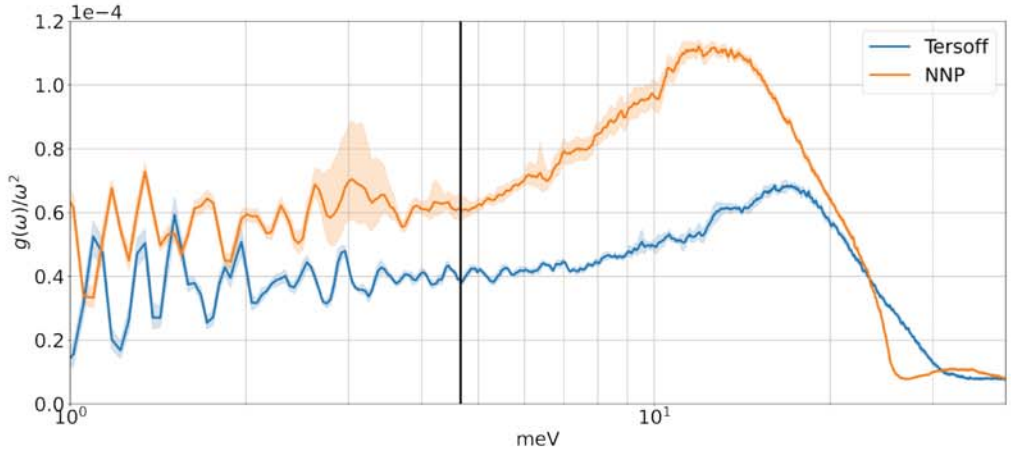


Figure 42: $\frac{g(\omega)}{\omega^2}$ in the low frequency spectrum for Panna and Tersoff potentials. Extracted form figure 30. The black line is an estimate of the lower frequency that can fit in the simulated cell. Points far below such line can be regarded as artifacts. Oscillation are due to the number of poitns in the reciprocal space used when evaluating the v-DOS. Such sampling has been performed on a regular mesh grid, and as a consequence at low frequency where the modes have privileged directions are poorly samapled. Samples generated with the NN potential follow the expected boson peak pattern. They start flat (4-5 meV and probably below) and then increas up to $\approx$ 12 meV. Tersoff has also a similar pattern.

Optic modes and $\frac{\Gamma}{2}$

In the upper part of the spectrum the main quantity of interest is the Raman TO-peak half-width $\Gamma/2$. The interest behind this quantity is given by its hypothesized linear correlation with the sample angular distortion[165, 166, 150], and the possibility to find a lower and an upper bound to the number of defects[11]. In a preliminary investigation we decided to first verify the linear correlation between $\Gamma/2$ and the angular distortion.

The computation of $\Gamma/2$ requires the Raman spectrum, a quantity that at the moment we don't have. However, accepting some approximations we can never the less have a good insight. Since our attention now is on a very limited part of the spectrum that has a Gaussian decay, it is not unreasonable to assume that the interaction between the light and the vibrational modes as a constant value. With this assumption we can simply convert the v-DOS to the Raman spectrum by dividing with a $1/\omega$ factor [167]. It is then possible to inquire the existence of the hypothesized relation. Fig. 43 shows that, with our crude approximations such relation does not seem to hold.

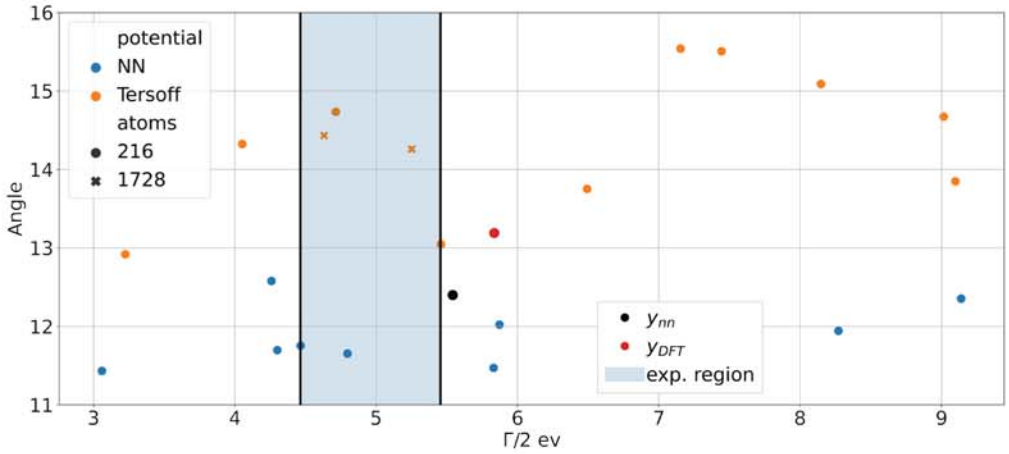


Figure 43: Distribution of the angle as a function of $\Gamma/2$, black lines are experimental region where $\Gamma/2$ is expected to live.

Bibliography

[1] P. Batude et al. “Advances, challenges and opportunities in 3D CMOS sequential integration”. In: *2011 International Electron Devices Meeting*. IEEE, Dec. 2011. DOI: 10.1109/iedm.2011.6131506.

[2] Perrine Batude et al. “Enabling 3D Monolithic Integration”. In: *ECS Trans.* 16.8 (Oct. 2008), pp. 47–54. DOI: 10.1149/1.2982853.

[3] Laurent Brunet et al. “(Invited) 3D Monolithic Integration”. In: *ECS Trans.* 85.8 (Apr. 2018), pp. 125–130. DOI: 10.1149/08508.0125ecst.

[4] G.L. Olson and J.A. Roth. “Kinetics of solid phase crystallization in amorphous silicon”. In: *Mater. Sci. Rep.* 3.1 (Jan. 1988), pp. 1–77. DOI: 10.1016/s0920-2307(88)80005-7.

[5] S. Roorda et al. “Structural relaxation and defect annihilation in pure amorphous silicon”. In: *Phys. Rev. B* 44.8 (Aug. 1991), pp. 3702–3725. DOI: 10.1103/physrevb.44.3702.

[6] Francesco Priolo and Emanuele Rimini. “Ion-beam-induced epitaxial crystallization and amorphization in silicon”. In: *Mater. Sci. Rep.* 5.6 (Jan. 1990), pp. 319–379. DOI: 10.1016/0920-2307(90)90001-j.

