universität wien

DISSERTATION

Titel der Dissertation

Ab initio study of defects in amorphous silicon nitride

Verfasser

Mag. Leif Eric Hintzsche

angestrebter akademischer Grad

Doktor der Naturwissenschaften (Dr. rer. nat.)

Wien, 2014

Studienkennzahl lt. Studienblatt:	A 791 411
Dissertationsgebiet lt. Studienblatt:	Physik
Betreuer:	Univ.-Prof. Dipl.-Ing. Dr Georg Kresse

Zusammenfassung

Amorphes Siliziumnitrid (a-SiN:H) wird üblicherweise aus zwei Gründen auf der Oberseite von kommerziellen Solarzellen verwendet: es dient als Antireflektionsschicht und es passiviert den darunterliegenden kristallinen Siliziumwafer (c-Si). Während die erste Eigenschaft hauptsächlich durch die Stickstoffkonzentration beeinflusst wird, ist die zweite stark von der Defektdichte im Bulk (a-SiN:H) und an der Grenzfläche (c-Si/a-SiN:H) abhängig. Zum Beispiel können elektronische Defektzustände in der Mitte der Bandlücke zu Verlusten von Ladungsträgern führen, indem sie Ladungen einfangen oder zu Shockley-Read-Hall-Rekombinationen führen. Diese Einfang- und Rekombinationszentren reduzieren die Qualität der Passivierung an der Grenzfläche und damit die gesamte Effizienz der Solarzelle. Die Passivierung wird andererseits durch eine fixierte, positive Ladung an der Grenzfläche erhöht, die Elektronen anzieht und Löcher abstößt. Diese Passivierung durch ein elektrisches Feld wird in der Regel durch positiv geladene, dreifach koordinierte Si Atome erklärt (K^+ Defekte). Es ist klar, dass ein gutes Verständnis von Defekten zwingend notwendig ist, um die Effizienz von Solarzellen zu optimieren. Jedoch is die Untersuchung der Defekte aus zwei Gründen besonders herausfordernd. Erstens werden elektronische Defekte durch quantenmechanische Effekte verursacht. Zweitens sind die strukturellen Eigenschaften auf atomarer Ebene durch die ausschließliche Verwendung von experimenteller Methoden nicht direkt erfaßbar.

In der vorliegenden Arbeit untersuchen wir die strukturellen und elektronischen Eigenschaften des a-SiN:H Bulk und der c-Si/a-SiN:H Grenzfläche mit einem groß angelegtem *ab initio* Molekulardynamik Ansatz. Dabei werden große Ensemble von Bulk- und Grenzflächenproben erzeugt, indem viele unterschiedliche Modellstrukturen von der Schmelze in den amorphen Zustand gekühlt werden. Die strukturellen und elektronischen Eigenschaften werden gemittelt und die Defektzustände werden mittels einer automatischen Routine ausgewertet. Unsere Arbeit konzentriert sich auf die Materialzusammensetzungen a-$Si_3N_{3.5}$:H and c-Si/a-$Si_3N_{3.5}$:H, von denen wir galuben, dass sie am häufigsten für Solarzellen verwendet werden. Wir haben aber auch die Zusammensetzung der Bulkmaterialen und die Wasserstoffkonzentration variiert. Schließlich haben wir verschiedene Kompromisse zwischen Systemgröße, Ensemblegröße und Abkühlrate angewendet, um die Konsistenz unserer Resultate zu überprüfen.

Im Allgemeinen stellen wir fest das die Defektkonzentration in der Bandlücke von drei Faktoren abhängt: von der Materialzusammensetzung, von der thermodynamischen Geschichte und von der Wasserstoffkonzentration. Die wichtigsten Defektklassen sind dreifach koordinierte Si Atome (K Defekte) und ideal koordinierte Si Atome mit verzerrten Si-Si Bindungen. Koordinationsde-

fekte sind dominant, aber in gut passivierte Proben, die nahe dem stöchiometrischen Fall sind (a-Si_3N_4), werden die Defekte, die von lokalen Verzerrungen verursacht werden, gleichermaßen wichtig. Das hat auch mit der Beobachtung zu tun, dass Wasserstoff Koordinationsdefekte heilen kann, aber kaum Zustände, die auf ideal koordinierten Atomen lokalisiert sind. An der Grenzfläche befinden sich die meisten Zustände in der Bandlücke innerhalb der Übergangsregion zwischen c-Si und a-SiN:H. Auf der kristallinen Seite sind die Zustände Leitungsbandzustände vom c-Si, welche über mehrerer verzerrte Si-Si Bindungen delokalisiert sind. Auf der amorphen Seite, sind die Zustände andererseits stark auf K Defekten lokalisiert. Wie auch im Bulk können diese Koordinationsdefekte von Wasserstoff geheilt werden. Zum Schluss schlagen wir das folgende Erklärungsmodell für die Bildung der fixierten, positiven Ladungsdichte an der c-Si/a-SiN:H Grenzfläche vor: Da das lokale Fermilevel von ungedoptem c-Si deutlich unter dem von a-SiN:H liegt, passt sich das Fermilevel des Heteroübergangs durch ein Ladungstransfer vom amorphen Teil zum kristallinen Teil an. Dies führt zu einem Überschuss an K^+ Defekten an der Grenzfläche.

Abstract

Hydrogenated amorphous silicon nitride (a-SiN:H) is commonly deposited on top of commercially available solar cells for two reasons: it serves as an anti-reflection coating (ARC) and it passivates the underlying crystalline silicon (c-Si) waver. While the first property is mainly influenced by the nitrogen concentration, the second is, additionally, highly dependent on the defect density in the a-SiN:H bulk and at the c-Si/a-SiN:H interface. For example, electronic defect states in the middle of the band gap can lead to carrier losses through charge trapping or Shockley, Read and Hall (SRH) recombination. These trapping and recombination centers reduce the passivation quality and thereby the overall cell efficiency. On the other hand, the passivation is improved by a fixed, positive charge density at the interface that attracts electrons and repels holes. This field effect passivation is usually explained by positively charged, threefold coordinated Si atoms (K^+ defects). It is clear that a good understanding of defects is mandatory in order to optimize the efficiency of solar cells. However, there are two facts that make defect investigations particularly challenging. First, electronic states are caused by quantum mechanical effects. Second, the structural properties at the atomic scale are not directly accessible using experimental methods only.

In the present work, we examine the structural and the electronic properties of the a-SiN:H bulk and the c-Si/a-SiN:H interface using an *ab initio* molecular dynamics (MD) approach. Large ensembles of bulk and interface samples are generated by cooling many different model structures from the melt to the amorphous state. The structural and the electronic properties are averaged and the defect states are examined by applying an automatic routine. Our work mainly focuses on the hydrogenated, nitrogen deficient compositions a-$Si_3N_{3.5}$:H and c-Si/a-$Si_3N_{3.5}$:H, which we deem are most representative for commercial solar cell applications. However, we have varied the bulk composition and the hydrogen (H) concentration as well. Finally, we investigated different compromises between system size, ensemble size, and cooling rate to verify the consistence of our results.

In general, we find that the concentration of defect states in the band gap depends on three factors: the system composition, the thermodynamic history, and the H concentration. The most important defect classes are threefold coordinated Si atoms (K defects) and ideally coordinated Si atoms with distorted Si-Si bonds. K defects are dominant, but, for well passivated samples that are close to the stoichiometric composition (a-Si_3N_4), defects due to local distortions become equally important. This is related to the observation that H can cure K defects, but hardly states localized at ideally coordinated, but geometrically distorted, atoms. At the interface, most gap

states are located within the transition region between c-Si and a-SiN:H. On the crystalline side, these states are c-Si "conduction band like" states that are delocalized over several distorted Si-Si bonds. On the amorphous side, on the other hand, the gap states are strongly localized at K defects. As in bulk a-$Si_3N_{3.5}$:H, these K defects can be cured by H. For the formation of the fixed, positive charge at the c-Si/a-SiN:H interface, we suggest the following explanation. Since the local Fermi level of undoped c-Si lies well below that of a-SiN:H, the Fermi level of the heterojunction aligns by a charge transfer from a-SiN:H to c-Si. This results in an abundance of K^+ defects at the interface.

Contents

1 Motivation

1.1 Singlejunction solar cells based on crystalline silicon (c-Si)

In this section, we briefly discuss the general importance of solar energy and the function of popular singlejunction solar cells that are based on c-Si. Specifically, we describe the transformation process from solar energy into electrical energy, the reasons for losses during this process, and the methods that are used to increase the efficiency of solar cells. Finally, we motivate the main objective of this thesis: the structural and electronic investigation of defects in hydrogenated amorphous silicon nitride (a-SiN:H).

1.1.1 Importance of solar cell technology

Alternative energies are becoming increasingly important for the developed as well as for the emerging countries. Although most of todays energy thirst that drives worldwide economy is nursed by fossil fuel and nuclear power, those conventional energy sources have some important disadvantages. Oil, gas, coal and even wood increase the green house effect by emitting a large amount of carbon dioxide, and nuclear energy is a dangerous alternative, as it can harm environment and people. Moreover, the disposed radioactive waste will definitely stay a problem for generations.

Besides the pollution of the environment, there is always the political relevance of energy resources. Although nearly all countries rely on non-renewable energy to keep their economy going, very few of them own non-renewable energy resources by themselves. For this reason, most countries need to import this kind of energy, and this need can easily be abused to exert political influence. It is, furthermore, well known that the price of non-renewable energy, such as the oil price, is highly dependent on the political situation in the producing countries. Since the aim for self-reliance and independence is strong, most countries would like to introduce a energy-political change towards renewable wind, water and solar energy.

Wind and water energy systems are both powerful tools to bring an alternative energy-politic into practice, but both systems have common disadvantages. They are big, stationary, and have high acquisition and maintenance costs. Todays solar cell devices are, on the other hand, adjustable in size, mobile, and cheap, what makes them especially important whenever flexible energy sources are needed. Although most prominent applications for solar cell devices are found in telecommunication and aerospace (e.g. satellites, space ships, space stations), the cells can in principle be applied for any device with a small power consumption to make it independent from a stationary energy grid.

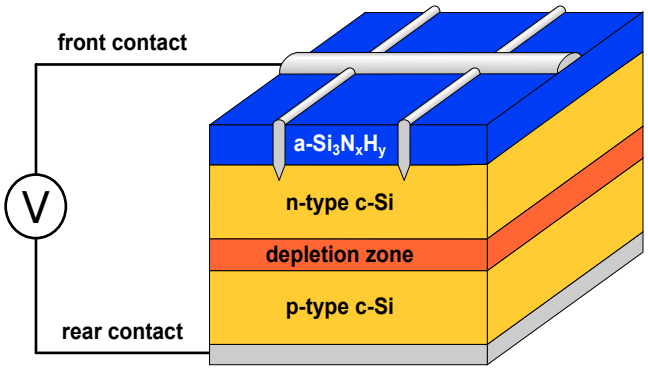


Figure 1: An illustration of a singlejunction solar cell based on c-Si. On top of the p-n junction, a-SiN:H serves as an anti-reflection coating (ARC) and passivation layer.

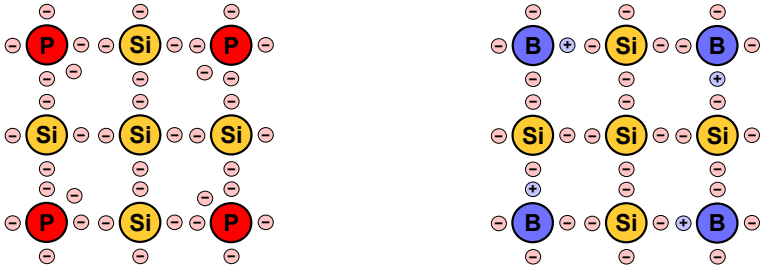
As the price of solar cells has significantly decreased, solar energy is even affordable for private households nowadays. Driven by financial reasons and/or ecological awareness, more and more people are investing in stationary solar panels to generate electricity for their own consumption or to sell it to the local electricity distributor. Apart from the initial outlay, the maintenance costs and possible public subsidies, the success of such investments mainly depend on the efficiency and the durability of the cells.

In summary, the importance of solar cells is obvious and ongoing. There are many possible applications for solar cells, with a constant demand for cheaper, more efficient, and more durable devices. For this reason, investors such as governments, companies, and private persons have great ecological, political, and financial interest in solar cell technology.

1.1.2 The function of a singlejunction solar cell based on c-Si

There are many materials used for photovoltaic (PV) devices such as gallium arsenide (GaAs), cadmium tellurid (CdTe), indium phosphide (InP) and copper indium gallium selenide (CIGS), but over 90 % of todays solar cells are based on silicon (Si) [1,2]. A plausible explanation for this market situation is the availability of nontoxic Si and the variety of its processing possibilities. For example, solar cells can be based on low efficient, but cheap, amorphous silicon (a-Si), cost-effective poly crystalline silicon (pc-Si) or highly-efficient, but more expensive, mono crystalline silicon (mc-Si). Under laboratory conditions, the efficiency of amorphous cells is about 10 %, while the efficiency of multi-crystalline and mono-crystalline cells is about 20 % and 25 %, respectively [3]. Altogether, nowadays most solar cells are produced with c-Si.

Fig. 1 illustrates the basic structure of a singlejunction solar cell that consists of a c-Si waver,



(a) n-type doping (b) p-type doping

Figure 2: While n-type doping with phosphorus (P) increases the number of electrons in the c-Si matrix, p-type doping with boron (B) decreases the number of electrons by introducing holes.

covered by an anti-reflection coating (ARC) and connected by two metal contacts [3, 4]. The waver consists of a p-doped and n-doped c-Si layer, whereas typical doping elements are phosphorus (P) or arsenic (As), for the n-type layer, and boron (B) or gallium (Ga), for the p-type layer [5]. As shown in Fig. 2(b), the n-type doping adds additional electrons to the conduction band (CB) of c-Si, while the p-type doping creates holes by removing electrons from the valence band (VB) of c-Si. As a result, there is a larger electron density in the n-type layer and a larger hole density in the p-type layer. This difference causes an electric field that lets the electrons diffuse into the p-type layer and the holes into the n-type layer until the p-n junction reaches a equilibrium state [5]. In this state the recombination of electrons and holes has formed a depletion region between both layer types, and the electric field within the depletion region effectively separates electrons from holes.

An excited state is, on the other hand, created in the p-n junction, when an external incident beam of light provides photons with enough energy to excite electrons from the valence band (VB) to the conduction band (CB). As illustrated in Fig. 3 [3,4], this process generates additional electron hole pairs by leaving holes in the CB behind. The corresponding energy that the photons need to create electron hole pairs is thereby naturally given by the band gap energy of c-Si (1.12 eV). Since the excited electrons and the holes recombine after a while again, both carriers have a finite life time. However, under permanent light exposure some electrons in the p-type layer and some holes in the n-type layer can travel through the depletion region, so that a voltage develops between both doped c-Si layers. When the electrodes are connected, both carrier types can recombine through the electric connection again. This results in a current that can be used for electric devices or for the injection of electricity into the local power supply

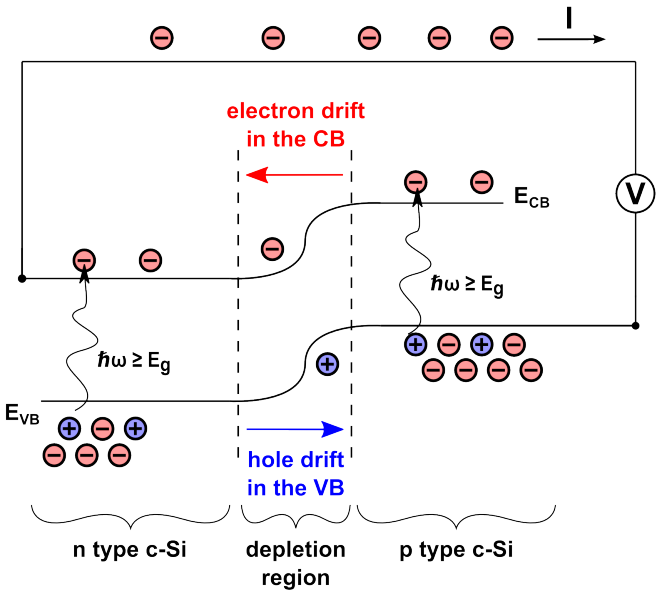


Figure 3: The photovoltaic (PV) effect in a heterojunction formed by p-type and n-type c-Si. Photons with energies $\hbar\omega$ greater than the band gap energy E_g create electron hole pairs. The potential difference between both bulk layers drives the electrons in the conduction band (CB) towards the n-type layer and the holes in the valence band (VB) towards the p-type layer. As a result, a depletion region and a voltage V develop between both layers. If both layers are connected by electrodes, the electrons will recombine with the holes by generating the current I.

system. This transformation process of solar energy into electrical energy in solar cells is called the PV effect.

1.1.3 The efficiency of solar cells and natural limitations

The efficiency η of the energy conversation of a solar cell is generally defined by the ratio between the maximal power output $P_{\max}$ of the cell and the total power input P_{in} of the incident beam of light.

$$\eta = \frac{P_{\max}}{P_{\text{in}}} = \frac{V_{\text{oc}} I_{\text{sc}} FF}{P_{\text{in}}}. \tag{1}$$

However, for technical reasons, it is more convenient to relate η to the open-circuit voltage V_{oc}, the short-circuit current I_{oc}, and the fill factor FF [6]. The open-circuit voltage is the electrical potential between the unconnected electrodes of the cell, thus the largest possible voltage from a solar cell. The short-circuit current, in contrast, is the largest possible current that runs through

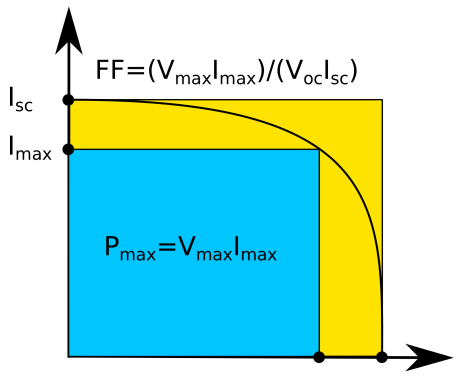


Figure 4: The electric current output I of a solar cell as a function of the voltage V. The fill factor FF of the cell is given by the ratio of the maximal power output $P_{\max}$ of the cell and the fictive power calculated by the product of the open-circuit voltage V_{oc} and the short-circuit current I_{sc}.

the directly connected electrodes. As a matter of fact, there is no current when $V = V_{\mathrm{oc}}$ and no voltage when $I = I_{\mathrm{sc}}$, and in both cases the power output is equal to zero. Therefore $P_{\max}$ must be given by a product of a voltage $V_{\max}$, with $V_{\max} < V_{\mathrm{oc}}$, and a current $I_{\max}$, with $I_{\max} < I_{\mathrm{sc}}$. The fill factor FF is then defined by

$$FF = (V_{\max} I_{\max})/(V_{\mathrm{oc}} I_{\mathrm{sc}}). \tag{2}$$

As illustrated in Fig. 4, the fill factor basically describes the ratio between the maximal power output of the cell, which is determined by the optimal working point of the solar cell, and a fictive power output that is given by the product of V_{oc} and I_{sc}.

There are two theoretical factors that naturally limit the possible efficiency of singlejunction solar cells. The first factor is that the voltage V and the current I contrarily depend on the size of the band gap. For example, the V_{oc} increases with increasing gap sizes, but I_{sc} decreases. This is because the increased band gap effectively reduces the fraction of photons with sufficient energy to generate electron hole pairs. It turns out that there is an optimal band gap value which lies between 1.4 eV and 1.6 eV [6]. Examples of suitable materials for p-n junctions are therefore c-Si (1.12 eV), InP (1.35 eV), GaAs (1.42 eV), CdTe (1.45 eV), and a-Si (1.75 eV) [4]. Although compounds such as GaAs and CdTe have nearly perfect band gap values, c-Si is usually preferred for large scale manufacturing because of its cost effectiveness and non-toxicity.

The second factor that limits the solar cell efficiency concerns the convertibility of the light spectrum. Since the spectrum also consists of photons with less or greater energy than the band gap energy, there is much photo energy that remains unused. Specifically, electron hole pairs

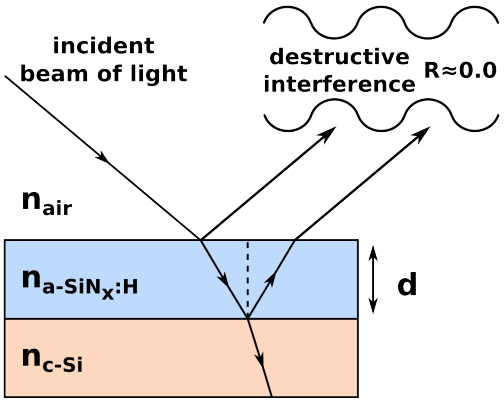


Figure 5: An incident beam of light reaches a singlejunction solar cell based on c-Si and a-SiN:H. The light is reflected two times, at the a-SiN:H surface and at the c-Si/a-SiN:H interface. The total reflectivity of the cell is minimal if the phase difference between both reflected light waves is equal to half of the wavelength. The phase difference is determined by the refractive index $n_{\text{a-SiN:H}}$ and the thickness d of the anti-reflection coating (ARC). As the refractive index of a-SiN:H can be controlled by changing the N concentration during plasma enhanced chemical vapor deposition (PECVD), a-SiN:H is often used as ARC. Furthermore, graded anti-reflective layers allow to obtain the extinction of different wavelengths.

are not created by photons with too low energy, and the excess energy of photons with too high energy is transformed into heat and therefore lost as well [6]. Because of this, a large part of the spectrum of the sunlight remains unused and the theoretically possible efficiency is limited to about 44 % [7]. Note that this limit only applies to singlejunction cells. There are also other types of solar cells that achieve larger practical and theoretical efficiency values, but those are not considered in this thesis.

1.1.4 Optical losses and optimization techniques

Besides the natural efficiency limit of singlejunction solar cells, there are additional optical losses that must be minimized to optimize the solar cell performance. First of all, there is the problem that not all the light reaches the p-n junction. For example, $5 - 15$ % of the incoming light is blocked by the contact grid and over 30 % of the incoming light is reflected on the surface of bare c-Si [1, 6]. For this reason, the contact grid must be designed to increase the amount of light that reaches the c-Si and losses due to the reflectivity must be decreased by deposing

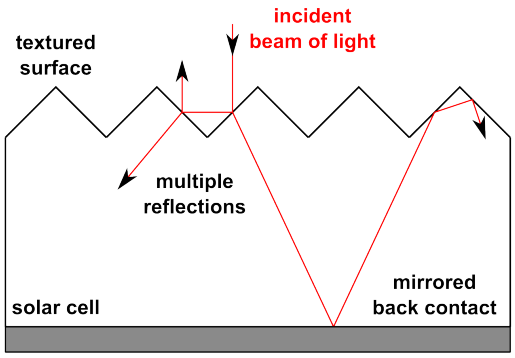


Figure 6: Textured surfaces and mirrored back contacts are used to generate several light reflections within the solar cell. This effectively increases the probability for photons to create electron hole pairs.

an anti-reflection coating (ARC) on c-Si. Typical ARCs consist of a dielectric material such as silicon dioxide (SiO_2), titanium dioxide (TiO_2), or silicon nitride (Si_3N_4) [8]. As explained in Fig. 5, the refractive index and the thickness of these additional anti-reflection layers can then be tuned to reduce the reflectivity to about 10 % [9].

Other optical losses are caused by photons that enter the p-n junction, but leave the device again without creating an electron hole pair. For this reason, reflective rear contacts and textured surfaces are often applied to trap the light in solar cells [4]. As shown in Fig. 6, both techniques lead to multiple reflections, which increase the probability of creating electron-hole pairs within the cell. Light trapping techniques are especially important for solar cells with very thin c-Si layers, as the probability for the absorption of a photon decreases with the thickness of the layer. This is also known as the Beer-Lambert's law [10]. On the other hand, very thin c-Si layers are highly desirable because c-Si wavers can cause over 30 % of the costs of todays solar cells [11].

1.1.5 Electron hole recombinations and defect states

Another factor that heavily influences the solar cell efficiency are electron hole recombinations. When an electron hole pair is created by a photon the system is in an excited state and, since the system wants to relax back into its equilibrium state, the electron and the hole recombine after a while again. For this reason minority carriers have only a finite lifetime τ in solar cells, which is defined by the ratio between the number of generated minority carriers Δn in the excited state and their recombination rate R.

$$\tau = \frac{\Delta n}{R}. \tag{3}$$

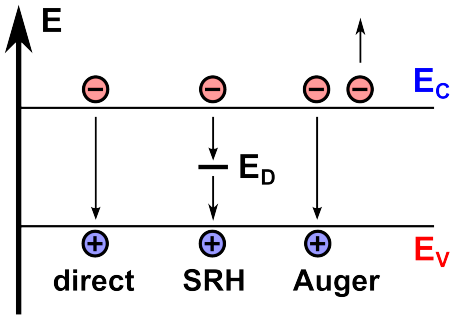


Figure 7: The three processes that lead to recombinations of electron hole pairs are (i) direct recombination, (ii) defect mediated Shockley, Read and Hall recombinations and (iii) Auger recombinations.

As shown in Fig. 7, it is possible to distinguish three different types of recombination processes that reduce the minority carrier lifetime in bulk materials. Firstly, there are direct recombinations, where electrons recombine with holes at the same point in reciprocal space. Since these processes emit photons, they are also called radiative recombinations. As radiative transitions are more likely for materials with a direct band gap, minority carrier lifetimes are typically much smaller in direct semiconductor such as GaAs (1 ns to 100 ns) than in indirect semiconductors such as c-Si (100 ns to 1 ms) [6, 12]. Secondly, there are Auger recombinations, where the excess energy of the electron recombination excites another electron in the valence band (VB) or conduction band (CB). This excited electron can then relax back into an equilibrium state by emitting phonons. Auger recombinations are very important in highly doped materials with a small band gap [6, 12]. Finally, there are defect assisted transitions, which allow the electrons to relax first from the VB to a defect level within the band gap and from there further to the CB. Since the statistics of this two step process was described by Shockley, Read and Hall (SRH) [13,14], recombinations via defect levels are also called SRH recombinations. It turns out that the defect levels in the middle of the band gap are the most effective SRH recombination centers because they effectively halve the size of the band gap [6, 12].

Direct, Auger, and SRH recombinations add all to the total recombination rate of electron hole pairs in a bulk material. The bulk minority carrier lifetime τ_{bulk} can hence be written as [15]

$$\frac{1}{\tau_{\text{bulk}}} = \frac{1}{\tau_{\text{band}}} + \frac{1}{\tau_{\text{Auger}}} + \frac{1}{\tau_{\text{SRH}}}. \tag{4}$$

Since additional electronic levels within the band gap are caused by structural defects such as broken or strained bonds, there is an increased electronic defect density everywhere where the structural proprieties of crystalline or amorphous structures are disturbed [4]. This is for

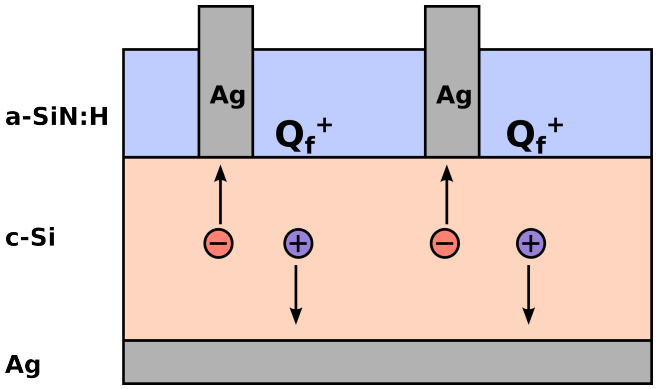


Figure 8: An illustration of the field effect passivation by the positive, fixed charge density Q_f^+ at the interface between c-Si and a-SiN:H. While Q_f^+ repels holes from the interface, it simultaneously attracts electrons to the silver contact of the solar cell. This separates electron hole pairs close to the interface and increases the lifetime of both carrier types in the solar cell.

example the case at surfaces, interfaces, and grain boundaries, but also close to impurities and precipitates. Note that all these terms can in principle be considered as surfaces, or as concepts that create surfaces. The total effective recombination rate in presence of a surface is then the recombination rate in the bulk material plus the recombination rate due to surfaces. The effective minority carrier lifetime τ_{eff} of solar cells is hence often given by

$$\frac{1}{\tau_{\text{eff}}} = \frac{1}{\tau_{\text{bulk}}} + \frac{1}{\tau_{\text{surf}}}. \tag{5}$$

Here the effective surface lifetime τ_{surf} is mainly influenced by the thickness of the waver, the area of the surface and the velocities of the minority carriers towards the front and the back surface of the solar cell (i.e. surface recombination velocity) [4, 16]. A fundamental challenge in PV industry is the reduction of losses due to recombination centers through the passivation of the structural defects at surfaces and in the bulk material. This implies to saturate dangling bonds and to reduce the electronic states in the band gap by minimizing the number of structural defects. As a side effect, this process also reduces the reaction probability of the p-n junction with other particles, which are omnipresent in our atmosphere.

1.2 Hydrogenated amorphous silicon nitride (a-SiN:H) as anti-reflection coating (ARC) and passivation layer

To reduce the reflectivity and the number of recombination centers in solar cells an ARC and passivation layer is needed, respectively. Materials which are used for this purpose are for

example silicon dioxide (SiO_2), titanium dioxide (TiO_2), zinc sulphide (ZnS), aluminium oxide (Al_2O_3) and silicon nitride (Si_3N_4) [8, 9, 11, 17]. Among these candidates, Si_3N_4 is the most frequently used material because it offers an excellent trade-off between anti-reflection and defect-passivation properties [18, 19]. Furthermore, it generates a positive, fixed charge density at the interface between c-Si and a-SiN:H that repels holes and attracts electrons, compare Fig. 8 [9, 20–22]. Since the electric field of the fixed charge separates electrons and holes at the interface, it increases the lifetime of the minority carriers and thereby the efficiency of solar cells.

A-SiN:H is usually deposited on top of c-Si wavers by plasma enhanced chemical vapor deposition (PECVD) [20, 23–25], though other deposition methods like sputtering are used as well [26, 27]. PECVD is particularly suitable for industrial production, because of its high deposition rates (a few minutes process time), and low deposition temperatures (300–400 °C) [19]. Additionally, the a-SiN:H composition can be adjusted by changing the gas flow of silane (SiH_4) and nitrogen/ammonia (N_2/NH_3) during the deposition process. As the refractive index is highly dependent on the nitrogen concentration [24], the reflectivity of the passivation film can be tuned in terms of the deposition parameters.

The passivation quality of a-SiN:H is improved by reducing the number of defect states at the c-Si/a-SiN:H interface (surface passivation) as well as in the underlying c-Si (bulk passivation) [22]. In particular, for the passivation of the bulk c-Si, H atoms play an important role. Since H_2 is immediately released into the atmosphere, Si-H and N-H bonds are the natural source of H. These bonds can break during the firing process (600-800 °C) and release single H atoms [11,22]. These atoms can then diffuse into the c-Si bulk to attach to under coordinated Si atoms. Although, best optical properties are achieved for N rich samples, the quality of the surface passivation is better for Si rich samples [22]. For this reason, optimal optical properties and passivation properties can not be realized simultaneously. However, commercial solar cells with good efficiencies (about 16.5 %) and low reflectivity (under 10 %) can still be fabricated by using properly adjusted a-Si_3N_x:H compositions [19, 22].

1.2.1 Experimental investigations on a-SiN:H

Early experimental studies by Aiyama *et al.* and Misawa *et al.* concentrated exclusively on the structural properties of a-SiN:H [28, 29]. Particularly, they determined many structural properties such as the structure factor (STF), the pair correlation function (PCF), the bond lengths, and the bond angles by using x-ray and neutron scattering. Both studies basically agreed that a-SiN:H mainly consist of fourfold coordinated Si atoms and threefold coordinated N atoms. As shown in Fig. 9, the Si atoms bond to N by forming a tetrahedral structure with

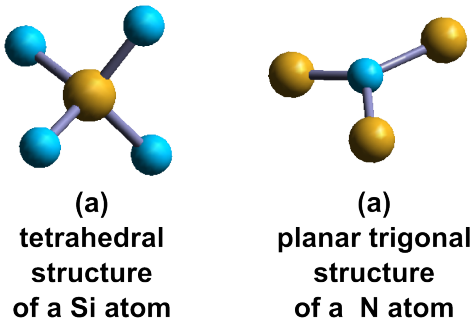


Figure 9: Ideally coordinated atoms in a-SiN:H. Most atoms in a-SiN:H are (a) fourfold coordinated Si atoms or (b) threefold coordinated N atoms.

ideal N-Si-N bonding angles of 109.5°, and the N atoms bond to Si by forming a planar trigonal structures with ideal Si-N-Si bonding angles of 120°. Although they mostly found Si-N bonds with a bonding distance of about 1.73 Å, they could also observe some Si-Si and N-N bonds with a typical bonding length of about 3.0 Å and 2.8 Å, respectively. Small deviations from this ideal a-SiN:H structure were explained by the effects due to structural defects, voids and surfaces.

Additional information on the bond densities, the density of states (DOS), the optical band gap and the refractive index were obtained by Giorgis *et al.* and Guraya *et al.* [23, 24, 27, 30, 31]. By changing the deposition rate and the gas environment during plasma enhanced chemical vapor deposition (PECVD) or during radio frequency sputtering (RF sputtering), they prepared many samples of hydrogenated amorphous silicon nitride with variable nitrogen concentration (a-Si_3N_x:H). Afterwards, they examined the samples using several optical and electronic measurement methods. They observed that the size of the optical band gap and the refractive index of a-SiN:H are strongly dependent on the nitrogen concentration of the samples. By changing the composition from pure amorphous silicon (a-Si) towards a-Si_3N_4, the band gap increases progressively from 1.9 eV to 5.2 eV and the refractive index decreases linearly from 3.5 to 1.8. It turns out that the compositions close to a-Si_3N_4 have the best optical properties for solar cell applications and that the bond densities of these compositions are in good agreement with the previous structural investigations, meaning that nearly all atoms are ideally coordinated and most bonds are Si-N bonds.

First experimental evidence for the existence of carrier traps that can lead to SRH recombinations within the band gap was given by Lenahan *et al.* and Warren *et al.* [32, 33]. Both studies prepared N rich a-SiN:H samples using plasma enhanced chemical vapor deposition (PECVD)

and high-purity a-Si_3N_4 powders using low pressure chemical vapor deposition (LPCVD) in order to identified parametric resonance centers through electron spin resonance (ESR) measurements. These centers could be attributed to singly occupied threefold coordinated Si atoms (K^0 defects) and singly occupied twofold coordinated N atoms (N^0 defects), whereas the K^0 defects are considered to be dominant in a-SiN:H.

Although we could image other defects such as doubly occupied threefold coordinated Si atoms (K^- defects), unoccupied threefold coordinated Si atoms (K^+ defects), doubly occupied twofold coordinated N atoms (N^- defects), and unoccupied twofold coordinated N atoms (N^+ defects), only K^0 defects and N^0 defects are paramagnetic and hence detectable by ESR measurements. To overcome this issue, K^+ defects, K^- defects, N^+ defects, and N^- defects are often convert to K^0 defects and N^0 defects by ultraviolet (UV) illumination [34]:

$$K^+ + K^- = 2K^0 \tag{6}$$

$$N^+ + N^- = 2N^0. \tag{7}$$

This method increases the number of detectable defect states, but we must consider that there might be still a considerable number of unconverted, undetected defects and that the detected defects might have changed their electronic properties through the conversation.

As solar cells have become a mass product, the cost-effective production of c-Si/a-SiN:H interfaces has increased in importance. This trend is, for instance, reflected in the publications by Aberle *et al.*, who investigated the deposition process of a-SiN:H on solar cells and emphasized the good passivation quality of the material [11, 20, 35]. Likewise, Schmidt *et al.*, Sopori *et al.*, and Soppe *et al.* examined the influence of the a-SiN:H deposition process on important solar cell properties such as the effective minority carrier lifetime, the surface recombination velocity, and the overall solar cell efficiency [17–19]. For example, Sopori *et al.* stressed that a-SiN:H does not only passivate the defects at the surface of solar cells (i.e. surface passivation) but also the defects in the underlying c-Si wavers (i.e. bulk passivation). As already mentioned before, the bulk passivation is reasoned by H atoms that are provided by a-SiN:H. During the metal-contact firing process, some of the H atoms diffuse into the bulk region, where they can passivate defects and impurities. This phenomena is called hydrogen passivation [4, 36–38]. The common model is that H atoms form a bond to the Si atoms with a dangling bond. Consequently, they reduce the amount of structural defects and of electronic defects simultaneously. Furthermore, Schmidt investigated the effective surface recombination velocity, the positive, fixed charge, and the density of interface states for different c-Si/a-Si_3N_x:H interfaces. He basically supported the model that the positive, fixed charge density is created by K^+ defects, which are located within

a 20 nm thick region of the a-SiN:H. These defects lie very close to the CBM of c-Si (less than 0.1 eV) and can capture electrons from the CB of c-Si.

1.2.2 Experimental limitations and computer simulations

Since electronic defect states can be localized on particular atoms, but also extended over several atoms, it is very challenging to explain the relationship between geometrical defects and electronic states in the band gap. Furthermore, structural properties at the atomic scale cannot easily be investigated by experimental methods. For instance, the structural information from x-ray or neutron diffraction is measured indirectly in reciprocal space, and measurements using transmission electron microscopy (TEM), scanning electron microscopy (SEM), or atomic force microscopy (AFM) are limited to microscopic and macroscopic investigations at the surface or close to the surface [39–41]. Moreover, some of these methods can heavily influence the materials properities of the samples. For instance, sample preparation as well as radiation beams might easily create artifacts by changing the atomic structure [40, 41].

The electronic properties are also challenging to examine by experiments. For example, the valence band (VB) and the conduction band (CB) of the density of states (DOS) are sometimes examined by photoemission spectroscopy (PES) and by inverse photoemission spectroscopy (IPES), respectively [42, 43]. Furthermore, the size of the optical band gap is often measured in terms of the changes of the absorption spectra [24, 27]. These methods give acceptable results of general electronic features, but they are still not completely satisfactory. Especially in amorphous structures there are many intermingled electronic contributions to the VB tail, to the CB tail and to the defect density in the middle of the gap, so that contributions from particular defects cannot directly be identified.

Because of these experimental limitations, computer simulations are often applied to obtain additional knowledge on specific defect contributions in a-SiN:H. In comparison with experiment, computer investigations have three main advantages. First, they provide direct access to atomic-scale structural informations, and *ab initio* methods offer insight into quantum mechanical properties. Second, the structural and electronic information can be evaluated at once and without disturbing the sample. Third, they have the potential to explain, to confirm, or to contradict experimental observations.

To build amorphous model structures by computer simulations different strategies can be adopted. A common strategy is to assemble small subunits and clusters with proper short range order (SRO) into a continuous random network (CRN) [44] or Bethe lattice [45,46]. Furthermore, bond switching methods can be applied to the crystalline samples until they become reasonably

amorphous [47,48]. Although very small defect concentrations are achievable by these strategies [47], the most common and maybe most unbiased method to determine amorphous structures is by cooling the samples from the melt. This can be done by either Monte Carlo (MC) [49–51] or molecular dynamics (MD) [48, 52–58] simulations.

Ultimately, all simulations need to rely on a potential energy surface to calculate jump probabilities and forces during the cooling and the relaxation. *Ab initio* methods are very desirable, but incur great computational cost [48, 53–56]. For this reason, many simulations use computationally cheaper force field methods such as the Tersoff, the tight binding, the Keating, the Busing, or the Born-Mayer-Huggins method [44, 46, 49, 50, 52, 57, 58]. *Ab initio* and force field methods can also be combined in order to save computation time and/or to simulate larger systems. In this case, the samples are annealed using the cheaper potentials, and the electronic properties are calculated with more accurate *ab initio* methods [51, 58] afterwards.

1.2.3 Computer simulations on a-SiN:H

In recent computer simulations, typically a small ensemble of a-SiN:H configurations were produced in order to investigate the structural, electronic, optical as well as vibrational properties [44, 46, 48–57]. These studies basically confirm the experimental bonding lengths, bonding angles, and coordination numbers, and they illustrated the dependency of the optical band gap on the composition of the system, the density of the sample, and the density of structural defects. However, all these studies have only considered a limited number of model structures for each a-SiN:H composition, making a statistically meaningful analysis very difficult. Especially because not only density fluctuations but also structural defects heavily influence the materials properties of a single structure (e.g. charge density, electrostatic potential, band gap), results based only on a few samples are quite ambiguous.

Besides, very few of these recent studies have explicitly investigated the electronic structure of Si and N dangling bonds like Robertson *et al.* did in the 90s [36, 59]. Robertson *et al.* applied tight binding methods to small, essentially crystalline model systems and proposed that K defects are the most important recombination centers in a-SiN:H. These defects outnumber N defects, lie in the middle of the band gap, and trap electrons or holes. Furthermore, they found out that the bonding and anti-bonding states of Si-Si bonds strongly contribute to the VB and the CB tail of a-SiN:H and that N defects only contribute to the VBM. Since they also observed that N_3SiH units do not contribute to the defect density within the band gap, it is commonly assumed that H increases the band gap by curing K defects. All these observations were later supported by Shu-Ya Lin, who also examined the electronic structure of the defects using the

tight binding method [46]. We must nevertheless note that, with the advances in *ab initio* density functional theory (DFT), it is now possible to study the defects in bulk a-SiN:H in much greater detail and using better founded approximations.

Since the defect concentrations at interfaces are always larger than in bulk materials, several other authors recently examined c-Si/a-SiN:H interfaces [22, 60–66]. For example, Butler *et al.* performed a topological analysis and compared the density of structural defects for different c-Si/a-SiN:H junctions [62, 63]. By introducing a transition region between c-Si and a-SiN:H, they observed an increasing concentration of K defects and N defects towards the a-SiN:H bulk. Furthermore, they estimated the recombination rate of charge carriers at the structural defects for different thicknesses of the transition region, and suggested an optimal value of 2 nm. Their studies relied on computationally cheap Tersoff potentials, which allowed them to calculate relatively thick transition regions. However, like other studies using force field methods [60], their work is limited to investigations of structural properties. Specifically, their conclusions on electronic properties are based on the assumption that mainly under coordinated atoms contribute as electronic defects. As the electronic properties are, nevertheless, important for many applications, Pham *et al.* generated c-Si/a-SiN:H interface structures by combining classical MD and *ab initio* methods [66]. Afterwards, they examined the G_0W_0 corrected band offsets and the change of the dielectric constant perpendicular to the interface. However, their focus clearly lied on the application of a-SiN:H as high K dielectric in metal oxide semiconductor field effect transistor (MOS-FET) devices and recombination centers were not examined in their work.

In other words, the current results on the defects in the a-SiN:H bulk and at the c-Si/a-SiN:H interface have been achieved with somewhat out-dated methods and the relationship between the electronic states and the structural defects has rather been implied than demonstrated. Specifically, it often seems that the structural information (i.e. geometrical defects) and the electronic information (i.e. electronic defect levels) were investigated separately and were not properly related to each other. Modern *ab initio* methods have contributed relatively little to this topic and force field simulations were restricted to the analysis of structural properties, with no electronic information at all. Moreover, all results obtained by *ab initio* methods were based on a small number of model structures, making the interpretation of these results very biased. For these reasons, we revise the corresponding results by applying state of the art methods to a large dataset of a-SiN:H samples.

1.3 The outline of this Ph.D. thesis

The present Ph.D. thesis is part of a European project that investigates contacting and passivation properties of high performance solar cells: the HiperSol project. This project is a collaboration between universities (i.e. University of Sheffield, University of Vienna), research institutes (i.e. SINTEF, the Energy Research Center of the Netherlands, the International Solar Energy Research Center in Konstanz, and the Royal Institute of Technology in Stockholm) and industry (i.e. Sunways and Isofoton) that accumulates various expert knowledge by bringing theoreticians, experimentalists and producers together. The selection of cooperation partners guarantees the availability of research facilities for computer simulations, experimental measurements, and manufacturing. The project is funded by the European Union and the ambitious objective of this framework is to combine results from computer simulations and experiments in order to understand the interface between the a-SiN:H passivation layer and the c-Si substrate in solar cell devices. Specifically, the aim of this thesis is to analyse the a-SiN:H bulk and the c-Si/a-SiN:H interface with state of the art computational methods. This includes the generation of large ensembles of a-SiN:H structures using *ab initio* MD techniques followed by the examination of defects and defect passivation.

This pure *ab initio* strategy has one major advantage in comparison to other approaches that rely on semi-empirical potentials. The method, which is based on density functional theory (DFT), includes quantum mechanical and electronic properties already during the annealing process. We are convinced that the resulting amorphous structures are far more reliable than structures generated by semi-empirical potential methods. For example, the formation and the breaking of bonds affect the total energy landscape during the simulation. These changes are invisible for semi-empirical potentials, but crucial to generate realistic model structures. Additionally, we profit from our many samples strategy, which has hardly been applied in literature yet. Since we generate many different model structures, it is possible to calculated important materials properties such as the pair correlation function (PCF), the angular distribution function (ADF), the structure factor (STF), and the density of states (DOS) in terms of ensemble averages. By relating the structural and electronic information automatically with each other, we finally perform a statistical analysis on defect states in the a-SiN:H bulk and at the c-Si/a-SiN:H interface.

2 Theoretical background

2.1 The time-independent Schrödinger equation in the Born-Oppenheimer approximation

In quantum mechanics, the classic equation of motion for particles is replaced by an equation of states, where a wave function characterizes the probability of finding particles in a certain quantum mechanic state. For a stationary system, this quantum mechanic equation is given by the time-independent Schrödinger equation [67, 68]

$$\mathrm{H}\Psi = E\Psi. \tag{8}$$

The wave functions Ψ are defined by all possible solutions of the eigenvalue problem with the Hamilton operator H, and the corresponding energies E are the eigenvalues of the operator.

In solid state physics and chemistry, the many body problem is given by atoms consisting of nuclei and electrons. The Born-Oppenheimer approximation splits the solution of this problem into two parts. The quantum mechanical solution for interacting electrons in an external potential and the solution for interacting nuclei. The interaction of the nuclei can thereby be treated quantum mechanically, although in the present work the nuclei are described by the classical Newton equation. The justification for this separation is the large difference in mass between the electrons and the nuclei [69, 70]. As a consequence, the movement of the nuclei is much slower than the movement of the electrons, so that the electrons feel the nuclei only as an approximately constant external potential.

The stationary Schrödinger equation in time for N_e interacting electrons in the external potential of N_n nuclei is therefore given by [68]

$$(\mathrm{T_e} + U_\mathrm{ee} + U_\mathrm{en})\Psi = E\Psi. \tag{9}$$

In this case, the many electron wave function $\Psi = \Psi(\mathbf{x}_1, \ldots, \mathbf{x}_{N_\mathrm{e}})$ depends on the electron coordinates $\mathbf{x}_i = (\mathbf{r}_i, s_i)$, which consist of the electron position $\mathbf{r}_i$ and a quantum mechanic state such as the electron spin s_i. The Hamilton operator is split up into the kinetic energy operator $\mathrm{T_e}$ of the electrons, the electron-electron repulsion U_ee, and the electron-nucleus attraction U_ne.

These quantities are defined by [68]

$$\mathrm{T_e} = \sum_{i=1}^{N_\mathrm{e}} (-\frac{1}{2}\nabla_i^2) \tag{10}$$

$$U_\mathrm{ee} = \sum_{i}^{N_\mathrm{e}} \sum_{j<i}^{N_\mathrm{e}} \frac{1}{\mathbf{r}_i - \mathbf{r}_j} \tag{11}$$

$$U_\mathrm{en} = \sum_{i}^{N_\mathrm{e}} \underbrace{(\sum_{\alpha}^{N_\mathrm{n}} \frac{-Z_\alpha}{|\mathbf{r}_i - \mathbf{R}_\alpha|})}_{V_\mathrm{ext}(\mathbf{r}_i)}, \tag{12}$$

whereas ∇_i is the nabla operator acting on the electron coordinate $\mathbf{r}_i$, Z_α is the charge of the nucleus α, R_α is the position of the nucleus α, and V_ext is the external potential caused by the nuclei.

If nucleus-nucleus interaction is treated classically, the total energy can simply be obtained by adding the nucleus-nucleus interaction [68]

$$U_\mathrm{nn} = \sum_{\alpha}^{N_\mathrm{n}} \sum_{\beta<\alpha}^{N_\mathrm{n}} \frac{Z_\alpha Z_\beta}{\mathbf{R}_\alpha - \mathbf{R}_\beta}. \tag{13}$$

to the energy of the N_e electron system

$$E_\mathrm{tot} = E + U_\mathrm{nn}. \tag{14}$$

The Hamilton operator H of the time independent Schrödinger equation, compare Eq. (9), is linear and Hermitian, and hence the eigen functions form an orthogonal set observing [68]

$$\int d\mathbf{x}^{N_\mathrm{e}}\, \Psi_k^*(\mathbf{x}_1, \ldots, \mathbf{x}_{N_\mathrm{e}}) \Psi_l(\mathbf{x}_1, \ldots, \mathbf{x}_{N_\mathrm{e}}) = \langle \Psi_k | \Psi_l \rangle = \delta_{kl}. \tag{15}$$

Here $\int d\mathbf{x}^{N_\mathrm{e}} = \int d\mathbf{x}_1 \cdots \int d\mathbf{x}_{N_\mathrm{e}} = \sum_{s_i} \int d\mathbf{r}_1 \cdots \sum_{s_{N_\mathrm{e}}} \int d\mathbf{r}_{N_\mathrm{e}}$ means the integration over all possible electron positions and the summation over all possible quantum states. By using an orthonormal basis set, the expectation value [68]

$$A[\Psi] = \langle A \rangle = \frac{\int d\mathbf{x}^{N_\mathrm{e}}\, \Psi^*(\mathbf{x}_1, \ldots, \mathbf{x}_{N_\mathrm{e}}) \mathrm{A} \Psi(\mathbf{x}_1, \ldots, \mathbf{x}_{N_\mathrm{e}})}{\int d\mathbf{x}^{N_\mathrm{e}}\, \Psi^*(\mathbf{x}_1, \ldots, \mathbf{x}_{N_\mathrm{e}}) \Psi(\mathbf{x}_1, \ldots, \mathbf{x}_{N_\mathrm{e}})} = \frac{\langle \Psi | \mathrm{A} | \Psi \rangle}{\langle \Psi | \Psi \rangle}, \tag{16}$$

of an observable A is simplified to $A[\Psi] = \langle \Psi | \mathrm{A} \, | \Psi \rangle$. Since the expectation values do not only depend on the operator but also on the wave function, they are functionals of Ψ. Specifically, it is possible to define the general energy functional of the many electron system by [68]

$$E[\Psi] = \langle \Psi | \mathrm{H} | \Psi \rangle. \tag{17}$$

The ground state of the system is given by the eigenstate Ψ_0 that fulfills the Schrödinger equation and minimizes the energy functional [68].

$$E_0 = E[\Psi_0] = \min_{\Psi} E[\Psi]. \tag{18}$$

Since the other eigenstates Ψ_k of the stationary Schrödinger equation are also stable points of the energy function $E[\Psi]$, all eigenstates can be identified by finding the solutions of the variational principle $\delta E[\Psi] = 0$.