[7] Brett C. Johnson, Jeffrey C. McCallum, and Michael J. Aziz. “Solid-Phase Epitaxy”. In: *Handbook of Crystal Growth*. Elsevier, 2015, pp. 317–363. DOI: 10.1016/b978-0-444-63304-0.00007-x.

[8] M. R. J. Gibbs, J. E. Evetts, and J. A. Leake. “Activation energy spectra and relaxation in amorphous materials”. In: *J. Mater. Sci.* 18.1 (Jan. 1983), pp. 278–288. DOI: 10.1007/bf00543836.

[9] J.S. Im and Harry A. Atwater. “Kinetic and thermodynamic enhancement of crystal nucleation and growth rates in amorphous Si film during ion irradiation”. In: *Nucl. Instrum. Methods Phys. Res., Sect. B* 59-60 (July 1991), pp. 422–426. DOI: 10.1016/0168-583x(91)95251-8.

[10] Corrado Spinella, Salvatore Lombardo, and Francesco Priolo. “Crystal grain nucleation in amorphous silicon”. In: *J. Appl. Phys.* 84.10 (Nov. 1998), pp. 5383–5414. DOI: 10.1063/1.368873.

[11] P. A. Stolk et al. “Contribution of defects to electronic, structural, and thermodynamic properties of amorphous silicon”. In: *J. Appl. Phys.* 75.11 (June 1994), pp. 7266–7286. DOI: 10.1063/1.356662.

[12] E. P. Donovan et al. “Heat of crystallization and melting point of amorphous silicon”. In: *Appl. Phys. Lett.* 42.8 (Apr. 1983), pp. 698–700. DOI: 10.1063/1.94077.

[13] S. Roorda et al. “Point defect populations in amorphous and crystalline silicon”. In: *Solid State Commun.* 75.3 (July 1990), pp. 197–200. DOI: 10.1016/0038-1098(90)90268-g.

[14] Srikanth Sastry and C. Austen Angell. “Liquid–liquid phase transition in supercooled silicon”. In: *Nat. Mater.* 2.11 (Oct. 2003), pp. 739–743. DOI: 10.1038/nmat994.

[15] P. Ganesh and M. Widom. “Liquid-Liquid Transition in Supercooled Silicon Determined by First-Principles Simulation”. In: *Phys. Rev. Lett.* 102.7 (Feb. 2009). DOI: 10.1103/physrevlett.102.075701.

[16] E. P. Donovan et al. “Calorimetric studies of crystallization and relaxation of amorphous Si and Ge prepared by ion implantation”. In: *J. Appl. Phys.* 57.6 (Mar. 1985), pp. 1795–1804. DOI: 10.1063/1.334406.

[17] L. Csepregi et al. “Substrate-orientation dependence of the epitaxial regrowth rate from Si-implanted amorphous Si”. In: *J. Appl. Phys.* 49.7 (July 1978), pp. 3906–3911. DOI: 10.1063/1.325397.

[18] B. Drosd and J. Washburn. “A new technique for observing the amorphous to crystalline transformation in thin surface layers on silicon wafers”. In: *J. Appl. Phys.* 51.8 (Aug. 1980), pp. 4106–4110. DOI: 10.1063/1.328220.

[19] G.L. Olson et al. “Laser-Induced Solid Phase Crystallization in Amorphous Silicon Films”. In: *MRS Proc.* 13 (1982). DOI: 10.1557/proc-13-141.

[20] J. S. Williams and R. G. Elliman. “Role of Electronic Processes in Epitaxial Recrystallization of Amorphous Semiconductors”. In: *Phys. Rev. Lett.* 51.12 (Sept. 1983), pp. 1069–1072. DOI: 10.1103/physrevlett.51.1069.

[21] A. Lietoila et al. “Epitaxial regrowth of intrinsic, 31P-doped and compensated (31P+11B-doped) amorphous Si”. In: *J. Appl. Phys.* 53.6 (June 1982), pp. 4399–4405. DOI: 10.1063/1.331222.

[22] Guo-Quan Lu, Eric Nygren, and Michael J. Aziz. “Pressure-enhanced crystallization kinetics of amorphous Si and Ge: Implications for point-defect mechanisms”. In: *J. Appl. Phys.* 70.10 (Nov. 1991), pp. 5323–5345. DOI: 10.1063/1.350243.

[23] J. A. Roth et al. “The Effect of Hydrogen on the Kinetics of Solid Phase Epitaxy in Amorphous Silicon”. In: *MRS Proc.* 205 (1990). DOI: 10.1557/proc-205-45.

[24] Michael J. Aziz, Paul C. Sabin, and Guo-Quan Lu. “The activation strain tensor: Nonhydrostatic stress effects on crystal-growth kinetics”. In: *Phys. Rev. B* 44.18 (Nov. 1991), pp. 9812–9816. DOI: 10.1103/physrevb.44.9812.

[25] Frans Spaepen. “A structural model for the interface between amorphous and crystalline Si or Ge”. In: *Acta Metall. Mater.* 26.7 (July 1978), pp. 1167–1177. DOI: 10.1016/0001-6160(78)90145-1.

[26] J.S. Williams and J.M. Poate. “Introduction to Implantation and Beam Processing”. In: *Ion Implantation and Beam Processing*. Elsevier, 1984, pp. 1–11. DOI: 10.1016/b978-0-12-756980-2.50005-9.

[27] R. Drosd and J. Washburn. “Some observations on the amorphous to crystalline transformation in silicon”. In: *J. Appl. Phys.* 53.1 (Jan. 1982), pp. 397–403. DOI: 10.1063/1.329901.

[28] Synopsys. *Sentaurus Process User Guide*. O-2018.06. Synopsys. June 2018.

[29] Ignacio Martin-Bragado and Victor Moroz. “Facet formation during solid phase epitaxy regrowth: A lattice kinetic Monte Carlo model”. In: *Appl. Phys. Lett.* 95.12 (Sept. 2009), p. 123123. DOI: 10.1063/1.3236535.

[30] Noam Bernstein, Michael J. Aziz, and Efthimios Kaxiras. “Atomistic simulations of solid-phase epitaxial growth in silicon”. In: *Phys. Rev. B* 61.10 (Mar. 2000), pp. 6696–6700. DOI: 10.1103/physrevb.61.6696.

[31] T. Motooka et al. “Molecular-dynamics simulations of solid-phase epitaxy of Si: Growth mechanisms”. In: *Phys. Rev. B* 61.12 (Mar. 2000), pp. 8537–8540. DOI: 10.1103/physrevb.61.8537.

[32] K. Gärtner and B. Weber. “Molecular dynamics simulations of solid-phase epitaxial growth in silicon”. In: *Nucl. Instrum. Methods Phys. Res., Sect. B* 202 (Apr. 2003), pp. 255–260. DOI: 10.1016/s0168-583x(02)01866-9.

[33] A. Mattoni and L. Colombo. "Boron ripening during solid-phase epitaxy of amorphous silicon". In: *Phys. Rev. B* 69.4 (Jan. 2004). DOI: 10.1103/physrevb.69.045204.

[34] C. Krzeminski et al. "Molecular dynamics simulation of the recrystallization of amorphous Si layers: Comprehensive study of the dependence of the recrystallization velocity on the interatomic potential". In: *J. Appl. Phys.* 101.12 (June 2007), p. 123506. DOI: 10.1063/1.2743089.

[35] E. Lampin and C. Krzeminski. "Molecular dynamics simulations of the solid phase epitaxy of Si: Growth mechanism and orientation effects". In: *J. Appl. Phys.* 106.6 (Sept. 2009), p. 063519. DOI: 10.1063/1.3211972.

[36] B.A. Gillespie and H.N.G. Wadley. "Atomistic examinations of the solid-phase epitaxial growth of silicon". In: *J. Cryst. Growth* 311.11 (May 2009), pp. 3195–3203. DOI: 10.1016/j.jcrysgro.2009.02.050.

[37] M. Posselt, H. Bracht, and D. Radić. "Atomistic simulations on the relationship between solid-phase epitaxial recrystallization and self-diffusion in amorphous silicon". In: *J. Appl. Phys.* 131.3 (Jan. 2022), p. 035102. DOI: 10.1063/5.0078015.

[38] J. Tersoff. "Empirical interatomic potential for silicon with improved elastic properties". In: *Phys. Rev. B* 38.14 (Nov. 1988), pp. 9902–9905. DOI: 10.1103/physrevb.38.9902.

[39] J. F. Justo et al. "Interatomic Potential for Condensed Phases and Bulk Defects in Silicon". In: *MRS Proc.* 469 (1997). DOI: 10.1557/proc-469-217.

[40] Martin Z. Bazant, Efthimios Kaxiras, and J. F. Justo. "Environment-dependent interatomic potential for bulk silicon". In: *Phys. Rev. B* 56.14 (Oct. 1997), pp. 8542–8552. DOI: 10.1103/physrevb.56.8542.

[41] 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.

[42] Erik J. Albenze and Paulette Clancy. "Interface Response Functions for Amorphous and Crystalline Si and the Implications for Explosive Crystallization". In: *Mol. Simulat.* 31.1 (Jan. 2005), pp. 11–24. DOI: 10.1080/08927020412331298658.

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

[44] Richard M. Martin. *Electronic Structure. Basic Theory and Practical Methods*. Cambridge University Press, Apr. 2004. DOI: 10.1017/cbo9780511805769.

[45] D.M. Ceperley and B.J. Alder. *Ground state of the electron gas by a stochastic method*. Tech. rep. Office of Scientific and Technical Information (OSTI), May 1980. DOI: 10.2172/5208790.

[46] John P. Perdew, Kieron Burke, and Yue Wang. "Generalized gradient approximation for the exchange-correlation hole of a many-electron system". In: *Phys. Rev. B* 54.23 (Dec. 1996), pp. 16533–16539. DOI: 10.1103/physrevb.54.16533.

[47] Claudia Filippi, Xavier Gonze, and C.J. Umrigar. "Generalized gradient approximations to density functional theory: Comparison with exact results". In: *Theoretical and Computational Chemistry*. Elsevier, 1996, pp. 295–326. DOI: 10.1016/s1380-7323(96)80090-2.

[48] John P. Perdew et al. "Restoring the Density-Gradient Expansion for Exchange in Solids and Surfaces". In: *Phys. Rev. Lett.* 100.13 (Apr. 2008). DOI: 10.1103/physrevlett.100.136406.

[49] M. Methfessel and A. T. Paxton. "High-precision sampling for {Brillouin}-zone integration in metals". In: *Phys. Rev. B* 40.6 (Aug. 1989), pp. 3616–3621. DOI: 10.1103/physrevb.40.3616.

[50] Nicola Marzari et al. "Thermal Contraction and Disordering of the Al(110) Surface". In: *Phys. Rev. Lett.* 82.16 (Apr. 1999), pp. 3296–3299. DOI: 10.1103/physrevlett.82.3296.

[51] R. Car and M. Parrinello. "Unified Approach for Molecular Dynamics and Density-Functional Theory". In: *Phys. Rev. Lett.* 55.22 (Nov. 1985), pp. 2471–2474. DOI: 10.1103/physrevlett.55.2471.

[52] J. Tersoff. "New empirical model for the structural properties of silicon". In: *Phys. Rev. Lett.* 56.6 (Feb. 1986), pp. 632–635. DOI: 10.1103/physrevlett.56.632.

[53] J. Tersoff. "New empirical approach for the structure and energy of covalent systems". In: *Phys. Rev. B* 37.12 (Apr. 1988), pp. 6991–7000. DOI: 10.1103/physrevb.37.6991.

[54] J. Tersoff. "Modeling solid-state chemistry: Interatomic potentials for multicomponent systems". In: *Phys. Rev. B* 39.8 (Mar. 1989), pp. 5566–5568. DOI: 10.1103/physrevb.39.5566.

[55] Philip M. Morse. “Diatomic Molecules According to the Wave Mechanics. II. Vibrational Levels”. In: *Physical Review* 34.1 (July 1929), pp. 57–64. DOI: 10.1103/physrev.34.57.

[56] Loup Verlet. “Computer ”Experiments” on Classical Fluids. I. Thermodynamical Properties of Lennard-Jones Molecules”. In: 159.1 (July 1967), pp. 98–103. DOI: 10.1103/physrev.159.98.

[57] Steve Plimpton. “Fast Parallel Algorithms for Short-Range Molecular Dynamics”. In: *J. Comput. Phys.* 117.1 (Mar. 1995), pp. 1–19. DOI: 10.1006/jcph.1995.1039.

[58] David J. C. MacKay. *Information Theory, Inference and Learning Algorithms*. Copyright Cambridge University Press, 2003.

[59] Jörg Behler and Michele Parrinello. “Generalized Neural-Network Representation of High-Dimensional Potential-Energy Surfaces”. In: *Phys. Rev. Lett.* 98.14 (Apr. 2007). DOI: 10.1103/physrevlett.98.146401.