2.2 The Hartree-Fock (HF) approximation

In the HF approximation, the wave function of the many electron system Ψ_{HF} is described by the Slater-Determinant [68, 71, 72], which is a sum of products of one-electron spin orbitals $\psi_i(\mathbf{x}_i)$.

$$\Psi_{\mathrm{HF}} = \frac{1}{\sqrt{N_{\mathrm{e}}!}} \begin{vmatrix} \psi_1(\mathbf{x}_1) & \psi_2(\mathbf{x}_1) & \cdots & \psi_{N_{\mathrm{e}}}(\mathbf{x}_1) \\ \psi_1(\mathbf{x}_2) & \psi_2(\mathbf{x}_2) & \cdots & \psi_{N_{\mathrm{e}}}(\mathbf{x}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \psi_1(\mathbf{x}_{N_{\mathrm{e}}}) & \psi_2(\mathbf{x}_{N_{\mathrm{e}}}) & \cdots & \psi_{N_{\mathrm{e}}}(\mathbf{x}_{N_{\mathrm{e}}}) \end{vmatrix} \tag{19}$$

The prefactor $\frac{1}{\sqrt{N_{\mathrm{e}}!}}$ normalizes the many electron wave function, and the spin orbitals consist of a spatial orbital $\psi(\mathbf{r})$ and, depending on the electron spin, of a spin function $\alpha(s)$ or $\beta(s)$. Specifically for $k = 1 \ldots N_{\mathrm{e}}/2$ and an even number of electrons, the N_{e} spin orbitals are constructed by

$$\psi_{2k-1}(\mathbf{x}) = \alpha(s)\psi_i(\mathbf{r}) \tag{20}$$

$$\psi_{2k}(\mathbf{x}) = \beta(s)\psi_i(\mathbf{r}). \tag{21}$$

The properties of the determinant correctly reproduces the anti-symmetry of the electrons and Pauli's exclusion principle [70]. In particular, this means that the determinant changes its sign whenever two electrons are exchanged

$$\Psi_{\mathrm{HF}}(\mathbf{x}_1, \ldots, \mathbf{x}_i, \ldots, \mathbf{x}_j, \ldots, \mathbf{x}_{N_{\mathrm{e}}}) = -\Psi_{\mathrm{HF}}(\mathbf{x}_1, \ldots, \mathbf{x}_j, \ldots, \mathbf{x}_i, \ldots, \mathbf{x}_{N_{\mathrm{e}}}) \tag{22}$$

and that it is equal to zero whenever two electrons are at the same position and in the same quantum state

$$\Psi_{\mathrm{HF}}(\mathbf{x}_1, \ldots, \mathbf{x}_i, \ldots, \mathbf{x}_j, \ldots, \mathbf{x}_{N_{\mathrm{e}}}) = 0 \quad \text{if } \mathbf{x}_i = \mathbf{x}_j. \tag{23}$$

When the Slater-Determinant is insert in Eq. (17), the total energy of the many electron system in the HF approximation is given by [68]

$$E_{\mathrm{HF}}[\psi] = \langle \Psi_{\mathrm{HF}} | \mathrm{H} | \Psi_{\mathrm{HF}} \rangle = \sum_{i=1}^{N_{\mathrm{e}}} H_i + \frac{1}{2} \sum_{i=1}^{N_{\mathrm{e}}} \sum_{j=1}^{N_{\mathrm{e}}} (J_{ij} - K_{ij}), \tag{24}$$

where

$$H_i = \int d\mathbf{x}\, \psi_i^*(\mathbf{x}) \Big[-\frac{1}{2}\nabla_i^2 + V_{\text{ext}}(\mathbf{r}) \Big] \psi_i(\mathbf{x}) \tag{25}$$

$$J_{ij} = \int d\mathbf{x}\, d\mathbf{x}'\, \psi_i(\mathbf{x}) \psi_i^*(\mathbf{x}) \frac{1}{|\mathbf{r}-\mathbf{r}'|} \psi_j^*(\mathbf{x}') \psi_j(\mathbf{x}') \tag{26}$$

$$K_{ij} = \int d\mathbf{x}\, d\mathbf{x}'\, \psi_i^*(\mathbf{x}) \psi_j(\mathbf{x}) \frac{1}{|\mathbf{r}-\mathbf{r}'|} \psi_i(\mathbf{x}') \psi_j^*(\mathbf{x}'). \tag{27}$$

Note that the H_i are one-electron integrals, while the Coulomb integrals J_{ij} and exchange integrals K_{ij} are two-electron integrals.

The minimization of Eq. (24) under the constraints of orthonormal one-electron wave functions $\int \psi_i(\mathbf{x}) \psi_j(\mathbf{x}) d\mathbf{x} = \delta_{ij}$ then leads to the set of one-electron HF differential equations [68, 70]

$$\mathrm{F} \psi_i(\mathbf{x}) = \sum_{j=1}^{N_\mathrm{e}} \varepsilon_{ij}^{\mathrm{HF}} \psi_j(\mathbf{x}). \tag{28}$$

Here the Fock operator F is given by

$$\mathrm{F} = -\frac{1}{2}\nabla^2 + V_{\text{ext}}(\mathbf{r}) + \mathrm{j} - \mathrm{k}, \tag{29}$$

and the Coulomb operator j and the exchange operator k are defined by [68]

$$\mathrm{j}(\mathbf{x}) f(\mathbf{x}) = \sum_{k=1}^{N_\mathrm{e}} \int d\mathbf{x}'\, \psi_k^*(\mathbf{x}') \psi_k(\mathbf{x}') \frac{1}{|\mathbf{r}-\mathbf{r}'|} f(\mathbf{x}) \tag{30}$$

$$\mathrm{k}(\mathbf{x}) f(\mathbf{x}) = \sum_{k=1}^{N_\mathrm{e}} \int d\mathbf{x}'\, \psi_k^*(\mathbf{x}') f(\mathbf{x}') \frac{1}{|\mathbf{r}-\mathbf{r}'|} \psi_k(\mathbf{x}). \tag{31}$$

The operator k is considered to be non-local, since the evaluation for a particular coordinate $\mathbf{x}$ also depends on the integration over all other coordinates $\mathbf{x}'$.

Since there are many possible solution of the HF equations, it is always possible to find a unitary transformation of the orbitals [73, 74]

$$\psi_i'(\mathbf{x}) = \sum_i U_{ij} \psi_i(\mathbf{x}) \text{ with } \sum_j U_{ij}^* U_{jk} = \delta_{ik}, \tag{32}$$

which does not change the HF many electron wave function Ψ_{HF}, but diagonalizes the matrix ϵ_{ij} and leads to the canonical HF equations [73, 74]

$$\mathrm{F} \psi_i'(\mathbf{x}) = \epsilon_{ii}' \psi_i'(\mathbf{x}). \tag{33}$$

Ultimately, the total energy of the many electron system in the HF approximation can be calculated in terms of the eigenvalues of the canonical HF equation minus one halve of the Coulomb and the exchange integrals [68]

$$E_{\mathrm{HF}} = \sum_i^{N_\mathrm{e}} \epsilon_{ii} - \frac{1}{2} \sum_{i=1}^{N_\mathrm{e}} \sum_{j=1}^{N_\mathrm{e}} (J_{ij} - K_{ij}). \tag{34}$$

According to Koopmans' theorem [75], the eigenvalues $\epsilon_{\rm v}$ of the valence band (VB) can be interpreted as negative ionization energies I. These are the energies needed to remove one-electron from a occupied one-electron wave function $\psi_i'(\mathbf{x})$, while the other $N-1$ electron states remain frozen [73]

$$I = E_{\rm v}^{N-1} - E_0^N = -\epsilon_{\rm v}. \tag{35}$$

Likewise, eigenvalues $\epsilon_{\rm c}$ of the conduction band (CB) can be considered as negative electron affinities A, when one-electron is add to an unoccupied one-electron wave function, while the other N electron states remain frozen [73]

$$A = E_0^N - E_{\rm c}^{N+1} = -\epsilon_{\rm c}. \tag{36}$$

Although the HF approximation explicitly treats the exchange effects by anti-symmetric wave functions, the method does not incorporate all electron-electron correlation effects such as correlation between electrons with opposite spins. Furthermore, the HF ansatz is much too reestrictive, since general many electron wave functions $\Psi(\mathbf{x}_1, \ldots, \mathbf{x}_n)$ cannot be represented by a single slater determinant. The difference between the exact total energy E of the many electron system and the HF energy $E^{\rm HF}$ is defined as the correlation energy [68, 73]

$$E_{\rm c}^{\rm HF} = E - E_{\rm HF}. \tag{37}$$

2.3 Density functional theory (DFT)

In 1964, Hohenberg and Kohn laid the foundation of density functional theory (DFT) by introducing two fundamental theorems. Let us consider a system of $N_{\rm e}$ electrons in an external potential $V_{\rm ext}$. Then Hohenberg and Kohn have shown (i) that the external potential $V_{\rm ext}$ is an unique functional of the electronic density $n(\mathbf{r})$ and (ii) that there exists a universal functional of the density $F[n]$, so that the exact ground state energy E_0 of the $N_{\rm e}$ electron system can be obtained by minimizing the energy functional [76]

$$E[n] = \int d\mathbf{r}\, V_{\rm ext}(\mathbf{r}) n(\mathbf{r}) + F[n]. \tag{38}$$

In DFT, the electronic density $n(\mathbf{r})$ is introduced as natural variable, as it determines the number of electrons $N_{\rm e} = \int d\mathbf{r}\, n(\mathbf{r})$ and the external potential $V_{\rm ext}$, which in turn determine all electronic properties of the systems. Furthermore, it is possible to to define the Hartree potential by

$$V_{\rm H}(\mathbf{r}) = \int d\mathbf{r}'\, \frac{n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \tag{39}$$

and to write

$$
\begin{aligned}
E_{\mathrm{en}}[n] &= \int d\mathbf{r}\, V_{\mathrm{ext}}(\mathbf{r}) n(\mathbf{r}) && (40)\\
F[n] &= E_{\mathrm{ee}}[n] + G[n] && (41)\\
E_{\mathrm{ee}}[n] &= \frac{1}{2}\int d\mathbf{r}\, V_{\mathrm{H}}(\mathbf{r}) n(\mathbf{r}). && (42)
\end{aligned}
$$

The energy functional is now given by [68]

$$
E[n] = E_{\mathrm{en}}[n] + E_{\mathrm{ee}}[n] + G[n], \tag{43}
$$

where $E_{\mathrm{en}}[n]$ is the electron-nucleus interaction functional, $E_{\mathrm{ee}}[n]$ is the electrostatic energy, which is also called Hartree energy, and $G[n]$ is a still unknown functional of the electronic density. Since $G[n]$ describes all interactions beyond the electrostatic interaction, the functional must account for the kinetic energy of the electrons, as well as exchange and correlation effects. Minimization of energy functional $E[n]$ yields the exact ground state density n_0. However, the exact functional $G[n]$ is not known and approximations must be found [68].

2.3.1 Kohn-Sham density functional theory (KS-DFT)

In 1965, Kohn and Sham proposed an approximation for the functional $G[n]$, which was based on an independent particle picture. They defined $G[n]$ as

$$
G[n] = T_{\mathrm{s}}[n] + E_{\mathrm{xc}}[n], \tag{44}
$$

where $T_{\mathrm{s}}[n]$ is the single particle kinetic energy functional of non-interacting particles and $E_{\mathrm{xc}}[n]$ is the exchange correlation functional of the interacting system. With this ansatz the KS energy functional can be written as [77]

$$
E^{\mathrm{KS}}[n] = T_{\mathrm{s}}[n] + E_{\mathrm{ne}}[n] + E_{\mathrm{ee}}[n] + E_{\mathrm{xc}}[n] \tag{45}
$$

Furthermore, they defined the electronic density in terms of non-interacting wave functions [77] and, for a non-spin polarized system, this density is given by

$$
n(\mathbf{r}) = \sum_n f_n\, \psi_n^*(\mathbf{r}) \psi_n(\mathbf{r}). \tag{46}
$$

The one-electron wave functions $\psi_i(\mathbf{r})$ are called KS orbitals, and f_n are the occupation numbers of each electronic state i varying between 0 and 2. Applying these definitions for the functional G and the density $n(\mathbf{r})$, one can apply the variational principle to the KS energy function E^{KS} to obtain the so-called KS equations [77]:

$$
\begin{aligned}
\epsilon_n\, \psi_n(\mathbf{r}) &= \left(-\frac{1}{2}\nabla^2 + V_{\mathrm{eff}}(\mathbf{r}) \right) \psi_n(\mathbf{r})\\
V_{\mathrm{eff}}(\mathbf{r}) &= V_{\mathrm{ext}}(\mathbf{r}) + V_{\mathrm{H}}(\mathbf{r}) + V_{\mathrm{xc}}(\mathbf{r})
\end{aligned} \tag{47}
$$

Similar to the HF ansatz, these equations are a set of one-electron "Schrödinger-like" equations with a Hamiltonian that contains a kinetic energy operator and an effective potential V_{eff}. The effective potential is the sum of the external potential V_{ext}, the Hartree-potential V_{H} and the so-called exchange-correlation potential defined by

$$V_{\text{xc}}(\mathbf{r}) = \frac{\delta E_{\text{xc}}[n]}{\delta n(\mathbf{r})}. \tag{48}$$

The energy values ϵ_n corresponding to the KS orbitals are called KS eigenvalues.

Since V_{eff} depends on $\psi_n(\mathbf{r})$ through the definition of $n(\mathbf{r})$, it is necessary to solve this set of equations selfconsistently, using some approximation for E_{xc}. The iterative solution proceeds usually along the following lines. Firstly, the electron density can be constructed from overlapping atomic densities. Then the KS Hamiltonian $\text{H}^{\text{KS}} = -\frac{1}{2}\nabla^2 + V_{\text{eff}}$ is constructed and the eigenvalue problem is solved by variation of $\psi_n(\mathbf{r})$ or by digitalization of H^{KS}. This results in a new set of KS orbitals from which one can construct a new density and thus a new Hamiltonian. These steps are repeated until $\psi_n(\mathbf{r})$, $n(\mathbf{r})$, and H^{KS} stay constant within a certain limit of accuracy.

Once the KS orbitals, which give the correct ground state density n_0 are found, Eq. (124) can be rewritten as [77]

$$E = \sum_i f_n\, \epsilon_n - E_{\text{ee}}[n] \tag{49}$$

since [68]

$$\sum_n f_n\, \epsilon_n = \sum_n f_n\, \langle \psi_n | \frac{1}{2}\nabla^2 + V_{\text{eff}} | \psi_n \rangle = T_{\text{s}}[n] + E_{\text{ne}}[n] + 2E_{\text{ee}}[n] + E_{\text{xc}}[n]. \tag{50}$$

All in all, the KS orbitals $\psi_n(\mathbf{r})$ and the associated eigenvalues ϵ_n are expedient tools to calculate the ground state energy E_0 and the corresponding ground state electronic density $n_0(\mathbf{r})$. In general, the orbitals and the eigenvalues have no physical interpretation. However for isolators and semiconductors, the maximum occupied KS orbital energy ϵ_{max} can be interpreted as the negative ionization potential IP or the negative electron affinity EA [78–80].

$$\epsilon_{\text{max}}(X) = -I(N_{\text{e}}), \quad \text{if } N_{\text{e}} - 1 < X < N_{\text{e}} \tag{51}$$

$$\epsilon_{\text{max}}(X) = -A(N_{\text{e}}), \quad \text{if } N_{\text{e}} < X < N_{\text{e}} + 1 \tag{52}$$

For metals, ϵ_{max} is the negative of the work function. Since the interpretation of the eigenvalues is in this case similar to the HF eigenvalues, this relationship is often called Koopmans' theorem for DFT. However, we must keep in mind that Koopmans' theorem applies for all occupied HF eigenvalues, while the DFT counterpart applies only for the highest occupied KS eigenvalue ϵ_{max}.

The DFT band gap is defined as the energy difference between the eigenvalues of the highest occupied molecular orbital (HOMO) and the lowest unoccupied molecular orbital (LUMO), or as the difference between the ionization energy and the electron affinity. Though it turns out that DFT strongly underestimates the band gap in comparison to experimental measurements [81]. This band gap problem is typically explained in terms of the derivative discontinuity C of the exact exchange correlation function, which occurs when the number of electrons is increased from $N_\mathrm{e} - \varepsilon$ to $N_\mathrm{e} + \varepsilon$, where ε is an infinitesimal non-integer value [79]

$$\left.\frac{\delta E_\mathrm{xc}[n]}{\delta n(\mathbf{r})}\right|_{N_\mathrm{e}-\varepsilon} - \left.\frac{\delta E_\mathrm{xc}[n]}{\delta n(\mathbf{r})}\right|_{N_\mathrm{e}+\varepsilon} = C. \tag{53}$$

Since the derivative discontinuity C must enter in the calculation of the band gap

$$E_\mathrm{g} = \epsilon_\mathrm{CBM}(N_\mathrm{e}) - \epsilon_\mathrm{VBM}(N_\mathrm{e}) + C, \tag{54}$$

and approximate functionals do not allow to determine C, the KS band gap underestimates the real band gap. However, it is sometimes reasoned that the gap center relative to the vacuum level is correctly reproduced by KS-DFT [79].

2.3.2 The exchange correlation functional: from LDA to PBEsol

In KS-DFT, the system of N_e interacting electrons is mapped to a system of N_e non-interacting electrons with the same density as the original many body problem. DFT uses severe simplifications in comparison to the original many body equations. Firstly, the DFT wave function

$$\Psi^\mathrm{KS}(\mathbf{x}_1, \ldots, \mathbf{x}_{N_\mathrm{e}}) = \prod_n \psi_n(\mathbf{r}) \tag{55}$$

is not antisymmetric with respect to the exchange of two particles. Secondly, the movement of the electrons is uncorrelated at the first sight. Moreover, the real kinetic energy operator T and the single particle kinetic energy operator $\mathrm{T_s}$ of non-interacting electrons differ. All these effects and differences are accounted for by the exchange correlation functional E_xc. If we knew E_xc, KS-DFT would describe the original many body problem exactly.

Unfortunately, the exact form of the exchange correlation functional is unknown and approximations are needed. Assuming the local density $n(\mathbf{r})$ to be locally uniform, we can, for example, approximate E_xc in the local density approximation (LDA) [76, 77]

$$E_\mathrm{xc}^\mathrm{LDA}[n] = \int d\mathbf{r}\ \epsilon_\mathrm{xc}^\mathrm{HEG}\big[n(\mathbf{r})\big]\ n(\mathbf{r}). \tag{56}$$

The exchange correlation energy per electron $\epsilon_\mathrm{xc}^\mathrm{HEG}$ is thereby derived from a homogeneous electron gas (HEG) with a density equal the local density $n(\mathbf{r})$. Usually, the local energy

contribution is split up into an exchange and correlation part [68, 74]

$$\epsilon_{\rm xc}^{\rm HEG}\big[n(\mathbf{r})\big] = \epsilon_{\rm x}^{\rm HEG}\big[n(\mathbf{r})\big] + \epsilon_{\rm c}^{\rm HEG}\big[n(\mathbf{r})\big], \tag{57}$$

where $\epsilon_{\rm x}^{\rm HEG}$ is given by the Dirac exchange energy functional [82]

$$\epsilon_{\rm x}^{\rm HEG}\big[n(\mathbf{r})\big] = -\frac{3}{4}(\frac{3}{\pi})^{\frac{1}{3}}\Big(n(\mathbf{r})\Big)^{\frac{1}{3}}, \tag{58}$$

and $\epsilon_{\rm c}^{\rm HEG}$ by an analytic fit to quantum Monte Carlo simulations [83–85].

Although the LDA can be applied to systems with nearly homogeneous local densities, more accurate methods are needed for systems with inhomogeneous local densities. Such inhomogeneities are, for instance, located near atoms, defects, interfaces, and surfaces.

The exact exchange correlation energy can be represented in terms of the exchange correlation hole [86]:

$$E_{\rm xc}[n] = \frac{1}{2}\int d\mathbf{r}d\mathbf{r}'\frac{n(\mathbf{r})n_{\rm xc}(\mathbf{r},\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}. \tag{59}$$

The exchange correlation hole $n_{\rm xc}(\mathbf{r},\mathbf{r}')$ accounts for an electronic depletion zone in the vicinity of negative charges due to Coulomb and Pauli exclusion principle. As for the exchange correlation energy per atom, the hole can be split up into an exchange and correlation part $n_{\rm xc} = n_{\rm x} + n_{\rm c}$. Both parts satisfy the following important properties. By integration over the depletion zone, the exchange part compensates exactly one-electron $\int d\mathbf{r}' n_{\rm x}(\mathbf{r},\mathbf{r}') = -1$, while the correlation part vanishes $\int d\mathbf{r}' n_{\rm c}(\mathbf{r},\mathbf{r}') = 0$. Furthermore, the exchange part is alway less or equal zero $n_{\rm x}(\mathbf{r},\mathbf{r}') \leq 0$.

Perdew *et al.* parametrized $n_{\rm xc}$ and introduced a real space cut off that restores some properties of the exchange correlation hole. This approach is called the generalized gradient approximation (GGA) and is the next practical ansatz beyond the LDA. The resulting functional depends on the electronic density and its local gradient [87]:

$$E_{\rm xc}^{\rm PBE}[n] = E_{\rm x}^{\rm PBE}[n] + E_{\rm c}^{\rm PBE}[n] \tag{60}$$

$$E_{\rm x}^{\rm PBE}[n] = \int d\mathbf{r}\,\epsilon_{\rm x}^{\rm HEG}\big[n(\mathbf{r})\big]\; F_{\rm x}\big[n(\mathbf{r}),\nabla n(\mathbf{r})\big]\; n(\mathbf{r}) \tag{61}$$

$$E_{\rm c}^{\rm PBE}[n] = \int d\mathbf{r}\left(\epsilon_{\rm c}^{\rm HEG}\big[n(\mathbf{r})\big] + H\big[n(\mathbf{r}),\nabla n(\mathbf{r})\big]\right) n(\mathbf{r}). \tag{62}$$

Here the gradient contributions of the enhancement factor F_x and the factor H are given by a simple set of parameters for the spin unpolarized-case that are completely defined in Ref. 87–89.

Note that there are several other GGA functionals in literature, but Perdew *et al.* showed that there are "natural" limitations on the accuracy of the GGA [89]. For example, corresponding

functionals can accurately calculate either total energies or lattice parameters, but not both. This is related to the gradient expansion for slow varying densities. The expansion can be either violated in favor of accurate energies or restored in favor of accurate lattice parameters and surface energies. A good comparison between modern GGAs is, for example, given by Csonka *et al.* [90].

Furthermore, there are two heavily used functionals in literature that are based on the functional form suggested by Perdew, Burke, and Ernzerhof (PBE). The first one is optimized for atoms and molecules and is the so-called PBE functional [87]. The second is optimized for solids and surfaces and is the so-called PBEsol functional [89]. Although both are optimized for different conditions, they both are based on the formulas in Eq. (61) and Eq. (62) and differ only in few parameters. In the present study, we exclusively use the PBEsol functional for solids and surfaces.

2.3.3 Heyd Scuseria and Ernzerhof (HSE) hybrid functional

The HF approximation treats the exchange effects of the electrons directly using anti-symmetric wave-functions, but it completely neglects the correlation of the electron movement. KS-DFT, on the other hand, corrects for exchange and correlation effects through an exchange-correlation functional E_{xc}. Since the HF approximation overestimates the band gap, while KS-DFT underestimates it, both methods are sometimes combined to build hybrid exchange-correlation functionals [70]

$$E_{\mathrm{xc}}^{\mathrm{hybrid}} \quad = \quad \alpha E_{\mathrm{x}}^{\mathrm{HF}} + (1-\alpha) E_{\mathrm{x}}^{\mathrm{DFT}} + E_{\mathrm{c}}^{\mathrm{DFT}}, \tag{63}$$

which can often improve upon both methods [91]. The Fock exchange energy E_{x} is given by [91]

$$E_{\mathrm{x}}^{\mathrm{HF}} = -\frac{1}{2} \sum_{i,j} \int d\mathbf{x}\, d\mathbf{x}'\, f_i\, f_j\, \psi_i^*(\mathbf{x}) \psi_j(\mathbf{x}) \frac{1}{|\mathbf{r}-\mathbf{r}'|} \psi_j^*(\mathbf{x}') \psi_i(\mathbf{x}'). \tag{64}$$

The summation goes over occupied states i with one-electron wave functions ψ_i and ψ_j and occupation numbers f_i and f_j. As in section 2.2, the integration implies the integration over the spatial coordinates $\mathbf{r}$ and $\mathbf{r}'$ and the summation over the electron spins s and s'. The mixing between the HF approximation and the KS-DFT in the exchange part is controlled by the parameter α, but the correlation part is always calculated using KS-DFT. The PBE0 hybrid functional, e.g., uses a mixing parameter of $\alpha = 1/4$ and the PBE functional that we have defined in Eq. (60) to Eq. (62). [91, 92].

Since the calculation of the long range exchange interactions can be computational demanding [91], Heyd Scuseria and Ernzerhof (HSE) suggested a hybrid functional where the exchange

interaction is split into a short range (SR) and long range (LR) part [91, 93]. This so-called "range-separated" hybrid functional is generated by splitting the Coulomb kernel like

$$\frac{1}{|\mathbf{r}-\mathbf{r}'|} = \underbrace{\frac{\mathrm{erfc}(\mu|\mathbf{r}-\mathbf{r}'|)}{|\mathbf{r}-\mathbf{r}'|}}_{\mathrm{SR}} + \underbrace{\frac{\mathrm{erf}(\mu|\mathbf{r}-\mathbf{r}'|)}{|\mathbf{r}-\mathbf{r}'|}}_{\mathrm{LR}}. \tag{65}$$

Here μ defines the characteristic length of the SR exchange interaction [94–96]. Only the SR exchange part is now calculated by mixing the HF and PBE exchange energies $E_{\mathrm{x}}^{\mathrm{HF,SR}}$ and $E_{\mathrm{x}}^{\mathrm{PBE,SR}}$, while the LR exchange part and the correlation energy are calculated in KS-DFT using the PBE functionals $E_{\mathrm{x}}^{\mathrm{PBE,LR}}$ and $E_{\mathrm{c}}^{\mathrm{PBE}}$.

$$E_{\mathrm{xc}}^{\mathrm{HSE}} = \alpha E_{\mathrm{x}}^{\mathrm{HF,SR}}(\mu) + (1-\alpha)E_{\mathrm{x}}^{\mathrm{PBE,SR}}(\mu) + E_{\mathrm{x}}^{\mathrm{PBE,LR}}(\mu) + E_{\mathrm{c}}^{\mathrm{PBE}} \tag{66}$$

As the HSE functional depends on two parameters α and μ, the proportion of the mixing and the boundary between the SR and LR interaction regime can in principle be controlled. For $\alpha = 1/4$ and $\mu = 0$ the HSE functional is equivalent to the PBE0 functional, and, for $\alpha = 1/4$ and $\mu \to \infty$, it equals PBE. A good compromise between calculation speed and accuracy is, for example, achieved for $\alpha = 1/4$ and $\mu = 0.2$ Å^{-1} [91, 93]. For a closed shell system, $E_{\mathrm{x}}^{\mathrm{HF,SR}}$ is then calculated by [91]

$$E_{\mathrm{x}}^{\mathrm{HF,SR}} = -\frac{1}{2}\sum_{i,j}\int d\mathbf{r}\, d\mathbf{r}'\, f_i\, f_j\, \psi_i^*(\mathbf{r})\psi_j(\mathbf{r})\frac{\mathrm{erfc}(\mu|\mathbf{r}-\mathbf{r}'|)}{|\mathbf{r}-\mathbf{r}'|}\psi_j^*(\mathbf{r}')\psi_i(\mathbf{r}'). \tag{67}$$

To construct $E_{\mathrm{x}}^{\mathrm{PBE,SR}}$ Heyd *et al.* modified the enhancement factor F_x, compare Eq. (61). The resulting PBE functional for short-range exchange interaction is

$$E_{\mathrm{x}}^{\mathrm{PBE,SR}}[n] = \int d\mathbf{r}\, \epsilon_{\mathrm{x}}^{\mathrm{HEG}}\big[n(\mathbf{r})\big]\; F_{\mathrm{x}}^{\mathrm{SR}}\big[n(\mathbf{r}), \nabla n(\mathbf{r})\big]\; n(\mathbf{r}), \tag{68}$$

whereas $F_{\mathrm{x}}^{\mathrm{SR}}$ is constructed by integrating a screened PBE exchange hole. The LR PBE functional is, finally, constructed by the difference between the total and the SR exchange PBE functional

$$E_{\mathrm{x}}^{\mathrm{PBE,LR}}[n] = \int d\mathbf{r}\, \epsilon_{\mathrm{x}}^{\mathrm{HEG}}\big[n(\mathbf{r})\big]\; \Big(F_{\mathrm{x}}\big[n(\mathbf{r}), \nabla n(\mathbf{r})\big] - F_{\mathrm{x}}^{\mathrm{SR}}\big[n(\mathbf{r}), \nabla n(\mathbf{r})\big]\Big)\; n(\mathbf{r}). \tag{69}$$

The detail derivation of $F_{\mathrm{x}}^{\mathrm{SR}}$ is given in Ref. 93 and in Ref. 97.

2.4 The GW method

The processes in the photoemission spectroscopy (PES) and the inverse photoemission spectroscopy (IPES) generate holes in the valence band (VB) and electrons in the conduction band (CB), respectively. In PES, an incident beam of photons hits the surface of the material and

knocks electrons out of the VB. Since the energy of the photons ω is known and the kinetic energy of the emitted electrons E_{kin} is measured, the corresponding energy levels of the VB states ϵ_n^{VB} can be obtained from the law of energy conservation.

$$\epsilon_n^{\text{VB}} = \omega - E_{\text{kin}} \tag{70}$$

Similarly, the energy levels of the CB states ϵ_n^{CB} can be obtained by IPES. Here a beam of electrons with well-defined kinetic energy E_{kin} hits the material and the impinging electrons relax into unoccupied CB states by emitting photons with energy ω. The energy levels can then be calculated by

$$\epsilon_n^{\text{CB}} = \omega - E_{\text{kin}}. \tag{71}$$

Informations from PES and IPES can in principle be combined to measure the total density of states $D(\omega)$, which can, for a translationally invariant system, be calculated by the integration of the spectral function $A(\mathbf{k}, \omega)$ [98].

$$D(\omega) = \int d\mathbf{k}\ A(\mathbf{k}, \omega) \tag{72}$$

Were both processes of the measurements to be without any further electronic interactions, the spectral function would have the form of a delta distribution function $\delta(\omega - \epsilon_{\mathbf{k}})$. With interacting electrons, however, the function has the form of a Lorentzian energy distribution

$$A(\mathbf{k}, \omega) = \frac{1}{\pi} \frac{\Gamma_{\mathbf{k}}}{(\omega - \tilde{\epsilon}_{\mathbf{k}})^2 + \Gamma_{\mathbf{k}}^2}. \tag{73}$$

Compared to the delta function of the non-interacting system, the distribution has now a width of $2\Gamma_k$ and shifted peak positions $\tilde{\epsilon}_{\mathbf{k}}$ [98,99]. In the next two sections, we discuss how the spectral function can be expressed in terms of a one particle Green's function.

2.4.1 The one particle causal Green's function

In an N electron system at ground state ψ_0, the one particle Green's function describes how one extra electron/hole propagates through the system. The probability where the extra particle can be found later depends on the interactions during the propagation. The formula of the one particle causal (i.e. time-ordered) Green's function G defines the probability amplitudes of electrons and holes at once [99, 100].

$$G(\mathbf{r}, t, \mathbf{r}', t') = \quad - \quad i\langle\psi_0|\, \mathrm{T}\, \hat{\psi}(\mathbf{r}, t)\hat{\psi}^\dagger(\mathbf{r}', t')\, |\psi_0\rangle \tag{74}$$

For $t' < t$, G defines the probability amplitude of creating an electron at $(\mathbf{r}', t')$ and finding it later at $(\mathbf{r}, t)$ and, for $t < t'$, it defines the probability amplitude of creating a hole at $(\mathbf{r}, t)$

and finding it later at $(\mathbf{r}', t')$. In particular, the Green's function is the expectation value of a time-ordered product of the creation operator $\hat{\psi}^\dagger(\mathbf{r}, t) = e^{i\mathrm{H}t}\hat{\psi}^\dagger(\mathbf{r})e^{-i\mathrm{H}t}$ and the destruction operator $\hat{\psi}(\mathbf{r}, t) = e^{i\mathrm{H}t}\hat{\psi}(\mathbf{r})e^{-i\mathrm{H}t}$ in the Heisenberg picture. For causal reasons, the time ordering operator T always shifts the operators that must be executed first to the front. For example, for two operators $\mathrm{A}(\mathbf{r}, t)$ and $\mathrm{B}(\mathbf{r}', t')$ this means [100]

$$\mathrm{T}\,\mathrm{A}(\mathbf{r}, t)\mathrm{B}(\mathbf{r}', t') = \Theta(t - t')\,\mathrm{A}(\mathbf{r}, t)\mathrm{B}(\mathbf{r}', t') - \Theta(t' - t)\,\mathrm{B}(\mathbf{r}', t')\mathrm{A}(\mathbf{r}, t). \tag{75}$$

As a consequence, Eq. (91) can be written as [100]

$$G(\mathbf{r}, t, \mathbf{r}', t') = - \; i\,\Theta(t - t')\,\langle\psi_0|\,\hat{\psi}(\mathbf{r}, t)\hat{\psi}^\dagger(\mathbf{r}', t')\,|\psi_0\rangle \tag{76}$$

$$+ \; i\,\Theta(t' - t)\,\langle\psi_0|\,\hat{\psi}^\dagger(\mathbf{r}', t')\hat{\psi}(\mathbf{r}, t)\,|\psi_0\rangle. \tag{77}$$

In a many electron Fermi system, the sign changes according to the number of permutations that are needed to shift all operators in the proper order.

2.4.2 The Lehman/spectral representation for non-interacting particles

We can distinguish two kind of probability amplitudes. First, the probability amplitudes for transitions of the ground state ψ_0^N, with N-particles, to excited states ψ_n^{N-1}, with $N-1$-particles. Second, the probability amplitudes for transitions of excited states ψ_n^{N+1}, with $N+1$-particles, to the ground state ψ_0^N, with N-particles. The corresponding energy differences between the initial and the final states are $\epsilon_n^{\mathrm{VB}} = E_n^{N-1} - E_0^N$ and $\epsilon_n^{\mathrm{CB}} = E_n^{N+1} - E_0^N$, respectively. These energies in principle correspond to the VB and the CB energies measured in PES and IPES measurements. By inserting a complete set of corresponding eigenstates $\sum_n |\psi_n^{N\pm1}\rangle\langle\psi_n^{N\pm1}|$ into the one particle Green's function of non-interacting particles G_0 and by using

$$\langle\psi_0^N|\hat{\psi}(\mathbf{r}, t)|\psi_n^{N\pm1}\rangle = \langle\psi_0^N|\hat{\psi}(\mathbf{r})|\psi_n^{N\pm1}\rangle\, e^{i(E_0^N - E_n^{N\pm1})t}, \tag{78}$$

we obtain [101]

$$G_0(\mathbf{r}, \mathbf{r}', t, t') = - \; i\sum_n \langle\psi_0^N|\,\hat{\psi}(\mathbf{r})|\psi_n^{N+1}\rangle\langle\psi_n^{N+1}|\hat{\psi}^\dagger(\mathbf{r}')\,|\psi_0^N\rangle e^{-i(E_n^{N+1} - E_0^N)(t-t')}\Theta(t - t') \tag{79}$$

$$+ \; i\sum_n \langle\psi_0^N|\,\hat{\psi}^\dagger(\mathbf{r}')|\psi_n^{N-1}\rangle\langle\psi_n^{N-1}|\hat{\psi}(\mathbf{r})\,|\psi_0^N\rangle e^{-i(E_0^N - E_n^{N-1})(t'-t)}\Theta(t' - t). \tag{80}$$

We can now introduce the chemical potentials μ^{N+1} and μ^N and the excitation energies ϵ_n^{N+1} and ϵ_n^{N-1} to rewrite the energy differences in the exponential functions [99].

$$E_n^{N+1} - E_0^N = \underbrace{(E_n^{N+1} - E_0^{N+1})}_{\epsilon_n^{N+1}} + \underbrace{(E_0^{N+1} - E_0^N)}_{\mu^{N+1} = \mu^N + O(\frac{1}{N})} \tag{81}$$

$$E_0^N - E_n^{N-1} = \underbrace{(E_0^N - E_0^{N-1})}_{\mu^N} - \underbrace{(E_n^{N-1} - E_0^{N-1})}_{\epsilon_n^{N-1}} \tag{82}$$

Since the chemical potentials μ^{N+1} and μ^N coincide for a large number of particles and the Green's function only depends on the time difference $\tau = t - t'$, we may define the real energy levels

$$\epsilon_n = \begin{cases} \mu - \epsilon_n^{N-1} \text{ for } & \epsilon_n \leq \mu \\ \mu + \epsilon_n^{N+1} \text{ for } & \epsilon_n > \mu \end{cases}, \tag{83}$$

in order to write

$$G_0(\mathbf{r}, \mathbf{r}', \tau) = - i \sum_n \langle \psi_0^N | \hat{\psi}(\mathbf{r}) | \psi_n^{N+1} \rangle \langle \psi_n^{N+1} | \hat{\psi}^\dagger(\mathbf{r}') | \psi_0^N \rangle e^{-i\epsilon_n \tau} \Theta(\tau) \tag{84}$$

$$+ i \sum_n \langle \psi_0^N | \hat{\psi}^\dagger(\mathbf{r}') | \psi_n^{N-1} \rangle \langle \psi_n^{N-1} | \hat{\psi}(\mathbf{r}) | \psi_0^N \rangle e^{-i\epsilon_n \tau} \Theta(-\tau). \tag{85}$$

Furthermore, we can use the identity [101]

$$\int_{-\infty}^{+\infty} d\tau \, e^{i\omega\tau} e^{-i\epsilon_n \tau} \Theta(\pm\tau) = \pm \frac{i}{\omega - \epsilon_n \pm i\eta} = \pm \frac{i}{\omega - \bar{\epsilon}_\alpha} \tag{86}$$

to Fourier transform $G(\mathbf{r}, \mathbf{r}', \tau)$. Here η is a infinitesimal number derived from a contour integration [101]. In the complex plane, it shifts the occupied states with $\epsilon_n \leq \mu$ a little bit over the real axis and the unoccupied states with $\epsilon_n > \mu$ a little bit under the real axis. The shifted energy values can then be written together using the complex energy values $\bar{\epsilon}_n = \epsilon_n \mp i\eta$. Applying this identity, we finally obtain the Lehman representation (or spectral representation) of the causal Green's function [101–105]

$$G_0(\mathbf{r}, \mathbf{r}', \omega) = \int_{-\infty}^{\mu_F} d\omega' \, \frac{A_0(\mathbf{r}, \mathbf{r}', \omega')}{\omega - \omega' - i\eta} + \int_{\mu_F}^{\infty} d\omega' \, \frac{A_0(\mathbf{r}, \mathbf{r}', \omega')}{\omega - \omega' + i\eta} = \sum_n \frac{\bar{\phi}_n(\mathbf{r}) \bar{\phi}_n^*(\mathbf{r}')}{\omega - \bar{\epsilon}_n}, \tag{87}$$

where

$$A_0(\mathbf{r}, \mathbf{r}', \omega) = \sum_n \bar{\phi}_n(\mathbf{r}) \bar{\phi}_n^*(\mathbf{r}') \delta(\omega - \epsilon_n) \tag{88}$$

is the spectral function of the non-interacting system and

$$\bar{\phi}_n(\mathbf{r}) = \begin{cases} \langle \psi_n^{N-1} | \hat{\psi}(\mathbf{r}) | \psi_0^N \rangle \text{ for } & \epsilon_n \leq \mu \\ \langle \psi_0^N | \hat{\psi}(\mathbf{r}) | \psi_n^{N+1} \rangle \text{ for } & \epsilon_n > \mu \end{cases} \tag{89}$$

are the probability amplitudes for adding and removing one particle. Now we have a relationship between the one particle causal Green's function of non-interacting particles $G_0(\mathbf{r}, \mathbf{r}, \omega)$ and a spectral function $A_0(\mathbf{r}, \mathbf{r}', \omega)$. In general, the spectral function can be obtained by applying the Kramers-Kronig relation $\frac{1}{\omega \pm i\eta} = P\frac{1}{\omega} \mp i\pi\delta(\omega)$ to Eq. (91) [101].

$$A_0(\mathbf{r}, \mathbf{r}', \omega) = \frac{-\operatorname{sgn}(\omega - \mu)}{\pi} \operatorname{Im} G_0(\mathbf{r}, \mathbf{r}', \omega) \tag{90}$$

Both functions are generally linked by the so-called quasiparticle wave functions $\bar{\phi}_n(\mathbf{r})$, compare Eq. (87) and Eq. (88). Describing rather one-particle transitions between the ground state and excited states, the orbitals are considered as quasi particles states with corresponding quasiparticle energies $\bar{\epsilon}_n = \epsilon_n \mp i\eta$.

2.4.3 The quasiparticle equation

For a system with interacting particles, the one particle Green's function can again be expressed in terms of quasiparticle orbitals $\bar{\psi}_n$ and quasiparticle energies $\bar{E}_n$, if the quasiparticle orbitals are complete, but unnormalized, and linear dependent. Specifically, the spectral representation is then [105]

$$G(\mathbf{r}, \mathbf{r}', \omega) = \sum_n \frac{\bar{\psi}_n(\mathbf{r})\bar{\psi}_n^*(\mathbf{r}')}{\omega - \bar{E}_n}, \tag{91}$$

and the spectral function can again derived by

$$A(\mathbf{r}, \mathbf{r}', \omega) = \frac{-\operatorname{sgn}(\omega - \mu)}{\pi} \operatorname{Im} G(\mathbf{r}, \mathbf{r}', \omega). \tag{92}$$

However, in comparison to a non-interacting system, the quasiparticle wave functions are different and the quasiparticle energies are complex poles with real parts that give the energy levels of the quasiparticles and imaginary parts that give the widths of the quasiparticle peaks. The imaginary parts are furthermore inverse proportional to the lifetimes of the quasiparticles. For the special case of non-interacting particles, the widths are infinitesimal small and the lifetimes are infinite [99]. The time development of the interacting Green's function is given by the so-called Dyson equation [105]

$$G(\mathbf{r}, \mathbf{r}', \omega) = G_0(\mathbf{r}, \mathbf{r}', \omega) + \int d\mathbf{r}_1\, d\mathbf{r}_2\, G_0(\mathbf{r}, \mathbf{r}_1, \omega)\Sigma(\mathbf{r}_1, \mathbf{r}_2, \omega)G(\mathbf{r}_2, \mathbf{r}', \omega), \tag{93}$$

which relates the interacting Green's function G with the non-interacting Green's function G_0 and the so-called self energy Σ. The self energy defines the general differences between interacting and non-interacting systems $\Sigma = G^{-1} - G_0^{-1}$. When the spectral representations of the Green's function are plugged into the Dyson equation, the so-called quasiparticle equation can be derived [105]

$$\bar{E}_n\bar{\psi}_n(\mathbf{r}) = \left(\frac{1}{2}\nabla^2(\mathbf{r}) + V_{\text{eff}}(\mathbf{r})\right)\bar{\psi}_n(\mathbf{r}) + \int d\mathbf{r}'\, \Sigma(\mathbf{r}, \mathbf{r}', \bar{E}_n)\, \bar{\psi}_n(\mathbf{r}') \tag{94}$$

$$V_{\text{eff}}(\mathbf{r}) = V_{\text{ext}}(\mathbf{r}) + V_{\text{H}}(\mathbf{r}). \tag{95}$$

Similar to the Kohn-Sham (KS) equation, compare Eq. (47), this equation is a kind of effective one particle "Schrödinger-like" equation. It consists of a kinetic energy operator $\frac{1}{2}\nabla^2$ and an effective potential V_{eff}, which contains the external potential V_{ext} and the Hartree potential V_{H}. In contrast to the KS equation, the exchange and correlation effects are now included through the self energy Σ. It turns out that the quasiparticle orbitals $\bar{\phi}_n(\mathbf{r})$ and energies $\bar{E}_n$ can be obtained solving the quasiparticle equation [102–104], but approximations for Σ are needed to

solve this problem. In the GW approximation with neglected vertex corrections in Σ, the self energy is given in terms of an one particle Green's function $G(\mathbf{r}, \mathbf{r}', \omega)$ and a screened Coulomb interaction $W(\mathbf{r}, \mathbf{r}', \omega)$ [74, 102]:

$$\Sigma(\mathbf{r}, \mathbf{r}', \omega) = \frac{i}{4\pi} \int d\omega' \, e^{i\eta\,\omega'} G(\mathbf{r}, \mathbf{r}', \omega') W(\mathbf{r}, \mathbf{r}', \omega - \omega'). \tag{96}$$

The screened interaction W is typically approximated by means of a "Dyson-like" equation with the Coulomb kernel, $V(\mathbf{r}, \mathbf{r}') = 1/|\mathbf{r} - \mathbf{r}'|$, and the irreducible polarizability $\tilde{\chi}$ [74].

$$W(\mathbf{r}, \mathbf{r}', \omega) = V(\mathbf{r}, \mathbf{r}') + \int d\mathbf{r}'' \, d\mathbf{r}''' \, V(\mathbf{r}, \mathbf{r}'') \tilde{\chi}(\mathbf{r}'', \mathbf{r}''', \omega) W(\mathbf{r}''', \mathbf{r}', \omega) \tag{97}$$

$$\tilde{\chi}(\mathbf{r}, \mathbf{r}', \omega) = -iG(\mathbf{r}, \mathbf{r}', \omega) G(\mathbf{r}, \mathbf{r}', \omega) \tag{98}$$

This is the so-called GW approximation. Since the Green's function is completely defined by the quasiparticle orbitals $\bar{\psi}_n(\mathbf{r})$ and quasiparticle energies $\bar{E}_n$, compare Eq. (91), the quasi particle equation can be solved iteratively, using Eq. (91) and Eq. (94) to (98). In practice, the screened potential W is implemented by calculating [74]

$$W(\mathbf{r}, \mathbf{r}', \omega) = \int d\mathbf{r}'' \varepsilon^{-1}(\mathbf{r}, \mathbf{r}'', \omega) V(\mathbf{r}'', \mathbf{r}'), \tag{99}$$

where the dielectric function is obtained applying the random phase approximation (RPA) or the plasmon pole approximation [106–108]. In our work, we use the symmetrized form of the dielectric function [108]

$$\varepsilon^{\mathrm{RPA}}_{\mathbf{q}}(\mathbf{G}, \mathbf{G}', \omega) = \delta_{\mathbf{G},\mathbf{G}'} - \frac{4\pi}{|\mathbf{q} + \mathbf{G}||\mathbf{q} + \mathbf{G}'|} \chi^{\mathrm{KS}}_{\mathbf{q}}(\mathbf{G}, \mathbf{G}', \omega), \tag{100}$$

where [108]

$$\chi^{\mathrm{KS}}_{\mathbf{q}}(\mathbf{G}, \mathbf{G}', \omega) = \lim_{\eta \to 0} \sum_{n,n',\mathbf{k}} (f_{n',\mathbf{k}-\mathbf{q}} - f_{n,\mathbf{k}}) \frac{\langle \psi_{n',\mathbf{k}-\mathbf{q}} | e^{-i(\mathbf{q}+\mathbf{G})\mathbf{r}} | \psi_{n,\mathbf{k}} \rangle \langle \psi_{n,\mathbf{k}} | e^{-i(\mathbf{q}+\mathbf{G}')\mathbf{r}'} | \psi_{n',\mathbf{k}-\mathbf{q}} \rangle}{\omega - (\epsilon_{n,\mathbf{k}} - \epsilon_{n',\mathbf{k}-\mathbf{q}}) + i\eta\,\mathrm{sgn}(\epsilon_{n',\mathbf{k}-\mathbf{q}} - \epsilon_{n,\mathbf{k}})} \tag{101}$$

is the time ordered KS polarizability. More details on the dielectric response function can be found in section 7.4. Results obtained by Kohn-Sham density functional theory (KS-DFT), are usually taken as first approximation to the quasiparticle energies and orbitals. After solving the quasiparticle equation, see Eq. (94), the self energy can then be updated in terms of the Green's function G and the screened Coulomb kernel W. In present work we use the GW_0 approach as described in Ref. 108 and Ref. 109. First, the KS-DFT orbitals and eigenvalues are used to calculate G_0 and W_0 for non-interacting particles. Afterwards, the eingenvalues are iterated to optain the quasiparticle energies and to update G, but W_0 is kept fix.

2.5 Single-electron wave functions in a periodic potential

Todays simulation methods can calculate up to several thousand atoms, but that are still too small system sizes in order to simulate bulk and interface systems, which contain in reality over 10^{23} atoms. Moreover, the system sizes of the simulations are usually so small that there are surface effects nearly everywhere in the sample. PBCs are thus applied to overcome these problems [110–112]. Under PBCs the system is represented by a unit cell, that is periodically repeated in space. The simulation cell is spanned by the cell vectors $\mathbf{a}_1$, $\mathbf{a}_2$, and $\mathbf{a}_3$, which also span a Bravais lattice by the following linear combination [5]:

$$\mathbf{R} = n_1\mathbf{a}_1 + n_2\mathbf{a}_2 + n_3\mathbf{a}_3. \tag{102}$$

Every lattice point $\mathbf{R}$ can then be constructed by this linear combination using the three integer coefficients n_1, n_2, and n_3. Associated with the Bravais lattice, there is a reciprocal lattice

$$\mathbf{G} = k_1\mathbf{b}_1 + k_2\mathbf{b}_2 + k_3\mathbf{b}_3, \tag{103}$$

where every reciprocal lattice vector $\mathbf{G}$ is given by the superposition of the reciprocal basis vectors $\mathbf{b}_1$, $\mathbf{b}_2$, and $\mathbf{b}_3$. Here k_1, k_2, and k_3 are again the corresponding integer numbers. The reciprocal basis set is calculated from its real space counterpart $\mathbf{b}_i = 2\pi(\mathbf{a}_j \times \mathbf{a}_k)/(\mathbf{a}_i\langle\mathbf{a}_j|\mathbf{a}_k\rangle)$, and both basis sets are orthogonal to each other $\mathbf{b}_i \cdot \mathbf{a}_j = 2\pi\delta_{ij}$.

The Schrödinger equation of a single electron under PBC is given by [5]

$$\left(\frac{1}{2}\nabla^2 + U(\mathbf{r})\right)\psi(\mathbf{r}) = E\psi(\mathbf{r}), \tag{104}$$

where $U(\mathbf{r})$ is a periodic potential

$$U(\mathbf{r} + \mathbf{R}) = U(\mathbf{r}). \tag{105}$$

For this periodic case, the Bloch theorem states that all eigenstates ψ of the one-electron Hamiltonian can be expressed as a product of a plane wave phase factor $e^{i\mathbf{kr}}$ and a function $u_{n,\mathbf{k}}(\mathbf{r})$ that has the same periodicity as the Bravais lattice. Specifically, this means that [5]

$$u_{n,\mathbf{k}}(\mathbf{r} + \mathbf{R}) = u_{n,\mathbf{k}}(\mathbf{r}) \tag{106}$$

and

$$\psi_{n,\mathbf{k}}(\mathbf{r}) = e^{i\mathbf{kr}}u_{n,\mathbf{k}}(\mathbf{r}). \tag{107}$$

Here n is the so-called band index, and k is a vector in reciprocal space, which is called Bloch vector. The Bloch vector can be considered as an additional quantum number introduced by the PBC.

Since any periodic function can be expanded into a plane wave basis set [111]

$$u_{n,\mathbf{k}}(\mathbf{r}) = \sum_{\mathbf{G}} C_{n,\mathbf{k},\mathbf{G}}\ e^{i\mathbf{G}\mathbf{r}}, \tag{108}$$

it follows that, for every band index n and Bloch vector $\mathbf{k}$, the one-electron wave function is given by the following plane wave expansion:

$$\psi_{n,\mathbf{k}}(\mathbf{r}) = \sum_{\mathbf{G}} C_{n,\mathbf{k},\mathbf{G}}\ e^{i(\mathbf{k}+\mathbf{G})\mathbf{r}}. \tag{109}$$

This expansion is usually limited to the plane waves with a kinetic energy $|\mathbf{k}+\mathbf{G}|^2/2$ smaller than a particular energy cut off E_c.

Furthermore, the Bloch vectors $\mathbf{k}$ can be restricted to the first Brillouin zone (1^{st} BZ), which is the first Wigner Seitz cell in reciprocal space. This is because all other Bloch states $\mathbf{k}$ can be calculated by a Bloch state of the 1^{st} BZ $\mathbf{k}_\mathrm{BZ}$ and an additional reciprocal lattice vector $\mathbf{G}$ [5].

$$\mathbf{k} = \mathbf{k}_\mathrm{BZ} + \mathbf{G} \tag{110}$$

This means that the Bloch states are periodic in reciprocal space, and that, for every band index n, the eigenstates $\psi_{n,\mathbf{k}}$ and the eigenvalues $\epsilon_{n,\mathbf{k}}$ of the Schrödinger equation are periodic in reciprocal space as well [5].

$$\psi_{n,\mathbf{k}} = \psi_{n,\mathbf{k}_\mathrm{BZ}+\mathbf{G}} \tag{111}$$

$$\epsilon_{n,\mathbf{k}} = \epsilon_{n,\mathbf{k}_\mathrm{BZ}+\mathbf{G}} \tag{112}$$

The band structure of solids is defined by considering the eigenvalues $\epsilon_{n,\mathbf{k}}$ as a set of periodic functions $\epsilon_n(\mathbf{k})$, and the density of states (DOS) is, for each band n, given by integrating over the 1^{st} BZ [5]

$$D_n(\epsilon) = \int_\mathrm{BZ} \frac{d\mathbf{k}}{4\pi^3} \delta\Big(\epsilon - \epsilon_n(\mathbf{k})\Big). \tag{113}$$

The total DOS, including the contributions of all bands, is given by $D(\epsilon) = \sum_n D_n(\epsilon)$. For practical reasons, the integration over the 1^{st} BZ is typically replaced by a summation over a discrete set of $\mathbf{k}$ points with appropriate weighting factors $w_\mathbf{k}$

$$D_n(\epsilon) = \sum_{\mathbf{k}} w_\mathbf{k}\, \delta\Big(\epsilon - \epsilon_{n,\mathbf{k}})\Big). \tag{114}$$

In present work, the summation is performed by using the $\mathbf{k}$-point grids and the weighting factors that are suggested by Monkhorst and Pack [113].

2.6 The frozen core approximation and pseudo (PS) potential methods

In section 2.5, we discussed that it is possible, and also practical, to expand one-electron wave functions into a plane wave basis set. Additionally, we mentioned that the expansion is truncated. This means that the kinetic energy $(\mathbf{k}+\mathbf{G})^2/2$ of a plane wave, remains always smaller than a certain energy cut off E_{cut} [112]. In Kohn-Sham density functional theory (KS-DFT), both the number of considered electrons and the size of the plane wave basis set have a huge impact on the computational costs of the calculations. It is therefore a major concern to reduce the number of electrons that must be calculated and to make the plane wave basis set for the electronic states as small as the required accuracy allows.

The frozen core approximation is typically applied to reduce the number of calculated electrons. The idea of this approximation is that only the valence electrons of the outmost electron shell participate in bonding and that the inner shells can be considered as constant or frozen. This limits the quantum-mechanical calculation to the wave functions of the valence electrons. However, the corresponding potentials and wave functions are still computational demanding. Since the wave functions are orthogonal to the core electrons, they oscillate strongly and/or their derivatives are large. As a result, a large basis set of plane waves is needed to correctly describe the wave functions of the valence electrons in the core region.

Pseudo (PS) potentials are, therefore, used to decrease the number of plane waves that are needed in the bonding region. As shown in Fig. 10, the PS potential $\tilde{V}$ replaces the exact potential V within the core region. Here the core region is defined by a sphere $\Omega_{R_\alpha}(r_\text{c})$ around the atomic position R_α with cut off radius r_c. The PS potential is constructed so that the PS wave functions $\tilde{\phi}$ are nodeless within the core region and reproduce the scattering properties of the exact wave functions ϕ [112]. Furthermore, the PS potential and the PS wave functions are equal to their exact counterparts outside the core region, and, at the boundary $r = r_\text{c}$, the PS wave functions coincide with the exact wave functions up to their second radial derivatives. Since the PS wave functions are designed to be much smoother than their exact counterparts, PS potential methods can reduce the necessary size of plane wave basis sets considerably.

2.7 The projector augmented wave (PAW) method

The PAW method is also a pseudo (PS) potential method, but it restores the correct features of the exact all electron (AE) wave function. This is done by adding correction terms within the core regions of the atoms. For each atom, the correction terms are basically differences between radial PS wave functions and radial AE wave functions. Within the core regions, the radial

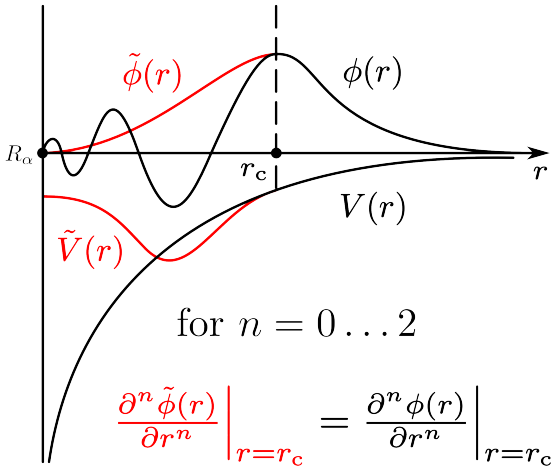


Figure 10: An illustration of the pseudo (PS) potential method inside and outside the core region of an atom α at position R_α. Inside the core region, the exact potential V and the exact wave function ϕ (black lines) are substituted by their PS counterparts $\tilde{V}$ and $\tilde{\phi}$ (red lines). The boundary conditions at r_c ensure a smooth transition between the PS functions and the exact functions.

wave functions are calculated by solving the radial Schrödinger equation of a single atom for each atom type. The AE wave function $|\psi_n\rangle$ is then given by [114, 115]

$$|\psi_n\rangle = |\tilde{\psi}_n\rangle + \sum_i a_{n,i}\Big(|\phi_i\rangle - |\tilde{\phi}_i\rangle\Big), \tag{115}$$

where $|\tilde{\psi}_n\rangle$ is the PS wave function, $|\phi_i\rangle$ and $|\tilde{\phi}_i\rangle$ are the AE and the PS partial wave functions of the reference atoms, and $a_{n,i}$ are the coefficients of the correction term. The PS wave function can be expressed in terms of a plane wave expansion, while the partial wave functions are usually given by a linear combination of atomic orbitals (LCAO) that is stored on a radial grid [112,114,115]. In the correction term, the sum over i is the summation over atomic positions $\mathbf{R}_\alpha$ and all quantum numbers. Note that the correction of the PS wave functions only applies within the core regions of the atoms Ω_r, since the PS wave functions coincide with the AE wave functions outside of the core region.

This framework is now augmented by introducing projector functions $|\tilde{p}_i\rangle$ that are dual to the PS partial waves

$$\delta_{i,j} = \langle \tilde{p}_i | \tilde{\phi}_j \rangle. \tag{116}$$

The coefficients of the correction term are defined by

$$a_{n,i} = \langle \tilde{p}_i | \tilde{\psi}_n \rangle, \tag{117}$$

and the transformation that restores the AE wave function is given by [114]

$$\mathrm{T} = 1 + \sum_i \Big(|\phi_i\rangle - |\tilde{\phi}_i\rangle \Big) \langle \tilde{p}_i |. \tag{118}$$

This transformation is the heart of the PAW method and allows to calculate expectation values of quasi local operator A. Specifically, the expectation value $\langle A \rangle$ of a quasi local observable A can be defined in terms of normalized wave functions $|\psi_n\rangle$ or in term of normalized PS wave functions $|\tilde{\psi}_n\rangle$.

$$\langle A \rangle = \sum_n f_n \langle \psi_n | \, \mathrm{A} \, | \psi_n \rangle \tag{119}$$

$$= \sum_n f_n \langle \tilde{\psi}_n | \, \tilde{\mathrm{A}} \, | \tilde{\psi}_n \rangle \tag{120}$$

Here the operators $\tilde{\mathrm{A}}$ and A are related by $\tilde{\mathrm{A}} = \mathrm{T}^\dagger \mathrm{A} \mathrm{T}$. By defining the occupation coefficients $\lambda_{ij} = \sum_n f_n \, a_{n,i} \, a_{n,j}$, the expectation value is, for a complete set of PS wave functions, approximated by

$$\langle A \rangle = \underbrace{\sum_n f_n \langle \tilde{\psi}_n | \mathrm{A} | \tilde{\psi}_n \rangle}_{\tilde{A}} + \Big(\underbrace{\sum_{i,j} \lambda_{ij} \langle \phi_i | \mathrm{A} | \phi_j \rangle}_{A^1} - \underbrace{\sum_{i,j} \lambda_{ij} \langle \tilde{\phi}_i | \mathrm{A} | \tilde{\phi}_j \rangle}_{\tilde{A}^1} \Big). \tag{121}$$

The all atom PS contribution to the expectation value is given by $\tilde{A}$, while the AE one centered contributions and the PS one centered contributions are given by A^1 and $\tilde{A}^1$, respectively. Note that $\tilde{A}$ is the expectation value of the operator A and the PS wave functions $|\tilde{\psi}_n\rangle$, while $\tilde{\mathrm{A}}$ is the PS counterpart of the operator A. The crucial point that makes the PAW method highly efficient is that the terms with the all atom PS wave functions $\tilde{\psi}_n$ and the partial wave functions $\tilde{\phi}_i$ are separated. In the calculation of quasi local expectation values, mixed terms that involve $|\tilde{\psi}_n\rangle$ and $|\phi_i\rangle$ are absent.

For instance, as the kinetic energy operator $-\frac{1}{2}\nabla^2$ and also the position operator $|\mathbf{r}\rangle\langle\mathbf{r}|$ are quasi local, it is possible to split the kinetic energy into

$$T_\mathrm{s}[n] = \underbrace{\sum_n f_n \langle \tilde{\psi}_n | - \frac{1}{2}\nabla^2 | \tilde{\psi}_n \rangle}_{\tilde{T}_\mathrm{s}} + \Big(\underbrace{\sum_{i,j} \lambda_{ij} \langle \phi_i | - \frac{1}{2}\nabla^2 | \phi_j \rangle}_{T_\mathrm{s}^1} - \underbrace{\sum_{i,j} \lambda_{ij} \langle \tilde{\phi}_i | - \frac{1}{2}\nabla^2 | \tilde{\phi}_j \rangle}_{\tilde{T}_\mathrm{s}^1} \Big) \tag{122}$$

and the electronic charge density into

$$n(\mathbf{r}) = \underbrace{\sum_n f_n \langle \tilde{\psi}_n | \mathbf{r} \rangle \langle \mathbf{r} | \tilde{\psi}_n \rangle}_{\tilde{n}(\mathbf{r})} + \Big(\underbrace{\sum_{i,j} \lambda_{ij} \langle \phi_i | \mathbf{r} \rangle \langle \mathbf{r} | \phi_j \rangle}_{n^1(\mathbf{r})} - \underbrace{\sum_{i,j} \lambda_{ij} \langle \tilde{\phi}_i | \mathbf{r} \rangle \langle \mathbf{r} | \tilde{\phi}_j \rangle}_{\tilde{n}^1(\mathbf{r})} \Big). \tag{123}$$

These equations for the kinetic energy and the electronic density can then plugged into

$$E^\mathrm{KS}[n] = T_\mathrm{s}[n] + E_\mathrm{ne}[n] + E_\mathrm{ee}[n] + E_\mathrm{xc}[n] \tag{124}$$

to calculate the KS energy in the PAW method, compare section 2.3.1.

In order to incorporate the frozen core electrons into the KS formalism the total electronic density n_{tot} can be split into the contributions from the valence electrons $n_{\mathrm{v}}(\mathbf{r})$, the contributions from the frozen electrons $n_{\mathrm{c}}(\mathbf{r})$ and the contributions from the nuclei $n_Z(\mathbf{r}) = \sum_\alpha Z_\alpha\, \delta(\mathbf{r}-\mathbf{R}_\alpha)$.

$$n_{\mathrm{tot}}(\mathbf{r}) = n_{\mathrm{v}}(\mathbf{r}) + \underbrace{n_{\mathrm{c}}(\mathbf{r}) + n_{\mathrm{Z}}(\mathbf{r})}_{n_{\mathrm{Zc}}(\mathbf{r})} \tag{125}$$

By introducing the compensation charge density $\hat{n}$, the total charge density of the PAW method can further be split into [115]

$$\underbrace{\Big(n_{\mathrm{v}}(\mathbf{r}) + n_{\mathrm{Zc}}(\mathbf{r})\Big)}_{n_{\mathrm{tot}}(\mathbf{r})} = \underbrace{\Big(\tilde{n}_{\mathrm{v}}(\mathbf{r}) + \tilde{n}_{\mathrm{Zc}}(\mathbf{r}) + \hat{n}(\mathbf{r})\Big)}_{\tilde{n}_{\mathrm{tot}}(\mathbf{r})} + \Bigg(\underbrace{\Big(n^1_{\mathrm{v}}(\mathbf{r}) + n_{\mathrm{Zc}}(\mathbf{r})\Big)}_{n^1_{\mathrm{tot}}(\mathbf{r})} - \underbrace{\Big(\tilde{n}^1_{\mathrm{v}}(\mathbf{r}) + \tilde{n}_{\mathrm{Zc}}(\mathbf{r}) - \hat{n}(\mathbf{r})\Big)}_{\tilde{n}^1_{\mathrm{tot}}(\mathbf{r})} \Bigg). \tag{126}$$

The compensation charge density is added to the PS charge densities $\tilde{n} + \tilde{n}_{\mathrm{Zc}}$ and $\tilde{n}^1 + \tilde{n}^1_{\mathrm{Zc}}$ so that the one centered PS charge density $\tilde{n}^1 + \hat{n}$ has the same multi pole moments as the one centered AE charge density n^1. [115] Then non-local Coulomb interactions can be evaluated. In practice, this means that

$$\int_{\Omega_r} d\mathbf{r} \left(n^1 - (\tilde{n}^1 + \hat{n})\right)\, |\mathbf{r}-\mathbf{R}|^l\; Y_L\Big(\vartheta(\mathbf{r}-\mathbf{R}), \varphi(\mathbf{r}-\mathbf{R})\Big) = 0 \tag{127}$$

must be fulfilled, where Y_L are spherical harmonics. The compensation charge can finally be written as [115]

$$\hat{n}(\mathbf{r}) = \sum_{i,j} \sum_L \lambda_{ij} \hat{Q}_{ijL}(\mathbf{r}-\mathbf{R}). \tag{128}$$

Here Q_{ijL} is, within each core region centered at $\mathbf{R}$, defined by a multiplication of multi pole moments q_i, spherical harmonics Y_L and a linear combination of Bessel functions, compare Ref. 115. Correspondingly, the sum over L means the summation over all angular momentum numbers.

Knowing the contributions to the total electronic density n_{tot}, it is possible to express the exact KS energy functional of all electrons by the kinetic energy functional $T_{\mathrm{s}}[n_{\mathrm{v}} + n_{\mathrm{c}}]$, the exchange correlation functional $E_{\mathrm{xc}}[n_{\mathrm{v}} + n_{\mathrm{c}}]$, and a Hartree functional $E_{\mathrm{H}}[n_{\mathrm{v}} + n_{\mathrm{Zc}}]$ [115].

$$E^{\mathrm{KS}}[n_{\mathrm{v}} + n_{\mathrm{c}}] = T_{\mathrm{s}}[n_{\mathrm{v}} + n_{\mathrm{c}}] + E_{\mathrm{H}}[n_{\mathrm{v}} + n_{\mathrm{Zc}}] + E_{\mathrm{xc}}[n_{\mathrm{v}} + n_{\mathrm{c}}] \tag{129}$$

Note that this form of the KS energy functional already contains the interaction between the nucleii

$$E_{\mathrm{H}}[n_Z] = \frac{1}{2} \int d\mathbf{r}\, \frac{n_Z(\mathbf{r}) n_Z(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} = \frac{1}{2} \sum_{\alpha,\beta} \frac{Z_\alpha Z_\beta}{|\mathbf{R}_\alpha - \mathbf{R}_\beta|}. \tag{130}$$

Using the PAW decomposition for the kinetic energy operator and the total density, it can be shown that the total energy functional can also be split into an all atom PS contribution and two one center energy terms.

$$E = \tilde{E} + \left(E^1 - \tilde{E}^1\right) \tag{131}$$

These contributions are defined by [115]

$$\begin{aligned}
\tilde{E} &= \sum_n f_n \langle\tilde{\psi}_n| - \frac{1}{2}\nabla^2|\tilde{\psi}_n\rangle && (132)\\
&+ \int d\mathbf{r}\, V_{\mathrm{H}}[\tilde{n}_{\mathrm{v}} + \hat{n}](\mathbf{r}) \left(\tilde{n}_{\mathrm{v}}(\mathbf{r}) + \hat{n}(\mathbf{r})\right) && (133)\\
&+ \int d\mathbf{r}\, V_{\mathrm{H}}[\tilde{n}_{\mathrm{Zc}}](\mathbf{r}) \left(\tilde{n}_{\mathrm{v}}(\mathbf{r}) + \hat{n}(\mathbf{r})\right) && (134)\\
&+ \int d\mathbf{r}\, V_{\mathrm{xc}}[\tilde{n}_{\mathrm{v}} + \tilde{n}_{\mathrm{c}} + \hat{n}](\mathbf{r}) \left(\tilde{n}_{\mathrm{v}}(\mathbf{r}) + \tilde{n}_{\mathrm{c}}(\mathbf{r}) + \hat{n}(\mathbf{r})\right) && (135)\\
&+ V_{nn}(\{\mathbf{R}_\alpha\}, \{Z_\alpha\}) && (136)\\
E^1 &= \sum_{i,j} \lambda_{ij} \langle\tilde{\phi}_i| - \frac{1}{2}\nabla^2|\tilde{\phi}_j\rangle && (137)\\
&+ \int_{\Omega_r} d\mathbf{r}\, V_{\mathrm{H}}[n^1_{\mathrm{v}}](\mathbf{r}) \left(n^1_{\mathrm{v}}(\mathbf{r})\right) && (138)\\
&+ \int_{\Omega_r} d\mathbf{r}\, V_{\mathrm{H}}[n_{\mathrm{Zc}}](\mathbf{r}) \left(n^1_{\mathrm{v}}(\mathbf{r})\right) && (139)\\
&+ \int_{\Omega_r} d\mathbf{r}\, V_{\mathrm{xc}}[n^1_{\mathrm{v}} + n_{\mathrm{c}}](\mathbf{r}) \left(n^1_{\mathrm{v}}(\mathbf{r}) + n_{\mathrm{c}}(\mathbf{r})\right) && (140)\\
\tilde{E}^1 &= \sum_{i,j} \lambda_{ij} \langle\tilde{\phi}_i| - \frac{1}{2}\nabla^2|\tilde{\phi}_j\rangle && (141)\\
&+ \int_{\Omega_r} d\mathbf{r}\, V_{\mathrm{H}}[\tilde{n}^1_{\mathrm{v}} + \hat{n}](\mathbf{r}) \left(\tilde{n}^1_{\mathrm{v}}(\mathbf{r}) + \hat{n}(\mathbf{r})\right) && (142)\\
&+ \int_{\Omega_r} d\mathbf{r}\, V_{\mathrm{H}}[\tilde{n}_{\mathrm{Zc}}](\mathbf{r}) \left(\tilde{n}^1_{\mathrm{v}}(\mathbf{r}) + \hat{n}(\mathbf{r})\right) && (143)\\
&+ \int_{\Omega_r} d\mathbf{r}\, V_{\mathrm{xc}}[\tilde{n}^1_{\mathrm{v}} + \tilde{n}_{\mathrm{c}} + \hat{n}](\mathbf{r}) \left(\tilde{n}^1_{\mathrm{v}}(\mathbf{r}) + \tilde{n}_{\mathrm{c}}(\mathbf{r}) + \hat{n}(\mathbf{r})\right). && (144)
\end{aligned}$$

Here $V_{\mathrm{H}}[n](\mathbf{r}) = \int d\mathbf{r}' \frac{n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}$ is the Hartree potential for a charge density n, and V_{nn} is the electrostatic energy of the point charges Z_α located at $\mathbf{R}_\alpha$ [115].

By varying the PAW total energy functional with respect to the PS density operator $\tilde{\rho} = \sum_n f_n |\tilde{\psi}_n\rangle\langle\tilde{\psi}_n|$, it is possible derive the Hamilton operator $\mathrm{H}^{\mathrm{PAW}}$ of the PAW method [114,115].

$$\mathrm{H}^{\mathrm{PAW}}=\tilde{\mathrm{H}}+(\mathrm{H}^1-\tilde{\mathrm{H}}^1)+\hat{\mathrm{H}} \tag{145}$$

$$\tilde{\mathrm{H}} \quad =-\frac{1}{2}\nabla^2+\tilde{V}_{\mathrm{eff}}(\mathbf{r}) \tag{146}$$

$$\mathrm{H}^1 \quad =\sum_{ij}|\tilde{p}_i\rangle\langle\tilde{p}_j|\ \langle\phi_i|-\frac{1}{2}\nabla^2+V^1_{\mathrm{eff}}(\mathbf{r})|\phi_j\rangle \tag{147}$$

$$\tilde{\mathrm{H}}^1 \quad =\sum_{ij}|\tilde{p}_i\rangle\langle\tilde{p}_j|\ \langle\tilde{\phi}_i|-\frac{1}{2}\nabla^2+\tilde{V}^1_{\mathrm{eff}}(\mathbf{r})|\tilde{\phi}_j\rangle \tag{148}$$

$$\hat{\mathrm{H}} \quad =\sum_{ij}|\tilde{p}_i\rangle\langle\tilde{p}_j|\ \left(\sum_L\Big(\int d\mathbf{r}\,\tilde{V}_{\mathrm{eff}}(\mathbf{r})Q_{i,j,L}(\mathbf{r})\Big)-\Big(\int_{\Omega_r}d\mathbf{r}\,\tilde{V}^1_{\mathrm{eff}}(\mathbf{r})Q_{i,j,L}(\mathbf{r})\Big)\right) \tag{149}$$

Here the effective potentials are given by [115]

$$\tilde{V}_{\mathrm{eff}}(\mathbf{r})=V_{\mathrm{H}}[\tilde{n}_{\mathrm{v}}+\tilde{n}_{\mathrm{Zc}}+\hat{n}]+V_{\mathrm{xc}}[\tilde{n}_{\mathrm{v}}+\tilde{n}_{\mathrm{c}}+\hat{n}] \tag{150}$$

$$V^1_{\mathrm{eff}}(\mathbf{r})=V_{\mathrm{H}}[n^1_{\mathrm{v}}+n_{\mathrm{Zc}}]+V_{\mathrm{xc}}[n^1_{\mathrm{v}}+n_{\mathrm{c}}] \tag{151}$$

$$\tilde{V}^1_{\mathrm{eff}}(\mathbf{r})=V_{\mathrm{H}}[\tilde{n}^1_{\mathrm{v}}+n_{\mathrm{Zc}}+\hat{n}]+V_{\mathrm{xc}}[\tilde{n}^1_{\mathrm{v}}+n_{\mathrm{c}}+\hat{n}]. \tag{152}$$

The PAW wave functions fulfill the generalized eigenvalue equation

$$\Big(\mathrm{H}^{\mathrm{PAW}}-\epsilon_n\mathrm{S}\Big)|\tilde{\psi}_n\rangle=0 \tag{153}$$

with the overlap operator

$$\mathrm{S} \;=\; \mathbf{1}+\sum_{ij}4\pi q^0_{ij}|\tilde{p}_i\rangle\langle\tilde{p}_j| \tag{154}$$

$$\langle\tilde{\psi}_m|\mathrm{S}|\tilde{\psi}_n\rangle \;=\; \delta_{mn}. \tag{155}$$

If the PS wave functions are expanded into a plane wave basis set $|\tilde{\psi}_n\rangle=\sum_{\mathbf{G}}c_{n,\mathbf{G}}|\mathbf{G}\rangle$, the eigenvalue equation can be rewritten into a matrix equation in the following way:

$$\langle\psi_m|\mathrm{H}^{\mathrm{PAW}}-\epsilon_n\mathrm{S}|\psi_n\rangle \;=\; 0 \tag{156}$$

$$\sum_{\mathbf{G},\mathbf{G}'}c^*_{m\mathbf{G}}\,c_{n\mathbf{G}'}\left(\mathrm{H}^{\mathrm{PAW}}_{\mathbf{G},\mathbf{G}'}-\epsilon_n\mathrm{S}_{\mathbf{G},\mathbf{G}'}\right) \;=\; 0 \tag{157}$$

$$(\mathbf{C}^{\mathrm{T}})^{-1}\times|\ \ \mathbf{C}^{\mathrm{T}}\mathbf{H}^{\mathrm{PAW}}\mathbf{C}-E\mathbf{C}^{\mathrm{T}}\mathbf{S}\mathbf{C} \;=\; 0 \tag{158}$$

$$\mathbf{H}^{\mathrm{PAW}}\mathbf{C}-E\mathbf{S}\mathbf{C} \;=\; 0 \tag{159}$$

Using this identity, the Hamilton matrix can be diagonalized in order to determine the expansion coefficients, the wave functions, and the electronic densities.

2.8 Ab initio molecular dynamics (MD)

In MD simulations, the trajectories of the atoms are calculated following Newton's dynamics. Materials properties are then calculated by a phase space integration [116].

$$\langle \mathrm{A} \rangle_{\mathrm{MC}} = \lim_{T\to\infty} \frac{1}{T} \int dt\, \mathrm{A}(\{\mathbf{R}_\alpha(t)\}, \{\dot{\mathbf{R}}_\alpha(t)\}) \tag{160}$$

The underlying assumption for this integration is that the ensemble averaged properties coincide with the time averaged properties for sufficient long simulation times T. This assumption is the so-called ergodic hypothesis. Although this hypothesis innately applies to micro-canonical ensembles only, Nose has extended this approach to canonical ensembles [117].

2.8.1 The calculation of the inter-atomic forces

Since we use the Born-Oppenheimer approximation and we treat only the electrons quantum-mechanically, the total energy of the atoms is given by

$$E_{\mathrm{tot}}(\{\mathbf{R}_\alpha\}, \{\dot{\mathbf{R}}_\alpha\}) = T(\{\dot{\mathbf{R}}_\alpha\}) + U_{\mathrm{eff}}(\{\mathbf{R}_\alpha\}). \tag{161}$$

Here $T = \sum_\alpha \frac{1}{2} m_\alpha \dot{\mathbf{R}}_\alpha^2$ is the kinetic energy of all atoms, but the effective potential U_{eff} includes, besides the electrostatic potential of the atoms, also the ground state total energy of the PAW energy functional. Specifically, in the PAW method the effective potential is given by

$$U_{\mathrm{eff}}(\{\mathbf{R}_\alpha\}) = \sum f_n \epsilon_n(\{\mathbf{R}_\alpha\}) - E_{\mathrm{ee}}[n] + E_{\mathrm{xc}}[n] - \int d\mathbf{r}\, V_{\mathrm{xc}}(\mathbf{r}) n(\mathbf{r}) + U_{\mathrm{nn}}(\{\mathbf{R}_\alpha\}), \tag{162}$$

where f_n are the occupation numbers of each PS wave function $\tilde{\psi}_n$, ϵ_n are the eigenvalues of the PAW method, E_{ee} is the Hartree potential, and U_{nn} is the electrostatic potential of the atoms. While the eigenvalues and the Hartree potential are selfconsistently calculated solving the generalized eigenvalue problem described in section 2.7, the electrostatic potential is evaluated by a Ewald summation [118]. Since only the eigenvalues and the electrostatic potential depend on the atomic positions, the force that acts on each atom α consists of two terms [115].

$$\mathbf{F}_\alpha = -\nabla_\alpha U_{\mathrm{eff}}(\{\mathbf{R}_\alpha\}) = -\Big(\sum f_n \nabla_\alpha \epsilon_n(\{\mathbf{R}_\alpha\}) + \nabla_\alpha U_{\mathrm{nn}}(\{\mathbf{R}_\alpha\}) \Big) \tag{163}$$

The eigen values ϵ_n can generally be written by

$$\epsilon_n(\{\mathbf{R}_\alpha\}) = \frac{\langle \tilde{\psi}_n | \mathrm{H}^{\mathrm{PAW}}(\{\mathbf{R}_\alpha\}) | \tilde{\psi}_n \rangle}{\langle \tilde{\psi}_n | \mathrm{S}(\{\mathbf{R}_\alpha\}) | \tilde{\psi}_n \rangle}. \tag{164}$$

Using first the usual formula for derivatives $\frac{f}{g} = \frac{f'g - fg'}{g^2}$ and next the orthonormality condition $\langle \tilde{\psi}_n | \mathrm{S} | \tilde{\psi}_n \rangle = 1$, it can be shown that [119]

$$\nabla_\alpha \epsilon_n(\{\mathbf{R}_\alpha\}) = \nabla_\alpha \langle \tilde{\psi}_n | \mathrm{H}^{\mathrm{PAW}}(\{\mathbf{R}_\alpha\}) | \tilde{\psi}_n \rangle - \underbrace{\langle \tilde{\psi}_n | \mathrm{H}^{\mathrm{PAW}}(\{\mathbf{R}_\alpha\}) | \tilde{\psi}_n \rangle}_{\epsilon_n} \nabla_\alpha \langle \tilde{\psi}_n | \mathrm{S}(\{\mathbf{R}_\alpha\}) | \tilde{\psi}_n \rangle. \tag{165}$$

Since the PS wave functions are expressed in terms of plane waves, which are in turn independent of the atomic positions $\mathbf{R}_\alpha$, the final expression of the forces is [115]

$$\mathbf{F}_\alpha = \sum_n f_n \langle \tilde{\psi}_n | \nabla_\alpha \left(\mathrm{H}^{\mathrm{PAW}}(\{\mathbf{R}_\alpha\}) - \epsilon_n(\{\mathbf{R}_\alpha\}) \mathrm{S}(\{\mathbf{R}_\alpha\}) \right) | \tilde{\psi}_n \rangle + \nabla_\alpha U_{\mathrm{nn}}(\{\mathbf{R}_\alpha\}). \quad (166)$$

This expression is very similar to the Hellman-Feynman theorem [120], but the Hellman-Feynman theorem applies for orthonormal wave functions only. The present equation with an overlap of the PS wave functions was first derived by Goedecker and Maschke [119]. An alternative way to derive this expression using the Lagrange formalism is demonstrated by Kresse *et al.* in Ref. 121. The explicit formulas of the PAW forces are given in Ref. 115.

2.8.2 The integration of the equation of motion

Once the forces have been calculated, the atomic positions have been given as an initial value, and some random initial velocities have been assigned to the ions, the trajectory of the atoms can be calculated by integrating the equation of motion for each atomic position $\mathbf{R}_\alpha(t)$. This can be done, for example, using the Verlet algorithm [122]. The algorithm can be derived easily by a Taylor expansion of the position vectors $\mathbf{R}_\alpha(t - \Delta t)$ and $\mathbf{R}_\alpha(t + \Delta t)$ around $t = 0$ [70, 116]

$$\mathbf{R}_\alpha(t - \Delta t) = \mathbf{R}_\alpha(t) - \dot{\mathbf{R}}_\alpha(t)\Delta t + \frac{1}{2}\ddot{\mathbf{R}}_\alpha(t)\Delta t^2 - \frac{1}{6}\dddot{\mathbf{R}}_\alpha(t)\Delta t^3 + O(\Delta t^4) \quad (167)$$

$$\mathbf{R}_\alpha(t + \Delta t) = \mathbf{R}_\alpha(t) + \dot{\mathbf{R}}_\alpha(t)\Delta t + \frac{1}{2}\ddot{\mathbf{R}}_\alpha(t)\Delta t^2 + \frac{1}{6}\dddot{\mathbf{R}}_\alpha(t)\Delta t^3 + O(\Delta t^4). \quad (168)$$

If both equations are added and the fourth order terms are neglected, the velocities $\dot{\mathbf{R}}_\alpha$ drop out and the atomic positions at $t + \Delta t$ are calculated by

$$\mathbf{R}_\alpha(t + \Delta t) = 2\mathbf{R}_\alpha(t) - \mathbf{R}_\alpha(t - \Delta t) + \ddot{\mathbf{R}}_i(t)\Delta t^2 + O(\Delta t^4). \quad (169)$$

Now in turn the equation depends on two atom positions at time t and $t - \Delta t$, and on the second derivative at time t. The derivative can be calculated from the forces.

$$\ddot{\mathbf{R}}_\alpha(t) = \frac{\mathbf{F}_\alpha(t)}{m_i} = -\frac{\nabla_\alpha E(t)}{m_i} \quad (170)$$

Although the velocities are not explicitly needed for the integration, they can be approximated by

$$\dot{\mathbf{R}}_\alpha(t) = \frac{\mathbf{R}_\alpha(t + \Delta t) - \mathbf{R}_\alpha(t - \Delta t)}{2\Delta t} + O(\Delta t^2). \quad (171)$$

Besides the initial positions and initial velocities, the time step Δt has to be chosen before the integration. Its choice depends preliminary on two factors. Firstly, it must be large enough so that the simulation can progress sufficiently fast. Secondly, it must be small enough so that integrating the equation of motion gives a reasonably accurate trajectory.

2.8.3 The ionic relaxation at ground state

The equilibrium configuration can be found by following the trajectory related to a damped MD simulation until the forces vanish within a certain tolerance value. In order to prevent oscillations of the atoms around their equilibrium positions in this relaxation process, the forces must be damped using a friction term ξ and a control parameter μ [70, 123]

$$\ddot{\mathbf{R}}_\alpha(t) = -2\mu\mathbf{F}_\alpha(t) - \xi\dot{\mathbf{R}}_\alpha(t). \tag{172}$$

As discussed in the previous section 2.8.2, the corresponding trajectory may be obtained by integrating the equation of motion with the Verlet algorithm.

$$\mathbf{R}_\alpha(t+\Delta t) \;=\; 2\mathbf{R}_\alpha(t) - \mathbf{R}_\alpha(t-\Delta t) + \Big(-2\mu\mathbf{F}_\alpha(t) - \xi\dot{\mathbf{R}}_\alpha(t)\Big)\Delta t^2. \tag{173}$$

The only difference is the additional dependency on the velocities. Inserting Eq. (171) in Eq. (173), the integration scheme for the damped MD simulation is given by

$$\mathbf{R}_\alpha(t+\Delta t) = \frac{2\mathbf{R}_\alpha(t) - (1-\xi\Delta t/2)\mathbf{R}_\alpha(t-\Delta t) - 2\mu\mathbf{F}(t)\Delta t^2}{(1+\xi\Delta t/2)}. \tag{174}$$

Since the Verlet algorithm is exactly equivalent to the Leap-Frog Verlet algorithm the final integration scheme can be rewritten as [70, 116]

$$\dot{\mathbf{R}}_\alpha(t+\Delta t/2) \;=\; \dot{\mathbf{R}}_\alpha(t-\Delta t/2) + \Big(-2\mu\mathbf{F}_\alpha(t) - \xi\dot{\mathbf{R}}_\alpha(t)\Big)\Delta t \tag{175}$$

$$\mathbf{R}_\alpha(t+\Delta t) \;=\; \mathbf{R}_\alpha(t) + \dot{\mathbf{R}}_\alpha(t+\Delta t/2)\Delta t. \tag{176}$$

In practice, the Leap-Frog Verlet algorithm is preferred because it is more stable than the Verlet algorithm. The main difference between both algorithms is that the positions and the velocities are not defined at the same time in Leap-Frog, so that intermediate half steps $\Delta t/2$ must be calculated. The trajectory of the Leap-Frog algorithm is the same as the trajectory of the Verlet algorithm with an local error proportional to the fourth order in Δt [124]. Note that, besides this damped MD scheme, there are also other methods used for relaxation. Popular methods in literature are for example the steepest descent method and the conjugate gradient method. [125]

2.8.4 The simulation at finite temperature

The canonical ensemble can be simulated using a Nose Hoover thermostat [117, 126]. In this case, the Hamiltonian of the atoms is given by [116]

$$H(\{\mathbf{R}_\alpha\},\{\mathbf{p}_\alpha\},s,p_s) = \sum_\alpha \frac{(\mathbf{p}_\alpha/s)^2}{2m_\alpha} + U_{\text{eff}}(\{\mathbf{R}_\alpha\}) + 3Nk_BT\ln(s) + \frac{p_s^2}{2Q}. \tag{177}$$

On the one hand, the Hamiltonian, similarly to Eq. (161), consists of a kinetic energy term $T(\{\mathbf{p}_\alpha/\mathbf{s}\})$, with scaled atomic momenta $\mathbf{p}_\alpha = m_\alpha \dot{\mathbf{R}}_\alpha$, and a effective potential $U_{\text{eff}}(\{\mathbf{R}_\alpha\})$. On the other hand, the thermostat couples the energy calculation to a heat bath at temperature T by introducing two new terms and the time-dependent scaling factors s and p_s. Both factors are related by $p_s = Q\dot{s}$, where Q is, similarly to the relationship between the atomic momenta and the atomic positions, a mass parameter. As a consequence, the ordinary equations of motion are extended by two more equations for s and p_s [127]:

$$\frac{d\mathbf{R}_\alpha(t)}{dt} = \frac{\mathbf{p}_\alpha(t)}{m_\alpha s^2(t)} \tag{178}$$

$$\frac{d\mathbf{p}_\alpha(t)}{dt} = \mathbf{F}_\alpha(t) \tag{179}$$

$$\frac{ds(t)}{dt} = \frac{p_s(t)}{Q} \tag{180}$$

$$\frac{dp_s(t)}{dt} = \frac{1}{s(t)}\Big(\sum_\alpha \frac{\mathbf{p}_\alpha^2(t)}{m_\alpha s^2(t)} - 3NkT\Big). \tag{181}$$

These equations can further be simplified introducing the friction function $\xi = \dot{s}/s$ and rescaling the momenta and the time by $\mathbf{p}'_\alpha = \mathbf{p}_\alpha/s$ and $dt' = dt/s$, respectively. With these substitutions, the equations of motion are [116, 127]:

$$\frac{d\mathbf{R}_\alpha(t')}{dt'} = \frac{\mathbf{p}'_\alpha(t')}{m_\alpha} \tag{182}$$

$$\frac{d\mathbf{p}'_\alpha(t')}{dt'} = \mathbf{F}_\alpha(t') - \xi(t')\mathbf{p}'_\alpha(t') \tag{183}$$

$$\frac{d\xi(t')}{dt'} = \frac{2}{Q}\Big(\sum_\alpha \frac{\mathbf{p}'^2_\alpha(t')}{2m_\alpha} - \frac{3}{2}NkT\Big). \tag{184}$$

Eq. (183) shows that the temperature dependence of the trajectory is incorporated by constraint forces, which are added to the conventional inter-atomic forces.

$$\ddot{\mathbf{R}}_\alpha(t') = \frac{1}{m_\alpha}\mathbf{F}_\alpha(t') - \xi(t')\dot{\mathbf{R}}_\alpha(t') \tag{185}$$

The strength of the constraint forces are mainly controlled by the friction function $\xi(t')$, which becomes constant for total kinetic energy values equal to $\frac{3}{2}NkT$. As a result, the kinetic energy of the atoms oscillate around $\frac{3}{2}NkT$ with a characteristic frequency of $\omega_N = 6NkT/Q$ [117]. The equations of motion with the friction functions ξ are integrated using a damped MD simulation, as discussed in section 2.8.3

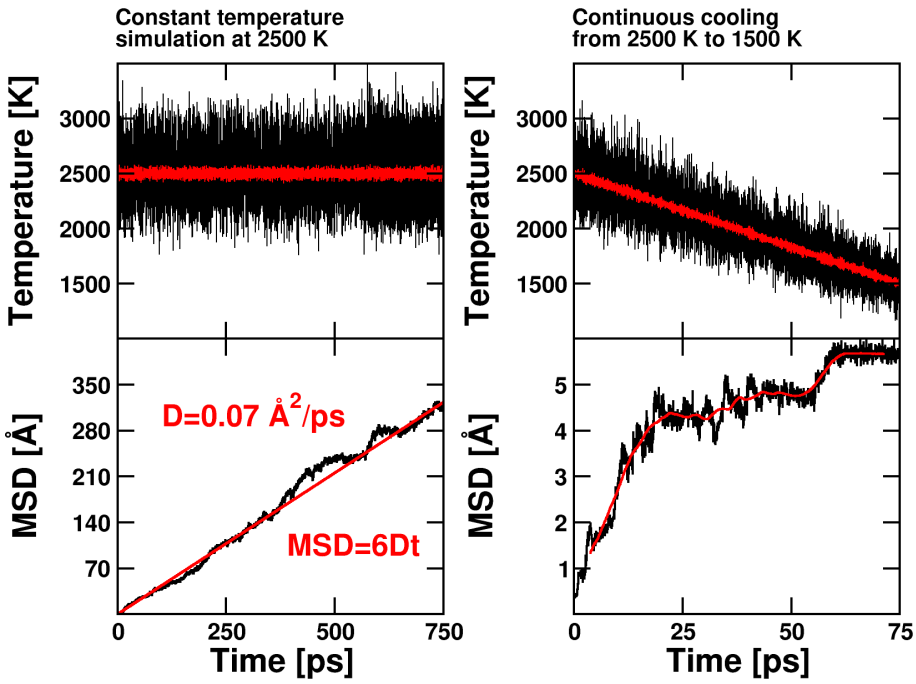


Figure 11: The temperature (upper panels) and the MSD (lower panels) of an amorphous sample during a constant temperature simulation (panels on the left hand side) and during continuous cooling (panels on the right hand side). At constant temperature, the diffusion constant is given by the slope of the MSD. After cooling to 1500 K, the atom diffusion in the amorphous sample has stopped and the MSD has reached a constant value (right hand side, lower panel).

3 Important observables and analysis methods

Starting from a c-Si_3N_4 structure plus some randomly added and removed Si, N, and H atoms, all amorphous samples were initially heated until they became liquid. Then they were stepwise cooled, slowly cooled, or rapidly quenched to freeze them into an amorphous network. The transition to the amorphous state was checked by investigating the atom diffusion by means of the mean square displacement (MSD) and by examining the structural properties by means of the pair correlation function (PCF), angular distribution function (ADF), and the structure factor (STF). Finally, the correlation between the samples was estimated in terms of the auto-correlation function and the statistical inefficiency. In this section, we give a short overview of these observables and control parameters.

3.1 Mean square displacement (MSD)

The MSD is defined by the average squared displacement of all atom positions $\mathbf{R}_\alpha$ [128]

$$\Delta R^2(t) = \frac{1}{N_n} \sum_{\alpha}^{N_n} \Big(\mathbf{R}_\alpha(t) - \mathbf{R}_\alpha(0) \Big)^2. \tag{186}$$

Fig. 11 illustrates the properties of the MSD during the simulation of an amorphous sample for two different temperature treatments. In the first case, the average temperature is constant at 2500 K for 750 ps (left hand side, upper panel), and the sample is liquid with an approximately constant atom diffusion. This can be seen from the nearly constant slope of the MSD (left hand side, lower panel). For long simulation times, the relationship between the MSD ΔR^2 and the diffusion constant D is given by the Einstein diffusion equation [128]

$$D = \lim_{t \to \infty} \frac{\Delta R^2(t)}{6t}. \tag{187}$$

In the figure, this relation is demonstrated by the fit of the linear function $\Delta R^2(t) = 6Dt$ to the data points of the simulation (left hand side, lower panel, red line). In the second case, the average temperature is decreased continuously from 2500 K to 1500 K in 75 ps (right hand side, upper panel), and the atom diffusion diminishes until it stops completely after 70 ps. Likewise, the slope of the MSD decreases until it becomes approximalty zero (right hand side, lower panel). Since the MSD always reaches a constant value when the atom diffusion stops, the MSD is very suitable to investigate the atom diffusion at low temperatures. Additionally, it gives some information on the atom diffusion even for short time intervals.

3.2 Pair correlation function (PCF)

To investigate bonding partners and bonding distances in amorphous materials one defines the PCF by [129]

$$g(r) = \frac{1}{\rho_{h_j}} \frac{\langle \sum_{j \neq i} \delta(r_{ij} - r) \rangle}{4\pi r^2}, \tag{188}$$

where δ is the Kronecker delta function, r_{ij} is the distance between the atoms i and j, and ρ_{h_j} is the homogeneous reference density of atoms j in the volume of the supercell. Note that the nominator $\langle \sum_{j \neq i} \delta(r_{ij} - r) \rangle$ gives the average number of atoms j within the volume $dV = dr\, 4\pi r^2$ of the infinitesimal spherical shell with radius r around the atoms i. Since the density of the spherical shell approaches ρ_{h_j} for $r \to \infty$, the PCF generally reaches the value of 1 for long distances. Consequently, the PCF gives the deviation from the homogeneous density for all r. When i and j refer to different atomic types, the partial PCF $g_{ij}(r)$ gives the average fraction

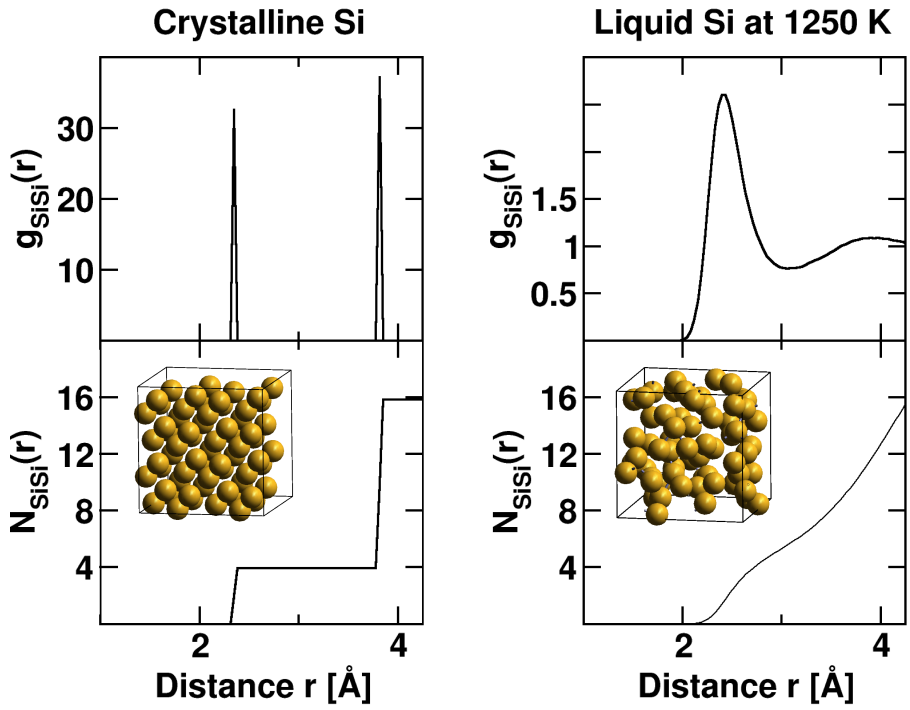


Figure 12: The PCF g_{SiSi} and the coordination number N_{SiSi} of crystalline and liquid Si. While the crystalline sample shows distinct, sharp maxima in the PCF (left hand side, upper panel) and a stepwise increase of the coordination number (left hand side, lower panel), the PCF of the liquid sample has much broader peak densities (right hand side, upper panel) and the coordination number increases smoothly (right hand side, lower panel).

of atoms with type j around atoms with type i at distance r. Thus one can derive the nearest neighbor distances (i.e. the bonding distances) directly from the first peak of the PCF and the next nearest distances from its second peak. Furthermore, the integration of the PCF up to its first minimum r_{min},

$$N_{ij} = \rho_{h_j} \int_0^{r_{\mathrm{min}}} dr\ 4\pi r^2\ g_{ij}(r), \tag{189}$$

defines the partial coordination numbers N_{ij} giving additional information on the composition of the nearest neighbor environment.

Fig. 12 illustrates the characteristic features of the PCF on the basis of crystalline and liquid Si. For the crystalline sample (panels on the left hand side), the PCF has sharp, distinct maxima at the nearest and the next nearest neighbor distances 2.345 Å and 3.815 Å, respectively (left hand side, upper panel). Likewise, we observe at both distances an immediate jump of the coordination number, from 0 to 4 and from 4 to 16, meaning that each Si atom has 4 nearest

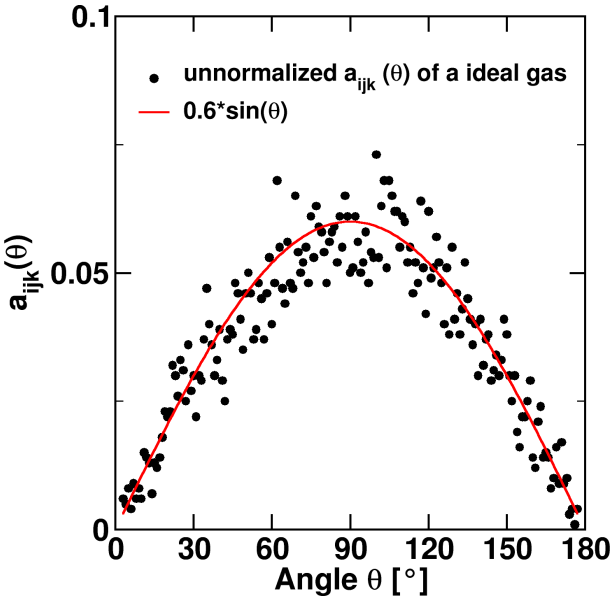


Figure 13: The unnormalized ADF of a ideal gas is proportional to $\sin(\Theta)$. In order to describe the deviation from the ideal gas correctly, the ADF of amorphous and liquid structures must be divided by $\sin(\Theta)$.

neighbor atoms and 16 next nearest neighbor atoms (left hand side, lower panel). In the liquid sample (panels on the right hand side), the peaks of the PCF are much less pronounced than in the crystalline structure and the peak positions are slightly shifted to 2.42 Å and 3.93 Å (right hand side, upper panel). Since the number of neighboring atoms increases now from 0 to 16 continuously, the number of atoms in the nearest neighbor shell is usually estimated at the first minimum of the PCF.