[60] Jörg Behler. “Neural network potential-energy surfaces in chemistry: A tool for large-scale simulations”. In: *Phys. Chem. Chem. Phys.* 13.40 (2011), p. 17930. DOI: 10.1039/c1cp21668f.

[61] 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.

[62] Albert P. Bartók, Risi Kondor, and Gábor Csányi. “On representing chemical environments”. In: *Phys. Rev. B* 87.18 (May 2013). DOI: 10.1103/physrevb.87.184115.

[63] Ralf Drautz. “Atomic cluster expansion for accurate and transferable interatomic potentials”. In: *Phys. Rev. B* 99.1 (Jan. 2019). DOI: 10.1103/physrevb.99.014104.

[64] Ruggero Lot et al. “PANNA: Properties from Artificial Neural Network Architectures”. In: *Comput. Phys. Commun.* 256 (Nov. 2020), p. 107402. DOI: 10.1016/j.cpc.2020.107402.

[65] Warren S. McCulloch and Walter Pitts. “A logical calculus of the ideas immanent in nervous activity”. In: *B. Math. Biol.* 52.1-2 (Jan. 1990), pp. 99–115. DOI: 10.1007/bf02459570.

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

[67] Kingma et al. “Adam: A Method for Stochastic Optimization”. In: *CoRR* (2014). arXiv: 1412.6980 [cs.LG].

[68] Martin Abadi et al. *TensorFlow: Large-scale Machine Learning on Heterogeneous Systems*. Software available from tensorflow.org. 2015.

[69] Charles R. Harris et al. "Array programming with NumPy". In: *Nature* 585.7825 (Sept. 2020), pp. 357–362. DOI: 10.1038/s41586-020-2649-2.

[70] The Pandas Development Team. *pandas-dev/pandas: Pandas*. 2022. DOI: 10.5281/ZENODO.3509134.

[71] John D. Hunter. "Matplotlib: A 2D Graphics Environment". In: *Computing in Science & Engineering* 9.3 (2007), pp. 90–95. DOI: 10.1109/mcse.2007.55.

[72] Michael Waskom. "seaborn: Statistical data visualization". In: *Journal of Open Source Software* 6.60 (Apr. 2021), p. 3021. DOI: 10.21105/joss.03021.

[73] Aidan P. Thompson et al. "LAMMPS - a flexible simulation tool for particle-based materials modeling at the atomic, meso, and continuum scales". In: *Comput. Phys. Commun.* 271 (Feb. 2022), p. 108171. DOI: 10.1016/j.cpc.2021.108171.

[74] Alexander Stukowski. "Visualization and analysis of atomistic simulation data with OVITO–the Open Visualization Tool". In: *Model. Simul. Mater. Sc.* 18.1 (Dec. 2009), p. 015012. DOI: 10.1088/0965-0393/18/1/015012.

[75] Paolo Giannozzi et al. "Quantum ESPRESSO toward the exascale". In: *J. Chem. Phys.* 152.15 (Apr. 2020), p. 154105. DOI: 10.1063/5.0005082.

[76] Yusuf Shaidu et al. "A systematic approach to generating accurate neural network potentials: The case of carbon". In: *npj Comput. Mater.* 7.1 (Apr. 2021). DOI: 10.1038/s41524-021-00508-6.

[77] Nongnuch Artrith and Alexander Urban. "An implementation of artificial neural-network potentials for atomistic materials simulations: Performance for TiO2". In: *Nato. Sc. S. Ss. Iii. C. S.* 114 (Mar. 2016), pp. 135–150. DOI: 10.1016/j.commatsci.2015.11.047.

[78] Fedwa El-Mellouhi, Normand Mousseau, and Laurent J. Lewis. "Kinetic activation-relaxation technique: An off-lattice self-learning kinetic Monte Carlo algorithm". In: *Physical Review B* 78.15 (Oct. 2008). DOI: 10.1103/physrevb.78.153202.

[79] Graeme Henkelman, Blas P. Uberuaga, and Hannes Jónsson. "A climbing image nudged elastic band method for finding saddle points and minimum energy paths". In: *The Journal of Chemical Physics* 113.22 (Dec. 2000), pp. 9901–9904. DOI: 10.1063/1.1329672.

[80] Albert P. Bartók et al. "Machine Learning a General-Purpose Interatomic Potential for Silicon". In: *Phys. Rev. X* 8.4 (Dec. 2018). DOI: 10.1103/physrevx.8.041048.

[81] John P. Perdew, Kieron Burke, and Matthias Ernzerhof. "Generalized Gradient Approximation Made Simple". In: *Phys. Rev. Lett.* 77.18 (Oct. 1996), pp. 3865–3868. DOI: 10.1103/physrevlett.77.3865.

[82] John P. Perdew, Kieron Burke, and Matthias Ernzerhof. "Generalized Gradient Approximation Made Simple [Phys. Rev. Lett. 77, 3865 (1996)]". In: *Phys. Rev. Lett.* 78.7 (Feb. 1997), pp. 1396–1396. DOI: 10.1103/physrevlett.78.1396.

[83] Paolo Giannozzi et al. "Ab initio calculation of phonon dispersions in semiconductors". In: *Phys. Rev. B* 43.9 (Mar. 1991), pp. 7231–7242. DOI: 10.1103/physrevb.43.7231.

[84] Osamu Sugino and Roberto Car. "Ab Initio Molecular Dynamics Study of First-Order Phase Transitions: Melting of Silicon". In: *Phys. Rev. Lett.* 74.10 (Mar. 1995), pp. 1823–1826. DOI: 10.1103/physrevlett.74.1823.

[85] Soohaeng Yoo, Sotiris S. Xantheas, and Xiao Cheng Zeng. "The melting temperature of bulk silicon from ab initio molecular dynamics simulations". In: *Chem. Phys. Lett.* 481.1-3 (Oct. 2009), pp. 88–90. DOI: 10.1016/j.cplett.2009.09.075.

[86] Andreas Pedersen, Laurent Pizzagalli, and Hannes Jónsson. "Optimal atomic structure of amorphous silicon obtained from density functional theory calculations". In: *New J. Phys.* 19.6 (June 2017), p. 063018. DOI: 10.1088/1367-2630/aa732e.