3.3 Angular distribution function (ADF)

Similar to the PCF, one can define the ADF by [130, 131]

$$a(\theta) = \frac{< \sum_{i \neq j} \sum_{k<i, \neq j} \delta(\theta_{ijk} - \theta) \Theta(r_{ji}^{\max} - r_{ji}) \Theta(r_{jk}^{\max} - r_{jk}) >}{\sin(\theta)}. \tag{190}$$

Here the Kronecker Delta δ gives the distribution of the bonding angles θ_{ijk} between atoms i, j, and k averaged over all centering atoms j. Furthermore, the Heaviside step function Θ limits the number of neighboring atoms i and k, around the centering atom j, to those with bonding distances r_{ji} and r_{jk} smaller than $r_{ji}^{\max}$ and $r_{ji}^{\max}$, respectively. The ADF is normalized by the form of the ADF of an ideal gas, see Fig. 13. When the indices i, j, and k represent atom types, the partial ADF $a_{ijk}(\theta)$ tells us the average number of bonding angles between the considered

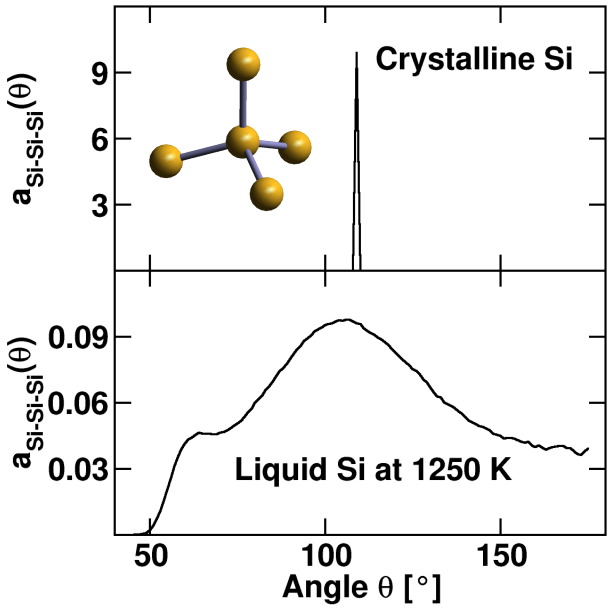


Figure 14: The ADF of crystalline and liquid Si. While the crystalline structure shows only tetrahedral bonding angles of 109.5° (upper panel), the bonding angles in the liquid can basically vary between 50° and 180° (lower panel). Note that the contributions from angles that are greater than 150° are amplified due to the normalization of the ADF (i.e. division by $\sin(\Theta)$), compare also Fig. 13.

atom types. In this case, the maximum bonding distances, $r_{ji}^{\max}$ and $r_{ji}^{\max}$, are usually estimated from the first minima of the corresponding partial PCFs.

Fig. 14 illustrates the main differences between the ADF of crystalline and liquid Si. In crystalline Si (upper panel), every Si atom is bonded to four Si atoms by forming a perfectly tetrahedral structure. For this reason, the ADF of c-Si has only one bonding angle at 109.5°. In contrast, the ADF of the liquid sample (lower panel) shows a variety of possible bonding angles between different Si atoms. Although many of the bonding angles are still close to the tetrahedral angle, the contributions from lower and larger angles have significantly increased in comparison to c-Si.

3.4 Structure factor (STF)

Long range order and density fluctuations can be investigated in terms of an ensemble averaged Faber-Ziman structure factor (STF)

$$S(\mathbf{k}) = \frac{1}{N} \left\langle \sum_i \sum_j \exp[-i\mathbf{k}(\mathbf{R}_i - \mathbf{R}_j)] \right\rangle . \tag{191}$$

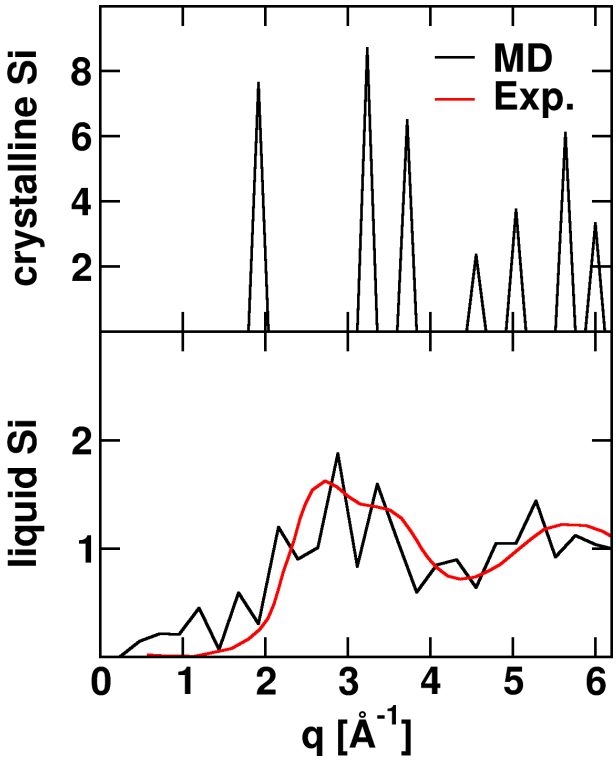


Figure 15: The static STF of crystalline and liquid Si (black lines). We can see that the STF of c-Si has distinct and sharp peak positions due to the long range range order of the crystalline structure (upper panel). However, there is no long range order in liquid Si so that the corresponding STF is a jagged, oscillating curve (lower panel). The STF is often calculated to compare with x-ray and neutron scattering experiments (red line). Here the experimental x-ray data for molten Si is obtained by Waseda and Suzuki [132]. The difference between the theory (black line) and the experiment (red line) can be explained by the insufficient size of the supercell.

Here $\mathbf{k}$ are wave vectors, N is the total number of atoms, and $\mathbf{R}_i$ are the corresponding atomic positions [133–136]. To obtain the isotropic STF one can calculate $S(\mathbf{k})$ for each point of the reciprocal lattice, which is built from the reciprocal basis vectors of the supercell. Afterwards, one can integrate all values of $|S(\mathbf{k})|^2$ within the infinitesimal sphere between q and $q + dq$ to obtain

$$S(q) = \frac{\int |S(\mathbf{k})|^2 \, \delta(|\mathbf{k}| - q) \, d^3k}{4\pi q^2}. \tag{192}$$

As with the PCF, the normalization factor comes from the volume element $dV = dq \, 4\pi q^2$ of an infinitesimal spherical shell, but in reciprocal space. For only one atom type, the partial STF $S_{\alpha\alpha}(q)$ can be obtained by a simple summation over all atoms of type α in Eq. (191). In this case, $S_{\alpha\alpha}(q)$ reaches the value 1 for $q \to \infty$. For mixed atom types, the partial STF $S_{\alpha\beta}(q)$ must be derived from the implicit equation

$$S_{\alpha+\beta}(q) = \sum_{\alpha\beta} (c_\alpha c_\beta)^{\frac{1}{2}} S_{\alpha\beta}(q), \tag{193}$$

where $S_{\alpha+\beta}$ is the total STF of the two atom types α and β, and c_α and c_β are the corresponding concentrations. [135]. In consequence of this implicit calculation, $S_{\alpha\beta}(q)$ has, in contrast to $S_{\alpha\alpha}(q)$, also negative values and reaches the value of 0 for $q \to \infty$.

Fig. 15 compares the STFs of crystalline and liquid Si with each other. The STF of c-Si has sharp and distinct peak positions due to the long range order in c-Si (upper panel, black line). However, in liquid Si, there is no long range order and we observe a jagged, oscillating curve (lower panel, black line). The jagged line is also related to insufficient sampling, since large supercells and many different structures are needed to obtain a good quality of the structure factor. Therefore, the PCF is mostly more suitable to compare the structural properties obtained by computer simulations. Nevertheless, the STF is often calculated to compare with experimental data. This is because experimentalists must calculate the PCF using the Fourier transformation of the STF. For example, the figure shows the x-ray scattering STF of molten Si that was calculated by Waseda and Suzuki (lower panel, red line) [132]. Another reason to calculate the STF is to examine density fluctuations. An indication for such fluctuations is, for example, the increase of the STF for $q \to 0$ (not shown in this figure).

3.5 Correlation and statistical inefficiency

In order to investigate the correlation between different amorphous sample structures, one can calculate the normalized autocorrelation function and the statistical inefficiency for the instantaneous potential energy. This energy value is defined by the total energy E_{tot} of a sample

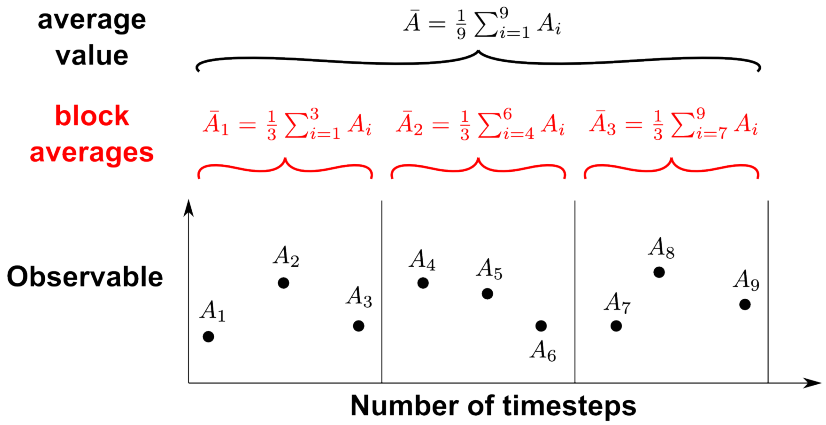


Figure 16: Referring to a series of measured observables, one has to calculate the variance of the average value and the variances of the block average values to estimate the statistical inefficiency. The figure illustrates the difference between both types of average values for 3 blocks with block length 3.

without the kinetic energy T of the ions

$$U_{\text{eff}}(\{\mathbf{R}_\alpha\}) = E_{\text{tot}}(\{\mathbf{R}_\alpha\}, \{\dot{\mathbf{R}}_\alpha\}) - T(\{\dot{\mathbf{R}}_\alpha\}). \tag{194}$$

For a simulation run with simulation length T, the autocorrelation function of an arbitrary observable $A(\tau)$ is furthermore given by [137]

$$C_{AA}(\tau) = \langle A(0)A(\tau)\rangle = \lim_{T\to\infty} \frac{1}{T}\int_0^T dt\, A(t)A(t+\tau). \tag{195}$$

The correlation function C_{AA} is equal to $\langle A^2\rangle$, for $\tau = 0$, and equal to $\langle A\rangle^2$, for $\tau \to \infty$, if the information at $\tau = 0$ and at $\tau = \infty$ are entirely uncorrelated. For this reason, the correlation function can be normalized by writing [137]

$$C^n_{AA}(\tau) = \frac{C_{AA}(\tau) - \langle A\rangle^2}{C_{AA}(0) - \langle A\rangle^2}. \tag{196}$$

This normalized version has the advantage that it is equal to 1, for $\tau = 0$, and equal to 0, for $\tau \to \infty$. For many systems, the normalized autocorrelation function decays exponentially $e^{-\tau/\tau_c}$ with the correlation time τ_c.

The statistical inefficiency is closely related to the autocorrelation function and it is basically given by the ratio between the variances $\sigma^2_{\langle A\rangle_b}$ and $\sigma^2_{\langle A\rangle}$ of the two different average values $\langle A\rangle_b$ and $\langle A\rangle$ [110].

$$s = \lim_{T_b\to\infty} T_b \frac{\sigma^2_{\langle A\rangle_b}}{\sigma^2_{\langle A\rangle}} \tag{197}$$

Here $\langle A \rangle$ is the usual average value of a observable A. If A_τ is a series of measured observables during the simulation run with simulation length T, this average value is given by

$$\langle A \rangle = \frac{1}{T} \sum_{\tau=1}^{T} A_\tau. \tag{198}$$

The corresponding variance of the measured observables A_τ is

$$\sigma^2 = \frac{1}{T} \sum_{\tau=1}^{T} \Big(A_\tau - \langle A \rangle \Big)^2. \tag{199}$$

If the series A_τ is now split into N_b blocks with block length T_b, the average value of each block b is [110]

$$\langle A \rangle_b = \frac{1}{T_b} \sum_{\tau=1}^{T_b} A_\tau, \tag{200}$$

and the corresponding variances of the block averages $\langle A \rangle_b$ are [110]

$$\sigma_b^2 = \frac{1}{N_b} \sum_{b=1}^{N_b} \Big(\langle A \rangle_b - \langle A \rangle \Big)^2. \tag{201}$$

In Fig. 16, the difference between the average value $\langle A \rangle$ and the block average values $\langle A \rangle_b$ is illustrated for $T = 9$ and $N_b = T_b = 3$. If the block length T_b in Eq. (197) is finally increased from 1 to T, the plateau value of the statistical inefficiency is twice the correlation time τ_c of the autocorrelation function. [110]

The relationship between the instantaneous potential energy, the autocorrelation function, and the statistical inefficiency is further explained in Fig. 17 and Fig. 18. Fig. 17 shows the energy distribution during the simulation of liquid Si at 1500 K. A normalized Gaussian function (red line) is fitted to the data points of the simulation (black dots). The average value of the instantaneous potential energy lies at -326.8 ± 1.3 eV. The normalized autocorrelation function and the statistical inefficiency of the energy values is shown in Fig. 18. After 700 time steps (i.e. about 1 ps) the energy values of the samples are approximately uncorrelated because the autocorrelation function has dropped down to 0 (upper panel, black line). The correlation time is then estimated in terms of timesteps by fitting a exponential function (upper panel, red line) to the decreasing values of the correlation function. It can be seen that $2\tau_\mathrm{c} = 240$ (lower panel, red line) is indeed close to the plateau value of the statistical inefficiency, which lies between 200 and 250 (lower panel, black line). The plateau value tells us that only every 240th structure contributes completely new information to the statistical average value of about -327 eV. [110] The statistical inefficiency can be decreased using less correlated samples for the calculation of the average instantaneous potential energy. For example, this can be achieved by selecting only low correlated, approximately equidistant snapshots from the trajectory of the simulation.

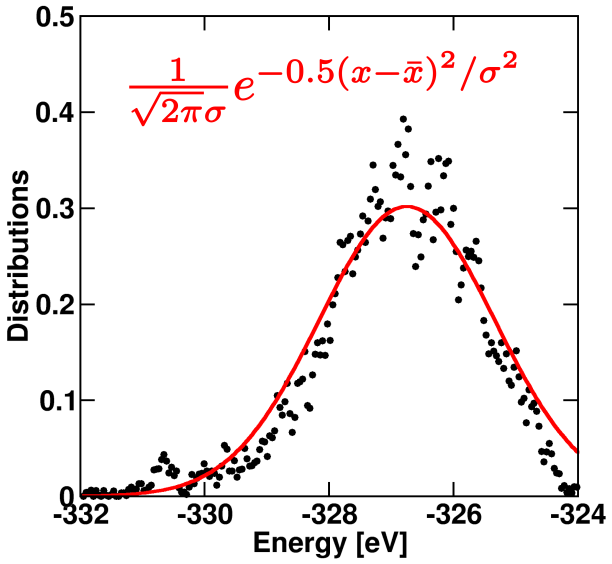


Figure 17: The distribution of the instantaneous potential energies during the simulation of liquid Si at 1500 K (black dots). A normalized Gaussian function is fitted to the data points (red line). The average instantaneous potential energy is about 327 eV ($\bar{x}$) with a standard deviation of ± 1.32 eV (σ).

4 Simulation setup

"For the presentation of our simulation setup, we closely follow the content of our previous publications about bulk amorphous silicon nitrides [138, 139]".

In all of our *ab initio* MD simulations, we applied periodic boundary condition (PBC) and the projector augmented wave (PAW) method [114, 140] as implemented by Kresse and Joubert in the Vienna *ab initio* simulation package (VASP) [115, 140, 141]. Furthermore, we normally used the PBEsol functional to describe exchange and correlation effects [89]. Only for final structures at the ionic ground state, which was determined using the PBEsol functional, the HSE hybrid functional and the GW method were sometimes applied to correct for the underestimated band gaps in Kohn-Sham density functional theory (KS-DFT). For the finite temperature simulations, we restricted the Brillouine sampling to the Γ-point, used a time step of 1.5 fs, and applied a Nose thermostat [117, 126] to model a canonical ensemble. The mass of hydrogen was always increased to 10 atomic mass units. This mass increase will change the dynamics of the system, but has no influence on the explored configuration space. This is because the partition function factorizes into a momentum and configuration dependent part, and the latter one is independent of the masses. In the following sections, we explain the construction of our calculation cells and

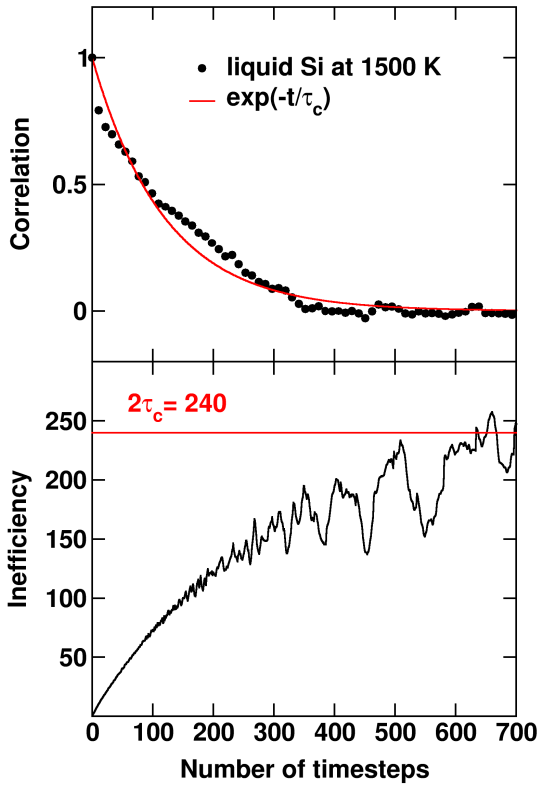


Figure 18: The normalized autocorrelation function (upper panel, black dots) and the statistical inefficiency (lower panel, black line) for the instantaneous potential energy values of the simulation of liquid Si at 1500 K. An exponential function is fitted to the data points of the correlation function (upper panel, red lines). The plateau value of the statistical inefficiency is twice the correlation length τ_c and tells us that only every 240th value out of 7000 values contributes completely new information to the average instantaneous potential energy in Fig. 17.

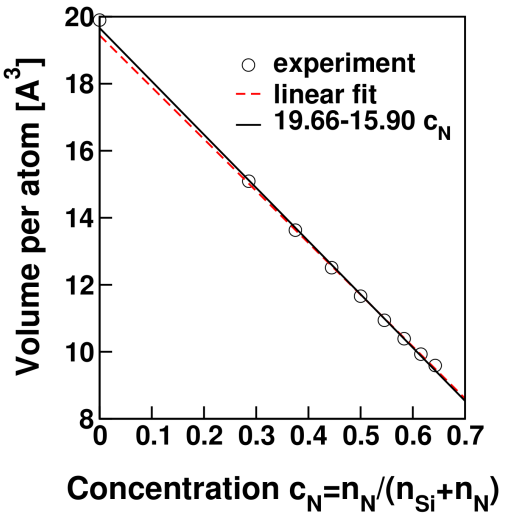


Figure 19: Volume per atom as a function of the N concentration $c_{\mathrm{N}} = n_{\mathrm{N}}/(n_{\mathrm{N}} + n_{\mathrm{Si}})$. The broken line shows the best linear fit to the experimental values (circles). The full line was obtained by readjusting the fit such that the density of crystalline Si is more accurately reproduced. The experimental data points for Si_3N_x:H are from M. Guraya *et al.* [27], and the density for crystalline Si is from N. W. Ashcroft and N. D. Mermin [5]. Note that c_N is estimated by neglecting the H content in the samples.

the applied MD treatment in more detail.

4.1 The system sizes and the a-Si_3N_x:H composition of the bulk and the interface samples

Non-stoichiometric amorphous silicon nitrides exhibit large density variations that seemingly depend strongly on the applied preparation conditions [23, 142, 143]. In this work, we mainly focus on amorphous silicon nitrides as used in solar cell industry and decided to determine the densities from the data published by Ippolito and Meloni [57]. We found it helpful to plot the average experimental volume per atom versus the nitrogen concentration, as shown in Fig. 19. The volumes would lie on a straight line, if Vegard's law was observed [144]. Indeed, within the experimental uncertainties, this relation seems to hold very well, but the fit deviates from the volume of crystalline silicon in the Si rich case by more than 2 % (red dashed line). Increasing the slope improves this behavior for the Si rich case (black full line) without deteriorating the root mean square error for the amorphous samples significantly. The final densities and volumes, as obtained from the fitted experimental data, are summarized in Tab. 1.

Table 1: The densities ρ and the volumes per atom of the Si_3N_x compositions as studied in the present work. The densities were derived from experimental values obtained by Guraya *et al.* [27], see Fig. 19. The table also shows the supercell volumes and the number of Si n_{Si} and N atoms n_{N} in the considered supercells.

	$Si_3N_{4.5}$	Si_3N_4	$Si_3N_{3.5}$	Si_3N_3
ρ [g/cm^3]	3.22	3.14	3.06	2.98
volume per atom [Å^3]	10.12	10.57	11.10	11.71
$n_N/(n_{\mathrm{Si}} + n_{\mathrm{N}})$	0.60	0.57	0.54	0.50
large supercell vol. [Å^3]	1973	1999	2020	2014
n_{Si}	78	81	84	86
n_{N}	117	108	98	86
small supercell vol. [Å^3]	1012	1033	1020	1077
n_{Si}	40	42	42	46
n_{N}	60	55	50	46

The second issue to address is the influence of H on the volume. The experimental data were almost always measured in the presence of substantial amounts of H approaching up to 25 at.% [27]. Remarkably, however, H seems to influence the volume only very little. For instance, although, for the data shown in Fig. 19, the H content varies from $15 - 25$ at.%, the deviation from Vegard's law is hardly noticeable, and, for Si_3N_4, the volume of the amorphous structure agrees within few percent with the volume of crystalline $\beta-Si_3N_4$. This observation is reinforced by considering the small covalent radius of H. Even if we double the covalent radius, $d_{\mathrm{H}} = 2r_{\mathrm{covalent}} = 0.62$ Å (this value corresponds to the H-H bond length), and determine the corresponding volume, $V_{\mathrm{H}} = 4\pi d_{\mathrm{H}}^3/3 \approx 1$ Å^3, we obtain only small corrections to the volume. If the concentration were 20 at.% H, it would take up at most 3 % of the entire volume. This is certainly a rather crude consideration, since one could argue that H disrupts the usual bonding topology of the network and therefore might yield a significantly larger volume increase, but neither the experimental data nor any of our results seem to support this idea. Most likely, H simply accommodates in the network at defective or highly strained sites and hence causes only negligible changes in the volume.

The relationship between composition and nitrogen concentration in a-Si_3N_x:H was provided by the Energy Research Center of the Netherlands (ECN). Our cooperation partner had prepared and analyzed over 80 different amorphous samples to extract the atomic concentrations that are shown in Fig. 20. Specifically, plasma enhanced chemical vapor deposition (PECVD) was used for the preparation and Fourier transform infrared (FTIR) spectroscopy and ellipsometry for the analysis. A more detailed description of the experimental methods is given by Lamers *et al.*, Soppe *et al.*, and Bustarret *et al.* [19, 22, 145]. Going from pure a-Si, $c_N \approx 0.0$, to a-Si_3N_4, $c_N \approx 0.57$, the data shows some distinct changes in element and bond concentrations. The number of Si atoms and Si-Si bonds decreases while the number of N atoms and Si-N bonds increases. The concentration of H atoms stays between 10 at.% and 15 at.% and the number of Si-H bonds usually exceeds the number of N-H bonds. Close to a-Si_3N_4, there are mainly Si-N bonds and the number of Si-H bonds drops down to the level of N-H bonds. These results are in good agreement with previous studies [27] and hence a good basis for our simulations.

As pointed out by Lamers *et al.* [22], compositions with a Si-N bond density between 120×10^{21} cm^{-3} and 130×10^{21} cm^{-3} are optimal for the bulk passivation in solar cells. On the other hand, the largest band gap and the lowest defect concentration is observed for nearly stoichiometric samples [27]. As a result, we concentrated our investigations on a-Si_3N_3:H, a-$Si_3N_{3.5}$:H, a-Si_3N_4. As shown in Fig. 20 (red symbols), the under stoichiometric samples were usually prepared with about 13 at.% H to obtain information on the defect passivation caused

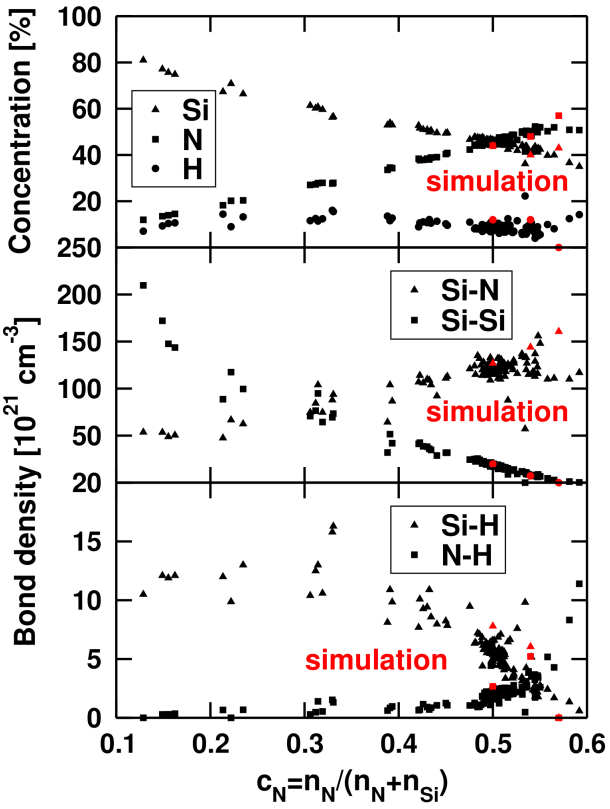
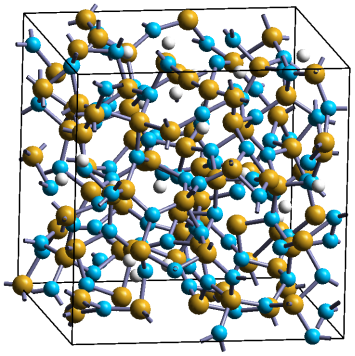
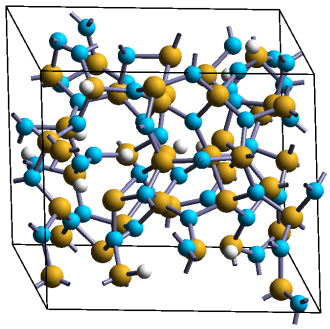


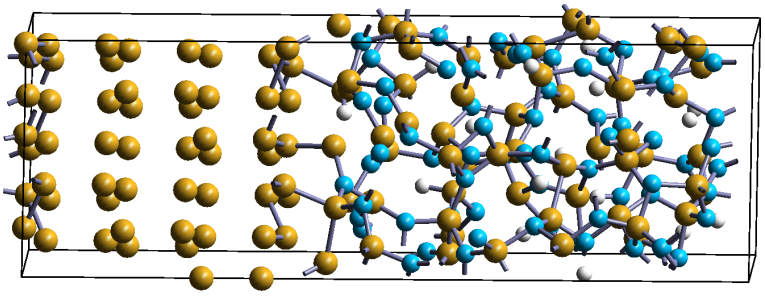
Figure 20: Experimental element compositions and bond densities in a-Si_3N_x:H for N concentrations between $c_N = 0.1$ and $c_N = 0.6$. For the simulations, we concentrated on the technologically relevant systems a-$Si_3N_3H_{0.8}$, a-$Si_3N_{3.5}H_{0.8}$, and a-Si_3N_4, with a N concentrations of $c_N = 0.50$, 0.54, and 0.57, respectively.



(a) A large bulk a-$Si_3N_{3.5}H_{0.8}$ structure with about 200 atoms in a cubic unit cell.



(b) A small bulk a-$Si_3N_{3.5}H_{0.8}$ structure with about 100 atoms in a hexagonal unit cell.



(c) A c-Si/a-$Si_3N_{3.5}H_{0.8}$ interface structure with about 200 atoms in a long hexagonal unit cell.

Figure 21: Typical bulk and interface structures in our simulations.

by H. In contrast, the a-Si_3N_4 samples were always simulated without H, because we expected a low defect concentration in the stoichiometric case. For the reason of completeness, we prepared some a-$Si_3N_{4.5}$ samples without H and some a-$Si_3N_{3.5}$ samples with 6 at.% and 20 at.% H. Thereby we cover also the case of too much N (i.e. over stoichiometry) and the case of a varying H concentration.

Fig. 21 shows example structures of our final bulk and interface samples. For bulk a-Si_3N_x:H, we simulated small and large sample sizes. The supercells of the large samples contained about 200 atoms in a cubic volume of about 2000 $Å^3$, compare Fig. 21(a) and the data for large supercells in Tab. 1. Note that the values shown in Tab. 1 apply for the a-Si_3N_x:H compositions with and without H, because the H atoms were added without changing the volumes. According to the fit in Fig. 19, the density of each system was between 2.98 and 3.22 $g\,cm^{-3}$ depending on the a-Si_3N_x:H composition. For the small sample structures, we used the same densities as for our large samples, see Tab. 1. The small samples contained about 100 atoms in a hexagonal unit cell of 1000 $Å^3$, compare Fig. 21(b). Finally, our c-Si/a-$Si_3N_{3.5}$:H interface systems were

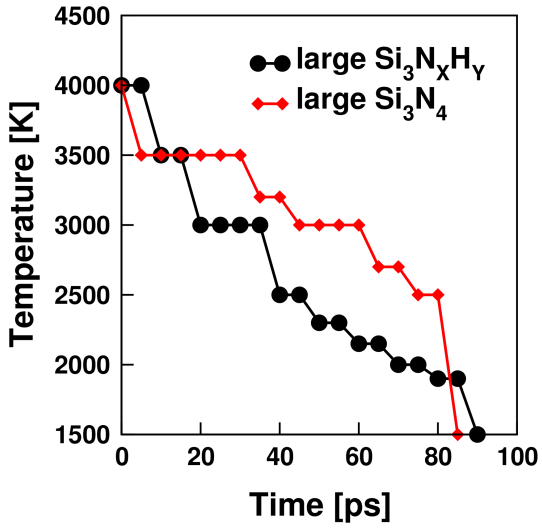


Figure 22: The stepwise cooling scheme of the large a-Si_3N_{4-x}:H structures. The temperature is stepwise decreased during the MD simulation run. The simulation of a-Si_3N_4 stayed longer at higher temperatures to prevent the stoichiometric samples from freezing too early.

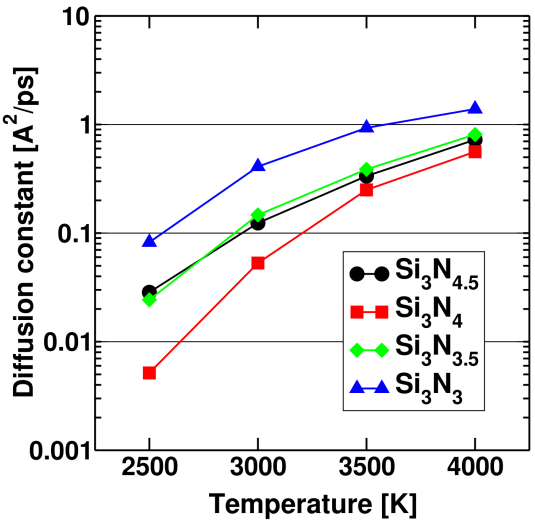


Figure 23: Diffusion constants averaged over both atom types (Si and N) in the large a-Si_3N_x sample structures at different temperatures. Stoichiometric a-Si_3N_4 typically freezes at higher temperatures and shows the lowest diffusion constant at all temperatures.

built by combining c-Si with a-$Si_3N_{3.5}$:H, see Fig. 21(c). The a-$Si_3N_{3.5}$:H parts were based on a hexagonal unit cell with 59 Si, 69 N and 17 or 1 H atoms in a volume of 1429 Å^3. The resulting densities of 3.06 or 3.04 g cm^{-3} were the same as used for the bulk samples, compare Tab. 1. The c-Si part was modeled by a $\sqrt{7} \times \sqrt{7}$ unit cell containing 4 double layers of 7 Si atoms in (111)-direction (i.e. 56 Si atoms). The density of this part was 2.32 g cm^{-3} with a cell volume of 1122 Å^3.

4.2 Stepwise cooled a-Si_3N_x:H bulk samples

Large cubic a-Si_3N_x:H bulk systems were generated using a stepwise cooling scheme. For each composition, the starting configuration was initially molten at 4000 K and then equilibrated for about 15 ps at 4000 K. Then the temperature was gradually decreased until the atomic configuration started to freeze in (i.e. constant MSD). To prepare a set of samples, we extended the simulation runs at 3500 K and 3000 K for a-Si_3N_4 and a-Si_3N_{4-x}, respectively. In this manner, three samples were generated for a-Si_3N_4, a-$Si_3N_{3.5}$, and a-Si_3N_3. For less relevant a-

$Si_3N_{4.5}$ and the hydrogenated systems only a single large representative structure was generated.

The adopted temperature profiles are shown in Fig. 22. Each constant temperature step lasted between 5 and 20 ps with a time step of 1.5 fs, and the total simulation time was between 80 and 100 ps. In order to prevent the systems from freezing at too high temperatures, we elongated the simulation from the last snapshot of the previous temperature step whenever we felt that the system was not well equilibrated. The main objective was to keep the MSD roughly similar for each system composition. As atom diffusion is defect mediated [146] and stoichiometric compositions have in general a lower defect concentration than under and over stoichiometric samples, a-Si_3N_4 shows a very low atom diffusion at low temperatures. For this reason, we simulated a-Si_3N_4 without H longer at higher temperatures, see red line in Fig. 22.

In general, the "freezing" temperature was strongly dependent on the system composition. This is demonstrated in Fig. 23, where the diffusion constant of different a-Si_3N_x compositions is averaged over both atom types (Si and N) and plotted versus the temperature. The highest 'freezing' temperature was observed for a-Si_3N_4 and the lowest for a-Si_3N_3. Stoichiometric a-Si_3N_4 freezes in at about $3000-3100$ K on a timescale of $20-40$ ps, which we observed repeatedly for several simulations at different system sizes, and the freezing goes along with the geometrical defects being progressively removed during the annealing. Usually the system locks in when two or zero defects are present in the sample. Note that defects in Si_3N_4 always come in pairs for topological reasons.

The observed 'freezing' temperature seems to be significantly larger than the experimental estimates of about 2000 K [147, 148]. We note though that the melting temperature of silicon nitride can not be measured, since the solid decomposes into atomic and molecular Si and N before melting sets in (gas pressure of Si). Matter of fact, our simulations are also hampered because of the short timescales and small system sizes that are accessible to us. In our simulations, a system freezes once all defects are cured. In real experimental samples, however, the density of mobile defects will only exponentially approach zero, sustaining diffusion at much lower temperatures. Furthermore, we believe that the experimental samples are never perfectly stoichiometric. Even being slightly off-stoichiometric, diffusion is enhanced since some defective bonds (e.g. Si-Si bonds) or coordination defects prevail. For a-$Si_3N_{3.5}$ and a-$Si_3N_{4.5}$, freezing occurs at roughly 2600 K, whereas, for Si_3N_3, the freezing occurs at roughly 2300 K. Finally, addition of 10 at.% H generally lowers the 'freezing' temperature by another 200 K (not shown). We also note that jump diffusion is well possible below these temperatures, but these events are very rare on the timescales accessible to us.

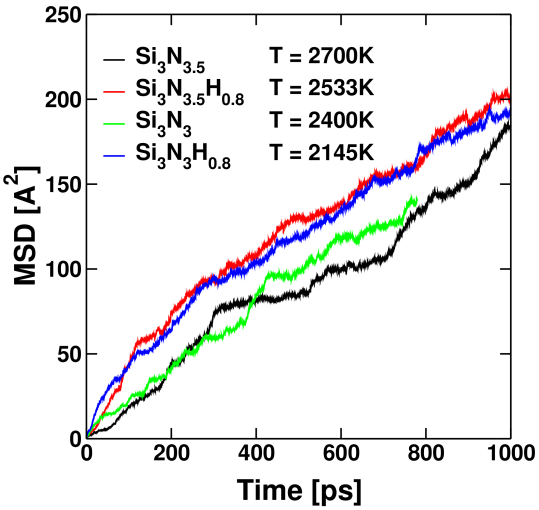


Figure 24: The MSD of different a-Si_3N_x:H structures during a 1 ns MD simulation. The temperature was adjusted so that the slopes of the MSDs were nearly equally steep. This corresponds to a similar diffusion constant for all a-Si_3N_x:H compositions.

4.3 Rapidly quenched and slowly cooled a-Si_3N_x:H bulk samples

We generated many configurations of small hexagonal sample structures for several a-Si_3N_x:H compositions with H (i.e. a-$Si_3N_{3.5}H_{0.4}$, a-$Si_3N_{3.5}H_{0.8}$, a-$Si_3N_{3.5}H_{1.7}$, and a-$Si_3N_3H_{0.8}$) and without H (i.e. a-Si_3N_3, a-$Si_3N_{3.5}$, a-Si_3N_4, and a-$Si_3N_{4.5}$). The liquid phase for all these compositions was initially equilibrated for 15 ps at 4000 K. The samples were then immediately quenched to about 200 K above their typical 'freezing temperature' and next equilibrated for 15 ps at the same temperature. Depending on the a-Si_3N_x:H composition, this temperature was typically between 2700 K and 2100 K, compare also Fig 23 of the previous section. Note that we have actually removed 4 atoms from the a-Si_3N_4 composition to introduce a few more defects that keep the diffusion in the nearly stoichiometric composition sizeable at around 3000 K. Thereafter, a 1 ns *ab initio* MD run was performed at approximately the same temperature as before. In order to obtain a comparable protocol, the temperatures were slightly adjusted during the simulation so that the diffusion was approximately constant and equal for all system compositions. For instance, Fig. 24 shows the similar slopes of the MSDs for four of our a-Si_3N_x:H compositions.

As illustrated in Fig. 25, we finally applied both a rapid and a slow annealing scheme. On the one hand, for all a-Si_3N_x:H composition, 1000 structures were generated by rapidly quenching (i.e. relaxing) every pico second one structure into the closest local minimum. We refer to these structures as 'rapidly quenched'. On the other hand, for a-Si_3N_3 and a-$Si_3N_{3.5}$ with and

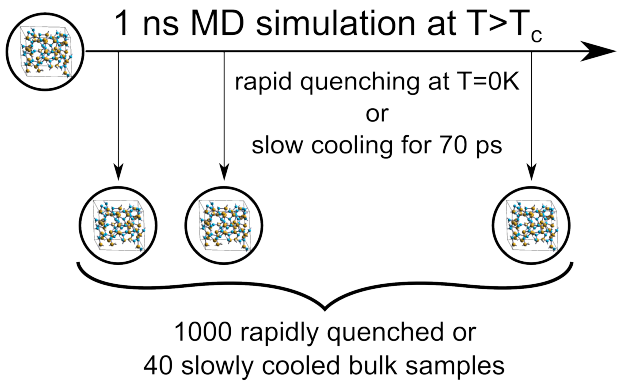


Figure 25: The creation scheme of the small a-Si_3N_x:H bulk structures. We performed a 1 ns *ab initio* MD simulation right above the freezing temperature T_c of the samples. Afterwards, we drew 1000 or 40 snapshots out of the trajectory of the large scale MD run. These snapshots were then rapidly quenched (i.e. relaxed at $T = 0$ K) or slowly cooled, respectively.

without H, 40 models were generated by slowly cooling selected structures from the long MD run. The initial starting structures for the slow cooling were taken from the long trajectory at roughly equal spacing. Then they were cooled slowly from the melt into the amorphous state by lowering the temperature by 1000 K in 75 ps. We refer to these structures as 'slowly cooled'. At the lowest temperature, diffusion was absent in all samples (i.e. constant MSD). For 4 out of the 40 slowly cooled a-$Si_3N_{3.5}$:H structures, the cooling protocols are shown in Fig. 26

To gain some insight on the difference between the rapidly quenched and slowly cooled structures, we investigated the distribution of the instantaneous potential energy in Fig. 27. The figure shows that the energy distribution is much narrower for the slowly annealed samples than for the rapidly quenched samples. However, except for the a-$Si_3N_{3.5}$ samples with and without H, the slowly cooled ensembles hardly exhibited lower energy structures than the rapidly quenched ensembles. For slowly cooled a-$Si_3N_{3.5}H_{0.8}$ and a-$Si_3N_{3.5}$, the structures with lowest energy are a little bit shifted in comparison to the rapidly quenched samples.

Since all structures, for a specific a-Si_3N_x:H composition, were produced by one single MD run, they might be to some extent correlated. For this reason, we investigated the normalized time correlation function of the instantaneous potential energy, as described in section 3.5. In summary, we determined that every 40th sample of the rapidly quenched ensembles contributes completely new information to the average instantaneous potential energy, while the 40 slowly cooled sample structures can be considered as independent, since we essentially chose uncorrelated start configurations from the long MD run.

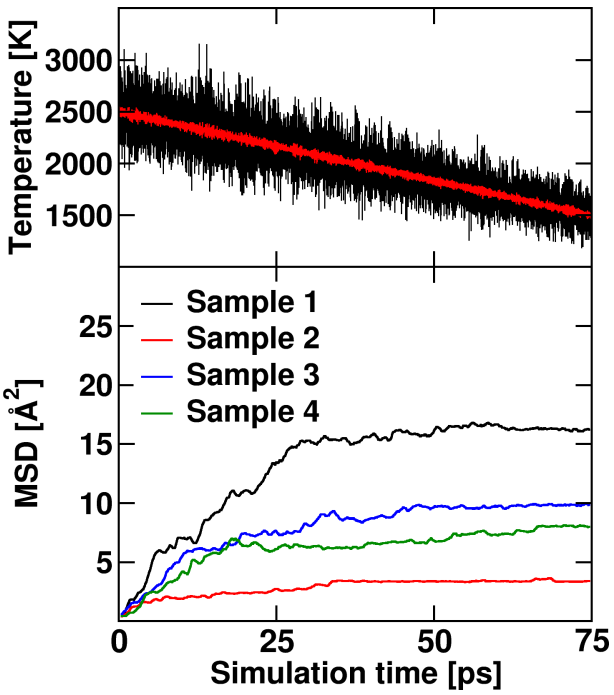


Figure 26: The temperature and the MSD of 4 of our a-$Si_3N_{3.5}H_{0.8}$ samples during a 75 ps cooling run. The MSD is averaged over all Si, N, and H atoms. After 40 ps the MSDs of all samples reach an approximately constant value, and the structures can be considered as completely frozen.

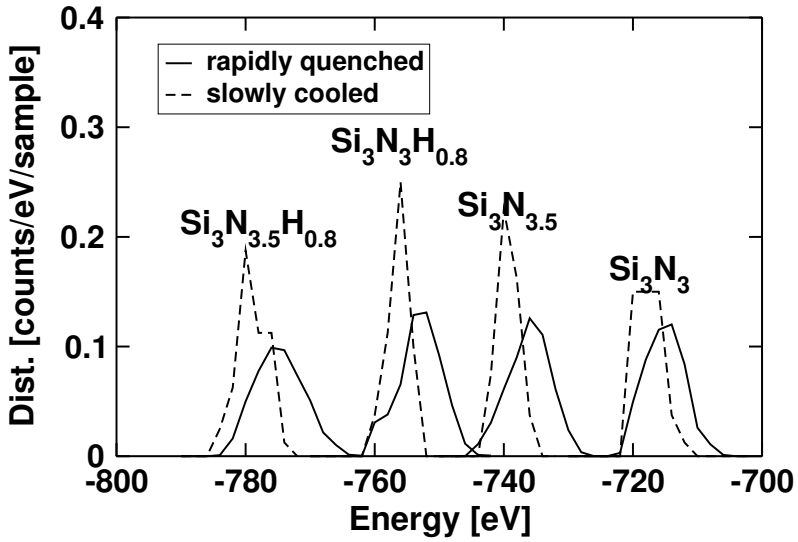


Figure 27: The instantaneous potential energy distribution of the rapidly quenched and the slowly cooled a-Si_3N_x:H bulk samples. For the slowly cooled samples (broken line), we find a significantly narrower energy distribution than for the rapidly quenched samples (full line).

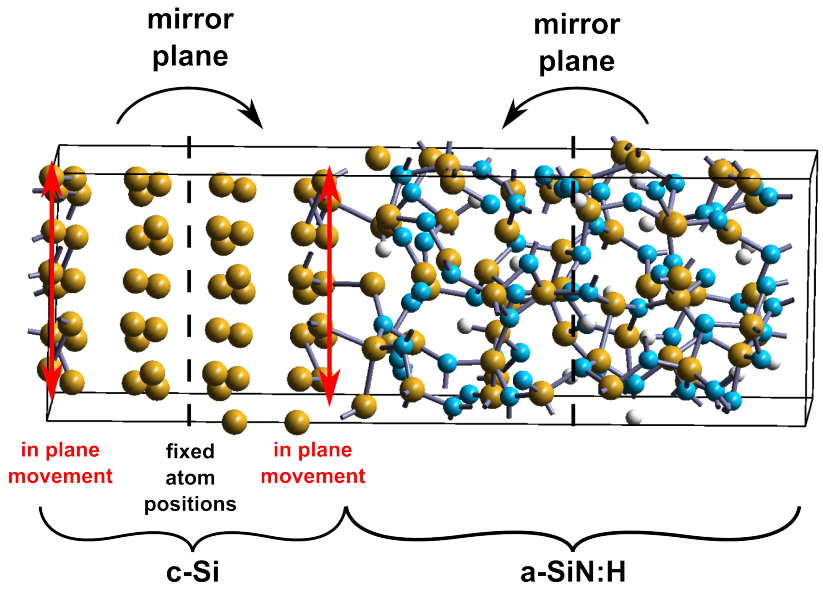


Figure 28: A simulation cell of a c-Si/a-$Si_3N_{3.5}H_{0.8}$ interface structure with about 200 atoms. During the MD simulation at finite temperature the c-Si atoms were partially frozen. Specifically, the positions of the c-Si atoms in the middle were always fixed, while the atom positions within the first c-Si layer at the interface were allowed to move parallel to the interface only. In the last relaxation step, the atoms of the first two double layers of c-Si were freed so that these atoms could adjust to the a-Si_3N_x:H structure. To investigate the spatial location of gap states close to the interface, we mirrored the number densities of the elements and the densities of defect states in the middle of the c-Si part and in the middle of the a-Si_3N_x:H part.

4.4 Rapidly quenched and slowly cooled c-Si/a-Si_3N_x:H interface samples

As summarized in Fig. 29, the generation process of our c-Si/a-Si_3N_x:H interfaces was divided into several steps. First, we generated a melt of the bulk a-$Si_3N_{3.5}$:H parts, without the c-Si bulk, by staying 15 ps at 4000 K. Then we immediately reduced the temperature to 2500 K or 2700 K, which is still above the freezing temperature of the amorphous samples with 11 or 1 at.% H (compare the temperatures in Fig. 23 and Fig. 24). After 7.5 ps equilibration time, we next started a long simulation run at the same temperature lasting for 225 ps. From the resulting trajectory we then drew 30 bulk a-$Si_3N_{3.5}$:H samples with an effective time difference of 7.5 ps between each sample. The extracted a-Si_3N_x:H samples were combined with bulk c-Si to form c-Si/a-Si_3N_x:H interface structures. As illustrated in Fig. 28, we locked the inner c-Si atoms for the following finite temperature simulations of the interface. Furthermore, the c-Si atoms of

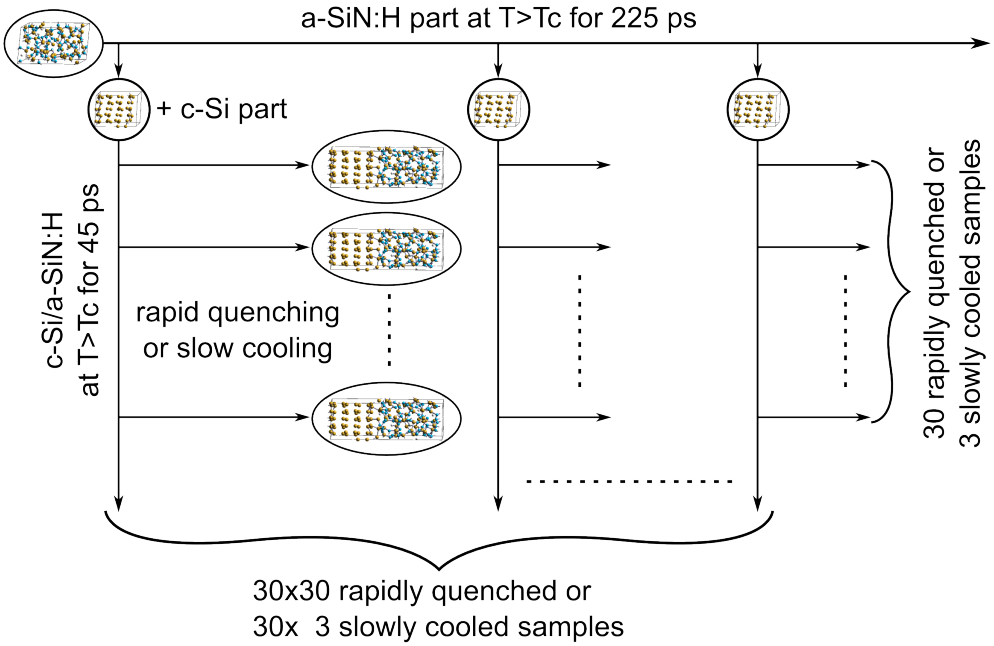


Figure 29: The creation scheme of the c-Si/a-Si_3N_x:H interface structures. Firstly, a-Si_3N_x:H bulk structures were simulated at a temperature T right above the freezing temperature T_c for 225 ps. Then we selected 30 equidistant snapshots of the bulk trajectory and combined them with the c-Si part to form c-Si/a-Si_3N_x:H interfaces. The interfaces were then simulated for 45 ps. We afterwards selected 30 or 3 snapshots from each interface trajectory, and these snapshots were rapidly quenched (i.e. relaxed at $T = 0$ K) or slowly cooled, respectively.

the first layers of the crystalline parts were allowed to move only parallel to the interface (i.e. sharp interface models). Note that the crystalline parts would have been destroyed by partial melting and diffusion of N into the c-Si part, if the c-Si layers had not at least partially been locked. In the next step, we equilibrated the interfaces for 4.5 ps at 2500 or 2700 K. After this second equilibration step, another long simulation run was started, staying 45 ps at the same temperature. From each trajectory of the 30 combined interfaces, we eventually extracted 30 interface samples with an effective time difference of 1.5 ps between each sample.

In this way, we obtained 900 (i.e. 30×30) interface configurations for the c-Si/a-$Si_3N_{3.5}$:H interfaces with 17 or 1 H atoms. From then on two different cooling schemes were applied. While we immediately relaxed all 900 structures (i.e. rapidly quenched samples), we also selected 90 structures and slowly cooled them from 2500 or 2700 K to 1500 K in 37.5 ps and relaxed them afterwards (i.e. slowly cooled samples). In all final relaxation steps, we removed the restriction for the first two double layers of c-Si and let the atoms in these layers move freely in all possible directions.

Ultimately, the ground state energies of the rapidly quenched and the slowly cooled structures were recalculated with a $2\times2\times1$ k-point grid. In contrast to the calculations on bulk a-$Si_3N_{3.5}$:H, this last refinement step was necessary to accurately sample the band structure of c-Si. As the band dispersion in the c-Si part is large and the band gap minimum of Si is close to the X-point, not all states fold back to the Γ-point for the considered interface cells. Additional k-points were needed to capture these states.

In Fig. 30 the correlation of the instantaneous potential energy of our final interface structures was again investigated by means of the normalized autocorrelation function [70] and the statistical inefficiency [110]. The figure shows that the energy values were only weakly correlated and that every 7th structure gives completely new information to the statistical average value of the energy. The slowly cooled structures, on the other side, were approximately uncorrelated.

We also performed calculations with interfaces where the first two double layers of c-Si were allowed to move freely during the slow cooling processes (i.e. graded interface models) [149]. In Fig. 31, the number densities of the sharp and the graded interface models are compared with each other along the z-axis. In the graded models, the amorphous silicon nitride parts basically moved into the c-Si and thereby shifted the position of the interface. The original interface, i.e. the interface position before the atom positions of the first two double layers were freed, lied at -3 Å. Although the main peak positions of all element concentrations stayed approximately at the same distance from the interface, we observe a decreased N concentration towards the interface in the graded model structures. The N concentration vanishes where the interface has

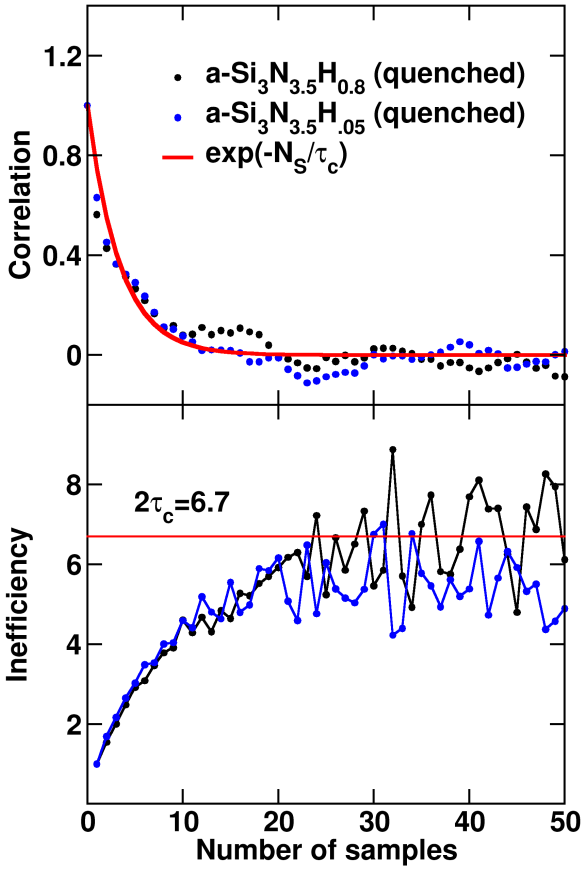


Figure 30: The normalized autocorrelation function and the statistical inefficiency of two rapidly quenched c-Si/a-Si_3N_x:H interface systems with different H concentrations. The correlation length τ_c is directly related with the plateau value of the statistical inefficiency [110].

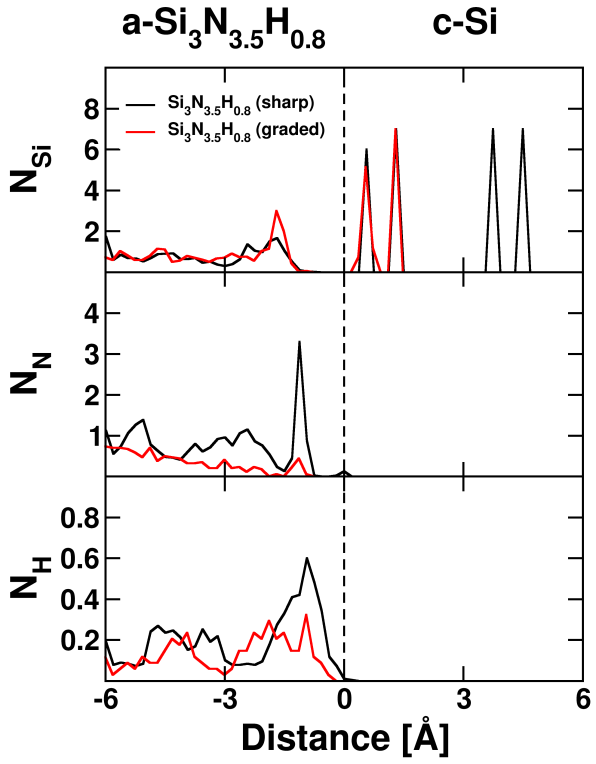


Figure 31: The number densities of Si, N, and H for sharp and graded c-Si/a-$Si_3N_{3.5}H_{0.8}$ interface models (black/red lines). The interface of the graded models has effectively shifted from its original position at -3 Å to 0 Å. Outgoing from the new interface position, the graded models show the same peak positions in the densities as the sharp models, but a decreased N concentration towards the interface.

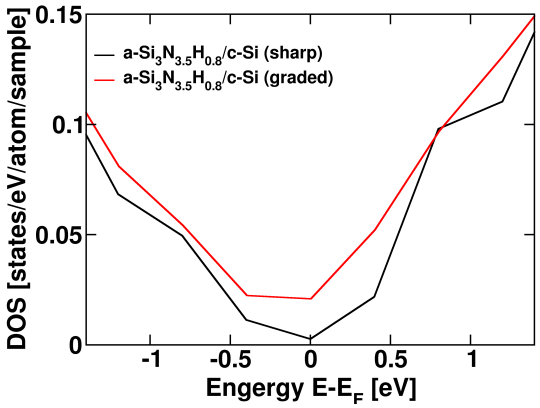


Figure 32: The DOS close to the Fermi level of sharp and graded c-Si/a-$Si_3N_{3.5}H_{0.8}$ interface models (black/red line) . In comparison, the defect density is generally larger for the graded interface models.

stopped to move into c-Si (i.e just before 0 Å). For both the graded and the sharp interface models, the H concentration is increased close to the interface. On the electronic level, the graded interface models show a larger defect concentration than the sharp models. This is for example shown in Fig. 32, where the DOS of the sharp and the graded models are compared close to the Fermi level. Since we found out that the overall defect properties are basically similar for both interface types and sharp interfaces are in fact quite commonly observed in TEM studies [149], we only investigate sharp interfaces models from now on.

4.5 A few comments on our simulation setup

The PAW potentials of Si and N were softened to reduce the plane wave basis set to a value of 150 eV [138]. In comparison to more accurate PAW potentials, the losses in accuracy were negligible small [138]. For instance, Fig. 33 shows that only small differences are found for the PCF and ADF of liquid $Si_3N_{3.5}$ at 3000 K . This makes us confident that the softened potentials give accurate results and are certainly far superior to MD simulations that use semi empirical potential models such as the Tersoff potential. The last statement is based on two observations. Firstly, our cooperation partners in Sheffield used a Tersoff potential with the parameters suggested by de Brito Mota *et al.* [49, 50] and had difficulties to generate structural properties that can be used as a input for our *ab initio* calculations. Specifically, the a-Si_3N_x:H samples based on their input structures were metallic, i.e. without band gap. These data are not shown in this work, as we focus only on our *ab initio* approach here. Secondly, Ippolito and

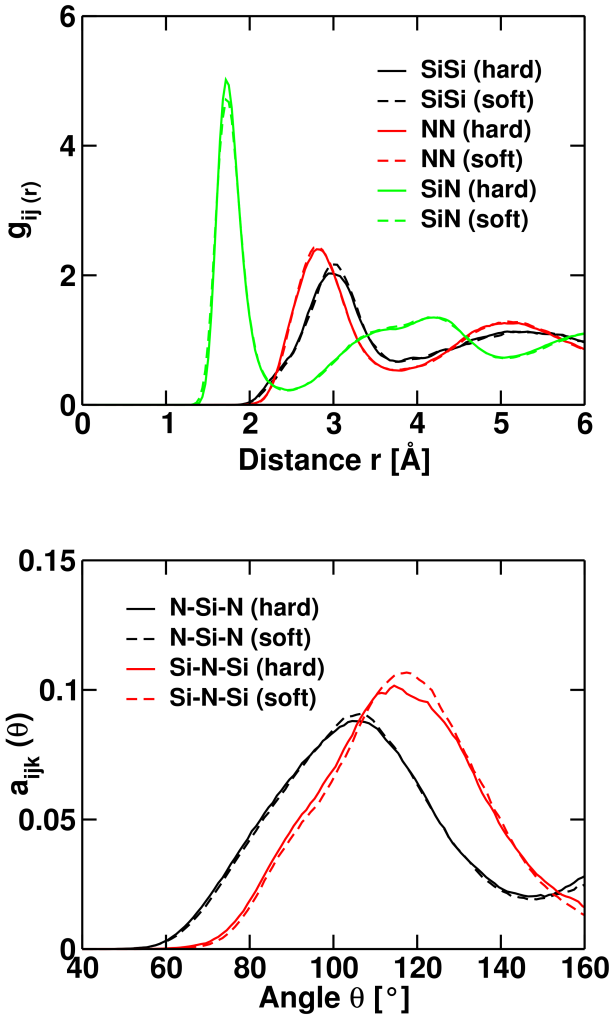


Figure 33: The PCFs $g_{ij}(r)$ and the ADFs $a_{ijk}(\Theta)$ of a-$Si_3N_{3.5}$ for different PAW potentials. Both functions are calculated by averaging over 6000 configurations at 3000 K. The soft potential used in our calculations (150 eV) gives within statistical uncertainties the same results as the harder reference (300 eV).

Meloni also used a "Tersoff-like" potential and observed a N-N peak in the PCF at distance of 1.3 Å [57]. Since this peak was not observed in experiments [28, 29], we suppose that they had similar problems as our cooperation partners.

The present simulations were all performed non spin-polarized with an even number of electrons. This is a compromise, but one that we believe corresponds well with the experimental situation. In a macroscopic sample, the Fermi level is pinned system wide. All orbitals well below the Fermi level will be doubly occupied, whereas all well above the Fermi level will be unoccupied. This corresponds to the situation we have simulated in our work. Singly occupied state are very rare in our simulations so that the neglect of spin polarization is reasonably well founded. In fact, in most of the final models, defects came in pairs with one defect level well below the Fermi level being doubly charged and the other one well above the Fermi level being unoccupied. In a macroscopic sample, a tiny fraction of the defects close to the macroscopic Fermi level will be ultimately singly occupied, and, according to our results, these defects are rare. The study of those important defects will be left for future work.

5 Results

"For the presentation of our results, we closely follow the content of our previous publications about bulk amorphous silicon nitrides [138, 139]".

5.1 Structural properties of the a-Si_3N_x:H bulk

Statements on structural properties such as bonding, ordering and density fluctuation are typically be derived from the pair correlation function (PCF), the angular distribution function (ADF) and the structure factor (STF). Hence we calculated all three functions for our a-Si_3N_x:H model structures to judge their quality and to examine the influence of stoichiometry, system size and thermodynamic treatment.

5.1.1 Pair correlation function (PCF) of bulk a-Si_3N_x:H

Fig. 34 shows the PCFs of the small and large a-Si_3N_x model structures with and without H. Corresponding mean bonding distances $\bar{d}_{ij}$, standard deviations σ_{ij}, and first and second minima $r_{\min}$ are given in Table 2. Furthermore, the partial coordination numbers N_{ij} are summarized in Tab. 3. Our first observation is that the small instantaneously quenched model structures (black solid lines) correspond very well with the large stepwise cooled reference structures (black broken lines). Only the nearest neighbor peaks of the reference samples are slightly narrower and higher. This is most likely caused by the stepwise cooling of the reference samples. However, the peak positions and the long range behaviour are identical and we see an overall excellent agreement between the small and the large systems at the level of the PCF.

For stoichiometric a-Si_3N_4, Si and N form a nearly perfect network of amorphous silicon nitride. As their partial coordination numbers are $N_{SiN} = 3.95$ and $N_{NSi} = 3.02$, see Tab 3, nearly all Si atoms are fourfold coordinated, and nearly all N atoms are threefold coordinated. The corresponding nearest neighbor distance between Si and N is $d_{SiN} = 1.75$ Å and the next nearest neighbor distances are $d_{SiSi} = 2.99$ Å and $d_{NN} = 2.89$ Å, see Tab 2. Additionally, we can see a small shoulder at about 2.5 Å due to non bonding Si atoms in Si-N-Si arrangements. For a-Si_3N_{4-x}, the loss of N forces Si to replace N bonds by weaker and longer Si bonds. Although Si still remains fourfold coordinated ($N_{SiN} + N_{SiSi} \approx 4$), the percentage of its Si neighbors increases from 0 %, in a-Si_3N_4, to about 10 %, in a-$Si_3N_{3.5}$, and upto about 25 %, in a-Si_3N_3. This increase is visible by the change of corresponding partial coordination numbers N_{SiN} and N_{SiSi}, and by the emerging additional nearest neighbor peak in the Si-Si PCF at 2.4 Å. In contrast to Si, N remains perfectly threefold coordinated to Si. For a-$Si_3N_{4.5}$, the case is similar to a-$Si_3N_{3.5}$,

though with exchanged roles of Si and N. N remains threefold coordinated ($N_{\mathrm{NSi}}+N_{\mathrm{NN}} \approx 3$), but the percentage of its N neighbors increases to about 10 %. The N-N PCF shows an additional nearest neighbor peak at 1.62 Å. Additional H in a-Si_3N_{4-x} can bond to Si, N or H, with its mean bonding distances of 1.50 ± 0.07 Å, 1.11 ± 0.04 Å and 0.77 ± 0.01 Å, respectively. By comparing the partial coordination numbers N_{HSi}, N_{HN}, and N_{HH} for different compositions and H concentrations, we observe that most H atoms bond to Si (between 47 % and 70 %), some to N (between 24 % and 48 %), and few to H (between 4 % and 13 %). One might assume that additional H would cure the samples from the substitutional Si-Si bonds, but the partial coordination numbers and the PCFs do not support this idea. While N_{SiN}, N_{SiSi}, and N_{NSi} are mostly somewhat reduced, the PCFs are hardly affected by H, compare Fig. 34 (red crosses). This finding suggests that we find Si-Si bonds even in hydrogenated a-Si_3N_{4-x}.

Present findings agree very well with previous experiments and *ab initio* simulations. For the stoichiometric samples, we can compare our mean atomic distances $\bar{d}_{\mathrm{SiSi}} = 3.00$ Å, $\bar{d}_{\mathrm{NN}} = 2.90$ Å, and $\bar{d}_{\mathrm{SiN}} = 1.75$ Å with reference values from experiments and other *ab initio* calculations (Tab. 4). The agreement of those values lies well within the accuracy of the measurement. Likewise, the emerging Si-Si peak between 2.40 Å and 2.50 Å can be validated in literature. For example, the small shoulder of the Si-Si PCF for stoichiometric samples, which is caused by edge sharing Si atoms, is also reported by Giacomazzi *et al.* and Jarolimek *et al.*, and the increasing peak intensity for under stoichiometric samples, due to substitutional Si-Si bonds, is also visible in the publications from Alvarez *et al.* and Ippolito *et al.* [53–55, 57]. Simulations on over stoichiometric samples are rare so that we can only compare our results with force field models obtained by Ippolito *et al.* [57]. While in our simulations the N-N peak develops at 1.62 Å only in over stoichiometric a-$Si_3N_{4.5}$ and shortens to 1.45 Å after relaxation with harder potentials, the N-N peak of the force field model lies at 1.3 Å, even in slightly under stoichiometric a-Si_3N_{4-x}. Since neither experiments nor other *ab initio* models show N-N bonds in a-Si_3N_4 samples, we think that this discrepancy is an artifact caused by the limitations of force field model.

5.1.2 Angular distribution function (ADF) of bulk a-Si_3N_x:H

Fig. 35 shows the ADFs of the small rapidly quenched and the large stepwise cooled a-Si_3N_x model structures with and without H. Despite more pronounced peak intensities for the small model structures, the general agreement of the ADFs is good. The increased intensities are most probable artificial finite size effects caused by the applied periodic boundary conditions. Nevertheless, the errors are small and have negligible effects on other properties such as the mean position of the peak. The peak positions are located at the mean angles of about 110°, for $\measuredangle_{\mathrm{NSiN}}$,

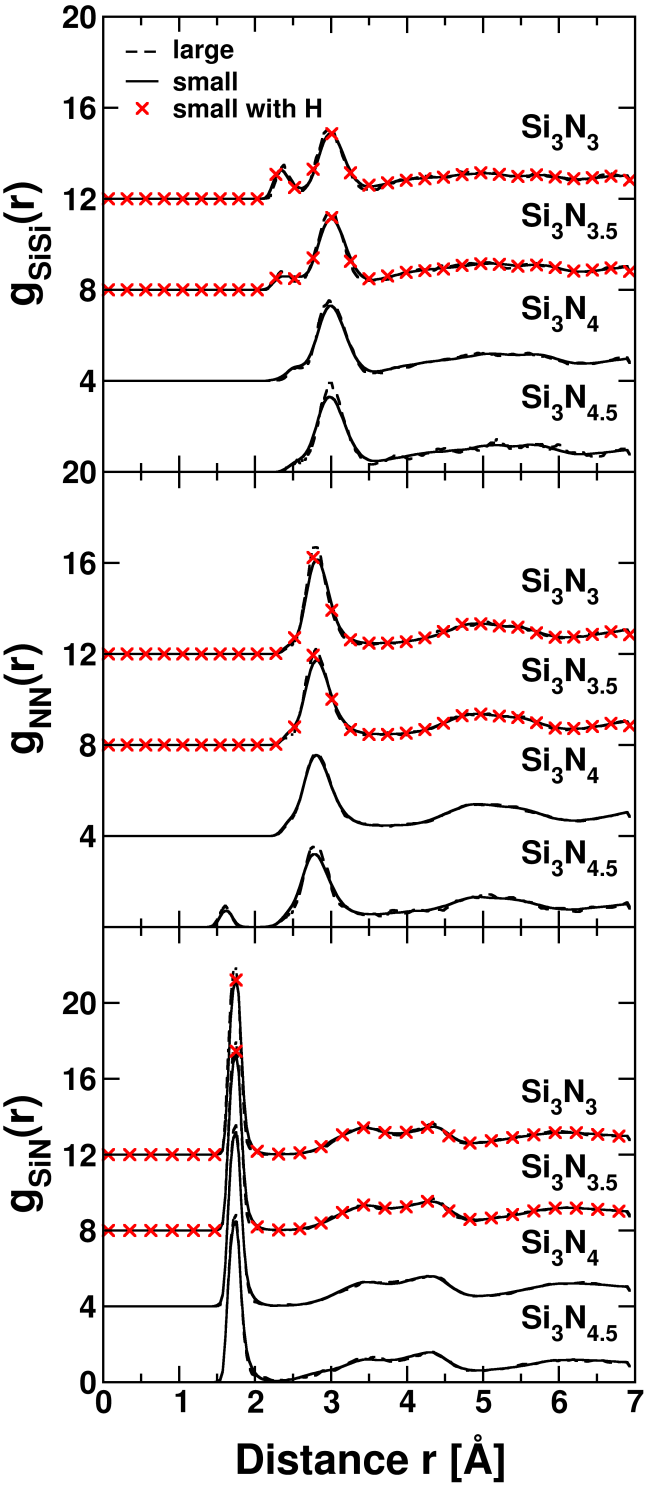


Figure 34: The PCFs of the small rapidly quenched and the large stepwise cooled a-Si_3N_x:H bulk samples. Additional Si-Si and N-N nearest neighbor peaks develop for Si rich and N rich compositions, respectively. The function is nearly independent of the considered system sizes and of the H concentrations.

Table 2: The mean bonding distances, the standard deviations, and the positions of the minima evaluated for the PCFs of the small rapidly quenched a-Si_3N_x bulk samples with and without H. All values are in Å. In a-$Si_3N_{4.5}$, a-$Si_3N_{3.5}$, and a-Si_3N_3 additional peaks emerge due to Si-Si and N-N neighbor atoms at 2.4 Å and 1.6 Å, respectively.

Si-Si	$Si_3N_{4.5}$	Si_3N_4	$Si_3N_{3.5}$	Si_3N_3
1st $\bar{d}_{SiSi}$	-	-	2.40	2.38
1st σ_{SiSi}	-	-	0.10	0.10
1st r_{min}	-	-	2.57	2.59
2nd $\bar{d}_{SiSi}$	3.00	2.99	3.02	3.02
2nd σ_{SiSi}	0.22	0.24	0.19	0.19
2nd r_{min}	3.56	3.56	3.53	3.49
N-N	$Si_3N_{4.5}$	Si_3N_4	$Si_3N_{3.5}$	Si_3N_3
1st $\bar{d}_{NN}$	1.62	-	-	-
1st σ_{NN}	0.06	-	-	-
1st r_{min}	1.98	-	-	-
2nd $\bar{d}_{NN}$	2.86	2.89	2.90	2.89
2nd σ_{NN}	0.26	0.26	0.25	0.23
2nd r_{min}	3.54	3.62	3.63	3.58
Si-N	$Si_3N_{4.5}$	Si_3N_4	$Si_3N_{3.5}$	Si_3N_3
1st $\bar{d}_{SiN}$	1.75	1.75	1.75	1.75
1st σ_{SiN}	0.09	0.08	0.07	0.07
1st r_{min}	2.26	2.30	2.30	2.30

Table 3: The partial coordination numbers N_{ij} evaluated for the small rapidly quenched a-Si_3N_x bulk samples. For a particular atom type i, N_{ij} specifies the number of nearest neighbors of type j.

	$Si_3N_{4.5}$	Si_3N_4	$Si_3N_{3.5}$	Si_3N_3
N_{SiN}	4.16	3.95	3.57	2.99
N_{SiSi}	-	-	0.38	1.01
N_{NSi}	2.77	3.02	3.00	2.99
N_{NN}	0.30	-	-	-
	$Si_3N_{3.5}H_{0.4}$	$Si_3N_{3.5}H_{0.8}$	$Si_3N_{3.5}H_{1.7}$	$Si_3N_3H_{0.8}$
N_{SiN}	3.51	3.47	3.40	2.97
N_{SiSi}	0.48	0.35	0.33	0.93
N_{NSi}	2.94	2.91	2.85	2.97
N_{NN}	-	-	-	-
N_{HSi}	0.51	0.51	0.47	0.70
N_{HN}	0.48	0.44	0.38	0.24
N_{HH}	0.04	0.06	0.13	0.06

and about 120°, for $\measuredangle_{\mathrm{SiNSi}}$, regardless of stoichiometry. The bonding angles perfectly reflect the sp^3 hybridization of Si, with an ideal tetragonal angle of 109.47°, and the sp^2 hybridization of N, with a planar trigonal angle of 120°. In both ADFs, we also find additional peaks at 90° with intensities decreasing for under and over stoichiometric samples. As illustrated in Fig. 35, these peaks are evidence for Si-N-Si-N square configurations which most frequently occur in stoichiometric Si_3N_4. As with the PCF, additional H has nearly no effect on the overall properties of the ADFs. Tab. 4 shows reference values for the main bond angles in a-Si_3N_4, which are all in good agreement with our results. Additionally, several authors, such as Giacomazzi *et al.* and Vedula *et al.*, also observe peak contributions at 90° caused by the edge sharing tetrahedra in the network of silicon nitride [54, 58]. That is also in agreement with our findings.

Table 4: Reference values for bond lengths and bond angles in a-Si_3N_4 obtained by experiments and simulations. All distances are in Å.

Ref.	1^{st} $\bar{d}_{\mathrm{SiSi}}$	2^{nd} $\bar{d}_{\mathrm{SiSi}}$	2^{nd} $\bar{d}_{\mathrm{NN}}$	1^{nd} $\bar{d}_{\mathrm{SiN}}$	$\measuredangle_{\mathrm{NSiN}}$	$\measuredangle_{\mathrm{SiNSi}}$
Aiyama *et al.* [28]	–	3.00	3.00	1.75	–	–
Misawa *et al.* [29]	–	3.01	2.83	1.73	109.8°	121.0°
Giacomazzi *et al.* [54]	2.42	3.03	2.76	1.73	109.1°	117.2°
Jarolimek *et al.* [150]	2.35-2.41	3.10	2.90	1.76	–	–
Hintzsche *et al.* [138]	–	2.99	2.89	1.75	108.6°	119.6°

5.1.3 Structure factor (STF) of bulk a-Si_3N_x:H

Fig. 36 shows the partial STFs of the large stepwise cooled a-Si_3N_x structures with and without H. Since the larger supercells have a finer grid in reciprocal space and the STF is thus resolvable even for small q values, we concentrated on the evaluation of the larger cells. Generally, the STFs are nearly independent of stoichiometry and H concentration. The main peak of all STFs lie at about 2.2 $\mathrm{Å}^{-1}$, which corresponds to a wavelength of 2.86 Å. Since in real space the second nearest neighbor distances of Si-Si and N-N are between 2.86 Å and 3.02 Å, the position of the main peaks most probable comes from the Fourier-transformed PCFs. At the same location we also see a strong anti-correlation in the Si-N STF, which might be caused by alternating Si-N-Si-N planes. The distances between Si-Si and N-N planes are roughly 2.8 Å. The general behaviour of the STFs is quite similar, although we can observe an slight indication of ordering.

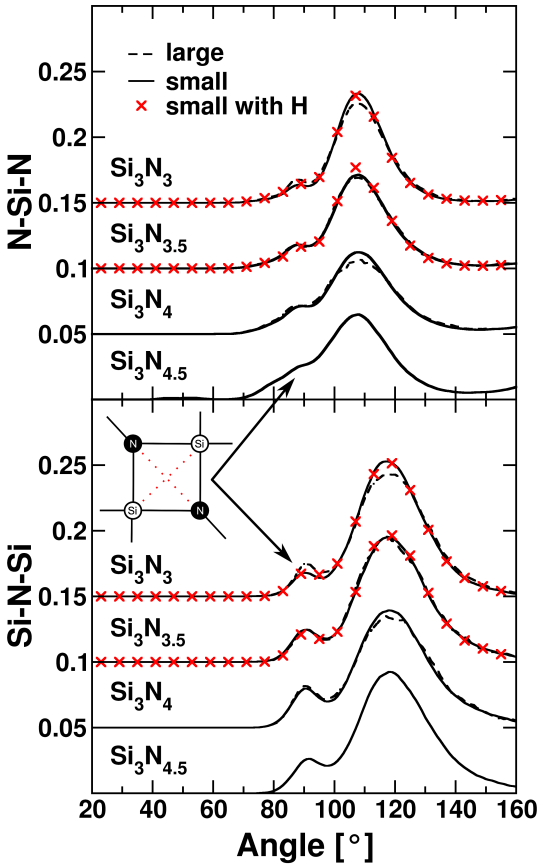


Figure 35: The ADFs of the small rapidly quenched and the large stepwise cooled a-Si_3N_x:H bulk samples with and without H. The small shoulder at about 90° depends on the number of square structures in a-Si_3N_x:H and has its highest intensity in a-Si_3N_4. The H concentration and the system size have only minor effects on the properties of the function.

In particular, for small wave vectors q, the values of S_{NN} rise in $Si_3N_{3.5}$ and Si_3N_3. The onset is best explained by N density fluctuations due to an increasing number of interconnected Si-Si bonds. The system thus ends up with nitrogen free Si networks that are visible as nitrogen density fluctuations. This idea is also supported by previous work on silicon nitride. For example, Robertson *et al.* reported that Si clusters start to interconnect at a stoichiometry of $Si_3N_{x=3}$ [151]. This percolation threshold corresponds with our observations.

5.1.4 Phase separation in Si rich a-Si_3N_{4-x} samples

The PCFs as well as the STFs in silicon nitride indicate an increasing number of Si-Si bonds in silicon rich a-Si_3N_{4-x}. In addition, manual inspection of a-Si_3N_3 samples shows spaces which are predominantly occupied by Si and by whole networks of Si-Si bonds. Since we do not believe that this observation is caused "randomly", we assume that this effect is related to phase separation. To study this effect in more details, we compare the local bonding topology for a-Si_3N_{4-x} without H. For Si rich samples without H, this approach utilizes the following conditions: (i) Si can bond to Si or N and (ii) all N atoms bond exclusively to Si. As one can see in Tab. 3, these conditions are approximately fulfilled.

For Si_3N_{4-x}, we now approximate the partial coordination number N_{SiSi} by x. If the formation of Si-Si bonds had been random, Si atoms with n N and $(4-n)$ Si neighbors would have occur with the probability of a binomial distribution.

$$p(n) = \frac{4!}{n!(4-n)!}\left(\frac{x}{4}\right)^{(4-n)}\left(\frac{4-x}{4}\right)^{n}. \tag{202}$$

However, in Fig. 37 we see that the actual distribution of $Si_{4-n}N_n$ units does not follow this expectation. Instead, we observe significantly more Si-Si_3N and Si-Si_4 units in a-$Si_3N_{3.5}$ and a-Si_3N_3, respectively. This observation supports the assumption that, in a-Si_3N_{4-x}, phase separation sets in for x between 0.5 and 1.0. Since Robertson also reported on a percolation threshold of Si-Si bonds that is slightly under $x = 1.0$ [36], our findings are in reasonable agreement with literature.

To provide some additional evidence for the phase separation in silicon nitrides with very low N concentration, we equilibrated a single a-$Si_3N_{1.5}$ structure for 10 ps at 2000 K, compare Fig. 38. As shown in Fig. 39, we subdivided the final model structure along the x-axis into four segments (S1–S4), and we calculated the diffusion constant and the number of Si and N atoms for each segment. The figure shows that the system indeed separates into three Si rich segments (S1,S2, and S4) and one N rich segment (S3). Thereby the N rich segment contains over 51 % of all N atoms. The corresponding variation in the a-Si_3N_x:H composition is from $x = 0.58$ to

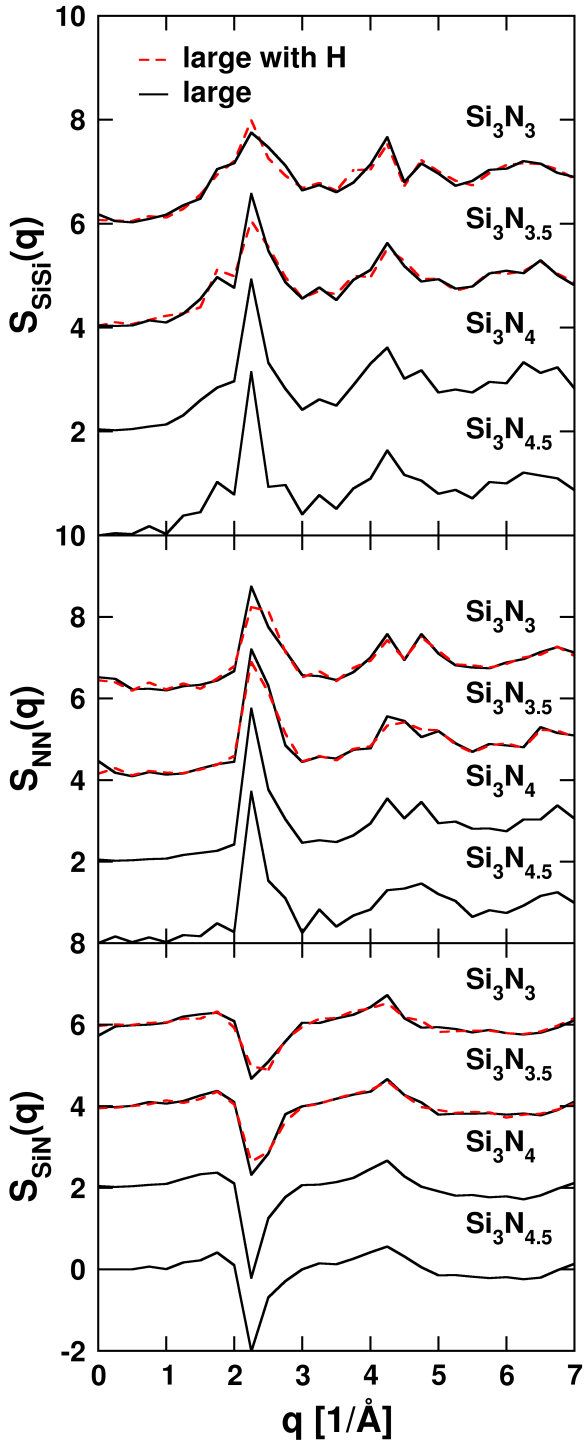


Figure 36: The STFs of the large stepwise cooled a-Si_3N_x bulk samples with and without H. We see sharp correlation peaks at 2.2 $Å^{-1}$ that indicate periodic structures at a corresponding distance of 2.86 Å. In a-Si_3N_3 and a-$Si_3N_{3.5}$, S_{NN} increases slightly for $q \to 0$. This change in long range order might be related to an increasing number of Si clusters with N deficiency.

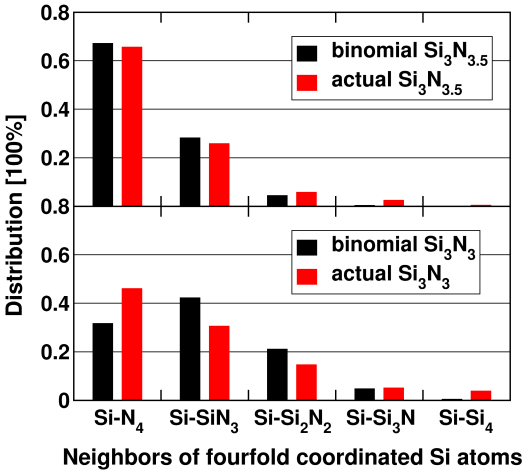


Figure 37: A detailed analysis of the bonding topology of fourfold coordinated Si atoms. The actual probability of finding a Si atom coordinated to n N and $(4-n)$ Si atoms is shown. Also shown is the distribution that we would have expected if the formation of a Si-Si bond had been entirely random. However, excess probabilities are found towards the two end points, in particular for a-Si_3N_3. This observation suggests a tendency towards Si clustering and "phase separation".

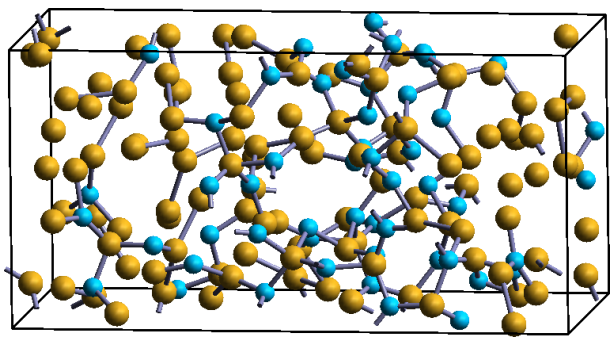


Figure 38: A a-$Si_3N_{1.5}$ model structure after 10 ps equilibration at 2000 K.

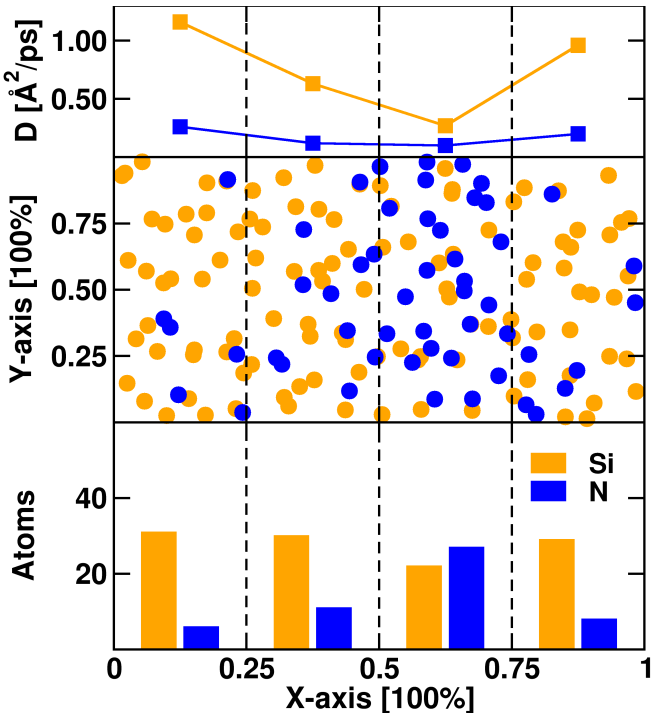


Figure 39: We subdivided the simulation cell of the a-$Si_3N_{1.5}$ model structure shown in Fig. 38 into four segments. For each segment (S1–S4) and for each atom type (Si and N), the diffusion constant (upper panel), the projected atom positions (middle panel), and the number densities (lower panel) are shown. The properties for Si are orange and the properties for N are blue.

$x = 3.68$. Especially noticeable is that the N rich segment also exhibits the lowest diffusion of all segments.

5.2 The electronic properties of the a-Si_3N_x:H bulk

In this section, we investigate general features of the density of states (DOS) of bulk a-Si_3N_x:H. We discuss the differences between the electronic properties of amorphous and crystalline materials, and we examine the characteristic changes of the DOS due to different atomic compositions. The results are presented on a large energy scale, whereas a more detailed examination of the electronic properties close to the Fermi level is given in section 5.3.3.

5.2.1 Density of states (DOS)

The most general features of the DOS of a-SiN:H can be derived directly from the bonding properties of Si and N. In the following, we shall refer to a perfectly bonded Si_3N_4 network when it consists exclusively of fourfold coordinated Si atoms, threefold coordinated N atoms and Si-N bonds. As discussed by Robertson *et al.*, Si and N atoms ideally bond by tetragonal sp3 and planar trigonal sp2 hybrid orbitals, respectively [151]. Electronically, the hybrid orbitals split up into occupied bonding σ and unoccupied anti-bonding σ^* states. Both form the main part of the valence band (VB) and the conduction band (CB), respectively. Additional occupied π states, which are caused by two lone pair N electrons, can be found at the valence band maximum (VBM). Note that we have shown in Tab. 3 of section 5.1.1 that most atoms in our a-Si_3N_x:H samples, which are in general close to the stoichiometric case $x = 4$, fulfill this condition.

In a-Si_3N_x:H, additional electronic states are naturally introduced by disorder. The network of Si-N bonds becomes for example disordered if bond lengths and bond angles are distorted; if Si-Si, N-N, Si-H, or N-H bonds form; or if coordination defects emerge. These effects are strongly intertwined and are known reasons for electronic states at the band edges and in the band gap. As a consequence, the VBM and the conduction band minimum (CBM) of a-Si_3N_x:H tail exponentially into the band gap while the band edges in crystalline silicon nitride are sharp [152].

Projecting the electronic states onto atom-centered spherical harmonics gives the site and orbital projected DOS. This projection can be used to distinguish contributions to the total DOS from different atom and orbital types. We have calculated and averaged the site and orbital projected DOS for over 1000 small rapidly quenched a-Si_3N_4 structures. The Si and N contributions to the total DOS with s or p orbital character are compared in Fig. 40. We observe that N contributions clearly dominate the whole VB between -21 and 0 eV. Between -21 and -15 eV, we mainly find N($2s$) orbital states and, between -12 and 0, there are mainly

N($2p$) orbital states. On the other hand, we observe Si contributions that slightly dominate the CB between 0 and 5 eV. There are Si($3s$) and Si($3p$) states in all mentioned energy intervals, but they only exceed the number of N states in the CB. In the VB, the main peak of the Si($3s$) states always lies slightly under the main peak of the Si($3p$) states. We suppose that N states dominate the VB because of the high electro negativity of N and its lone pair electrons. This assumption is also supported by experimental and theoretical evidence for an electron transfer from Si to N [153]. The calculated contributions to the total DOS correspond very well with literature [54], and they represent the general electronic features of all a-Si_3N_x:H compositions.

However, there are still some electronic changes for different a-SiN:H compositions. In Fig. 41, we have highlighted these changes by means of the averaged site projected DOS. The figure shows the Si and N contribution in a-Si_3N_{4-x} and a-Si_3N_{4+x} in comparison to Si_3N_4 (solid lines). Additionally, we show the contributions in a-Si_3N_{4-x}:H (red crosses) and in β-Si_3N_4 (broken line). The electronic states in the figure have been aligned with respect to the N core levels. We have chosen this strategy, since all N atoms remain approximately threefold coordinated in all considered structures, compare Tab. 3 in section 5.1.1. The main differences between the structures can be seen at the VBM and the CBM. Firstly, by comparing a-Si_3N_4 with β-Si_3N_4 (solid vs dashed lines), we observe the typical band tailing in amorphous materials. Especially at the CBM, the tail states of Si and N decrease the band gap by about 2 eV. Secondly, there is an increased number of Si states at the VBM of a-$Si_3N_{3.5}$ and a-Si_3N_3. These states might be related to the increased number of Si-Si bonds in Si rich samples, compare Fig. 34 in section 5.1.1. Analogously, we find additional N states at the CBM of a-$Si_3N_{4.5}$. Finally, H does not change any of these results. This behavior is unexpected, since it is experimentally known that additional H effectively increases the band gap by defect passivation. This issue shall be resolved in in section 5.3.3, where we inspect the electronic structure close to the gap in much more detail.

5.2.2 Inverse participation ratio (IPR)

In this section, we investigate the degree of localization of electronic states in the vicinity of the band gap by means of the IPR. The IPR is defined as

$$p^{-1}_{n\mathbf{k}} = \frac{N \sum_{\alpha,L} \left| c_{L,n\mathbf{k}}(\mathbf{R}_\alpha) \right|^4}{\left(\sum_{\alpha,L} \left| c_{L,n\mathbf{k}}(\mathbf{R}_\alpha) \right|^2 \right)^2}, \tag{203}$$

where $c_{L,n\mathbf{k}}(\mathbf{R}_j) = \langle Y_L(\mathbf{R}_\alpha) | \psi_{n\mathbf{k}} \rangle$ is the projection of the electronic state $n\mathbf{k}$ onto spherical harmonics Y_L centered at $\mathbf{R}_\alpha$. The sum is performed over all atoms α and all angular momenta

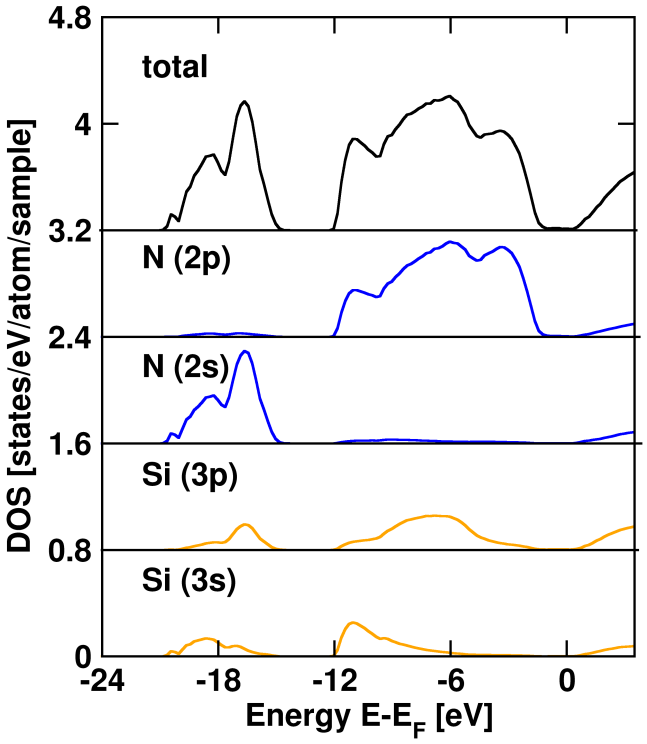


Figure 40: The averaged site and orbital projected DOS of the small rapidly quenched a-$Si_3N_{3.9}$ bulk samples. The comparison with the total DOS shows that the valence band (VB) is clearly dominated by N(2p) states and the conduction band (CB) is slightly dominated by Si(3p) states. These are the general electronic features of silicon nitrides that apply to under and over stoichiometric structures as well.

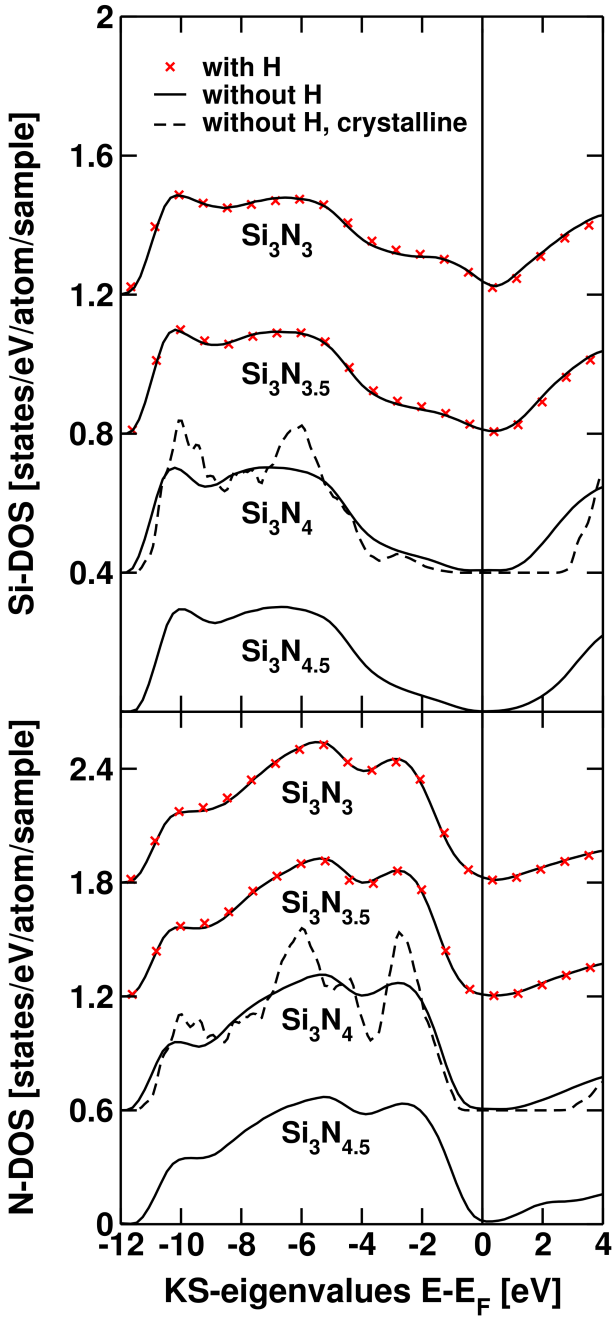


Figure 41: The averaged Si and N projected DOS of the small rapidly quenched a-Si_3N_x:H bulk samples with and without H (red crosses and solid red lines). Also shown is the site projected DOS of β-Si_3N_4 (dashed black lines). The crystalline structure exhibits the largest band gap and no defects. In comparison, we see close to the Fermi level an increased number of Si states at the VBM of a-Si_3N_{4-x}:H. Likewise, the number of N states increases at the CBM of a-Si_3N_{4+x}.

L, whereas N is the total number of atoms. Depending on the local structure, electronic states vary between being completely localized on particular atoms and being completely extended over all atoms. For example, states in the VB or in the CB are usually extended over all atoms, with IPR values equal to 1, while defect states in the band gap rather localize on particular atoms, with IPR values approaching N.

The comparison of the site projected DOS with the IPR directly shows the relationship between electronic states and their localization. We thus examined the evolution of both properties in the large stepwise cooled a-Si_3N_x samples under the influence of a decreasing N concentration. Specifically, Fig. 42 shows the Si and N site projected DOS of a large a-Si_3N_4, a-$Si_3N_{3.5}$, and a-Si_3N_3 sample without H. In addition, it shows the corresponding values of the IPR. The energy levels have been calculated using the HSE functional, which has led to increased band gaps in comparison to plain Kohn-Sham density functional theory (KS-DFT) calculations. We observe that the electronic states are strongly localized at the VBM of stoichiometric a-Si_3N_4 and mainly consist of N states. The band gap is quite distinct with no defect states within. However, going towards a-$Si_3N_{3.5}$ and a-Si_3N_3, the number of the Si states increases at both band edges and the states become progressively delocalized at the VBM. In contrast to the stoichiometric case, the band gap is reduced and filled with localized and partially localized defect states. Robertson *et al.* explained the growth of extended Si states in under stoichiometric samples with the increasing number of Si bonds [151]. As our structural analysis has already suggested, Si bonds firstly form Si cluster and then interconnect to larger Si networks. For this reason, we can support Robertson's expectation that the Si states are partially localized on Si clusters until the clusters percolate. Due to this interconnection, the Si states can spread over the entire network of Si bonds. The present examination of the IPR supports this explanation as well.

As partially localized states can be considered as an indication for tail states, the IPR can also be used to roughly estimate the size of the band gap, as suggested by Justo *et al.* [51]. Thereby, a threshold value s is defined and the band gap is given by the energy difference between the first state of the VBM with an energy value $\epsilon_i > s$ and the last state of the CBM with an energy value $\epsilon_i < s$. Since the IPR values in Fig. 42 show a considerable spread, we have some freedom for the choice of the threshold value s. In our case, we have chosen $s = 4.26$ because this value gives a minimal deviation from the experimental reference values [27]. Using this threshold value, the measured band gap values are 5.71 eV, for a-Si_3N_4, 5.10 eV, for a-$Si_3N_{3.5}$, and 3.21 eV, for a-Si_3N_3. Although these values give a good agreement with optical band gap measurements [27], we have to consider that this method is ambiguous and somewhat vague.