[87] C.R. Miranda et al. "Vacancy-like defects in a-Si: A first principles study". In: *J. Non-Cryst. Solids* 338-340 (June 2004), pp. 400–402. DOI: 10.1016/j.jnoncrysol.2004.03.006.

[88] Peter Scherpelz and Giulia Galli. "Optimizing surface defects for atomic-scale electronics: Si dangling bonds". In: *Phys. Rev. Materials* 1.2 (July 2017). DOI: 10.1103/physrevmaterials.1.021602.

[89] Bert Voigtländer. “Fundamental processes in Si/Si and Ge/Si epitaxy studied by scanning tunneling microscopy during growth”. In: *Surf. Sci. Rep.* 43.5-8 (Sept. 2001), pp. 127–254. DOI: 10.1016/s0167-5729(01)00012-7.

[90] Y.-W. Mo et al. “Growth and equilibrium structures in the epitaxy of Si on Si(001)”. In: *Phys. Rev. Lett.* 63.21 (Nov. 1989), pp. 2393–2396. DOI: 10.1103/physrevlett.63.2393.

[91] Cong Li et al. “A comprehensive study of the high-pressure–temperature phase diagram of silicon”. In: *J. Mater. Sci.* 53.10 (Feb. 2018), pp. 7475–7485. DOI: 10.1007/s10853-018-2087-9.

[92] Dipak Prasad and Nilanjan Mitra. “An atomistic study of phase transition in cubic diamond Si single crystal subjected to static compression”. In: *Nato. Sc. S. Ss. Iii. C. S.* 156 (Jan. 2019), pp. 232–240. DOI: 10.1016/j.commatsci.2018.09.037.

[93] F. Wooten, K. Winer, and D. Weaire. “Computer Generation of Structural Models of Amorphous Si and Ge”. In: *Phys. Rev. Lett.* 54.13 (Apr. 1985), pp. 1392–1395. DOI: 10.1103/physrevlett.54.1392.

[94] F. Wooten and D. Weaire. “Generation of random network models with periodic boundary conditions”. In: *J. Non-Cryst. Solids* 64.3 (May 1984), pp. 325–334. DOI: 10.1016/0022-3093(84)90186-8.

[95] D.E. Polk. “Structural model for amorphous silicon and germanium”. In: *J. Non-Cryst. Solids* 5.5 (May 1971), pp. 365–376. DOI: 10.1016/0022-3093(71)90038-x.

[96] Richard Zallen. *The Physics of Amorphous Solids*. Wiley-VCH Verlag GmbH & Co. KGaA, Nov. 1983. DOI: 10.1002/3527602798.

[97] J. Q. Broughton and X. P. Li. “Phase diagram of silicon by molecular dynamics”. In: *Phys. Rev. B* 35.17 (June 1987), pp. 9120–9127. DOI: 10.1103/physrevb.35.9120.

[98] W. D. Luedtke and Uzi Landman. “Preparation and melting of amorphous silicon by molecular-dynamics simulations”. In: *Phys. Rev. B* 37.9 (Mar. 1988), pp. 4656–4663. DOI: 10.1103/physrevb.37.4656.

[99] Stephen J. Cook and Paulette Clancy. “Comparison of semi-empirical potential functions for silicon and germanium”. In: *Phys. Rev. B* 47.13 (Apr. 1993), pp. 7686–7699. DOI: 10.1103/physrevb.47.7686.

[100] Manabu Ishimaru, Shinji Munetoh, and Teruaki Motooka. "Generation of amorphous silicon structures by rapid quenching: A molecular-dynamics study". In: *Phys. Rev. B* 56.23 (Dec. 1997), pp. 15133–15138. DOI: 10.1103/physrevb.56.15133.

[101] A. Ramstad, G. Brocks, and P. J. Kelly. "Theoretical study of the Si(100) surface reconstruction". In: *Phys. Rev. B* 51.20 (May 1995), pp. 14504–14523. DOI: 10.1103/physrevb.51.14504.

[102] G. P. Purja Pun and Y. Mishin. "Optimized interatomic potential for silicon and its application to thermal stability of silicene". In: *Phys. Rev. B* 95.22 (June 2017). DOI: 10.1103/physrevb.95.224103.

[103] Thomas J Lenosky et al. "Highly optimized empirical potential model of silicon". In: *Model. Simul. Mater. Sc.* 8.6 (Oct. 2000), pp. 825–841. DOI: 10.1088/0965-0393/8/6/305.

[104] Anant D. Kulkarni et al. "Oxygen Interactions with Silica Surfaces: Coupled Cluster and Density Functional Investigation and the Development of a New ReaxFF Potential". In: *The Journal of Physical Chemistry C* 117.1 (Dec. 2012), pp. 258–269. DOI: 10.1021/jp3086649.

[105] K. Hata, Y. Sainoo, and H. Shigekawa. "Atomically Resolved Local Variation of the Barrier Height of the Flip-Flop Motion of Single Buckled Dimers of Si(100)". In: *Phys. Rev. Lett.* 86.14 (Apr. 2001), pp. 3084–3087. DOI: 10.1103/physrevlett.86.3084.

[106] Mario Valle and Artem R. Oganov. "Crystal fingerprint space – a novel paradigm for studying crystal-structure sets". In: *Acta Crystallographica Section A Foundations of Crystallography* 66.5 (Aug. 2010), pp. 507–517. DOI: 10.1107/s0108767310026395.

[107] Heike Hofmann, Karen Kafadar, and Hadley Wickham. *Letter-value plots: Boxplots for large data*. Tech. rep. had.co.nz, 2011.

[108] 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). DOI: 10.1103/physrevb.85.045439.

[109] Matthew A. Hopcroft, William D. Nix, and Thomas W. Kenny. "What is the Young's Modulus of Silicon?" In: *J. Microelectromech. S.* 19.2 (Apr. 2010), pp. 229–238. DOI: 10.1109/jmems.2009.2039697.

[110] M. Råsander and M. A. Moram. "On the accuracy of commonly used density functional approximations in determining the elastic constants of insulators and semiconductors". In: *The Journal of Chemical Physics* 143.14 (Oct. 2015), p. 144104. DOI: 10.1063/1.4932334.

[111] John J. Hall. "Electronic Effects in the Elastic Constants of n-Type Silicon". In: *Phys. Rev.* 161.3 (Sept. 1967), pp. 756–761. DOI: 10.1103/physrev.161.756.

[112] J. R. Morris et al. "Melting line of aluminum from simulations of coexisting phases". In: *Phys. Rev. B* 49.5 (Feb. 1994), pp. 3109–3115. DOI: 10.1103/physrevb.49.3109.