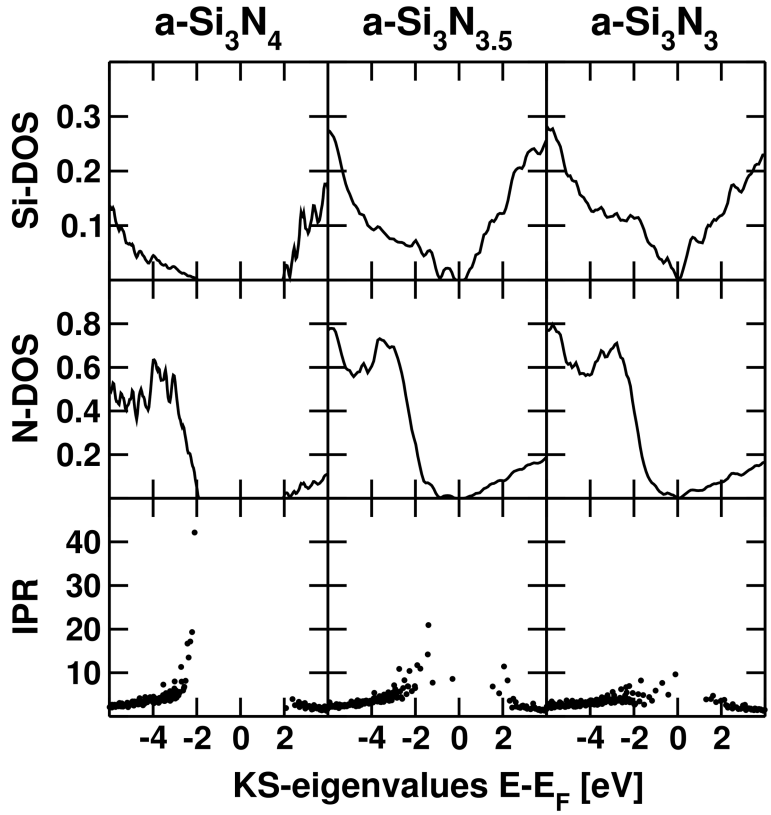


Figure 42: The Si and N site projected DOS near the band gap of the large stepwise cooled a-Si_3N_{4-x} bulk samples. For each state, the degree of localization is shown by means of the IPR. For decreasing N concentrations, the number of Si states in the band gap increases and the Si states at the valence band maximum (VBM) delocalize. The rise of delocalized Si states is most likely related to the growth of Si-Si bonds in Si rich samples.

To avoid large inaccuracies in the band gap measurement, we also followed a more sophisticated approach in the next section.

5.2.3 Optical band gap

In the previous section, the investigation of the IPR has shown that VB and CB states are typically extended, whereas tail states or defect states are rather localized. However, we have not yet derived reliable band gap values from the IPR. Experimentally, the optical band gap of amorphous semiconductors can be measured using the so-called Tauc method. In this section, we first yield a short overview of this method, and, afterwards, we describe how a similar approach can be applied for model structures generated by computer simulations.

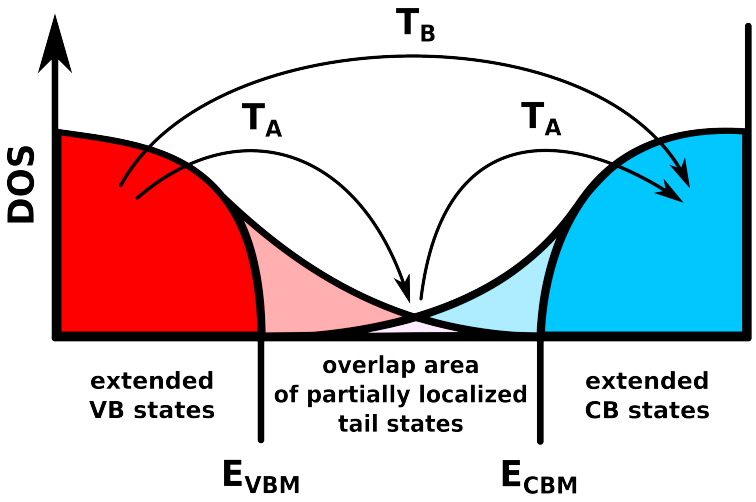


Figure 43: The absorption of light can cause two different types of electronic transitions: (a) between extended states and partially localized states T_A and (b) between extended band states only T_B. The onset of transition type T_B is characteristic for the size of the optical band gap $E_g = E_{CBM} - E_{VBM}$.

Tauc *et al.* have introduced a model that links the properties of the absorption spectra to electronic transitions [154–156]. As transitions between localized states are very unlikely, we can expect that transitions occur predominantly in combination with extended states. For example, between localized states and extended band states or between extended band states only. Both transition types are illustrated in Fig. 43, and we have named them T_A and T_B, respectively. In addition, we suppose that transitions of type T_A need much less activation energy than transitions of type T_B.

For this reason, Tauc *et al.* have divided the onset of light absorption in amorphous materials into three regimes: A low and medium energy regime due to transitions of type T_A and a high energy regime due to transitions of type T_B [155,156]. While the first regime is characteristic for the onset of absorption, the second regime is typically extended until the absorption coefficient $\alpha(E)$ reaches the threshold value of $\alpha(E) = 10^4$ cm^{-1}.

The first two regimes are often treated together because of their similar exponential form [155]

$$\alpha(E) \propto e^{E/E_U}. \tag{204}$$

The second regime is most often called Urbach tail, named after Franz Urbach who has found a similar regime for crystalline structures [155, 157]. For this reason, E_U is called Urbach energy if the exponential function refers to the second absorption regime. Note that the first and the second regime are sometimes fitted as superposition of two exponential functions [156], or, if

they are indistinguishable, they are simply treated as only one regime [158, 159].

The third energy regime is typically found for $\alpha(E) > 10^4 \mathrm{cm}^{-1}$, and it obeys in contrast to the first two regimes a power law [155]

$$\alpha(E) \propto \frac{(E - E_\mathrm{T})^2}{E}, \tag{205}$$

where E_T is the so-called Tauc band gap. The change from the exponential law to the power law is then supposed to be caused by the onset of transitions of type $\mathrm{T_B}$.

In our work, the absorption spectrum $\alpha(E)$ is derived from the frequency dependent dielectric matrix in the independent particle approximation. By neglecting excitonic effect and local-field corrections, the imaginary part of the dielectric matrix is calculated using [160]

$$\mathrm{Im}\,\epsilon_{ij}(E) = \frac{4\pi e^2}{\Omega} \lim_{q\to 0} \frac{1}{q^2} \sum_{c,v,\mathbf{k}} 2\, w_\mathbf{k} \delta(\epsilon_{c\mathbf{k}} - \epsilon_{v\mathbf{k}} - E) \langle u_{c,\mathbf{k}+q\mathbf{e}_i} | u_{v,\mathbf{k}} \rangle \langle u_{v,\mathbf{k}} | u_{c,\mathbf{k}+q\mathbf{e}_j} . \rangle. \tag{206}$$

Here c and v refer to the summation over the CB and the VB states, respectively. The summation runs over the 1st BZ, and it is weighted by $w_\mathbf{k}$. For each k-point, $u_{v\mathbf{k}}$ and $u_{c\mathbf{k}}$ are the cell periodic parts of the Bloch waves $\psi_{n\mathbf{k}} = u_{n\mathbf{k}} e^{i\mathbf{kr}}$. The infinitesimal small external perturbation of the system enters by the Cartesian direction vectors $q\mathbf{e}_i$ for $q \to 0$. The real part of the dielectric matrix is then calculated using the Kramer-Kronig relation [160]

$$\mathrm{Re}\,\epsilon_{ij}(E) = 1 + \frac{2}{\pi} P \int_0^\infty d\omega' \, \frac{\mathrm{Im}\,\epsilon_{ij}(\omega')\omega'}{\omega'^2 - E^2 + i\eta}. \tag{207}$$

Finally, the aborption spectrum is evaluated by [161]

$$\alpha(E) = \frac{1}{3} \sum_{i=1}^{3} \frac{2E}{c} \sqrt{\frac{|\epsilon_{ii}(E)| - \mathrm{Re}\,\epsilon_{ii}(E)}{2}}. \tag{208}$$

As illustrated in Fig. 44, the change of the absorption regimes makes it possible to define the optical band gap in two ways [151, 155]. On the one hand, we can define the band gap E_{04} as the energy value when $\alpha(E)$ reaches $10^4 \mathrm{cm}^{-1}$. One the other hand, we can define the Tauc band gap E_T by means of the formula [151, 155]

$$\sqrt{\alpha E} = B(E - E_\mathrm{T}). \tag{209}$$

Here extrapolation of the linear regime of the function $\sqrt{\alpha E}$ gives a line with slope B that crosses the energy axis just at E_T. In Fig. 44, the absorption coefficient $\alpha(E)$ and the function $\sqrt{\alpha E}$ are averaged over several independent a-Si_3N_{4-x} sample structures. Specifically, we averaged over 40 small continuously cooled a-Si_3N_3 and a-$Si_3N_{3.5}$ structures and over 4 stepwise cooled a-Si_3N_4 structures.

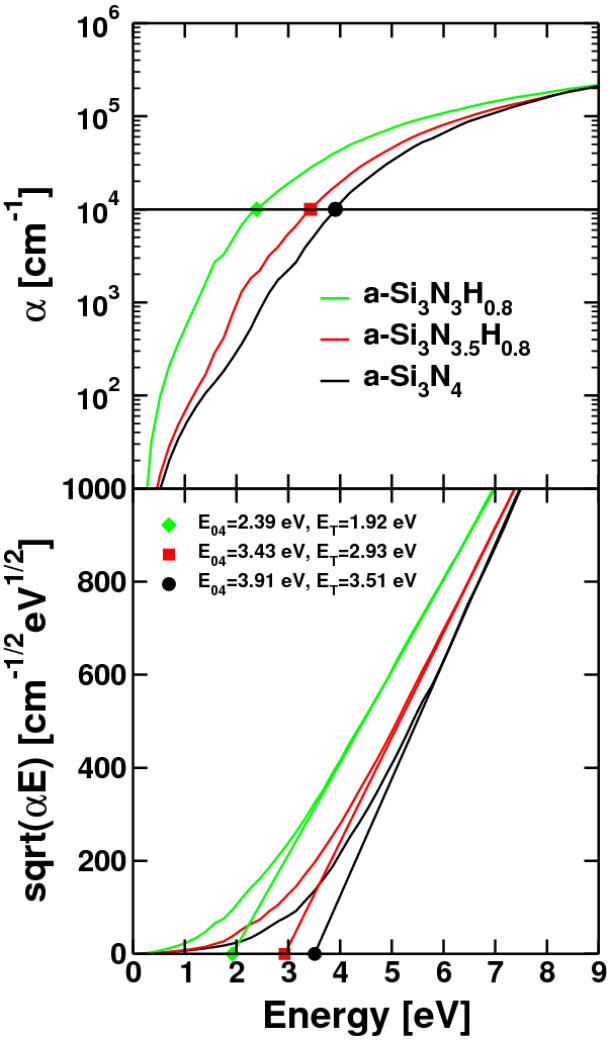


Figure 44: The two band gaps E_{04} and E_{T} from KS-DFT calculations are derived from the absorption coefficients $\alpha(E)$ of the a-Si_3N_{4-x} samples. E_{04} is defined as the energy when α reaches the threshold of 10^4 cm^{-1}. According to Tauc [154], E_{T} is defined by the intersection between the energy axis and the extrapolation of the linear regime of the function $\sqrt{\alpha E}$.

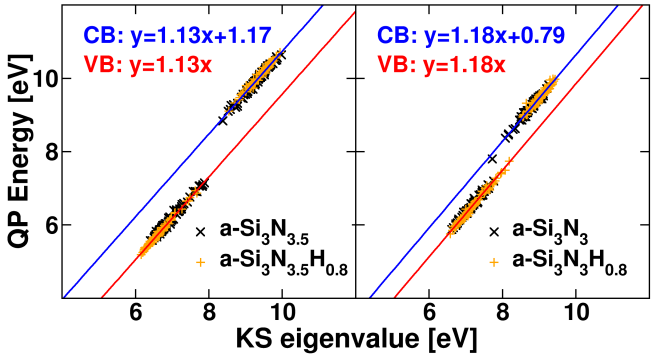


Figure 45: The KS eigenvalues versus the corresponding quasiparticle energies in a-Si_3N_3:H and a-$Si_3N_{3.5}$:H reference systems. The KS and GW0 band gaps are related by a simple up-shift of the conduction band (CB) states with respect to the valence band (VB) states. The offset values of all considered systems are summarized in Tab 5.

Table 5: The parameters obtained by fitting the linear function $y = kx + \Delta$ to the GW_0 quasiparticle energies as a function of the KS one-electron energies. The calculated scissor correction Δ is added to the optical KS band gap values determined in Fig. 44.

	$Si_3N_3H_{0\backslash 0.8}$	$Si_3N_{3.5}H_{0\backslash 0.8}$	Si_3N_4
k	1.18	1.13	1.12
Δ	0.79 eV	1.17 eV	1.19 eV

Since it is a known problem that KS-DFT calculations underestimate the band gap, we also apply GW_0 corrections to the KS gaps. The method is illustrated in Fig. 45, where the KS eigenvalues of a-Si_3N_x:H are plotted versus the corresponding quasiparticle energies. Since the electronic states lie nearly on two parallel lines, we have fitted two linear functions to the occupied VB states and to the unoccupied CB states, respectively. The GW_0 corrections are then given by a simple up-shift of the CB states. The fitting parameters for each considered reference system are summarized explicitly in Tab 5. Note that we use the same parameters for the systems with and without H. In the following we assume that the GW_0 corrections for the reference systems apply to all samples with the same stoichiometry.

Using this analysis scheme, we can compare the GW_0 corrected band gap values E_{04} and E_T with experimental results obtained by the Energy Research Center of the Netherlands and

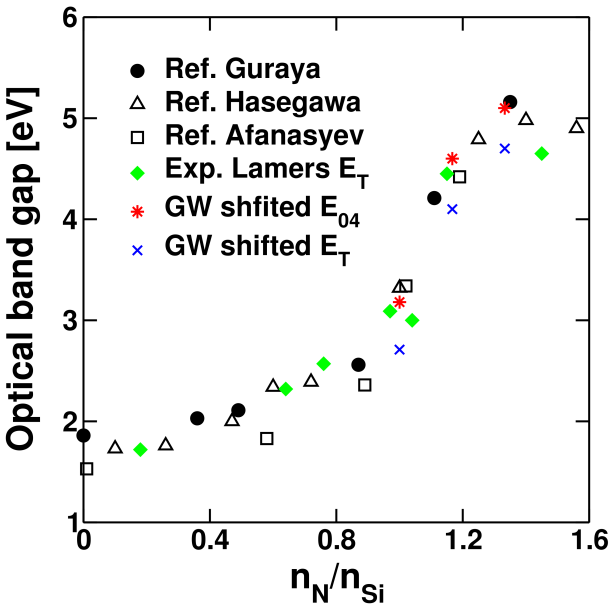


Figure 46: GW_0 corrected band gap values in comparison with experimental results obtained by the Energy Research Center of the Netherlands and other reference values [27, 162, 163].

others [27,162,163]. The comparison is shown in Fig. 46 and our values are summarized in Tab. 6. Typically, for one particular stoichiometry, the E_{04} gaps are between 0.4 and 0.5 eV larger than the Tauc gaps. As a consequence the values for E_{T} lie a little bit below the experimental results, while the values for E_{04} meet them perfectly. In spite of this difference, the relative change depending on stoichiometry obviously remains constant for both definitions.

5.3 The defect states in the a-Si_3N_x:H bulk

To examine electronic defect states in a-Si_3N_x:H, we inspect the electronic and the structural properties of the materials and link both properties with each other. Furthermore, since hundreds of different sample structures are inspected, we automatized this task. In this section, we first discuss the computer program that identifies the electronic defect states and the corresponding local structures. Afterwards, we show the most important defect classes in our bulk samples, and, finally, we perform a quantitative defect analysis in the vicinity of the band gap.

5.3.1 Analysis framework

We define electronic defects as localized states in the region of the band gap, which is given by the energy difference between the valence band maximum (VBM) and the conduction band

Table 6: The calculated optical band gaps obtained by using the threshold value of $10^4 \mathrm{cm}^{-1}$ for the absorption coefficient α (E_{04}) and by extrapolating the Tauc linear regime (E_T), compare Fig. 44. The values obtained by KS-DFT are furthermore compared before and after the GW_0 correction, compare Fig. 45.

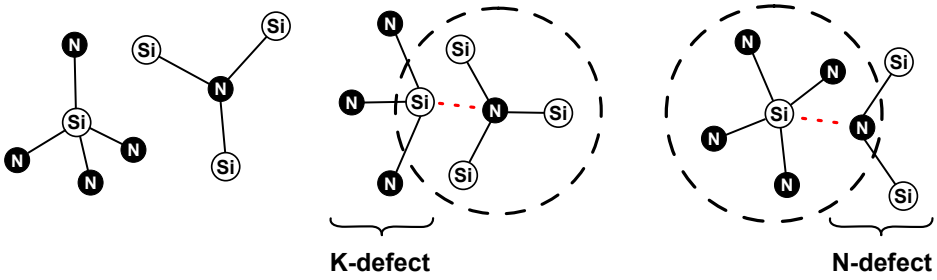
	E_{04}^{KS}	$E_{\mathrm{T}}^{\mathrm{KS}}$	$E_{04}^{\mathrm{GW}_0}$	$E_{\mathrm{T}}^{\mathrm{GW}_0}$
a-$Si_3N_3H_{0.8}$	2.4	1.9	3.2	2.7
a-$Si_3N_{3.5}H_{0.8}$	3.4	2.9	4.6	4.1
a-Si_3N_4	3.9	3.5	5.1	4.7

minimum (CBM). The VBM and the CBM are identified by a threshold value for the occupation numbers f_i, which can vary between 0.0 and 2.0. Usually, we identify the VBM with the last state having a occupation greater equal 1.0 and the CBM with the first state having a occupation less than 1.0. Although this distinction might seem arbitrary, it turns out that, in a system with only even number of electrons, the states at the band edges are rather doubly occupied or empty. Therefore, our threshold value leads to clear cut between the occupied and unoccupied states. The energetic location of the defects within the DOS is determined by their KS eigenvalues. As the electronic states in the vicinity of the band gap are most important for our work, we restrict our analysis to the last six occupied and first six unoccupied states. In contrast to delocalized states in the CB or in the VB, localized states can be attributed to specific atoms. These are identified by projecting the orbitals $\psi_{n\mathbf{k}}$ onto spherical harmonics $Y_L(\mathbf{R}_\alpha)$ and by summing over all angular momentum numbers L.

$$e_{n\mathbf{k}}(\mathbf{R}_\alpha) = \sum_L |\langle Y_L(\mathbf{R}_\alpha)|\psi_{n\mathbf{k}}\rangle|^2. \tag{210}$$

For a particular band number n and k-point $\mathbf{k}$, the largest value of $e_{n\mathbf{k}}(\mathbf{R}_\alpha)$ then reveals the atom position $\mathbf{R}_\alpha$ where the electronic state is predominantly localized. For our investigations, the four atoms where the defect localizes most strongly are stored in a list for a subsequent structural analysis.

The position of geometrical defects in each structural model are determined by first building a nearest neighbor list using a fairly large radial cutoff (i.e. larger than the first minima of the partial PCFs). In the second step, for over coordinated atoms, the longest bond is removed from the list until all atoms are either ideally coordinated or under coordinated. In this step, over coordinated atoms (e.g. fivefold coordinated Si atoms and fourfold coordinated N atoms)



(a) ideal bonding (b) K-defect or fourfold N (c) fivefold Si or N-defect

Figure 47: The common geometrical bonding topologies of Si and N atoms in Si_3N_{4-x}: (b) The indicated defect can be either counted as a K defect or a fourfold coordinated N atom. (c) The indicated defect can be either counted as a N defect or a fivefold coordinated Si atom. To remove these ambiguities, a nearest neighbor list is built applying large radial cutoffs. Then the bonds resulting in over coordinated atoms are removed. This does not modify the total number of defects.

are converted into neighboring under coordinated atoms, leaving the total defect density unmodified. This defect conversion is illustrated in Fig.47(b) and (c). Here a ideally coordinated Si atom (N atom) is excluded from the nearest neighbor list of a over coordinated N atom (Si atom) so that the over coordinated atom becomes ideally coordinated and the neighboring atom becomes under coordinated. Removing over coordinated atoms, also corresponds to the common experimental notion that the dominant defects are K defects and N defects, and, as shown in Fig. 47, it is anyway not unambiguously possible to distinguish between over coordinated and under coordinated geometrical defects.

At this point of our analysis, we know at which atoms the electronic states in the band gap are geometrically localized and the nearest neighbor environment (i.e. neighboring atoms, bond distances, bond angles, coordination numbers) of this atoms. However, under coordinated atoms often affect the electronic properties beyond their nearest neighbor environment, and electronic states are sometimes spread over several atoms. For this reason, we have further adjusted our analysis by investigating all four atoms where an electronic state is mostly localized. If at least one of the corresponding atoms is under coordinated, we consider the whole electronic state to be caused by the under coordinated atom where the state is most strongly localized. In contrast, if all four atoms are ideally coordinated, we consider the state to be caused by other reasons than under coordination. These alternatives are explicitly specified in the next section. Finally, we are able to assign every electronic state in the band gap to one particular atom site. Manual

inspection of the electronic and the structural properties confirms that, in this way, over 80 % of all electronic defect states can be correctly assigned to structural defects (coordination defects, distorted bonds). Although the present analysis scheme cannot replace detailed investigations by hand, it represents a consistent, effective method to investigate defects for large numbers of samples.

5.3.2 Defect classification

For selected prototype defects, the charge density and the local electronic DOS are shown in Fig. 48. For each example, the DOS is aligned with respect to the Fermi level which is chosen to lie mid gap.

In our simulations, the most common defect is the K-defect, which corresponds to a threefold coordinated Si atom [see Fig. 48 (a)]. The electronic defect state is a dangling sp^3 orbital pointing towards the missing N atom. The defect state can be either occupied (red line) or unoccupied (blue line), and the Kohn-Sham eigenvalues can be located in the gap, or gradually merge into the conduction or valence band tails, or even overlap with the valence or conduction band. Only mid gap defect states are well localized, whereas states that are close or in the valence or conduction band are less well localized. In average, only one out of two geometrical K-defects will result in an electronic defect state in the band gap, and therefore, a strict one-to-one correspondence between geometrical defects and electronic defect states in the gap does not exist.

The second class of defect states are related to Si-Si bonds or larger Si networks or Si chains. Compared to the Si-N bond, the Si-Si bond results in a much reduced bonding anti-bonding splitting. If the local tetrahedral Si environment is distorted or a Si-Si bond is particularly short or long, electronic states in the band gap can be introduced. Examples for such defect states are shown in Fig. 48 (b). For strong local distortions, sp^3 orbitals hybridize less efficiently with neighboring atoms and the charge distribution is sometimes reminiscent of K-defects. Furthermore, as the nitrogen content decreases, the number of homopolar Si-Si bonds grows, and larger Si percolation networks develop. [55, 138] Within such a Si network, some bonding or anti-bonding Si $3p$ linear combinations can move into the gap [Fig. 48 (b) left]. Using an automatic procedure, Si-Si states are difficult to distinguish from other states related to a distorted Si environment, as both defect classes tend to be intermingled. Visual inspection of the charge density, however, suggests that the majority (over 70 %) of the gap states that cannot be assigned to under-coordinated atoms are localized on Si-Si bonds. This observation is similar to a recent finding of Khomyakov *et al.* for amorphous silicon (a-Si). [164]

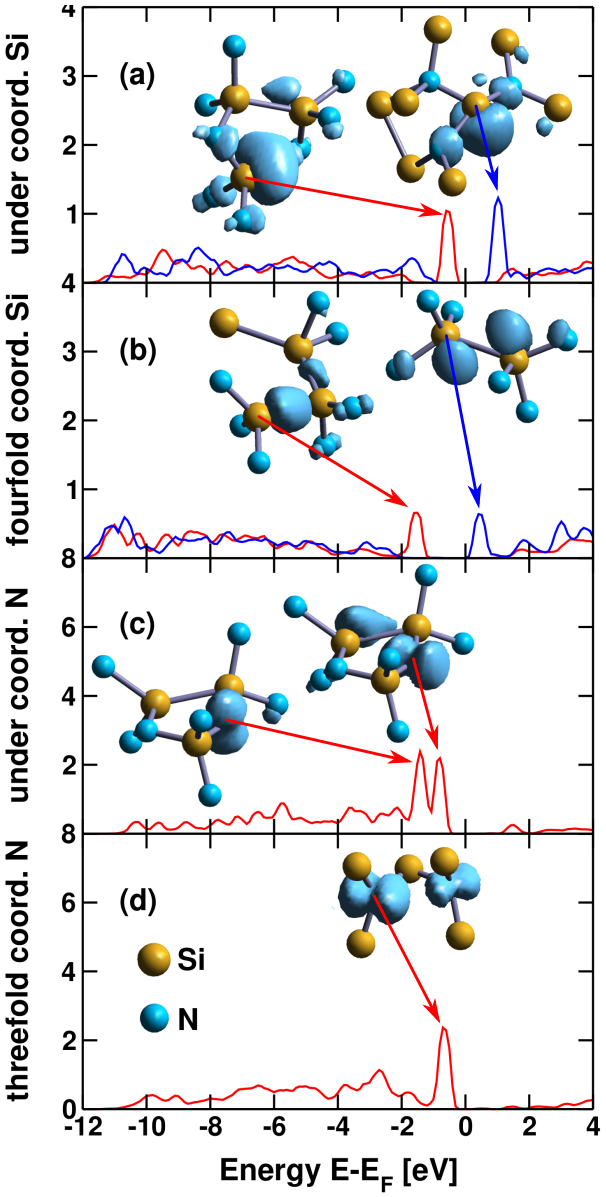


Figure 48: The local structure around exemplary geometrical defects and the electronic density of states (DOS) evaluated at the corresponding sites (origin of arrow): (a) Si dangling bond (K-defect), (b) bonding and anti-bonding states on fourfold coordinated Si, (c) N lone pair and dangling bond (N-defect), and (d) lone pair on threefold coordinated N. The charge density corresponding to the peak in the DOS is additionally shown. The contributions to the DOS can be either occupied (red line) or unoccupied (blue line).

N-defects, twofold coordinated N atoms, are rarely found and they are entirely absent in the slowly cooled nitrogen deficient samples: in the nitrogen-deficient case, a surplus of Si atoms exists, and Si atoms grab any existing N-dangling bond. Although N-defects are only present in small number in the rapidly quenched ensembles, we, nevertheless, show their electronic properties in Fig. 48 (c). N-defects are characterized by two localized doubly occupied states close to the valence band edge. Normally, nitrogen in Si_3N_{4-x} adopts a flat coordination with three Si neighbors forming sp^2 hybrid orbitals that develop bonding and anti-bonding linear combinations with the three neighboring Si atoms. The remaining two electrons form a lone non-bonding pair perpendicular to the NSi_3 plane. This state corresponds to the left peak. The second peak is the dangling bond state related to the broken N-Si bond. The energetic ordering of these two states depends on the local geometry and can interchange. Nevertheless, every two fold coordinated N atom is characterized by two sharp resonances in the electronic DOS, both close to the valence band maximum.

The final class of defects are distorted N centers (N-Si_3) shown in Fig. 48 (d). A strong local distortion or unfavorable electronic environment can shift the N lone pair states up into the band gap. This is again a fairly rare defect species, always doubly occupied and energetically close to the valence band edge.

In summary, the dominant geometrical defects resulting in gap states are three fold coordinated Si atoms (K-defects) and tetrahedrally coordinated, but distorted, Si-Si bonds or Si percolation networks [Fig. 48 (a)-(b)]. In our simulations the K-defects are mostly doubly occupied K^- (below the average Fermi-level) or unoccupied K^+. Since the defect density reduces towards the Fermi-level, only a small fraction of geometrical defects qualifies to become singly occupied, namely those that are close to the Fermi-level. Ultimately, only singly occupied K-defects are easy to measure, for instance in paramagnetic resonance experiments. As already mentioned, the study of singly charged K^0 defects is left for future studies. The important question we address in the remainder is how the defect concentration changes with stoichiometry and annealing history.

5.3.3 The defect contributions to the density of states (DOS)

Fig. 49 shows the contributions to the average DOS from states that are localized at under coordinated atoms and from states that are localized at ideally coordinated atoms. Close to the intrinsic Fermi level, these contributions are caused by either K defects (black lines) or by distorted Si-Si bonds (red lines), respectively. The results are shown for the rapidly quenched and slowly cooled bulk systems (first/second column) of a-$Si_3N_{3.5}$ and a-Si_3N_3 (first/second

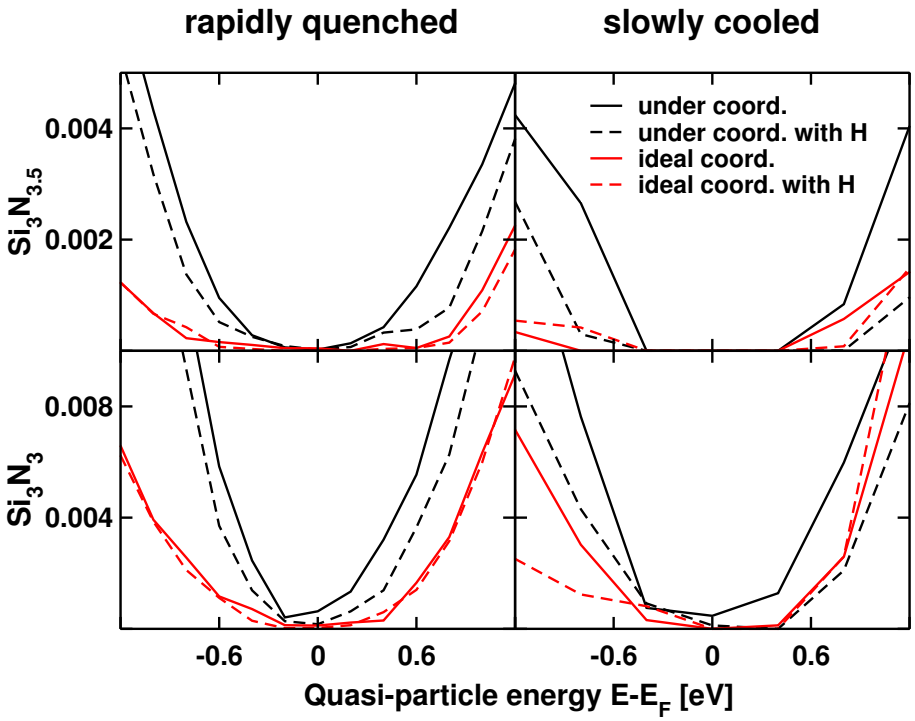


Figure 49: The DOS of a-$Si_3N_{3.5}$ and a-Si_3N_3 (first/second row) close to the Fermi level of the rapidly quenched and the slowly annealed ensembles (first/second column). The states are assigned to atoms with under or ideal coordination (black/red lines). Additional H reduces the amount of electronic states originating from under coordinated atoms (red solid/dashed lines). However, H has little effect on states that are located at ideally coordinated atoms (red solid/dashed lines).

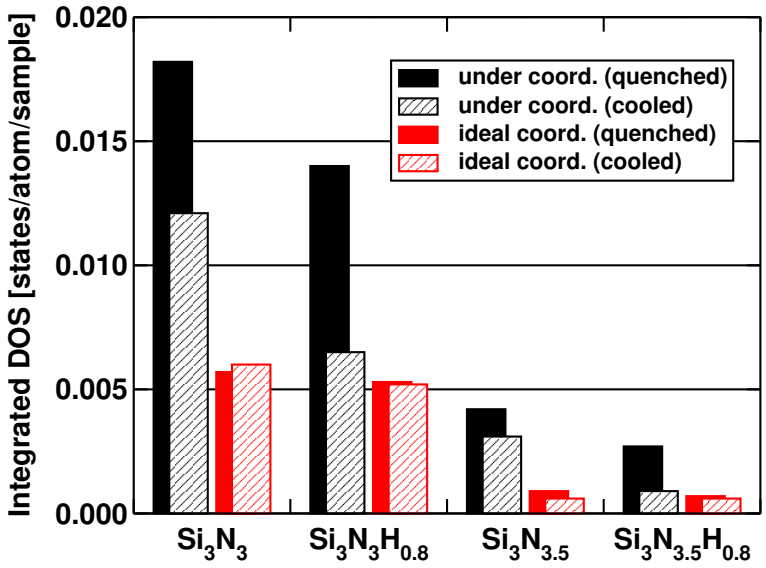


Figure 50: Defect contributions due to atoms with under and ideal coordination (black/red bars) integrated within the range of -1.2 eV and 1.2 eV around the Fermi level for the rapidly quenched and the slowly cooled a-Si_3N_x:H samples (solid/hatched bars). Most defects are located at under coordinated atoms, which are actually K defects. However, for well passivated, slowly cooled a-$Si_3N_{3.5}H_{0.8}$, gap states due to ideally coordinated atoms have become equally numerous as K defects. Additional H reduces the amount of electronic defect states originating from under coordinated atoms, but the number of electronic defects localized at ideally coordinated atoms remains roughly constant.

row). Additionally, we compared the systems with and without H (solid/dashed lines). Note that the quasi-particle energies have been obtained from the KS eigenvalues by applying GW_0 corrections as described in section 5.2.3.

Comparing the four plots in Fig. 49, we make the following observations. Firstly, for comparable cooling conditions and H concentrations, the defect concentration in a-$Si_3N_{3.5}$ (first row) is generally much smaller than in a-Si_3N_3 (second row). Note that the y-axes of the corresponding plots show a different range of values. This observation is clear, as we already know from the structural analysis that K defects and Si-Si bonds are rare in compositions close to a-Si_3N_4. In contrast, these geometrical defects more often occur in a-Si_3N_3.

Secondly, the total defect concentration is generally reduced by slower cooling conditions, and the cooling condition influences the fraction of coordination defects strongly. For instance, for all rapidly quenched systems (first column), the total defect density is made up of 70-80 % coordination defects (solid black lines) and 20-30 % Si-Si states (solid red lines). Hence the clear majority of electronic gap states are related to coordination defects. However, for slower cooling (second column), the concentration of coordination defects is decreased by 20-60 % (solid black lines). All defect contributions tail exponentially from the VB as well as from the CB side into the band gap. Though this tailing is especially obvious in the rapidly quenched systems (first column), it is, despite inferior statistics, still visible in the slowly cooled systems (second column).

Finally, we find that coordination defects can also be passivated by additional H. For example, in the rapidly quenched systems (first column), H cures 20-30 % of all coordination defects (solid/dashed black lines) and, in the slowly cooled systems (second column), it even cures up to 70 % of them (solid/dashed black lines). The effectiveness of the passivation consequently depends on the cooling rate as well. As H is very light and small, it is still very mobile at low temperatures. Slower cooling rates are, therefore, needed to give the H atoms enough time to reach the coordination defects so that they can passivate the dangling bonds. Obviously, H hardly cures defects related to ideally coordinated atoms such as most Si-Si bonding and anti bonding states. We believe that such "defects" are caused by strong local distortions, and they might be partially related to local strain caused by the small supercell sizes.

To quantify the amount of defects in the samples, we have then integrated the defect contributions within the energy range between -1.2 eV and 1.2 eV around the Fermi level Fig. 50. The results are again shown for under and ideally coordinated defects (black/red bars), for the rapidly quenched and slowly cooled samples (solid/hatched bars), and for all system compositions with and without H. In summary, we observe the highest defect concentrations in rapidly

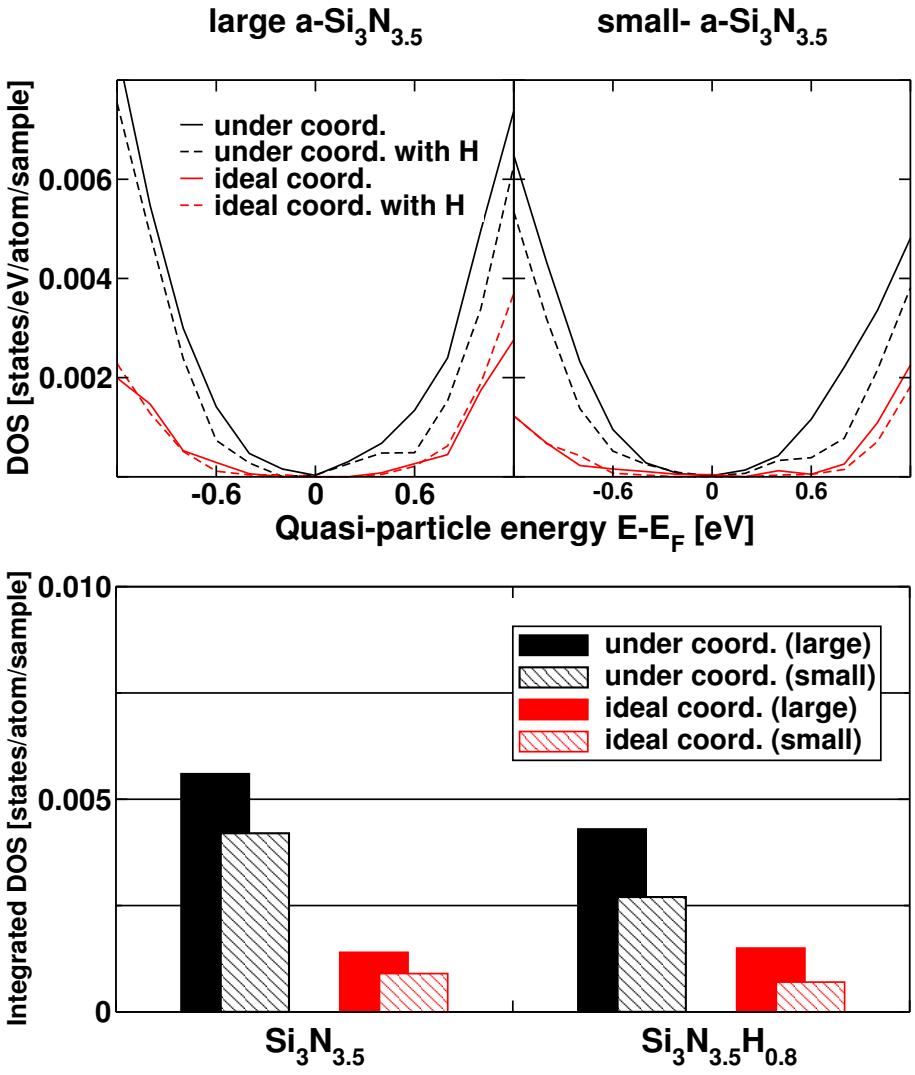


Figure 51: A comparison of the defect contributions to the total DOS in a-$Si_3N_{3.5}$ for two different system sizes. Although the larger samples show in general a larger defect density, the contributions from under coordinated atoms remain dominant and H still mainly cures coordination defects.

quenched a-Si_3N_3 without H (leftmost bars) and the lowest in slowly cooled a-$Si_3N_{3.5}$ with H (rightmost bars), and defect concentrations depend on three factors: (i) the system composition, (ii) the cooling rate, and (iii) the H concentration. In all systems K defects are dominant. However, in slowly cooled a-$Si_3N_{3.5}$ with H, the concentration of defects due to distorted Si-Si bonds is nearly on par with the concentration of coordination defects (rightmost hatched bars). This is because H almost exclusively passivates coordination defects (solid/hatched black bars).

Since finite size effects can be important, we have additionally created 500 rapidly quenched models for a-$Si_3N_{3.5}$ and a-$Si_3N_{3.5}H_{0.8}$ containing twice as many atoms (about 200 atoms) as the small structures (about 100 atoms) considered hereto. The unit cells of these structures are based on our large cubic reference structures, compare Tab. 1 in section 4.3. These calculations are about for times more expensive, than the calculations for the smaller ensembles, and possibly not as well equilibrated. In Fig. 51 the defect contributions between -1.2 eV and 1.2 eV around the Fermi level are again compared. We find more defects in the larger samples, but our essential results are unmodified. Under coordinated Si atoms are dominant in the rapidly quenched sample structures and they can be effectively passivated by H.

5.4 The gap states at the c-Si/a-$Si_3N_{3.5}$:H interface

To investigate the structural and electronic properties of the c-Si/a-$Si_3N_{3.5}$:H interfaces, we use the same analysis framework as for the a-$Si_3N_{3.5}$:H bulk samples. As described in section 5.3.1, the states in the gap are first identified on an electronic level and then assigned to particular atoms where the states are mainly localized. The gap states that are localized at Si atoms can then further be distinguished by their spacial location (i.e. bulk and interface states), their energy level (i.e. VB and CB states), and their participation ratio (i.e. localized and delocalized states). Like in the bulk samples, our most important classes of gap states arise from K defects and states that located at formally ideally coordinated Si atoms, but with distorted Si-Si bonds. There are almost no N defects, and ideally coordinated N atoms are negligibly contributing to states in the middle of the band gap, as they contribute only near the VBM. Since we use non spin polarized calculations, nearly all states are doubly occupied or fully unoccupied. However, a few partially occupied states can be found close to the intrinsic Fermi level of the semiconductor/insulator heterojunction.

5.4.1 The spacial localization of the gap states at c-Si/a-$Si_3N_{3.5}$:H interfaces

To examine the changes of the structural and the electronic properties as a function of the distance from the interface, Fig. 52 shows the number densities of Si, N, and H and the density

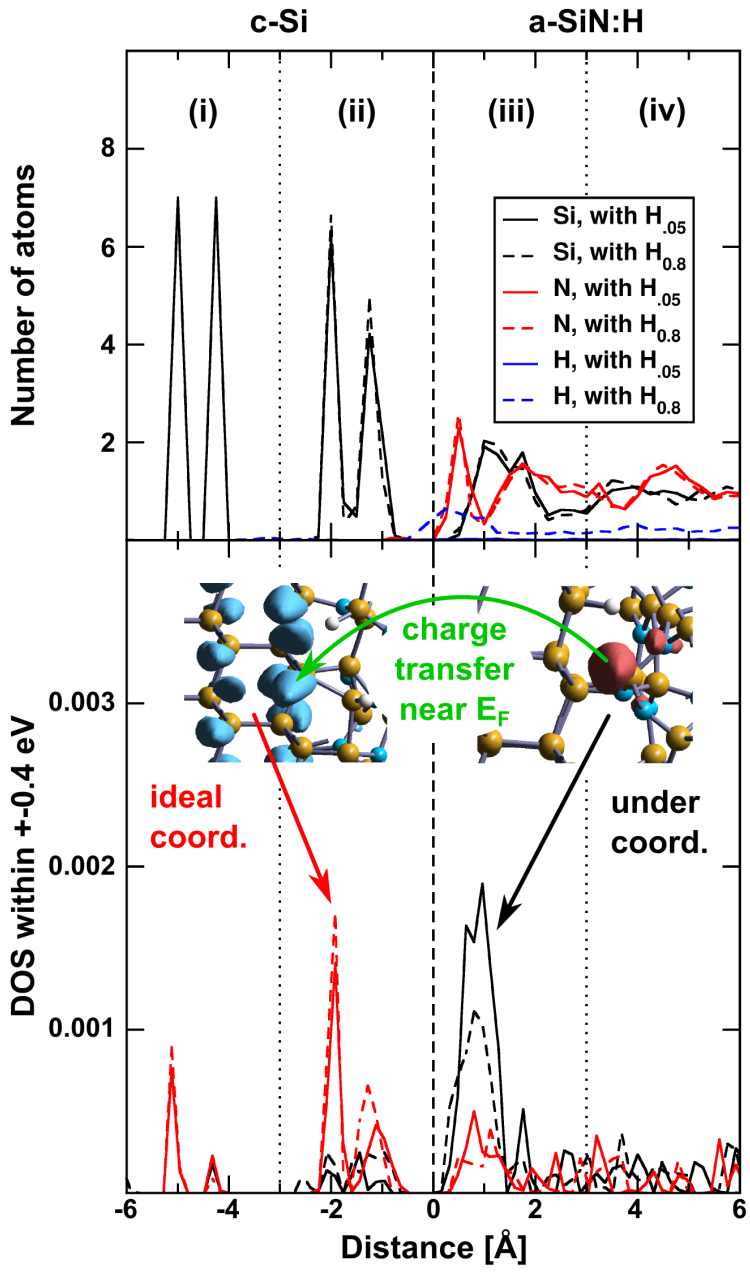


Figure 52: The upper panel shows the number densities of Si, N and H (black, red, and blue lines in the upper panel) at the slowly cooled c-Si/a-$Si_3N_{3.5}$:H interfaces (red and black lines in the lower panel) with 1 and 11 at.% H (solid/dashed lines in both panels). Furthermore, the electronic defect contribution to the total DOS are compared in the lower panel. The considered electronic states lie close to the Fermi level (±0.4 eV) and are localized either at ideally or under coordinated atoms (red/black lines in the lower panel). These states are mostly unoccupied conduction band (CB) states or occupied dangling bonds in a-$Si_3N_{3.5}H_{0.8}$ (blue/red isosurface in the lower panel). In 7 % of our samples, charge is transfered from a K^+ defect on the a-$Si_3N_{3.5}$:H side to "CB-like" states on the c-Si side (green arrow).

of gap states along the z-direction of the supercells. The results are shown only for the slowly cooled interfaces with 1 and 11 at.% H (solid/dashed lines), as these samples are well passivated and the rapidly quenched interfaces basically give similar results. Due to the applied periodic boundary condition (PBC) we have in principle two interfaces. To treat both interfaces together, we mirrored all densities and averaged them over all samples. The mirror planes are illustrated in Fig. 28 of section 4.4. As most effective SRH recombination centers lie in the middle of the gap (i.e. deep defect centers), the investigation of the gap states is limited to the states within the energy window between -0.4 and 0.4 eV. Note that the colors in the upper panel of Fig. 52 refer to the element types (i.e. Si, N, and H), while the colors in the lower panel refer to the coordination of the defects (i.e. under and ideal coordination). The density profiles in Fig. 52 can be divided into four distinct regions depending on the distance from the interface: (i) a bulk c-Si region, between -6 and -3 Å; (ii) an interface c-Si region, between -3 and 0 Å; (iii) an interface a-$Si_3N_{3.5}$:H region, between 0 and 3 Å; and (iv) a bulk a-$Si_3N_{3.5}$:H region, between 3 and 6 Å.

By comparing the samples with less and more H in Fig. 52, we generally find no differences in the overall number densities (upper panel, solid/dashed lines), but in the electronic defect densities (lower panel, solid/dashed lines). The additional H reduces the number of coordination defects (lower panel, black solid/dashed lines), but it increases slightly the number of electronic states related to ideally coordinated Si atoms (lower panel, red solid/dashed lines). This is because most coordination defects have simply been converted to ideally coordinated atoms by H passivation. The following detailed observations of the interface regions are valid for both H concentrations.

We start with the examination of the structural properties shown in the upper panel of Fig. 52. Although the Si structures in region (i) and (ii) are both crystalline with 7 Si atoms per layer, they also show characteristic differences. While the structure in region (i) has remained perfectly crystalline during the simulation, due to the fixed atom positions, the structure in region (ii) has broadened during the final relaxation step. This is noticeable by the decreased and broadened peak intensities of the Si number density (upper panel, black lines). Obviously, the relaxation has introduced distortions to match the c-Si structure in region (ii) to the a-$Si_3N_{3.5}$:H part at the interface. For the samples with 11 at.% H, we find a very small fraction of H atoms between the first two double layers of c-Si (upper panel, blue dashed line), and, for the samples with 1 at.% H, there is no discernible H content (upper panel, blue solid line). The silicon nitride structures in region (iii) and (iv) are both amorphous, but the influence of the interface is likewise noticeable in the number densities of Si, N and H. In region (iii), there are pronounced

density modulations for all element types, and the maxima are at the typical bulk distances d_{SiSi}=2.5 Å, d_{SiN}=1.75 Å, and d_{SiH}=1.5 Å, compare Tab. 2 in section 5.1.1. These modulations are quite a robust features of the c-Si/a-$Si_3N_{3.5}$:H interface. For example, we tried to get rid of these modulations by varying the width of the a-$Si_3N_{3.5}$:H part in other simulations, but the modulations always remained pronounced close to the interface. In region (iv), the concentration of Si and N are then alternating with less pronounced maxima (upper panel, black and red lines), and the H concentration decreases to a lower, but nearly constant value, in the samples with 11 at.% H (upper panel, blue dashed line). We always observe the largest H concentration close to the interface.

Like the number profile of the elements, the spatially resolved electronic DOS in the lower panel of Fig. 52 also shows two important characteristics of the interfaces. First, most electronic gap states are localized within the interface regions (ii) and (iii). Second, while electronic states in region (ii) are primarily located at atoms with ideal coordination (lower panel, red lines), states in region (iii) are rather located at atoms with under coordination (lower panel, black lines). In region (ii), nearly all Si atoms have ideal coordination, as the movement of the first two c-Si layers was limited during the finite temperature simulation. Therefore, states that are localized in region (ii) are mostly related to the strong variations in bond lengths and bond angles of distorted Si-Si bonds at the interface. We shall see later that these states are rather delocalized "CB like" states (lower panel, red lines). Only the atoms in the first c-Si layer have a higher chance to be under coordinated as they bond to the a-$Si_3N_{3.5}$:H part. In region (iii), most states are related to under coordinated Si atoms (lower panel, black lines). This observation agrees very well with our previous investigation on bulk a-$Si_3N_{3.5}$:H and supports the idea that K defects are the most important defect class in a-Si_3N_{4-x}:H structures. In the "bulk-like" regions (i) and (iv), the concentration of electronic gap states is considerably lower than at the interface. Although region (i) is perfectly crystalline, we also find a few electronic states to be localized here. These are caused by the structural relaxation in region (ii). As electronic states due to distorted Si-Si bonds delocalize in a network of Si bonds, they sometimes spread into the c-Si part.

In short, we see that the density of states in the gap is significantly increased at the c-Si/a-$Si_3N_{3.5}$:H interface compared to a-$Si_3N_{3.5}$:H. This observation is strongly related to the structural mismatch of c-Si and a-$Si_3N_{3.5}$:H, which causes distortions in the first c-Si double layer and long range density fluctuations on the amorphous side. On the c-Si side, the states counted as "gap states" are in fact CB like c-Si states (lower panel, red arrow), whereas the states on the a-$Si_3N_{3.5}$:H side are usually localized at under coordinated Si atoms (lower panel,

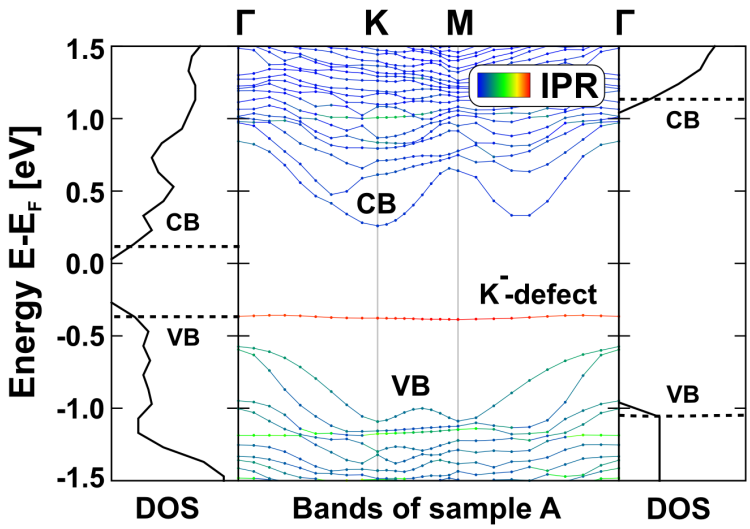


Figure 53: The band structure and the DOS of a single c-Si/a-$Si_3N_{3.5}$:H interface sample in the presence of a defect state. For each state, the degree of localization is given in terms of the IPR. The corresponding color gradient goes from blue, for delocalized states, to red, for localized states. Additionally, we show the DOS of two defect free c-Si and a-$Si_3N_{3.5}$:H bulk samples. All states have been aligned with respect to the average Fermi level of all interface samples. The band offsets are indicated by horizontal dashed lines in the panels of bulk c-Si and bulk a-$Si_3N_{3.5}$:H.

black arrow). The latter are similar to the K defects in bulk a-$Si_3N_{3.5}$:H [139]. A central finding is that charge is partially transfered from originally occupied K defects to the c-Si CB in 7 % of our microscopic models (lower panel, green arrow). This result, which is substantiated in the next section, explains the existence of the positive, fixed charge at c-Si/a-$Si_3N_{3.5}$:H interfaces. When H is present, the H concentration is clearly increased on the a-$Si_3N_{3.5}$:H side close to the interface (upper panel, blue dashed lines), but the N and Si concentrations are hardly changed. Although H reduces the number of K defects at the interface (lower panel, solid/dashed black line), our general observations remain unchanged.

5.4.2 The band structure of the gap states at c-Si/a-$Si_3N_{3.5}$:H interfaces

We now investigate three microscopic models of the c-Si/a-$Si_3N_{3.5}$:H interface in more detail. The first sample A is shown in Fig. 53. It contains a single localized K^- defect in the middle of the gap. Likewise, most coordination defects in our models are doubly occupied K^- defects or doubly unoccupied K^+ defects. Though we also find a few singly occupied K^0 defects close to

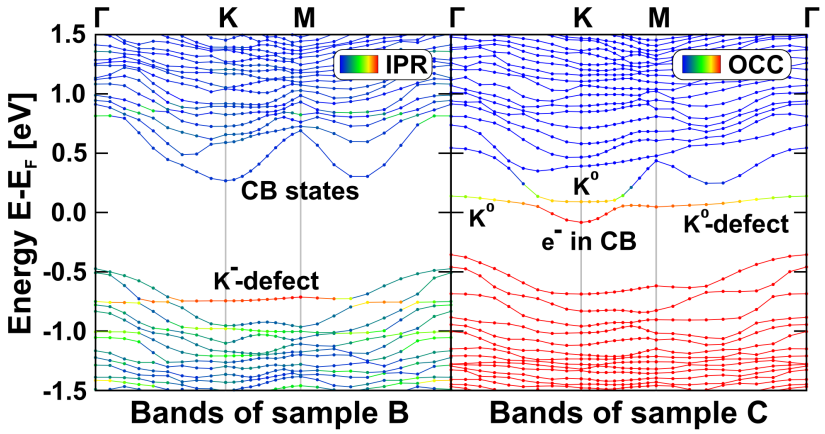


Figure 54: Two examples for forbidden band crossing at a c-Si/a-$Si_3N_{3.5}$:H interface. In sample A, a K^- defect overlaps with the valence band (VB), and, in sample B, a K^- defect overlaps with the conduction band (CB). As demonstrated by means of the IPR in sample A, both defects exchange their localized character with the delocalized character of the VB and the CB, respectively. In sample B, the occupation (OCC) of the defect is partially transfered to the CB.

the Fermi level. In addition to the band structure of sample A, we show the DOSs of bulk c-Si and a defect free bulk a-$Si_3N_{3.5}$:H model structure in the left and right most panel of Fig. 53, respectively. Furthermore, the VB and the CB edges of the bulk samples are indicated by horizontal dashed lines.

The bulk states have been aligned at the electrostatic potentials of the Si and the N atoms in the middle of the c-Si and a-$Si_3N_{3.5}$:H part, respectively. The comparison of the bulk DOSs shows that the indirect band gap of c-Si (about 0.5 eV) is much smaller than the band gap of a-$Si_3N_{3.5}$:H (about 2.3 eV), and the CB offset (about 1.0 eV) is clearly larger than the VB offset (about 0.8 eV). All these findings are consistent with photo emission experiments and previous simulations on c-Si/c-Si_3N_4 interfaces [42, 64]. The band structure furthermore demonstrates that the in plane dispersion of the defect level is negligible. This indicates that the lateral supercell size is generally sufficient in the a-$Si_3N_{3.5}$:H part, as there is no interaction between the periodically repeated defect images. However, since the VB and the CB of c-Si are strongly dispersed, not all electronic states in the c-Si/a-$Si_3N_{3.5}$:H interface can be resolved using a k-point sampling of the Γ point only. For this reason, we have recalculated the ground state of our final interface structures using a 8×8 k-point sampling parallel to the interface.

Since defect related states in the amorphous part can be situated anywhere within the gap

of a-$Si_3N_{3.5}$:H [139], they often overlap with the CB or VB of c-Si. To give examples for such situations Fig. 54 shows the band structure of two samples B and C. In sample B, a doubly occupied K^- defect overlaps with the VB of c-Si, and, in sample C, a K defect overlaps with the CB of c-Si. In both samples, the defect states are again strongly localized and practically flat. Nevertheless, since a small hybridization between the c-Si and the defects exists, forbidden crossings are observed. The most interesting case is sample C. Here the average defect level is slightly above the CBM of c-Si, causing a partial transfer of charge from the defect level into the CB of c-Si. This has reduced the occupancy of the K defect and relates to the main finding mentioned above: occupied K^- defects and K^0 defects are converted to unoccupied K^+ defects at the c-Si/a-$Si_3N_{3.5}$:H interface. The opposite, a partial transfer from the VB of c-Si to a K^+ defect, is never observed.

5.4.3 The density of states in the band gap of c-Si/a-$Si_3N_{3.5}$:H interfaces

To obtain a statistically more meaningful result, we explore the average DOS evaluated for both sides of c-Si/a-$Si_3N_{3.5}$:H interfaces, close to the Fermi level. Fig. 55 shows the electronic contributions from states located on the c-Si side (red lines) and from states located on the a-$Si_3N_{3.5}$:H side (black lines). Furthermore, the figure shows the difference between the interfaces with 1 and 11 at.% H (solid/dashed lines) and the intrinsic Fermi levels of bulk c-Si and bulk a-$Si_3N_{3.5}$:H (red and black vertical lines).

On the c-Si side, we observe a slight band tailing introduced by local disorder and distorted Si bonds (red lines). However, the majority of the gap states are caused by K defects on the a-$Si_3N_{3.5}$:H side (black lines). The most important observation is that the minimum of the contributions from c-Si states is shifted towards lower energy values (left) in comparison to the minimum of the contributions from a-$Si_3N_{3.5}$:H states. Likewise, the intrinsic Fermi level of bulk c-Si lies about 0.16 eV below the average Fermi level of bulk a-$Si_3N_{3.5}$:H, compare red and black vertical lines in Fig. 55. Impaired by the issue of too small KS band gaps, we believe that we observe only the lower bound of this offset here. For the alignment of the bulk Fermi levels, we used the electrostatic potentials of "bulk-like" Si atoms in the c-Si part of the interface and of "bulk-like" N atom in the a-$Si_3N_{3.5}$:H part of the interface. The alignment is illustrated in Fig. 56 and the values used for the alignment are summarized in Tab. 7. Note that all these observations basically apply independently of the H concentration. As for our previous studies on bulk a-$Si_3N_{3.5}$:H [138,139], H mainly reduces the number of defects. For the interface samples, it has reduced the misalignment of the Fermi levels from 0.24 eV to 0.16 eV though.

Fig. 57 illustrates the c-Si/a-SiN:H heterojunction before and after band alignment. On the

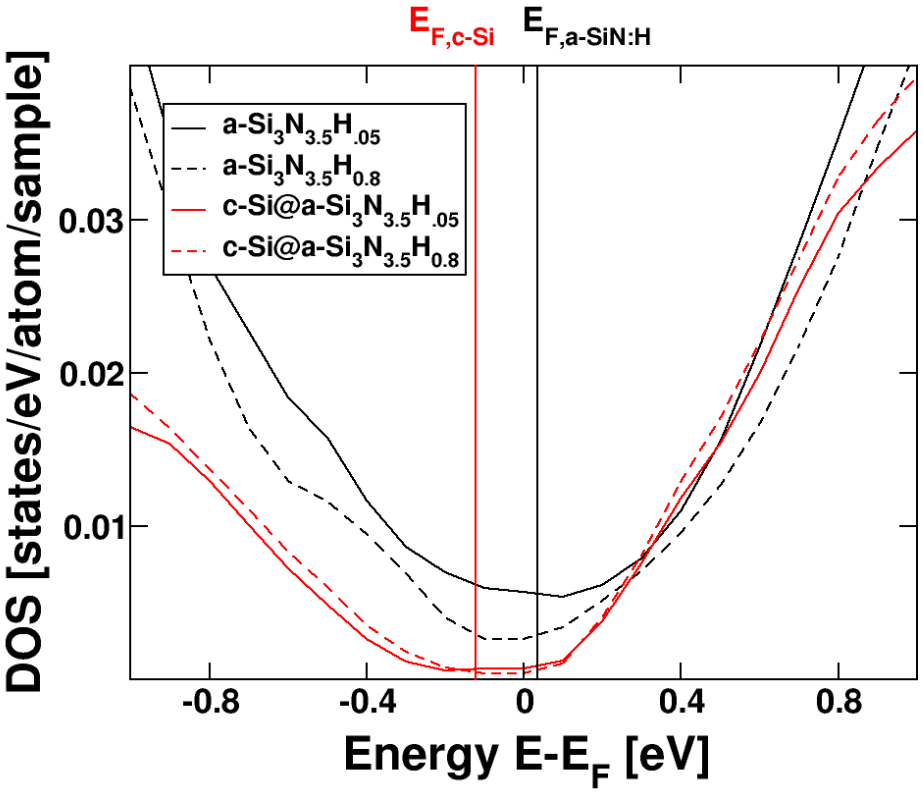


Figure 55: Defect contributions from the c-Si side (red lines) and the a-$Si_3N_{3.5}$:H side (black lines) of the c-Si/a-$Si_3N_{3.5}$:H interfaces with 1 and 17 H atoms (solid/dashed lines). The Fermi level and the band edges of bulk c-Si and bulk a-$Si_3N_{3.5}H_{0.8}$ are indicated (red/black vertical lines).

Table 7: The intrinsic Fermi levels of the c-Si bulk, the a-$Si_3N_{3.5}H_{0.8}$ bulk, and the c-Si/a-$Si_3N_{3.5}H_{0.8}$ interface are calculated as $E_F = (E_{VBM} + E_{CBM})/2$. To align the Fermi levels of the bulk samples with the Fermi level of the interface, we used the average electrostatic potential of bulk Si atoms in c-Si $U_{Si}^{c\text{-}Si}$ and of bulk N atoms in a-$Si_3N_{3.5}H_{0.8}$ $U_N^{a\text{-}SiN:H}$. For the interface system, the potentials are estimated from "bulk-like" atoms located in the middle of the c-Si part and in the middle of the a-$Si_3N_{3.5}H_{0.8}$ part, respectively.

	c-Si	c-Si/a-$Si_3N_{3.5}H_{0.8}$	a-$Si_3N_{3.5}H_{0.8}$
E_{VBM}	5.065	5.023	6.500
E_{CBM}	5.540	5.826	8.827
E_F	5.302	5.424	7.664
$U_{Si}^{c\text{-}Si}$	-55.147	-55.147	–
$U_N^{a\text{-}SiN:H}$	–	-59.605	-57.400

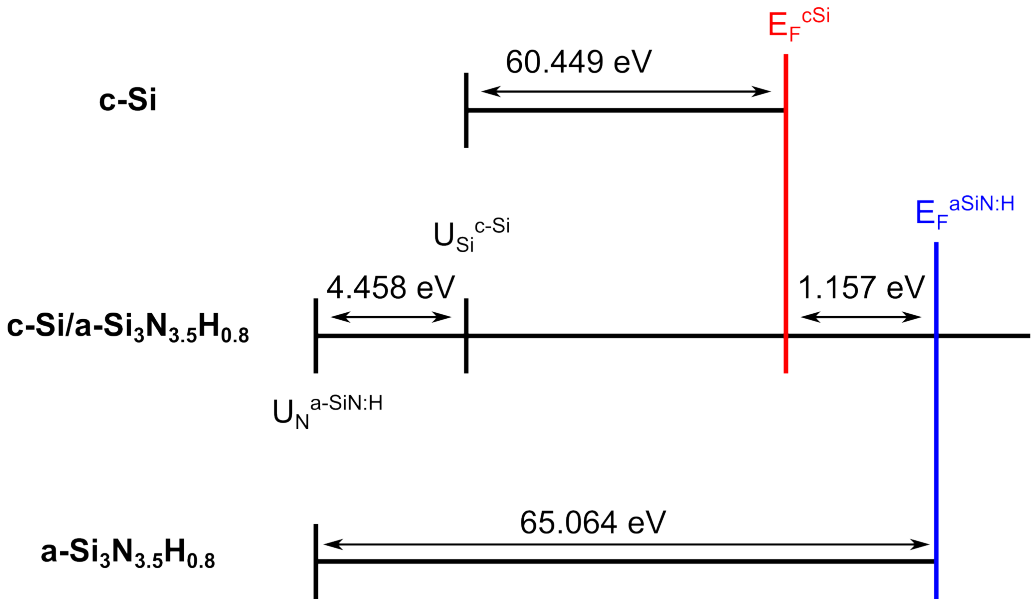


Figure 56: The alignment of the Fermi levels of the c-Si bulk $E_F^{c\text{-}Si}$ and the a-$Si_3N_{3.5}H_{0.8}$ bulk $E_F^{a\text{-}SiN:H}$ at the c-Si/a-$Si_3N_{3.5}H_{0.8}$ interface. For the alignment of $E_F^{c\text{-}Si}$, we used the electrostatic potential of the Si atoms in the c-Si part of the interface $U_{Si}^{c\text{-}Si}$, and, for the alignment of $E_F^{a\text{-}SiN:H}$, we used the electrostatic potential of the N atoms in the a-$Si_3N_{3.5}H_{0.8}$ part of the interface $U_N^{a\text{-}SiN:H}$.

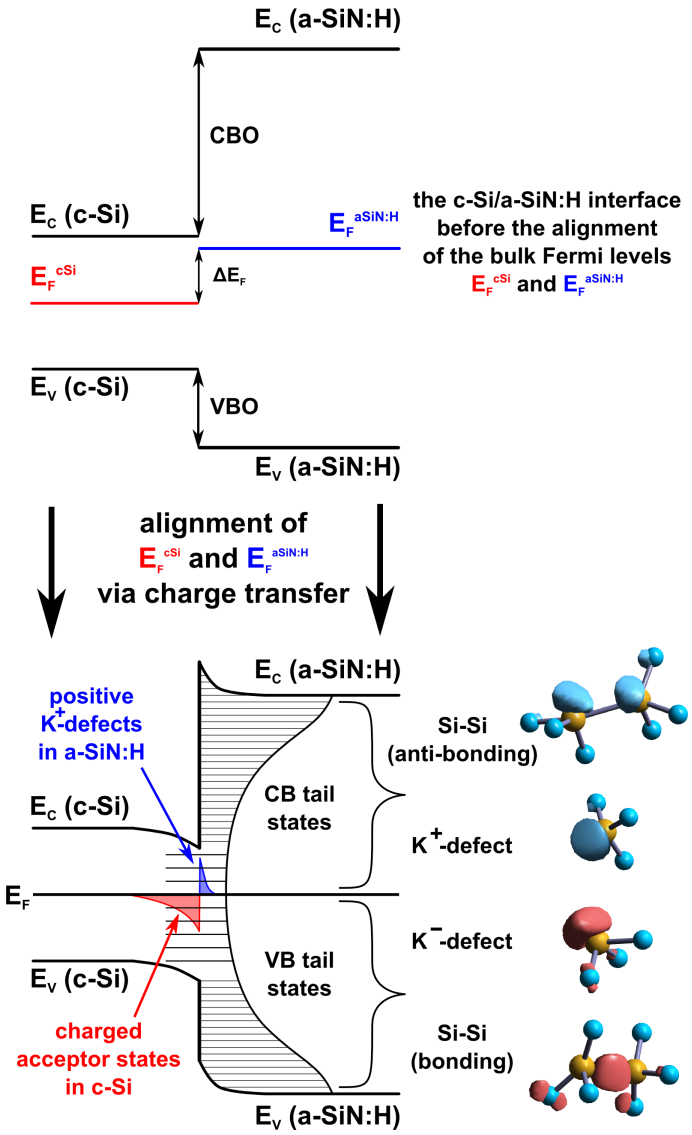


Figure 57: The band offsets and the intrinsic Fermi levels in a c-Si/a-SiN:H heterojunction before and after band alignment. Before the alignment (upper part), the difference between the Fermi levels of bulk c-Si and of bulk a-SiN:H is ΔE_{F}. After the alignment (lower part), both Fermi levels have aligned to only one common Fermi level, but the conduction band offsets have remained unchanged. In general, localized defect states at c-Si/a-$Si_3N_{3.5}$:H interfaces have the same character as defect states in the band gap of bulk a-$Si_3N_{3.5}$:H. They are mostly K^- defects and K^+ defects or they are related to distorted Si-Si bonds. In the lower part of the figure, occupied and unoccupied orbitals are indicated via red and blue isosurfaces, respectively. We suppose that the alignment of the Fermi levels is related to a charge transfer from K^- defects on the a-$Si_3N_{3.5}$:H side to acceptor states on the c-Si side. The resulting K^+ defects form a positive charge density (lower part, blue area) that is screened by the negatively charged acceptor states (lower part, red area).

left hand side, we see the intrinsic Fermi level, the VBM and the CBM of bulk c-Si, and, on the right hand side, we see the same for bulk a-SiN:H. In the middle, the offset between the Fermi level of bulk c-Si and bulk a-SiN:H is indicated by the difference between the red and the blue lines (upper part of the figure). In a macroscopic sample, the bulk Fermi levels align to one interface Fermi level, as indicated by the straight black line going from the c-Si side to the a-SiN:H side (lower part of the figure). As already mentioned in section 5.4.1 and section 5.4.2, we assume that this alignment is done via a charge transfer from doubly occupied K^- defects and singly occupied K^0 defects to acceptor levels in c-Si. This assumption is based on the fact that most defects within the band gap area of a-SiN:H are actually K defects, as indicated by the prototype defects in the lower part of Fig. 57. After the charge transfer, the K^- defects and K^0 defects are converted to K^+ defects. Therefore, we consider the offset of the bulk Fermi levels as the reason for the generation of the positive, fixed charge at the c-Si/a-SiN:H interface. The fixed charge is located on the a-SiN:H side of the interface (lower part of the figure, blue arrow) and is screened by the charged acceptor states on the c-Si side (lower part of the figure, red arrow).

However, in our microscopic realizations, there are no acceptor states on the c-Si side of the interface except for the CBM. For this reason the bulk Fermi levels are not aligned and the microscopic Fermi level of the c-Si part can be located anywhere within the band gap area. Furthermore, only if an occupied K defect is located above the CBM of c-Si, a partial charge transfer to the c-Si part is observed, like in sample C in Fig. 54. This partial charge transfer shifts the c-Si states upwards and the a-$Si_3N_{3.5}$:H states downwards. This line of thought indicates that it is not quite straightforward to determine highly accurate CB and VB offsets between amorphous model structures and crystalline samples using small model structures, even though we average over many microscopic models. That the average Fermi level of c-Si lies well below that of a-$Si_3N_{3.5}$:H is, however, a robust feature corroborated by three independent evidences: (i) the observation of microscopic models with a charge transfer to c-Si, (ii) the visual misalignment of the minima in the electronic DOSs, and (iii) the misalignment of the calculated bulk Fermi levels.

6 Conclusions

Electronic and structural properties of amorphous materials can effectively be examined by using exclusively *ab initio* methods. We generated many configurations of bulk a-Si_3N_x:H and of the interface between c-Si and a-Si_3N_x:H by cooling many different structures from the liquid to the amorphous state. The materials properties (e.g. PCF, ADF, STF, DOS, band gap) were then averaged and an automatic routine was applied to obtain statistical information on structural and electronic defects. Although system sizes, cooling rates, and system sizes were still very limited in comparison to simulations based on force fields, our methodology shows how to use todays computational recourses effectively and how to identify electronic defects consistently on the electronic level.

The investigation of the bulk properties of a-Si_3N_3, a-$Si_3N_{3.5}$, a-Si_3N_4, and a-$Si_3N_{4.5}$ has showed that all a-Si_3N_x structures are dominated by Si-N bonds. These bonds form either tetrahedral SiN_4 units or planar, triangular NSi_3 units. The lowest concentration of coordination defects, Si-Si bonds, and N-N bonds are thereby found in the a-Si_3N_4 samples. These samples also show the lowest atom diffusion and the largest optical band gap of all systems. More geometric defects and smaller band gaps are, in contrast, observed in the a-Si_3N_{4-x} and the a-Si_3N_{4+x} samples. For Si rich a-Si_3N_{4-x}, we furthermore found out that Si-Si bonds interconnect to whole Si clusters. In the samples with additional H, H reduces the number of coordination defects, but it has little effect on the general structural properties. All these results are in excellent agreement with scattering and absorption measurements. Therefore, we believe that the generated ensembles of many different model structures are a good basis for a detail investigation on structural and electronic defects.

Defects are minority species, but they still have a impact on the DOS in the vicinity of the band gap. We observed that they strongly influence the position of the microscopic Fermi level and effectively reduce the size of the average band gap. Performing a detail defect analysis, K defects and Si-Si bonds were identified to be the most important defects in a-Si_3N_3 and a-$Si_3N_{3.5}$, whereas the coordination defects are predominant close to the intrinsic Fermi level. In contrast to Si atoms, N atoms have mostly ideal coordination and always lone pair electron states that contribute only at the VBM. In principle those defect contributions have been known since 1984 [59], but computer investigations on electronic defects have lacked in accuracy and significance so far. This might be the reason why other studies often create the impression that both defect types are associated with particular energy levels. Specifically, it seems that K defect are always located close to the Fermi level and Si-Si states are always located at the band

edges. However, those relations are not true in general because we detected both defect types anywhere in the band gap. We cannot judge whether this finding is just unwritten knowledge in the scientific community or not, but we can clarify this point right now for future studies.

The most important parameters that control the defect density in the band gap are identified by a statistical analysis. In agreement with experiments and other simulations, we found out that the defect density is strongly influenced by the system composition and the cooling rate. While more defects are detected in very off stoichiometric, rapidly quenched samples, less defects are found in nearly stoichiometric, slowly cooled samples. This result highlights the importance of a careful sample preparation not only in simulations, but also in experiments. We further observed that additional H enhances the reduction of defects by attaching to K defects. This process is in line with the typical model for the passivation of dangling bonds by H. On the other hand, defects related to ideally coordinated atoms cannot be cured by H, and we hence believe that there is a natural limitation for the defect passivation by H in amorphous silicon nitride. Although the incurable defects become equally important as coordination defects in well passivated samples, they have been hardly mentioned in literature yet.

A electronic defect state is in general either localized on a particular atom or delocalized over several atoms. For this reason, a isolated electronic defect state, with no other defects in its immediate spatial vicinity, is normally strongly localized. However, if its associated structural defect is interconnected to a network of Si-Si bonds or close to other geometrical defects, the electronic defect state often spreads over several atoms and delocalizes more. As a matter of fact, the localization of electronic states has important consequences for the electronic defect properties as it affects the interaction with charge carriers and other electronic states. For example, Tauc *et al.* reported that electronic transitions between the VB and the CB often involve the interaction between localized and delocalized states [154]. Transitions between two localized states are, in contrast, rather unlikely. We can hence consider isolated defects as pure carrier traps, while delocalized defect states are a natural requirement for many carrier recombinations. Since we found more localized states in the stoichiometric samples and more delocalized states in the under stoichiometric samples, we conclude that not only the number of defects, but also the manner of electronic transitions, can be changed by controlling the silicon nitride composition.

At the c-Si/a-$Si_3N_{3.5}$:H interface, we basically detected the same defect types as in the a-$Si_3N_{3.5}$:H bulk. The states in the band gap are again predominantly localized at K defects or at distorted Si-Si bonds. However, we also observed some differences and additional features. Firstly, both type of gap states are mainly localized close to the interface. In this region, there

is also the largest structural mismatch between the c-Si and a-$Si_3N_{3.5}$:H part of the interface. Secondly, the Si-Si states are localized on the c-Si side, while the coordination defects are localized on the a-$Si_3N_{3.5}$:H side. Thirdly, the states in the a-$Si_3N_{3.5}$:H part are strongly localized on coordination defects, while the c-Si state are rather delocalized with features of the band structure that are inherited from the CB of bulk c-Si. Finally, there is a energy offset between the defect contributions from the c-Si side and the contributions from the a-SiN:H side, and, similarly, the Fermi level of bulk c-Si lies under that of bulk a-SiN:H.

We believe that, in a macroscopic device, the Fermi levels of c-Si and a-SiN:H always align to one common Fermi level by a charge transfer from the a-SiN:H side to the c-Si side. In our microscopic simulation cells, however, a energy offset of the Fermi levels is still visible due to the lack of acceptor states in the c-Si part. Although the alignment of the Fermi levels is limited in our simulations, we found a few model structures that give evidence for the charge transfer in a real device. In those samples, coordination defects on the a-SiN:H side donate their charge partially to CB states located on the c-Si side. We conclude that this charge transfer explains the formation of the positive, fixed charge at c-Si/a-SiN:H interfaces. Since the dangling electrons of K^0 defects and of K^- defects do not participate in a bonding, these defects are preferred donors. Once the electrons are transfered to the c-Si side, the occupied coordination defects are converted into unoccupied K^+ defects. This charge transfer creates a positive charge density that is localized on the a-SiN:H side close to the interface. Coordination defects can, therefore, act not only as trapping and recombination centers, but also as positive, fixed charges that introduce field-effect passivation. For this reason the passivation using amorphous silicon nitride is always a compromise between those properties.

Although our results can easily be rationalized and are in good agreement with the present experimental evidence and prevailing assumptions, there is still the possibility for further improvements. All our calculations were performed non spin polarized and with a even number of electrons. Consequently, nearly all states below the Fermi level are doubly occupied, and nearly all states above the Fermi level are unoccupied. Only close to the Fermi level, there are a few singly occupied states. This situation mostly coincides with the experimental situation, but, for more detail information on defects with a specific occupation, spin polarization is nonetheless very desirable. Furthermore, we must note that the mobility of H might not only support the passivation of the defects, but also the creation of the defects. For example, $N_3 \equiv Si-H$ centers, which are omnipresent in our H rich samples, might easily be converted into $N_3 \equiv Si\bullet$ (K^0 defects) by impinging light or thermal degradation. This situation could also be modeled on the basis of our samples by simply removing particular H atoms from the final silicon nitride

structures. There is also the possibility to apply large scale MD simulations to other amorphous structures. A good candidate is, for example, amorphous silicon oxide as it is also used for solar cells and for many other semiconductor devices.

Ultimately, we want to emphasize that defect states in amorphous silicon nitride cannot be investigated on the basis of only a few model structures. As defects heavily influence the electronic properties of a single sample structure, meaningful conclusions can only be obtained by averaging over many independent sample structures. The applied pure *ab initio* method demonstrates that large scale MD simulation can effectively be utilized to sample the configuration space of amorphous structures without the need of extra force field methods. We think that this is one of the most accurate and consistent ways to examine electronic defect states nowadays. Since our defect analysis is in principle applicable to every amorphous structure, we hope that we can inspire researchers to adapt our method also for other materials. This would definitely increase the significance of modern defect investigations in the field of computational materials science.

7 Appendix

7.1 Fourier transformation

A periodic function $f(x+L) = f(x)$ can be expanded into its discrete Fourier coefficients F_{k_n}, with $k_n = n\frac{2\pi}{\lambda}$, by

$$f(x) = \sum_{n=-\infty}^{\infty} e^{-ik_n x} F_{k_n} \tag{211}$$

$$F_{k_n} = \frac{1}{L}\int_{-L/2}^{L/2} dx\, e^{ik_n x} f(x). \tag{212}$$

Thereby the coefficients are average values of the function $f(x)$. The function is weighted by the phase factor $e^{-ik_n x}$ for each x within the considered interval of length L.

Similarly, a non-periodic function $f(t)$ can be expanded into its continuous Fourier spectrum $F(\omega)$, with $\omega = \frac{2\pi}{T}$, by

$$f(t) = \int_{-\infty}^{\infty} d\omega\, e^{i\omega t} F(\omega) \tag{213}$$

$$F(\omega) = \int_{-\infty}^{\infty} dt\, e^{-i\omega t} f(t). \tag{214}$$

7.1.1 The Fourier transformation of the Coulomb kernel

For every radial symmetric function $f(r)$, the Fourier transformation can be performed by using spherical coordinates.

$$g(k) = \int d\mathbf{r} e^{-i\mathbf{qr}} f(r) \tag{215}$$

$$g(k) = \int_0^{\infty} dr\, r^2 \int_0^{2\pi} d\varphi \int_0^{\pi} d\theta\, \sin(\theta) e^{-i\,q\,r\cos(\theta)} f(r) \tag{216}$$

$$g(k) = 2\pi \int_0^{\infty} dr\, r^2 f(r) \int_0^{\pi} d\theta\, \sin(\theta) e^{-i\,q\,r\cos(\theta)} \tag{217}$$

$$\tag{218}$$

Using the variable transformation $u = \cos(\theta)$ and $du = -\sin(\theta)d\theta$, we can furthermore write

$$g(k) = 2\pi \int_0^{\infty} dr\, r^2 f(r) \int_1^{-1} (-1)\, du\, e^{-i\,q\,r\,u} \tag{219}$$

$$g(k) = 2\pi \int_0^{\infty} dr\, r^2 f(r) \int_{-1}^{1} du\, e^{-i\,q\,r\,u} \tag{220}$$

$$g(k) = 2\pi \int_0^{\infty} dr\, r^2 f(r) \frac{e^{i\,q\,r} - e^{-i\,q\,r}}{i\,q\,r} \tag{221}$$

$$g(k) = 2\pi \int_0^{\infty} dr\, r^2 f(r) \frac{2\sin(qr)}{qr}. \tag{222}$$

Choosing the specific identity $f(r) = 1/r = \lim_{\eta\to 0} e^{-\eta r}/r$, we finally obtain the Fourier transformed Coulomb kernel.

$$
\begin{aligned}
g(k) &= \frac{4\pi}{q} \lim_{\eta\to 0} \int_0^\infty dr\, e^{-\eta r} \sin(qr) && (223)\\
g(k) &= \frac{4\pi}{q} \lim_{\eta\to 0} \mathrm{Im} \int_0^\infty dr\, e^{iqr-\eta r} && (224)\\
g(k) &= \lim_{\eta\to 0} \frac{4\pi}{q^2+\eta^2} = \frac{4\pi}{q^2} && (225)\\
& && (226)
\end{aligned}
$$

7.1.2 The convolution theorem

In reciprocal space, the convolution h of two functions r and f can be expressed as a simple product using the Fourier spectra H, R, and F.

$$
\begin{aligned}
h(t) &= \int dt'\, r(t-t') f(t') && (227)\\
H(\omega') &= R(\omega') F(\omega') && (228)
\end{aligned}
$$

It can be shown that the same rule also applies in the case of multiple convolutions in one equation.

$$
\begin{aligned}
h(t) &= \int dt' dt'' r_1(t-t') r_2(t'-t'') f(t'') && (229)\\
H(\omega') &= R_1(\omega') R_2(\omega') f(\omega') && (230)
\end{aligned}
$$

In the following, we show the proof for the case of a double convolution:

$$
\begin{aligned}
h(t) &= \int dt' dt'' r_1(t-t') r_2(t'-t'') f(t'') && (231)\\
\int d\omega' e^{-i\omega' t} H(\omega') &= \int dt' dt'' && (232)\\
&\times \Big(\int d\omega' e^{-i\omega'(t-t')} R_1(\omega') \Big) && (233)\\
&\times \Big(\int d\omega'' e^{-i\omega''(t'-t'')} R_2(\omega'') \Big) && (234)\\
&\times \Big(\int d\omega''' e^{-i\omega'''(t''-t''')} f(\omega''') \Big) && (235)\\
\int d\omega' e^{-i\omega' t} H(\omega') &= \int d\omega' d\omega'' d\omega''' e^{-i\omega' t} R_1(\omega') R_2(\omega'') f(\omega''') && (236)\\
&\times \underbrace{\Big(\int dt' e^{-it'(\omega''-\omega')} \Big)}_{\delta(\omega'-\omega'')} \underbrace{\Big(\int dt'' e^{-it''(\omega'''-\omega'')} \Big)}_{\delta(\omega'''-\omega'')} && (237)\\
\int d\omega' e^{-i\omega' t} H(\omega') &= \int d\omega' e^{-i\omega' t} R_1(\omega') R_2(\omega') f(\omega') && (238)
\end{aligned}
$$

For higher order convolutions, the rule can be demonstrated analogously.

7.2 Functional derivative

If we consider a functional F depending on a function g, then g can be varied by adding a function h times a scalar ϵ. The functional derivative with respect to g is then defined in terms of a Taylor expansion [165].

$$F[g+\epsilon h] = F[g] + \epsilon h \frac{\partial F[g]}{\partial g} + O(\epsilon^2) \tag{239}$$

$$h\frac{\partial F[g]}{\partial n} = \lim_{\epsilon\to 0}\frac{(F[g+\epsilon h]-F[g])}{\epsilon} \tag{240}$$

Let us for example consider the Hartree potential as a functional of the electronic density n

$$F[n](\mathbf{r},t) = \int d\boldsymbol{\xi}\,\frac{n(\boldsymbol{\xi},t)}{|\mathbf{r}-\boldsymbol{\xi}|} = \int d\boldsymbol{\xi} d\tau\,\frac{n(\boldsymbol{\xi},\tau)\delta(\tau-t)}{|\mathbf{r}-\boldsymbol{\xi}|}. \tag{241}$$

For the special choice of $h = \delta(\xi-\mathbf{r}')\delta(\tau-t')$, the functional derivative of the Hartree potential with respect to the electronic density is then given by

$$\frac{\partial F[n](\mathbf{r},t)}{\partial n(\mathbf{r}',t')} = \lim_{\epsilon\to 0}\frac{F[n+\epsilon\delta(\boldsymbol{\xi}-\mathbf{r}')\delta(\tau-t')]-F[n]}{\epsilon} \tag{242}$$

$$= \int d\boldsymbol{\xi} d\tau \frac{\delta(\tau-t)\delta(\tau-t')\delta(\boldsymbol{\xi}-\mathbf{r}')}{|\mathbf{r}-\boldsymbol{\xi}|} \tag{243}$$

$$= \frac{\delta(t-t')}{|\mathbf{r}-\mathbf{r}'|}. \tag{244}$$

7.3 Linear response as first order Taylor expansion

Let us consider a functional g that depends on a function f. If a perturbation term δf is add to g, $g[f+\delta f]$ can be expressed as a first order Taylor expansion in f by [166]

$$g[f+\delta f] = g[f] + \frac{\delta g[f]}{\delta f}\delta f + O\Big[(\delta f)^2\Big]. \tag{245}$$

If we consider only small variations of δf and we define the consequent variation of g by $\delta g = g[f+\delta f]-g[f]$, we can neglect all higher order terms and obtain $\delta g = \frac{\delta g[f]}{\delta f}\delta f$. Finally, if g and f depend on a variable x, and g is affected by all variations of δf due to x, we find the usual expression for a response function R by

$$\delta g(x) = \int dx' \underbrace{\frac{\delta g(x)}{\delta f(x')}}_{R(x,x')}\delta f(x'). \tag{246}$$

The δ symbol that indicates the variation of the functional g through the function f is sometimes neglected in literature.

$$g(x) = \int dx' \underbrace{\frac{\delta g(x)}{\delta f(x')}}_{R(x,x')} f(x'). \tag{247}$$

However, we must keep in mind that the linear response is the linear part of a Taylor expansion that gives the variation of a functional due to a small perturbation term.

7.4 Dielectric response

Let us consider a material in the presence of an external charge density $n_{\text{ext}}(\mathbf{r})$ that causes an external field $\mathbf{E}_{\text{ext}}(\mathbf{r}) = \mathbf{D}(\mathbf{r})$. The material responds to this external perturbation by introducing an induced charge density $n_{\text{ind}}(\mathbf{r})$ and an electrical field $\mathbf{E}_{\text{ind}}(\mathbf{r}) = -4\pi\mathbf{P}(\mathbf{r})$. As a result, the external electrical field within the material is reduced to $\mathbf{E}_{\text{tot}}(\mathbf{r}) = \mathbf{E}_{\text{ext}}(\mathbf{r}) + \mathbf{E}_{\text{ind}}(\mathbf{r}) = \mathbf{D}(\mathbf{r}) - 4\pi\mathbf{P}(\mathbf{r})$ [167]. To characterize the response of the material, we want to find explicit relationships between external and introduced fields. For the very simple example of macroscopic, homogeneous, isotropic materials that are exposed to a constant external field, such a relationship is given by $\mathbf{E} = \epsilon^{-1}\mathbf{D}$. Here the fields are connected by the dielectric constant $\epsilon^{-1} = E/D$.

At microscopic scale, materials are usually inhomogeneous and the response to external perturbations depends on space and time [167, 168]:

$$E_i(\mathbf{r}, t) \quad = \quad \int d\mathbf{r}'\, dt'\, \epsilon_{ij}^{-1}(\mathbf{r}, \mathbf{r}', t - t')\, D_j(\mathbf{r}', t'), \tag{248}$$

where

$$\epsilon_{ij}^{-1}(\mathbf{r}, \mathbf{r}', t - t') = \delta\mathbf{E}_i(\mathbf{r}, t)/\delta\mathbf{D}_j(\mathbf{r}', t') \tag{249}$$

is commonly known as the inverse dielectric matrix [167, 168]. Although this form gives the natural description of a response to an external perturbation, the external field is, in literature, most often expressed in terms of the total field and the dielectric matrix ϵ_{ij}.

$$D_i(\mathbf{r}, t) = \int d\mathbf{r}'\, dt'\, \epsilon_{ij}(\mathbf{r}, \mathbf{r}', t - t')\, E_j(\mathbf{r}', t'). \tag{250}$$

Since we want to investigate periodic structures and ϵ_{ij} and ϵ_{ij}^{-1} depend on a time difference $t - t'$ only, it is often more convenient to transform the formulas into reciprocal space and frequency space.

$$E_i(\mathbf{q}, \omega) \quad = \quad \int d\mathbf{q}'\, \epsilon_{ij}^{-1}(\mathbf{q}, \mathbf{q}', \omega)\, D_j(\mathbf{q}', \omega) \tag{251}$$

$$D_i(\mathbf{q}, \omega) \quad = \quad \int d\mathbf{q}'\, \epsilon_{ij}(\mathbf{q}, \mathbf{q}', \omega)\, E_j(\mathbf{q}', \omega) \tag{252}$$

From now on, we are going to omit the frequency dependency in reciprocal space to keep the equations compact, but we keep in mind that it exists implicitly and that we can add it when needed. Since the matrices are inverse to each other, they are connected by

$$\int d\mathbf{q}''\epsilon_{ij}^{-1}(\mathbf{q}, \mathbf{q}'')\epsilon_{ij}(\mathbf{q}'', \mathbf{q}') = \delta_{ij}\delta(\mathbf{q} - \mathbf{q}'). \tag{253}$$

7.4.1 The relationship between the longitudinal dielectric matrix and the longitudinal dielectric functions

In the case of pure longitudinal fields, the Gauss law relates the fields $\mathbf{E}(\mathbf{r})$, $\mathbf{D}(\mathbf{r})$, and $\mathbf{P}(\mathbf{r})$ to the corresponding charge densities n_{tot}, n_{ext}, and n_{ind} and to the electrostatic potentials V_{tot}, V_{ext}, and V_{ind} [134, 169]:

$$\begin{aligned} \nabla\,\mathbf{E}(\mathbf{r}) &= -4\pi\, n_{\mathrm{tot}}(\mathbf{r}), & \nabla\,\mathbf{D}(\mathbf{r}) &= -4\pi\, n_{\mathrm{ext}}(\mathbf{r}), & \nabla\,\mathbf{P}(\mathbf{r}) &= -4\pi\, n_{\mathrm{ind}}(\mathbf{r}) \\ \mathbf{E}(\mathbf{r}) &= -\nabla V_{\mathrm{tot}}(\mathbf{r}), & \mathbf{D}(\mathbf{r}) &= -\nabla V_{\mathrm{ext}}(\mathbf{r}), & \mathbf{P}(\mathbf{r}) &= -\nabla V_{\mathrm{ind}}(\mathbf{r}) \end{aligned}$$

Analogously to the equations of the fields in section 7.4, the total charge density and the total electrostatic potential of the material are given by

$$n_{\mathrm{tot}}(\mathbf{r}) = n_{\mathrm{ext}}(\mathbf{r}) + n_{\mathrm{ind}}(\mathbf{r}) \tag{254}$$

$$V_{\mathrm{tot}}(\mathbf{r}) = V_{ext}(\mathbf{r}) + V_{\mathrm{ind}}(\mathbf{r}). \tag{255}$$

Furthermore, the solution of the Poisson equation gives the dependencies of the electrostatic potentials from the corresponding charge densities

$$V(\mathbf{r}) = \int d\mathbf{r}'\, v(\mathbf{r}-\mathbf{r}')n(\mathbf{r}'). \tag{256}$$

In reciprocal space, this relationship is given by

$$V(\mathbf{q}) = v(\mathbf{q})n(\mathbf{q}). \tag{257}$$

Here the Coulomb kernel is given by $v(\mathbf{r}-\mathbf{r}') = 1/|\mathbf{r}-\mathbf{r}'|$ or $v(\mathbf{q}) = 4\pi/\mathbf{q}^2$, respectively.

Using the Gauss law and Eq. (250), the relationship between the inverse dielectric matrix and the inverse dielectric function can be shown in four steps. In the first step, the components of the gradients of the external and the total potentials are inserted for the components of the fields.

$$E_i(\mathbf{r}) = \int d\mathbf{r}'\epsilon^{-1}_{ij}(\mathbf{r},\mathbf{r}')D_j(\mathbf{r}') \tag{258}$$

$$\Big(\nabla_i V_{tot}(\mathbf{r})\Big) = \int d\mathbf{r}'\epsilon^{-1}_{ij}(\mathbf{r},\mathbf{r}')\Big(\nabla'_j V_{ext}(\mathbf{r}')\Big) \tag{259}$$

In the next step, the potentials are expanded into their Fourier series in order to apply the partial derivatives to the phase factors $e^{i\mathbf{qr}}$.

$$\Big(\nabla_i\Big(\int d\mathbf{q}'\, e^{-i\mathbf{q}'\mathbf{r}}V_{tot}(\mathbf{q}')\Big)\Big) = \int d\mathbf{r}'\,\epsilon^{-1}_{ij}(\mathbf{r},\mathbf{r}')\Big(\nabla'_j\Big(\int d\mathbf{q}'\, e^{-i\mathbf{q}'\mathbf{r}'}V_{ext}(\mathbf{q}')\Big)\Big) \tag{260}$$

$$\Big(\int d\mathbf{q}' e^{-i\mathbf{q}'\mathbf{r}}q'_i V_{tot}(\mathbf{q}')\Big) = \int d\mathbf{r}'\,\epsilon^{-1}_{ij}(\mathbf{r},\mathbf{r}')\Big(\int d\mathbf{q}'\, e^{-i\mathbf{q}'\mathbf{r}'}\, q'_j\, V_{ext}(\mathbf{q}')\Big) \tag{261}$$

By a proper change of the order of the integrals, we can now identify a delta function and the Fourier transformed, inverse dielectric matrix.

$$\int d\mathbf{r}\, e^{i\mathbf{q}\mathbf{r}} \left(\int d\mathbf{q}' e^{-i\mathbf{q}'\mathbf{r}} q_i' V_{tot}(\mathbf{q}') \right) = \int d\mathbf{r}\, e^{i\mathbf{q}\mathbf{r}} \left(\int d\mathbf{q}' d\mathbf{r}' e^{-i\mathbf{q}'\mathbf{r}'} \epsilon_{ij}^{-1}(\mathbf{r},\mathbf{r}') q_j' V_{ext}(\mathbf{q}') \right) \quad (262)$$

$$\int d\mathbf{q}' \underbrace{\left(\int d\mathbf{r}\, e^{i(\mathbf{q}'-\mathbf{q})\mathbf{r}} \right)}_{\delta(\mathbf{q}'-\mathbf{q})} q_i' V_{tot}(\mathbf{q}') = \int d\mathbf{q}' \underbrace{\left(\int d\mathbf{r} d\mathbf{r}' e^{i\mathbf{q}\mathbf{r}} e^{-i\mathbf{q}'\mathbf{r}'} \epsilon_{ij}^{-1}(\mathbf{r},\mathbf{r}') \right)}_{\epsilon_{ij}^{-1}(\mathbf{q},\mathbf{q}')} q_j' V_{ext}(\mathbf{q}') \quad (263)$$

If we calculate now the scalar product with $\mathbf{q}$, we can bring the components of the wave vectors to the right hand side to define the inverse dielectric function ϵ_V^{-1}.

$$q_i V_{tot}(\mathbf{q}) = \int d\mathbf{q}' q_j' \epsilon_{ij}^{-1}(\mathbf{q},\mathbf{q}') V_{ext}(\mathbf{q}') \quad (264)$$

$$V_{tot}(\mathbf{q}) = \int d\mathbf{q}' \underbrace{\frac{q_i q_j'}{\mathbf{q}^2} \epsilon_{ij}^{-1}(\mathbf{q},\mathbf{q}')}_{\epsilon_V^{-1}(\mathbf{q},\mathbf{q}')} V_{ext}(\mathbf{q}'). \quad (265)$$

To define the corresponding inverse dielectric function ϵ_n^{-1} that links the external with the total charge density, we use the solution of the Poisson equation in reciprocal space $V(\mathbf{q}) = v(\mathbf{q}) n(\mathbf{q})$.

$$v(\mathbf{q}) n_{tot}(\mathbf{q}) = \int d\mathbf{q}'\, \epsilon_V^{-1}(\mathbf{q},\mathbf{q}') v(\mathbf{q}') n_{ext}(\mathbf{q}') \quad (266)$$

$$n_{tot}(\mathbf{q}) = \int d\mathbf{q}' \underbrace{v^{-1}(\mathbf{q}) \epsilon_V^{-1}(\mathbf{q},\mathbf{q}') v(\mathbf{q}')}_{\epsilon_n^{-1}(\mathbf{q},\mathbf{q}')} n_{ext}(\mathbf{q}') \quad (267)$$

The relationships between the inverse dielectric matrices and the inverse dielectric functions are finally given by [167]

$$\epsilon_V^{-1}(\mathbf{q},\mathbf{q}') = \frac{q_i q_j'}{\mathbf{q}^2} \epsilon_{ij}^{-1}(\mathbf{q},\mathbf{q}') \quad (268)$$

$$\epsilon_n^{-1}(\mathbf{q},\mathbf{q}') = \frac{\mathbf{q}^2}{\mathbf{q}'^2} \epsilon_V^{-1}(\mathbf{q},\mathbf{q}'). \quad (269)$$

Note that, Eq. (259) to (272) would have remained valid, if we had exchanged the external perturbation and the response by starting with

$$D_i(\mathbf{r}) = \int d\mathbf{r}' \epsilon_{ij}(\mathbf{r},\mathbf{r}') E_j(\mathbf{r}'). \quad (270)$$

This would have lead to

$$\epsilon_V(\mathbf{q},\mathbf{q}') = \frac{q_i q_j'}{\mathbf{q}^2} \epsilon_{ij}(\mathbf{q},\mathbf{q}') \quad (271)$$

$$\epsilon_n(\mathbf{q},\mathbf{q}') = \frac{\mathbf{q}^2}{\mathbf{q}'^2} \epsilon_V(\mathbf{q},\mathbf{q}'). \quad (272)$$

The reason why we have derived the expressions for the inverse dielectric function first is that it describes the response to an external field for a homogeneous system, whereas the dielectric

function is not a causal response function in general [167,170–172]. Only for the limit $q \to 0$, the dielectric function can be considered as a response function as well [172], and the macroscopic dielectric function ϵ_{M} is given by [173]

$$\epsilon_{\mathrm{M}} = \lim_{\mathbf{q}\to 0}\lim_{\mathbf{q}'\to 0} \epsilon_V(\mathbf{q},\mathbf{q}') = 1/\lim_{q\to 0}\lim_{q'\to 0} \epsilon_V^{-1}(\mathbf{q},\mathbf{q}'). \tag{273}$$

7.4.2 The relationship between the longitudinal dielectric functions and the reducible and irreducible polarizability

Previous dielectric functions have only related similar properties, i.e. external with internal densities and external with internal potentials. In contrast, the reducible polarizability $\chi = \delta n_{\mathrm{ind}}/\delta V_{\mathrm{ext}}$ and the irreducible polarizability $\tilde{\chi} = \delta n_{\mathrm{ind}}/\delta V_{\mathrm{tot}}$ relate densities to potentials. In principle, they describe how the induced charge density n_{ind} changes with respect to the electrostatic potentials V_{ext} and V_{tot}, respectively [166,174]. If we Fourier transform both equations (see section 7.1), apply the deconvolution theorem with respect to $t - t'$ (see section 7.1.2), and omit the frequency dependency (compare section 7.4), we obtain the compact formulas

$$n_{\mathrm{ind}}(\mathbf{q}) = \int d\mathbf{q}'\, \chi(\mathbf{q},\mathbf{q}') V_{\mathrm{ext}}(\mathbf{q}') \tag{274}$$

$$n_{\mathrm{ind}}(\mathbf{q}) = \int d\mathbf{q}'\, \tilde{\chi}(\mathbf{q},\mathbf{q}') V_{\mathrm{tot}}(\mathbf{q}). \tag{275}$$

Both polarizabilities can in principle be used to calculate the longitudinal dielectric functions. Let us first consider the case of the reducible polarizability. Since we know that the total density consists of the external and the induced density we can write

$$n_{\mathrm{tot}}(\mathbf{q}) = n_{\mathrm{ext}}(\mathbf{q}) + n_{\mathrm{ind}}(\mathbf{q}) \tag{276}$$

$$n_{\mathrm{tot}}(\mathbf{q}) = n_{\mathrm{ext}}(\mathbf{q}) + \int d\mathbf{q}'\, \chi(\mathbf{q},\mathbf{q}') V_{\mathrm{ext}}(\mathbf{q}'). \tag{277}$$

Then the solution of the Poisson equation $V_{\mathrm{ext}}(\mathbf{q}) = v(\mathbf{q}) n_{\mathrm{ext}}(\mathbf{q})$ can be insert to identify the inverse dielectric function that connects the total with the external charge density.

$$n_{\mathrm{tot}}(\mathbf{q}) = n_{\mathrm{ext}}(\mathbf{q}) + \int d\mathbf{q}'\, \chi(\mathbf{q},\mathbf{q}') v(\mathbf{q}') n_{\mathrm{ext}}(\mathbf{q}') \tag{278}$$

$$n_{\mathrm{tot}}(\mathbf{q}) = \int d\mathbf{q}' \underbrace{\Big(\delta(\mathbf{q}-\mathbf{q}') + \chi(\mathbf{q},\mathbf{q}') v(\mathbf{q}')\Big)}_{\epsilon_n^{-1}(\mathbf{q},\mathbf{q}')} n_{\mathrm{ext}}(\mathbf{q}') \tag{279}$$

Furthermore, using $\epsilon_V^{-1}(\mathbf{q},\mathbf{q}') = v(\mathbf{q}) v^{-1}(\mathbf{q}') \epsilon_n^{-1}(\mathbf{q},\mathbf{q}')$, we obtain

$$\epsilon_V^{-1}(\mathbf{q},\mathbf{q}') = v(\mathbf{q}) v^{-1}(\mathbf{q}') \Big(\delta(\mathbf{q}-\mathbf{q}') + \chi(\mathbf{q},\mathbf{q}') v(\mathbf{q}')\Big) \tag{280}$$

$$\epsilon_V^{-1}(\mathbf{q},\mathbf{q}') = \delta(\mathbf{q}-\mathbf{q}') + v(\mathbf{q}) \chi(\mathbf{q},\mathbf{q}'). \tag{281}$$

Finally, we can analogously derive the relationship between the longitudinal dielectric functions and the irreducible polarizability.

$$\epsilon_n(\mathbf{q},\mathbf{q}') = \delta(\mathbf{q}-\mathbf{q}') - \tilde{\chi}(\mathbf{q},\mathbf{q}')v(\mathbf{q}') \tag{282}$$

$$\epsilon_V(\mathbf{q},\mathbf{q}') = \delta(\mathbf{q}-\mathbf{q}') - v(\mathbf{q})\tilde{\chi}(\mathbf{q},\mathbf{q}'). \tag{283}$$

7.4.3 The relationship between the reducible and the irreducible polarizability

Since the dielectric functions ϵ_V and ϵ_V^{-1} are inverse, it is possible to find a connection between the reducible and the irreducible polarizability. Therefore, we insert the formulas of the dielectric functions into the inverse relationship (see Eq. 253).

$$\delta(\mathbf{q}-\mathbf{q}'') = \int d\mathbf{q}'\epsilon^{-1}(\mathbf{q},\mathbf{q}')\epsilon(\mathbf{q}',\mathbf{q}'') \tag{284}$$

$$\delta(\mathbf{q}-\mathbf{q}'') = \int d\mathbf{q}'\Big(\delta(\mathbf{q}-\mathbf{q}') + v(\mathbf{q})\chi(\mathbf{q},\mathbf{q}')\Big)\Big(\delta(\mathbf{q}'-\mathbf{q}'') - v(\mathbf{q}')\tilde{\chi}(\mathbf{q}',\mathbf{q}'')\Big) \tag{285}$$

The expansion of the product then simply yields

$$\delta(\mathbf{q}-\mathbf{q}'') = \delta(\mathbf{q}-\mathbf{q}'') \tag{286}$$

$$+ \; v(\mathbf{q})\chi(\mathbf{q},\mathbf{q}'') - v(\mathbf{q})\tilde{\chi}(\mathbf{q},\mathbf{q}'') - \int d\mathbf{q}'\tilde{V}(\mathbf{q})\chi(\mathbf{q},\mathbf{q}')v(\mathbf{q}')\tilde{\chi}(\mathbf{q}',\mathbf{q}'') \tag{287}$$

$$0 = v(\mathbf{q})\Big(\chi(\mathbf{q},\mathbf{q}'') - \tilde{\chi}(\mathbf{q},\mathbf{q}'') - \int d\mathbf{q}'\chi(\mathbf{q},\mathbf{q}')v(\mathbf{q}')\tilde{\chi}(\mathbf{q}',\mathbf{q}'')\Big) \tag{288}$$

. Finally, we obtain a "Dyson-like" equation that relates both polarizabilities with each other.

$$\chi(\mathbf{q},\mathbf{q}'') = \tilde{\chi}(\mathbf{q},\mathbf{q}'') + \int d\mathbf{q}'\chi(\mathbf{q},\mathbf{q}')v(\mathbf{q}')\tilde{\chi}(\mathbf{q}',\mathbf{q}'') \tag{289}$$

7.4.4 The Kohn-Sham (KS) single particle polarizability and the random phase approximation (RPA)

Neither the reducible polarizability χ nor the irreducible polarizability $\tilde{\chi}$ are exactly computable in our calculations. We can however compute the change in the charge density due to changes in the effective KS potential and then the single particle KS polarizability χ^{KS}. The KS polarizability links the variation of an effective KS potential V_{eff} with the induced density n_{ind} [166,168,174].

$$n_{\mathrm{ind}}(\mathbf{r},t) = \int d\mathbf{r}'\,dt'\chi(\mathbf{r},\mathbf{r}',t-t')V_{\mathrm{ext}}(\mathbf{r}',t') \tag{290}$$

$$n_{\mathrm{ind}}(\mathbf{r},t) = \int d\mathbf{r}'\,dt'\chi^{\mathrm{KS}}(\mathbf{r},\mathbf{r}',t-t')V_{\mathrm{eff}}(\mathbf{r}',t'). \tag{291}$$

The effective potential commonly consists of the external potential V_{ext}, the Hartree potential V_{H}, and the exchange correlation potential V_{xc}. Both equations give, in principle, the same induced electronic density, and we can relate both response functions by

$$\int d\mathbf{r}'\chi(\mathbf{r},\mathbf{r}',\omega)V_{\mathrm{ext}}(\mathbf{r}',\omega) = \int d\mathbf{r}'\chi^{\mathrm{KS}}(\mathbf{r},\mathbf{r}',\omega)\Big(V_{\mathrm{ext}}(\mathbf{r}',\omega) + V_{\mathrm{H}}(\mathbf{r}',\omega) + V_{\mathrm{xc}}(\mathbf{r}',\omega)\Big). \tag{292}$$

Since we can also define response functions $\frac{\partial V_{\rm H}(\mathbf{r},t)}{\partial V_{\rm ext}(\mathbf{r}',t')}$ and $\frac{\partial V_{\rm xc}(\mathbf{r},t)}{\partial V_{\rm ext}(\mathbf{r}',t')}$, which link the Hartree potential and the exchange correlation potential with the external potential, we can express Eq. (292) in terms of the external potential only.

$$V_{\rm H}(\mathbf{r},t) = \int d\mathbf{r}'d\mathbf{r}''dt'dt'' \underbrace{\frac{\partial V_{\rm H}(\mathbf{r},t)}{\partial n(\mathbf{r}',t')}}_{v(\mathbf{r}-\mathbf{r}')\delta(t-t')} \underbrace{\frac{\partial n(\mathbf{r}',t')}{\partial V_{\rm ext}(\mathbf{r}'',t'')}}_{\chi(\mathbf{r}',\mathbf{r}'',t'-t'')} V_{\rm ext}(\mathbf{r}'',t'') \tag{293}$$

$$V_{\rm xc}(\mathbf{r},t) = \int d\mathbf{r}'d\mathbf{r}''dt'dt'' \underbrace{\frac{\partial V_{\rm xc}(\mathbf{r},t)}{\partial n(\mathbf{r}',t')}}_{f_{\rm xc}(\mathbf{r},\mathbf{r}',t-t')} \underbrace{\frac{\partial n(\mathbf{r}',t')}{\partial V_{\rm ext}(\mathbf{r}'',t'')}}_{\chi(\mathbf{r}',\mathbf{r}'',t'-t'')} V_{\rm ext}(\mathbf{r}'',t'') \tag{294}$$

Here we used the chain rule of functional derivatives to identify the Coulomb kernel $v(\mathbf{r}-\mathbf{r}') = 1/|\mathbf{r}-\mathbf{r}'|$ and the reducible polarizability χ. Furthermore, we introduced the exchange correlation response function $f_{\rm xc}$ [166, 175].