[113] James R. Morris and Xueyu Song. "The melting lines of model systems calculated from coexistence simulations". In: *J. Chem. Phys.* 116.21 (June 2002), pp. 9352–9358. DOI: 10.1063/1.1474581.

[114] Ryosuke Jinnouchi, Ferenc Karsai, and Georg Kresse. "On-the-fly machine learning force field generation: Application to melting points". In: *Phys. Rev. B* 100.1 (July 2019). DOI: 10.1103/physrevb.100.014105.

[115] Khalid Laaziri et al. "High-energy x-ray diffraction study of pure amorphous silicon". In: *Phys. Rev. B* 60.19 (Nov. 1999), pp. 13520–13533. DOI: 10.1103/physrevb.60.13520.

[116] Khalid Laaziri et al. "High Resolution Radial Distribution Function of Pure Amorphous Silicon". In: *Phys. Rev. Lett.* 82.17 (Apr. 1999), pp. 3460–3463. DOI: 10.1103/physrevlett.82.3460.

[117] M. G. Grimaldi, P. Baeri, and M. A. Malvezzi. "Melting temperature of unrelaxed amorphous silicon". In: *Phys. Rev. B* 44.4 (July 1991), pp. 1546–1553. DOI: 10.1103/physrevb.44.1546.

[118] W. Sinke et al. "Transient structural relaxation of amorphous silicon". In: *J. Non-Cryst. Solids* 99.2-3 (Feb. 1988), pp. 308–323. DOI: 10.1016/0022-3093(88)90439-5.

[119] Robert H. Swendsen and Jian-Sheng Wang. "Replica Monte Carlo Simulation of Spin-Glasses". In: *Phys. Rev. Lett.* 57.21 (Nov. 1986), pp. 2607–2609. DOI: 10.1103/physrevlett.57.2607.

[120] G. Zhao, Y. J. Yu, and X. M. Tan. "Nature of the first-order liquid-liquid phase transition in supercooled silicon". In: *J. Chem. Phys.* 143.5 (Aug. 2015), p. 054508. DOI: 10.1063/1.4928194.

[121] André Hedler, Siegfried Ludwig Klaumünzer, and Werner Wesch. "Amorphous silicon exhibits a glass transition". In: *Nat. Mater.* 3.11 (Oct. 2004), pp. 804–809. DOI: 10.1038/nmat1241.

[122] Philippe Beaucage and Normand Mousseau. "Liquid–liquid phase transition in Stillinger–Weber silicon". In: *J. Phys. Condens. Matter* 17.15 (Apr. 2005), pp. 2269–2279. DOI: 10.1088/0953-8984/17/15/002.

[123] S. De Gironcol M. Peressi M. Fornari. “Coordination defects in amorphous silicon and hydrogenated amorphous silicon: A characterization from first-principles calculations”. In: *Philos. Mag. B* 80.4 (Apr. 2000), pp. 515–521. DOI: 10.1080/014186300255168.

[124] David Adler. “Density of States in the Gap of Tetrahedrally Bonded Amorphous Semiconductors”. In: *Phys. Rev. Lett.* 41.25 (Dec. 1978), pp. 1755–1758. DOI: 10.1103/physrevlett.41.1755.

[125] Sébastien Le Roux and Philippe Jund. “Ring statistics analysis of topological networks: New approach and application to amorphous GeS2 and SiO2 systems”. In: *Nato. Sc. S. Ss. Iii. C. S.* 49.1 (June 2010), pp. 70–83. DOI: 10.1016/j.commatsci.2010.04.023.

[126] Lester Guttman. “Ring structure of the crystalline and amorphous forms of silicon dioxide”. In: *J. Non-Cryst. Solids* 116.2-3 (Feb. 1990), pp. 145–147. DOI: 10.1016/0022-3093(90)90686-g.

[127] M. M. J. Treacy and K. B. Borisenko. “The Local Structure of Amorphous Silicon”. In: *Science* 335.6071 (Feb. 2012), pp. 950–953. DOI: 10.1126/science.1214780.

[128] F. Kail et al. “The configurational energy gap between amorphous and crystalline silicon”. In: *physica status solidi (RRL) - Rapid Research Letters* 5.10-11 (Sept. 2011), pp. 361–363. DOI: 10.1002/pssr.201105333.

[129] “Structural Properties of Amorphous Silicon and Amorphous Chalcogenides”. In: *Amorphous Semiconductors*. John Wiley & Sons, Ltd, Dec. 2016. Chap. Structural Properties of Amorphous Silicon and Amorphous Chalcogenides, pp. 11–42. DOI: 10.1002/9781118758236.ch3.

[130] Frank H. Allen, Olga Kennard, and Robin Taylor. “Systematic analysis of structural data as a research technique in organic chemistry”. In: *Accounts Chem. Res.* 16.5 (May 1983), pp. 146–153. DOI: 10.1021/ar00089a001.

[131] S. Kugler and Z. Várallyay. “Possible unusual atomic arrangements in the structure of amorphous silicon”. In: *Phil. Mag. Lett.* 81.8 (Aug. 2001), pp. 569–574. DOI: 10.1080/09500830110055355.

[132] Shirley V. King. “Ring Configurations in a Random Network Model of Vitreous Silica”. In: *Nature* 213.5081 (Mar. 1967), pp. 1112–1113. DOI: 10.1038/2131112a0.

[133] D. A. Drabold. “Silicon: The gulf between crystalline and amorphous”. In: *physica status solidi (RRL) - Rapid Research Letters* 5.10-11 (Oct. 2011), pp. 359–360. DOI: 10.1002/pssr.201105444.

[134] J. S. Custer et al. “Density of amorphous Si”. In: *Appl. Phys. Lett.* 64.4 (Jan. 1994), pp. 437–439. DOI: 10.1063/1.111121.

[135] K. Laaziri, S. Roorda, and J.M Baribeau. “Density of amorphous SixGe1−x alloys prepared by high-energy ion implantation”. In: *J. Non-Cryst. Solids* 191.1-2 (Nov. 1995), pp. 193–199. DOI: 10.1016/0022-3093(95)00310-x.

[136] Dario Alfè. “PHON: A program to calculate phonons using the small displacement method”. In: *Comput. Phys. Commun.* 180.12 (Dec. 2009), pp. 2622–2633. DOI: 10.1016/j.cpc.2009.03.010.

[137] R.J. Nemanich et al. “Raman scattering from microcrystalline Si films: Considerations of composite structures with different optical absorption properties”. In: *J. Non-Cryst. Solids* 114 (Dec. 1989). Proceedings of the Thirteenth International Conference on Amorphous and Liquid Semiconductors, pp. 813–815. DOI: 10.1016/0022-3093(89)90729-1.

[138] Z Iqbal and S Veprek. “Raman scattering from hydrogenated microcrystalline and amorphous silicon”. In: *J. Phys. C: Solid State Phys.* 15.2 (Jan. 1982), pp. 377–392. DOI: 10.1088/0022-3719/15/2/019.

[139] Michael O. Thompson et al. “Melting Temperature and Explosive Crystallization of Amorphous Silicon during Pulsed Laser Irradiation”. In: *Phys. Rev. Lett.* 52.26 (June 1984), pp. 2360–2363. DOI: 10.1103/physrevlett.52.2360.

[140] J. Dana. Honeycutt and Hans C. Andersen. “Molecular dynamics study of melting and freezing of small Lennard}-{Jones clusters”. In: *J Phys Chem* 91.19 (Sept. 1987), pp. 4950–4963. DOI: 10.1021/j100303a014.

[141] Andrew S. Clarke and Hannes Jónsson. “Structural changes accompanying densification of random hard-sphere packings”. In: *Phys. Rev. E* 47.6 (June 1993), pp. 3975–3984. DOI: 10.1103/physreve.47.3975.

[142] E. Maras et al. “Global transition path search for dislocation formation in Ge on Si(001)”. In: *Comput. Phys. Commun.* 205 (Aug. 2016), pp. 13–21. DOI: 10.1016/j.cpc.2016.04.001.

[143] M. Posselt and A. Gabriel. “Atomistic simulation of amorphous germanium and its solid phase epitaxial recrystallization”. In: *Phys. Rev. B* 80.4 (July 2009). DOI: 10.1103/physrevb.80.045202.

[144] N. Jakse and A. Pasturel. “Liquid-Liquid Phase Transformation in Silicon: Evidence from First-Principles Molecular Dynamics Simulations”. In: *Phys. Rev. Lett.* 99.20 (Nov. 2007). DOI: 10.1103/physrevlett.99.205702.

[145] G.L. Olson et al. “Kinetic Competition During Solid Phase Crystallization in Ion–Implanted Silicon”. In: *MRS Proc.* 23 (1983). DOI: 10.1557/proc-23-375.

[146] G. J. Galvin, J. W. Mayer, and P. S. Peercy. “Solidification kinetics of pulsed laser melted silicon based on thermodynamic considerations”. In: *Appl. Phys. Lett.* 46.7 (Apr. 1985), pp. 644–646. DOI: 10.1063/1.95514.

[147] P. N. Keating. “Effect of Invariance Requirements on the Elastic Strain Energy of Crystals with Application to the Diamond Structure”. In: *Phys. Rev.* 145.2 (May 1966), pp. 637–645. DOI: 10.1103/physrev.145.637.

[148] Jinping Luo et al. “Assessing the EDIP potential for atomic simulation of carbon diffusion, segregation and solubility in silicon melt”. In: *J. Cryst. Growth* 546 (Sept. 2020), p. 125785. DOI: 10.1016/j.jcrysgro.2020.125785.

[149] Luigi Brambilla et al. “Solid–liquid interface velocity and diffusivity in laser-melt amorphous silicon”. In: *Appl. Phys. Lett.* 77.15 (Oct. 2000), pp. 2337–2339. DOI: 10.1063/1.1317535.

[150] R. L. C. Vink, G. T. Barkema, and W. F. van der Weg. “Raman spectra and structure of amorphous Si”. In: *Phys. Rev. B* 63.11 (Mar. 2001). DOI: 10.1103/physrevb.63.115210.

[151] Andreas Pedersen, Laurent Pizzagalli, and Hannes Jónsson. “Atomic and electronic structures of a vacancy in amorphous silicon”. In: *Phys. Rev. B* 101.5 (Feb. 2020). DOI: 10.1103/physrevb.101.054204.

[152] R. J. Bell and P. Dean. “The structure of vitreous silica: Validity of the random network theory”. In: *Philos. Mag.* 25.6 (June 1972), pp. 1381–1398. DOI: 10.1080/14786437208223861.

[153] R. J. Bell and P. Dean. “Atomic vibrations in vitreous silica”. In: *Discuss. Faraday Soc.* 50 (1970), p. 55. DOI: 10.1039/df9705000055.

[154] S. S. Schoenholz et al. “Understanding Plastic Deformation in Thermal Glasses from Single-Soft-Spot Dynamics”. In: *Phys. Rev. X* 4.3 (July 2014). DOI: 10.1103/physrevx.4.031014.

[155] Lijin Wang et al. “Low-frequency vibrational modes of stable glasses”. In: *Nat. Commun.* 10.1 (Jan. 2019). DOI: 10.1038/s41467-018-07978-1.

[156] Yong Zhang and Edward J. Maginn. “A comparison of methods for melting point calculation using molecular dynamics simulations”. In: *J. Chem. Phys.* 136.14 (Apr. 2012), p. 144116. DOI: 10.1063/1.3702587.

[157] G. Pompe and E. Hegenbarth. “Thermal Conductivity of Amorphous Si at Low Temperatures”. In: *Phys. Status Solidi B* 147.1 (May 1988), pp. 103–108. DOI: 10.1002/pssb.2221470109.

[158] V. K Malinovsky et al. “Universal Form of the Low-Energy (2 to 10 meV) Vibrational Spectrum of Glasses”. In: *Europhys. Lett.* 11.1 (Jan. 1990), pp. 43–47. DOI: 10.1209/0295-5075/11/1/008.

[159] Hideyuki Mizuno, Hayato Shiba, and Atsushi Ikeda. “Continuum limit of the vibrational properties of amorphous solids”. In: *Proc. Natl. Acad. Sci.* 114.46 (Oct. 2017), E9767–E9774. DOI: 10.1073/pnas.1709015114.

[160] Asaph Widmer-Cooper et al. “Irreversible reorganization in a supercooled liquid originates from localized soft modes”. In: *Nat. Phys.* 4.9 (July 2008), pp. 711–715. DOI: 10.1038/nphys1025.

[161] C. Laermans and M. Coeck. “Low-frequency Raman scattering in bulk neutron-disordered silicon”. In: *Physica B* 263-264 (Mar. 1999), pp. 280–282. DOI: 10.1016/s0921-4526(98)01226-5.