If we then plug these identities in Eq. (292) and perform a Fourier transformation, we find the relationship between the reducible polarizability and the KS polarizability.

$$\chi(\mathbf{r},\mathbf{r}',\omega) = \chi^{\rm KS}(\mathbf{r},\mathbf{r}',\omega) \tag{295}$$

$$+ \int d\mathbf{r}''d\mathbf{r}'''\chi^{\rm KS}(\mathbf{r},\mathbf{r}',\omega)v(\mathbf{r}'-\mathbf{r}'')\chi(\mathbf{r}'',\mathbf{r}''',\omega) \tag{296}$$

$$+ \int d\mathbf{r}''d\mathbf{r}'''\chi^{\rm KS}(\mathbf{r},\mathbf{r}',\omega)f_{\rm xc}(\mathbf{r}',\mathbf{r}'',\omega)\chi(\mathbf{r}'',\mathbf{r}''',\omega) \tag{297}$$

We can further simplify this equation by neglecting the exchange correlation kernel $f_{\rm xc}$ to obtain the polarizability of the RPA [176].

$$\chi^{\rm RPA}(\mathbf{r},\mathbf{r}',\omega) = \chi^{\rm KS}(\mathbf{r},\mathbf{r}',\omega) + \int d\mathbf{r}''d\mathbf{r}'''\chi^{\rm KS}(\mathbf{r},\mathbf{r}',\omega)v(\mathbf{r}'-\mathbf{r}'')\chi^{\rm RPA}(\mathbf{r}'',\mathbf{r}''',\omega) \tag{298}$$

Finally, the KS response function can be expressed in terms of Green's functions and, within the Lehman representation, we have [174]

$$\chi^{\rm KS}(\mathbf{r},\mathbf{r}',\omega) = \lim_{\eta\to 0}\sum_{ij}(f_i - f_j)\left(\frac{\psi_i^*(\mathbf{r})\psi_j(\mathbf{r})\psi_i(\mathbf{r}')\psi_j^*(\mathbf{r}')}{\omega-(\epsilon_j-\epsilon_i)+i\eta} - \frac{\psi_i(\mathbf{r})\psi_j^*(\mathbf{r})\psi_j(\mathbf{r}')\psi_i^*(\mathbf{r}')}{\omega-(\epsilon_j-\epsilon_i)+i\eta}\right). \tag{299}$$

Here $\psi_i(\mathbf{r})$ and $\psi_j(\mathbf{r})$ are the one particle KS orbitals, and ϵ_i and ϵ_j are the corresponding eigenvalues. Note that exchange correlation effects are included in the KS orbitals.

7.4.5 The calculation of the longitudinal dielectric function applying the RPA

The longitudinal inverse dielectric function in the RPA is straightforwardly given by

$$\epsilon_V^{-1}(\mathbf{q},\mathbf{q}') = \delta(\mathbf{q},\mathbf{q}') + v(\mathbf{q})\chi^{\rm RPA}(\mathbf{q},\mathbf{q}'). \tag{300}$$

Demanding that the corresponding inverse function exists and that it is related to the irreducible polarizability, the reducible and the irreducible polarizability are related in the usual way, compare Eq. (289) in section 7.4.3. For this case, the comparison with the reducible polarizability, see Eq. (298), suggests a close connection between the irreducible polarizability and the KS response function.

$$\chi^{\mathrm{RPA}}(\mathbf{q},\mathbf{q}'') = \tilde{\chi}(\mathbf{q},\mathbf{q}'') + \int d\mathbf{q}'\chi^{\mathrm{RPA}}(\mathbf{q},\mathbf{q}')v(\mathbf{q}')\tilde{\chi}(\mathbf{q}',\mathbf{q}'') \tag{301}$$

$$\chi^{\mathrm{RPA}}(\mathbf{q},\mathbf{q}'') = \chi^{\mathrm{KS}}(\mathbf{q},\mathbf{q}'') + \int d\mathbf{q}'\chi^{\mathrm{KS}}(\mathbf{q},\mathbf{q}')v(\mathbf{q}')\chi^{\mathrm{RPA}}(\mathbf{q}',\mathbf{q}'') \tag{302}$$

Specifically, we can multiply the last equation by $(\chi^{\mathrm{RPA}})^{-1}$ from the right hand side and by $(\chi^{\mathrm{KS}})^{-1}$ from the left hand side to obtain

$$\delta(\mathbf{q}-\mathbf{k}'') = \int d\mathbf{q}''\,\chi^{\mathrm{KS}}(\mathbf{q},\mathbf{q}'')(\chi^{\mathrm{RPA}})^{-1}(\mathbf{q}'',\mathbf{k}'') + \chi^{\mathrm{KS}}(\mathbf{q},\mathbf{k}'')v(\mathbf{k}'') \tag{303}$$

$$(\chi^{\mathrm{KS}})^{-1}(\mathbf{k},\mathbf{k}'') = (\chi^{\mathrm{RPA}})^{-1}(\mathbf{k},\mathbf{k}'') + \delta(\mathbf{k}-\mathbf{k}'')v(\mathbf{k}''). \tag{304}$$

If we continue to multiply the last result by χ^{KS} from the right hand side and χ^{RPA} from the left hand side, the equations can be written as

$$\delta(\mathbf{k}-\mathbf{q}'') = \int d\mathbf{k}''\,(\chi^{\mathrm{RPA}})^{-1}(\mathbf{k},\mathbf{k}'')\chi^{\mathrm{KS}}(\mathbf{k}'',\mathbf{q}'') + v(\mathbf{k})\chi^{\mathrm{KS}}(\mathbf{k},\mathbf{q}'') \tag{305}$$

$$\chi^{\mathrm{RPA}}(\mathbf{q},\mathbf{q}'') = \chi^{\mathrm{KS}}(\mathbf{q},\mathbf{q}'') + \int d\mathbf{k}\,\chi^{\mathrm{RPA}}(\mathbf{q},\mathbf{k})v(\mathbf{k})\chi^{\mathrm{KS}}(\mathbf{k},\mathbf{q}''),. \tag{306}$$

Comparing Eq. 302 with Eq. 306, we eventually see that the order of χ^{KS} and χ^{RPA} in the integration is exchangeable and that Eq. 301 has exactly the same form as Eq. 306. For this reason, we can approximate the irreducible polarizability $\tilde{\chi}$ by the KS response function χ^{KS} and the longitudinal dielectric function in Eq. 283 by

$$\epsilon_V(\mathbf{q},\mathbf{q}') = \delta(\mathbf{q},\mathbf{q}') - v(\mathbf{q})\chi^{\mathrm{KS}}(\mathbf{q},\mathbf{q}'). \tag{307}$$

References

[1] A. Luque and S. Hegedus. *Handbook of photovoltaic science and engineering*. Wiley, 2003.

[2] A. Goetzberger, C. Hebling, and H.W. Schock. Photovoltaic materials, history, status and outlook. *Mater. Sci. Eng. R*, 40(1):1 – 46, 2003.

[3] M.A. Green, K. Emery, Y. Hishikawa, W. Warta, and E.D. Dunlop. Solar cell efficiency tables (version 39). *Prog. Photovoltaics: research and applications*, 20(1):12–20, 2012.

[4] T. Markvart and L. Castaner. *Solar cells: materials, manufacture and operation*. Elsevier Science & Technology, 2005.

[5] N.W. Ashcroft and N.D. Mermin. *Solid state physics*. Brooks Cole, 1976.

[6] M.A. Green. *Solar cells: operating principles, technology, and system applications*. Prentice Hall, 1981.

[7] W. Shockley and H.J. Queisser. Detailed balance limit of efficiency of p-n junction solar cells. *J. Appl. Phys.*, 32(3):510–519, 1961.

[8] H.K. Raut, V.A. Ganesh, A.S. Nair, and S. Ramakrishna. Anti-reflective coatings: a critical, in-depth review. *Energy Environ. Sci.*, 4:3779–3804, 2011.

[9] B.S. Richards. Comparison of TiO_2 and other dielectric coatings for buried-contact solar cells: a review. *Prog. Photovoltaics: research and applications*, 12(4):253–281, 2004.

[10] S. J. Fonash. *Solar cell device physics*. Academic Press, 2009.

[11] A.G. Aberle. Surface passivation of crystalline silicon solar cells: a review. *Prog. Photovoltaics: research and applications*, 8(5):473–487, 2000.

[12] R.H. Bube. *Fundamentals of solar cells: photovoltaic solar energy conversion*. Academic Press, 1983.

[13] W. Shockley and W.T. Read. Statistics of the recombinations of holes and electrons. *Phys. Rev.*, 87:835–842, 1952.

[14] R.N. Hall. Electron-hole recombination in germanium. *Phys. Rev.*, 87:387–387, 1952.

[15] A. McEvoy, T. Markvart, L. Castaner, T. Markvart, and L. Castaner. *Practical handbook of photovoltaics: fundamentals and applications*. Elsevier Science, 2003.

[16] A.B. Sproul. Dimensionless solution of the equation describing the effect of surface recombination on carrier decay in semiconductors. *J. of Appl. Phys.*, 76(5):2851–2854, 1994.

[17] J. Schmidt and A.G. Aberle. Carrier recombination at silicon–silicon nitride interfaces fabricated by plasma-enhanced chemical vapor deposition. *J. Appl. Phys.*, 85(7):3626–3633, 1999.

[18] B. Sopori. Silicon nitride processing for control of optical and electronic properties of silicon solar cells. *J. Electron. Mater.*, 32(10):1034–1042, 2003.

[19] W. Soppe, H. Rieffe, and A. Weeber. Bulk and surface passivation of silicon solar cells accomplished by silicon nitride deposited on industrial scale by microwave PECVD. *Prog. Photovoltaics: research and applications*, 13(7):551–569, 2005.

[20] A.G. Aberle. Overview on SiN surface passivation of crystalline silicon solar cells. *Sol. Energy Mater. Sol. Cells*, 65(14):239 – 248, 2001.

[21] S. Dauwe, J. Schmidt, A. Metz, and R. Hezel. Fixed charge density in silicon nitride films on crystalline silicon surfaces under illumination. In *Photovoltaic Specialists Conference, 2002. Conference Record of the Twenty-Ninth IEEE*, pages 162–165, 2002.

[22] M.W.P.E. Lamers, K.T. Butler, P.E. Vullum, J.H. Harding, and A. Weeber. Characterization of a-SiN_x:H layer: bulk properties, interface with Si and solar cell efficiency. *Phys. Status Solidi (a)*, 210(4):658–668, 2013.

[23] F. Giorgis, C.F. Pirri, and E. Tresso. Structural properties of a-$Si_{1-x}N_x$:H films grown by plasma enhanced chemical vapour deposition by SiH_4 + NH_3 + H_2 gas mixtures. *Thin Solid Films*, 307(12):298 – 305, 1997.

[24] F. Giorgis, F. Giuliani, C.F. Pirri, E. Tresso, C. Summonte, R. Rizzoli, R. Galloni, A. Desalvo, and P. Rava. Optical, structural and electrical properties of device-quality hydrogenated amorphous silicon-nitrogen films deposited by plasma-enhanced chemical vapour deposition. *Philos. Mag. B*, 77(4):925–944, 1998.

[25] F. Duerinckx and J. Szlufcik. Defect passivation of industrial multicrystalline solar cells based on PECVD silicon nitride. *Sol. Energy Mater. Sol. Cells*, 72(14):231 – 246, 2002.

[26] C.J. Mogab and E. Lugujjo. Backscattering analysis of the composition of silicon-nitride films deposited by rf reactive sputtering. *J. Appl. Phys.*, 47(4):1302–1309, 1976.

[27] M.M. Guraya, H. Ascolani, G. Zampieri, J.I. Cisneros, J.H. Dias da Silva, and M.P. Cantão. Bond densities and electronic structure of amorphous SiN_x:H. *Phys. Rev. B*, 42(9):5677–5684, 1990.

[28] T. Aiyama, T. Fukunaga, K. Niihara, T. Hirai, and K. Suzuki. An x-ray diffraction study of the amorphous structure of chemically vapor-deposited silicon nitride. *J. Non-Cryst. Solids*, 33(2):131 – 139, 1979.

[29] M. Misawa, T. Fukunaga, K. Niihara, T. Hirai, and K. Suzuki. Structure characterization of CVD amorphous Si_3N_4 by pulsed neutron total scattering. *J. Non-Cryst. Solids*, 34(3):313 – 321, 1979.

[30] M.M. Guraya, H. Ascolani, G. Zampieri, J.H. Dias da Silva, M.P. Cantão, and J.I. Cisneros. Electronic structure of amorphous Si-N compounds. *Phys. Rev. B*, 49:13446–13451, 1994.

[31] F. Giorgis, C. Vinegoni, and L. Pavesi. Optical absorption and photoluminescence properties of a-$Si_{1-x}N_x$:H films deposited by plasma-enhanced CVD. *Phys. Rev. B*, 61:4693–4698, 2000.

[32] P.M. Lenahan and S.E. Curry. First observation of the ^{29} Si hyperfine spectra of silicon dangling bond centers in silicon nitride. *Appl. Phys. Lett.*, 56(2):157–159, 1990.

[33] W.L. Warren, P.M. Lenahan, and S.E. Curry. First observation of paramagnetic nitrogen dangling-bond centers in silicon nitride. *Phys. Rev. Lett.*, 65:207–210, 1990.

[34] J. Robertson, W.L. Warren, and J. Kanicki. Nature of the Si and N dangling bonds in silicon nitride. *J. Non-Cryst. Solids*, 187(0):297 – 300, 1995.

[35] A.G. Aberle and R. Hezel. Progress in low-temperature surface passivation of silicon solar cells using remote-plasma silicon nitride. *Prog. Photovoltaics: research and applications*, 5(1):29–50, 1997.

[36] J. Robertson. Defects and hydrogen in amorphous silicon nitride. *Philos. Mag. B*, 69(2):307–326, 1994.

[37] R. Tsu, D. Martin, J. Gonzalez-Hernandez, and S.R. Ovshinsky. Passivation of dangling bonds in amorphous Si and Ge by gas adsorption. *Phys. Rev. B*, 35:2385–2390, 1987.

[38] J. Yoo, S.K. Dhungel, and J. Yi. Annealing optimization of silicon nitride film for solar cell application. *Thin Solid Films*, 515(19):7611 – 7614, 2007.

[39] M. De Graef and M.E. McHenry. *Structure of materials: an introduction to crystallography, diffraction and symmetry* . Cambridge University Press, 2007.

[40] D.B. Williams and C.B. Carter. *Transmission electron microscopy: a textbook for materials science*. Springer, 2009.

[41] P. Eaton and P. West. *Atomic force microscopy* . Oxford University Press, 2010.

[42] R. Kärcher, L. Ley, and R.L. Johnson. Electronic structure of hydrogenated and unhydrogenated amorphous SiN_x ($0 \leq x \leq 1.6$): a photoemission study. *Phys. Rev. B*, 30:1896–1910, 1984.

[43] A. Iqbal, W.B. Jackson, C.C. Tsai, J.W. Allen, and C.W. Bates. Electronic structure of silicon nitride and amorphous silicon/silicon nitride band offsets by electron spectroscopy. *J. Appl. Phys.*, 61(8):2947 –2954, 1987.

[44] L. Ouyang and W.Y. Ching. Systematic approach to generate near-perfect periodic continuous random network models: application to amorphous Si_3N_4. *Phys. Rev. B*, 54:R15594–R15597, 1996.

[45] E.C. Ferreira and C.E.T. Gonçalves da Silva. Electronic structure of amorphous Si_3N_4 in the cluster-Bethe-lattice approximation. *Phys. Rev. B*, 32:8332–8337, 1985.

[46] S.Y. Lin. Electronic defect states of amorphous silicon nitride. *Opt. Mater.*, 23(12):93 – 98, 2003.

[47] P. Kroll. Structure and reactivity of amorphous silicon nitride investigated with density-functional methods. *J. Non-Cryst. Solids*, 293295(0):238 – 243, 2001.

[48] S.Zh. Karazhanov, P. Kroll, A. Holt, A. Bentzen, and A. Ulyashin. Comparative analysis of electronic structure and optical properties of crystalline and amorphous silicon nitrides. *J. Appl. Phys.*, 106(5):053717–053717–6, 2009.

[49] F. De Brito Mota, J.F. Justo, and A. Fazzio. Structural and electronic properties of silicon nitride materials. *Int. J. Quantum Chem.*, 70(4-5):973–980, 1998.

[50] F. de Brito Mota, J.F. Justo, and A. Fazzio. Hydrogen role on the properties of amorphous silicon nitride. *J. Appl. Phys.*, 86(4):1843 –1847, 1999.

[51] J. F. Justo, F. de Brito Mota, and A. Fazzio. First-principles investigation of a-SiN_x:H. *Phys. Rev. B*, 65:073202, 2002.

[52] N. Umesaki, N. Hirosaki, and K. Hirao. Structural characterization of amorphous silicon nitride by molecular dynamics simulation. *J. Non-Cryst. Solids*, 150(13):120 – 125, 1992.

[53] F. Alvarez and A.A. Valladares. First-principles simulations of atomic networks and optical properties of amorphous SiN_x alloys. *Phys. Rev. B*, 68:205203, 2003.

[54] L. Giacomazzi and P. Umari. First-principles investigation of electronic, structural, and vibrational properties of a-Si_3N_4. *Phys. Rev. B*, 80(14), 2009.

[55] K. Jarolimek, R.A. de Groot, G.A. de Wijs, and M. Zeman. Atomistic models of hydrogenated amorphous silicon nitride from first principles. *Phys. Rev. B*, 82:205201, 2010.

[56] S. Nekrashevich, A. Shaposhnikov, and V. Gritsenko. Study of the atomic and electronic structures of amorphous silicon nitride and defects in it. *JETP Letters*, 94:202–205, 2011.

[57] M. Ippolito and S. Meloni. Atomistic structure of amorphous silicon nitride from classical molecular dynamics simulations. *Phys. Rev. B*, 83:165209, 2011.

[58] R.P. Vedula, N.L. Anderson, and A. Strachan. Effect of topological disorder on structural, mechanical, and electronic properties of amorphous silicon nitride: an atomistic study. *Phys. Rev. B*, 85:205209, 2012.

[59] J. Robertson and M.J. Powell. Gap states in silicon nitride. *Appl. Phys. Lett.*, 44(4):415 –417, 1984.

[60] A. Omeltchenko, M.E. Bachlechner, A. Nakano, R.K. Kalia, P. Vashishta, I. Ebbsjö, A. Madhukar, and P. Messina. Stress Domains in Si(111)/a-Si_3N_4 nanopixel: ten-million-atom molecular dynamics simulations on parallel computers. *Phys. Rev. Lett.*, 84:318–321, Jan 2000.

[61] M. Yang, R.Q. Wu, W.S. Deng, L. Shen, Z.D. Sha, Y.Q. Cai, Y.P. Feng, and S.J. Wang. Electronic structures of β-Si_3N_4(0001)/Si(111) interfaces: perfect bonding and dangling bond effects. *J. of Appl. Phys.*, 105(2):–, 2009.

[62] K.T. Butler, M.P.W.E. Lamers, A.W. Weeber, and J.H. Harding. Molecular dynamics studies of the bonding properties of amorphous silicon nitride coatings on crystalline silicon. *J. Appl. Phys.*, 110(12):124905, 2011.

[63] K.T. Butler, J.H. Harding, M.P.W.E. Lamers, and A.W. Weeber. Stoichiometrically graded sinx for improved surface passivation in high performance solar cells. *J. of Appl. Phys.*, 112(9):–, 2012.

[64] Flage˘Larsen, E. and Løvvik, O.M. and Fang, C.M. and Kresse, G. β-Si_3N_4(0001)/Si(111) interface: Phosphorus defects, valence band offsets, and their role of passivating the interface states. *Phys. Rev. B*, 88:165310, Oct 2013.

[65] M.W.P.E. Lamers, L.E. Hintzsche, K.T. Butler, P.E. Vullum, C.M. Fang, M. Marsman, G. Jordan, J.H. Harding, G. Kresse, and A. Weeber. The interface of a-SiN_x:H and Si: linking the nano-scale structure to passivation quality. *Sol. Energy Mater. Sol. Cells*, 2013.

[66] T. Anh Pham, T. Li, H.V. Nguyen, S. Shankar, F. Gygi, and G. Galli. Band offsets and dielectric properties of the amorphous Si_3N_4/Si(100) interface: a first-principles study. *Applied Physics Letters*, 102(24):–, 2013.

[67] C. Cohen-Tannoudji, B. Diu, and F. Laloe. *Quantum mechanics: volume 1*. John Wiley and Sons, 1977.

[68] R.G. Parr and Y. Weitao. *Density functional theory of atoms and molecules*. Oxford University Press, 1989.

[69] M. Born and R. Oppenheimer. Zur Quantentheorie der Molekeln. *Annalen der Physik*, 389(20):457–484, 1927.

[70] J. Thijssen. *Computational physics*. Cambridge University Press, 2007.

[71] J.C. Slater. The theory of complex spectra. *Phys. Rev.*, 34:1293–1322, Nov 1929.

[72] V. Fock. Näherungsmethode zur Lösung des quantenmechanischen Mehrkörperproblems. *Zeitschrift für Physik*, 61(1-2):126–148, 1930.

[73] A. Szabo and N.S. Ostlund. *Modern quantum chemistry*. Dover Pubn Inc, 1996.

[74] J. Kohanoff. *Electronic structure calculations for solids and molecules*. Cambridge University Press, 2006.

[75] T. Koopmans. Über die Zuordnung von Wellenfunktionen und Eigenwerten zu den einzelnen Elektronen eines Atoms. *Physica*, 1(16):104 – 113, 1934.

[76] P. Hohenberg and W. Kohn. Inhomogeneous electron gas. *Phys. Rev.*, 136:B864–B871, 1964.

[77] W. Kohn and L.J. Sham. Self-consistent equations including exchange and correlation effects. *Phys. Rev.*, 140:A1133–A1138, 1965.

[78] J.P. Perdew, R.G. Parr, M. Levy, and J.L. Balduz. Density-functional theory for fractional particle number: derivative discontinuities of the energy. *Phys. Rev. Lett.*, 49:1691–1694, 1982.

[79] J.P. Perdew and M. Levy. Physical content of the exact Kohn-Sham orbital energies: band gaps and derivative discontinuities. *Phys. Rev. Lett.*, 51:1884–1887, Nov 1983.

[80] J.P. Perdew and M. Levy. Comment on "Significance of the highest occupied Kohn-Sham eigenvalue". *Phys. Rev. B*, 56:16021–16028, 1997.

[81] L.J. Sham and M. Schlüter. Density functional theory of the energy gap. *Phys. Rev. Lett.*, 51:1888–1891, 1983.

[82] P.A.M. Dirac. Note on exchange phenomena in the thomas atom. *Math. Proc. Cambridge*, 26:376–385, 1930.

[83] D.M. Ceperley and B.J. Alder. Ground state of the electron gas by a stochastic method. *Phys. Rev. Lett.*, 45:566–569, 1980.

[84] S.H. Vosko, L. Wilk, and M. Nusair. Accurate spin-dependent electron liquid correlation energies for local spin density calculations: a critical analysis. *Can. J. Phys.*, 58(8):1200–1211, 1980.

[85] J. P. Perdew and A. Zunger. Self-interaction correction to density-functional approximations for many-electron systems. *Phys. Rev. B*, 23:5048–5079, 1981.

[86] O. Gunnarsson and B.I. Lundqvist. Exchange and correlation in atoms, molecules, and solids by the spin-density-functional formalism. *Phys. Rev. B*, 13:4274–4298, 1976.

[87] J.P. Perdew, K. Burke, and Y. Wang. Generalized gradient approximation for the exchange-correlation hole of a many-electron system. *Phys. Rev. B*, 54:16533–16539, 1996.

[88] J.P. Perdew, K. Burke, and M. Ernzerhof. Generalized gradient approximation made simple. *Phys. Rev. Lett.*, 77:3865–3868, 1996.

[89] J.P. Perdew, A. Ruzsinszky, G.I. Csonka, O.A. Vydrov, G.E. Scuseria, L.A. Constantin, X. Zhou, and K. Burke. Restoring the density-gradient expansion for exchange in solids and surfaces. *Phys. Rev. Lett.*, 100:136406, 2008.

[90] G.I. Csonka, J.P. Perdew, A. Ruzsinszky, P.H.T. Philipsen, S. Lebègue, J. Paier, O.A. Vydrov, and J.G. Ángyán. Assessing the performance of recent density functionals for bulk solids. *Phys. Rev. B*, 79:155107, 2009.

[91] J. Paier, M. Marsman, K. Hummer, and G. Kresse. Screened hybrid density functionals applied to solids. *J. Chem. Phys.*, 124:154709, 2006.

[92] C. Adamo and V. Barone. Toward reliable density functional methods without adjustable parameters: the PBE0 model. *J. Chem. Phys.*, 110:6158, 1999.

[93] J. Heyd, G.E. Scuseria, and M. Ernzerhof. Hybrid functionals based on a screened Coulomb potential. *J. Chem. Phys.*, 118:8207, 2003.

[94] R.D. Adamson, J.P. Dombroski, and P.M.W. Gill. Chemistry without Coulomb tails. *Chem. Phys. Lett.*, 254(56):329 – 336, 1996.

[95] P.M.W. Gill and R.D. Adamson. A family of attenuated Coulomb operators. *Chem. Phys. Lett.*, 261(12):105 – 110, 1996.

[96] T. Leininger, H. Stoll, H.J. Werner, and A. Savin. Combining long-range configuration interaction with short-range density functionals. *Chem. Phys. Lett.*, 275(34):151 – 160, 1997.

[97] M. Ernzerhof and J.P. Perdew. Generalized gradient approximation to the angle- and system-averaged exchange hole. *J. Chem. Phys.*, 109:3313, 1998.

[98] G. Rickayzen. *Green's functions and condensed matter.* Academic Press, 1980.

[99] F. Aryasetiawan and O. Gunnarsson. The GW method. *Rep. Prog. Phys.*, 61(3):237, 1998.

[100] D.N. Zubarev. Double-time Green functions in statistical physics. *Sov. Phys. Usp.*, 3(3):320, 1960.

[101] J.W. Negele and H. Orland. *Quantum many particle systems.* Westview Press, 1998.

[102] L. Hedin. New method for calculating the one-particle Green's function with application to the electron-gas problem. *Phys. Rev.*, 139:A796–A823, 1965.

[103] L. Hedin, B.I. Lundqvist, and S. Lundqvist. Beyond the one-electron approximation: density of states for interacting electrons. *J. Res. Nat. Bur. Stand. Sec. A: Phys. Ch.*, 74A(3):417, 1970.

[104] L. Hedin. On correlation effects in electron spectroscopies and the GW approximation. *J. Phys.: Condens. Matter*, 11(42):R489, 1999.

[105] W.G. Aulbur, L. Jnsson, and J.W. Wilkins. Quasiparticle calculations in solids. volume 54 of *Solid State Phys.*, pages 1 – 218. Academic Press, 1999.

[106] W. von der Linden and P. Horsch. Precise quasiparticle energies and hartree-fock bands of semiconductors and insulators. *Phys. Rev. B*, 37:8351–8362, 1988.

[107] G. Onida, L. Reining, and A. Rubio. Electronic excitations: density-functional versus many-body Green's-function approaches. *Rev. Mod. Phys.*, 74:601–659, 2002.

[108] M. Shishkin and G. Kresse. Implementation and performance of the frequency-dependent GW method within the PAW framework. *Phys. Rev. B*, 74:035101, 2006.

[109] M. Shishkin and G. Kresse. Self-consistent GW calculations for semiconductors and insulators. *Phys. Rev. B*, 75:235102, 2007.

[110] M.P. Allen and D.J. Tildesley. *Computer simulation of liquids.* Oxford University Press, 1989.

[111] M.C. Payne, M.P. Teter, D.C. Allan, T.A. Arias, and J.D. Joannopoulos. Iterative minimization techniques for ab initio total-energy calculations: molecular dynamics and conjugate gradients. *Rev. Mod. Phys.*, 64:1045–1097, 1992.

[112] M. Marsman. *Periodic systems, plane waves, the PAW method, and hybrid functionals*, pages 61–76. Wiley-VCH Verlag GmbH and Co. KGaA, 2009.

[113] H.J. Monkhorst and J.D. Pack. Special points for Brillouin zone integrations. *Phys. Rev. B*, 13:5188–5192, Jun 1976.

[114] P.E. Blöchl. Projector augmented-wave method. *Phys. Rev. B*, 50:17953–17979, 1994.

[115] G. Kresse and D. Joubert. From ultrasoft pseudopotentials to the projector augmented-wave method. *Phys. Rev. B*, 59:1758–1775, 1999.

[116] D. Frenkel and B. Smit. *Understanding molecular simulation.* Academic Press, 2001.

[117] S. Nosé. A unified formulation of the constant temperature molecular dynamics methods. *J. Chem. Phys.*, 81(1):511–519, 1984.

[118] A.Y. Toukmaji and J.A. Board Jr. Ewald summation techniques in perspective: a survey. *Computer Physics Communications*, 95(23):73 – 92, 1996.

[119] S. Goedecker and K. Maschke. Operator approach in the linearized augmented-plane-wave method: Efficient electronic-structure calculations including forces. *Phys. Rev. B*, 45:1597–1604, Jan 1992.

[120] R.P. Feynman. Forces in molecules. *Phys. Rev.*, 56:340–343, 1939.

[121] G. Kresse and J. Furthmller. Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set. *Computational Materials Science*, 6(1):15 – 50, 1996.

[122] L. Verlet. Computer "experiments" on classical fluids. I. Thermodynamical properties of Lennard-Jones molecules. *Phys. Rev.*, 159:98–103, 1967.

[123] M.I.J. Probert. Improved algorithm for geometry optimisation using damped molecular dynamics. *J. Comput. Phys.*, 191(1):130 – 146, 2003.

[124] W.F. van Gunsteren and H.J.C. Berendsen. Computer simulation of molecular dynamics: Methodology, applications, and perspectives in chemistry. *Angewandte Chemie International Edition in English*, 29(9):992–1023, 1990.

[125] F. Tassone, F. Mauri, and R. Car. Acceleration schemes for ab initio molecular-dynamics simulations and electronic-structure calculations. *Phys. Rev. B*, 50:10561–10573, Oct 1994.

[126] W.G. Hoover. Canonical dynamics: equilibrium phase-space distributions. *Phys. Rev. A*, 31:1695–1697, 1985.

[127] D.J. Evans and B.L. Holian. The Nose–Hoover thermostat. *J. Chem. Phys.*, 83(8):4069–4074, 1985.

[128] A.R. Leach. *Molecular modelling: principles and applications*. Prentice Hall, 2001.

[129] A.M. Ovrutsky, A.S. Prokhoda, and M.S. Rasshchupkyna. *Computational materials science: surfaces, interfaces, crystallization*. Elsevier, 2013.

[130] G. Kresse and J. Hafner. Ab initio molecular dynamics for liquid metals. *Phys. Rev. B*, 47:558–561, Jan 1993.

[131] Kopin Liu. *Quantum dynamical resonances in chemical reactions: from A + BC to polyatomic systems*, pages 1–46. John Wiley and Sons, Inc., 2012.

[132] Y. Waseda and K. Suzuki. Structure of molten silicon and germanium by x-ray diffraction. *Zeitschrift fr Physik B Condensed Matter*, 20(4):339–343, 1975.

[133] T.E. Faber and J.M. Ziman. A theory of the electrical properties of liquid metals. *Philos. Mag.*, 11(109):153–173, 1965.

[134] N.W. Ashcroft and D.C. Langreth. Structure of binary liquid mixtures. I. *Phys. Rev.*, 156:685–692, 1967.

[135] P. Vashishta, R.K. Kalia, J.P. Rino, and I. Ebbsjö. Interaction potential for SiO_2: a molecular-dynamics study of structural correlations. *Phys. Rev. B*, 41:12197–12209, 1990.

[136] J.P. Hansen and I.R. McDonald. *Theory of simple liquids*. Number November. Elsevier, third edition, 2005.

[137] C.H. Wang. *Spectroscopy of condensed media: dynamics of molecular interactions*. Academic Press Inc, 1985.

[138] L.E. Hintzsche, C. Fang, T. Watts, M. Marsman, G. Jordan, M.W.P.E. Lamers, A.W. Weeber, and G. Kresse. Density functional theory study of the structural and electronic properties of amorphous silicon nitrides: Si_3N_{4-x}:H. *Phys. Rev. B*, submitted 2012.

[139] L.E. Hintzsche, C.M. Fang, M. Marsman, G. Jordan, M.W.P.E. Lamers, A.W. Weeber, and G. Kresse. Defects and defect healing in amorphous $Si_3N_{4-x}H_y$: An *ab initio* density functional theory study. *Phys. Rev. B*, 88:155204, Oct 2013.

[140] J. Furthmüller, J. Hafner, and G. Kresse. *Ab initio* calculation of the structural and electronic properties of carbon and boron nitride using ultrasoft pseudopotentials. *Phys. Rev. B*, 50:15606–15622, 1994.

[141] G. Kresse and J. Furthmüller. Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Phys. Rev. B*, 54:11169–11186, 1996.

[142] T.M. Searle. *Properties of amorphous silicon and its alloys*. Institution of Engineering and Technology, 1998.

[143] R.E.I. Schropp, S. Nishizaki, Z.S. Houweling, V. Verlaan, C.H.M. van der Werf, and H. Matsumura. All hot wire CVD TFTs with high deposition rate silicon nitride (3nm/s). *Solid-State Electronics*, 52(3):427 – 431, 2008.

[144] A.R. Denton and N.W. Ashcroft. Vegard's law. *Phys. Rev. A*, 43:3161–3164, 1991.

[145] E. Bustarret, M. Bensouda, M.C. Habrard, J.C. Bruyère, S. Poulin, and S.C. Gujrathi. Configurational statistics in a-$Si_xN_yH_z$ alloys: a quantitative bonding analysis. *Phys. Rev. B*, 38(12):8171–8184, 1988.

[146] H. Mehrer. *Diffusion in solids.* Springer, 2007.

[147] R. Riedel and M. Seher. Crystallization behaviour of amorphous silicon nitride. *J. Eur. Ceram. Soc.*, 7(1):21 – 25, 1991.

[148] Y. Cerenius. Melting-temperature measurements on alpha-silicon nitride to a pressure of 37 GPa. *J. Am. Ceram. Soc.*, 82(2):380–386, 1999.

[149] M. Lamers, L.E. Hintzsche, K.T. Butler, P.E. Vullum, C.-M. Fang, M. Marsman, G. Jordan, J.H. Harding, G. Kresse, and A. Weeber. The interface of a-SiN_x:H and Si: linking the nano-scale structure to passivation quality. *Sol. Energy Mater. Sol. Cells*, 2013.

[150] K. Jarolimek, R.A. de Groot, G.A. de Wijs, and M. Zeman. First-principles study of hydrogenated amorphous silicon. *Phys. Rev. B*, 79:155206, 2009.

[151] J. Robertson. Electronic structure of silicon nitride. *Philos. Mag. B*, 63(1):47–77, 1991.

[152] M.H. Cohen. Review of the theory of amorphous semiconductors. *J. Non-Cryst. Solids*, 4(0):391 – 409, 1970.

[153] M. Yashima, Y. Ando, and Y. Tabira. Crystal structure and electron density of -silicon nitride:experimental and theoretical evidence for the covalent bonding and charge transfer. *J. Phys. Chem. B*, 111(14):3609–3613, 2007. PMID: 17388537.

[154] J. Tauc. Absorption edge and internal electric fields in amorphous semiconductors. *Mater. Res. Bull.*, 5(8):721 – 729, 1970.

[155] J. Tauc and A. Menth. States in the gap. *J. Non-Cryst. Solids*, 810(0):569 – 585, 1972.

[156] D.L. Wood and J. Tauc. Weak absorption tails in amorphous semiconductors. *Phys. Rev. B*, 5:3144–3151, 1972.

[157] F. Urbach. The long-wavelength edge of photographic sensitivity and of the electronic absorption of solids. *Phys. Rev.*, 92:1324–1324, 1953.

[158] K. Saito and A.J. Ikushima. Absorption edge in silica glass. *Phys. Rev. B*, 62:8584–8587, 2000.

[159] K.B.K. Teo, A.C. Ferrari, G. Fanchini, S.E. Rodil, J. Yuan, J.T.H. Tsai, E. Laurenti, A. Tagliaferro, J. Robertson, and W.I. Milne. Highest optical gap tetrahedral amorphous carbon. *Diamond and Related Materials*, 11(36):1086 – 1090, 2002.

[160] M. Gajdoš, K. Hummer, G. Kresse, J. Furthmüller, and F. Bechstedt. Linear optical properties in the projector-augmented wave methodology. *Phys. Rev. B*, 73:045112, Jan 2006.

[161] C. Ambrosch-Draxl and J.O. Sofo. Linear optical properties of solids within the full-potential linearized augmented planewave method. *Computer Physics Communications*, 175(1):1 – 14, 2006.

[162] S. Hasegawa, M. Matsuda, and Y. Kurata. Bonding configuration and defects in amorphous SiN_x:H films. *Appl. Phys. Lett.*, 58(7):741–743, 1991.

[163] I.V. Afanasyev-Charkin, L.G. Jacobsohn, R.D. Averitt, and M. Nastasi. Amorphous silicon nitride films of different composition deposited at room temperature by pulsed glow discharge plasma immersion ion implantation and deposition. *J. Vac. Sci. Technol. A: Vacuum, Surfaces, and Films*, 22(6):2342–2346, 2004.

[164] P.A. Khomyakov, W. Andreoni, N.D. Afify, and A. Curioni. Large-scale simulations of a-si:h: the origin of midgap states revisited. *Phys. Rev. Lett.*, 107:255502, 2011.

[165] R.A. Bertlmann. *Anomalies in quantum field theory.* Oxford University Press, 2001.

[166] M.A.L. Marques, C.A. Ullrich, F. Nogueira, A. Rubio, K. Burke, and E.K.U. Gross. *Time-dependent density functional theory.* Springer, 2006.

[167] L.V. Keldysh, D.A. Kirzhnitz, and A.A. Maradudin. *The dielectric function of condensed systems.* North Holland, 1989.

[168] V.M. Agranovich and V.L. Ginzburg. *Crystal optics with spatial dispersion, and excitons: an account of spatial dispersion.* Springer, 1984.

[169] G. Russakoff. A derivation of the macroscopic maxwell equations. *Am. J. Phys.*, 38(10):1188–1195, 1970.

[170] D.A. Kirzhnits. Are the Kramers-Kronig relations for the dielectric permittivity of a material always valid? *Soviet Physics Uspekhi*, 19(6):530, 1976.

[171] O.V. Dolgov, D.A. Kirzhnits, and E.G. Maksimov. On an admissible sign of the static dielectric function of matter. *Rev. Mod. Phys.*, 53:81–93, Jan 1981.

[172] R. Car, E. Tosatti, S. Baroni, and S. Leelaprute. Dielectric band structure of crystals: General properties and calculations for silicon. *Phys. Rev. B*, 24:985–999, Jul 1981.

[173] . Baroni and R. Resta. Ab initio. *Phys. Rev. B*, 33:7017–7021, May 1986.

[174] M.A.L. Marques, N. T. Maitra, F. M.S. Nogueira, E.K.U. Gross, and A. Rubio. *Fundamentals of time-dependent density functional theory.* Springer, 2012.

[175] R.F. Nalewajski, F.A. Buot, J.F. Dobson, R.M. Dreizler, E. Engel, M. Gross, E.K.U.and Petersilka, and A.K. Rajagopal. *Density functional theory II: relativistic and time dependent extensions.* Springer, 1996.

[176] M.S. Hybertsen and S.G. Louie. *Ab initio* static dielectric matrices from the density-functional approach. i. formulation and application to semiconductors and insulators. *Phys. Rev. B*, 35:5585–5601, 1987.

Acknowledgments

First of all, I thank my family who have strongly supported my studies of physics. Since I had always been interested in natural science and computer technology, they really gave me the chance to realize one of my personal dreams. Especially, I want to thank Doris Hohenecker and Julian Hohenecker for bearing with me during my Ph.D. time. Maybe it was not always easy-going with me.

Special thanks I owe to Georg Kresse who did a impressive job as my supervisor. After a vivid job interview, he introduced me to the Computational Materials Physics group, and he enabled me to work on the HyperSol project. In our regular discussions, I learned a lot from his expertise and his experience. His feedback was always clear, constructive and motivating. I think that his foresight contributed greatly to the success of the overall project and prevented me from doing many beginner's mistakes.

Likewise, the working atmosphere in the group was very pleasant and professional. This situation was definitely due to my friendly, cooperative group members. Not only I had a lot of rewarding discussions with my room mates Gerald Jordan and Changming Fang during the project, but also with Merzuk Kaltak, Jiri Klimes, Tobias Sander, and Ronald Starke who gladly shared their opinions on density functional theory, GW methods, and linear response theory with me. Furthermore, I would like to thank Martijn Marsman for his support whenever Georg was too busy to have a discussion. Maybe it is because Martijn is a also experienced trainee in martial arts, but he is a real expert in solving difficult situations. No matter what I had to edit, he always had good advices for me. For technical questions, on the other hand, there was one fix point in our group. Michael Pöltl was the man who managed all computational requirements. He helped me to control the mass of data that I had generated with my calculations, and he often gave me expert information on hardware and software issues. For these and many other reasons, I thank all of my colleges for the great and exciting time we had together.

Finally, my work benefited very much from the help from our cooperation partners in the HiperSol project. For example, I often had a very professional and confidential exchange of information with Machteld Lamers and Arthur Weeber, from the Energy Research Center of the Netherlands, and with Keith Butler and John Harding, from the University of Sheffield. Thereby, I have learned how important regular discussions are in order to verify one's own scientific understanding and to harmonize theoretical with experimental concepts. Finally, I thank the company SINTEF for initiating the HiperSol project and the European Union for their financial support.

Curriculum Vitae

Personal Data

Name	Mag Leif Eric Hintzsche
Address	Untere Kirchengasse 26, 2452 Mannersdorf, Austria
Telephone	+43 (0)660 5742 787
Email	leif.eric.hintzsche@googlemail.com
Date of birth	09.12.1982
Place of birth	Stuttgart-Bad Cannstatt, Germany

Education

02/11 – present **Research Assistant for the EU-Project: "High Performance Solar Cells"**

Institute of Computational Materials Physics, University of Vienna, Austria

- PhD thesis on the simulation of amorphous silicon nitride for solar cell applications, supervisor: Univ.-Prof. Dr. Georg Kresse

10/04 – 02/11 **Study of Physics**

University of Vienna, Austria

- Specialization in computational physics, statistical physics, and solid state physics
- Diploma thesis on a MEAM-potential to simulate force fields and atom diffusion, supervisor: Univ.-Prof. Dr. Wolfgang Püschl

07/03 **Abitur (A-levels)**

Gustav-Stresemann Gymnasium, Fellbach, Germany

Publications

10/13 L.E. Hintzsche, C.M. Fang, M. Marsman, G. Jordan, M.W.P.E. Lamers, A.W. Weeber, and G. Kresse, Phys. Rev. B 88, 155204: "Defects and defect healing in amorphous silicon nitrides: an ab initio density functional theory study"

05/13 M.W.P.E. Lamers, L.E. Hintzsche, K.T. Butler, P.E. Vullum, C.M. Fang, M. Marsman, G. Jordan, J.H. Harding, G. Kresse, and A.W. Weeber, Sol. Mater
Sol. Energy Mater. Sol. Cells 120, 311: "The interface of a-SiN_x:H and Si: linking the nano-scale structure to passivation quality"

12/12 L.E. Hintzsche, C.M. Fang, T. Watts, M. Marsman, G. Jordan, M.W.P.E. Lamers, A.W. Weeber, and G. Kresse, Phys. Rev. B 86, 235204: "Density functional theory study of the structural and electronic properties of amorphous silicon nitrides"

03/12 M. Marsman, G. Jordan, L.E. Hintzsche, Y.S. Kim, and G. Kresse, Phys. Rev. B 85, 115122: "Gaussian charge-transfer charge distributions for non-self-consistent electronic structure calculations"

01/12 G. Kresse, M. Marsman, L.E. Hintzsche, and E. Flage-Larsen, Phys. Rev. B 85, 045205: "Optical and electronic properties of Si_3N_4 and α-SiO_2"