[162] M. Ivanda, I. Hartmann, and W. Kiefer. “Boson peak in the Raman spectra of amorphous gallium arsenide: Generalization to amorphous tetrahedral semiconductors”. In: *Phys. Rev. B* 51.3 (Jan. 1995), pp. 1567–1574. DOI: 10.1103/physrevb.51.1567.

[163] Frank Finkemeier and Wolfgang von Niessen. “Phonons and phonon localization in $a-\mathrm{Si}$: Computational approaches and results for continuous-random-network-derived structures”. In: *Phys. Rev. B* 58.8 (Aug. 1998), pp. 4473–4484. DOI: 10.1103/physrevb.58.4473.

[164] Frank Finkemeier and Wolfgang von Niessen. “Boson peak in amorphous silicon: A numerical study”. In: *Phys. Rev. B* 63.23 (May 2001). DOI: 10.1103/physrevb.63.235204.

[165] R. Tsu, J.G. Hernandez, and F.H. Pollak. “Determination of energy barrier for structural relaxation in a-Si and a-Ge by Raman scattering”. In: *J. Non-Cryst. Solids* 66.1-2 (July 1984), pp. 109–114. DOI: 10.1016/0022-3093(84)90307-7.

[166] D. Beeman, R. Tsu, and M. F. Thorpe. “Structural information from the Raman spectrum of amorphous silicon”. In: *Phys. Rev. B* 32.2 (July 1985), pp. 874–878. DOI: 10.1103/physrevb.32.874.

[167] Reuben Shuker and Robert W. Gammon. “Raman-Scattering Selection-Rule Breaking and the Density of States in Amorphous Materials”. In: *Phys. Rev. Lett.* 25.4 (July 1970), pp. 222–225. DOI: 10.1103/physrevlett.25.222.

Summary

Evaluation of neural network methodologies for understanding the interface phenomena in integrated materials in microelectronics: application to the case of solid phase epitaxy in Silicon

Moore's Law is ended, but all is not lost. To maintain a higher transistor density, changing device geometry from planar to three-dimensional was one of the first successful approaches to take advantage of dimension z. In particular, recently, 3D monolithic integration has been proposed to align different wafers on top of each other. With this technique, a layer of amorphous silicon is deposited on top of the transistor, this amorphous layer is then crystallized by performing a solid phase epitaxy (SPE) at low temperature. Solid phase epitaxy in silicon material is a general phenomenon that consists in the regrowth of a perfect diamond structure from an amorphous material in direct contact with a crystalline substrate without the need to liquefy it. Industrialists use numerical simulations to optimize their SPE process, but the current codes do not capture all the microscopic events that occur within the material. This is due to the very poor knowledge we have of the microscopic phenomena that need further investigation.
This work achieves a complete ab-initio study of solid phase epitaxy in silicon based on state-of-the-art numerical approaches. First, an interatomic neural network potential is developed and fitted on a data set built from scratch for this specific task. The points of the data set are solved in energy and forces in the density functional theory framework. Particular interest is given to the construction and veracity of the dataset used to train the potential: choice of the functional, exhaustive sampling of each phase space region of interest, reproducibility of the physical properties necessary to model the physical phenomena of the SPE. This neural network was then used to study solid phase epitaxy with molecular dynamics. The simulations show a variety of phenomena that

were not previously accessible with less accurate potentials, and give a microscopic interpretation to the observed experimental activation barrier. In addition, a statistical approach comparing a series of structures with a state-of-the-art density functional theory structure was performed to discuss the relevance of the quality of the amorphous used in our SPE simulation. We determined that several metrics such as the geometry of our amorphous structures and the energy excess with respect to the diamond silicon are consistent with the literature.

Évaluation des méthodologies de réseaux neuronaux pour la compréhension des phénomènes d'interface dans les matériaux intégrés dans les domaines de la microélectronique : application au cas de l'épitaxie en phase solide dans le silicium

La loi de Moore est terminée mais tout n'est pas perdu. Pour maintenir une plus grande densité de transistors, la modification de la géométrie des dispositifs, qui passe du planaire au tridimensionnel, a été l'une des premières approches adoptées avec succès pour tirer parti de la dimension zéro. Récemment, l'intégration monolithique 3D a notamment été proposée pour aligner différents wafers l'un sur l'autre. Avec cette technique, une couche de silicium amorphe est déposée sur le dessus du transistor, cette couche amorphe est ensuite cristallisée par épitaxie en phase solide (SPE) à basse température. L'épitaxie en phase solide dans le Silicium est un phénomène général qui consiste à faire repousser une structure de type diamant parfaite à partir d'un matériau amorphe en contact direct avec un substrat cristallin sans qu'il soit nécessaire de le liquéfier. Les industriels s'aident alors de simulations numériques pour l'optimisation de leur procédé, mais les codes actuels ne capturent pas l'ensemble des événements microscopiques qui se produisent au sein du matériau. Cela est dû à la très faible connaissance que nous avons des phénomènes microscopiques qui doivent être étudiés plus en détails.
Ce travail réalise une étude ab-initio complète de l'épitaxie en phase solide dans le silicium en s'appuyant sur des approches numériques à l'état de l'art. Tout d'abord, un potentiel interatomique de réseau neuronal est développé et ajusté sur un ensemble de données construit de toutes pièces pour ce procédé spécifique. Les points de l'ensemble de données sont résolus en énergie et en forces avec la théorie de la fonctionnelle de la densité. Un intérêt particulier est donné sur la construction et la véracité du jeu de données utilisé pour entrainer le potentiel : choix de la fonctionnelle, échantillonnage exhaustif de chaque région d'espace de phase d'intérêt, reproductibilité des propriétés physiques nécessaires à la modélisation des phénomènes physiques de la SPE. Ce réseau de neurones a ensuite été utilisé pour étudier l'épitaxie en phase solide avec la dynamique moléculaire. Les simulations montrent une variété de phénomènes qui n'étaient pas accessibles auparavant avec des potentiels moins précis, et donnent une interprétation microscopique à la barrière d'activation expérimentale observée. En complément, une approche statistique à partir d'une série de structures issues de la DFT a été réalisée pour discuter la pertinence de la qualité de l'amorphe utilisé dans notre simulation SPE. Nous avons déterminé que plusieurs métriques comme la géométrie de nos structures amorphes et l'excès d'énergie par rapport au silicium diamant sont conformes avec la littérature.